



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

---

FACULTAD DE CIENCIAS

EVOLUCIÓN DE LOS AYUDANTES DE PROFESOR EN LA  
UNAM, A TRAVÉS DEL ANÁLISIS DE SERIES DE TIEMPO

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

**Actuaria**

PRESENTA:

**Xóchitl Adriana Hernández Luna**

TUTOR:

Dra. Ruth Selene Fuentes García

COTUTOR DE TESIS:

Dr. Víctor Hugo Jacobo Armendáriz



Ciudad Universitaria, CD. MX., 2019



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

1. Datos del alumno

Hernández  
Luna  
Xóchitl Adriana  
Universidad Nacional Autónoma de  
México  
Facultad de Ciencias  
Actuarial  
307264890

2. Datos del tutor

Dra.  
Ruth Selene  
Fuentes  
García

3. Datos del sinodal 1

Dr.  
Ricardo  
Ramírez  
Aldana

4. Datos del sinodal 2 - Cotutor

Dr.  
Víctor Hugo  
Jacobó  
Armendáriz

5. Datos del sinodal 3

Act.  
Jaime  
Vázquez  
Alamilla

6. Datos del sinodal 4

Mat.  
Margarita Elvira  
Chávez  
Cano

7. Datos del trabajo escrito.

Evolución de los ayudantes de profesor en la UNAM, a través del análisis de series de tiempo  
Número de páginas 82  
2019

*Para mi familia.*



# Agradecimientos

A la Dr. Ruth por se una guía, muchas gracias por tu paciencia.

Gracias al Dr. Víctor Hugo.

A la Universidad Nacional Autónoma de México por la educación que me ha dado.

A la Dirección General del Personal Académico por poder desarrollar este trabajo.

A los sinodales Dr. Ricardo Ramírez Aldana, Mat. Margarita Elvira Chávez Cano y Act. Jaime Vázquez Alamilla, por sus valiosas observaciones.



# Índice general

<b>Agradecimientos</b>	<b>V</b>
<b>Índice de figuras</b>	<b>IX</b>
<b>Índice de tablas</b>	<b>XI</b>
<b>Acronimos</b>	<b>XIII</b>
<b>Introducción</b>	<b>XV</b>
<b>1. Series de tiempo.</b>	<b>1</b>
1.1. Análisis clásico . . . . .	1
1.2. Series de tiempo estacionarias. . . . .	2
1.3. Modelos para series estacionarias . . . . .	3
1.4. Modelos de medias móviles . . . . .	5
1.5. Modelos autorregresivos . . . . .	7
1.6. Función de autocorrelación parcial . . . . .	9
1.7. Modelos autorregresivos y de promedios móviles . . . . .	10
1.8. Modelos autorregresivos integrados de promedios móviles . . . . .	10
1.9. Procesos ARIMA estacionales . . . . .	11
1.10. Pronóstico . . . . .	12
1.10.1. Esperanza condicional . . . . .	12
1.10.2. Proyección lineal óptima . . . . .	13
1.11. Construcción del modelo . . . . .	15
1.12. Estimación de parámetros . . . . .	17
1.12.1. Momentos . . . . .	17
1.12.2. Mínimos cuadrados . . . . .	17
1.12.3. Máxima verosimilitud . . . . .	18
1.13. Diagnóstico y selección del modelo . . . . .	20
1.13.1. Criterio de selección de modelos . . . . .	20
1.13.2. Prueba de correlación . . . . .	20
1.13.3. Prueba de normalidad . . . . .	21



<b>2. Descripción de la población de ayudantes de la UNAM</b>	<b>23</b>
2.1. Estatuto del Personal Académico . . . . .	23
2.2. Datos de los ayudantes de la UNAM . . . . .	24
<b>3. Evolución de la población de ayudantes de la UNAM de 1999 a 2014</b>	<b>31</b>
3.1. Análisis de la población de ayudantes de la UNAM de 2001 a 2014 . . . . .	40
3.2. Análisis de la población de ayudantes de la UNAM de 2009 a 2014 . . . . .	46
Otro modelo . . . . .	50
<b>4. Conclusiones</b>	<b>59</b>
<b>Bibliografía</b>	<b>65</b>

# Índice de figuras

2.1. Pirámides Poblacionales 1999, 2001, 2009 y 2014 . . . . .	27
2.2. Antigüedad 1999, 2001, 2009 y 2014 . . . . .	28
2.3. Número de horas contratadas por nombramiento de ayudante en 1999, 2001, 2009 y 2014 . . . . .	29
3.1. Población de ayudantes 1999-2014. . . . .	32
3.2. ACF y PACF de la población de ayudantes apartir de 1999. . . . .	33
3.3. Gráfica de $(1 - B)(1 - B^{24})Y_t$ . . . . .	34
3.4. Periodograma. . . . .	35
3.5. Gráfica de la subserie con ciclo estacional. . . . .	36
3.6. Serie con la diferencia $(1 - B)(1 - B^{12})Y_t$ . . . . .	37
3.7. P-valores de la prueba Ljung-Box para los residuales. . . . .	39
3.8. Gráficas para verificar los residuales del modelo ARIMA(1,1,1)(2,1,1)[12], para la serie de ayudantes a partir de 1999. . . . .	40
3.9. Pronósticos para un año, con la serie 1999 . . . . .	41
3.10. Pronósticos para cinco años, con la serie 1999 . . . . .	41
3.11. Gráfica del comportamiento estacional. . . . .	42
3.12. $(1 - B)(1 - B^{12})X_t$ para la serie a partir de 2001. . . . .	43
3.13. Análisis de residuales del modelo ARIMA(1,1,1)(0,1,1)[12] para 2001. . . . .	44
3.14. Análisis de residuales del modelo ARIMA(1,1,1)(1,1,1)[12] para 2001. . . . .	45
3.15. Pronósticos para un año, con la serie 2001 . . . . .	46
3.16. Pronósticos para cinco años, con la serie 2001 . . . . .	47
3.17. Comparación de pronósticos puntuales para las series 1999 vs 2001. . . . .	48
3.18. Comparación de pronósticos con intervalos de confianza para las series 1999 vs 2001. . . . .	49
3.19. Serie a partir de 2009 con la diferencia $(1 - B)(1 - B^{12})Z_t$ . . . . .	50
3.20. Análisis de residuales del modelo ARIMA(1,1,1)(0,1,1)[12] para la serie 2009. . . . .	51
3.21. Análisis de residuales del modelo ARIMA(1,1,1)(1,1,1)[12] para la serie 2009. . . . .	52
3.22. Pronósticos para un año, con la serie 2009 . . . . .	53
3.23. Pronósticos para cinco años, con la serie 2009 . . . . .	54
3.24. Comparación de pronósticos puntuales para las series 1999 vs 2001 vs 2009. . . . .	54
3.25. Comparación de pronósticos con intervalos de confianza para las series 1999 vs 2001 vs 2009. . . . .	55
3.26. Serie de residuales $v_t$ para modelar la serie a partir de 2009. . . . .	55
3.27. Serie de residuales diferenciada $(1 - B)v_t$ para modelar la serie a partir de 2009. . . . .	57
3.28. Análisis de residuales para el ajustes de la serie 2009. . . . .	58
3.29. Pronósticos $Z_t = \beta_0 + \sum_{i=1}^{23} \beta_i \delta_j(t) + \frac{(1 - \theta_1)}{(1 - B)(1 - \phi_1)(1 - \Phi_{12})} \varepsilon_t$ . . . . .	58

4.1. Número ayudantes contra los pronósticos realizados para el modelado en 1999, 2001, 2009 y 2009 con variables mudas . . . . .	60
4.2. Estadísticas oficiales para 2015 y 2016 . . . . .	62
4.3. Estadísticas oficiales para 2017 y 2018 . . . . .	63

# Índice de tablas

2.1. Número de nombramientos por ayudante . . . . .	25
2.2. Nombramientos por figura . . . . .	25
3.1. Desviación estándar de la serie y sus respectivas diferencias. . . . .	33
3.2. Modelos para la serie de ayudantes a partir de 1999. . . . .	37
3.3. Parámetros estimados para 1999 . . . . .	38
3.4. Parámetros estimados para 1999 . . . . .	38
3.5. Modelos para la serie 2001 . . . . .	42
3.6. Parámetros estimados para 2001 . . . . .	46
3.7. Modelos para la serie 2009. . . . .	47
3.8. Parámetros estimados para 2009 . . . . .	49
3.9. Coeficientes del modelo con variables dummies . . . . .	56
3.10. Parámetros estimados con su intervalo de confianza. . . . .	57
4.1. Comparación de las estadísticas oficiales con los pronósticos . . . . .	60



# Acrónimos

**ACF** Función de autocorrelación

**PACF** Función de autocorrelación parcial

**MA(q)** Modelo de media móvil de orden q

**AR(p)** Modelo autorregresivo de orden p

**ARMA(p,q)** Modelo autorregresivo y de promedios móviles

**ARIMA(p,d,q)** Modelo autorregresivo integrado de promedios móviles

**AIC** *Akaike Information Criterion*

**AICc** *Akaike Information Criterion Corrected*

**BIC** *Bayesian information Criterion*



# Introducción.

El personal académico de la UNAM está conformado por las siguientes figuras: Profesores e Investigadores, Técnicos académicos y Ayudantes de profesor o de investigador. Este estudio se enfoca en los ayudantes, definidos según el Estatuto del Personal Académico de la UNAM (EPA).

¿Porqué enfocarse en la población de ayudantes? generalmente los análisis estadísticos sobre los académicos en la UNAM se centran principalmente en el estudio del personal de carrera y el resto a profesores de asignatura. Así que esta tesis surge en la búsqueda de una mejor representación de la población de ayudantes.

Comenzando con un análisis descriptivo de los datos de la población de ayudantes que brindará una breve descripción de su composición por sexo y edad, su antigüedad, las horas que asisten en la enseñanza o investigación. Sin embargo, para examinar como ha evolucionado la población de ayudantes a lo largo de dieciséis años, se realizará un análisis de la serie de tiempo del número de ayudantes y se buscará predecir la población para los siguientes años.

¿Porqué tomar los datos quincenalmente? generalmente de las estadísticas disponibles se encuentran reportadas en forma anual, resultando en pocos datos para realizar un análisis de series de tiempo, pero se podría modelar con una regresión o interpolación, incluso con algún método demográfico. Sin embargo, los datos con los que se cuentan para este análisis es el número de ayudantes desde la primera quincena de 1999 hasta la quincena veinticuatro de 2014. Al observar la población de ayudantes quincenalmente, se logra apreciar el aumento de ayudantes al principio y final de cada semestre y los cambios en el crecimiento de la población de ayudantes, además de tener más datos para llevar a cabo el análisis de series de tiempo que seguirá una descomposición clásica, pues la serie posee tendencia y componente estacional.

La serie refleja algunos momentos en los cuales el comportamiento del personal no fue regular, por ejemplo, la huelga que se dio de 1999 a 2000. Estos eventos irregulares no permitirían un análisis de series de tiempo basado en un modelo lineal, como en los modelos ARMA.

Por ello se considera el periodo de 2001 a 2014, donde el comportamiento de la serie se muestra más regular después de registrarse un evento atípico, aunque al ajustar un modelo ARMA no se logra captar toda la relación de dependencia entre las observaciones.

Sin embargo, con el fin de disponer de un mejor pronóstico y teniendo en cuenta que el comportamiento de la población de ayudantes no se estabiliza completamente sino hasta 2009 y dado que la incidencia de algún evento atípico no se puede prever, se propone otro periodo de observación 2009 a 2014.

En el Capítulo 1 se presenta en resumen la teoría básica para el análisis de series de tiempo con modelos autorregresivos y de promedios móviles. En el Capítulo 2 los datos obtenidos del análisis descriptivo son contrastados con el EPA. Mientras en el Capítulo 3 se escudriña la evolución de los ayudantes aplicando la



teoría expuesta en el Capítulo uno, dando como resultado un modelo preferible, ésto aunado con las consideraciones descritas anteriormente, deriva en proponer diversos periodos de observación analizarlos de nueva cuenta y dar previsiones para los próximos cinco años y comparar los ajustes. Por ultimo en el Capítulo 4 se concluye con las previsiones del número de ayudantes y algunas consideraciones para tratar de mejorar el cumplimiento del EPA.

# Capítulo 1

## Series de tiempo.

Las series de tiempo son sucesiones de variables aleatorias, o bien procesos estocásticos. El análisis de la serie que se aplicará en este estudio, se centra en las relaciones o la estructura de relación de las variables en el tiempo, para proponer un modelo estadístico capaz de describir tal comportamiento dando lugar a una herramienta para realizar pronósticos.

Si una serie se puede predecir exactamente con funciones matemáticas se dice que la serie es determinista, mientras que si la serie es de alguna manera determinada por sus valores pasados se llama estocástica y los pronósticos estarán condicionados a los valores anteriores.

En este capítulo se presenta un resumen de los conceptos básicos, para un revisión exhaustiva ver Montgomery et al. (2011), Brockwell and Davis (2002), Wei (2006) y Hamilton (1994).<sup>1</sup>

### 1.1. Análisis clásico.

Siguiendo un enfoque clásico para el análisis de una serie de tiempo, el comportamiento de la serie se descompone en tendencia, variación estacional o componente estacional y el remanente de variación que no se puede explicar es llamado la componente irregular o aleatorio o de errores.

La **tendencia** se puede definir como un cambio sistemático, no periódico. Es decir, un movimiento suave y regular a lo largo del tiempo.

La **variación estacional** es un patrón de comportamiento que se repite en un periodo fijo, por ejemplo un año, un semestre. Si la serie tiene oscilaciones, pero estas no corresponden a una frecuencia fija, se tiene un **comportamiento cíclico** y se debe incluir otra componente  $C_t$

La **componente irregular** es la serie que resulta de quitar la tendencia, la componente estacional y los ciclos, esta serie puede o no ser aleatoria.

---

<sup>1</sup>Mientras que para el tratamiento de series de tiempo en R se recomienda consultar a Shumway and Stoffer (2006), Cowpertwait and Metcalfe (2009) y Hyndman and Athanasopoulos (2014)

En un esquema aditivo la serie se puede descomponer en:

$$Y_t = m_t + S_t + I_t. \quad (1.1)$$

donde  $Y_t$  es la serie observada en el tiempo  $t$ ,  $m_t$  es la tendencia,  $S_t$  la componente estacional e  $I_t$  la componente irregular. Este esquema se utiliza cuando la variación de la serie no muestra demasiados cambios en el tiempo.<sup>1</sup>

El objetivo de la descomposición clásica es estimar las componentes de tendencia y estacionalidad, para dejar una serie de residuales con un comportamiento más estable en el tiempo, que permita realizar pronósticos.

## 1.2. Series de tiempo estacionarias.

De manera formal, una serie de tiempo es una realización de una sucesión de variables aleatorias indexadas y definidas sobre un intervalo de tiempo  $\{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}\}$ . Es decir, una observación de un proceso estocástico, cuyo comportamiento está determinado por su función de distribución conjunta y todos los subconjuntos finitos de las funciones de distribución de las variables involucradas, ver Madsen (2007) y Cryer and Chan (2008). Sea

$$f_{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}}(y_1, y_2, \dots, y_n) = \mathbb{P}(Y_{t_1} \leq y_1, Y_{t_2} \leq y_2, \dots, Y_{t_n} \leq y_n).$$

la función de distribución conjunta.

**Definición 1.2.1.** Un proceso  $\{Y(t)\}$  se dice **fuertemente estacionario o estrictamente estacionario** si para todas las distribuciones de dimensión finita, éstas son invariantes a cambios en el tiempo, es decir, para cada  $n$ , cualquier conjunto de puntos en el tiempo  $\{t_1, t_2, \dots, t_n\}$  y para cualquier rezago  $h$  se tiene:

$$f_{Y_{t_1}, \dots, Y_{t_n}}(y_1, \dots, y_n) = f_{Y_{t_1+h}, \dots, Y_{t_n+h}}(y_1, \dots, y_n). \quad (1.2)$$

Conocer la función de distribución conjunta para datos no simulados no es el objetivo del análisis de la serie de tiempo. Las funciones de la media, varianza y covarianza dan un resumen de la información sobre la estructura de relación proporcionada por la función de distribución conjunta, lo cual se utilizará para definir una versión más accesible.

**Definición 1.2.2.** Un proceso  $\{Y(t)\}$  se dice **débilmente estacionario o estacionario de segundo orden o procesos de covarianza estacionaria** sí:

- I  $\mu_t$  es independiente de  $t$ ,
- II  $\gamma(t+h, t)$  es independiente de  $t$  para cada incremento  $h$ .

---

<sup>1</sup>Para ver los demás esquemas y en qué condiciones se utilizan se puede consultar Guerrero (2003) y González (2011)

Donde:

**Definición 1.2.3.** La **función esperanza o media** de  $\{Y(t)\}$  se define como:

$$\mu_{Y_t} = \mathbb{E}(Y_t), \quad \forall t \in \mathbb{Z}.$$

**Definición 1.2.4.** La **función de autocovarianza (ACVF)** para un proceso estacionario  $\{Y(t)\}$  se puede expresar como:

$$\gamma(h) = \text{Cov}(Y_t, Y_{t+h}) = \mathbb{E}[(Y_t - \mu)(Y_{t+h} - \mu)], \quad \forall h \in \mathbb{Z}. \quad (1.3)$$

**Definición 1.2.5.** La **función de autocorrelación (ACF)** para un proceso estacionario  $\{Y(t)\}$  se puede expresar como:

$$\rho(h) = \frac{\text{Cov}(Y_t, Y_{t+h})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_{t+h})}} = \frac{\gamma(h)}{\sqrt{\gamma(0)}\sqrt{\gamma(0)}} = \frac{\gamma(h)}{\gamma(0)}, \quad \forall h \in \mathbb{Z}. \quad (1.4)$$

En otras palabras para que una serie sea débilmente estacionaria es necesario que la media, así como la varianza sean constantes en el tiempo y que la covarianza no dependa del tiempo sino de la distancia temporal entre las observaciones. El número de pasos en el tiempo o **rezago** de la serie es conocido como **lag**.

Para realizar inferencia estadística sobre una serie de tiempo es necesario suponer que la serie es estacionaria, es decir, que la estructura se mantiene a lo largo del tiempo. En este trabajo, cuando se mencione una serie estacionaria se referirá a un proceso débilmente estacionario.

Las funciones de autocovarianza y autocorrelación miden el grado de asociación entre la serie con una copia de sí misma retrasada  $h$  puntos de tiempo Janacek (2001), solo que la autocorrelación es más fácil de interpretar, ya que la función de autocovarianza depende de la unidad de la medida de las observaciones resultando en una interpretación carente sentido de la relación entre los datos, mientras la autocorrelación evalúa los datos y los normaliza<sup>1</sup>, haciendo el rango de variación de  $[-1, 1]$ , para ver las diferentes interpretaciones de correlación positiva, negativa y nula ver Rincón (2007).

Dentro del análisis de series de tiempo la gráfica del ACF o correlograma jugará un papel muy importante pues se utilizará como un indicador de la existencia de la asociación entre las observaciones, además de emplearla como un indicador en la identificación del modelo para la serie.

### 1.3. Modelos para series estacionarias

Un proceso estacionario de gran interés es una sucesión de **ruido blanco**, se define como una sucesión de variables aleatorias no correlacionadas  $\{\varepsilon_t\}$ , con media cero y varianza finita  $\sigma_\varepsilon^2$ , denotado por

<sup>1</sup>Normalizar un dato es hacer lo de longitud uno. Y se refiere a restarle su media y dividirlo entre su desviación estándar.

$\{\varepsilon_t\} \sim WN(0, \sigma^2)$ , está caracterizado por:

$$\begin{aligned} \mu_\varepsilon &= \mathbb{E}[\varepsilon_t] = 0, \quad \forall t. \\ \gamma_\varepsilon(h) &= \text{Cov}(\varepsilon_t, \varepsilon_{t+h}) = \begin{cases} \sigma^2, & \text{si } h = 0, \\ 0, & \text{si } h \neq 0. \end{cases} \quad \forall t. \end{aligned} \quad (1.5)$$

Los ruidos blancos deben su importancia a las propiedades estadísticas que poseen y que los hacen un elemento esencial en el modelado ARMA, durante la estimación de parámetros, las pruebas de bondad de ajuste y hasta para realizar los pronósticos.

**Definición 1.3.1.** Un filtro lineal de un proceso de ruido blanco es un **proceso lineal**, denotado  $\{Y(t)\}$ , se define como una combinación lineal de perturbaciones aleatorias de un ruido blanco, es decir:

$$Y_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}, \quad \forall t. \quad (1.6)$$

donde  $\mu$  es la media común,  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$  y  $\{\psi_j\}$  es una sucesión de ponderadores tales que  $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ .

**Definición 1.3.2.** Un proceso lineal es **causal** si  $\psi_j = 0, \forall j < 0$ , es decir:

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad \forall t. \quad (1.7)$$

donde  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$  y  $\{\psi_j\}$  es una sucesión de constantes tales que  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ .

Si un proceso  $Y_t$  es causal entonces no depende del futuro, debido a que es una suma ponderada de perturbaciones aleatorias pasadas y la presente.

**Proposición 1.3.1.** *Si un proceso  $Y_t$  es causal entonces es estacionario.*<sup>1</sup>

*Demostración.* Como  $Y_t$  está caracterizado por:<sup>2</sup>

$$\begin{aligned} \mu_Y &= \mathbb{E}[Y_t] = \mu, \quad \forall t. \\ \gamma_Y(h) &= \text{Cov}(Y_t, Y_{t+h}) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \gamma_\varepsilon(i-j+h) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h}. \end{aligned} \quad (1.8)$$

Debido a que la media es independiente del tiempo y la función de covarianza también lo es, se cumple con las condiciones para ser estacionario de la definición 1.2.2, por lo tanto es estacionario. ■

<sup>1</sup>Recordando que cuando se mencione estacionario se habla en el sentido débilmente estacionario

<sup>2</sup>Ver Madsen (2007)

Por el Teorema de descomposición de Wold, cualquier serie de tiempo débilmente estacionaria  $Y_t$  puede ser representada de forma causal como en la expresión 1.7, siempre que  $\{\psi_j\}$  satisfagan  $\sum_{j=0}^{\infty} |\psi_j|^2 < \infty$ . Ver Hamilton (1994).

El **operador de rezago o backward shift** permite tener una notación simplificada en el análisis de series de tiempo, denotado por  $B$  se define como:

$$BY_t = Y_{t-1}, \quad \forall t \in \mathbb{Z}. \quad (1.9)$$

es decir, retrasa a la serie a un periodo, así como para cualquier retraso o lag  $j$  se tiene:

$$B^j Y_t = Y_{t-j}, \quad \forall t \in \mathbb{Z}. \quad (1.10)$$

Retomando la expresión causal 1.7 y suponiendo  $\mu = 0$  y  $\psi_0 = 1$ , se introduce a  $\psi(B)$ , un operador lineal talque:

$$\psi(B) = 1 + \sum_{j=1}^{\infty} \psi_j B^j. \quad (1.11)$$

entonces le expresión para un proceso causal 1.7, para  $\mu = 0$  se puede expresar como:

$$Y_t = \psi(B) \varepsilon_t. \quad (1.12)$$

donde  $\psi(B)$  es conocida como la función transferencia del proceso (transfer function).

