

Análisis de Series de Tiempo

Sofía Villers Gómez Carlos Fernando Vásquez Guerra
David Alberto Mateos Montes de Oca

Contents

I	Series de tiempo	6
1	Introducción	7
2	Descomposición temporal	9
2.1	Ejercicios	11
3	Procesos estocásticos	13
3.1	Proceso estocástico estacionario	13
3.2	Ruido blanco (“white noise”)	14
3.3	Caminata aleatoria	14
3.4	Series de tiempo	15
3.5	Ejercicios	17
4	Funciones de autocovarianza y autocorrelación	18
4.1	Función de autocorrelación parcial	18
4.2	Correlograma	19
4.3	Prueba de Ljung-Box	19
4.4	Ejercicios	20
5	Transformaciones	21
5.1	Suavizamiento por medias móviles	21
5.2	Suavizamiento por polinomios ajustados	21
5.3	Diferencias de Box-Jenkins	22
5.4	Ejercicios	22
II	Procesos lineales estacionarios	24
6	$AR(p)$: Proceso Autoregresivo	25
6.1	Proceso Autoregresivo de orden 1: $AR(1)$	25
6.1.1	Estacionario en Media	25
6.1.2	Estacionario en Covarianza	26
6.2	Proceso Autoregresivo de orden 2: $AR(2)$	28
6.2.1	Estacionario en Media	28

<i>CONTENTS</i>	2
6.2.2 Estacionario en Covarianza	29
6.3 Ejercicios	31
7 $MA(q)$: Proceso de Medias Móviles	32
7.1 Proceso de Medias Móviles de orden 1: $MA(1)$	32
7.2 Proceso de Medias Móviles de orden 2: $MA(2)$	34
7.3 Ejercicios	35
8 $ARMA(p, q)$: Proceso Autoregresivo de Medias Móviles	36
8.1 $ARMA(1, 1)$	37
8.1.1 Media	37
8.1.2 Covarianzas	37
8.2 Ejercicios	39
III Procesos lineales no estacionarios	40
9 $ARIMA(p, d, q)$: Proceso Autoregresivo Integrado y de Medias Móviles	41
9.1 Ejercicios	44
10 $IMA(d, q)$: Proceso Integrado y de Media Móvil	45
IV Otras consideraciones	48
11 Predictores Lineales	49
11.1 Ejercicios	50
12 Construcción de modelos para series de tiempo univariadas	51
12.1 Ejercicios	52
V Procesos no lineales	53
13 Introducción	54
14 $ARCH(q)$: Proceso Autoregresivo con Heterocedasticidad Condicional	55
15 $GARCH(p, q)$: Proceso Autoregresivo Generalizado con Heterocedasticidad Condicional	58
VI Tareas	60

Prefacio

Primera edición del bookdown *Análisis de Series de Tiempo* para su uso en la materia Análisis de Supervivencia y Series de tiempo y sus relacionadas impartidas por los autores, así como para aquellos estudiantes que deseen adquirir el conocimiento pertinente de tal tópico.

Objetivos

- Otorgar un material electrónico de calidad con el contenido referente al Análisis de Series de Tiempo como un esfuerzo de los autores para lograr un proceso de aprendizaje autodidacta por parte del alumno y así optimizar el tiempo, tanto de los profesores, como el de los alumnos.
- Plasmar las bases teóricas de esta rama de la estadística con el uso de ejemplos y contenido visual para un mejor entendimiento de cada subtema que se trate.
- Dar continuidad al material para el curso *Análisis de Supervivencia y Series de Tiempo*.

Estructura

Este libro se compone de 15 capítulos. Se recomienda que la consulta de los capítulos se realice de acuerdo al índice, ya que a medida que se avanza en índice, se asume el conocimiento de los capítulos previos.

Detalles técnicos

Para la creación de este material se hizo uso de varios sistemas de software como LaTeX y CSS para el diseño de ciertos elementos. Todos los cálculos y gráficas fue creado con el lenguaje de programación R ya sea con el uso del paquete `base` o algún otro de los paquetes que se mencionan a continuación.

```
R version 4.3.1 (2023-06-16)
Platform: aarch64-apple-darwin20 (64-bit)
Running under: macOS Sonoma 14.2.1

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRlapack.dylib; LAPACK v

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

time zone: America/Mexico_City
tzcode source: internal

attached base packages:
[1] stats      graphics  grDevices  utils      datasets  methods    base
```

other attached packages:

```
[1] TSA_1.3.1      forecast_8.21.1 kableExtra_1.3.4 knitr_1.45
[5] latex2exp_0.9.6 patchwork_1.1.3 lubridate_1.9.3 forcats_1.0.0
[9] stringr_1.5.1  dplyr_1.1.4     purrr_1.0.2     readr_2.1.4
[13] tidyr_1.3.0    tibble_3.2.1    ggplot2_3.4.4    tidyverse_2.0.0
```

loaded via a namespace (and not attached):

```
[1] gtable_0.3.4      xfun_0.41        lattice_0.22-5    tzdb_0.4.0
[5] quadprog_1.5-8    vctrs_0.6.5      tools_4.3.1       generics_0.1.3
[9] curl_5.2.0        parallel_4.3.1    fansi_1.0.6       xts_0.13.1
[13] pkgconfig_2.0.3    Matrix_1.6-4     webshot_0.5.5     lifecycle_1.0.4
[17] compiler_4.3.1     munsell_0.5.0     fontawesome_0.5.2 leaps_3.1
[21] htmltools_0.5.7    yaml_2.3.8        pillar_1.9.0      nlme_3.1-164
[25] fracdiff_1.5-2     locfit_1.5-9.8    tidyselect_1.2.0  rvest_1.0.3
[29] digest_0.6.33      stringi_1.8.3     bookdown_0.37     splines_4.3.1
[33] tseries_0.10-55    fastmap_1.1.1     grid_4.3.1        colorspace_2.1-0
[37] cli_3.6.2          magrittr_2.0.3    utf8_1.2.4        withr_2.5.2
[41] scales_1.3.0        timechange_0.2.0  TTR_0.24.4        rmarkdown_2.25
[45] httr_1.4.7         quantmod_0.4.25   nnet_7.3-19       timeDate_4032.109
[49] zoo_1.8-12         hms_1.1.3         urca_1.3-3        evaluate_0.23
[53] lmtest_0.9-40      viridisLite_0.4.2 mgcv_1.9-1         rlang_1.1.2
[57] Rcpp_1.0.11        glue_1.6.2        xml2_1.3.6         svglite_2.1.3
[61] rstudioapi_0.15.0 R6_2.5.1          systemfonts_1.0.5
```

Este libro fue escrito con bookdown usando RStudio.

Esta versión fue escrita con:

Finding R package dependencies ... Done!

```
setting  value
version  R version 4.3.1 (2023-06-16)
os       macOS Sonoma 14.2.1
system   aarch64, darwin20
ui       X11
language (EN)
collate  en_US.UTF-8
ctype    en_US.UTF-8
tz       America/Mexico_City
date     2024-08-19
pandoc   3.1.1 @ /Applications/RStudio.app/Contents/Resources/app/quarto/bin/tools/ (via rmarkdown)
```

Licencia

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

This is a human-readable summary of (and not a substitute for) the license. Please see <https://creativecommons.org/licenses/by-sa/4.0/legalcode> for the full legal text.

You are free to:

- **Share**—copy and redistribute the material in any medium or format
- **Remix**—remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

- **Attribution**—You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **ShareAlike**—If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.
- **No additional restrictions**—You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Part I

Series de tiempo

Chapter 1

Introducción

El objetivo en esta parte será responder preguntas como: *¿Cuál será el precio de las acciones de Facebook para el último bimestre del 2024?*, *¿Cuál será el nivel de partículas contaminantes en la CDMX para noviembre de 2024?*, *¿Cuál será la capacidad de un procesador intel para el año 2025?*, puede parecer, en primera instancia, una tarea complicada. Si bien no tenemos una “bola mágica” con la que podamos adivinar el futuro, disponemos de ciertos procesos estocásticos llamados *Series de Tiempo*, cuyo objetivo principal es el *pronóstico*.

Una serie tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y espaciados entre sí de manera uniforme (diario, semanal, semestral, anual, entre otros). Por esta razón, los datos usualmente son dependientes entre sí. El principal objetivo de una serie de tiempo (denotada por X_t , donde $t = 1, 2, \dots, n$) es realizar un análisis de los datos para conocer su comportamiento a través del tiempo y, a partir de dicho conocimiento, y asumiendo que mantendrá un comportamiento similar al observado, poder realizar predicciones en uno o más períodos de tiempo situados en el futuro.

Se pueden encontrar series de tiempo en diferentes campos de estudio, por ejemplo:

- Economía: Índices de precios mensuales, exportaciones totales mes a mes, tasa de interés semanal.
- Física: Nivel de precipitación diaria, temperatura diaria o mensual.
- Demografía: Tasa de natalidad, tasa de mortalidad, número de habitantes en cierta región.
- Marketing: Ventas diarias o mensuales de la compañía.

En el análisis de las series de tiempo es necesario señalar que, para llevarlo a cabo, hay que tener en cuenta los siguientes supuestos:

- Se considera que existe una cierta estabilidad en la estructura del fenómeno estudiado. Para que se cumpla este supuesto será necesario estudiar períodos lo más homogéneos posibles.
- Los datos deben ser homogéneos en el tiempo, es decir, se debe mantener la definición y la medición de la magnitud objeto de estudio. Este supuesto no se da en muchas de las series económicas, ya que es frecuente que las estadísticas se perfeccionen con el paso del tiempo, produciéndose saltos en la serie debidos a un cambio en la medición de la magnitud estudiada. Un caso particularmente frecuente es el cambio de base en los índices de precios, de producción, etc. Tales cambios de base implican cambios en los productos y las ponderaciones que entran en la elaboración del índice que repercuten considerablemente en la comparabilidad de la serie en el tiempo.

Como se mencionó previamente el objetivo es realizar un análisis de la serie de tiempo y construir un modelo matemático que refleje el comportamiento de los datos a través del tiempo. Dicho modelo servirá para proyectar los valores de la serie en el futuro, por lo tanto la calidad de las predicciones dependerán, en buena medida, del proceso generador de la serie: así, si la variable observada sigue algún tipo de esquema o patrón de comportamiento más o menos fijo (serie determinista) seguramente obtengamos predicciones más o menos fiables, con un grado de error bajo. Por el contrario, si la serie no sigue ningún

patrón de comportamiento específico (serie aleatoria), seguramente nuestras predicciones carecerán de validez por completo.

Dentro de los métodos de predicción cuantitativos, se pueden distinguir dos grandes enfoques alternativos:

- 1) Por un lado, el análisis univariante de series temporales mediante el cual se intenta realizar previsiones de valores futuros de una variable, utilizando como información la contenida en los valores pasados de la propia serie temporal. Dentro de esta metodología se incluyen los métodos de descomposición y la familia de modelos ARIMA univariantes que veremos más adelante.
- 2) El otro gran bloque dentro de los métodos cuantitativos estaría integrado por el análisis multivariante o de tipo causal, denominado así porque en la explicación de la variable o variables objeto de estudio intervienen otras adicionales a ella o ellas mismas.

En este curso nos centraremos en el análisis univariante de las series de tiempo.

En la mayoría de capítulos se adjuntarán al final diferentes enlaces al Post SeriesTCode de Rpubs, en el cual se da un enfoque aplicado a cada tema. Se recomienda que al final de cada capítulo se revise el contenido correspondiente para obtener un panorama completo del tema. En este caso en particular se solicita estudiar la sección Datos para conocer los objetos `ts` de R.

En caso de necesitar un repaso u obtener conocimiento del lenguaje de programación R, se recomienda revisar el siguiente post y si se desea obtener un conocimiento del lenguaje intermedio-avanzado se recomienda leer el post EfficientR.

