



Procesos estocásticos irregularmente espaciados: Una variación del modelo autorregresivo irregular de primer orden IAR

Wilmar Sepulveda Herrera

Universidad del valle
Facultad de Ingeniería, Escuela de Estadística
Cali, Valle del Cauca
2023

Procesos estocásticos irregularmente espaciados

by

Wilmar Sepulveda Herrera

Trabajo de investigación para optar por el título de:

Magister en

En

Estadística

Aceptamos esta tesis
Conforme a los requerimientos de la norma

Universidad del valle
Facultad de Ingeniería, Escuela de Estadística
Cali, Valle del Cauca
2023

Agradecimientos

Estoy agradecido con el director de esta tesis, el profesor Cesar Ojeda que siempre estuvo dispuesto a ayudarme, y con mi novia, que siempre estuvo apoyandome incluso cuando pensé en abandonar. Muchas gracias.

Resumen

Este trabajo se centra en el estudio de los procesos estocásticos irregularmente espaciados, los cuales son de gran importancia en diversos contextos pero a menudo pasan desapercibidos. Aunque se han propuesto enfoques previos para abordar este fenómeno, como el modelo Autorregresivo Irregular de primer orden (IAR), estos modelos solo permiten autocorrelaciones positivas. En esta investigación, se busca generalizar el modelo IAR extendiéndolo para incluir autocorrelaciones negativas.

Se propone una función específica que resuelve este problema y se deja abierta la posibilidad de proponer funciones más robustas en futuros estudios. Además, se desarrollaron estimadores basados en la máxima verosimilitud y bootstrap para el modelo propuesto, y se evaluaron sus propiedades mediante simulaciones.

Finalmente, se aplicó el modelo a un conjunto de datos reales para analizar su relevancia práctica. Los resultados demuestran que el modelo propuesto puede ser utilizado por investigadores para abordar una amplia gama de problemas relacionados con procesos estocásticos. **Palabras clave:** Procesos estocásticos, Series de tiempo, Tiempos irregulares, Máxima verosimilitud, Bootstrap, Simulación.

Abstract

This work focuses on the study of irregularly spaced stochastic processes, which are of great importance in various contexts but often go unnoticed. Although previous approaches have been proposed to address this phenomenon, such as the first-order Irregular Autoregressive (IAR) model, these models only allow for positive autocorrelations. In this research, we aim to generalize the IAR model by extending it to include negative autocorrelations.

A specific function is proposed to solve this problem, and the possibility of proposing more robust functions in future studies is left open. Furthermore, maximum likelihood and bootstrap estimators were developed for the proposed model, and their properties were evaluated through simulations.

Finally, the model was applied to a real dataset to assess its practical relevance. The

results demonstrate that the proposed model can be used by researchers to address a wide range of problems related to stochastic processes.

Keywords: Stochastic processes, Time series, Irregular times, Maximum likelihood, Bootstrap, Simulation.

Contents

Abstract	v
List of figures	ix
List of tables	x
1 Introducción	1
1.1 Objetivos	2
1.1.1 Objetivo general	2
1.1.2 Objetivos específicos	2
2 Marco teórico	3
2.1 Procesos estocásticos irregularmente espaciados	3
2.1.1 Formas de describir un proceso estocástico	4
2.2 El proceso de medias móviles irregular de primer orden (IMA)	4
2.2.1 El punto de vista distribucional	5
2.2.2 El punto de vista construccionista	5
2.3 El proceso Autorregresivo de primer orden irregular (IAR)	6
2.3.1 Estimación	7
2.4 Modelo autoregresivo complejo irregular (CIAR)	8
3 Modelo autorregresivo de primer orden irregularmente espaciado	10
3.1 Construcción del modelo	10
3.2 Implementando propuesta de ϕ_n	12
3.3 El modelo autorregresivo irregularmente espaciado de primer orden	12
3.4 Propiedades del proceso	13
3.5 Pronóstico en el modelo IAR	13
3.6 Estimación por máxima verosimilitud	14

3.7	Estimación Bootstrap	15
3.8	Estudio de Simulación Monte Carlo	16
3.9	Comparación del Modelo IAR vs el CIAR	21
3.10	Aplicación en datos astronómicos	23
4	Conclusiones	26
	References	27

List of Figures

3-1	simulación del modelo IAR con coeficientes positivos	16
3-2	simulación del modelo IAR con coeficientes negativos	17
3-3	Esquema general del estudio de Monte Carlo.	18
3-4	Resultados estimacion por máxima verosimilitud.	19
3-5	Resultados estimacion Bootstrap.	19
3-6	Comportamiento de la simulación, tomando como base el IAR. ECM	21
3-7	Comportamiento de la simulación, tomando como base el CIAR. ECM . . .	21
3-8	Comportamiento de la simulación, tomando como base el IAR. ϕ	22
3-9	Comportamiento de la simulación, tomando como base el CIAR. ϕ	22
3-10	Comportamiento de la serie de tiempo original.	23
3-11	Ajuste del modelo Loess.	23
3-12	Residuales del modelo Loess (eliminando tendencias).	23
3-13	Ajuste del modelo IAR a la serie transformada	23
3-14	Comportamiento de las predicciones y su intervalo de predicción.	24
3-15	qqplot de los residuos del modelo.	24
3-16	ACF de los residuos.	24
3-17	Ljung-Box test para los residuos del modelo.	24

List of Tables

3-1 Resultados del estudio de Monte Carlo 20

Chapter 1

Introducción

Los procesos estocásticos, y en particular el análisis de series de tiempo, son metodologías estadísticas ampliamente utilizadas en la actualidad. Estas técnicas se aplican comúnmente en el pronóstico de valores futuros y en el estudio de las características de un fenómeno, incluyendo su estructura de autocorrelación. La función de autocorrelación puede ser positiva, lo que es importante en muchos contextos, como la demografía (por ejemplo, Novoseltseva et al. (2019)) o la explicación de fenómenos ecológicos (por ejemplo, Yang et al. (2019)). También pueden darse autocorrelaciones negativas, que son relevantes en el sector financiero para entender el comportamiento de los rendimientos de acciones (por ejemplo, Kuttu & Bokpin (2017)) y en la biología (por ejemplo, Rindorf et al. (2020)). En este contexto, los procesos estocásticos son importantes en diversos campos del conocimiento para caracterizar la dependencia temporal de un evento de interés y tomar mejores decisiones.

La teoría de series temporales es ampliamente utilizada en la actualidad para modelar la autocorrelación de una secuencia de interés. Uno de los métodos más populares es el modelo autorregresivo integrado de medias móviles (ARIMA). Sin embargo, este y otros métodos similares asumen que las observaciones son tomadas regularmente en el tiempo, lo cual no siempre es el caso en la práctica. Esto puede deberse a la presencia de valores faltantes o a la naturaleza de las observaciones, lo que puede hacer que la serie de tiempo sea irregularmente espaciada (Elorrieta et al. 2019). De hecho, es común encontrar series de tiempo con este tipo de irregularidades en diversas aplicaciones, como en sistemas caóticos, medicina, imágenes satelitales y calidad del aire, entre otros (Shamsan et al. 2020, Liu et al. 2019, Ghaderpour & Vujadinovic 2020, Dilmaghani et al. 2007).

En la literatura actual, existen diversos métodos para modelar series de tiempo que no son tomadas regularmente en el tiempo. Algunos autores, como Adorf (1995), abordan

este problema transformando las series irregulares en regulares a través de la interpolación, pero este enfoque puede generar sesgos. Otros, como Robinson (1977), tratan la serie de tiempo como una realización discreta de un proceso estocástico de tiempo continuo, lo cual también presenta problemas. Recientemente, Ojeda et al. (2023) propusieron el modelo de medias móviles irregulares (*IMA*). Bajo normalidad, este modelo es estrictamente estacionario, mientras que, considerando otros supuestos distribucionales, se comporta como un proceso débilmente estacionario.

En esta línea, Eyheramendy et al. (2018) proponen el Modelo Autorregresivo Irregular (IAR) de primer orden. Sin embargo, este modelo tiene una limitación importante: no acepta valores negativos en su función de autocorrelación (ACF). Para solucionar este problema, los autores presentaron el Modelo Autorregresivo Complejo Irregular (CIAR) en un estudio posterior (Elorrieta et al. 2019). En ambos casos, se han evaluado las propiedades estadísticas de los modelos y se han propuesto métodos de estimación que han obtenido buenos resultados. No obstante, tanto el modelo IAR como el IMA tienen una desventaja: su función de autocorrelación solo puede ser positiva, lo que limita su capacidad para detectar autocorrelaciones negativas, que suelen ser importantes en la práctica. Este problema no está presente en los modelos ARIMA clásicos. Por lo tanto, el objetivo de este trabajo es modificar el modelo IAR para que pueda modelar tanto autocorrelaciones positivas como negativas sin necesidad de recurrir al plano complejo.

1.1 Objetivos

1.1.1 Objetivo general

Desarrollar un modelo autorregresivo irregular de primer orden capaz de modelar estructuras de autocorrelación subyacentes con valores tanto positivos como negativos.

1.1.2 Objetivos específicos

- Encontrar estimadores por máxima verosimilitud y bootstrap para el modelo propuesto.
- Evaluar las propiedades del modelo propuesto y sus estimaciones vía simulación.
- Aplicar el modelo propuesto a un conjunto de datos reales y evaluar su capacidad para capturar la dinámica de la serie de tiempo y realizar predicciones precisas.

Chapter 2

Marco teórico

Para el tratamiento de procesos estocásticos y series de tiempo, ha sido mucha la teoría estadística implementada y se ha convertido en una herramienta importante para tratar observaciones dependientes y entender la naturaleza de dicha dependencia. La mayoría de estos métodos asumen que las series de tiempo son regularmente espaciadas, sin embargo, este escenario no siempre se da y ahí es donde cobran importancia los procesos estocásticos irregularmente espaciados.