**Definición 1.3.3.** Un proceso lineal es **invertible** si puede expresarse como:

$$\pi(B)Y_t = \sum_{j=0}^{\infty} \pi_j Y_{t-j} = \varepsilon_t. \quad (1.13)$$

$$\text{donde } \pi(B) = \sum_{j=0}^{\infty} \pi_j B^j, \quad \text{y} \quad \sum_{j=0}^{\infty} |\pi_j| < \infty.$$

A  $\pi(B)$  se le conoce como la forma invertible del proceso. Claramente  $\psi(B)$  es el operador inverso  $\pi(B)$ , es decir:

$$\pi(B) \psi(B) = 1 \Leftrightarrow \pi(B) = \psi^{-1}(B). \quad (1.14)$$

*Sin embargo su existencia dependerá del dominio de  $\psi^{-1}(B)$ . Si existe  $\pi(B)$  se puede determinar usando la expansión en series de Taylor de  $\psi^{-1}(B)$ . Para una revisión más profunda de las propiedades de los procesos lineales, referirse a Madsen (2007).*

## 1.4. Modelos de medias móviles

**Definición 1.4.1.** Un proceso  $\{Y(t)\}$  sigue un modelo de **media móvil de orden  $q$** , denotado por  $MA(q)$ , para  $q = 1, 2, \dots$ , si  $\{Y(t)\}$  es un proceso estacionario de segundo orden y puede expresarse como:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad \forall q \in \mathbb{N} - \{0\}. \quad (1.15)$$

donde  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$  y  $\theta_1, \dots, \theta_q$  son constantes.

Utilizando el operador de rezago para expresar la ecuación (1.15) se obtiene:

$$\begin{aligned} Y_t &= \mu + \varepsilon_t + \theta_1 B \varepsilon_t + \dots + \theta_q B^q \varepsilon_t. \\ &= \mu + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t. \\ &= \mu + \theta_q(B) \varepsilon_t, \quad \forall q \in \mathbb{N} - \{0\}. \end{aligned} \quad (1.16)$$

donde  $\theta_q(B) = 1 + \theta_1 B + \dots + \theta_q B^q$  es un polinomio en  $B$  de grado  $q$ . La ecuación  $\theta_q(B) \varepsilon_t = 0$  es llamada **ecuación característica**, en donde  $B$  es una raíz del polinomio y puede ser un número real o complejo.

Un proceso  $MA(q)$  esta caracterizado por:

Sea  $\theta_0 = 1$

$$\mathbb{E}[Y_t] = \mu + \sum_{i=0}^q \theta_i \mathbb{E}[\varepsilon_{t-i}] = \mu. \quad (1.17)$$

$$Var(Y_t) = \gamma_Y(0) = Var\left(\mu + \sum_{i=0}^q \theta_i \varepsilon_{t-i}\right) = \sigma^2 (1 + \theta_1^2 + \dots + \theta_q^2). \quad (1.18)$$

$$\begin{aligned} \gamma_Y(h) &= Cov(Y_t, Y_{t+h}) = \mathbb{E}\left[\left(\sum_{i=0}^q \varepsilon_{t-i}\right)\left(\sum_{j=0}^q \varepsilon_{t+h-j}\right)\right] \\ &= \begin{cases} \sigma^2 (-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_{q-h} \theta_q), & \text{si } 0 \leq h \leq q, \\ 0, & \text{si } h > q. \end{cases} \end{aligned} \quad (1.19)$$

$$\rho_h = \begin{cases} \frac{-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_{q-h} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & \text{si } h = 1, 2, \dots, q, \\ 0, & \text{si } h > q. \end{cases} \quad (1.20)$$

Para consultar los pasos intermedios ver Brockwell and Davis (2002). Nótese que:

- Un modelo  $MA(q)$  es estacionario, puesto que es una suma finita de términos de ruido blanco estacionario.<sup>1</sup>
- Las funciones de autocovarianza y autocorrelación se hacen cero para rezagos mayores el parámetro  $q$ .

**Teorema 1.4.1.** *Un proceso  $MA(q)$ ,  $\{Y_t\}$  definido por  $Y_t = \mu + \theta_q(B) \varepsilon_t$ , se dice invertible sí solo sí las raíces de la ecuación  $\theta_q(B) = 1 + \theta_1 B + \dots + \theta_q B^q = 0$  están fuera círculo unitario.*

La prueba se encuentra en Madsen (2007). Para ejemplificar considere, dos procesos  $MA(1)$  dados por  $X_t = \varepsilon_t + \theta \varepsilon_{t-1}$ ,  $Y_t = \varepsilon_t + \frac{1}{\theta} \varepsilon_{t-1}$ , con  $\theta \in \mathbb{R} - \{0\}$ ,  $Var(X_t) = \sigma^2 (1 + \theta^2)$ ,

$$Var(Y_t) = \sigma^2 \left(1 + \frac{1}{\theta^2}\right), \quad \rho_{X_t}(h) = \rho_{Y_t}(h) = \begin{cases} 1 & \text{si } h = 0 \\ \frac{\theta}{1 + \theta^2} & \text{si } h = 1 \\ 0 & \text{si } h > 1 \end{cases}, \text{ sus respectivas funciones de auto-}$$

correlación se concluye que es la misma para ambos, por lo que al restringir el parámetro  $\theta$  para un modelo

<sup>1</sup>Visto de otra manera un modelo  $MA(q)$  es un proceso lineal con  $q$  ponderadores distintos de cero.

$MA(1)$  asegura una función de autocorrelación única y exclusiva. Sin embargo, más adelante se verá como cumpliéndose la condición de invertibilidad existe una expresión de la series en términos del pasado útil para el pronóstico.

## 1.5. Modelos autorregresivos

**Definición 1.5.1.** Un proceso  $\{Y(t)\}$  sigue un modelo **autorregresivo de orden  $p$** , denotado por **AR( $p$ )**, para  $q = 1, 2, \dots$ , si cumple que:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t. \quad (1.21)$$

donde  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ .

Utilizando el operador rezago para expresar (1.21)

$$\begin{aligned} Y_t - \phi_1 B^1 Y_t - \dots - \phi_p B^p Y_t &= \varepsilon_t. \\ (1 - \phi_1 B^1 - \dots - \phi_p B^p) Y_t &= \varepsilon_t. \\ \phi_p(B) Y_t &= \varepsilon_t. \end{aligned} \quad (1.22)$$

Nótese que cualquier modelo  $AR(p)$  es invertible por como está definido el modelo, sin embargo, no se puede garantizar que este sea estacionario, para ello se hará uso del siguiente resultado.

**Teorema 1.5.1.** *Un proceso  $AR(p)$ ,  $\{Y_t\}$  definido por  $\phi_p(B) Y_t = \varepsilon_t$ , es estacionario si solo si las raíces de la ecuación  $\phi_p(B) = 1 - \phi_1 B^1 - \dots - \phi_p B^p = 0$  están fuera del círculo unitario.*

La prueba se encuentra en Madsen (2007) o Brockwell and Davis (2009). Para ilustrar dicho resultado considere un modelo  $AR(1)$  dado por  $Y_t - \phi_1 Y_{t-1} = \varepsilon_t$ , se puede expresar como

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t \quad \Leftrightarrow \quad Y_t (1 - \phi_1 B) = \varepsilon_t.$$

Entonces:

$$Y_{t-1} = \phi_1 Y_{t-2} + \varepsilon_{t-1}.$$

sustituyendo en la ecuación anterior

$$Y_t = \phi_1^2 Y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t.$$

iterando  $k$  veces

$$Y_t = \phi_1^{k+1} Y_{t-k-1} + \phi_1^k \varepsilon_{t-k} + \dots + \phi_1 \varepsilon_{t-1} + \varepsilon_t,$$

Como  $Y_t$  es estacionario se tiene  $\mathbb{E}[Y_t^2]$  es finita e independiente de  $t$ , tomando el limite en media cuadrada de  $\mathbb{E}\left[Y_t - \sum_{j=0}^{\infty} \phi_1^k \varepsilon_{t-j}\right]^2$ , ver Brockwell and Davis (2002), se tiene;

$$Y_t = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-j}.$$



para ser estacionario es necesario que  $|\phi_1| < 1$ . Además

$$\mathbb{E}[Y_t] = \sum_{j=0}^{\infty} \mathbb{E}[\phi_1^j \varepsilon_{t-j}] = 0.$$

y

$$\begin{aligned} \text{Cov}(Y_{t+h}, Y_t) &= \lim_{n \rightarrow \infty} \mathbb{E} \left[ \left( \sum_{j=0}^n \phi_1^j \varepsilon_{t+h-j} \right) \left( \sum_{k=0}^n \phi_1^k \varepsilon_{t-k} \right) \right], \\ &= \sigma_\varepsilon^2 \phi_1^{|h|} \sum_{j=0}^{\infty} \phi_1^{2j}, \\ &= \sigma_\varepsilon^2 \phi_1^{|h|} \frac{1}{1 - \phi_1^2}. \end{aligned} \quad (1.23)$$

Así  $Y_t = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-k}$  es la expresión de un modelo  $AR(1)$  en un modelo  $MA(\infty)$ , que cumple con la ecuación autorregresiva característica del  $AR(1)$ , por lo tanto  $|\phi_1| < 1$  es la única solución estacionaria. ver Brockwell and Davis (2009). Con lo cual la raíz del polinomio asociado  $B = \frac{1}{\phi_1}$  está fuera del círculo unitario.

La función de autocovarianza para los modelos  $AR(p)$  satisface la siguiente ecuación diferencial lineal:

$$\gamma(k) + \phi_1 \gamma(k-1) + \dots + \phi_p \gamma(k-p) = 0, \quad \text{para } k = 1, 2, \dots, \quad (1.24)$$

con la condición inicial dada por

$$\gamma(0) + \phi_1 \gamma(-1) + \dots + \phi_p \gamma(-p) = \sigma_\varepsilon^2. \quad (1.25)$$

Para ello, se multiplica 1.21 por  $Y_{t-k}$  y se toma esperanza para obtener:

$$\mathbb{E}[Y_{t-k} Y_t + \phi_1 Y_{t-k} Y_{t-1} + \dots + \phi_p Y_{t-k} Y_{t-p}] = \mathbb{E}[Y_{t-k} \varepsilon_t].$$

Considerando que:

$$\mathbb{E}[Y_{t-k} \varepsilon_t] = \begin{cases} \sigma_\varepsilon^2, & k = 0, \\ 0 & k = 1, 2, \dots \end{cases},$$

y

$$\mathbb{E}[Y_{t-k} Y_t + \phi_1 Y_{t-k} Y_{t-1} + \dots + \phi_p Y_{t-k} Y_{t-p}] = \gamma(k) + \phi_1 \gamma(k-1) + \dots + \phi_p \gamma(k-p).$$

De sustituir  $k = 1, 2, \dots, p$  en 1.24 y dividir lo por  $\gamma(0)$  se obtiene las llamadas ecuaciones de Yule-Walker, dadas por:

$$\begin{pmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{pmatrix} = \begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \vdots & \vdots & \dots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix}, \quad (1.26)$$

La prueba de la solución explícita de la ecuación diferencial 1.24, se encuentra en Madsen (2007). Si se vuelve a dividir 1.26 por  $\gamma(0)$ , se tiene que la función de autocorrelación también cumple con 1.24.

## 1.6. Función de autocorrelación parcial

La función de autocorrelación parcial mide la correlación entre observaciones separadas  $k$  periodos de tiempo, sin los efectos de las observaciones intermedias, se denotará por  $\phi_{kk}$ . La definición formal está dada por 1.27, aunque hay diversas expresiones equivalentes, que buscan ayudar en el cálculo de la función.

**Definición 1.6.1.** La función de autocorrelación parcial (PACF) para un proceso estacionario  $\{Y(t)\}$  está definida como:

$$\phi_{kk} = \text{Corr}[Y_t, Y_{t+k} | Y_{t+1} \dots Y_{t+k-1}]. \quad (1.27)$$

Un enfoque es considerar la función de autocorrelación parcial como un modelo de regresión donde la variable dependiente  $Y_{t+k}$  de una proceso estacionario de media cero, está determinada por  $k$  variables rezagadas  $Y_{t+k-1}, Y_{t+k-2}, \dots, Y_t$ , es decir:

$$Y_{t+k} = \phi_{k1}Y_{t+k-1} + \phi_{k2}Y_{t+k-2} + \dots + \phi_{kk}Y_t + \varepsilon_{t+k}, \quad (1.28)$$

donde  $\phi_{k,i}$  denota el  $i$ -ésimo parámetro de regresión y  $\varepsilon_{t+k}$  es un término de error con media cero y no correlacionado con  $Y_{t+k-j}$  para  $j = 1, 2, \dots, k$ . Si se multiplica 1.28 por  $Y_{t+k-j}$  y tomando la esperanza se obtiene:

$$\gamma_j = \phi_{k1}\gamma_{j-1} + \phi_{k2}\gamma_{j-2} + \dots + \phi_{kk-1}\gamma_{j-k+1} + \phi_{kk}\gamma_{j-k} \quad 1 \leq j \leq k.$$

O su equivalente, más usada, dada por:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk-1}\rho_{j-k+1} + \phi_{kk}\rho_{j-k}, \quad 1 \leq j \leq k.$$

el anterior sistema de ecuaciones se puede expresar en términos de las ecuaciones de Yule-Walker como:

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_{k-1} \\ \rho_k \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-2} & \rho_{k-3} & \rho_{k-4} & \dots & 1 & \rho_1 \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{pmatrix} \begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk-1} \\ \phi_{kk} \end{pmatrix}, \quad (1.29)$$

donde  $\phi_{kk}$  es la correlación parcial al rezago  $k$ , la ecuación 1.29 se puede expresar de forma simplificada en términos de matrices dada por:

$$\mathbf{P}_k \phi_k = \rho_k. \quad (1.30)$$

La función de autocorrelación parcial (PACF) se obtiene resolviendo la ecuación 1.29, iteradamente con respecto a  $\phi_{kk}$  para  $k = 1, 2, \dots$ , usando la regla de Cramèr se tiene:

$$\phi_{11} = \rho_1, \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}, \dots, \phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}.$$

Para ver los diferentes enfoques para calcular la función de autocorrelación parcial, referirse a Wei (2006), Brockwell and Davis (2009) y Hamilton (1994).

## 1.7. Modelos autorregresivos y de promedios móviles

**Definición 1.7.1.** Un proceso o serie  $\{Y(t)\}$  sigue un modelo **autorregresivo y de promedios móviles**, denotado por **ARMA(p,q)**, para  $p = 1, 2, \dots$  y  $q = 1, 2, \dots$ , si  $\{Y(t)\}$  es estacionaria de segundo orden y puede expresarse como:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (1.31)$$

con  $\phi_p \neq 0$ ,  $\theta_q \neq 0$  donde  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ .

Utilizando el operador rezago para expresar (1.31)

$$\begin{aligned} Y_t - \phi_1 B Y_t - \dots - \phi_p B^p Y_t &= \varepsilon_t + \theta_1 B^1 \varepsilon_t + \dots + \theta_q B^q \varepsilon_t. \\ (1 - \phi_1 B^1 - \dots - \phi_p B^p) Y_t &= (1 + \theta_1 B^1 + \dots + \theta_q B^q) \varepsilon_t. \\ \phi_p(B) Y_t &= \theta_q(B) \varepsilon_t. \end{aligned} \quad (1.32)$$

Condiciones para  $\{Y_t\}$

- Condición de estacionariedad. Las raíces de  $\phi_p(B) = 1 - \phi_1 B^1 - \dots - \phi_p B^p$  deben estar fuera del círculo unitario, es decir, si  $\phi_p(z_i) = 0$ ,  $z_i$  debe cumplir  $|z_i| > 1$ .
- Condición de identificabilidad. Las raíces de  $\theta_q(B) = 1 + \theta_1 B^1 + \dots + \theta_q B^q$  deben estar fuera del círculo unitario, es decir, si  $\theta_q(z_i) = 0$ ,  $z_i$  debe cumplir  $|z_i| > 1$ .
- Condición de parsimonia, el modelo esta expresado en su forma reducida. sí  $\phi_p(B)$  y  $\theta_q(B)$  no tienen raíces comunes.

Para ver una prueba formal de cada una de las condiciones anteriores ver Brockwell and Davis (2009). El modelo ARMA puede ser visto como una regresión del valor presente sobre los últimos valores correlacionados con los errores.

Es necesario satisfacer el supuesto de estacionariedad de segundo orden para hacer un ajuste y análisis de una serie de tiempo observada. Algunos métodos para remover una tendencia y estacionalidad son: el uso de transformaciones, el uso del operador diferencia, el uso del operador de retraso o rezago, uso de polinomios ajustados, suavizamiento de las serie y suavizamiento exponencial. Por lo que un modelo estacionario puede ser adecuado para una serie de residuales que no contenga tendencia y ciclos estacionales.

## 1.8. Modelos autorregresivos integrados de promedios móviles

Una serie que posee tendencia y componente estacional es no estacionaria, ya que generalmente las funciones que originan tales variaciones dependen del tiempo, con lo cual no cumplen con la definición de

estacionariedad. La generalización de los modelos ARMA, utilizados para modelar series estacionarias, son los modelos autorregresivos integrados de promedios móviles (ARIMA) o ARMA integrado, en los cuales el operador diferencia se usa para convertir a una serie no estacionaria en otra que sea estacionaria, por lo cual antes de comenzar es necesario presentar el operador diferencia. Así la primera diferencia en el rezago o lag uno es:

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t. \quad (1.33)$$

donde  $B$  es el operador rezago definido en la expresión 1.9. En la practica solo se le denomina como la diferencia de orden uno. Aplicar el operador diferencia reiteradamente es equivalente a tomar potencias de  $\nabla = (1 - B)$ , para ejemplificarlo, se observará la diferencia de orden dos:

$$\nabla^2 X_t = \nabla(\nabla X_t) = (1 - B)(1 - B)X_{t-1} = (1 - 2B + B^2)X_t = X_t - 2X_{t-1} + X_{t-2}. \quad (1.34)$$

Como se mencionó anteriormente el operador diferencia se utiliza convertir una serie en estacionaria, por lo cual se puede establecer que si una serie posee una raíz unitaria entonces es no estacionaria.

**Definición 1.8.1.** Un proceso o serie  $\{Y(t)\}$  sigue un modelo **integrado autorregresivo y de promedios móviles**, denotado por **ARIMA(p,d,q)**, si la d-ésima diferencia,  $W_t = \nabla^d Y_t = (1 - B)^d Y_t$  es un proceso estacionario ARMA(p,q), se expresa como:

$$\phi_p(B) \nabla^d Y_t = \phi_p(B) (1 - B)^d Y_t = \theta_q(B) \varepsilon_t. \quad (1.35)$$

donde  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ .

Las mismas condiciones de estacionariedad e invertibilidad que se utilizan para los modelos autorregresivos y de media móvil también se aplican a un modelo ARIMA.

## 1.9. Procesos ARIMA estacionales

Como se mencionó en el enfoque clásico 1.1, en las serie de tiempo pueden existir efectos estacionales esta característica donde los datos experimentan variaciones regulares y previsibles que se repiten cada año o cada cierto periodo. Por ejemplo, si una empresa que vende helados verá sus ventas crecer en el verano, pero caer en el invierno.

El período estacional se puede entender como el tiempo que pasa para que se repita el comportamiento en una serie de tiempo, por ejemplo una serie que presenta periodicidad anual, pero si dicha serie consiste de observaciones mensuales, el período será 12, en cambio, si la serie es trimestral, el período será 4. Se denotará al período estacional como  $s$ .

Existen varias modificaciones dentro de los modelos ARIMA para poder explicar la conducta de alguna componente estacional y no estacional, sobretodo para las series donde las dependencias sobre las tendencias pasadas ocurren más marcadas en los múltiplos de un retraso de la temporalidad  $s$  de la componente estacional, por lo cual es apropiado introducir polinomios autorregresivos de medias móviles que identifiquen la

componente estacional.

Se define el operador diferencia estacional en la expresión 1.36 denotado por  $\nabla_s^D X_t$ .

$$\nabla_s^D X_t = (1 - B^s)^D X_t. \quad (1.36)$$

**Definición 1.9.1.** El modelo estacional autorregresivo de medias móviles puro, denotado por  $ARMA(P, Q)_s$ , se puede expresar como:

$$\Phi_P(B^s) Y_t = \Theta_Q(B^s) \varepsilon_t. \quad (1.37)$$

donde los polinomios estacionales asociados están dados por:

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}. \quad (1.38)$$

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \dots + \Theta_Q B^{Qs}. \quad (1.39)$$

En general se puede combinar modelos estacionales y no estacionales para modelar mejor una serie obteniendo modelos  $ARMA(p, q) \times (P, Q)_s$ , dados por

$$\Phi_P(B^s) \phi_p(B) Y_t = \Theta_Q(B^s) \theta_q(B) \varepsilon_t. \quad (1.40)$$

**Definición 1.9.2.** El modelo multiplicativo estacional autorregresivo integrado de medias móviles o  $SARIMA$ , se puede expresar como:

$$\Phi_P(B^s) \phi_p(B) \nabla_s^D \nabla^d Y_t = \Theta_Q(B^s) \theta_q(B) \varepsilon_t. \quad (1.41)$$

donde  $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ , el modelo en general esta denotado por  $ARIMA(p, d, q) \times (P, D, Q)_s$

## 1.10. Pronóstico

Supongasé que se necesita en predecir el valor de  $Y_{t+h}$  para  $h > 0$ , basado en un conjunto de variables  $\mathbf{X}_t$ , observadas al momento  $t$ . En particular la predicción de  $Y_{t+1}$  se le conoce como la predicción a un paso, en base a sus  $m$  valores más recientes. Sea  $\mathbf{X}_t$  es el conjunto de valores  $Y_t, Y_{t-1}, \dots, Y_{t-m+1}$  y  $Y_{t+1|t}^*$  el pronóstico de  $Y_{t+1}$  basado en  $\mathbf{X}_t$ , información presente y pasada de  $Y_t$ .

### 1.10.1. Esperanza condicional

Se asume que la función de pérdida es cuadrática, tal que, se minimice:

$$\mathbb{E} \left( Y_{t+1} - Y_{t+1|t}^* \right)^2 \quad (1.42)$$

La expresión 1.42 es conocida como el Error Cuadrático Medio (MSE) asociado al pronóstico  $Y_{t+1|t}^*$  de notado por:

$$MSE \left( Y_{t+1|t}^* \right) = \mathbb{E} \left( Y_{t+1} - Y_{t+1|t}^* \right)^2.$$

El pronóstico con el menor error cuadrático medio resulta ser la función de esperanza condicional de  $Y_{t+1}$  sobre  $\mathbf{X}_t$ , es decir:

$$Y_{t+1|t}^* = \mathbb{E}(Y_{t+1} | \mathbf{X}_t). \quad (1.43)$$

La verificación de 1.43 se puede ver en Hamilton (1994).

