



## **Análisis de series temporales**

*Trabajo Práctico II – Grupo 5*

Integrantes: Alberto Gir Erdmann, Agostina Misaña, Ciro Couste,  
Carla Menzella y Arnaldo Viera.

## Resumen Ejecutivo

El presente trabajo tiene como objetivo poner en práctica los conocimientos adquiridos en la materia Análisis de Series temporales de la Maestría en Ciencia de Datos de la Universidad Austral.

Para ello analizaremos dos series de temporales:

- 1) Cantidad de noches que los viajeros no residentes pernoctan en un complejo hotelero dentro de Argentina.
- 2) Variación porcentual del índice de precios al consumidor de la república Argentina.

A diferencia del primer práctico decidimos utilizar la segunda serie en variación porcentual y no en niveles porque pensamos que más allá del nivel de precios, lo importante es analizar la inflación.

Y debido al aprendizaje realizado en el primer práctico, pudimos observar que a medida que la serie incrementaba el valor de los niveles de precios se incrementaba el valor de la variación. Dicho dato pierde referencia si no lo comparamos en términos porcentuales. Ejemplificando una inflación mensual del 2% es comparable en términos de impacto económica con otro punto en la serie donde haya dado un valor similar, aunque si partimos de niveles de precios diferentes la variación en términos absolutos también será diferente.

## Índice de Contenido

Resumen Ejecutivo .....	2
Introducción.....	4
Marco Teórico .....	5
Análisis de Resultados .....	8
<b>Serie de tiempo 1</b> .....	8
1) Análisis exploratorio de la serie .....	9
2) Análisis de la estacionariedad.....	9
3) Cambios estructurales .....	11
4) Modelos .....	13
5) Análisis de residuos del modelo.....	14
6) Predicción .....	14
<b>Serie de Tiempo 2</b> .....	16
1) Analisis exploratorio .....	16
2) Análisis de la estacionariedad:.....	17
3) Cambios estructurales .....	18
4) Modelos .....	19
5) Análisis de los residuos.....	20
6) Predicción .....	20
Referencias bibliográficas .....	22

## Introducción

Analizaremos las dos series por separado. El procedimiento se describe a continuación:

### a) Identificación

Representar gráficamente la serie, además de su función de autocorrelación simple (ACF) y función de autocorrelación parcial (PACF).

La gráfica de la serie nos indica si la serie es estacionaria o no.

Según los motivos por los que la serie no es estacionaria, tendremos que aplicar los siguientes procedimientos hasta hacerla estacionaria.

- Si tiene tendencia: determinamos si tiene tendencia determinística o estocástica.
- Si la serie tiene estacionalidad: Tomaremos diferencias estacionales hasta que desaparezca el patrón estacional.
- Si es heterocedástica, es decir, no tiene varianza constante, habrá que transformar la serie.

### b) Estimación y verificación.

Luego representaremos las mismas a través de modelos ARIMA o SARIMA según corresponda.

Observando las dos gráficas del ACF y PACF de la serie transformada podemos hacernos una idea del modelo que describe nuestra serie, o al menos de cuáles son los primeros candidatos que debemos probar.

Para comprobar analíticamente (no visualmente) un modelo frecuentemente se ajusta varios modelos candidatos  $ARIMA(p,d,q)$  y escogeremos como un buen modelo aquel que tenga los residuales semejantes al de un ruido blanco, además que tenga los valores del AIC (Criterio de Información de Akaike) y BIC (Criterio de Información Bayesiana) menores con relación al resto de los modelos candidatos.

### c) Predicción.

Una de las razones de la popularidad del proceso de construcción de modelos es su éxito en la predicción. Los modelos son buenos para realizar predicciones a corto plazo. Una vez representadas estudiaremos el comportamiento de los residuos de los modelos para determinar si se comportan como un ruido blanco. O si por el contrario precisan transformaciones.

### d) Comparación de los resultados con los modelos determinados en el trabajo anterior.

Una vez obtenidos los Criterios de Información de los modelos de este trabajo práctico (a partir de la función accuracy), los compararemos con los resultados del trabajo práctico anterior, para determinar cuál modelo aplicado tiene una mejor probabilidad de ajuste.

## Marco Teórico

Un proceso puede ser no estacionario:

- en la media: caso donde tiene una tendencia creciente o decreciente. No es estable en el tiempo.
- en la varianza
- en las autocorrelaciones.
- en otras características de la distribución de las variables.

### 1) Procesos no estacionarios en **varianza**:

En estos casos se debe aplicar transformaciones que conviertan al proceso original en un proceso condicionalmente homocedástico. Las técnicas relacionadas con la teoría de los modelos lineales han sido desarrolladas mediante los siguientes supuestos:

- La estructura de la Esperanza Matemática Condicionada es aditiva.
- La varianza de las variables que componen el proceso es constante.
- Las variables del proceso están normalmente distribuidas.

Existen distintas transformaciones que se pueden aplicar, las cuales todas dependen de un mismo parámetro el cual debe ser determinado empíricamente a partir de la estimación del coeficiente  $\lambda$ .

Para la estimación de los coeficientes  $\lambda$  en las transformaciones dirigidas a la eliminación de la heterocedasticidad se suele emplear la Relación Rango-Mediana que se obtiene de dividir la serie de tiempo original en  $k$  subseries de  $n_k$  y calcular para cada una de ellas el Rango y la Mediana.

Distintos tipos de transformaciones:

- Box-Cox (diferentes variaciones en diferentes niveles de la serie)
- Tukey.
- Logarítmica, es la más ampliamente aplicada para reducir la simetría de la distribución (en particular series económicas).
- Wallis (1987).

### 2) Procesos no Estacionarios en la **Media**:

Una propiedad importante que diferencia a los procesos integrados de los estacionarios es la forma en que desaparece la dependencia con el tiempo.

En los procesos estacionarios ARMA las autocorrelaciones disminuyen geométricamente, y se hacen prácticamente cero a los pocos rezagos.

En los procesos integrados las autocorrelaciones disminuyen linealmente con el tiempo y es posible encontrar coeficientes distintos de cero hasta para rezagos muy altos.

Existe una clase de procesos estacionarios donde las autocorrelaciones disminuyen más lentamente con el tiempo que en el caso de los procesos ARMA, y también que en los procesos integrados (procesos de memoria larga).

Estos procesos son casos particulares de la representación de Wold para procesos estacionarios, pero donde los coeficientes decrecen muy lentamente, lo que implica que las autocorrelaciones entre las observaciones también decrecen muy lentamente.

Al generalizar se dirá que un proceso es integrado de orden  $d \geq 0$  y se representa como  $I(d)$  cuando al diferenciarlo  $d$  veces se obtiene un proceso estacionario.

