



| **UNR** Universidad
Nacional de Rosario

LICENCIATURA EN ESTADÍSTICA

Trabajo Práctico

*“Evaluación de las Tendencias de Ventas Mensuales en una Empresa de Logística
Mediante Series de Tiempo”*

Autor: Santini, Franco

Docentes: Mendez, Fernanda - Sigal, Facundo

2024

Tabla de contenidos

Introducción	1
Análisis descriptivo	1
Análisis de dispersion	3
Identificación del modelo	4
Diferenciación de la serie	4
Evaluación de los supuestos de los modelos	6
Evaluación de la capacidad predictiva	8
Anexo	9
Bibliografía	10

Introducción

En este trabajo se busca estudiar las unidades vendidas (bebidas con y sin alcohol) en Hectolitro¹ (HL) de una importante empresa de logística en la ciudad de Casilda, Santa Fe. El estudio se realiza con 80 observaciones, medidas mensualmente en el período (01/01/2018 - 31/08/2024).

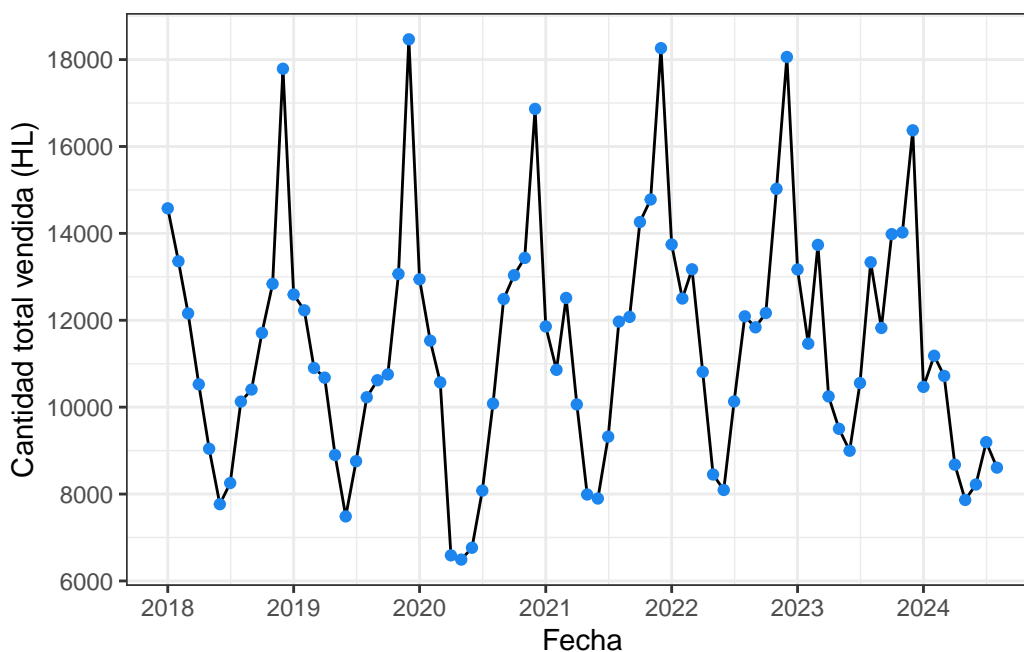
Cabe aclarar que en el período de estudio, atravesamos un evento muy catastrófico en todas las áreas, particularmente en el sector de logística, la pandemia del COVID-19. Al afectar principalmente a las personas, Argentina implementó el Aislamiento Social, Preventivo y Obligatorio (ASPO) como medida para combatir la pandemia de COVID-19. Debido a esta medida, a fines de marzo del 2020 - principios de abril del 2020, el sector de logística fue afectado y en general se produjo una caída en las ventas con respecto al año anterior.

Teniendo esto en cuenta, resulta de interés pronosticar a futuro las unidades vendidas en (HL) de la empresa, mediante modelos de *series de tiempo*.

Análisis descriptivo

En primera instancia, comenzamos con el gráfico de la serie para observar el comportamiento de las ventas a través del período en estudio.

Figura 1: Unidades vendidas mensuales en (HL)

