



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE CIENCIAS

ANÁLISIS DE INTERVENCIÓN Y FUNCIÓN DE
TRANSFERENCIA: UN ESTUDIO APLICADO A
SERIES DE TIEMPO ECONÓMICAS EN MÉXICO

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

PRESENTA:

JESÚS FRANCISCO LEYVA PICHARDO

DIRECTOR DE TESIS:

DR. RICARDO RAMÍREZ ALDANA



Ciudad Universitaria, Cd. Mx., 2018



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

*“Facts are stubborn things,
but statistics are pliable.”
Mark Twain*

Agradecimientos

A Dios todopoderoso, por haberme guiado por el camino correcto y darme las fuerzas necesarias para seguir adelante. Sin él nada de esto sería posible.

A mis padres: Eliezer Leyva y María Isabel Pichardo por su comprensión, paciencia y apoyo a lo largo de este sinuoso camino. Gracias por esforzarse para que nunca me falte nada.

A mis hermanos: María Fernanda y José Armando por ayudarme en los momentos complicados y por enseñarme que para lograr tus metas sólo necesitas tener un poco de fé y esperanza.

Al Dr. Ricardo Ramírez Aldana por su tiempo, compromiso y dedicación con este trabajo. Muchas gracias por su confianza y respaldo.

Al M. en C. José Salvador Zamora, al M. en C. Fernando Daniel Pérez Arriaga, al Act. Gerardo Sisniega Lira y al Dr. Alberto Contreras Cristán por su tiempo y empeño en esta tesis. Sus comentarios y observaciones perfeccionaron de gran manera este trabajo.

A la UNAM y en especial a la Facultad de Ciencias, por abrirme sus puertas y darme la oportunidad de aprender nuevas cosas. Sus aulas me permitieron crecer y madurar como persona.

Índice general

1. Modelos para Series de Tiempo univariadas	1
1.1. Introducción	1
1.2. Procesos Estacionarios	2
1.2.1. Ruido Blanco	4
1.2.2. Operadores de Retraso	5
1.3. Modelos de Promedios Móviles	7
1.3.1. Modelo MA(1)	8
1.3.2. Modelo MA(2)	10
1.3.3. Modelo MA(q)	13
1.4. Modelos Autorregresivos (AR)	15
1.4.1. Modelo AR(1)	15
1.4.2. Modelo AR(2)	18
1.4.3. Modelo AR(p)	21
1.5. Modelos ARMA	24
1.5.1. Modelo ARMA(1,1)	24
1.5.2. Modelo ARMA(p,q)	27
1.6. Modelos ARIMA	32
1.6.1. Identificación del Modelo	34

1.6.2.	Estimación de los parámetros	34
1.6.3.	Verificación de supuestos.	35
1.6.4.	Pronóstico	37
1.7.	Implementación en R	40
2.	Función de Transferencia	51
2.1.	Introducción	51
2.2.	Teoría preliminar	52
2.3.	Modelo de Función de Transferencia	64
2.4.	Función de Correlación Cruzada	65
2.5.	Especificación, Ajuste y Verificación	67
2.6.	Pronóstico	75
2.7.	Procedimiento alternativo	77
2.8.	Implementación en R e ITSM	78
2.8.1.	Implementación en R	78
2.8.2.	Implementación en ITSM	92
3.	Análisis de Intervención	107
3.1.	Introducción	107
3.2.	Análisis de Intervención: teoría	108
3.3.	Modelos de Intervención	110
3.4.	Identificación, Ajuste y Verificación	115
3.5.	Implementación en R e ITSM	116
3.5.1.	Implementación en R	116
3.5.2.	Implementación en ITSM	130

4. Outliers en Series de Tiempo **141**

4.1. Introducción	141
4.2. Tipos de Outliers	142
4.3. Teoría de acuerdo a Chang y Tiao	145
4.3.1. Outlier Innovativo y Outlier Aditivo	145
4.3.2. Criterio de razón de verosimilitud	148
4.3.3. Implementación en R	153
4.4. Teoría de acuerdo a Chen y Liu	174
4.4.1. Algoritmo de detección, estimación y pronóstico	174
4.4.2. Implementación en R	177
4.5. Punto de Cambio	187
4.5.1. Detección de Puntos de Cambio	187
4.5.2. Detección de Punto de Cambio Único	188
4.5.3. Detección de Punto de Cambio Múltiple	189
4.5.4. Modelo de Intervención	190
4.5.5. Implementación en R	190

A. Inflación Mexicana **203**

B. Tasa de crecimiento del PIB de México **207**

Índice de figuras

1.1. Ruido Blanco Gaussiano simulado y su ACF	6
1.2. Simulación de un proceso MA(1) con $\theta_1=0.4$ y su ACF	9
1.3. Simulación de un proceso MA(2) con $\theta_1 =-0.7, \theta_2 =0.1$ y su ACF .	12
1.4. Simulación de un proceso AR(1) con $\phi_1 =0.8$ y su ACF	19
1.5. Simulación de un proceso AR(2) con $\phi_1 =0.9, \phi_2 =-0.2$ y su ACF	22
1.6. Simulación de un proceso ARMA(1,1) con $\phi_1 =0.6, \theta_1 =0.5$ y su ACF	27
1.7. Simulación proceso ARIMA(2,1,2) con $\phi_1 =-0.3, \phi_2 =-0.8, \theta_1 =0.2,$ $\theta_2 =-0.1$ y su ACF	33
1.8. Precio del petróleo Mezcla Mexicana	41
1.9. ACF y PACF MME	42
1.10. ACF y PACF MME Primera Diferencia	43
1.11. p-values de la prueba de Ljung-Box	45
1.12. ACF y PACF de los residuos	46
1.13. Q-Q plot e histograma	47
1.14. Gráfico de valores ajustados contra residuos	48
1.15. Pronóstico Mezcla Mexicana	50
2.1. Funciones de respuesta al impulso, primera estructura	59

2.2. Funciones de respuesta al impulso, segunda estructura	60
2.3. Funciones de respuesta al impulso, tercera estructura	60
2.4. Funciones de respuesta al impulso, cuarta estructura	61
2.5. Funciones de respuesta al impulso, quinta estructura	61
2.6. Funciones de respuesta al impulso, sexta estructura	62
2.7. Funciones de respuesta al impulso, séptima estructura	63
2.8. Funciones de respuesta al impulso, octava estructura	63
2.9. Funciones de respuesta al impulso, novena estructura	64
2.10. Gráficos de las tasas de crecimiento anual del PIB de México y Brasil, respectivamente	80
2.11. ACF y PACF del input	80
2.12. Correlaciones cruzadas entre Y y Z	83
2.13. ACF y PACF de N_t ajustada	85
2.14. Gráfica de los residuos del modelo ajustado en conjunto	87
2.15. ACF y PACF de los residuos del modelo ajustado conjunto	88
2.16. Función de Correlación Cruzada entre Z_t y N_t	88
2.17. Histograma y q-q plot normal de los residuos del modelo conjunto	89
2.18. Pronóstico de la tasa de crecimiento del PIB de México	93
2.19. Gráfica de la tasa de crecimiento del PIB de Brasil	94
2.20. ACF y PACF de la tasa de crecimiento de la economía de Brasil	94
2.21. Gráfica de la tasa de crecimiento del PIB de México	96
2.22. Residuos de ambas series usando el mismo modelo ajustado	96
2.23. Función de correlación aplicada \hat{Z}_t y \hat{Y}_t	98
2.24. Series $X_{t,1}$ y $X_{t,2}$	100
2.25. N_t	100

2.26. ACF y PACF de N_t	101
2.27. Residuos del modelo ajustado conjunto	103
2.28. ACF de los residuos del modelo ajustado conjunto	103
2.29. Pronóstico a 5 años de la tasa de crecimiento del PIB de México .	105
3.1. Función Escalón y función Pulso	111
3.2. Primera forma de la función Escalón y función Pulso	112
3.3. Segunda forma de la función Escalón y función Pulso	112
3.4. Tercera forma de la función Escalón y función Pulso	113
3.5. Tercera forma de la función Escalón y función Pulso	114
3.6. Última forma funcional	114
3.7. Inflación anual Mexicana (1961 – 2017)	117
3.8. Inflación Anual Mexicana (1961 – 2017) con $1982 \leq t \leq 1988$. . .	118
3.9. Gráfico de la función escalón e Inflación	120
3.10. ACF y PACF de N_t	122
3.11. Valores ajustados (rojo) y valores reales (azul)	124
3.12. Residuos del modelo Intervención1	125
3.13. ACF y PACF de los residuos del modelo conjunto	126
3.14. Histograma y QQplot de los residuos	126
3.15. Pronóstico Inflación Mexicana Anual (1961 – 2016)	129
3.16. Gráfico del efecto dinámico y la serie de Inflación Anual de México (1961 – 2016)	130
3.17. Residuos (N_t)	132
3.18. ACF y PACF de N_t	132
3.19. Residuos del modelo conjunto final	136

3.20. ACF y PACF de los residuos del modelo conjunto	136
3.21. QQplot normal de los residuos del modelo conjunto	136
3.22. Pronóstico a 5 años de la Inflación Anual de México (1961-2016) .	138
4.1. Cambios temporales con distinta delta	143
4.2. Efecto del outlier aditivo	144
4.3. Efecto de un cambio de nivel	145
4.4. Efecto del outlier innovativo	146
4.5. Efecto de un cambio temporal	147
4.6. Gráfica de la Inflación mensual de México (1970 – 1985)	154
4.7. Residuos del primer y último modelo	168
4.8. Pronóstico a 12 meses de la Inflación mensual Mexicana	173
4.9. Inflación Mensual de México 1994 – 2017	179
4.10. Serie original (verde) y el segundo modelo ajustado sin los efectos de outliers (rojo)	184
4.11. Pronóstico a un año de la Inflación Mexicana mensual (1994-2017)	186
4.12. Punto de cambio encontrado en la serie de Inflación General	191
4.13. Inflación General con $T =$ diciembre de 1988	192
4.14. ACF y PACF muestral de la Inflación General Mexicana hasta el dato 143	193
4.15. Valores de la Inflación General y valores ajustados	195
4.16. Efecto de la Intervención	196
4.17. Pronóstico Inflación General	198
A.1. Inflación Mensual Mexicana (1970 – 2017)	205
B.1. Tasa de Crecimiento del PIB (1961 – 2015)	208

Índice de cuadros

1.1. Comportamiento de la ACF y PACF para los modelos AR y MA .	32
1.2. Comportamiento de la ACF y PACF para los modelos ARMA . .	32
1.3. Precios Históricos mensuales de la Mezcla Mexicana	40
2.1. Tasas de crecimiento anual del PIB de México y Brasil	106
3.1. Inflación Anual de México 1961 – 2016	139
4.1. Inflación Mensual de México (1970 – 1985)	153
4.2. Inflación Mensual de México (1994-2017)	178
4.3. Inflación General de México (1977-2017)	199

Introducción

El presente proyecto pone énfasis en los temas de Análisis de Intervención, Función de Transferencia y Análisis de datos atípicos en una serie de tiempo. Las series de tiempo, en la mayoría de los casos, se encuentran influenciadas por sucesos externos, políticos, económicos, normativos, entre otros. Estos sucesos externos se denominan comúnmente *intervenciones*. La técnica de Análisis de Intervención consiste en evaluar el efecto de las intervenciones en el proceso de comportamiento de una serie de tiempo. El Análisis de Intervención se caracteriza porque con antelación se sospecha que algún suceso externo ha influido considerablemente en la serie temporal. Pero hay ocasiones en que no conocemos la existencia de estos sucesos o cuándo empiezan. Así, si la causa y el tiempo de ocurrencia de los sucesos se desconocen, entonces haremos uso del Análisis de datos atípicos. Por otro lado, el tema de Función de Transferencia es un tema estadístico de gran importancia, ya que trata de explicar una serie temporal utilizando otra como variable explicativa. En el esquema general, nos referimos a la variable explicativa como *variable input* y a la variable respuesta como *variable output*.

Estos temas son de interés debido a la necesidad de saber y entender cómo diversos acontecimientos pueden afectar de manera importante los diversos indicadores o variables de la economía mexicana. La forma en cómo se comporten los datos de la variable serán determinantes para saber su comportamiento a futuro.

El primer objetivo de este trabajo es el de servir como material de apoyo en cualquier investigación económica o estadística.

El segundo objetivo es el de introducir al lector al uso de *software* estadístico, en este caso, R (Cryer (2008)) e ITSM (Brockwell (2016)). Estos dos programas, a lo largo del tiempo, han demostrado que son herramientas esenciales en el quehacer estadístico.

El tercer objetivo es el de modelar el comportamiento de la tasa de crecimiento

del PIB de México.

El cuarto objetivo será modelar el comportamiento de las tasas de inflación (mensual, anual y general) de México. Para nadie es una sorpresa que la inflación impacta en el bolsillo de los mexicanos. Debido a ello, se usarán las técnicas de Análisis de Intervención y de datos atípicos.

En el capítulo 1 se tratará la metodología de series de tiempo propuesta por Box & Jenkins. Se verá cómo surgen y qué importancia tienen los modelos autorregresivos $AR(p)$, de medias móviles $MA(q)$, autorregresivos de medias móviles $ARMA(p,q)$ y autorregresivos integrados de medias móviles $ARIMA(p,q)$. En la parte práctica, se usará una serie de tiempo de los precios por barril de mezcla mexicana. A esta serie se le ajustarán un modelo, se le verificarán los supuestos y por último, se hará un pronóstico.

En el capítulo 2 veremos el tema de Función de Transferencia donde se dará su definición y sus características. Se mencionará lo que es la Función de Correlación Cruzada y su importante papel en este tema. Posteriormente, se darán los pasos de especificación, ajuste, verificación y pronóstico. Finalmente, se utilizará ITSM y R para ajustar este tipo de modelo, los datos que usaremos son: como input la tasa de crecimiento del PIB de Brasil y output la tasa de crecimiento del PIB de México.

En el capítulo 3 se analizará el concepto de Análisis de Intervención, así como su desarrollo y su efecto en series temporales. En este caso, en la parte práctica usaremos la Inflación mexicana anual de 1961 a 2016.

Finalmente, en el capítulo 4 se hará referencia al Análisis de datos atípicos o *outliers* en series de tiempo. Primero, se presentará la teoría de Chang y Tiao, luego la teoría de Chen y Liu, y para terminar la teoría de puntos de cambio. En la parte de aplicación, en el primer caso usaremos la Inflación mexicana mensual de 1970 a 1985, en el segundo caso la Inflación mexicana mensual de 1990 a 2017 y en último modelo la tasa de Inflación mexicana general.

Capítulo 1

Modelos para Series de Tiempo univariadas

1.1. Introducción

En un mundo globalizado, se generan grandes cantidades de información acerca de diversos temas políticos, informáticos, mercadológicos, económicos, etc. La información generada tanto de individuos como de empresas o países, es comúnmente recopilada con fines de investigación, para posteriormente crear estudios y análisis que nos lleven a una correcta planeación y toma de decisiones. Al registro de observaciones numéricas, efectuada a intervalos de tiempo fijo de tales eventos económicos, es a lo que generalmente se conoce como *series de tiempo*.

Un enfoque estocástico que se utilizará para el análisis de series de tiempo se basa en los trabajos de Box Jenkins y la metodología ARIMA.

La estrategia basada en este enfoque consta de los siguientes pasos:

1. Identificación del modelo.
2. Estimación de los parámetros del modelo.
3. Verificación de supuestos.
4. Pronóstico.
5. Uso del modelo.

Se empezará con la teoría de Procesos Estacionarios, después se ahondará en algunos operadores útiles en Series de Tiempo y finalmente se llegará a los modelos autorregresivos, promedios móviles, ARMA y ARIMA.

1.2. Procesos Estacionarios

Definición 1.1 *Se define a una serie de tiempo como un proceso estocástico $\{X_t\}_{t \in \mathbb{Z}}$, es decir, una familia de variables aleatorias, sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.*

Definición 1.2 $\{X_t\}_{t \in \mathbb{Z}}$ es un proceso estocástico de segundo orden si:

$$\mathbb{E}(X_t^2) < \infty$$

para todo $t \in \mathbb{Z}$.

Definición 1.3 Un proceso $\{X_t\}_{t \in T}$ se llama estrictamente estacionario (en el sentido fuerte) si:

$$(X_{t_1}, \dots, X_{t_n}) =_D (X_{t_1+\tau}, \dots, X_{t_n+\tau})$$

para cualesquiera elementos $t_1, \dots, t_n, \tau \in \mathbb{Z}$ y con $n \geq 1$.

Donde “ $=_D$ ” significa igualdad en distribución.

Definición 1.4 Sea $\{X_t\}$ de 2º orden, un proceso $\{X_t\}_{t \in T}$ se llama estacionario de segundo orden (débilmente estacionario) si:

$$\mathbb{E}(X_t) = \mu_t = \mu$$

$$Cov(X_t X_s) = Cov(X_{t+\tau} X_{s+\tau})$$

para todo $t, s, \tau \in \mathbb{Z}$

En otras palabras, la media $\mathbb{E}(X_t)$ y la $Cov(X_t X_s) = \mathbb{E}(X_t X_s) - \mathbb{E}(X_t) \mathbb{E}(X_s)$, son las mismas, a pesar de movernos en τ unidades de tiempo, es decir, no dependen del tiempo.

Se puede comprobar directamente que si una serie de tiempo es estrictamente estacionaria y con varianza finita, entonces es estacionaria en el sentido débil. Por otro lado, también se tiene que tomar en cuenta que hay procesos con varianza infinita y que además son estrictamente estacionarios pero no son débilmente estacionarios.

Finalmente, como aclaración cuando se hable de estacionario en general se refiere a estacionario en el sentido débil.

Definición 1.5 Si $\{X_t\}$ es un proceso estocástico a tiempo discreto y de segundo orden, entonces $\mathbf{C}(t_1, t_2) = \mathbb{E}((X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2}))$; $t_1, t_2 \in \mathbb{Z}$, es la función de autocovarianza del proceso.

Definición 1.6 La función de autocorrelación se define como:

$$\rho(t_1, t_2) = \frac{\mathbf{C}(t_1, t_2)}{\sqrt{\mathbb{V}(X_{t_1})}\sqrt{\mathbb{V}(X_{t_2})}} = \frac{\mathbf{C}(t_1, t_2)}{\sqrt{\sigma_{t_1}^2}\sqrt{\sigma_{t_2}^2}}$$

Con $\mathbb{V}(X_t) = \sigma_t^2$.

Utilizando la Definición 1.4 se tiene:

$$\rho(t_1, t_2) = \frac{|\mathbf{C}(t_1, t_2)|}{\sqrt{\sigma_{t_1}^2}\sqrt{\sigma_{t_2}^2}} \leq \frac{\sqrt{\mathbb{E}(X'^2)}\sqrt{\mathbb{E}(Y'^2)}}{\sqrt{\sigma_{t_1}^2}\sqrt{\sigma_{t_2}^2}} = \frac{\mathbb{E}((X_{t_1} - \mu_{t_1})^2(X_{t_2} - \mu_{t_2})^2)}{\sqrt{\sigma_{t_1}^2}\sqrt{\sigma_{t_2}^2}} = 1$$

es decir, $|\rho(t_1, t_2)| \leq 1$.

Si $\{X_t\}$ es estacionario de segundo orden, entonces $\mathbf{C}(t_1, t_2)$ es una función que en realidad depende de la *distancia* entre los tiempos $t_1 - t_2$.

Tomar $\tau = -t_1$, entonces se tiene:

$$\mathbf{C}(t_1, t_2) = \mathbf{C}(t_1 - t_1, t_2 - t_1) = \mathbf{C}(0, t_2 - t_1)$$

y usando $\tau = t_2$, se tiene:

$$\mathbf{C}(t_1, t_2) = \mathbf{C}(t_1 - t_2, 0)$$

Así que depende de $|t_1 - t_2|$.

De esta manera se puede obtener una sucesión S_τ tal que

$$\mathbf{C}(t_1, t_2) = S(|t_1 - t_2|) = S_\tau$$

con $\tau = |t_1 - t_2|$ donde $\{S_\tau\}_\tau$ es la sucesión de autocovarianzas.

Similarmente puede obtenerse la sucesión de autocorrelaciones.

Definición 1.7 *La sucesión de autocorrelación (ACF por sus siglas en inglés) está definida por*

$$\{\rho_\tau\}_\tau = \left\{ \frac{S_\tau}{S_0} \right\}_\tau$$

donde $\tau \in \mathbb{Z}$, $S_0 = \mathbb{V}(X_t)$ para cualquier t y $\rho_\tau = \frac{S_\tau}{\sqrt{S_0^2}} = \frac{S_\tau}{S_0}$.

Como observación final es necesario mencionar que $\mathbf{C}(0, t) = \text{Cov}(X_0, X_t)$ y no $\text{Cov}(0, X_t)$.

1.2.1. Ruido Blanco

Definición 1.8 *Un ruido blanco es una sucesión de variables aleatorias no correlacionadas tal que $\mathbb{E}(X_t) = 0$ y $\mathbb{V}(X_t) = \sigma^2 \forall t$. La notación es WN (del inglés white noise).*

La sucesión de autocovarianzas es:

$$\text{Cov}(X_t, X_{t+\tau}) = S_\tau = \begin{cases} \text{Cov}(X_t, X_t) = \mathbb{V}(X_t) = \sigma^2 & \text{si } \tau = 0, \\ \text{Cov}(X_t, X_{t+\tau}) = 0 & \text{si } \tau \neq 0. \end{cases}$$

Y su sucesión de autocorrelaciones es:

$$\rho_\tau = \begin{cases} 1 & \text{si } \tau = 0, \\ 0 & \text{si } \tau \neq 0. \end{cases}$$

Se puede observar que un ruido blanco es un proceso estacionario, ya que $\mathbb{E}(X_t) = 0$, $\mathbb{V}(X_t) = \sigma^2$ y $Cov(X_t, X_{t+\tau})$ no dependen del tiempo.

Definición 1.9 *Un proceso estocástico $\{X_t\}_{t \in \mathbb{Z}}$ es un ruido blanco estricto (denotado como SWN por sus siglas en inglés) si es una sucesión de variables aleatorias independientes e idénticamente distribuidas con varianza finita.*

En la Figura 1.1 se tiene una simulación de 500 observaciones de un proceso de ruido blanco gaussiano con media cero y varianza igual a 1 con su respectivo ACF, función de autocorrelación, en la cual se observa que todas las autocorrelaciones caen en los intervalos de confianza, de esta manera se está en presencia de un ruido blanco. El código en R es el siguiente:

```
ruidoblanco=rnorm(500,0,1)
plot.ts(ruidoblanco, main="Ejemplo de Ruido Blanco Gaussiano (0,1)"
, xlab="Tiempo", ylab="X(t)", col="red")
acf(ruidoblanco, main="Correlograma", col="2", lag=50)
mean(ruidoblanco)
[1] 0.04600405
var(ruidoblanco)
[1] 1.037781
```

Analizando la media, varianza y el ACF se puede observar que la simulación cumple con las características de ruido blanco.

Como observación final, se debe notar que, la distribución del ruido blanco es Gaussiana. Pero hay que tener cuidado, no por este ejemplo se debe asumir que todo ruido blanco forzosamente involucra una distribución Gaussiana o Normal. Este ejemplo es un proceso de ruido blanco con distribución Normal, media 0 y varianza igual a σ_X^2 . Se denotará al ruido blanco como $\{X_t\} \sim WN(0, \sigma_X^2)$, en este caso por ser ruido blanco Gaussiano se escribirá como: $\{X_t\} \sim WN_G(0, \sigma_X^2)$.

1.2.2. Operadores de Retraso

El primer operador que se mencionará se llama *operador de retraso*, que se denotará por la letra B (del inglés backward). Este operador se define de la siguiente manera:

$$BZ_t = Z_{t-1}$$

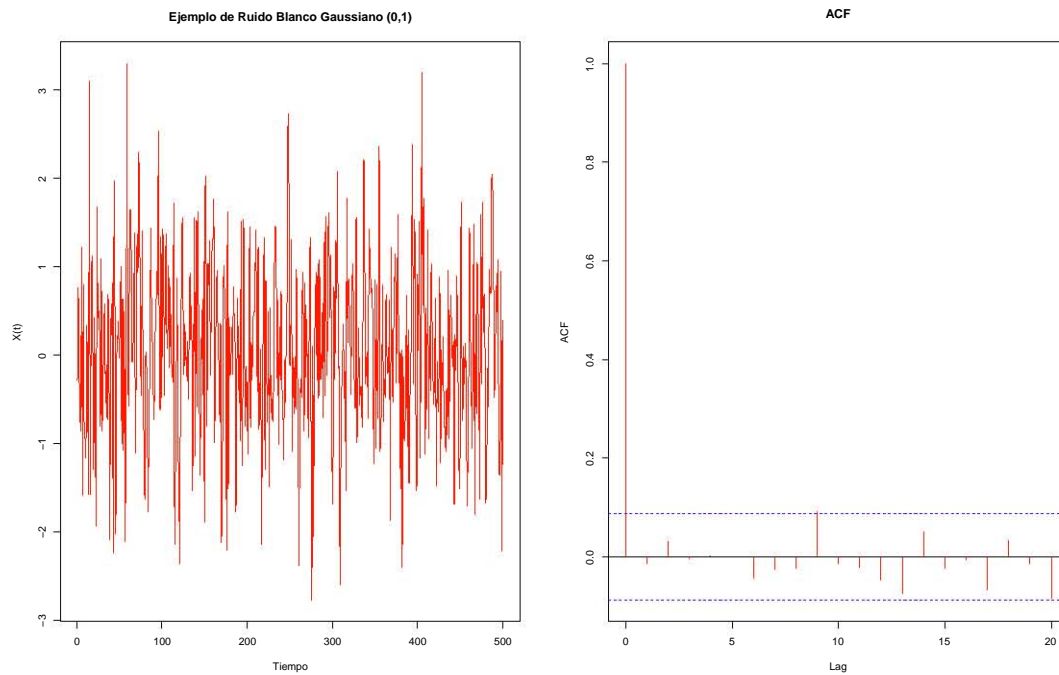


Figura 1.1: Ruido Blanco Gaussiano simulado y su ACF

Si se aplica sucesivamente el operador B se obtiene

$$B^2 Z_t = B(BZ_t) = Z_{t-2}$$

$$B^3 Z_t = B(B^2 Z_t) = Z_{t-3}$$

...

$$B^k Z_t = B(B^{k-1} Z_t) = Z_{t-k}$$

De forma general, la expresión a la que se llega es

$$B^k Z_t = Z_{t-k} \quad (1.1)$$

para $k = 0, 1, 2, \dots$ y $\forall t$.

Es necesario aclarar que al aplicar el operador B^k a Z_t se obtiene la variable retrasada k periodos y, debido a que $B^0 = 1$ se tiene $B^0 Z_t = Z_t$

Otro operador de uso frecuente es el operador diferencia ∇ , este operador está íntimamente ligado con B . Se escribe de la siguiente manera:

$$\nabla Z_t = Z_t - Z_{t-1} \text{ para toda } t$$

La relación que liga a ∇ con B es la siguiente

$$\nabla Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B) Z_t$$

de esta manera, se puede obtener una expresión general para ∇^k , así como se obtuvo para B^k . La expresión es la siguiente

$$\nabla^k Z_t = \sum_{j=0}^k \frac{k!}{j! (k-j)!} (-1)^j Z_{t-j} \quad \text{para } k = 0, 1, 2, \dots \text{ y toda } t$$

Usando el Teorema Binomio de Newton se tiene

$$\nabla^k Z_t = (1 - B)^k Z_t = \sum_{j=0}^k \binom{k}{j} (-B)^j Z_t. \quad (1.2)$$

1.3. Modelos de Promedios Móviles

Los modelos MA son llamados así por sus siglas en inglés (*Moving Average*) y fueron introducidos por Yule (1926) y Slutsky (1927). La idea de estos modelos es representar al proceso estocástico $\{X_t\}$ como una suma finita de choques aleatorios $\{\epsilon_t\}$, es decir

$$X_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \epsilon_t = \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i} = \theta(B) \epsilon_t$$

donde $\theta_1, \theta_2, \dots, \theta_q$ son los parámetros de promedios móviles y $\theta(B)$ es el polinomio de promedios móviles.

1.3.1. Modelo MA(1)

El proceso de promedios móviles de orden uno (MA(1)) se expresa de la siguiente manera

$$X_t = (1 + \theta_1 B) \epsilon_t = \epsilon_t + \theta_1 \epsilon_{t-1}$$

donde $\theta_1 \in \mathbb{R}$ y el proceso $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$.

Entonces, la esperanza y la varianza son

$$\mathbb{E}(X_t) = \mathbb{E}(\epsilon_t + \theta_1 \epsilon_{t-1}) = \mathbb{E}(\epsilon_t) + \theta_1 \mathbb{E}(\epsilon_{t-1}) = 0 \quad (1.3)$$

$$\begin{aligned} S_0 &= Var(X_t) \\ &= Var(\epsilon_t + \theta_1 \epsilon_{t-1}) \\ &= Var(\epsilon_t) + \theta_1^2 Var(\epsilon_{t-1}) \\ &= \sigma_\epsilon^2 (1 + \theta_1^2) < \infty \end{aligned} \quad (1.4)$$

La sucesión de autocovarianza está dada por

$$\begin{aligned} S_\tau &= Cov(X_t, X_{t+\tau}) \\ &= Cov(\epsilon_t + \theta_1 \epsilon_{t-1}, \epsilon_{t+\tau} + \theta_1 \epsilon_{t+\tau-1}) \\ &= Cov(\epsilon_t, \epsilon_{t+\tau}) + \theta_1 Cov(\epsilon_t, \epsilon_{t+\tau-1}) + \theta_1 Cov(\epsilon_{t-1}, \epsilon_{t+\tau}) + \theta_1^2 Cov(\epsilon_{t-1}, \epsilon_{t+\tau-1}) \end{aligned} \quad (1.5)$$

$$S_\tau = \begin{cases} \sigma_\epsilon^2 (1 + \theta_1^2) & \text{si } \tau = 0, \\ \sigma_\epsilon^2 \theta_1 & \text{si } |\tau| = 1, \\ 0 & \text{si } |\tau| > 1. \end{cases}$$

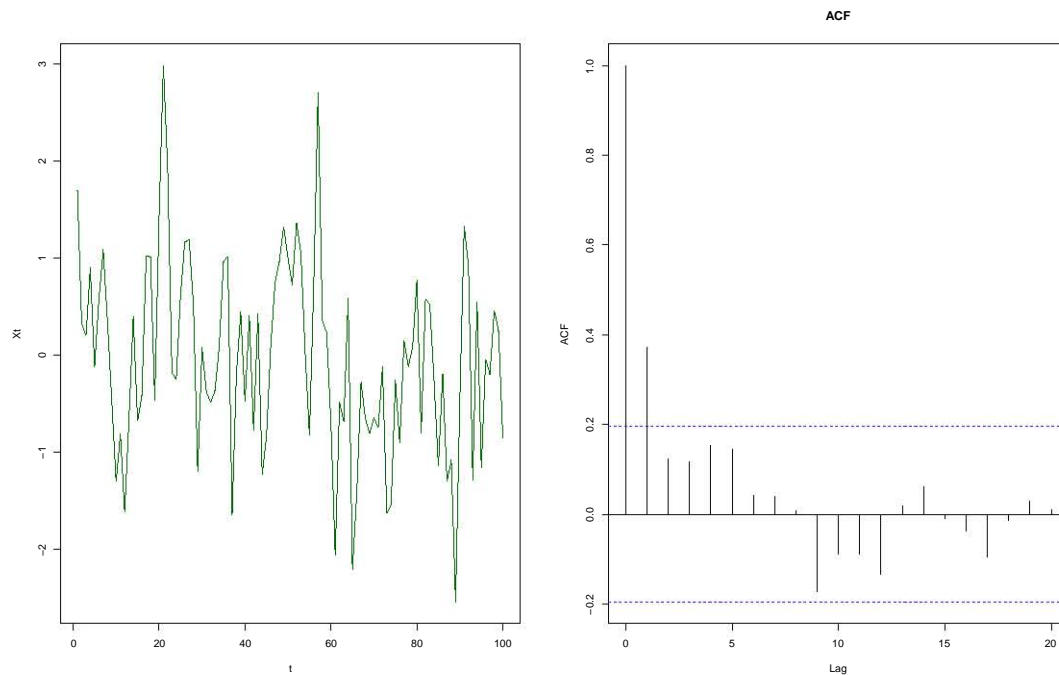


Figura 1.2: Simulación de un proceso MA(1) con $\theta_1=0.4$ y su ACF

Mientras que la sucesión de autocorrelaciones teórica (ACF) está dada por:

$$\rho_\tau = \begin{cases} 1 & \text{si } \tau = 0, \\ \frac{\theta_1}{1+\theta_1^2} & \text{si } |\tau| = 1, \\ 0 & \text{si } |\tau| > 1. \end{cases}$$

Se puede concluir que el proceso $\{X_t\}$ es estacionario, ya que $\mathbb{E}(X_t)$, $Var(X_t)$ y S_τ no dependen del tiempo. El siguiente código genera una simulación de 100 datos de un MA(1) con $\theta_1 = 0.4$. En la Figura 1.2 se puede observar la gráfica del proceso y su ACF.

```
ts.sim <- arima.sim(list(order = c(0,0,1), ma = 0.4), n = 100,
  sd = sqrt(1))
ts.plot(ts.sim, col="darkgreen",xlab="t",ylab="Xt")
acf(ts.sim,main="ACF")
```

Por lo mencionado anteriormente, se sabe que todo proceso MA es estacionario. Para comprobar que el proceso es invertible es necesario que la suma

$$\theta(x)^{-1} = 1 + \sum_{j=1}^{\infty} \theta^j x^j$$

converja dentro o sobre el círculo unitario, lo cual ocurre si $|\theta| < 1$. De forma sencilla, cuando un proceso se puede expresar mediante un modelo autorregresivo (sección 1.4.), se dirá que dicho proceso es invertible. Como consecuencia, la condición de invertibilidad para un proceso MA también se puede expresar en términos del polinomio $\theta(B)$ por el requerimiento de que las raíces de la ecuación

$$\theta(x) = 0$$

se encuentre fuera del círculo unitario. En la sección 1.5. se hablará más sobre el tema.

En este caso, el proceso está descrito como $X_t = \epsilon_t + 0.4\epsilon_{t-1}$ y tiene como raíz de la ecuación a $\theta(x) = 1 + 0.4x = 0$, lo que da $x = -2.5$, entonces se concluye que también es invertible, ya que está fuera del círculo unitario.

1.3.2. Modelo MA(2)

Un proceso MA(2) está denotado de la siguiente forma

$$X_t = (1 + \theta_1 B + \theta_2 B^2) \epsilon_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

donde $\theta_1, \theta_2 \in \mathbb{R}$ y $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$

El primer momento está descrito como

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}) \\ &= \mathbb{E}(\epsilon_t) + \theta_1 \mathbb{E}(\epsilon_{t-1}) + \theta_2 \mathbb{E}(\epsilon_{t-2}) \\ &= 0 \end{aligned} \tag{1.6}$$

La varianza se obtiene a partir de:

$$\begin{aligned}
\text{Var}(X_t) &= \text{Var}(\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2}) \\
&= \text{Var}(\epsilon_t) + \theta_1^2\text{Var}(\epsilon_{t-1}) + \theta_2^2\text{Var}(\epsilon_{t-2}) \\
&= \sigma_\epsilon^2 + \theta_1^2\sigma_\epsilon^2 + \theta_2^2\sigma_\epsilon^2 \\
&= \sigma_\epsilon^2(1 + \theta_1^2 + \theta_2^2)
\end{aligned} \tag{1.7}$$

A continuación, se construirá la sucesión de autocovarianzas teórica, ésta se define como:

$$\begin{aligned}
S_\tau &= \text{Cov}(X_t, X_{t+\tau}) \\
&= \text{Cov}(\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2}, \epsilon_{t+\tau} + \theta_1\epsilon_{t+\tau-1} + \theta_2\epsilon_{t+\tau-2}) \\
&= \text{Cov}(\epsilon_t, \epsilon_{t+\tau}) + \theta_1\text{Cov}(\epsilon_t, \epsilon_{t+\tau-1}) + \theta_2\text{Cov}(\epsilon_t, \epsilon_{t+\tau-2}) + \theta_1\text{Cov}(\epsilon_{t-1}, \epsilon_{t+\tau}) \\
&\quad + \theta_1^2\text{Cov}(\epsilon_{t-1}, \epsilon_{t+\tau-1}) + \theta_1\theta_2\text{Cov}(\epsilon_{t-1}, \epsilon_{t+\tau-2}) + \theta_2\text{Cov}(\epsilon_{t-2}, \epsilon_{t+\tau}) \\
&\quad + \theta_1\theta_2\text{Cov}(\epsilon_{t-2}, \epsilon_{t+\tau-1}) + \theta_2^2\text{Cov}(\epsilon_{t-2}, \epsilon_{t+\tau-2})
\end{aligned} \tag{1.8}$$

Entonces:

$$S_\tau = \begin{cases} \sigma_\epsilon^2(1 + \theta_1^2 + \theta_2^2) & \text{si } \tau = 0, \\ \sigma_\epsilon^2(\theta_1 + \theta_1\theta_2) & \text{si } \tau = 1, \\ \sigma_\epsilon^2\theta_2 & \text{si } \tau = 2, \\ 0 & \text{si } \tau \geq 3. \end{cases}$$

Por lo tanto, la sucesión de autocorrelaciones teórica (ACF) vendrá dada por:

$$\rho_\tau = \begin{cases} 1 & \text{si } \tau = 0, \\ \frac{(\theta_1 + \theta_1\theta_2)}{(1 + \theta_1^2 + \theta_2^2)} & \text{si } \tau = 1, \\ \frac{\theta_2}{(1 + \theta_1^2 + \theta_2^2)} & \text{si } \tau = 2, \\ 0 & \text{si } \tau \geq 3. \end{cases}$$

De los resultados anteriores resulta evidente que el proceso es estacionario, porque $\mathbb{E}(X_t)$, $\text{Var}(X_t)$ y S_τ no dependen del tiempo.

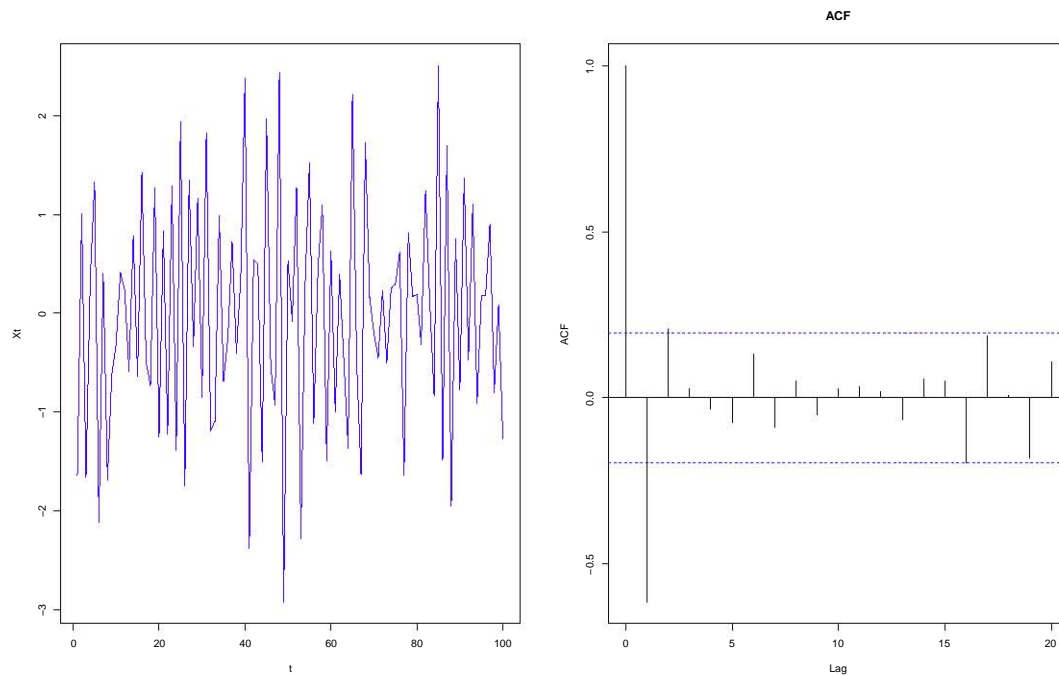


Figura 1.3: Simulación de un proceso MA(2) con $\theta_1 = -0.7$, $\theta_2 = 0.1$ y su ACF

El siguiente código genera una simulación de 100 datos de un MA(2) con $\theta_1 = -0.7$ y $\theta_2 = 0.1$. La Figura 1.3 muestra el proceso graficado con su respectivo ACF.

```
ts.sim <- arima.sim(list(order = c(0,0,2), ma = c(-0.7,0.1)),
  n = 100, sd = sqrt(1))
ts.plot(ts.sim, col="blue",xlab="t",ylab="Xt")
acf(ts.sim,main="ACF")
```

Para comprobar que el proceso simulado es también invertible se requiere que las raíces de la ecuación

$$1 + \theta_1 x + \theta_2 x^2 = 0 \quad (1.9)$$

con $\theta_1 = -0.7$ y $\theta_2 = 0.1$ se encuentren fuera del círculo unitario. Se obtienen que las raíces son: $x_1 = 5$ y $x_2 = 2$, entonces el proceso simulado es invertible.

Del teorema de Schur (sección 1.4.2.) se infiere que las condiciones de invertibilidad también se pueden obtener de la siguiente manera, usando:

$$|\theta_2| < 1 \quad \theta_2 + \theta_1 < 1 \quad \theta_2 - \theta_1 < 1 \quad (1.10)$$

Entonces, sustituyendo los valores de θ_1 y θ_2 en 1.10 se tiene

$$|0.1| < 1 \quad 0.1 - 0.7 < 1 \quad 0.1 + 0.7 < 1$$

Por lo tanto, el proceso simulado es estacionario e invertible.

1.3.3. Modelo MA(q)

En general, el proceso estocástico $\{X_t\}$ es de tipo promedios móviles de orden q si $\{X_t\}$ se puede representar mediante

$$X_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q} \quad \forall t \in \mathbb{Z}$$

$q \in \mathbb{N}$ y con $\theta_1, \theta_2, \dots, \theta_q \in \mathbb{R}$ donde el proceso $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$

Como regularmente se hace, se demostrará que el proceso MA(q) es estacionario, entonces primero se calculará la esperanza. La esperanza del proceso es de la siguiente manera

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q}) \\ &= \mathbb{E}(\epsilon_t) + \theta_1\mathbb{E}(\epsilon_{t-1}) + \theta_2\mathbb{E}(\epsilon_{t-2}) + \cdots + \theta_q\mathbb{E}(\epsilon_{t-q}) \\ &= 0 \end{aligned} \quad (1.11)$$

En el caso de la varianza se tiene:

$$\begin{aligned} \text{Var}(X_t) &= \text{Var}(\epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q}) \\ &= \text{Var}(\epsilon_t) + \theta_1^2\text{Var}(\epsilon_{t-1}) + \theta_2^2\text{Var}(\epsilon_{t-2}) + \cdots + \theta_q^2\text{Var}(\epsilon_{t-q}) \\ &= \sigma_\epsilon^2(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2) \end{aligned} \quad (1.12)$$

A continuación se obtiene la sucesión de autocovarianzas teórica. Dicha sucesión puede verse como:

$$S_\tau = Cov(X_t, X_{t+\tau}) = \mathbb{E}((X_t - \mu)(X_{t+\tau} - \mu)) = \mathbb{E}(X_t X_{t+\tau})$$

Suponiendo $\theta_0 = 1$ se tiene:

$$\begin{aligned} X_t X_{t+\tau} &= \left(\sum_{j=0}^q \theta_j \epsilon_{t-j} \right) \left(\sum_{l=0}^q \theta_l \epsilon_{t-l+\tau} \right) \\ &= \sum_{j=0}^q \sum_{l=0}^q \theta_j \theta_l \epsilon_{t-j} \epsilon_{t-l+\tau} \end{aligned} \quad (1.13)$$

Aplicando el operador esperanza y calculando se tiene:

$$\begin{aligned} \mathbb{E}(X_t X_{t+\tau}) &= \sum_{j=0}^q \sum_{l=0}^q \theta_j \theta_l \mathbb{E}(\epsilon_{t-j} \epsilon_{t-l+\tau}) \\ &= \sum_{j=0}^q \sum_{l=0}^q \theta_j \theta_l (\mathbb{E}(\epsilon_{t-j} \epsilon_{t-l+\tau}) - \mathbb{E}(\epsilon_{t-j}) \mathbb{E}(\epsilon_{t-l+\tau})) \\ &= \sum_{j=0}^q \sum_{l=0}^q \theta_j \theta_l S_{\tau+j-l} \end{aligned} \quad (1.14)$$

Recordar que S_τ es σ_ϵ^2 si $\tau = 0$ y cero en otro caso. En esta situación $\tau - l + j = 0$, o sea $l = \tau + j$. Como $l \leq q$ entonces $\tau + j \leq q$ si y sólo si $j \leq q - \tau$ donde $j \geq 0$, por lo tanto $0 \leq j \leq q - \tau$ y $\tau \leq q$. Entonces cuando $0 \leq \tau \leq q$, la covarianza no es cero, y se tiene que

$$S_\tau = \mathbb{E}(X_t X_{t+\tau}) \begin{cases} \left(\sum_{j=0}^{q-\tau} \theta_j \theta_l \right) \sigma_\epsilon^2 & \text{si } 0 \leq \tau \leq q, \\ 0 & \text{si } \tau > q. \end{cases}$$

de hecho $S_{-\tau}$ puede demostrarse que tiene la misma expresión y

$$S_{|\tau|} = \begin{cases} S_\tau & \text{si } \tau \geq 0, \\ S_{-\tau} & \text{si } \tau < 0. \end{cases}$$

con

$$S_{|\tau|} = \begin{cases} \left(\sum_{j=0}^{q-|\tau|} \theta_j \theta_{|\tau|+j} \right) \sigma_\epsilon^2 & \text{si } 0 \leq |\tau| \leq q, \\ 0 & \text{si } |\tau| > q. \end{cases}$$

Por otra parte $V(X_t) = S_0 = \sigma_\epsilon^2 \sum_{j=0}^q \theta_j^2$, entonces la sucesión de autocorrelación parcial es

$$\rho_{|\tau|} = \begin{cases} \left(\sum_{j=0}^{q-|\tau|} \theta_j \theta_{|\tau|+j} \right) \left(\sum_{j=0}^q \theta_j^2 \right)^{-1} & \text{si } 0 \leq |\tau| \leq q, \\ 0 & \text{si } |\tau| > q. \end{cases}$$

Es evidente que MA(q) es un proceso estacionario.

1.4. Modelos Autorregresivos (AR)

Se le da el nombre de *autorregresivo* porque el modelo se puede expresar como

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t$$

la cual es básicamente una ecuación de regresión lineal, con la característica de que el valor de la variable respuesta X en el periodo t depende no de los valores de un conjunto de variables independientes, sino de sus propios valores observados en periodos anteriores a t .

1.4.1. Modelo AR(1)

El caso más simple es el de un modelo autorregresivo de orden uno, o sea AR(1), que se representa como:

$$X_t = \phi_1 X_{t-1} + \epsilon_t$$

con $\phi_1 \in \mathbb{R}$ y $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$.

Para que dicha serie sea estacionaria se requiere, según el Teorema de Causalidad que se verá en la sección 1.5, que la raíz de la ecuación

$$1 - \phi_1 x = 0 \quad (1.15)$$

se encuentre fuera del círculo unitario; es decir, se requiere $|\phi| < 1$ para asegurar la estacionariedad del proceso AR(1).

El valor esperado del proceso $\{X_t\}$ será (si $|\phi_1| < 1$):

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}\left((1 - \phi_1 B)^{-1} \epsilon_t\right) \\ &= \mathbb{E}\left(\sum_{i=0}^{\infty} (\phi_1 B)^i \epsilon_t\right) \\ &= \mathbb{E}(\epsilon_t) + \phi_1 \mathbb{E}(\epsilon_{t-1}) + \phi_1^2 \mathbb{E}(\epsilon_{t-2}) + \dots \\ &= 0 \end{aligned} \quad (1.16)$$

donde $X_t - \phi_1 X_{t-1} = \epsilon_t$ es lo mismo que $X_t(1 - \phi_1 B) = \epsilon_t$ y despejando se tiene: $X_t = (1 - \phi_1 B)^{-1} \epsilon_t$. La serie es geométrica con razón $\phi_1 B$ y converge a $\frac{1}{1 - \phi_1 B}$ si y sólo si $|\phi_1| < 1$.

Se puede observar que un proceso AR(1) con coeficiente $|\phi_1| < 1$ puede escribirse como un proceso MA(∞), un proceso de medias móviles de orden infinito.

Para $\tau = 0$, la varianza del proceso es

$$\begin{aligned} S_0 &= Var(X_t) \\ &= Cov(X_t, X_t) \\ &= Cov(\phi_1 X_{t-1} + \epsilon_t, \phi_1 X_{t-1} + \epsilon_t) \\ &= \phi_1^2 Cov(X_{t-1}, X_{t-1}) + Cov(\epsilon_t, \epsilon_t) \\ &= \phi_1^2 Cov(X_0, X_0) + \sigma_\epsilon^2 \\ &= \phi_1^2 S_0 + \sigma_\epsilon^2 \end{aligned} \quad (1.17)$$

Donde, $Cov(X_{t-1}, X_{t-1}) = Cov(X_0, X_0)$ ya que el proceso $\{X_t\}$ es estacionario. Despejando S_0 se puede obtener la varianza del proceso, es decir, se tiene:

$$S_0 = \frac{\sigma_\epsilon^2}{1 - \phi_1^2}$$

Otra forma para obtener la varianza del proceso $\{X_t\}$ tomando en cuenta que $|\phi_1| < 1$ es la siguiente:

$$\begin{aligned}
\text{Var}(X_t) &= \text{Var}((1 - \phi_1 B)^{-1} \epsilon_t) \\
&= \text{Var}\left(\sum_{i=0}^{\infty} (\phi_1 B)^i \epsilon_t\right) \\
&= \text{Var}(\epsilon_t + \phi_1 \epsilon_{t-1} + \phi_1^2 \epsilon_{t-2} + \dots) \\
&= \sigma_\epsilon^2 + \phi_1^2 \sigma_\epsilon^2 + \phi_1^4 \sigma_\epsilon^2 + \dots \\
&= \sigma_\epsilon^2 (1 + \phi_1^2 + \phi_1^4 + \dots) \\
&= \frac{\sigma_\epsilon^2}{1 - \phi_1^2}
\end{aligned} \tag{1.18}$$

Ahora, para calcular la sucesión de autocovarianzas teórica, para $\tau > 1$ se tiene la siguiente relación:

$$\begin{aligned}
S_\tau &= \text{Cov}(X_t, X_{t+\tau}) \\
&= \text{Cov}(X_t, X_{t-\tau}) \\
&= \text{Cov}(\phi_1 X_{t-1} + \epsilon_t, X_{t-\tau}) \\
&= \phi_1 \text{Cov}(X_{t-1}, X_{t-\tau}) + \text{Cov}(\epsilon_t, X_{t-\tau}) \\
&= \phi_1 \text{Cov}(X_{t-1}, X_{t-\tau}) \\
&= \phi_1 S_{\tau-1} \\
&= \phi_1^\tau S_0 \\
&= \frac{\phi_1^\tau \sigma_\epsilon^2}{1 - \phi_1^2}
\end{aligned} \tag{1.19}$$

Donde, $\phi_1^\tau S_0$ se obtiene iterando el paso anterior y $\text{Cov}(\epsilon_t, X_{t-\tau}) = 0$, ya que el proceso $\{\epsilon_t\}$ no está correlacionado con el proceso X_s para cada $s < t$.

Se sigue que las autocorrelaciones deben ser de la forma

$$\begin{aligned}\rho_\tau &= \frac{S_\tau}{S_0} \\ &= \phi_1^\tau \\ \rho_{|\tau|} &= \phi_1^{|\tau|}, \quad \tau = 0, \pm 1, \pm 2, \dots\end{aligned}\tag{1.20}$$

lo cual indica que, conforme $\tau > 0$ crece, la función de autocorrelación (ACF) tiende a cero, con decaimiento exponencial cuando $0 < \phi_1 < 1$ y con signos alternados cuando $-1 < \phi_1 < 0$.

El siguiente código genera una simulación de 100 datos de un AR(1) con $\phi_1 = 0.8$. La Figura 1.4 muestra la imagen de la simulación con su ACF.

```
ts.sim <- arima.sim(list(order = c(1,0,0), ar = 0.8),
  n = 100, sd = sqrt(1))
ts.plot(ts.sim, col="orange3",xlab="t",ylab="Xt")
acf(ts.sim,main="ACF")
```

Todo proceso AR es invertible, para demostrar que también es estacionario se requiere que la raíz de la ecuación

$$1 - \phi_1 x = 0$$

se encuentre fuera del círculo unitario, sustituyendo $\phi_1 = 0.8$ en 1.15, se tiene que $1 - 0.8x = 0$, entonces $x = 1.25$. Por lo tanto el proceso simulado es estacionario e invertible.

1.4.2. Modelo AR(2)

El siguiente modelo que se considerará será el autorregresivo de segundo orden, el cual se define de la siguiente forma:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

con $\phi_1, \phi_2 \in \mathbb{R}$ y $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$

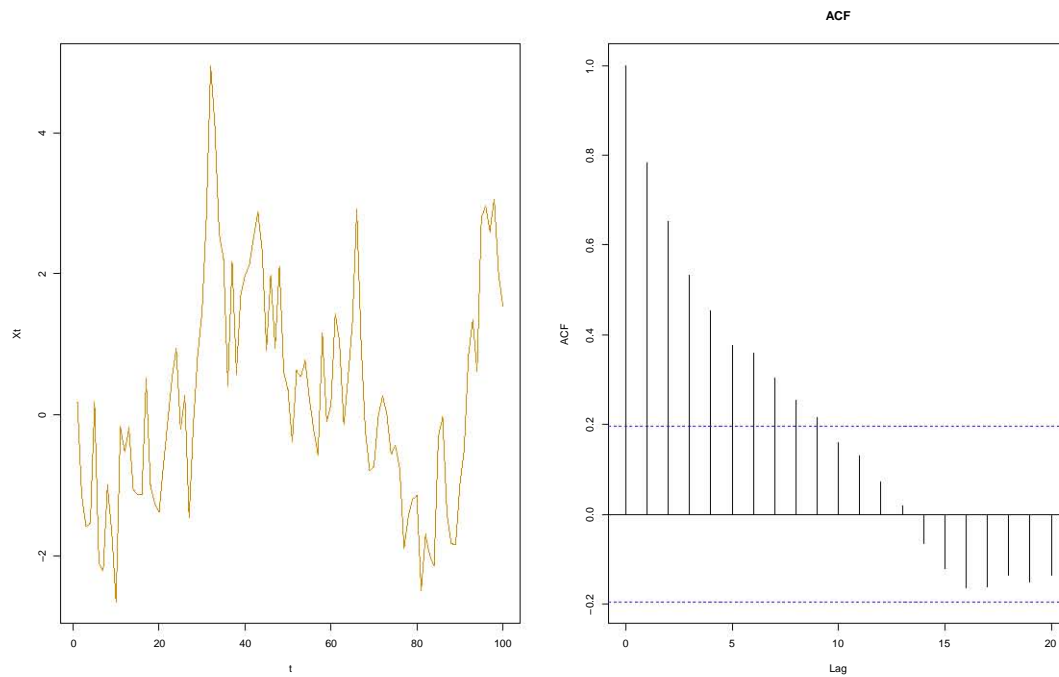


Figura 1.4: Simulación de un proceso AR(1) con $\phi_1 = 0.8$ y su ACF

Para que el proceso sea estacionario, según el Teorema de Causalidad, debe cumplirse que las raíces de

$$1 - \phi_1 x - \phi_2 x^2 = 0 \quad (1.21)$$

se encuentra fuera del círculo unitario.

Una forma alternativa para demostrar estacionariedad es a partir del Teorema de Schur el cual dice que el valor absoluto de las raíces del polinomio característico de orden p dado por la siguiente relación:

$$x^p - \alpha_1 x^{p-1} - \alpha_2 x^{p-2} - \dots - \alpha_{p-1} x - \alpha_p = 0$$

serán todos menores a uno, si y sólo si los p determinantes que se muestran a continuación son positivos:

$$D_1 = \begin{vmatrix} -1 & \alpha_p \\ \alpha_p & -1 \end{vmatrix} > 0$$

$$D_2 = \begin{vmatrix} -1 & 0 & \alpha_p & \alpha_{p-1} \\ \alpha_1 & -1 & 0 & \alpha_p \\ \alpha_p & 0 & -1 & \alpha_1 \\ \alpha_{p-1} & \alpha_p & 0 & -1 \end{vmatrix} > 0$$

$$D_p = \begin{vmatrix} -1 & 0 & \dots & 0 & \alpha_p & \alpha_{p-1} & \dots & \alpha_1 \\ \alpha_1 & -1 & \dots & 0 & 0 & \alpha_p & \dots & \alpha_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{p-1} & \alpha_{p-2} & \dots & -1 & \ddot{0} & 0 & \dots & \alpha_p \\ \alpha_p & 0 & \dots & 0 & -1 & \alpha_1 & \dots & \alpha_{p-1} \\ \alpha_{p-1} & \alpha_p & \dots & 0 & 0 & -1 & \dots & \alpha_{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_1 & \alpha_2 & \dots & \alpha_p & \ddot{0} & 0 & \dots & -1 \end{vmatrix} > 0$$

Para que un proceso AR(2) sea estacionario se debe cumplir que

$$D_1 = \begin{vmatrix} -1 & \phi_2 \\ \phi_2 & -1 \end{vmatrix} = 1 - \phi_2^2 > 0$$

$$D_2 = \begin{vmatrix} -1 & 0 & \phi_2 & \phi_1 \\ \phi_1 & -1 & 0 & \phi_2 \\ \phi_2 & 0 & -1 & \phi_1 \\ \phi_1 & \phi_2 & 0 & -1 \end{vmatrix} = (1 + \phi_2)^2 ((1 - \phi_2)^2 - \phi_1^2) > 0$$

La condición $D_1 > 0$ es equivalente a pedir que $|\phi_2| < 1$, mientras que la condición $D_2 > 0$ es equivalente a pedir que $(1 - \phi_2)^2 > \phi_1^2$, es decir, para que un proceso AR(2) sea estacionario es necesario que se cumplan las siguientes condiciones:

$$|\phi_2| < 1, \quad \phi_2 + \phi_1 < 1 \quad y \quad \phi_2 - \phi_1 < 1$$

Ahora, usando las ecuaciones de Yule-Walker, las cuales se verán con mayor detenimiento en la Sección 1.4.3, se tiene:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \tag{1.22}$$

las cuales permiten obtener los valores de ρ_1 y ρ_2 en función de los parámetros autorregresivos ϕ_1 y ϕ_2 . Resolviendo las ecuaciones de Yule-Walker para ρ_1 y ρ_2 se obtiene:

$$\begin{aligned} \rho_1 &= \frac{\phi_1}{1 - \phi_2} \\ \rho_2 &= \phi_2 + \frac{\phi_1^2}{1 - \phi_2} \end{aligned} \tag{1.23}$$

En general, el ACF satisface

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2}, \quad \tau \geq 3$$

Se pondrá especial atención al caso cuando $\phi_1^2 + 4\phi_2 \geq 0$, ya que las raíces de $\phi(z)$ son reales y la sucesión de autocorrelaciones teórica (ACF) decaerá rápidamente a cero (todas las autocorrelaciones serán positivas si la primera lo es y tendrán signos alternados si la primera autocorrelación es negativa), y si $\phi_1^2 + 4\phi_2 < 0$ las raíces serán complejas, así que el ACF correspondiente seguirá un comportamiento sinusoidal convergente a cero.