### 1.10.2. Proyección lineal óptima

Restringiendo el pronóstico a la clase de las funciones lineales de  $\mathbf{X}_t$ . La predicción es una combinación lineal de las observaciones pasadas, es decir:

$$P_{t+h}Y_t = Y_{t+h|t}^* = \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + \dots + \alpha_m Y_{t-m+1}. \quad (1.44)$$

Se necesita encontrar el valor de los coeficientes  $(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_m)$  que minimicen el error del pronostico:

$$S(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_m) = \mathbb{E}(P_{t+h}Y_t - \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + \dots + \alpha_m Y_{t-m+1})^2. \quad (1.45)$$

para ello se toman las diferencias parciales tales que:

$$\frac{\partial S(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_m)}{\partial \alpha_j} = 0 \quad \text{para } j = 0, \dots, m. \quad (1.46)$$

Según Brockwell and Davis (2002) se muestra como el cálculo de las ecuaciones 1.46 es equivalente a las siguientes ecuaciones:

$$\mathbb{E} \left[ Y_{t+h} - \alpha_0 - \sum_{i=1}^m \alpha_i Y_{t-i+1} \right] = 0. \quad (1.47)$$

$$\mathbb{E} \left[ \left( Y_{t+h} - \alpha_0 - \sum_{i=1}^m \alpha_i Y_{t-i+1} \right) (Y_{t-j+1}) \right] = 0 \quad \text{para } j = 1, \dots, m. \quad (1.48)$$

En Hamilton (1994), la ecuación 1.48 es la condición de que error del pronostico  $(P_{t+h}Y_t - \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + \dots + \alpha_m Y_{t-m+1})$  no este correlacionado con  $(Y_t, Y_{t-1}, \dots, Y_{t-m+1})$ . Las ecuaciones 1.47 y 1.48 se pueden expresar de forma compacta en notación vectorial donde:

$$\alpha_0 = \mu \left( 1 - \sum_{i=1}^m \alpha_i \right) \quad (1.49)$$

$$\Gamma_m \alpha_m = \gamma_m(h) \quad (1.50)$$

donde

$$\alpha_m = (\alpha_1, \dots, \alpha_m)', \quad (1.51)$$

$$\Gamma_m = [\gamma(i-j)]_{i,j=1}^m, \quad (1.52)$$

$$\gamma_m(h) = (\gamma(h), \gamma(h+1), \dots, \gamma(h+m-1))'. \quad (1.53)$$

Entonces

$$P_{t+h}Y_t = Y_{t+h|t}^* = \mu + \sum_{i=1}^m \alpha_i (Y_{t+1-i} - \mu). \quad (1.54)$$

Donde  $\alpha_m$  satisface 1.50. De la ecuación 1.54 se obtiene que el valor esperado del error de predicción es cero y el error de predicción en media cuadrada esta dado por:

$$\mathbb{E}(Y_{t+h} - P_{t+h}Y_t)^2 = \gamma(0) - 2 \sum_{i=1}^m \alpha_i \gamma(h+i) + \sum_{i=1}^m \sum_{j=1}^m \alpha_i \gamma(i-j) \alpha_j \quad (1.55)$$

$$= \gamma(0) - \alpha'_m \gamma_m(h). \quad (1.56)$$

La última igualdad se cumple por 1.50. Para ver las propiedades de la proyección lineal óptima consultar Brockwell and Davis (2002) y Hamilton (1994).

Nótese que encontrar proyección lineal óptima  $P_{t+h}Y_t$  en términos de  $(Y_t, Y_{t-1}, \dots, Y_{t-m+1})$ , requieren encontrar la solución para un sistema de  $m$  ecuaciones, si  $m$  es grande esto llevaría mucho tiempo y esfuerzo, pero debido a las propiedades de linealidad de la proyección a  $h$ -pasos  $P_{t+h}Y_t$  se puede aplicar un algoritmo recursivo para elaborar dichos cálculos.

En particular, se ocupará algoritmo de Durbin-Levinson. Se pide que  $\Gamma_m$  la matriz de covarianza, sea una matriz no singular, ver la proposición y su prueba en Brockwell and Davis (2009). Entonces

$$P_{t+1}Y_t = Y_{t+1|t}^* = \phi' \mathbf{X}_m = \phi_{m1}Y_m + \dots + \phi_{mm}Y_1. \quad (1.57)$$

donde

$$\phi_m = \Gamma_m^{-1} \gamma_m. \quad (1.58)$$

con  $\gamma_m = (\gamma(1), \dots, \gamma(m))'$ , se define su correspondiente error cuadrático medio como;

$$\nu_m := \mathbb{E}(Y_{t+h} - P_{t+h}Y_t)^2 = \gamma(0) - \phi'_m \gamma_m. \quad (1.59)$$

El *Algoritmo Durbin-Levinson* se define a continuación. Sea  $\{Y_t\}$  un proceso estacionario con media cero y función de autocovarianza  $\gamma(\cdot)$  tal que  $\gamma(0) > 0$  y  $\gamma(h) \rightarrow 0$  cuando  $h \rightarrow \infty$ , entonces los coeficientes  $\phi_{mj}$  y los errores cuadráticos medios  $\nu_m$  definido por 1.57 y 1.54 satisfacen  $\phi_{11} = \frac{\gamma(1)}{\gamma(0)}$  y  $\nu_0 = \gamma(0)$

$$\phi_{mm} = \left[ \gamma(m) - \sum_{j=1}^{m-1} \phi_{m-1,j} \gamma(m-j) \right] \quad (1.60)$$

$$\begin{bmatrix} \phi_{m1} \\ \vdots \\ \phi_{mm} \end{bmatrix} = \begin{bmatrix} \phi_{m-1,1} \\ \vdots \\ \phi_{m-1,m-1} \end{bmatrix} - \phi_{mm} \begin{bmatrix} \phi_{m-1,m-1} \\ \vdots \\ \phi_{m-1,1} \end{bmatrix} \quad (1.61)$$

$$\nu_m = \nu_{m-1} [1 - \phi_{mm}]. \quad (1.62)$$

La prueba del algoritmo se hace mediante el principio de inducción matemática sobre las 1.60 que se encuentra en Brockwell and Davis (2009) y Brockwell and Davis (2002).

Dado el enfoque que se utilizó en la sección 1.6, para definir la función de autocorrelación parcial los coeficientes  $\phi_{j,j}$  son equivalentes a los del algoritmo de Durbin-Levinson, lo cual muestra la relación entre las  $\phi_{j,j}$  y la reducción del error cuadrático medio del pronóstico a un paso, como el número de predictores se incrementa de  $n-1$  a  $n$

## 1.11. Construcción del modelo

Los pasos básicos para ajustar un modelo ARIMA son:

- Graficar los datos.
- Analizar posibles transformaciones de los datos.
- Identificación del orden de dependencia,  $(p, d, q)$ , del modelo.
- Estimación de parámetros.
- Diagnóstico y selección del modelo.

El primer paso en el análisis de cualquier serie de tiempo es graficarla, con el objetivo de identificar discontinuidades, valores atípicos, ver si es necesario aplicar una transformación lineal<sup>1</sup> a los datos o si se puede expresar conforme a la descomposición clásica 1.1. Después de decidir si las posible transformación de la serie, se debe identificar el orden de dependencia, se inicia indagando sobre si es necesario utilizar una diferencia su orden, una serie con tendencia muy probablemente necesitará una diferencia. Sin embargo, las gráficas de las ACF y PACF pueden ayudar a determinar el orden de la diferencia, un decaimiento lento en el ACF indica que es necesario una diferencia, si el PACF muestra una decaimiento ondulado muy marcado sugiere que es necesario una diferencia estacional. Posteriormente de conocer el grado de la diferencia aplicada a la serie, se propondrán los grados de dependencia del modelo  $p$  y  $q$ , para la parte autorregresiva y de media móvil. Una vez que se ha identificado el orden del modelo (es decir, los valores de  $p$ ,  $d$  y  $q$ , es necesario estimar los parámetros,  $\sigma_\varepsilon^2$ ,  $(\phi_1, \dots, \phi_p)$ ,  $(\theta_1, \dots, \theta_q)$ ). Finalmente el diagnóstico de un modelo, dice hasta que punto un modelo puede explicar las asociaciones en los datos y con ello el comportamiento de la serie, básicamente consiste en comprobar si las hipótesis realizadas sobre los residuos se cumplen.

### Gráficas ACF y PACF

Como se mencionó anteriormente las gráficas de las funciones autocorrelación y autocorrelación parcial ayudan a identificar las posibles transformaciones de los datos así como el orden del modelo hasta como indicadores de la bondad de ajuste, pero para ello se debe definir sus estimadores.

La función de autocovarianza estimada cuando no se conoce la media está dada por:

$$\hat{\gamma}_h = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n) (X_t - \bar{X}_n). \quad (1.63)$$

donde  $\bar{X}_n = \sum_{t=1}^n X_n$ , es el estimador de la media.

Mientras que la función de autocorrelación estimada está dada por:

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}. \quad (1.64)$$

---

<sup>1</sup>Para una mayor explicación sobre las transformaciones y en que circunstancias se utilizan ver González (2011), Shumway and Stoffer (2006).



Para ver con mas detalle las propiedades de los estimadores ver Madsen (2007) y Brockwell and Davis (2002).

El estimador de la función de autocorrelación parcial está dado por las ecuaciones de Durbin-Levinson estimadas:

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{k,k} \hat{\phi}_{k-1,k-j} \quad \text{para } k = 3, 4, \dots, \text{ y } j = 1, 2, \dots, k-1, \quad (1.65)$$

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j} \quad \text{donde } \hat{\phi}_{11} = \hat{\rho}_1. \quad (1.66)$$

Las gráficas de las funciones de autorrelación (ACF) y autocorrelación parcial (PACF) se comportarán de las siguiente manera dependiendo del tipo de modelo:

- Si se tiene un modelo  $AR(p)$ , la función de autocorrelación decae de manera exponencial o en forma de la función seno, mientras la función de autocorrelación parcial será cero para los rezagos mayores a  $p$ .
- Por otra parte, si se tiene un modelo  $MA(q)$  la función de autocorrelación será cero para para rezagos mayores a  $q$ , mientras que la función de autocorrelación parcial decaerá deforma exponencial o como función de seno.
- Para el modelo  $ARMA(p, q)$  ambas funciones decaerán deforma exponencial o como una función de seno, pero en la función de autocorrelación estará más marcada en el posible rezago, del cuál depende el orden de la parte de promedios móviles, de igual manera la función de autocorrelación parcial tendrá un valor más grande en el posible rezago del cuál depende el orden de la parte autorregresiva, de la combinación de ambos o de sus divisores, se propone el orden del modelo.
- En cuanto a los modelos estacionales  $AR(P)$ , la función de autocorrelación decae de manera exponencial o en forma de la función seno pero estarán más marcados los rezagos  $ks$ , donde  $s$  es el período estacional y  $k = 1, 2, \dots$ . Mientras la función de autocorrelación parcial será cero para los rezagos mayores a  $Ps$ .
- Para los modelos estacionales  $MA(Q)$ , la función de autocorrelación será cero para rezagos mayores a  $Qs$ , con  $s$  como el período estacional, mientras que la función de autocorrelación parcial decaerá deforma exponencial o como función de seno con un valor más grande para los rezagos  $ks$ , con  $k = 1, 2, \dots$
- Para los modelos estacionales  $ARMA(P, Q)$  ambas funciones decaerán de forma exponencial o como una función de seno y tendrán picos en rezagos  $ks$ , con  $s$  el período estacional y  $k = 1, 2, \dots$ . En la función de autocorrelación, del rezago  $ks$  con mayor magnitud se puede intuir el orden de la parte de promedios móviles estacional  $Q = k$ , de igual forma la función de autocorrelación parcial indicará el orden de la parte autorregresiva  $P = k$ , de la combinación de ambos o de sus divisores se propone el orden del modelo.
- Finalmente si se tiene un ruido blanco las autocorrelaciones muestrales serán próximas a cero. Si se supone normalidad e independencia, las autocorrelaciones asintóticamente se distribuyen  $\hat{\rho}_h \sim N(\rho_h, \frac{1}{n})$ , para una muestra grande,  $\rho_h = 0$ , sí  $|\rho_h| < \frac{2}{\sqrt{n}}$ . La prueba formal se encuentra en Shumway and Stoffer (2006).

## 1.12. Estimación de parámetros <sup>1</sup>

### 1.12.1. Momentos

Este método consiste en resolver las ecuaciones resultado de igualar los momentos muestrales con los momentos poblacionales para obtener los estimadores de parámetros desconocidos.

En particular para los modelos ARMA se hace uso de la relación de la función de autocorrelación con los parámetros  $\{\phi\}_{i=1}^p$ ,  $\{\theta\}_{i=1}^q$  y  $\sigma_\varepsilon^2$ . Básicamente la función de autocorrelación teórica se sustituye por su estimador muestral para resolver para los parámetros desconocidos.

Este método es adecuado para modelos  $AR(p)$ . Sin embargo, para modelos  $MA(q)$  y  $ARMA(p, q)$  ecuaciones devuelven un conjunto de ecuaciones no lineales, en general el estimador de momentos es difícil para un proceso que no sean un  $AR(p)$ . Aunque el estimador de momentos puede servir como un buen valor de inicio para otros métodos de estimación. Se tiene que para un proceso  $AR(p)$  el estimador por momentos es equivalente a resolver las ecuaciones de Yule-Walker estimadas, con respecto a  $\phi_1, \phi_2, \dots, \phi_p$ :

$$\begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(p) \end{pmatrix} = \begin{pmatrix} 1 & \hat{\rho}(1) & \dots & \hat{\rho}(p-1) \\ \hat{\rho}(1) & 1 & \dots & \hat{\rho}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}(p-1) & \hat{\rho}(p-2) & \dots & 1 \end{pmatrix} \begin{pmatrix} -\phi_1 \\ -\phi_2 \\ \vdots \\ -\phi_p \end{pmatrix}.$$

### 1.12.2. Mínimos cuadrados

Sea  $Y_t$  un modelo  $AR(p)$  entonces:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t, \quad t = p+1, p+2, \dots \quad (1.67)$$

donde  $\{\varepsilon_t\}$  es un ruido blanco, suponiendo que se tienen  $Y_1, \dots, Y_N$  observaciones. Sea:

$$\mathbf{X}_t' = (-Y_{t-1}, \dots, -Y_{t-p}), \quad t = p+1, p+2, \dots \quad \phi_t' = (\phi_1, \dots, \phi_p). \quad (1.68)$$

Entonces la ecuación 1.67 puede ser expresada como:

$$Y_t = \mathbf{X}_t' \phi + \varepsilon_t, \quad t = p+1, p+2, \dots \quad (1.69)$$

El modelo 1.69, se puede expresar para todos los datos de la serie, por lo que se define:

$$\mathbf{Y}' = (Y_{p+1}, \dots, Y_N), \quad (1.70)$$

$$\mathbf{X}' = (X_{p+1}, \dots, X_N), \quad (1.71)$$

$$\varepsilon' = (\varepsilon_{p+1}, \dots, \varepsilon_N), \quad (1.72)$$

Por lo tanto el modelo 1.67 se puede expresar de forma matricial para todos los datos observados como:

$$Y_t = \mathbf{X}\phi + \varepsilon_t. \quad (1.73)$$

El estimador por mínimos cuadrados se obtiene minimizando:

$$S(\phi) = (\mathbf{Y} - \mathbf{X}\phi)'(\mathbf{Y} - \mathbf{X}\phi). \quad (1.74)$$

<sup>1</sup>Basado en Shumway and Stoffer (2006) y Madsen (2007)

Donde  $S(\phi)$  es la suma de errores cuadrados y por definición los parámetros  $\phi$  que minimicen dicha suma son el estimador por mínimos cuadrados  $\hat{\phi}$ . Por las propiedades del estimador por mínimos cuadrados, se puede encontrar la solución del modelo lineal 1.73 para  $\phi$ , resolviendo:

$$(\mathbf{X}'\mathbf{X})\hat{\phi} = \mathbf{X}'\mathbf{Y}. \quad (1.75)$$

llamada ecuación normal, y como el estimador para varianza del ruido blanco  $\hat{\sigma}_\varepsilon^2$ :

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\hat{\theta})}{N-p} = \frac{\sum_{t=p+1}^N \varepsilon_t^2(\hat{\theta})}{N-p}. \quad (1.76)$$

Para ver a detalle las propiedades del estimador por mínimos cuadrados para los modelos ARIMA consultar Madsen (2007).

### 1.12.3. Máxima verosimilitud

Los estimadores de máxima verosimilitud son definidos como aquellos valores de los parámetros más probables dados los datos que se tengan. La función de verosimilitud  $L$  está definida por la función de distribución conjunta de los datos. En particular, éste método de estimación requiere suponer una distribución normal para la serie de ruido blanco  $\{\varepsilon_t\}$ , con media cero y varianza  $\sigma_\varepsilon^2$ .

Considerando un modelo  $ARMA(p, q)$  dado por 1.31:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (1.77)$$

Y sea:

$$\theta' = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) \mathbf{Y}'_t = (Y_t, Y_{t-1}, \dots, Y_1). \quad (1.78)$$

Donde  $\mathbf{Y}'_t$  son todas las observaciones hasta el tiempo  $t$ . Como se mencionó anteriormente la función de verosimilitud está definida por la función de distribución conjunta para todos las observaciones dados los valores de  $\theta$  y  $\sigma_\varepsilon^2$

$$\begin{aligned} L(\mathbf{Y}_N; \theta, \sigma_\varepsilon^2) &= f(\mathbf{Y}_N | \theta, \sigma_\varepsilon^2) \\ &= f(Y_N | \mathbf{Y}_{N-1}, \theta, \sigma_\varepsilon^2) f(\mathbf{Y}_{N-1} | \theta, \sigma_\varepsilon^2) \\ &= \left( \prod_{t=p+1}^N f(Y_t | \mathbf{Y}_{t-1}, \theta, \sigma_\varepsilon^2) \right) f(\mathbf{Y}_p | \theta, \sigma_\varepsilon^2). \end{aligned} \quad (1.79)$$

Esta es una formula general de la función de verosimilitud para series de tiempo. Suponiendo que se tiene toda la historia de los datos, se calcula la predicción a una paso como:

$$\hat{Y}_{t|t-1}(\theta) = \mathbb{E}[Y_t | \mathbf{Y}_{t-1}, \theta] = \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{i=1}^q \theta_i(\theta). \quad (1.80)$$

la cual, es la esperanza de  $(Y_t | \mathbf{Y}_{t-1}, \theta, \sigma_\varepsilon^2)$ . Así el error de predicción está dado por:

$$\varepsilon_t(\theta) = Y_t - \hat{Y}_{t|t-1}(\theta). \quad (1.81)$$

con varianza  $\sigma_\varepsilon^2$ . Dado que  $\varepsilon_t$  se distribuye normal, se tiene:

$$\begin{aligned} f(Y_t | \mathbf{Y}_{t-1}, \theta, \sigma_\varepsilon^2) &= \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left(-\frac{(Y_t | \mathbf{Y}_{t-1}(\theta) - \mathbb{E}[Y_t | \mathbf{Y}_{t-1}, \theta])^2}{2\sigma_\varepsilon^2}\right) \\ &= \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left(-\frac{\varepsilon_t^2(\theta)}{2\sigma_\varepsilon^2}\right). \end{aligned} \quad (1.82)$$

Se puede expresar la función de verosimilitud condicional sobre  $\mathbf{Y}_p$  como:

$$\begin{aligned} L(\mathbf{Y}_N; \theta, \sigma_\varepsilon^2) &= \prod_{t=p+1}^N f(Y_t | \mathbf{Y}_{t-1}, \theta, \sigma_\varepsilon^2), \\ &= (\sigma_\varepsilon^2 2\pi)^{-\frac{N-p}{2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=p+1}^N \varepsilon_t^2(\theta)\right). \end{aligned} \quad (1.83)$$

Tomando algoritmos se obtiene:

$$\log(L(\mathbf{Y}_N; \theta, \sigma_\varepsilon^2)) = -\frac{N-p}{2} \log(\sigma_\varepsilon^2) - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=p+1}^N \varepsilon_t^2(\theta) + c. \quad (1.84)$$

Donde  $c$  es una constante. Para encontrar el mínimo se aplica a la ecuación 1.84 una diferencia con respecto a  $\sigma_\varepsilon^2$ , se iguala a cero, dando:

$$\frac{\partial \log(L)}{\partial \sigma_\varepsilon^2} = -\frac{(N-p)}{2} \frac{1}{\sigma_\varepsilon^2} - \frac{1}{2\sigma_\varepsilon^4} \sum_{t=p+1}^N \varepsilon_t^2(\theta) = 0. \quad (1.85)$$

Entonces el mínimo obtenido es:

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=p+1}^N \varepsilon_t^2(\theta)}{N-p}, \quad (1.86)$$

Sustituyendo 1.86 en 1.84 se tiene:

$$\log(L(\mathbf{Y}_N; \theta, \hat{\sigma}_\varepsilon^2)) = -\frac{N-p}{2} \log\left(\sum_{t=p+1}^N \varepsilon_t^2(\theta)\right) + c. \quad (1.87)$$

Lo cual muestra que el estimador por máxima verosimilitud para  $\theta$  se obtiene minimizando:

$$S(\theta) = \sum_{t=p+1}^N \varepsilon_t^2(\theta). \quad (1.88)$$

Donde el el estimador por máxima verosimilitud para  $\hat{\sigma}_\varepsilon^2$  es:

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\hat{\theta})}{N-p}. \quad (1.89)$$

Para ver las propiedades de los estimadores por máxima verosimilitud y su generalización consultar Madsen (2007). Cuando R estima el modelo ARIMA, utiliza estimación de máxima verosimilitud (MLE). Esta técnica encuentra los valores de los parámetros que maximizan la probabilidad de obtener los datos que hemos observado.

Para los modelos ARIMA, MLE es muy similar a las estimaciones de mínimos cuadrados que se obtendrían al minimizar  $\sum_{t=1}^T e_t^2$ .

### 1.13. Diagnóstico y selección del modelo

El diagnóstico de un modelo o la prueba de bondad de ajuste, dice hasta que punto un modelo puede explicar las asociaciones en los datos y con ello el comportamiento de la serie, básicamente consiste en comprobar si las hipótesis realizadas sobre los residuos se cumplen. Lo cual lleva a determinar, si el modelo estimado es estadísticamente adecuado.

Siguiendo la metodología Box y Jenkins, después de estimar los parámetros del modelo ARIMA, se continua el diagnóstico del modelo. El resultado del diagnóstico está vinculado con el proceso de identificación del modelo, pues si el modelo es inadecuado, es necesario volver al paso de identificación y proponer otro modelo.

#### 1.13.1. Criterio de selección de modelos

Si resultado de las pruebas de bondad de ajuste se obtienen varios modelos, se pueden optar por utilizar los siguientes criterios, para elegir el más adecuado.

Medidas estadísticas de la bondad de ajuste del modelo son:

**AIC** o *Akaike Information Criterion*, se define como:

$$AIC(M) = -2\log(L) + 2M.$$

donde L corresponde a la verosimilitud y M la cantidad de parámetros estimados del modelo. En particular para un modelo ARMA se tiene:

$$\log(L) = -\frac{n}{2} \log 2\pi\sigma_\epsilon^2 - \frac{S(\phi, \mu, \theta)}{\sigma_\epsilon^2}.$$

**AICc** o *Akaike Information Criterion Corrected*, se define como:

$$AICc = -2\ln L(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) + 2(p+q) \frac{n}{n-p-q-2}.$$

**BIC** o *Bayesian information Criterion*

$$BIC = -2\ln L(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) + (p+q) \ln n,$$

Se prefiere el menor valor del criterio, pues es mejor el ajuste.