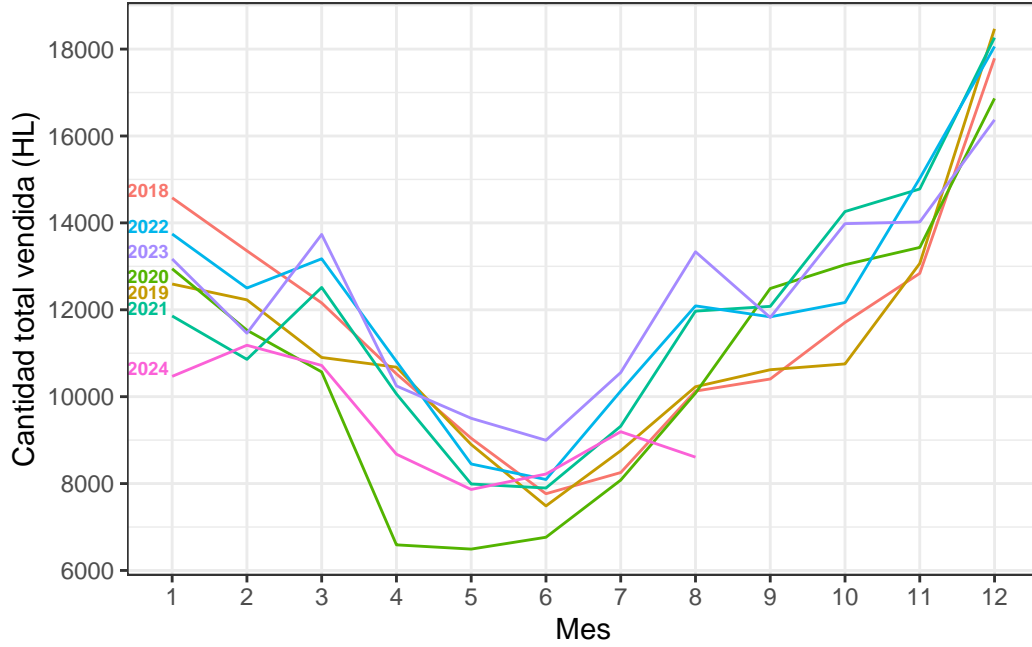


En la Figura 1 se puede observar un comportamiento estacional de las unidades vendidas, disminuye considerablemente en los meses (mayo, junio y julio) y luego aumenta considerablemente en los meses

¹Unidad de volumen equivalente a 100 litros.

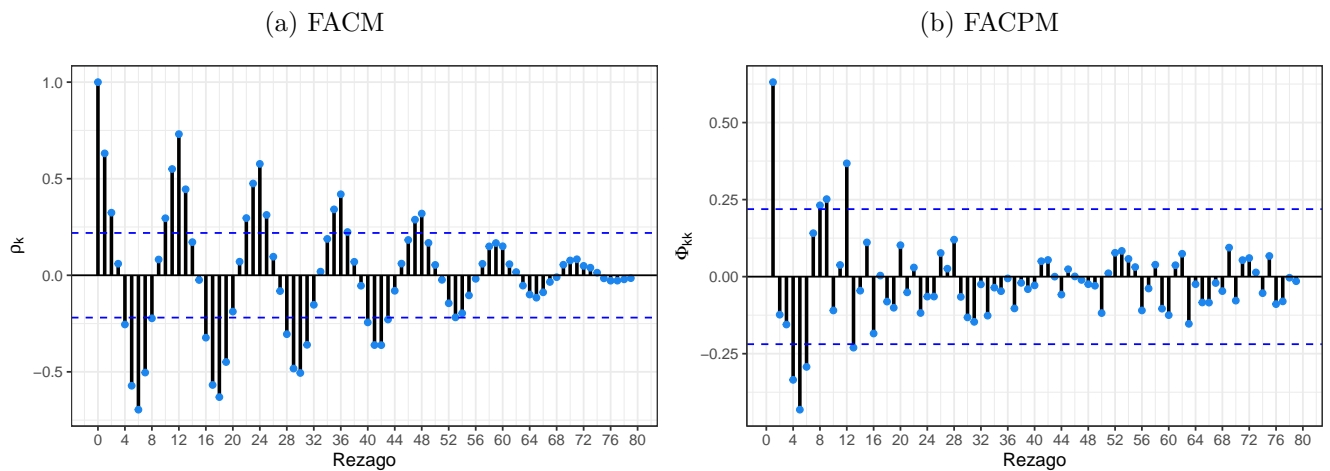
(noviembre - diciembre - enero) en casi todos los años, exceptuando el año 2020 que puede deberse a la pandemia.

Figura 2: Comportamiento anual de las unidades vendidas en (HL)



La Figura 2 podemos observar que claramente hay un comportamiento estacional de las unidades vendidas, pero se descarta cualquier tipo de tendencia creciente o decreciente de las mismas, dado que no se observa ni un crecimiento, ni un decrecimiento año a año de las unidades vendidas mensuales.