El siguiente código genera una simulación de 100 datos de un AR(2) con $\phi_1 = 0.9$ y $\phi_2 = -0.2$. La simulación se muestra en la Figura 1.5 con su respectivo ACF

```
ts.sim <- arima.sim(list(order = c(2,0,0), ar =c(0.9, -0.2)), n = 100,
  sd = sqrt(1))
ts.plot(ts.sim, col="navy",xlab="t",ylab="Xt")
acf(ts.sim,main="ACF")
```

La ecuación de este proceso, para verificar estacionariedad, es

$$1 - 0.9x + 0.2x^2 = 0$$

las raíces son: $x_1 = 2.5$ y $x_2 = 2$. Lo cual comprueba que el proceso simulado es estacionario e invertible.

1.4.3. Modelo AR(p)

Como caso general de un proceso autorregresivo se va a considerar el AR(p) que se describe de la siguiente forma:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t \quad \forall t \in \mathbb{Z}$$

con $\phi_1, \phi_2, \dots, \phi_p \in \mathbb{R}$ y $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$.

Un proceso AR(p) será estacionario si y sólo si las raíces de la ecuación característica

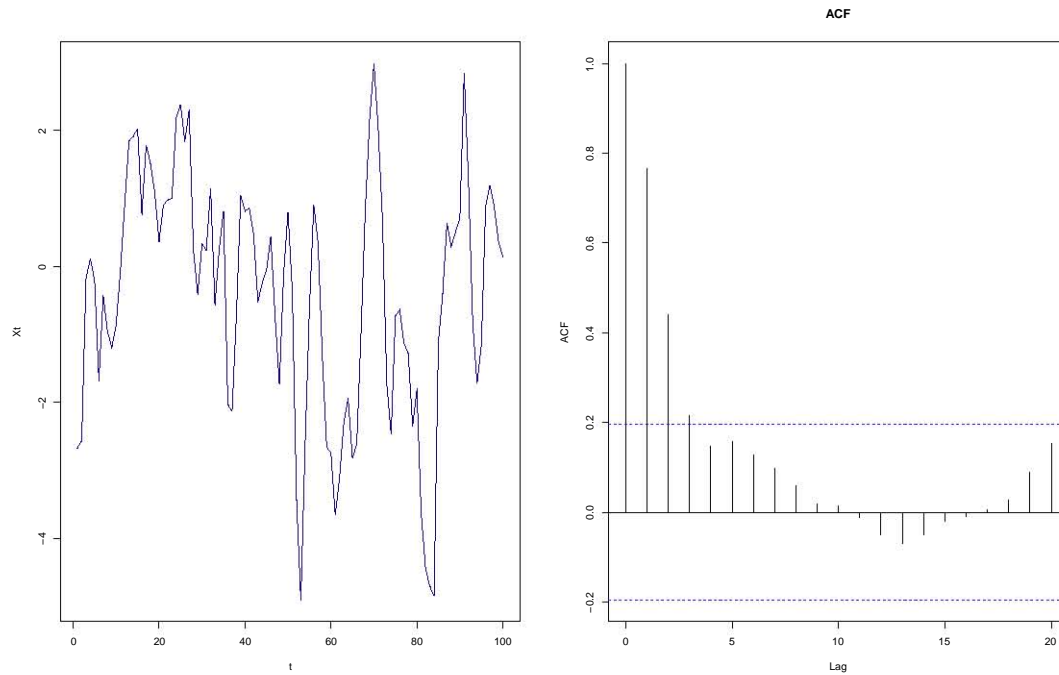


Figura 1.5: Simulación de un proceso AR(2) con $\phi_1 = 0.9$, $\phi_2 = -0.2$ y su ACF

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0$$

se encuentran fuera del círculo unitario. Como se vió en la sección 1.4.2 en la práctica el teorema de Schur es útil para encontrar las condiciones de estacionariedad en términos de los parámetros ϕ_1, \dots, ϕ_p . Aunque, hay otra manera de verificar el supuesto de estacionariedad y es utilizar las llamadas ecuaciones de Yule-Walker, correspondientes al proceso AR(p). La forma de obtenerlas es la siguiente, primero se hace uso de un proceso AR(p) y se multiplicará con el término $X_{t-\tau}$ con $1 \leq \tau \leq p$, posteriormente se aplicará el operador esperanza, haciendo todos estos cálculos entonces se tiene

$$\mathbb{E}(X_t X_{t-\tau}) = \phi_1 \mathbb{E}(X_{t-1} X_{t-\tau}) + \dots + \phi_p \mathbb{E}(X_{t-p} X_{t-\tau}) + \mathbb{E}(\epsilon_t X_{t-\tau})$$

Ahora, es posible demostrar a partir de la definición de causalidad (la cual se verá en la sección 1.5.) que la esperanza del proceso $\{X_t\}$ es cero y ver además a $\mathbb{E}(X_t X_{t-\tau})$ como:

$$\begin{aligned}
\mathbb{E}(X_t X_{t-\tau}) &= \mathbb{E}(X_t X_{t-\tau}) - \mathbb{E}(X_t) \mathbb{E}(X_{t-\tau}) \\
&= \text{Cov}(X_t, X_{t-\tau}) \\
&= S_\tau
\end{aligned} \tag{1.24}$$

Así como se obtuvo $\mathbb{E}(X_t X_{t-\tau}) = S_\tau$, se puede obtener también:

$$\begin{aligned}
\mathbb{E}(X_{t-1} X_{t-\tau}) &= S_{\tau-1} \\
\mathbb{E}(X_{t-2} X_{t-\tau}) &= S_{\tau-2} \\
&\dots \\
\mathbb{E}(X_{t-p} X_{t-\tau}) &= S_{\tau-p}
\end{aligned} \tag{1.25}$$

De manera que estas ecuaciones se transforman en:

$$S_\tau - \phi_1 S_{\tau-1} - \dots - \phi_p S_{\tau-p} = \mathbb{E}(\epsilon_t X_{t-\tau})$$

con ayuda de la definición de causalidad es posible ver que $\mathbb{E}(\epsilon_t X_{t-\tau}) = 0$. Al sustituir esta expresión en la ecuación anterior se obtiene, para $1 \leq \tau \leq p$, las siguientes expresiones

$$\begin{aligned}
\tau = 1 \quad S_1 - \phi_1 S_0 - \dots - \phi_p S_{p-1} &= 0 \\
\tau = 2 \quad S_2 - \phi_1 S_1 - \dots - \phi_p S_{p-2} &= 0 \\
&\vdots \\
\tau = p \quad S_p - \phi_1 S_{p-1} - \dots - \phi_p S_0 &= 0
\end{aligned} \tag{1.26}$$

Dividiendo entre S_0 , se obtienen finalmente las ecuaciones de Yule-Walker:

$$\begin{aligned}
\tau = 1 \quad \rho_1 - \phi_1 \rho_0 - \dots - \phi_p \rho_{p-1} &= 0 \\
\tau = 2 \quad \rho_2 - \phi_1 \rho_1 - \dots - \phi_p \rho_{p-2} &= 0 \\
&\vdots \\
\tau = p \quad \rho_p - \phi_1 \rho_{p-1} - \dots - \phi_p \rho_0 &= 0
\end{aligned} \tag{1.27}$$

de aquí se pueden obtener los valores de las p autocorrelaciones tomando siempre en cuenta los parámetros autorregresivos $\phi_1, \phi_2, \dots, \phi_p$.

De forma general las autocorrelaciones se pueden obtener de la siguiente igualdad

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \dots + \phi_p \rho_{\tau-p} \quad \tau \geq p + 1$$

Por último, es importante señalar que un proceso AR(p) estacionario tiene asociada un ACF que decae rápidamente a cero.

1.5. Modelos ARMA

A la combinación formada por procesos autorregresivos y de medias móviles se conoce como proceso ARMA. Dicha generalización surge debido a que muchas series de tiempo presentan características tanto de procesos AR como de procesos MA.

1.5.1. Modelo ARMA(1,1)

El proceso ARMA(1,1) está dado por la siguiente ecuación:

$$X_t - \phi_1 X_{t-1} = \epsilon_t + \theta_1 \epsilon_{t-1}$$

$\forall t \in \mathbb{Z}$, con $\phi_1, \theta_1 \in \mathbb{R}$ y $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$.

Sea un proceso ARMA(1,1) causal (en el siguiente subtema se hablará sobre un proceso causal) dado por $X_t = \phi_1 X_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}$. Se obtendrá la sucesión de autocovarianzas teórica y la sucesión de autocorrelaciones teórica asociadas al proceso X_t .

La varianza del proceso se obtiene de:

$$\mathbb{E}(X_t^2) = \phi_1 \mathbb{E}(X_t X_{t-1}) + \mathbb{E}(X_t \epsilon_t) + \theta_1 \mathbb{E}(X_t \epsilon_{t-1})$$

$$S_0 = \phi_1 \mathbb{E}(X_t X_{t-1}) + \mathbb{E}(X_t \epsilon_t) + \theta_1 \mathbb{E}(X_t \epsilon_{t-1})$$

donde

$$\begin{aligned}\mathbb{E}(X_t \epsilon_t) &= \phi_1 \mathbb{E}(X_{t-1} \epsilon_t) + \mathbb{E}(\epsilon_t^2) + \theta_1 \mathbb{E}(\epsilon_{t-1} \epsilon_t) \\ &= \sigma_\epsilon^2\end{aligned}\tag{1.28}$$

y

$$\begin{aligned}\mathbb{E}(X_t \epsilon_{t-1}) &= \phi_1 \mathbb{E}(X_{t-1} \epsilon_{t-1}) + \mathbb{E}(\epsilon_t \epsilon_{t-1}) + \theta_1 \mathbb{E}(\epsilon_{t-1}^2) \\ &= \phi_1 \sigma_\epsilon^2 + \theta_1 \sigma_\epsilon^2 \\ &= \sigma_\epsilon^2 (\phi_1 + \theta_1)\end{aligned}\tag{1.29}$$

Por lo tanto, la varianza del proceso está dada por:

$$S_0 = \phi_1 S_1 + [1 + \theta_1 (\theta_1 + \phi_1)] \sigma_\epsilon^2$$

De manera equivalente, la sucesión de autocovarianzas teórica del proceso está dada por:

$$S_\tau = \phi_1 \mathbb{E}(X_{t-1} X_{t-\tau}) + \mathbb{E}(X_{t-\tau} \epsilon_t) + \theta_1 \mathbb{E}(X_{t-\tau} \epsilon_{t-1})$$

es decir:

$$S_\tau = \begin{cases} \phi_1 S_0 + \theta_1 \sigma_\epsilon^2 & \text{si } \tau = 1, \\ \phi_1 S_{\tau-1} & \text{si } \tau \geq 2. \end{cases}$$

Ahora, se resolverá un sistema de ecuaciones entre S_0 y S_1 para encontrar sus soluciones y así poder expresar de mejor manera la sucesión de autocovarianzas teórica y la sucesión de autocorrelaciones teórica, entonces se tiene:

$$\begin{aligned}S_0 &= \phi_1 S_1 + [1 + \theta_1 (\phi_1 + \theta_1)] \sigma_\epsilon^2 \\ S_1 &= \phi_1 S_0 + \theta_1 \sigma_\epsilon^2\end{aligned}$$

Haciendo unos cuantos cálculos, al final se obtiene:

$$S_0 = \frac{[1 + 2\phi_1\theta_1 + \theta_1^2] \sigma_\epsilon^2}{1 - \phi_1^2}$$

$$S_1 = \frac{[(1 + \phi_1\theta_1)(\phi_1 + \theta_1)] \sigma_\epsilon^2}{1 - \phi_1^2}$$

Por lo tanto, la sucesión de autocovarianzas teórica está dada por:

$$S_\tau = \frac{\phi_1^{\tau-1} [(1 + \phi_1\theta_1)(\phi_1 + \theta_1)] \sigma_\epsilon^2}{1 - \phi_1^2} \quad \tau = 1, 2, \dots$$

Y la sucesión de autocorrelación teórica es:

$$\rho_\tau = \frac{\phi_1^{\tau-1} [(1 + \phi_1\theta_1)(\phi_1 + \theta_1)]}{1 + 2\phi_1\theta_1 + \theta_1^2} \quad \tau = 1, 2, \dots$$

El siguiente código genera una simulación de 100 datos de un ARMA(1,1) con $\phi_1 = 0.6$ y $\theta_1 = 0.5$. La simulación se puede observar en la Figura 1.6 con su ACF.

```
ts.sim <- arima.sim(list(order = c(1,0,1), ar = 0.6, ma = 0.5), n = 100,
  sd = sqrt(1))
ts.plot(ts.sim, col="brown",xlab="t",ylab="Xt")
acf(ts.sim,main="ACF")
```

El proceso simulado es de la forma: $X_t - 0.6X_{t-1} = \epsilon_t + 0.5\epsilon_{t-1}$ o escrito en términos de los operadores: $(1 - 0.6B)X_t = (1 + 0.5B)\epsilon_t$. Este modelo es estacionario porque la raíz de $\phi(x) = 1 - 0.6x = 0$ es $x = \frac{5}{3}$ el cual está fuera del círculo unitario.

El modelo también es invertible porque la raíz de $\theta(x) = 1 + 0.5x = 0$ es $x = -2$ el cual está también fuera del círculo unitario.

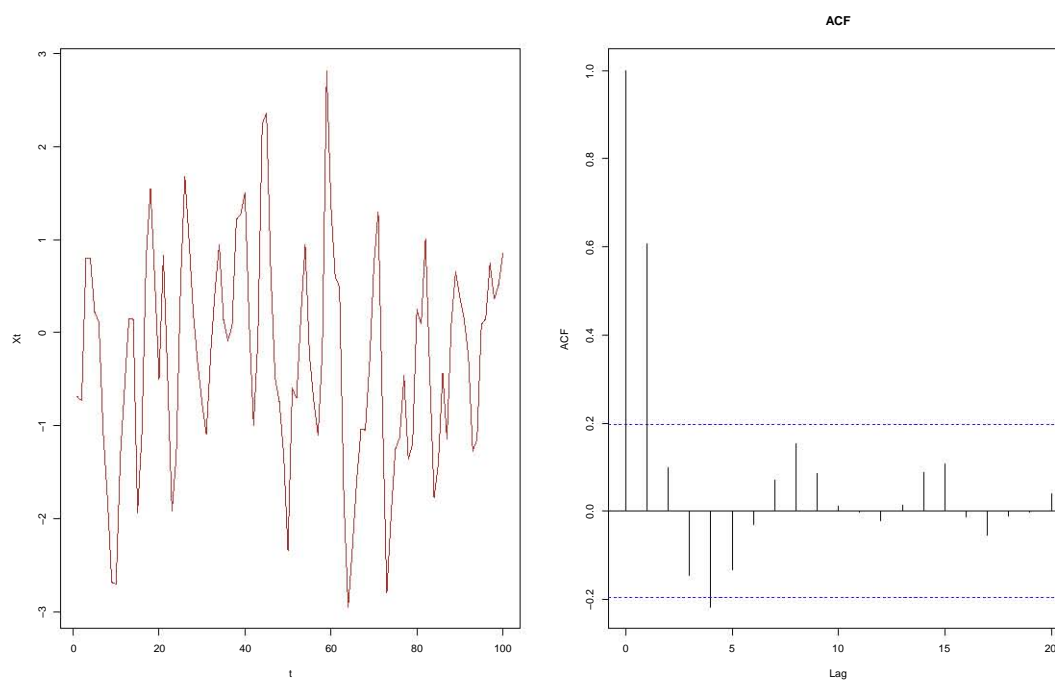


Figura 1.6: Simulación de un proceso ARMA(1,1) con $\phi_1 = 0.6$, $\theta_1 = 0.5$ y su ACF

1.5.2. Modelo ARMA(p,q)

Un proceso $\{X_t\}$ es de tipo ARMA(p,q) si $\{X_t\}$ es estacionario, se escribe de la siguiente forma

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

$\forall t \in \mathbb{Z}$, $p, q \in \mathbb{N}$, con $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q \in \mathbb{R}$ y $\{\epsilon_t\} \sim WN(0, \sigma_\epsilon^2)$, o bien, en función de los polinomios de retraso de orden p y q respectivamente $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ y $\theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)$ se tiene que

$$\phi(B) X_t = \theta(B) \epsilon_t$$

Para justificar el porqué de esta expresión, hay que recordar que

$$B X_t = X_{t-1}$$

Y de manera general

$$B^k X_t = X_{t-k}$$

Entonces se observa:

$$\begin{aligned}\phi(B) X_t &= (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t \\ &= X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p}\end{aligned}\quad (1.30)$$

De forma similar se tiene:

$$\begin{aligned}\theta(B) \epsilon_t &= (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \epsilon_t \\ &= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}\end{aligned}\quad (1.31)$$

Par que un proceso ARMA sea causal (y por lo tanto estacionario) se requiere que las raíces $\phi(z) = 0$ estén fuera del círculo unitario. De manera análoga será invertible si las raíces de $\theta(z) = 0$ también están fuera del círculo unitario.

Una propiedad deseable en un proceso ARMA(p,q) es la llamada causalidad que significa que el proceso $\{X_t\}$ depende únicamente del pasado. Esto significa que X_t puede escribirse como combinación lineal de los ϵ_t . A continuación se enunciará el teorema de causalidad e invertibilidad.

Definición 1.10 *Un proceso ARMA(p,q) es causal si existe una sucesión de pesos $\{\psi_j\}$ tales que $\sum_{j=0}^{\infty} |\psi_j| < \infty$ y $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \quad \forall t \in T$*

Lo anterior significa que un proceso MA es causal.

Teorema 1.5.1 (Causalidad) *Sea $\{X_t\}$ un proceso ARMA(p,q) para el cual los polinomios $\theta(\cdot)$ y $\phi(\cdot)$ no tiene ceros en común. Entonces $\{X_t\}$ es causal si y sólo si $\phi(z) \neq 0 \quad \forall z \in \mathbb{C}$ tal que $|z| \leq 1$. Los coeficientes $\{\psi_j\}$ en la ecuación $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, $t = 0, \pm 1, \pm 2, \pm 3, \dots$ quedan determinadas por la relación*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}$$

En palabras, el teorema dice que un proceso es causal si y sólo si las raíces del polinomio autorregresivo se encuentran fuera del círculo unitario.

Definición 1.11 Un proceso ARMA(p, q) definido por las ecuaciones $\phi(B)X_t = \theta(B)\epsilon_t$, $t \in \{0, \pm 1, \pm 2, \pm 3\}$, se llama invertible si existe una sucesión de pesos $\{\pi_j\}$ tales que $\sum_{j=0}^{\infty} |\pi_j| < \infty$ y $\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \quad \forall t \in T$

Es decir, significa despejar el error en términos del proceso. Es inmediato que un proceso AR es invertible.

Teorema 1.5.2 (Invertibilidad) Sea $\{X_t\}$ un proceso ARMA(p, q) para el cual los polinomios $\theta(\cdot)$ y $\phi(\cdot)$ no tiene ceros en común. Entonces $\{X_t\}$ es invertible si y sólo si $\theta(z) \neq 0 \quad \forall z \in \mathbb{C}$ tal que $|z| \leq 1$. Los coeficientes $\{\pi_j\}$ en la ecuación $\epsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$, $t = 0, \pm 1, \pm 2, \pm 3, \dots$ quedan determinadas por la relación

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}$$

Definición 1.12 La Función de Autocorrelación Parcial (PACF por sus siglas en inglés) mide el grado de asociación entre 2 variables sin considerar el efecto de otra variable o conjunto de variables.

$$\rho_{xy \cdot z} = \frac{\rho_{xy} - \rho_{xz}\rho_{yz}}{\sqrt{1 - \rho_{xz}^2} \sqrt{1 - \rho_{yz}^2}}$$

Formalmente, es la correlación entre los residuos R_x de la regresión de x con z y R_y la regresión de y con z .

Ejemplo: En un modelo AR(1) estacionario, el cual es $X_t = \phi X_{t-1} + \epsilon_t$, se había obtenido $\rho_\tau = \phi^\tau$ entonces

$$\begin{aligned} \rho_{X_t X_{t-2} \cdot X_{t-1}} &= \frac{\rho_{X_t X_{t-2}} - \rho_{X_t X_{t-1}} \rho_{X_{t-2} X_{t-1}}}{\sqrt{1 - \rho_{X_t X_{t-1}}^2} \sqrt{1 - \rho_{X_{t-2} X_{t-1}}^2}} \\ &= \frac{\rho_2 - \rho_1^2}{\sqrt{1 - \rho_1^2} \sqrt{1 - \rho_1^2}} \\ &= \frac{\phi^2 - \phi^2}{\sqrt{1 - \phi^2} \sqrt{1 - \phi^2}} \\ &= 0 \end{aligned} \tag{1.32}$$

Según las ecuaciones de Yule-Walker en un proceso AR(p)

$$\begin{aligned}
\rho_1 &= \phi_1 + \cdots + \phi_p \rho_{p-1} \\
\rho_2 &= \phi_1 \rho_1 + \cdots + \phi_p \rho_{p-2} \\
&\vdots \\
\rho_p &= \phi_1 \rho_{p-1} + \cdots + \phi_p
\end{aligned} \tag{1.33}$$

así que, en un AR(2) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$ la autocorrelación parcial a dos pasos es

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2 \rho_1 \\
\rho_2 &= \phi_1 \rho_1 + \phi_2
\end{aligned} \tag{1.34}$$

entonces, resolviendo el sistema de ecuaciones, se tiene

$$\begin{aligned}
\rho_1 &= \frac{\phi_1}{1 - \phi_2} \\
\rho_2 &= \frac{\phi_1^2}{1 - \phi_2} + \phi_2
\end{aligned} \tag{1.35}$$

de donde

$$\begin{aligned}
\rho_{X_t X_{t-2} \cdot X_{t-1}} &= \frac{\rho_2 - \rho_1^2}{\sqrt{1 - \rho_1^2} \sqrt{1 - \rho_1^2}} \\
&= \frac{\frac{\phi_1^2}{1 - \phi_2} + \phi_2 - \left(\frac{\phi_1}{1 - \phi_2}\right)^2}{1 - \frac{\phi_1^2}{(1 - \phi_2)^2}} \\
&= \frac{\phi_1^2 \left[\frac{1}{1 - \phi_2} - \frac{1}{(1 - \phi_1)^2} \right] + \phi_2}{1 - \frac{\phi_1^2}{(1 - \phi_2)^2}} \\
&= \frac{\left[\phi_1^2 \left[\frac{1 - \phi_2 - 1}{(1 - \phi_2)^2} \right] + \phi_2 \right]}{\left(1 - \frac{\phi_1^2}{(1 - \phi_2)^2} \right)} \\
&= \frac{\phi_2 \left(1 - \frac{\phi_1^2}{(1 - \phi_2)^2} \right)}{\left(1 - \frac{\phi_1^2}{(1 - \phi_2)^2} \right)} \\
&= \phi_2
\end{aligned} \tag{1.36}$$

en un modelo AR(2), la autocorrelación parcial a dos pasos no es cero.

La autocorrelación parcial puede generalizarse para cuando hay un retraso o 'lag' cualquiera usando la definición que se basa en regresiones.

Para un proceso $X_1, \dots, X_k, X_{k+1}, \dots$ consideramos para $k \in \{1, 2, \dots\}$ la regresión de X_{k+1} y X_1 con respecto a los valores intermedios

$$\begin{aligned}
X_1 &= \alpha_1 + \alpha_2 X_2 + \dots + \alpha_k X_k \\
X_{k+1} &= \beta_1 + \beta_2 X_2 + \dots + \beta_k X_k
\end{aligned} \tag{1.37}$$

Estimando por mínimos cuadrados

$$\begin{aligned}
R_1 &= X_1 - (\hat{\alpha}_1 + \hat{\alpha}_2 X_2 + \dots + \hat{\alpha}_k X_k) \\
R_{k+1} &= X_{k+1} - (\hat{\beta}_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_k X_k)
\end{aligned} \tag{1.38}$$

De donde la autocorrelación parcial a k pasos es $PACF(k) = Corr(R_{k+1}, R_1)$ y $PACF(1) = Corr(X_1, X_2)$ mide la asociación entre X_1 y X_{k+1} sin el efecto de X_2, \dots, X_k

Identificación de los parámetros p y q

Cuadro 1.1: Comportamiento de la ACF y PACF para los modelos AR y MA

	AR(p)	MA(q)
ACF	Disminuye lentamente	Se trunca después del retraso q
PACF	Se trunca después del retraso p	Disminuye lentamente

Cuadro 1.2: Comportamiento de la ACF y PACF para los modelos ARMA

ARMA(p,q)	
ACF	Disminuye lentamente después del retraso $(q-p)$
PACF	Disminuye lentamente después del retraso $(p-q)$

Las funciones de autocorrelación y autocorrelación parcial permiten identificar si las observaciones de una serie de tiempo tienen el comportamiento de un modelo AR(p), MA(q) o ARMA.

Como se muestra en los Cuadros 1.1 y 1.2 las funciones ACF y PACF para un proceso ARMA no se truncan sino que van disminuyendo. Esto se debe a que se supone al proceso como causal e invertible, entonces la parte autorregresiva tiene una representación media móvil infinita por lo que la función de autocorrelación no se trunca, del mismo modo la parte media móvil tiene una representación autorregresiva infinita y por tanto la función de autocorrelación parcial no se trunca.

1.6. Modelos ARIMA

Los modelos autorregresivos e integrados de promedios móviles (ARIMA) se pueden ver como una generalización de los modelos ARMA. Un proceso X_t es un proceso ARIMA si $(1 - B)^d X_t$ es un proceso ARMA.

Se dice que $\{X_t\}$ es un proceso ARIMA(p,d,q) si se satisface:

$$\phi(B)(1 - B)^d X_t = \theta(B)\epsilon_t$$

$$\forall t \in \mathbb{Z}, p, d, q \in \mathbb{N}$$

El uso del operador $(1 - B)^d$ surge de la posibilidad de que ciertas series de tiempo no muestran estacionariedad. Lo que se hace al aplicar este operador, es

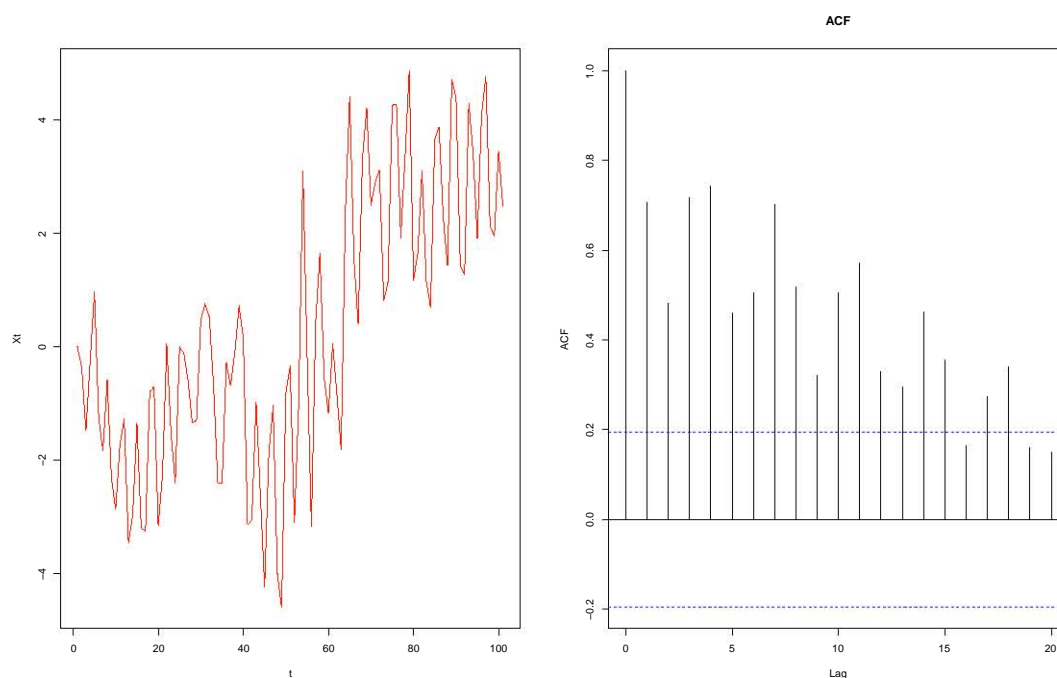


Figura 1.7: Simulación proceso ARIMA(2,1,2) con $\phi_1 = -0.3$, $\phi_2 = -0.8$, $\theta_1 = 0.2$, $\theta_2 = -0.1$ y su ACF

eliminar una posible tendencia polinomial de orden d que se presente en la serie de datos, de esta forma, la serie se transformará en un modelo ARMA y se podrá trabajar de mejor manera.

El siguiente código genera una simulación de 100 datos de un ARIMA(2, 1, 2) con $\phi_1 = -0.3$, $\phi_2 = -0.8$, $\theta_1 = 0.2$ y $\theta_2 = -0.1$. El proceso graficado es la Figura 1.7.

```
ts.sim <- arima.sim(list(order = c(2,1,2), ar = c(-0.3,-0.8),
  ma = c(0.2,-0.1)), n = 100, sd = sqrt(1))
ts.plot(ts.sim, col="red",xlab="t",ylab="Xt")
acf(ts.sim,main="ACF")
```

Construcción del modelo

El proceso para construir un modelo ARIMA consta de cuatro pasos principalmente. Primero, se identificará un modelo ARIMA haciendo uso de los datos que se lleven a análisis. Segundo, se estimarán los parámetros desconocidos. Tercero, se verificarán los supuestos para determinar si el modelo es apropiado. Cuarto y último, se hará un pronóstico para darse una idea de lo que podría venir a futuro.

1.6.1. Identificación del Modelo

Una de las primeras actividades que se hacen, es graficar los datos para darse una idea de su comportamiento, posteriormente será necesario aplicar el ACF muestral y PACF muestral a los datos históricos, esto con el fin de identificar un modelo de promedios móviles, autorregresivo ó ARMA. Para encontrar la parte de promedios móviles es necesario usar el ACF muestral, ver los intervalos de confianza y verificar la última línea que se sale, es en ese punto la elección de la q ; para la parte autorregresiva la identificación del modelo a partir del ACF es complicado. En general, el orden un proceso AR sólo se puede conseguir con la función de autocorrelación parcial (PACF), se hace un procedimiento similar al de MA, se observa el PACF y se ve la última línea que se sale de los intervalos y esa será la p que se usará. Si los datos no llegan a presentar estacionariedad, será necesario aplicar el operador diferencia a los datos (en la modelación estadística por lo regular se aplican a lo mucho dos diferencias, ya que si llegan a ser más, se corre el riesgo de perder mucha información) y posteriormente aplicar nuevamente el ACF y PACF muestral, en este caso se estará en presencia de un modelo ARIMA.

1.6.2. Estimación de los parámetros

Los modelos que se han analizado son modelos paramétricos, es decir, dependen de parámetros. Por lo tanto, se deben estimar los parámetros, esto es lo que se conoce como estimación del modelo.

Existen varios métodos de estimación, entre los más populares se encuentran el método de momentos, máxima verosimilitud o por mínimos cuadrados. En este caso se usará el método de máxima verosimilitud.

Método de máxima verosimilitud

Para este método se parte del supuesto de que $\{\epsilon_t\}$ es un proceso de ruido blanco con distribución normal, media cero y varianza σ_ϵ^2 . Ahora, consideremos un modelo ARMA(p,q) con p, q conocidas

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

Y suponiendo que se cuenta con la información del proceso

$$(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

Donde X_i es la variable aleatoria que denota el valor de la serie a tiempo i y x_i el valor que tomó.

Consideramos los siguientes vectores con el objetivo de hacer más sencilla la notación: $\Theta_q = (\theta_1, \dots, \theta_q)$ y $\Phi_p = (\phi_1, \dots, \phi_p)$. Entonces, se puede verificar en Brockwell and Davis (2016) que los estimadores máximo verosímiles son los siguientes:

$$\hat{\sigma}^2 = n^{-1}S(\hat{\Theta}_q, \hat{\Phi}_p)$$

donde $S(\hat{\Theta}_q, \hat{\Phi}_p) = \sum_{j=1}^n \frac{(X_j - \hat{X}_j)^2}{r_{j-1}}$; $\hat{\Theta}_q, \hat{\Phi}_p$ son los valores de Θ_q, Φ_p que minimizan la expresión

$$l(\Theta_q, \Phi_p) = \ln(n^{-1}S(\Theta_q, \Phi_p)) + n^{-1} \sum_{j=1}^n \ln(r_{j-1})$$

$$\text{y } r_n = \mathbb{E}(\epsilon_{n+1} - \hat{\epsilon}_{n+1})^2.$$

En algunos casos se suele usar el estimador insesgado $\hat{\sigma}^2 = \frac{S(\hat{\Theta}_q, \hat{\Phi}_p)}{n-p-q-1}$.

Estos procesos de minimización se tienen que hacer mediante cálculos numéricos. Hay diversos programas estadísticos que cumplen con esta tarea, entre ellos encontramos a: R e ITSM, en el caso de este capítulo nos enfocaremos en R.

1.6.3. Verificación de supuestos.

Después de que un modelo tentativo ha sido ajustado a los datos, se debe verificar si el modelo es adecuado, esto se logra a través del análisis de los residuos. Los residuos para un proceso ARMA(p,q) pueden obtenerse de

$$\hat{\epsilon}_t = X_t - \left(\sum_{i=1}^p \hat{\phi}_i X_{t-i} + \sum_{i=1}^q \hat{\theta}_i \hat{\epsilon}_{t-i} \right)$$

Si el modelo es adecuado, entonces los residuos $\hat{\epsilon}_t$ deberán comportarse como un proceso de ruido blanco, ya que ésta es nuestra hipótesis, esto se logra revisando que los residuales estén no correlacionados (independencia implica no correlación), que tengan varianza constante y que se distribuyan normal.

Para ver que los residuos están no correlacionados, entonces deberán estar dentro de las bandas de confianza de la función de autocorrelación muestral (ACF), en el caso de las autocorrelaciones parciales (PACF) también deben estar dentro de las bandas, de lo contrario aún falta modelar algo. Ahora, en vez de considerar las autocorrelaciones de los residuales individualmente, también se podrían considerar en conjunto, para esta tarea se tendrá que recurrir a la prueba de Ljung-Box.

Prueba de Ljung-Box.

Esta prueba estadística ayuda para saber si un grupo (k) de autocorrelaciones son iguales a cero, esto con el fin de saber si los errores no están correlacionados. La prueba de hipótesis es la siguiente

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0 \quad vs \quad H_1 : \text{al menos una } \rho_i \neq 0 \quad i = \overline{1, k}$$

El estadístico asociado es $Q(k) = n(n+1) \sum_{j=1}^k \frac{\hat{\rho}_j^2}{n-j}$.

donde:

- n = tamaño muestral.
- $\hat{\rho}_j$ = estimador muestral de la autocorrelación muestral.
- k = número de retrasos en H_0

$Q(k)$ se distribuye asintóticamente $\chi^2_{(k-p-q)}$ así que se rechaza H_0 a un nivel α si $Q(k) > \chi^2_{(k-p-q)}^{2(1-\alpha)}$.

En caso de que los residuos estén correlacionados, buscar otro candidato de modelo.

Para el caso de normalidad en los residuos, se podrá investigar ésta situación usando pruebas como: Kolmogorov-Smirnov, Anderson-Darling; por medio de histogramas y ver si hay normalidad o finalmente por un gráfico qq-plot. Afortunadamente el software R tiene todas estas pruebas.

En caso de que no se cumpla con este supuesto, se podría aplicar una transformación en los datos (Por ejemplo: $\frac{1}{X_t}, \log X_t$, entre otras).

Finalmente, para verificar que la varianza sea constante, es necesario hacer una gráfica de los residuos contra el tiempo y ver si la varianza parece cumplir con ese supuesto. En caso de que no lo sea, aplicar algún tipo de transformación.

1.6.4. Pronóstico

Una vez que el modelo ha sido ajustado y revisado de forma exhaustiva, se pueden pronosticar valores futuros. Si se denota el tiempo actual por T , entonces $X_{T+\tau}$ es el pronóstico τ periodos y se expresa como $\hat{X}_{T+\tau}^{(T)}$. El mejor criterio para obtener pronósticos deseables es por medio del error cuadrático medio, para lo cual el valor esperado de los errores de pronóstico al cuadrado, $\mathbb{E} \left[\left(X_{T+\tau} - \hat{X}_{T+\tau}^{(T)} \right)^2 \right] = \mathbb{E} \left(e_T^{(\tau)2} \right)$, se minimiza. Se puede demostrar que el mejor pronóstico en el sentido de media cuadrática, es la esperanza condicional de $X_{T+\tau}$ dada observaciones actuales y pasadas, es decir, X_T, X_{T-1}, \dots :

$$\hat{X}_{T+\tau}^{(T)} = \mathbb{E} [X_{T+\tau} \mid X_T, X_{T-1}, \dots] \quad (1.39)$$

Por ejemplo, considérese un proceso ARMA(p,q) estacionario a tiempo $T + \tau$, es decir, τ periodos en el futuro:

$$\phi(B) X_{T+\tau} = \theta(B) \epsilon_{T+\tau} \quad (1.40)$$

ya que el modelo es estacionario, se puede reescribir como una representación de promedio móviles, donde $\psi(B) = \sum_{i=0}^{\infty} \psi_i B^i = \frac{\theta(B)}{\phi(B)}$, entonces se tiene que:

$$X_{T+\tau} = \sum_{i=0}^{\infty} \psi_i \epsilon_{T+\tau-i} \quad (1.41)$$

Se puede partir la ecuación 1.41 en dos de la siguiente manera

$$X_{T+\tau} = \sum_{i=0}^{\tau-1} \psi_i \epsilon_{T+\tau-i} + \sum_{i=\tau}^{\infty} \psi_i \epsilon_{T+\tau-i} \quad (1.42)$$

En esta partición, se puede ver que la componente $\sum_{i=0}^{\tau-1} \psi_i \epsilon_{T+\tau-i}$ involucra los errores futuros, mientras que la componente $\sum_{i=\tau}^{\infty} \psi_i \epsilon_{T+\tau-i}$ involucra errores presentes y pasados. De la ecuación entre valores pasados y futuros y los correspondientes choques aleatorios, así como también el hecho de que los choques aleatorios tienen media cero y son independientes, se puede mostrar que el mejor pronóstico es

$$\begin{aligned} \hat{X}_{T+\tau}^{(T)} &= \mathbb{E} [X_{T+\tau} \mid X_T, X_{T-1}, \dots] \\ &= \sum_{i=\tau}^{\infty} \psi_i \epsilon_{T+\tau-i} \end{aligned} \quad (1.43)$$

donde

$$\mathbb{E} [\epsilon_{T+\tau-i} \mid X_T, X_{T-1}, \dots] = \begin{cases} 0 & \text{si } i < \tau, \\ \epsilon_{T+\tau-i} & \text{si } i \geq \tau. \end{cases}$$

El error de predicción se calcula como

$$e_T^{(\tau)} = X_{T+\tau} - \hat{X}_{T+\tau}^{(T)} = \sum_{i=0}^{\tau-1} \psi_i \epsilon_{T+\tau-i} \quad (1.44)$$

ya que el error de predicción en la ecuación 1.44 es una combinación de choques aleatorios, se tiene

$$\mathbb{E} (e_T^{(\tau)}) = 0 \quad (1.45)$$

$$\begin{aligned}
\text{Var} \left(e_T^{(\tau)} \right) &= \text{Var} \left[\sum_{i=0}^{\tau-1} \psi_i \epsilon_{T+\tau-i} \right] \\
&= \sum_{i=0}^{\tau-1} \psi_i^2 \text{Var} \left(\epsilon_{T+\tau-i} \right) \\
&= \sigma^2 \sum_{i=0}^{\tau-1} \psi_i^2 \\
&= \sigma^2(\tau), \quad \tau = 1, 2, \dots
\end{aligned} \tag{1.46}$$

Donde $\sigma^2(\tau)$ es simplemente la notación de la varianza del error de predicción. Por otro lado, es evidente que la varianza del error de predicción llega a ser enorme conforme τ es más grande, esto es notorio ya que se espera que haya más incertidumbre al querer pronosticar más valores futuros. Además, si los choques aleatorios se asumen normales, $N(0, \sigma^2)$, entonces los errores de predicción también serán normales con $N(0, \sigma^2(\tau))$. Se puede obtener el intervalo de predicción a un nivel de confianza del $100 \times (1 - \alpha)$ por ciento para las observaciones futuras de la siguiente forma

$$\mathbb{P} \left(\hat{X}_{T+\tau}^{(T)} - \mathbf{z}_{\frac{\alpha}{2}} \sigma(\tau) < X_{T+\tau} < \hat{X}_{T+\tau}^{(T)} + \mathbf{z}_{\frac{\alpha}{2}} \sigma(\tau) \right) = 1 - \alpha \tag{1.47}$$

donde \mathbf{z} es el percentil $\frac{\alpha}{2}$ de la distribución normal estándar. Por lo tanto, el intervalo de predicción a un nivel de confianza del $100 \times (1 - \alpha)$ por ciento para $X_{T+\tau}$ es

$$\hat{X}_{T+\tau}^{(T)} \pm \mathbf{z}_{\frac{\alpha}{2}} \sigma(\tau) \tag{1.48}$$

Como observación final, en caso de que se cuentan con más de un modelo, ¿Cómo elegir el mejor? En este caso se pueden usar dos criterios: el Criterio de Información Akaike y el Criterio de Información Bayesiano.

Este par de criterios penalizan la suma de residuos al cuadrado al incluir parámetros adicionales en el modelo. Los modelos que tienen pequeños valores de AIC o BIC son considerados como buenos modelos.

Los criterios son los siguientes:

$$AIC = \ln \left(\frac{\sum_{t=1}^T e_t^2}{T} \right) + \frac{2p}{T} \quad (1.49)$$

$$BIC = \ln \left(\frac{\sum_{t=1}^T e_t^2}{T} \right) + \frac{p \ln(T)}{T} \quad (1.50)$$

1.7. Implementación en R

Para la implementación en R, se usarán los precios históricos mensuales por barril de la Mezcla Mexicana (MME) en dólares. Los datos fueron obtenidos del portal del SIE (Sistema de Información Energética). Los precios fueron recabados desde el mes de Enero de 2005 al mes de Octubre de 2016.

Los datos se muestran en el Cuadro 1.3

AÑO	E	F	M	A	M	JN	JL	A	S	O	N	D
2005	31.7	33.7	38.4	39.9	40.1	44.7	46.6	50.3	50.7	48	43.8	45.1
2006	50.2	48.8	51.4	57.1	56.8	55.7	59.4	60.2	51.3	48.5	47.6	49.8
2007	44.4	48.4	50.5	54.6	56.2	60	64.5	63	67.5	72.1	79.8	80.2
2008	81.4	89.4	94.9	104.2	114.2	120.3	106.6	85.6	60.3	42.4	33.7	38
2009	38.2	42	47.8	56.4	64.4	61	67.2	64.7	68.9	72.5	69.7	72.2
2010	70.2	72.1	74.4	67.8	67.4	68.5	69.8	70.3	74.7	77.5	82.2	85.9
2011	90.5	102.4	109.5	104.4	103	104.8	98.7	100.1	101.1	107.6	106.3	108.5
2012	110.2	112.8	108	102.2	91.4	95.4	101.5	102.1	99.1	95.4	96.7	100.6
2013	105.4	103	99.1	98.7	97.9	101	100.8	99.7	95	89.5	91.7	90.7
2014	93.1	93.5	95.7	96.8	98.8	94.6	90.8	85.8	75.2	71.4	52.4	41.7
2015	47.3	47.4	50.7	54.1	53.9	46.6	39.7	37.9	37.5	35.4	28.7	23.9
2016	24.5	29.4	32.3	27.3	40.1	38.7	38.4	37.8	40.5	38.2		

Cuadro 1.3: Precios Históricos mensuales de la Mezcla Mexicana

Se cargan las paqueterías y se grafican los datos.

```
library(forecast)
library(lmtest)
library(stats)
library(LSTS)
setwd("C:/Users/Jesus Pichardo Leyva/Documents/TESIS")
datos<-read.table("MME.csv",sep="," ,header=FALSE)
MMEt<-ts(datos, frequency = 12, start = c(2005, 1))
```

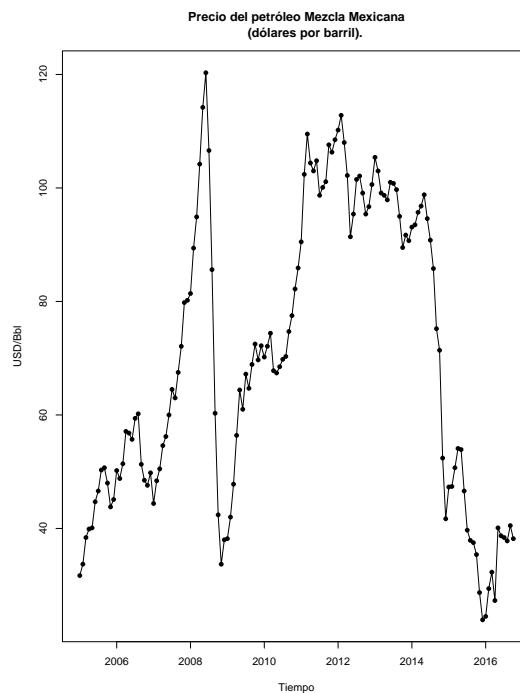


Figura 1.8: Precio del petróleo Mezcla Mexicana

```
plot(MMEt, type="o", pch=16, cex=0.5, xlab='Tiempo',
      ylab='USD/Bbl', main="Precio del petroleo Mezcla Mexicana\
      (dolares por barril).")
```

En la Figura 1.8 se muestran los datos graficados, donde se ven bajadas y subidas muy pronunciadas. Se puede observar que en 2009 hubo una bajada del precio del crudo mexicano debido a la Crisis Mundial. Después, hubo grandes ingresos por la venta de petróleo entre los años 2011 y 2014, posteriormente vuelve a bajar el precio del barril debido a la desaceleración de la economía de China (véase Egan (2016)).

Ahora, se calcularán los ACF y PACF muestrales para encontrar el mejor modelo que se ajuste a los datos

```
par(mfrow=c(1,2), oma=c(0,0,0,0))
acf(ts(MMEt,freq=1), lag.max=20, main="ACF MME")
pacf(ts(MMEt,freq=1), lag.max=20, main="PACF MME")
```

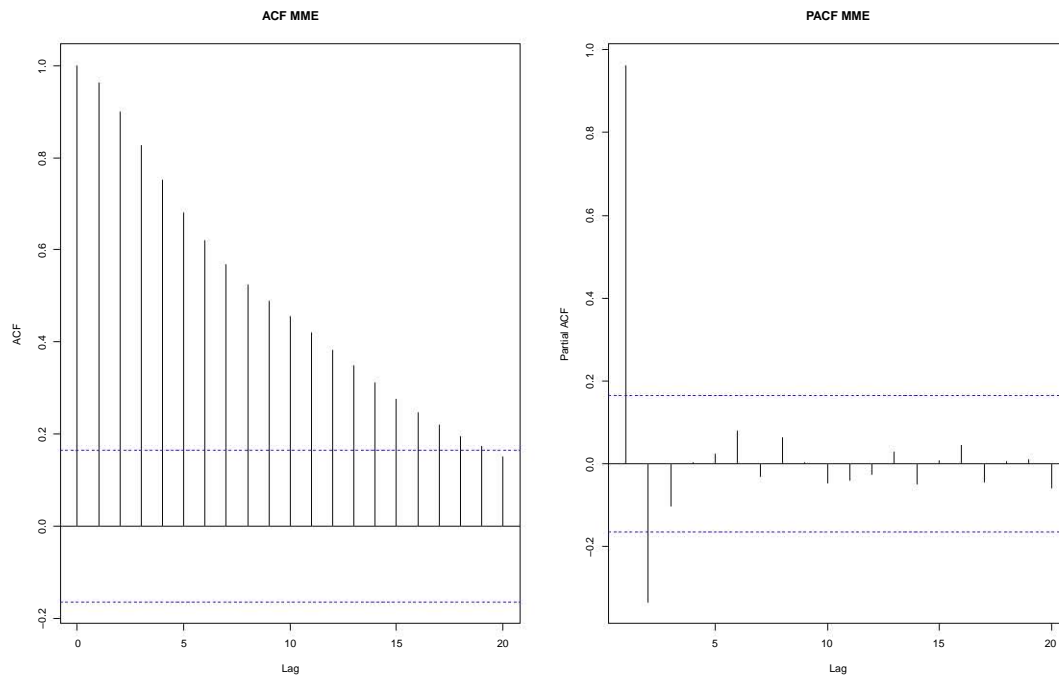


Figura 1.9: ACF y PACF MME

Se puede observar que el ACF (Figura 1.9) tiene un comportamiento decreciente, por otro lado el PACF se *corta* en 2, entonces en este caso se recurrirá a diferenciar la serie (ya que la serie no es estacionaria) y se volverán a calcular los ACF y PACF.

El código y los ACF y PACF con los datos diferenciados son:

```
acf(ts(diff(MMEt),freq=1), lag.max=20, main="ACF MME\
Primera Diferencia")
pacf(ts(diff(MMEt),freq=1), lag.max=20, main="PACF MME\
Primera Diferencia")
```

El PACF (Figura 1.10) sugiere un modelo AR(1), en el caso del ACF marca un MA(2). Para evitar elegir un modelo inadecuado se ajustarán tres modelos: ARIMA(1,1,0), ARIMA(1,1,1) y ARIMA(1,1,2); se elegirá el mejor de acuerdo a los criterios AIC y BIC.

```
fit1<-stats::arima(MMEt, order=c(1,1,0))
fit1
```

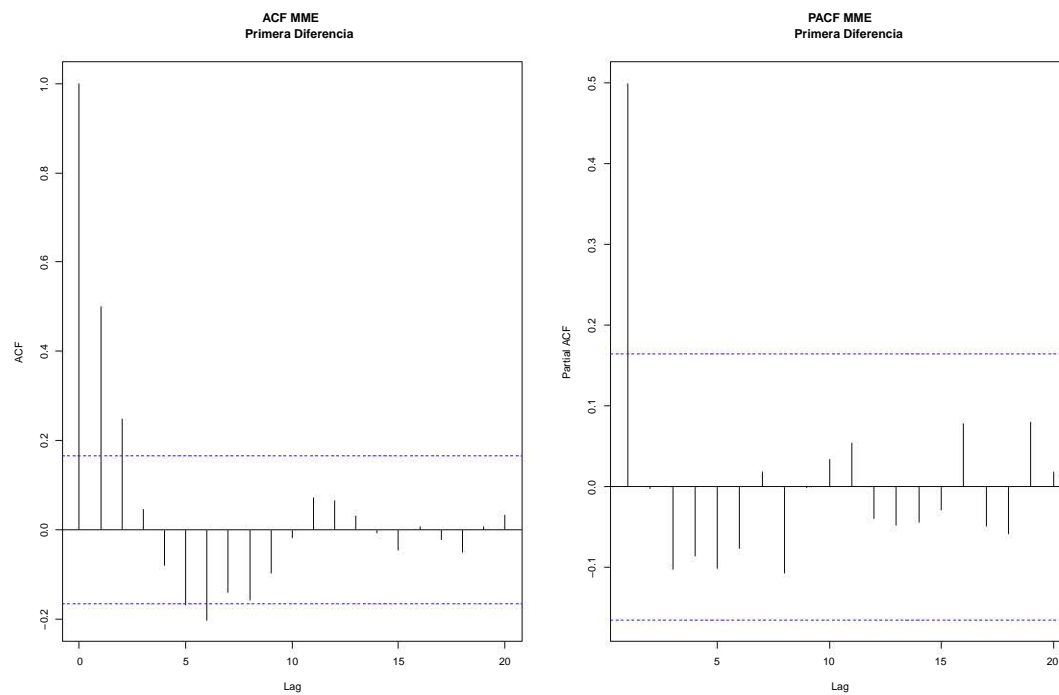


Figura 1.10: ACF y PACF MME Primera Diferencia

Call:

```
stats::arima(x = MMEt, order = c(1, 1, 0))
```

Coefficients:

```
      ar1
      0.4970
s.e.  0.0725
```

sigma² estimated as 25.29: log likelihood = -427.95, aic = 859.89

```
AIC(fit1)
```

```
[1] 859.8942
```

```
BIC(fit1)
```

```
[1] 865.7917
```

```
fit2<-stats::arima(MMEt, order=c(1,1,1))
```

```
fit2
```

Call:

```
stats::arima(x = MMEt, order = c(1, 1, 1))
```

Coefficients:

	ar1	ma1
	0.4941	0.0038
s.e.	0.1294	0.1427

sigma^2 estimated as 25.29: log likelihood = -427.95, aic = 861.89

AIC(fit2)

[1] 861.8935

BIC(fit2)

[1] 870.7397

```
fit3<-stats::arima(MMEt, order=c(1,1,2))
```

fit3

Call:

```
stats::arima(x = MMEt, order = c(1, 1, 2))
```

Coefficients:

	ar1	ma1	ma2
	0.3438	0.1495	0.1185
s.e.	0.2488	0.2521	0.1267

sigma^2 estimated as 25.13: log likelihood = -427.52, aic = 863.04

AIC(fit3)

[1] 863.0384

BIC(fit3)

[1] 874.8335

Tomando en cuenta los criterios AIC y BIC se tomará el modelo ARIMA(1,1,0) como el mejor modelo. El modelo ajustado es de la siguiente forma:

$$MME_t(1 - 0.4970B)(1 - B) = \epsilon_t$$

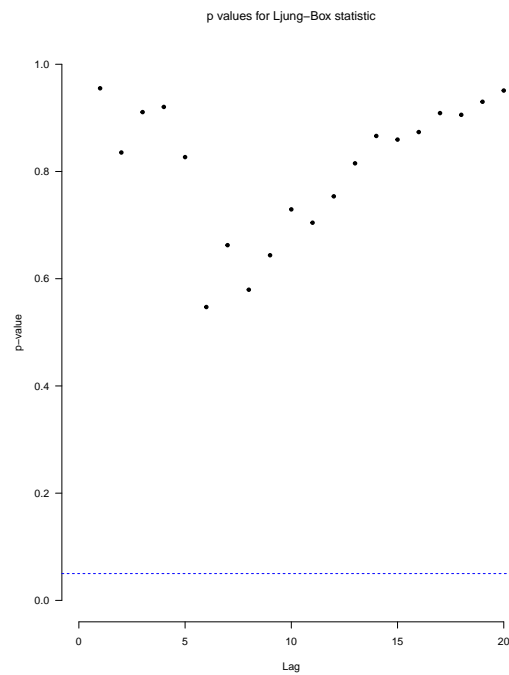


Figura 1.11: p-values de la prueba de Ljung-Box

Donde MME_t son los datos de la Mezcla Mexicana.

