UNIVERSIDAD NACIONAL DE COLOMBIA

Departamento de Estadística
Series de tiempo univariadas
Tarea 2
01 de febrero - 2025

Jhon Alejandro Ramírez Daza jramirezda@unal.edu.co

---

1. Punto 1

Consiga dos series de diferente periodicidad (por ejemplo: Una mensual y otra trimestral, series semanales, diarias o intradía son aceptables aunque más difíciles de modelar en la práctica). Documentar en dos párrafos la forma exacta de construcción de cada serie, esto es fuente, procedimiento, unidades de medida y otra información que consideren relevante. Brevemente aclare el interés suyo o de algún usuario en modelar y/o pronosticar dichas series. Presente las gráficas de cada una de estas dos series y describa brevemente la dinámica de cada serie. El número de datos debe incluir al menos desde 1970 para series anules, desde 1980 para bimestrales y para trimestrales, desde enero de 1986 en el caso de mensuales; desde 1990 en caso semanal y desde 2001 en el diario.

Para este ejercicio se usaron las siguientes series de tiempo:

Producción anual de Ferrari, en la que se recopila anualmente cuantas unidades se producían de automóviles Ferrari al año, empezando desde 1947 y hasta el año 2023, siendo Ferrari una importantísima empereza de autos de lujo y ser tan escasas las unidades producidas de forma anual puede resultar importante saber a posibles compradores cuantas unidades disponibles podrían producirse en años posteriores, estos datos fueron recopilados por el Ferrari Club España con información dada por Ferrari.

El gráfico de la serie de tiempo es el siguiente:

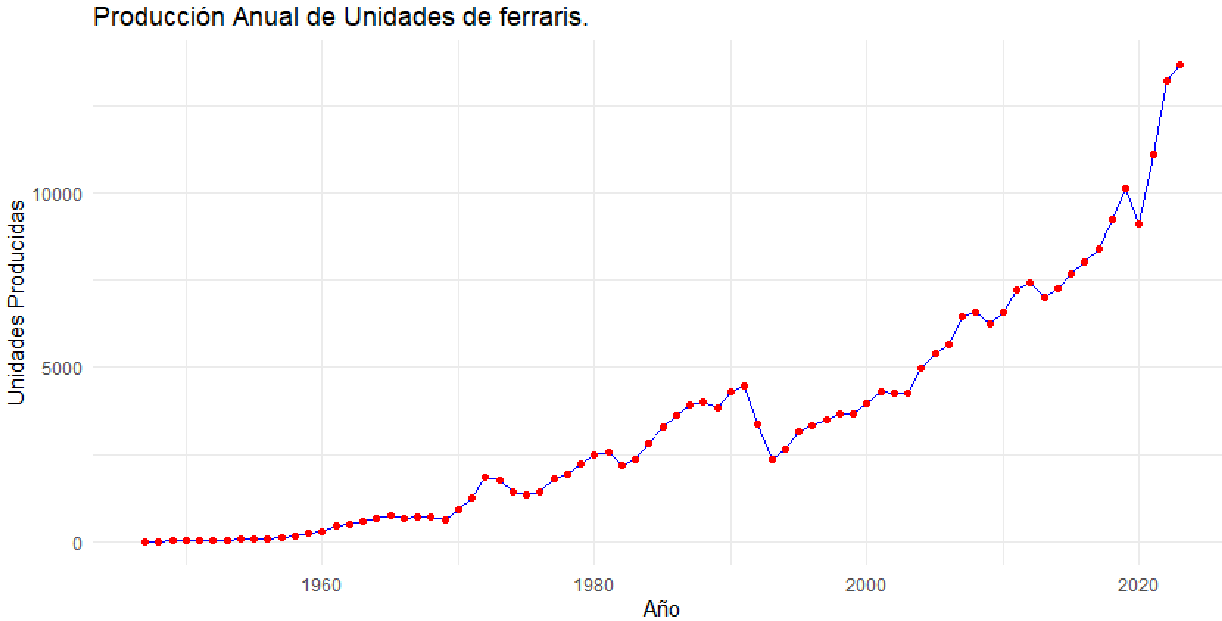

Figura 1: Producción anual Ferrari

En la figura 1 notamos como la cuenta con una tendencia que parece a simplista estocástica, no presenta a simple vista problemas de volatilidad, ni comportamientos estacionales.

Cantidad diaria de manchas solares contadas por SILSO, Observatorio Real de Bélgica, Bruselas. originalmente este dataset cuanta con datos desde 1850 hasta 2024 pero solo vamos a tener en cuanta los valores desde el año 2000, ya que al ser datos diarios son una gran cantidad de observaciones, por lo que nos concentraremos en los datos mas resientes, puede resultar de mucha utilidad conocer cual es el comportamiento de estas manchas solares para conocer como se relacionan estas con la actividad solar por ejemplo, y con esta información tener predicciones de el clima solar proximo.

la gráfica de esta serie de tiempo es la siguiente:

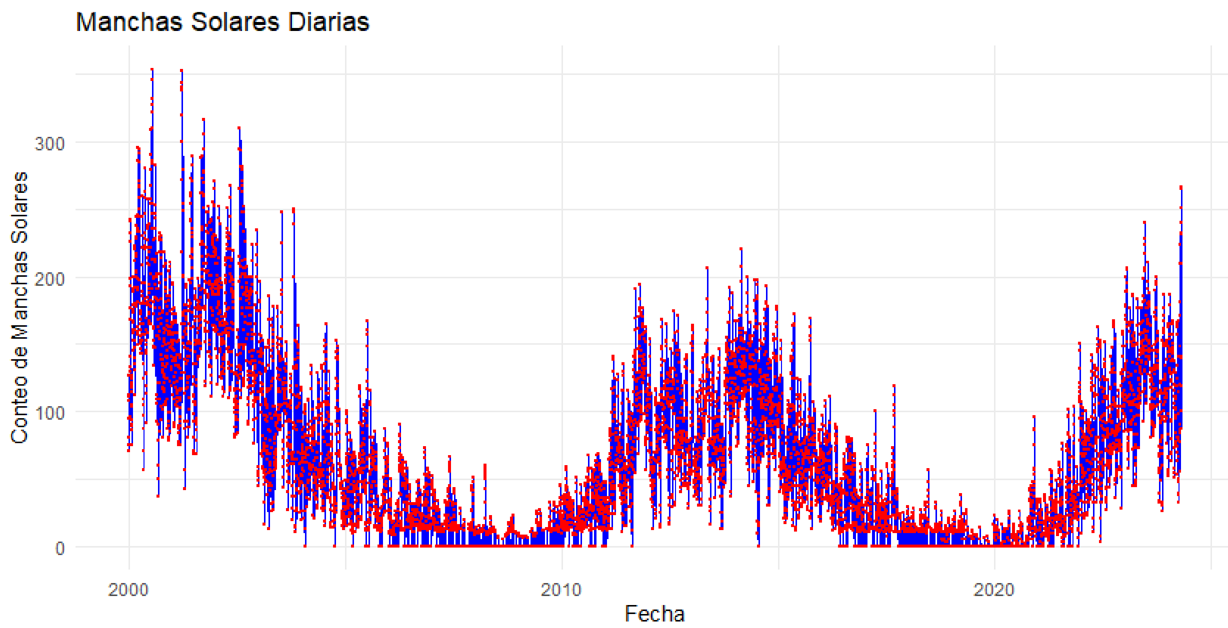

Figura 2: Conteo de manchas solares diarias

De la gráfica 2 vemos que no es muy claro a simple vista si existe un nivel base, esto dado a que en la gráfica se puede ver algún estacionalidad o un ciclo, tampoco se ven problemas con volatilidad.

2. Punto 2

Trabajando en Excel u otro programa (R o Python), calcule y presente gráficamente dos promedios móviles de algún orden para cada una de las series de la pregunta 1. Comente brevemente las diferencias entre ellas y la serie original.

Empecemos con los datos de producción de Ferrari, se usaron ventanas de 3,5 y 10 años para la creación de los promedios móviles en la gráfica 3.

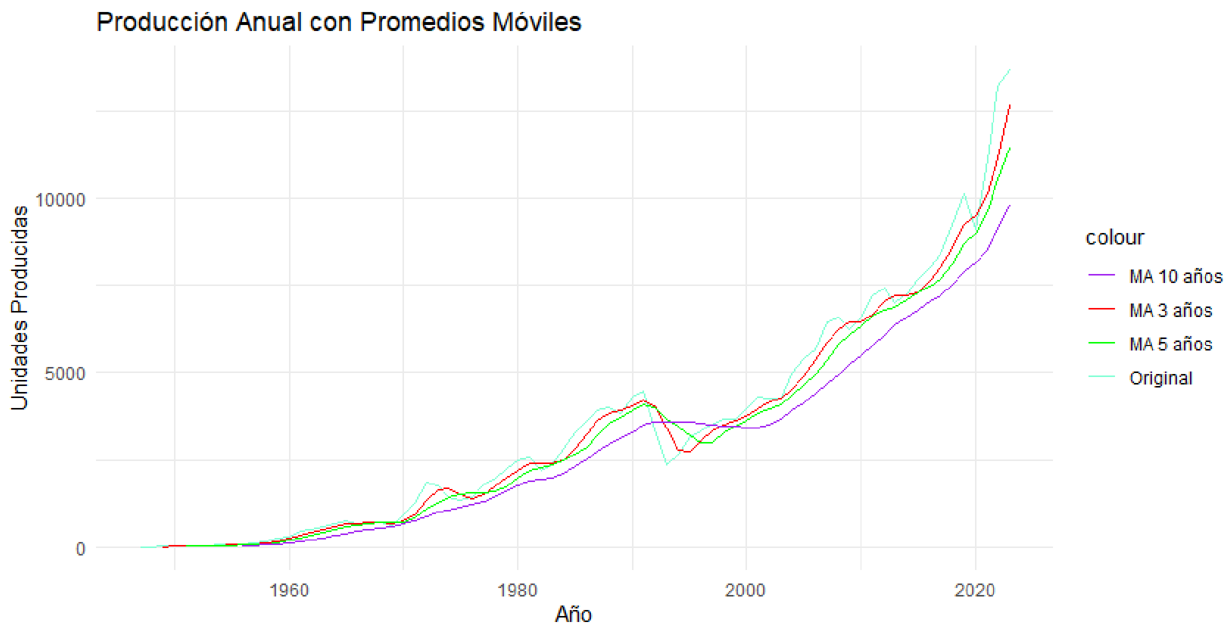

Figura 3: Promedios móviles producción Ferrari

De el gráfico 3 vemos que perdemos menos datos al usar 3 años pero la gráfica conserva mucha de la dinámica de la serie original, por otro lado el promedio móvil usando 10 años muestra de forma mas clara la tendencia que posee la serie original sin embargo se pierden bastantes datos al no contar con demasiados por ser una serie anual.