Este capítulo se estructura de la siguiente manera: en la sección 2.1, se presenta la definición matemática de un proceso estocástico irregularmente espaciado. Luego, en la sección 2.2, se aborda el proceso de medias móviles irregularmente espaciado. A continuación, en la sección 2.3, se analiza un proceso estocástico autorregresivo irregularmente espaciado. Finalmente, en la sección 2.4, se define el proceso autorregresivo complejo irregularmente espaciado, conocido como *CIAR*.

2.1 Procesos estocásticos irregularmente espaciados

Sea (Ω, \mathcal{B}, P) un espacio de probabilidad. Si se define un proceso estocástico como una medida que mapea $x : \Omega \mapsto \mathbb{R}^{\mathbb{T}}$, donde

$$x(\omega) = \{X_{\tau}(\omega), \tau \in \mathbb{T}\}$$

\mathbb{T} es llamado conjunto de índices y la variable aleatoria $X_{\tau}(\omega)$ es llamada coordenada del proceso o trayectoria. Consideremos ahora $\mathbb{T}' = \{\tau_1, \tau_2, \tau_3, \dots\}$, con $\Delta_{n+1} = \tau_{n+1} - \tau_n$, para $n \geq 1$

$$x' = \{X_{\tau}, \tau \in \mathbb{T}'\} \tag{2-1}$$

x' es un proceso estocástico irregularmente espaciado. Una serie de tiempo irregularmente espaciada es una realización finita de un proceso estocástico irregularmente espaciado. Note que si $\Delta_{n+1} = \tau_{n+1} - \tau_n = 1$ para $n \geq 1$, el proceso x' es un proceso regularmente espaciado.

2.1.1 Formas de describir un proceso estocástico

Una forma de describir un proceso estocástico x , es especificar la función de distribución conjunta de $\{X_{\tau_1}, X_{\tau_2}, \dots, X_{\tau_n}\}$ para todo n , este es llamado *punto de vista distribucional*. Por otra parte, se puede describir el proceso proporcionando una formula para el valor X_τ para cada punto τ en términos de una familia de variables aleatorias con comportamiento probabilístico conocido, esto hace que podamos ver el proceso como una función de otros procesos (o cómo familias de procesos iid); esta forma es llamada *punto de vista constructorista*.

2.2 El proceso de medias móviles irregular de primer orden (IMA)

Teniendo en cuenta el conjunto $\mathbb{T}' = \{\tau_1, \tau_2, \tau_3, \dots\}$ propuesto anteriormente, tal que sus diferencias Δ_{n+1} para $n \geq 1$ están acotados uniformemente lejos de cero. Ahora, sea $m : \mathbb{T}' \mapsto \mathbb{R}$ una función tal que $m(\tau_n) = 0$, para cualquier $\tau_n \in \mathbb{T}'$. A continuación, sea $\Gamma : \mathbb{T}' \times \mathbb{T}' \mapsto \mathbb{R}$ una función tal que para cualquier pareja $\tau_n, \tau_s \in \mathbb{T}'$,

$$\Gamma(\tau_n, \tau_s) = \begin{cases} \gamma_0 & |n - s| = 0 \\ \gamma_1, \Delta_{\max(n,s)} & |n - s| = 1 \\ 0 & |n - s| \geq 1 \end{cases}$$

Note que Γ puede ser representada cómo una matriz diagonal así

$$\Gamma = \begin{bmatrix} \gamma_0 & \gamma_1, \Delta_2 & 0 & 0 & \dots \\ \gamma_1, \Delta_2 & \gamma_0 & \gamma_1, \Delta_3 & 0 & \\ 0 & \gamma_1, \Delta_3 & \gamma_0 & \gamma_1, \Delta_4 & \\ \cdot & & & \cdot & \\ \cdot & & & & \cdot \\ \cdot & & & & & \cdot \end{bmatrix} \quad (2-2)$$

2.2. EL PROCESO DE MEDIAS MÓVILES IRREGULAR DE PRIMER ORDEN (IMA) 5

Ser Γ_n la truncación $n \times n$ de Γ y asumiendo $\gamma_1, \Delta_j \neq 0$, para $j = 2, \dots, n$, Γ_n es definida positiva si $\gamma_0 \geq 0$ y $(\frac{\gamma_1, \Delta_{n+1}}{\gamma_0})^2 \leq 1/4$, para $j = 2, \dots, n$

Esto implica que existe un proceso Gaussiano estacionario $\{X_{\tau_n}, \tau_n, \tau \in \mathbb{T}\}$, único hasta la equivalencia, con media 0 y covarianza Γ . Este proceso es llamado, proceso de medias móviles de primer orden irregularmente espaciado de forma general. A continuación se darán las expresiones particulares de este proceso desde los dos puntos de vista antes definidos.

2.2.1 El punto de vista distribucional

En (2-2), γ_0 y γ_1, Δ_{n+1} , para $n \geq 1$, representa la varianza y las covarianzas de primer orden respectivamente. Definimos la varianza como $\gamma_0 = \sigma^2(1 + \theta^2)$ y las covarianzas de primer orden como $\gamma_1, \Delta_{n+1} = \sigma^2 \theta^{\Delta_{n+1}}$, $\sigma^2 > 0$ y $0 < \theta < 1$. Por tanto, obtenemos el proceso estocástico irregularmente espaciado de primer orden con matriz de covarianzas

$$\Gamma = \begin{bmatrix} 1 + \theta^2 & \theta^{\Delta_2} & 0 & 0 & \dots \\ \theta^{\Delta_2} & 1 + \theta^2 & \theta^{\Delta_3} & 0 & \\ 0 & \theta^{\Delta_3} & 1 + \theta^2 & \theta^{\Delta_4} & \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \end{bmatrix} \quad (2-3)$$

el cual contiene el modelo de medias móviles convencional como caso especial. Este es llamado proceso Gaussiano irregular de medias móviles de primer orden.

2.2.2 El punto de vista constructorista

Ahora, como es usual, especificaremos el proceso IMA como función de otros procesos estocásticos. Sea $\{\epsilon_{\tau_n}\}_{n \geq 1}$ variables aleatorias independientes que siguen una distribución normal $N(0, \sigma^2 c_n(\theta))$ con $\sigma^2 > 0, 0 < \theta < 1, c_1(\theta) = 1 + \theta^2$ y

$$c_n(\theta) = 1 + \theta^2 - \frac{\theta^{2\Delta_n}}{c_{n-1}(\theta)} \text{ para } n \geq 2$$

donde $\Delta_n = \tau_n - \tau_{n-1}$. El proceso $\{X_{\tau_n}, \tau_n \in \mathbb{T}\}$, es decirse tiene un proceso IMA si $X_{\tau_1} = \epsilon_{\tau_1}$ y para $n \geq 2$

$$X_{\tau_n} = \epsilon_{\tau_n} + \frac{\theta^{\Delta_n}}{c_{n-1}(\theta)} \epsilon_{\tau_{n-1}} \quad (2-4)$$

Decimos que $\{X_{\tau_n}, \tau_n \in \mathbb{T}'\}$ es un proceso IMA con media μ si $\{X_{\tau_n} - \mu, \tau_n \in \mathbb{T}'\}$ es un proceso IMA

2.3 El proceso Autorregresivo de primer orden irregular (IAR)

Se denota X_{τ_n} a una observación medida en el tiempo τ_n y se define el proceso IAR de primer orden así:

$$X_{\tau_n} = \phi^{\tau_n - \tau_{n-1}} X_{\tau_{n-1}} + \sigma \sqrt{1 - \phi^{2(\tau_n - \tau_{n-1})}} \varepsilon_{\tau_n} \quad (2-5)$$

Donde ε_{τ_n} son variables aleatorias independientes con media cero y varianza unitaria, observe que $E(X_{\tau_n}) = 0$ y $Var(X_{\tau_n}) = \sigma^2$ para todo X_{τ_n} .

Así, la covarianza entre X_{τ_k} y X_{τ_j} es $E(X_{\tau_k}, X_{\tau_j}) = \sigma^2 \phi^{\tau_k - \tau_j}$ para $k \geq j$. La matriz de varianzas y autocovarianzas se muestra en 2-6.

$$\begin{bmatrix} 1 & \phi^{\tau_2 - \tau_1} & \phi^{\tau_3 - \tau_1} & \cdot & \cdot & \cdot & \cdot & \phi^{\tau_n - \tau_1} \\ & 1 & \phi^{\tau_3 - \tau_2} & \phi^{\tau_4 - \tau_2} & \cdot & \cdot & \cdot & \phi^{\tau_n - \tau_2} \\ & & 1 & \phi^{\tau_4 - \tau_3} & \phi^{\tau_5 - \tau_3} & \cdot & \cdot & \phi^{\tau_n - \tau_3} \\ & & & 1 & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & \cdot & \cdot & \cdot \\ & & & & & & \phi^{\tau_n - \tau_{n-1}} & \cdot \\ & & & & & & & 1 \end{bmatrix} \quad (2-6)$$

Así, para cualquier par de tiempos observacionales $s < t$, podemos definir la función autocovarianza cómo:

$$\gamma(t-s) = E(X_t, X_s) = \sigma^2 \phi^{t-s}$$

También, la función de autocorrelación (ACF) está dada por $\rho(t-s) = \frac{\gamma(t-s)}{\gamma(0)} = \phi^{t-s}$

La secuencia X_{τ_n} corresponde a un proceso debilmente estacionario de segundo orden, también se puede demostrar que bajo algunas condiciones el proceso es estacionario y

ergódico. Ver (Eyheramendy et al. 2018).

