

Trabajo Práctico N° 2

Análisis de Series Temporales

Maestría en Explotación de Datos y Gestión del Conocimiento

Layla Scheli, Franco Lianza, Lucio Scalzo, Ignacio Mujica, Alexis Walker

09 December 2021

Índice

1	Introducción	4
2	Marco Teórico	5
2.1	Tests de estacionariedad	5
2.2	Modelos SARIMA	5
3	Análisis de Resultados	6
3.1	Índice de Precios al Consumidor	6
3.2	Índice de Salarios	8
4	Conclusiones	11
5	Referencias bibliográficas	12
6	Apéndices	13
6.1	Modelo $ARIMA(0, 2, 6)$ para Índice de Precios al Consumidor	13
6.2	Modelo $SARIMA(0, 2, 0)(0, 2, 6)[12]$ para Índice de Precios al Consumidor	13
6.3	Modelo $SARIMA(1, 0, 0)(1, 0, 0)[12]$ para Índice de Salarios	13

Abstract

El presente trabajo tiene por objetivo estudiar la relación entre dos variables económicas fundamentales, el IPC (Índice de precios al Consumidor) y el índice de Salarios.

Ambos indicadores son considerados fundamentales para el desarrollo de un país ya que conllevan, en la relación entre ambos, dos efectos antagónicos según se comporten.

Por un lado, cómo evoluciona en nivel de vida de la población. Si los salarios se incrementan en mayor proporción que los bienes y servicios, consiguientemente se mejora el bienestar además del incremento de la producción ocasionando mayor demanda.

Por otro, como se demostrará en el desarrollo del presente. Si los indicadores de incremento de bienes y servicios crecen a mayor ritmo que los salarios, disminuye el bienestar de la masa asalariada.

Como resultado de este proceso inflacionario y de baja demanda interna hay que mencionar que esto perjudica a los sectores productivos internos.

Las causas que originan el fenómeno de inflación no serán tratadas por lo complejo y extenso que es el tema.

1 Introducción

Las series de tiempo son el resultado de observar valores de una variable aleatoria a lo largo del tiempo en intervalos iguales o desiguales. En el análisis de las series temporales se utilizan métodos que permiten interpretarlas y extraer información sobre las relaciones entre los datos de la serie u otras series.

Según Mochón and Beker (2008), *“La inflación es el crecimiento generalizado y continuo de los precios de los bienes y los servicios”* (p. 496), y el índice de precios al consumidor (IPC) representa el costo de la canasta de bienes y servicios consumidos por la economía local.

El fenómeno de la inflación se produce por múltiples causas. Las teorías tradicionales se pueden resumir en: inflación de demanda, que se produce cuando los gastos superan la capacidad de producción de la economía, por causas monetarias, cuando se incrementa la emisión monetaria, la inflación de costos, cuando los precios suben por grupos económicos de presión y la inflación estructural, que es la que se origina debido a los problemas estructurales de las economías en vías de desarrollo.

La percepción individual sugiere que los motivos, en la coyuntura actual, son variados: especulación, emisión desmedida, presión por devaluación, bajas exportaciones, monopolios formadores de precios, elevado gasto público, mercado laboral altamente regulado, etc.

Analizar como evolucionan conjuntamente en el tiempo el Índice de precios al consumidor y el Índice de salarios puede ayudar a clarificar los problemas y soluciones posibles.

Además, esta comparativa muestra cómo se comporta el poder adquisitivo de un empleado. Es posible determinar si se recompone conforme aumentan los precios, acompaña o se abre una brecha que da cuenta de su disminución.