Por otro lado para la serie de observaciones de manchas solares se usaron 30, 180 y 365 días como ventana para realizar estos promedios móviles, dando como resultado la gráfica 4:

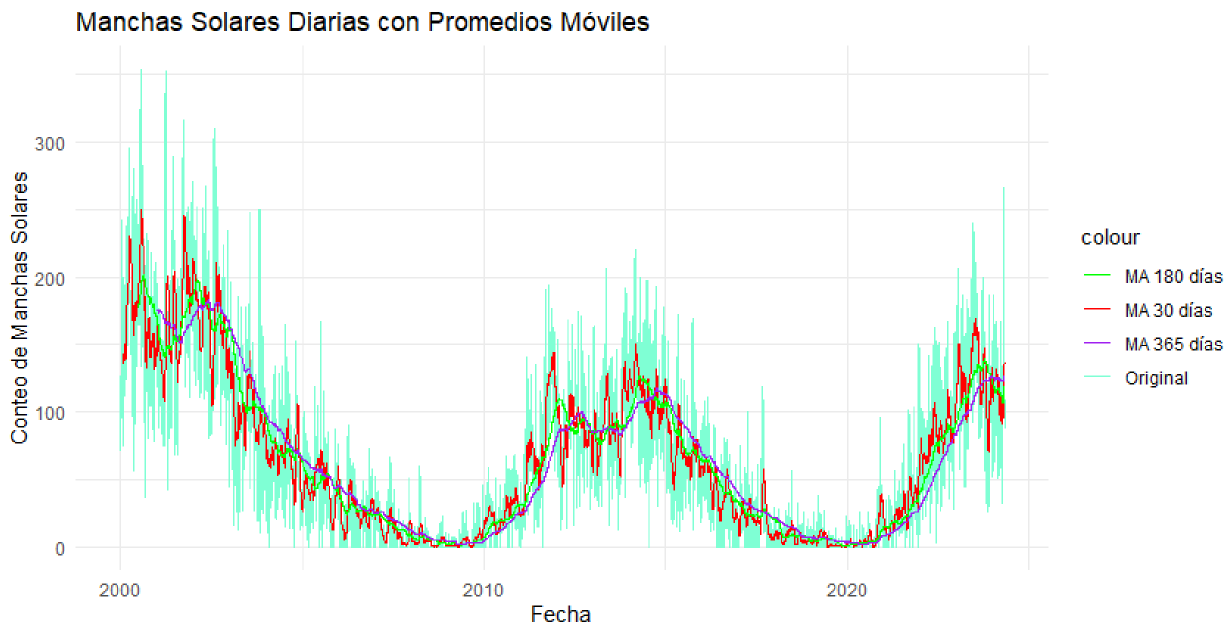

Figura 4: Promedios móviles manchas solares observadas diarias

De la gráfica 4 podemos ver que con tan solo 30 días se ajusta fuertemente a la curva aunque con menor dispersión, mientras que con 180 y 365 días se puede ver de forma mas suave la tendencia que presenta la serie, también es posible ver de forma más clara el ciclo o estacionalidad.

3. Punto 3

Repita el siguiente ejercicio numérico tres veces. Genere $n = 500$ observaciones a partir del modelo ARMA dado por:

$$x_t = 0{,}9x_{t-1} + w_t - 0{,}9w_{t-1}$$

con $w_t \sim iidN(0, 1)$. Dibuje los datos simulados, calcule la ACF y la PACF de muestra de los datos simulados y ajuste un modelo ARMA(1, 1) a los datos. ¿Qué sucedió y cómo explica los resultados?

los resultados graficos se presentan a continuación:

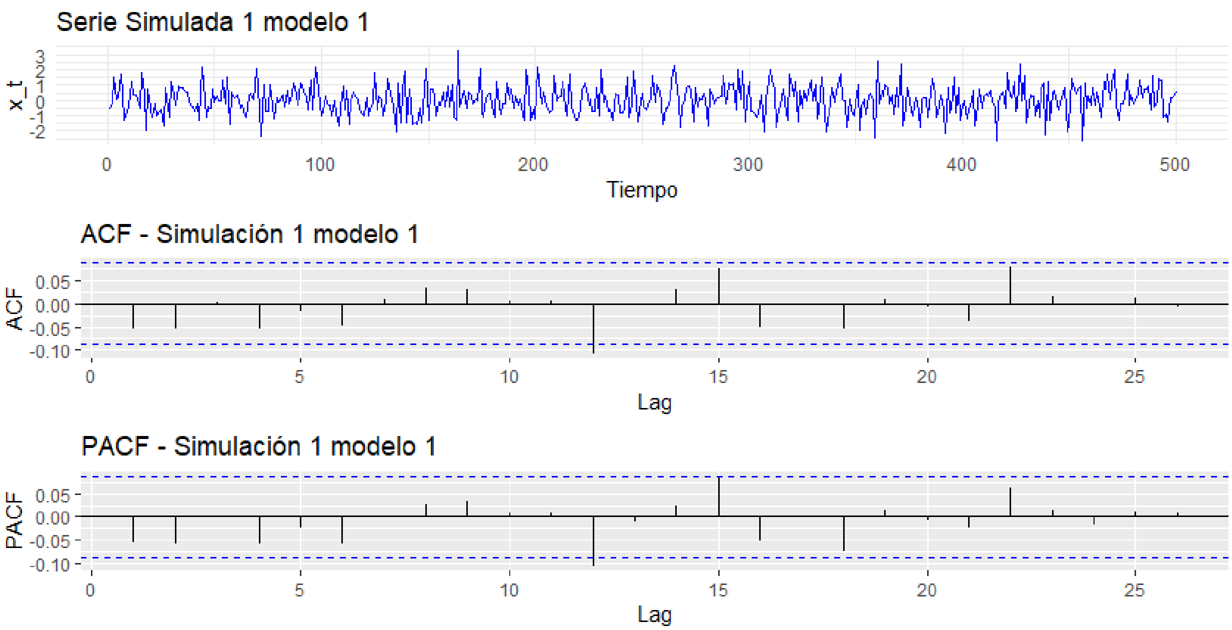

Figura 5: Simulación 1 modelo 1

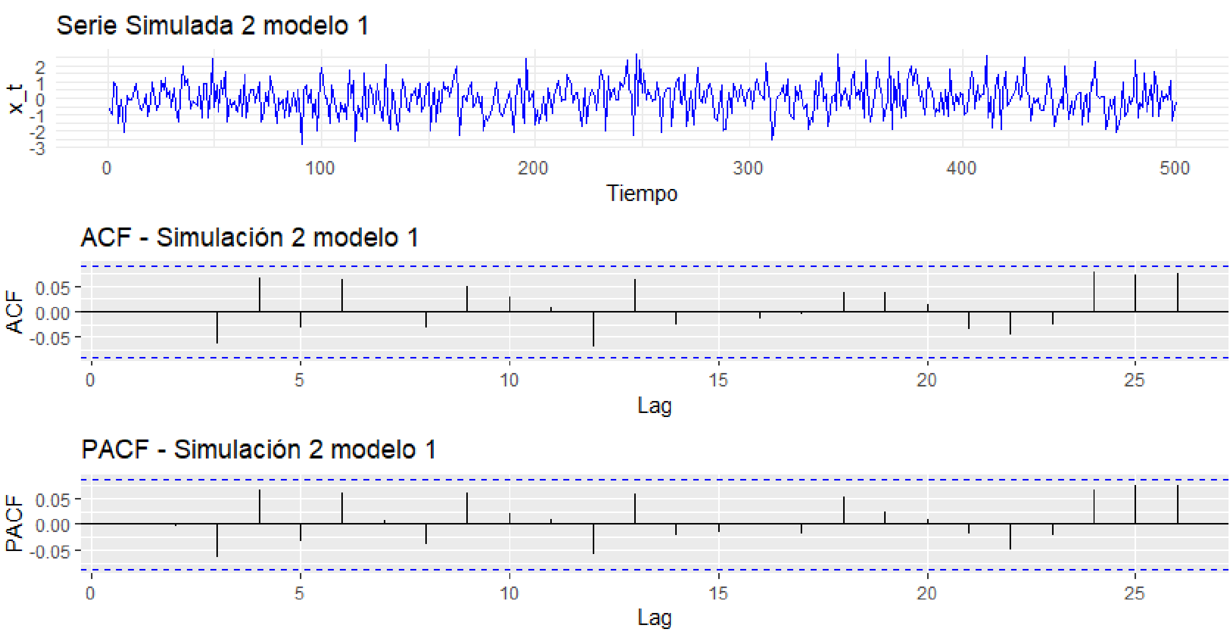

Figura 6: Simulación 2 modelo 1

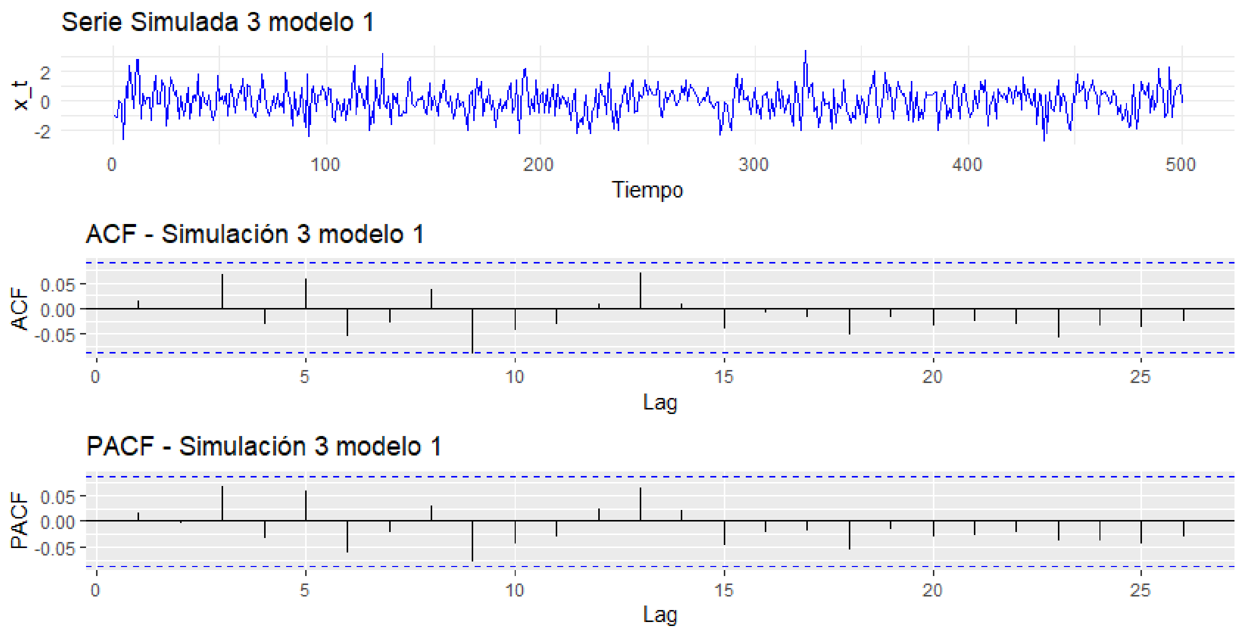

Figura 7: Simulación 3 modelo 1

Con las 3 simulaciones se realizaron los graficos fijándonos en los graficos de las series vemos que como es esperable por la simulación las 3 tienen como nivel base el 0 dispersiones similares aunque en el gráfico 6 se tiene un eje y más extenso esto por algunos datos mas grandes que se simularon.

Por otro lado vemos que tanto para los graficos ACF como PACF no se ve ningún comportamiento, como era de esperarse, solo se dio que en la simulación 1 se presento un resago apenas significativo en el resago 12 tan para el gráfico ACF como el PACF, como se ve en 5.

por otro lado los resultados del ajuste del modelo $ARMA(1,1)$ para cada una de las simulaciones se ve resumido en la siguiente tabla 1

Simulación	AR1 Estimado	MA1 Estimado	σ^2 Estimado	AIC
1	0.7135	-0.7794	0.936	1393.91
2	-0.0012	-0.0012	1.02	1436.78
3	0.6188	-0.5996	0.9832	1418.46

Cuadro 1: Resumen de las estimaciones de los modelos ARMA(1,1) para las tres simulaciones.