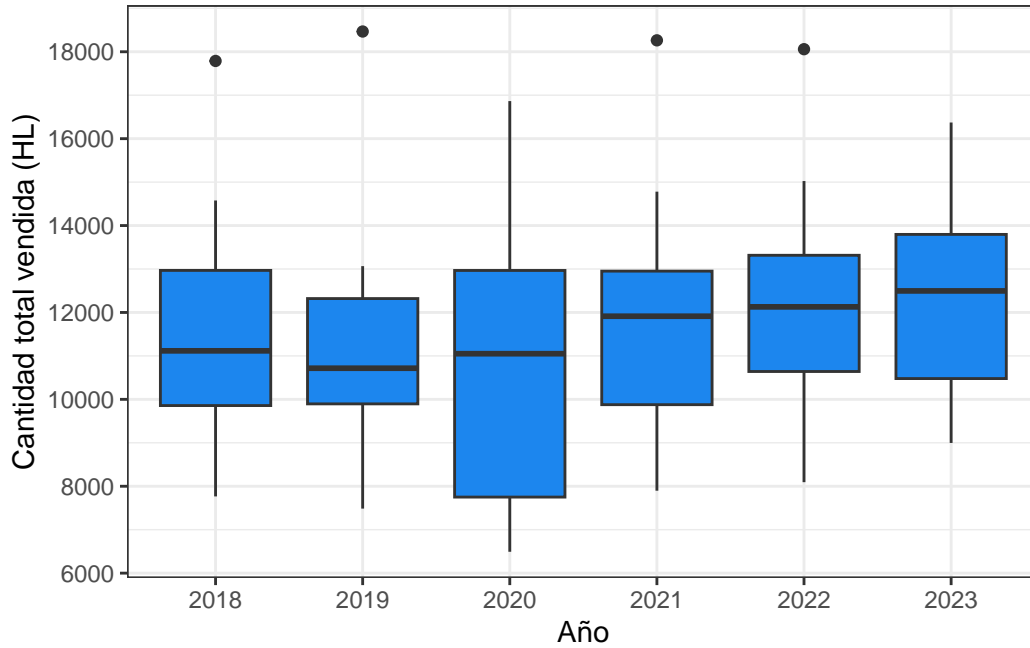
Figura 3: Funciones de autocorrelación muestral



De acuerdo con lo mostrado en Figura 3, se puede observar que la FACM tiene un decrecimiento lento y sinuoso, notando también que en los rezagos 12, 24, 36 y 48 son significativos esto da un indicio de que la serie tiene un comportamiento estacional, mientras que la FACPM no parece tener un patrón claro.

Analisis de dispersion

Figura 4: Boxplots de las unidades vendidas en (HL) por año



Se puede observar en la Figura 4 que la variabilidad no parece ser constante año a año, por lo que puede estar dando una advertencia de que la serie no es estacionaria en variancia, además se puede apreciar la presencia de outliers en los años 2018, 2019, 2021 y 2022. Se excluyó el año 2024 de la Figura 4 porque no estaban los datos completos.

Una vez observados los Boxplots, procederemos a realizar la transformación de Box-Cox. Dependiendo del valor de λ obtenido, determinaremos si es necesario transformar la serie o si puede mantenerse en su forma original. Además, puede ser de utilidad ajustar el coeficiente de variación según los valores de λ arrojados por la transformación de Box-Cox, de esta manera, podremos identificar si una transformación de potencia es adecuada para estabilizar la varianza de la serie. Aquel valor de λ para el cuál se minimice el coeficiente de variación, será la transformación que tendremos que realizar.