Índice de Precios al Consumidor versus Índice de Salarios acumulados

Con base Diciembre 2016

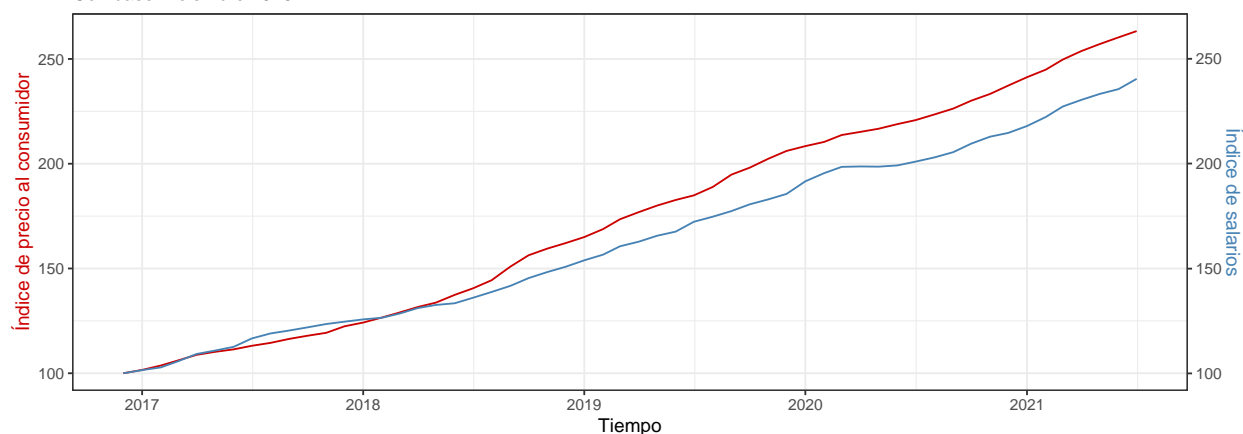


Fig. 1: Inflación versus Salarios

La información fue extraída de la página oficial del Instituto Nacional de Estadística y Censos - INDEC Argentina:

- Índice de precios al consumidor
- Índice de salarios

Se considera la variación mensual acumulada estableciendo a Diciembre 2016 como el 100% y se pretende predecir los cambios de las series para Enero 2022.

2 Marco Teórico

Los modelos autorregresivos y de medias móviles son validos para series de tiempo que son *estacionarias*. Se define que una serie es estacionaria de orden 2 cuando sus segundos momentos son constantes a lo largo del tiempo. Es decir, cuando su media y varianza se mantienen constantes.

Por lo contrario, una serie es *no estacionaria* cuando la tendencia y/o la variabilidad cambian a lo largo del tiempo.

El presente trabajo práctico tiene como objetivo el análisis de la estacionariedad de las series presentadas y su modelado mediante modelos SARIMA.

2.1 Tests de estacionariedad

La estacionariedad puede ser puesta a prueba mediante diferentes tests. Se detallan los dos a utilizar.

2.1.1 Dickey-Fuller Aumentado

En la ecuación de Dickey-Fuller aumentada (1):

$$\nabla Y_t = a_0 + a_1 t + (\phi_1 - 1)Y_{t-1} + \epsilon_t \quad (1)$$

Se pone a prueba el coeficiente que acompaña al rezago del período anterior: $(\phi_1 - 1)$.

El estadístico de este test es:

$$\tau^e = \frac{\hat{\phi}_1 - \phi_1}{\sigma(\hat{\phi}_1)} \quad (2)$$

Se plantea la regla de decisión:

Si $\tau \leq \tau^e < 0 \Rightarrow$ No Rechazar H_0

Si $\tau^e < \tau \Rightarrow$ Rechazar H_0

Como se explica en Peña Sanchez de Rivera (2010), la hipótesis nula establece que la raíz mayor de un $AR(p+1)$ es igual a uno, por lo que el proceso es no estacionario.

2.1.2 Kwiatkowski-Phillips-Schmidt-Shin

Kwiatkowski et al. (1992) definen otra prueba para raíces unitarias en la cual la hipótesis nula es contraria al test ADF, es decir, plantea que la serie temporal observable es estacionaria.