Un proceso estacionario es, por tanto, siempre  $I(0)$ . En la práctica la mayoría de las series no estacionarias que son integradas tienen un orden  $d \leq 3$ .

### Modelos ARIMA(p, d, q)

Se puede generalizar para cualquier proceso ARMA, permitiendo que una o varias raíces el operador AR sean la unidad. Se obtienen procesos del tipo,

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) (1 - B)^d Y_t = \phi_0 + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t$$

$$\Phi(B) (1 - B)^d Y_t = \phi_0 + \Theta(B) \epsilon_t$$

Donde:

- $p$  es el orden la parte autorregresiva estacionaria
- $d$  es el número de raíces unitarias (orden de integración del proceso)
- $q$  es el orden de la parte media móvil

La persistencia de valores positivos en el correlograma (aunque sean pequeños) y de decrecimiento lineal caracteriza a los procesos no estacionarios.

El procedimiento a seguir es el siguiente:

- Detectar Raíces Unitarias
- Construir FAC y analizar el decrecimiento

### Test de Raíces Unitarias

Ecuación de Dickey Fuller y se pone a prueba el coeficiente que acompaña al rezago del período anterior. Las hipótesis de esta prueba son,

$H_0 : \phi_1 - 1 = 0$  "No Estacionario"       $H_1 : \phi_1 - 1 < 0$  "Estacionario"

### Componentes Determinísticos

- No incluye término independiente ni lineal: ("None")  $\nabla Y_t = (\phi_1 - 1) Y_{t-1} + \epsilon_t$
- Incluya término independiente pero no incluye término lineal: ("Drift")
  - $\nabla Y_t = a_0 + (\phi_1 - 1) Y_{t-1} + \epsilon_t$
- Incluya ambos términos: ("Trend")  $\nabla Y_t = a_0 + a_1 t + (\phi_1 - 1) Y_{t-1} + \epsilon_t$

A esta última ecuación se denomina Ecuación de Dickey Fuller Aumentada.

Existen otras pruebas de hipótesis para poner a prueba la conjetura de Raíz Unitaria. Por ejemplo, Phillips-Perron KPSS Said-Dickey Zivot & Andrews.

### Modelos SARIMA(p,d,q)(P,D,Q)[s]

Un modelo de media móvil integrada autorregresiva estacional (SARIMA) es la incorporación de tendencias estacionales al modelo ARIMA.

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

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

Algunas series de tiempo pueden representar FAC que decaen lentamente en los rezagos múltiples del período estacional. Por lo tanto, el modelo SARIMA multiplicativo se puede expresar en una expresión similar al modelo ARIMA:

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

Un lento decaimiento en el gráfico de auto correlaciones (FAC) es una señal de que puede ser necesaria la diferenciación.

# Análisis de Resultados

## Serie de tiempo 1

En la serie de tiempo 1 de pernoctes no residentes desde el año 2017 hasta el 2021, se observa un componente de ciclos (seasonal), sin tendencia aparente hasta que decrece abruptamente al inicio de 2020. Dado este suceso extremo u outlier (inicio cuarentenas por COVID-19) se decide construir una nueva serie sin el año 2020: Pernoctes no residentes – serie acotada (Ilustración 1.1).

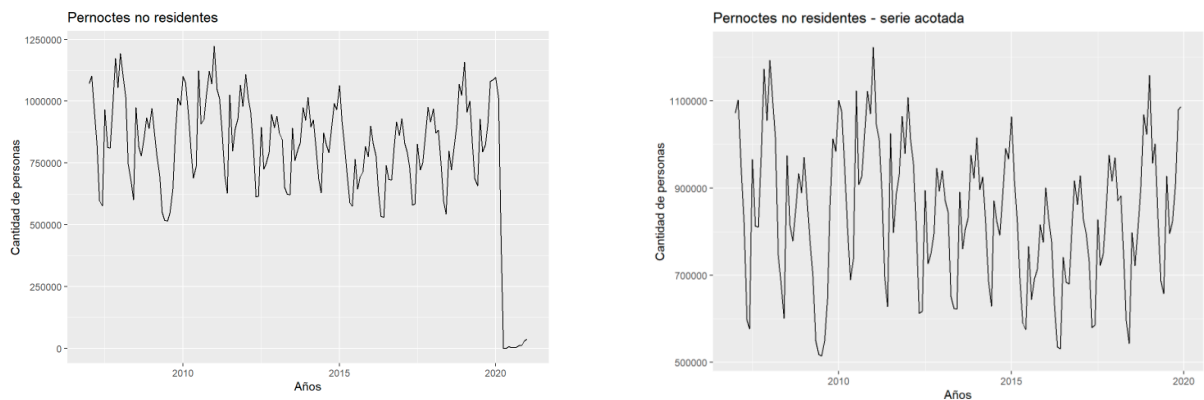


Ilustración 1.1

Se procede a graficar la descomposición con el rango la serie ajustado (Ilustración 1.2):

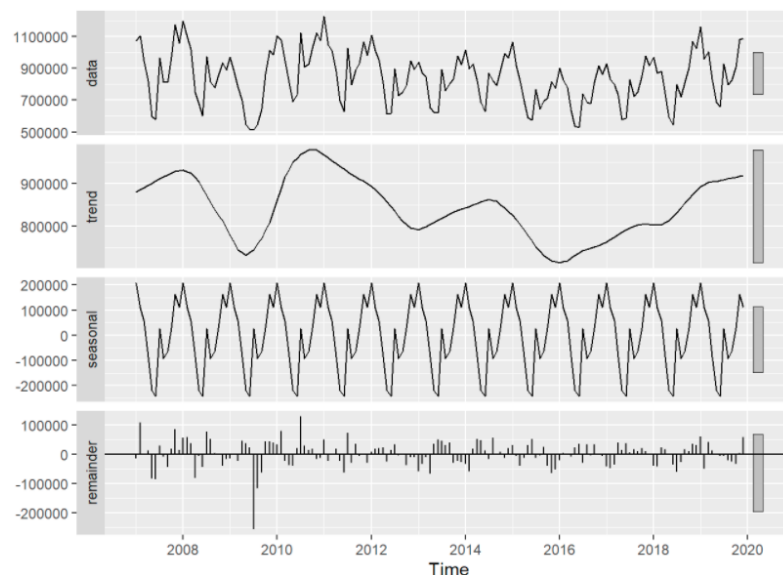


Ilustración 1.2

Se puede comprobar que no hay eventos o intervenciones relevantes, se identifican una estacionalidad y una tendencia variable.



## 1) Análisis exploratorio de la serie

En una primera aproximación se obtiene que la distribución de la serie se aproxima a la distribución normal a grandes rasgos (Tabla 1.1).

