

Tarea 2

EYP3907 - Series de Tiempo

Sebastián Celaya

Camila Echeverría

Francisca Vilca

Introducción

Utilizando una base de datos que contiene información sobre el ancho de los anillos de árboles pertenecientes a la especie *Pino Silvestre*, que puede encontrarse en el siguiente [link](#), ajustaremos un modelo ARMA y realizaremos diversos procedimientos para comprobar su ajuste.

Análisis exploratorio

La figura 1a muestra los valores del ancho del anillo registrados entre los años 1721 y 1889.

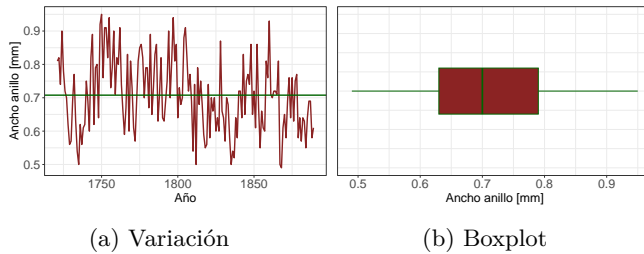


Figura 1: Sobre la serie de tiempo

Gracias a la figura 1b, es posible ver que la mediana de estos datos se encuentra cercana a 0.7 y que no tenemos datos atípicos, aunque la segunda mitad de las observaciones parecieran estar ligeramente más dispersa que la primera.

Luego, en los gráficos de la figura 2 podemos ver la estructura de correlación de los datos. Si bien no se puede detectar estacionalidad a simple vista, las observaciones sí presentan altos niveles de correlación.

Ajuste de un modelo ARMA

A simple vista, de los gráficos de ACF y PACF, apreciamos que nuestro modelo tiene estructura de un ARMA. Por lo que, ayudándonos de la función `auto.arima()`, se nos propone usar un MA(1), pero notamos que este no es

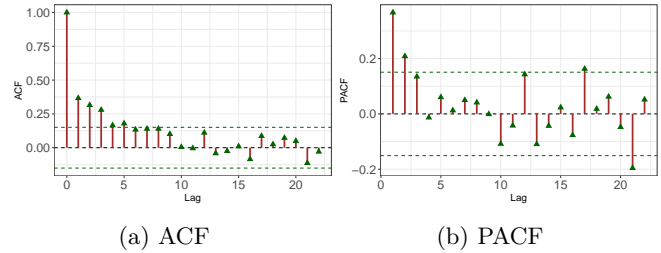


Figura 2: Gráficos de autocorrelación

capaz de capturar toda la estructura de la serie temporal. Luego, de probar distintos modelos, optamos por ajustar un modelo ARMA(1,1), ya que, este presenta el mejor ajuste como se presentará a continuación. Además, el modelo ajustado no considera diferenciación pues, como veremos más adelante, la serie no es estacional.

a) Significancia estadística de los coeficientes del modelo

Al revisar el valor-p asociado a cada coeficiente del modelo, es claro notar que todos son significativos, tal como se muestra en la Tabla 1:

Tabla 1: Resumen de estimaciones

| | Estimación | SE | valor-p |
|------------|------------|--------|---------|
| AR | 0.8367 | 0.0816 | 0.0000 |
| MA | -0.5698 | 0.1209 | 0.0000 |
| Intercepto | 0.7079 | 0.0196 | 0.0000 |

b) Estacionaridad e invertibilidad del modelo ARMA

Una forma sencilla de comprobar la estacionaridad en los datos es con el test de Dickey-Fuller, el cual nos da un valor-p de 0.01 que rechaza la hipótesis nula de que los datos son estacionales. Por otro lado, una forma sencilla de verificar invertibilidad del modelo ARMA(1,1), es de

forma gráfica, comprobando que los coeficientes se encuentren al interior de la circunferencia unitaria, lo que puede ser apreciado en la figura 3:

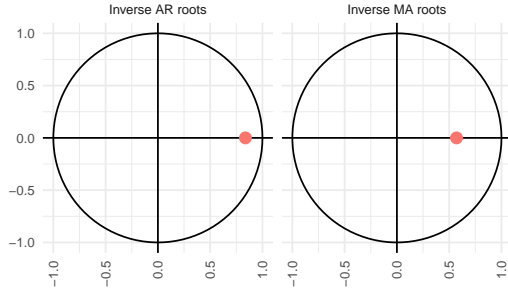
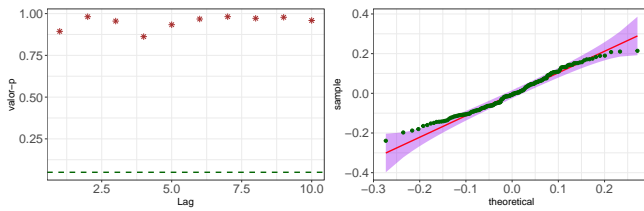


Figura 3: Gráfico de raíces unitarias

c) Test de blancura - homocedasticidad y normalidad de los residuos

Los residuos del modelo deben cumplir estas propiedades. Para ello veremos diferentes test que se le pueden aplicar para comprobar ello:

- La figura 4a, correspondiente a los valores-p del estadístico Ljung-Box nos muestra que los residuos efectivamente corresponden a ruido blanco, es decir, no están correlacionados.
- Luego, al realizar el test de Kolmogorov-Smirnov para evaluar la normalidad, el valor-p obtenido es de 0.27 aproximadamente, por lo que no se rechaza la hipótesis nula: los residuos provienen de una distribución normal. Este ajuste se puede observar en la figura 4b.



(a) Valores-p de Ljung-Box

(b) Normalidad

Figura 4: Revisión de supuestos

- De la misma manera, se obtiene un valor-p de 0.24 al realizar el test Breusch-Pagan para evaluar homocedasticidad, por lo que podemos concluir a favor de esta.

d) ¿Es necesario realizar una transformación de Box-Cox?

Tras visualizar el gráfico de Ljung-Box de figura 4a se observa que *no* es necesario realizar una transformación de Box-Cox. Además si comparamos un modelo con y sin la transformación Box-Cox se puede observar que por AIC y BIC el mejor modelo es el que no utiliza la transformación de Box-Cox, como se logra visualizar en Tabla 2

Tabla 2: Comparación modelos

| Modelo | AIC | BIC |
|-------------|-----------|-----------|
| Sin Box-Cox | -294.0414 | -281.5218 |
| Con Box-Cox | -128.8216 | -116.302 |

Predicciones

a) Predicciones a un paso por el algoritmo Durbin-Levinson (D-L)

El algoritmo Durbin-Levinson, empleado para predicciones a un paso en series temporales, ha arrojado una estimación puntual de 0.6610. Este valor representa la proyección para el próximo periodo en la serie temporal. Acompañada por un intervalo de confianza que oscila entre 0.6067 y 0.7152, esta estimación proporciona una medida de la incertidumbre asociada, permitiendo una evaluación más completa de la fiabilidad de la predicción a corto plazo.

b) ACF empírico vs teórico

Calculando el ACF de un ARMA(1, 1) y generando bandas de confianza utilizando el método Bartlett se generó la figura 5. El análisis del ACF revela notables similitudes entre las autocorrelaciones empíricas y el patrón característico de un modelo ARMA (1,1) expuesto en figura 5.

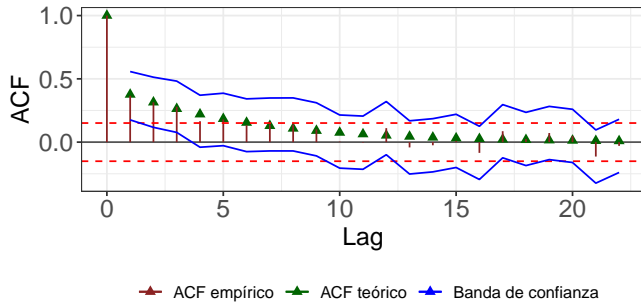


Figura 5: ACF empírico vs teórico

c) Periodograma vs Densidad espectral

En la evaluación del modelo ARMA(1,1) en series temporales, se destacó una notable correspondencia entre las características espectrales predichas y la estructura observada en el Periodograma. La similitud entre los picos de frecuencia en la Densidad Espectral del modelo y las elevaciones en el Periodograma respalda la capacidad del ARMA(1,1) para capturar eficazmente las componentes temporales. Esta convergencia entre la Densidad Espectral y el Periodograma subraya la precisión del modelo al reflejar la variabilidad en los datos, proporcionando una herramienta valiosa para el análisis de la dinámica temporal.