2.2 Modelos SARIMA

Los modelos *ARIMA* no logran captar los efectos estacionales de una serie mientras que los *SARIMA* captan el comportamiento puramente estacional de una serie.

La combinación de los modelos *ARMA*(p, d) estacionales y no estacionales se expresa como *ARMA*(p, q)(P, Q) s donde:

- P es el componente autorregresivo estacional
- Q el componente de media móvil estacional
- s representa el período estacional

Por lo que se llega a la ecuación (3).

$$\Theta_P(B^s)\phi_p(B)Y_t = \Theta_Q(B^s)\theta_q(B)\epsilon_t \quad (3)$$

Al agregar la parte integrada a la ecuación (3), se llega a la expresión del modelo *SARIMA*.

$$\Theta_P(B^s)\phi_p(B)Y_t \nabla_s^D \nabla^d = \phi_0 + \Theta_Q(B^s)\theta_q(B)\epsilon_t \quad (4)$$

3 Análisis de Resultados

3.1 Índice de Precios al Consumidor

3.1.1 Analisis de la serie

En la Figura 2, se observa que la serie de IPC es *no estacionaria* ya que su media varia a lo largo del tiempo y su función de autocorrelación decrece en forma lineal lentamente.

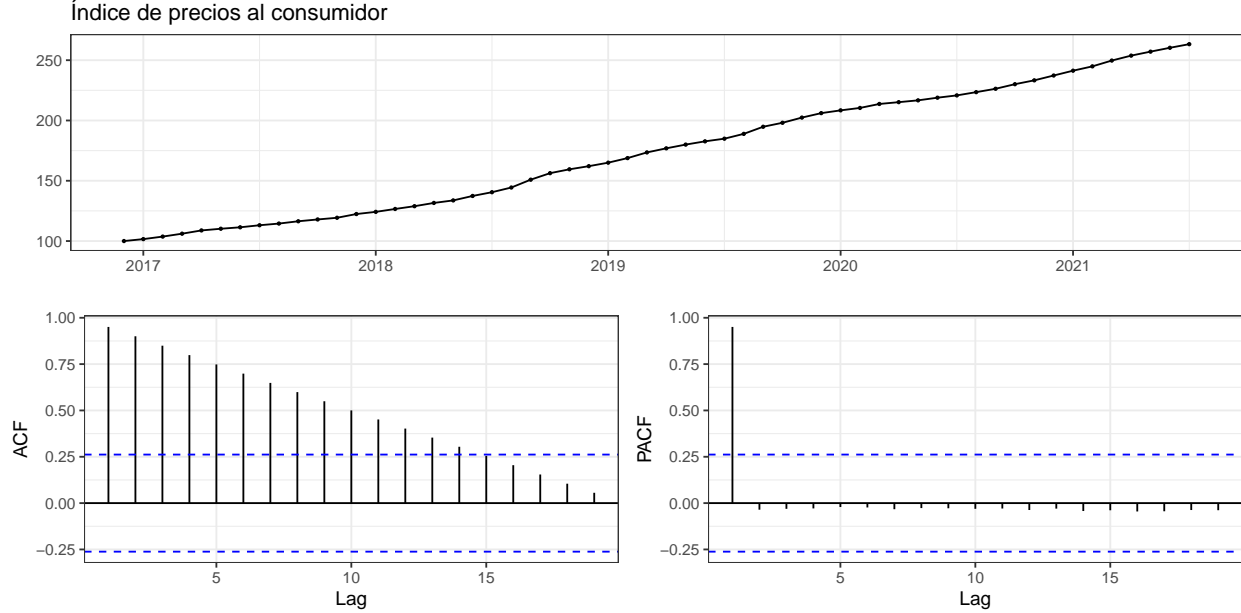


Fig. 2: Índice de precios al consumidor

Es posible someter la evidencia gráfica a distintas pruebas estadísticas para ratificar o rectificar lo observado.