Teorema 2.3.1. *Considere el proceso definido por la ecuación 2-5 y supongamos que el ruido de entrada es una secuencia iid de variables aleatorias con media cero y varianza unitaria. Además, suponga que $\tau_j - \tau_{j-n} \geq C \log n$ como $n \rightarrow \infty, 0 \leq \phi \leq 1$, donde C es una constante positiva que satisface $C \log \phi^2 \geq -1$. Luego, existe una solución al proceso definido por la ecuación 2-5 y la secuencia $\{X_{\tau_j}\}$ es estacionaria y ergódica.*

Note también que si $\tau_n - \tau_{n-1} = 1$ para todo n , la ecuación 2-5 se convierte en:

$$X_{\tau_n} = \phi X_{\tau_{n-1}} + \sigma \sqrt{1 - \phi^2} \varepsilon_{\tau_n} \quad (2-7)$$

La ecuación 2-7 corresponde al modelo autorregresivo de primer orden para datos regularmente espaciados AR(1). Por lo tanto, el modelo IAR es una extensión del modelo autorregresivo regular

2.3.1 Estimación

La verosimilitud de una muestra aleatoria $X_{\tau_1} \dots X_{\tau_n}$ puede ser expresada como

$$f(X_{\tau_1} \dots X_{\tau_n}; \theta) = f(X_{\tau_1}; \theta) * f(X_{\tau_2} | X_{\tau_1}; \theta) * \dots * f(X_{\tau_n} | X_{\tau_{n-1}}; \theta)$$

donde $\theta = (\sigma^2, \phi)$ es el vector de parámetros del modelo, para poder describir el proceso de estimación, se asume que las distribuciones condicionales son Gaussianas, es decir, se asume que

$f(X_{\tau_1}) \sim N(0, \sigma^2)$ y $f(X_{\tau_n} | X_{\tau_{n-1}}; \theta) \sim N(\phi^{\tau_j - \tau_{j-1}} X_{\tau_{j-1}}, \sigma^2(1 - \phi^{2(\tau_j - \tau_{j-1})}))$ para $j = 1, 2, 3 \dots n$

basado en la ecuación 2-5, la log-verosimilitud del proceso puede ser escrita como

$$l(\theta) = \frac{n}{2} \log(2\pi) + \frac{1}{2} \sum_{j=1}^n \log v_{\tau_j} + \frac{1}{2} \sum_{j=1}^n \frac{e_{\tau_j}}{v_{\tau_j}} \quad (2-8)$$

se define $e_{\tau_1} = X_{\tau_1}$, $e_{\tau_j} = X_{\tau_j} - \phi^{\tau_j - \tau_{j-1}} X_{\tau_{j-1}}$ para $j > 1$ y sus varianzas $v_{\tau_1} = \text{Var}(e_{\tau_1})$.

Observe que el predictor pasado de tiempo finito del proceso en el tiempo τ_j es dado por $\hat{X}_{\tau_1} = 0$, y $\hat{X}_{\tau_j} = \phi^{\tau_j - \tau_{j-1}} X_{\tau_{j-1}}$, para $j = 2, \dots, n$

Por tanto, $e_{\tau_j} = X_{\tau_j} - \hat{X}_{\tau_j}$ es el error de predicción con varianza $v_{\tau_1} = \text{Var}(e_{\tau_1} = \sigma^2)$ y $v_{\tau_j} = \text{var}(e_{\tau_j}) = \sigma^2[1 - \phi^{2(\tau_j - \tau_{j-1})}]$, para $j = 2, \dots, n$

Por maximización directa de la función de log-verosimilitud (2-8), se puede obtener el estimador de máxima verosimilitud de σ^2

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n \frac{(X_{\tau_j} - \hat{X}_{\tau_j})}{\tau_{\tau_j}}, \text{ donde } \tau_{\tau_j} = v_{\tau_j}/\sigma^2. \quad (2-9)$$

Pero no es posible encontrar $\hat{\phi}$, el estimador por máxima verosimilitud de ϕ , por maximización directa de la verosimilitud, es necesario recurrir a metodos iterativos.

2.4 Modelo autoregresivo complejo irregular (CIAR)

Para derivar una extensión compleja del modelo 2-5, se sigue el mismo proceso que para series de tiempo regulares. Suponiendo que x_{τ_j} es una secuencia de valores complejos, tal que $x_{\tau_j} = y_{\tau_j} + iz_{\tau_j} \forall j = 1, \dots, n$, y de manera similar, $\phi = \phi^R + i\phi^I$ es el coeficiente complejo del modelo y $\varepsilon_{\tau_j} = \varepsilon_{\tau_j}^R + i\varepsilon_{\tau_j}^I$ es un ruido blanco complejo, definimos el proceso CIAR como

$$y_{\tau_j} + iz_{\tau_j} = (\phi^R + i\phi^I)^{\tau_j - \tau_{j-1}} (y_{\tau_{j-1}} + iz_{\tau_{j-1}}) + \sigma_{\tau_j} (\varepsilon_{\tau_j}^R + i\varepsilon_{\tau_j}^I) \quad (2-10)$$

donde $\sigma_{\tau_j} = \sigma \sqrt{1 - |\phi^{\tau_j - \tau_{j-1}}|^2}$ y $|\cdot|$ es el módulo de un número complejo. Suponemos que solo se observa la parte real y_{τ_j} y que la parte imaginaria z_{τ_j} es un proceso latente. Además, se asume que $\varepsilon_{\tau_j}^R$ y $\varepsilon_{\tau_j}^I$, la parte real e imaginaria de ε_{τ_j} , respectivamente, son independientes con media cero y varianza positiva $\mathbb{V}(\varepsilon_t^R) = 1$ y $\mathbb{V}(\varepsilon_t^I) = c$, respectivamente, donde c es un parámetro fijo que toma valores en \mathbb{R}^+ . Generalmente, se supone que $c = 1$, y los valores iniciales son $y_{\tau_1} = \sigma \varepsilon_{\tau_1}^R$ y $z_{\tau_1} = \sigma \varepsilon_{\tau_1}^I$. En el siguiente lema, se presentan algunas de las propiedades de este proceso.

Lemma 2.4.1. *Considere el proceso CIAR x_{τ_j} descrito por la ecuación 2-10. Defina $\gamma_0 = \mathbb{E}(\bar{x}_{tj}x_{\tau_j})$, $\gamma_k = \mathbb{E}(\bar{x}_{tj+k}x_{\tau_j})$ y ρ_k como la varianza, autocovarianza y autocorrelación, respectivamente, del proceso x_{τ_j} . A continuación, el valor esperado, la varianza, la autocovarianza y la autocorrelación del proceso*

- $\mathbb{E}(x_{\tau_j}) = 0$
- $\mathbb{V}(x_{\tau_j}) = \gamma_0 = \mathbb{E}(\bar{x}_{tj}x_{\tau_j}) = \sigma^2(1 + c)$,
- $\gamma_k = \mathbb{E}(\bar{x}_{tj+k}x_{\tau_j}) = \overline{\phi^{\Delta_k}} \sigma^2(1 + c)$,

- $\rho_k = \overline{\phi^{\Delta_k}}$

donde $\Delta_k = \tau_{j+k} - \tau_j$ denota las diferencias de tiempo entre los tiempos de observación τ_{j+k} y τ_j . Además, \bar{x}_{tj} es el conjugado complejo de x_{τ_j} .

Si $|\phi| = |\phi^R + i\phi^I| < 1$, los resultados en el Lema 2.4.1 muestran que la secuencia compleja x_{τ_j} es un proceso débilmente estacionario. Observamos que la ACF ρ_k del proceso CIAR decae a una tasa $\phi^{\tau_{j+k} - \tau_j}$ (la llamada decaimiento exponencial). Esta estructura de autocorrelación es diferente a la de un proceso CARFIMA de memoria antipersistente o de memoria intermedia, ya que la ACF de estos últimos decae más lentamente que un decaimiento exponencial. Por lo tanto, aunque ambos modelos pueden ajustar series de tiempo irregulares con valores negativos de la ACF, el uso apropiado de estos modelos dependerá de la estructura de correlación de los datos. Para tomar una decisión sobre el modelo más adecuado para los datos, se pueden utilizar técnicas basadas en la verosimilitud de los datos, como AIC, BIC, y así sucesivamente.

Chapter 3

Modelo autorregresivo de primer orden irregularmente espaciado

En este capítulo, proponemos un modelo basado en la propuesta de Eyheramendy et al. (2018), la cual permite modelar unicamente estructuras de autocorrelacion positivas, el llamado $IAR(1)$. Nosotros generalizamos esta propuesta para que se permitan estructuras tanto positivas como negativas, teniendo en cuenta que nuestra novedosa propuesta, debe tener como caso particular el modelo $IAR(1)$ y el clasico modelo $AR(1)$

3.1 Construcción del modelo

En esta sección, construiremos un proceso estocástico estacionario que tiene una estructura autorregresiva y que además considera espacios de tiempo irregulares. Suponemos que el comportamiento irregularmente espaciado es independiente de las propiedades estocásticas del proceso.

Sea $\{\varepsilon_\tau, \tau \in \mathbb{T}\}$ una secuencia iid de variables aleatorias cada una con distribución $N(0, 1)$ y defino:

$$\begin{aligned} X_{\tau_1} &= W_1^{1/2} \varepsilon_{\tau_1} \\ X_{\tau_{n+1}} &= \phi_{n+1} X_{\tau_n} + W_{n+1}^{1/2} \varepsilon_{\tau_{n+1}} \end{aligned} \tag{3-1}$$

Donde, $0 \leq \phi \leq 1$, $E(X_{\tau_n}) = 0$ y $\{W_n\}_{n \geq 1}$ es una secuencia que caracteriza los momentos del proceso, ϕ_n es una funcion de Δ_n y del parametro ϕ del modelo autorregresivo que definiremos mas adelante. así, para $n \geq 1$, tenemos $E(X_{\tau_n}) = 0$,

$$V(X_{\tau_1}) = W_1$$

$$V(X_{\tau_{n+1}}) = \phi_{n+1}^2 V(X_{\tau_n}) + W_{n+1}$$

Ahora, bajo este modelo, calculemos $Cov(X_{\tau_n}, X_{\tau_{n+k}})$, esto puede hacerse mirando casos particulares

- si $k=1$

$$Cov(X_{\tau_n}, X_{\tau_{n+1}}) = Cov(X_{\tau_n}, \phi_{n+1} X_{\tau_n} + W_{n+1}^{1/2}) = \phi_{n+1} Cov(X_{\tau_n}, X_{\tau_n}) = \phi_{n+1} Var(X_{\tau_n})$$