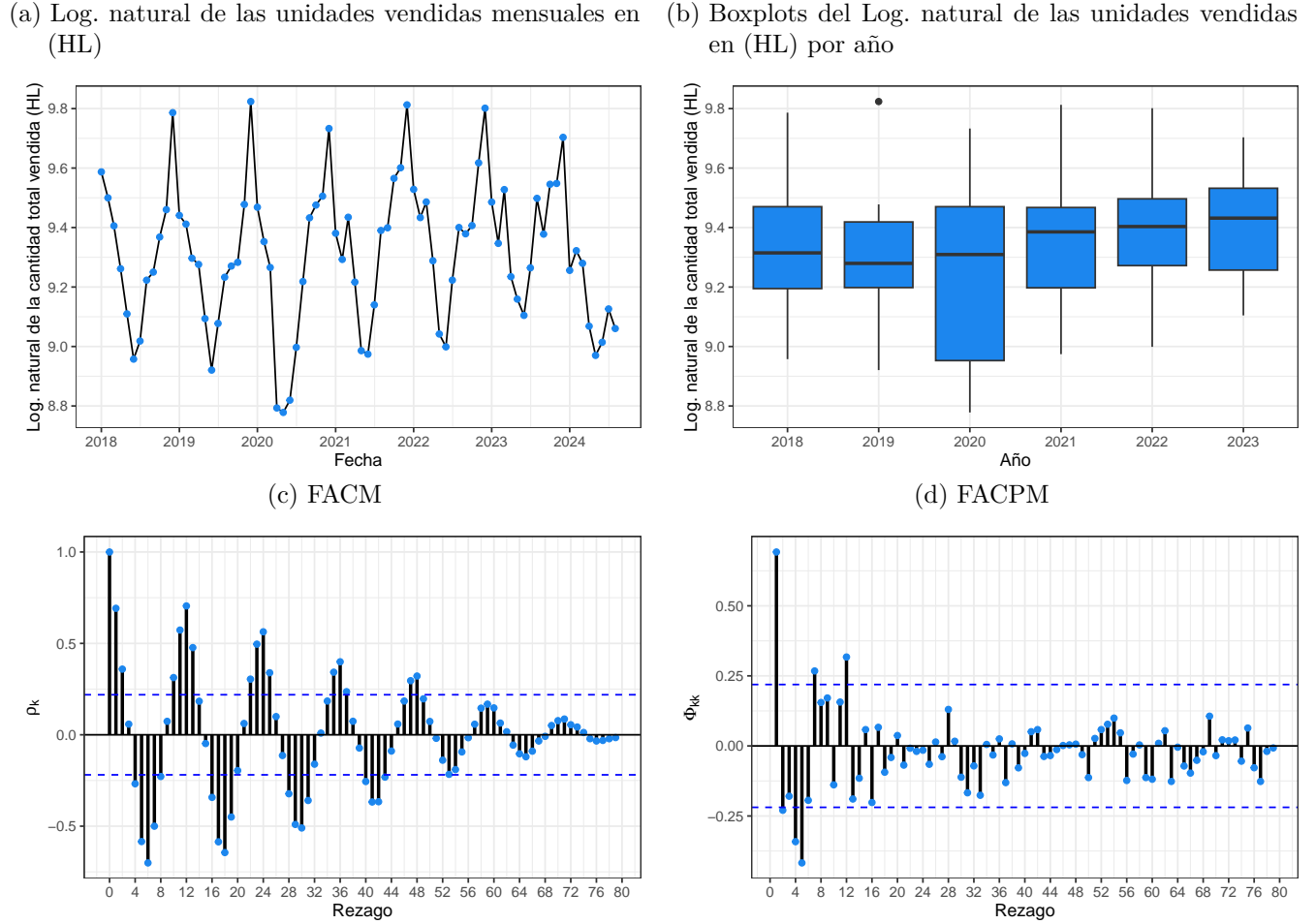
Tabla 1: Coeficiente de variación

λ	Coeficiente de variación
-2.0	0.5001
-1.0	0.2433
-0.5	0.1206
0.0	0.0258
0.5	0.1198
1.0	0.2405
2.0	0.4909

Observando la Tabla 1, el coeficiente de variación mínimo esta asociado a un $\lambda = 0$, por lo que es apropiado aplicar la transformacion $y^{(\lambda)} = \ln(y)$.

Identificación del modelo

Figura 5: Resumen de la Serie transformada



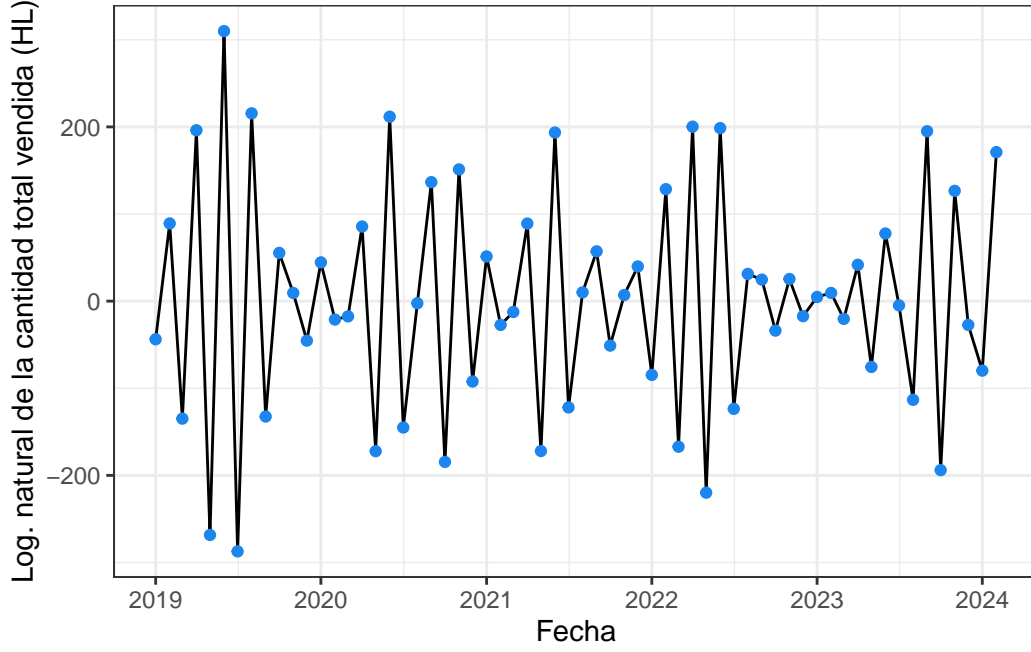
Luego de transformar a la serie, se realizó la Figura 5 que contiene un resumen acerca del comportamiento de la serie. Notando, que ahora en el gráfico (b) se puede observar que la variabilidad parece ser un poco más constante año a año.