Tabla 1.1

vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis
1	156	843496.7	163645.9	838108	844112.1	174844.5	514856	1223545	708689	-0.03604553	-0.718023

De los datos se extrae que la media es levemente superior a la mediana. Lo cual se corrobora con los cuantiles.

Se verifican supuestos de normalidad por distintos métodos, que plantean que la variable evaluada se distribuye como si fuera una distribución normal:

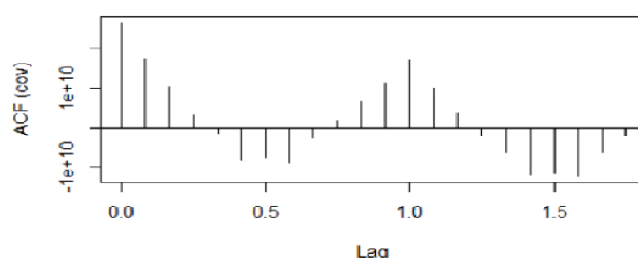
- El test de Jarque-Bera arroja un p-value = 0.2106.
- El test de Anderson-Darling arroja p-value = 0.3081.
- El test de Shapiro-Wilk arroja p-value = 0.1193.

En todos los casos el p-value es mayor a una significancia de 0.05. Esto nos indica que NO se rechaza la  $H_0$  de normalidad. Serie 1 cumple con supuesto de Normalidad.

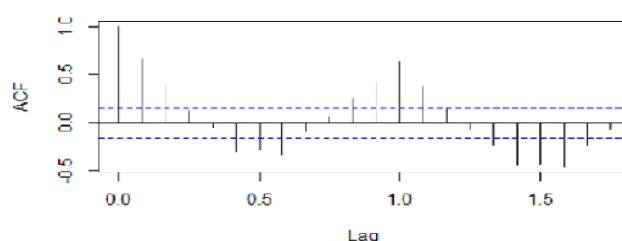
## 2) Análisis de la estacionariedad

Interpretación de los gráficos FAS, FAC y FACP (Ilustración 1.3)

- Grafica FAS: se observan lags que se destacan cíclicamente con cambio de signo.



- En la gráfica de FAC, se nota un descenso rápido de los coeficientes de correlación generando una forma sinusoidal con varios lags significativos.



- Grafico FACP: muestra lags significativos que indicarían un AR de órdenes 4, 6, 7, 10 y 13.

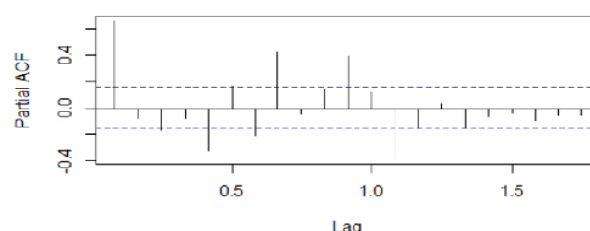


Ilustración 1.3

## 2.1 Análisis de estacionalidad en la varianza:

Se verifica la estacionalidad en la varianza y ver la necesidad de aplicar una transformación, a través del valor de lambda propuesta por la función "powerTransform" (Tabla 1.2):

Tabla 1.2

	LRT	df	pval
LR test, lambda = (0)	6.409315	1	0.011352
LR test, lambda = (1)	0.00574365	1	0.93959

- Lambda=0 → p-value es muy chico. (Se rechaza la H0)
- Lambda=1 → p-value grande.

La función arroja un valor de lambda de 0.9707, por lo tanto se define que no es necesaria una transformación.

Se decide tomar como referencia el resultado de Jarque-Bera de normalidad y no se transforma la serie original.

## 2.2 Análisis de la estacionalidad en la media:

- Test LJUNG – BOX: rechazo  $H_0$  significa que hay coeficientes de autocorrelación distintos a cero (Gráfico 1.4).

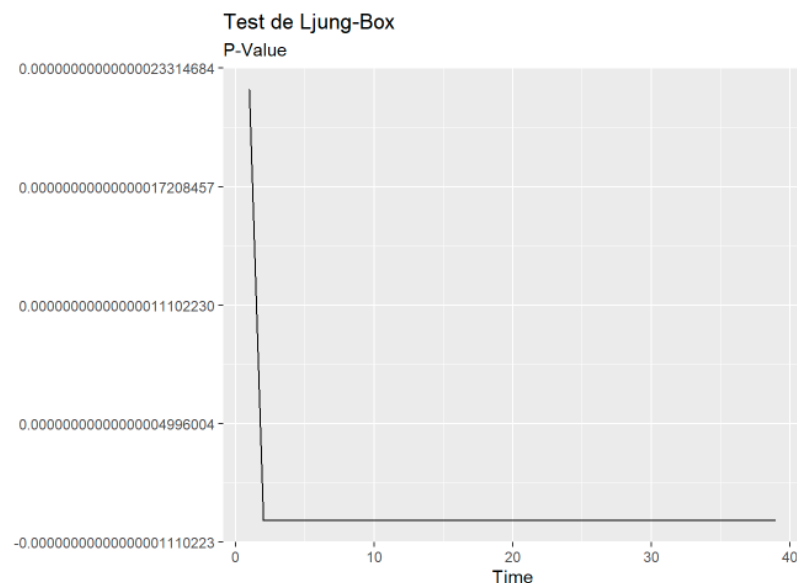


Ilustración 1.4

Al resultar p-values muy pequeños, se rechaza la  $H_0$ , por lo que hay coeficientes de autocorrelación distintos a cero.

Test de raíces unitarias: se ejecutan los test y se resumen en Tabla 1.3

- Dickey-Fuller: indica que no hay raíz unitaria. Estacionario.

Dickey-Fuller = -5.2349, Lag order = 5, p-value = 0.01. Alternative hypothesis: stationary.

- Se aplican las pruebas de Dickey-Fuller Aumentado con nivel de significancia de 0.05 con sus tres posibilidades:
  - TIPO 0– None(Sin constante ni tendencia): serie no es estacionaria.
  - TIPO 1 – Drift (con constante): serie es estacionaria.
  - TIPO 2 – Trend (tendencia y constante): serie es estacionaria con tendencia y drift (pendiente e intercepto).
- Ndiffs: utilizando los métodos “adf”, “pp”, y “kpss”, no precisa diferenciar para un nivel de significancia de 0.05.
- Test KPSS: no se rechaza  $H_0$ , lo que indica que la serie es estacionaria con tendencia determinística.
- Test de Zivot-Andrews: indica que la serie es estacionaria con cambio estructural.