Con respecto a la table 1 vemos que con los datos simulados 1 se estimaron de forma más cercana a la realidad los parámetros tanto de la parte AR como de la parte MA de igual forma la varianza estimada fue bastante cerca, por otro lado los datos simulados 2 llevaron a unos parámetros bastante alejados de los valores reales tanto de el parámetro AR como del parámetro MA mientras que la estimación de la varianza también estuvo muy aceptada.

En este caso no resulta de utilidad comparar los AIC ya que los 3 son igual modelo, pero los datos simulados 1 llevaron a un AIC más bajo.

Repita el ejercicio con los dos modelos siguientes:

$$x_t = 0{,}2x_{t-1} + 0{,}55x_{t-2} + w_t w_t \sim iidN(0, \sigma^2 = 2{,}25)$$

$$x_t = w_t + 0{,}9w_{t-1} - 0{,}8w_{t-2} - 0{,}8w_{t-3} w_t \sim iidN(0, \sigma^2 = 9)$$

Como vemos el segundo modelo corresponde a un modelo $AR(2)$ puro y el tercer modelo a un $MA(3)$ puro, realizamos 3 simulaciones de cada uno igual que en el caso anterior, empecemos por el modelo 2:

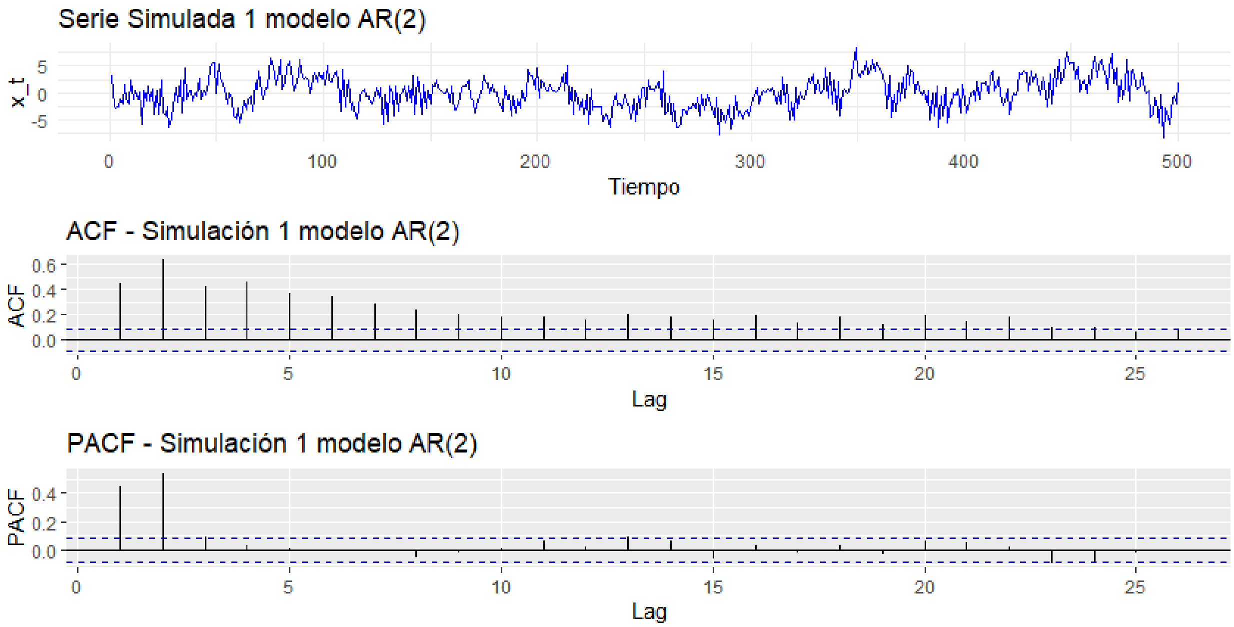

Figura 8: Simulación 1 modelo 2

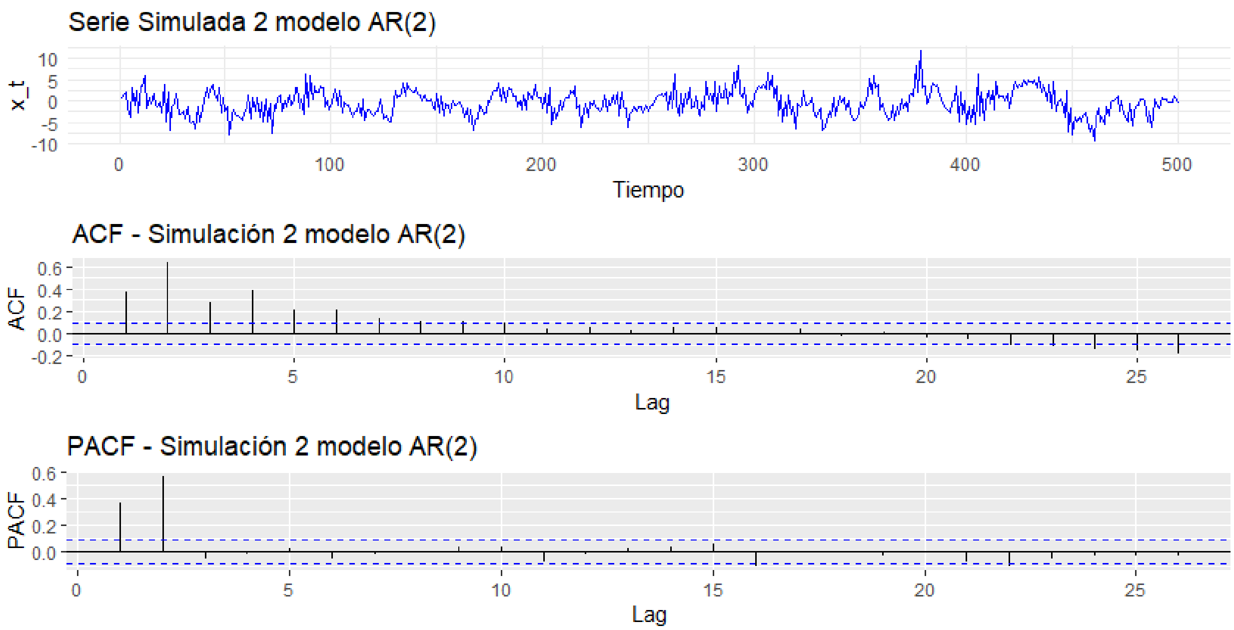

Figura 9: Simulación 2 modelo 2

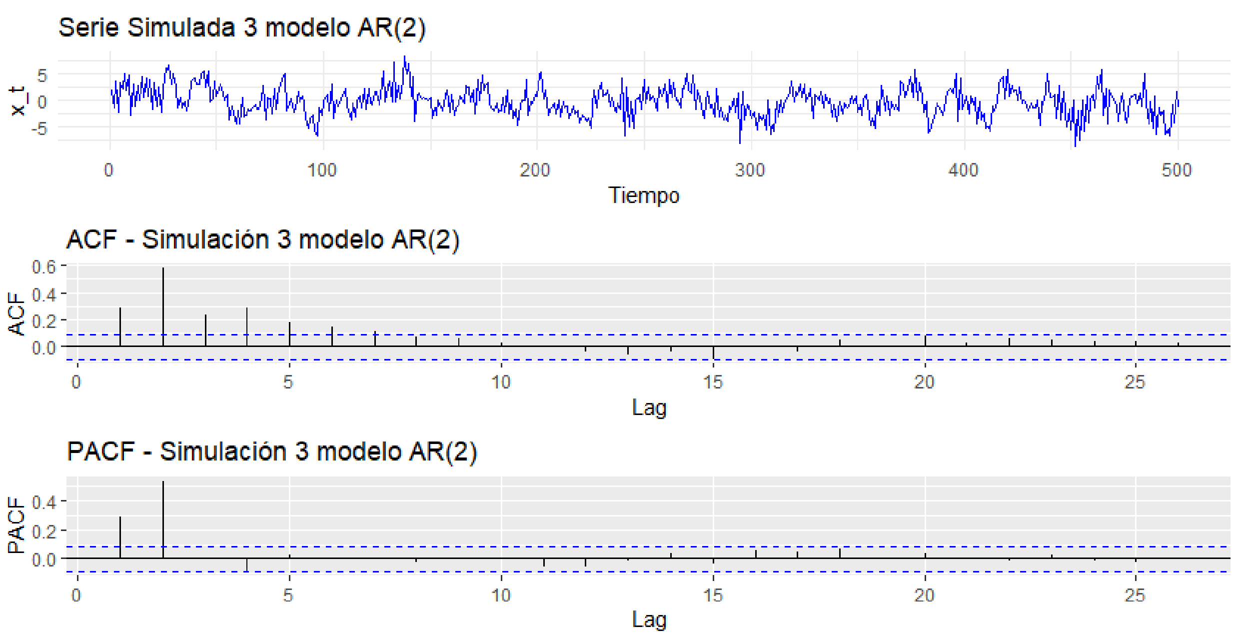

Figura 10: Simulación 3 modelo 2

Los graficos de las 3 simulaciones resultan sumamente parecidos, no notandoce una gran diferencia entre ellos, de igual manera en los 3 graficos ACF se nota un descreimiento lento, mientras que en los tres PACF se ve claramente un truncamiento en 2 lo que corresponde con lo esperado para un modelo $AR(2)$, de hecho por los graficos resultaría posible pensar en el valor estimado de los coeficientes auto regresivos.

Ahora los resultados del ajuste del modelo se presentan a continuación :

Simulación	AR1 Estimado	AR2 Estimado	σ^2 Estimado	AIC
1	0.1951	0.5301	2.075	1792.71
2	0.1536	0.5355	2.318	1848.13
3	0.1756	0.5770	2.367	1858.64

Cuadro 2: Resumen de las estimaciones de los modelos ARMA(2,0) para las tres simulaciones.

En este caso como vemos en la tabla **??** los tres modelos resultantes para cada muestra resulta sumamente parecido en términos de todas las estimaciones tanto de los coeficientes autor regresivos como la varianza.