Test de Estacionariedad: ADF

En la Tabla 1 se observa τ^e entre el valor crítico y cero, lo que determina que no hay evidencia para rechazar H_0 (serie no estacionaria).

	Estadístico	1pct	5pct	10pct
tau3	-2.469	-4.04	-3.45	-3.15
phi2	7.057	6.50	4.88	4.16
phi3	3.819	8.73	6.49	5.47

Tabla 1: Prueba de ADF para IPC

Test de Estacionariedad: KPSS

Kwiatkowski et al. (1992) definen otra prueba en la cual la hipótesis nula determina que una serie temporal observable es estacionaria. Los resultados de la misma se encuentran detallados en la Tabla 2.

Estadístico	P-Value	Lag
1.5	0.01	3

Tabla 2: Test KPSS para IPC

En este caso el p-value es menor al grado de significación 0.05, por lo que se rechaza H_0 .

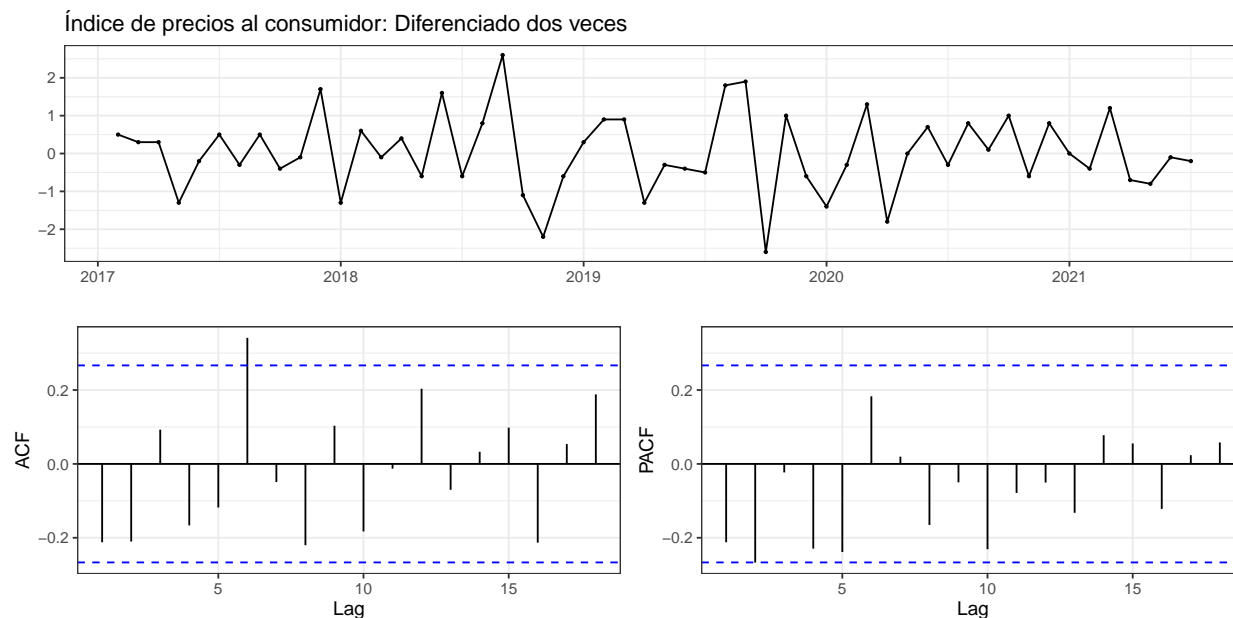
En conclusión, ambos tests concluyen que la serie de IPC es *no estacionaria*.

3.1.2 Diferenciación

Como se puede asumir que la tendencia de una serie en el instante t es muy próxima a la del instante $t - 1$, se puede formular una nueva serie de la manera $Y_t = y_t - y_{t-1}$. Este proceso se define como *diferenciación* y permite transformar una serie no estacionaria a estacionaria.