- si $k=2$

$$Cov(X_{\tau_n}, X_{\tau_{n+2}}) = Cov(X_{\tau_n}, \phi_{n+2} X_{\tau_{n+1}} + W_{n+2}^{1/2}) = \phi_{n+2} Cov(X_{\tau_n}, X_{\tau_{n+1}}) = \phi_{n+2} \phi_{n+1} Var(X_{\tau_n})$$

- si $k=3$

$$Cov(X_{\tau_n}, X_{\tau_{n+3}}) = Cov(X_{\tau_n}, \phi_{n+3} X_{\tau_{n+2}} + W_{n+3}^{1/2}) = \phi_{n+3} Cov(X_{\tau_n}, X_{\tau_{n+2}}) = \phi_{n+3} \phi_{n+2} \phi_{n+1} Var(X_{\tau_n})$$

- Reemplazando recursivamente hasta $k = n$ tenemos:

$$Cov(X_{\tau_n}, X_{\tau_{n+k}}) = \prod_{i=1}^k \phi_{n+i} Var(X_{\tau_n}) \quad (3-2)$$

Para que este proceso sea estacionario, debemos demostrar que la varianza es constante en cualquier momento τ , por tanto, debemos mostrar que, para $n \geq 1$; $Var(X_{\tau_{n+1}}) = Var(X_{\tau_1}) = \gamma_0$, es decir:

$$\phi_{n+1}^2 \gamma_0 + W_{n+1} = W_1 = \gamma_0 \quad (3-3)$$

despejando W_{n+1} de 3-3 tenemos:

$$W_{n+1} = \gamma_0(1 - \phi_{n+1}^2) \quad (3-4)$$

Finalmente, para que nuestro modelo coincida con el modelo $AR(1)$ en el momento en que $\Delta_{n+1} = 1$ para todo n , hacemos $\gamma_0 = \frac{\sigma^2}{1-\phi^2}$

3.2 Implementando propuesta de ϕ_n

El foco de la investigación está en generar una propuesta estable y teóricamente coherente para la secuencia ϕ_{n+1} , después de mucho estudiarlo, llegamos a esta expresión.

$$\phi_{n+1} = \text{sign}(\phi)|\phi|^{\Delta_{n+1}} \quad (3-5)$$

Si reemplazamos 3-5 en 3-4 obtenemos:

$$W_{n+1} = \frac{\sigma^2}{1-\phi^2} \left(1 - \text{sign}(\phi)^2 |\phi|^{2\Delta_{n+1}}\right) \quad (3-6)$$

El modelo final, reemplazando 3-5 y 3-6 en 3-1 tiene la expresión:

$$X_{\tau_{n+1}} = \text{sign}(\phi)|\phi|^{\Delta_{n+1}} X_{\tau_n} + \left[\frac{\sigma^2}{1-\phi^2} \left(1 - \text{sign}(\phi)^2 |\phi|^{2\Delta_{n+1}}\right) \right]^{1/2} \varepsilon_{\tau_{n+1}} \quad (3-7)$$

por definición la varianza de la expresión 3-7 es:

$$V(X_{\tau_{n+1}}) = \frac{\sigma^2}{1-\phi^2} \text{ y } Cov(X_{\tau_n}, X_{\tau_{n+k}}) = \left(\frac{\sigma^2}{1-\phi^2} \right) \text{sign}(\phi)^k |\phi|^{\sum_{i=1}^k \Delta_{n+i}}$$

3.3 El modelo autorregresivo irregularmente espaciado de primer orden

Basado en los calculos y definiciones anteriores, esta nueva propuesta de modelo autorregresivo irregularmente espaciado IAR es realizada desde un punto de vista constructorista. Más adelante en este documento presentamos las propiedades del proceso.

Definición 3.3.1 (Proceso IAR - Autocorrelaciones positivas y negativas). Sea $\{\varepsilon_{\tau_n}\}_{n \geq 1}$ variables aleatorias, cada una con distribución $N(0,1)$, siendo $\sigma^2 > 0$ y $-1 < \phi < 1$, se dice es un proceso IAR Si $X_1 = \left(\frac{\sigma^2}{1-\phi^2} \right)^{1/2} \varepsilon_{\tau_1}$ y para $n \geq 2$.

$$X_{\tau_n} = \text{sign}(\phi)|\phi|^{\Delta_n} X_{\tau_{n-1}} + \left[\frac{\sigma^2}{1-\phi^2} \left(1 - \text{sign}(\phi)^2 |\phi|^{2\Delta_n}\right) \right]^{1/2} \varepsilon_{\tau_n} \quad (3-8)$$

Donde $\Delta_n = \tau_n - \tau_{n-1}$

3.4 Propiedades del proceso

Si $X_n = [X_{\tau_1}, \dots, X_{\tau_n}]'$ es un vector aleatorio de un proceso IAR, luego X_n es un vector aleatorio Gaussiano con $m_n = 0$ y matriz de varianzas y covarianza dada por

$$\Gamma_n = \begin{bmatrix} \gamma_0 & & & & & \\ \phi_2 \gamma_0 & \gamma_0 & & & & \\ \phi_2 \phi_3 \gamma_0 & \phi_3 \gamma_0 & \ddots & & & \\ \vdots & \vdots & & \gamma_0 & & \\ \prod_{i=2}^{n-1} \phi_i \gamma_0 & \prod_{i=3}^{n-1} \phi_i \gamma_0 & \cdots & \phi_{n-1} \gamma_0 & \gamma_0 & \\ \prod_{i=2}^n \phi_i \gamma_0 & \prod_{i=3}^n \phi_i \gamma_0 & \cdots & \phi_{n-1} \phi_n \gamma_0 & \phi_n \gamma_0 & \gamma_0 \end{bmatrix},$$

donde $\phi_n = \text{sign}(\phi)|\phi|^{\Delta_n}$ y $\gamma_0 = \frac{\sigma^2}{1-\phi^2}$, para $n \geq 2$. Así, el proceso IAR es, por definicion un proceso Gaussiano debilmente estacionario y por tanto estrictamente estacionario. Además, note que cuando $\Delta_n = 1$, para $n \geq 2$, obtenemos el proceso AR convencional.

3.5 Pronóstico en el modelo IAR

El objetivo es predecir los valores futuros de la serie $X_{\tau_{n+m}}$, donde $m = 1, 2, \dots$, basados en la información recolectada $X_{\tau_{1:n}} = X_{\tau_1}, \dots, X_{\tau_n}$. El predictor de error cuadrático medio mínimo para $X_{\tau_{n+m}}$ es:

$$\hat{X}_{\tau_{n+m}} = E(X_{\tau_{n+m}} | X_{\tau_{1:n}}) \quad (3-9)$$

Esto se debe a que la esperanza condicional minimiza el error cuadrático medio, como se demuestra en el apéndice.

$$E[X_{\tau_{n+m}} - g(X_{\tau_{1:n}})]. \quad (3-10)$$

Si consideramos la predicción a un paso, $\hat{X}_{\tau_{n+1}}$, basada en el modelo 3-7, tenemos que:

$$\hat{X}_{n+1} = E(X_{\tau_{n+1}} | X_{\tau_{1:n}}) = \text{sign}(\phi)|\phi|^{\Delta_{n+1}} X_{\tau_n}. \quad (3-11)$$

Además, $X_{\tau_1} = W_1^{1/2} \varepsilon_{\tau_1}$, por lo tanto $\hat{X}_{\tau_1} = 0$. Con estos resultados, el error de predicción es:

$$(X_{\tau_k} - \hat{X}_{\tau_k}) = W_k^{1/2} \varepsilon_{\tau_k}, \quad k = 1, \dots, n \quad (3-12)$$

Podemos ver que, debido a la forma en que se define el modelo, el error de predicción sigue una distribución normal con media 0 y varianza W_k .

3.6 Estimación por máxima verosimilitud

Sea X_{τ} observado en los puntos τ_1, \dots, τ_n , la función de verosimilitud para estimar los parámetros σ y ϕ es la siguiente:

$$L(\phi, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi W_i}} e^{-\frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{2W_i}} \quad (3-13)$$

Ahora, aplicamos logaritmo para calcular la log-verosimilitud $l(\phi, \sigma)$

$$\begin{aligned} l(\phi, \sigma) &= \sum_{i=1}^n \log\left(\frac{1}{\sqrt{2\pi W_i}}\right) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i} \\ &= -\frac{1}{2} \sum_{i=1}^n \log(2\pi W_i) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i} \\ &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log(W_i) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i} \end{aligned} \quad (3-14)$$

Donde W_i es definido por la ecuación 3-4. Se puede minimizar así:

$$W_i = \frac{\sigma^2(1 - \phi^{2\Delta_i})}{1 - \phi^2}$$

donde $W_1 = \frac{\sigma^2}{1 - \phi^2}$

Reemplazando W_i en 3-14 obtenemos:

$$-\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log\left(\frac{\sigma^2(1 - \phi^{2\Delta_i})}{1 - \phi^2}\right) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{\left(\frac{\sigma^2(1 - \phi^{2\Delta_i})}{1 - \phi^2}\right)} \quad (3-15)$$

Por maximización directa de la función de log verosimilitud (3-15), podemos obtener el estimador de σ^2

$$\frac{\partial(l)}{\partial\sigma^2} = \frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma^2} - \frac{1}{2(\sigma^2)^2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{\frac{1 - \phi^{2\Delta_i}}{1 - \phi^2}} = 0 \quad (3-16)$$

Despejando σ^2

$$\sigma_n^2(\phi) = \frac{1}{n} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{\frac{1-\phi^{2\Delta_i}}{1-\phi^2}} \quad (3-17)$$

Ahora si reemplazamos la expresion encontrada para $\sigma_n^2(\phi)$ en 3-15, la nueva funcion objetivo para estimar ϕ será:

$$q_n(\phi) = -\frac{1}{2} \log \left(\sigma_n^2(\phi) \left[\frac{1-\phi^{2\Delta_i}}{1-\phi^2} \right] \right) \quad (3-18)$$