Ahora veamos el modelo 3:

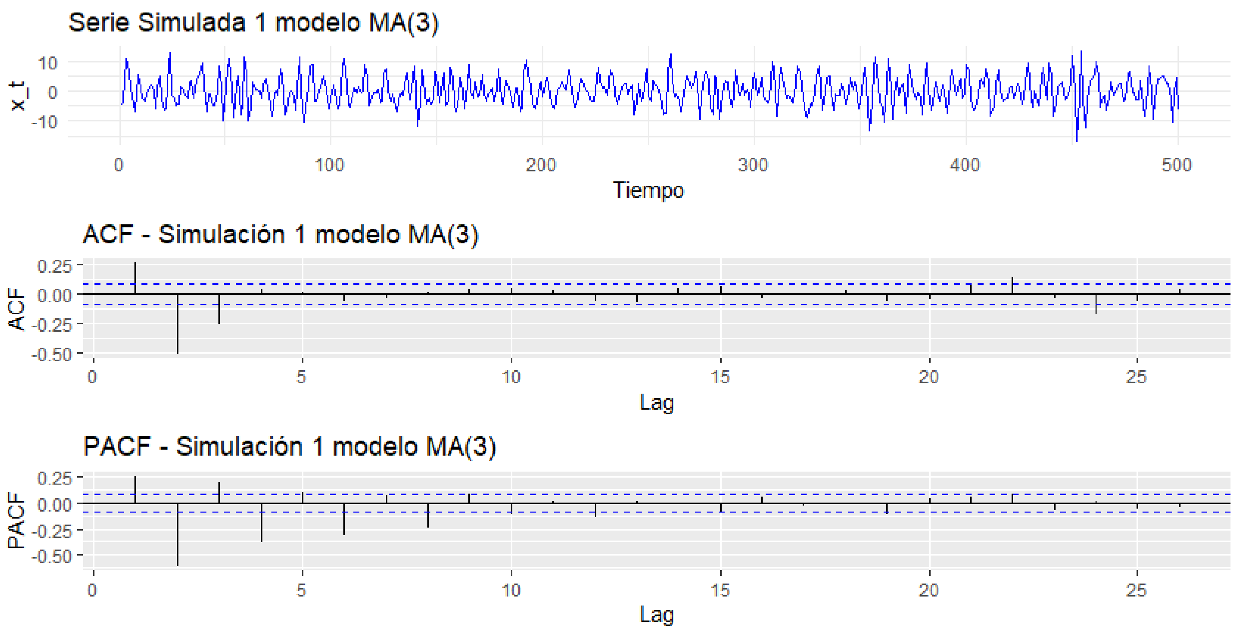

Figura 11: Simulación 1 modelo 3

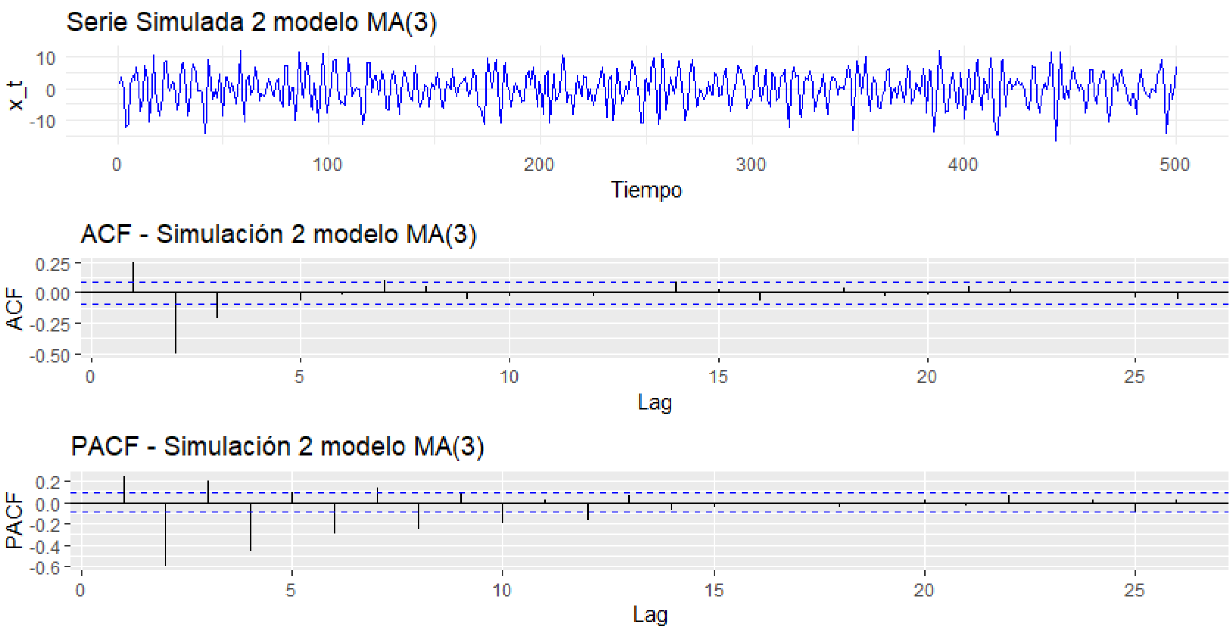

Figura 12: Simulación 2 modelo 3

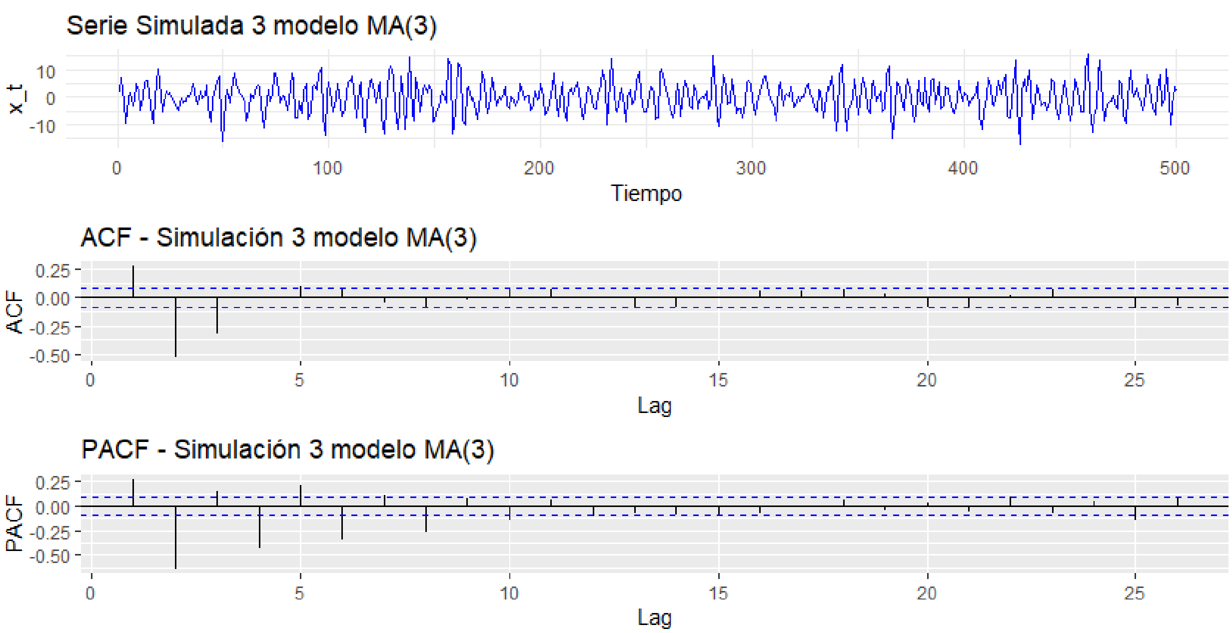

Figura 13: Simulación 3 modelo 3

De los anteriores graficos notamos un comportamiento similar de forma completamente esperable, por otra parte en los graficos ACF se puede notar el mismo comportamiento un truncamiento en 3 que corresponde con el comportamiento de un modelo $MA(3)$ puro, mientras que en las PACF se denota un decaimiento.

El resumen del comportamiento del ajuste del modelo se ve a continuación en la siguiente tabla:

Simulación	MA1 Estimado	MA2 Estimado	MA3 Estimado	σ^2 Estimado	AIC
1	0.8959	-0.7792	-0.7772	8.252	2489.02
2	0.8634	-0.8100	-0.7678	9.581	2563.62
3	0.8879	-0.8068	-0.8008	9.582	2563.99

Cuadro 3: Resumen de las estimaciones de los modelos ARMA(0,3) para las tres simulaciones.

Vemos en la table 3 los valores estimados de los coeficientes y la varianza resultan muy buenos y muy cercanos sin importar la muestra con la que fueron estimados.

4. Punto 4

Para cada una de las cuatro series de tiempo de la Hoja1 del libro archivo Excel llamado "HW02-DatosPunto4.xls", haciendo uso de la metodología de Box-Jenkins: (i) estime el modelo ARMA correspondiente, (ii) evalue los residuales del modelo y (iii) realice pronósticos hasta el horizonte $h = 6$. Por último, determine si el proceso estimado es estacionario e invertible.

Para lo siguiente presentamos para cada una de las series su gráfico junto con las gráficas ACF y PACF:

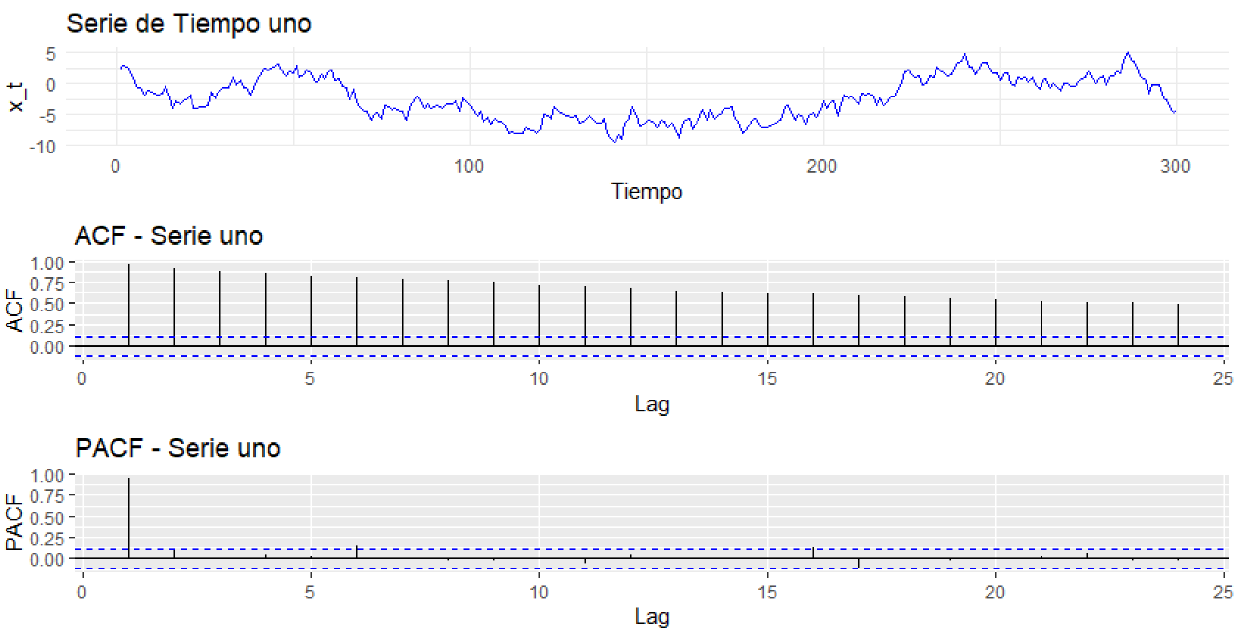

Figura 14: Serie uno

Para esta primera serie como podemos por en el gráfico 14 no se evidencian problemas de volatilidad y parece estar sobre un valor constante, también es distinguible un comportamiento en la función ACF donde decrece lentamente, mientras que en la función PACF se ve un bastante claro truncamiento en 1 lo que nos da el indicio de que podríamos estar ante un modelo $AR(1)$ pero eso se determinara más adelante.