Se calcularán los valores ajustados y los residuos. Posteriormente, se verificarán los supuestos, sólo para comprobar si el modelo elegido es adecuado.

```
residuales<-as.vector(residuals(fit1))
ajustados<-as.vector(fitted(fit1))
```

Primero, se revisará que los residuos estén no correlacionados, para ello se calcularán la prueba de Ljung-Box y los ACF y PACF para los residuos.

```
Box.Ljung.Test(residuales, lag = 20, main = NULL)
Box.test(residuales, lag=20, fitdf=1, type="Ljung")
Box-Ljung test
```

```
data: residuales
X-squared = 10.811, df = 19, p-value = 0.93
```

En ambas pruebas de Ljung-Box el p-value es mayor a 0.05, entonces no se rechaza que todas las autocorrelaciones son iguales a 0. Por lo tanto el primer

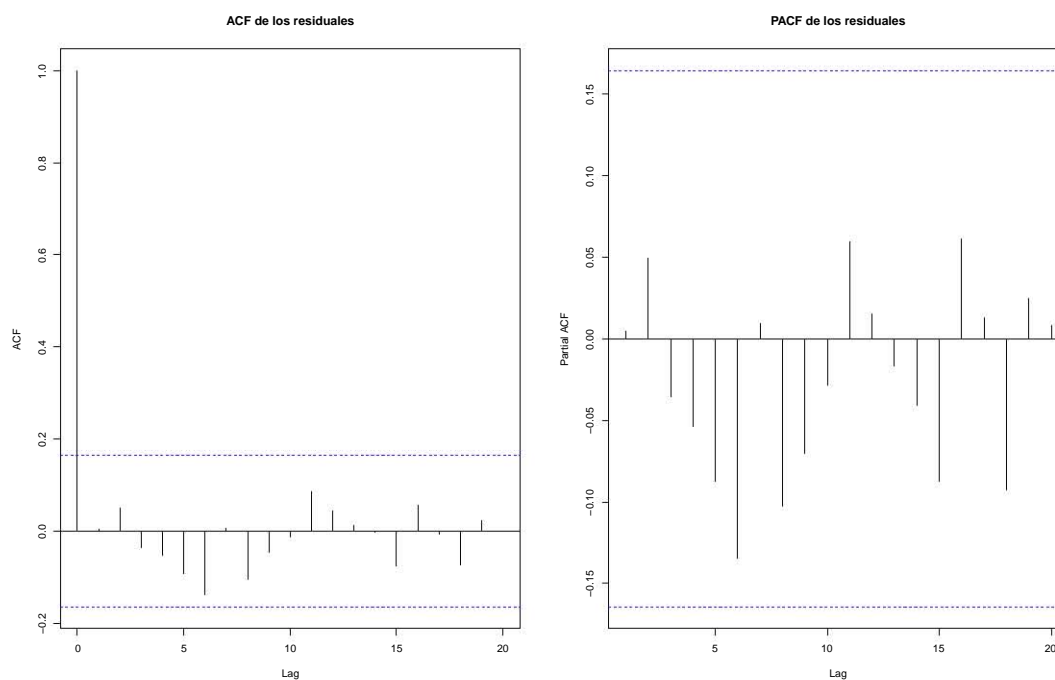


Figura 1.12: ACF y PACF de los residuos

supuesto se cumple. Ahora, se calculará el ACF y PACF de los residuos para buscar el mismo objetivo que con Ljung-Box, que es encontrar no autocorrelación en los residuos.

```
acf(ts(residuales,freq=1), lag.max=20, main="ACF de los residuales")
pacf(ts(residuales,freq=1), lag.max=20, main="PACF de los residuales")
```

Los ACF y PACF de los residuos (Figura 1.12) están por debajo de los intervalos, entonces se puede afirmar que no hay autocorrelación en los residuos. Ahora, se comprobará si los residuales tienen distribución normal, para ello, se calculará el q-q plot, el histograma y principalmente las pruebas de Lilliefors y Shapiro-Wilks.

Primero, el q-q plot y el histograma

```
qqnorm(residuales,datax = T, pch=16, xlab = 'Residuales')
qqline(residuales, datax = T)
hist(residuales,col="gray",xlab='Residual',main='')
```

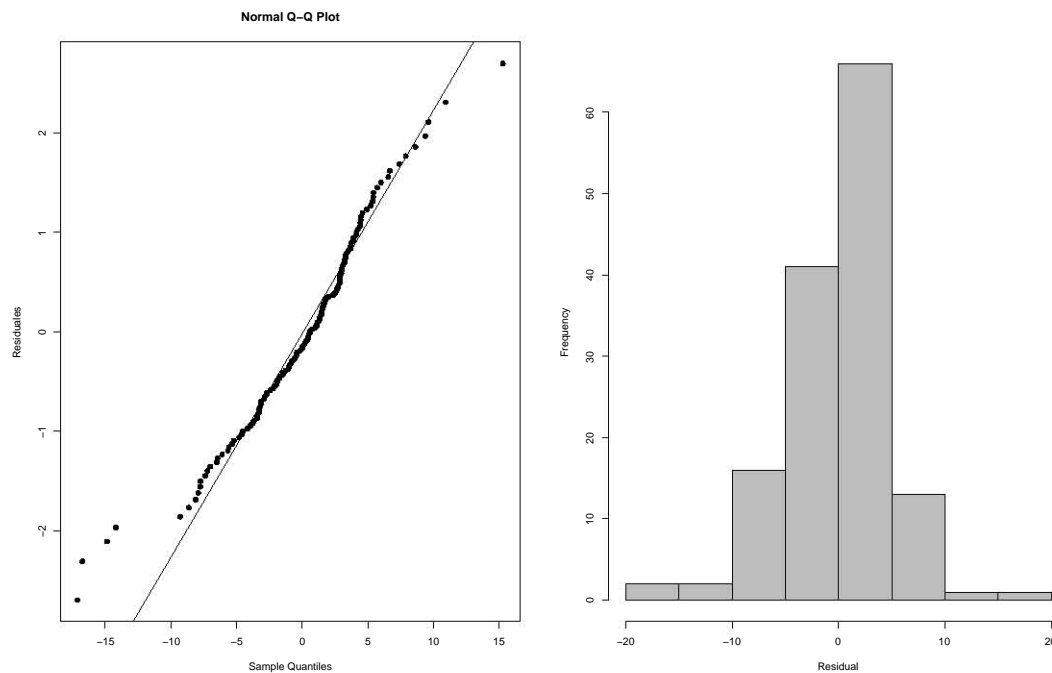


Figura 1.13: Q-Q plot e histograma

Cualitativamente se ve que hay cierto comportamiento normal. En el q-q plot (Figura 1.13) se puede observar que hay colas pesadas y eso podría afectar a que no se cumpla con la hipótesis de normalidad; en el histograma se puede observar la misma situación que con el q-q plot. Sin embargo, los dos gráficos dan certeza de cumplir con el supuesto de normalidad. Para tratar de resolver este problema, se usará la prueba de Lilliefors y Shapiro-Wilk, estas pruebas ayudarán a saber si los residuos son gaussianos con mayor precisión.

```
lillie.test(residuales)
```

```
Lilliefors (Kolmogorov-Smirnov) normality test
```

```
data: residuales
```

```
D = 0.069883, p-value = 0.08623
```

```
shapiro.test(residuales)
```

```
Shapiro-Wilk normality test
```

```
data: residuales
```

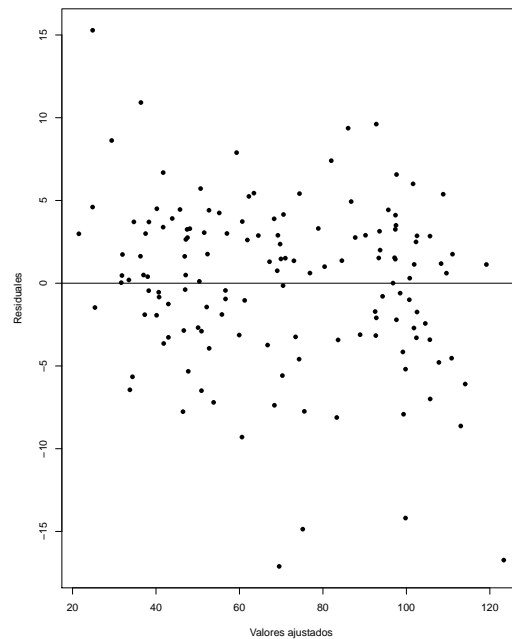


Figura 1.14: Gráfico de valores ajustados contra residuos

$W = 0.96136$, $p\text{-value} = 0.0004971$

En el caso de la prueba de Lilliefors, los residuos son normales, ya que el $p\text{-value}$ es mayor a 0.05. Por otro lado, el $p\text{-value}$ con la prueba de Shapiro-Wilk es menor a 0.05, es decir, se rechaza que los residuos sean normales. Por último, se verá si los residuos tienen varianza constante, para ello se utilizará el gráfico de valores ajustados contra residuos.

```
plot(ajustados,residuales,pch=16, xlab='Valores ajustados',
     ylab='Residuales')
abline(h=0)
```

El gráfico (Figura 1.14) no muestra una forma curva o en abanico; sin embargo, los residuos se ven dispersos. Por lo tanto, se puede afirmar que no hay varianza constante y también se observa que hay datos atípicos. Por ahora no se profundizará en el tema, ya que esto se verá con más atención en los capítulos 3 y 4. Finalmente, se hará el pronóstico a un año.

```
forecast<-as.array(forecast(fit1,h=12))
```

```
forecast
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Nov 2016	37.05683	30.612423	43.50124	27.2009606	46.91270
Dec 2016	36.48864	24.886738	48.09054	18.7450625	54.23222
Jan 2017	36.20623	20.052892	52.35957	11.5018311	60.91063
Feb 2017	36.06587	15.924615	56.20712	5.2624807	66.86925
Mar 2017	35.99610	12.331199	59.66100	-0.1962422	72.18844
Apr 2017	35.96142	9.142018	62.78083	-5.0553174	76.97817
May 2017	35.94419	6.262627	65.62575	-9.4498438	81.33822
Jun 2017	35.93562	3.625557	68.24569	-13.4783583	85.34960
Jul 2017	35.93137	1.182281	70.68045	-17.2127732	89.07550
Aug 2017	35.92925	-1.102665	72.96116	-20.7061764	92.56467
Sep 2017	35.92820	-3.255706	75.11210	-23.9984122	95.85481
Oct 2017	35.92767	-5.296947	77.15230	-27.1199440	98.97529

Se puede ver que los pronósticos se mueven entre 35 y 37 dólares y tienen cierto grado de validez, ya que los datos originales han mostrado que el precio del barril de petróleo nacional se mueve entre los 35 dólares y 45 dólares. Por otro lado, los intervalos de confianza son muy grandes debido en gran parte a la varianza no constante de los datos y a los potenciales outliers que puede llegar a tener ésta serie de tiempo.

El gráfico 1.15 corresponde al pronóstico

```
plot(forecast,main='Pronstico Mezcla \
      Mexicana', xlab='Mes',ylab='USD/Bbl')
```

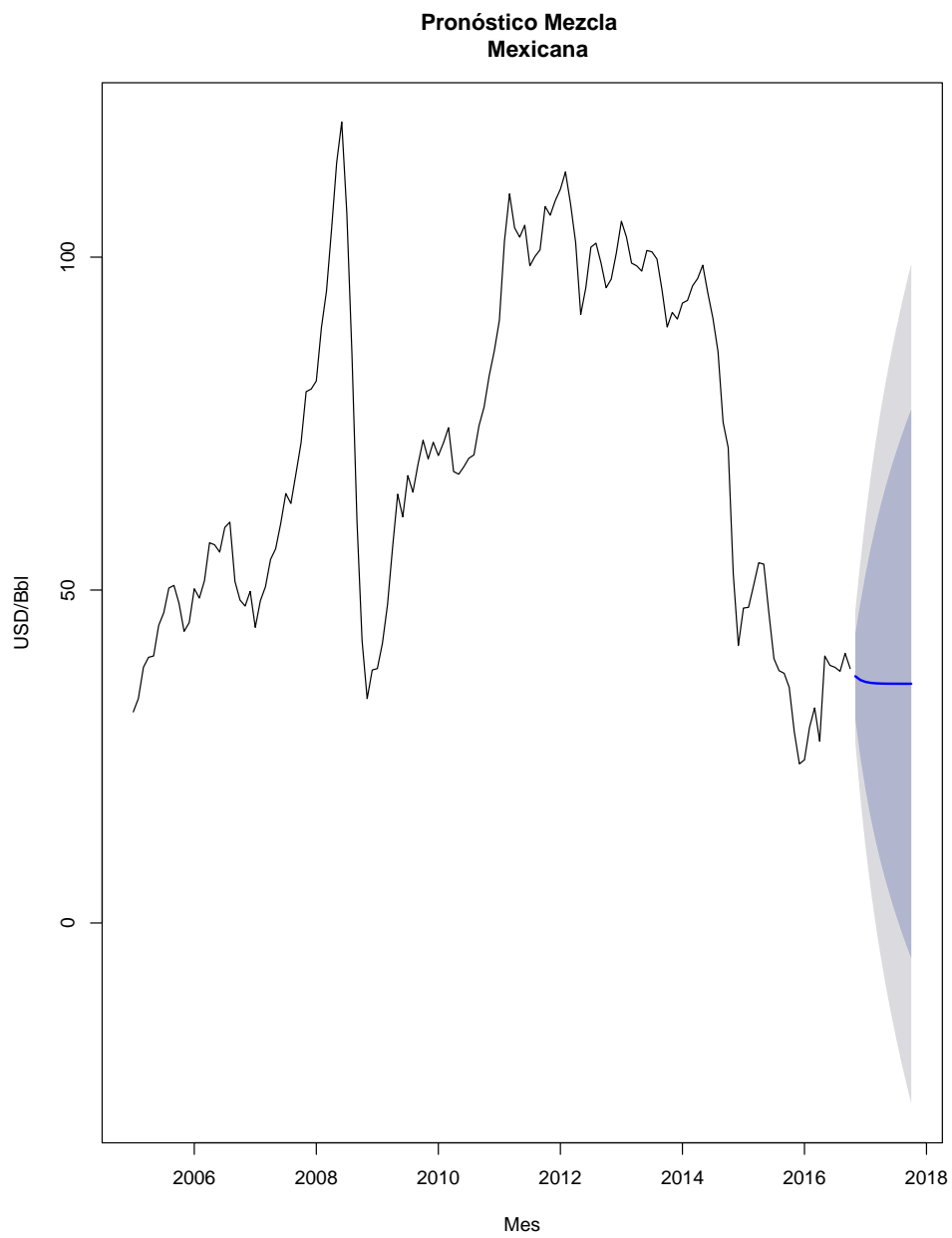


Figura 1.15: Pronóstico Mezcla Mexicana

Capítulo 2

Función de Transferencia

2.1. Introducción

El modelo de función de transferencia es un modelo estadístico que describe la relación entre una variable output X_2 y una o más variables input X_1 . Este modelo tiene muchas aplicaciones en economía, negocios, física, biología, entre otras.

Por simplicidad, el capítulo se enfocará en modelos lineales discretos. Por tiempo discreto, se entiende que los datos son observados en intervalos de tiempo discretos de manera ordenada, aunque los procesos puedan ser continuos en el tiempo.

Conviene especificar un modelo de función de transferencia cuando:

1. Se espera que la relación entre el input y output tenga una estructura dinámica suficientemente rica como para que la representación econométrica requiriese de un número elevado de parámetros.
2. Se cree que una representación adecuada de la estructura estocástica del término de error resultante de la relación de X_1 hacia X_2 necesita de una modelización ARIMA, pues no estaría siendo capturado por los sencillos esquemas utilizados en el modelo lineal general.

Un modelo de función de transferencia tiene ciertas ventajas adicionales:

1. Tiene una representación parsimoniosa, es decir, basada en un número re-

ducido de coeficientes.

2. Dispone de una estrategia sencilla y gradual para la especificación de un modelo dinámico de relación que capture adecuadamente el efecto del input sobre el output.

2.2. Teoría preliminar

Se define el filtro lineal como

$$\begin{aligned}
 X_{t,2} &= L(X_{t,1}) \\
 &= \sum_{j=0}^{\infty} \tau_j B^j X_{t,1} \\
 &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1} \\
 &= T(B) X_{t,1}
 \end{aligned} \tag{2.1}$$

donde $\sum_{j=0}^{\infty} \tau_j B^j$ se denomina función de respuesta al impulso, pues sus sucesivos coeficientes describen el efecto que, a través del tiempo, tendría sobre X_2 un impulso en la variable X_1 . A 2.1 se le conoce como regresión retrasada (*lagged regression* por su nombre en inglés) porque relaciona la variable dependiente $X_{t,2}$ con la variable input $X_{t,1}$ retrasada, o sea vemos, cómo importan los valores retrasados del input sobre el output. Tomando en cuenta lo anterior, 2.1 es:

1. Invariante en el tiempo: los coeficientes $\{\tau_j\}$ no dependen del tiempo.
2. Físicamente realizable. Si $\tau_j = 0$ para $j < 0$, es decir, el output $X_{t,2}$ es una función lineal de los valores actuales y pasados del input.

$$\begin{aligned}
 X_{t,2} &= \tau_0 X_{t,1} + \tau_1 X_{t-1,1} + \dots \\
 &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1}
 \end{aligned} \tag{2.2}$$

3. Estable. Si y sólo si $\sum_{j=0}^{\infty} |\tau_j| < \infty$

Hay dos casos especiales para el input:

Función de respuesta al impulso. Si $X_{t,1}$ es un impulso unitario en $t = 0$, es decir,

$$X_{t,1} = \begin{cases} 1 & \text{si } t = 0, \\ 0 & \text{si } t \neq 0. \end{cases} \quad (2.3)$$

entonces el output $X_{t,2}$ es

$$\begin{aligned} X_{t,2} &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1} \\ &= \tau_t \end{aligned} \quad (2.4)$$

Por lo tanto, los coeficientes τ_j en 2.2 son llamados, como anteriormente se dijo, función de respuesta al impulso.

Función de respuesta al escalón. Si $X_{t,1}$ es un escalón unitario, es decir,

$$X_{t,1} = \begin{cases} 0 & \text{si } t < 0, \\ 1 & \text{si } t \geq 0. \end{cases} \quad (2.5)$$

entonces el output $X_{t,2}$ es

$$\begin{aligned} X_{t,2} &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1} \\ &= \sum_{j=0}^t \tau_j \end{aligned} \quad (2.6)$$

Por lo tanto, los coeficientes τ_j son llamados función de respuesta al escalón.

Una generalización de la función de respuesta al escalón se obtiene cuando la ecuación 2.5 es modificada tal que $X_{t,1}$ se mantiene en un cierto valor X después de $t \geq 0$; es decir,

$$X_{t,1} = \begin{cases} 0 & \text{si } t < 0, \\ X & \text{si } t \geq 0. \end{cases} \quad (2.7)$$

entonces, se tiene que el output $X_{t,2}$ es

$$\begin{aligned} X_{t,2} &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1} \\ &= \left(\sum_{j=0}^t \tau_j \right) X \\ &= gX \end{aligned} \quad (2.8)$$

donde g es llamada la ganancia de la relación.

Para obtener una representación más real se le suele añadir un término de ruido a la ecuación 2.2, esto con el fin de tomar en cuenta factores ignorados que pudieran tener un efecto en la variable output. Por lo tanto, la representación final del modelo es

$$\begin{aligned} X_{t,2} &= T(B) X_{t,1} + N_t \\ &= \sum_{j=0}^{\infty} \tau_j X_{t-j,1} + N_t \end{aligned} \quad (2.9)$$

$\{N_t\}$ es un proceso estacionario con media cero no correlacionado con el proceso de entrada $\{X_{t,1}\}$.

Por otro lado, se puede suponer que los coeficientes $\{\tau_j\}$ tienen una estructura que puede representarse como

$$\begin{aligned} T(B) &= \sum_{j=0}^{\infty} \tau_j B^j \\ &= \frac{\omega(B)}{v(B)} \\ &= \frac{\omega_0 + \omega_1 B + \dots + \omega_s B^s}{1 - v_1 B - \dots - v_r B^r} \end{aligned} \quad (2.10)$$

El denominador de la ecuación 2.10 resumen una cantidad infinita de coeficientes con una estructura determinada por $\{v_i\}$ como si se tratara de la parte AR en un modelo ARMA y el numerador representa el ajuste que se le pueden hacer a los coeficientes $\{\omega_i\}$ como si fuera la parte MA de un modelo ARMA.

Entonces, el modelo de función de transferencia con ruido de la ecuación 2.9, puede ser reescrito como:

$$X_{t,2} = \frac{\omega(B)}{v(B)} X_{t,1} + N_t \quad (2.11)$$

En algunos procesos, puede haber un retraso sobre el input $X_{t,1}$, lo que ocasiona, como se dijo anteriormente, un impacto retrasado sobre la variable $X_{t,2}$. Entonces, se asume que hay b unidades de tiempo de retraso entre la respuesta y el input, esto da como resultado el siguiente modelo de función de transferencia con ruido.

$$\begin{aligned} X_{t,2} &= \frac{\omega(B)}{v(B)} X_{t-b,1} + N_t \\ &= \frac{\omega(B)}{v(B)} B^b X_{t,1} + N_t \\ &= T(B) X_{t,1} + N_t \end{aligned} \quad (2.12)$$

donde $T(B) = \frac{\omega(B)}{v(B)} B^b$.

La estabilidad de $T(B)$ depende de $v(B)$, de hecho, se sabe que $T(B)$ es estable si todas las raíces de $1 - v_1 m - \dots - v_r m^r$ son menores que 1 en valor absoluto.

Una vez que $\omega(B)$ y $v(B)$ son estimados, $T(B)$ se puede calcular recursivamente de

$$v(B) T(B) = \omega(B) B^b \quad (2.13)$$

Expandiendo 2.13 se obtiene

$$(1 - v_1B - v_2B^2 - \cdots - v_rB^r) (\tau_0 + \tau_1B + \tau_2B^2 + \tau_3B^3 + \cdots) = (\omega_0 + \omega_1B + \omega_2B^2 + \cdots + \omega_sB^s) B^b \quad (2.14)$$

Desarrollando 2.14 y agrupando términos

$$\tau_0 + (\tau_1 - v_1\tau_0) B + (\tau_2 - \tau_1v_1) B^2 + (\tau_3 - \tau_1v_2) B^3 + \cdots + (\tau_j - v_r\tau_{j-r}) B^j + \cdots = \omega_0B^b + \omega_1B^{b+1} + \omega_2B^{b+2} + \cdots + \omega_sB^{b+s} \quad (2.15)$$

Sustituyendo $j = b + s$ en 2.15 , por lo tanto se tiene

$$\tau_0 + (\tau_1 - v_1\tau_0) B + (\tau_2 - \tau_1v_1) B^2 + (\tau_3 - \tau_1v_2) B^3 + \cdots + (\tau_j - v_r\tau_{j-r}) B^j + \cdots = \omega_0B^b + \omega_1B^{b+1} + \omega_2B^{b+2} + \cdots + \omega_{j-b}B^j \quad (2.16)$$

Ahora, se igualan los coeficientes de potencias similares de B de 2.16 , entonces

$$\begin{aligned} \tau_0 &= 0 \\ \tau_1 - v_1\tau_0 &= 0 \\ \tau_2 - \tau_1v_1 &= 0 \\ \tau_3 - \tau_1v_2 &= 0 \\ \tau_b - \tau_{b-r}v_r &= \omega_0 \\ &\vdots \\ \tau_j - \tau_{j-r}v_r &= \omega_{j-b} \end{aligned}$$

Por lo tanto, se tiene

$$\begin{aligned}\tau_0 &= 0 \\ \tau_1 &= 0 \\ \tau_2 &= 0 \\ \tau_3 &= 0 \\ &\vdots \\ \tau_j &= \omega_{j-b}, \quad j = b, b+1, \dots, b+s\end{aligned}$$

De esta manera se llega a 2.17

$$\tau_j - v_1\tau_{j-1} - v_2\tau_{j-2} - \dots - v_r\tau_{j-r} = \begin{cases} \omega_{j-b} & \text{si } j = b+1, \dots, b+s \\ 0 & \text{si } j > b+s \end{cases} \quad (2.17)$$

con $\tau_b = \omega_0$ y $\tau_j = 0$ para $j < b$. Por lo tanto, de forma general se obtiene:

$$\tau_j = \begin{cases} 0 & \text{si } j < b \\ v_1\tau_{j-1} + v_2\tau_{j-2} + \dots + v_r\tau_{j-r} + \omega_0 & \text{si } j = b \\ v_1\tau_{j-1} + v_2\tau_{j-2} + \dots + v_r\tau_{j-r} + \omega_{j-b} & \text{si } j = b+1, \dots, b+s \\ v_1\tau_{j-1} + v_2\tau_{j-2} + \dots + v_r\tau_{j-r} & \text{si } j > b+s \end{cases} \quad (2.18)$$

Las características de los coeficientes $\{\tau_j\}$ son determinados por los valores b , r y s .

Ejemplo.

CASO 1. Se considerará el caso para $b = 2$, $r = 1$ y $s = 2$. Entonces, se tiene el siguiente modelo

$$X_{t,2} = \frac{\omega_0 + \omega_1 B + \omega_2 B^2}{1 - v_1 B} X_{t-2,1}$$

La estabilidad de este modelo se logra cuando $|v_1| < 1$, ya que así las raíces del polinomio $v(B)$ se encuentran fuera del círculo unitario. Usando 2.18 se tiene:

$$\begin{aligned}
\tau_0 &= \tau_1 = 0 \\
\tau_2 &= \omega_0 \\
\tau_3 &= v_1\omega_0 + \omega_1 \\
\tau_4 &= v_1^2\omega_0 + v_1\omega_1 + \omega_2 \\
\tau_j &= v_1\tau_{j-1}, \quad j > 4
\end{aligned}$$

CASO 2. En este caso se tomará el modelo cuando $r = 2$, $s = 2$ y $b = 2$. Entonces

$$X_{t,2} = \frac{\omega_0 + \omega_1 B + \omega_2 B^2}{1 - v_1 B - v_2 B^2} X_{t-2,1}$$

La estabilidad de la función de transferencia depende de la raíces asociadas al polinomio $1 - v_1 m - v_2 m^2$. Para verificar la estabilidad, las raíces se obtienen usando

$$m_1, m_2 = \frac{v_1 \pm \sqrt{v_1^2 + 4v_2}}{2}$$

Para asegurar dicha característica, se debe cumplir que $|m_1|, |m_2| < 1$. Esto significa que

$$\begin{aligned}
v_2 - v_1 &< 1 \\
v_2 + v_1 &< 1 \\
-1 &< v_2 < 1
\end{aligned}$$

Si $v_1^2 + 4v_2 < 0$ se exhibirá un comportamiento senoidal, ya que las raíces serán complejas, por el contrario, en caso de que $v_1^2 + 4v_2 > 0$ se tendrá un patrón de decaimiento exponencial. A continuación se muestran las distintas formas o estructuras de la función de respuesta al impulso con $b = 2$, $r \leq 2$ y $s \leq 2$. Estas distintas estructuras servirán para elegir el modelo.

En la Figura 2.1 es un modelo de la forma $X_{t,2} = \omega_0 X_{t-2,1}$ con $\omega_0 = 0.5$ en el primer caso y $\omega_0 = -0.5$ en el segundo caso.

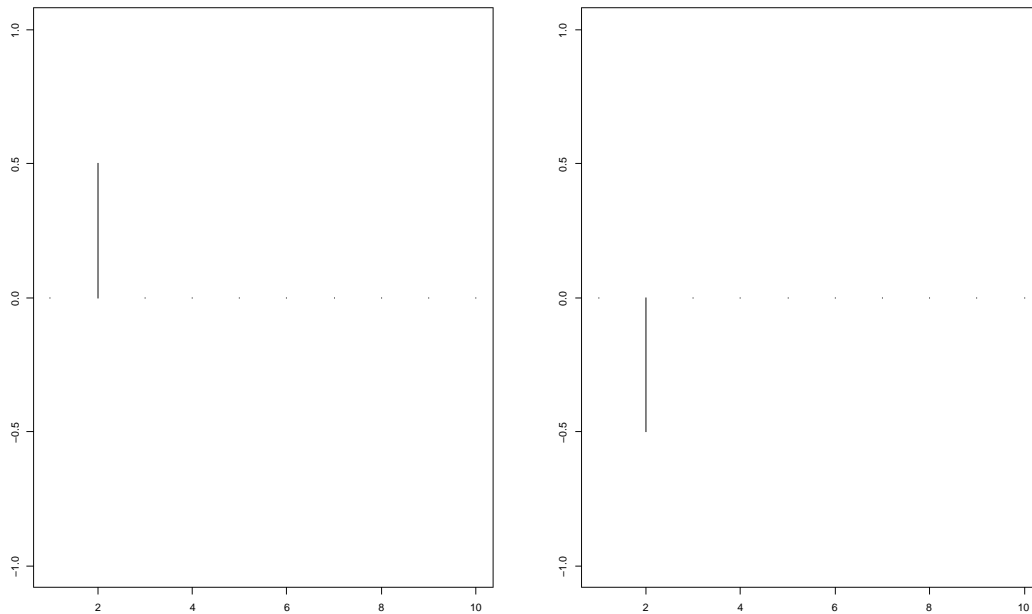


Figura 2.1: Funciones de respuesta al impulso, primera estructura

En la Figura 2.2 es un modelo de la forma $X_{t,2} = (\omega_0 + \omega_1 B) X_{t-2,1}$ con $\omega_0 = 0.5$ y $\omega_1 = 0.4$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$ y $\omega_1 = -0.4$.

En la Figura 2.3 es un modelo de la forma $X_{t,2} = (\omega_0 + \omega_1 B + \omega_2 B^2) X_{t-2,1}$ con $\omega_0 = 0.5$, $\omega_1 = 0.4$ y $\omega_2 = 0.6$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$, $\omega_1 = -0.4$ y $\omega_2 = 0.6$.

En la Figura 2.4 es un modelo de la forma $X_{t,2} = \frac{\omega_0}{1-v_1 B} X_{t-2,1}$ con $\omega_0 = 0.5$ y $v_1 = 0.6$ en el primer caso, mientras que en el segundo caso $\omega_0 = -0.5$ y $v_1 = 0.6$.

En la Figura 2.5 es un modelo de la forma $X_{t,2} = \frac{\omega_0 + \omega_1 B}{1-v_1 B} X_{t-2,1}$ con $\omega_0 = 0.5$, $\omega_1 = 0.4$ y $v_1 = 0.6$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$, $\omega_1 = -0.4$ y $v_1 = 0.6$.

En la Figura 2.6 es un modelo de la forma $X_{t,2} = \frac{\omega_0 + \omega_1 B + \omega_2 B^2}{1-v_1 B} X_{t-2,1}$ con $\omega_0 = 0.5$, $\omega_1 = 0.4$, $\omega_2 = 0.6$ y $v_1 = 0.6$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$, $\omega_1 = -0.4$, $\omega_2 = 0.6$ y $v_1 = 0.6$.

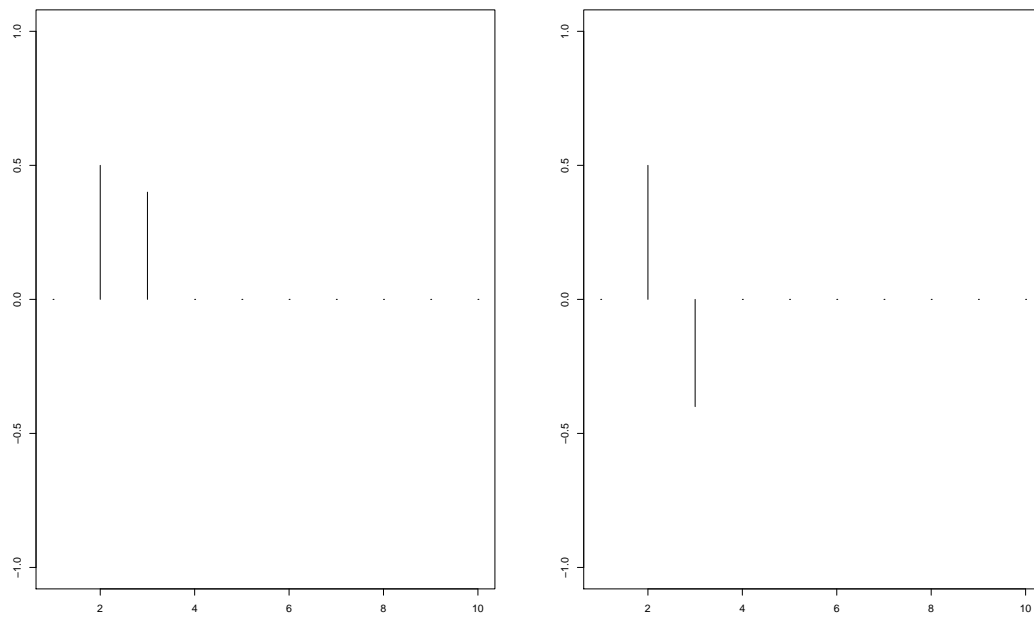


Figura 2.2: Funciones de respuesta al impulso, segunda estructura

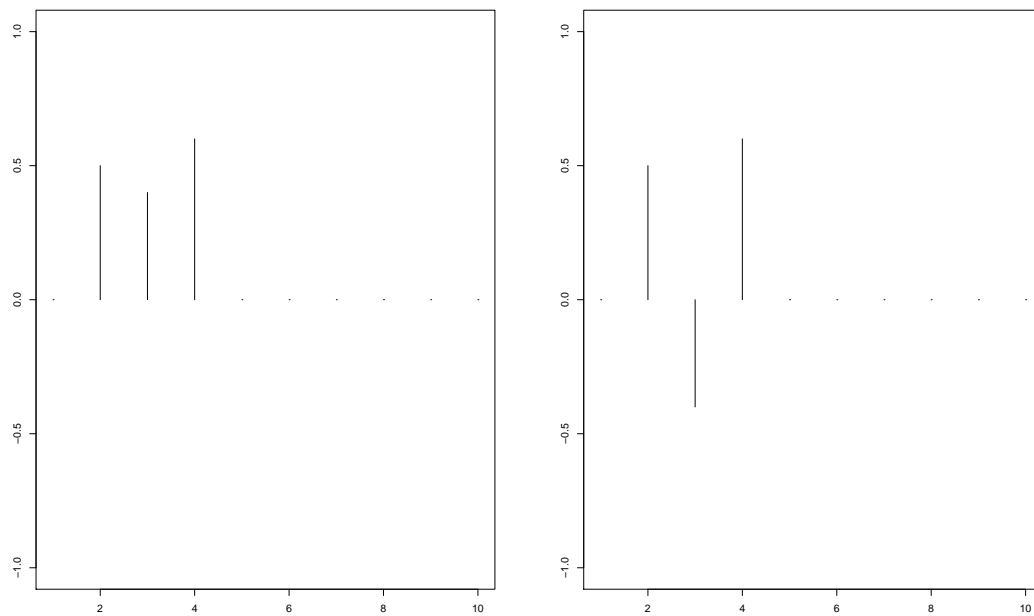


Figura 2.3: Funciones de respuesta al impulso, tercera estructura

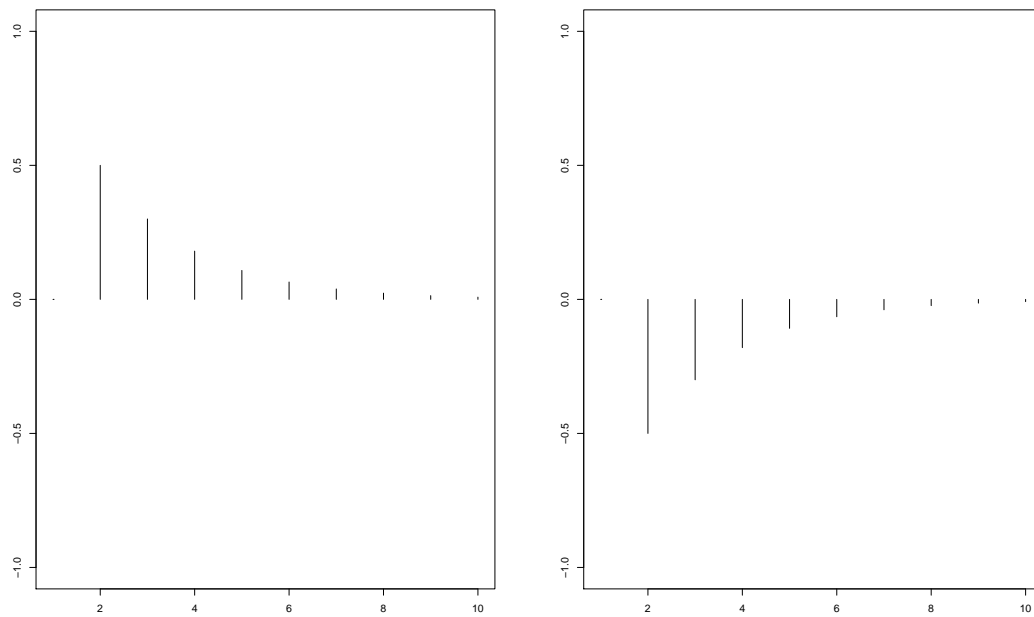


Figura 2.4: Funciones de respuesta al impulso, cuarta estructura

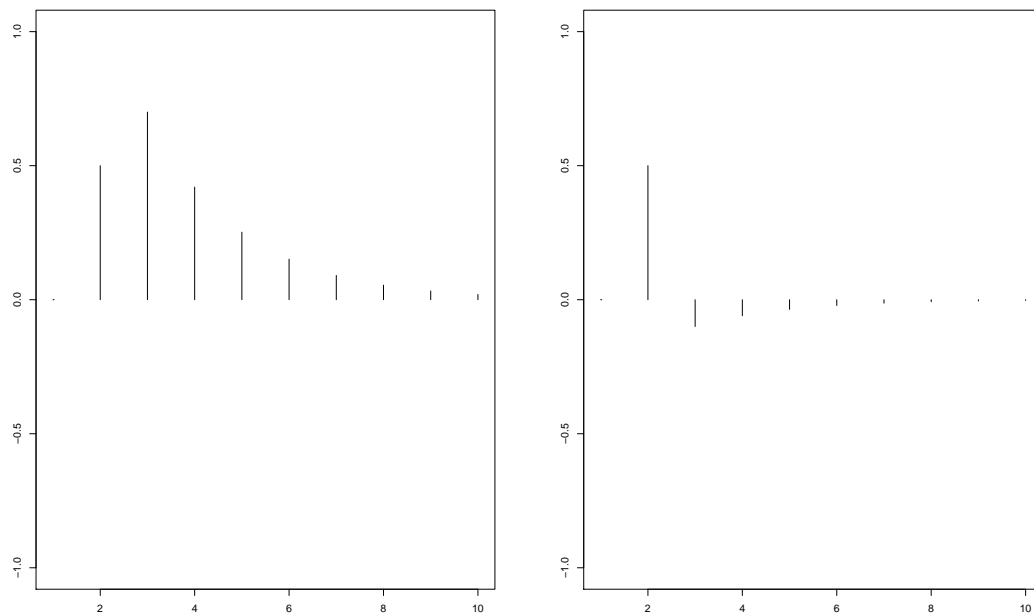


Figura 2.5: Funciones de respuesta al impulso, quinta estructura

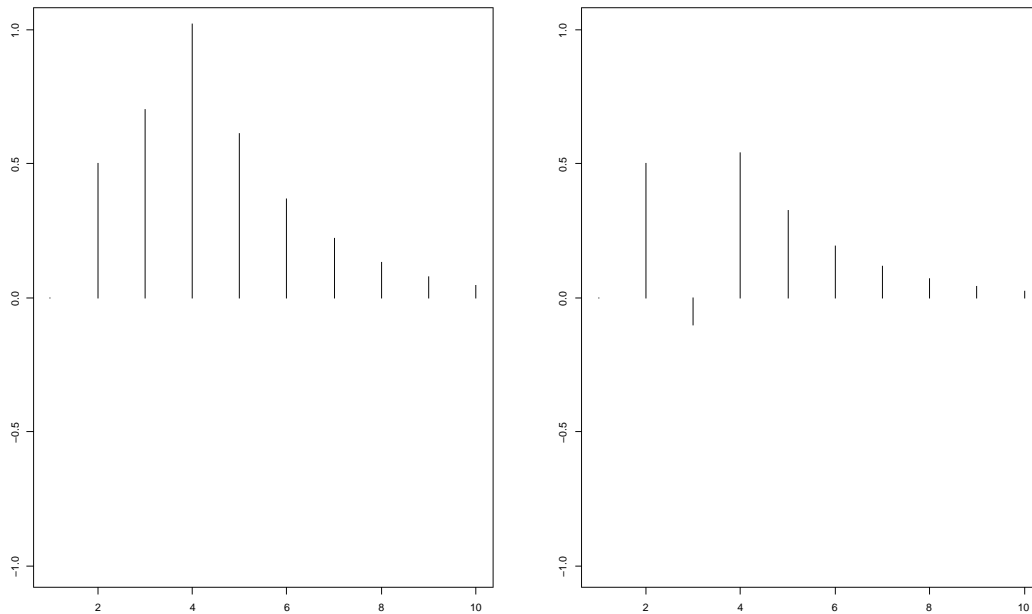


Figura 2.6: Funciones de respuesta al impulso, sexta estructura

En la Figura 2.7 es un modelo de la forma $X_{t,2} = \frac{\omega_0}{1-v_1B-v_2B^2}X_{t-2,1}$ con $\omega_0 = 0.5$, $v_1 = 0.6$ y $v_2 = 0.3$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$, $v_1 = 0.6$ y $v_2 = -0.7$.

En la Figura 2.8 es un modelo de la forma $X_{t,2} = \frac{\omega_0+\omega_1B}{1-v_1B-v_2B^2}X_{t-2,1}$ con $\omega_0 = 0.5$, $\omega_1 = 0.4$, $v_1 = 0.6$ y $v_2 = 0.3$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$, $\omega_1 = 0.4$, $v_1 = 0.6$ y $v_2 = -0.7$.

Por último, la Figura 2.9 es un modelo de la forma $X_{t,2} = \frac{\omega_0+\omega_1B+\omega_2B^2}{1-v_1B-v_2B^2}X_{t-2,1}$ con $\omega_0 = 0.5$, $\omega_1 = 0.4$, $\omega_2 = 0.6$, $v_1 = 0.6$ y $v_2 = 0.3$ en el primer caso, mientras que en el segundo caso $\omega_0 = 0.5$, $\omega_1 = 0.4$, $\omega_2 = 0.6$, $v_1 = 0.6$ y $v_2 = -0.7$.

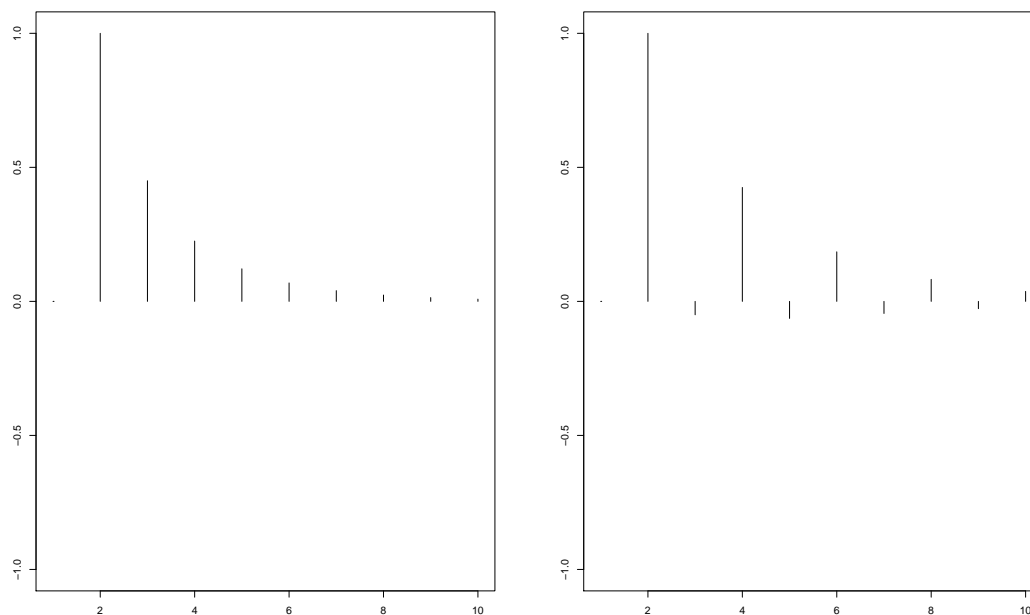


Figura 2.7: Funciones de respuesta al impulso, séptima estructura

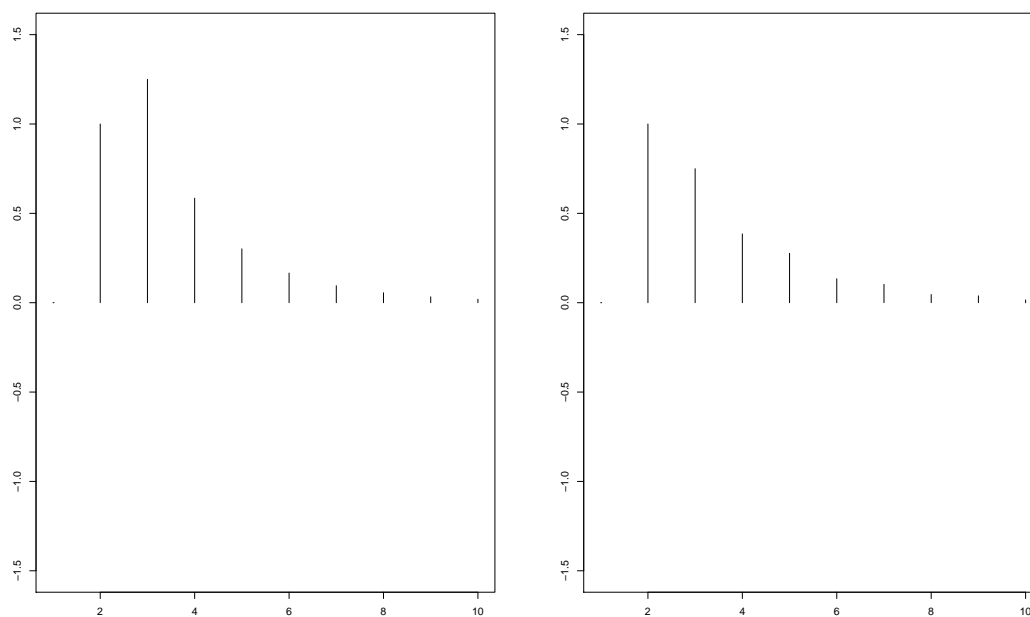


Figura 2.8: Funciones de respuesta al impulso, octava estructura

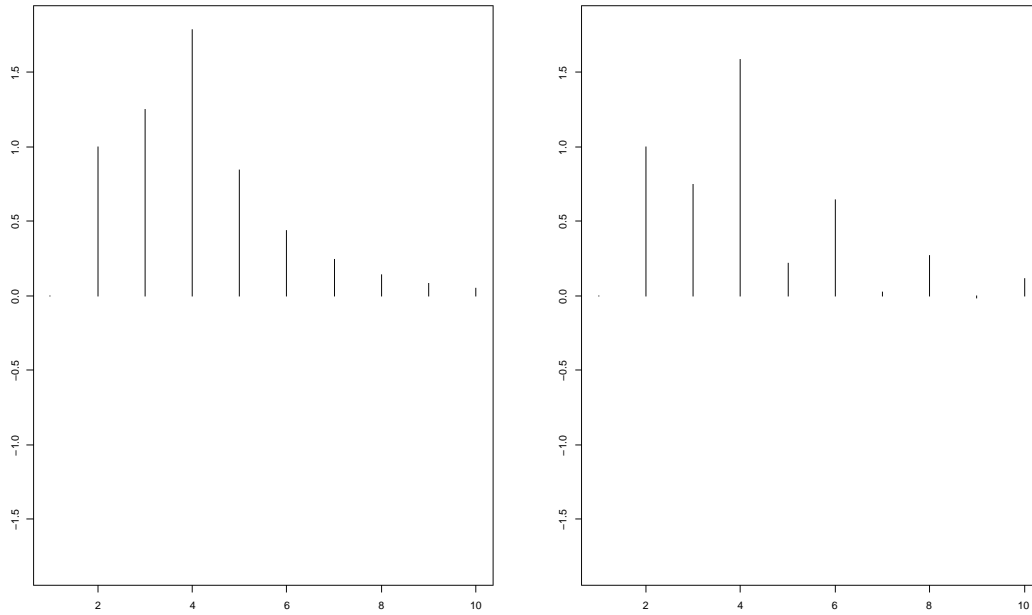


Figura 2.9: Funciones de respuesta al impulso, novena estructura

2.3. Modelo de Función de Transferencia

El modelo de función de transferencia, usando 2.12 se escribe como

$$\begin{aligned} X_{t,2} &= T(B) X_{t,1} + \psi_N(B) W_t \\ &= \frac{\omega(B)}{v(B)} X_{t-b,1} + \frac{\theta_N(B)}{\varphi_N(B)} W_t \end{aligned} \quad (2.19)$$

donde se observa que es un modelo de regresión retrasada con errores (N_t) ARIMA(p,d,q), la cual es

$$\underbrace{\phi_N(B) (1-B)^d}_{\varphi_N(B)} N_t = \theta_N(B) W_t$$

donde $\phi_N(B)$ y $\theta_N(B)$ son los polinomios autoregresivo y de promedios móviles asociados al proceso ARIMA, además N_t es independiente de $X_{t,1}$ y $\{W_t\}$ es ruido blanco con $\mathbb{E}(W_t) = 0$.

Después de reordenar la ecuación 2.19, se tiene:

$$\underbrace{v(B) \varphi_N(B)}_{=v^*(B)} X_{t,2} = \underbrace{\varphi_N(B) \omega(B)}_{=\omega^*(B)} X_{t-b,1} + \underbrace{v(B) \theta_N(B)}_{=\theta^*(B)} W_t \quad (2.20)$$

$$v^*(B) X_{t,2} = \omega^*(B) X_{t-b,1} + \theta^*(B) W_t \quad (2.21)$$

O bien,

$$X_{t,2} - \sum_{j=1}^{r^*} v_j^* X_{t-j,2} = \omega_0^* X_{t-b,1} + \sum_{j=1}^{s^*} \omega_j^* X_{t-b-j,1} + W_t + \sum_{j=1}^{q^*} \theta_j^* W_{t-j} \quad (2.22)$$

donde $v^*(B) = 1 - v_1^* B - \dots - v_{r^*}^* B^{r^*}$, $\omega^*(B) = \omega_0^* + \omega_1^* B + \dots + \omega_{s^*}^* B^{s^*}$ y $\theta^*(B) = 1 + \theta_1^* B + \dots + \theta_{q^*}^* B^{q^*}$ son los polinomios obtenidos al simplificar la expresión 2.20.

Debido a la adición de $X_{t,1}$, el modelo 2.22 es también llamado en la literatura econométrica como un modelo ARMAX. Por lo tanto, el modelo de función de transferencia dado en 2.22 se puede interpretar como un modelo ARMA con una variable exógena adicional $X_{t,1}$.

2.4. Función de Correlación Cruzada

Para la serie de tiempo bivariada $(X_{t,1}, X_{t,2})$ se define la covarianza cruzada como

$$S(t,1, s,2) = Cov(X_{t,1}, X_{s,2})$$

Asumiendo que $(X_{t,1}, X_{s,2})$ es (débilmente) estacionario, se tiene

$$\mathbb{E}(X_{t,1}) = \mu_1, \text{ constante para toda } t$$

$$\mathbb{E}(X_{t,2}) = \mu_2, \text{ constante para toda } t$$

$$\text{Cov}(X_{t,1}, X_{t+j,1}) = S_j^1, \text{ depende sólo de } j$$

$$\text{Cov}(X_{t,2}, X_{t+j,2}) = S_j^2, \text{ depende sólo de } j$$

y

$$\text{Cov}(X_{t,1}, X_{t+j,2}) = S_j^{12}, \text{ depende sólo de } j \text{ para } j = 0, \pm 1, \pm 2, \dots$$

La función de correlación cruzada es una medida muy útil de correlación entre dos variables aleatorias.

La función de correlación cruzada (CCF) está definida como

$$\rho_j^{12} = \text{corr}(X_{t,1}, X_{t+j,2}) = \frac{S_j^{12}}{\sqrt{S_0^1 S_0^2}} = \frac{S_j^{12}}{\sigma_{X1} \sigma_{X2}} \text{ para } j = 0, \pm 1, \pm 2, \dots$$

Observación. $\rho_j^{12} \neq \rho_{-j}^{12}$, pero $\rho_j^{12} = \rho_{-j}^{21}$

Así, la función de correlación cruzada no sólo mide la magnitud de una asociación, sino también su dirección. A diferencia de la función de autocorrelación (FAC) y de la función de autocorrelación parcial (FACP), ésta no es simétrica y su valor central ρ_0^{12} es el coeficiente de correlación habitual entre ambas variables.

La matriz de correlación al retraso j se define como:

$$\rho_j = \begin{bmatrix} \rho_j^1 & \rho_j^{12} \\ \rho_j^{21} & \rho_j^2 \end{bmatrix}$$

Para una muestra de n observaciones, la covarianza cruzada muestral se estima de la siguiente forma

$$\hat{S}_j^{12} = \frac{1}{n} \sum_{t=1}^{n-j} (X_{t,1} - \bar{X}_1) (X_{t+j,2} - \bar{X}_2) \text{ para } j = 0, 1, 2, \dots$$

y

$$\hat{S}_{-j}^{12} = \frac{1}{n} \sum_{t=1}^{n+j} (X_{t-j,1} - \bar{X}_1) (X_{t,2} - \bar{X}_2) \text{ para } j = -1, -2, \dots$$

Donde \bar{X}_1 y \bar{X}_2 son las medias muestrales de las series $X_{t,1}$ y $X_{t,2}$, respectivamente. De manera similar, la correlación cruzada muestral se estima de la siguiente manera

$$r_j^{12} = \hat{\rho}_j^{12} = \frac{\hat{S}_j^{12}}{\sqrt{\hat{S}_0^1 \hat{S}_0^2}} \text{ para } j = 0, \pm 1, \pm 2, \dots$$

donde $\hat{S}_0^1 = \frac{1}{n} \sum_{t=1}^n (X_{t,1} - \bar{X}_1)^2$ y $\hat{S}_0^2 = \frac{1}{n} \sum_{t=1}^n (X_{t,2} - \bar{X}_2)^2$

2.5. Especificación, Ajuste y Verificación

Antes de empezar se considera el modelo general de función de transferencia dado anteriormente y denotado como:

$$\begin{aligned} X_{t,2} &= T(B) X_{t,1} + N_t \\ &= \frac{\omega(B)}{v(B)} X_{t-b,1} + \frac{\theta_N(B)}{\varphi_N(B)} W_t \end{aligned} \quad (2.23)$$

El objetivo en esta parte, es obtener alguna idea de los órdenes r y s de los operadores que se tienen en el numerador como el denominador del modelo de función de transferencia, además de estimar los parámetros v , ω y b (parámetro de retraso). Después, se ajustará un proceso al ruido N_t y se obtendrán estimaciones iniciales de los parámetros Φ y Θ asociadas al proceso ARIMA correspondiente al proceso N_t . Posteriormente, se reestimaré todo el modelo ya en su conjunto,

es decir, se estimará nuevamente el modelo juntando la parte input y el ruido. Finalmente, se verificarán los supuestos del modelo, para saber si es adecuado para su uso. El proceso consta de 6 pasos y se explican de la siguiente forma

PASO 1. Obtener estimaciones preliminares de los coeficientes de T(B).

Identificación de T(B) usando mínimos cuadrados.

Escribimos 2.23 en su forma desarrollada, es decir:

$$X_{t,2} = \tau_0 X_{t,1} + \tau_1 X_{t-1,1} + \tau_2 X_{t-2,1} + \cdots + N_t \quad (2.24)$$

Donde $\{X_{t,2}\}$, $\{X_{t,1}\}$ y N_t son procesos estacionarios con media cero. Entonces, al multiplicar 2.24 por $X_{t-k,1}$ para $k \geq 0$, se obtiene

$$X_{t-k,1} X_{t,2} = \tau_0 X_{t-k,1} X_{t,1} + \tau_1 X_{t-k,1} X_{t-1,1} + \tau_2 X_{t-k,1} X_{t-2,1} + \cdots + X_{t-k,1} N_t \quad (2.25)$$

Si se hace una suposición que $X_{t-k,1}$ no está correlacionado con N_t para toda t , aplicando el operador esperanza en 2.25 se produce el siguiente conjunto de ecuaciones

$$S_k^{21} = \tau_0 S_k^{11} + \tau_1 S_{k-1}^{11} + \tau_2 S_{k-2}^{11} + \cdots \text{ para } k = 0, 1, 2, \dots \quad (2.26)$$

Supóngase que los pesos τ_j son cero hasta $k = K$ donde $K \in \mathbb{Z}$. Entonces las primeras $K + 1$ ecuaciones de 2.26 se pueden escribir como

$$S^{21} = \Gamma^{11} \tau \quad (2.27)$$

$$S^{21} = \begin{bmatrix} S_0^{21} \\ S_1^{21} \\ S_2^{21} \\ \vdots \\ S_K^{21} \end{bmatrix}$$

$$\tau = \begin{bmatrix} \tau_0 \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_K \end{bmatrix}$$

$$\Gamma^{11} = \begin{bmatrix} S_0^{11} & S_1^{11} & \cdots & S_K^{11} \\ S_1^{11} & S_0^{11} & \cdots & S_{K-1}^{11} \\ \vdots & \vdots & \ddots & \vdots \\ S_K^{11} & S_{K-1}^{11} & \cdots & S_0^{11} \end{bmatrix}$$

Sustituyendo estimaciones \hat{S}_K^{11} de la función de autocovarianza del input $X_{t,1}$ y las estimaciones \hat{S}_K^{21} de la función de autocovarianza cruzada entre el input $X_{t,1}$ y $X_{t,2}$, la ecuación 2.27 da $K + 1$ ecuaciones lineales para los $K + 1$ pesos. Sin embargo, este enfoque puede traer estimaciones equivocadas, $X_{t,1}$ puede tener fuerte autocorrelación, además el procedimiento se ve demasiado engorroso.

La versión muestral de la ecuación 2.27, representa las ecuaciones por mínimos cuadrados de una regresión lineal entre $X_{t,2}$ sobre $X_{t,1}, X_{t-1,1}, \dots, X_{t-k,1}$. Por lo tanto, es mejor utilizar el método de *preblanqueamiento* del input.

Método de Preblanqueamiento.

Suponga que el input $X_{t,1}$ de la ecuación 2.23 sigue un modelo ARIMA como

$$\underbrace{\phi_X(B)(1-B)^d}_{\varphi_X(B)} X_{t,1} = \theta_X(B) Z_t$$

donde $\{Z_t\}$ es un ruido blanco con varianza σ_Z^2 . De manera equivalente se tiene que

$$Z_t = \theta_X(B)^{-1} \varphi_X(B) X_{t,1}$$

Siempre que $\{X_{t,1}\}$ sea invertible. Usando esta notación, se puede observar que $\pi_X(B) = \theta_X(B)^{-1} \varphi_X(B)$ es visto como un filtro que cuando se aplica a $X_{t,1}$

se genera una serie de tiempo de ruido blanco, a esto se le debe el nombre de *preblanqueamiento*.

Cuando se aplica este filtro al modelo de función de transferencia 2.12, se obtiene

$$\underbrace{\theta_X(B)^{-1} \varphi_X(B) X_{t,2}}_{Y_t} = \theta_X(B)^{-1} \varphi_X(B) T(B) X_{t,1} + \underbrace{\theta_X(B)^{-1} \varphi_X(B) N_t}_{N_t^*}$$

$$Y_t = T(B) Z_t + N_t^*$$

La covarianza cruzada entre las series filtradas Z_t y Y_t es:

$$\begin{aligned} S_j^{YZ} &= Cov(Y_t, Z_{t-j}) \\ &= \mathbb{E}(Y_t Z_{t-j}) \\ &= \mathbb{E}(Z_{t-j} (\tau_0 Z_t + \tau_1 Z_{t-1} + \tau_2 Z_{t-2} + \tau_3 Z_{t-3} + \cdots + N_t^*)) \\ &= \mathbb{E}(\tau_0 Z_{t-j} Z_t + \tau_1 Z_{t-j} Z_{t-1} + \tau_2 Z_{t-j} Z_{t-2} + \tau_3 Z_{t-j} Z_{t-3} + \cdots + Z_{t-j} N_t^*) \\ &= \sum_{i=0}^{\infty} \mathbb{E}(\tau_i Z_{t-j} Z_{t-i}) \end{aligned} \quad (2.28)$$

$\mathbb{E}(Z_{t-j} N_t^*) = 0$, porque se asume que $X_{t,1}$ y N_t son independientes y N_t^* es función de N_t y Z_t de $X_{t,1}$.

Como las τ convergen absolutamente, entonces la suma se puede intercambiar con la esperanza.

Supóngase que Z_t es ruido blanco

$$\mathbb{E}(Z_{t-j} Z_{t-i}) = \begin{cases} 0 & \text{si } j \neq i \\ \sigma_Z^2 & \text{si } j = i \end{cases} \quad (2.29)$$

$$\begin{aligned} \text{Entonces, } S_j^{YZ} &= \tau_j Var(Z_t) \\ &= \tau_j \sigma_Z^2 \end{aligned} \quad (2.30)$$

De la ecuación 2.30, se tiene que $S_j^{YZ} = \tau_j \sigma_Z^2$ y por lo tanto

$$\tau_j = \frac{S_j^{YZ}}{\sigma_Z^2} = \frac{\rho_j^{YZ} \sigma_Y \sigma_Z}{\sigma_Z^2} = \rho_j^{YZ} \frac{\sigma_Y}{\sigma_Z}$$

donde $\rho_j^{YZ} = \text{corr}(Y_t, Z_{t-j})$ es la correlación cruzada entre $\{Y_t\}$ y $\{Z_t\}$ con retraso j , $\sigma_Z^2 = \text{Var}(Z_t)$, $\sigma_Y^2 = \text{Var}(Y_t)$. Entonces, a través de estimaciones muestrales se pueden obtener estimaciones iniciales para τ_j

$$\hat{\tau}_j = r_j^{YZ} \frac{\hat{\sigma}_Y}{\hat{\sigma}_Z} \quad (2.31)$$

Dado los valores iniciales de τ_j , se puede determinar la estructura de la función de respuesta al impulso, incluyendo el parámetro de retraso.

En resumen, primero se ajustará un proceso ARIMA(p,d,q) al input $\{X_{t,1}\}$ obteniendo una estimación $\{\hat{Z}_t\}$ del proceso $\{Z_t\}$. Se obtienen los valores $(\hat{Z}_1, \dots, \hat{Z}_n)$ con n el número de observaciones. Los valores $\hat{\Phi}$, $\hat{\Theta}$ y σ_Z^2 se estiman por máxima verosimilitud.

Segundo, se aplica el operador $\hat{\pi}(B) = \hat{\theta}_X(B)^{-1} \hat{\varphi}_X(B)$ a la serie output $\{X_{t,2}\}$ obteniendo la serie $(\hat{Y}_1, \dots, \hat{Y}_n)$. Finalmente, se calculan las correlaciones y correlaciones cruzadas para $\{Z_t\}$ y $\{Y_t\}$, en este caso $\{\hat{Z}_t\}$ y $\{\hat{Y}_t\}$ como una serie bivariada. Obtener cuáles correlaciones ρ_j^{YZ} son significativas según las bandas de confianza $\frac{z^{1-\alpha/2}}{\sqrt{n}}$. Recordar que $\rho_j^{YZ} = \text{corr}(Z_t, Y_{t-j})$, así que este valor necesariamente debe ser cero, pues el input actual no puede impactar al output retrasado.

PASO 2. Especificar los órdenes de r y s.

Una vez que las estimaciones iniciales de τ_j de la ecuación 2.31 se obtienen, se pueden usar para especificar los órdenes r y s en

$$\begin{aligned} T(B) &= \frac{\omega(B)}{v(B)} B^b \\ &= \frac{\omega_0 + \omega_1 B + \dots + \omega_s B^s}{1 - v_1 B - \dots - v_r B^r} B^b \end{aligned} \quad (2.32)$$

La especificación de los órdenes r y s se puede lograr al graficar $\hat{\tau}_j$.

En particular, si $\hat{\tau}_j = 0$ para $j < b$ y $\hat{\tau}_j = \omega_0 v_1^{j-b}$ para $j \geq b$, entonces $\hat{T}(B) = \sum_{j=b}^{\infty} \tau_j B^j = \sum_{j=b}^{\infty} \omega_0 v_1^{j-b} B^j = \sum_{k=0}^{\infty} \omega_0 v_1^k B^{k+b} = \omega_0 B^b \sum_{k=0}^{\infty} (v_1 B)^k$. Suponiendo $|v_1| < 1$. Entonces,

$$\hat{T}(B) = \frac{\omega_0}{1 - v_1 B} B^b \quad (2.33)$$

En otras palabras, cuando las τ_j forman una serie geométrica, el operador $\hat{T}(B)$ tiene la forma de 2.32 con $r = 1$ y $s = 0$.

PASO 3. Obtener las estimaciones de v_i y ω_i .

Cuando el modelo especificado es como 2.33, ω_0 y v_1 se estiman a partir de las τ_j . Si $j = b$, se tiene que $\hat{\tau}_b = \omega_0 v_1^0 = \omega_0$, entonces $\hat{\tau}_b = \hat{\omega}_0$. Para $j = b + 1$, $\hat{\tau}_{b+1} = \omega_0 v_1^{b+1-b} = \omega_0 v_1$, entonces $\hat{v}_1 = \frac{\hat{\tau}_{b+1}}{\hat{\tau}_b}$,

$$\hat{v}_1 = \frac{\hat{\tau}_{b+1}}{\hat{\tau}_b}$$

En el párrafo anterior se dio de forma breve la estimación para el modelo más simple, en caso de que el modelo sea más complejo, ω_i y v_i se pueden estimar usando 2.18

$$\tau_j - v_1 \tau_{j-1} - v_2 \tau_{j-2} - \cdots - v_r \tau_{j-r} = \begin{cases} \omega_{j-b} & \text{si } j = b + 1, \dots, b + s \\ 0 & \text{si } j > b + s \end{cases} \quad (2.34)$$

con $\tau_b = \omega_0$ y $\tau_j = 0$ para $j < b$.

PASO 4. Modelar el ruido.

La estimación del ruido se obtiene a partir de

$$\begin{aligned} \hat{N}_t &= X_{t,2} - \frac{\hat{\omega}(B)}{\hat{v}(B)} X_{t-\hat{b},1} \\ &= X_{t,2} - \hat{X}_{t,2} \end{aligned} \quad (2.35)$$

Para modelar el ruido se hace uso del ACF y PACF, de esta manera se contará con suficiente información para ajustarle un proceso ARIMA al ruido.

$$\phi_N(B)(1-B)^d N_t = \theta_N(B) W_t$$

donde W_t es ruido blanco, o lo que es lo mismo

$$\varphi_N(B) N_t = \theta_N(B) W_t$$

PASO 5. Ajusta el modelo en conjunto.

Lo que se hizo del paso 1 al paso 4 fue especificar el modelo y las estimaciones iniciales de los parámetros, ahora, se ajustará todo el modelo en conjunto. Así que

$$\begin{aligned} X_{t,2} &= \frac{\omega(B)}{v(B)} B^b X_{t,1} + N_t \\ &= \frac{\omega(B)}{v(B)} X_{t-b,1} + \frac{\theta_N(B)}{\varphi_N(B)} \end{aligned} \quad (2.36)$$

donde $\varphi_N(B) = (1-B)^d \phi_N(B)$.

Entonces, $v(B) X_{t,2} = B^b \omega(B) X_{t,1} + \frac{\theta_N(B)}{\varphi_N(B)} v(B) W_t$

$$\varphi_N(B) v(B) X_{t,2} = B^b \omega(B) \varphi_N(B) X_{t,1} + \theta_N(B) v(B) W_t \quad (2.37)$$

Este será el modelo que se reestimaré y se obtendrá $\underline{\omega} = (\omega_0, \omega_1, \dots, \omega_s)'$, $\underline{v} = (v_1, \dots, v_r)'$, $\Phi_N(B) = (\phi_1, \dots, \phi_p)'$, $\Theta_N(B) = (\theta_1, \dots, \theta_q)'$ y σ_W^2 .

Para obtener la reestimación, la ecuación 2.37 se puede escribir como:

$$c(B) X_{t,2} = d(B) X_{t-b,1} + e(B) W_t \quad (2.38)$$

donde

$$\begin{aligned} c(B) &= v(B) \varphi_N(B) \\ &= (1 - v_1 B - \dots - v_r B^r) (1 - \varphi_1 B - \dots - \varphi_p B^p) \\ &= (1 - c_1 B - c_2 B^2 - \dots - c_{p+r} B^{p+r}) \end{aligned}$$

$$\begin{aligned}
d(B) &= \omega(B) \varphi_N(B) \\
&= (\omega_0 + \omega_1 B + \dots + \omega_s B^s) (1 - \varphi_1 B - \dots - \varphi_p B^p) \\
&= (d_0 - d_1 B - d_2 B^2 - \dots - d_{p+s} B^{p+s})
\end{aligned}$$

y

$$\begin{aligned}
e(B) &= \theta_N(B) v(B) \\
&= (1 + \theta_1 B + \dots + \theta_q B^q) (1 - v_1 B - \dots - v_r B^r) \\
&= (1 - e_1 B - e_2 B^2 - \dots - e_{r+q} B^{r+q})
\end{aligned}$$

Así, aplicando cada uno de los polinomios definidos sobre 2.38 y despejando, se tiene

$$W_t = X_{t,2} - c_1 X_{t-1,2} - \dots - c_{p+r} X_{t-p-r,2} - d_0 X_{t-b,1} + d_1 X_{t-b-1,1} + \dots + d_{p+s} X_{t-b-p-s,1} + e_1 W_{t-1} + \dots + e_{r+q} W_{t-r-q}$$

donde c_i , d_j y e_k son funciones de v_i , ω_j , ϕ_k y θ_l . Bajo la suposición de que las W_t 's son gaussianas, se tiene la siguiente función de verosimilitud condicional

$$L(v, \omega, \Phi_N, \Theta_N, \sigma_W^2 | b, X_{t,1}, X_{t,2}, X_{0,1}, X_{0,2}, W_0) = (2\pi\sigma_W^2)^{-n/2} \exp \left[-\frac{1}{2\sigma_W^2} \sum_{t=1}^n W_t^2 \right]$$

donde $X_{0,1}$, $X_{0,2}$ y W_0 son valores iniciales para calcular W_t . En general, la estimación de los parámetros se puede lograr minimizando la suma de cuadrados condicional.

$$S \left(v, \omega, \Phi_N, \Theta_N | b \right) = \sum_{t=t_0}^n W_t^2$$

Donde $t_0 = \max \{p + r + 1, b + p + s + 1\}$, donde b (parámetro de retraso) se asume como un valor conocido.

Afortunadamente, existe software con capacidad comprobada para ajustar este tipo de modelos complejos. Por ejemplo: R, SAS, Stata, ITSM, SCA, etc.

En caso de que se tengan más de un candidato de modelo, es recomendable usar los criterios de Información Akaike (AIC) y de Información Bayesiano (BIC) para seleccionar el modelo más adecuado.

PASO 6. Verificar los supuestos del modelo.

1. El supuesto de que el ruido $\{W_t\}$ es ruido blanco requiere la examinación de los residuos \hat{W}_t . Para hacer esta tarea de forma adecuada, se hace uso del ACF y PACF de los residuos. Para que se cumpla el supuesto, las ACF y PACF para todos los lags deben estar dentro de los intervalos de confianza. En caso de que no sea ruido blanco, verificar el ajuste de N_t y cambiar de modelo.
2. El supuesto de independencia entre $X_{t,1}$ y N_t . Para tal tarea, se observa la función de correlación cruzada muestral entre \hat{Z}_t y \hat{W}_t . Las autocorrelaciones cruzadas para todos los lags deben estar dentro de los intervalos de confianza. En caso de que sea así, empezar de nuevo y ajustar un nuevo modelo de función de transferencia.
3. Finalmente, revisar el supuesto de normalidad del modelo ajustado en conjunto.

2.6. Pronóstico

Para obtener un pronóstico óptimo usando información de ambas series $X_{t,1}$ y $X_{t,2}$, primero se construye un modelo de función de transferencia conectando las series $X_{t,1}$ y $X_{t,2}$ de la forma 2.23, es decir

$$X_{t,2} = v^{-1}(B) \omega(B) X_{t-b,1} + \varphi_N^{-1}(B) \theta_N(B) W_t \quad b \geq 0 \quad (2.39)$$

De forma general, el componente de ruido del modelo 2.39, se asume que es estadísticamente independiente del input $X_{t,1}$. También, se asume que un modelo estocástico adecuado para el input $X_{t,1}$ es

$$X_{t,1} = \varphi_X^{-1}(B) \theta_X(B) Z_t \quad (2.40)$$

Ahora, 2.39 se puede escribir como

$$X_{t,2} = T(B) X_{t,1} + \psi_N(B) W_t \quad (2.41)$$

donde $\psi_N(B) = \frac{\theta_N(B)}{\varphi_N(B)}$ y $T(B) = \frac{\omega(B)}{v(B)} B^b$. Eliminando $X_{t,1}$ de la expresión 2.41 y utilizando 2.40 se tiene

$$X_{t,2} = \sum_{j=0}^{\infty} \alpha_j Z_{t-j} + \sum_{j=0}^{\infty} \psi_j W_{t-j} \quad (2.42)$$

donde $\alpha(B) = T(B) \theta_X(B) \varphi_X^{-1}(B)$.

Nótese que cada límite de la combinación lineal de $\{X_{t,1}, X_{t,2}, -\infty < t \leq n\}$ es un límite de la combinación lineal de $\{Z_t, W_t, -\infty < t \leq n\}$, además $\{Z_t\}$ y $\{W_t\}$ no están correlacionados. Usando 2.42 se tiene

$$\tilde{P}_n X_{n+h,2} = \sum_{j=h}^{\infty} \alpha_j Z_{n+h-j} + \sum_{j=h}^{\infty} \psi_j W_{n+h-j} \quad (2.43)$$

donde $\tilde{P}_n X_{n+h,2}$ es el mínimo error cuadrático medio.

Fijando $t = n + h$ en 2.42 y restando 2.43 se tiene el error cuadrático medio

$$\begin{aligned} \mathbb{E} \left(X_{n+h,2} - \tilde{P}_n X_{n+h,2} \right)^2 &= \sigma_Z^2 \sum_{j=0}^{h-1} \alpha_j^2 + \sigma_W^2 \sum_{j=0}^{h-1} \psi_j^2 \\ &= \sigma^2(h) \end{aligned} \quad (2.44)$$

el cual servirá para calcular los límites de predicción.

Para calcular las predicciones $\tilde{P}_n X_{n+h,2}$ se procede a lo siguiente. Se reescribe 2.41 como

$$A(B) X_{t,2} = B^b U(B) X_{t,1} + V(B) W_t \quad (2.45)$$

donde A , U y V son polinomios de la forma

$$A(B) = 1 - A_1B - \dots - A_qB^q$$

$$U(B) = U_0 + U_1B + \dots + U_uB^u$$

$$V(B) = 1 + V_1B + \dots + V_vB^v$$

Aplicando el operador \tilde{P}_n a la ecuación 2.45 con $t = n + h$, se obtiene

$$\tilde{P}_n X_{n+h,2} = \sum_{j=1}^a A_j \tilde{P}_n X_{n+h-j,2} + \sum_{j=0}^u U_j \tilde{P}_n X_{n+h-b-j,1} + \sum_{j=h}^v V_j W_{n+h-j} \quad (2.46)$$

donde la última suma es cero si $h > v$.

Ya que $\{X_{t,1}\}$ no está correlacionado con $\{W_t\}$, las predicciones que aparecen en la segunda suma en 2.46 son obtenidas al predecir la serie univariada $\{X_{t,1}\}$ (para más detalles ver Brockwell (1991), Sección 13.1).

2.7. Procedimiento alternativo

Otro procedimiento para el paso 3, 4 y 5 es el siguiente:

PASO 3. Se obtiene $\hat{\beta} = (\hat{\omega}_0, \hat{\omega}_1, \dots, \hat{\omega}_q, \hat{v}_1, \dots, \hat{v}_p)$ ajustando la regresión

$$X_{t,2} = \sum_{k=1}^r v_k X_{t-k,2} + \sum_{k=0}^s \omega_k X_{t-b-k,1} + U_t \quad (2.47)$$

donde $U_t = v(B) N_t$.

El modelo de regresión se justifica, ya que, como se sabe, también $X_{t,2} = \frac{\omega(B)B^b}{v(B)}X_{t,1} + N_t$, entonces al multiplicar toda la ecuación por $v(B)$ y desarrollando los polinomios se llega a 2.47. La regresión, mencionada arriba, llega a ser una opción en caso de que no se cuente con un software suficientemente potente para ajustar *directamente*.

Guardar los residuos \hat{U}_t para el siguiente paso.

PASO 4. Aplicar la transformación de promedios móviles $v(B)$ a \hat{U}_t para obtener \hat{N}_t y ajustar un modelo ARIMA al ruido, obteniendo los coeficientes estimados $\hat{\Theta}_N(B)$ y $\hat{\Phi}_N(B)$ en un proceso $\varphi_N(B)N_t = \theta_N(B)W_t$.

PASO 5. Usar mínimos cuadrados para minimizar $\sum W_t^2$ y obtener $\underline{\omega}$, \underline{v} , $\hat{\Phi}_N$, $\hat{\Theta}_N$ reestimados a partir del modelo

$$\varphi_N(B)v(B)X_{t,2} = B^b\varphi_N(B)\omega(B)X_{t,1} + \varphi_N(B)v(B)W_t$$

2.8. Implementación en R e ITSM

2.8.1. Implementación en R

En este caso se usarán dos programas para ocupar los dos tipos de procedimientos existentes, se usará ITSM para el primer procedimiento y R para el procedimiento alternativo. Los datos que se ocuparán son las tasas de crecimiento anual del PIB de México y Brasil, los datos fueron recabados del portal del Banco Mundial. Las tasas fueron recogidas desde el año de 1961 a 2015.

Los datos se muestran en el Cuadro 2.1.

Se comenzará usando software R.

PASO 1.

Se cargan las paqueterías y se grafican los datos.