También se dejan los siguientes enlaces relacionados al aprendizaje de este lenguaje de programación:

- Programación en R
- R for Data Science (2e)
- Advanced R

Chapter 2

Descomposición temporal

El análisis clásico de las series temporales se basa en la suposición de que los valores que toma la variable de observación es la consecuencia de tres componentes, cuya actuación conjunta da como resultado los valores medidos.

Los componentes de una serie de tiempo son:

1. **Tendencia:** La tendencia o tendencia a largo plazo de una serie es por lo común el resultado de factores a largo plazo. En términos intuitivos, la tendencia de una serie de tiempo caracteriza el patrón gradual y consistente de las variaciones de la propia serie, que se consideran consecuencias de fuerzas persistentes que afectan el crecimiento o la reducción de la misma, tales como: cambios en la población, en las características demográficas de la misma, cambios en los ingresos, en la salud, en el nivel de educación y tecnología. Las tendencias a largo plazo se ajustan a diversos esquemas. Algunas se mueven continuamente hacia arriba, otras declinan, y otras más permanecen igual en un cierto período o intervalo de tiempo.
2. **Estacionalidad:** El componente de la serie de tiempo que representa la variabilidad en los datos debida a influencias de las estaciones, se llama componente estacional. Esta variación corresponde a los movimientos de la serie que ocurren año tras año en los mismos meses (o en los mismos trimestres) del año poco más o menos con la misma intensidad. Por ejemplo: Un fabricante de albercas inflables espera poca actividad de ventas durante los meses de otoño e invierno y tiene ventas máximas en los de primavera y verano, mientras que los fabricantes de equipo para la nieve y ropa de abrigo esperan un comportamiento anual opuesto al del fabricante de albercas.
3. **Componente aleatoria:** Esta se debe a factores a corto plazo, imprevisibles y no recurrentes que afectan a la serie de tiempo. Como este componente explica la variabilidad aleatoria de la serie, es impredecible, es decir, no se puede esperar predecir su impacto sobre la serie de tiempo. Existen dos tipos de variación irregular:
 - a) Las variaciones que son provocadas por acontecimientos especiales, fácilmente identificables, como las elecciones, inundaciones, huelgas, terremotos.
 - b) Variaciones aleatorias o por casualidad, cuyas causas no se pueden señalar en forma exacta, pero que tienden a equilibrarse a la larga.

Se puede observar que de los tres componentes, los dos primeros son determinísticos, mientras que la última es aleatoria.

Adicionalmente, la asociación de estos componentes en una serie temporal puede ser de tipo aditivo, multiplicativo o combinación de los dos anteriores

Entonces la serie de tiempo en un esquema aditivo se puede expresar como:

$$X_t = T_t + E_t + I_t$$

Mientras que la serie de tiempo en un esquema multiplicativo se expresaría como:

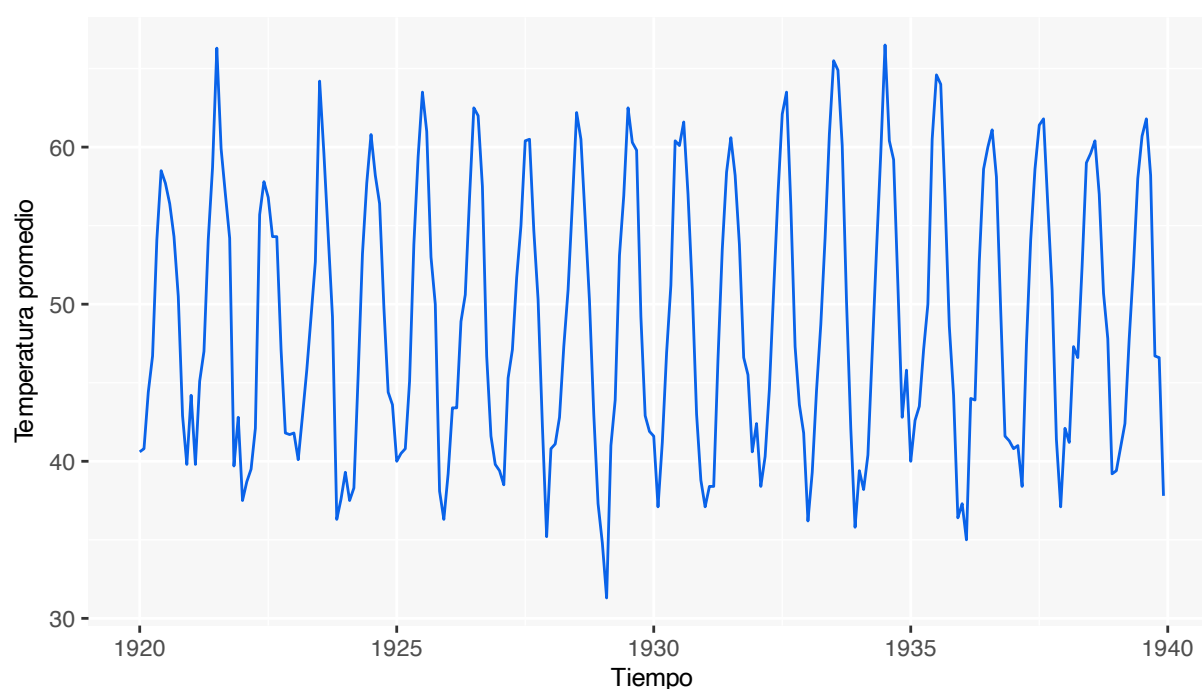
$$X_t = T_t * E_t * I_t$$

Y un ejemplo de combinación se expresaría como:

$$X_t = (T_t + E_t) * I_t$$

donde T_t es el componente de **tendencia**, E_t el componente de **estacionalidad** y I_t la componente **aleatoria**.

Las siguientes gráficas son un ejemplo basado en el conjunto de datos **nottem**, en el cual se tienen las temperaturas mensuales promedio en Nottingham entre 1920 y 1939. Esta serie de tiempo tiene el siguiente comportamiento a través de los años.



En la gráfica se puede observar que los datos parecen tener un comportamiento cíclico, lo cual es esperado ya que los datos corresponden a temperaturas mensuales. Así mismo el gráfico no muestra que exista una tendencia en los datos es decir no podemos decir que las temperaturas incrementan o decrecientan al pasar del tiempo.

La función R `decompose()`, obtiene las componentes de tendencia, estacionalidad y aleatoria de una serie temporal a través de medias móviles, y además permite obtener los componentes en base a un esquema aditivo ó multiplicativo. Es una función genérica de R, lo que significa que no requiere de la instalación de ninguna librería.

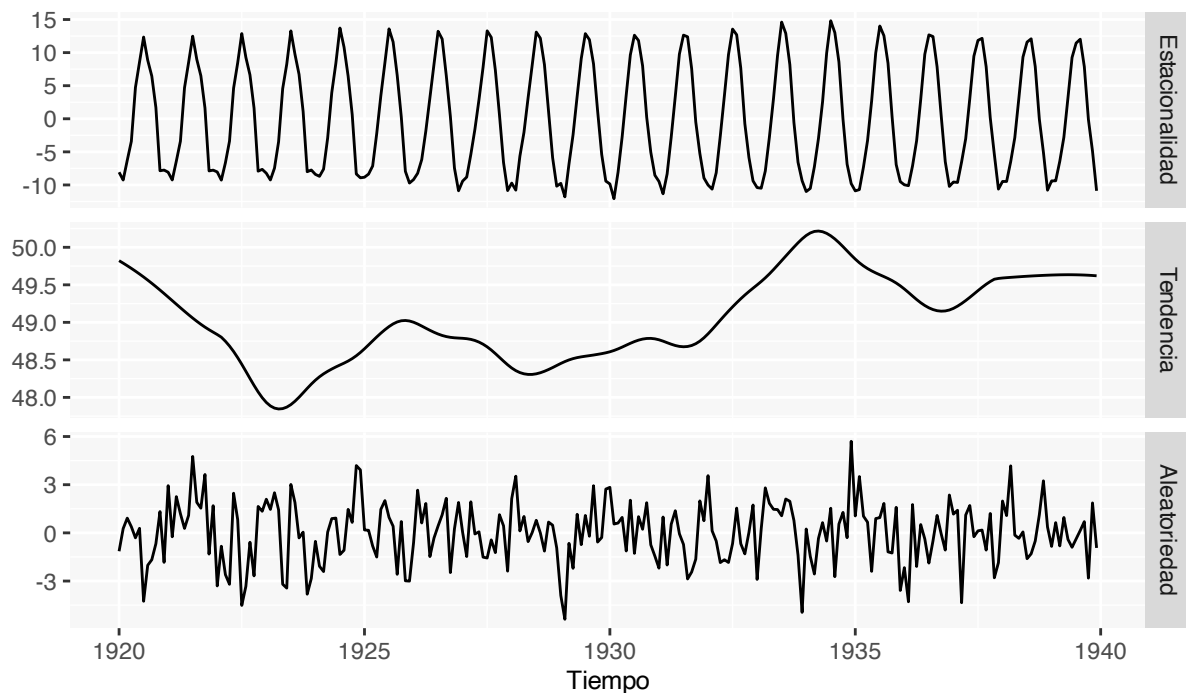
La forma de utilizar esta función es la siguiente:

```
decompose(x, type = c("additive", "multiplicative"), filter = NULL)
```

La función calcula el componente de tendencia utilizando medias móviles, (si `filter = NULL`, se utilizan medias móviles simétricas), los índices de estacionalidad son promedios de los índices de estacionalidad que se obtienen al desestacionalizar la serie por el modelo elegido, por último, el componente irregular se obtiene eliminando la tendencia y estacionalidad de la serie temporal.

Nota importante es que la función requiere que los datos tengan forma de serie temporal.

Ahora aplicaremos la función `decompose()` a la serie de tiempo graficada anteriormente. Las siguientes gráficas representan las componentes antes mencionadas de la serie de tiempo correspondiente a las temperaturas mensuales promedio en Nottingham entre 1920 y 1939.



La descomposición nos corrobora la presencia de una estacionalidad en los datos y detecta una ligera tendencia a la alza, aunque la escala en el gráfico de la tendencia es muy pequeña muestra que al pasar el tiempo las temperaturas promedio parecen estarse incrementando. En cuanto al gráfico de aleatoriedad la descomposición efectivamente lo presenta como una serie de datos aleatorios.

Se recomienda la lectura de la sección Descomposición para ver un ejemplo aplicado de la descomposición de series de tiempo, ya sea con la función `decompose()`, o la famosa función `stl()`. Además, pueden existir casos en que no se tenga suficiente información para descomponer una serie, por lo que se pueden utilizar otras alternativas como la mostrada en el apartado Un caso curioso del mismo blog, en donde se utilizan las transformaciones de Fourier para obtener una descomposición temporal.

2.1 Ejercicios

1. ¿Qué características debe tener un conjunto de datos para considerarse una serie de tiempo?
2. Una serie de tiempo se puede descomponer en: (Seleccione la opción correcta)
 - a) Parte Multiplicativa, Parte Aditiva y Parte Aleatoria
 - b) Tendencia, Estacionalidad y Parte Aleatoria
 - c) Parte Univariante y Parte Multivariante
3. Utilizar los datos en el objeto “AirPassengers” que corresponden a 144 mediciones del número de pasajeros mensuales en líneas aéreas y:
 - a) Grafique la serie de tiempo y describa que patrón observa. Comente su código.
 - b) Aplique la función `decompose()` a los datos y describa los hallazgos. ¿Qué puede concluir además de los hallazgos técnicos?
 - c) Grafique la serie de los logaritmos de los datos y describa el cambio observado respecto a lo visto en el inciso a)
4. Seleccione un conjunto de datos de algún tema de su interés, plantee un problema/pregunta que quiera responder con sus datos utilizando análisis de series de tiempo.

- a. Limpie, prepare y transforme los datos a formato de series de tiempo.
- b. Grafique la serie de tiempo y describa que patrón observa.