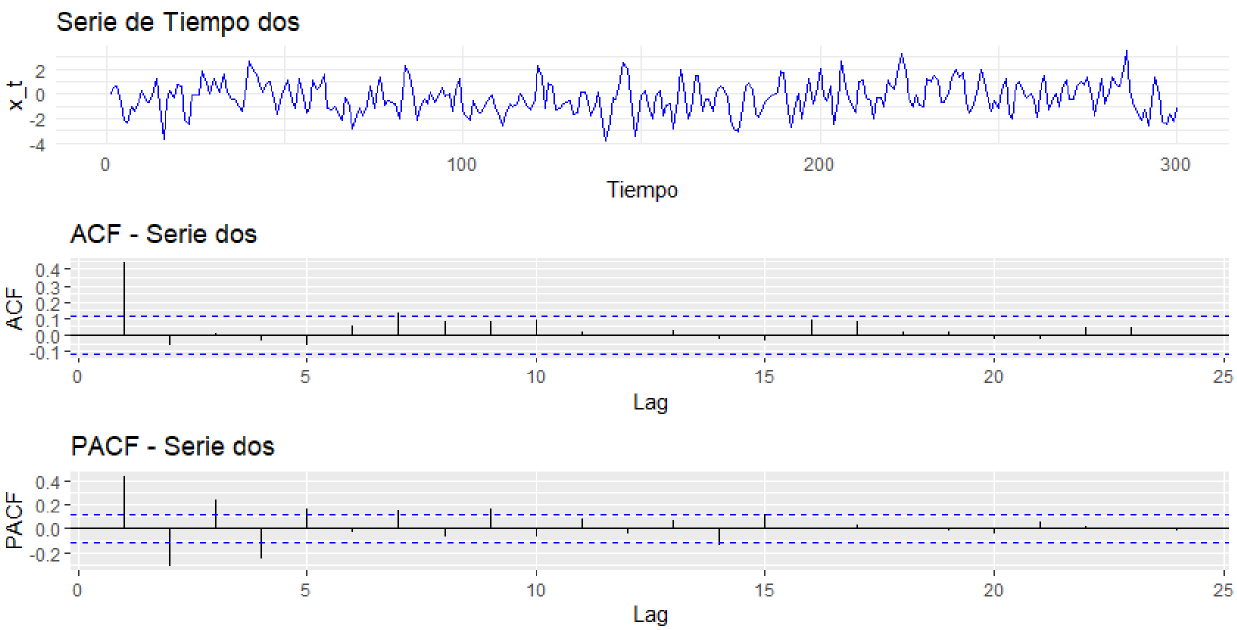

Figura 15: Serie dos

Para esta primera serie como podemos por en el gráfico 15 no se evidencian problemas de volatilidad y parece estar sobre un valor constante, también es distinguible un comportamiento en la función PACF donde decrece lentamente, mientras que en la función ACF se ve un bastante claro truncamiento en 1 lo que nos da el indicio de que podríamos estar ante un modelo $MA(1)$ pero eso se determinara más adelante.

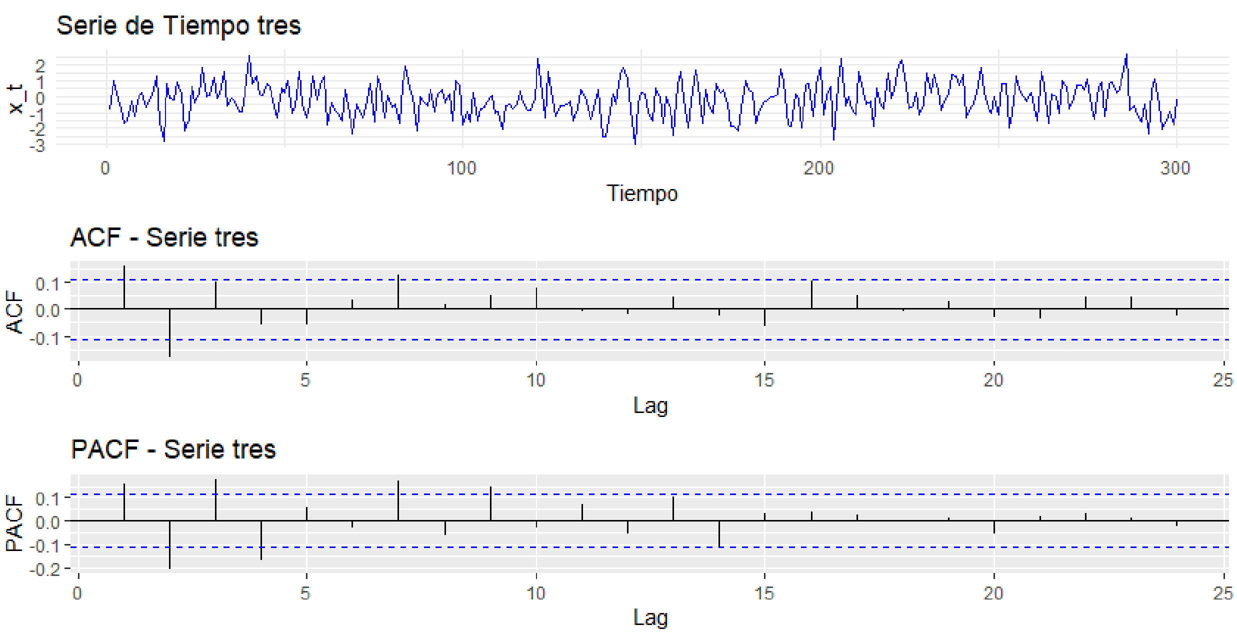

Figura 16: Serie tres

Para esta primera serie como podemos por en el gráfico 16 no se evidencian problemas de volatilidad y parece estar sobre un valor constante, en este caso no se puede decir con seguridad si existe algún comportamiento en las funciones ACF y PACF por lo que desde este gráfico aun no podemos aventurarnos a tener algún candidato a modelo $ARMA$.

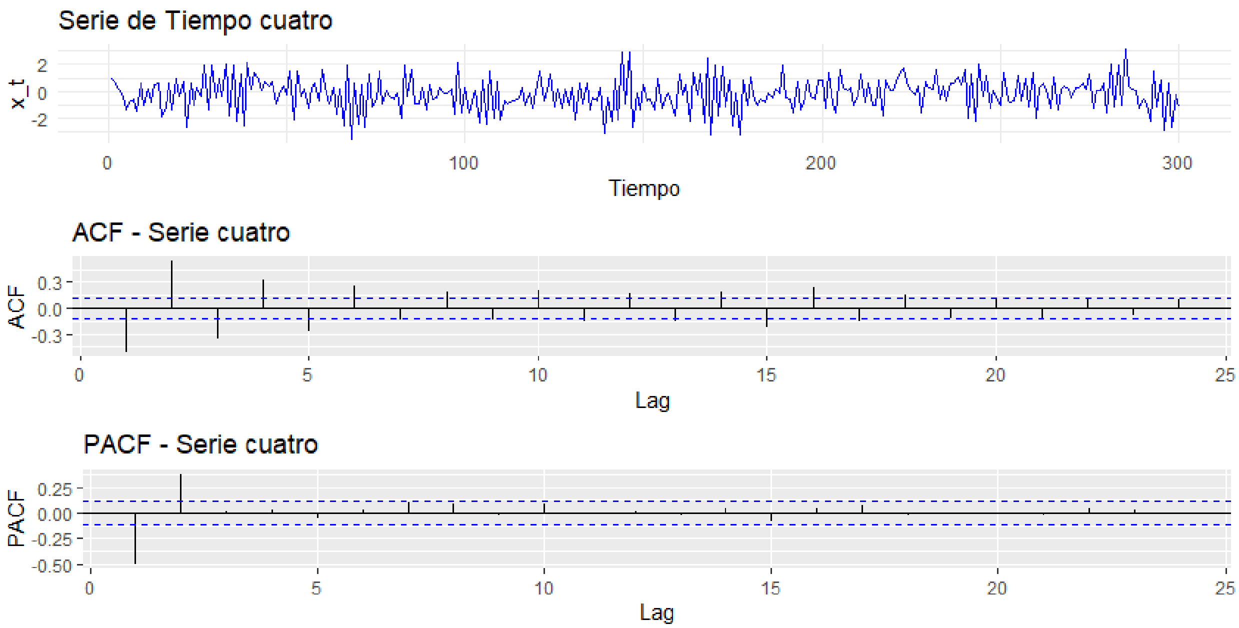

Figura 17: Serie cuatro

Para esta primera serie como podemos por en el gráfico 17 no se evidencian problemas de volatilidad y parece estar sobre un valor constante, también es distinguible un comportamiento en la función ACF donde decrece lentamente, mientras que en la función PACF se ve un bastante claro truncamiento en 2 lo que nos da el indicio de que podríamos estar ante un modelo $AR(2)$ pero eso se determinara más adelante.

Con lo anterior con cada una de las 4 series se utilizo la función EACF para ver por este método el grado con el que poder ajustar en modelo $ARMA(p, q)$ a cada una de estas series, por otro lado también se utilizo la función autoarima para realizar las misma tarea.

- Los resultados de estos dos métodos para para la serie uno es un modelo $ARMA(1, 1)$ esto tanto gráficamente por medio de la función EACF como por medio de la función autoarima restringiendo el valor de la diferencia regular a 0, bajo este modelo se estimaron el siguiente modelo:

$$x_t = -2{,}2012 + 0{,}9669x_{t-1} + w_t - 0{,}1393w_{t-1}$$

  con esto vemos que el modelo cuenta con media, en $-2{,}2012$ y con estos modelos podemos, pasar a la evaluación de residuales donde al aplicarles la pruna Ljung-Box llegamos a un p-valor de 0,42 por lo que no tenemos suficiente evidencia estadística para rechazar la hipótesis de diferencia de correlaciones temporales entre los resagos con lo que verificamos el supuesto de covarianza entre los errores. por otro lado veamos la siguiente gráfica 18 de evaluación de los residuales:

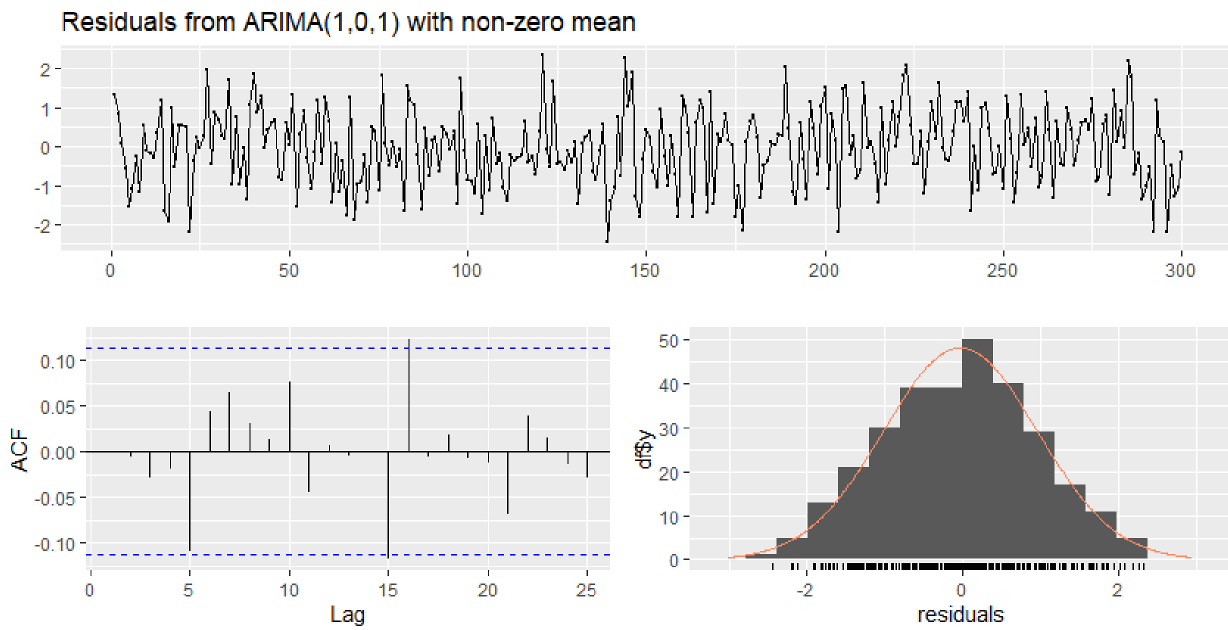

Figura 18: Evaluación residuales serie uno

Con la anterior gráfica 18 vemos como los residuales parecen tener un buen valor medio en 0 y comparados con la distribución normal estándar cumplen de forma razonable el supuesto de normalidad que se usa para la estimación de los parámetros por máxima verosimilitud.