#### 1.13.2. Prueba de correlación

**Prueba Box-Pierce**  $H_0: \{\hat{\epsilon}_t\}$  es una serie de ruido blanco. vs  $H_a: \{\hat{\epsilon}_t\}$  no es una serie de ruido blanco.

Estadístico de prueba

$$Q = n \sum_{k=1}^K \hat{\rho}_k^2$$

donde  $\hat{\rho}_k$  es la correlación muestral de la serie  $\{\hat{\epsilon}_t\}$ ,  $K$  es el máximo lag a considerar, en se consideran apropiado  $K = 10$  para datos sin efecto estacional y  $K = 2m$  para datos con efecto estacional, pues  $m$  es el periodo estacional.  $Q \sim \chi_{k-p-q}^2$ .

**Prueba Box-Ljung**  $H_0: \{\hat{\epsilon}_t\}$  es una serie de ruido blanco. vs  $H_a: \{\hat{\epsilon}_t\}$  no es una serie de ruido blanco.

Estadístico de prueba

$$\tilde{Q} = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-K}.$$

donde  $\hat{\rho}_k$  es la correlación muestral de la serie  $\{\hat{\epsilon}_t\}$ ,  $p+q < K \ll n$ ,  $\tilde{Q} \sim \chi_{k-p-q}^2$ .

- Sí  $p\text{-valor} \leq 0.05$ : Se puede rechazar la hipótesis nula asumiendo una probabilidad del 5% de cometer un error. Así que se puede asumir que los valores están mostrando dependencia el uno del otro.
- Si  $p\text{-valor} > 0.05$ : No tiene suficiente evidencia estadística para rechazar la hipótesis nula. Así que no se puede asumir que los valores son dependientes. Esto podría significar que los valores son dependientes de todos modos o puede significar que los valores son independientes.

En particular, no se puede probar la independencia de los valores de series de tiempo utilizando la prueba de Ljung-Box. Sólo puede probar la dependencia.

### 1.13.3. Prueba de normalidad

**Prueba Anderson-Darling para normalidad** Prueba no paramétrica basada en la función de distribución empírica para la hipótesis compuesta de normalidad.

$H_0$  : Los datos siguen una distribución normal. vs  $H_a$  : Los datos no siguen una distribución normal.

El estadístico de prueba es

$$A = -n - \left( \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln(p_{(i)}) + \ln(1-p_{(n-i+1)})] \right).$$

donde  $p_{(i)} = \Phi\left(\frac{[x_{(i)} - \bar{x}]}{s}\right)$ ,  $\Phi$  es la función de distribución acumulativa de una distribución normal,  $\bar{x}$  y  $s$  son la media y desviación muestral. Entonces el estadístico de bondad de ajuste de Anderson-Darling  $A$  mide el área entre la línea ajustada de la distribución normal, y la función de distribución empírica. El estadístico de Anderson-Darling es una distancia elevada al cuadrado que tiene mayor ponderación en las colas de la distribución.

Para determinar si los datos no siguen una distribución normal, se compara el p-valor con el nivel de significancia. Generalmente, se usa un nivel de significancia del 5% e indica que el riesgo de concluir que los datos no siguen una distribución normal, cuando en realidad los datos siguen una distribución normal.

- Si  $p\text{-valor} \leq 0.05$ : La decisión es rechazar la hipótesis nula asumiendo una probabilidad del 5% de cometer un error y concluir que sus datos no siguen una distribución normal.
- Si  $p\text{-valor} > 0.05$ : La decisión es que no se puede rechazar la hipótesis nula, es decir, no se tiene suficiente evidencia para concluir que los datos no siguen una distribución normal.

En el ajuste de series de tiempo para datos reales, los residuales pueden ser un ruido blanco y no distribuirse normal, pese a que la normalidad un requisito en la estimación de los parámetros por máxima verosimilitud y sus intervalos de confianza, en estos casos tener una gran la cantidad de datos es esencial, pues aunque no se cumpla la normalidad se puede calcular los estimadores como normales de acuerdo a la ley de los grandes números.

## Capítulo 2

# Descripción de la población de ayudantes de la UNAM

En este capítulo se presenta una descripción de los ayudantes de profesor basado en algunas estadísticas y el marco legal. Dentro del personal académico de la UNAM, la figura de ayudante de profesor o de investigador surge ante la necesidad de preparar a futuros académicos para el desempeño de funciones docentes o de investigación.

### 2.1. Estatuto del Personal Académico

El Estatuto del Personal Académico (EPA) forma parte del marco jurídico de la relación laboral de la universidad con el personal académico. Es el instrumento que rige y preserva la vida académica. En el EPA se señalan tanto lineamientos laborales como académicos como son los procedimientos de ingreso, promoción, permanencia, evaluación académica, entre otros. En particular los artículos que se tomarán como referencia para la figura de ayudante de profesor son el 4 y del 20 al 28 del EPA. Con los cuales se contrastará la información obtenida.

Un nombramiento es la forma en la cual se contrata al personal, que puede ser interino, definitivo o por contrato de prestación de servicios, dependiendo de la figura académica en que se emplee. En el caso particular del nombramiento de ayudante se clasifica en ayudantes de profesor o de investigador con categorías A, B o C.

Los ayudantes de profesor o investigador de medio tiempo y de tiempo completo, laboran 20 y 40 horas semanales respectivamente, pueden ocupar los niveles A, B o C; los ayudantes por horas que siempre son de profesor: solo pueden ocupar los niveles A o B y auxiliar en una materia o curso específico sin exceder de doce horas por semana



Otra característica del nombramiento de ayudante, es que se otorga por un plazo no mayor de un año y puede renovarse hasta por cuatro veces, es decir, la duración máxima del nombramiento de ayudante es de cinco años. La excepción es que existe la posibilidad de que el consejo técnico, apruebe una prórroga de los nombramientos de ayudante por un número mayor de años.

Se debe tener en consideración que una persona puede poseer uno o más nombramientos sin exceder las 48 horas semanales.<sup>1</sup> Por ejemplo, una persona x, tiene un nombramiento de profesor de asignatura por 20 horas y otro de ayudante de profesor por 5 en la facultad a, pero también cuenta con otro nombramiento de ayudante de profesor por 8 horas en la facultad b, es decir una persona puede tener más de un nombramiento siempre y cuando cumpla con el EPA. Cabe señalar que en el momento de contar el número de ayudantes solo se consideró a la persona y no los nombramientos que posee.

## 2.2. Datos de los ayudantes de la UNAM

En este trabajo se trató a todos los ayudantes como una sola población sin importar su categoría, resultado de examinar el comportamiento de los ayudantes por horas, A y B, el cual es similar, mientras que la población ayudantes de investigador es muy pequeña y al incluirla no cambia los resultados.

Además si se consideran los requisitos para ascender de una categoría a otra, el porcentaje de créditos y la experiencia adquirida mientras se permanezca como ayudante, no son discriminatorios por lo que cualquiera puede alcanzar la siguiente categoría.

Para tener una perspectiva general de la evolución de la población de ayudantes se darán datos estadísticos para los años 1990, 2001, 2009 y 2014 tomando como referencia la quincena 12. Se hace referencia a una quincena específica pues los datos tienen variaciones quincenales y en particular la elección se realizó en base a la quincena con menor cantidad de datos omitidos.

Analizando la tabla 2.1, vemos que la población de ayudantes en 1999 estaba conformada por 3,220 personas con algún nombramiento de ayudante, el 23.8% poseía más de un nombramiento, para cubrir un total de 4,167 nombramientos. Para 2001 se dio una disminución en el número de ayudantes, al ser 2,883 personas con al menos un nombramiento de ayudante relacionadas con 3,771 nombramientos, de dicha población el 24.7% tiene más de un nombramiento. En 2009 se tenían 5,592 nombramientos adscritos a 4,158 personas con algún nombramiento de ayudante, donde el 26.5% de ayudantes poseen más de un nombramiento. Para finalizar en 2014 había 4,763 personas con algún nombramiento de ayudante adscritos a 6,424 nombramientos, y el 27.41% tenía más de un nombramiento.

En la tabla 2.2 se observan los nombramientos en las diferentes figuras para las personas que poseen al menos un nombramiento de ayudante para los años seleccionados. Las figuras académicas con más adscripción en la población de estudio son ayudantes y profesores de asignatura.

La población de ayudantes está conformada principalmente por hombres, dado que por cada 100 mujeres hay 115, 116, 112, y 116 hombres para los años de estudio. En cuanto a la edad en 1999, la edad con mayor

---

<sup>1</sup> Ver EPA artículo 6 fracción VIII

**Tabla 2.1:** Número de nombramientos por ayudante

Nombramientos	1999	2001	2009	2014
1	2,454	2,183	3,056	3,457
2	620	551	836	1,009
3	119	121	214	251
4	20	18	39	36
5	6	9	12	8
6	1	1	1	2
Personas	3,220	2,883	4,158	4,763

**Tabla 2.2:** Nombramientos por figura

Figura Académica	1999	2001	2009	2014
Ayudante de profesor	3,297	2,940	4,258	4,882
Investigador	3	1	0	0
Profesor de asignatura	786	749	1,271	1,482
Profesor de carrera	7	9	9	11
Técnico académico	74	72	54	49
Nombramientos	4,167	3,771	5,592	6,424

frecuencia es de 27 y 28 años para las mujeres y 26 para hombres, en 2001 es de 27 años para las mujeres y 28 años para hombres, mientras en 2009 fueron 24 y 26 años para mujeres y hombres respectivamente, en 2014 fueron 24 años para ambos sexos.

Al examinar las pirámides de la población de ayudantes [2.1](#), es claro que la mayor parte de los ayudantes son jóvenes entre los 20 y 30 años, no obstante para 1999, 2001, 2009 y 2014 la edad promedio fue de 32.99, 34.58, 34.49, 34.87 años respectivamente, ya sea por un aumento en incorporaciones de personas con mayor edad o individuos que se mantienen en su nombramiento.

La proporción de ayudantes ubicada entre los 18 y 19 años es pequeña resultado del requisito de créditos así como de la edad promedio en la cual un individuo inicia su instrucción universitaria.

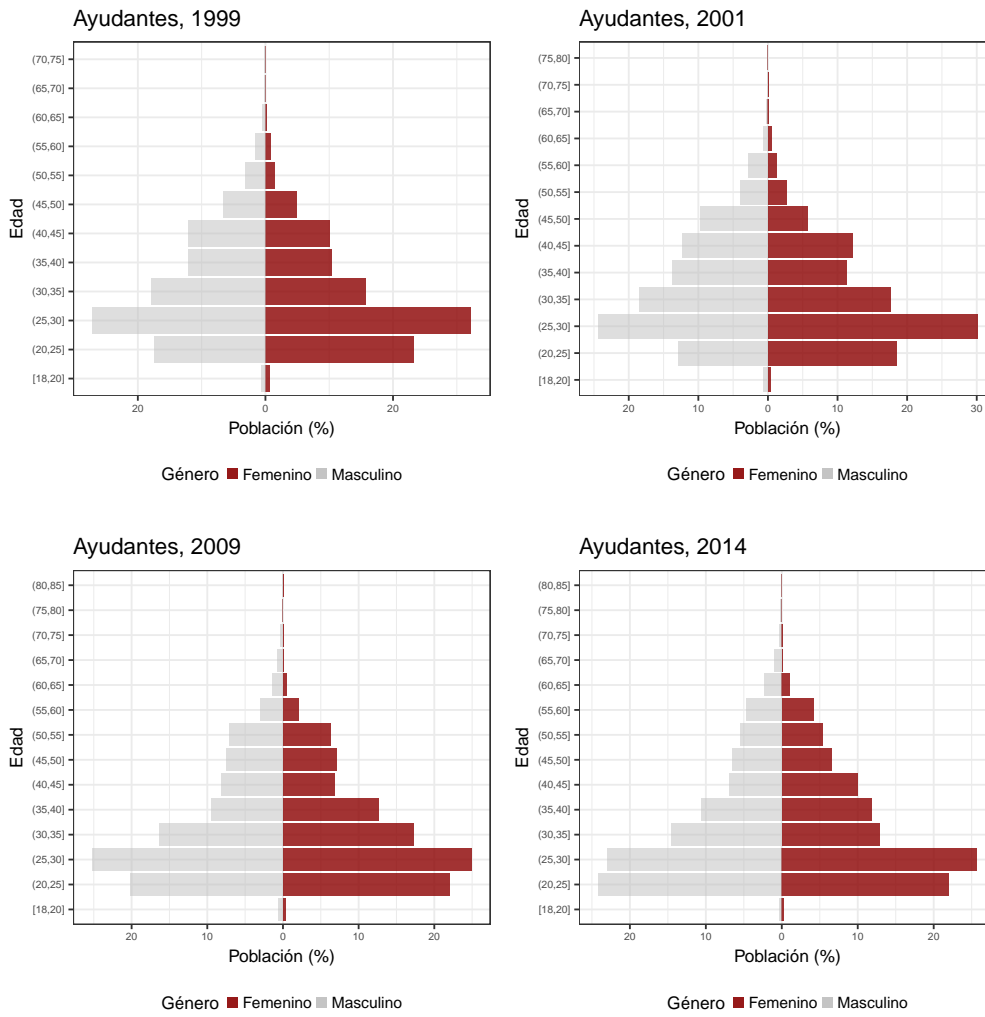
Analizando los cambios en la proporción de las pirámides poblacionales en 2009 se presenta una proporción muy similar entre los grupos de edad de 40 a 45, 45 a 50 y 50 a 55, mientras que las restantes tienen

una forma más regular. Sin embargo la edad máxima de los ayudantes sigue aumentando.

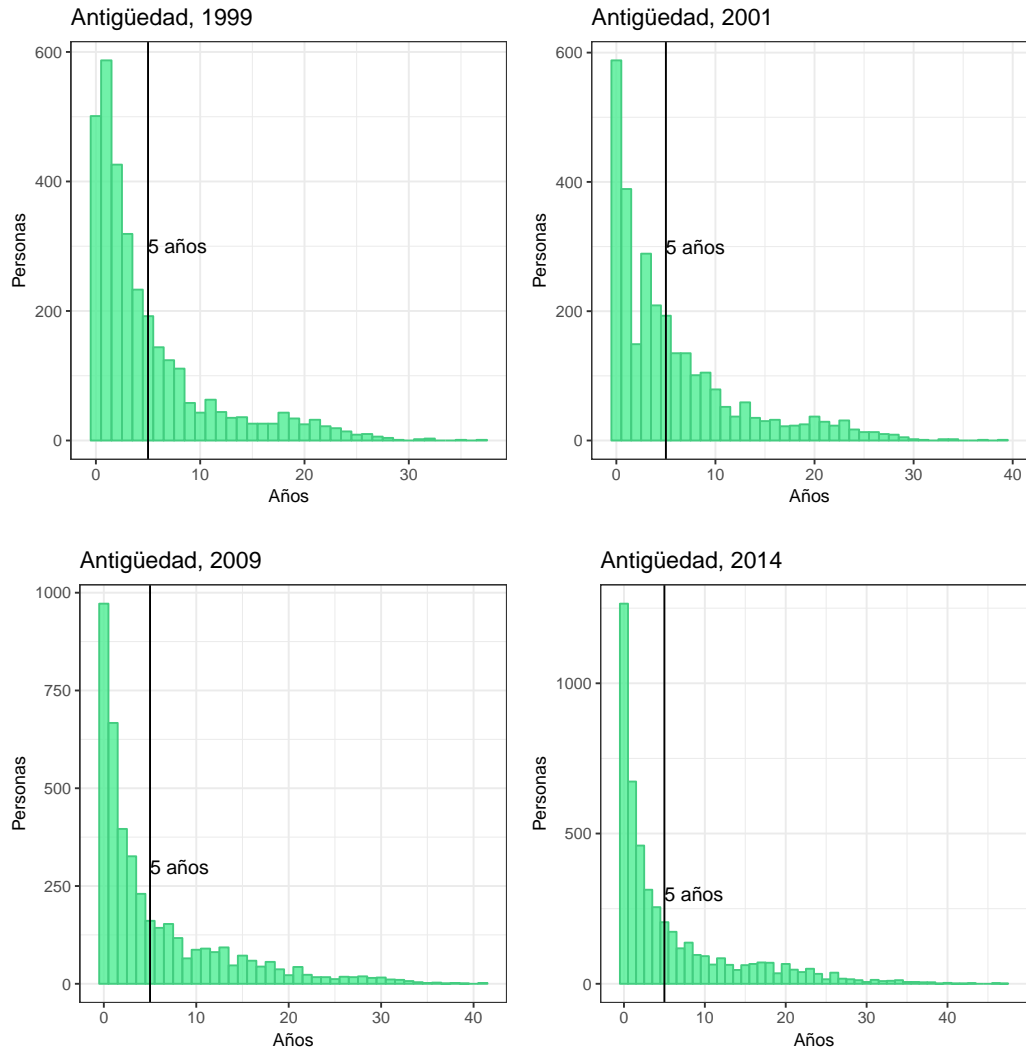
Analizando la antigüedad la tabla 2.2 se percibe que al menos mitad de la población tiene entre cero y cuatro años de antigüedad en el periodo de estudio, sin embargo la antigüedad promedio es 5.01, 5.81, 5.66 y 5.19 años, arriba del límite para permanecer en un nombramiento.

Para obtener estadísticas sobre las horas contratadas, hay que tener en consideración que una persona puede tener varios nombramientos con diversas figuras académicas, siempre que cumpla con los requisitos y necesidades de la institución, es por ello que se sólo se tomaron en consideración las horas de servicio por nombramiento de ayudante, ver figura 2.3

En 1999 los ayudantes registraron un total de 36,306 horas para 3,297 nombramientos de ayudante, lo cual equivaldría a 11.01 horas por nombramiento. Para 2001 fueron 31,811 horas totales en 2,940 nombramientos obteniendo 10.82 horas por nombramiento. En 2009 fueron 38,994 horas totales para 4,258 nombramientos obteniendo 9.157 horas por nombramiento. Mientras para 2014 fueron 43,092 horas totales en 4,882 nombramientos obteniendo 10.12 horas por nombramiento.



**Figura 2.1:** Pirámides Poblacionales 1999, 2001, 2009 y 2014



**Figura 2.2:** Antigüedad 1999, 2001, 2009 y 2014

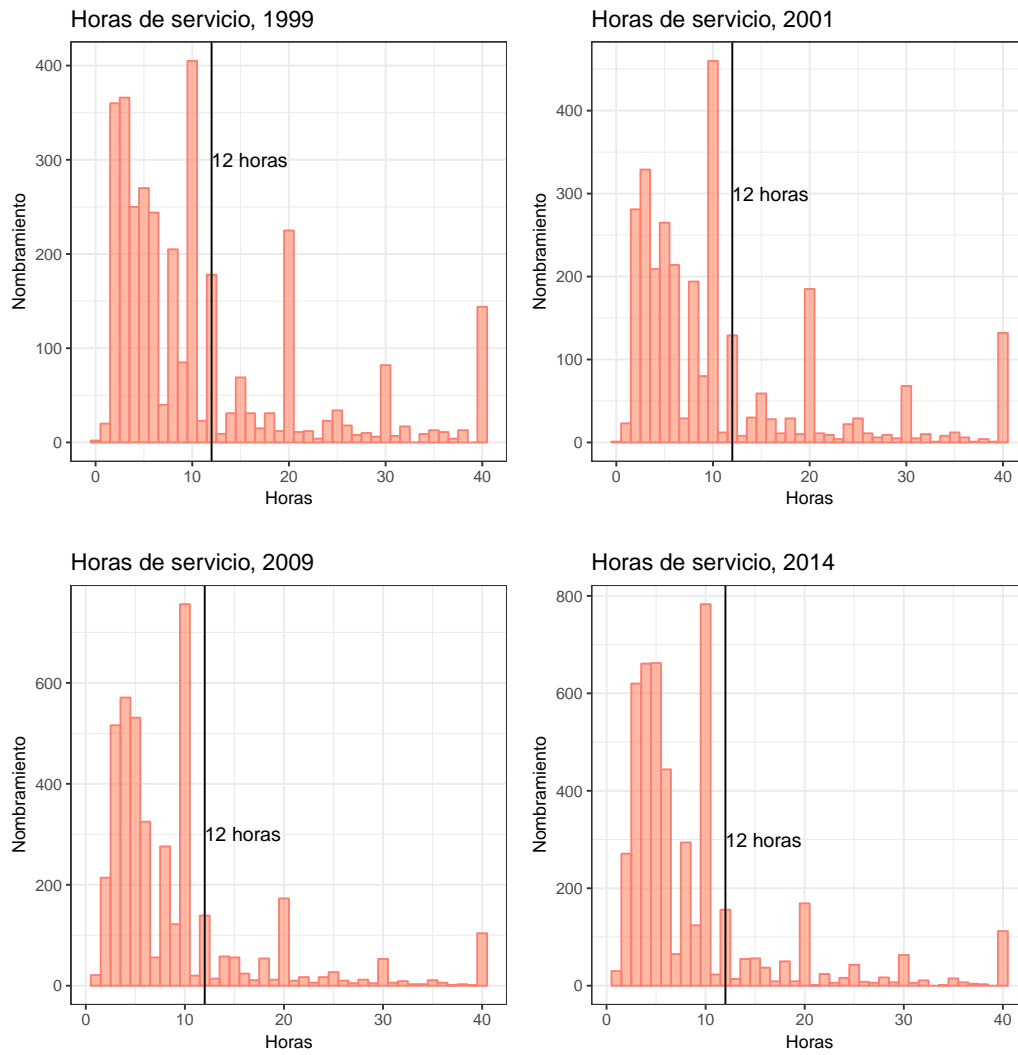


Figura 2.3: Número de horas contratadas por nombramiento de ayudante en 1999, 2001, 2009 y 2014



## Capítulo 3

# Evolución de la población de ayudantes de la UNAM de 1999 a 2014

En este capítulo se examina a la población de ayudantes mediante el análisis de series de tiempo. Para este estudio se hará uso del programa estadístico R, donde se realizaron las gráficas, estimaciones proyecciones y pruebas de bondad de ajuste. En particular de las paqueterías *seasonal*, *tseries*, *forecast*, *nortest*, *ggplot2*. Para dudas con respecto al uso del programa en series de tiempo consultar Hyndman and Athanasopoulos (2014), Cowpertwait and Metcalfe (2009), Cryer and Chan (2008) y Shumway and Stoffer (2006).