Diferenciación de la serie

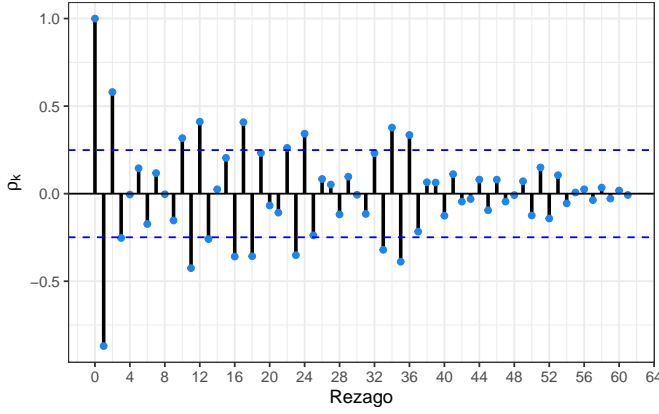
Una vez realizado el análisis descriptivo, se notó el comportamiento estacional de la serie, por lo qué, se realiza una diferenciación en la parte estacional de la misma, dado que lo rezagos estacionales en la FACM observada en la Figura 5 parecen decrecer lentamente de forma lineal, pasando a trabajar con 62 observaciones.

Figura 6: Resumen de la Serie diferenciada en la parte estacional

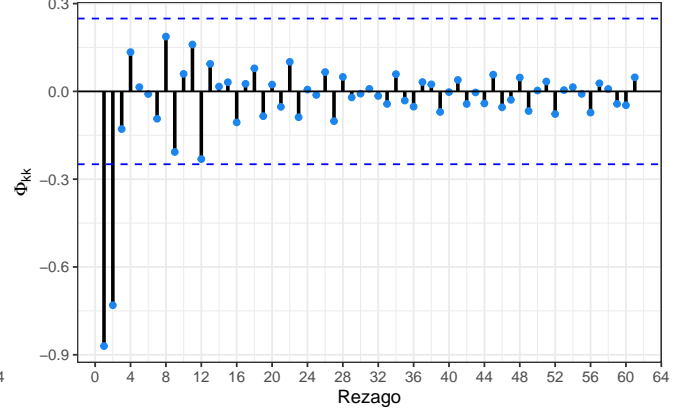
(a) Log. natural de las unidades vendidas mensuales en (HL)



(b) FACM



(c) FACPM



En la Figura 6 se observa, el gráfico de la serie diferenciada en la parte estacional y sus respectivas funciones de autocorrelación muestral y de autocorrelación parcial muestral. En el gráfico (a) se puede observar que la serie parece ser estacionaria, dado que varía de manera oscilante sobre un valor constante. Luego para realizar la identificación de los posibles modelos que se pueden ajustar a esta serie, utilizaremos los gráficos (b) y (c). En el gráfico de la FACM se puede observar que 2 rezagos son significativos, y el tercer rezago está en el límite de ser significativo, viendo la parte estacional parece que los rezagos 12, 24 y 36 son significativos, presentando un decrecimiento exponencial. En el gráfico de la FACPM se puede observar que solo los primeros 2 rezagos son significativos, y el rezago 12 correspondiente a la parte estacional, parece estar al límite de ser significativo.

Por lo que teniendo esto en cuenta, planteamos los siguientes modelos que se pueden ajustar a nuestra serie:

- Modelo 1: $SARIMA(0, 0, 3)(1, 1, 0)_{12}$
- Modelo 2: $SARIMA(0, 0, 2)(1, 1, 0)_{12}$
- Modelo 3: $SARIMA(2, 0, 0)(1, 1, 0)_{12}$

Una vez planteado los modelos que vamos a utilizar para el pronóstico de la serie, se procede a estimarlos y compararlos.