La idea es primero estimar ϕ por métodos numéricos y luego estimar σ^2 con la expresion 3-17

3.7 Estimación Bootstrap

Ahora, aplicamos el método bootstrap para estimar ϕ_0 en el proceso IAR con $\sigma_0^2 = 1$. Sea X_τ observado en los puntos τ_1, \dots, τ_N y considere $\hat{\phi}_N$ como la estimación MLE respectiva. Las innovaciones estandarizadas estimadas son

$$e_{\tau_n}^s = \frac{X_{\tau_n} - \hat{X}_{\phi_n}(\hat{\phi}_N)}{\sqrt{W_n(\hat{\phi}_N)}},$$

para $n = 2, \dots, N$. El llamado remuestreo basado en el modelo podría proceder por muestreo equiprobable con reemplazo de los residuos centrados $e_{t_2}^s - \bar{e}, \dots, e_{t_N}^s - \bar{e}$, donde $\bar{e} = \sum_{n=2}^N e_{t_n}^s / (N-1)$, para obtener innovaciones simuladas $\zeta_{t_1}^*, \dots, \zeta_{t_N}^*$, y luego establecer:

$$X_{\tau_1}^* = \sqrt{W(\hat{\phi}_N)} \zeta_{\tau_1}^*, \quad (3-19)$$

$$X_{\tau_n}^* = \hat{\phi}_n X_{\tau_{n-1}}^* + \sqrt{W(\hat{\phi}_N)} \zeta_{\tau_n}^* \quad (3-20)$$

A continuación, estimamos los parámetros a través de MLE suponiendo que los datos son $X_{t_n}^*$. Por lo tanto, podemos repetir este proceso un gran número, B, de veces generando una colección de estimaciones de parámetros bootstrap. A continuación, podemos aproximar la distribución de muestra finita del estimador, $\hat{\phi}_N$, a partir de los valores de los parámetros bootstrap.

3.8 Estudio de Simulación Monte Carlo

Para probar el modelo 3-8, se genera un escenario de simulación donde se analizan sus propiedades y las de los estimadores por máxima verosimilitud y bootstrap. La simulación se lleva a cabo considerando valores específicos para los parámetros: $\sigma_0 = 1$, $\phi_0 \in \{\pm 0.1, \pm 0.5, \pm 0.9\}$, y $n \in \{20, 50, 100\}$. Estos valores permiten observar el comportamiento irregular del modelo.

En el escenario de simulación, los tiempos τ_1, \dots, τ_n pueden ser regulares o irregulares. En este caso, consideramos $\Delta_n = \tau_n - \tau_{n-1} \sim 1 + \text{poisson}(\lambda = 2)$, asegurando así que no haya valores iguales a cero.

La Figura 3-1 muestra la trayectoria simulada para cada combinación de parámetros establecidos cuando $\phi > 0$. Las marcas rojas en la parte superior del gráfico resaltan de manera más clara el comportamiento irregular. Por otro lado, la Figura 3-2 muestra el mismo comportamiento para el caso de $\phi < 0$.

Con fines de simulación de tipo Montecarlo, se generan $M = 1000$ trayectorias $\{S_m\}_{m=1}^M$ y se estima ϕ_0 para cada configuración.

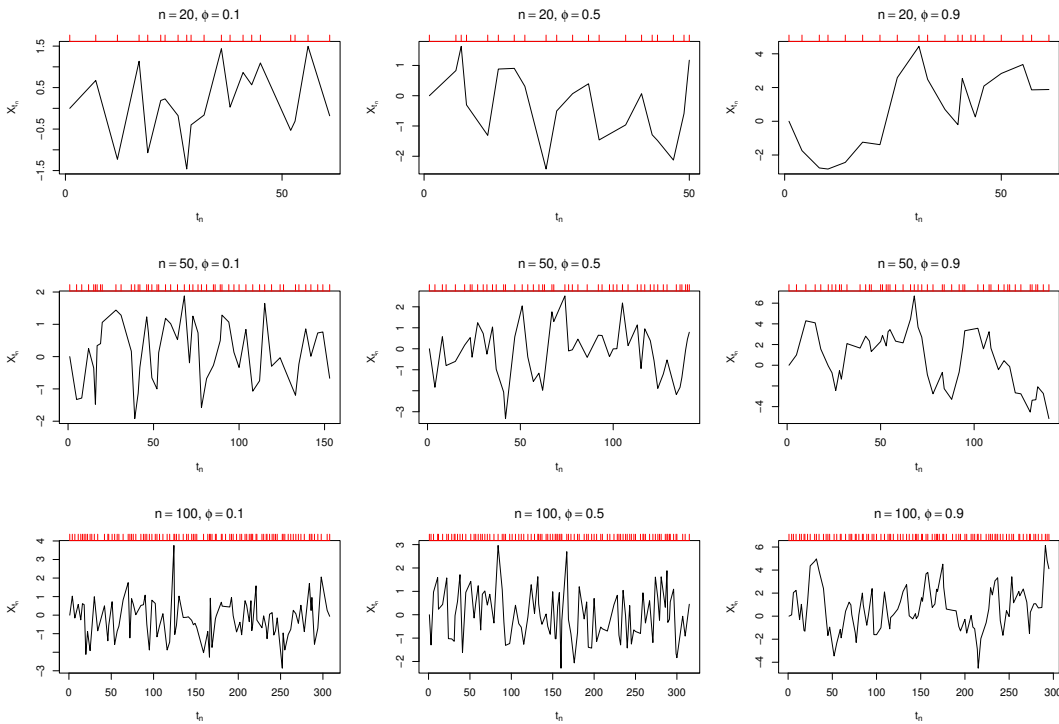


Figure 3-1: simulación del modelo IAR con coeficientes positivos

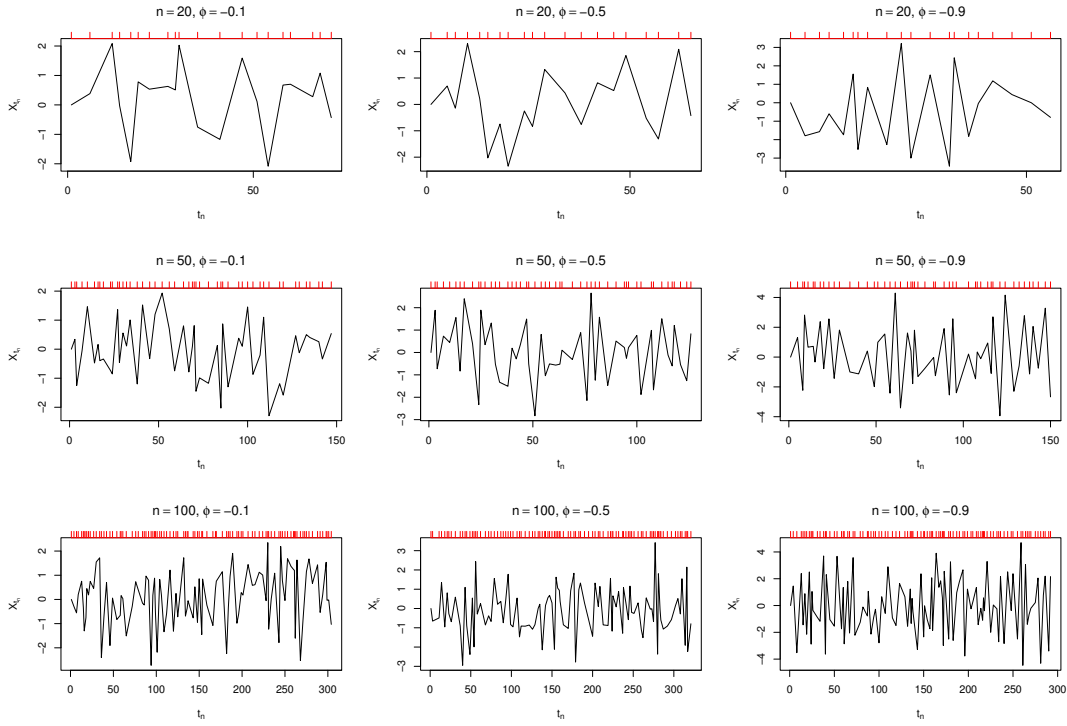


Figure 3-2: simulación del modelo IAR con coeficientes negativos

Sea $\hat{\phi}_m^{\text{MLE}}$ la estimación ML y $\widehat{\text{se}}(\hat{\phi}_m^{\text{MLE}})$ el error estandar estimado para la m -ésima trayectoria. El error estandar se estima por la curvatura de la superficie de verosimilitud en $\hat{\phi}_m^{\text{MLE}}$. Resumimos las estimaciones de máxima verosimilitud M mediante

$$\hat{\phi}^{\text{MLE}} = \frac{1}{M} \sum_{m=1}^M \hat{\phi}_m^{\text{MLE}} \quad \text{y} \quad \widehat{\text{se}}(\hat{\phi}^{\text{MLE}}) = \frac{1}{M} \sum_{m=1}^M \widehat{\text{se}}(\hat{\phi}_m^{\text{MLE}}).$$

Por otra parte, para cada trayectoria, hemos simulado $B = 500$ trayectorias bootstrap representadas por $\{S_{m,b}\}_{b=1}^B$. Entonces, obtuvimos $\left\{ \{\hat{\phi}_{m,b}^b\}_{b=1}^B \right\}_{m=1}^M$ (las estimaciones ML). La estimación bootstrap y su error estándar estimado se definen como

$$\hat{\phi}_m^b = \frac{1}{B} \sum_{b=1}^B \hat{\phi}_{m,b}^b \quad \text{and} \quad \widehat{\text{se}}^2(\hat{\phi}_m^b) = \frac{1}{B-1} \sum_{b=1}^B (\hat{\phi}_{m,b}^b - \hat{\phi}_m^b)^2,$$

para $m = 1, \dots, M$. Por último, resumimos las M estimaciones bootstrap mediante

$$\hat{\phi}^b = \frac{1}{M} \sum_{m=1}^M \hat{\phi}_m^b \quad \text{and} \quad \widehat{\text{se}}(\hat{\phi}^b) = \frac{1}{M} \sum_{m=1}^M \widehat{\text{se}}(\hat{\phi}_m^b).$$