TEST	SERIE ESTACIONARIA	SERIE NO ESTACIONARIA
Dickey-Fuller	X	
D-F Aumentado TIPO 0– None		X
D-F Aumentado TIPO 1 – Drift	X	
D-F Aumentado TIPO 2 – Trend	X	
Ndiffs	X	
KPSS	X	
Zivot-Andrews	X	

Tabla 1.3

El estadístico de test de la serie es menor al valor crítico (inclusive para  $\alpha = 0.01$ ), por lo que SE RECHAZA LA  $H_0$  de existencia de una raíz unitaria por cambio estructural. Esto indica que la serie es estacionaria con cambio estructural. El test Z-A plantea un quiebre en el mes 98 (Gráfico 1.5).

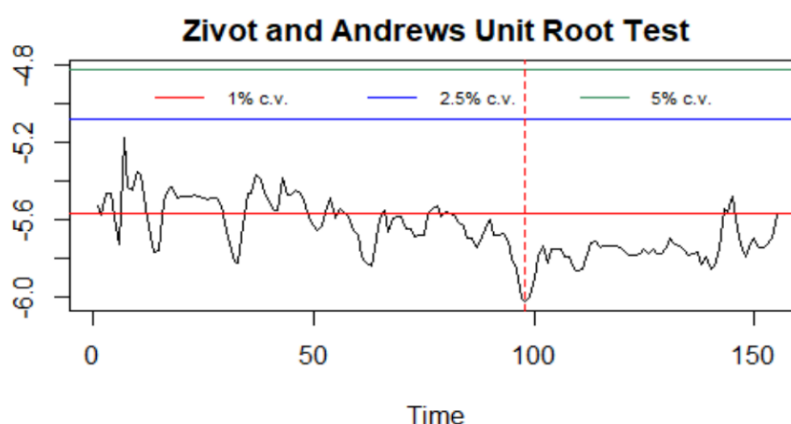


Ilustración 1.5

### 3) Cambios estructurales

### 3.1 Cambios estructurales con tendencia ( $SdT1sinout \sim timeline1$ )

A partir del hallazgo de estacionariedad con tendencia en test anterior, se busca determinar los cambios estructurales y las rectas de regresión lineal asociadas.

El p-value del intercepto indica que es significativo (RECHAZA  $H_0$ ), mientras que el p-value de la tendencia no sería significativo (NO SE RECHAZA LA  $H_0$ ), siendo entonces este coeficiente despreciable en la ecuación.

Se prueban varios modelos alternando las ventanas de tiempo, cuyos resultados siempre devuelven coeficientes no significativos.

### 3.2 Cambio estructural de nivel/drift ( $SdT1sinout \sim 1$ )

Se prueban otras ventanas de tiempo, cuyo resultado siempre devuelve coeficientes significativos:

Tabla 1.3

```
Coefficients:
      Estimate Std. Error t value      Pr(>|t|)
(Intercept)  843497     13102   64.38 <0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 163600 on 155 degrees of freedom

t test of coefficients:

      Estimate Std. Error t value      Pr(>|t|)
(Intercept)  843497     13102  64.378 < 0.00000000000000022 ***
```

Determinamos el punto de corte y luego los visualizamos. Identificamos fechas y coeficientes en los puntos de corte.

A partir de los  $m = 1$  a 9 puntos de ruptura con las fechas asociadas y las métricas {RSS, BIC}, el valor mínimo de BIC ocurre en  $m = 4$ . Trazamos la salida de la función breakpoint() para obtener una comprensión visual (Ilustración 1.6):

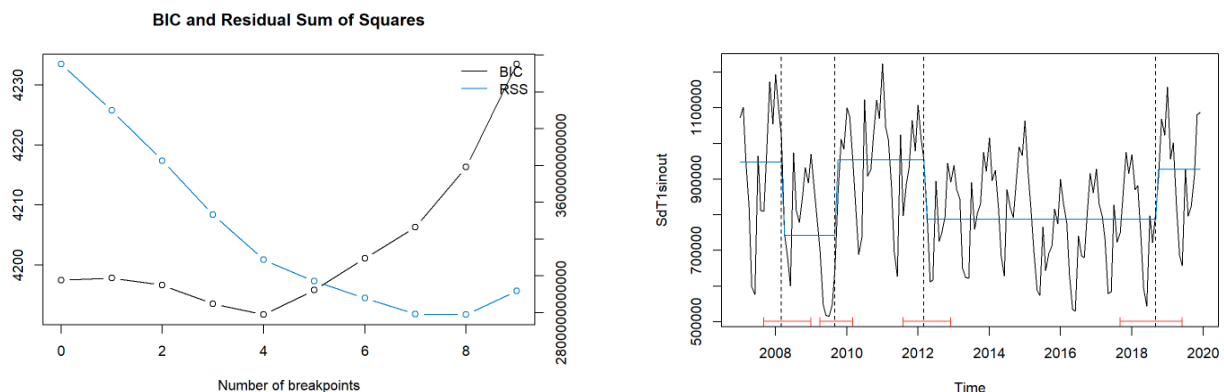
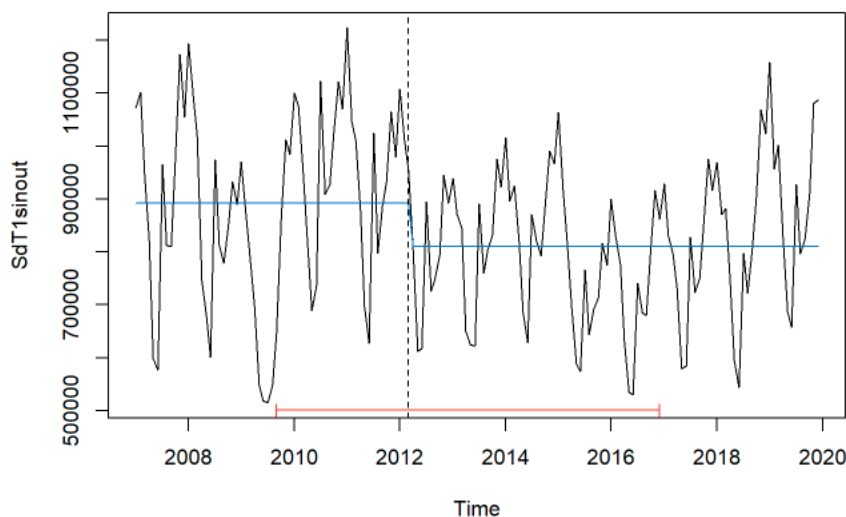


Ilustración 1.6

Se elige optar por un solo punto de quiebre para trabajar con mayor cantidad de datos. Los valores BIC para los primeros 4 breakpoints no son muy diferentes.

#### Serie con 1 break

Se generará el modelo de la serie temporal con el intervalo a partir de 04.2012, según el punto de quiebre encontrado (Gráfico 1.7).