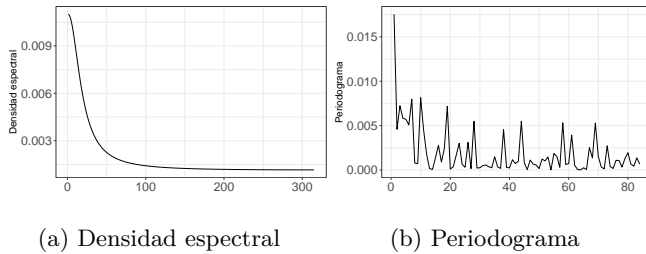


Figura 6: Comparación

Método de Whittle

Ahora si se ajustan los datos con el método de Whittle, también con un modelo ARMA(1,1), obtenemos coeficientes significativos. Además, si usamos la función `forecast()` obtenemos que la predicción de la dimensión del anillo para el siguiente año es de 0.661 y construyendo un intervalo de confianza al 95% es de [0.605,0.716]. Los cuales al compararlos con los obtenidos con el método de Durbin-Levinson vemos que las diferencias son mínimas.

Una forma sencilla de comparar los ajustes es mediante un gráfico de líneas para visualizar que tanto logran extraer la información de los datos originales como en la figura 7

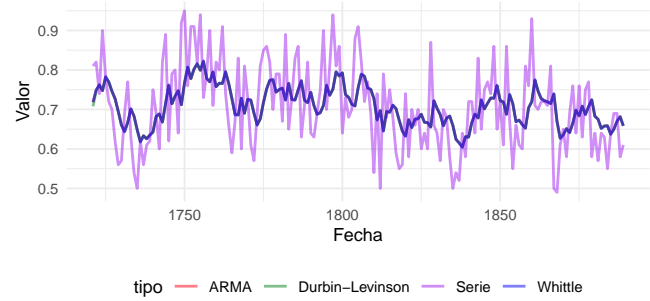


Figura 7: Compara

De la figura 7, vemos claramente que el ajuste mediante el algoritmo de Durbin-Levinson, el de Whittle y el del modelo ARMA son bastantes similares, por lo mismo sus predicciones tambien lo son.

Comparación de ajustes

Finalmente, una forma bastante sencilla de comparar los modelos es mediante medidas para evaluar el error las cuales se presentan en la Tabla 3:

| | ME | RMSE | MAE | MPE | MAPE | MASE |
|---------|---------|------|------|-------|--------|------|
| ARMA | -0.001 | 0.09 | 0.08 | -2.14 | 11.90 | 0.81 |
| D-L | -0.0009 | 0.09 | 0.08 | -2.17 | 11.91 | 0.81 |
| Whittle | -0.001 | 0.09 | 0.08 | 27.09 | 173.63 | 0.81 |

Tabla 3: Medidas de calidad de ajuste

Analizando las medidas de calidad de ajuste, vemos claramente que el modelo ARMA(1,1) es el que comete menos errores para cada una de las medidas. Sin embargo, el modelo propuesto con el algoritmo de Durbin-Levinson, es bastante similar en los errores que comete, excepto que comete un error mayor en el Error Medio. Lamentablemente, el modelo proveniente del método Whittle tiene fallas en las medidas que son asociadas al porcentaje, como el MAPE y el MPE. Lo que nos da indicios para no seleccionarlo como modelo de predicción, pues tiene fallas de más del 100%, para el caso del MAPE.

En conclusión, creemos que un buen modelo para predecir el ancho del anillo de los *Pinos Silvestres* es un modelo ARMA(1,1), puesto que sus métricas de error son menores y posee una óptima calidad de ajuste.