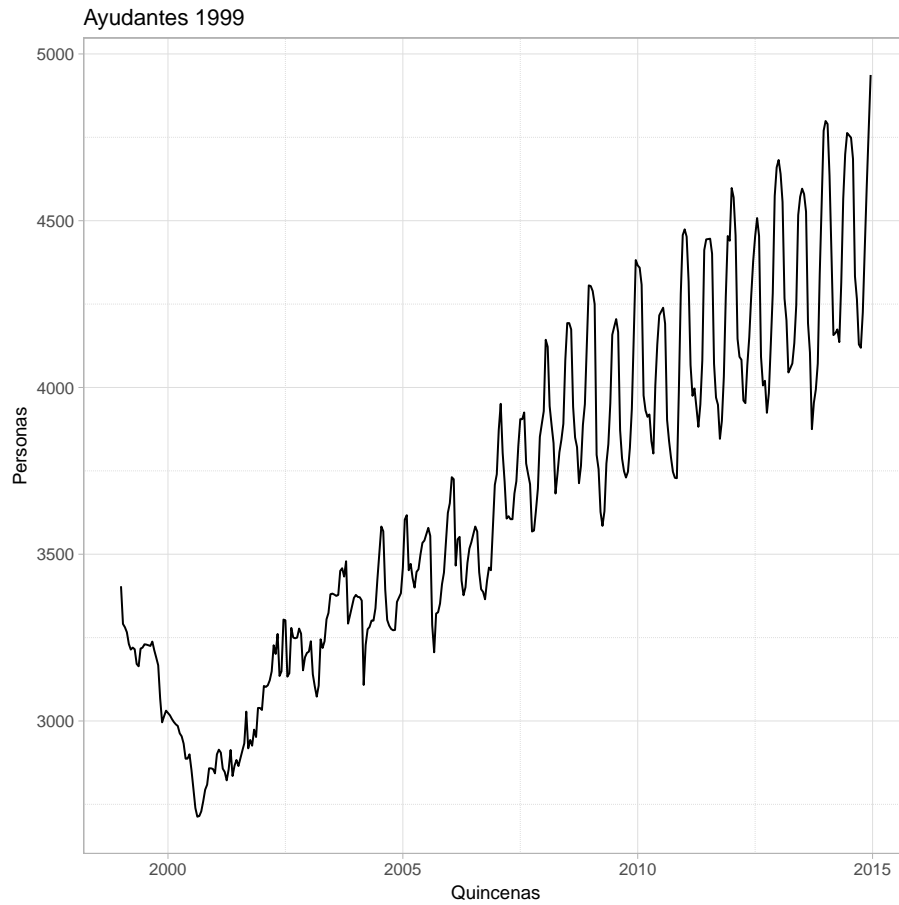
La información con la que se cuenta es el número de ayudantes por quincena a partir de 1999 hasta la última quincena de 2014. Se puede ver en la figura 3.1. El comportamiento de la población de ayudantes no es estacionario, pues crece a partir de agosto de 2000, a una tasa del 2.3% anual, es decir tiene tendencia, además al inicio hay un comportamiento irregular cuyo contexto histórico es la ocurrencia de la huelga de 1999, no obstante al final se aprecia un marcado patrón estacional.

Además de la gráfica de la serie, otra herramienta para evidenciar la no estacionariedad de la serie, son las gráficas de las funciones de autocorrelación y autocorrelación parcial figura 3.2. La ACF presenta un decaimiento lento indicando la existencia de una tendencia, mientras que el decrecimiento ondulado y los picos marcados en los rezagos 2, 12, 25 en la PACF apuntan a la presencia de una componente estacional.

En la gráfica de la población de ayudantes figura 3.1 no hay evidencia de un cambio en la varianza, por lo que no se aplicará ninguna transformación de Box-Cox. Después de descartar una transformación de Box-Cox, se analizó la tendencia mediante modelos de regresión, el mejor ajuste fue una tendencia lineal, es decir el crecimiento en la población de ayudantes es lineal.

En consecuencia, se buscará modelar una componente estacional con una tendencia lineal, el método más sencillo es aplicar una diferencia combinada, es decir, emplear una diferencia lineal del tipo  $(1 - B)^d$ ,





**Figura 3.1:** Población de ayudantes 1999-2014.

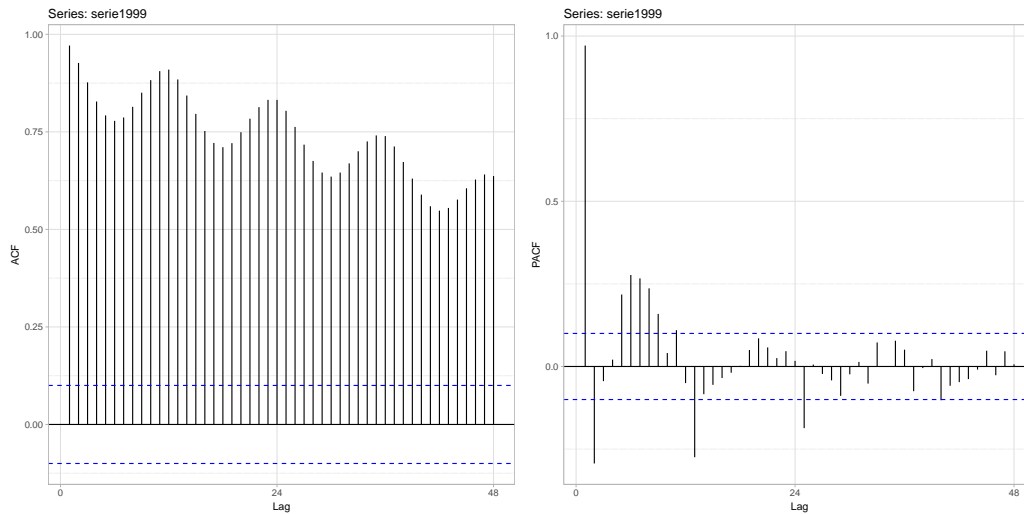
para quitar la tendencia y una diferencia estacional del tipo  $(1 - B^s)^d$  para remover el efecto estacional.<sup>1</sup>

Generalmente cuando se aplican diferencias estacionales se toma el periodo como la frecuencia anual de los datos, en este caso sería 24,<sup>2</sup> debido a que los datos son quincenales. Además sí se considera el operador diferencia como un filtro lineal para disminuir la varianza de los datos, la diferencia combinada con la diferencia estacional en el rezago 24,  $(1 - B)(1 - B^{24})$ , es la que reduce más la varianza de la serie como se puede apreciar en la tabla 3.1.

Sin embargo, a la serie que resulta de la operación diferencia  $W_t = (1 - B)(1 - B^{24})Y_t$ , al ajustarle un modelo ARMA los órdenes p,q, eran muy grandes y las gráficas de las funciones ACF y PACF figura 3.3,

<sup>1</sup>Aunque para quitar el efecto estacional se trató con diferentes enfoques funciones armónicas, filtros de promedios móviles, funciones nulas.

<sup>2</sup>aplicar una diferencia 24 es quitar un ciclo anual



**Figura 3.2:** ACF y PACF de la población de ayudantes apartir de 1999.

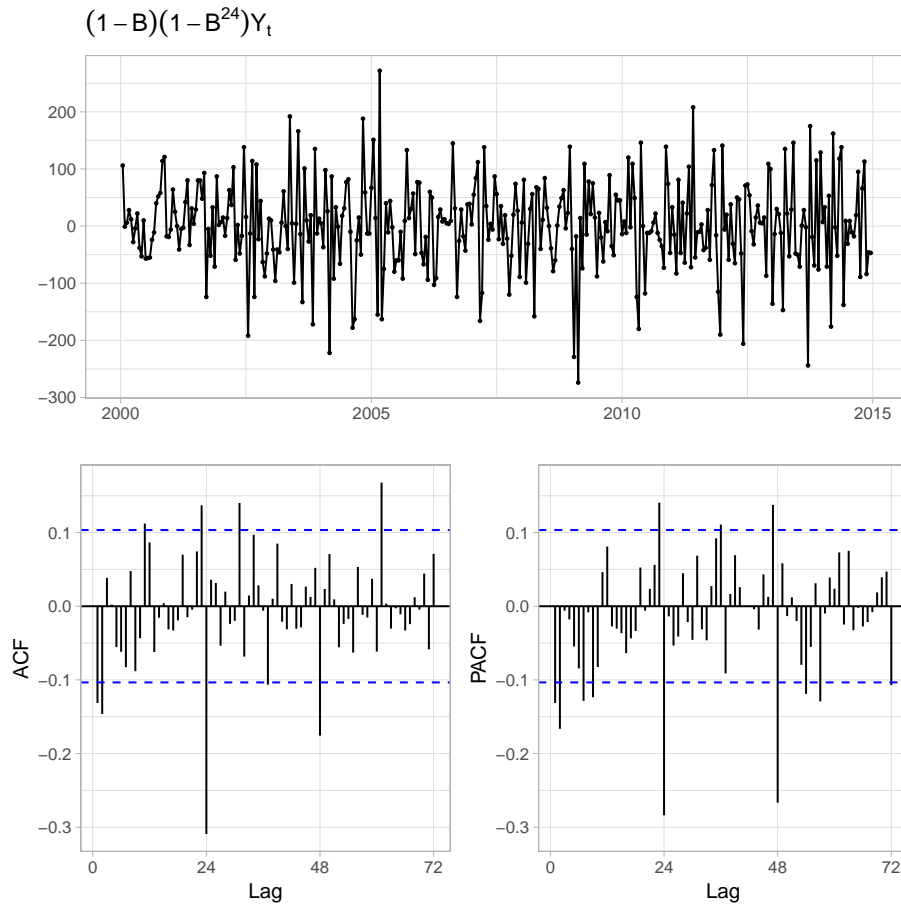
**Tabla 3.1:** Desviación estándar de la serie y sus respectivas diferencias.

Serie	$Y_t$	$(1-B)Y_t$	$(1-B^{12})Y_t$	$(1-B^{12})^2 Y_t$	$(1-B^{24})Y_t$	$(1-B)(1-B^{24})Y_t$	$(1-B)(1-B^{12})Y_t$
Desviación estándar	524.80	107.06	138.05	223.56	164.03	79.46	78.51

revelan la existencia de una dependencia estacional, por lo que era necesario ajustar modelos con componente estacional. Sin embargo, incluso si se modela con modelos SARIMA hay dependencias que no se pueden explicar y la bondad de ajuste de estos modelos es pobre, no cumplen con la no correlación, ni con la normalidad.

Es evidente que la serie posee una variación estacional, pero el periodo propuesto no es el que corresponde con la frecuencia de la serie, para ello se hará uso del periodograma, el cual descompone la estructura de dependencias de una serie en términos de su frecuencia. Muestra en que frecuencias o periodos la serie tiene un ciclo, generalmente se utiliza cuando la serie posee ciclos o variaciones estacionales y no se conoce el periodo, aunque se puede modelar una serie de tiempo a través del análisis en términos de las frecuencias siguiendo un enfoque denominado análisis espectral. Para una explicación formal del periodograma consultar Brockwell and Davis (2002), Cryer and Chan (2008), Shumway and Stoffer (2006).

El periodograma de la serie sin tendencia figura 3.4, indica el mayor valor en 2 ciclos por cada año de observación correspondiente a la frecuencia  $\frac{1}{12}$ , es decir cada 12 quincenas hay un ciclo equivalente a un ciclo por semestre, seguido por 8, 6 y 4 ciclos por año, equivalente a las frecuencias  $\frac{1}{3}$ ,  $\frac{1}{4}$  y  $\frac{1}{6}$  respectivamente. Pero se tomará la frecuencia del periodo estacional como 12, por ser el mayor valor.

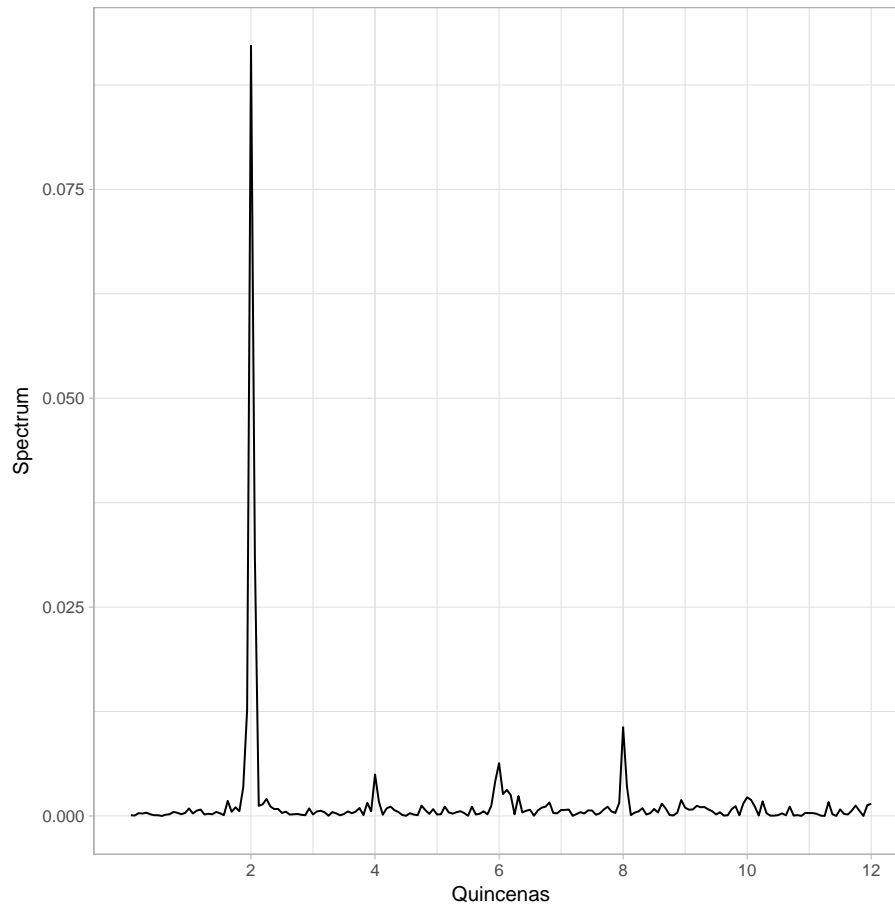


**Figura 3.3:** Gráfica de  $(1-B)(1-B^{24})Y_t$ .

Otra herramienta para analizar el comportamiento estacional de la población de ayudantes es la gráfica de la distribución por quincenas, figura 3.5, donde se representa la frecuencia quincenal de la población de ayudantes durante el periodo observado, las líneas horizontales simbolizan el número promedio de ayudantes por quincena, en consecuencia se detecta como el promedio de ayudantes aumenta en las quincenas 1 a 3, 13 a 15 y 23 a 24, coincidiendo con el inicio y final de los semestres.

Después conocer que el periodo estacional es 12 y la tendencia es lineal, se continuarán aplicando transformaciones para obtener una serie estacionaria y ajustarle un modelo ARMA, algunos enfoques adoptados fueron:

- Se buscó quitar la tendencia y la componente estacional mediante filtros de promedios móviles pero la serie de residuales mantenía la estructura de dependencia de la serie original.
- Otro enfoque fue modelar la serie como una regresión del tiempo y funciones de seno y coseno en



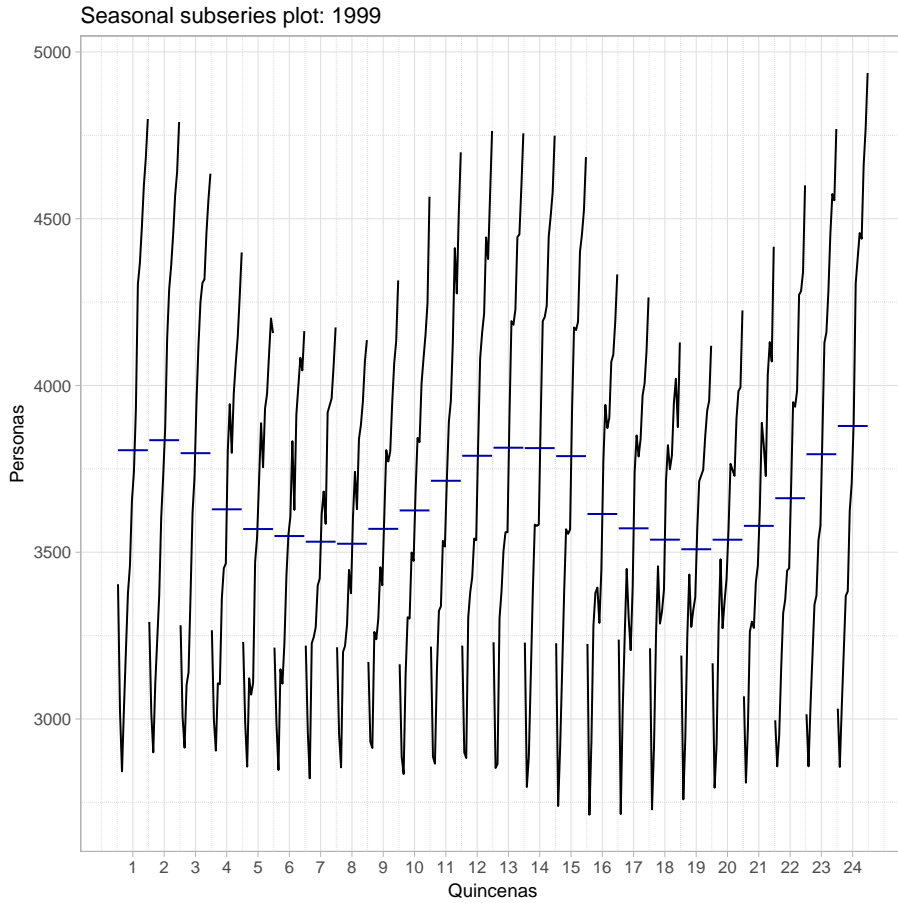
**Figura 3.4:** Periodograma.

las frecuencias más importantes mostradas en el periodograma, pero la serie de residuales todavía mantenía una estructura de correlación considerable.

- También se trató de modelar como función del tiempo y variables mudas, aunque la reducción de la estructura de dependencia, los ajustes obtenidos fallaban en la bondad de ajuste.

Al final, se aplicó una diferencia combinada con diferencia estacional en el rezago 12,  $(1 - B)(1 - B^{12})$ , resultando una serie con menor dependencia y esto se puede observar en las gráficas de las funciones ACF y PACF, figura 3.6.

Para modelar la serie diferenciada las gráficas de las funciones ACF y PACF figura 3.6 muestran que además de ajustar un modelo ARMA es necesario considerar modelos estacionales, pues la ACF muestra valores significativos en los rezagos 1, 3, 12, por lo que se propone modelos con la parte de promedios móviles de orden  $MA(1)$  y  $MA_{12}(1)$ . Mientras que en la PACF se observan picos en los rezagos 1, 2, 12,

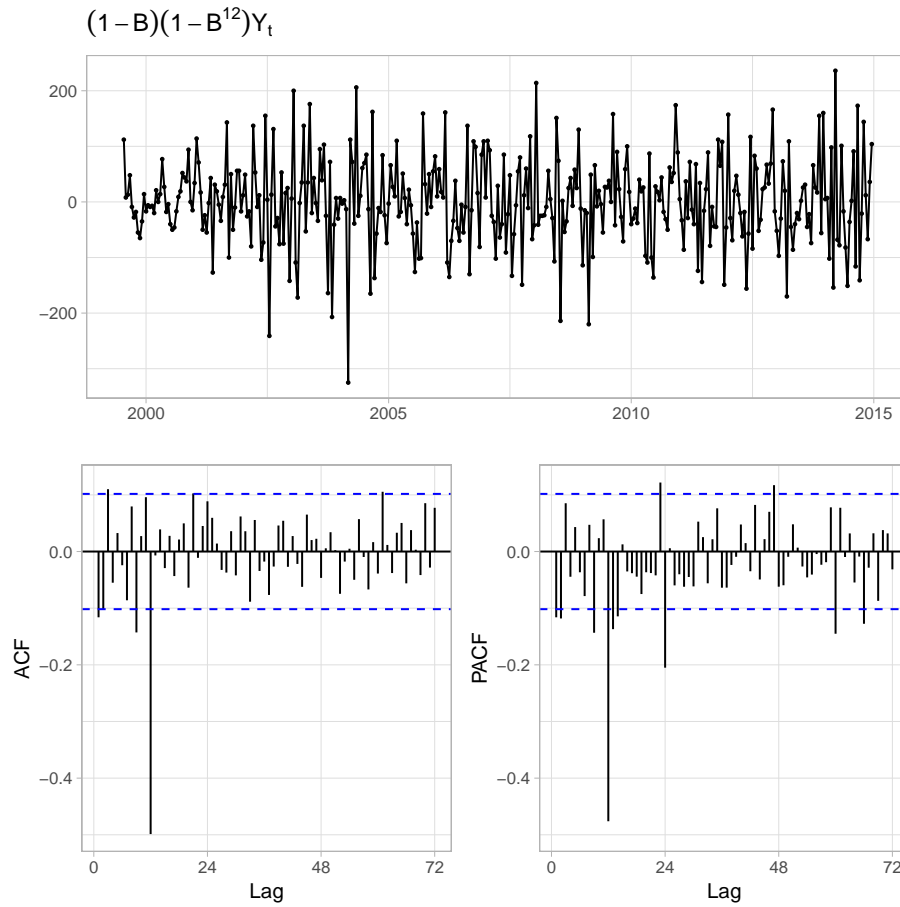


**Figura 3.5:** Gráfica de la subserie con ciclo estacional.

23, 24, es decir, se puede proponer modelos con la parte autorregresiva de orden  $AR(1)$ ,  $AR(2)$ ,  $AR_{12}(1)$  y  $AR_{12}(2)$ .

Después de sugerir y probar la bondad de ajuste de diferentes ajustes, los modelos a considerar están en la tabla 3.2. Donde la última columna de dicha tabla es p-valor de la prueba Ljung-Box sobre los residuales del modelo, para los primeros 48 rezagos. Cabe señalar que los modelos propuestos no lograron captar toda la estructura de correlación entre los datos, tienen fallos principalmente con las dependencias de cada 12 quincenas y cada 24. Sin embargo se buscará el modelo que capte mejor las relaciones de dependencia entre las observaciones.

Hay diferentes criterios para elegir entre diferentes modelos, si se basa en el AICc seleccionaría el modelo  $ARIMA(1,1,1)(2,1,1)[12]$ , además si se considera el mayor p-valor para la prueba de Ljung-Box al lag 48 también se elegiría. Sin embargo, si uno se basa en el criterio del BIC se elegiría el modelo



**Figura 3.6:** Serie con la diferencia  $(1-B)(1-B^{12})Y_t$

**Tabla 3.2:** Modelos para la serie de ayudantes a partir de 1999.

Modelo	Orden	AICc	BIC	p-valor LB
1	ARIMA(1,1,1)(2,1,1)[12]	4155.1375	4178.4040	0.0864
2	ARIMA(1,1,1)(0,1,2)[12]	4156.5473	4175.9639	0.0693
3	ARIMA(1,1,1)(0,1,1)[12]	4155.1792	4170.7347	0.0645
4	ARIMA(1,1,1)(1,1,2)[12]	4157.1123	4180.3788	0.0638

ARIMA(1,1,1)(0,1,1)[12]. Tomando como criterio de elección p-valor para la prueba de Ljung-Box, se elige el modelo en el cual a un cierto número de rezagos, no se rechace que las autocorrelaciones conjuntamente sean cero.

Por lo anterior se elige el modelo ARIMA(1,1,1)(2,1,1)[12]. En la tabla 3.3 se muestra los parámetros estimados y sus respectivos intervalos de confianza. *Revisando la significancia de dichos parámetros se observa que  $\Phi_{1_{12}}$  y  $\Phi_{2_{12}}$  pueden ser ceros y por lo tanto son poco significativos. Dando como resultado que se prefiera el modelo ARIMA(1,1,1)(0,1,1)[12]*

**Tabla 3.3:** Parámetros estimados y su intervalo de confianza para el modelo ARIMA(2,1,1)(0,1,1)[12], serie de ayudantes a partir de 1999.