```
#Establece el directorio trabajo
library(forecast)
library(stats)
library(MASS)
library(TSA)
```

```

library(trend)
setwd("C:/Users/Jesus Pichardo Leyva/Documents/TESIS")

#Lee los archivos .csv y los almacena

datos<-read.table("Crecimiento PIB.csv",sep="," ,header=FALSE)
datos

#Objetos para cada una de las dos series

Xt_1<-datos[,1]
Xt_2<-datos[,2]

#Como series de tiempo

Xt_1<-ts(Xt_1, frequency = 1, start = c(1961, 1))
Xt_2<-ts(Xt_2, frequency = 1, start = c(1961, 1))
plot(Xt_1,type="l",xlab='Time',ylab=expression(italic(X[italic(t_1)])))
plot(Xt_2,type="l",xlab='Time',ylab=expression(italic(X[italic(t_2)])))

```

Como se puede observar en la Figura 2.10 las tasas de crecimiento son similares. A pesar de ser las dos economías más importantes de América Latina, han tenido bajadas y subidas muy pronunciadas debido a crisis económicas internas y externas. La variable output será la tasa de México y la variable input la tasa de Brasil.

Lo primero que se va a hacer será ajustar un modelo ARIMA(p,d,q) al input $\{X_{t,1}\}$ produciendo el proceso $\{Z_t\}$, para conseguir esa tarea se aplicará el ACF y PACF a la serie input.

```

acf(ts(Xt_1,freq=1), lag.max=20, main="ACF Xt_1 ")
pacf(ts(Xt_1,freq=1), lag.max=20, main="PACF Xt_1")

```

El ACF y PACF (Figura 2.11) proponen un modelo ARMA(1, 2). Sin embargo, se ajustará un modelo AR(1) debido a que es mejor modelo de acuerdo a AIC.

```

fit<-arima(Xt_1, order=c(1, 0, 0))
ar1<-as.numeric(fit$coef[1])
ar1
[1] 0.5120055

```

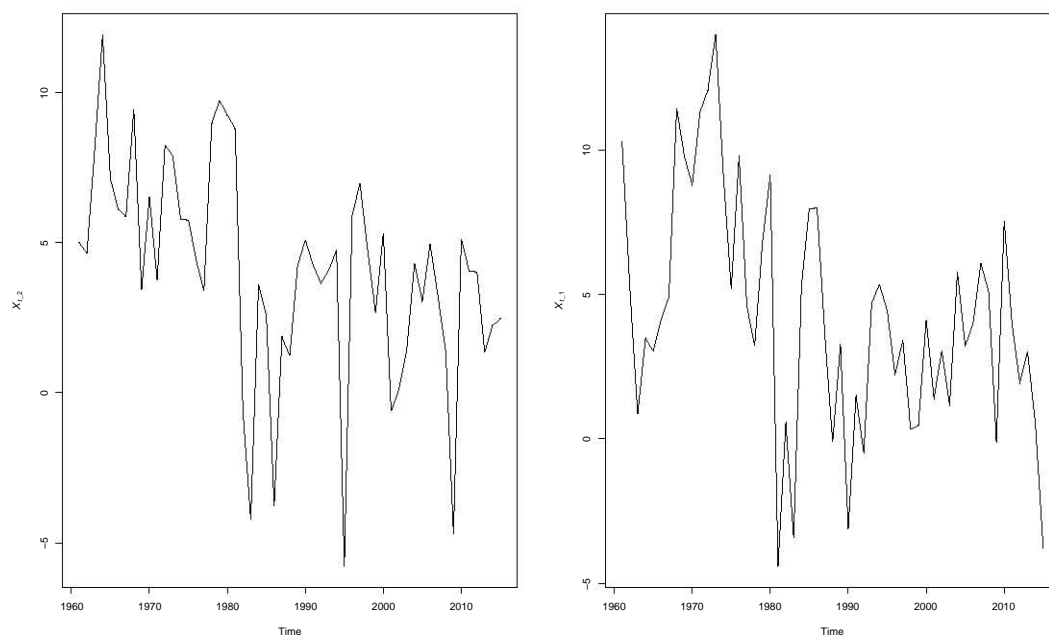


Figura 2.10: Gráficos de las tasas de crecimiento anual del PIB de México y Brasil, respectivamente

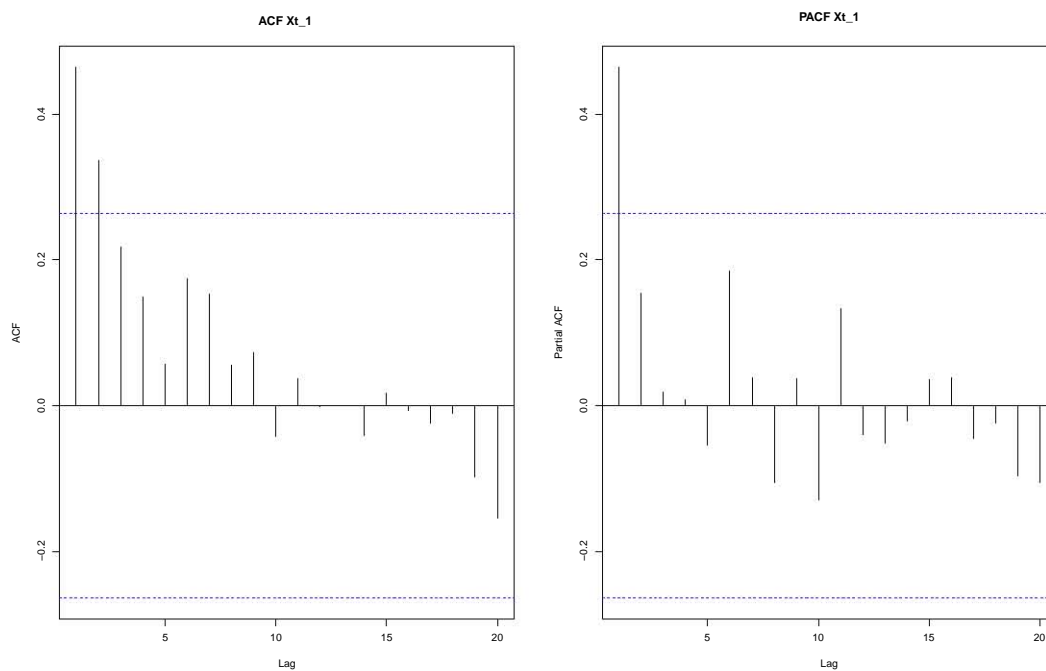


Figura 2.11: ACF y PACF del input

El valor del coeficiente es 0.5120055.

Se guarda los residuos del modelo AR(1) sobre el input (tasa de Brasil), en una nueva variable Xt.Z

```
Xt.Z<-resid(fit)
```

Esta serie de residuos es la que se conoce como \hat{Z}_t . Se tiene el valor del filtro $\pi(B) = (1 - 0.5120055B)$ (pues el polinomio de promedios móviles es 1 al ajustar el proceso como un AR(1)).

Ahora, se aplicará al output el operador $\pi(B)$, para obtener $Y_t = \pi(B) X_{t,2}$, se pretende trabajar sin tendencia en el output, para ello se aplicará la prueba de Mann-Kendall (véase Pohlert (2018)) en el output para saber si existe tendencia en la serie.

```
test<-mk.test(Xt_2)
test1<-mk.test(Xt_1)
summary.trend.test(test)
Mann-Kendall Test

two-sided homogeneity test
H0: S = 0 (no trend)
HA: S != 0 (monotonic trend)

Statistics for total series
  S varS  Z  tau  pvalue
1 -511 18975 -3.7 -0.344 0.0002136
summary.trend.test(test1)
Mann-Kendall Test

two-sided homogeneity test
H0: S = 0 (no trend)
HA: S != 0 (monotonic trend)

Statistics for total series
  S varS  Z  tau  pvalue
1 -416 18974 -3 -0.28 0.0025886
```

Por el momento sólo se tomará en cuenta el output, se observa que el p-value es menor a 0.05, entonces se rechaza la hipótesis nula de no tendencia, entonces

existe tendencia en el output. Se eliminará la tendencia lineal al output a través de una regresión lineal simple con el tiempo y con ese proceso se trabajará.

```
Yt.d<-resid(lm(Xt_2~time(Xt_2), na.action=NULL))
```

Entonces, la serie output con la que se va a trabajar se le aplica el filtro $\pi(B)$. Esto corresponde al paso 1.

Dentro de filter va primero la serie sobre la que se quiere aplicar $\pi(B)$, que es Yt.d en este caso.

```
Yt.fil = filter(Yt.d, filter=c(1,-ar1), method="convolution",sides=1)
```

Ésta es la serie \hat{Y}_t . Ahora, nos fijaremos en las correlaciones cruzadas entre \hat{Y}_t y \hat{Z}_t . Se introduce primero el residuo del output (Y) y luego el del input (Z). En la Figura 2.12 muestra que de cero en adelante son las correlaciones entre Y y Z (en el orden correcto como pide el cálculo). En los valores negativos queda la correlación entre Z y Y. La que interesa es la de 0 en adelante.

```
ccf(Yt.fil,Xt.Z, main="", ylab="Correlaciones cruzadas"
    , na.action=na.omit)
abline(v=0,col='blue')
```

PASO 2 Y 3.

En la Figura 2.12 se observa que, de lag cero en adelante, sólo el lag 5 está fuera de las bandas. Del lado contrario todas están dentro, como debe de ser. De acuerdo al $\hat{T}(B)$ elegido y a la Figura 2.7, se tiene que $b = 5$, $s = 0$, $r = 2$ y hay tendencia tanto en output como en el input, entonces, hay que ajustar la regresión de la siguiente forma (sustituyendo $X_{t,2} = Y_{t,d}$ y $X_{t,1} = X_{t,d}$ dado que hay tendencia)

$$Y_{t,d} = v_1 Y_{t-1,d} + v_2 Y_{t-2,d} + w_0 X_{t-5,d} + U_t$$

$$\text{con } U_t = v(B)N_t = (1 - v_1 B - v_2 B^2)N_t$$

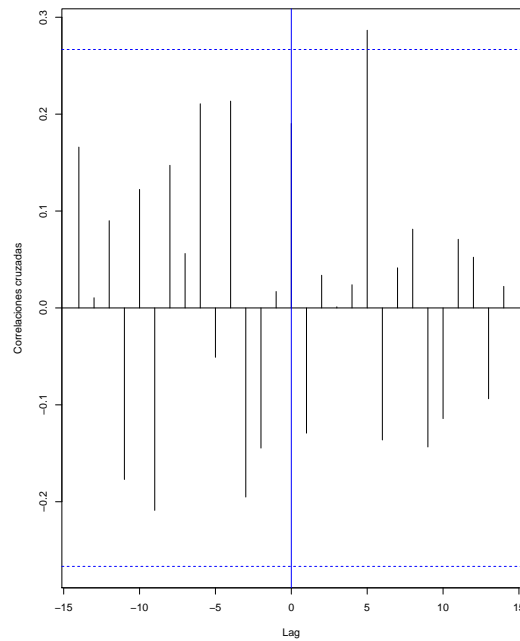


Figura 2.12: Correlaciones cruzadas entre Y y Z

$Y_{t,d}$ y $X_{t,d}$ se obtienen de la siguiente forma

```
Yt.d = resid(lm(Xt_2~time(Xt_2), na.action=NULL))
Xt.d = resid(lm(Xt_1~time(Xt_1), na.action=NULL))
```

Ahora, se creará *base retrasos* uniendo las dos series en los tiempos que tengan en común ambas series, usando los retrasos de que se requieren para el input y output (en este caso) $X_{t,d}$ y $Y_{t,d}$, respectivamente. En el caso del output, se requiere que esté retrasada una y dos unidades, además de la serie sin retraso; mientras que para la serie input, se requiere que esté retrasada 5 unidades. Esta base servirá para ajustar la regresión.

```
base_retrasos = ts.intersect(Yt.d=as.ts(Yt.d), Yt.1.d=lag(as.ts(Yt.d),-1),
  Yt.2.d=lag(as.ts(Yt.d),-2),Xt.5.d=lag(as.ts(Xt.d),-5),dframe=TRUE)
```

Usando la base de datos generada se ajusta la regresión, sin término constante (por eso hay un cero).

```

base.fit <- lm(Yt.d~0+Yt.1.d+Yt.2.d+Xt.5.d, data=base_retrasos)
summary(base.fit)

Call:
lm(formula = Yt.d ~ 0 + Yt.1.d + Yt.2.d + Xt.5.d, data = base_retrasos)

Residuals:
    Min       1Q   Median       3Q      Max
-7.3477 -1.0551  0.0201  1.8735  5.1631

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
Yt.1.d     0.1319     0.1393   0.947  0.3483
Yt.2.d    -0.1273     0.1337  -0.952  0.3458
Xt.5.d     0.2997     0.1181   2.538  0.0145 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.032 on 47 degrees of freedom
Multiple R-squared:  0.158, Adjusted R-squared:  0.1042
F-statistic: 2.939 on 3 and 47 DF,  p-value: 0.04271

```

El valor de v_1 se estimaría como 0.1319, v_2 como -0.1273 y el de ω_0 como 0.2997.

PASO 4.

Como $U_t = v(B)N_t$, hay que obtener el N_t (despejarlo) donde $v(B) = 1 - v_1B - v_2B^2$, así que aprovechamos el valor de v_1 y v_2 obtenidos arriba.

```

v1 = as.numeric(base.fit$coef[1])
v1
[1] 0.1319318
v2 = as.numeric(base.fit$coef[2])
v2
[1] -0.1272911

```

El proceso $U_t = v(B)N_t = (1 - v_1B - v_2B^2)N_t$ es como un proceso AR(2), así como arriba, se usa `filter()` para poder estimar N_t , o sea despejamos N_t como $N_t = v(B)^{-1}U_t$ y entonces `filter` es sobre los residuos \hat{U}_t obtenidos de la regresión de arriba.

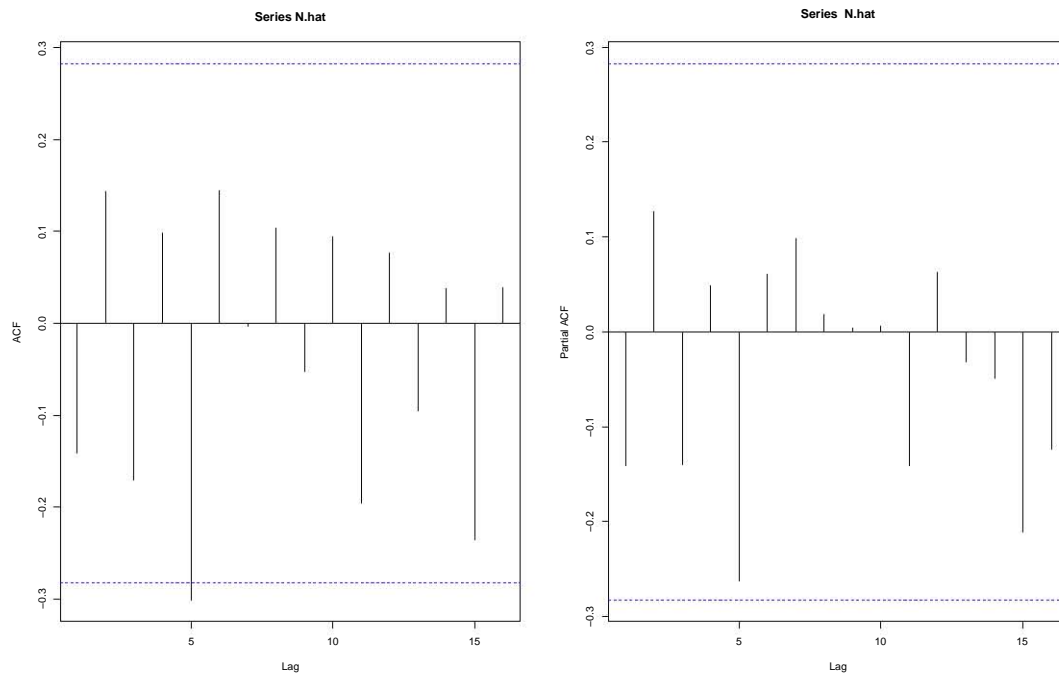


Figura 2.13: ACF y PACF de N_t ajustada

```
N.hat = filter(resid(base.fit), filter=c(1,-v1,-v2),
  method="convolution",sides=1)
```

Ahora, se observa el ACF y PACF del proceso N_t estimado (primero se omiten los NA's en caso de existir).

```
N.hat<-na.omit(N.hat)
```

```
acf(N.hat)
pacf(N.hat)
```

Tomando en cuenta la Figura 2.13 se ajustará un modelo MA(5) a N_t .


```
N.fit <- arima(N.hat, order=c(0,0,5),include.mean = F)
N.fit
```

Call:

```
arima(x = N.hat, order = c(0, 0, 5), include.mean = F)
```

Coefficients:

	ma1	ma2	ma3	ma4	ma5
	-0.0461	0.1213	-0.0697	0.0869	-0.4240
s.e.	0.1478	0.1496	0.1712	0.1144	0.2374

sigma² estimated as 7.733: log likelihood = -117.72, aic = 245.44

PASO 5 y 6.

Finalmente, se ajustará el modelo conjunto y se verificarán el supuesto de no autocorrelación de los residuos, independencia entre Z_t (input) y N_t y normalidad de los residuos.

Para terminar, el ajuste conjunto y pronóstico.

```
ts.xt<-ts(Xt_1)
lag5.x<-lag(ts.xt,-5)
ts.yt<-ts(Xt_2)
dat3<-cbind(ts.xt,lag5.x,ts.yt)
dimnames(dat3)[[2]]<-c("xt","lag5x","yt")
data2<-na.omit(as.data.frame(dat3))
modint<-arimax(data2$yt, order=c(0,0,5), xtransf=data.frame(data2$lag5x),
               transfer=list(c(2,0)), include.mean = FALSE)
modint
Call:
arimax(x = data2$yt, order = c(0, 0, 5), include.mean = FALSE,
       xtransf = data.frame(data2$lag5x), transfer = list(c(2, 0)))
```

Coefficients:

	ma1	ma2	ma3	ma4	ma5	data2.lag5x-AR1
	0.4922	0.0597	-0.1692	0.2324	-0.3650	0.0260
s.e.	0.1573	0.1515	0.1831	0.1525	0.1455	0.1288
	data2.lag5x-AR2		data2.lag5x-MA0			
	0.2445		0.5244			
s.e.	0.1441		0.0938			

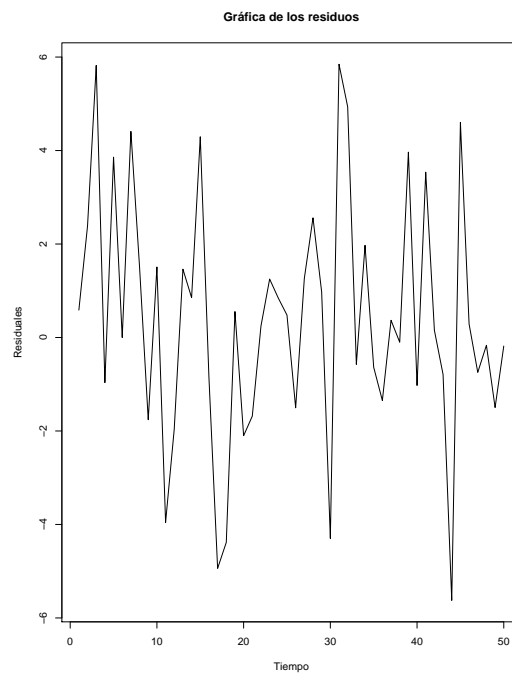


Figura 2.14: Gráfica de los residuos del modelo ajustado en conjunto

σ^2 estimated as 7.175: log likelihood = -122.78, aic = 261.56

```
residuos<-modint$residuals
plot(residuos,xlab="Tiempo",ylab="Residuales",
      main="Grafica de los residuos",pch=19)
acf(ts(residuos,freq=1), lag.max=20, main="ACF de los residuos")
pacf(ts(residuos,freq=1), lag.max=20, main="PACF de los residuos")
modint.res<-as.vector(residuals(modint))
Zt<-as.vector(Xt.Z)
ccf(modint.res,Zt,main='CCF de Z(t) y residuales de TF-N')
abline(v=0,col='blue')
hist(residuos,xlab = "Residuos",ylab = "Frecuencia",
      main = "Histograma de los residuos",pch=19)
residuost<-rstandard(modint)
qqnorm(residuost,ylab = "Standardized residuals",xlab = "Normal Scores",
        main = "Plot probabilidad Normal")
qqline(residuost)
```

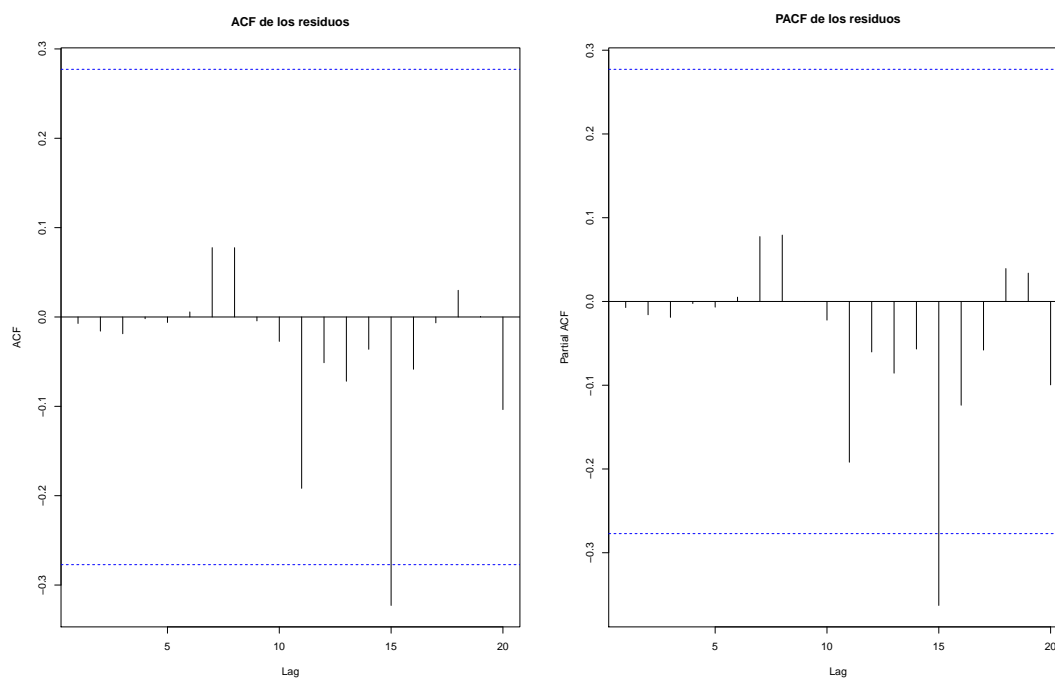


Figura 2.15: ACF y PACF de los residuos del modelo ajustado conjunto

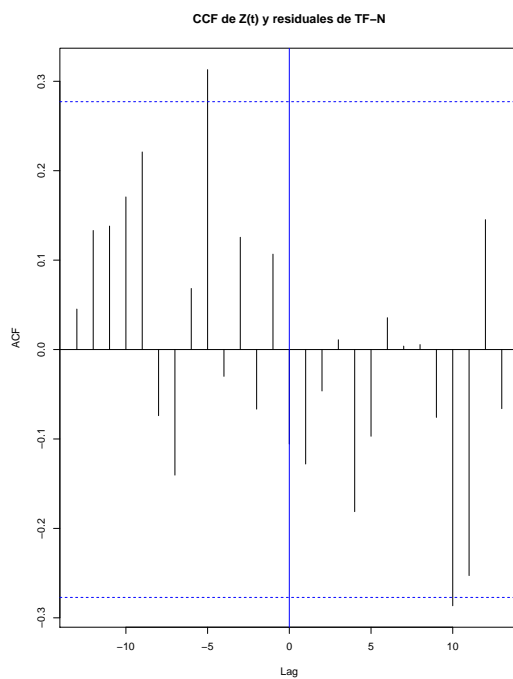


Figura 2.16: Función de Correlación Cruzada entre Z_t y N_t

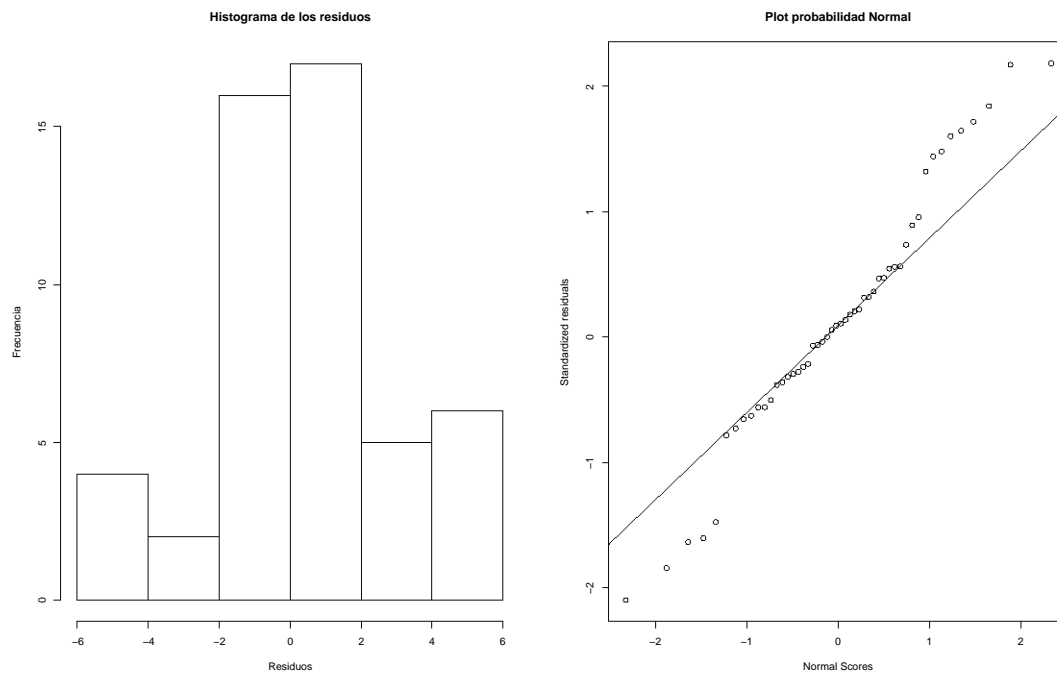


Figura 2.17: Histograma y q-q plot normal de los residuos del modelo conjunto

La Figura 2.14 muestra los residuos del modelo ajustado final. Al ver la Figura 2.15 se observa que no hay autocorrelación en los residuos.

Al observar la Figura 2.16 se puede afirmar que hay independencia, ya que no son graves las *salidas* de dos lags. Por otro lado, la Figura 2.17 muestra histograma de los residuos y q-q plot, el cual se puede observar que los residuos tienen cierto comportamiento normal, habrá que poner atención en las colas. Para ello, se usarán las pruebas de Lilliefors y Shapiro-Wilks.

```
library(nortest)
lillie.test(residuos)
```

Lilliefors (Kolmogorov-Smirnov) normality test

```
data: residuos
D = 0.097828, p-value = 0.2711
```

```
shapiro.test(residuos)
```

Shapiro-Wilk normality test

```
data:  residuos
W = 0.97239, p-value = 0.2888
```

En ambas pruebas el p-value es mayor a 0.05, entonces los residuos son normales. El modelo conjunto ajustado es adecuado y listo para realizar el pronóstico.

```
tau<-5
Xt.pronostico<-predict(fit,n.ahead=tau)

Xt.new<-c(Xt_1,fit$mean)
modint.res<-as.vector(residuals(modint))
res.tf.new<-c(rep(0,5),modint.res,rep(0,tau))
Yt.new<-c(Xt_2,rep(0,tau))

theta1<-modint[[1]][1]
theta1
      ma1
0.4921825
theta2<-modint[[1]][2]
theta2
      ma2
0.05965816
theta3<-modint[[1]][3]
theta3
      ma3
-0.1691669
theta4<-modint[[1]][4]
theta4
      ma4
0.2324412
theta5<-modint[[1]][5]
theta5
      ma5
-0.3650116
v1<-modint[[1]][6]
v1
data2.lag5x-AR1
      0.02599989
v2<-modint[[1]][7]
```

```

v2
data2.lag5x-AR2
    0.2445153
w0<-modint[[1]][8]
w0
data2.lag5x-MA0
    0.5243913

L<-length(Xt_2)
for(i in (L+1):(L+tau)){
  Yt.new[i]<-v1*Yt.new[i-1]+v2*Yt.new[i-2]+w0*Xt.new[i-5]
  +res.tf.new[i]+(theta1-v1)*res.tf.new[i-1]
  +(-(v1*theta1)+theta2-v2)*res.tf.new[i-2]
  +(-(v1*theta2)+theta3-(v2*theta1))*res.tf.new[i-3]
  +(-(v1*theta3)+theta4-(v2*theta2))*res.tf.new[i-4]
  +(-(v1*theta4)+theta5-(v2*theta3))*res.tf.new[i-5]
  -((v1*theta5)+(v2*theta4))*res.tf.new[i-6]
  -(v2*theta5)*res.tf.new[i-7]
}

#Limites de prediccion

phix<-fit[[1]][1]
phix
    ar1
0.5120055
t0star<-w0
tstar<-c(t0star)
psi<-phix^(0:(tau-1))
sig2.z<-fit$sigma2
sig2.z
[1] 12.41854
sig2.err<-modint$sigma2
sig2.err
[1] 7.174541
sig2.tfn<-rep(0,5)
b<-5
for(i in 1:5){
  if((i-b)<=0){

```

```

    sig2.tfn[i]<-sig2.err*sum(psi[1:i]^2)
  }
  else{
    sig2.tfn[i]<-sig2.z*sum(tstar[1:(i-b)]^2)
    +sig2.err*sum(psi[1:i]^2)
  }
}

plot(Yt.new[1:L],type="l",pch=16,cex=0.5,xlab='Time',
      ylab='PIB Mxico',xlim=c(1,60),ylim = c(-10,15))
lines(56:60,Yt.new[56:60],col="blue")
lines(56:60,Yt.new[56:60]+1.96*sqrt(sig2.tfn),col="red",lty=2)
lines(56:60,Yt.new[56:60]-1.96*sqrt(sig2.tfn),col="red",lty=2)
legend("bottomleft", legend = c("Xt_2", "Pronostico","Ints.Confianza (95%)"),
      lty = rep(1,3), col = c("black", "blue", "red"), bty = "n")
Yt.new[56:60]
[1] 2.6959524 1.6784333 2.2760144 0.7317743 -1.4014088

```

Para empezar tau representa el número de valores que se van a pronosticar, se definieron dos nuevas variables `res.tf.new` y `Yt.new` que agregarán ceros al final de los residuos del modelo ajustado en conjunto y a $X_{t,2}$, esos nuevos ceros agregados son en igual cantidad a los valores que se van a pronosticar, esto con el fin de usar el `for` programado. Posteriormente con esa información, se definen los límites de predicción que servirán para construir los intervalos de confianza. Finalmente, se obtienen los pronósticos y los intervalos de confianza al 95 por ciento.

El gráfico del pronóstico es la Figura 2.18. Se ve que hay un ligero crecimiento y luego decrece, si se observan detenidamente los pronósticos son acertados, la economía mexicana anualmente en promedio crece al 2 por ciento.

2.8.2. Implementación en ITSM

Para el caso de ITSM se utilizará la misma base de datos que en R.

PASO 1.

Lo primero que se hará es graficar la variable input $X_{t,1}$ (tasa de crecimiento del PIB de Brasil).

File > Project > Open > Univariate > PIB Brasil.CSV

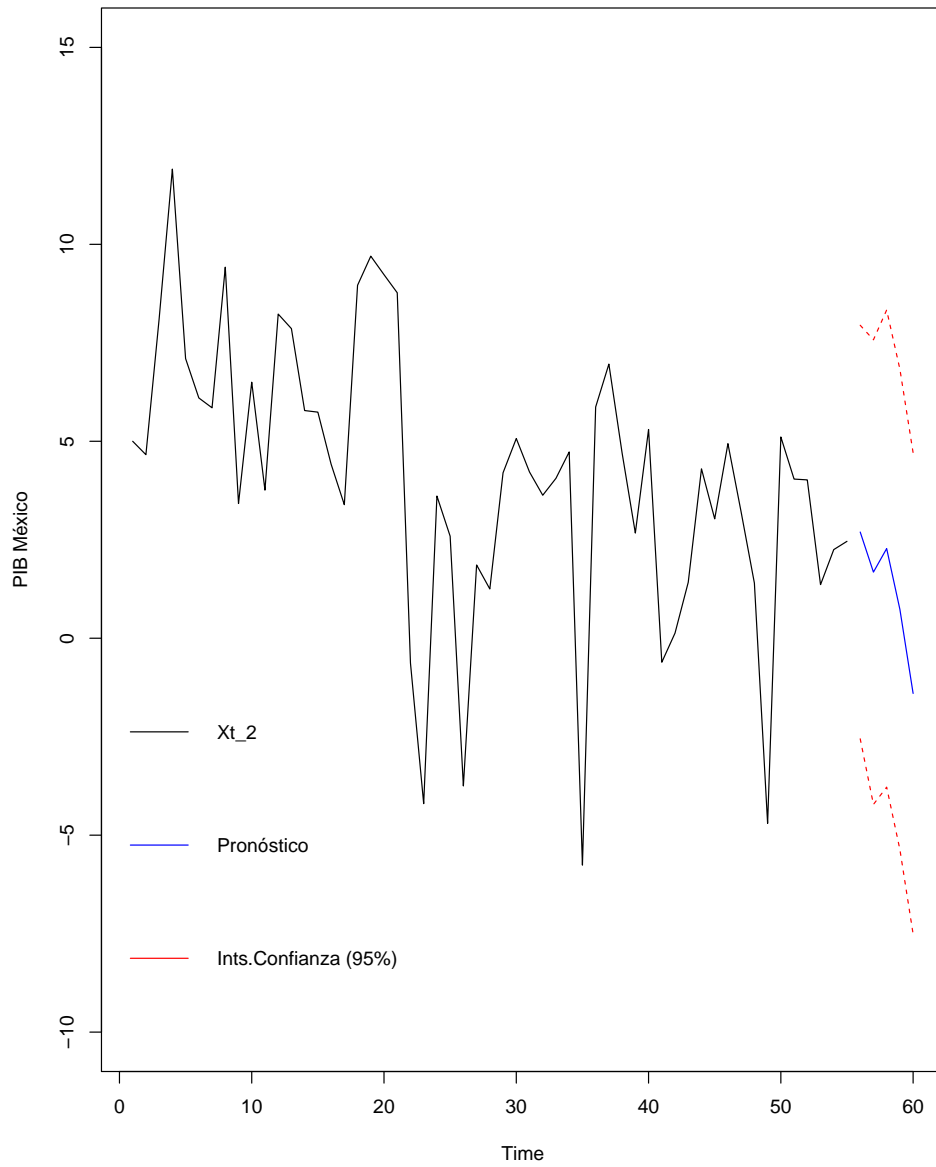


Figura 2.18: Pronóstico de la tasa de crecimiento del PIB de México

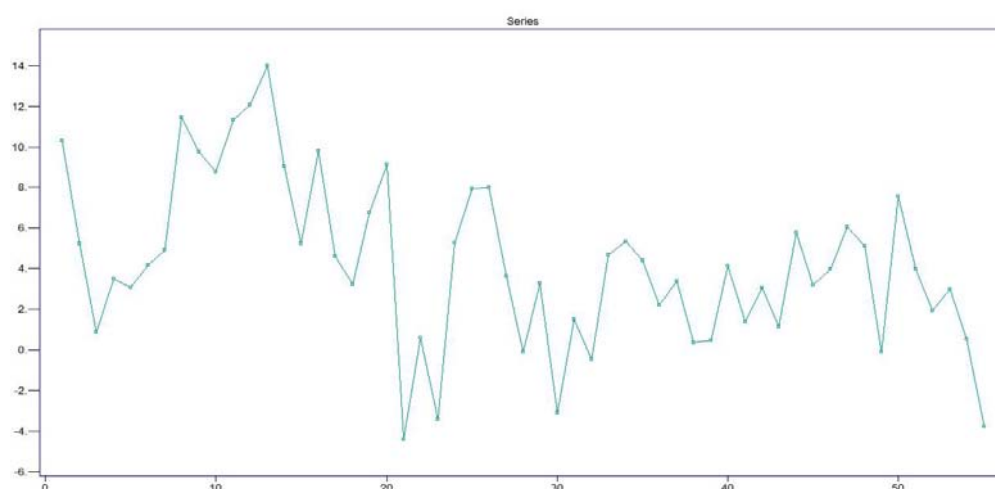


Figura 2.19: Gráfica de la tasa de crecimiento del PIB de Brasil

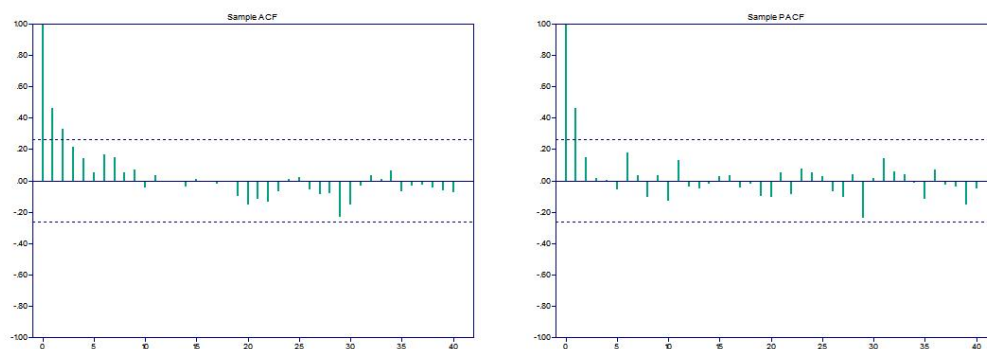


Figura 2.20: ACF y PACF de la tasa de crecimiento de la economía de Brasil

En la gráfica 2.19 se puede observar el gráfico de los porcentajes de crecimiento de la economía brasileña. Lo que se quiere hacer, como en R, es ajustar un proceso ARIMA a éstos datos por máxima verosimilitud y obtener los residuales del modelo, que en este caso será el proceso Z_t . Ahora, se obtendrán el ACF y PACF muestral del input, de este modo se podrá elegir un modelo adecuado.

Statistics > ACF/PACF > Sample

La Figura 2.20 propone un modelo ARMA(1,2). Sin embargo, se repetirá lo hecho en R y se ajustará un AR(1).