Mediante la función `ndiffs` de R, se estima que el numero de diferenciaciones necesarias para convertir a la serie IPC en estacionaria (2 veces).

La Figura 3 muestra el índice de precios al consumidor diferenciado.



Test de Estacionariedad: ADF

Se corrobora que la serie de IPC diferenciada es estacionaria mediante el test Dickey-Fuller aumentado (Tabla 3). El valor de τ^e se encuentra dentro de la zona de rechazo, por lo que se concluye que se rechaza H_0 .

	Estadístico	1pct	5pct	10pct
tau3	-7.139	-4.04	-3.45	-3.15
phi2	16.994	6.50	4.88	4.16
phi3	25.486	8.73	6.49	5.47

Tabla 3: Test de ADF para IPC diferenciada

Test de Estacionariedad: KPSS

El test KPSS coincide con ADF. El p-value es mayor al grado de significación de 0.05, por lo que no se rechaza la hipótesis nula de estacionariedad.

Estadístico	P-Value	Lag
0.055	0.1	3

Tabla 4: Test KPSS para IPC diferenciada

Se confirma entonces que la serie diferenciada es *estacionaria*.

3.1.3 Modelado

En el trabajo practico anterior, se obtuvo un modelo $ARIMA(0, 2, 6)$ para modelar la serie IPC (ver summary en Apendice). Si bien la serie no parece presentar algún componente de estacionalidad, se propone un modelo $SARIMA(0, 2, 0)(0, 2, 6)[12]$ (ver summary en Apendice) para compararlo con el modelo anterior.

3.1.4 Métricas de performance

Se comparan ambos modelos mediante AIC y BIC en la Tabla 5.

Modelo	AIC	BIC
ARIMA(0,2,6)	154.470	158.448
SARIMA(0,2,0)(0,2,6)[12]	144.063	148.267

Tabla 5: Metricas modelos propuestos para IPC

Parece que el modelo $SARIMA(0, 2, 0)(0, 2, 6)[12]$ es mejor al $ARIMA(0, 2, 6)$ propuesto en el trabajo anterior. Es necesario realizar un análisis de los residuos del modelo para verificar que presenta un buen ajuste.

3.1.5 Análisis de diagnóstico

Si bien el modelo SARIMA presenta mejores métricas de información que el modelo baseline del trabajo practico anterior, el modelo no ajusta de manera correcta los datos. Al graficar los residuos en la Figura 4, se puede observar que los residuos presentan autocorrelacion y no se distribuyen normalmente.