Chapter 3

Procesos estocásticos

La palabra estocástico, que tiene origen Griego, era usada bajo el significado de *perteneciente al azar*. En ese sentido un proceso estocástico se define como un conjunto de variables aleatorias ordenadas según el tiempo (o el espacio que corresponda), el cual puede ser continuo o discreto. Se denota la variable aleatoria en el tiempo t por $X(t)$ o X_t con $-\infty < t < \infty$ en caso de ser continua o bien $t = (0, \pm 1, \pm 2, \dots)$ en caso de ser discreta.

Como un ejemplo de proceso estocástico tenemos, el número de personas que esperan en una fila en el instante t del tiempo.

3.1 Proceso estocástico estacionario

Al trabajar con series de tiempo el escenario ideal sería trabajar con series que posean la característica de tener la media y la autocovarianza constantes a lo largo del tiempo, es decir, que sus valores oscilen dentro de un rango de valores y no muestren tendencia clara, creciente o decreciente, teóricamente conocidos como procesos estacionarios. Para definir un proceso estocástico como un proceso estacionario es sumamente necesario conocer la función de densidad conjunta de las variables aleatorias que conforman el proceso, no obstante, en la práctica no es común que se logre. Es por esto por lo que los procesos estacionarios se pueden definir de la siguiente forma:

- **Procesos estrictamente estacionarios:** Se dice que un proceso $X_t, t \in Z$ es estrictamente estacionario, si sus funciones de densidad para un conjunto arbitrario de variables $X_t, X_{t+1}, \dots, X_{t+m}$ son invariantes respecto a desplazamientos en el tiempo, es decir, que cumplen:

$$f(X_t, X_{t+1}, \dots, X_{t+m}) = f(X_{t+k}, X_{t+k+1}, \dots, X_{t+k+m}) \quad \forall t, m, k \in Z$$

- **Procesos débilmente estacionarios:** Un proceso $X_t, t \in Z$ se dice que es débilmente estacionario de orden k si los primeros k momentos son invariantes a través del tiempo. Podemos definir un proceso débilmente estacionario de segundo orden si cumple lo siguiente:

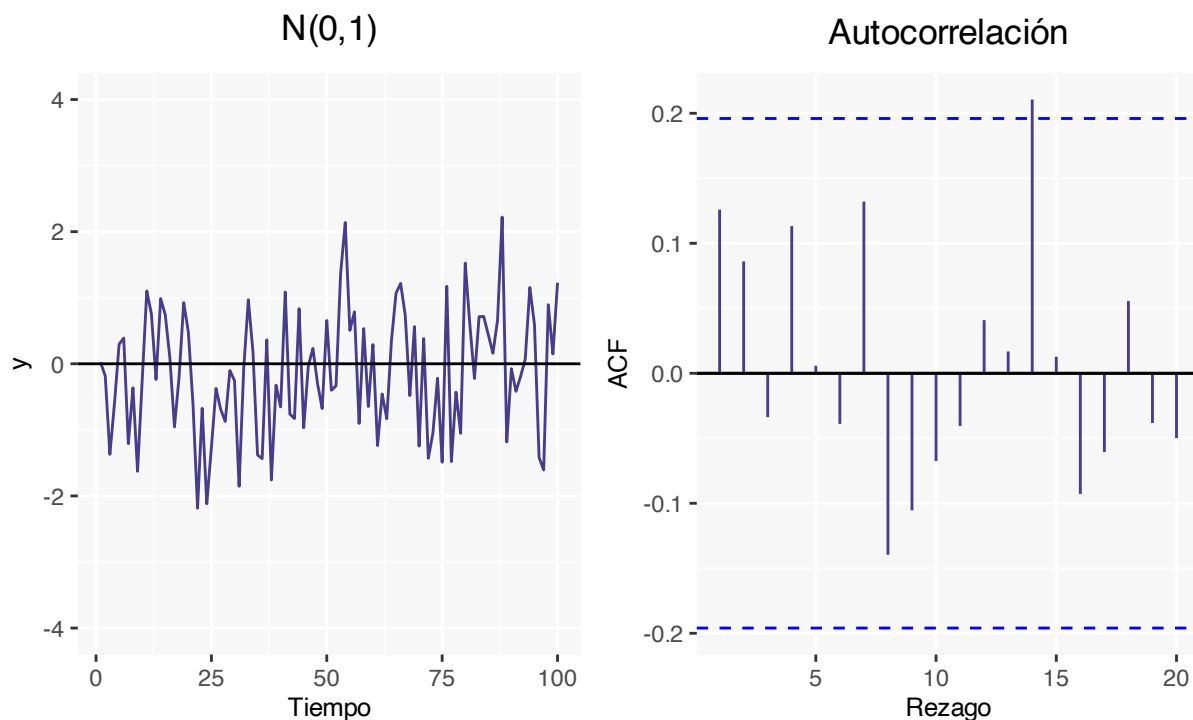
1. $E[|X_t|^2] < \infty \forall t \in Z$
2. $E[X_t] = \mu \quad \forall t \in Z$, lo cual quiere decir que las esperanzas de las variables aleatorias son independientes del tiempo por lo cual permanecen constantes.
3. $Cov(X_t, X_s) = Cov(X_{t+m}, X_{s+m}) \forall t, s, m \in Z$, es decir las covarianzas de dos variables aleatorias del proceso que se encuentran en distintos puntos del tiempo dependen solamente del lapso transcurrido entre cada una de ellas.

3.2 Ruido blanco (“white noise”)

Un ruido blanco¹ es un caso simple de los procesos estocásticos, donde los valores son independientes e idénticamente distribuidos a lo largo del tiempo con media cero e igual varianza, se denota por ϵ_t .

$$\epsilon_t \sim N(0, \sigma^2)$$

$$Cov(\epsilon_{t_i}, \epsilon_{t_j}) = 0 \forall t_i \neq t_j$$

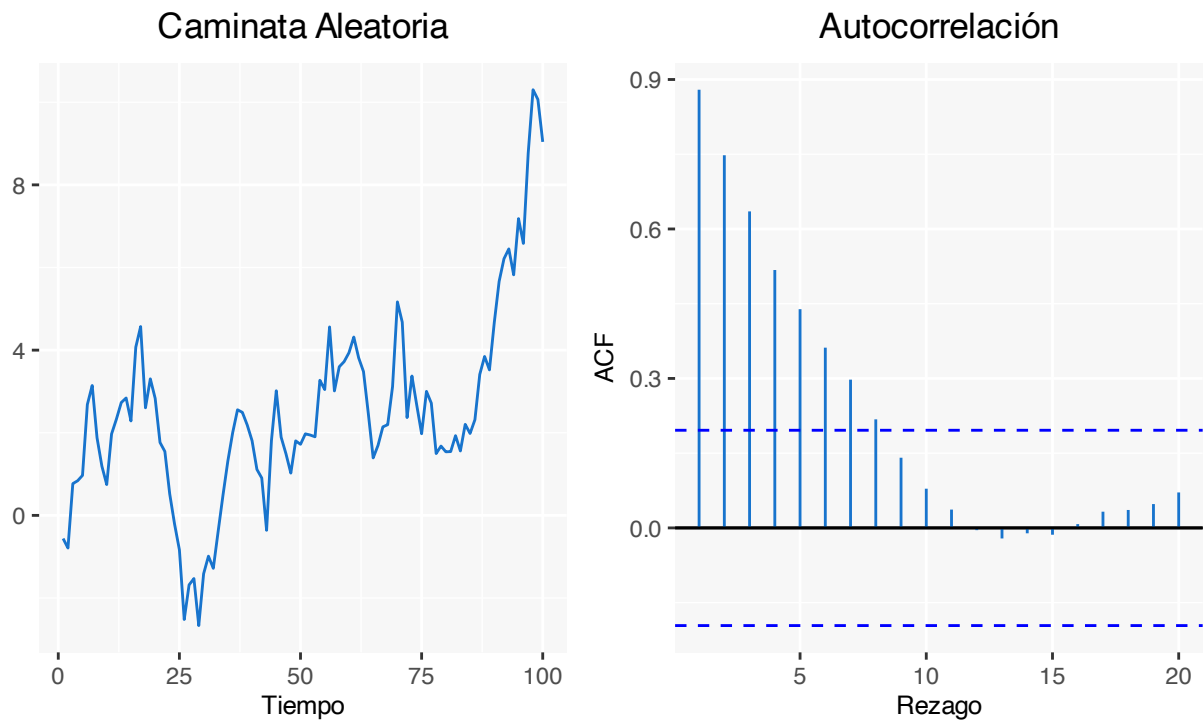


La segunda gráfica lleva el nombre de **correlograma**, la cual esta creada con los valores de autocorrelación, los cuales se verán en el siguiente capítulo.

3.3 Caminata aleatoria

Es un proceso estocástico X_t donde la primera diferencia de este proceso es un ruido blanco: $\nabla X_t = \epsilon_t$

¹El nombre de ruido blanco puede explicarse fácilmente en el análisis espectral en series de tiempo, el cual hace un estudio sobre la frecuencia y no tanto sobre el tiempo. Este no se verá en este material pero se puede estudiar a detalle esto en el libro Cryer and Chan (2008). Bajo un análisis espectral, en un proceso de ruido, su función de densidad espectral (se puede pensar la densidad espectral como la cantidad de varianza en un intervalo $(\omega, \omega + d\omega)$ que aporta un modelo basado en una serie de Fourier) es constante, es decir que esto sucede para todas las frecuencias. Esto es análogo al espectro de luz blanca en el ámbito físico; ya que todos los colores se distribuyen uniformemente para formar la luz blanca.



La anterior gráfica esta basada en 100 simulaciones de una normal estándar bajo con una semilla en 123.

3.4 Series de tiempo

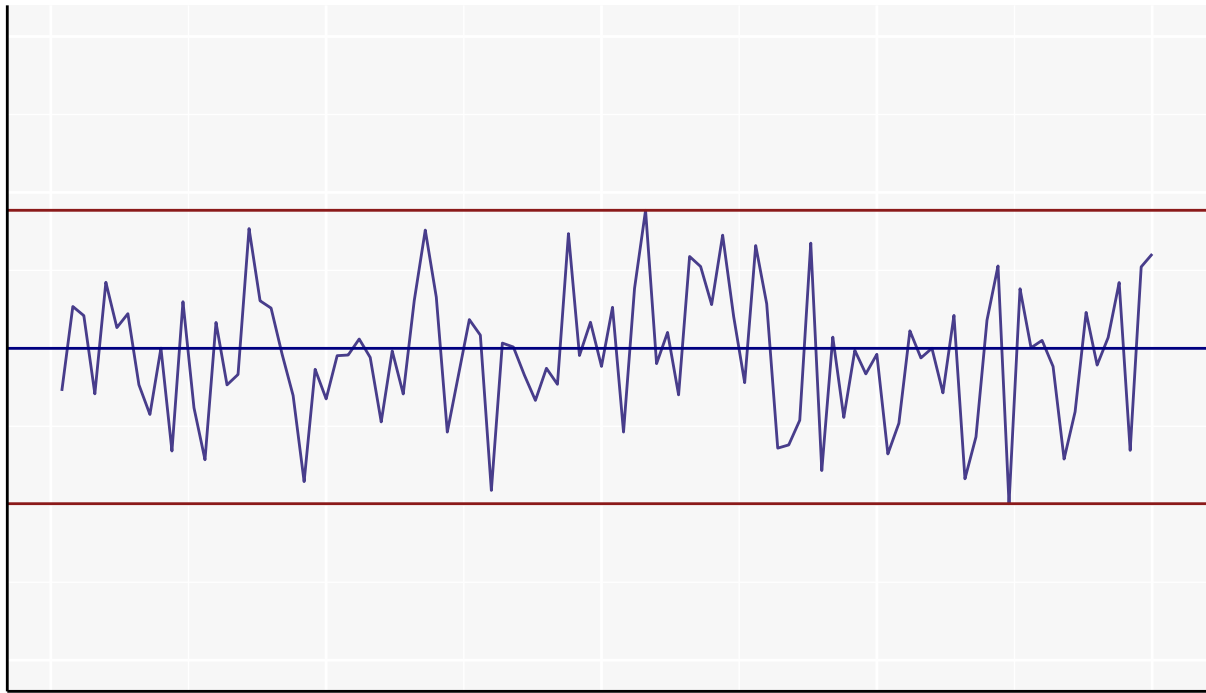
Según las características de las series de tiempo estas pueden clasificarse en:

Estacionarias: Una serie de tiempo estacionaria es estable a lo largo del tiempo, es decir su media, varianza y **autocovarianza** (en diferentes rezagos o diferentes tiempos) son constantes en el tiempo.