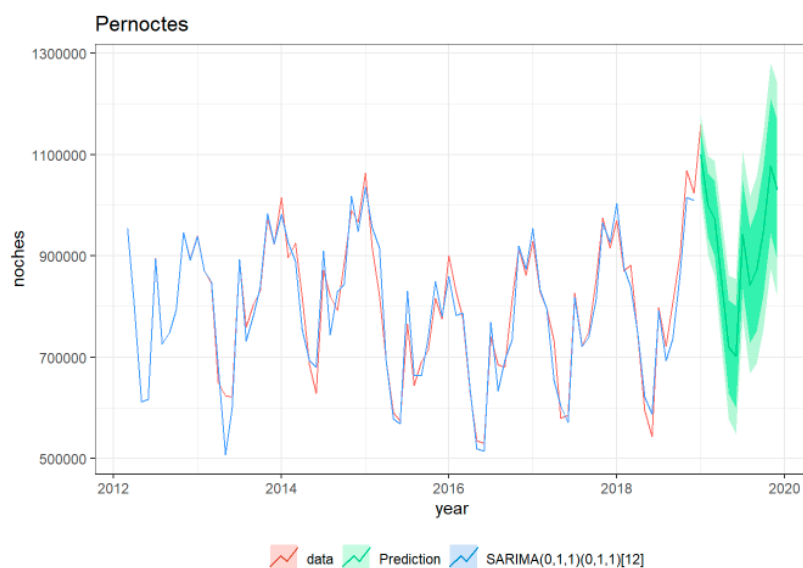
*Ilustración 1.7*

#### 4) Modelos

Se definen los subconjuntos de train y test para la serie final correspondiente al intervalo definido (04.2012 - 12.2019).

Se utiliza la función `auto.arima` para comprobar los modelos sugeridos por el algoritmo en el intervalo de serie acotado. Best model:  $ARIMA(0,1,1)(0,1,1)[12]$

Modelo 1 ajustado solo con datos Train: se generan los gráficos de serie original, valores ajustados en período Train y predicción en período Test (Gráfico 1.8):



*Ilustración 1.8*

Se generan los criterios de información del modelo para Train y Test (Tabla 1.4):

Tabla 1.4

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	5132.335	34338.69	24107.09	0.6115581	3.120256	0.3791619	-0.005373753	NA
Test set	-11224.144	39602.58	35655.60	-1.8119730	4.039798	0.5607993	0.022963250	0.2454371

## 5) Análisis de residuos del modelo

Se analizan los residuos del modelo 1 (Train) y se aplica Ljung-Box (Gráfico 1.9):

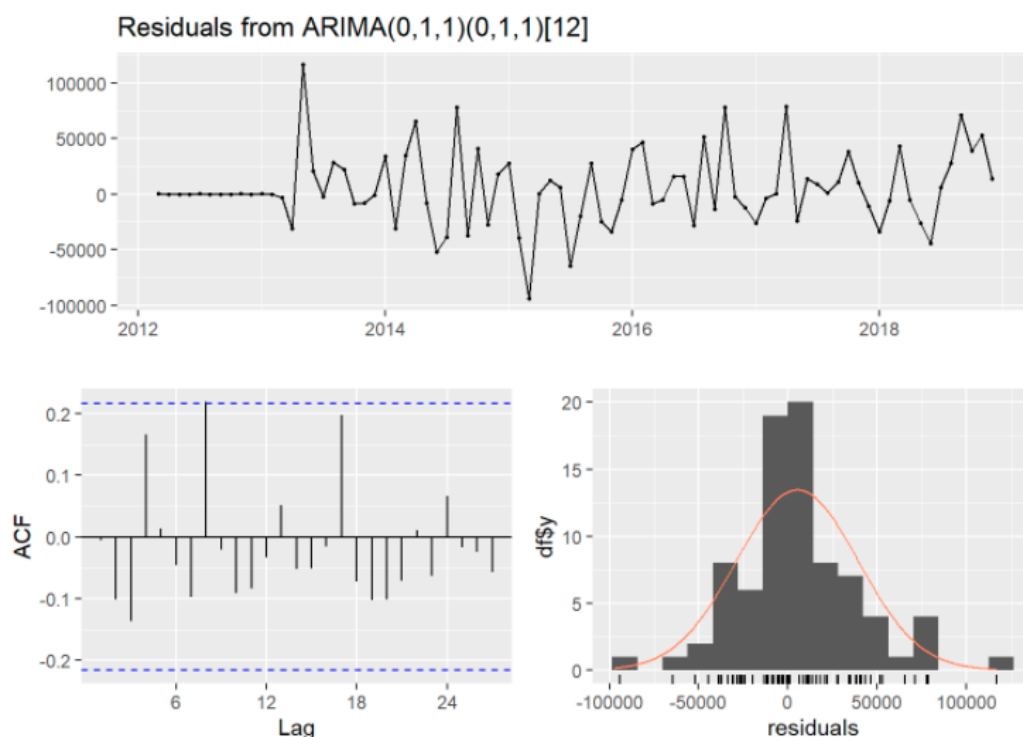


Ilustración 1.9

Resultan p-values mayores que nivel de significancia 0.05 (el menor encontrado es 0.084), por lo que NO se rechaza la  $H_0$  de Ljung-Box test ( $H_0$ : Los residuos están distribuidos independientemente). El gráfico de PAC solo muestra apenas significativo un lag (8), pero se lo considera incluido por el error del intervalo de confianza.

## 6) Predicción

Tomando la serie completa, se hace un pronóstico de los siguientes 12 meses a partir de 01-2020 (Gráfico 1.10):

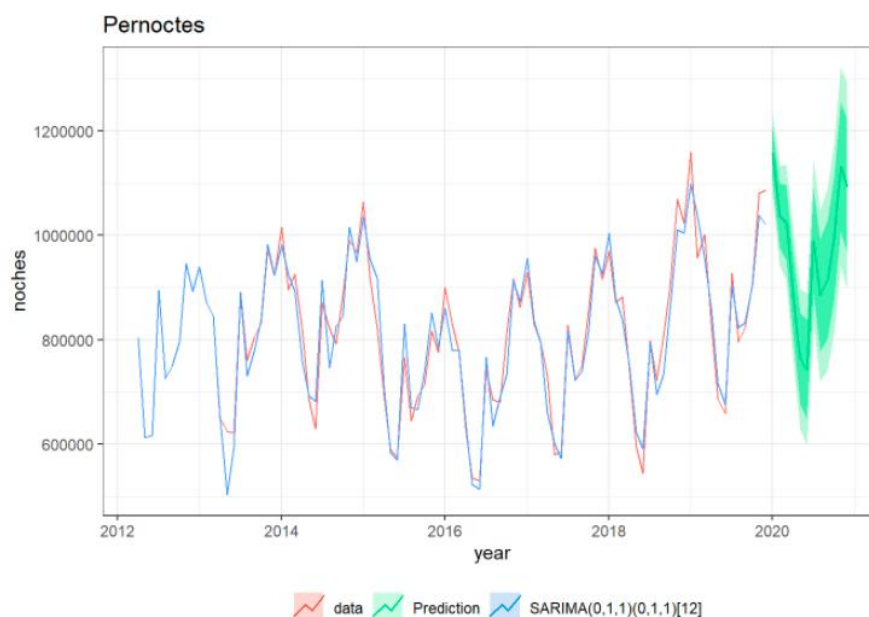


Ilustración 1.10

El modelo se construye y se utiliza para predecir en la circunstancia de no existencia del fenómeno COVID-19, que como se identificó al inicio del análisis, impacta a partir de los primeros meses de 2021.