Además, como medida del rendimiento del estimador, utilizamos el error cuadrático medio (RMSE), y el coeficiente de variación (CV) estimado definido por

$$\widehat{\text{RMSE}}_{\hat{\phi}^{\text{MLE}}} = \left(\widehat{\text{se}}(\hat{\phi}^{\text{MLE}})^2 + \widehat{\text{bias}}_{\hat{\phi}^{\text{MLE}}}^2 \right)^{1/2}, \text{ and}$$

$$\widehat{\text{CV}}_{\hat{\phi}^{\text{MLE}}} = \frac{\widehat{\text{se}}(\hat{\phi}^{\text{MLE}})}{|\hat{\phi}^{\text{MLE}}|},$$

donde $\widehat{\text{bias}}_{\hat{\phi}^{\text{MLE}}} = \hat{\phi}^{\text{MLE}} - \phi$. Por último, estimamos la varianza del estimador mediante

$$\widetilde{\text{se}}^2(\hat{\phi}^{\text{MLE}}) = \frac{1}{M-1} \sum_{m=1}^M (\hat{\phi}_m^{\text{MLE}} - \hat{\phi}^{\text{MLE}})^2$$

De la misma manera, definimos $\widehat{\text{RMSE}}_{\hat{\phi}^{\text{b}}}$, $\widehat{\text{CV}}_{\hat{\phi}^{\text{b}}}$, $\widehat{\text{bias}}_{\hat{\phi}^{\text{b}}}$ y $\widetilde{\text{se}}^2(\hat{\phi}^{\text{b}})$ para el caso Bootstrap. La figura 3-3 muestra flujo de trabajo de este estudio de simulación

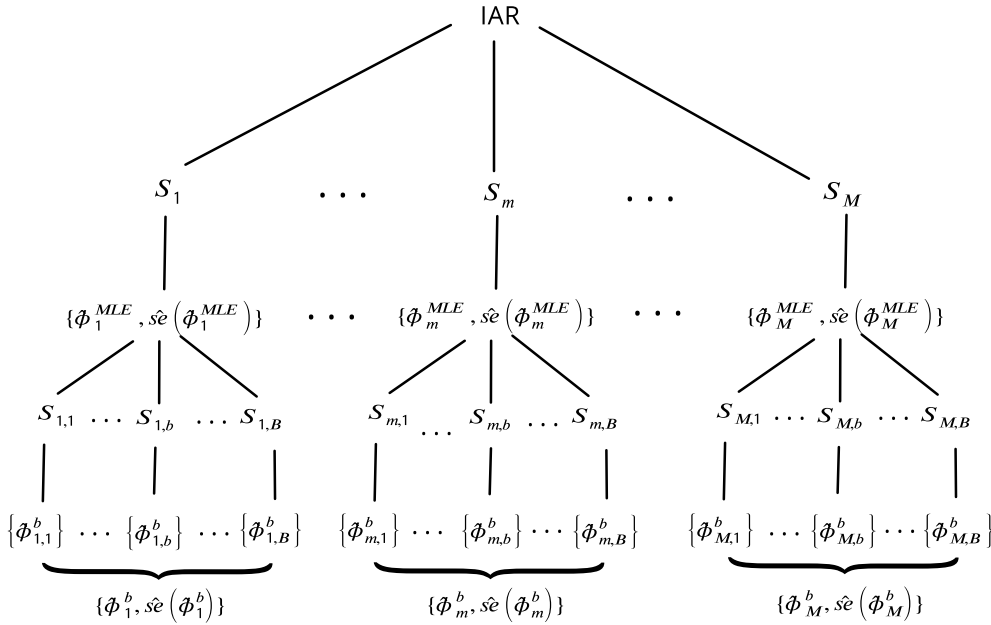


Figure 3-3: Esquema general del estudio de Monte Carlo.

En este estudio, se evaluaron los resultados del modelo en el caso de tiempo irregularmente espaciado, aunque también se comprobó que el modelo es válido para tiempos regulares con $\Delta_n = 1$ para todos los valores de n .

Para analizar el comportamiento de los estimadores de máxima verosimilitud y bootstrap, se realizaron simulaciones de distribuciones de muestras finitas. Los resultados de estas simulaciones se muestran en las figuras 3-4 y 3-5. Se puede observar que ambos estimadores presentan un comportamiento satisfactorio, parecen ser insesgados y consistentes en sus resultados.

Note que para valores pequeños de n y para valores cercanos a cero de ϕ , los estimadores tienden a tener una mayor variabilidad. Este hallazgo es importante para el análisis de los resultados, ya que sugiere que la precisión de los estimadores puede verse afectada por factores como la cantidad de datos disponibles y la magnitud del parámetro de interés. En segundo lugar, para ambos estimadores, presentamos las medidas de rendimiento para ambos métodos de estimación las cuales se ven reflejadas en la tabla 3-1

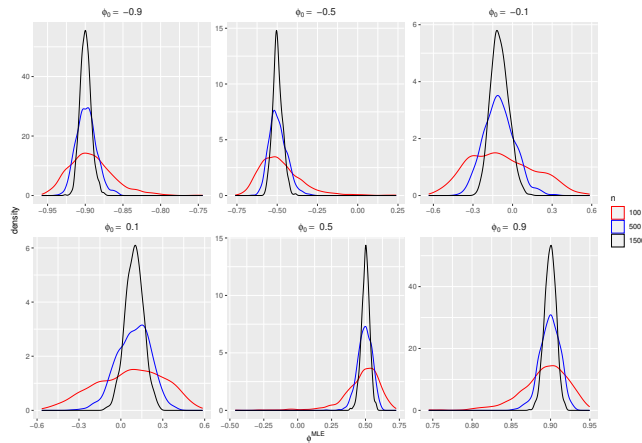


Figure 3-4: Resultados estimacion por máxima verosimilitud.

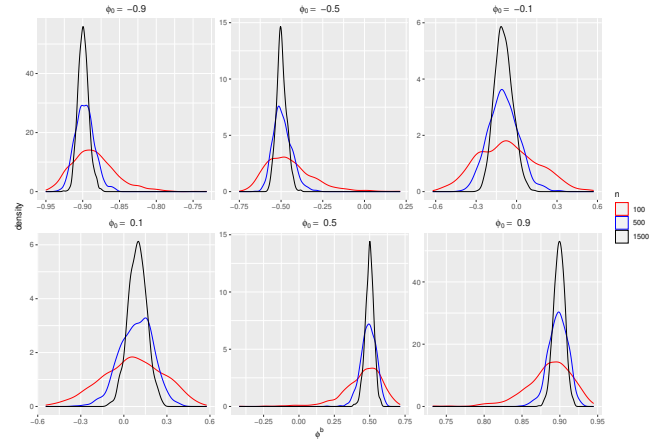


Figure 3-5: Resultados estimacion Bootstrap.

Table 3-1: Resultados del estudio de Monte Carlo

n	Máxima Verosimilitud						Bootstrap				
	ϕ	$\hat{\phi}^{MLE}$	$\hat{se}(\phi^{MLE})$	$\widehat{bias}_{\hat{\phi}^{MLE}}$	$\widehat{RMSE}_{\hat{\phi}^{MLE}}$	$\widehat{CV}_{\hat{\phi}^{MLE}}$	$\hat{\phi}^b$	$\hat{se}(\phi^b)$	$\widehat{bias}_{\hat{\phi}^b}$	$\widehat{RMSE}_{\hat{\phi}^b}$	$\widehat{CV}_{\hat{\phi}^b}$
100	-0.9	-0.892	0.029	-0.008	0.031	0.033	-0.885	0.032	-0.015	0.035	0.036
	-0.5	-0.479	0.126	-0.021	0.128	0.263	-0.450	0.153	-0.050	0.161	0.340
	-0.1	-0.081	0.210	-0.019	0.211	2.586	-0.068	0.222	-0.032	0.224	3.267
	0.1	0.062	0.211	0.038	0.214	3.388	0.060	0.222	0.040	0.226	3.726
	0.5	0.473	0.126	0.027	0.129	0.267	0.450	0.145	0.050	0.154	0.323
	0.9	0.892	0.029	0.008	0.030	0.033	0.886	0.032	0.014	0.035	0.036
500	-0.9	-0.898	0.013	-0.002	0.013	0.014	-0.897	0.013	-0.003	0.013	0.014
	-0.5	-0.495	0.052	-0.005	0.053	0.106	-0.491	0.054	-0.009	0.055	0.110
	-0.1	-0.101	0.110	0.001	0.110	1.095	-0.097	0.111	-0.003	0.111	1.153
	0.1	0.091	0.111	0.009	0.112	1.227	0.087	0.111	0.013	0.112	1.274
	0.5	0.495	0.052	0.005	0.052	0.105	0.491	0.054	0.009	0.054	0.109
	0.9	0.899	0.013	0.001	0.013	0.014	0.898	0.013	0.002	0.013	0.014
1500	-0.9	-0.900	0.007	0.000	0.007	0.008	-0.899	0.007	-0.001	0.007	0.008
	-0.5	-0.498	0.030	-0.002	0.030	0.060	-0.497	0.030	-0.003	0.030	0.060
	-0.1	-0.098	0.066	-0.002	0.066	0.673	-0.096	0.066	-0.004	0.066	0.688
	0.1	0.096	0.067	0.004	0.067	0.695	0.094	0.066	0.006	0.067	0.705
	0.5	0.500	0.030	0.000	0.030	0.059	0.498	0.030	0.002	0.030	0.060
	0.9	0.900	0.007	0.000	0.007	0.008	0.899	0.007	0.001	0.007	0.008

3.9 Comparación del Modelo IAR vs el CIAR

En este apartado, nos disponemos a comparar nuestra solución, con la solución propuesta en Elorrieta et al. (2019) descrita en la sección 2.4 (El modelo CIAR), para ello, se plantea un escenario de simulación similar al de la sección anterior, donde, $\sigma_0 = 1$, $\phi_0 \in \{-0.1, -0.5, -0.9\}$, $n \in \{100, 5000, 1500\}$ y $\Delta_n = \tau_n - \tau_{n-1} \sim 1 + \text{poisson}(\lambda = 2)$. En este caso solo simulamos muestras aleatorias con parámetro negativo ya que es el foco de este estudio. En el caso del IAR, se asumen que $\phi^I = 0$ (la parte imaginaria es cero) y que $c = 0$. Para comparar nuestro modelo con el CIAR, utilizaremos el *Error cuadrático medio (ECM)*, que definiremos así

$$ECM = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n}}$$

donde $e_i = X_{t_i} - \hat{X}_{\tau_i}$ para $i = 1, \dots, n$.