Es decir que sin importar el momento en que se midan (invariantes respecto al tiempo) se debe cumplir lo siguiente:

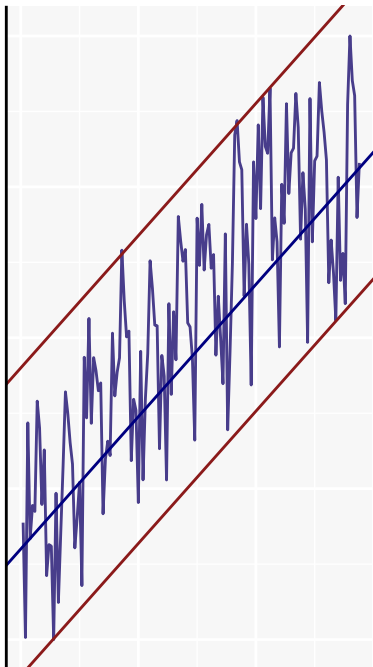
- Media: $\mathbb{E}(X_t) = \mathbb{E}(X_{t+k}) = \mu$
- Varianza: $Var(X_t) = Var(X_{t+k}) = \sigma^2$
- Covarianza: $\mathbb{E}[(X_t - \mu)(X_{t+k} - \mu)] = \gamma_k$

Aquí se muestra un ejemplo en el cual la varianza no cambia a lo largo del tiempo y la media permanece constante

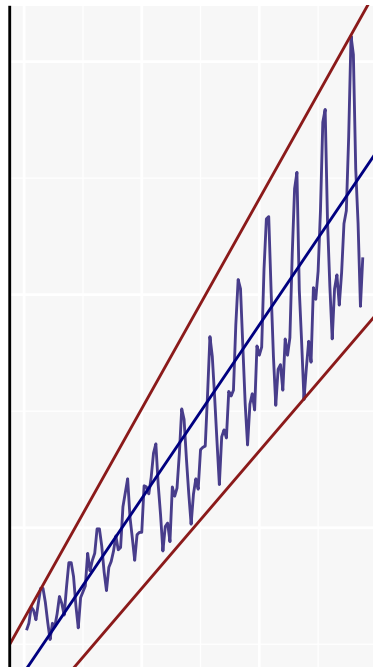


Varianza cte. tendencia cte.

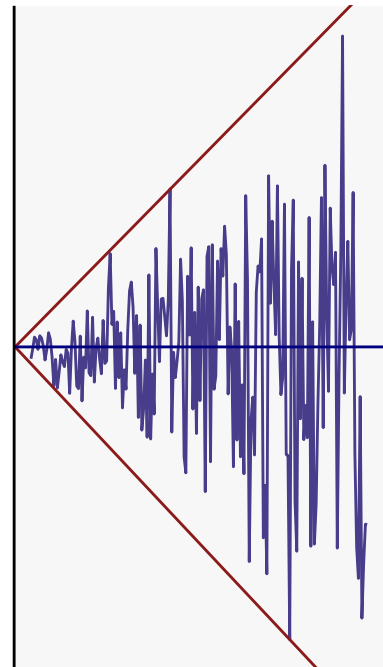
No estacionaria: Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante. Existen muchos casos en este tipo de series, aquí se muestran algunos ejemplos.



Varianza cte. tendencia no cte.



Varianza no cte. tendencia no cte.



Varianza no cte. tendencia cte.

En resumen, si una serie es estacionaria, su media, su varianza y su autocovarianza (en diferentes rezagos) permanecen iguales sin importar el momento en el cual se midan; es decir, son invariantes respecto al tiempo.

3.5 Ejercicios

1. Reproducir la caminata aleatoria con semilla (123) mostrada en la gráfica anterior. Ver ¿qué sucede al aumentar el número de simulaciones ($n = 1000$)? ¿Por qué sucede este nuevo comportamiento?
2. Un proceso es estacionario débil en orden 2 si tiene media y varianzas constantes (Seleccione la opción correcta)
 - a) Cierto
 - b) Falso
3. Utilizar los datos en el objeto “AirPassengers” que corresponden a 144 mediciones del número de pasajeros mensuales en líneas aéreas y grafique la serie de la diferencia de los logaritmos de los datos. ¿El gráfico sugiere que un modelo estacionario podría ser apropiada para las diferencias de los logaritmos?

Chapter 4

Funciones de autocovarianza y autocorrelación

Como se mencionó en el apartado anterior, tenemos que para un proceso estacionario X_t , su esperanza es $E[X_t] = \mu$ y su varianza $Var[X_t] = E[X_t - \mu]^2 = \sigma^2$ las cuales son constantes; así como las covarianzas $Cov(X_t, X_s)$, que son funciones que solamente dependen del tiempo que transcurre entre t y s ; entonces en este caso podemos escribir la **covarianza** entre X_t y X_{t+k} como:

$$\gamma_k = Cov(X_t, X_{t+k}) = E(X_t - \mu)(X_{t+k} - \mu)$$

Y su **correlación** como:

$$\rho_k = \frac{Cov(X_t, X_{t+k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+k})}} = \frac{\gamma_k}{\gamma_0}$$

donde, $Var(X_t) = Var(X_{t+k}) = \gamma_0$.

También conocida como la **función de autocorrelación (ACF)**, la cual es una medida de la relación para los valores de la serie respecto a los valores de esta misma, observados $k - t$ unidades de tiempo.

La función de autocorrelación tiene las siguientes propiedades:

- $\rho_0 = 1$
- $-1 \leq \rho_j \leq 1$
- $\rho_j = \rho_{-j}$

En general, se utiliza la función de autocorrelación muestral r_k , la cual queda determinada por la siguiente expresión

$$r_k = \frac{\sum_{t=k+1}^n (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}$$

Para $k = 1, 2, \dots$ y donde \bar{X} es la media de las observaciones. Se aprecia que los elementos de la suma en el denominador corresponden a las covarianzas entre todas las observaciones que tienen un rezago de k unidades.

4.1 Función de autocorrelación parcial

La función de correlación parcial estima la correlación entre una observación k tiempos después de la observación actual removiendo los efectos de las correlaciones de las observaciones intermedias; es decir:

$$\pi_j = \text{corr}(X_j, X_{j-k} | X_{j-1}, X_{j-2}, \dots, X_{j-k-1})$$

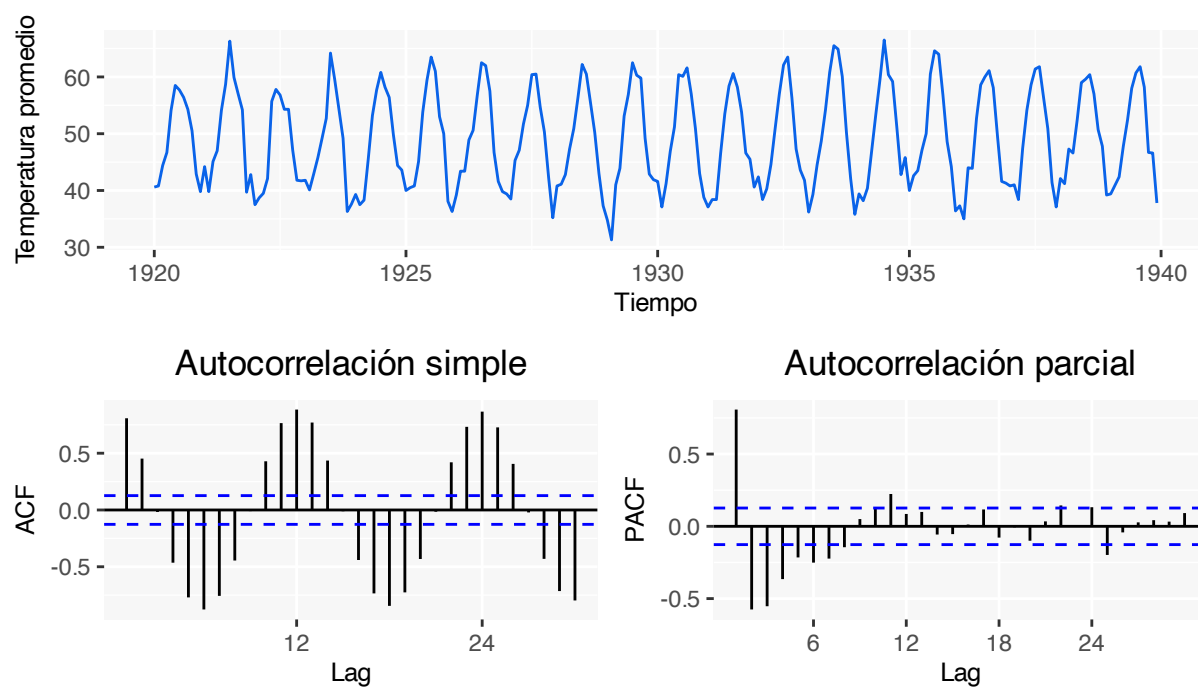
Para este caso, el estimador apropiado es la función de autocorrelación parcial muestral Cryer and Chan (2008), el cual puede obtenerse utilizando r_k para la siguiente expresión recursiva.

$$\pi_j = \frac{\rho_j - \sum_{k=1}^{j-1} \pi_{j-1,k} \rho_{j-k}}{1 - \sum_{k=1}^{j-1} \pi_{j-1,k} \rho_k}$$

donde $\pi_{j,k} = \pi_{j-1,k} - \pi_j \pi_{j-1,j-k}$ para $k = 1, 2, \dots, j-1$.

4.2 Correlograma

Una vez calculadas las funciones de autocorrelación y autocorrelación parcial, se grafican contra los diferentes rezagos para obtener los correlogramas. Estos nos serán de gran utilidad para la identificación del modelo, tal como se menciona en el capítulo 12 y en el complemento a este material SeriesTCode. Como ejemplo, se presentan los gráficos correspondientes a los datos de `nottem`.



4.3 Prueba de Ljung-Box

Esta prueba permite probar en forma conjunta que todos los coeficientes de autocorrelación son simultáneamente iguales a cero. La prueba está definida como

$$LB = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k} \sim \chi_{(m)}^2$$

donde n es el tamaño de la muestra, m la longitud del rezago.

Las hipótesis de esta prueba son:

H_0 : Los coeficientes de autocorrelación son simultáneamente iguales a cero

H_1 : Alguno de los coeficientes de autocorrelación es distinto de cero

Entonces si LB excede el valor crítico de la tabla Ji cuadrada al nivel de significancia seleccionado, no se acepta la hipótesis de que todos los coeficientes de autocorrelación son iguales a cero, por lo tanto al menos algunos de ellos deben ser diferentes de cero¹.

En la sección Autocorrelación del blog puedes encontrar gráficas interesantes sobre este tema y aplicaciones directas en R.