Comparación de Criterios de Información entre modelo ARIMA de TP1 y el aplicado en este TP2 (SARIMA):

Se compara el Accuracy obtenido con el resultante del modelo ARIMA(2,1,1) aplicado en el trabajo práctico anterior (Tabla 1.5):

Tabla 1.5

Test	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
TP1 ARMA(2,1,1)	22092.9	70937.2	60256.4	1.9697	7.0854	0.5880	0.2736	1.0304
TP2 SARIMA(0,1,1) (0,1,1) [12]	-1224.1	39602.5	35655.6	-1.8119	4.0397	0.5607	0.0229	0.2454

A partir de estos valores, se concluye que el modelo aplicado a la serie 1 que considera la ESTACIONALIDAD en la serie temporal, obtuvo mejores valores (menores) en los Criterios de Información resultantes.

## Serie de Tiempo 2

### 1) Analisis exploratorio

“El Índice de Precios al Consumidor (NdA: IPC) es un indicador que mide la evolución promedio de los precios de un conjunto de bienes y servicios representativos del gasto de consumo de los hogares residentes en un área determinada” (INDEC, 2016).

En este caso se tomó la serie de tiempo del IPC desde enero 2017 hasta octubre 2021 pero transformando la serie original: se tomó la variación en puntos porcentuales en vez del valor absoluto.

La unidad es % y no AR\$. En el gráfico 2.1 siguiente se puede ver la evolución de la serie. Entre mediados de 2018 y mediados de 2019 se percibe un alza en las variaciones porcentuales de precios (IPC).

En la ilustración 2.1 de la derecha se puede ver la descomposición de la serie donde queda en evidencia el cambio de tendencia.

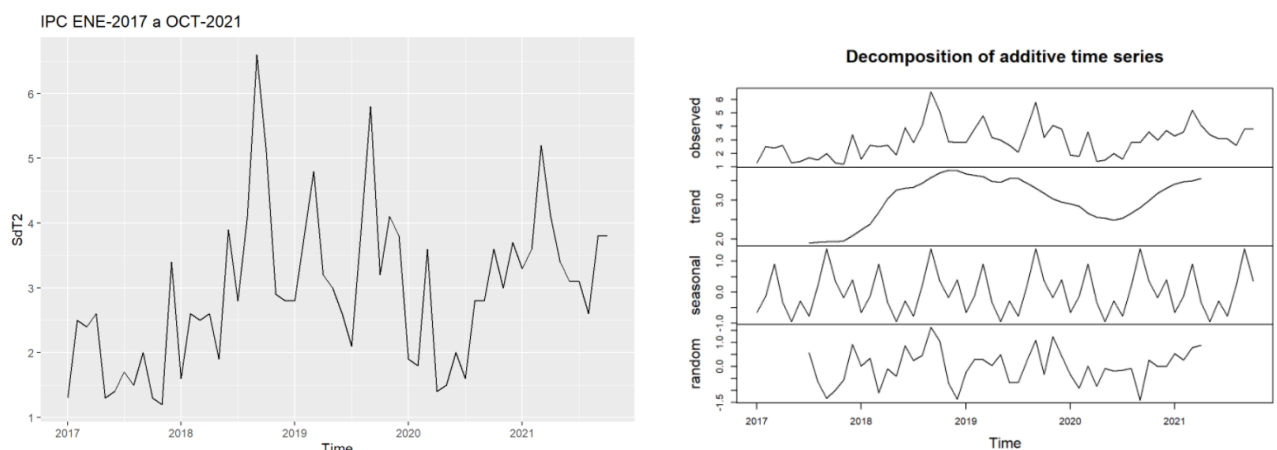


Ilustración 2.1

Se comprueba el supuesto de normalidad de la serie original:

- El test de Jarque-Bera arroja un p-value = 0.059.
- El test de Anderson-Darling arroja p-value = 0.1177.
- El test de Shapiro-Wilk arroja p-value = 0.018.

El valor de p-value es mayor al  $\alpha = 0.05$ , por lo que NO SE RECHAZA la  $H_0$  de normalidad de la serie data2 para Jarque-Bera y Anderson-Darling. Sin embargo, el test Shapiro-Wilk indica que se debe RECHAZAR la  $H_0$  de normalidad, por lo tanto se analiza la variancia de la serie antes de avanzar con el modelo.



## 2) Análisis de la estacionariedad:

### 2.1 Estacionariedad en la varianza

A priori, se observan períodos con variabilidad cambiante y cambios estructurales en la tendencia. También se observan tendencias variantes a lo largo de la serie. Los coeficientes de autocorrelación decrecen rápidamente en el gráfico de FAC (ACF).

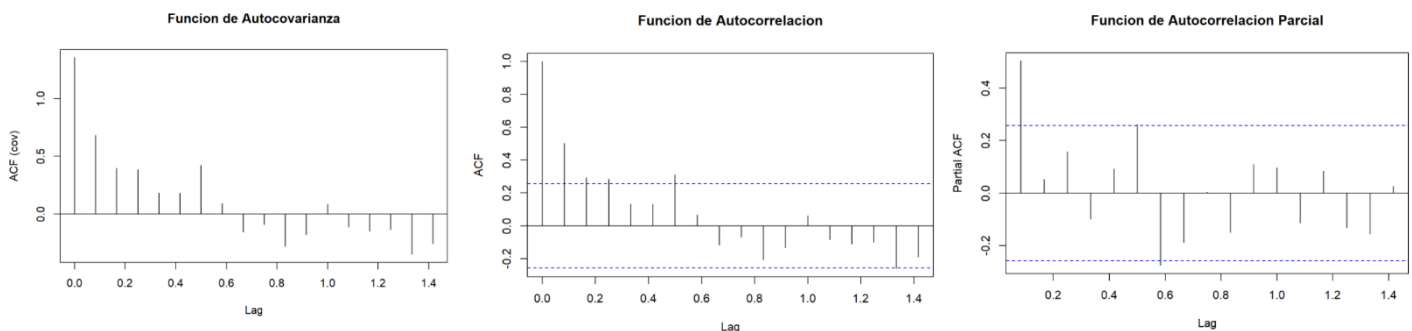


Ilustración 2.2

Al rechazar normalidad, se define la transformación de potencia adecuada, mediante PowerTransform.