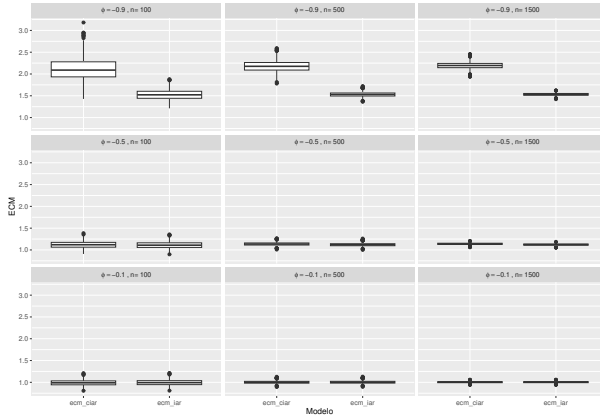


Figure 3-6: Comportamiento de la simulación, tomando como base el IAR. ECM

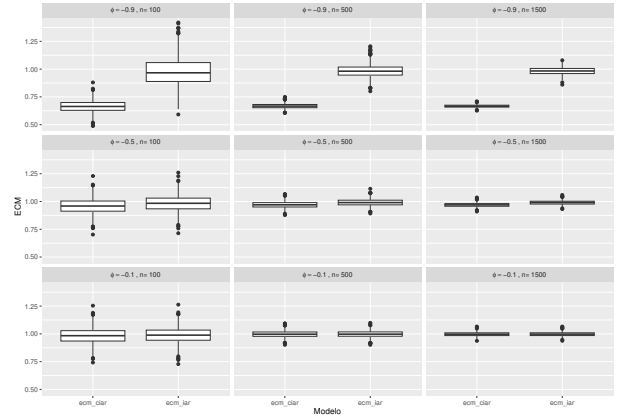


Figure 3-7: Comportamiento de la simulación, tomando como base el CIAR. ECM

la figura 3-6 muestra el resultado de la simulación de un IAR con las condiciones dadas, estimando tanto el IAR como el CIAR con los mismos datos simulados y midiendo el ECM de la predicción. Por otra parte la figura 3-7 muestra los mismos resultados pero esta vez, los datos fueron simulados a partir del CIAR.

Esta simulación permite llevarnos a la conclusión de que, cuando el parámetro $\phi_n \rightarrow -1$ cada modelo es mejor en su contexto independiente del tamaño de muestra, cuando el valor de ϕ disminuye, el ECM tiende a ser similar para ambos modelos independiente del contexto de la simulación y el tamaño de muestra. Lo que nos permite concluir que estos

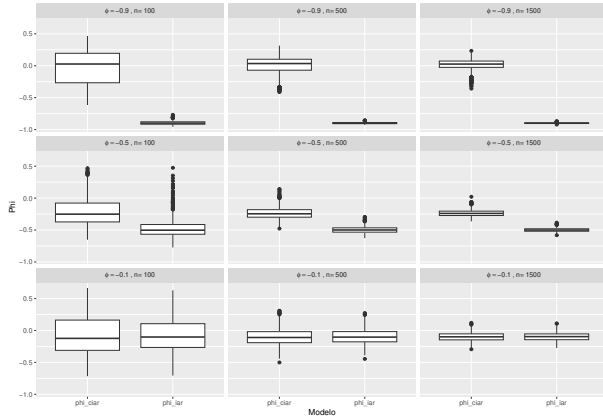


Figure 3-8: Comportamiento de la simulación, tomando como base el IAR. ϕ

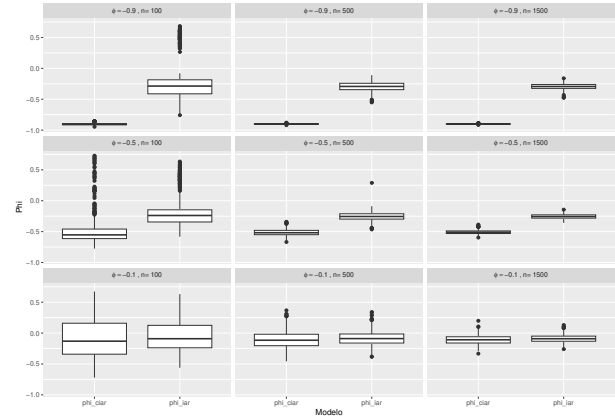


Figure 3-9: Comportamiento de la simulación, tomando como base el CIAR. ϕ

modelos son diferentes y cada uno modela estructuras diferentes. Como trabajo futuro se podría jugar con el parámetro imaginario del CIAR. Esto, se sale de los objetivos de este trabajo.

Además las figuras 3-8 y 3-9 nos muestran que los valores estimados de ϕ son totalmente diferentes, lo que nos afirma la conclusión de que ambos modelos son diferentes, sobre todo en valores cercanos a -1, habría que encontrar un punto donde sean comparables.

3.10 Aplicación en datos astronómicos

Estos datos de Lira Teillery et al. (2015), describen una serie temporal de la cantidad de energía en ergios que llega por cm^2 en un ancho de banda de un Amstrong ($10^{(-15)}ergs/s/cm^2/A$) desde la galaxia activa (AGN) MCG-6-30-15 medida en la banda K entre agosto de 2006 y julio de 2011. La medición se realizó utilizando la cámara ANDICAM, la cual estaba montada en el telescopio de 1.3 metros ubicado en el Observatorio Interamericano Cerro Tololo (CTIO). El tiempo t_n está medido en días julianos heliocentricos.

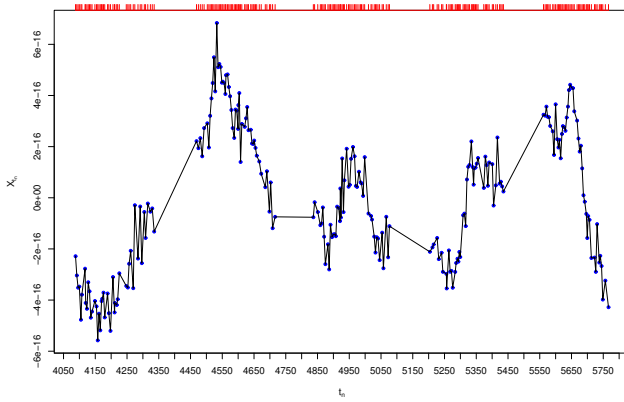


Figure 3-10: Comportamiento de la serie de tiempo original.

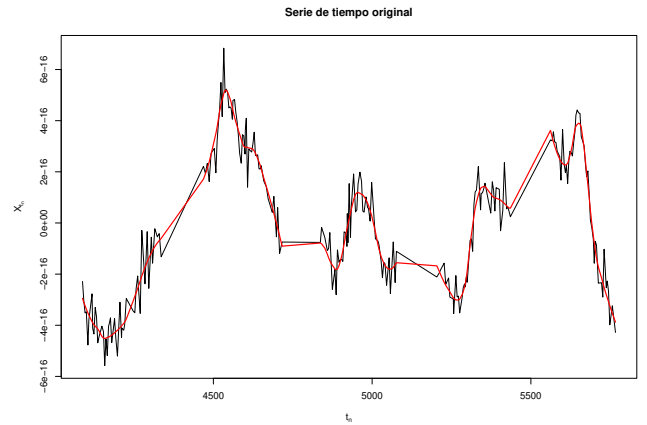


Figure 3-11: Ajuste del modelo Loess.

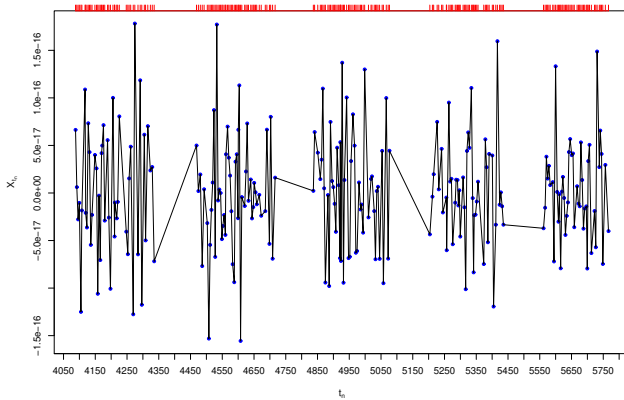


Figure 3-12: Residuales del modelo Loess (eliminando tendencias).

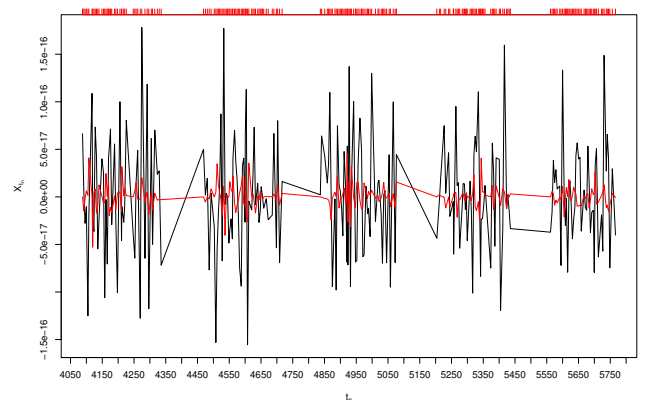


Figure 3-13: Ajuste del modelo IAR a la serie transformada .