```
Model > Specify
Model > Estimation > Preliminary
Model > Estimation > Max Likelihood
```

ARMA Model:

$$X(t) = .5120 X(t-1) \\ + Z(t)$$

WN Variance = 12.418878

AR Coefficients
.511966

Standard Error of AR Coefficients
.122286

(Residual SS)/N = 12.4189

AICC = .299175E+03
BIC = .298500E+03
FPE = 12.878836

-2Log(Likelihood) = .294944E+03

Accuracy parameter = .00465000

Number of iterations = 2

Number of function evaluations = 14

Optimization stopped within accuracy level.

Se observa que el valor ϕ_1 es igual a 0.5120 como en R. De este modo, se obtienen los residuos Z_t

```
File > Export > Residuals > Clipboard
```

Después, se abre la tasa de crecimiento del PIB mexicano y con el mismo modelo que con Brasil se obtienen los residuos del output. Los residuos de ambos

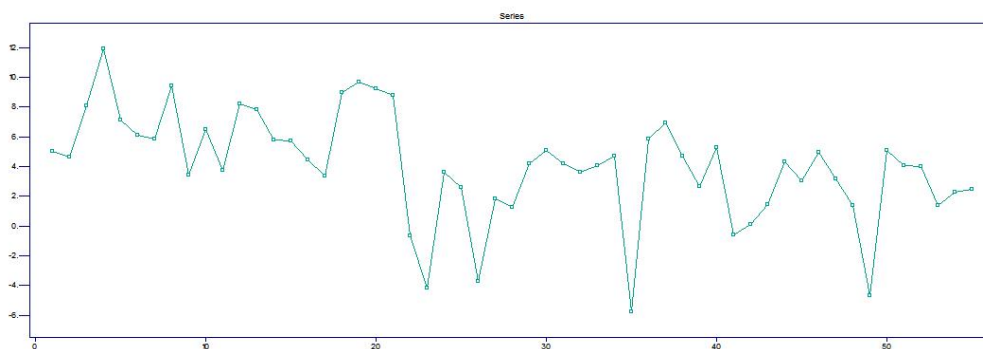


Figura 2.21: Gráfica de la tasa de crecimiento del PIB de México

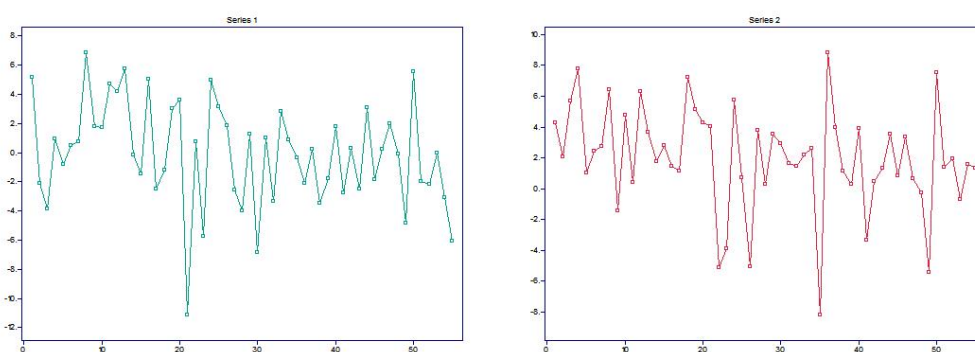


Figura 2.22: Residuos de ambas series usando el mismo modelo ajustado

modelos se pegan en un Excel en formato CSV (en este caso el documento se llama residuos) y se abren en el paquete estadístico.

```
File > Import File > PIB Mexico.CSV
File > Export > Residuals > Clipboard
File > Project > Open > Multivariate > Residuos.CSV
```

En la Figura 2.22 se tienen los residuos de ambas series, en este caso son \hat{Z}_t y \hat{Y}_t , respectivamente.

Posteriormente, se calculará la función de correlación cruzada de la siguiente forma

```
Statistics > Cross-correlation > Sample
```

=====

ITSM::Multivariate(CCF)

=====

of Lags = 20

Sample Cross-correlations:

Cor[X_1(t+h),X_1(t)]

1.0000	-.0907	.1226	.0350	.0773
-.1113	.1148	.1204	-.0625	.0987
-.1538	.0960	-.0554	.0416	-.0828
.0697	-.0175	-.0189	.0803	-.0583
-.0891				

Cor[X_1(t+h),X_2(t)]

.2514	.0654	-.0859	-.1189	.2627
.0163	.2637	.1240	.1876	-.1540
.1595	-.1356	.1092	.0112	.1543
-.1145	.1394	-.0430	-.1711	-.0281
-.0174				

Cor[X_2(t+h),X_1(t)]

.2514	-.0814	.0957	.0819	.0437
.2992	-.0869	.1011	.0744	-.0842
-.0914	.1275	.0951	-.0638	.0545
-.3032	-.1092	.0992	.0218	-.0173
.1698				

Cor[X_2(t+h),X_2(t)]

1.0000	-.1445	-.0534	.0184	.1195
-.1716	.2046	.0369	.0046	.0868
-.0014	-.1097	.1365	.0370	.0280
-.0733	.0651	.0386	-.2966E-05	-.0026
-.0399				

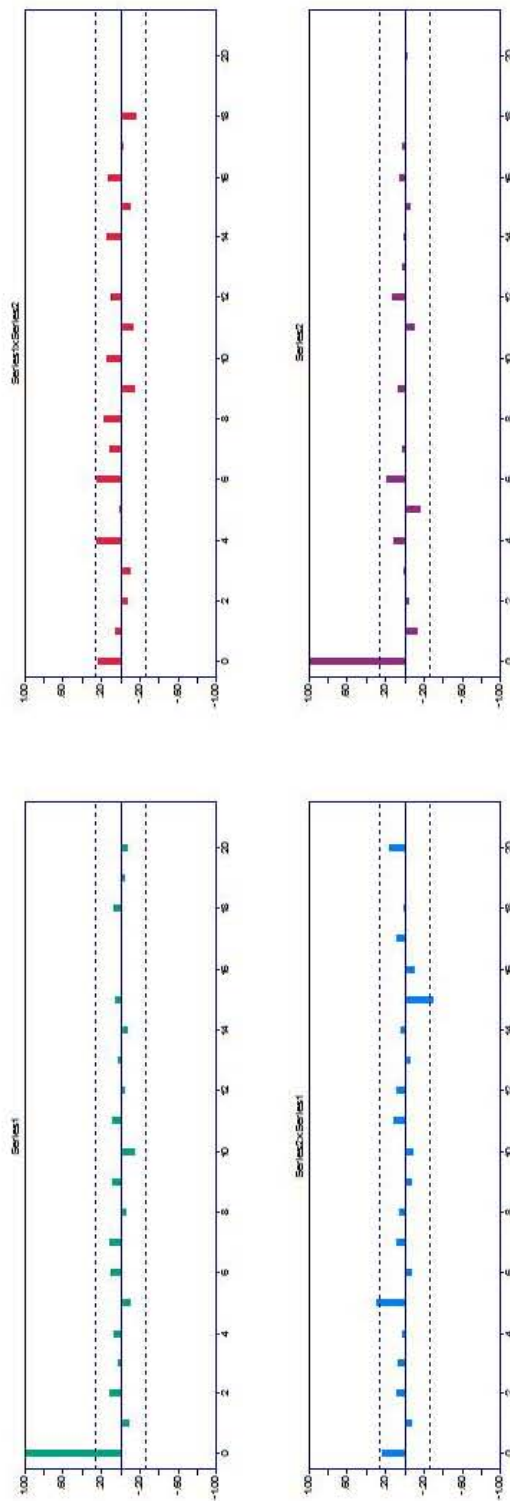


Figura 2.23: Función de correlación aplicada \hat{Z}_t y \hat{Y}_t

Se observa en la Figura 2.23 que hay cuatro cuadros de correlación. Las series 1 y 2 (verde y morado, respectivamente) se refieren a la correlación por separado de los residuos de Y_t y Z_t , en estas dos gráficas sólo se puede comprobar que las series $X_{t,1}$ y $X_{t,2}$ son ruido blanco. Por otro lado, las otras dos gráficas son las que interesan, lo que hace R en un gráfico ITSM lo hace en dos; Serie2xSerie1 es la correlación cruzada entre $\{Y_t\}$ y $\{Z_t\}$ y es la que interesa, si se observa como en R sería la correlación cruzada de 0 en adelante y Serie1xSerie2 es la es la correlación cruzada entre $\{Z_t\}$ y $\{Y_t\}$ (de 0 hacia atrás en R). Se observa el gráfico que interesa (Serie2xSerie1) y se ve que que $b = 5$ será el valor del retraso que se tomará y se hará lo mismo que en R donde $s = 0$ y $r = 2$, esto debido a la Figura 2.7.

PASO 2 y 3.

Ahora, se abrirá un CSV con las series originales que son $X_{t,1}$ (Brasil) y $X_{t,2}$ (México) y se especificará el modelo.

```
File > Project > Open > Multivariate > Crecimiento PIB.CSV
Transfer > Specify model > Transfer model
Transfer > Estimation
```

```
=====
ITSM::(Transfer Function Model Estimates)
=====
```

$$X(t,2) = T(B) X(t,1) + N(t)$$

$$T(B) = \frac{B^5(.2900)}{(1 -.1150 B^1 -.4600 B^2)}$$

$$X(t,1) = Z(t)$$

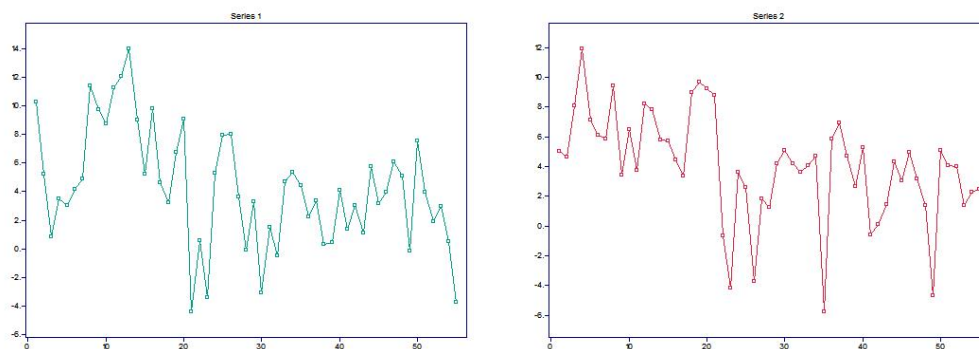
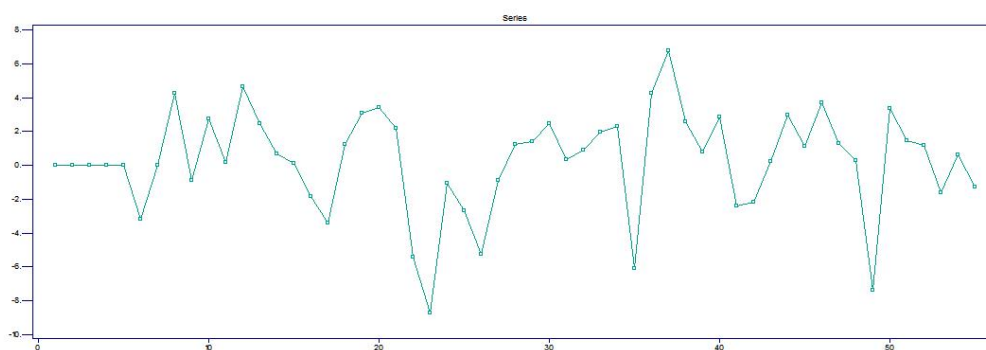
{Z(t)} is WN(0,1.000000)

$$N(t) = W(t)$$

{W(t)} is WN(0,10.743982)

AICC = .231445E+04

Accuracy Parameter .010000

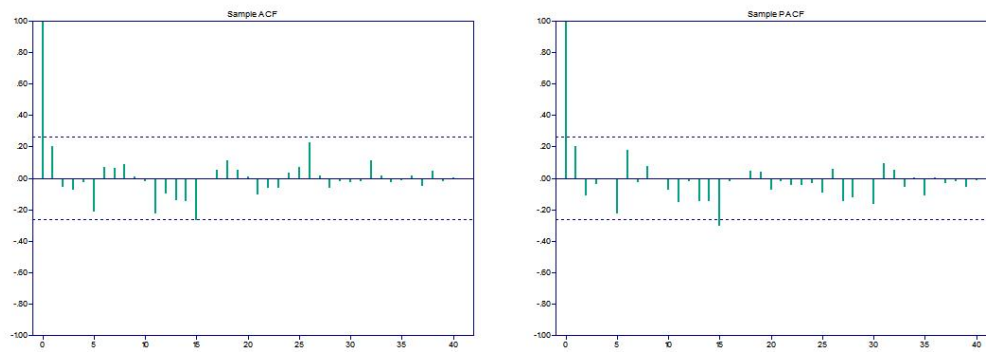
Figura 2.24: Series $X_{t,1}$ y $X_{t,2}$ Figura 2.25: N_t **PASO 4.**

Se extraerán los residuos del modelo (N_t) y se guardarán en un formato tsm con el nombre $N_t.tsm$

```
Transfer > Residual Analysis > Plot Residuals
File > Export > Residuals > File
File > Project > Open > Univariate > Nt.tsm
```

Ahora, se va a aplicar el ACF y PACF a N_t para ajustarle un modelo.

```
Statistics > ACF/PACF > Sample
```

Figura 2.26: ACF y PACF de N_t

De acuerdo a la Figura 2.26 no hay ningún lag que se salga. Sin embargo, se le ajustará un MA(5) a N_t como en R.

```
Model > Specify
Model > Estimation > Preliminary
Model > Estimation > Max Likelihood
```

```
=====
ITSM::(Maximum likelihood estimates)
=====
```

Method: Maximum Likelihood

ARMA Model:

$$X(t) = Z(t) + .3572 Z(t-1) - .05042 Z(t-2) - .2080 Z(t-3) \\ + .1913 Z(t-4) - .4670 Z(t-5)$$

WN Variance = 6.955763

MA Coefficients

.357216	-.050421	-.208037	.191330
-.466962			

Standard Error of MA Coefficients

.119236	.125983	.123008	.125983
.119236			

(Residual SS)/N = 6.95576

AICC = .281877E+03

BIC = .273176E+03

-2Log(Likelihood) = .268127E+03

Accuracy parameter = .000560000

Number of iterations = 23

Number of function evaluations = 213

PASO 5.

Los valores son similares a los conseguidos en R. Ahora, regresamos a la serie bivariada y se llenará cada parte de la especificación del modelo.

Transfer > Specify model > Input Model > Input AR(1) e Input MA(0)
 Transfer > Specify model > Transfer Model > Denominator (2) y Numerator (0)
 Transfer > Specify model > Residual Model > Residual AR(0) y Residual MA(5)
 Transfer > Estimation

```
=====
ITSM::(Transfer Function Model Estimates)
=====
```

$$X(t,2) = T(B) X(t,1) + N(t)$$

$$T(B) = \frac{B^5(.4208)}{(1 - .1379 B^1 - .2766 B^2)}$$

$$X(t,1) = .0000 X(t-1,1) + Z(t)$$

{Z(t)} is WN(0,1.000000)

$$N(t) = W(t) + .2907 W(t-1) + .03370 W(t-2) - .07435 W(t-3) + .1325 W(t-4) - .3130 W(t-5)$$

{W(t)} is WN(0,10.983497)

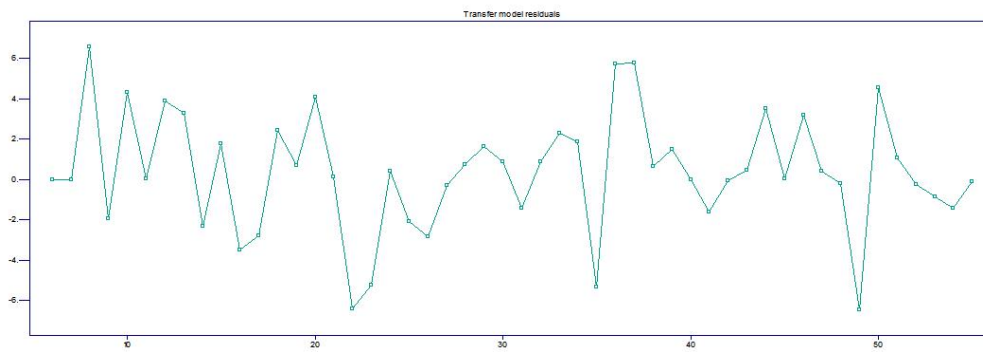


Figura 2.27: Residuos del modelo ajustado conjunto

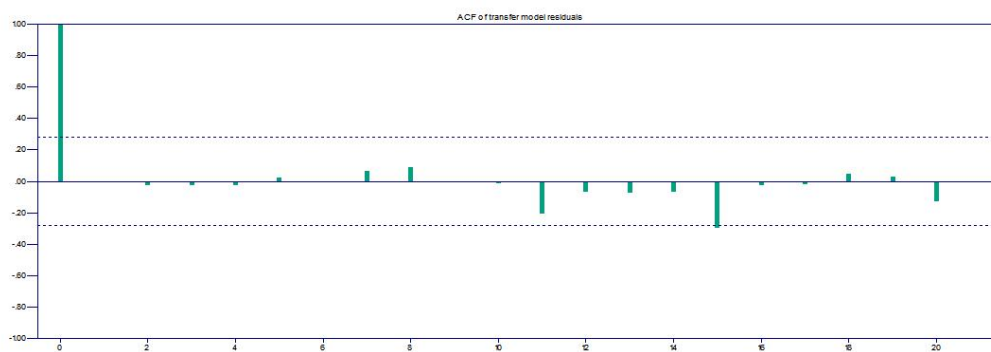


Figura 2.28: ACF de los residuos del modelo ajustado conjunto

AICC = .230807E+04

Accuracy Parameter .100000

Las estimaciones de R e ITSM no se alejan mucho, son muy cercanas entre sí.

Por último, a estudiar los residuos del ajuste conjunto y el pronóstico.

```
Transfer > Residual Analysis > Plot Residuals
```

```
Transfer > Residual Analysis > Plot ACF
```

```
Transfer > Residual Analysis > Plot Residuals
```

De acuerdo con las Figuras 2.27 y 2.28 los residuos del modelo final son ruido blanco. Entonces, el modelo es bueno. El pronóstico (Figura 2.29) se obtiene de la siguiente manera

```
Forecasting > Transfer function
```

Time	Prediction	sqrt(MSE)	Approximate 95 Percent Prediction Bounds	
			Lower	Upper
56	2.26975	3.31414	-4.22584	8.76534
57	2.42822	3.45134	-4.33628	9.19271
58	2.42188	3.45314	-4.34616	9.18992
59	1.60959	3.46193	-5.17566	8.39484
60	-.73894	3.48965	-7.57853	6.10065

Los valores pronosticados en ITSM son muy similares a los obtenidos en R.

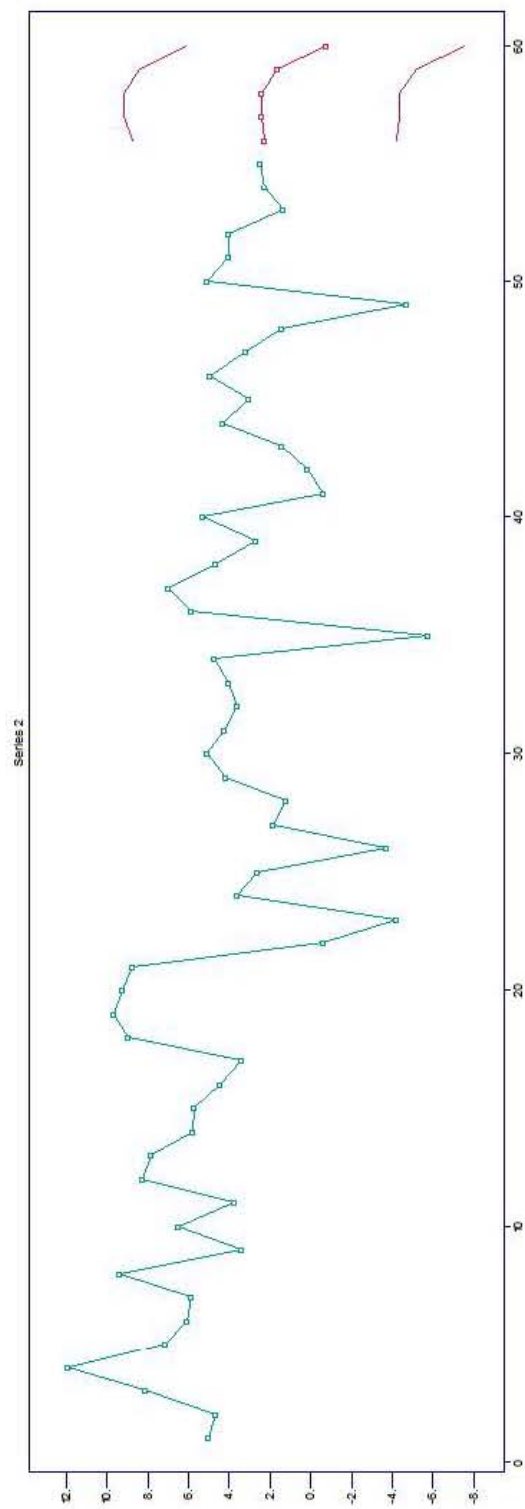


Figura 2.29: Pronóstico a 5 años de la tasa de crecimiento del PIB de México

AÑO	México	Brasil
1961	5	10.28
1962	4.66	5.22
1963	8.11	0.87
1964	11.91	3.49
1965	7.1	3.05
1966	6.1	4.15
1967	5.85	4.92
1968	9.42	11.43
1969	3.42	9.74
1970	6.5	8.77
1971	3.76	11.3
1972	8.23	12.05
1973	7.86	13.98
1974	5.78	9.04
1975	5.74	5.21
1976	4.42	9.79
1977	3.39	4.61
1978	8.96	3.23
1979	9.7	6.77
1980	9.23	9.11
1981	8.77	-4.39
1982	-0.63	0.58
1983	-4.2	-3.41
1984	3.61	5.27
1985	2.59	7.95
1986	-3.75	7.99
1987	1.86	3.6
1988	1.25	-0.1
1989	4.2	3.28
1990	5.07	-3.1
1991	4.22	1.51
1992	3.63	-0.47
1993	4.06	4.67
1994	4.73	5.33
1995	-5.76	4.42
1996	5.87	2.21
1997	6.96	3.4
1998	4.7	0.34
1999	2.67	0.47
2000	5.3	4.11
2001	-0.61	1.39
2002	0.13	3.05
2003	1.42	1.14
2004	4.3	5.76
2005	3.03	3.2
2006	4.94	3.96
2007	3.2	6.07
2008	1.4	5.09
2009	-4.7	-0.13
2010	5.11	7.53
2011	4.04	3.97
2012	4.02	1.92
2013	1.36	3
2014	2.25	0.5
2015	2.46	-3.77

Cuadro 2.1: Tasas de crecimiento anual del PIB de México y Brasil

Capítulo 3

Análisis de Intervención

3.1. Introducción

El Análisis de Intervención es una herramienta estadística usada para evaluar el impacto de un evento especial, anormal o externo en una serie de tiempo determinada; tal evento podría ser:

- Cambio en la política económica.
- Crisis financiera.
- La aparición de una nueva ley ó regulación.
- Vacaciones, aumentos salariales.
- Desastres naturales: terremotos, huracanes, tornados, sequías, inundaciones.
- Guerras, huelgas, terrorismo, movimientos socio-políticos, etc.
- Algún hecho anormal en procesos químicos o físicos.

Es importante mencionar que el evento ó *intervención* (T) se asume conocido.

La técnica de Análisis de Intervención que aquí se presenta, está basada en el artículo de Box y Tiao (1975), en el cual aparece, por primera vez este tipo de metodología.

3.2. Análisis de Intervención: teoría

El modelo es igual al de Función de Transferencia

$$Y_t = \sum_{j=0}^{\infty} \tau_j X_{t-j} + N_t \quad (3.1)$$

donde $\{N_t\}$ se le asocia con un modelo ARIMA, que representa la parte estocástica de la serie y el cual puede expresarse como

$$\begin{aligned} \phi_N(B) (1 - B)^d N_t &= \theta_N(B) W_t \\ \varphi_N(B) N_t &= \theta_N(B) W_t \end{aligned} \quad (3.2)$$

Las raíces de $\theta_N(B)$ están fuera del círculo unitario y $\varphi_N(B)$ pueden estar sobre o fuera del círculo unitario.

$\{X_t\}$ no es una función aleatoria, sino una función determinística de t . Calculando la esperanza a la expresión 3.1, se tiene

$$\begin{aligned} \mathbb{E}(Y_t) &= \sum_{j=0}^{\infty} \tau_j X_{t-j} + \mathbb{E}(N_t) \\ &= \sum_{j=0}^{\infty} \tau_j X_{t-j} \end{aligned} \quad (3.3)$$

En otras palabras, la media del proceso $\{Y_t\}$ está dada por $\sum_{j=0}^{\infty} \tau_j X_{t-j}$. Se eligen $\{\tau_t\}$ y $\{X_t\}$ de tal modo que el cambio de nivel esté bien representado por $\sum_{j=0}^{\infty} \tau_j X_{t-j}$.

Si la serie tiene $\mathbb{E}(Y_t) = 0$ para $t \leq T$ y $\mathbb{E}(Y_t) \rightarrow 0$ cuando $t \rightarrow \infty$, lo cual quiere decir que la serie está en cero antes de T , después salta y a futuro recupera su nivel original, usar como input $\{X_t\}$

$$X_t = P_t(T) = \begin{cases} 1 & \text{si } t = T \\ 0 & \text{si } t \neq T \end{cases} \quad (3.4)$$

conocida como **función pulso**.

En este ejemplo, $\sum_{j=0}^{\infty} \tau_j X_{t-j} = \tau_{t-T} X_T$ pues sólo si $j = t - T$, $X_{t-j} \neq 0$. O sea, se tiene una constante τ_{t-T} multiplicado por X_T y hay un salto en T .

En una serie para la cual $\mathbb{E}(Y_t) = 0$ para $t \leq T$ y $\mathbb{E}(Y_t) \rightarrow a \neq 0$ para $t \rightarrow \infty$ (la serie oscila alrededor del cero y luego salta a “ a ” y no regresa). Se usa de input:

$$X_t = S_t(T) = \sum_{k=T}^{\infty} P_t(k) = \begin{cases} 1 & \text{si } t \geq T \\ 0 & \text{si } t < T \end{cases} \quad (3.5)$$

a ésta se le conoce como **función escalón**.

Notar que la función pulso puede ser obtenida diferenciando la función escalón, es decir, $P_t(T) = (1 - B) S_t(T) = S_t(T) - S_{t-1}(T)$.

El objetivo del análisis de intervención es estimar el efecto de intervención dado por $\sum_{j=0}^{\infty} \tau_j X_{t-j}$ y usar el modelo resultante

$$Y_t = \sum_{j=0}^{\infty} \tau_j X_{t-j} + N_t \quad (3.6)$$

para pronosticar.

Usualmente las funciones elegidas son del tipo 3.4, 3.5 ó variantes; eso dependerá del criterio del investigador.

El operador $T(B) = \sum_{j=0}^{\infty} \tau_j B^j$ usado para obtener $\sum_{j=0}^{\infty} \tau_j X_{t-j} = \left(\sum_{j=0}^{\infty} \tau_j B^j \right) X_t$ es simplificado aproximando el operador $T(B)$ del filtro lineal $\{\tau_j\}$ con un operador de razón

$$T(B) = \frac{\omega(B)}{v(B)} B^b \quad (3.7)$$

con b parámetro de tardanza,

$$\omega(B) = \omega_0 + \omega_1 B + \cdots + \omega_s B^s$$

y

$$v(B) = 1 - v_1 B - v_2 B^2 - \cdots - v_r B^r$$

.

El polinomio $v(B)$ mide el comportamiento del efecto permanente de la intervención.

Usando 3.1, 3.2 y 3.7 el modelo de intervención de forma general que definido como

$$\begin{aligned} Y_t &= \frac{\omega(B)}{v(B)} B^b X_t + \frac{\theta_N(B)}{\varphi_N(B)} W_t \\ &= T(B) X_t + N_t \end{aligned} \quad (3.8)$$

3.3. Modelos de Intervención

Al escoger adecuadamente b, r, s y los coeficientes ω_i y v_j , el término $T(B) X_t$ puede tener gran variedad de formas funcionales.

Véase algunos operadores $T(B)$ y modelos.

1. **Cambio constante permanente.** Usar $S_t(T)$ como X_t y el operador $T(B) = \omega_0$, o sea, $T(B) X_t = \omega_0 S_t(T)$. En el esquema general es $b = 0$, $r = 0$ y $s = 0$.
2. **Cambio constante temporal.** Corresponde a la función

$$X_t = \begin{cases} 1 & \text{si } T \leq t \leq T + d \\ 0 & \text{si } o.c. \end{cases}$$

donde d es la cantidad de periodos en las cuales la serie salta (si $d = 0$, se obtiene $P_t(T)$). Usar $T(B) = \omega_0$. En el esquema general $b = 0$, $r = 0$ y $s = 0$.

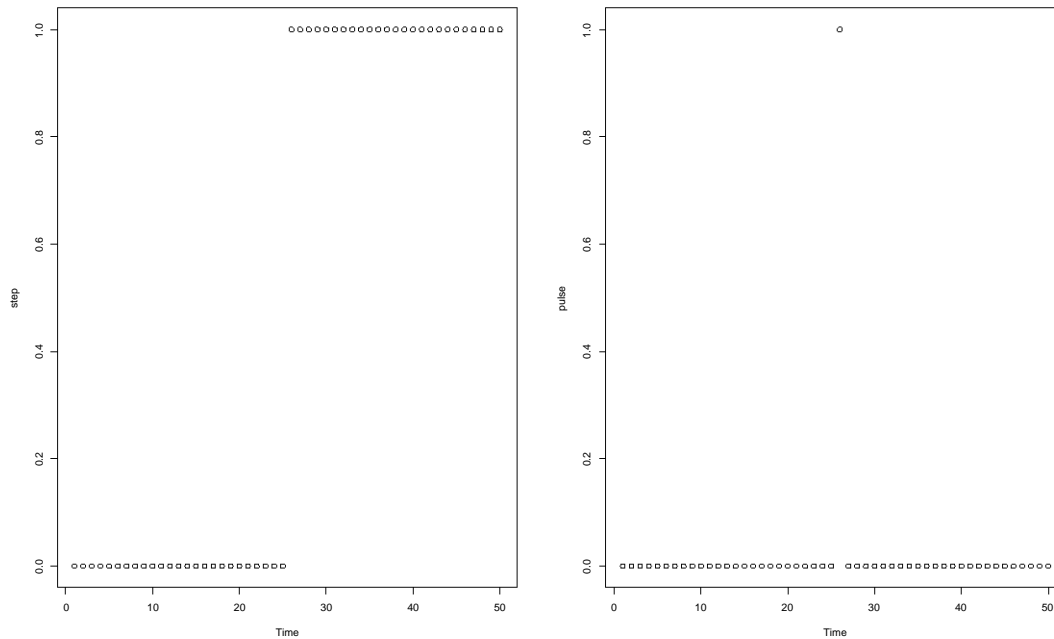


Figura 3.1: Función Escalón y función Pulso

3. **Cambio gradual de un nivel a otro de la serie.** Usar $X_t = S_t(T)$ y el operador $T(B) = \frac{\omega_0}{1-v_1B}$, para que el cambio sea gradual, se necesita que $|v_1| < 1$, entonces

$$T(B) = \frac{\omega_0}{1-v_1B} S_t(T) = \sum_{j=0}^{\infty} (v_1B)^j \omega_0 S_t(T) = \sum_{j=0}^{\infty} v_1^j \omega_0 S_{t-j}(T)$$

se ve que existe un cambio gradual a través de una serie geométrica con razón v_1 .

La Figura 3.1 muestra la Función Escalón y función Pulso.

La Figuras 3.2, 3.3, 3.4, 3.5 y 3.6 muestran distintas formas funcionales usando la Función Escalón y función Pulso con $b = 0$ y $0 < v_1 < 1$.

La Figura 3.2 corresponde a las formas $T(B) X_t = \omega_0 S_t(T)$ (que se conoce como cambio constante permanente) y $T(B) X_t = \omega_0 P_t(T)$, respectivamente.

La Figura 3.3 corresponde a las formas $T(B) X_t = \frac{\omega_0}{1-v_1B} S_t(T)$ (que se conoce como cambio gradual de un nivel a otro de la serie) y $T(B) X_t = \frac{\omega_0}{1-v_1B} P_t(T)$, respectivamente.

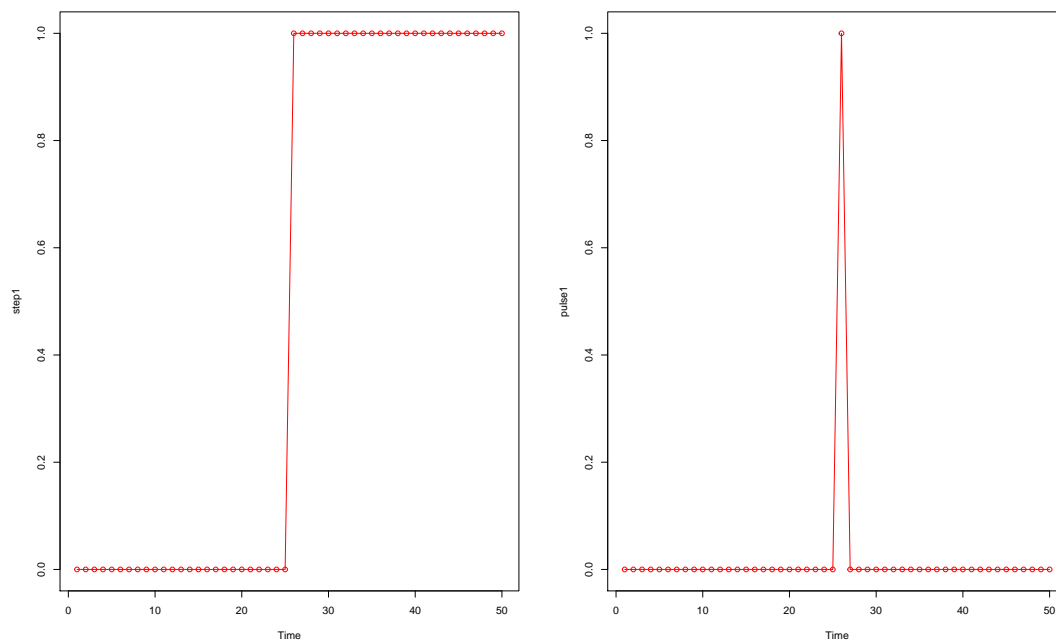


Figura 3.2: Primera forma de la función Escalón y función Pulso

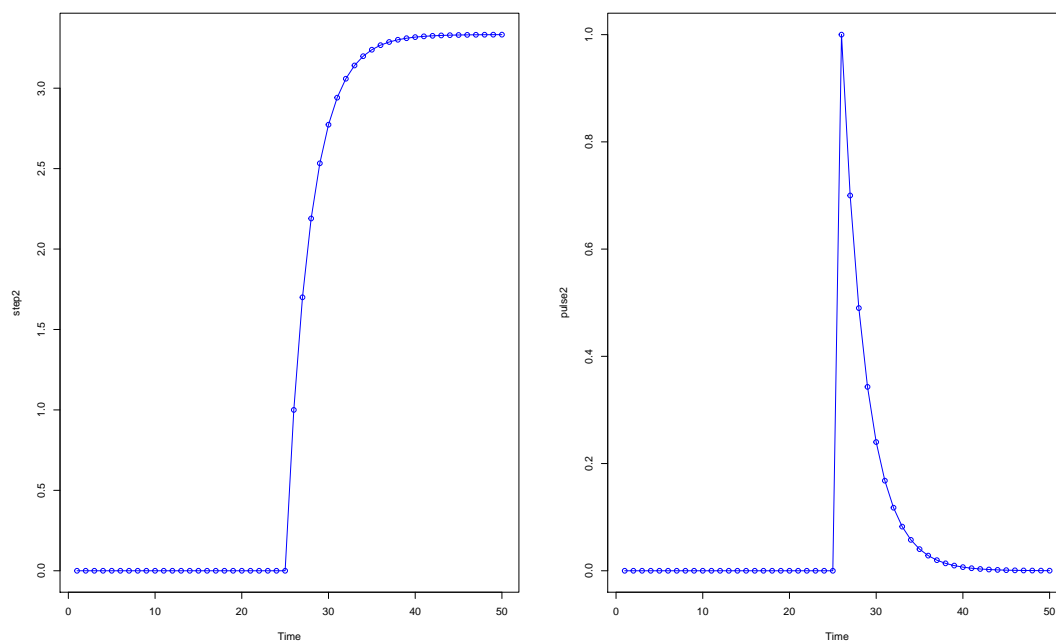


Figura 3.3: Segunda forma de la función Escalón y función Pulso

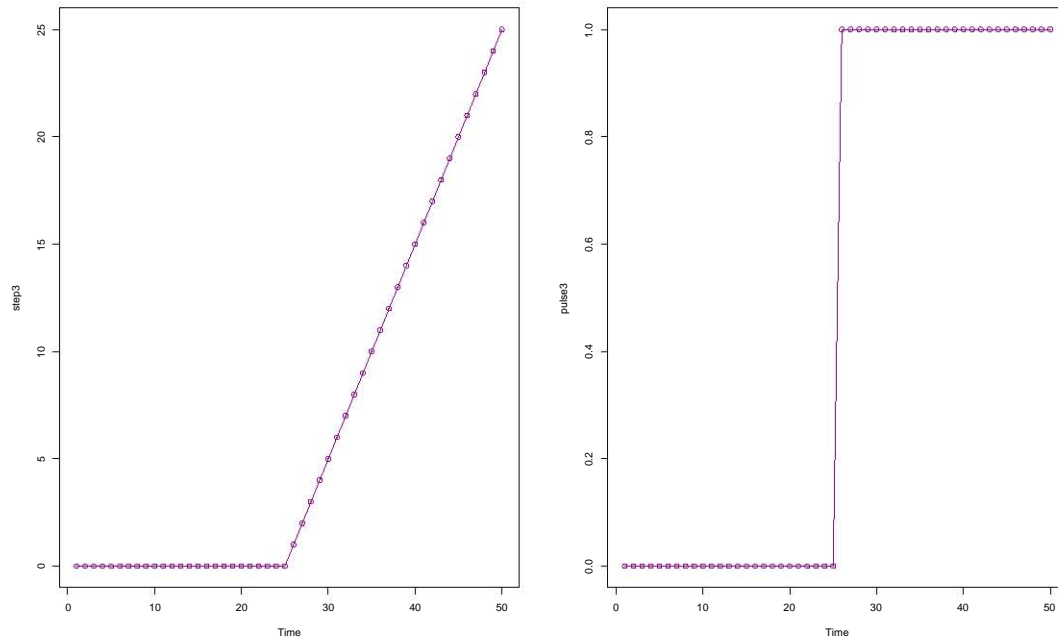


Figura 3.4: Tercera forma de la función Escalón y función Pulso

La Figura 3.4 corresponde a las formas $T(B)X_t = \frac{\omega_0}{1-B}S_t(T)$ (efecto rampa) y $T(B)X_t = \frac{\omega_0}{1-B}P_t(T)$, respectivamente.

La Figura 3.5 corresponde a las formas $T(B)X_t = \left\{ \frac{\omega_0}{1-v_1B} + \frac{\omega_1}{1-B} \right\} S_t(T)$ y $T(B)X_t = \left\{ \frac{\omega_0}{1-v_1B} + \frac{\omega_1}{1-B} \right\} P_t(T)$, respectivamente.

La Figura 3.6 describe la forma de cambio constante temporal.

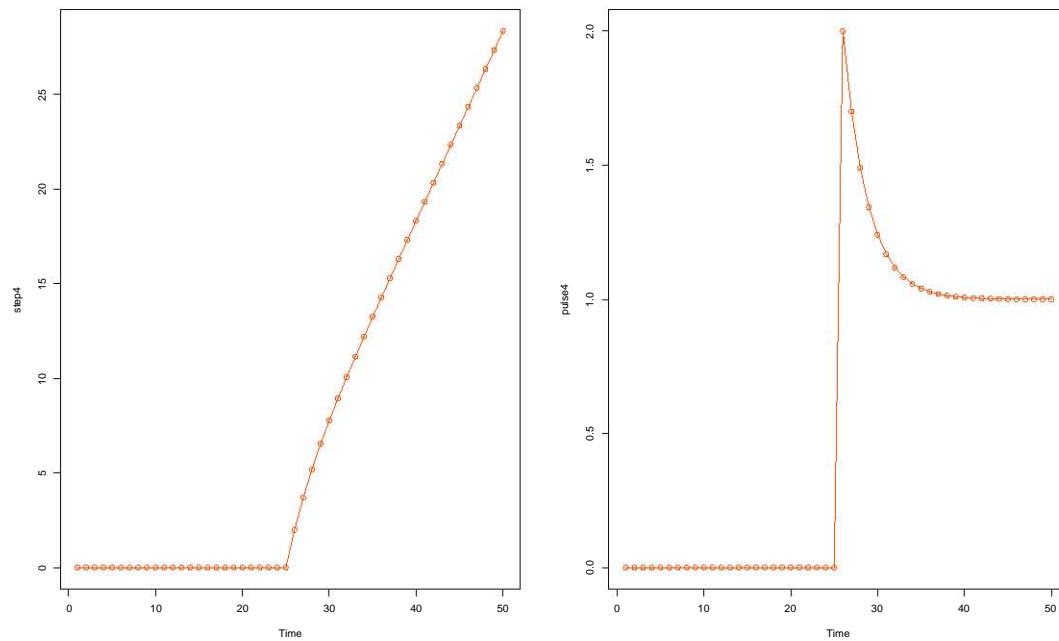


Figura 3.5: Tercera forma de la función Escalón y función Pulso

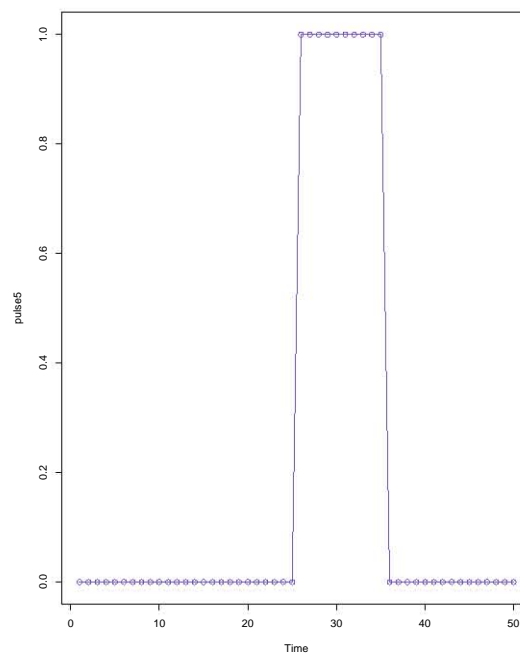


Figura 3.6: Última forma funcional

3.4. Identificación, Ajuste y Verificación

Identificación.

Construir un modelo que describa el cambio de nivel de la serie, para ello se seleccionarán el tipo de función con la que se trabajará, b, p, q , además de $\underline{\omega} = (\omega_0, \omega_1, \dots, \omega_s)$ y $\underline{v} = (v_1, v_2, \dots, v_r)$; éstas se podrán elegir de acuerdo al criterio del analista.

Ajuste.

Después de la etapa de identificación, se sigue con la modelación del ruido $\{N_t\}$ donde $\hat{N}_t = Y_t - \hat{T}(B)X_t$, para lograr esto se estima un modelo ARIMA a $\{N_t\}$ y se obtienen los parámetros $\underline{\omega}, \underline{v}, \Theta_N, \Phi_N$ por mínimos cuadrados.

Otra opción, (como el modelo de función de transferencia alterno), es ajustar una regresión, después obtener los residuos $\{N_t\}$ de esta regresión y modelarlos con un proceso ARIMA.

Posteriormente, se hace un reajuste conjunto, es decir, juntar la parte dinámica y estocástica del modelo.

Supóngase que se tiene

$$Y_t = \frac{\omega(B)}{v(B)} X_{t-b} + N_t \quad (3.9)$$

El proceso $\{N_t\}$, como se vió anteriormente, es $\phi_N(B)(1-B)^d N_t = \theta_N(B)W_t$. Entonces, la verosimilitud se puede obtener en términos de un vector n dimensional \underline{a} cuyo t -ésimo elemento es $a_t = (1-B)^d \left(Y_t - \frac{\omega(B)}{v(B)} X_{t-b} \right)$.

El correspondiente proceso asociado a a_t

$$a_t = \frac{\theta_N(B)}{\phi_N(B)} W_t \quad (3.10)$$

es estacionario. Por lo tanto, la función de verosimilitud es escrita como

$$L(\Delta, (\sigma_W)^2 | Y) = (2\pi(\sigma_W)^2)^{-\frac{n}{2}} |M|^{\frac{1}{2}} \exp \left\{ \frac{-S(\Delta)}{2(\sigma_W)^2} \right\} \quad (3.11)$$

donde $\Delta = (\omega_0, \dots, \omega_s, v_1, v_2, \dots, v_r, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ es el vector de los parámetros de la parte dinámica (ω_i y v_j) y la parte estocástica (ϕ_i y θ_j) y $M^{-1}(\sigma_W)^2$ es la matriz de covarianzas del vector $\underline{\mathbf{a}}$ y

$$\begin{aligned} S(\Delta) &= \underline{\mathbf{a}}' M \underline{\mathbf{a}} \\ &= \sum_{t=-\infty}^n [W_t | Y, \Delta]^2 \end{aligned} \quad (3.12)$$

con $[W_t | Y, \Delta]$ como el valor esperado condicional de W_t dado Δ y Y .

Los valores de los elementos de Δ minimizando 3.12 es lo que se conoce como mínimos cuadrados.

En la práctica, se obtiene $\hat{\Delta}$, $V(\hat{\Delta})$ y $(\sigma_W)^2$ usando programas computacionales que ajusten mínimos cuadrados no lineales, esto con el fin de minimizar numéricamente a $S(\Delta)$. Afortunadamente, R e ITSM cumplen con este trabajo. Finalmente, se hará el pronóstico.

Verificación de los Supuestos.

Comprobar que $\{\hat{W}_t\}$ es ruido blanco. Para ello, se usarán el ACF y PACF muestral, también se utilizará la prueba de aleatoriedad.

3.5. Implementación en R e ITSM

3.5.1. Implementación en R

Para la implementación de Análisis de Intervención en R, se usarán los datos anuales de la Inflación de México, los datos fueron obtenidos de la página del Banco Mundial. Los datos fueron registrados desde el año de 1961 al año de 2016.

Identificación.

Se cargan las paqueterías correspondiente, se cargan los datos y se les aplica una transformación (logaritmo) para tratar de mejorar el supuesto de varianza y normalidad en los datos, posteriormente se grafican

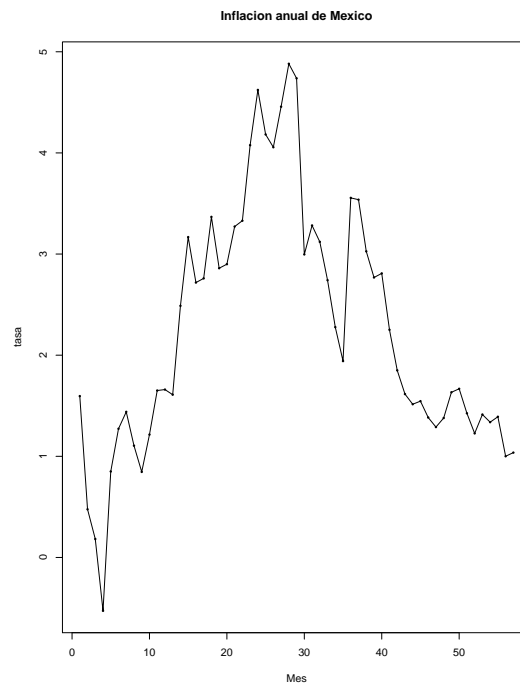


Figura 3.7: Inflación anual Mexicana (1961 – 2017)

```

library(lmtest)
library(forecast)
library(TSA)
library(stats)
library(dynlm)
setwd("C:/Users/Jesus Pichardo Leyva/Documents/TESIS")
datos<-read.table("Inflacion Anual Mexico.csv",sep="," ,header=FALSE)
datos<-datos$V2
datos<-log(datos)
datos

plot(datos,type="o",pch=16,cex=0.5,xlab='Mes',ylab='tasa',
      main="Inflacion anual de Mexico")
abline(v=23,col="2")
abline(v=29,col="2")
mtext("1982",side=3,at=23)
mtext("1988",side=3,at=29)

```

Se puede observar en la Figura 3.7 que hubo una aparente estabilidad entre

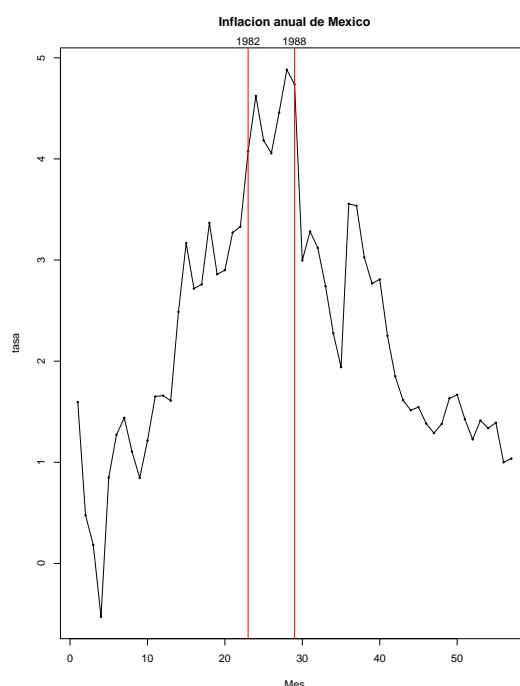


Figura 3.8: Inflación Anual Mexicana (1961 – 2017) con $1982 \leq t \leq 1988$

1961 y 1972 en los cuales no hubo movimientos bruscos. Sin embargo, a partir de 1973 la inflación a tasa anual empieza a incrementar, hasta que en 1976 hubo una devaluación del peso debido a un mal manejo de las finanzas públicas mexicanas. Posteriormente, en 1982 se presentan dos devaluaciones, la primera hizo que la inflación aumentara fuertemente, en esta ocasión fue debido a una crisis de divisas y la segunda, esta situación se dio debido a que trascendió que México no podía pagar su deuda externa. Al término de 1983 el dólar llegó a 161 pesos; a 210 en 1984. En 1985 la inflación vuelve a tener un movimiento ascendente debido a un ajuste en el control de precios de diversos productos. Durante 1986 la inflación aumentó debido a la caída de los precios del petróleo, esta caída modificó la tasa de tipo de cambio y finalmente impactó a la inflación. En 1987 hubo aumentos salariales y ajustes a los precios de algunos bienes, lo cual justifica un incremento en la tasa inflacionaria. Para 1988 hay una disminución en la inflación con respecto al año anterior debido a la entrada en vigor del PSE (Pacto de Solidaridad Económica). Para mayor información véase Guerrero (1981, 1984, 2009)

Para hacer el análisis de Intervención se tomará el efecto dinámico de 1982 a 1988 como un cambio constante temporal (Figura 3.8).

Se tomará la siguiente función X_t donde

$$X_t = \begin{cases} 1 & \text{si } 23 \leq t \leq 29 \\ 0 & \text{si o.c.} \end{cases}$$

Donde 23 es el índice con el dato de 1982 y 29 pertenece a 1988.

```
Yt<-datos
L<-length(Yt)
Xt<-c(rep(0,22),rep(1,L-50),rep(0,L-29))
```

Ahora, volvemos a la función y a la serie de datos como series de tiempo y se grafican (Figura 3.9).

```
Xt<-ts(Xt, frequency = 1, start = c(1961, 1))
Yt<-ts(Yt, frequency = 1, start = c(1961, 1))
```

```
plot(Xt)
plot(Yt)
```

La función será de la forma $b = 0$, $s = 0$ y $r = 0$ (cambio constante temporal).

Ajuste.

En el caso de R, se ajustará la siguiente regresión

$$Y_t = \omega_0 X_t$$

En la regresión Y_t son los datos. Después se obtendrán los residuos (N_t) y se modelarán como un proceso ARIMA.

```
est<-dynlm(Yt ~ 0 + Xt)
summary(est)
```

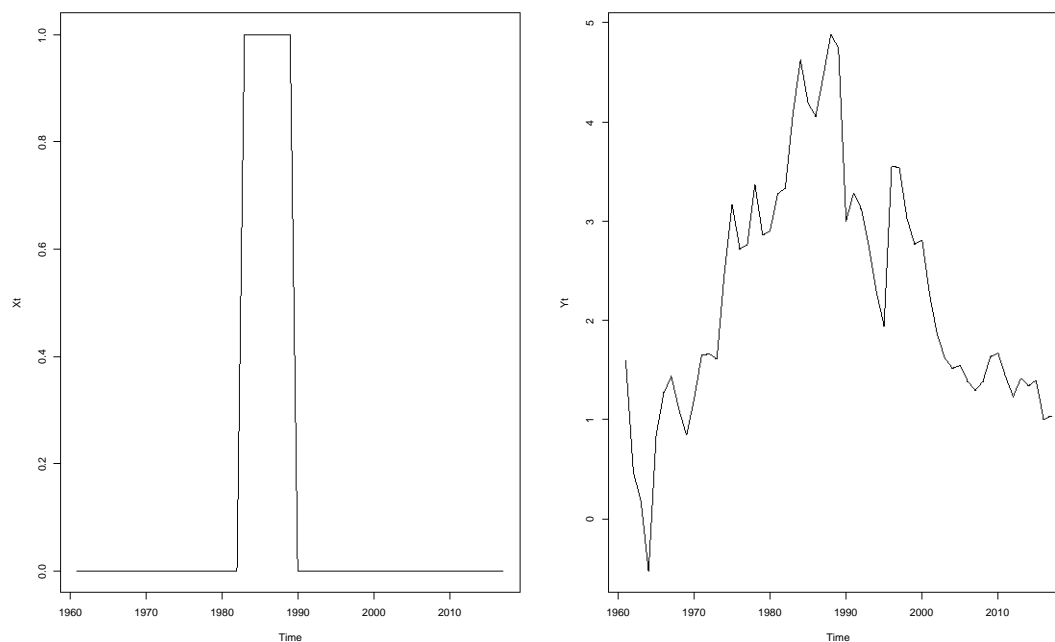


Figura 3.9: Gráfico de la función escalón e Inflación

Time series regression with "ts" data:
Start = 1961, End = 2017

Call:
dynlm(formula = Yt ~ 0 + Xt)

Residuals:

Min	1Q	Median	3Q	Max
-0.5276	1.1053	1.5953	2.7594	3.5553

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Xt	4.4305	0.7734	5.729	4.18e-07 ***

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 2.046 on 56 degrees of freedom
Multiple R-squared: 0.3695, Adjusted R-squared: 0.3582
F-statistic: 32.82 on 1 and 56 DF, p-value: 4.182e-07

```
coeftest(est)
```

```
t test of coefficients:
```

```
      Estimate Std. Error t value Pr(>|t|)
Xt  4.43053    0.77339  5.7287 4.182e-07 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
```

```
w0hat<-as.numeric(estcoef[1])
```

```
w0hat
```

```
[1] 4.430531
```

El valor estimado es significativo. Sin embargo, todavía falta agregar la parte estocástica del modelo, eso se logrará agregando el ruido o disturbio N_t . Para obtener N_t se hace lo siguiente

```
Nt<-residuals(est)
```

Ahora, se va a modelar este ruido con un proceso ARIMA, primero se obtendrán el ACF y PACF para tratar de elegir un proceso acorde a N_t

```
acf(ts(Nt,freq=1), lag.max=20, main="ACF Inflacion Anual Nt")
pacf(ts(Nt,freq=1), lag.max=20, main="PACF Inflacion Anual Nt ")
```

Analizando la Figura 3.10 y usando el criterio AIC se elegirá el mejor modelo. Nuestros tres candidatos son los procesos ARMA(1, 2), ARMA(1, 1) y ARMA(1, 3) para modelar N_t . Ahora sí, se ajustará todo el modelo en conjunto, enlazando la parte dinámica y la parte estocástica.