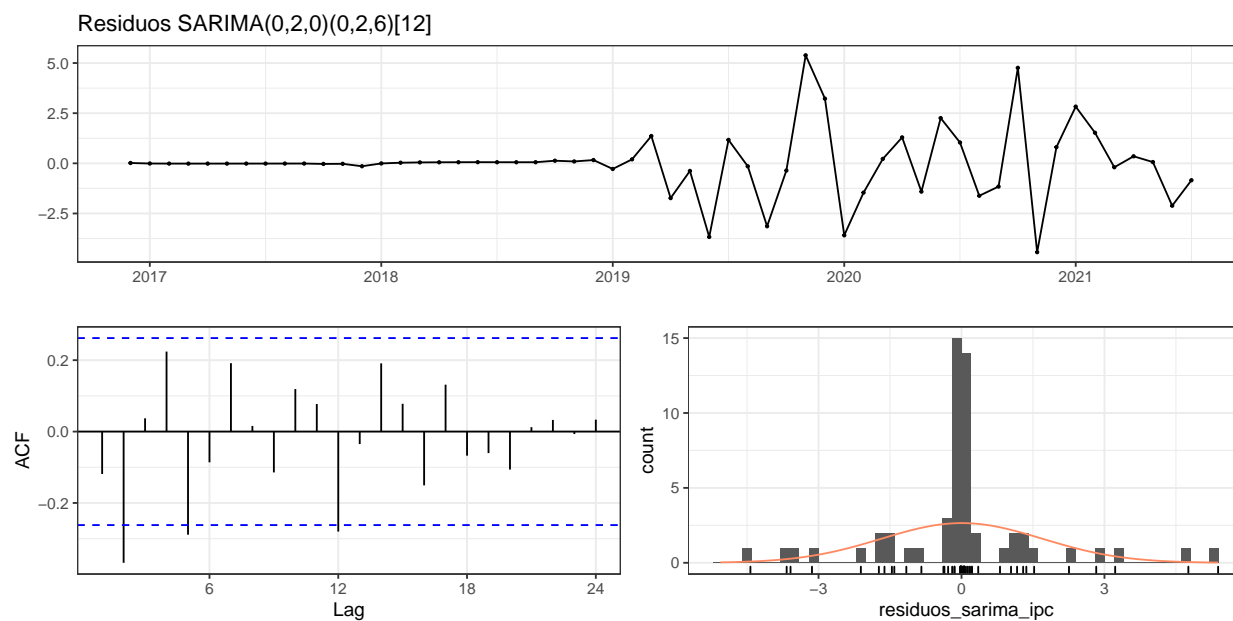


Fig. 4: Residuos modelo SARIMA

3.2 Índice de Salarios

La Figura 5 muestra la serie de Índice de Salarios.

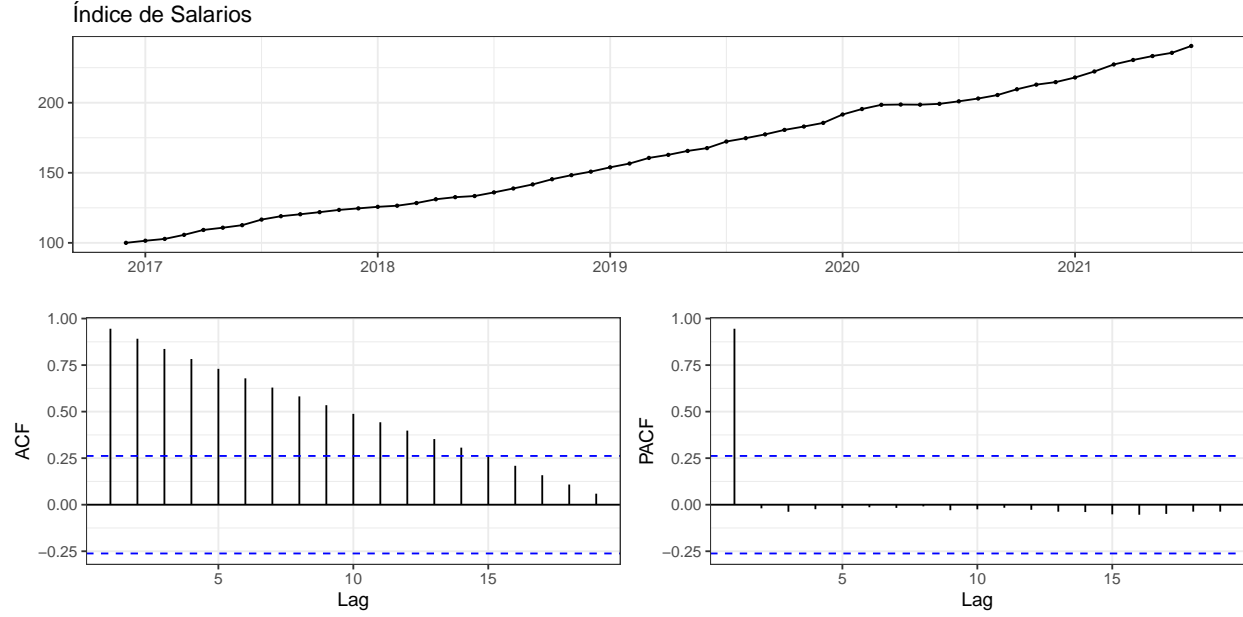


Fig. 5: Índice de Salarios

3.2.1 Análisis de la serie

Test de Estacionariedad: ADF

En la Tabla 6 se observa τ^e entre el valor crítico y cero, lo que determina que no hay evidencia para rechazar H_0 .