Parámetro	Estimador	2.5 %	97.5 %
$\phi_1$	0.7944	0.67	0.92
$\theta_1$	-0.9217	-1.00	-0.84
$\Phi_{1_{12}}$	0.0451	-0.15	0.24
$\Phi_{2_{12}}$	0.0451	-0.01	0.29
$\Theta_{1_{12}}$	-0.7111	-0.87	-0.56

Así, la población de ayudantes sigue un comportamiento de la forma:

$$(1 - B)(1 - B^{12})(1 - \phi_1)Y_t = (1 - \theta_1)(1 - \Theta_{1_{12}})\varepsilon_t$$

donde los parámetros  $\phi_1$ ,  $\theta_1$ ,  $\Theta_{1_{12}}$  con sus respectivos intervalos de confianza están dados en la tabla 3.4. Se debe de examinar la bondad de ajuste, es decir, que tan bueno es el ajuste.

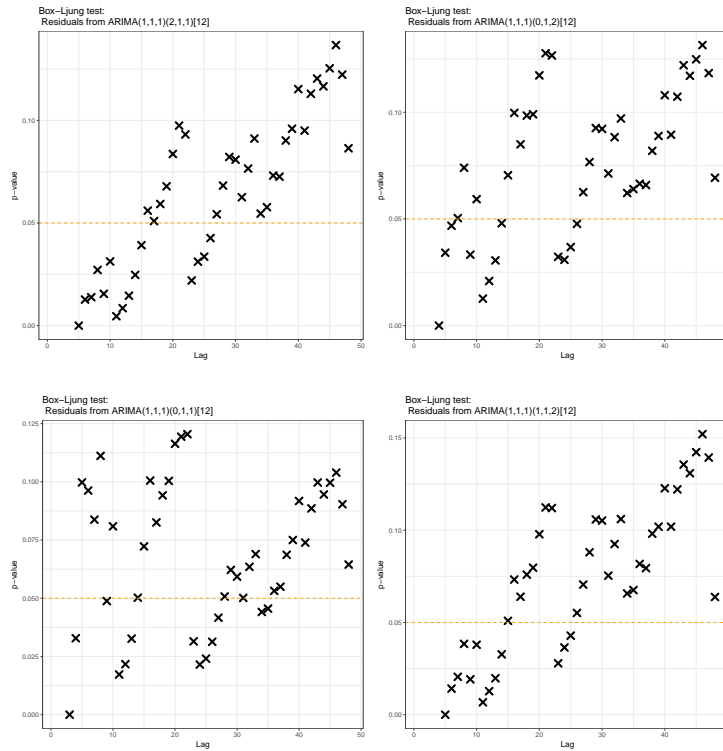
**Tabla 3.4:** Parámetros estimados y su intervalo de confianza para el modelo ARIMA(1,1,1)(0,1,1)[12], serie de ayudantes a partir de 1999.

Parámetro	Estimador	2.5 %	97.5 %
$\phi_1$	0.81	0.69	0.93
$\theta_1$	-0.93	-1.01	-0.85
$\Theta_{1_{12}}$	-0.64	-0.72	-0.56

De las gráficas del ACF y PACF para los residuales y la gráfica del p-valor de la prueba Ljung-Box se muestra en la figura 3.8, se concluye que los residuales mantienen una dependencia entre las observaciones separadas por 11 y 23 periodos, lo cual implica que los residuales no son un ruido blanco, además al realizar la prueba Anderson-Darling para normalidad y examinando el histograma de los residuales dieron como resultado que no son normales, es decir, el modelo propuesto no posee una buena bondad de ajuste.

Anderson-Darling normality test

data: fit03\$residuals



**Figura 3.7:** P-valores de la prueba Ljung-Box para los residuales.

A = 2.4324, p-value = 3.691e-06

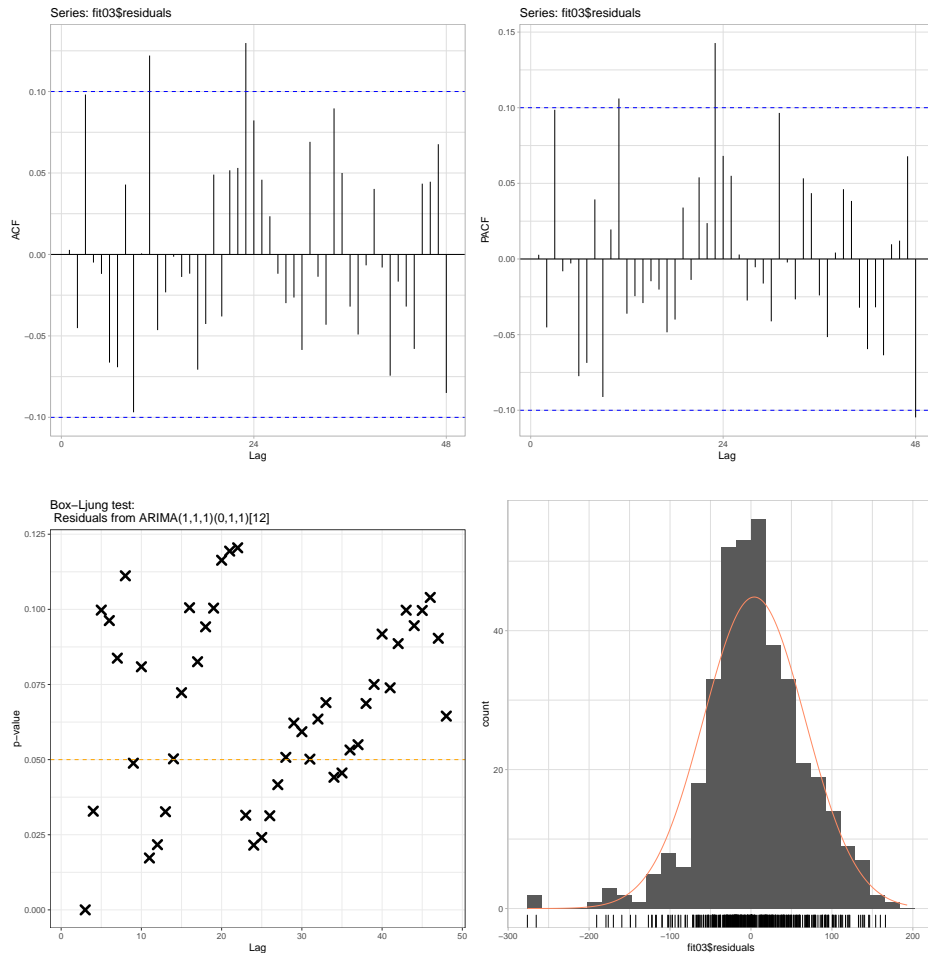
Lilliefors (Kolmogorov-Smirnov) normality test

data: fit03\$residuals

D = 0.06689, p-value = 0.0002932

Sin embargo, debido a que uno de los propósitos es poder estimar la población futura, se realizará la estimación de los pronósticos, los cuales dan una idea del comportamiento de los ayudantes, pero no serán tan confiables. En las figuras 3.9 y 3.10, se puede observar los pronósticos para un año y cinco años respectivamente. Si se plantea un horizonte de pronóstico más grande la previsión terminará siendo la media y el intervalo de confianza seguirá creciendo como lo hace para el caso de cinco años.



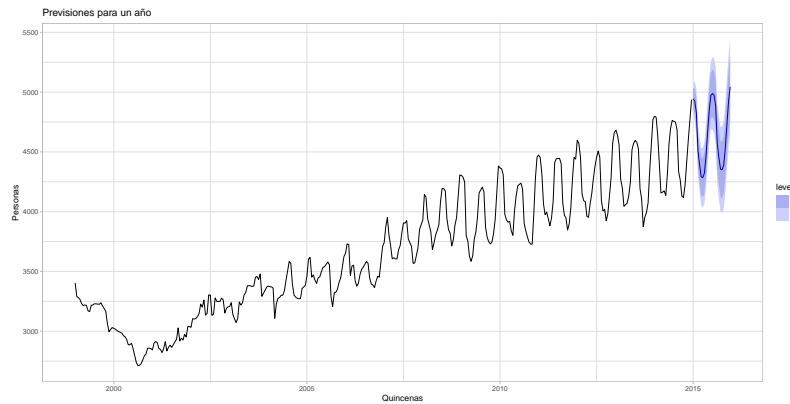


**Figura 3.8:** Gráficas para verificar los residuales del modelo  $ARIMA(1,1,1)(2,1,1)[12]$ , para la serie de ayudantes a partir de 1999.

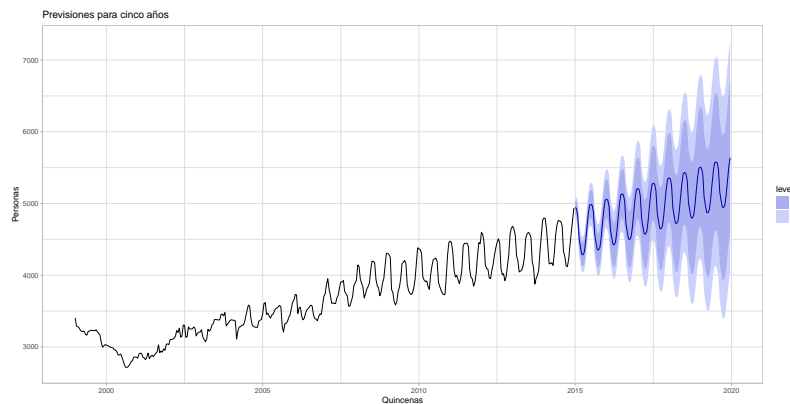
### 3.1. Análisis de la población de ayudantes de la UNAM de 2001 a 2014

Debido a que las primeras observaciones surgen de un evento atípico, la huelga de 1990, y el comportamiento estable la serie, la cantidad de los datos con los que se cuenta, se buscó reducir los efectos de tal evento quitando las primeras observaciones y analizando la serie a partir de 2001, la cual se denominará  $X_t$ .

Además de tomar en cuenta el contexto histórico para explicar los cambios en las series se puede utilizar la gráfica del comportamiento estacional, que es la descomposición de los datos en su frecuencia del patrón estacional por cada año de observación y señala más claramente los cambios en el comportamiento de las



**Figura 3.9:** Gráfica de los pronósticos para un año del modelo ARIMA(1,1,1)(0,1,1)[12].

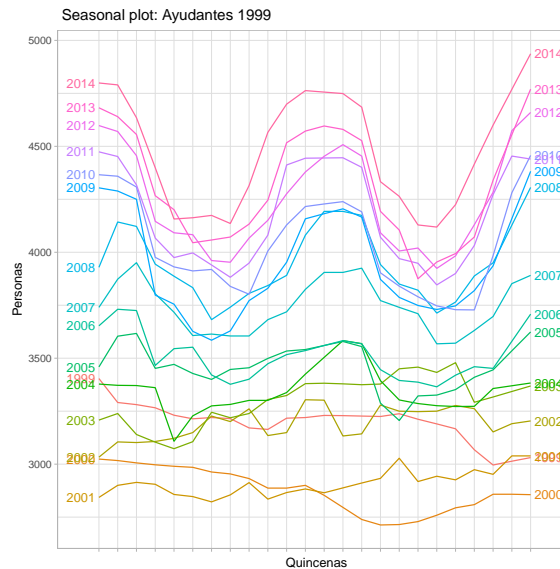


**Figura 3.10:** Gráfica de los pronósticos para cinco años del modelo ARIMA(1,1,1)(0,1,1)[12].

series.

Para el análisis de los cambios en el crecimiento de la población de ayudantes, la figura 3.11 es la gráfica del comportamiento estacional, muestra como los primeros dos años la población tiene un comportamiento diferente al resto. Sin embargo, a partir de 2007 la serie comienza a tener un patrón de comportamiento estacional semejante sin que el crecimiento de la población disminuya. No se aplicará una transformación de Box-Cox.

Se mantiene la variación estacional con periodo 12, para ello se volvió a graficar el periodograma la serie diferenciada y se mantuvo el resultado. De nueva cuenta, se intentaron diversos enfoques para hacer estacionaria la serie, al final la transformación con la menor estructura de dependencia fue la diferencia combinada  $(1 - B)(1 - B^{12})$ , para proponer ajustes hay que tomar en consideración la gráfica del ACF y PACF de la serie diferenciada que se muestra en la figura 3.12.



**Figura 3.11:** Gráfica del comportamiento estacional.

Resultado de la inspección las gráficas de las funciones ACF y PACF se sabe que la diferencia no logra explicar completamente la variación estacional, por lo que es necesario el uso de un modelo con términos estacionales, por la presencia de un valor grande al rezago 12 en la ACF y en los rezagos 12 y 24 para la PACF, se puede proponer modelos con factores estacionales  $MA_{12}(1)$  y  $AR_{12}(1)$  ó  $AR_{12}(2)$ . Mientras que para la parte no estacional se puede proponer los órdenes  $MA(1)$  o  $AR(2)$  y las combinaciones entre estos valores.

Luego de analizar los diferentes modelos propuestos, los expuestos en la tabla 3.5 son los que tienen el menor valor en los criterios de información, en las gráficas de sus respectivas funciones de ACF y PACF muestran menores asociaciones entre los residuales; por tanto son tomados en consideración. Se examinará la bondad de ajuste de cada uno para elegir el mejor.

**Tabla 3.5:** Modelos para la serie 2001

Modelo	Orden	AICc	BIC	p-valor LB
1	ARIMA(1,1,1)(0,1,1)[12]	3633.0765	3648.0613	0.1289
2	ARIMA(1,1,1)(1,1,1)[12]	3633.3040	3652.0030	0.1891

Si se considera el BIC o AICc se elegiría el modelo ARIMA(1,1,1)(0,1,1)[12], pero la diferencia entre los dos los criterios de selección es muy pequeña, ver tabla 3.5, sin embargo, si se considera la prueba de correlación o de ruido blanco para el rezago 48 se consideraría el modelo ARIMA(1,1,1)(1,1,1)[12].

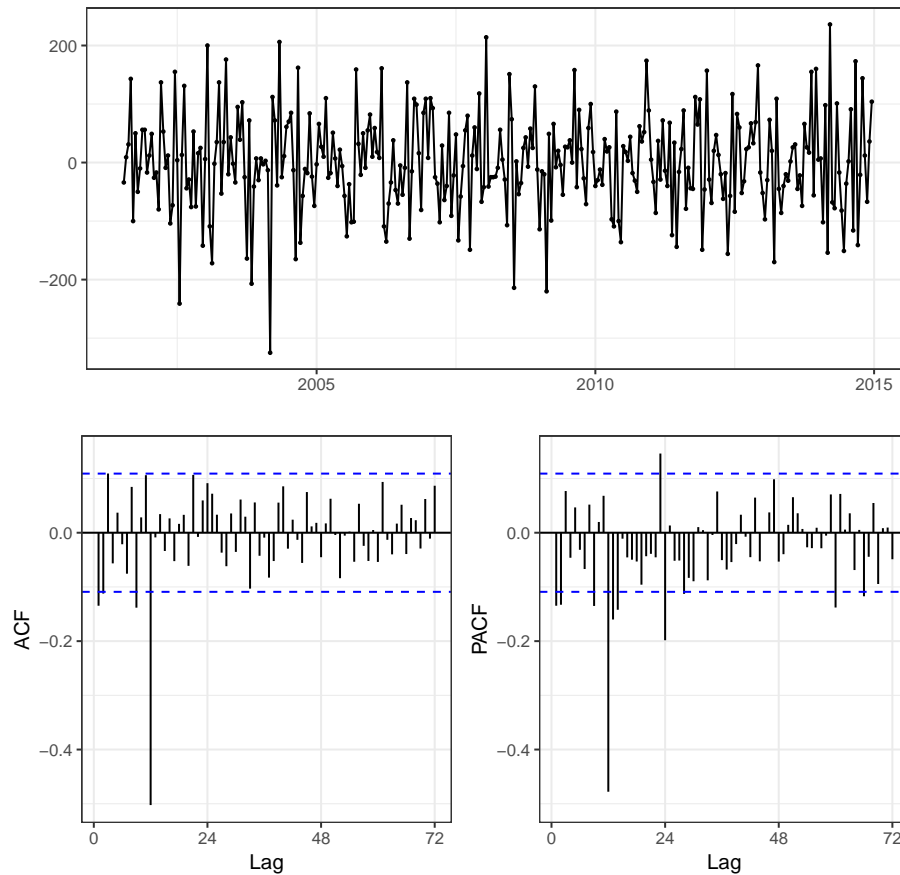
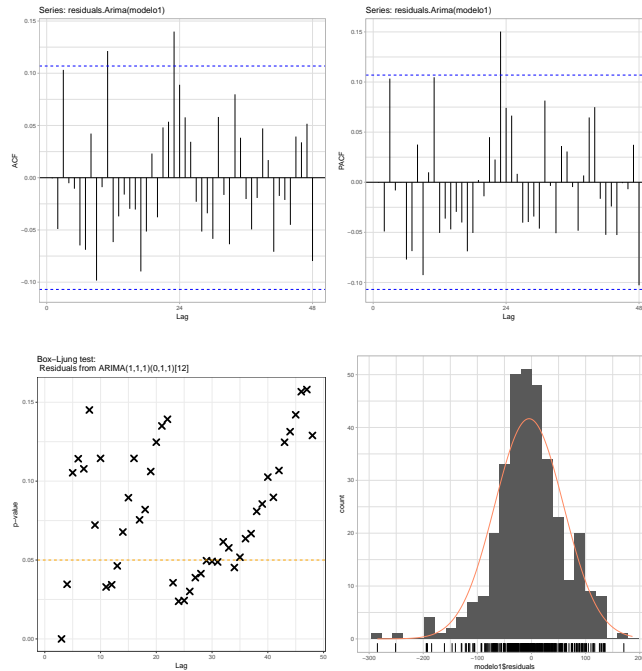


Figura 3.12:  $(1 - B)(1 - B^{12})X_t$  para la serie a partir de 2001.

Si siguiendo con las pruebas de normalidad para ambos modelos se tiene que el  $p$ -valor es menor al 5%, por lo que se concluye que sus datos no siguen una distribución normal con una confianza del 95%. Como se mencionó se puede tener un ruido blanco que no sea normal, aunque la normalidad de los residuales, es la base de la estimación de los parámetros, la cantidad de datos que se tiene es suficiente para que aunque no se cumpla la normalidad los estimadores sean calculados como variables normales de acuerdo a la ley de los grandes números.

```
> ad.test(m1_2001$residuals)
Anderson-Darling normality test
data:  m1_2001$residuals
A = 2.0833, p-value = 2.632e-05
```



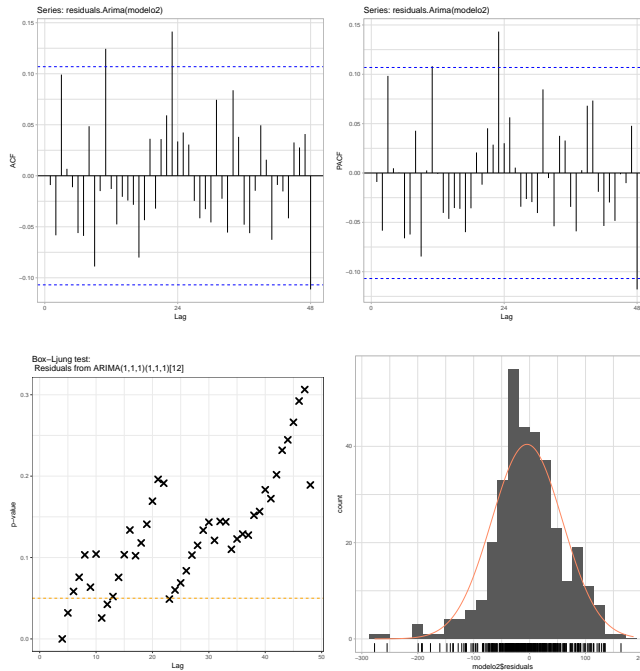
**Figura 3.13:** Análisis de residuales del modelo  $ARIMA(1,1,1)(0,1,1)[12]$  para 2001.

```
> ad.test(m2_2001$residuals)
Anderson-Darling normality test
data:  m2_2001$residuals
A = 2.0618, p-value = 2.971e-05
```

Considerando las gráficas de las funciones ACF y PACF para ambos modelos, figuras 3.13 y 3.14, son muy similares entre sí, pero cambia la latitud de los rezagos que rebasan el umbral de confianza, por lo cual se elegiría el modelo  $ARIMA(1,1,1)(0,1,1)[12]$ .

Sin embargo, al comparar las gráficas del p-valor de la prueba Ljung-Box, para los primeros 48 rezagos de ambos modelos, se observa como el modelo  $ARIMA(1,1,1)(1,1,1)[12]$  logra captar mejor las asociaciones entre los datos, *al tener una menor cantidad rezagos por debajo del umbral de rechazo de la hipótesis nula de autocorrelación cero*. Pero los residuales de ambos modelos están mostrando dependencia.

El ajuste seleccionado es el modelo  $ARIMA(1,1,1)(1,1,1)[12]$ , es decir, la evolución de la población de ayudantes analizados a partir de 2001 sigue un comportamiento de la forma:



**Figura 3.14:** Análisis de residuales del modelo ARIMA(1,1,1)(1,1,1)[12] para 2001.

$$(1 - B)(1 - B^{12})(1 - \phi_1)(1 - \Phi_{12})X_t = (1 - \theta_1)(1 - \Theta_{12})\varepsilon_t$$

donde los parámetros  $\phi_1, \theta_1, \Phi_{12}, \Theta_{12}$  con sus respectivos intervalos de confianza están dados en la tabla 3.6.

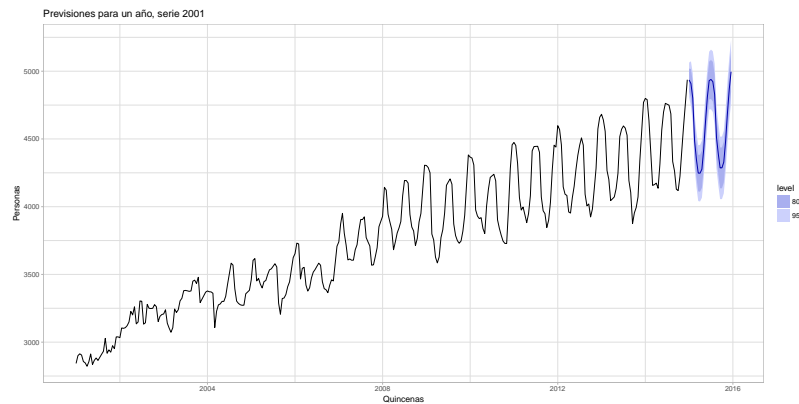
Aunque el modelo propuesto falle en la bondad de ajuste, el modelo logra explicar mejor las asociaciones entre los datos a partir de 2001, donde la serie tiene menos variaciones inusuales, pero también se observa como el efecto de la huelga sigue afectando la evolución de la población de ayudantes.