```
Intervencion1<-arimax(Yt,order = c(1,0,2), xtransf=Xt,
                      transfer=list(c(0,0)),include.mean=F)
Intervencion1
```

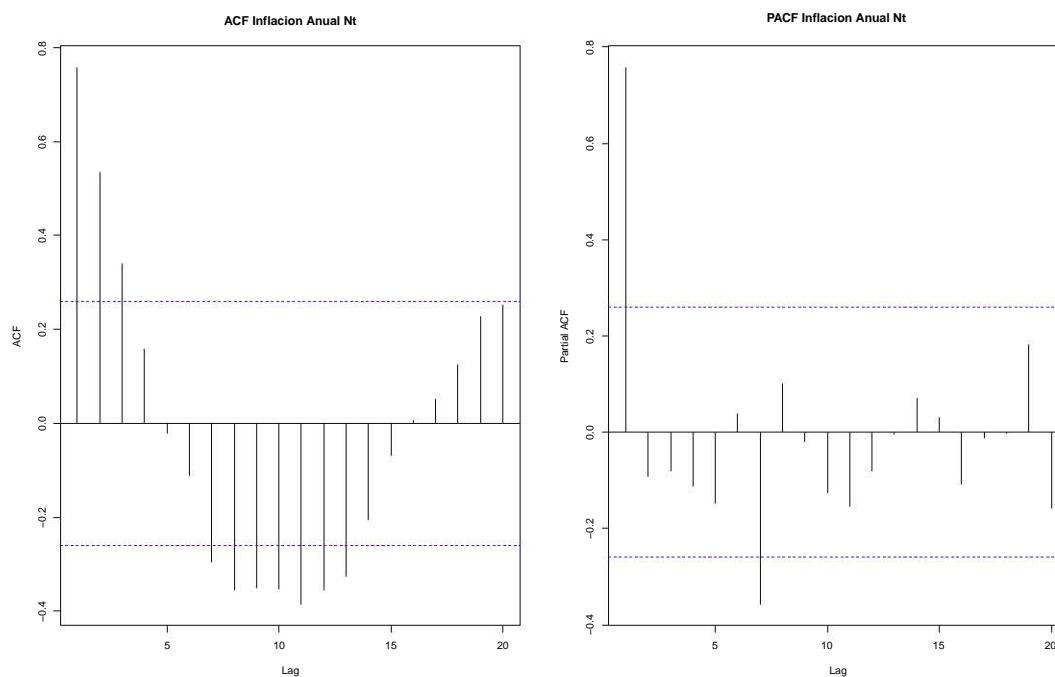


Figura 3.10: ACF y PACF de N_t

Call:

```
arimax(x = Yt, order = c(1, 0, 2), include.mean = F, xtransf = Xt,
       transfer = list(c(0, 0)))
```

Coefficients:

	ar1	ma1	ma2	T1-MA0
	0.9835	-0.0656	-0.2140	1.2406
s.e.	0.0161	0.1405	0.1257	0.3219

sigma² estimated as 0.2091: log likelihood = -37.72, aic = 83.44

AIC(Intervencion1)

[1] 85.43889

```
Intervencion2<-arimax(Yt,order = c(1,0,1), xtransf=Xt,
                    transfer=list(c(0,0)),include.mean=F)
```

Intervencion2

Call:

```
arimax(x = Yt, order = c(1, 0, 1), include.mean = F, xtransf = Xt,
       transfer = list(c(0, 0)))
```

Coefficients:

	ar1	ma1	T1-MA0
	0.9721	0.0070	1.2511
s.e.	0.0238	0.1962	0.3381

sigma^2 estimated as 0.2188: log likelihood = -39.02, aic = 84.04
AIC(Intervencion2)

[1] 86.04315

```
Intervencion3<-arimax(Yt,order = c(1,0,3), xtransf=Xt,
                      transfer=list(c(0,0)),include.mean=F)
```

Intervencion3

Call:

```
arimax(x = Yt, order = c(1, 0, 3), include.mean = F, xtransf = Xt,
       transfer = list(c(0, 0)))
```

Coefficients:

	ar1	ma1	ma2	ma3	T1-MA0
	0.9854	-0.0069	-0.1939	-0.1333	1.2547
s.e.	0.0147	0.1491	0.1168	0.1467	0.3195

sigma^2 estimated as 0.2061: log likelihood = -37.32, aic = 84.64
AIC(Intervencion3)

[1] 86.64201

Se elige el proceso ARMA(1, 2) para modelar N_t .

Analizando las estimaciones finales se nota que tienen mucho sentido; porque $\omega_0 = 1.2406$, lo que quiere decir esta estimación es que a partir del punto conocido hay un salto de ese porcentaje, posteriormente se mantendrá ese efecto hasta 1988. El modelo final es de la siguiente forma

$$\begin{aligned} Y_t^* &= 1.2406X_t + N_t \\ &= 1.2406X_t + \frac{(1 - 0.0656B - 0.2140B^2)}{1 - 0.9835B}W_t \end{aligned}$$

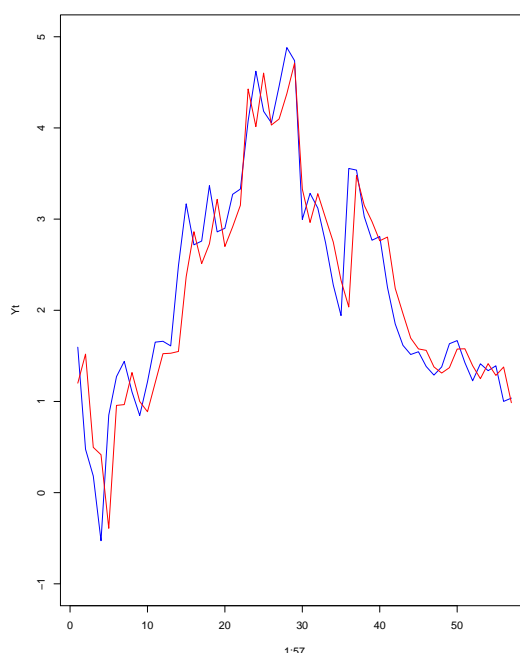


Figura 3.11: Valores ajustados (rojo) y valores reales (azul)

donde $(1 - 0.9835B) N_t = (1 - 0.0656B - 0.2140B^2) W_t$ y $Y_t^* = \log(Y_t)$. Ahora, se calcularán los residuos, es decir, la parte que no explica el modelo. Posteriormente se computarán los valores ajustados y se compararán con los valores reales.

```
resint<-residuals(Intervencion1)

full.pred<-Yt-resint
length(Yt)
plot(y=Yt,x=1:57,type="l",col="blue",ylim=c(-1,5))
lines(y=full.pred,x=1:57,col="red")
```

La traslación entre valores ajustados y reales de la Figura 3.11 se debe a que en la mayoría de los casos es complicado encontrar modelos que se ajusten perfectamente a los datos. Sin embargo, se observa que el modelo explica gran parte de la serie temporal adecuadamente, en conclusión es un buen modelo. Ahora, se muestra la gráfica de los residuos

```
residuales<-Intervencion1residuals
```

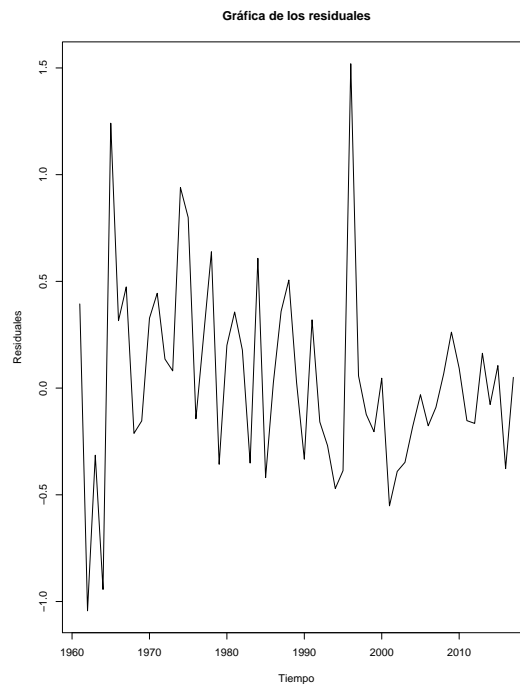


Figura 3.12: Residuos del modelo Intervención1

```
plot(residuales,xlab="Tiempo",ylab="Residuales",
     main="Grafica de los residuales",pch=19)
```

Antes de hacer el pronóstico se hará la verificación de los supuestos. Se verá si los residuos del modelo conjunto están no correlacionados.

```
acf(residuales, lag.max=10, main="ACF Residuales")
pacf(residuales, lag.max=10, main="PACF Residuales")
```

Analizando la Figura 3.13 no hay correlación entre los residuos. Ahora, se analizará la normalidad de los residuos.

```
hist(residuales,xlab = "Residuales",ylab = "Frecuencia",
     main = "Histograma de los residuales",pch=19)
residualest<-rstandard(Intervencion1)
qqnorm(residualest,ylab = "Standardized residuals",xlab = "Normal Scores",
     main = "Plot probabilidad Normal",col="blue")
qqline(residualest)
```

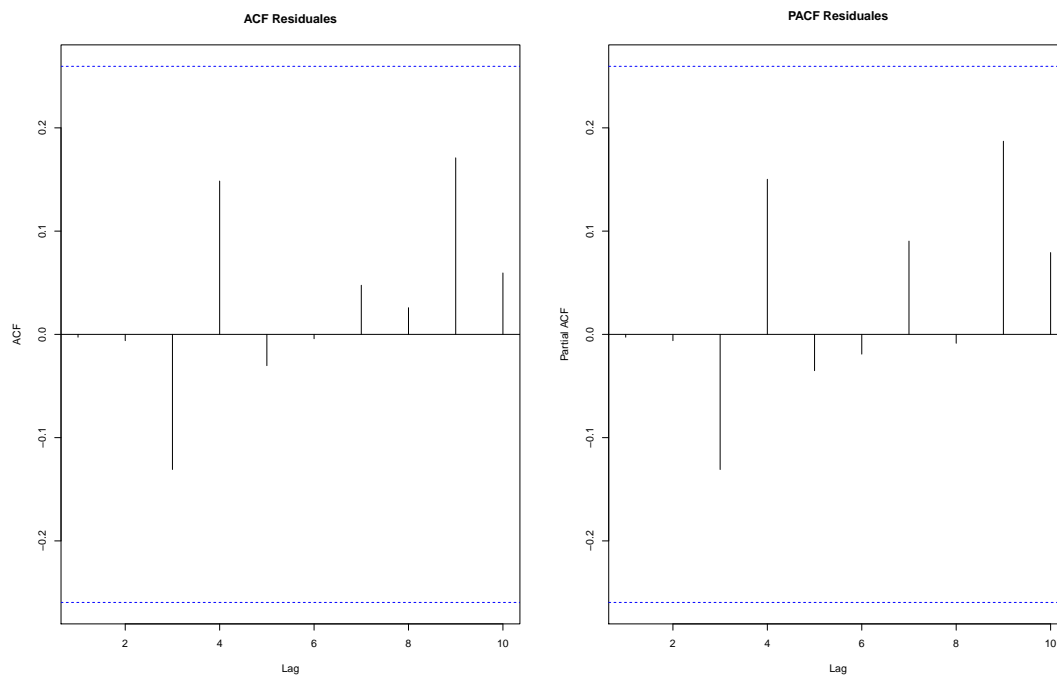



Figura 3.13: ACF y PACF de los residuos del modelo conjunto

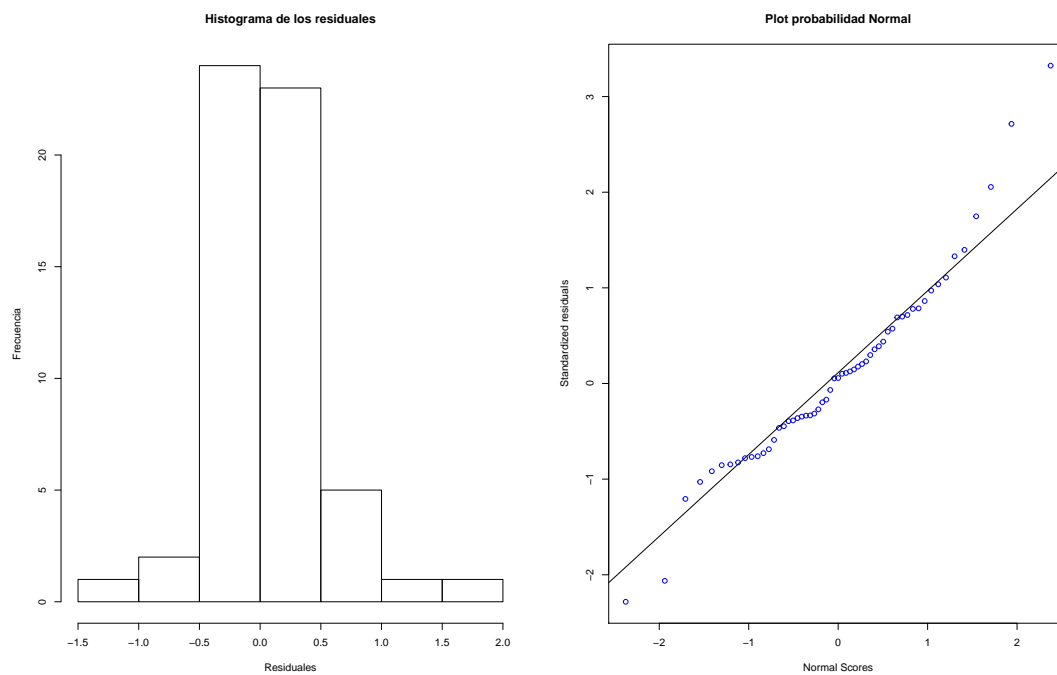


Figura 3.14: Histograma y QQplot de los residuos

```

library(nortest)
lillie.test(residuales)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  residuales
D = 0.085708, p-value = 0.373
shapiro.test(residuales)

Shapiro-Wilk normality test

data:  residuales
W = 0.9572, p-value = 0.04213

```

Analizando la Figura 3.14 se observa cierta normalidad en los residuos. Usando la prueba de Lilliefors se observa que el p-valor es igual a 0.373 entonces no se rechaza la hipótesis nula de que los residuos se distribuyen normal; por otro lado, Shapiro rechaza la hipótesis nula de que los residuos tienen distribución normal por muy poco, en este caso si usáramos un nivel de significancia de 0.01 ambas pruebas demostrarían normalidad en los residuos. Entonces, se puede asumir que los residuos son ruido blanco.

Pronóstico.

Ahora se hará el pronóstico a 5 valores y además agregando los intervalos de confianza a 95 por ciento y 99 por ciento.

```

h<-5
phi1<-Intervencion1[[1]][1]
phi1
      ar1
0.9835285
theta1<-Intervencion1[[1]][2]
theta1
      ma1
-0.06562588
theta2<-Intervencion1[[1]][3]
theta2
      ma2
-0.2139849
w0<-Intervencion1[[1]][4]

```

```

w0
  T1-MA0
1.240595
res<-as.vector(residuals(Intervencion1))
res.Int<-c(res,rep(0,h))
L<-length(datos)
L
[1] 57
Yt<-c(Yt,rep(0,h))
Xt<-c(rep(0,22),rep(1,L-50),rep(0,L-29+5))

for(i in (L+1):(L+h)){
  Yt[i]<-(phi1)*Yt[i-1]+w0*Xt[i]-(phi1*w0)*Xt[i-1]
  +res.Int[i]+(theta1)*res.Int[i-1]+(theta2)*res.Int[i-2]
}

plot(Yt[1:L],type="l",pch=16,cex=0.5,xlim=c(0,65),ylim=c(-1,5),
      xlab='Time',ylab='Inflacin Anual')
lines(58:62,Yt[58:62],col="blue")
lines(58:62,Yt[58:62]+ 1.96*sqrt(Intervencion1$sigma2),
      lty=2,col="red")
lines(58:62,Yt[58:62]- 1.96*sqrt(Intervencion1$sigma2),
      lty=2,col="red")
lines(58:62,Yt[58:62]+ 2.576*sqrt(Intervencion1$sigma2),
      lty=2,col="purple")
lines(58:62,Yt[58:62]- 2.576*sqrt(Intervencion1$sigma2),
      lty=2,col="purple")
Yt[58:62]
[1] 1.0196603 1.0028650 0.9863463 0.9700998 0.9541208
exp(Yt[58:62])
[1] 2.772253 2.726081 2.681420 2.638208 2.596387

```

La Figura 3.15 muestra el pronóstico del modelo, se puede ver que los datos de Inflación pronosticados son valores que están entre dos y tres por ciento (agregando la función exponencial a los pronósticos), lo cual es bueno ya que está dentro del objetivo del INEGI y Banco de México. En términos generales, el efecto de intervención fue eficiente y de alcances reales.

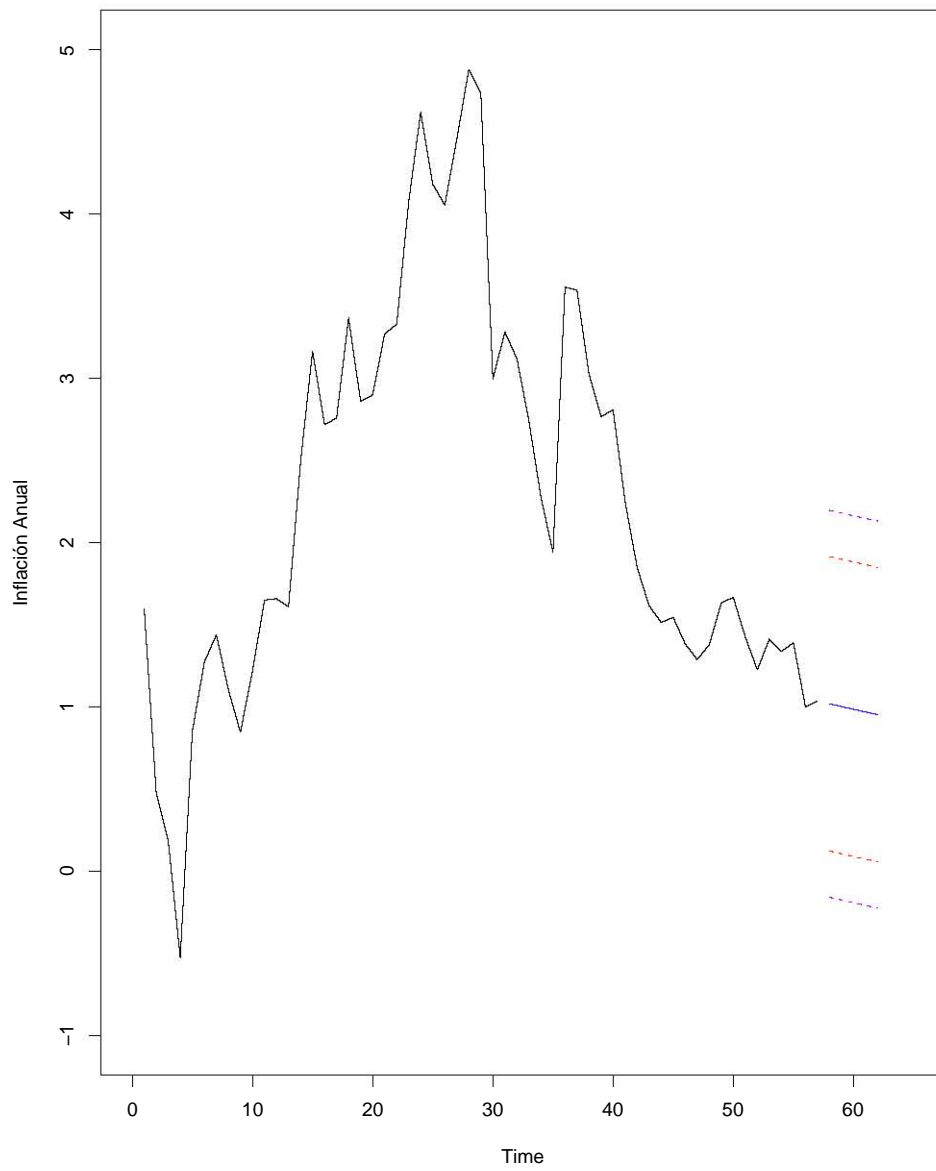


Figura 3.15: Pronóstico Inflación Mexicana Anual (1961 – 2016)

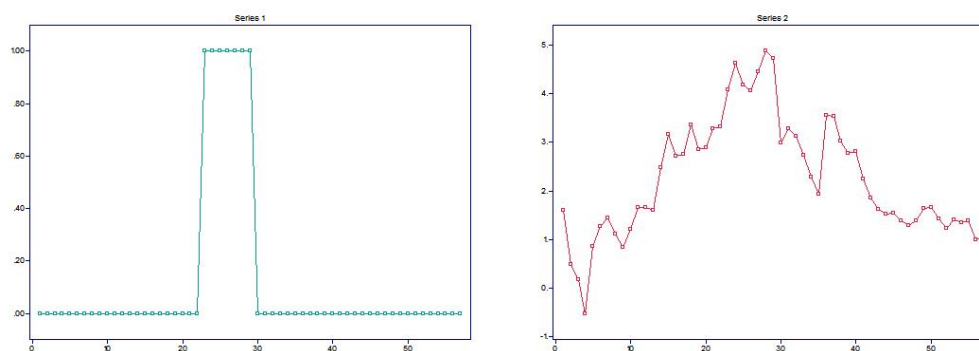


Figura 3.16: Gráfico del efecto dinámico y la serie de Inflación Anual de México (1961 – 2016)

3.5.2. Implementación en ITSM

Identificación.

Se usará la misma base de datos de R. Se empezará abriendo los datos, en este caso será una serie multivariada, ya que introduce el efecto dinámico y la serie temporal (Figura 3.16).

File > Project > Open > Multivariate > Inflacion Anual Mexico itsm.CSV

Ahora, se especificará la parte dinámica del modelo y se extraerán los residuos (N_t)

Transfer > Specify Model > Transfer Model > Denominator (cero),
 Numerator (cero) y Delay (cero)

Transfer > Estimation

```
=====
ITSM::(Transfer Function Model Estimates)
=====
```

```
X(t,2) = T(B) X(t,1) + N(t)
```

```
T(B) = B^0(4.430)
```

```
X(t,1) = Z(t)
  {Z(t)} is WN(0,1.000000)
```

```
N(t) = W(t)
  {W(t)} is WN(0,4.263059)
```

```
AICC = .358275E+03
```

```
Accuracy Parameter .010000
```

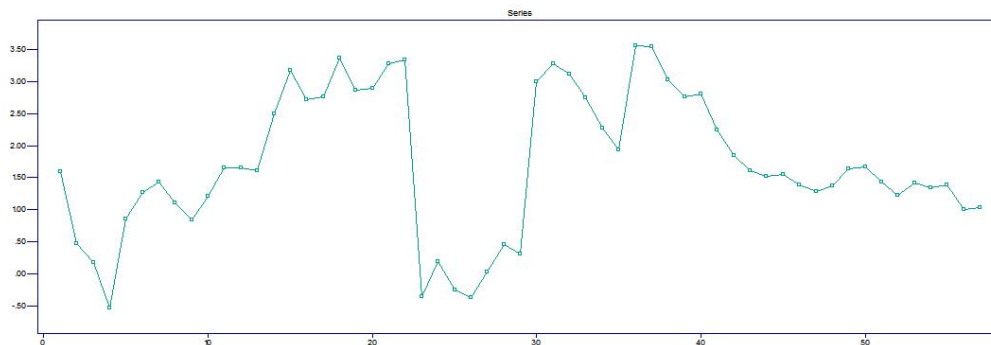
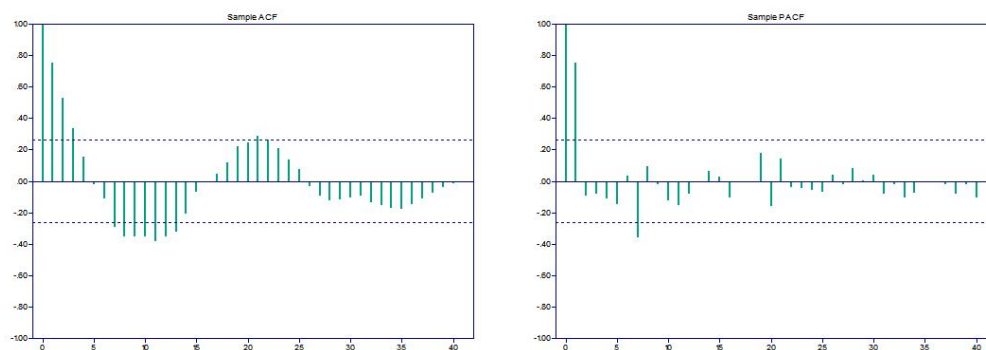
```
File > Export > Residuals > File > Nt_intervention.tsm
```

Se observa que la estimación es casi igual a la obtenida con la regresión en R. Ahora que ya se extrajo los residuos (Figura 3.17), es momento de abrirlos y ajustarles un proceso ARIMA.

```
File > Project > Open > Univariate > Nt_intervention.tsm
Statistics > ACF/PACF > Sample
```

Para tratar de obtener resultados similares se ajustará un ARMA(1,2) a N_t como en R. Se muestra a continuación un ajuste preliminar del modelo ARIMA:

```
Model > Estimation > Preliminary > AR order (uno) y MA order (dos)
> Method Hannan-Rissanen
```

Figura 3.17: Residuos (N_t)Figura 3.18: ACF y PACF de N_t

```

=====
ITSM::(Preliminary estimates)
=====

Method: Hannan-Rissanen

ARMA Model:
X(t) = .7348 X(t-1)
      + Z(t) + .7160 Z(t-1) - .4106 Z(t-2)

WN Variance = 2.086037

AR Coefficients
    .734768

MA Coefficients
    .715985    -.410550

Ratio of AR coeff. to 1.96 * (standard error)
    3.503548

Ratio of MA coeff. to 1.96 * (standard error)
    1.514836    -.839250

(Residual SS)/N = 2.08604

-2Log(Like) = .216244E+03
AICC = .225013E+03

```

Nos regresamos a la serie bivariada y se hace el ajuste juntando la parte dinámica y estocástica

```

Transfer > Specify Model > Transfer Model > Denominator (cero),
                                           Numerator (cero)
                                           y Delay (cero)
Transfer > Specify Model > Residual Model > Residual AR (uno) y
                                           Residual MA (dos)
Transfer > Estimation

```



```
=====
ITSM::(Transfer Function Model Estimates)
=====
```

$$X(t,2) = T(B) X(t,1) + N(t)$$

$$T(B) = B^{0(1.215)}$$

$$X(t,1) = Z(t)$$

{Z(t)} is WN(0,1.000000)

$$N(t) = .9950 N(t-1)$$

$$+ W(t) - .03000 W(t-1) - .2550 W(t-2)$$

{W(t)} is WN(0,.205984)

$$AICC = .219041E+03$$

Accuracy Parameter .010000

Los parámetros ajustados del modelo final son similares a los ajustados en R.

Verificación de Supuestos.

Ahora, se verificarán que los residuos cumplan con ser ruido blanco.

```
File > Export > Residual > File > residuales.tsm
File > Open > Project > Univariate > residuales.tsm
Statistics > Residual Analysis > ACF/PACF
Statistics > Residual Analysis > Tests of Randomness
```

```

=====
ITSM::(Tests of randomness on residuals)
=====

Ljung - Box statistic = 16.977 Chi-Square ( 20 ),
p-value = .65449

McLeod - Li statistic = 4.5863 Chi-Square ( 20 ),
p-value = .99986

# Turning points = 31.000~AN(36.667,sd = 3.1323),
p-value = .07043

# Diff sign points = 28.000~AN(28.000,sd = 2.1985),
p-value = 1.00000

Rank test statistic = .67400E+03~AN(.79800E+03,sd = 72.634),
p-value = .08779

Jarque-Bera test statistic (for normality) = 20.688 Chi-Square (2),
p-value = .00003

Order of Min AICC YW Model for Residuals = 0

```

Analizando la prueba de Ljung-Box se tiene que el p-valor es mayor a 0.05 entonces no se rechaza H_0 y se afirma que todas las correlaciones son iguales a cero. Por lo tanto, los residuos son ruido blanco. Lo mismo se observa en la Figura 3.20. Analizando la Figura 3.21 se observa cierto comportamiento normal en los residuos.

Pronóstico.

Por último, la Figura 3.22 muestra los pronósticos a 5 años (regresando a la serie bivariada).

```
Forecasting > Transfer-function
```

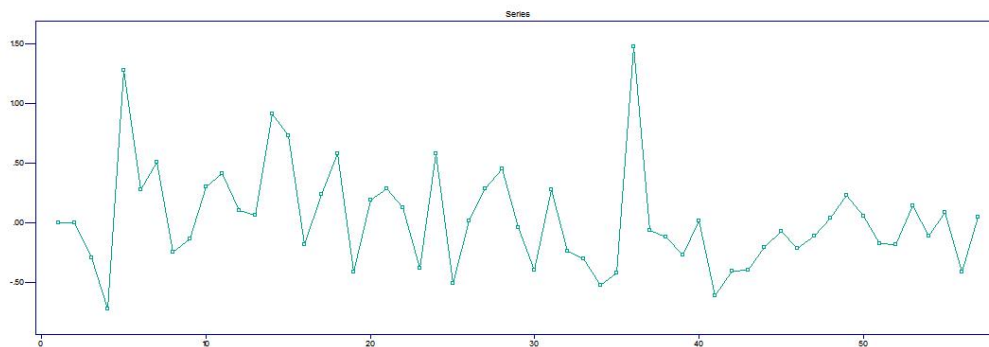


Figura 3.19: Residuos del modelo conjunto final

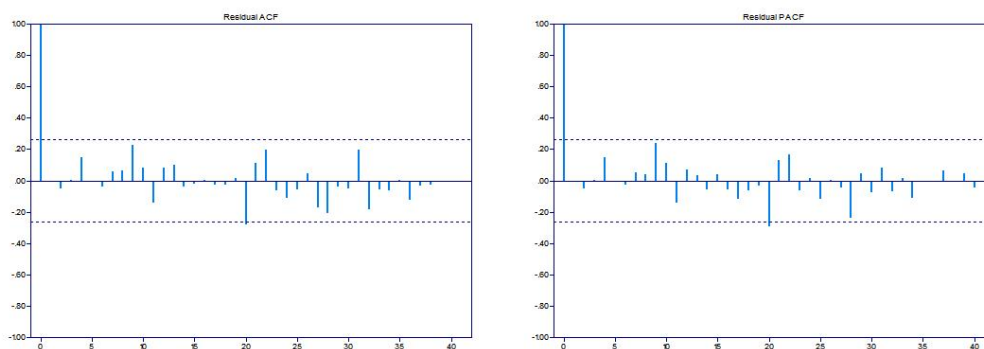


Figura 3.20: ACF y PACF de los residuos del modelo conjunto

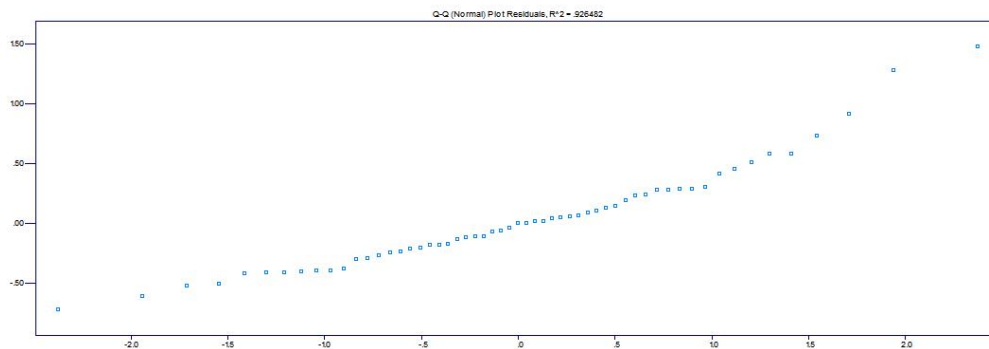


Figura 3.21: QQplot normal de los residuos del modelo conjunto

```

=====
ITSM::(Transfer Function Model Prediction)
=====

X(t,2) = T(B) X(t,1) + N(t)

T(B) = B^0(1.215)

X(t,1) = Z(t)
  {Z(t)} is WN(0,1.000000)

N(t) = .9950 N(t-1)
      + W(t) - .03000 W(t-1) - .2550 W(t-2)
  {W(t)} is WN(0,.205984)

X(t,2)

                                     Approximate 95 Percent
                                     Prediction Bounds
Time      Prediction      sqrt(MSE)      Lower      Upper
  58      1.23235      1.29700      -1.30972      3.77443
  59      1.20205      1.36895      -1.48104      3.88515
  60      1.19604      1.40586      -1.55940      3.95149
  61      1.19006      1.44148      -1.63519      4.01531
  62      1.18411      1.47589      -1.70859      4.07681

```

Los valores pronosticados son muy similares a los obtenidos en R.

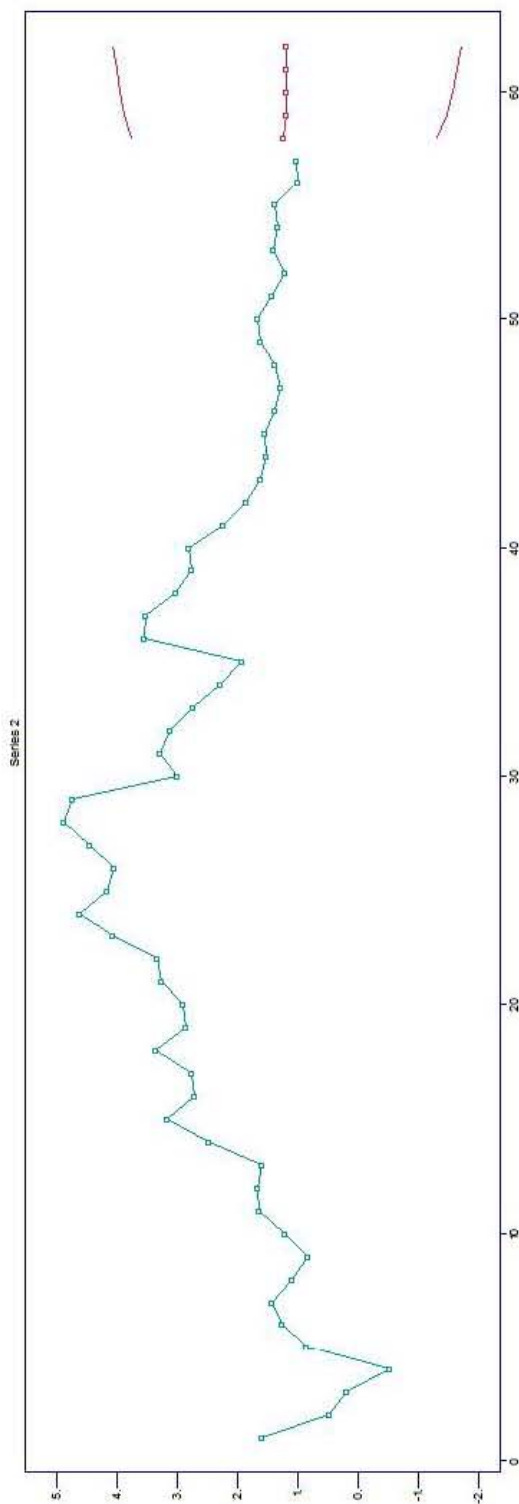


Figura 3.22: Pronóstico a 5 años de la Inflación Anual de México (1961-2016)

AÑO	Tasa
1961	3.25
1962	2.75
1963	2.93
1964	5.9
1965	1.56
1966	5.08
1967	3.08
1968	1.05
1969	7.07
1970	2.63
1971	6.3
1972	6.49
1973	13.42
1974	23.11
1975	15.63
1976	19.36
1977	30.46
1978	16.01
1979	19.63
1980	33.41
1981	26.01
1982	60.92
1983	90.47
1984	59.09
1985	56.74
1986	73.62
1987	139.66
1988	112.71
1989	26.53
1990	28.13
1991	23.25
1992	14.41
1993	34.08
1994	8.23
1995	31.58
1996	29.25
1997	17.8
1998	15.11
1999	17.65
2000	10.82
2001	5.37
2002	5.62
2003	5.97
2004	8.31
2005	5.41
2006	6.18
2007	5.02
2008	6
2009	3.54
2010	4.49
2011	5.29
2012	3.25
2013	1.76
2014	4.71
2015	2.99
2016	4.61

Cuadro 3.1: Inflación Anual de México 1961 – 2016

Capítulo 4

Outliers en Series de Tiempo

4.1. Introducción

Como se mencionó en el capítulo anterior, las observaciones de una serie de tiempo son influenciadas por algunos eventos como: cambios repentinos en la política económica o social, crisis financieras, desastres naturales, huelgas, marchas o simplemente errores no reportados en la realización de la serie temporal.

Estos eventos crean observaciones que no son consistentes con el resto de la serie. Tales observaciones son conocidas como *outliers*. La presencia de outliers en una serie de tiempo puede traer efectos sustanciales en el comportamiento de las autocorrelaciones muestrales (selección de modelo errónea), las estimaciones de los parámetros del modelo ARIMA (introducción de sesgo en los parámetros), los pronósticos (predicciones equivocadas) e intervalos de confianza (intervalos más grandes).

Cuando el tiempo y las causas son conocidas, sus efectos pueden ser cuantificados usando teoría de Análisis de Intervención. Sin embargo, en la práctica, el tiempo de estos eventos es por lo regular desconocido. Un enfoque común es identificar esos outliers y posteriormente usar teoría de Análisis de Intervención.

Se consideran cuatro tipos de outliers: outlier innovativo, outlier aditivo, cambio de nivel y cambio temporal.

Para este capítulo se utilizarán los trabajos de Chang y Tiao (1988) y Chen y Liu (1993) principalmente.

4.2. Tipos de Outliers

La representación general de un outlier se puede dar de la siguiente forma: $L(B)I(t_j)$, donde $L(B)$ es un polinomio de operadores de retraso e $I(t_j)$ es una variable indicadora que toma el valor de 1 cuando $t = j$, es decir cuando el outlier 'salta' y el valor de 0 en otro caso.

En la teoría, se consideran cuatro tipos de outliers: outlier innovativo (innovational outlier, IO), outlier aditivo (additive outlier, AO), cambio de nivel (level shift, LS) y cambio temporal (temporary change, TC).

El polinomio $L(B)$ para cada tipo de outlier se define como:

$$\text{IO} : L(B) = \frac{\theta(B)}{\varphi(B)} \quad (4.1)$$

$$\text{AO} : L(B) = 1 \quad (4.2)$$

$$\text{LS} : L(B) = \frac{1}{1-B} \quad (4.3)$$

$$\text{TC} : L(B) = \frac{1}{1-\delta B} \quad (4.4)$$

El valor de δ es por lo regular igual a 0.7. Cada uno de ellos tiene una estructura específica en influencia en los datos, como se puede observar en las Figuras 4.2, 4.3, 4.4 y 4.5 .

En la Figura 4.1 se muestran distintos tipos de cambios temporales, esto depende de la δ . En el caso de $\delta = 1$, la estructura cambia y se convierte en un cambio de nivel.

Se define Y_t^* como la serie observada sujeta a m outliers con pesos ω_j , donde ω_j representa la magnitud de los atípicos.

$$Y_t^* = \sum_{j=1}^m \omega_j L_j(B) I_t(t_j) + \frac{\theta(B)}{\varphi(B)} W_t \quad (4.5)$$

donde $\varphi(B) = (1-B)^d \phi(B)$ y $\varphi(B) Z_t = \theta(B) W_t$

Los residuos estimados, que están contaminados con los outliers, están dados por:

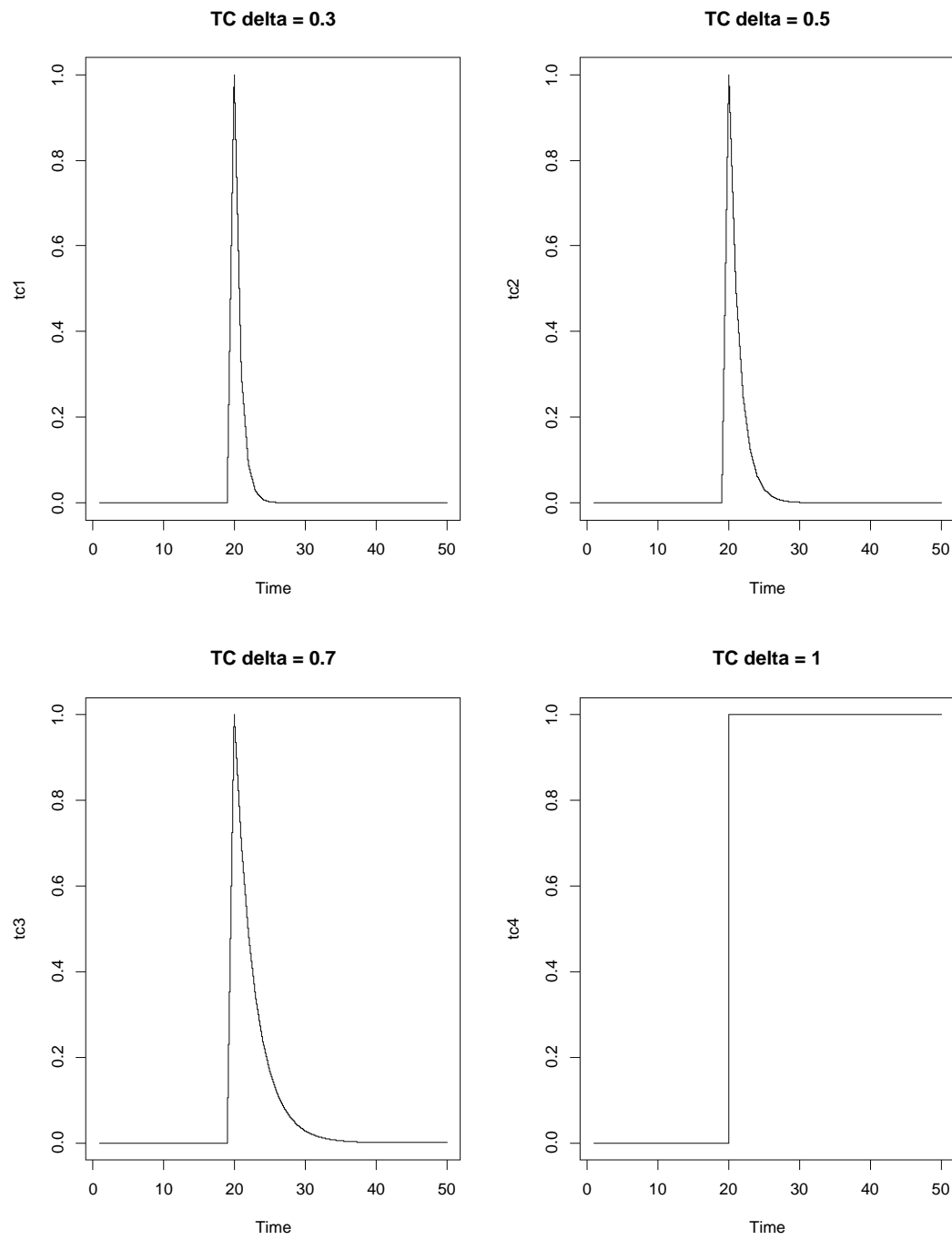


Figura 4.1: Cambios temporales con distinta delta

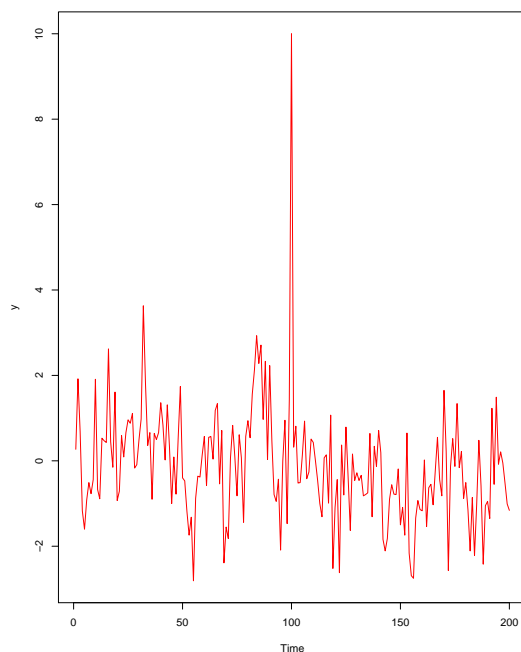


Figura 4.2: Efecto del outlier aditivo

$$\begin{aligned} \pi(B) Y_t^* &\equiv \hat{e}_t \\ &= \sum_{j=1}^m \omega_j \pi(B) L_j(B) I_t(t_j) + W_t \end{aligned} \quad (4.6)$$

donde los coeficientes de $\pi(B) = \sum_{i=0}^{\infty} \pi_i B^i$ se puede determinar de la relación:

$$\begin{aligned} \pi(B) &= \frac{\varphi(B)}{\theta(B)} \\ &= 1 - \pi_1 B - \pi_2 B^2 - \dots \end{aligned} \quad (4.7)$$

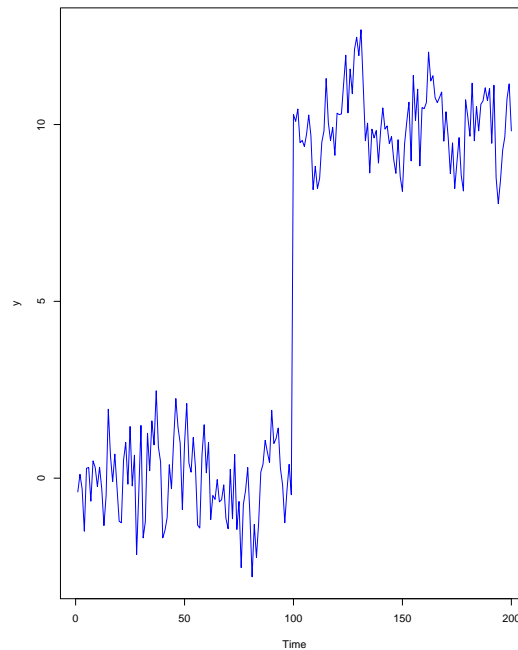


Figura 4.3: Efecto de un cambio de nivel

4.3. Teoría de acuerdo a Chang y Tiao

En esta parte del capítulo se hablará de la teoría de Chang y Tiao, la cual sólo presenta los outliers aditivo e innovativo. Se propone el criterio de razón de verosimilitud para probar la existencia de datos atípicos y el método para saber distinguirlos. Por último, un procedimiento iterativo para estimar los parámetros de la parte dinámica y la parte estocástica.

4.3.1. Outlier Innovativo y Outlier Aditivo

Sea Z_t el proceso de serie de tiempo libre del impacto de outliers, y sea Y_t la serie de tiempo observada. Se asume que Z_t sigue un modelo ARMA $\varphi(B)Z_t = \theta(B)W_t$. Entonces, usando 4.5 con $m = 1$, un outlier aditivo o 'outlier de observación' a tiempo T se modela como:

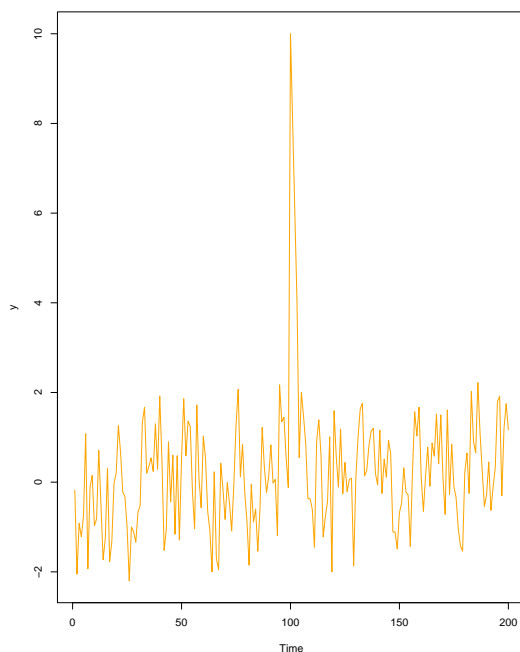


Figura 4.4: Efecto del outlier innovativo

$$\begin{aligned}
 Y_t &= \omega I_t(T) + Z_t \\
 &= \omega I_t(T) + \frac{\theta(B)}{\varphi(B)} W_t
 \end{aligned} \tag{4.8}$$

donde

$$I_t(T) = \begin{cases} 1 & \text{si } t = T \\ 0 & \text{si } t \neq T \end{cases} \tag{4.9}$$

denota el indicador pulso a tiempo T , así el caso AO puede ser llamado también modelo de error bruto, ya que sólo el nivel de la T -ésima observación es afectada.

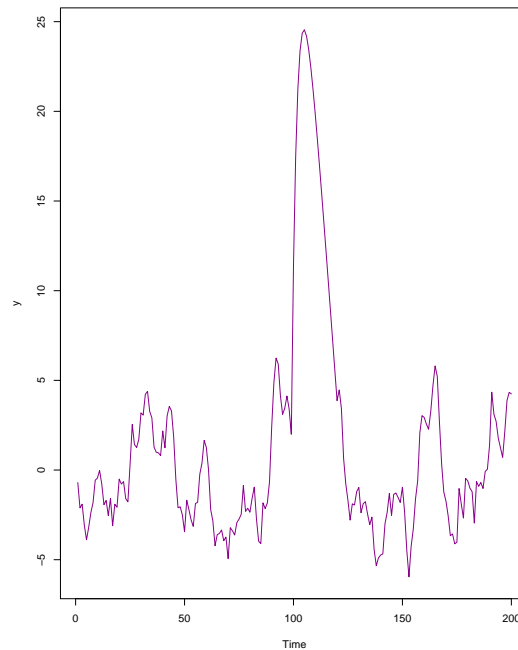


Figura 4.5: Efecto de un cambio temporal

Un outlier innovativo a tiempo T ó impacto innovativo es modelado como:

$$\begin{aligned}
 Y_t &= \frac{\theta(B)}{\varphi(B)} (\omega I_t(T) + W_t) \\
 &= \omega \frac{\theta(B)}{\varphi(B)} I_t(T) + Z_t
 \end{aligned} \tag{4.10}$$

Así, un AO afecta el nivel de la serie observada sólo a tiempo T , por una cantidad ω . Mientras que un IO representa un impacto aleatorio a tiempo T , que afecta a todas las observaciones sucesivas $(Y_T, Y_{T+1}, Y_{T+2}, \dots)$, a través del sistema dinámico descrito por $\psi(B) = \frac{\theta(B)}{\varphi(B)}$, tal que $Y_t = \omega \psi_i + Z_t$ para $t = T + i \geq T$. Para una serie estacionaria, el efecto de IO es temporal, ya que ψ_i decae exponencialmente a 0, pero para una serie no estacionaria con $d \geq 1$, puede haber efectos permanentes.

4.3.2. Criterio de razón de verosimilitud

El Criterio de razón de verosimilitud sirve para detectar un outlier o dato atípico y además el tipo de observación atípica, ya sea outlier innovativo o outlier aditivo. Se tienen dos casos, uno es cuando los parámetros ARIMA y T son conocidos y otro cuando ninguno de los dos es conocido.

Parámetros ARIMA y T conocidos

Usando 4.6 y 4.7 con $m = 1$, se tiene que:

$$IO : e_t = \omega I_t(T) + W_t \quad (4.11)$$

$$AO : e_t = \omega \pi(B) I_t(T) + W_t \quad (4.12)$$

Para n observaciones disponibles, el modelo AO en 4.12 se escribe como

$$\begin{bmatrix} e_1 \\ \vdots \\ e_{T-1} \\ e_T \\ e_{T+1} \\ e_{T+2} \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ -\pi_1 \\ -\pi_2 \\ \vdots \\ -\pi_{n-T} \end{bmatrix} + \begin{bmatrix} W_1 \\ \vdots \\ W_{T-1} \\ W_T \\ W_{T+1} \\ W_{T+2} \\ \vdots \\ W_n \end{bmatrix}$$

Se puede observar que para un AO la información está repartida sobre los residuos $e_T, e_{T+1}, e_{T+2}, \dots$, con pesos decrecientes $1, -\pi_1, -\pi_2, \dots$

Sea $\hat{\omega}_{A,T}$ el estimador por mínimos cuadrados de ω para el modelo AO, se tiene que

$$\begin{aligned}
AO : \hat{\omega}_{A,T} &= \frac{e_T - \sum_{i=1}^{n-T} \pi_i e_{T+i}}{\sum_{i=0}^{n-T} \pi_i^2} \\
&= \frac{\pi^*(F) e_t}{\tau^2}
\end{aligned} \tag{4.13}$$

se puede observar que el modelo AO es la combinación lineal de e_T, e_{T+1}, \dots ; $\pi^*(F) = (1 - \pi_1 F - \pi_2 F^2 - \dots - \pi_{n-T} F^{n-T}) = 1 - \sum_{i=1}^{n-T} \pi_i F^i$, F es el operador adelanto tal que $F e_t = e_{t+1}$, y $\tau^2 = \sum_{i=0}^{n-T} \pi_i^2$. Haciendo uso de la teoría de regresión, la varianza del estimador es

$$\begin{aligned}
Var(\hat{\omega}_{A,T}) &= Var\left(\frac{\pi^*(F) e_t}{\tau^2}\right) \\
&= \frac{1}{\tau^4} Var(\pi^*(F) e_t) \\
&= \frac{1}{\tau^4} Var\left(e_T - \sum_{i=1}^{n-T} \pi_i e_{T+i}\right) \\
&= \frac{\tau^2 \sigma_W^2}{\tau^4} \\
&= \frac{\sigma_W^2}{\tau^2}
\end{aligned} \tag{4.14}$$

de manera similar, sea $\hat{\omega}_{I,T}$ el estimador por mínimos cuadrados de ω para el modelo IO, se tiene:

$$IO : \hat{\omega}_{I,T} = e_t \tag{4.15}$$

es decir, el impacto ω es simplemente el residuo a tiempo T. La varianza es

$$\begin{aligned}
Var(\hat{\omega}_{I,T}) &= Var(e_T) \\
&= Var(\omega I_t(T) + W_t) \\
&= \sigma_W^2
\end{aligned} \tag{4.16}$$

Ya que $\tau^2 \geq 1$, se puede concluir que $Var(\hat{\omega}_{A,T}) \leq Var(\hat{\omega}_{I,T}) = \sigma_W^2$, y en algunos casos $Var(\hat{\omega}_{A,T})$ puede llegar a ser mucho más pequeña que σ_W^2 .

Varios métodos se han llevado a cabo para comprobar la presencia de un outlier tipo AO o IO en un tiempo dado T . La prueba puede denotar la hipótesis de que $\omega = 0$ en cualquier modelo, ya sea 4.8 ó 4.10, contra $\omega \neq 0$.

La prueba de verosimilitud se puede derivar para ambas situaciones

$$\lambda_{I,T} = \frac{\hat{\omega}_{I,T}}{\sigma_W} \quad (4.17)$$

y

$$\lambda_{A,T} = \frac{\tau \hat{\omega}_{A,T}}{\sigma_W} \quad (4.18)$$

respectivamente, para IO y AO. Bajo la hipótesis nula ($\omega = 0$) ambas estadísticas en 4.17 y 4.18 tendrán distribución normal estándar.

Por lo tanto, la prueba se puede escribir de la siguiente manera:

$$H_0 : \omega = 0 \quad vs \quad H_1 : \lambda_{I,T} = \frac{\hat{\omega}_{I,T}}{\sigma_W}$$

$$H_0 : \omega = 0 \quad vs \quad H_1 : \lambda_{A,T} = \frac{\tau \hat{\omega}_{A,T}}{\sigma_W}$$

Parámetros ARIMA y T desconocidos

En la práctica, el tiempo T de un posible outlier es desconocido. Para enfrentar el problema de detección de outliers en tiempos desconocidos, se procede a calcular $\lambda_{I,t}$ y $\lambda_{A,t}$ para cada $t = 1, 2, \dots, n$ y luego se toman decisiones basadas en los resultados muestrales obtenidos. Por lo regular, también los parámetros ϕ_j , θ_j , π_j y σ_W^2 son desconocidos y tienen que ser estimados. Se sabe que la existencia de outliers hace que las estimaciones de los parámetros seas seriamente sesgadas.

Chang y Tiao (1988), propusieron un procedimiento iterativo de detección de outliers.

Procedimiento propuesto

PASO 1. Modelar la serie $\{Z_t\}$ asumiendo que no hay outliers, el mejor modelo se elegirá de acuerdo al Criterio de Información Bayesiano (BIC). Calcular los residuos del modelo estimado, es decir

$$\begin{aligned}\hat{e}_t &= \hat{\pi}(B) Y_t \\ &= \frac{\hat{\varphi}(B)}{\hat{\theta}(B)} Y_t\end{aligned}\quad (4.19)$$

donde $\hat{\varphi}(B) = (1 - \hat{\varphi}_1 B - \dots - \hat{\varphi}_p B^p)$ y $\hat{\theta}(B) = (1 + \hat{\theta}_1 B + \dots + \hat{\theta}_q B^q)$ y $\hat{\sigma}_W^2 = \frac{1}{n} \sum_{t=1}^n \hat{e}_t^2$.

PASO 2. Calcular $\hat{\lambda}_{I,t} = \frac{\hat{\omega}_{I,t}}{\hat{\sigma}_W}$ y $\hat{\lambda}_{A,t} = \frac{\hat{\tau}\hat{\omega}_{A,t}}{\hat{\sigma}_W}$ usando el modelo estimado. Se define

$$\hat{\lambda}_T = \max_t \left[\max \left(|\hat{\lambda}_{I,t}|, |\hat{\lambda}_{A,t}| \right) \right] \quad (4.20)$$

donde T denota el tiempo cuando éste máximo ocurre.

La posibilidad de encontrar un outlier de tipo AO es identificando a tiempo T si $|\hat{\lambda}_T| > C$, donde C es una constante predeterminada con valores típicos de 3.0, 3.5, 4.0. El efecto de este IO se puede remover al modificar los datos usando 4.10 como

$$\tilde{Y}_t = Y_t - \hat{\omega}_{I,T} \frac{\hat{\theta}(B)}{\hat{\varphi}(B)} I_t(T) \quad (4.21)$$

y definiendo los nuevos modelos residuos, usando 4.11, como

$$\tilde{e}_t = \hat{e}_t - \hat{\omega}_{I,T} I_t(T) \quad (4.22)$$

Una nueva estimación de $\tilde{\sigma}_W^2$ es calculada de los residuos modificados.

Si $\hat{\lambda}_T = |\hat{\lambda}_{A,T}| > C$, la posibilidad de un AO es identificado en T. El efecto de este AO se puede remover usando 4.8

$$\tilde{Y}_t = Y_t - \hat{\omega}_{A,T} I_t(T) \quad (4.23)$$

y se define los nuevos residuos usando 4.12, es decir,

$$\tilde{e}_t = \hat{e}_t - \hat{\omega}_{A,T} \hat{\pi}(B) I_t(T) \quad (4.24)$$

PASO 3. Recalcular $\lambda_{I,T}$ y $\lambda_{A,T}$ basándose en los residuos modificados y σ_W^2 , repetir paso 2 hasta que todos los outliers sean identificados. Las estimaciones iniciales en $\pi(B)$ permanecen sin cambio.

PASO 4. Supóngase que el paso 3 está terminado y k outliers han sido identificados a tiempo T_1, T_2, \dots, T_k . Tratar estos outliers como conocidos, estimar los parámetros $\omega_1, \omega_2, \dots, \omega_k$ y los parámetros de la serie de tiempo usando el modelo

$$Y_t = \sum_{j=1}^k \omega_j L_j(B) I_t(T_j) + \frac{\theta(B)}{\varphi(B)} W_t \quad (4.25)$$

donde $L_j(B) = 1$ para un AO y $L_j(B) = \frac{\theta(B)}{\varphi(B)}$ para un IO a tiempo $t = T_j$. Este resultado tiene como nuevos residuos a

$$\hat{e}_t = \hat{\pi}(B) \left[Y_t - \sum_{j=1}^k \hat{\omega}_j \hat{L}_j(B) I_t(T_j) \right] \quad (4.26)$$

y una nueva σ_W^2 se obtiene.

Del paso 2 al paso 4 se repite hasta que todos los outliers son identificados y sus impactos estimados. Así, se tiene el modelo final ajustado

$$Y_t = \sum_{j=1}^k \hat{\omega}_j \hat{L}_j I_t(T_j) + \frac{\hat{\theta}(B)}{\hat{\phi}(B)} W_t \quad (4.27)$$

donde $\hat{\omega}_j = (\hat{\omega}_1, \dots, \hat{\omega}_k)$, $\hat{\phi}(B) = (1 - \hat{\phi}_1 B - \dots - \hat{\phi}_p B^p)$ y finalmente $\hat{\theta}(B) = (1 + \hat{\theta}_1 B + \dots + \hat{\theta}_q B^q)$ se obtienen en la última iteración.

4.3.3. Implementación en R

Para la implementación de datos atípicos en Series de Tiempo en R, se usarán los datos mensuales de Inflación de México. La inflación muestra la tasa de variación de precios en la economía en general. Los datos fueron obtenidos del Sistema de datos del INEGI y fueron registrados desde el año de 1970 al año de 1985.

Los datos se muestran en el Cuadro 4.1

AÑO	E	F	M	A	M	JN	JL	A	S	O	N	D
1970	0.76	-0.01	0.29	0.13	0.21	0.61	0.49	0.46	0.24	0.03	0.54	0.85
1971	0.99	0.41	0.38	0.51	0.21	0.45	-0.08	0.91	0.33	0.1	0.16	0.48
1972	0.44	0.31	0.55	0.63	0.2	0.74	0.38	0.66	0.46	0.07	0.65	0.34
1973	1.45	0.83	0.88	1.58	1.06	0.82	2.56	1.6	2.38	1.28	1.23	3.88
1974	3.58	2.26	0.77	1.36	0.79	0.99	1.45	1.06	1.13	1.98	2.78	0.78
1975	1.28	0.55	0.63	0.84	1.34	1.7	0.8	0.87	0.73	0.51	0.7	0.82
1976	1.93	1.87	0.98	0.7	0.7	0.4	0.85	0.96	3.41	5.63	4.52	2.51
1977	3.19	2.21	1.74	1.51	0.88	1.23	1.13	2.05	1.77	0.76	1.09	1.38
1978	2.22	1.44	1.04	1.11	0.98	1.38	1.7	1	1.14	1.21	1.03	0.85
1979	3.55	1.44	1.36	0.9	1.31	1.11	1.21	1.51	1.23	1.75	1.29	1.77
1980	4.88	2.31	2.06	1.75	1.63	1.98	2.79	2.07	1.11	1.51	1.73	2.62
1981	3.22	2.46	2.14	2.26	1.51	1.4	1.76	2.06	1.86	2.22	1.92	2.69
1982	4.97	3.93	3.65	5.42	5.62	4.82	5.15	11.22	5.34	5.18	5.06	10.68
1983	10.88	5.37	4.84	6.33	4.34	3.79	4.94	3.88	3.08	3.32	5.87	4.28
1984	6.35	5.28	4.27	4.33	3.32	3.62	3.28	2.84	2.98	3.49	3.43	4.25
1985	7.42	4.15	3.87	3.08	2.37	2.5	3.48	4.37	3.99	3.8	4.61	6.81

Cuadro 4.1: Inflación Mensual de México (1970 – 1985)

Se cargan las paqueterías y se grafican los datos

```
library(TSA)
library(forecast)
library(lmtest)
```

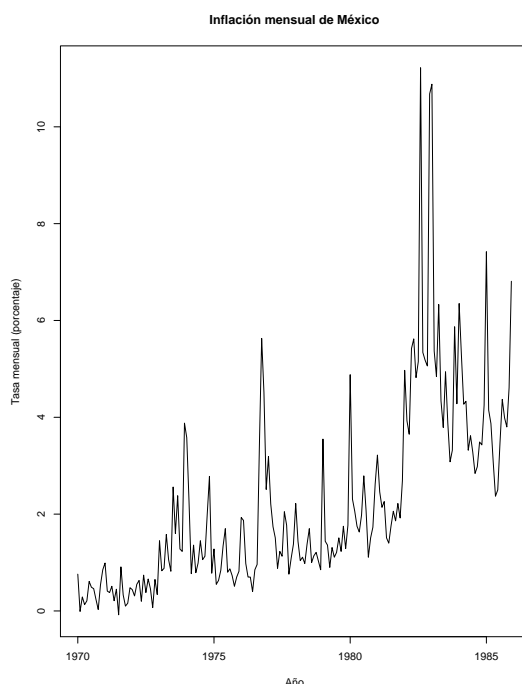


Figura 4.6: Gráfica de la Inflación mensual de México (1970 – 1985)

```
setwd("C:/Users/Jesus Pichardo Leyva/Documents/TESIS")
datos<-read.table("Inflacion Mensual Mexico.csv",sep=" ",header=FALSE)
datos<-datosV1
datos
datos<-ts(datos, frequency = 12, start = c(1970, 1))
datos

plot(datos,main="Inflacion mensual de Mexico",xlab="Año",
      ylab="Tasa mensual (porcentaje)")
```

Se puede observar en la Figura 4.6 que hubo una aparente estabilidad entre 1970 y 1975 en los cuales no hubo movimientos bruscos. Sin embargo, se encuentra que en septiembre de 1976 hubo una devaluación del peso. En enero de 1980 entró en vigor la Ley del impuesto al valor agregado (IVA), el cual pudiera tener algún efecto en la inflación. Posteriormente, en febrero de 1982 se presenta nuevamente otra devaluación, que sin duda hizo que la inflación aumentara fuertemente, en esta ocasión fue debido a una crisis de divisas. Finalmente, en agosto de 1982 se presentó otra devaluación, este evento se dio debido a que trascendió que México no podía pagar su deuda externa. Al término de 1983 el dólar llegó a 161 pesos;

a 210 en 1984. Sin duda, fueron años complicados.

Ahora, se ajustará el mejor modelo ARIMA a los datos cronológicos de acuerdo al Criterio de Información Bayesiano. De esta manera se empezará con el algoritmo

```
fit<-forecast::auto.arima(datos,allowdrift = F,allowmean = F,
                          seasonal = F, ic="bic")
fit
Series: datos
ARIMA(1,1,1)

Coefficients:
          ar1          ma1
      0.3961   -0.8223
s.e.  0.1201    0.0802

sigma^2 estimated as 1.287:  log likelihood=-294.39
AIC=594.78   AICc=594.91   BIC=604.54
```

Como se puede observar, el mejor modelo estimado es un ARIMA(1,1,1). Sin embargo, la sigma cuadrada estimada se puede considerar alta, esta situación puede ocasionar que los pronósticos no sean tan buenos y que los intervalos de confianza sean más grandes.

Se empezará con el modelo ajustado, posteriormente se usarán los comandos detectAO y detectIO para encontrar posibles outliers o datos atípicos de este estilo.

El código es el siguiente

```
datos.aj<-arima(datos,order = c(1,1,1),include.mean = F)
datos.aj

Call:
arima(x = datos, order = c(1, 1, 1), include.mean = F)

Coefficients:
          ar1          ma1
      0.3961   -0.8223
s.e.  0.1201    0.0802
```

```

sigma^2 estimated as 1.274: log likelihood = -294.39, aic = 592.78
detectAO(datos.aj)
      [,1]      [,2]      [,3]      [,4]      [,5]
ind    145.000000 148.0000 152.000000 156.000000 158.000000
lambda2 4.222532 3.8994 5.538398 3.940036 -4.83548
detectIO(datos.aj)
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
ind    82.000000 121.000000 152.000000 156.000000 158.000000 181.000000
lambda1 3.844425 3.831123 7.682284 6.416787 -4.15297 3.998698

```

Se detectan ocho outliers. Sin embargo, sólo se tomará el que tenga la lambda más grande en valor absoluto; en este caso será el dato 152 y corresponde a un dato atípico innovativo, ya que fue el más grande en ambos casos (IO y AO). Se volverá a estimar el modelo, pero ya con este dato atípico incluido. Posteriormente, se aplicará la detección de outliers nuevamente.