4.4 Ejercicios

1. ¿Qué mide la auto-correlación? (Seleccione la opción correcta)
 - (a) Dependencia Lineal entre múltiples puntos de series de tiempo diferentes observadas a tiempos distintos.
 - (b) Dependencia lineal entre dos puntos de la misma serie observada a tiempos distintos.
 - (c) Dependencia cuadrática entre dos puntos de una misma serie observada a tiempo distintos.
 - (d) Dependencia lineal entre dos puntos de series distintas observadas al mismo tiempo.
2. Suponga que el correlograma de una serie de tiempo con 100 observaciones tiene: $\hat{\rho}_1 = 0.31$, $\hat{\rho}_2 = 0.07$, $\hat{\rho}_3 = -0.05$, $\hat{\rho}_4 = 0.06$, $\hat{\rho}_5 = -0.03$, $\hat{\rho}_6 = 0.27$, $\hat{\rho}_7 = 0.08$, $\hat{\rho}_8 = 0.05$, $\hat{\rho}_9 = 0.02$, $\hat{\rho}_{10} = -0.01$. ¿Dada esta información que se debería tener en cuenta al proponer un modelo de predicción para esta serie?
3. El precio diario de oro durante 252 días de trading en 2005 están en el objeto gold del paquete TSA.
 - (a) Grafique la serie de tiempo. ¿Que patrón se observa?
 - (b) Grafique la serie de las diferencias de los logaritmos de los datos. ¿El gráfico sugiere que un modelo estacionario podría ser apropiado para las diferencias de los logaritmos? Explique brevemente.
 - (c) Utilice la función ACF para las diferencias de los logaritmos de los datos. ¿Es evidencia suficiente de que los log-precios del oro siguen un modelo de caminata aleatoria?, ¿por qué?

¹Originalmente la prueba esta diseñada para comprobar estadísticamente si las observaciones son independientes entre sí, bajo esta hipótesis, los coeficientes de correlación serán cero y el estadístico LB será cercano a cero. Esta prueba es de gran ayuda para ver la calidad del modelo que se aplique y para validar las hipótesis que se puedan requerir en la construcción de intervalos de predicción. Se hablará más de su uso en el capítulo 12.

Chapter 5

Transformaciones

La mayoría de las series de tiempo no son débilmente estacionarias, es decir que no muestran una media ni una varianza constantes a lo largo del tiempo y pueden mostrar tendencias crecientes o bien, decrecientes. Para poder trabajar con este tipo de series de una forma más sencilla y manejable, existen métodos para transformarlas y verlas como realizaciones de una serie de tiempo débilmente estacionaria. Las transformaciones más comunes son el suavizamiento a través de medias móviles o ajustando polinomios a la serie en cuestión, la diferenciación.

5.1 Suavizamiento por medias móviles

Se mencionó que las series de tiempo se pueden ver como la suma de tres componentes: una tendencia, una estacionalidad y una componente aleatorio o irregular. Ahora bien, en este método de suavizamiento el objetivo es estimar y extraer **la tendencia** (T_t) del modelo. Lo anterior se puede realizar estimando la tendencia con:

$$\hat{T}_t = (2q + 1)^{-1} \sum_{j=-q}^q X_{t-j}, q + 1 \leq t \leq n - q$$

Este es uno de muchos filtros lineales que podrían aplicarse ($\hat{T}_t = \sum_{j=-\infty}^{\infty} a_j X_{t-j}$). Obsérvese que para valores grandes de q , $(2q + 1)^{-1} \sum_{j=-q}^q [X_{t-j} - \hat{T}_t] \approx 0$, lo cual no sólo atenúa el ruido, también permite que la tendencia lineal $T_t = c_0 + c_1 t$ pase sin distorsión. Sin embargo hay que tener cuidado en la selección de q , ya que valores muy grandes, si T_t no es lineal, entonces se suavizará la serie pero la estimación de la tendencia será mala.

5.2 Suavizamiento por polinomios ajustados

En este método de suavizamiento el objetivo es estimar y extraer la **tendencia** (T_t) y la **estacionalidad** (E_t) del modelo. Si $E_t = 0$, se tiene un caso de no estacionariedad simple, por lo que el proceso tiene un comportamiento estacionario alrededor de la tendencia y para estimar T_t se supone que tiene la siguiente forma:

$$T_t = a_0 + a_1 t + \dots + a_p t^p$$

Si se tiene $p = 1$ la tendencia es lineal, si $p = 2$ la tendencia es cuadrática.

Los parámetros a_i se estiman mediante mínimos cuadrados ordinarios, es decir que minimicen $\sum_{i=1}^n (x_i - T_i)^2$.

5.3 Diferencias de Box-Jenkins

Consiste en aplicar diferencias a la serie de tiempo estudiada hasta que las observaciones se perciban como componentes de una serie débilmente estacionaria. Se definen los siguientes operadores para este método:

El **operador de retraso** se denota con una letra B o \mathcal{L} y se define como el valor retrasado de una serie de tiempo temporal, indicado con el exponente del operador la cantidad de retrasos. De manera particular se tiene $BX_t = X_{t-1}$, por lo que si se aplica varias veces el operador, la serie se retardaría k -unidades temporales:

$$B^k X_t = X_{t-k}$$

En particular $B^0 X_t = X_t$. Por lo que se podría utilizar notación polinómica para expresar el operador de retraso:

$$\phi(B)X_t = (\phi_0 + \phi_1 B + \phi_2 B^2) X_t = \phi_0 X_t + \phi_1 X_{t-1} + \phi_2 X_{t-2}$$

El **operador diferencia** se expresa con ∇ y se define como la diferencia entre el valor al periodo t y valor rezagado k periodos. Para $k = 1$ se define como $\nabla X_t = X_t - X_{t-1}$. Para $k = 2$ se tendría la diferencia de las diferencias de un rezago.

$$\nabla^2 X_t = \nabla X_t - \nabla X_{t-1} = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2})$$

Y así sucesivamente según vaya aumentando el valor de k .

Ambos operadores se relacionan de la siguiente manera:

$$\nabla Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t$$

Teniendo que $\nabla = (1 - B)$

Si aplicamos el operador diferencia sucesivamente entonces se obtiene:

$$\nabla^k = (1 - B)^k; \quad \nabla^k X_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j X_{t-j}$$

Ejemplos del uso de estas y otras transformaciones se puede ver en el la sección Transformaciones del blog, aunque se recomienda su lectura después de estudiar los modelos ARIMA. En el capítulo 9 se hará mención de este contenido nuevamente.

5.4 Ejercicios

1. ¿Cuáles son las razones por las que se sugiere aplicar transformaciones como suavizamiento a través de medias móviles o diferenciación a las series de tiempo?
2. Realiza los siguientes incisos:
 - a) En el siguiente enlace podrás obtener datos con el número de nacimientos por mes en la ciudad de Nueva York, desde enero de 1946 hasta diciembre de 1959, así como el código para la creación de un objeto de serie de tiempo en R adecuado para estos datos.
 - b) Compara la serie de tiempo de los nacimientos, gráficamente, con las diferencias de los valores de la serie en 1 rezago. Para esto se sugiere utilizar la función `diff()`. ¿Qué puedes decir de las series de tiempo? ¿Cuál de las series tiene un comportamiento más estacionario? ¿Qué harías para que la serie original aumentará su estacionariedad? ¿Qué pasa si aumentamos el número de rezagos en las diferencias?

- c) Utiliza la función `decompose()` en ambas series de tiempo y anota tus conclusiones (en el enlace de descarga de los datos vienen algunas conclusiones técnicas sobre la serie original, puedes tomarlas como referencia y dar otros puntos relevantes. Trata de dar conclusiones técnicas y no técnicas). ¿Existen periodos interesantes en los que se dan ciertos patrones?
 - d) Muestra visualmente la autocorrelación y la autocorrelación parcial de ambas series de tiempo y determina otorga comentarios generales. ¿Las funciones de autocorrelación muestran que al utilizar diferencias obtenemos un comportamiento estacionario? ¿Tenemos en algún caso un comportamiento similar a una caminata aleatoria o al ruido blanco? ¿Existe alguna diferencia relevante si aumentamos el número de rezagos en las diferencias?
 - e) Con la función `forecast::ma()` puedes suavizar la serie de tiempo con el método de medias móviles. Realiza el análisis que necesites y concluye si tiene un gran impacto a favor el suavizar la serie de tiempo original. ¿Es mejor suavizar que al utilizar la diferencia entre las observaciones?
- Recomendaciones: Utiliza la función `autoplot()` del paquete `forecast` en objetos de series de tiempo y el resultado de la función `decompose()` como si fuera la función `plot()`. Obtendrás un objeto `ggplot` que tendrá mayor versatilidad para su edición. Además, te recomendamos utilizar las funciones `ggAcf()` y `ggPacf()` del mismo paquete para el análisis de autocorrelación.
3. Con los datos seleccionados para las prácticas
- c. Aplique la función `decompose` a los datos y describa los hallazgos. ¿Qué puede concluir además de los hallazgos técnicos?
 - d. Grafique la serie de los logaritmos de los datos y describa el cambio observado respecto a lo visto en el inciso b) de la práctica 1 ¿El gráfico sugiere que un modelo estacionario podría ser apropiado para los logaritmos?

Part II

Procesos lineales estacionarios

Chapter 6

$AR(p)$: Proceso Autoregresivo

Los modelos autoregresivos se basan en la idea de que el valor actual de la serie X_t , puede explicarse en función de p valores pasados $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, donde p determina el número de rezagos necesarios para pronosticar un valor actual.

El modelo autoregresivo de orden p está dado por:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \epsilon_t$$

Re acomodando en ambos lados de la igualdad, obtenemos

$$X_t - \phi_0 - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \epsilon_t$$

Expresado en términos del operador de retardos para que quede todo en función de X_t :

$$-\phi_0 + (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \dots - \phi_p B^p) X_t = \epsilon_t \phi_p(B) X_t = \epsilon_t + \phi_0$$

donde ϵ_t es un proceso de ruido blanco y $\phi_0, \phi_1, \phi_2, \dots, \phi_p$ son parámetros del modelo.

En las siguientes secciones se verán los casos particulares de $p=1$ y $p=2$. En la práctica por interpretabilidad y reducción de los modelos se busca que los datos ajusten a modelos autoregresivos con valores pequeños del parámetro p .

6.1 Proceso Autoregresivo de orden 1: $AR(1)$

En los procesos $AR(1)$ la variable X_t está determinada únicamente por el valor pasado, esto es X_{t-1} .

$$X_t = \phi_0 + \phi_1 X_{t-1} + \epsilon_t$$

Donde ϵ_t es un ruido blanco con media 0 y con varianza σ^2 e independiente de X_t . Para verificar que sea estacionario (débilmente) se debe verificar la estacionariedad en media y covarianza.

6.1.1 Estacionario en Media

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(\phi_0 + \phi_1 X_{t-1} + \epsilon_t) = \phi_0 + \phi_1 \mathbb{E}(X_{t-1}) + \mathbb{E}(\epsilon_t) \\ &= \phi_0 + \phi_1 \mathbb{E}(X_{t-1}) \end{aligned}$$

Para que X_t sea estacionario en la media se debe cumplir que $E(X_t) = E(X_{t-1})$

Entonces:

$$(1 - \phi_1) \mathbb{E}(X_t) = \phi_0 \implies E(X_t) = \frac{\phi_0}{1 - \phi_1}$$

por lo tanto $E(X_t) = \frac{\phi_0}{1 - \phi_1}$ y $\phi_1 \neq 1$.