Con lo anterior veamos un gráfico de los pronósticos:

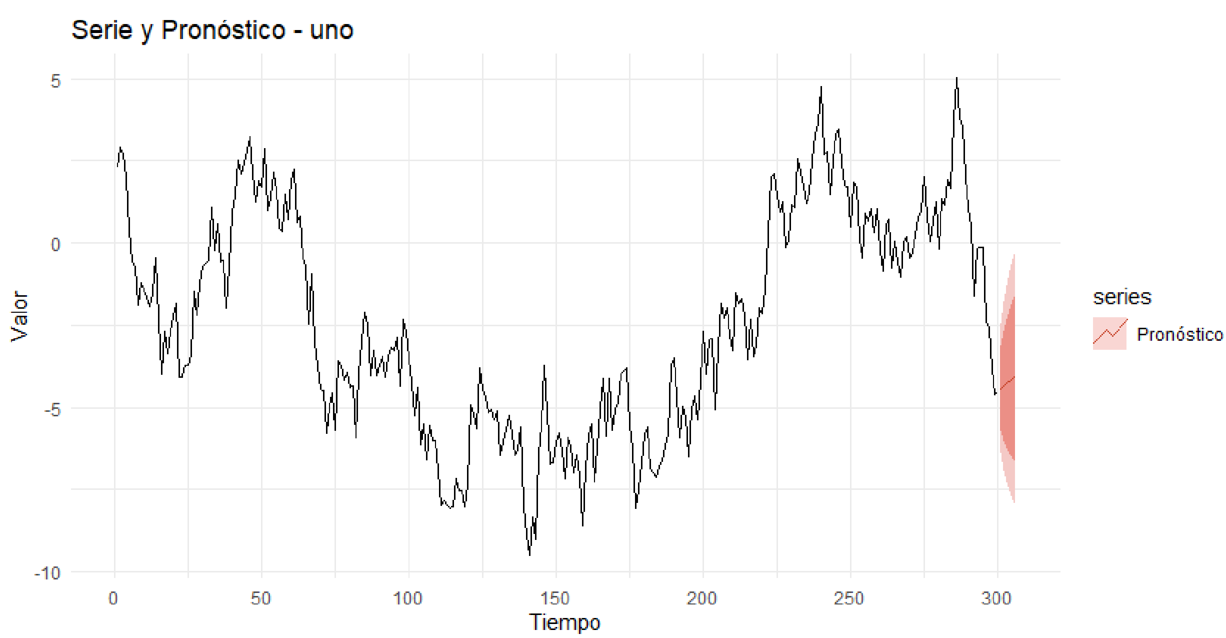

Figura 19: Pronósticos serie uno

por ultimo al usar el polinomio al que se llego anteriormente y calculando sus raíces llegamos a la conclusión de que el modelo es tanto estacionario como invertible.

- Los resultados de estos dos métodos para para la serie dos es un modelo $MA(1)$ esto tanto gráficamente por medio de la función EACF como por medio de la función autoarima restringiendo el valor de la diferencia regular a 0, bajo este modelo se estimaron el siguiente modelo:

$$x_t = -0{,}2767 + 0{,}9099 w_{t-1} + w_t$$

con esto vemos que el modelo cuenta con media, en $-0{,}2767$ y con estos modelos podemos, pasar a la evaluación de residuales donde al aplicarles la pruna Ljung-Box llegamos a un p-valor de 0,327 por lo que no tenemos suficiente evidencia estadística para rechazar la hipótesis de diferencia de correlaciones temporales entre los resagos con lo que verificamos el supuesto de covarianza entre los errores. por otro lado veamos la siguiente gráfica 20 de evaluación de los residuales:

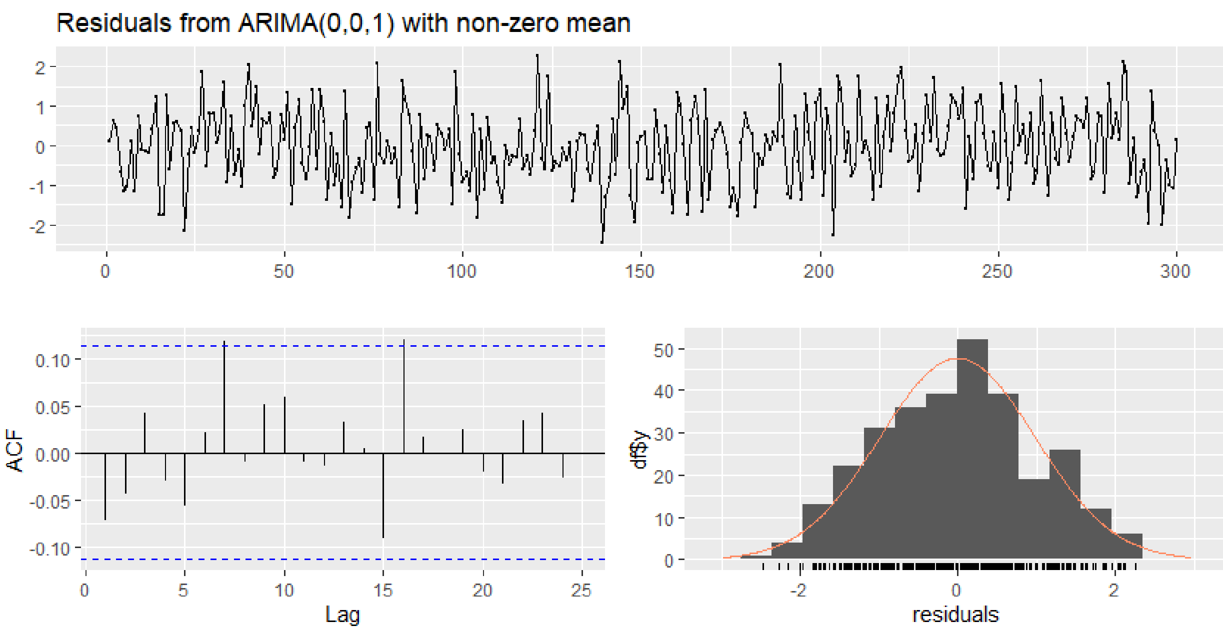

Figura 20: Evaluación residuales serie dos

Con la anterior gráfica 20 vemos como los residuales parecen tener un buen valor medio en 0 y comparados con la distribución normal estándar cumplen de forma razonable el supuesto de normalidad que se usa para la estimación de los parámetros por máxima verosimilitud.

Con lo anterior veamos un gráfico de los pronósticos:

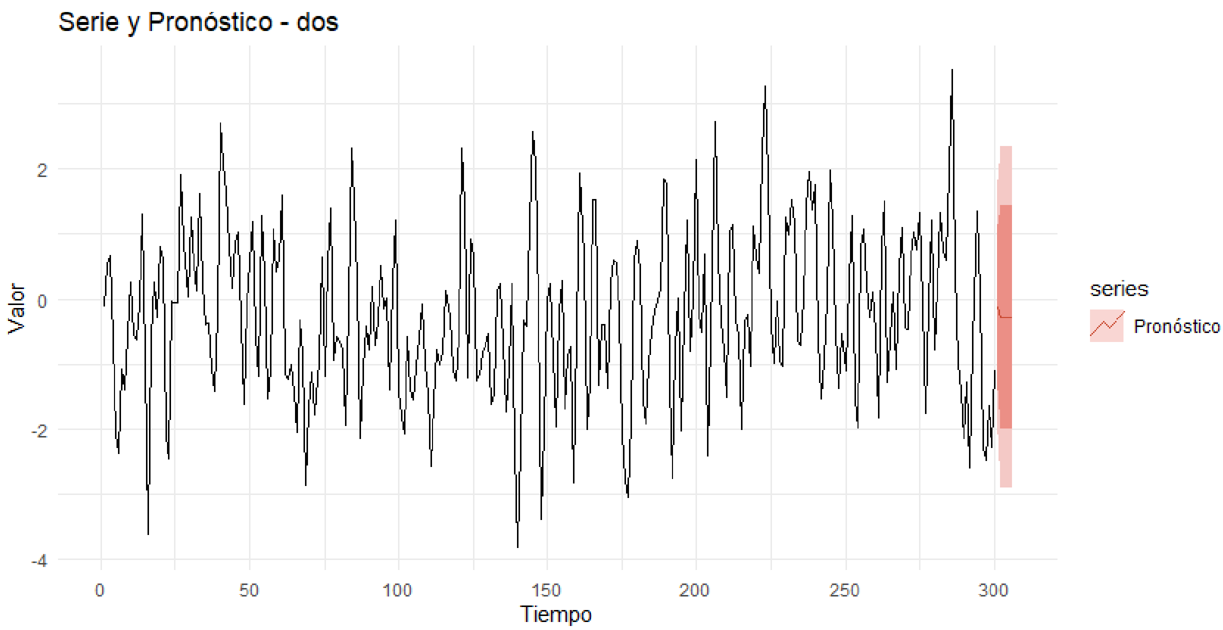

Figura 21: Pronósticos serie dos

por ultimo al usar el polinomio al que se llego anteriormente y calculando sus raíces llegamos a la conclusión de que el modelo es tanto estacionario como invertible.

- Los resultados de estos dos métodos para para la serie tres es un modelo $ARMA(1,2)$ esto tanto gráficamente por medio de la función EACF como por medio de la función autoarima restringiendo el valor de la diferencia regular a 0, bajo este modelo se estimaron el siguiente modelo:

$$x_t = -0{,}1876 - 0{,}4194x_{t-1} + w_t - 0{,}7560w_{t-1} - 0{,}1444w_{t-2}$$

con esto vemos que el modelo cuenta con media, en $-0{,}1876$ y con estos modelos podemos, pasar a la evaluación de residuales donde al aplicarles la pruna Ljung-Box llegamos a un p-valor de 0,3226 por lo que no tenemos suficiente evidencia estadística para rechazar la hipótesis de diferencia de correlaciones temporales entre los resagos con lo que verificamos el supuesto de covarianza entre los errores. por otro lado veamos la siguiente gráfica 22 de evaluación de los residuales:

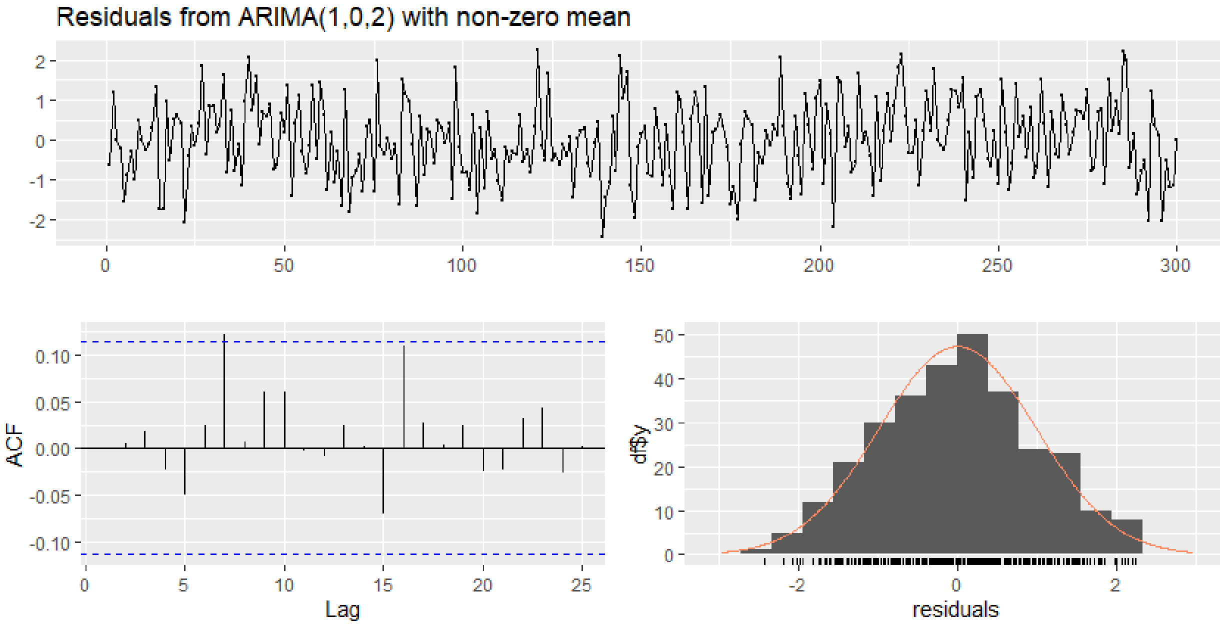

Figura 22: Evaluación residuales serie tres

Con la anterior gráfica 22 vemos como los residuales parecen tener un buen valor medio en 0 y comparados con la distribución normal estándar cumplen de forma razonable el supuesto de normalidad que se usa para la estimación de los parámetros por máxima verosimilitud.