Tabla 2: Comparación de los modelos

Modelos	$\hat{\sigma}^2$	Log_veros.	AIC	AICc	BIC
$SARIMA(1, 0, 0)(0, 1, 1)_{12}$	0.0087	57.1397	-108.2794	-107.8656	-101.8980
$SARIMA(2, 0, 0)(1, 1, 0)_{12}$	0.0100	55.0957	-102.1914	-101.4897	-93.6829
$SARIMA(0, 0, 2)(1, 1, 0)_{12}$	0.0104	53.9543	-99.9087	-99.2069	-91.4001
$SARIMA(0, 0, 3)(1, 1, 0)_{12}$	0.0104	54.5016	-99.0032	-97.9318	-88.3676

Se puede observar, que los modelos con menor AIC son:

1. $SARIMA(2, 0, 0)(1, 1, 0)_{12}$
2. $SARIMA(1, 0, 0)(0, 1, 1)_{12}$

El modelo 1 tiene 3 parámetros de los cuáles ϕ_2 no es significativo, por lo que a dicho modelo le sacamos un parámetro en la parte regular. Por lo que el cambiamos el modelo 1 por un modelo $SARIMA(1, 0, 0)(1, 1, 0)_{12}$. Con respecto al modelo 2, sus dos parámetros son significativos, teniendo además una estimación de $\hat{\sigma}^2$ menor que la del modelo 1. De igual manera, seguiré con los modelos, definiendo cuál usaré para la predicción, aquél que tenga una mejor capacidad predictiva, teniendo en cuenta la verificación de los supuestos de los mismos.

Evaluación de los supuestos de los modelos

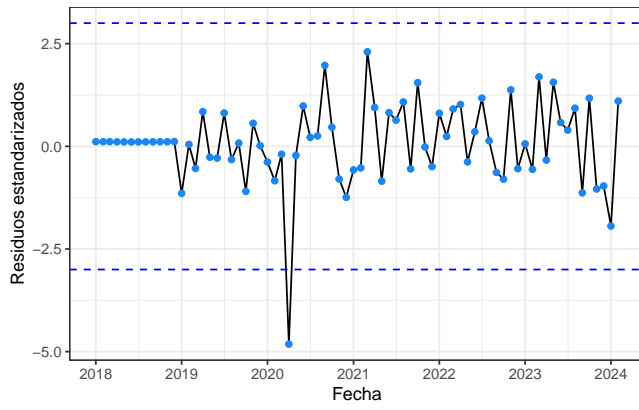
Esta instancia se centra en la verificación de los supuestos de ambos modelos mencionados anteriormente. Se hará un análisis exhaustivo para el modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$, mientras que para el modelo $SARIMA(1, 0, 0)(1, 1, 0)_{12}$, al ser un procedimiento análogo, se dejarán en el anexo la Figura 8.

En la [Figura 7](#) se puede observar en los gráficos (a) y (b) que hay una observación atípica, correspondiente al año 2020, lo cual se puede deber a alguna situación inusual ocurrida en pandemia, además los residuos varían alrededor del cero y una variancia constante. También se puede apreciar, en el gráfico (b) que los residuos parecen ser normales, obviamente sin incluir al outlier. Luego en los gráficos (c) y (d), da la impresión de que los residuos son incorrelados. Por lo que podemos concluir que el modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ cumple con los supuestos.

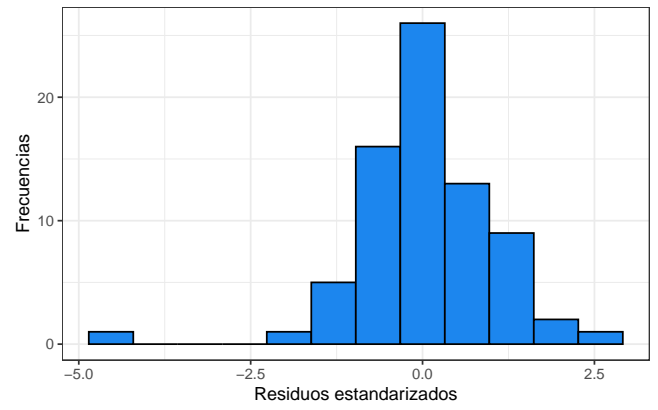
Para el modelo $SARIMA(1, 0, 0)(1, 1, 0)_{12}$ se realizó lo mismo obteniendo resultados análogos al modelo anterior.