6.1.2 Estacionario en Covarianza

Para que $AR(1)$ sea estacionario, la varianza tiene que ser constante y finita en el tiempo.

$$\begin{aligned}
 \gamma_0 &= Var(X_t) = \mathbb{E}[(X_t - \mathbb{E}(X_t))^2] \\
 &= \mathbb{E}[(\phi_0 + \phi_1 X_{t-1} + \epsilon_t - \phi_0 - \phi_1 \mathbb{E}(X_{t-1}))^2] \\
 &= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \epsilon_t)^2] \\
 &= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})))^2 + 2\phi_1(X_{t-1} - \mathbb{E}(X_{t-1}))\epsilon_t + \epsilon_t^2] \\
 &= \phi_1^2 \mathbb{E}[(X_{t-1} - \mathbb{E}(X_{t-1}))^2] + 2\phi_1 \mathbb{E}[X_{t-1} - \mathbb{E}(X_{t-1})] \mathbb{E}(\epsilon_t) + \mathbb{E}[\epsilon_t^2] \\
 &= \phi_1^2 Var(X_{t-1}) + \sigma^2
 \end{aligned}$$

Buscamos que X_t sea estacionario en la varianza, por lo que bajo el supuesto de proceso estacionario:

$$\begin{aligned}
 &\implies Var(X_t) = Var(X_{t-1}) \\
 &\implies Var(X_t) = \gamma_0 = \phi_1^2 Var(X_{t-1}) + \sigma^2 \\
 &\implies \gamma_0 = \phi_1^2 \gamma_0 + \sigma^2 \implies (1 - \phi_1^2) \gamma_0 = \sigma^2 \\
 &\implies \gamma_0 = \frac{\sigma^2}{1 - \phi_1^2}
 \end{aligned}$$

Véase que para que sea estacionario, con varianza constante y finita es necesario que $|\phi_1| < 1$. En resumen

$$Var(X_t) = \gamma_0 = \frac{\sigma^2}{1 - \phi_1^2}$$

Respecto a la covarianza $Cov(X_t, X_{t-k})$ para $k = 1, \dots$, se tiene lo siguiente

$$\begin{aligned}
 Cov(X_t, X_{t-k}) &= \gamma_k = \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
 &= \mathbb{E}[(\phi_0 + \phi_1 X_{t-1} + \epsilon_t - \mathbb{E}(X_t))(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
 &= \mathbb{E}[(\phi_0 + \phi_1 X_{t-1} + \epsilon_t - \phi_0 - \phi_1 \mathbb{E}(X_{t-1}))(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
 &= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \epsilon_t)(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
 &= \mathbb{E}[\phi_1(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-k} - \mathbb{E}(X_{t-k}))] + \mathbb{E}[\epsilon_t(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
 &= \phi_1 \mathbb{E}[(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-k} - \mathbb{E}(X_{t-k}))] + \mathbb{E}[X_{t-k} - \mathbb{E}(X_{t-k})] \mathbb{E}[\epsilon_t] \\
 &= \phi_1 \gamma_{k-1} + \gamma_0 0 = \phi_1 \gamma_{k-1}
 \end{aligned}$$

$$\begin{aligned}
 \gamma_1 &= \gamma_0 \phi_1 \\
 \text{Entonces: } \gamma_2 &= \gamma_1 \phi_1 \\
 &\vdots
 \end{aligned}$$

$\therefore AR(1)$ es estacionario en covarianza si $|\phi_1| < 1$ y la **función de covarianza** será:

$$\gamma_k = \begin{cases} \frac{\sigma^2}{1 - \phi_1^2} & k = 0 \\ \phi_1 \gamma_{k-1} & k > 0 \end{cases}$$

Los **coeficientes de correlación** quedan determinados por la siguiente expresión:

$$\rho_k = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)} \sqrt{Var(X_{t-k})}} = \frac{\gamma_k}{\gamma_0} = \frac{\phi_1 \gamma_{k-1}}{\gamma_0} = \phi_1 \frac{\gamma_{k-1}}{\gamma_0} = \phi_1 \rho_{k-1}$$

Por lo que la **función de autocorrelación** para $AR(1)$ es :

$$\rho_k = \begin{cases} 1 & k = 0 \\ \phi \rho_{k-1} & k \geq 1 \end{cases} = \begin{cases} 1 & k = 0 \\ \phi_1^{\frac{\gamma_k - 1}{\gamma_0}} & k \geq 1 \end{cases}$$

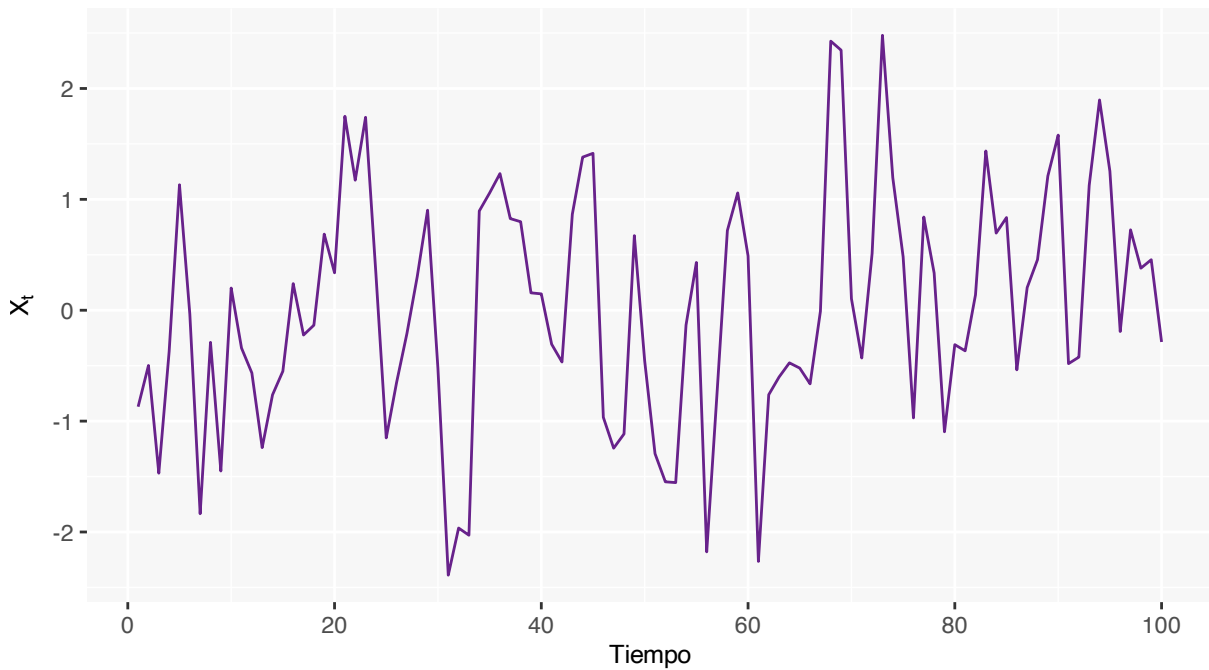
Observemos que para el modelo $AR(1)$, la función de autocorrelación es exponencial

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= \phi_1 \rho_0 = \phi_1 \\ \rho_2 &= \phi_1 \rho_1 = \phi_1^2 \implies \rho_k = \begin{cases} 1 & k = 0 \\ \phi_1^k & k \geq 1 \end{cases} \\ &\vdots \\ \rho_k &= \phi_1^k \end{aligned}$$

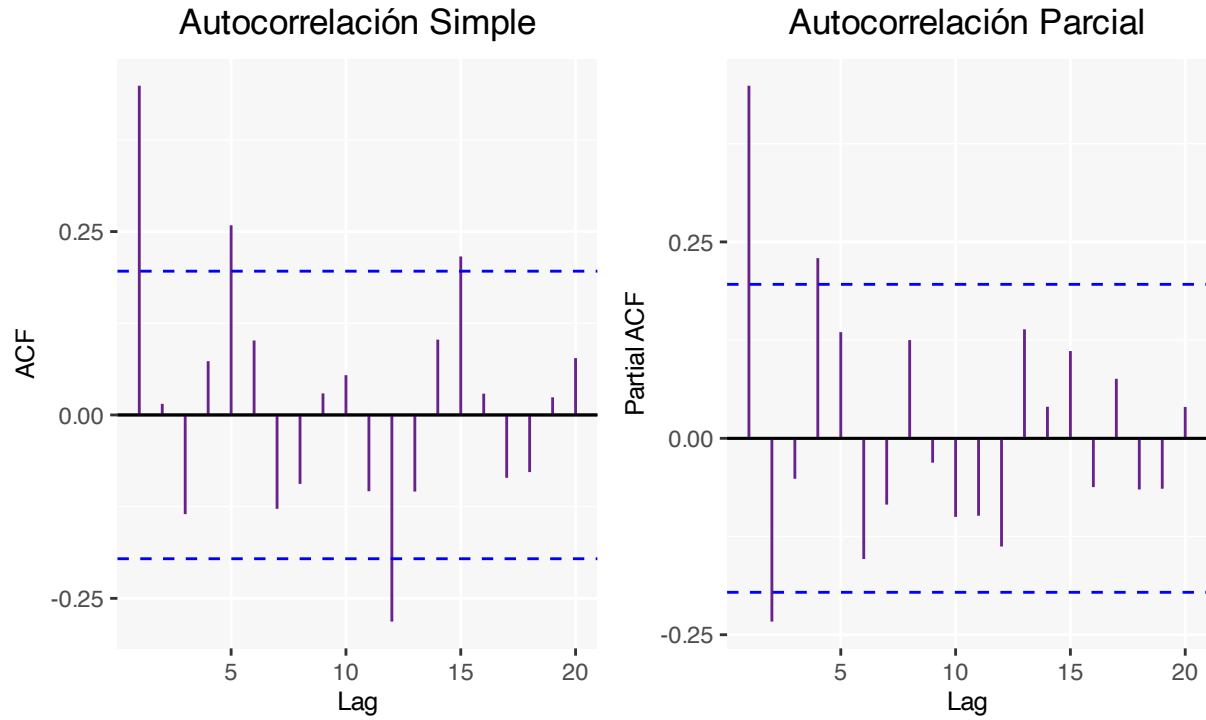
- Caso particular: $AR(1) : X_t = \phi_1 X_{t-1} + \epsilon_t$, es decir que $\phi_0 = 0$. Para este caso se tiene $\mathbb{E}(X_t) = \phi_1 \mathbb{E}(X_{t-1}) \implies (1 - \phi_1) \mathbb{E}(X_t) = 0 \implies \mathbb{E}(X_t) = \frac{0}{(1 - \phi_1)} = 0$.

A continuación se muestran los resultados para un modelo $AR(1)$ de la siguiente forma $X_t = 0.35X_{t-1} + \epsilon_t$

Ejemplo de modelo $AR(1)$



Además de las gráficas de Autocorrelación simple y parcial.



6.2 Proceso Autoregresivo de orden 2: $AR(2)$

En los procesos $AR(2)$ la variable X_t está determinada por el valor pasado y el anterior a este.

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

Donde ϵ_t es un ruido blanco (media cero y varianza σ^2). Asumiendo estacionariedad débil, se tiene que la media y la varianza del proceso serán las siguientes.

6.2.1 Estacionario en Media

Bajo el supuesto de estacionariedad:

$$\begin{aligned} \mathbb{E}(X_t) &= \mathbb{E}(X_{t-1}) = \mathbb{E}(X_{t-2}) \\ \implies (1 - \phi_1 - \phi_2)\mathbb{E}(X_t) &= \phi_0 \\ \implies \mathbb{E}(X_t) &= \frac{\phi_0}{1 - \phi_1 - \phi_2} \end{aligned}$$

\therefore Para que sea estacionario, se tiene que cumplir que $1 - \phi_1 - \phi_2 \neq 0$.