Con lo anterior veamos un gráfico de los pronósticos:

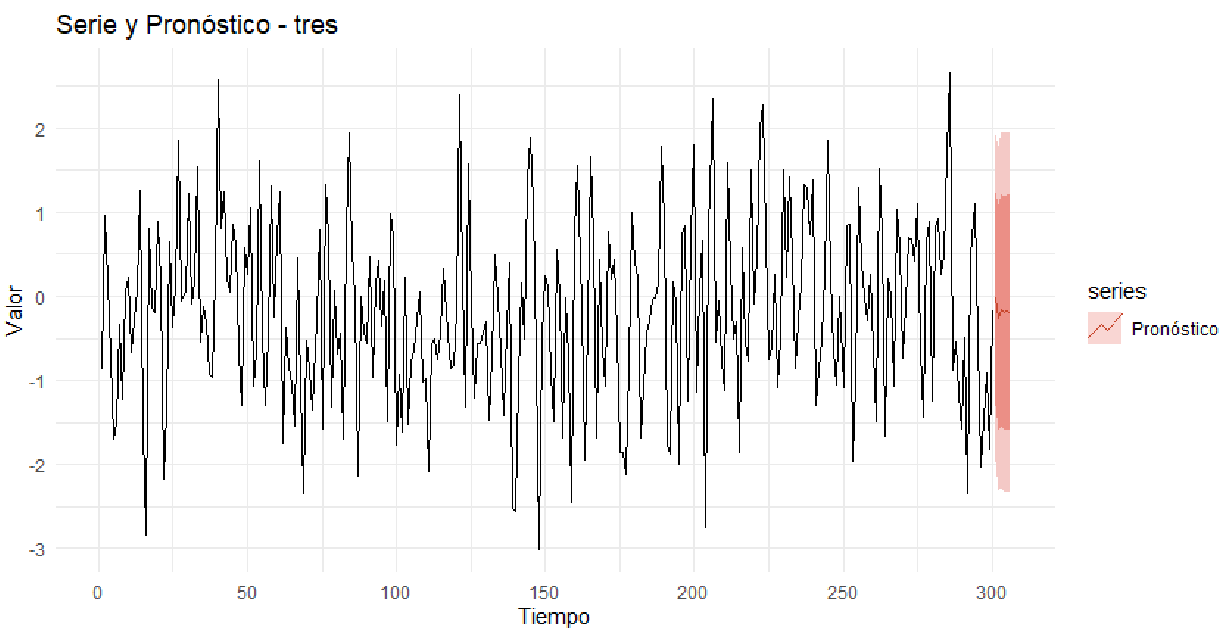

Figura 23: Pronósticos serie tres

por ultimo al usar el polinomio al que se llego anteriormente y calculando sus raíces llegamos a la conclusión de que el modelo es tanto estacionario como invertible.

- Los resultados de estos dos métodos para para la serie cuatro es un modelo $AR(2)$ esto tanto gráficamente por medio de la función EACF como por medio de la función autoarima restringiendo el valor de la diferencia regular a 0, bajo este modelo se estimaron el siguiente modelo:

$$x_t = -0{,}1714 + -0{,}3131x_{t-1} + 0{,}3763x_{t-2} + w_t$$

con esto vemos que el modelo cuenta con media, en $-0{,}1714$ y con estos modelos podemos, pasar a la evaluación de residuales donde al aplicarles la pruna Ljung-Box llegamos a un p-valor de 0,4069 por lo que no tenemos suficiente evidencia estadística para rechazar la hipótesis de diferencia de correlaciones temporales entre los resagos con lo que verificamos el supuesto de covarianza entre los errores. por otro lado veamos la siguiente gráfica 24 de evaluación de los residuales:

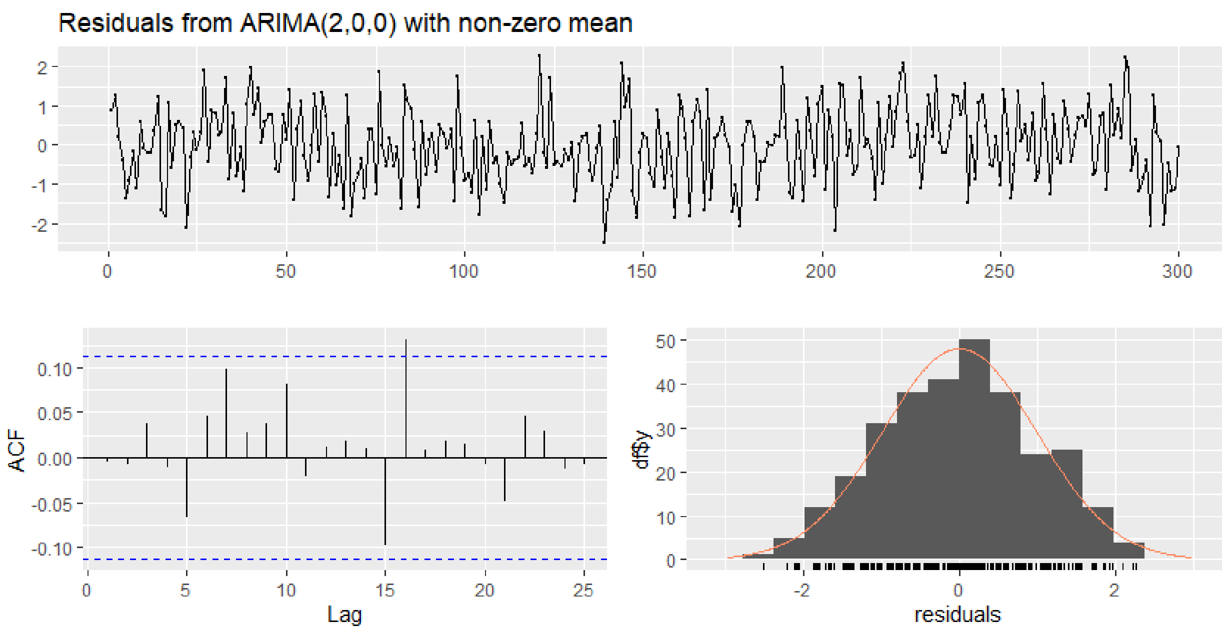

Figura 24: Evaluación residuales serie cuatro

Con la anterior gráfica 24 vemos como los residuales parecen tener un buen valor medio en 0 y comparados con la distribución normal estándar cumplen de forma razonable el supuesto de normalidad que se usa para la estimación de los parámetros por máxima verosimilitud.

Con lo anterior veamos un gráfico de los pronósticos:

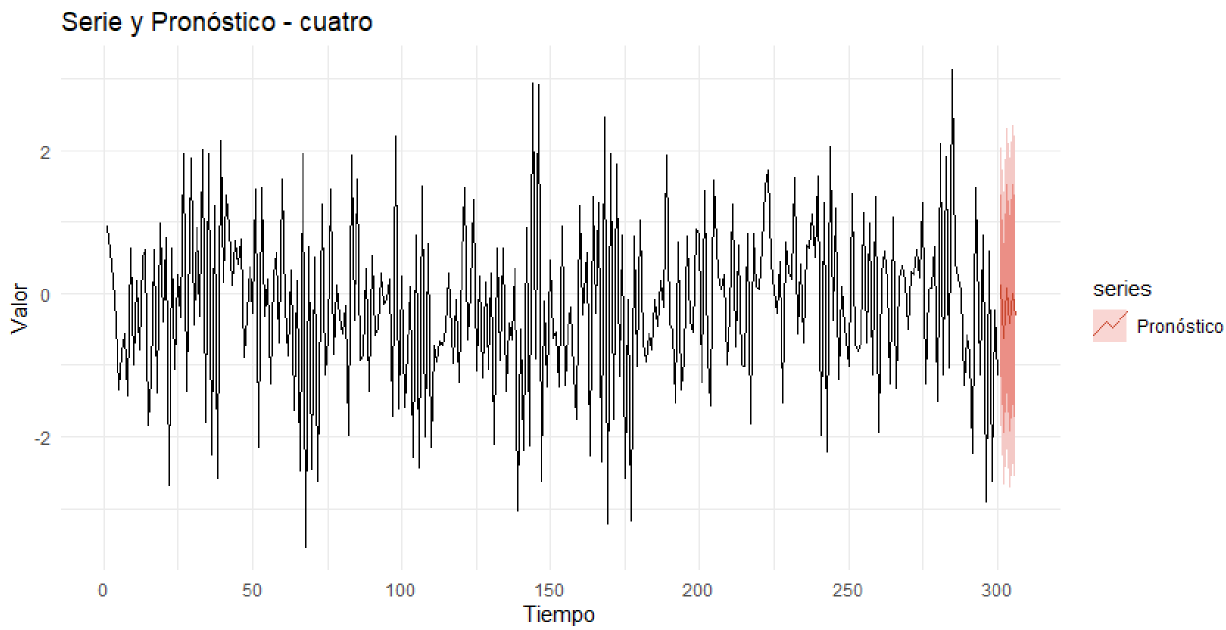

Figura 25: Pronósticos serie cuatro

por ultimo al usar el polinomio al que se llego anteriormente y calculando sus raíces llegamos a la conclusión de que el modelo es tanto estacionario como invertible.