Debido a que el modelo ARIMA(1,1,1)(1,1,1)[12] logra captar mejor la variación de la población de ayudantes, también se realizó la estimación de los pronósticos, en las figuras 3.15 y 3.16 se puede observar los pronósticos para un año y cinco años respectivamente.

Comparando los pronósticos de los modelos ARIMA(1,1,1)(0,1,1)[12] y ARIMA(1,1,1)(1,1,1)[12] para la serie tomada a partir de 1999 y 2001 respectivamente, se observa como las estimaciones puntuales se parecen, figura 3.17, y como los intervalos de confianza del primer modelo son más grandes, figura 3.18, debido en el primer ajuste no se logra captar completamente la estructura de correlación.

**Tabla 3.6:** Parámetros estimados y su intervalo de confianza para el modelo ARIMA(1,1,1)(0,1,1)[12] serie tomada a partir de 2001.

Parámetro	Estimado	2.5 %	97.5 %
$\phi_1$	0.80	0.73	0.86
$\theta_1$	-1.00	-1.04	-0.96
$\Phi_{1_{12}}$	-0.14	-0.35	0.07
$\theta_{1_{12}}$	-0.53	-0.73	-0.33



**Figura 3.15:** Gráfica de los pronósticos para un año del modelo ARIMA(1,1,1)(1,1,1)[12] para la serie tomada a partir de 2001.

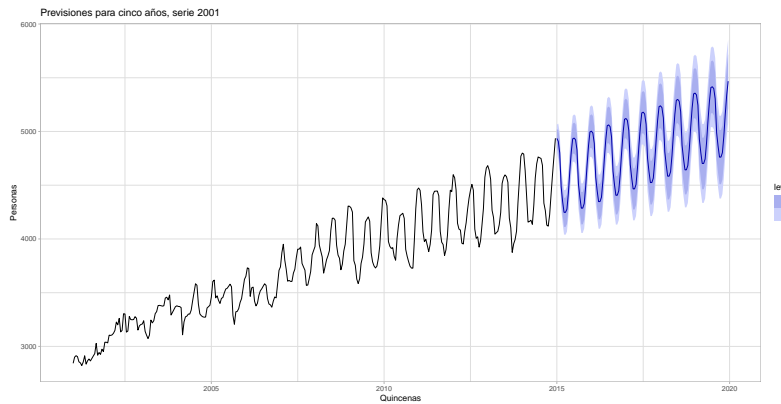
### 3.2. Análisis de la población de ayudantes de la UNAM de 2009 a 2014

Como resultado de los análisis de la serie desde 1999 y la considerada a partir de 2001, ambos siguen sin proporcionar un modelo capaz de explicar las correlaciones entre las observaciones de la serie, se propuso de nueva cuenta analizar la serie pero a partir de 2009, denotada por  $Z_t$ , para proporcionar pronósticos más confiables.

Ya que la serie observada desde 2009 tiene una componente estacional muy marcada, primero se verificó que la serie mantenía el mismo periodo a través de su periodograma, además persiste la tendencia lineal, una vez tomado en consideración lo anterior, se trató de modelar mediante una regresión sobre el tiempo y las funciones armónicas en la frecuencia  $\frac{1}{12}$ , es decir:

$$Z_t = \beta_0 + \beta_1 t + \beta_2 \text{sen} \left( 2\pi t \frac{1}{12} \right) + \beta_3 \text{cos} \left( 2\pi t \frac{1}{12} \right) + \varepsilon_t,$$

también se intentó con las frecuencias  $\frac{1}{3}$ ,  $\frac{1}{4}$ ,  $\frac{1}{6}$ ,  $\frac{5}{12}$ ,  $\frac{1}{2}$  y  $\frac{1}{24}$ , pero los residuales de dichos ajustes conservaban



**Figura 3.16:** Gráfica de los pronósticos para cinco años del modelo ARIMA(1,1,1)(1,1,1)[12] para la serie tomada a partir de 2001.

bastantes asociaciones entre las observaciones y al tratar de modelar con procesos ARMA, dichos modelos carecían de bondad de ajuste.

Una vez más, se utiliza el enfoque de la diferencia combinada para quitar la tendencia y el patrón estacionario, dada por  $(1 - B)(1 - B^{12})$ , para proponer un ajuste hay que tomar en consideración la gráfica de las ACF y PACF de la serie diferenciada, figura 3.19.

Después de analizar varios modelos, los que tienen mejores propiedades se exhiben en la tabla 3.7, si se observa son los mismos modelos examinados para la serie anterior, sin embargo hay que verificar la bondad de ajuste sin que influya en el resultado anterior.

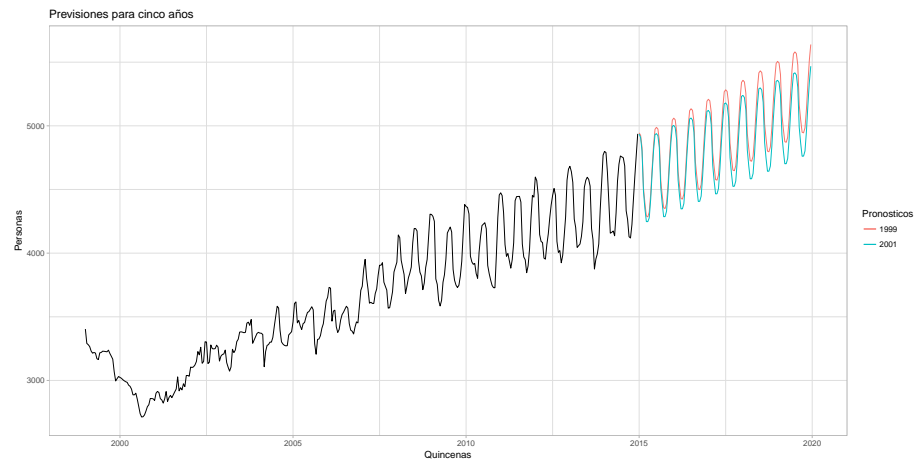
**Tabla 3.7:** Modelos para la serie 2009.

Modelo	Orden	AICc	BIC	p-valor LB
1	ARIMA(1,1,1)(0,1,1)[12]	1473.1436	1484.3269	0.4677
2	ARIMA(1,1,1)(1,1,1)[12]	1474.9116	1488.8076	0.4328

Si se considera el BIC o AICc se elegiría el modelo ARIMA(1,1,1)(0,1,1)[12], pero la diferencia entre ambos modelos es muy pequeña, ver tabla 3.7, si se considera la prueba de no correlación o de ruido blanco para el rezago 48, también se consideraría el modelo ARIMA(1,1,1)(0,1,1)[12].

De nueva cuenta, ambos modelos fallan la prueba de normalidad, no podrá tomarse como criterio de selección esta prueba. Observando las gráficas de ACF y PACF de ambos modelos, figuras 3.20 y 3.14 respectivamente, parece que existe una asociación entre las observaciones separadas a 12 rezagos no explicada por los modelos, esta correlación significativa hace dudar que los residuales sean un ruido blanco.





**Figura 3.17:** Comparación de pronósticos puntuales para las series 1999 vs 2001.

Sin embargo, examinado las gráficas del  $p$ -valor de la prueba Ljung-Box para los modelos, figuras 3.20 y 3.14 respectivamente, la mayoría de rezagos tienen un  $p$ -valor  $> 0,05$ , por lo cual, no se tiene suficiente evidencia estadística para rechazar la hipótesis nula. Lo cual implica que no se puede suponer que los residuales están autocorrelacionados.

Box-Ljung test

```
data: residuals.Arima(aj1_2009)
```

```
X-squared = 6.0838, df = 9, p-value = 0.7315
```

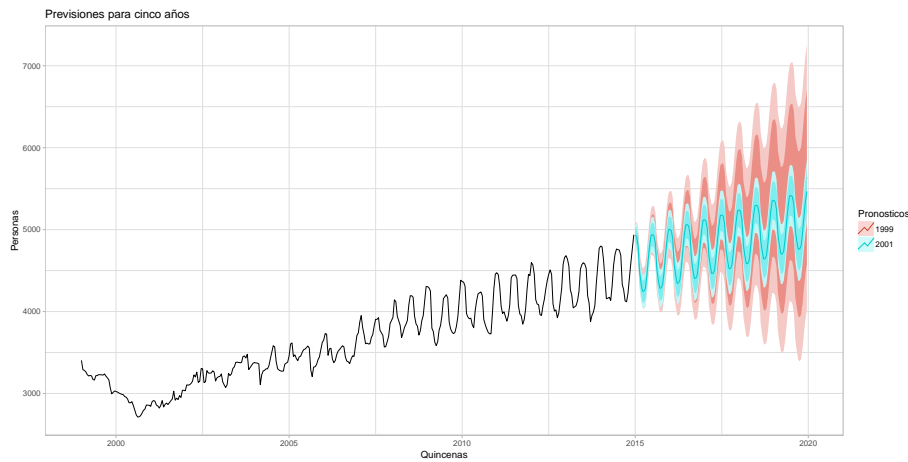
Box-Ljung test

```
data: residuals.Arima(aj2_2009)
```

```
X-squared = 5.3862, df = 8, p-value = 0.7156
```

Como resultado de examinar las gráficas de las ACF y PACF de los residuales de ambos modelos, donde se muestra una dependencia en los residuales separados por 12 periodos, además de no poder concluir la no dependencia para los residuales de ambos modelos, se realizará la prueba Box-Ljung para el rezago 12, la cual tiene un  $p$ -valor  $> 0.05$ , se concluye que no se tiene suficiente evidencia estadística para afirmar la dependencia de los valores, tampoco puede afirmar la independencia de los residuales hasta el rezago 12.

Debido a que los modelos tienen características estadísticas similares, la elección se tomará por parsimonia y el modelo ARIMA(1,1,1)(0,1,1)[12] es el seleccionado, la evolución de la población de ayudantes analizados a partir de 2009 sigue un comportamiento de la forma:



**Figura 3.18:** Comparción de pronósticos con intervalos de confianza para las series 1999 vs 2001.

$$(1 - B)(1 - B^{12})(1 - \phi_1)Z_t = (1 - \theta_1)(1 - \Theta_{12})\varepsilon_t,$$

donde los parámetros  $\phi_1, \theta_1, \Theta_{12}$  con sus respectivos intervalos de confianza están dados en la tabla 3.8.

**Tabla 3.8:** Parámetros estimados y su intervalo de confianza para el modelo ARIMA(1,1,1)(0,1,1)[12].

Parámetro	Estimado	2.5 %	97.5 %
$\phi_1$	0.71	0.58	0.84
$\theta_1$	-1.00	-1.08	-0.92
$\Theta_{12}$	-0.79	-1.01	-0.56

Los pronósticos, realizados por medio de este ajuste serán más confiables ya que este modelo logra captar mejor las relaciones de asociación de los datos, los datos tienen menos variaciones. En las figuras 3.22 y 3.23, se muestran en los pronósticos la población de ayudantes para uno y cinco años.

Comparando los pronósticos de los modelos ARIMA(1,1,1)(0,1,1)[12], ARIMA(1,1,1)(1,1,1)[12] y ARIMA(1,1,1)(0,1,1)[12] para la serie tomada a partir de 1999, 2001 y 2009 respectivamente, las estimaciones puntuales se parecen aunque las previsiones realizadas con base en la serie de 1999 proyecta un mayor crecimiento, figura 3.24, y en el caso de los pronósticos con sus respectivos intervalos de confianza se muestran en la figura 3.18, observese como reduciendo la variabilidad de los datos, la anchura de los intervalos de confianza disminuyen, en el primer ajuste no se logra captar completamente la estructura de correlación.

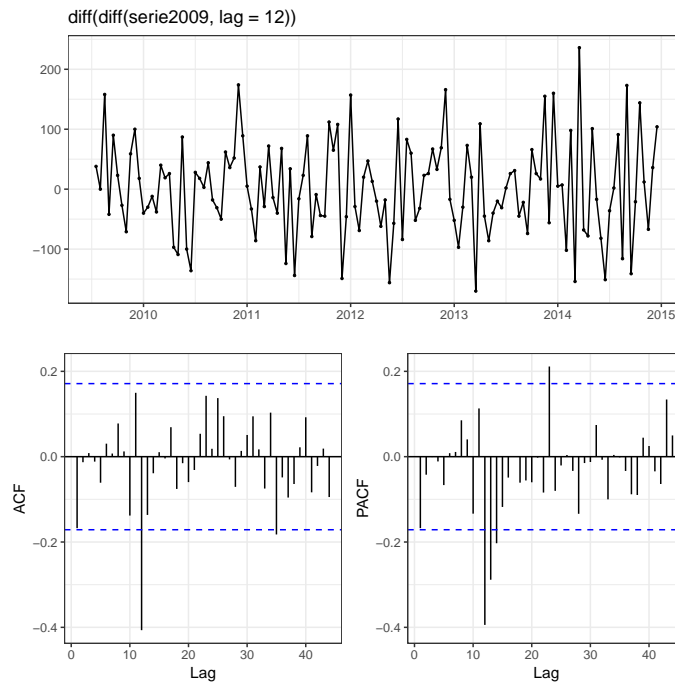


Figura 3.19: Serie a partir de 2009 con la diferencia  $(1 - B)(1 - B^{12})Z_t$ .

### Otro modelo

Al final, se encontró un modelo que cumple con todas las pruebas de bondad de ajuste, solo que es más complejo de explicar que los modelos anteriores.

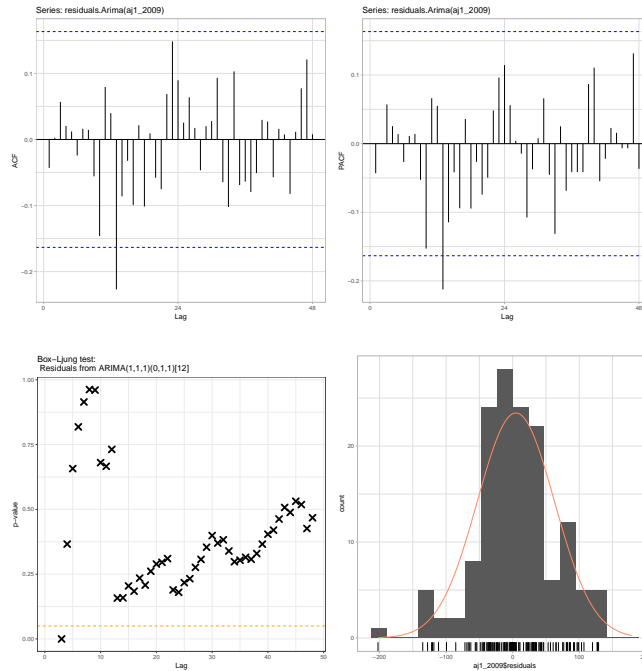
En este enfoque primero se trató la variación estacional por medio de variables mudas y se realizó una regresión con ellas, los ponderadores asociados con este filtro se pueden observar en la tabla 3.9.

Después de modelar la componente estacional falta explicar la tendencia lineal que aun existe en los residuales, ver figura 3.26, para lo cual se aplicó una diferencia de orden uno, ver figura 3.27.

Para proponer ajustes hay que tomar en consideración la gráfica del ACF y PACF de la serie diferenciada de residuales que se muestra en la figura 3.27, un valor grande cerca del rezago 12 para ambas gráficas hacen sospechar que la regresión con variables mudas no logra explicar completamente la variación estacional, por lo que es necesario el uso de un modelo con términos estacionales, se puede proponer modelos con factores estacionales MA(1) o AR(1), mientras que para la parte no estacional se puede proponer los órdenes MA(1) o AR(2) y las combinaciones entre estos valores.

Al final el modelo propuesto para la serie de residuales está dado por:  $(1 - B)(1 - \phi_1)(1 - \Phi_{12})v_t = (1 - \theta_1)\varepsilon_t$ , los valores estimados de los parámetros se muestran en la tabla 3.10.

Se probará la bondad de ajuste para este modelo, primero se inspeccionó la gráficas de ACF y PACF, figura 3.28, las cuales son semejantes a las gráficas respectivas de un ruido blanco, puesto que los valores de



**Figura 3.20:** Análisis de residuales del modelo  $ARIMA(1,1,1)(0,1,1)[12]$  para la serie 2009.

los rezagos no son significativos, es decir no rebasan el umbral de confianza.

De la gráfica del  $p$ -valor de la prueba Ljung-Box, figura 3.28, parece que no hay información suficiente para demostrar la existencia de dependencias en los residuales para los diferentes rezagos, más adelante se muestra la prueba de hipótesis de no dependencia en el rezago 48, cuya decisión es no en rechazar  $H_0$  pues  $p\text{-valor} > 0.05$ , entonces no se puede concluir que los datos no mantengan una estructura de dependencia.

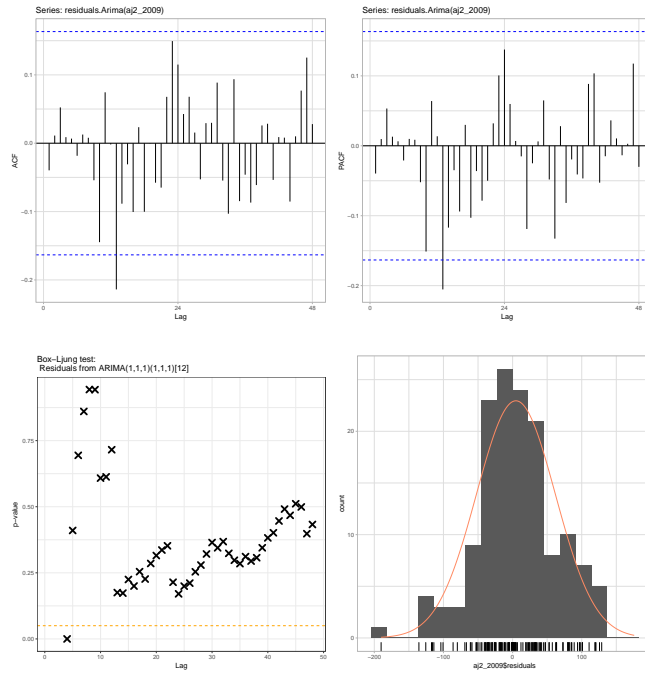
Para los resultados de las pruebas de normalidad Anderson-Darling y Lilliefors, para ambas se tiene el  $p\text{-valor} > 0.05$ , por lo cual no se rechaza  $H_0$ , es decir no se tiene suficiente evidencia estadística para concluir los residuales no siguen una distribución normal.

Anderson-Darling normality test

```
data: residuals.Arima(r.ajuste.fit)
```

A = 0.50473, p-value = 0.2001

Lilliefors (Kolmogorov-Smirnov) normality test



**Figura 3.21:** Análisis de residuales del modelo  $ARIMA(1,1,1)(1,1,1)[12]$  para la serie 2009.

```
data: residuals.Arima(r.ajuste.fit)
```

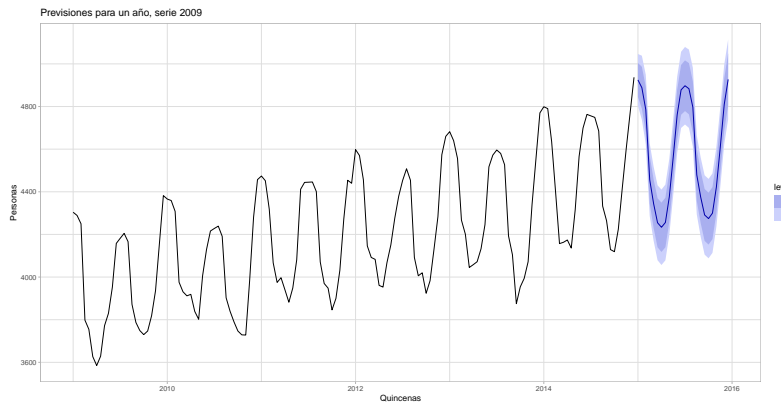
D = 0.067098, p-value = 0.1162

Ljung-Box test

```
data: Residuals from ARIMA(1,1,1)(1,0,0)[12]
```

Q\* = 33.004, df = 45, p-value = 0.9076

Model df: 3. Total rezagos used: 48



**Figura 3.22:** Gráfica de los pronósticos para un año del modelo ARIMA(1,1,1)(0,1,1)[12] para la serie tomada a partir de 2009.

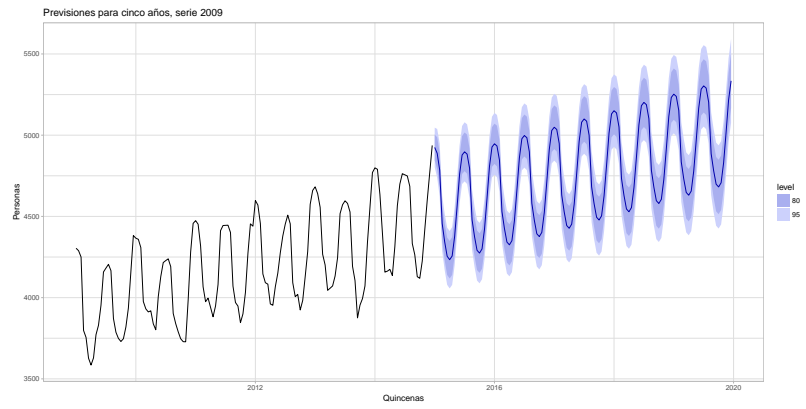
El modelo completo se puede expresar como:

$$Z_t = \beta_0 + \sum_{i=1}^{23} \beta_i \delta_j(t) + v_t, \tag{3.1}$$

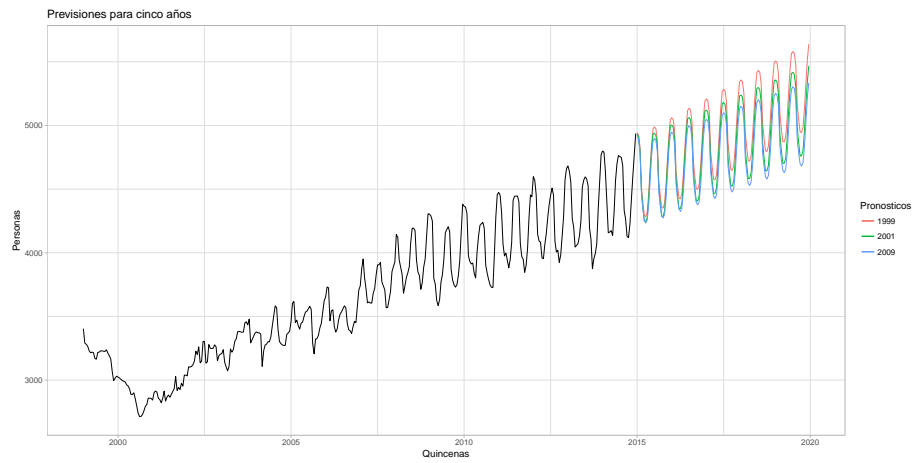
$$(1 - B)(1 - \phi_1)(1 - \Phi_{12})v_t = (1 - \theta_1)\varepsilon_t$$

donde  $\delta_j(t) = \begin{cases} 1 & \text{si } t = j, j + s, j + 2s, \dots \text{ para } j = 1, \dots, 23 \text{ y } s = 24, \\ 0 & \text{en otro caso,} \end{cases}$

El uso de las variables mudas se puede observar como un filtro lineal que quita una memoria larga donde los sucesos que afecta hoy estarán impactando el valor de la serie durante 24 quincenas, es decir un año es el efecto que no logran captar solo a través de modelos ARIMA.