Figura 7: Análisis de residuos modelo $SARIMA(1, 0, 0)(0, 1, 1)_{12}$

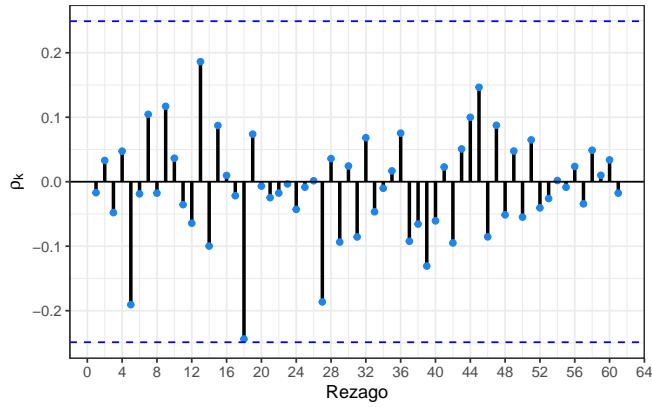
(a) Residuos estandarizados vs fecha



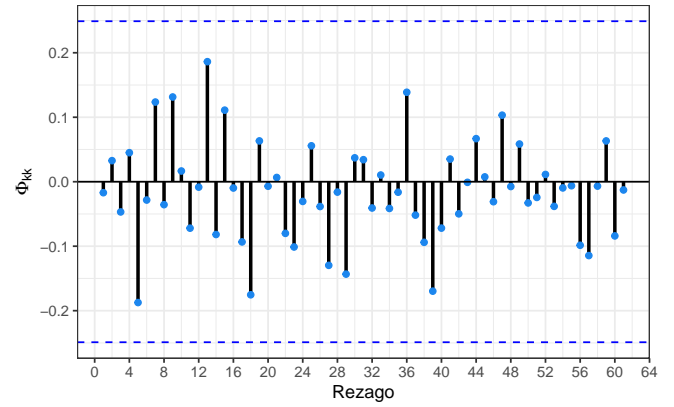
(b) Histograma de los residuos estandarizados