Anexos

```
##taller 2 jramirezda
#cargar paquetes
# Crear un dataframe con los datos
produccion <- data.frame(
  A o = 1947:2023,
  Unidades = c(3, 5, 21, 26, 33, 44, 57, 58, 61, 81, 113, 183, 248, 306, 441, 493, 589, 654, 740, 665, 706, 729, 619, 928,
               1246, 1844, 1772, 1436, 1337, 1426, 1798, 1939, 2221, 2470, 2565, 2209, 2366, 2842, 3288, 3640, 3902, 4001,
               3821, 4293, 4487, 3384, 2345, 2671, 3144, 3350, 3518, 3652, 3669, 3946, 4305, 4276, 4238, 4975, 5409, 5671,
               6465, 6587, 6250, 6573, 7195, 7405, 7000, 7255, 7664, 8014, 8398, 9251, 10131, 9119, 11115, 13221, 13663)
)

# Crear un objeto de serie de tiempo
ts_produccion <- ts(produccion$Unidades, start = 1947, frequency = 1)

# Graficar la serie de tiempo
library(ggplot2)
ggplot(produccion, aes(x = A o , y = Unidades)) +
  geom_line(color = "blue") +
  geom_point(color = "red") +
```

```r
  labs(title = "Producci n Anual de Unidades de ferraris.",
       x = "A o",
       y = "Unidades Producidas") +
  theme_minimal()

# Cargar y filtrar los datos de manchas solares
library(readr)
manchas_solars <- read_csv("daily_sunspots_time_series_1850-01_2024-05.csv")
manchas_solars <- manchas_solars %>% select(date, counts) %>% filter(date >= "2000-01-01")

# Convertir a serie de tiempo
ts_manchas <- ts(manchas_solars$counts, start = c(2000, 1), frequency = 365)

# Graficar la serie de tiempo de manchas solares
ggplot(manchas_solars, aes(x = date, y = counts)) +
  geom_line(color = "blue") +
  geom_point(color = "red", size = 0.5) +
  labs(title = "Manchas Solares Diarias",
       x = "Fecha",
       y = "Conteo de Manchas Solares") +
  theme_minimal()


###promedios moviles

# Calcular promedios m viles
manchas_solars <- manchas_solars %>%
  mutate(
    MA5 = rollmean(counts, k = 30, fill = NA, align = "right"),
    MA10 = rollmean(counts, k = 180, fill = NA, align = "right"),
    MA15 = rollmean(counts, k = 365, fill = NA, align = "right")
  )
produccion <- produccion %>%
  mutate(
    MA3 = rollmean(Unidades, k = 3, fill = NA, align = "right"),
    MA4 = rollmean(Unidades, k = 5, fill = NA, align = "right"),
    MA5 = rollmean(Unidades, k = 10, fill = NA, align = "right")
  )

# Graficar la serie de tiempo de manchas solares con promedios m viles
ggplot(manchas_solars, aes(x = date)) +
  geom_line(aes(y = counts, color = "Original")) +
  geom_line(aes(y = MA5, color = "MA 30 d as")) +
  geom_line(aes(y = MA10, color = "MA 180 d as")) +
  geom_line(aes(y = MA15, color = "MA 365 d as")) +
  labs(title = "Manchas Solares Diarias con Promedios M viles",
       x = "Fecha",
       y = "Conteo de Manchas Solares") +
  theme_minimal() +
  scale_color_manual(values = c("Original" = "aquamarine", "MA 30 d as" = "red", "MA 180
    d as" = "green", "MA 365 d as" = "purple"))

# Graficar la serie de tiempo de producci n con promedios m viles
ggplot(produccion, aes(x = A o)) +
```

```
  geom_line(aes(y = Unidades, color = "Original")) +
  geom_line(aes(y = MA3, color = "MA 3 a os")) +
  geom_line(aes(y = MA4, color = "MA 5 a os")) +
  geom_line(aes(y = MA5, color = "MA 10 a os")) +
  labs(title = "Producci n Anual con Promedios M viles",
       x = "A o",
       y = "Unidades Producidas") +
  theme_minimal() +
  scale_color_manual(values = c("Original" = "aquamarine", "MA 3 a os" = "red", "MA 5 a os" = "green", "MA 10 a os" = "purple"))

# punto 3
library(forecast)
library(ggplot2)
library(tseries)
library(gridExtra)

set.seed(123)  # Fijar semilla para reproducibilidad
n <- 500       # N mero de observaciones
simulations <- list()
#arma1,1
for (i in 1:3) {
  # Simulaci n del modelo ARMA(1,1)
  ar1 <- 0.9
  ma1 <- -0.9
  Wt <- rnorm(n, mean = 0, sd = 1)
  xt <- numeric(n)
  xt[1] <- Wt[1]

  for (t in 2:n) {
    xt[t] <- ar1 * xt[t - 1] + Wt[t] + ma1 * Wt[t - 1]
  }

  # Guardar resultados
  simulations[[i]] <- xt

  # Gr fico de la serie simulada
  p1 <- ggplot(data.frame(t = 1:n, x = xt), aes(x = t, y = x)) +
    geom_line(color = "blue") +
    labs(title = paste("Serie Simulada", i,"modelo 1"), x = "Tiempo", y = "x_t") +
    theme_minimal()

  # ACF y PACF
  p2 <- ggAcf(xt) + ggtitle(paste("ACF - Simulaci n", i,"modelo 1"))
  p3 <- ggPacf(xt) + ggtitle(paste("PACF - Simulaci n", i, "modelo 1"))

  # Ajuste del modelo ARMA(1,1)
  modelo_ajustado <- arima(xt, order = c(1, 0, 1))
  print(summary(modelo_ajustado))

  # Mostrar gr ficos
  grid.arrange(p1, p2, p3, nrow = 3)
}

```

```
#ar 2
for (i in 1:3) {
  # Simulaci n del modelo AR(2)
  ar1 <- 0.2
  ar2 <- 0.55
  Wt <- rnorm(n, mean = 0, sd = sqrt(2.25))
  xt <- numeric(n)
  xt[1] <- Wt[1]  # Correcci n del  ndice  inicial
  xt[2] <- ar1 * xt[1] + Wt[2]  # Definir segundo valor

  for (t in 3:n) {
    xt[t] <- ar1 * xt[t - 1] + ar2 * xt[t - 2] + Wt[t]
  }

  # Guardar resultados
  simulations[[i]] <- xt

  # Gr fico de la serie simulada
  p1 <- ggplot(data.frame(t = 1:n, x = xt), aes(x = t, y = x)) +
    geom_line(color = "blue") +
    labs(title = paste("Serie Simulada", i, "modelo AR(2)"), x = "Tiempo", y = "x_t") +
    theme_minimal()

  # ACF y PACF
  p2 <- ggAcf(xt) + ggtitle(paste("ACF - Simulaci n", i, "modelo AR(2)"))
  p3 <- ggPacf(xt) + ggtitle(paste("PACF - Simulaci n", i, "modelo AR(2)"))

  # Ajuste del modelo AR(2)
  modelo_ajustado <- arima(xt, order = c(2, 0, 0))
  print(summary(modelo_ajustado))

  # Mostrar gr ficos
  grid.arrange(p1, p2, p3, nrow = 3)
}

#ma 3
for (i in 1:3) {
  # Simulaci n del modelo AR(2)
  ma1 <- 0.9
  ma2 <- -0.8
  ma3 <- -0.8
  Wt <- rnorm(n, mean = 0, sd = sqrt(9))
  xt <- numeric(n)
  xt[1] <- Wt[1]  # Correcci n del  ndice  inicial
  xt[2] <- ma1 * Wt[1] + Wt[2]  # Definir segundo valor
  xt[3] <- ma1 * Wt[2] +ma2* Wt[1] +Wt[3]

  for (t in 4:n) {
    xt[t] <- ma1 * Wt[t - 1] + ma2 * Wt[t - 2] +ma3* Wt[t-3] + Wt[t]
  }

  # Guardar resultados
  simulations[[i]] <- xt

```

```r
  # Gr fico de la serie simulada
  p1 <- ggplot(data.frame(t = 1:n, x = xt), aes(x = t, y = x)) +
    geom_line(color = "blue") +
    labs(title = paste("Serie Simulada", i, "modelo MA(3)"), x = "Tiempo", y = "x_t") +
    theme_minimal()

  # ACF y PACF
  p2 <- ggAcf(xt) + ggtitle(paste("ACF - Simulaci n", i, "modelo MA(3)"))
  p3 <- ggPacf(xt) + ggtitle(paste("PACF - Simulaci n", i, "modelo MA(3)"))

  # Ajuste del modelo AR(2)
  modelo_ajustado <- arima(xt, order = c(0, 0, 3))
  print(summary(modelo_ajustado))

  # Mostrar gr ficos
  grid.arrange(p1, p2, p3, nrow = 3)
}


#punto 3
library(forecast)
library(ggplot2)
library(tseries)
library(gridExtra)
library(readxl)
library(TSA)  # Para la funci n EACF

# Cargar datos desde Excel
Datos <- read_excel("HW02-DatosPunto4.xls")

# Convertir las series en objetos de serie de tiempo
timeseries_list <- list(
  uno = ts(Datos$serie1, frequency = 1),
  dos = ts(Datos$serie2, frequency = 1),
  tres = ts(Datos$serie3, frequency = 1),
  cuatro = ts(Datos$serie4, frequency = 1)
)

plot_series <- function(series, series_name) {
  n <- length(series)
  df <- data.frame(t = 1:n, x = series)

  # Gr fico de la serie de tiempo
  p1 <- ggplot(df, aes(x = t, y = x)) +
    geom_line(color = "blue") +
    labs(title = paste("Serie de Tiempo", series_name), x = "Tiempo", y = "x_t") +
    theme_minimal()

  # Gr fico de ACF
  p2 <- ggAcf(series) + ggtitle(paste("ACF - Serie", series_name))

  # Gr fico de PACF
  p3 <- ggPacf(series) + ggtitle(paste("PACF - Serie", series_name))

```

```r
  # Combinar los gr ficos en una sola figura
  grid.arrange(p1, p2, p3, nrow = 3)
}

# Generar los gr ficos para cada serie de tiempo
for (i in seq_along(timeseries_list)) {
  series_name <- names(timeseries_list)[i]
  plot_series(timeseries_list[[i]], series_name)
}

##arma
timeseries_results <- list()

for (name in names(timeseries_list)) {
  ts_data <- timeseries_list[[name]]

  # Identificaci n del modelo ARMA con EACF
  print(paste("EACF para la serie:", name))
  eacf_result <- eacf(ts_data)
  print(eacf_result)

  # Ajuste del modelo ARMA (sin diferenciaci n regular)
  modelo_ajustado <- auto.arima(ts_data, d = 0, seasonal = FALSE)
  print(paste("Modelo ajustado para:", name))
  print(summary(modelo_ajustado))

  # Evaluaci n de residuales
  checkresiduals(modelo_ajustado)

  # Pron sticos h=6
  forecast_h <- forecast(modelo_ajustado, h = 6)
  print(forecast_h)

  # Graficar la serie original y el pron stico
  p1 <- autoplot(ts_data) +
    autolayer(forecast_h, series = "Pron stico", PI = TRUE) +
    labs(title = paste("Serie y Pron stico -", name), x = "Tiempo", y = "Valor") +
    theme_minimal()

  # Evaluaci n de estacionariedad e invertibilidad
  estacionario <- all(Mod(polyroot(c(1, -modelo_ajustado$coef[grepl("ar", names(modelo_
    ajustado$coef))]))) > 1)
  invertible <- all(Mod(polyroot(c(1, modelo_ajustado$coef[grepl("ma", names(modelo_ajustado
    $coef))]))) > 1)

  print(paste("La serie", name, "es estacionaria:", estacionario))
  print(paste("La serie", name, "es invertible:", invertible))

  # Guardar resultados
  timeseries_results[[name]] <- list(
    eacf = eacf_result,
    modelo = modelo_ajustado,
    forecast = forecast_h,
    estacionario = estacionario,
```

```r
    invertible = invertible,
    plot = p1
  )

  print(timeseries_results[[name]]$plot)
}
```

Listing 1: Código de R para series de tiempo tarea 2