6.2.2 Estacionario en Covarianza

$$\begin{aligned}
\gamma_0 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))^2] \\
&= \mathbb{E}[(\phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t - \phi_0 - \phi_1 \mathbb{E}(X_{t-1}) - \phi_2 \mathbb{E}(X_{t-2}))^2] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2})) + \epsilon_t)^2] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2})))^2 \\
&\quad + 2\epsilon_t[\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2}))] + \epsilon_t^2] \\
&= \mathbb{E}[\phi_1^2(X_{t-1} - \mathbb{E}(X_{t-1}))^2 + 2\phi_1\phi_2(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-2} - \mathbb{E}(X_{t-2})) \\
&\quad + \phi_2^2(X_{t-2} - \mathbb{E}(X_{t-2}))^2 + 2\epsilon_t[\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2}))] + \epsilon_t^2] \\
&= \mathbb{E}[\phi_1^2(X_{t-1} - \mathbb{E}(X_{t-1}))^2] + 2\phi_1\phi_2\mathbb{E}[(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-2} - \mathbb{E}(X_{t-2}))] \\
&\quad + \phi_2^2\mathbb{E}[(X_{t-2} - \mathbb{E}(X_{t-2}))^2] + 2\mathbb{E}[\epsilon_t]\mathbb{E}[\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2}))] + \mathbb{E}[\epsilon_t^2] \\
&= \phi_1^2\gamma_0 + 2\phi_1\phi_2\gamma_1 + \phi_2^2\gamma_0 + \sigma^2
\end{aligned}$$

Pero, véase lo siguiente

$$\begin{aligned}
\gamma_1 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-1} - \mathbb{E}(X_{t-1}))] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2})) + \epsilon_t)(X_{t-1} - \mathbb{E}(X_{t-1}))] \\
&= \mathbb{E}[\phi_1(X_{t-1} - \mathbb{E}(X_{t-1}))^2 + \phi_2(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-2} - \mathbb{E}(X_{t-2})) + \epsilon_t(X_{t-1} - \mathbb{E}(X_{t-1}))] \\
&= \phi_1\gamma_0 + \phi_2\gamma_1 \\
\Rightarrow \gamma_1 &= \frac{\phi_1}{(1 - \phi_2)}\gamma_0 = \phi_1\gamma_0 + \phi_2\gamma_1
\end{aligned}$$

En general la autocovarianza de orden k , para $k > 1$ será:

$$\begin{aligned}
\gamma_k &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \phi_2(X_{t-2} - \mathbb{E}(X_{t-2})) + \epsilon_t)(X_{t-k} - \mathbb{E}(X_{t-k}))] \\
&= \mathbb{E}\{\phi_1(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-k} - \mathbb{E}(X_{t-k}))\} + \mathbb{E}\{\phi_2(X_{t-2} - \mathbb{E}(X_{t-2}))(X_{t-k} - \mathbb{E}(X_{t-k}))\} \\
&\quad + \mathbb{E}\{\epsilon_t(X_{t-k} - \mathbb{E}(X_{t-k}))\} \\
&= \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}
\end{aligned}$$

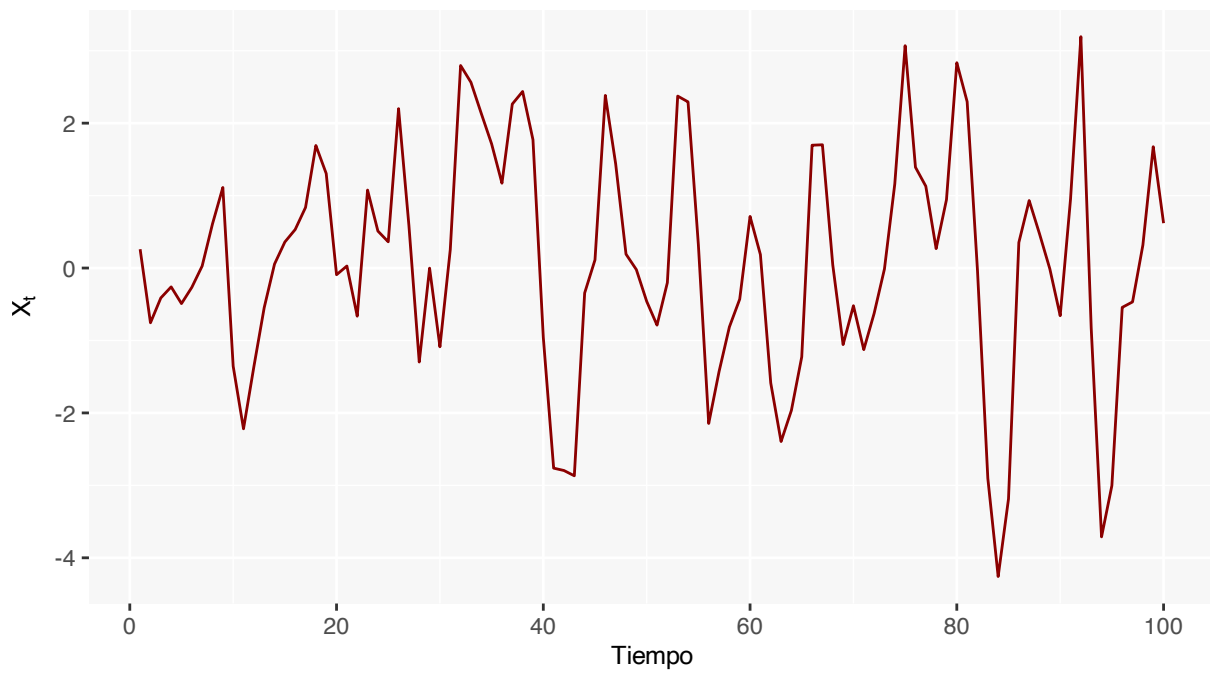
Entonces la **función de autocovarianza** de un modelo $AR(2)$ es la siguiente

$$\gamma_k = \begin{cases} \gamma_0 & \text{si } k = 0 \\ \gamma_1 & \text{si } k = 1 \\ \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} & \text{si } k > 1 \end{cases}$$

Y la correspondiente **función de autocorrelación** de un modelo $AR(2)$ es:

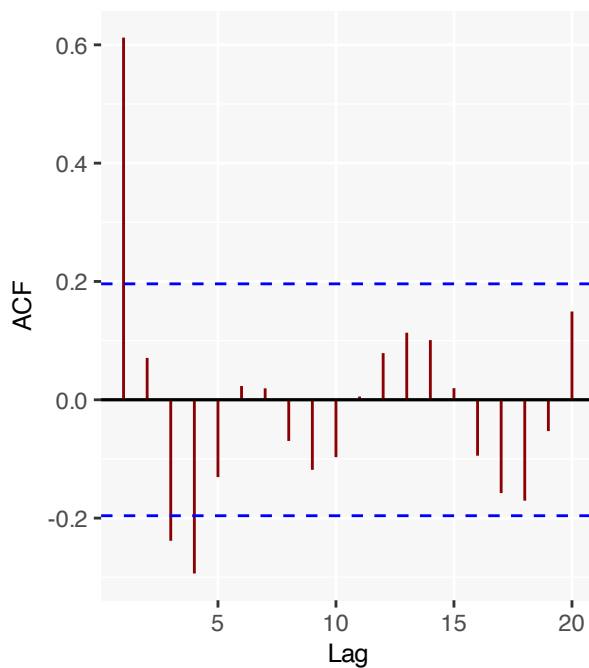
$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ \frac{\gamma_1}{\gamma_0} = \frac{\phi_1}{1-\phi_2} & \text{si } k = 1 \\ \phi_1\rho_{k-1} + \phi_2\rho_{k-2} & \text{si } k > 1 \end{cases}$$

A continuación se da un ejemplo de un modelo $AR(2)$ de la siguiente forma $X_t = 0.8X_{t-1} - 0.4X_{t-2} + \epsilon_t$

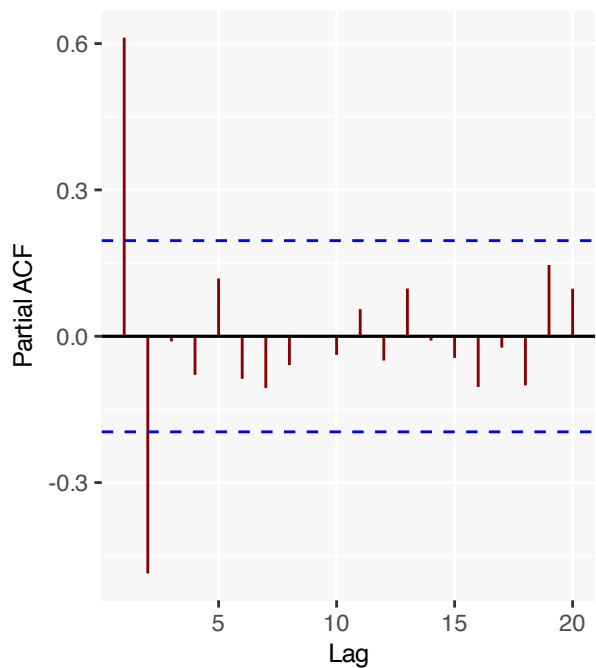
Ejemplo de modelo $AR(2)$ 

Y al igual que para el ejemplo del modelo $AR(1)$, se agregan las gráficas de autocorrelación simple y parcial.

Autocorrelación Simple



Autocorrelación Parcial



Algo interesante de tener los procesos en términos de operadores de retardos es que podemos obtener las condiciones de estacionariedad en AR por las raíces del polinomio que se deriva de esta notación:

$$\begin{aligned}
 AR(1) : X_t &= \phi_1 X_{t-1} + \epsilon_t \\
 \Rightarrow (1 - \phi_1 B) X_t &= \epsilon_t \\
 \Rightarrow B &= \frac{1}{\phi_1} \Rightarrow |\phi_1| < 1
 \end{aligned}$$

$$\begin{aligned} AR(2) : X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t \\ \implies (1 - \phi_1 B - \phi_2 B^2) X_t &= \epsilon_t \end{aligned}$$

Por lo que las raíces del polinomio $(1 - \phi_1 B - \phi_2 B^2)$ serán $B^1, B^2 = \frac{\theta_1 \pm \sqrt{\theta_1^2 + 4\theta_2}}{-2\theta_2}$.

Y gracias a todo esto, podemos interpretar que para que los procesos sean estacionarios, se solicita que las raíces estén fuera del círculo unitario, es decir: $|B^1| > 1$ y $|B^2| > 1$.

En la sección Identificación y modelación lineal se muestra el uso de R para la creación de modelos de lineales en series de tiempo y en particular en el apartado AR se estudia el comportamiento del proceso autoregresivo.

6.3 Ejercicios

1. Encuentre la función de covarianza γ_k para el modelo AR(1) en la sección 6.1, definido como $X_t = 0.35X_{t-1} + \epsilon_t$ donde $\epsilon_t \sim N(0, 1)$
2. En el modelo $Y_t = 0.3 - 0.5Y_{t-1} + 0.2\epsilon_t$ con $\epsilon_t \sim N(0, 1)$ la $COV(Y_t, Y_{t-1})$ es cero, es decir $\gamma_1 = 0$ (el valor de la variable al tiempo t no está afectada por el valor que tomó la variable 2 tiempos atrás)
3. Con los datos seleccionados para las prácticas
 - e. Grafique la serie de las diferencias de los logaritmos de los datos. ¿El gráfico sugiere que un modelo estacionario podría ser apropiado para las diferencias de los logaritmos? Explique brevemente.
 - f. Utilice la función ACF para las diferencias de los logaritmos de los datos para determinar si los datos transformados siguen un modelo de caminata aleatoria. Sí/No ¿por qué?

Chapter 7

$MA(q)$: Proceso de Medias Móviles

Estos modelos se puede decir son determinados por una fuente externa, además de que suponen linealidad. Es decir que el valor actual de la serie X_t esta influenciado por los valores de la fuente externa.

El modelo de medias móviles de orden q está dado por:

$$MA(q) : X_t - \mu = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}$$

Expresado en términos del operador de retardos

$$\begin{aligned} X_t &= \mu + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t \\ &= \mu + \theta_q(B) \epsilon_t \end{aligned}$$

donde ϵ_t es un proceso de ruido blanco, $\theta_1, \theta_2, \dots, \theta_q$ son los parámetros del modelo.

Antes de continuar, véase que es posible escribir un modelo AR con un MA . Supongamos un modelo $AR(2)$, en el cual se tiene lo siguiente

$$\begin{aligned} AR(2) : X_t &= \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t \\ \implies \phi(B) X_t &= (1 - \phi_1 B - \phi_2 B^2) X_t = \phi_0 + \epsilon_t \\ \therefore X_t &= \frac{\phi_0}{\phi(B)} + \frac{\epsilon_t}{\phi(B)} = \mu + (1 - \psi_1 B - \psi_2 B^2 - \dots) \epsilon_t \end{aligned}$$

Sin olvidar que las condiciones de estacionariedad para un proceso $AR(1)$ y $AR(2)$ son $|\phi_1| < 1$ y $|1 - \phi_1 - \phi_2| > 0$ respectivamente¹.