(c) FACM de los residuos



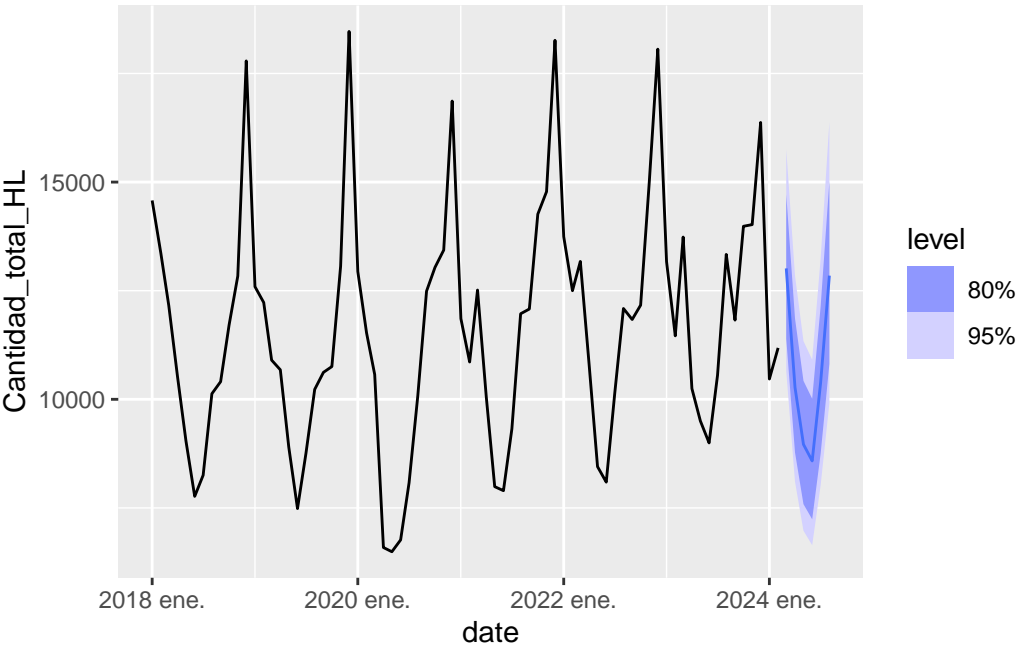
(d) FACPM de los residuos



Evaluación de la capacidad predictiva

```
# A tibble: 2 x 5
  nombre_modelos      RMSE  MAE  MPE  MAPE
  <chr>            <dbl> <dbl> <dbl> <dbl>
1 $SARIMA(1,0,0)(0,1,1)_{12}$ 1720. 1291. -14.3  14.4
2 $SARIMA(2,0,0)(1,1,0)_{12}$ 2181. 1797. -20.1  20.1

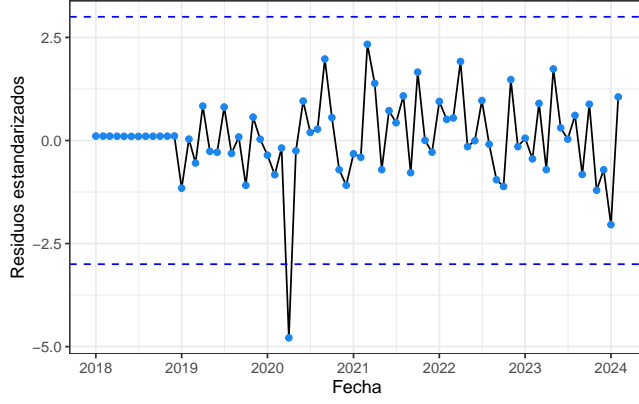
# A fable: 12 x 5 [1M]
# Key:      key, .model [2]
  key .model      date Cantidad_total_HL  .mean
  <dbl> <chr>      <mt> <dist> <dbl>
1 1 auto      2024 mar.  t(N(9.4, 0.0087)) 12460.
2 1 auto      2024 abr.  t(N(9.2, 0.012))  9830.
3 1 auto      2024 may.   t(N(9, 0.014))   8572.
4 1 auto      2024 jun.   t(N(9, 0.014))   8203.
5 1 auto      2024 jul.   t(N(9.2, 0.015))  9781.
6 1 auto      2024 ago.   t(N(9.4, 0.015)) 12150.
7 1 sarima200110 2024 mar.   t(N(9.5, 0.01)) 13013.
8 1 sarima200110 2024 abr.   t(N(9.2, 0.014)) 10273.
9 1 sarima200110 2024 may.   t(N(9.1, 0.015))  8963.
10 1 sarima200110 2024 jun.   t(N(9, 0.016))   8582.
11 1 sarima200110 2024 jul.   t(N(9.2, 0.016)) 10386.
12 1 sarima200110 2024 ago.   t(N(9.5, 0.016)) 12847.
```



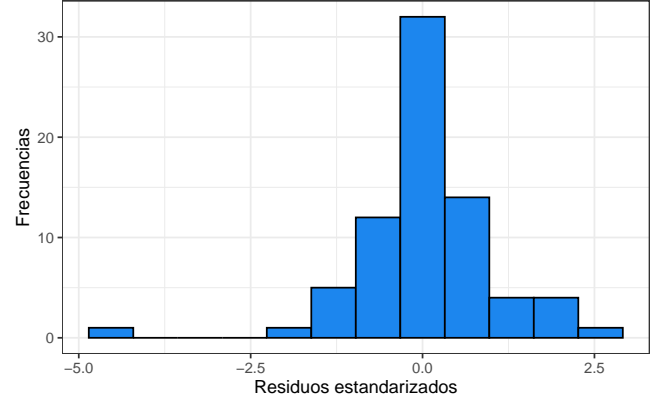
Anexo

Figura 8: Análisis de residuos modelo $SARIMA(1, 0, 0)(1, 1, 0)_{12}$

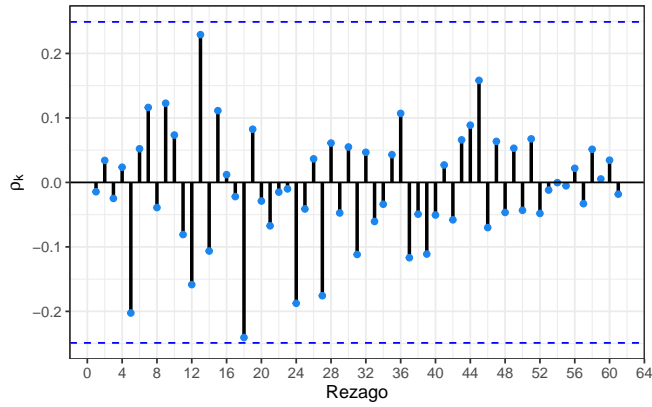
(a) Residuos estandarizados vs fecha



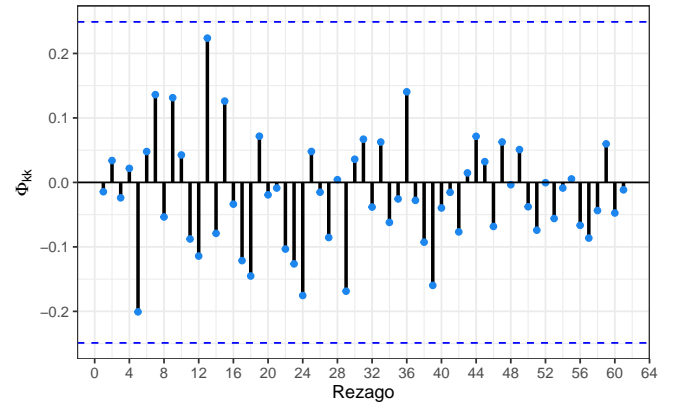
(b) Histograma de los residuos estandarizados



(c) FACM de los residuos



(d) FACPM de los residuos



Bibliografía

- Wei, W.S. (2006). Time Series Analysis: Univariate and Multivariate Methods, 2nd edition
- Hyndman, R.J., & Athanasopoulos, G. (2021). Forecasting: principles and practice, 3rd edition
- Guerrero, V.M. (1993). Time series analysis supported by power transformation