Como el resultando de esta función, indica un p-value significativo para un  $\lambda=0$ , se aplica la transformación de potencia a la serie.

Luego se comprueba el supuesto de normalidad de la serie transformada:

- El test de Jarque-Bera arroja un p-value = 0.7199.
- El test de Anderson-Darling arroja p-value = 0.1735.
- El test de Shapiro-Wilk arroja p-value = 0.2195.

En este caso, la  $H_0$  de normalidad se confirma (NO SE RECHAZA) en todos los test de normalidad. Se construye ahora la serie de tiempo a partir de la variable transformada.

### 2.1. Estacionariedad en la media

- Test LJUNG – BOX: Todos los p-values dan muy bajos, por lo que SE RECHAZA la  $H_0$ , y la serie tiene coeficientes de autocorrelación distintos a cero.

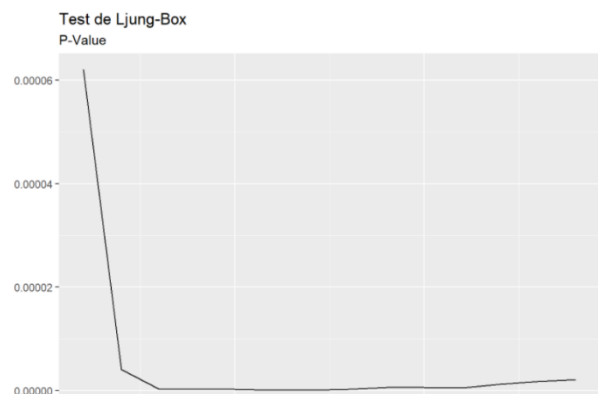


Ilustración 2.3

## Test de raíces unitarias

En la tabla 2.1 se observan los resultados arrojados de los distintos test probados.

Tabla 2.1

TEST	SERIE ESTACIONARIA	SERIE NO ESTACIONARIA
Dickey-Fuller		X
D-F Aumentado TIPO 0 – None		X
D-F Aumentado TIPO 1 – Drift	X	
D-F Aumentado TIPO 2 – Trend	X	
Ndiffs	X	
KPSS	X	
Zivot-Andrews	X	

Se plantea un quiebre en el mes 36, equivalente con dic-2019 (Ilustracion 2.4).

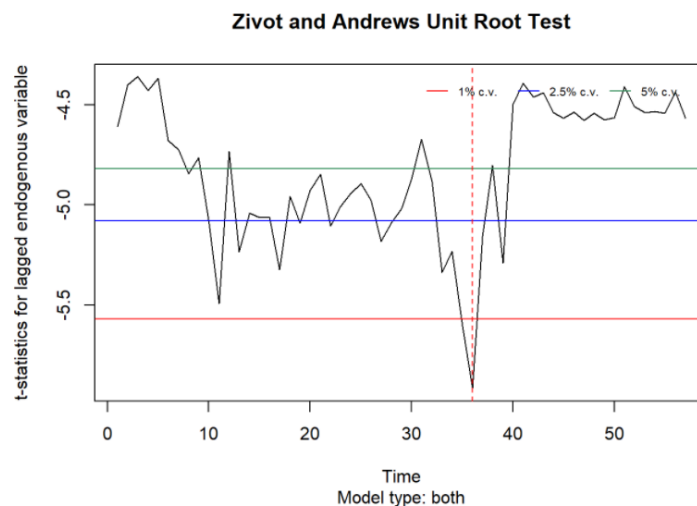


Ilustración 2.4

## 3) Cambios estructurales

### 3.1. Cambios estructurales con tendencia (SdT2 ~ timeline)

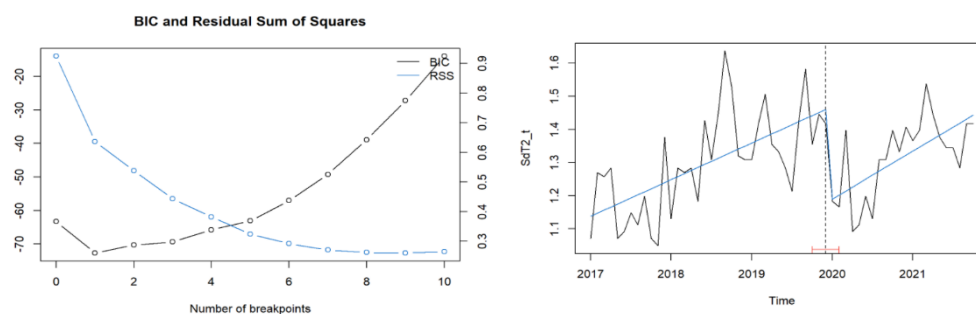


Ilustración 2.5

Se busca determinar los cambios estructurales y las rectas de regresión lineal asociadas con el siguiente modelado:  $SdT2 \sim \text{timeline}$ . Como los p-values de intercepto y tendencia son muy pequeños, rechazamos la  $H_0$  que los considera no significativos. Se identifican los posible puntos de quiebre de la tendencia.

El mejor Criterio de información se da para  $m=1$  (Ilustración 2.5).

#### 4) Modelos

Se utiliza la función `auto.arima`, para gener el modelo de la mejor performance. Cuyo resultado arroja como → Best model:  $ARIMA(0,1,1)(1,0,0)[12]$

Modelo 1 ajustado solo con datos Train: se generan los gráficos de serie original, valores ajustados en período Train y predicción en período Test (Gráfico 1.8):

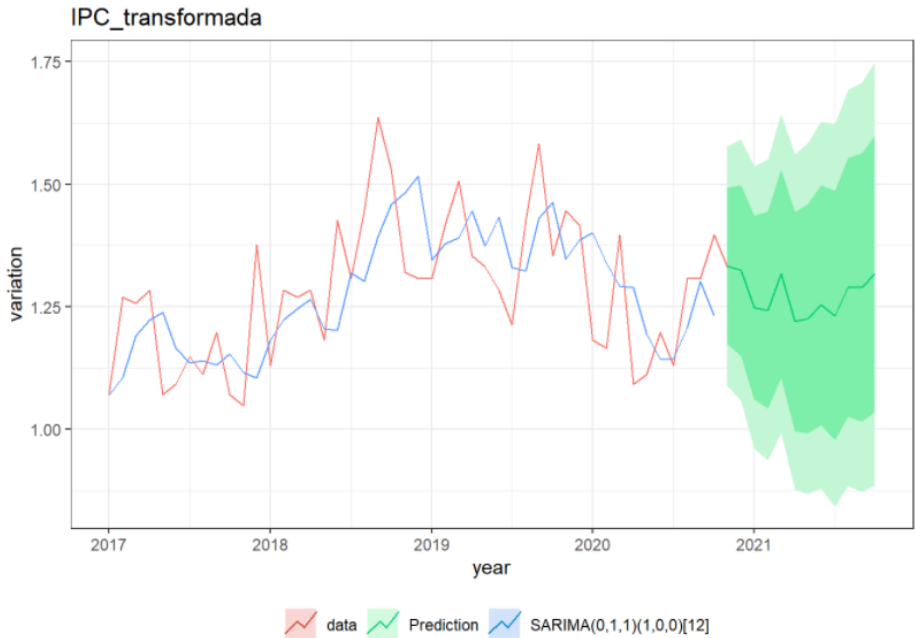


Ilustración 2.6

Se estiman métricas de performance del modelo seleccionado:

Tabla 2.3

Training set	0.006302009	0.1203768	0.09796848	-0.1306514	7.560753	0.6257020	0.08382275	NA
Test set	0.114156209	0.1331057	0.11535321	8.0571662	8.149884	0.7367342	0.37899888	1.768742

## 5) Análisis de los residuos

Resultan p-values mayores que nivel de significancia 0.05 (el menor encontrado es 0.381), por lo que NO se rechaza la  $H_0$  de Ljung-Box test ( $H_0$ : Los residuos están distribuidos independientemente). El gráfico de PAC no muestra lags significativos.

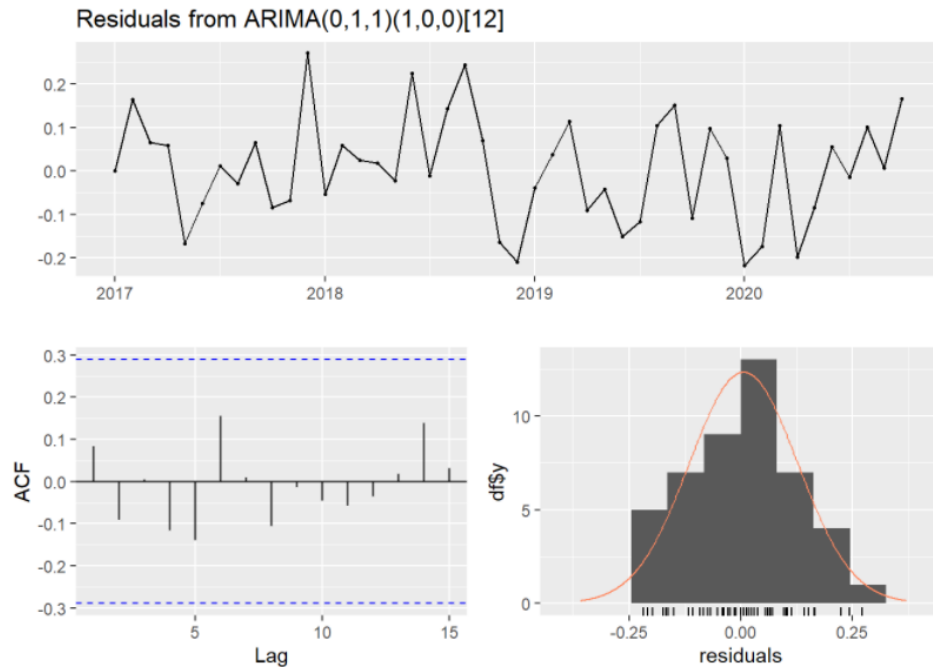


Ilustración 2.7

## 6) Predicción

Se generaliza el modelo para toda la serie (Train+Test) y se realiza la predicción para 12 meses. Se procede luego a des-transformar la serie. Se genera una serie de tiempo a partir de los valores predichos de la variable destransformada.

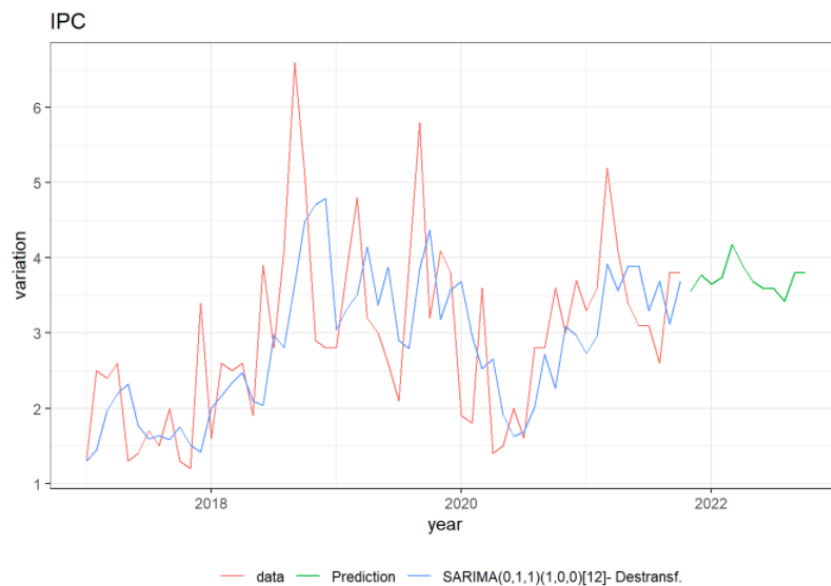


Ilustración 2.8

Se compara con el Accuracy obtenido en el trabajo práctico anterior con el modelo ARIMA(0,2,1):

*Tabla 2.4*

Test	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
<b>TP1 ARMA(0,2,1)</b>	43.1657	52.7700	43.1657	9.9392	9.9392	0.6546	0.7744	3.5354
<b>TP2 SARIMA (0,1,1) (1,0,0)[12]</b>	0.1141	0.1331	0.1153	8.0571	8.1498	0.7367	0.3789	1.7687

A partir de estos valores, se concluye que el modelo aplicado a la serie 2 que considera la ESTACIONALIDAD en la serie temporal, obtuvo mejores valores (menores) en casi todos los Criterios de Información resultantes (a excepción de MASE: Mean Absolute Scaled Error).

## Referencias bibliográficas

- Uriel. Análisis de Series Temporales. Modelos ARIMA. Capítulo 4 destinado a Procesos Estocástico.
- Peña. Análisis de Series Temporales. Alianza Editorial. Capítulos 6.
- Chatfield et al.. The Analysis of Time Series. An Introduction with R. Seventh Edition. CRC Press. Leer Capítulo 5 destinado a Forecasting.
- Hyndman. Forecasting: Principles and Practice. Second Edition. Monash University, Australia. Leer Capítulo 3 destinado a herramientas de Forecasting y medidas de performance.