```

datos.aj.out<-arimax(datos,order = c(1,1,1),include.mean=F,io=c(152))
datos.aj.out

```

Call:

```

arimax(x = datos, order = c(1, 1, 1), include.mean = F, io = c(152))

```

Coefficients:

```

      ar1      ma1 IO-152
      0.4661 -0.9000 6.8592
s.e. 0.0893 0.0455 1.0372

```

```

sigma^2 estimated as 1.036: log likelihood = -274.82, aic = 555.64
detectAO(datos.aj.out)
      [,1]      [,2]      [,3]
ind    145.000000 156.000000 158.000000
lambda2 4.01167 4.127616 -4.021041
detectIO(datos.aj.out)
      [,1]      [,2]      [,3]      [,4]      [,5]
ind    82.000000 121.000000 156.000000 158.000000 181.000000
lambda1 4.056687 4.060888 7.299513 -3.75124 4.045913

```

Dado que el dato 152 tuvo el valor lambda mayor en valor absoluto en ambos métodos, ya fue incluido en el modelo. Ahora se incluirá al dato que tenga el

índice 156 como outlier innovativo, se estima nuevamente el modelo y se aplica la detección de datos atípicos.

```
datos.aj.out1<-arimax(datos,order = c(1,1,1),include.mean=F,io=c(152,156))
datos.aj.out1
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), include.mean = F, io = c(152, 156))
```

Coefficients:

	ar1	ma1	I0-152	I0-156
	0.4912	-0.8637	6.0739	3.2771
s.e.	0.1120	0.0677	0.8163	0.8432

sigma² estimated as 0.9305: log likelihood = -264.41, aic = 536.81

```
detectA0(datos.aj.out1)
```

	[,1]	[,2]
ind	145.000000	158.000000
lambda2	4.253708	-4.228355

```
detectI0(datos.aj.out1)
```

	[,1]	[,2]	[,3]	[,4]	[,5]
ind	82.000000	121.000000	157.000000	158.000000	181.000000
lambda1	3.999971	4.146828	5.835907	-4.165012	4.246665

Ahora, en este caso, se tomará el dato 157 como un dato atípico innovativo. Se ajusta nuevamente los datos y se agrega este outlier al modelo. Posteriormente, se aplica la detección hasta que no haya más outliers.

```
datos.aj.out2<-arimax(datos,order = c(1,1,1),include.mean=F,io=c(152,156,157))
datos.aj.out2
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), include.mean = F, io = c(152, 156, 157))
```

Coefficients:

	ar1	ma1	I0-152	I0-156	I0-157
	0.5191	-0.8752	6.1278	5.7089	5.8163
s.e.	0.1011	0.0591	0.7157	0.7928	0.7927

sigma² estimated as 0.7247: log likelihood = -240.54, aic = 491.09


```

detectA0(datos.aj.out2)
      [,1]      [,2]
ind    145.000000 148.000000
lambda2  4.434543  3.657684
detectI0(datos.aj.out2)
      [,1]      [,2]      [,3]      [,4]
ind    48.000000 82.000000 121.000000 181.000000
lambda1  3.700692 4.328375  4.543022  4.653392

```

Se agrega el dato 181 como dato atípico de tipo innovativo al modelo, se ajusta nuevamente y se aplican los comandos para la detección

```

datos.aj.out3<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181))

```

```
datos.aj.out3
```

Call:

```

arimax(x = datos, order = c(1, 1, 1), include.mean = F,
       io = c(152, 156, 157, 181))

```

Coefficients:

	ar1	ma1	I0-152	I0-156	I0-157	I0-181
	0.5367	-0.8641	6.1022	5.6762	5.7829	3.2662
s.e.	0.1171	0.0750	0.6661	0.7414	0.7412	0.6670

sigma² estimated as 0.6445: log likelihood = -229.3, aic = 470.61

```

detectA0(datos.aj.out3)
      [,1]      [,2]      [,3]
ind    81.000000 145.000000 192.000000
lambda2  3.692554  4.445476  3.707554
detectI0(datos.aj.out3)
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
ind    48.000000 82.000000 109.000000 121.000000 145.000000 192.000000
lambda1  3.883151 4.467222  3.771659  4.775946  3.708053  3.707554

```

Se toma la lambda más grande en valor absoluto, en este caso corresponde al dato 121. Se considera éste dentro del modelo como dato atípico innovativo y se vuelve a estimar.

```
datos.aj.out4<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121))
```

```
datos.aj.out4
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), include.mean = F,
       io = c(152, 156, 157, 181, 121))
```

Coefficients:

	ar1	ma1	I0-152	I0-156	I0-157	I0-181	I0-121
	0.5662	-0.8636	6.0820	5.6532	5.7582	3.2581	2.8724
s.e.	0.1191	0.0772	0.6235	0.6976	0.6975	0.6244	0.6229

sigma² estimated as 0.5807: log likelihood = -219.31, aic = 452.62

```
detectA0(datos.aj.out4)
```

	[,1]	[,2]	[,3]
ind	81.000000	145.0000	192.00000
lambda2	3.795068	4.4733	3.81076

```
detectI0(datos.aj.out4)
```

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
ind	48.00000	81.000000	82.000000	109.000000	145.000000	192.00000
lambda1	4.03886	3.766475	4.531251	3.941863	3.829504	3.81076

Se observa los valores de las distintas lambda y se ve que la lambda más grande es la que tiene el índice 82 y corresponde a un dato atípico innovativo. Se agrega al modelo la identificación de este outlier y se reestima el modelo.

```
datos.aj.out5<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82))
```

```
datos.aj.out5
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), include.mean = F,
       io = c(152, 156, 157, 181, 121, 82))
```

Coefficients:

	ar1	ma1	I0-152	I0-156	I0-157	I0-181	I0-121	I0-82
	0.4851	-0.8172	6.0889	5.6415	5.7525	3.2995	2.8924	2.0472
s.e.	0.1446	0.0997	0.6162	0.6821	0.6816	0.6206	0.6155	0.6453

```

sigma^2 estimated as 0.5501:  log likelihood = -214.13,  aic = 444.26
> detectA0(datos.aj.out5)
      [,1]      [,2]      [,3]
ind    81.000000 145.000000 192.000000
lambda2 3.826635 4.475055 3.918855
> detectIO(datos.aj.out5)
      [,1]      [,2]      [,3]      [,4]      [,5]
ind    48.000000 81.000000 109.000000 145.000000 192.000000
lambda1 4.014384 3.840971 3.998568 3.914797 3.918855

```

Ahora, se incluirá al dato que tenga el índice 145 como outlier aditivo, se estima nuevamente el modelo y se aplica la detección de datos atípicos.

```

datos.aj.out6<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82),
xreg = data.frame(A0=seq(datos)==145))
datos.aj.out6

```

Call:

```

arimax(x = datos, order = c(1, 1, 1),
       xreg = data.frame(A0 = seq(datos) == 145),
       include.mean = F, io = c(152, 156, 157, 181, 121, 82))

```

Coefficients:

```

      ar1      ma1      A0  IO-152  IO-156  IO-157  IO-181  IO-121
0.5153 -0.8278 1.6431 6.0881 5.6434 5.7517 3.2843 2.8852
s.e. 0.1471 0.1015 0.5978 0.5994 0.6669 0.6664 0.6031 0.5987

      IO-82
2.0004
s.e 0.6260

```

```

sigma^2 estimated as 0.5294:  log likelihood = -210.46,  aic = 438.92
detectA0(datos.aj.out6)
      [,1]      [,2]      [,3]
ind    81.000000 148.000000 192.000000
lambda2 3.785489 3.746262 3.900247

```

```
detectIO(datos.aj.out6)
      [,1]      [,2]      [,3]      [,4]
ind      48.000000 81.000000 109.000000 192.000000
lambda1  4.045222  3.844888  4.016593  3.900247
```

Se analizan los outliers y se agrega el dato 48, como dato atípico innovativo, al modelo. Posteriormente, se vuelve a cargar el algoritmo de detección.

```
datos.aj.out7<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82,48),
                      xreg = data.frame(A0=seq(datos)==145))
datos.aj.out7
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), xreg = data.frame(A0 = seq(datos) == 145),
       include.mean = F, io = c(152, 156, 157, 181, 121, 82, 48))
```

Coefficients:

```
      ar1      ma1      A0  IO-152  IO-156  IO-157  IO-181  IO-121  IO-82
      0.5007 -0.8174 1.6422  6.0876  5.6387  5.7481  3.2927  2.8885  2.0169
s.e.  0.1537  0.1073 0.5873  0.5889  0.6542  0.6536  0.5936  0.5883  0.6191

      IO-48
      1.617
s.e  0.591
```

sigma² estimated as 0.5093: log likelihood = -206.75, aic = 433.49

```
detectA0(datos.aj.out7)
      [,1]      [,2]      [,3]
ind      81.000000 148.000000 192.000000
lambda2  3.852518  3.748013  3.936633
detectIO(datos.aj.out7)
      [,1]      [,2]      [,3]
ind      81.000000 109.000000 192.000000
lambda1  3.880081  4.049875  3.936633
```

Ahora, se incluirá al dato que tenga el índice 109 como outlier innovativo, se estima nuevamente el modelo y se aplica la detección de datos atípicos.

```
datos.aj.out8<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82,48,109),
xreg = data.frame(A0=seq(datos)==145))
datos.aj.out8
```

Call:

```
arimax(x = datos, order = c(1, 1, 1),
       xreg = data.frame(A0 = seq(datos) == 145),
       include.mean = F, io = c(152, 156, 157, 181, 121, 82, 48, 109))
```

Coefficients:

	ar1	ma1	A0	I0-152	I0-156	I0-157	I0-181	I0-121
	0.5617	-0.8346	1.6467	6.0659	5.6235	5.7288	3.2666	2.8779
s.e.	0.1531	0.1067	0.5501	0.5513	0.6180	0.6178	0.5543	0.5509
	I0-82	I0-48	I0-109					
	1.9269	1.5828	2.3863					
s.e.	0.5744	0.5531	0.5501					

sigma² estimated as 0.4645: log likelihood = -197.95, aic = 417.9

```
detectA0(datos.aj.out8)
```

	[,1]	[,2]	[,3]
ind	81.00000	148.00000	192.00000
lambda2	3.89703	3.751102	4.005915

```
detectI0(datos.aj.out8)
```

	[,1]	[,2]	[,3]
ind	81.00000	167.00000	192.00000
lambda1	4.004958	3.670463	4.005915

Se observa los valores de las distintas lambda y se ve que la lambda más grande es la que tiene el índice 192 y corresponde a un dato atípico innovativo. Se agrega al modelo la identificación de este outlier y se reestima el modelo.

```
datos.aj.out9<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82,48,109,192),
xreg = data.frame(A0=seq(datos)==145))
datos.aj.out9
```

Call:

```
arimax(x = datos, order = c(1, 1, 1),
       xreg = data.frame(A0 = seq(datos) == 145),
       include.mean = F, io = c(152, 156, 157, 181, 121, 82, 48, 109, 192))
```

Coefficients:

	ar1	ma1	A0	I0-152	I0-156	I0-157	I0-181	I0-121	I0-82
	0.5179	-0.8185	1.6439	6.0782	5.6304	5.7386	3.3004	2.8878	1.9863
s.e.	0.1361	0.0921	0.5373	0.5387	0.5998	0.5995	0.5415	0.5380	0.5632
	I0-192	I0-48	I0-109						
	2.4905	1.6047	2.3849						
s.e.	0.6610	0.5404	0.5373						

sigma² estimated as 0.4324: log likelihood = -191.11, aic = 406.21

detectA0(datos.aj.out9)

[,1] [,2]

ind 81.000000 148.000000

lambda2 4.085486 3.92351

detectI0(datos.aj.out9)

[,1] [,2]

ind 81.000000 167.000000

lambda1 4.132457 3.741842

Es momento de incluir al dato que tenga el índice 81 como outlier innovativo, se estima nuevamente el modelo y se aplica la detección de datos atípicos.

```
datos.aj.out10<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82,48,109,192,81),
                      xreg = data.frame(A0=seq(datos)==145))
datos.aj.out10
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), xreg = data.frame(A0 = seq(datos) == 145),
       include.mean = F, io = c(152, 156, 157, 181, 121, 82, 48, 109, 192, 81))
```

Coefficients:

	ar1	ma1	A0	I0-152	I0-156	I0-157	I0-181	I0-121	I0-82
	0.4868	-0.8028	1.6422	6.0831	5.6290	5.7396	3.3164	2.8950	2.6632
s.e.	0.1351	0.0927	0.5321	0.5337	0.5915	0.5911	0.5370	0.5329	0.6079

	I0-192	I0-81	I0-48	I0-109
	2.5049	1.4855	1.6190	2.3849
s.e.	0.6503	0.5929	0.5356	0.5321

sigma² estimated as 0.4186: log likelihood = -188, aic = 402

detectA0(datos.aj.out10)

[,1]

ind 148.000000

lambda2 3.970685

detectI0(datos.aj.out10)

[,1]

ind 167.000000

lambda1 3.771872

Se toma la lambda más grande en valor absoluto, en este caso corresponde al dato 148. Se considera esto dentro del modelo como dato atípico aditivo y se vuelve a estimar.

```
datos.aj.out11<-arimax(datos,order = c(1,1,1),include.mean=F,
                        io=c(152,156,157,181,121,82,48,109,192,81),
xreg = data.frame(A0=seq(datos)==145,A0=seq(datos)==148))
datos.aj.out11
```

Call:

```
arimax(x = datos, order = c(1, 1, 1),
       xreg = data.frame(A0 = seq(datos) == 145,
                        A0 = seq(datos) == 148), include.mean = F, io = c(152, 156, 157,
181, 121, 82, 48, 109, 192, 81))
```

Coefficients:

	ar1	ma1	A0	A0.1	I0-152	I0-156	I0-157	I0-181
	0.4799	-0.8000	1.6526	0.9015	6.0932	5.6348	5.7458	3.3202
s.e.	0.1354	0.0927	0.5290	0.5310	0.5308	0.5876	0.5872	0.5342
	I0-109	I0-192	I0-81	I0-121	I0-82	I0-48		
	2.3849	2.5085	1.4898	2.8968	2.6722	1.6226		
s.e.	0.5290	0.6455	0.5888	0.5299	0.6045	0.5327		

sigma² estimated as 0.4123: log likelihood = -186.56, aic = 401.12

```
detectA0(datos.aj.out11)
[1] "No A0 detected"
detectIO(datos.aj.out11)
      [,1]      [,2]
ind      83.000000 167.000000
lambda1  3.657378  3.779798
```

Se observa que ya no hay outliers aditivos. Por otro lado, todavía hay datos atípicos innovativos, ahora se añadirá el índice 167, como dato atípico innovativo, al modelo.

```
datos.aj.out12<-arimax(datos,order = c(1,1,1),include.mean=F,
                      io=c(152,156,157,181,121,82,48,109,192,81,167),
                      xreg = data.frame(A0=seq(datos)==145,A0=seq(datos)==148))
datos.aj.out12
```

Call:

```
arimax(x = datos, order = c(1, 1, 1), xreg = data.frame(A0 = seq(datos) == 145,
              A0 = seq(datos) == 148), include.mean = F, io = c(152, 156, 157, 181, 121,
              82, 48, 109, 192, 81, 167))
```

Coefficients:

	ar1	ma1	A0	A0.1	IO-152	IO-156	IO-157	IO-181	IO-121
	0.5739	-0.8413	1.6535	0.8703	6.0708	5.6349	5.7397	3.2778	2.8755
s.e.	0.1242	0.0819	0.4968	0.4982	0.4983	0.5591	0.5590	0.4994	0.4974
	IO-109	IO-192	IO-81	IO-167	IO-82	IO-48			
	2.3865	2.4595	1.4339	2.0155	2.5605	1.5782			
s.e.	0.4969	0.6207	0.5611	0.4986	0.5714	0.4993			

sigma² estimated as 0.3808: log likelihood = -178.96, aic = 387.93

```
detectA0(datos.aj.out12)
[1] "No A0 detected"
detectIO(datos.aj.out12)
      [,1]
ind      169.000000
lambda1  3.661568
```

Se agrega al modelo el índice 169 que corresponde a un dato atípico innovativo. Se considera éste dentro del modelo y se vuelve a estimar.


```

datos.aj.out13<-arimax(datos,order = c(1,1,1),include.mean=F,
                        io=c(152,156,157,181,121,82,48,109,192,81,167,169),
                        xreg = data.frame(A0=seq(datos)==145,A0=seq(datos)==148))
datos.aj.out13

```

Call:

```

arimax(x = datos, order = c(1, 1, 1), xreg = data.frame(A0 = seq(datos) == 145,
  A0 = seq(datos) == 148), include.mean = F,
  io = c(152, 156, 157, 181, 121,
    82, 48, 109, 192, 81, 167, 169))

```

Coefficients:

	ar1	ma1	A0	A0.1	I0-152	I0-156	I0-157	I0-181
	0.5601	-0.8263	1.6539	0.8704	6.0684	5.6291	5.7352	3.2859
s.e.	0.1376	0.0940	0.4818	0.4833	0.4833	0.5419	0.5417	0.4853
	I0-121	I0-82	I0-48	I0-109	I0-192	I0-81	I0-167	I0-169
	2.8779	2.5671	1.5797	2.3880	2.4600	1.4354	2.0255	1.6573
s.e.	0.4825	0.5563	0.4847	0.4819	0.6026	0.5440	0.4828	0.4854

sigma² estimated as 0.3587: log likelihood = -173.25, aic = 378.49

```
detectA0(datos.aj.out13)
```

```
[1] "No A0 detected"
```

```
detectI0(datos.aj.out13)
```

```
[1] "No I0 detected"
```

Ya no hay más outliers, el modelo ajustado es de la siguiente forma

$$\begin{aligned}
Y_t = & \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (1.5797) I_t(48) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (1.4354) I_t(81) + \\
& \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (2.5671) I_t(82) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (2.3880) I_t(109) + \\
& \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (2.8779) I_t(121) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (6.0684) I_t(152) + \\
& \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (5.6291) I_t(156) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (5.7352) I_t(157) + \\
& \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (2.0255) I_t(167) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (1.6573) I_t(169) + \\
& \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (3.2859) I_t(181) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} (2.4600) I_t(192) + \\
& (1.6539) I_t(145) + (0.8704) I_t(148) + \frac{(1 - 0.8263B)}{(1 - B)(1 - 0.5601B)} W_t
\end{aligned}$$

Se puede notar que el outlier 48 se debe a diciembre de 1973 y esta se debió a una elevación de los precios internacionales de diversos productos, la cual elevó la inflación. Los datos atípicos 81 y 82 son los datos de septiembre y octubre de 1976 y se deben a un anuncio del gobierno de devaluar la moneda nacional, la cual provocó un aumento de la inflación. El outlier 109 corresponde al mes de enero de 1979, en ese mes hubo un aumento de la inflación debido a una política económica de aumentar los salarios de los trabajadores. El dato atípico 121 pertenece a enero de 1980, en ese mes entró en vigor el Impuesto al Valor Agregado (IVA) y esto ocasionó un fuerte incremento inflacionario. Los datos atípicos de 145 y 148 están ligados a los meses de enero y abril de 1981, en enero hubo un nuevo aumento salarial a los trabajadores, lo más probable es que ese aumento salarial no sólo afectó a la inflación del mes mencionado, sino también al mes de abril debido a que su efecto puede permanecer todavía activo, aunque de forma más ligera. Los outliers 152, 156 y 157 se deben a las devaluaciones de agosto y diciembre de 1982, el outlier 157 se debe al efecto que tuvieron esas devaluaciones y que afectaron al dato inflacionario de enero de 1983, la inflación en esos meses tuvo un aumento de tipo momentáneo; sin embargo, fueron muy altos esos aumentos. Los datos atípicos 167 y 169 corresponden a los meses de noviembre de 1983 y enero de 1984, respectivamente, en estos meses hubo un aumento en los precios de los productos básicos, lo cual ocasionó un aumento inflacionario pero de baja intensidad. Por último, el outlier 192 se refiere al mes de diciembre de 1985, en

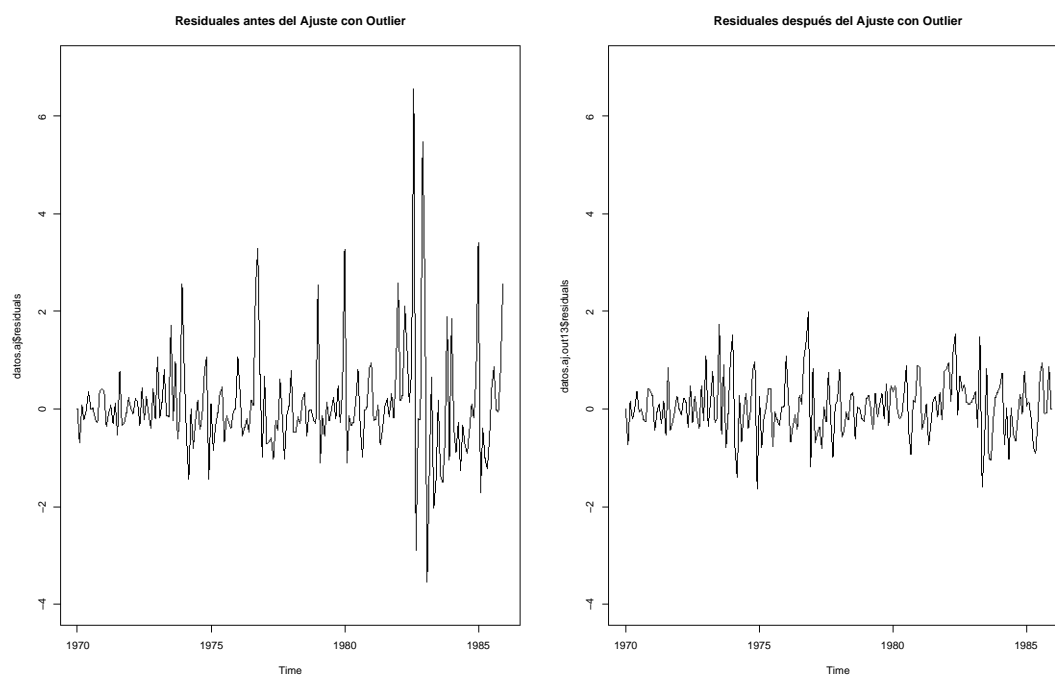


Figura 4.7: Residuos del primer y último modelo

este caso se debió a un ajuste de los precios, la inflación subió, pero no fue de gran impacto.

Ahora, se calcularán y graficarán los residuos del primer y último modelo para hacer la comparación.

```
plot(datos.aj$residuals,ylim=c(-4,7),
     main='Residuales antes del Ajuste con Outlier')
plot(datos.aj.out13$residuals,ylim=c(-4,7),
     main='Residuales despues del Ajuste con Outlier')
```

De acuerdo a la Figura 4.7 la variabilidad de los residuos es menor después de considerar la presencia de datos atípicos, lo cual es muy bueno, por que nos permite tener pronósticos más reales e intervalos de confianza más pequeños. Es momento del pronóstico

```
L<-length(datos)
L
[1] 192
```

```
It_48<-c(rep(0,47),rep(1,1),rep(0,L-48))
It_81<-c(rep(0,80),rep(1,1),rep(0,L-81))
It_82<-c(rep(0,81),rep(1,1),rep(0,L-82))
It_109<-c(rep(0,108),rep(1,1),rep(0,L-109))
It_121<-c(rep(0,120),rep(1,1),rep(0,L-121))
It_145<-c(rep(0,144),rep(1,1),rep(0,L-145))
It_148<-c(rep(0,147),rep(1,1),rep(0,L-148))
It_152<-c(rep(0,151),rep(1,1),rep(0,L-152))
It_156<-c(rep(0,155),rep(1,1),rep(0,L-156))
It_157<-c(rep(0,156),rep(1,1),rep(0,L-157))
It_167<-c(rep(0,166),rep(1,1),rep(0,L-167))
It_169<-c(rep(0,168),rep(1,1),rep(0,L-169))
It_181<-c(rep(0,180),rep(1,1),rep(0,L-181))
It_192<-c(rep(0,191),rep(1,1),rep(0,L-192))

h<-12
phi1<-datos.aj.out13[[1]][1]
phi1
  ar1
0.5600978
theta1<-datos.aj.out13[[1]][2]
theta1
  ma1
-0.8262929
ao_145<-datos.aj.out13[[1]][3]
ao_145
  A0
1.653866
ao_148<-datos.aj.out13[[1]][4]
ao_148
  A0.1
0.8704384
io_152<-datos.aj.out13[[1]][5]
io_152
  I0-152
6.068361
io_156<-datos.aj.out13[[1]][6]
io_156
  I0-156
5.62914
```

```
io_157<-datos.aj.out13[[1]][7]
io_157
IO-157
5.73515
io_181<-datos.aj.out13[[1]][8]
io_181
IO-181
3.285884
io_121<-datos.aj.out13[[1]][9]
io_121
IO-121
2.877866
io_82<-datos.aj.out13[[1]][10]
io_82
IO-82
2.567074
io_48<-datos.aj.out13[[1]][11]
io_48
IO-48
1.579689
io_109<-datos.aj.out13[[1]][12]
io_109
IO-109
2.38803
io_192<-datos.aj.out13[[1]][13]
io_192
IO-192
2.459966
io_81<-datos.aj.out13[[1]][14]
io_81
IO-81
1.435443
io_167<-datos.aj.out13[[1]][15]
io_167
IO-167
2.025507
io_169<-datos.aj.out13[[1]][16]
io_169
IO-169
1.657323
res<-as.vector(residuals(datos.aj.out12))
```

```

resout<-c(res,rep(0,h))
Yt<-c(datos,rep(0,h))
L<-length(datos)

It_48<-c(rep(0,47),rep(1,1),rep(0,L-48+12))
It_81<-c(rep(0,80),rep(1,1),rep(0,L-81+12))
It_82<-c(rep(0,81),rep(1,1),rep(0,L-82+12))
It_109<-c(rep(0,108),rep(1,1),rep(0,L-109+12))
It_121<-c(rep(0,120),rep(1,1),rep(0,L-121+12))
It_145<-c(rep(0,144),rep(1,1),rep(0,L-145+12))
It_148<-c(rep(0,147),rep(1,1),rep(0,L-148+12))
It_152<-c(rep(0,151),rep(1,1),rep(0,L-152+12))
It_156<-c(rep(0,155),rep(1,1),rep(0,L-156+12))
It_157<-c(rep(0,156),rep(1,1),rep(0,L-157+12))
It_167<-c(rep(0,166),rep(1,1),rep(0,L-167+12))
It_169<-c(rep(0,168),rep(1,1),rep(0,L-169+12))
It_181<-c(rep(0,180),rep(1,1),rep(0,L-181+12))
It_192<-c(rep(0,191),rep(1,1),rep(0,L-192+12))

for(i in (L+1):(L+h)){
  Yt[i]<-(phi1+1)*Yt[i-1]-phi1*Yt[i-2]+io_48*It_48[i]+io_81*It_81[i]+
    io_82*It_82[i]+io_109*It_109[i]+io_121*It_121[i]+
io_152*It_152[i]+io_156*It_156[i]+ io_157*It_157[i]+
io_167*It_167[i]+io_169*It_169[i]+io_181*It_181[i]+
io_192*It_192[i]+(io_48*theta1)*It_48[i-1]+
(io_81*theta1)*It_81[i-1]+(io_82*theta1)*It_82[i-1]+
(io_109*theta1)*It_109[i-1]+(io_121*theta1)*It_121[i-1]+
(io_152*theta1)*It_152[i-1]+(io_156*theta1)*It_156[i-1]+
(io_157*theta1)*It_157[i-1]+(io_167*theta1)*It_167[i-1]+
    (io_169*theta1)*It_169[i-1]+(io_181*theta1)*It_181[i-1]+
(io_192*theta1)*It_192[i-1]+ao_145*It_145[i]+
ao_148*It_148[i]-(phi1+1)*ao_145*It_145[i-1]-
(phi1+1)*ao_148*It_148[i-1]+(ao_145*phi1)*It_148[i-2]-
(phi1*ao_148)*It_145[i-2]+resout[i]+theta1*resout[i-1]
}

plot(Yt[1:L],type="l",pch=16,cex=0.5,xlim = c(0,205),ylim=c(0,12),
      xlab='Time',ylab='Tasa de Inflacion (porcentaje)',
      main="Pronostico Inflacion mensual de Mexico")

```

```

lines(193:204,Yt[193:204],col="blue")
lines(193:204,Yt[193:204]+ 1.96*sqrt(datos.aj.out3$sigma2),lty=2,col="red")
lines(193:204,Yt[193:204]- 1.96*sqrt(datos.aj.out3$sigma2),lty=2,col="red")
lines(193:204,Yt[193:204]+ 2.576*sqrt(datos.aj.out3$sigma2),lty=2,col="purple")
lines(193:204,Yt[193:204]- 2.576*sqrt(datos.aj.out3$sigma2),lty=2,col="purple")

```

```
Yt[193:204]
```

```

[1] 6.009622 5.561332 5.310246 5.169614 5.090845 5.046728 5.022017 5.008177
[11] 5.000425 4.996083 4.993652 4.992290

```

De acuerdo al pronóstico (Figura 4.8), se observa que la inflación estará entre 6 y 5 por ciento por mes, tendrá un ligero decaimiento. Si comparamos con los datos originales, los cuales son

```
8.84 4.45 4.65 5.22 5.56 6.42 4.99 7.97 6 5.72 6.76 7.9
```

se observa cierta similitud usando los pronósticos y los intervalos de confianza. Principalmente, la finalidad de esta serie fue la de cuantificar los efectos que pueden tener decisiones gubernamentales en la inflación y encontrar un modelo que ayude a predecir lo mejor posible.

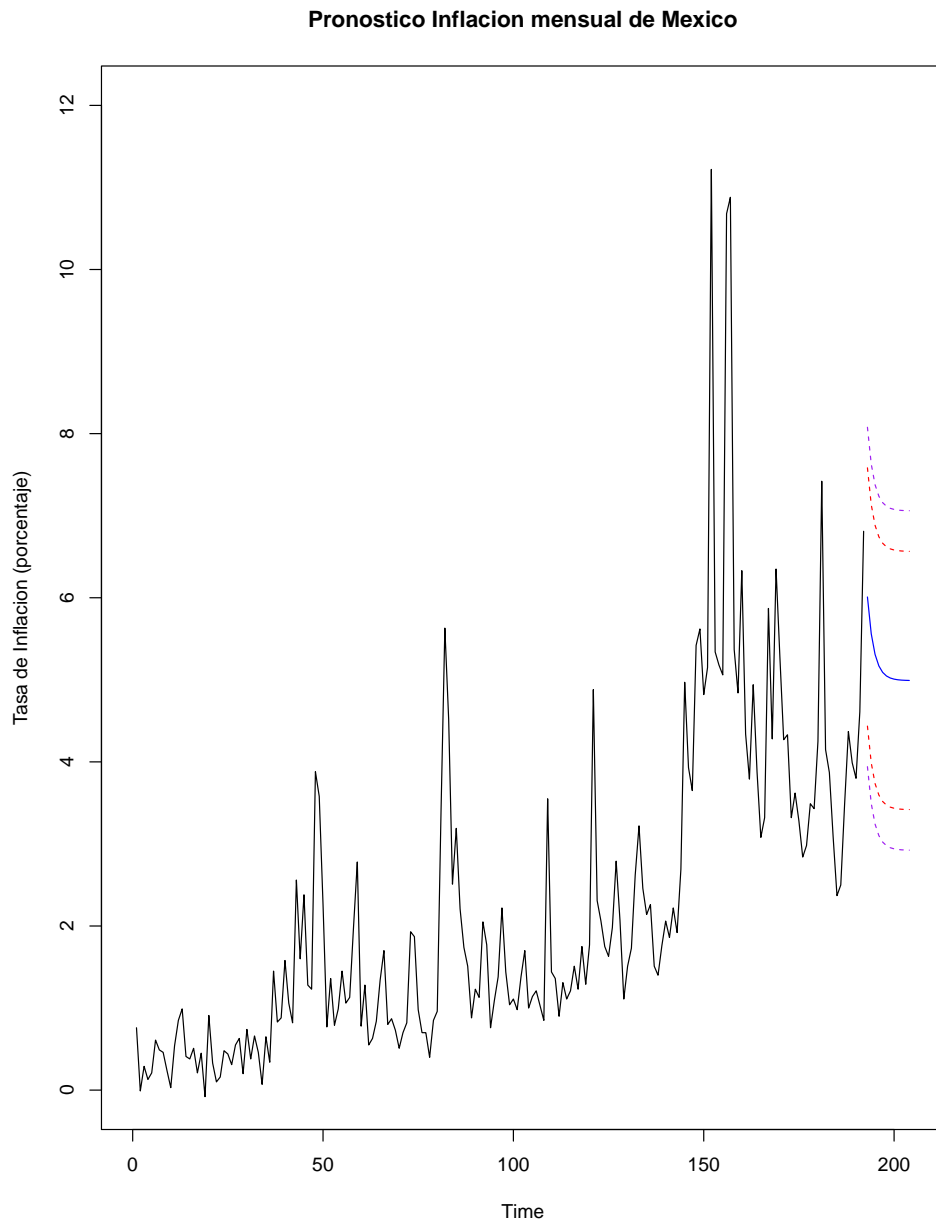


Figura 4.8: Pronóstico a 12 meses de la Inflación mensual Mexicana

4.4. Teoría de acuerdo a Chen y Liu

La teoría de Chen y Liu (1993) se basa en un procedimiento iterativo de detección de outliers y ajuste para obtener estimaciones conjuntas de los parámetros del modelo y de los efectos de los outliers. Una de las principales diferencias de este método con otros, como el de Chang y Tiao (1988) y Tsay (1986) son:

- Los efectos de los outliers son estimados simultáneamente usando teoría de regresión.
- Los parámetros del modelo y los efectos de los outliers son estimados conjuntamente.

Se ha demostrado mediante técnicas de simulación que este procedimiento mejora la estimaciones, ya que evita efectos de propagación (la presencia de algunos atípicos produce que observaciones no atípicas parezcan serlo) y de enmascaramiento (la presencia de algunos atípicos enmascara la presencia de otros).

Para lograr la detección de outliers y la estimación de los parámetros conjuntamente el procedimiento propuesto consta de tres pasos. Usando 4.5 y 4.6 se tiene

4.4.1. Algoritmo de detección, estimación y pronóstico

PASO 1. Localizar outliers.

1.1. Ajustar un modelo ARIMA a los datos y obtener los residuos usando 4.6.

1.2. Para $t = 1, 2, \dots, n$, calcular $\hat{\tau}_{IO}(t)$, $\hat{\tau}_{AO}(t)$, $\hat{\tau}_{LS}(t)$ y $\hat{\tau}_{TC}(t)$ en 4.32 usando los residuos obtenidos en 1.1, entonces sea

$$\eta_t = \max \{ |\hat{\tau}_{IO}(t)|, |\hat{\tau}_{AO}(t)|, |\hat{\tau}_{LS}(t)|, |\hat{\tau}_{TC}(t)| \}$$

Si $\max_t \eta_t = |\hat{\tau}_k(t_j)| > C$, donde C es un valor crítico predeterminado, entonces hay posibilidad de un outlier tipo k , donde $k = IO, AO, LS$ ó TC .

PASO 2. Remover outliers.

2.1. Dado un conjunto de posibles outliers encontrados en 1.2, un modelo ARIMA es elegido y ajustado de acuerdo a la ecuación de regresión 4.5. La significancia de los outliers es evaluada nuevamente en el nuevo modelo ajustado. Los outliers que no sean significativos son removidos del modelo, es decir, si $\min_j |\hat{\tau}_j| = \hat{\tau}_v \leq C$ para $j = 1, \dots, m$, donde $\hat{\tau}_j = \frac{\hat{\omega}_j}{std(\hat{\omega}_j)}$ y C es el mismo valor crítico usado en 1.2.

PASO 3.

3.1. Ajustar la serie removiendo los outliers significativos basándose en 2.1.

3.2. Iterar, primero para la serie original y luego para la serie ajustada en 3.1. Esto se realiza con el fin de ver si hay más outliers. En caso de que ya no haya atípicos, acabamos con la iteración y el modelo ajustado en 2.1 es el correcto y listo para hacer el pronóstico. Por el contrario, si existieran algunos outliers, éstos son añadidos al conjunto de outliers encontrados en la iteración previa.

Localizar Outliers

Sustituyendo en la ecuación 4.6 $L(B)$, por la definición de cada tipo de outlier y multiplicando por $\pi(B)$, se puede ver que la ecuación para el residuo llega a ser, respectivamente para cada tipo de outlier.

$$IO : \hat{e}_t = \omega I_t(t_j) + W_t \quad (4.28)$$

$$AO : \hat{e}_t = \omega \pi(B) I_t(t_j) + W_t \quad (4.29)$$

$$TC : \hat{e}_t = \omega \frac{\pi(B)}{1 - \delta B} I_t(t_j) + W_t \quad (4.30)$$

y

$$LS : \hat{e}_t = \omega \frac{\pi(B)}{1 - B} I_t(t_j) + W_t \quad (4.31)$$

Ahora, un posible enfoque para detectar outliers, es examinar el valor máximo de la siguiente estadística t

$$\hat{\tau}_k = \frac{\hat{\omega}_k(t_j)}{\hat{\sigma}_W \left(\sum_{i=t_j}^n x_{k,i}^2 \right)} \quad (4.32)$$

con

$$\hat{\omega}_k = \frac{\sum_{i=t_j}^n \hat{e}_t x_{k,i}}{\sum_{i=t_j}^n x_{k,i}^2} \quad (4.33)$$

donde k significa para cada tipo de outlier, $k = \text{IO, AO, TC y LS}$. Por otro lado, el regresor (en este caso, outlier) $x_k = \pi(B) L_j(B) I_t(t_j)$.

En presencia de outliers, los residuos están contaminados; por lo tanto, σ_W puede estar sobreestimado si se usa la desviación típica estándar usual. En este caso, se usará la desviación media absoluta (MAD) definida como:

$$\hat{\sigma}_W = 1.483 \times \text{media}(|\hat{e} - \tilde{e}|)$$

con

$$\tilde{e} = \text{media}(\hat{e})$$

El valor crítico C se obtiene de acuerdo al número de datos de la serie temporal original (n), $\text{cval}=3$ si $n \leq 50$; $\text{cval}=4$ si $n \geq 450$; $\text{cval}=3+0.0025*(n-50)$ en otro caso.

Remove outliers

Aunque la detección de outliers fue hecha en el paso 1, se debe estar consciente que la selección inicial del modelo ARIMA puede ser afectada por la presencia de outliers.

Para evitar encontrar outliers espurios causados por una elección impropia del modelo, la ecuación 4.5 es ajustada a la serie observada incluyendo los regresores (outliers) encontrados en el paso 1, si alguno de los outliers llega a ser no significativo, entonces será removido del conjunto de posible outliers.

Iteración Final

Aquí, se iteran la serie original y la serie ajustada del paso 2 sin los efectos de los outliers significativos. Como se mencionó, en esta parte sólo sirve para inspeccionar si no hay más outliers potenciales. En caso de que no haya más atípicos, el modelo ajustado en el paso 2 es el adecuado.

Como comentario final, ya que en la primera parte de este capítulo se habló del AO e IO, entonces en este subcapítulo sólo se trabajará con TC y LS en el ejemplo. Esto con el único fin de evitar reiterar mucho en el tema y también de explotar lo máximo posible el software R.

4.4.2. Implementación en R

Para la implementación de datos atípicos en Series de Tiempo de acuerdo a Chen y Liu, se usarán los datos mensuales de la Inflación de México. Los datos fueron obtenidos del Sistema de Información Económica del Banco de México y fueron registrados desde el mes de Enero de 1994 al mes de Enero de 2017.

Los datos se muestran en el Cuadro 4.2

Se cargan las paqueterías y se grafican los datos

```
library(forecast)
library(stats)
library(tsoutliers)
setwd("C:/Users/Jesus Pichardo Leyva/Documents/TESIS")
datos<-read.table("Inflacin Mensual Mexico.csv",sep="," ,header=FALSE)
datos<-datos[,2]
datos

yt<-ts(datos, frequency = 12, start = c(1994, 1))
plot(yt)
```

AÑO	E	F	M	A	M	JN	JL	A	S	O	N	D
1994	0.78	0.51	0.51	0.49	0.48	0.5	0.44	0.47	0.71	0.52	0.53	0.88
1995	3.76	4.24	5.9	7.97	4.18	3.17	2.04	1.66	2.07	2.06	2.47	3.26
1996	3.59	2.33	2.2	2.84	1.82	1.63	1.42	1.33	1.6	1.25	1.52	3.2
1997	2.57	1.68	1.24	1.08	0.91	0.89	0.87	0.89	1.25	0.8	1.12	1.4
1998	2.18	1.75	1.17	0.94	0.8	1.18	0.96	0.96	1.62	1.43	1.77	2.44
1999	2.53	1.34	0.93	0.92	0.6	0.66	0.66	0.56	0.97	0.63	0.89	1
2000	1.34	0.89	0.55	0.57	0.37	0.59	0.39	0.55	0.73	0.69	0.86	1.08
2001	0.55	-0.07	0.63	0.5	0.23	0.24	-0.26	0.59	0.93	0.45	0.38	0.14
2002	0.92	-0.06	0.51	0.55	0.2	0.49	0.29	0.38	0.6	0.44	0.81	0.44
2003	0.4	0.28	0.63	0.17	-0.32	0.08	0.14	0.3	0.6	0.37	0.83	0.43
2004	0.62	0.6	0.34	0.15	-0.25	0.16	0.26	0.62	0.83	0.69	0.85	0.21
2005	0	0.33	0.45	0.36	-0.25	-0.1	0.39	0.12	0.4	0.25	0.72	0.61
2006	0.59	0.15	0.13	0.15	-0.45	0.09	0.27	0.51	1.01	0.44	0.52	0.58
2007	0.52	0.28	0.22	-0.06	-0.49	0.12	0.42	0.41	0.78	0.39	0.71	0.41
2008	0.46	0.3	0.72	0.23	-0.11	0.41	0.56	0.58	0.68	0.68	1.14	0.69
2009	0.23	0.22	0.58	0.35	-0.29	0.18	0.27	0.24	0.5	0.3	0.52	0.41
2010	1.09	0.58	0.71	-0.32	-0.63	-0.03	0.22	0.28	0.52	0.62	0.8	0.5
2011	0.49	0.38	0.19	-0.01	-0.74	0	0.48	0.16	0.25	0.67	1.08	0.82
2012	0.71	0.2	0.06	-0.31	-0.32	0.46	0.56	0.3	0.44	0.51	0.68	0.23
2013	0.4	0.49	0.73	0.07	-0.33	-0.06	-0.03	0.28	0.38	0.48	0.93	0.57
2014	0.89	0.25	0.27	-0.19	-0.32	0.17	0.28	0.36	0.44	0.55	0.81	0.49
2015	-0.09	0.19	0.41	-0.26	-0.5	0.17	0.15	0.21	0.37	0.51	0.55	0.41
2016	0.38	0.44	0.15	-0.32	-0.45	0.11	0.26	0.28	0.61	0.61	0.78	0.46
2017	1.7											

Cuadro 4.2: Inflación Mensual de México (1994-2017)

La Figura 4.9 que es la Inflación Mexicana mensual de 1994 a 2017 muestra que hay diversos cambios temporales y de nivel también. Ahora, se ajustará el mejor modelo ARIMA a los datos, esto con el fin de empezar el análisis. Posteriormente, se calculan los residuos y se encuentran los valores atípicos.

```
length(yt)
[1] 277
fit<-forecast::auto.arima(yt,seasonal = F,allowdrift = F,allowmean = F,
                           ic="bic")

fit
Series: yt
ARIMA(1,1,1)

Coefficients:
          ar1          ma1
          0.7710    -0.9671
s.e.      0.0477     0.0175
```

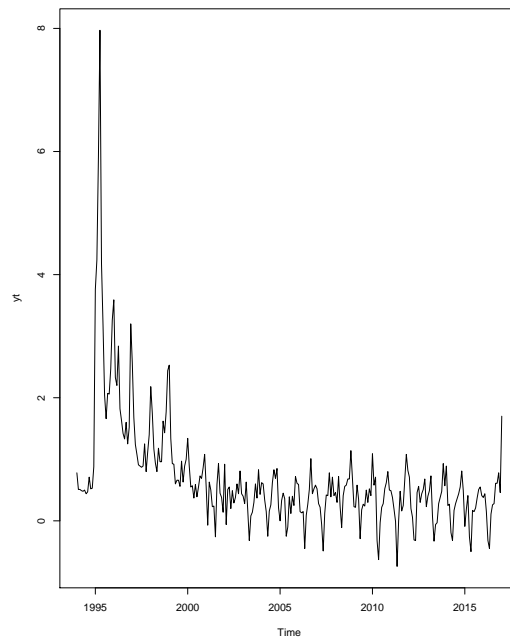


Figura 4.9: Inflación Mensual de México 1994 – 2017

```

sigma^2 estimated as 0.2447:  log likelihood=-196.82
AIC=399.64  AICc=399.73  BIC=410.5
pars<-coefs2poly(fit)
resid<-residuals(fit)

cval<-3.5675
maxit.iloop<-6
n<-277
locate.outliers(resid, pars, cval, types = c("LS","TC"))
  type ind  coefhat    tstat
4   LS  17 -2.001286 -7.151335
5   LS  26 -1.012812 -3.619153
6   LS  62 -1.126369 -4.024932
7   LS 277  1.273135  3.601591
8   TC  13  2.559200  7.359541
9   TC  15  2.173616  6.250709
10  TC  16  3.190397  9.174687
12  TC  36  1.758755  5.057687
stage1<-locate.outliers.oloop(yt, fit, types = c("LS","TC"),
                              cval = cval, maxit.iloop = maxit.iloop)

```

```

stage1
$fit
$fit$coefs
      ar1      ma1
0.7254058 -0.7776169

$fit$pars
$arcoefs
[1] 1.7254058 -0.7254058

$macoefs
[1] -0.7776169

attr(,"class")
[1] "ArimaPars"

$fit$n
[1] 277
$outliers
  type ind   coefhat   tstat
4    LS  17 -2.0012861 -7.151335
5    LS  26 -1.0128122 -3.619153
6    LS  62 -1.1263688 -4.024932
7    LS 277  1.2731347  3.601591
8    TC  13  2.5592000  7.359541
10   TC  16  3.1903973  9.174687
12   TC  36  1.7587553  5.057687
1    LS   1  0.9631222  3.813579
14   LS  15  2.5701584 10.176800
28   LS  33  1.0062922  3.984515
29   LS  35  0.9493094  3.758886
30   LS  45  0.9431208  3.734381
31   LS  47  0.9083884  3.596855
32   LS  49  1.1118602  4.402522
33   LS  57  0.9898398  3.919370
34   LS  60  0.9467239  3.748648
3    LS  10 -1.2472808 -3.681041
101  LS  22 -1.3642454 -4.026233
65   LS  79 -1.3272533 -3.917058
67   LS  86 -1.3424138 -3.961800
35   TC  28  1.3874358  3.742992

```

```
24    TC  24  1.2561947  3.586294
2     TC  14  1.3464617  4.028040
```

```
iter
[1] 2
```

En el código se puso un máximo de 6 iteraciones para encontrar el número máximo de outliers y sus tipos, el número permitido de iteraciones depende del investigador (entre 4 y 6). Ahora se irá al segundo paso, este paso trata de quedarse con los atípicos más significativos ajustando un nuevo modelo, en este caso el modelo ARIMA puede cambiar, a diferencia de la teoría de Chang y Tiao donde el modelo ARIMA era siempre el mismo. En este caso, hay más flexibilidad.

```
outliers<-stage1$outliers
xreg<-outliers.effects(outliers, n , weights = F, pars = pars)
iter<-0
while(TRUE)
{
  fit<-forecast::auto.arima(x=yt,xreg=xreg,allowdrift = F,allowmean = F,
    seasonal = F, ic="bic")
  xregcoefs<-coef(fit)
  ind<-match(colnames(xreg),names(xregcoefs))
  xregcoefs<-xregcoefs[ind]
  tstats<-xregcoefs/sqrt(diag(fit$var.coef)[ind])
  ref<-which(abs(tstats)<cval)
  if(length(ref)>0)
  {
    outliers<-outliers[-ref,]
    xreg<-data.matrix(xreg[,-ref])
  } else
  break
  if(nrow(outliers)==0)
  break
  iter<-iter+1
}
print(fit)
Series:
Regression with ARIMA(1,0,0) errors
```


Coefficients:

	ar1	LS17	LS62	LS277	TC13	TC16	TC36	LS15
	0.6647	-1.9008	-0.9940	1.2672	3.1606	2.8348	1.7001	3.2738
s.e.	0.0458	0.2808	0.1472	0.3284	0.3284	0.3412	0.3334	0.2791
	TC14							
	1.6074							
s.e.	0.3277							

sigma² estimated as 0.111: log likelihood=-84.29

AIC=188.59 AICc=189.41 BIC=224.83

print(outliers)

	type	ind	coefhat	tstat
4	LS	17	-2.001286	-7.151335
6	LS	62	-1.126369	-4.024932
7	LS	277	1.273135	3.601591
8	TC	13	2.559200	7.359541
10	TC	16	3.190397	9.174687
12	TC	36	1.758755	5.057687
14	LS	15	2.570158	10.176800
2	TC	14	1.346462	4.028040

En un principio se ajustó un proceso ARIMA(1,1,1) a los datos, después de agregar los atípicos el proceso ha cambiado a un AR(1). La sigma estimada disminuyó considerablemente del primer al segundo modelo, al igual el AIC y BIC disminuyeron casi a la mitad.

Ahora, se va a iterar el último modelo ajustado sin los efectos de los datos atípicos con la serie de tiempo original. Esto se hace con el fin de ver si hay más atípicos que se pudieran agregar al modelo. En caso de que no haya más atípicos, el modelo ajustado con los outliers es adecuado para hacer el pronóstico.