	Estadístico	1pct	5pct	10pct
tau3	-1.449	-4.04	-3.45	-3.15
phi2	7.416	6.50	4.88	4.16
phi3	2.343	8.73	6.49	5.47

Tabla 6: Prueba de ADF para Índice de Salarios

Test de Estacionariedad: KPSS

Los resultados de la prueba se encuentran detallados en la Tabla 7.

Estadístico	P-Value	Lag
1.5	0.01	3

Tabla 7: Test KPSS para Índice de Salarios

Ambos tests concluyen que la serie de Índice de Salarios es *no estacionaria*.

3.2.2 Diferenciación

En la Figura 5 se observa que, al igual que la serie de IPC, el Índice de Salarios tampoco es estacionario. Su media varía a lo largo del tiempo y su función de autocorrelación también decrece lentamente en forma lineal.

Mediante la función `ndiffs` de R, se estima que el número de diferenciaciones necesarias para convertir la serie en estacionaria. El resultado obtenido es 1.

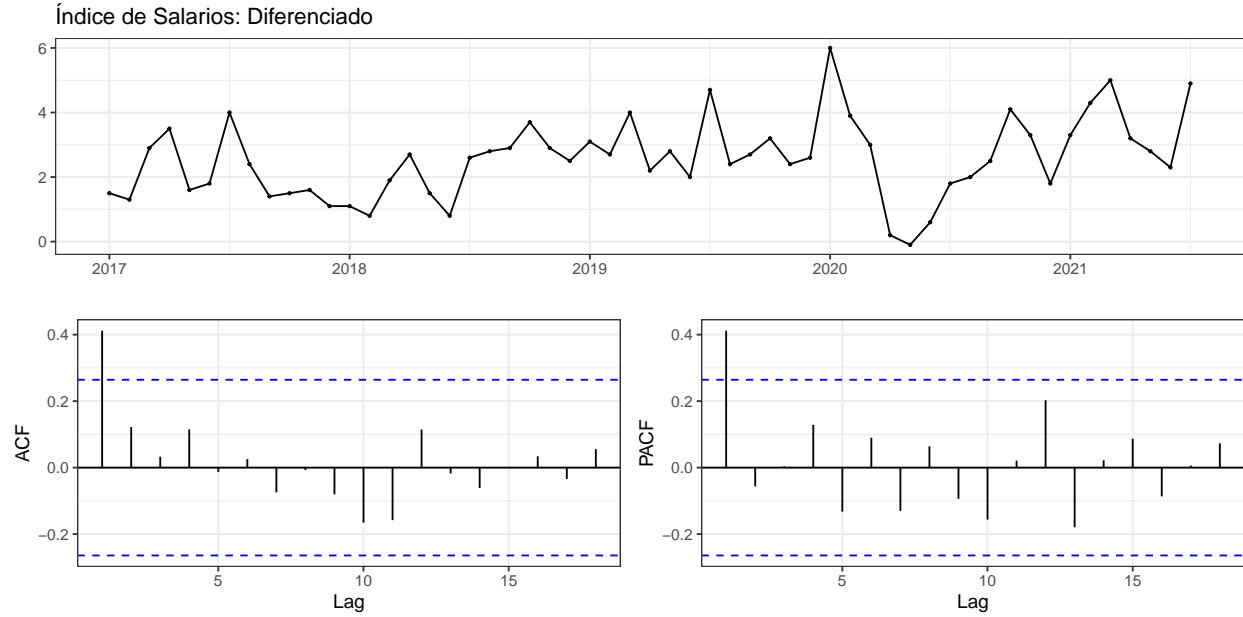


Fig. 6: Índice de Salarios diferenciado

Test de Estacionariedad: ADF

Al aplicar el test ADF a la serie diferenciada, el valor de τ^e se encuentra en la zona de rechazo de H_0 , por lo que se confirma que la serie diferenciada es estacionaria.