La Figura 3-10 ilustra el comportamiento de la misma, donde los ticks rojos indican pa-

trones irregulares en la serie. Se observan tendencias en ciertos segmentos donde la media no es constante. Para modelar adecuadamente esta serie, es necesario eliminar estas tendencias utilizando un modelo Loess. El resultado de aplicar el modelo Loess se muestra en la Figura 3-11.

Posteriormente, en la Figura 3-12, se presenta el comportamiento de los residuos del modelo Loess. Se puede apreciar que hemos logrado eliminar los cambios en la media de la serie. Con base en esta serie transformada, realizaremos las estimaciones del modelo IAR. El ajuste del modelo IAR se muestra en la Figura 3-13.

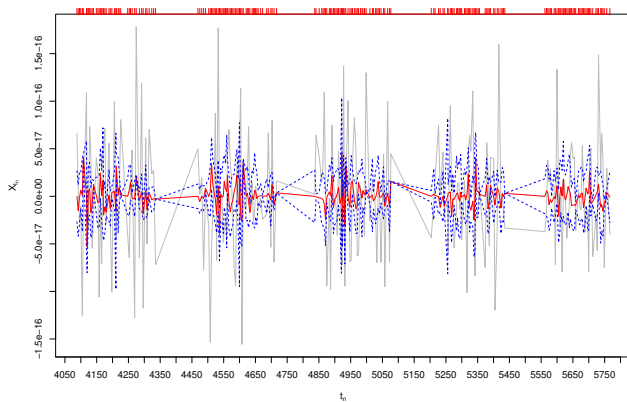


Figure 3-14: Comportamiento de las predicciones y su intervalo de predicción.

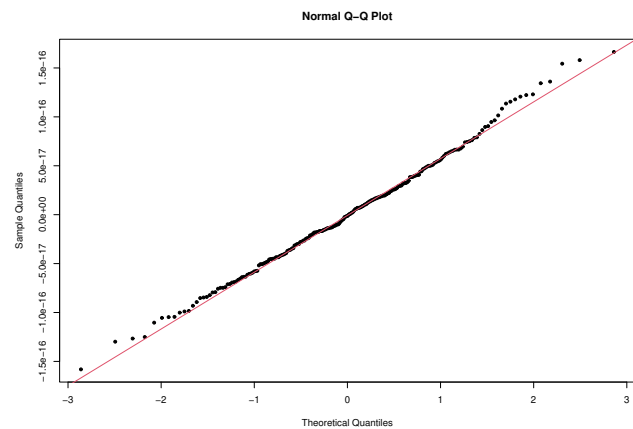


Figure 3-15: qqplot de los residuos del modelo.

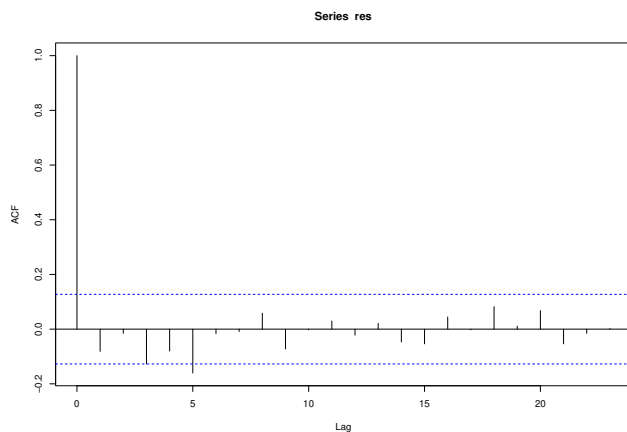


Figure 3-16: ACF de los residuos.

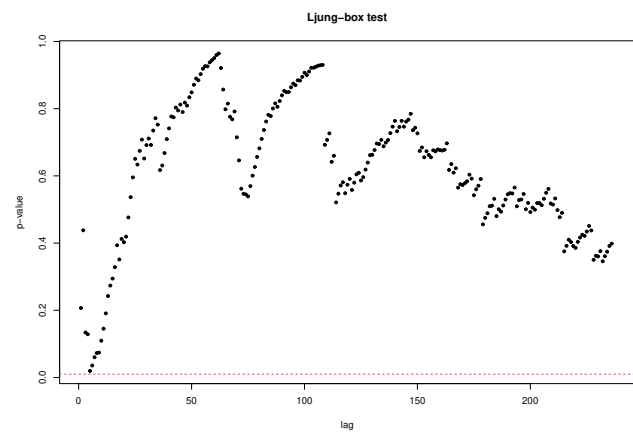


Figure 3-17: Ljung-Box test para los residuos del modelo.

Los parametros estimados para la serie transformada son los siguientes:

$$\begin{aligned}\hat{\phi}^{MLE} &= -0.69 & \hat{se}(\phi^{MLE}) &= 0.0613 \\ \hat{\phi}^b &= -0.67 & \hat{se}(\phi^b) &= 0.1012\end{aligned}$$

En este problema, se puede ver que el parametro estimado es negativo, cosa que no podriamos haber logrado con el modelo IAR original. Ahora, miremos como se comportan los supuestos de este modelo. Las figuras 3-15, 3-16 y 3-17 nos permite estar tranquilos con los supuestos sobre los errores del modelo, puesto que no hay problemas de normalidad ni dependencia en los mismos.

Chapter 4

Conclusiones

A lo largo de este trabajo, se ha enfatizado la importancia tanto práctica como teórica de abordar los procesos estocásticos con espaciamiento irregular. Se ha desarrollado un modelo autorregresivo irregular (IAR) que permite estructuras de autocorrelación tanto negativas como positivas, lo cual amplía las aplicaciones potenciales de estos modelos. El modelo IAR propuesto ha demostrado satisfacer estas características de manera efectiva.

Además, se han encontrado estimadores de máxima verosimilitud y bootstrap para el modelo propuesto, los cuales se han demostrado insesgados, consistentes y de fácil implementación. Estas estimaciones han sido comparadas con un modelo existente llamado modelo CIAR, y se ha observado que los estimadores propuestos son igual de buenos o incluso mejores en algunos casos particulares.

También se destaca la relevancia práctica de este modelo al analizar un escenario de aplicación en el cual se obtiene un parámetro negativo, que fue el enfoque principal de este proyecto. Esto resalta la importancia práctica del modelo propuesto y abre la puerta para que otros investigadores lo utilicen en sus problemas específicos.

Un trabajo futuro para este tipo de modelos es proponer estimaciones bayesianas, las cuales podrían mejorar la precisión de los estimadores. Además, se sugiere realizar una comparación más exhaustiva con otros modelos, como el CIAR, para encontrar posibles equivalencias o demostrar definitivamente que son modelos diferentes.

Los investigadores pueden seguir esta metodología para explorar funciones más complejas de ϕ_n y generar soluciones adicionales para este interesante y apasionante problema.

References

- Adorf, H.-M. (1995), Interpolation of irregularly sampled data series—a survey, in ‘Astronomical Data Analysis Software and Systems IV’, Vol. 77, p. 460.
- Dilmaghani, S., Henry, I. C., Soonthornnonda, P., Christensen, E. R. & Henry, R. C. (2007), ‘Harmonic analysis of environmental time series with missing data or irregular sample spacing’, *Environmental science & technology* **41**(20), 7030–7038.
- Elorrieta, F., Eyheramendy, S. & Palma, W. (2019), ‘Discrete-time autoregressive model for unequally spaced time-series observations’, *Astronomy & Astrophysics* **627**, A120.
- Eyheramendy, S., Elorrieta, F. & Palma, W. (2018), ‘An irregular discrete time series model to identify residuals with autocorrelation in astronomical light curves’, *Monthly Notices of the Royal Astronomical Society* **481**(4), 4311–4322.
- Ghaderpour, E. & Vujadinovic, T. (2020), ‘Change detection within remotely sensed satellite image time series via spectral analysis’, *Remote Sensing* **12**(23), 4001.
- Kuttu, S. & Bokpin, G. A. (2017), ‘Feedback trading and autocorrelation patterns in sub-saharan african equity markets’, *Emerging Markets Finance and Trade* **53**(1), 213–225.
- Lira Teillery, P., Arévalo, P., Uttley, P., McHardy, I. & Videla, L. (2015), ‘Long-term monitoring of the archetype seyfert galaxy mcg-6-30-15: X-ray, optical and near-ir variability of the corona, disc and torus’.
- Liu, M., Stella, F., Hommersom, A., Lucas, P. J., Boer, L. & Bischoff, E. (2019), ‘A comparison between discrete and continuous time bayesian networks in learning from clinical time series data with irregularity’, *Artificial intelligence in medicine* **95**, 104–117.
- Novoseltseva, M., Gutova, S. & Glinchikov, K. (2019), Spatial and econometric modeling of the demographic situation in the kemerovo region, in ‘IOP Conference Series: Earth and Environmental Science’, Vol. 272, IOP Publishing, p. 032155.

- Ojeda, C., Palma, W., Eyheramendy, S. & Elorrieta, F. (2023), A novel first-order autoregressive moving average model to analyze discrete-time series irregularly observed, *in* O. Valenzuela, F. Rojas, L. J. Herrera, H. Pomares & I. Rojas, eds, 'Theory and Applications of Time Series Analysis and Forecasting', Springer International Publishing, Cham, pp. 91–103.
- Rindorf, A., Cadigan, N., Howell, D., Eero, M. & Gislason, H. (2020), 'Periodic fluctuations in recruitment success of atlantic cod', *Canadian Journal of Fisheries and Aquatic Sciences* **77**(2), 236–246.
- Robinson, P. (1977), 'Estimation of a time series model from unequally spaced data', *Stochastic Processes and their Applications* **6**(1), 9–24.
- Shamsan, A., Wu, X., Liu, P. & Cheng, C. (2020), 'Intrinsic recurrence quantification analysis of nonlinear and nonstationary short-term time series', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30**(9), 093104.
- Yang, Q., Fowler, M. S., Jackson, A. L. & Donohue, I. (2019), 'The predictability of ecological stability in a noisy world', *Nature ecology & evolution* **3**(2), 251–259.