```
tmp <- coefs2poly(fit)
xreg <- outliers.effects(outliers, n, weights = FALSE, pars = tmp)
xregcoefs <- coef(fit)
ind <- match(colnames(xreg), names(xregcoefs))
xregcoefs <- xregcoefs[ind]
oeff <- xreg %*% cbind(xregcoefs)
yadj <- yt - oeff
plot(yt, col = "forestgreen", ylim=c(-1,8))
lines(yadj,col="red")
```

```

fit1<-forecast::auto.arima(x=yadj,allowdrift = F,allowmean = F,
    seasonal = F, ic="bic")
tmp<-coefs2poly(fit1)
resid<-residuals(fit1)
sigma<-1.483*quantile(abs(resid-quantile(resid,probs = 0.5)),probs = 0.5)
locate.outliers(resid,pars,cval,types = c("LS","TC"))
[1] type      ind      coefhat tstat
<0 rows> (or 0-length row.names)

```

No hay más outliers, entonces el modelo ajustado es de la siguiente forma

$$\begin{aligned}
 Y_t = & \frac{3.1606}{(1 - 0.7B)} I_t(13) + \frac{3.2738}{(1 - B)} I_t(15) + \frac{2.8348}{(1 - 0.7B)} I_t(16) - \frac{1.9008}{(1 - B)} I_t(17) \\
 & + \frac{1.7001}{(1 - 0.7B)} I_t(36) - \frac{0.9940}{(1 - B)} I_t(62) + \frac{1.2672}{(1 - B)} I_t(277) + \frac{W_t}{(1 - 0.6647B)}
 \end{aligned}$$

Si se analizan estos datos atípicos 13, 15, 16 y 17, se refieren a los primeros meses del año 1995 y como se sabe en esos meses México enfrentó una de las más virulentas crisis económicas que se tenga en la historia, hubo una devaluación de la moneda y esto trajo un aumento en la inflación. Luego, se tiene el datos atípico 36, éste corresponde a Diciembre de 1996, aumentó la inflación debido a un incremento al salario mínimo y precio del transporte público en el país. El atípico 62 está ligado al mes de Febrero de 1999, disminuyó la inflación debido a una correcta política monetaria, también se debe a a la apreciación del tipo de cambio y por el comportamiento favorable en el precio de frutas, verduras y electricidad. Finalmente, el atípico 277 se refiere al último mes registrado (Enero de 2017), después de muchos años de tener la inflación estacionaria, en este mes hubo un incremento en la inflación debido al alza en el precio de las gasolinas, esta subida significó el incremento más grande en 18 años.

Ahora, se realizará el pronóstico

En la Figura 4.10 se puede observar que más allá de los outliers que se le quitaron al modelo ajustado en el segundo paso, no hay otros más que se puedan ver. Por supuesto que hay residuales grandes, pero no al grado de convertirse en outliers. Por último, el pronóstico.

```

npred <- 12
newxreg <- outliers.effects(outliers, length(yt) + npred)

```

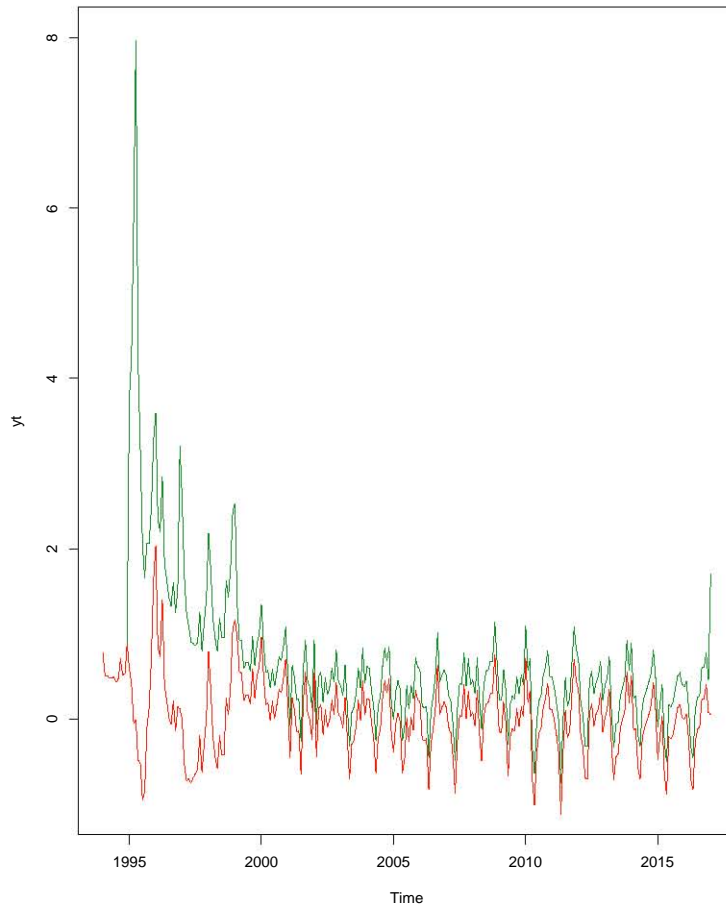


Figura 4.10: Serie original (verde) y el segundo modelo ajustado sin los efectos de outliers (rojo)

```

newxreg <- ts(newxreg[-seq_along(yt),], start = c(2017, 2))
p <- predict(fit, n.ahead=npred, newxreg=newxreg)

plot(cbind(yt, p$pred), plot.type = "single", ylab = "", type = "n",
      ylim=c(-1,8))
lines(yt)
lines(p$pred, type = "l", col = "blue")
lines(p$pred + 1.96 * p$se, type = "l", col = "red", lty = 2)
lines(p$pred - 1.96 * p$se, type = "l", col = "red", lty = 2)
legend("topright",
       legend = c("Serie observada", "Pronstico",
                  "Intervalos de confianza 95%"),
       lty = c(1,1,2,2), col = c("black", "blue", "red", "red"),
       bty = "n")
ppred

```

	Jan	Feb	Mar	Apr	May	Jun	Jul
2017		1.681949	1.669949	1.661972	1.656670	1.653145	1.650802
2018	1.646557						
	Aug	Sep	Oct	Nov	Dec		
2017	1.649244	1.648209	1.647521	1.647063	1.646759		
2018							

Los datos reales de esos meses son y el gráfico del pronóstico está dado por la Figura 4.11

0.58 0.61 0.12 -0.12 0.25 0.38 0.49 0.31 0.63 1.03

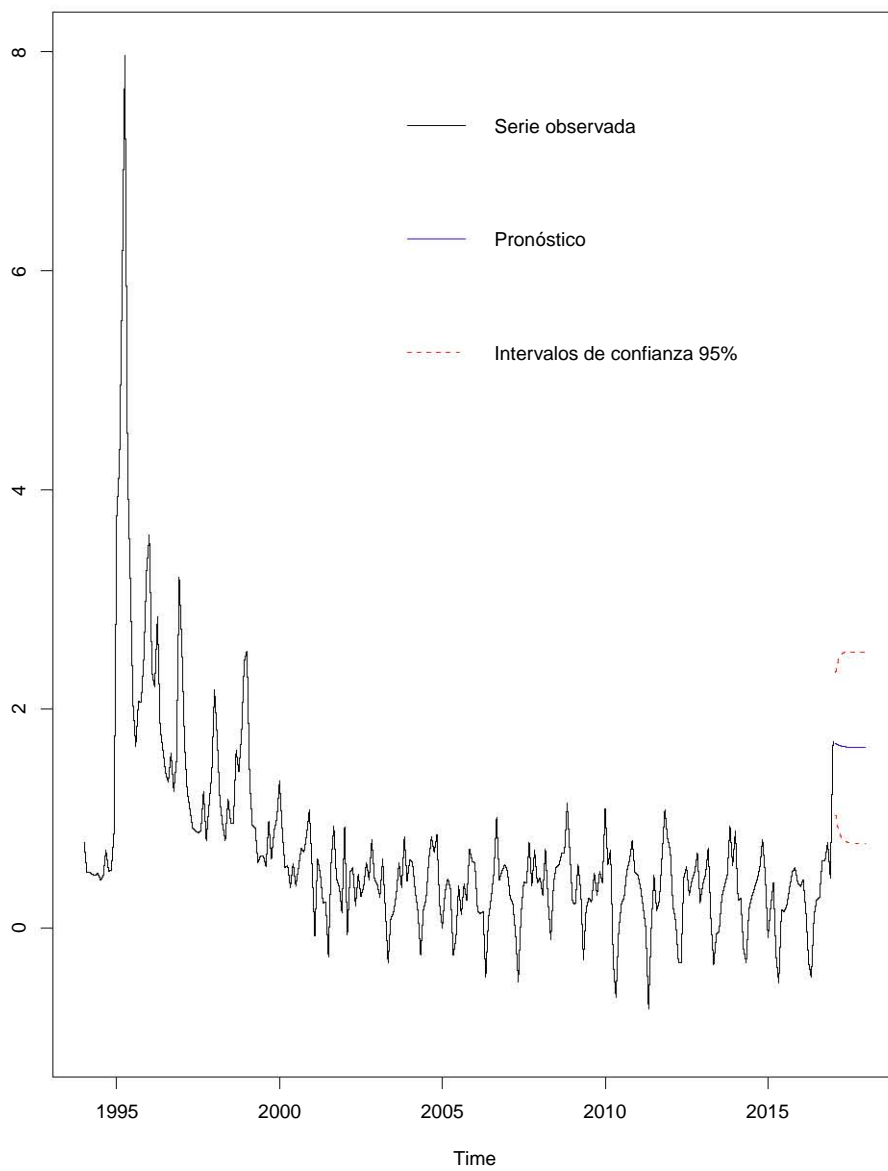


Figura 4.11: Pronóstico a un año de la Inflación Mexicana mensual (1994-2017)

4.5. Punto de Cambio

En esta parte de la tesis, se usará otra técnica para detectar cambios de nivel en los datos que se están trabajando, este método se conoce como punto de cambio. Uno de los desafíos clave en el análisis de punto de cambio es la habilidad para detectar múltiples cambios en una serie de tiempo o secuencia dada. A través de los años varios algoritmos de detección se han propuesto para superar este problema; entre ellos se tiene el algoritmo de segmentación binaria, el algoritmo de vecindad del segmento y más reciente el algoritmo PELT.

En términos generales lo que se hará es encontrar el punto de cambio y ese punto será el evento que se asumirá como conocido, posteriormente se ajustará un modelo de Intervención similar al del capítulo 3. El artículo de apoyo que se tomará, para llevar a cabo esta tarea, será el de Nathan Balke (1993), Rebecca Killick (2014) y el libro de Víctor Guerrero (2009).

4.5.1. Detección de Puntos de Cambio

En su forma más simple, la detección de puntos de cambio es el nombre dado al problema de estimar el punto en el cual las propiedades estadísticas de una secuencia de observaciones cambian. Detectar tales cambios es importante en una gran variedad de áreas. Ejemplos recientes incluyen a la climatología, informática, biología, finanzas, medicina, etc.

Formalmente, se define una secuencia de datos como $Y_{1:n} = (Y_1, \dots, Y_n)$. El punto de cambio se dice que ocurre dentro de este conjunto cuando existe un tiempo $\tau \in \{1, \dots, n-1\}$ tal que las propiedades estadísticas de $\{Y_1, \dots, Y_\tau\}$ y $\{Y_{\tau+1}, \dots, Y_n\}$ son diferentes de alguna manera. Extendiendo esta idea de un punto de cambio a múltiples cambios, se tendrá un número de puntos de cambio, m , y de acuerdo a sus posiciones en el tiempo, $\tau_{1:m} = (\tau_1, \dots, \tau_m)$. Se define $\tau_0 = 0$ y $\tau_{m+1} = n$, se dice que los puntos de cambio son ordenados tal que $\tau_i < \tau_j$ si y sólo si $i < j$. Por consiguiente los m puntos de cambio dividirán los datos en $m+1$ segmentos, con el i -ésimo segmento teniendo los datos $Y_{(\tau_{i-1}+1):\tau_i}$. Cada segmento será resumido por un conjunto de parámetros. Los parámetros asociados con el i -ésimo segmento será denotado por (ζ_i, ξ_i) donde ξ_i es un conjunto de *parámetros de acompañamiento* (no son tan importantes; sin embargo, son esenciales para explicar la teoría) y ζ_i es el conjunto de parámetros que contienen los cambios.

4.5.2. Detección de Punto de Cambio Único

Antes de considerar el problema general de indentificar $\tau_{1:m}$ puntos de cambio, se tomará en cuenta la identificación de un solo punto de cambio. La detección de un punto de cambio único se puede plantear como una prueba de hipótesis. La hipótesis nula, H_0 , corresponde a que no hay punto de cambio ($m = 0$) y la hipótesis alternativa, H_1 , a que hay un punto de cambio ($m = 1$).

Para llevar a cabo la tarea, es necesario introducir la razón general de verosimilitud para probar esta hipótesis. El método de la razón de verosimilitud requiere del cálculo de la máxima log-verosimilitud bajo ambas hipótesis (nula y alternativa). Para la hipótesis nula la máxima log-verosimilitud es $\log p(Y_{1:n}|\hat{\zeta})$, donde $p(\cdot)$ es la función de densidad de probabilidad asociada con la distribución de los datos y $\hat{\zeta}$ es la estimación por máxima verosimilitud de los parámetros.

Bajo la hipótesis nula, se considera un modelo con un punto de cambio en τ_1 , con $\tau_1 \in \{1, 2, \dots, n-1\}$. Entonces, la máxima log-verosimilitud para un τ_1 dado es

$$ML(\tau_1) = \log p(Y_{1:\tau_1}|\hat{\zeta}_1) + \log p(Y_{(\tau_1+1):n}|\hat{\zeta}_2) \quad (4.34)$$

por otro lado, dada la naturaleza discreta de la localización del punto de cambio, el valor máximo log-verosímil bajo la hipótesis alternativa es simplemente $\max_{\tau_1} ML(\tau_1)$, donde el máximo se toma sobre todos los posibles ubicaciones de puntos de cambio. La prueba estadística es

$$\lambda = 2 \left[\max_{\tau_1} ML(\tau_1) - \log p(Y_{1:n}|\hat{\zeta}) \right] \quad (4.35)$$

La prueba involucra elegir un umbral, c , tal que se rechaza la hipótesis nula si $\lambda > c$. Si se rechaza la hipótesis nula, es decir, se detecta un punto de cambio, entonces se estimará su posición como $\hat{\tau}_1$ (el valor de τ_1 que maximiza $ML(\tau_1)$). El valor apropiado de c será criterio del investigador.

Es claro que la prueba se puede extender a múltiples cambios simplemente al sumar la verosimilitud para cada uno de los m segmentos. El problema que se tiene es al identificar el máximo de $ML(\tau_1 : m)$ para todas las posibles combinaciones de $\tau_{1:m}$.

4.5.3. Detección de Punto de Cambio Múltiple

Con el incremento de teoría de series de tiempo, hay una creciente necesidad de localizar eficientemente múltiples puntos de cambio. El enfoque más común para identificar varios puntos de cambio en la literatura es minimizar

$$\sum_{i=1}^{m+1} [\mathbf{C}(Y_{(\tau_{i-1}+1):\tau_i})] + \kappa f(m) \quad (4.36)$$

donde \mathbf{C} es una función de costo para un segmento y $\kappa f(m)$ es una penalización para protegerse contra el sobreajuste (esto podría ser una versión de punto de cambio múltiple del umbral c). El paquete **changepoint** de R implementa tres algoritmos de punto de cambio múltiple que minimizan 4.36; segmentación binaria, vecindad del segmento y el propuesto recientemente podado lineal exacto del tiempo (PELT).

Segmentación binaria es posiblemente el método más usado para encontrar múltiples puntos de cambio. En términos sencillos, segmentación binaria primero aplica una prueba estadística de punto de cambio simple a todos los datos, si un punto de cambio es encontrado, el conjunto de datos es dividido en dos a la altura de punto de cambio. El procedimiento de punto de cambio simple es repetido con los dos nuevos conjuntos de datos, antes y después del cambio. Si más puntos de cambio son encontrados en alguno de los dos conjuntos, entonces los datos se seguirán dividiendo. Este proceso continúa hasta que no haya más puntos de cambio. Este procedimiento es una minimización aproximada de 4.36 con $f(m) = m$. Segmentación binaria es un algoritmo aproximado pero es computacionalmente rápido, ya que sólo considera un subconjunto de las 2^{n-1} posibles soluciones.

El algoritmo de vecindad del segmento minimiza la expresión dada por la ecuación 4.36 usando una técnica de programación dinámica para obtener la segmentación óptima para $m + 1$ puntos de cambio reusando la información que fue calculada para m puntos de cambio. Mientras que este algoritmo es exacto, la complejidad computacional es considerablemente más alta que segmentación binaria.

Por último, el algoritmo PELT es muy similar al algoritmo de vecindad. Sin embargo, PELT se puede mostrar más eficiente computacionalmente hablando, debido a su construcción dinámica y de podado, lo cual lo hace un algoritmo no oneroso. La principal característica es el incremento lineal de puntos de cambio conforme el número de datos crece.

Los tres algoritmos están disponibles en R.

4.5.4. Modelo de Intervención

En este caso, la metodología del análisis de intervención consiste primero, en determinar el momento en el que hubo un cambio de nivel, para ello se utilizará la teoría de punto de cambio anteriormente explicada. Luego, se contruirá un modelo estocástico para representar a la serie que comprende desde la observación inicial hasta la observación inmediata anterior a la intervención. Posteriormente se procede a postular un modelo dinámico para representar al efecto de la intervención y acto seguido se estima el modelo completo para todo el periodo muestral de observación, es decir, se estiman tanto los parámetros que aparecen en la función de intervención como los del modelo determinado previamente para la parte estocástica. Por último, se verifica que se satisfagan los supuestos del modelo. Finalmente, se hace el pronóstico y los intervalos de confianza.

Esta metodología es relativamente distinta a la mostrada en el capítulo 3, ya que en ese caso se contruye la parte estocástica y dinámica para toda la serie, en este caso la parte estocástica sólo será para la serie antes de la intervención y la parte dinámica para la serie con la intervención y después de ésta.

En realidad, ambas metodologías logran los mismos fines, pero en lugar de ser competidoras entre sí resultan ser complementarias dentro del contexto de construcción de un modelo para una serie temporal determinada. Por este motivo no es necesario comparar la efectividad de una respecto a otra.

4.5.5. Implementación en R

Para esta última parte de este capítulo, se usarán los datos de la Inflación General de México, la información fue obtenida del Sistema de Información Económica del Banco de México y fueron registrados desde el mes de Enero de 1977 al mes de Abril de 2017.

Los datos se muestran en el Cuadro 4.3

Se cargan las paqueterías y se grafican los datos

```
library(lmtest)
library(forecast)
```

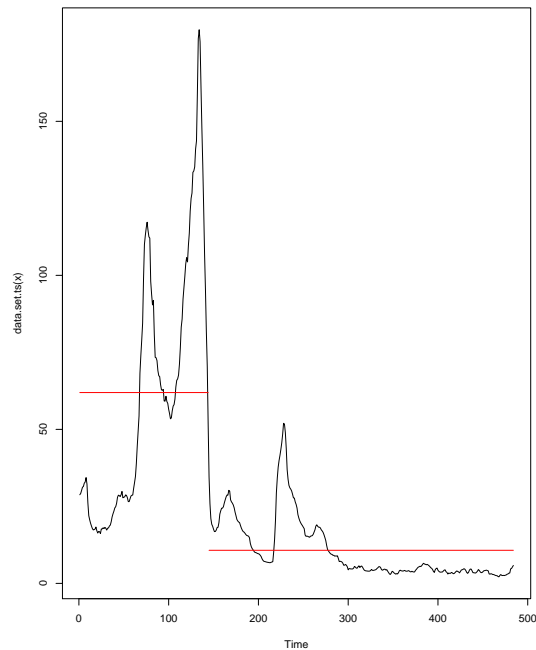


Figura 4.12: Punto de cambio encontrado en la serie de Inflación General

```
library(changepoint)
library(TSA)
library(stats)
setwd("C:/Users/Jesus Pichardo Leyva/Documents/TESIS")
datos<-read.table("Inflacion General Mexico.csv",sep=" ",header=FALSE)
datos<-datos[,2]
datos
```

Se usará el método de segmentación binaria para detectar el punto de cambio de la serie temporal.

```
cambio<-cpt.mean(datos, penalty = "BIC",pen.value = 0, method="BinSeg",Q=1)
cpts(cambio)
[1] 144
plot(cambio)
```

Donde Q es el número de puntos de cambio.

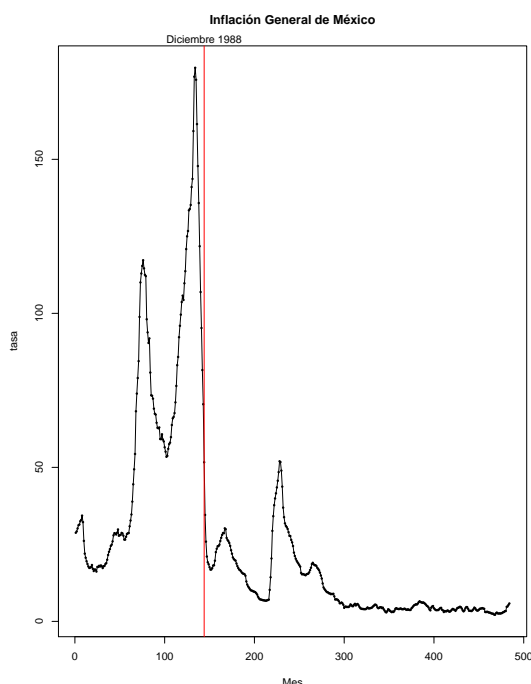


Figura 4.13: Inflación General con T= diciembre de 1988

Se encontró en la Figura 4.12 que a partir del índice 144 la serie cambió, se puede decir que hubo un cambio de nivel. Si se busca la fecha de ese índice, se encontrará que el dato 144 corresponde a diciembre de 1988, la cual marca el inicio de un nuevo gobierno presidencial (Figura 4.13).

```
plot(datos,type="o",pch=16,cex=0.5,xlab='Mes',ylab='tasa',
      main="Inflacion General de Mexico")
abline(v=144,col="2")
mtext("Diciembre 1988",side=3,at=144)
```

Ahora, se ajusta la parte estocástica del modelo de intervención, es decir, sólo nos tomaremos los datos desde el primer dato hasta el dato 143 (uno antes de la *intervención*). Para ello, se utilizarán el PACF y ACF muestral hasta el dato 143.

```
datos<-ts(datos, frequency = 12, start = c(1977, 1))
acf(ts(datos[1:143],freq=1), lag.max=20,
     main="ACF Inflacin General Mxico hasta el dato 143")
```

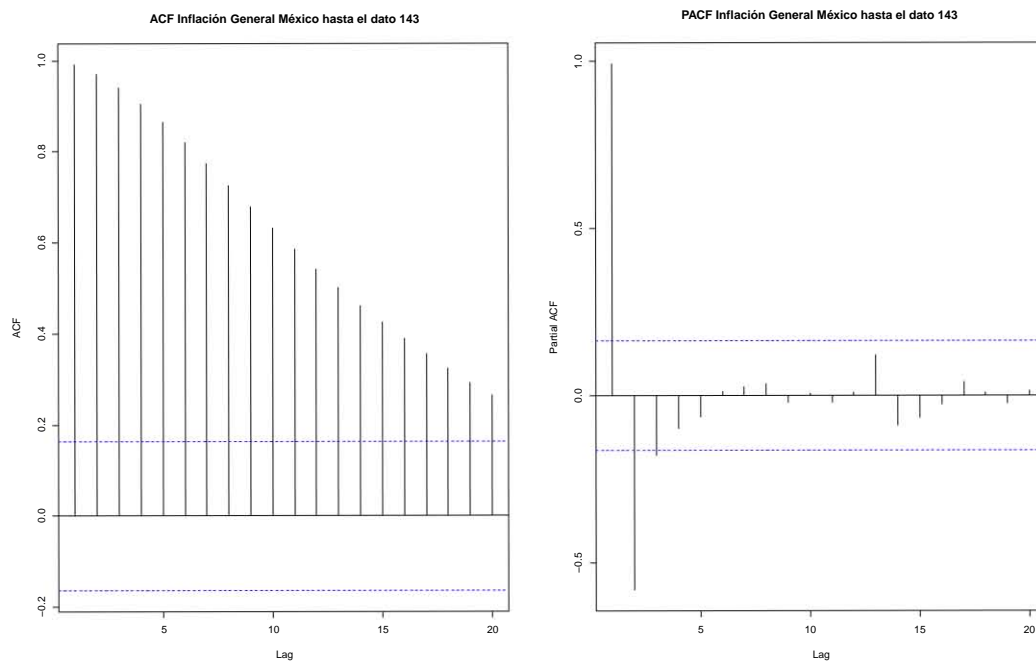


Figura 4.14: ACF y PACF muestral de la Inflación General Mexicana hasta el dato 143

```
pacf(ts(datos[1:143],freq=1), lag.max=20,
      main="PACF Inflacin General Mxico hasta el dato 143 ")
```

De acuerdo a la Figura 4.14 se tomó la decisión de ajustar un AR(1).

Como se observa, el modelo de intervención parece ser con un cambio de nivel, entonces se usará una función escalón donde

$$S_t(144) = \begin{cases} 1 & \text{si } t \geq 144 \\ 0 & \text{si } o.c. \end{cases}$$

Ajustando el modelo se tiene

```
L<-length(datos)
St<-c(rep(0,143),rep(1,L-143))
Ajuste<-arimax(datos,order=c(1,0,0),xtransf=St,transfer =list(c(1,0)),
              include.mean=F)
Ajuste
```

Call:

```
arimax(x = datos, order = c(1, 0, 0), include.mean = F, xtransf = St,
       transfer = list(c(1, 0)))
```

Coefficients:

	ar1	T1-AR1	T1-MA0
	0.9993	0.6288	-20.9188
s.e.	0.0013	0.0567	2.5903

```
sigma^2 estimated as 8.557: log likelihood = -1209.58, aic = 2425.17
coefstest(Ajuste)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)		
ar1	0.9993134	0.0013004	768.4921	< 2.2e-16 ***		
T1-AR1	0.6287891	0.0567178	11.0863	< 2.2e-16 ***		
T1-MA0	-20.9187983	2.5902986	-8.0758	6.702e-16 ***		

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1	1

El modelo ajustado es de la forma

$$Y_t = \frac{-20.9188}{1 - 0.6288B} S_t(144) + \frac{W_t}{1 - 0.9993B}$$

Lo que quiere decir, la parte dinámica del modelo, es ω_0 es negativa entonces hay un cambio de nivel hacia abajo y el valor v_1 que el cambio es gradual hacia abajo. Si se observa la Inflación General mexicana se notará que tiene un descenso lento, no es una rampa o algo parecido, hasta llegar a ser constante.

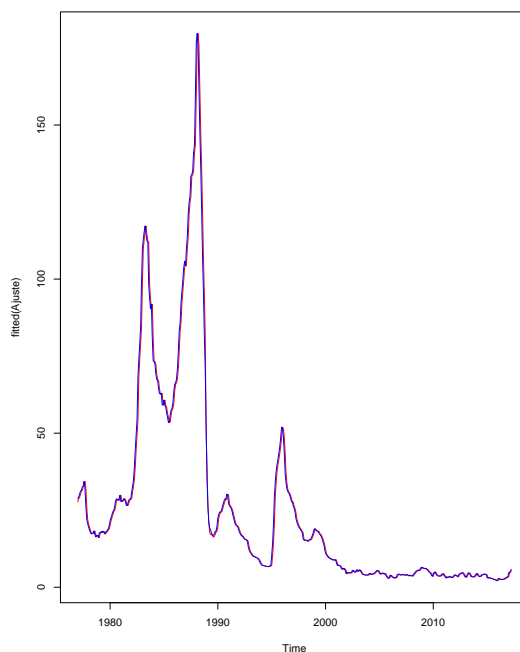


Figura 4.15: Valores de la Inflación General y valores ajustados

Ahora, se graficarán los datos de la serie de tiempo (Y_t) pegados con los valores ajustados.

```
plot(fitted(Ajuste),col="red", type="l")
lines(datos,type="l",col="blue")
```

Si se observa con detenimiento la Figura 4.15 se notará que los valores reales y los valores ajustados prácticamente son iguales, hasta pareciera una simulación, entonces el modelo ajustado es muy bueno.

Por un lado se construirá la parte dinámica y por el otro la parte estocástica, esto con el fin de poder realizar eficientemente el pronóstico. Primero la parte dinámica

```
St<-c(rep(0,143),rep(1,L-143+48))
intv.effect <- St
intv.effect <- ts(filter(intv.effect, filter = 0.6288,
                        method = "rec", sides = 1) *(-20.9188) )
```

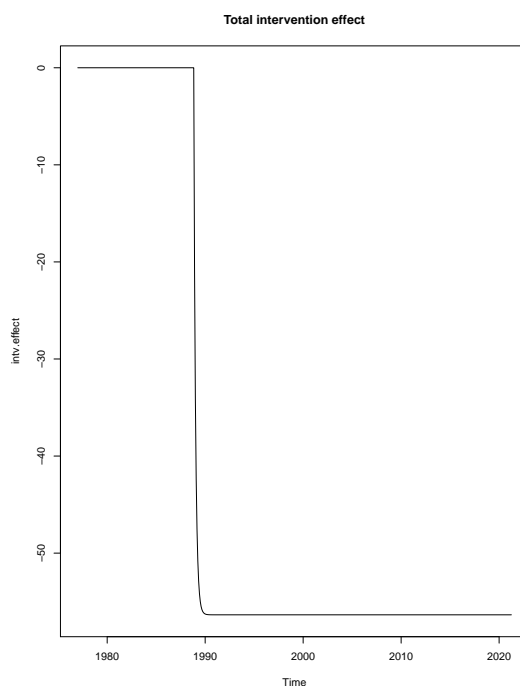


Figura 4.16: Efecto de la Intervención

```
intv.effect <- ts(intv.effect, start = start(datos),
                 frequency = frequency(datos))
```

Y luego la parte estocástica

```
estoc<-arima(window(datos, end = c(1988,12)), order = c(1,0,0),
             xreg = window(intv.effect, end = c(1988,12)))
```

El efecto de intervención está dado por la Figura 4.16

```
plot(intv.effect, type = "l", main = "Efecto de la Intervencion")
```

Para terminar, el pronóstico (Figura 4.17). Lo que se puede observar es que la parte dinámica del modelo captura de excelente forma los datos de inflación.

```
p <- predict(estoc, n.ahead = 388, newxreg=window(start=c(1989,01),
```

```
intv.effect))

plot(cbind(window(datos, end = c(1988,12)),p$pred), plot.type = "single",
      ylab = "", type = "n", ylim=c(-60, 190))
lines(window(datos, end = c(1988,12)), type = "l")
lines(window(datos, start = c(1988,12)), col = "forestgreen", type = "l")
lines(p$pred, type = "l", lty=2, col = "blue4")
legend("topright", legend = c("Observaciones antes de intervencion",
                              "Observaciones despues de intervencion",
                              "Pronosticos","Intervalos de Confianza"),
      lty = rep(1, 4),
      col = c("black", "forestgreen", "blue4","firebrick"),
      bty = "n",cex=0.8)
lines(p$pred + 1.96 * p$se, lty = 2, col = "firebrick")
lines(p$pred - 1.96 * p$se, lty = 2, col = "firebrick")
```

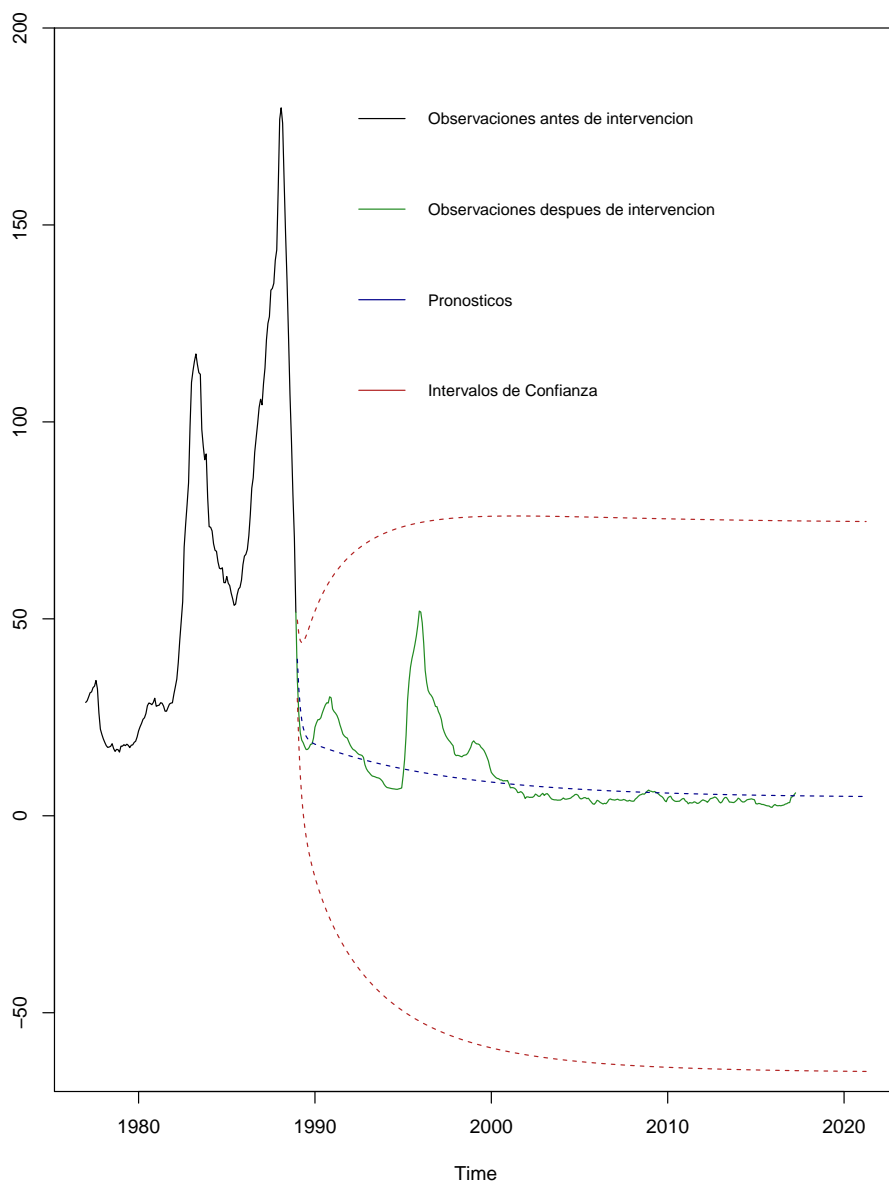



Figura 4.17: Pronóstico Inflación General

AÑO	E	F	M	A	M	JN	JL	A	S	O	N	D
1977	28.77	29.19	30.17	31.22	31.45	32.53	32.91	34.35	32.22	26.13	22	20.66
1978	19.53	18.63	17.81	17.35	17.46	17.64	18.3	17.07	16.35	16.86	16.79	16.17
1979	17.68	17.68	18.05	17.8	18.18	17.87	17.31	17.91	18.01	18.63	18.93	20.02
1980	21.56	22.6	23.45	24.49	24.89	25.97	27.94	28.64	28.49	28.2	28.77	29.85
1981	27.8	27.98	28.08	28.72	28.57	27.83	26.55	26.54	27.47	28.36	28.6	28.68
1982	30.86	32.74	34.71	38.88	44.5	49.37	54.35	68.21	73.95	79	84.49	98.84
1983	110.04	112.95	115.39	117.25	114.61	112.5	112.08	98.08	93.83	90.39	91.87	80.78
1984	73.4	73.25	72.31	69.06	67.41	67.14	64.49	62.84	62.68	62.96	59.2	59.16
1985	60.75	59.04	58.43	56.53	55.1	53.43	53.73	56.02	57.55	58.02	59.82	63.75
1986	65.92	66.38	67.62	71.11	76.44	83.17	85.84	92.25	95.96	99.58	103.67	105.75
1987	104.34	109.76	113.69	120.86	125	126.73	133.44	133.88	135.18	141	143.65	159.17
1988	176.83	179.73	175.83	161.44	147.82	135.81	121.79	106.92	95.24	81.59	70.5	51.66
1989	34.56	25.89	21.05	19.2	18.54	17.58	16.81	16.85	17.3	18.13	18.21	19.7
1990	22.48	23.57	24.4	24.43	24.89	26.11	27.13	28.08	28.68	28.62	30.21	29.93
1991	27.11	26.46	26.04	25.45	24.51	23.1	21.97	20.76	20.25	19.92	19.72	18.79
1992	17.95	17.3	16.82	16.64	16.28	15.85	15.56	15.46	15.32	14.82	12.96	11.94
1993	11.32	10.91	10.44	10.09	10	9.87	9.7	9.62	9.48	9.14	8.72	8.01
1994	7.5	7.18	7.1	7.01	6.92	6.85	6.81	6.74	6.71	6.83	6.93	7.05
1995	10.23	14.31	20.43	29.39	34.15	37.72	39.91	41.57	43.48	45.66	48.46	51.97
1996	51.72	48.95	43.75	36.93	33.83	31.82	31.03	30.6	30	28.97	27.77	27.7
1997	26.44	25.64	24.46	22.33	21.23	20.35	19.7	19.18	18.76	18.24	17.77	15.72
1998	15.27	15.35	15.27	15.1	14.97	15.31	15.41	15.5	15.92	16.65	17.41	18.61
1999	19.02	18.54	18.26	18.23	18.01	17.39	17.04	16.58	15.83	14.91	13.92	12.32
2000	11.02	10.52	10.11	9.73	9.48	9.41	9.12	9.1	8.85	8.91	8.87	8.96
2001	8.11	7.09	7.17	7.11	6.95	6.57	5.88	5.93	6.14	5.89	5.39	4.4
2002	4.79	4.79	4.66	4.7	4.68	4.94	5.51	5.29	4.95	4.94	5.39	5.7
2003	5.16	5.52	5.64	5.25	4.7	4.27	4.13	4.04	4.04	3.96	3.98	3.98
2004	4.2	4.53	4.23	4.21	4.29	4.37	4.49	4.82	5.06	5.4	5.43	5.19
2005	4.54	4.27	4.39	4.6	4.6	4.33	4.47	3.95	3.51	3.05	2.91	3.33
2006	3.94	3.75	3.41	3.2	3	3.18	3.06	3.47	4.09	4.29	4.09	4.05
2007	3.98	4.11	4.21	3.99	3.95	3.98	4.14	4.03	3.79	3.74	3.93	3.76
2008	3.7	3.72	4.25	4.55	4.95	5.26	5.39	5.57	5.47	5.78	6.23	6.53
2009	6.28	6.2	6.04	6.17	5.98	5.74	5.44	5.08	4.89	4.5	3.86	3.57
2010	4.46	4.83	4.97	4.27	3.92	3.69	3.64	3.68	3.7	4.02	4.32	4.4
2011	3.78	3.57	3.04	3.36	3.25	3.28	3.55	3.42	3.14	3.2	3.48	3.82
2012	4.05	3.87	3.73	3.41	3.85	4.34	4.42	4.57	4.77	4.6	4.18	3.57
2013	3.25	3.55	4.25	4.65	4.63	4.09	3.47	3.46	3.39	3.36	3.62	3.97
2014	4.48	4.23	3.76	3.5	3.51	3.75	4.07	4.15	4.22	4.3	4.17	4.08
2015	3.07	3	3.14	3.06	2.88	2.87	2.74	2.59	2.52	2.48	2.21	2.13
2016	2.61	2.87	2.6	2.54	2.6	2.54	2.65	2.73	2.97	3.06	3.31	3.36
2017	4.72	4.86	5.35	5.82								

Cuadro 4.3: Inflación General de México (1977-2017)

Conclusiones

A lo largo de esta tesis se analizaron diversos modelos importantes en el quehacer estadístico, los cuales sirven para tratar de entender mejor la realidad y percibir con más claridad el futuro, ya que, en muchos casos, por desconocer el futuro, se puede caer en graves riesgos de toda índole. Se sabe que cada modelo está restringido a diversos supuestos, estos supuestos son esenciales para tener pronósticos más apegados a la realidad.

En todo modelo estadístico, es necesario conocer el comportamiento de los datos, y la mejor manera de hacerlo es graficando, de esta forma se podrá elegir un modelo acorde a los registros que se tengan. Una vez establecido el modelo, debe buscarse no hacer una sobreparametrización, es decir, se debe buscar la forma más sencilla posible para hacer un ajuste adecuado.

Se ha analizado en el capítulo uno teoría básica sobre procesos estacionarios, algunos operadores importantes en series de tiempo y finalmente, los modelos autorregresivos, promedios móviles, ARMA y ARIMA. En la parte aplicada, se usaron los datos del precio del barril de mezcla mexicana y luego se ajustó un modelo ARIMA(1, 1, 0), este modelo fue seleccionado gracias a los criterios AIC y BIC, posteriormente se revisaron los supuestos de normalidad y de no correlación, resultando un modelo adecuado para pronosticar. Sin embargo, cuando se analizó la variabilidad de los residuos resultó que la varianza era no constante ya que los residuos se veían dispersos y también se encontró una fuerte presencia de outliers, debido a esta situación los intervalos de confianza son muy grandes. Por otro lado, los pronósticos son adecuados a la realidad del precio de barril de petróleo. El precio de barril de petróleo (MME), a lo largo de este año anduvo moviéndose entre 35 dólares y 50 dólares.

En el capítulo dos, se vio teoría y se ajustó un modelo de función de transferencia. El modelo tiene la ventaja de explicar una serie de tiempo usando otra. En el modelo se usó la tasa de crecimiento del PIB de México como variable output y la tasa de crecimiento del PIB de Brasil como input. Los supuestos del modelo

se cumplieron y el pronóstico fue adecuado ya que, como se mencionó, la tasa de crecimiento de la economía mexicana ha sido, en los últimos años, del dos por ciento en promedio.

En los capítulos tres y cuatro, se abordó el tema de Análisis de Intervención con punto conocido y con punto o puntos desconocidos. Estos puntos del tiempo son eventos que hacen que una serie de tiempo tenga un movimiento distinto. En algunos casos, como se dijo, son conocidos, en otros casos, debido a la falta de información, son desconocidos y se tienen que encontrar por medio de métodos estadísticos. En el primer caso se usó la inflación anual mexicana, el modelo fue eficiente ya que nos da pronósticos apegados a la realidad, los pronósticos fueron de dos por ciento por año, con intervalos que abren desde uno a tres por ciento. En el segundo caso se usó la inflación mensual mexicana de 1970 – 1985, el modelo fue también altamente predictivo, ya que se ingresaron diversos acontecimientos al modelo para tener predicciones e intervalos de confianza más precisos, se compararon los pronósticos que se obtuvieron del modelo con los datos reales y se observa que las diferencias son mínimas. En el tercer caso, se usó inflación mensual del país de 1994 – 2017, este modelo tuvo mayor flexibilidad ya que el modelo ARIMA puede cambiar ya con las intervenciones incluidas, a diferencia del modelo anterior donde el modelo ARIMA siempre es el mismo, en esta serie se obtuvieron pronósticos de uno por ciento mensual con intervalos de cero por ciento a dos por ciento, comparando con datos reales parece que la diferencia es mínima. Estos modelos son muy importantes en economía, no sólo porque estiman los efectos que distorsionan la forma de la serie temporal, sino también porque dan pronósticos más apegados a la realidad e intervalos pequeños.

Para finalizar este trabajo, se puede mencionar que se ha dado un pronóstico confiable a cada una de las series de tiempo usadas, por lo que expertos en esta materia (estadística, economía y econometría) pueden basarse en este desarrollo estadístico e incluir, si así lo creen necesario, en algún proyecto de investigación económica.

Apéndice A

Inflación Mexicana

La inflación es el aumento generalizado y sostenido del nivel de precios de bienes y servicios existentes en el mercado en un país durante un período de tiempo, frecuentemente un año. Cuando el nivel general de precios sube, con cada unidad de moneda se adquieren menos bienes y servicios. Es decir, que la inflación refleja la disminución del poder adquisitivo de la moneda.

Los efectos negativos de la inflación incluyen la disminución del valor real de la moneda a través del tiempo, el desaliento del ahorro y un descenso en la inversión debido a la incertidumbre sobre el valor futuro del dinero, escasez de bienes y finalmente golpea al desarrollo económico.

Las causas de la inflación pueden ser diversas, entre las que se encuentran:

- **Inflación por consumo o demanda.** Esta inflación obedece a la ley de la oferta y la demanda. Si la demanda de bienes excede la capacidad de producción o importación de bienes, los precios tienden a aumentar.
- **Inflación por costos.** Esta inflación ocurre cuando el precio de las materias primas (cobre, petróleo, energía, etc) aumenta, lo que hace que el productor, buscando mantener su margen de ganancia, incremente sus precios.
- **Inflación autoconstruida.** Esta inflación ocurre cuando se prevee un fuerte incremento futuro de precios, y entonces se comienzan a ajustar éstos desde antes para que el aumento sea gradual.
- **Inflación generada por un excesivo incremento en la oferta de dinero.** Esto es típico en países con alta inflación donde los trabajadores piden

aumentos de salarios para contrarrestar los efectos inflacionarios, lo cual da pie al aumento en los precios por parte de los empresarios, originando un círculo peligroso de inflación.

Por otro lado, es necesario mencionar que existe otro escenario llamado *deflación*, que es la disminución generalizada de los precios de bienes y servicios. Esto ocurre cuando hay una disminución en la demanda y un exceso en la productividad (oferta). Esta situación es peligrosa, ya que los precios de los bienes y servicios bajan debido a que la gente ya no compra como antes (disminuye la demanda), esto ocasiona que las empresas obtengan menos ganancias y finalmente tengan que recortar empleos (aumento del desempleo). Si la situación permanece así, ocasionaría que la economía del país se desplome y finalmente crearía una crisis.

En México, el Banco Central (Banco de México) e INEGI proponen un intervalo de tolerancia de inflación, este está entre 2 y 4 por ciento.

Por lo tanto, se puede entender porqué es importante cuantificar la inflación y tratar de protegerse de sus efectos adversos.

En nuestro país, la inflación se calcula usando el Índice Nacional de Precios al Consumidor (INPC), el cual se elabora dando seguimiento a los precios de una canasta de bienes y servicios representativa del consumo de los hogares urbanos mexicanos. Para construir la canasta de consumo se utiliza como principal fuente de información la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) que recaba el INEGI. La información que provee esta Encuesta, se incorporan como ponderadores de las variaciones de precios en el cálculo del INPC.

Finalmente, la serie mensual de inflación $\{INF_t\}$, se obtiene a partir del Índice Nacional de Precios al Consumidor $\{INPC_t\}$, a través de la relación

$$INF_t = 100 \left(\frac{INPC_t - INPC_{t-1}}{INPC_{t-1}} \right) \quad \text{para toda } t$$

es decir, la inflación es el porcentaje de crecimiento mensual en los precios al consumidor.

En la Figura B.1 se muestra que la inflación tiene grandes momentos de volatilidad y a partir del ao 2000 parece ser que la inflación se comporta como un proceso estacionario.

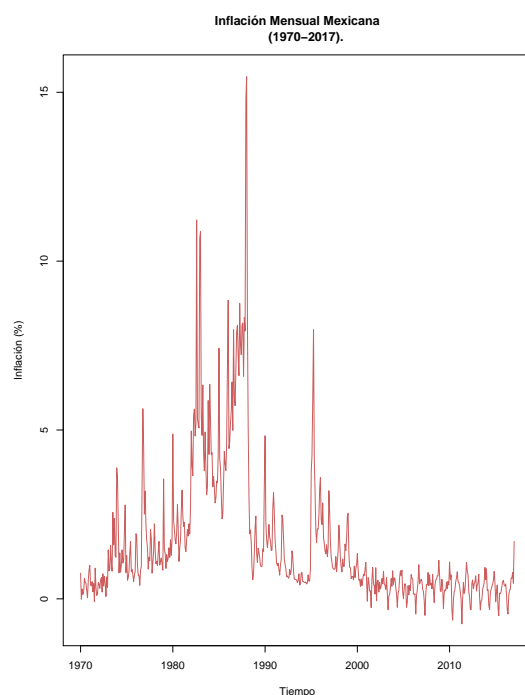


Figura A.1: Inflación Mensual Mexicana (1970 – 2017)

La serie de tiempo comienza en el año de 1970 y termina en 2017. Empezaremos haciendo un análisis histórico del porqué hubo esos incrementos. En los años 70's, un hecho provocó un auge inflacionario importante, el aumento y súbito aumento de los precios del petróleo entre 1975 y 1979. Como se explicó en el capítulo 3, en 1976 se devaluó el peso en un 25% frente al dólar. A partir de ese año las devaluaciones sistemáticas no se detendrían provocando altas tasas de inflación.

Los años 80's se empiezan con un aumento en la inflación, debido a que México no podía pagar su deuda externa en 1982. Esto se debió a una política monetaria restrictiva aplicada por el Departamento del Tesoro de Estados Unidos que fortaleció al dólar y elevó las tasas de interés (en dólares), por lo que el pago de los intereses de la deuda se volvieron impagables. El gobierno al caer en estado de insolvencia, devaluó el peso y estatizó el sistema bancario. Estas acciones provocan que la inflación creciera de forma alarmante. También en los años ochenta y principios de los noventa, tuvieron lugar los llamados Pactos Económicos que a partir de 1987 lograron reducir y estabilizar la inflación que pasó de un histórico 150% a un 7% en 1994, hasta que la devaluación de diciembre de 1994 y la posterior crisis económica provocaron un incremento muy fuerte en la inflación.

Toda la década pasada (2000–2010) y hasta lo que llevamos de esta década, el objetivo de estabilidad económica, reflejada en una tasa de inflación de un dígito, ha sido alcanzado.

Apéndice B

Tasa de crecimiento del PIB de México

La tasa de crecimiento del PIB (Producto Interno Bruto) es una medición frecuente empleada en el estudio económico para estudiar los incrementos y disminuciones que experimenta el PIB en determinados periodos de tiempo.

Los periodos comprendidos a la hora de realizar los cálculos de las tasas suelen ser de un año generalmente, o también por trimestres. Básicamente esta tasa es bastante útil a la hora de observar y medir el crecimiento económico que un país o territorio ha experimentado en dicho periodo y otorgado a este cambio un valor porcentual.

El cálculo de esta tasa se realiza dividiendo un valor PIB menos otro valor PIB anterior entre este mismo número y multiplicado por 100 para obtener el resultado expresado como porcentaje.

$$CPIB_t = 100 \left(\frac{PIB_t - PIB_{t-1}}{PIB_{t-1}} \right)$$

Mediante el empleo de este simple y rápido cálculo se puede conocer fácilmente la variación entre dos momentos determinados, comprobando si ha existido un incremento de la variable o una disminución de la misma en dicho periodo de tiempo.

En el caso de México, la tasa de crecimiento del PIB ha sido baja a lo largo de los últimos 40 años.

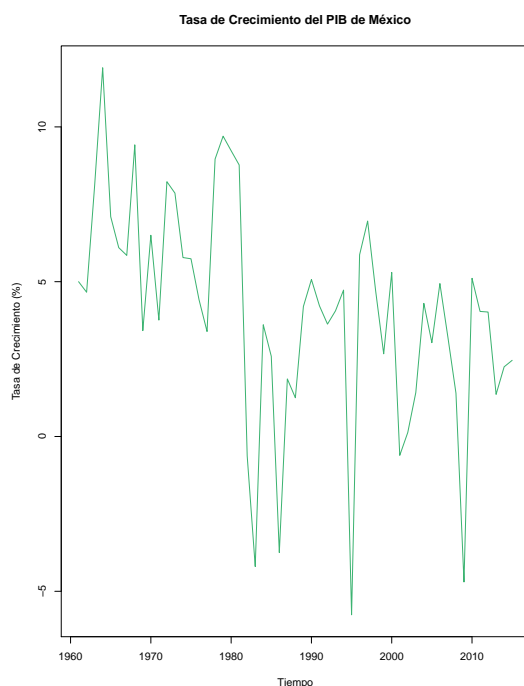


Figura B.1: Tasa de Crecimiento del PIB (1961 – 2015)

Repasamos algunos acontecimientos importantes.

Entre 1950 y 1970 la economía mexicana funcionaba relativamente bien. El crecimiento del PIB fluctuaba entre el 3 y 4 por ciento anual. Este fue un período al que se le llamó de *desarrollo estabilizador*.

Para mediados de 1970, como se mencionó, las variables macroeconómicas de México empezaron a cambiar de manera alarmante. La disminución del crecimiento de la economía fue en parte causada por el deterioro de los precios del petróleo y por la alta regulación y nacionalización de importantes sectores de la economía nacional.

Después, en 1982 hubo otra crisis económica en el país, los hechos más importantes de este suceso fueron: la drástica devaluación del peso, decrecimiento del PIB de -0.6 por ciento, una inflación anual de casi el 100 por ciento y finalmente, un caos en los mercados financieros. La actividad económica siguió en picada lo que ocasionó que el PIB disminuyera de -0.6 por ciento en 1982 a -4.1 por ciento en 1983.

A finales de los años ochenta y principios de los noventa, en pleno gobierno

de Carlos Salinas de Gortari, la economía creció a tasas del 3 por ciento anual en promedio y una inflación controlada. Muchos analistas económicos lo deducen al Pacto de Solidaridad Económica implementado en ese sexenio.

Para 1994 y 1995, una crisis económica en México, que se le conoció a nivel mundial como *Efecto Tequila*, ocasiona que la tasa de crecimiento del PIB disminuyera 9 por ciento.

De 1995 a 2000, el crecimiento económico se recuperó. Sin embargo, no fue al ritmo que se esperaba, a pesar de que la inflación siguió descendiendo.

Finalmente, del año 2000 a la actualidad, el PIB creció en promedio 2 por ciento por año (2.35 por ciento, para ser exactos), debido en gran parte a la corrupción, inseguridad, baja productividad laboral y bajo nivel educativo que tiene el país. En 2009, se registró la peor caída en lo que va del siglo, que fue de -4.7 por ciento, debido a la crisis económica que azotó a todo el mundo.

Bibliografía

- [1] ATKINS, M.S. (1979), *A Case Study on the Use of Intervention Analysis Applied to Traffic Accidents*, The Journal of the Operational Research Society, **30**, 651–659.
- [2] BALKE, N.S. (1993), *Detecting Level Shifts in Time Series*, Journal of Business & Economic Statistics, **11**, 81–92.
- [3] BOX, G.E.P. ET. AL. (2016), *Time Series Analysis-Forecasting and Control*, quinta edición, USA. John Wiley & Sons.
- [4] BOX, G.E.P. y TIAO, G.C. (1975), *Intervention Analysis with Applications to Economic and Enviromental Problems*, Journal of the American Statistical Association, **70**, 70–79.
- [5] BROCKWELL, P.J. y DAVIS, R.A. (2016), *Introduction to Time Series and Forecasting*, tercera edición, USA. Springer.
- [6] BROCKWELL, P.J. y DAVIS, R.A. (1991), *Time Series: Theory and Methods*, segunda edición, USA. Springer.
- [7] CHANG, I., TIAO, G.C. y CHEN, C. (1988), *Estimation of Time Series Parameters in the Presence of Outliers*, Technometrics, **30**, 193–204.
- [8] CHEN, C. y LIU, L. (1993), *Joint Estimation of Model Parameters and Outliers Effects in Time Series*, Journal of the American Statistical Association, **88**, 284–297.
- [9] CHUNG, R.C.P., IP, W.H. y CHAN, S.L. (2009), *An ARIMA-Intervention Analysis Model for the Financial Crisis in China's Manufacturing Industry*, International Journal of Engineering Business Management, **1**, 15–18.
- [10] CRYER, J.D. y CHAN, K. (2008), *Time Series Analysis With Applications in R*, segunda edición, USA. Springer.

- [11] EGAN, M. (2016, Enero 20), *5 razones del desplome del precio del petróleo*. CNN en Español. Recuperado de <http://cnnespanol.cnn.com/2016/01/20/5-razones-del-desplome-del-precio-del-petroleo/0>
- [12] GUERRERO, V.M. (2009), *Análisis Estadístico y Pronóstico de Series de Tiempo Económicas*, tercera edición, México. Jit Press.
- [13] GUERRERO, V.M. (1984), *Análisis, Evaluación y Pronóstico de la Inflación en México, mediante un modelo univariado de Series de Tiempo*, Banco de México S.A., México.
- [14] GUERRERO, V.M. (1982), *Medición de los Efectos Inflacionarios causados por algunas decisiones gubernamentales: teoría y aplicaciones de Análisis de Intervención*, Centro de Estudios Económicos, El Colegio de México, A.C., México.
- [15] GUERRERO, V.M. y VERA, G. (1981), *Una aplicación del Análisis de Intervención a Series de Tiempo de la Economía Mexicana*, Banco de México S.A., México.
- [16] HARVEY, A.C. y DURBIN J. (1986), *The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling*, Journal of the Royal Statistical Society, **149**, 187–227.
- [17] JARRET, J.E. y KYPER E. (2011), *ARIMA Modeling with Intervention to Forecast and Analyze Chinese Stock Prices*, International Journal of Engineering Business Management, **3(3)**, 53–58.
- [18] KILLICK, R. y ECKLEY, I.A. (2014), *changepoint: An R Package for Changepoint Analysis*, Journal of Statistical Software, **50**, issue 3.
- [19] LÓPEZ-DE-LACALLE, J. (2017), *tsoutliers R Package for Detection of Outliers in Time Series*.
- [20] MCDOWELL, A. (2002), *From the help desk: Transfer functions*, The Stata Journal, **2**, 71–85.
- [21] MONTGOMERY, D.C., JENNINGS C.L. y KULAHCI M. (2015), *Introduction to Time Series Analysis and Forecasting*, segunda edición, USA. John Wiley & Sons.
- [22] NOVALES CINCA, A. (1993), *Econometría*, segunda edición, España. McGraw-Hill.

- [23] PANKRATZ, A. (1991), *Forecasting with Dynamic Regression Models*, primera edición, USA. John Wiley & Sons.
- [24] POHLERT, T. (2018), *Package trend*.
- [25] TSAY, R.S. (1988), *Outliers, Level Shifts, and Variance Changes in Time Series*, Journal of Forecasting, **7**, 1–20.
- [26] TSAY, R.S. (1986), *Time Series Model Specification in the Presence of Outliers*, Journal of the American Statistical Association, **81**, 132–141.
- [27] WEI, W.W.S. (2006), *Time Series Analysis: Univariate and Multivariate Methods*, segunda edición, USA. Pearson.