	Estadístico	1pct	5pct	10pct
tau3	-4.341	-4.04	-3.45	-3.15
phi2	6.361	6.50	4.88	4.16
phi3	9.435	8.73	6.49	5.47

Tabla 8: Prueba de ADF para Índice de Salarios diferenciado

Test de Estacionariedad: KPSS

El test KPSS sobre la serie diferenciada concuerda con el test ADF, concluyendo que la serie diferenciada es *estacionaria*.

Estadístico	P-Value	Lag
0.367	0.091	3

Tabla 9: Test KPSS para Índice de Salarios diferenciado

3.2.3 Modelado

En este trabajo practico no se propone otro modelo ya que en el anterior el seleccionado fue un $SARIMA(1,0,0)(1,0,0)[12]$ cuyo summary se encuentra en el Apendice.

4 Conclusiones

TODO:

5 Referencias bibliográficas

- Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. 1992. “Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” *Journal of Econometrics* 54 (1): 159–78.
- Mochón, F., and V. Beker. 2008. *Economía, Principios y Aplicaciones*. McGraw Hill.
- Peña Sanchez de Rivera, D. 2010. *Análisis de Series Temporales*. Alianza Editorial S.A.

6 Apéndices

6.1 Modelo $ARIMA(0, 2, 6)$ para Índice de Precios al Consumidor

```
##
## Call:
## arima(x = ts_indice_precios_consumidor, order = c(0, 2, 6), fixed = c(0, 0,
##      0, 0, 0, NA))
##
## Coefficients:
##      ma1  ma2  ma3  ma4  ma5      ma6
##      0    0    0    0    0  0.3161
## s.e.    0    0    0    0    0  0.1356
##
## sigma^2 estimated as 0.9389:  log likelihood = -75.24,  aic = 154.47
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 0.01006946 0.951667 0.7470765 0.01516551 0.4501471 0.2516179
##              ACF1
## Training set -0.2047803
```

6.2 Modelo $SARIMA(0, 2, 0)(0, 2, 6)[12]$ para Índice de Precios al Consumidor

```
##
## Call:
## arima(x = ts_indice_precios_consumidor, order = c(0, 2, 6), seasonal = list(order = c(0,
##      2, 6), period = 12), fixed = c(0, 0, 0, 0, 0, NA, 0, 0, 0, 0, 0, NA))
##
## Coefficients:
##      ma1  ma2  ma3  ma4  ma5      ma6  sma1  sma2  sma3  sma4  sma5  sma6
##      0    0    0    0    0  0.2624    0    0    0    0    0    0
## s.e.    0    0    0    0    0  0.3439    0    0    0    0    0  NaN
##
## sigma^2 estimated as 5.226:  log likelihood = -69.03,  aic = 144.06
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 0.007990367 1.673686 0.9665561 -0.0006645426 0.4608221 0.3255394
##              ACF1
## Training set -0.118806
```

6.3 Modelo $SARIMA(1, 0, 0)(1, 0, 0)[12]$ para Índice de Salarios

```
## Series: ts_indice_salarios
## ARIMA(1,1,0)(1,0,0)[12] with drift
##
## Coefficients:
##      ar1      sar1      drift
##      0.5295  0.3589  2.5043
## s.e.  0.1207  0.1385  0.4065
##
## sigma^2 estimated as 1.079:  log likelihood=-75.21
## AIC=158.42  AICc=159.28  BIC=166.23
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 0.03277458 0.9989346 0.7664073 -0.001050329 0.493962 0.02621617
```

```
## ACF1
## Training set 0.06995948
```