**Figura 3.23:** Gráfica de los pronósticos para cinco años del modelo  $ARIMA(1,1,1)(0,1,1)[12]$  para la serie tomada a partir de 2009.



**Figura 3.24:** Comparación de pronósticos puntuales para las series 1999 vs 2001 vs 2009.

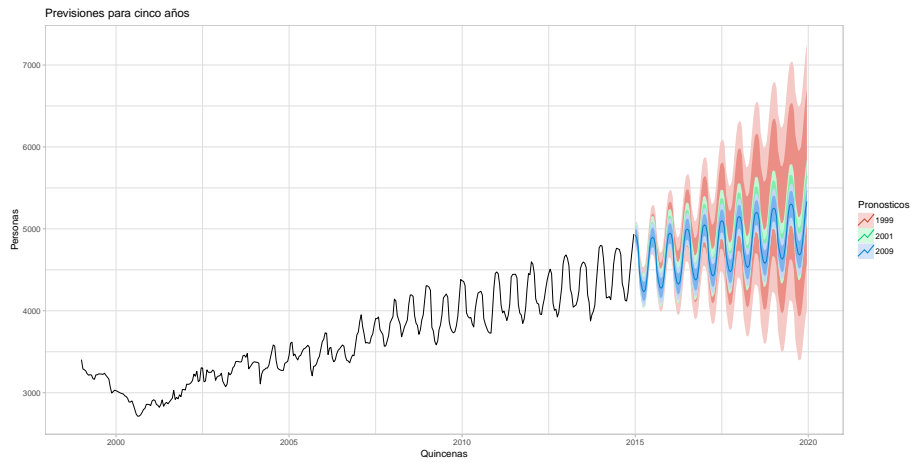


Figura 3.25: Comparación de pronósticos con intervalos de confianza para las series 1999 vs 2001 vs 2009.

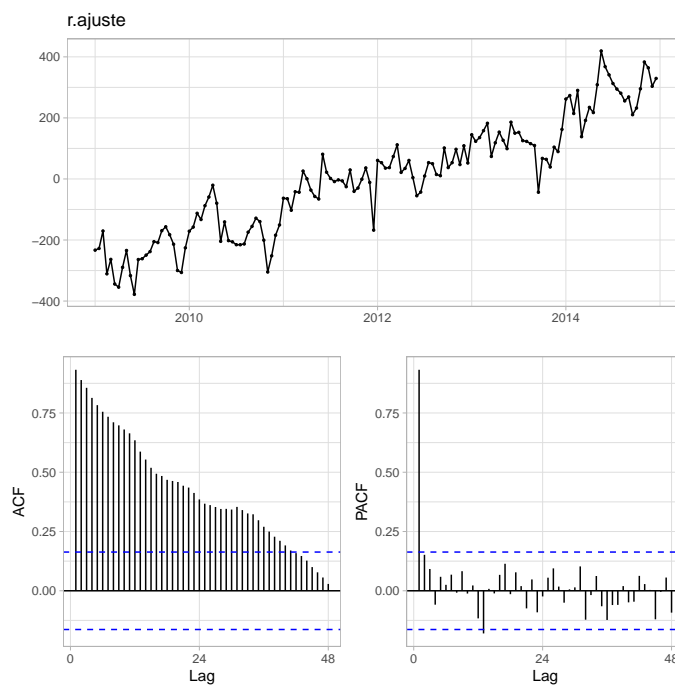
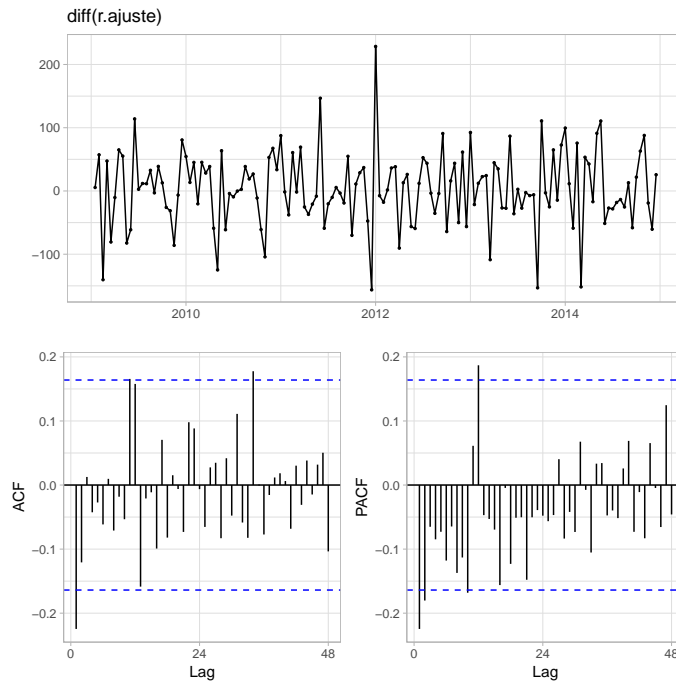


Figura 3.26: Serie de residuales  $v_t$  para modelar la serie a partir de 2009.



**Tabla 3.9:** Coeficientes de la :  $Z_t = \beta_0 + \sum_{i=1}^{23} \beta_i \delta_j(t) + v_t$ .

	Estimate	Std. Error	t value	Pr(>  t )
$\beta_0$	4607.6667	82.6602	55.74	0.0000
$\beta_1$	-70.5000	116.8992	-0.60	0.5476
$\beta_2$	-91.0000	116.8992	-0.78	0.4378
$\beta_3$	-187.1667	116.8992	-1.60	0.1120
$\beta_4$	-498.8333	116.8992	-4.27	0.0000
$\beta_5$	-589.1667	116.8992	-5.04	0.0000
$\beta_6$	-636.5000	116.8992	-5.44	0.0000
$\beta_7$	-668.1667	116.8992	-5.72	0.0000
$\beta_8$	-689.1667	116.8992	-5.90	0.0000
$\beta_9$	-601.3333	116.8992	-5.14	0.0000
$\beta_{10}$	-461.0000	116.8992	-3.94	0.0001
$\beta_{11}$	-276.6667	116.8992	-2.37	0.0195
$\beta_{12}$	-185.6667	116.8992	-1.59	0.1149
$\beta_{13}$	-164.3333	116.8992	-1.41	0.1624
$\beta_{14}$	-153.1667	116.8992	-1.31	0.1926
$\beta_{15}$	-203.6667	116.8992	-1.74	0.0840
$\beta_{16}$	-530.3333	116.8992	-4.54	0.0000
$\beta_{17}$	-612.3333	116.8992	-5.24	0.0000
$\beta_{18}$	-689.1667	116.8992	-5.90	0.0000
$\beta_{19}$	-721.0000	116.8992	-6.17	0.0000
$\beta_{20}$	-678.0000	116.8992	-5.80	0.0000
$\beta_{21}$	-574.8333	116.8992	-4.92	0.0000
$\beta_{22}$	-371.8333	116.8992	-3.18	0.0019
$\beta_{23}$	-142.3333	116.8992	-1.22	0.2258



**Figura 3.27:** Serie de residuales diferenciada  $(1 - B) v_t$  para modelar la serie a partir de 2009.

**Tabla 3.10:** Parámetros estimados con su intervalo de confianza.

Parámetro	Estimador	2.5 %	97.5 %
$\phi_1$	0.51	0.12	0.90
$\theta_1$	-0.79	-1.09	-0.50
$\Phi_{1 2}$	0.18	0.01	0.35

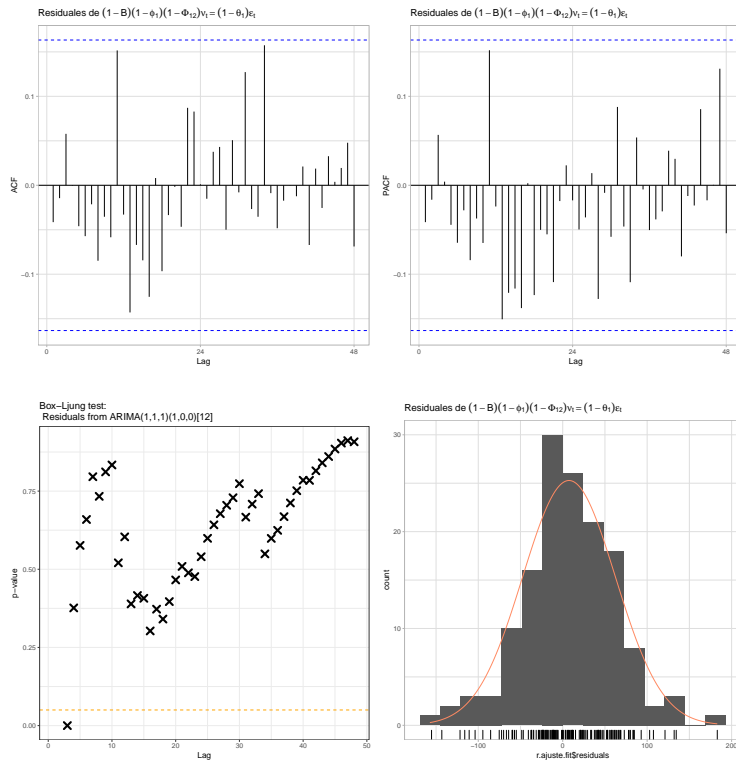


Figura 3.28: Análisis de residuales para el ajustes de la serie 2009.

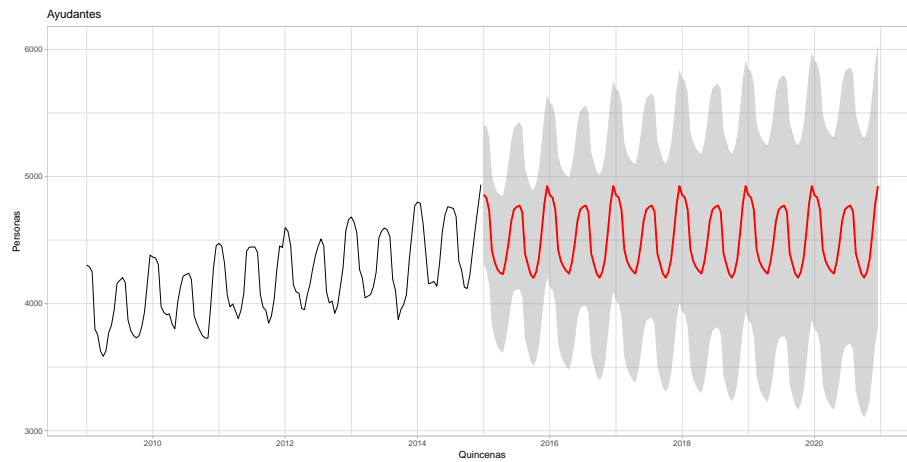


Figura 3.29: Pronósticos  $Z_t = \beta_0 + \sum_{i=1}^{23} \beta_i \delta_j(t) + \frac{(1-\theta_1)}{(1-B)(1-\phi_1)(1-\Phi_{12})} \epsilon_t$ .

## Capítulo 4

# Conclusiones

La mayoría de la población de ayudantes son jóvenes, aunque existen sus excepciones, lo cual también influye en la antigüedad donde se puede observar, como la mayoría de los ayudantes cumplen con el límite de cinco años, pues al final de cuentas la ayudantía es solo un proceso de formación.

También es previsible que entre más ayudantes menor serán las horas de servicio que cada persona pueda brindar. El comportamiento ideal sería, que la población de ayudantes creciera junto con su rotación, es decir que se cumpliera con el límite de cinco años, haciendo a la ayudantía más accesible para quienes cumplan con los requisitos.

Cada paso del proceso es importante desde, ¿cuál es el problema a resolver?, ¿cuál es el objetivo?, ¿con qué información se cuenta?, ¿cómo recolectar los datos?, ¿es necesario tratar los datos?, ¿qué procedimiento o técnica se debe utilizar?, verificar los resultados.

En este caso particular, el tomar los datos de manera quincenal se utilizó como recurso para tener más datos para elaborar el análisis, es claro que esta decisión también involucró la existencia de un patrón estacional en la serie. Se puede pensar e incluso sucede que en varias observaciones la serie se mantenía igual de una quincena a otra o aumentaba mínimamente pero también se observan grandes diferencias en las quincenas del principio y final del semestre, que no se hubiesen notado de no tener los datos, y que puede ser útil en la planeación de recursos.

El comportamiento de la población de ayudantes cambió abruptamente desde la huelga de 1999, este cambio repercutió a lo largo de muchos años, haciendo que el enfoque ARIMA no fuese suficiente para analizar la evolución de ayudantes.

La segmentación de la serie, con la que se buscaba una serie más regular, dejó modelos muy parecidos con deficiencias en la bondad de ajuste, es decir, todos ellos tienen un remanente de la estructura de relación. La aplicación de la función logaritmo es necesaria para datos provenientes de conteos.

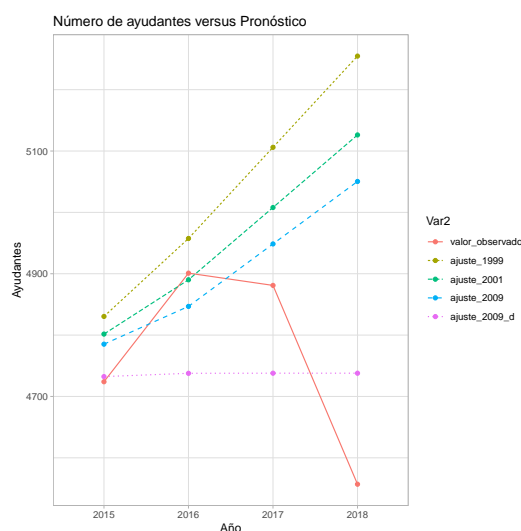
En cuanto a las previsiones o pronósticos, el modelo basado en la serie de 1999 prevé un crecimiento sostenido, mientras el modelo con variables indicadoras anticipa un tope de crecimiento y un comportamiento estacional más suavizado.

Como ya se cuenta con los valores pronosticados, es un buen ejercicio ver qué tan alejados se encontraron los valores proporcionados por los modelos propuestos de los valores reales. Dicha comparación se puede ver en la tabla 4.1 y en la figura 4.1.

El modelo basado en la serie de 2009 con variables indicadoras en cuestión de pronósticos no logra capturar todas las asociaciones, aunque para el año 2015 y 2018 tiene los valores más cercanos a los reales. En definitiva el ajuste con base en la serie completa sobre estima la población de ayudantes. Mientras que los pronósticos para la serie segmenta en 2001 y 2009 siguen un comportamiento muy parecido, aunque la previsión para el 2016 fue mejor para el modelo con base en la serie de 2001.

Año	Valor real	modelado 1999	modelado 2001	modelado 2009	modelado 2009 va. indicadoras
2015 <sup>1</sup>	4,724	4,830.396	4,801.676	4,785.212	4,732.375
2016 <sup>2</sup>	4,901	4,957.476	4,890.147	4,846.785	4,737.729
2017 <sup>3</sup>	4,881	5,106.054	5,007.968	4,948.595	4,737.903
2018 <sup>4</sup>	4,557	5,254.751	5,126.227	5,050.417	4,737.908

**Tabla 4.1:** Comparación de las estadísticas oficiales con los pronósticos



**Figura 4.1:** Número ayudantes contra los pronósticos realizados para el modelado en 1999, 2001, 2009 y 2009 con variables mudas

<sup>1</sup> FUENTE: Nómina de la quincena 03 de 2015, Dirección General de Personal, UNAM.

<sup>2</sup> FUENTE: Nómina de la quincena 03 de 2016, Dirección General de Personal, UNAM.

<sup>3</sup> FUENTE: Nómina de la quincena 03 de 2017, Dirección General de Personal, UNAM.

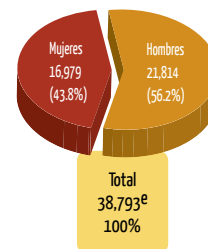
<sup>4</sup> FUENTE: Nómina de la quincena 03 de 2018, Dirección General de Personal, UNAM.

En la actualidad ya existen un sin número de paquetes estadísticos y enfoques para manejar series de tiempo, en particular el análisis de intervención se ha popularizado para estas, pues al momento de ajustar dan un modelo con un orden reducido con los outliers como regresores de la serie, otro enfoque se da con el apogeo big data, el tratamiento de las series de tiempo mediante las redes neuronales .

## 2 Universidad Nacional Autónoma de México

Personal académico | 2015  
Personal académico por figura

Figura académica	Personas <sup>a</sup>		
	Hombres	Mujeres	Total
Investigador	1,647	904	2,551
Profesor de Carrera	3,031	2,390	5,421
Técnico Académico	2,053	2,198	4,251
Profesor de Asignatura	13,810	10,525	24,335
Ayudantes <sup>b</sup>	2,550	2,174	4,724
Otros <sup>c</sup>	163	58	221



a Se refiere al número de académicos dentro de una misma figura o subsistema, sin duplicidad alguna. Puede ocurrir duplicidad de académicos entre figuras o subsistemas (véase personal académico en el Glosario).

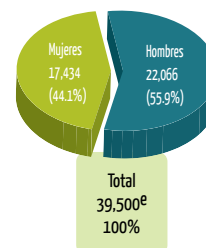
b Incluye a las figuras de Ayudante de Profesor de Asignatura y Ayudante de Investigador.

c Incluye a profesores e investigadores visitantes y eméritos, a jubilados docentes en activo y a jubilados eméritos en activo.

## 2 Universidad Nacional Autónoma de México

Personal académico | 2016  
Personas por figura académica

Figura académica	Personas <sup>a</sup>		
	Hombres	Mujeres	Total
Investigador	1,672	920	2,592
Profesor de Carrera	3,034	2,428	5,462
Técnico Académico	2,116	2,242	4,358
Profesor de Asignatura	13,980	10,789	24,769
Ayudantes <sup>b</sup>	2,589	2,312	4,901
Otros <sup>c</sup>	151	54	205



a Se refiere al número de académicos dentro de una misma figura o subsistema, sin duplicidad alguna. Puede ocurrir duplicidad de académicos entre figuras o subsistemas (véase personal académico en el Glosario).

b Incluye a las figuras de Ayudante de Profesor de Asignatura y Ayudante de Investigador.

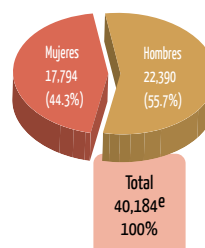
c Incluye a profesores e investigadores visitantes y eméritos, a jubilados docentes en activo y a jubilados eméritos en activo.

**Figura 4.2:** Estadísticas oficiales para 2015 y 2016

### Personal académico | 2017

Personas por figura académica

Figura académica	Personas <sup>a</sup>		
	Hombres	Mujeres	Total
Investigador	1,680	935	2,615
Profesor de Carrera	3,040	2,447	5,487
Técnico Académico	2,159	2,264	4,423
Profesor de Asignatura	14,209	11,132	25,341
Ayudantes <sup>b</sup>	2,620	2,261	4,881
Otros <sup>c</sup>	148	52	200



a Se refiere al número de académicos dentro de una misma figura o subsistema sin duplicidad alguna (véase *personal académico* en el Glosario).

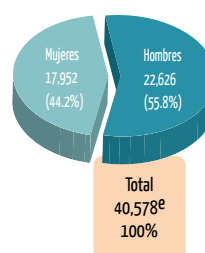
b Incluye a las figuras de Ayudante de Profesor de Asignatura y Ayudante de Investigador.

c Incluye a profesores e investigadores visitantes y eméritos, a jubilados docentes en activo y a jubilados eméritos en activo.

### Personal académico | 2018

Personas por figura académica

Figura académica	Personas <sup>a</sup>		
	Hombres	Mujeres	Total
Investigador	1,706	953	2,659
Profesor de Carrera	3,054	2,449	5,503
Técnico Académico	2,180	2,289	4,469
Profesor de Asignatura	14,484	11,407	25,891
Ayudantes <sup>b</sup>	2,488	2,069	4,557
Otros <sup>c</sup>	137	45	182



a Se refiere al número de académicos dentro de una misma figura o subsistema sin duplicidad alguna (véase *personal académico* en el Glosario).

b Incluye a las figuras de Ayudante de Profesor de Asignatura y Ayudante de Investigador.

c Incluye a profesores e investigadores visitantes y eméritos, jubilados docentes y eméritos en activo.

**Figura 4.3:** Estadísticas oficiales para 2017 y 2018





# Bibliografía

- Bisgaard, S. and Kulahci, M. (2011). *Time Series Analysis and Forecasting by Example*. Wiley Series in Probability and Statistics. Wiley.
- Brockwell, P. and Davis, R. (2002). *Introduction to Time Series and Forecasting*. Number v. 1 in Introduction to Time Series and Forecasting. Springer. [1](#), [6](#), [7](#), [13](#), [14](#), [16](#), [33](#)
- Brockwell, P. and Davis, R. (2009). *Time Series: Theory and Methods*. Springer Series in Statistics. Springer. [7](#), [8](#), [10](#), [14](#)
- Chatfield, C. (2004). *The Analysis of Time Series: An Introduction*. CRC Press, Florida, US, 6th edition.
- Cowpertwait, P. and Metcalfe, A. (2009). *Introductory Time Series with R*. Use R! Springer New York. [1](#), [31](#)
- Cryer, J. and Chan, K. (2008). *Time Series Analysis: With Applications in R*. Springer Texts in Statistics. Springer New York. [2](#), [31](#), [33](#)
- Douc, R., Moulines, E., and Stoffer, D. (2014). *Nonlinear Time Series: Theory, Methods and Applications with R Examples*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.
- González, V. M. C. (2011). *Pronosticos: Metodología de Box-Jenkins*. [2](#), [15](#)
- Guerrero, V. (2003). *Análisis Estadístico de Series de Tiempo Económicas*. International Thomson Editores, S. A. de C. V. [2](#)
- Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press. [1](#), [5](#), [10](#), [13](#), [14](#)
- Hyndman, R. and Athanasopoulos, G. (2014). *Forecasting: principles and practice*. OTexts. [1](#), [31](#)
- Janacek, G. (2001). *Practical Time Series*. Arnold texts in statistics. Arnold. [3](#)
- Kirchgässner, G. and Wolters, J. (2008). *Introduction to Modern Time Series Analysis*. Springer.
- Li, T. (2013). *Time Series with Mixed Spectra*. Taylor & Francis.
- Madsen, H. (2007). *Time Series Analysis*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis. [2](#), [4](#), [5](#), [6](#), [7](#), [8](#), [16](#), [17](#), [18](#), [19](#)
- Montgomery, D., Jennings, C., and Kulahci, M. (2011). *Introduction to Time Series Analysis and Forecasting*. Wiley Series in Probability and Statistics. Wiley. [1](#)

- Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. Use R! Springer, New York, second edition.
- Rincón, L. (2007). *Curso intermedio de probabilidad*. Las prensas de ciencias. México : UNAM, Facultad de Ciencias, 2007. 3
- Shumway, R. and Stoffer, D. (2006). *Time Series Analysis and Its Applications: With R Examples*. Springer Texts in Statistics. Springer New York. 1, 15, 16, 17, 31, 33
- Wei, W. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. Pearson Addison Wesley. 1, 10