7.1 Proceso de Medias Móviles de orden 1: $MA(1)$

Este modelo determina el valor de X_t en función de la innovación actual y su primer retardo, es decir:

$$\begin{aligned} X_t &= \mu - \theta_1 \epsilon_{t-1} - \epsilon_t = \mu - \theta_1 \epsilon_{t-1} - \epsilon_t \\ \implies X_t &= \mu - (1 + \theta_1 B) \epsilon_t \end{aligned}$$

donde ϵ_t es un proceso de ruido blanco y θ es el parámetro. Además, véase que

$$\mathbb{E}(X_t) = \mu - \theta_1 \mathbb{E}(\epsilon_{t-1}) - \mathbb{E}(\epsilon_t) = \mu$$

¹Una pregunta interesante es: ¿Se puede escribir un proceso MA como un proceso AR ? La respuesta es sí. A esto se le conoce como **invertibilidad** y al igual que la estacionariedad, esta sucede bajo algunas condiciones. Para más información puede consultarse Cryer and Chan (2008).

Por lo que este proceso ya es estacionario en su media. Sólo resta estudiar el caso para la covarianza. Asumiendo estacionariedad débil, se tiene lo siguiente

$$\begin{aligned}
 \gamma_0 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))^2] = \mathbb{E}[(\mu - \theta_1 \epsilon_{t-1} - \epsilon_t - \mu)^2] \\
 &= \mathbb{E}[\theta_1^2 \epsilon_{t-1}^2 + 2\theta_1 \epsilon_{t-1} \epsilon_t + \epsilon_t^2] \\
 &= \mathbb{E}[\theta_1^2 \epsilon_{t-1}^2] + 2\theta_1 \mathbb{E}[\epsilon_{t-1} \epsilon_t] + \mathbb{E}[\epsilon_t^2] \\
 &= \theta_1^2 \sigma^2 + \sigma^2 \\
 &= (1 + \theta_1^2) \sigma^2
 \end{aligned}$$

$$\begin{aligned}
 \gamma_1 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-1} - \mathbb{E}(X_{t-1}))] \\
 &= \mathbb{E}[(-\epsilon_t - \theta_1 \epsilon_{t-1})(-\epsilon_{t-1} - \theta_1 \epsilon_{t-2})] \\
 &= \mathbb{E}[\epsilon_t \epsilon_{t-1} + \theta_1 \epsilon_t \epsilon_{t-2} + \theta_1 \epsilon_{t-1}^2 + \theta_1^2 \epsilon_{t-1} \epsilon_{t-2}] \\
 &= \theta_1 \mathbb{E}[\epsilon_{t-1}^2] \\
 &= \theta_1 \sigma^2
 \end{aligned}$$

$$\gamma_2 = \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-2} - \mathbb{E}(X_{t-2}))] = 0$$

\vdots

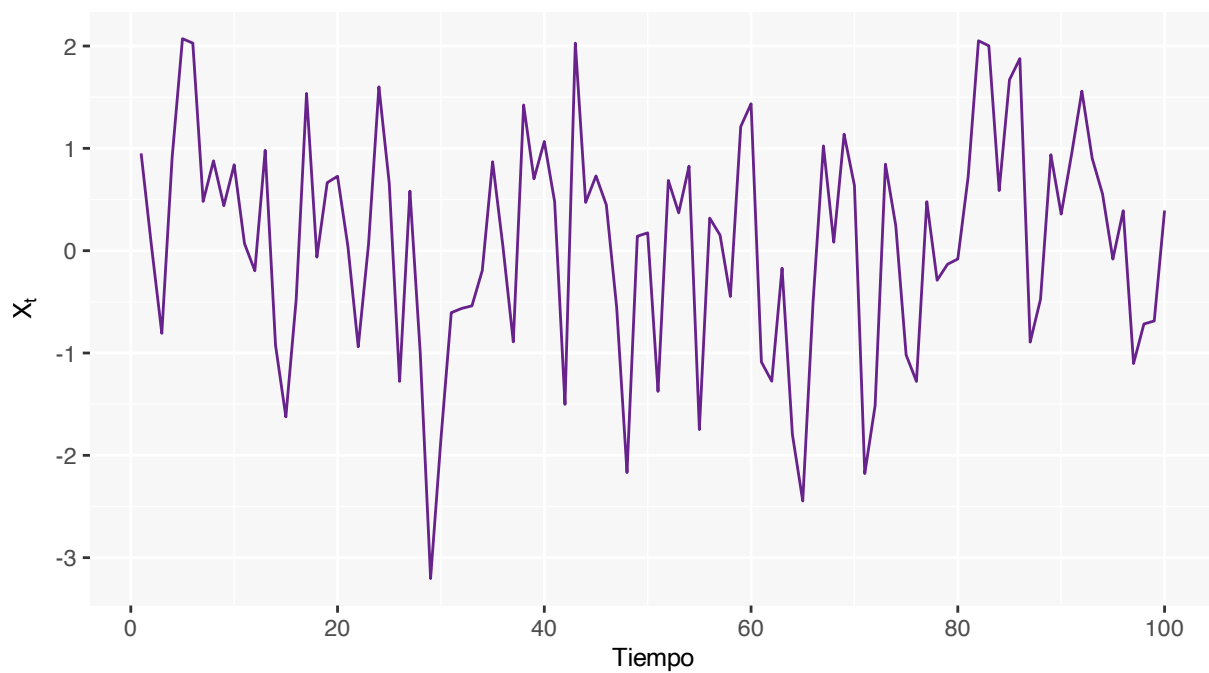
Entonces la **función de autocovarianza** de un modelo $MA(1)$ es:

$$\gamma_k = \begin{cases} (1 + \theta_1^2) \sigma^2 & \text{si } k = 0 \\ \theta_1 \sigma^2 & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases}$$

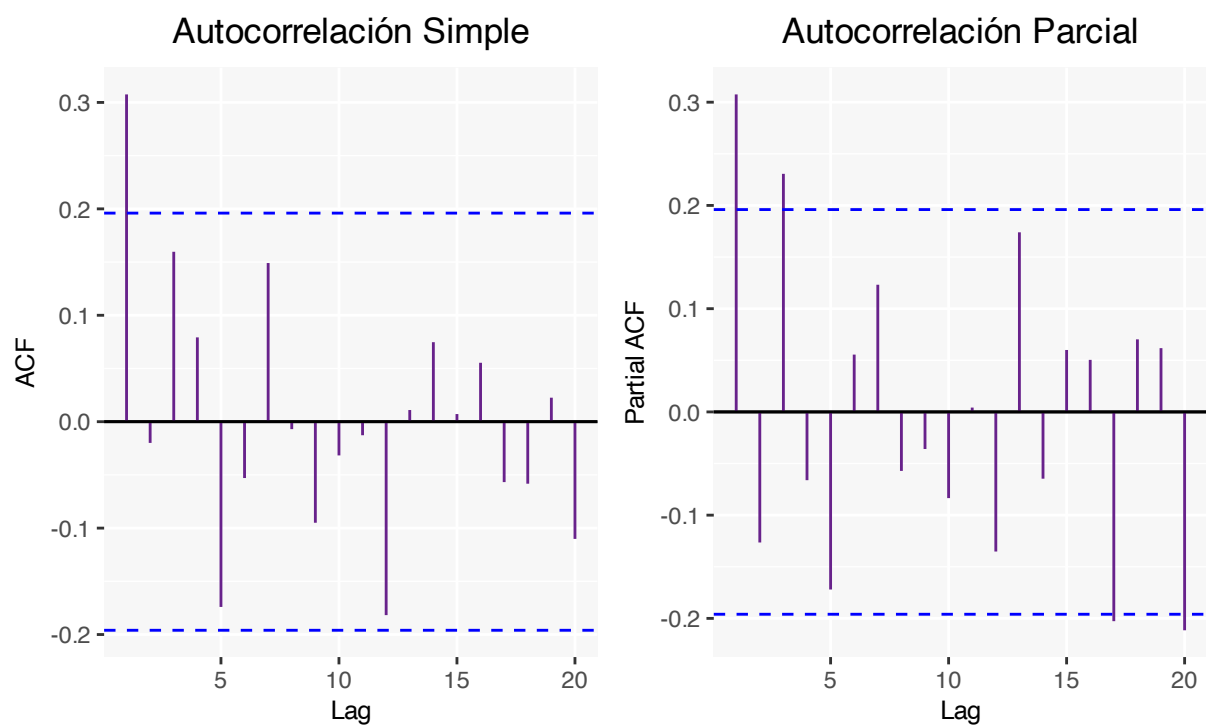
Y la correspondiente **función de autocorrelación** de un modelo $MA(1)$ es:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1 & \text{si } k = 0 \\ \frac{\theta_1}{1 + \theta_1^2} & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases}$$

A continuación se muestran los resultados para un modelo $AR(1)$ de la siguiente forma $X_t = 0.45\epsilon_{t-1}$

Ejemplo de modelo $MA(1)$ 

Además de las gráficas de Autocorrelación simple y parcial.

7.2 Proceso de Medias Móviles de orden 2: $MA(2)$

Este modelo esta determinado por:

$$MA(2) : X_t = \theta_0 - \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$$

$$X_t = \theta_0 - (1 + \theta_1 B + \theta_2 B^2) \epsilon_t$$

donde ϵ_t es un proceso de ruido blanco y θ_1, θ_2 son los parámetros del modelo. Asumiendo estacionariedad débil, es fácil comprobar que $\mathbb{E}(X_t) = \theta_0$. Respecto a la covarianza se tiene lo siguiente

$$\begin{aligned}\gamma_0 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))^2] = \mathbb{E}[(-\epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2})^2] \\ &= \mathbb{E}[\epsilon_t^2] + \theta_1^2\mathbb{E}[\epsilon_{t-1}^2] + \theta_2^2\mathbb{E}[\epsilon_{t-2}^2] \\ &= (1 + \theta_1^2 + \theta_2^2)\sigma^2 \\ \gamma_1 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-1} - \mathbb{E}(X_{t-1}))] \\ &= \mathbb{E}[(-\epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2})(-\epsilon_{t-1} - \theta_1\epsilon_{t-2} - \theta_2\epsilon_{t-3})] \\ &= \theta_1\mathbb{E}[\epsilon_{t-1}^2] + \theta_1\theta_2\mathbb{E}[\epsilon_{t-2}^2] \\ &= \theta_1(1 + \theta_2)\sigma^2 \\ \gamma_2 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-2} - \mathbb{E}(X_{t-2}))] \\ &= \mathbb{E}[(-\epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2})(-\epsilon_{t-2} - \theta_1\epsilon_{t-3} - \theta_2\epsilon_{t-4})] \\ &= \theta_2\mathbb{E}[\epsilon_{t-2}^2] \\ &= \theta_2\sigma^2\end{aligned}$$

Entonces la **función de autocovarianza** de un modelo $MA(2)$ es:

$$\gamma_k = \begin{cases} (1 + \theta_1^2 + \theta_2^2)\sigma^2 & \text{si } k = 0 \\ \theta_1(1 + \theta_2)\sigma^2 & \text{si } k = 1 \\ \theta_2\sigma^2 & \text{si } k = 2 \\ 0 & \text{si } k > 2 \end{cases}$$

Y la correspondiente **función de autocorrelación** de un modelo $MA(2)$ es:

$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ \frac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{si } k = 1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2} & \text{si } k = 2 \\ 0 & \text{si } k > 2 \end{cases}$$

La sección especializada a este tema en el blog es el apartado MA.

7.3 Ejercicios

- 1) Sea $Z_t \sim N(0, 1)$. Calcular la función de autocovarianza γ_k para el modelo $X_t = 0.3 + 0.1Z_t - 0.2Z_{t-1}$
- 2) Muestre que los procesos $MA(1)$ siguientes
 - a. $Z_t = a_t + \theta a_{t-1}$
 - b. $Y_t = a_t + \frac{1}{\theta}a_{t-1}$ donde $0 < |\theta| < 1$ y $a_t \sim N(0, 1)$, tienen las mismas funciones de autocorrelación

Chapter 8

$ARMA(p, q)$: Proceso Autoregresivo de Medias Móviles

Es muy probable que una serie de tiempo X_t , tenga características de un proceso AR y de un proceso MA al mismo tiempo, por lo que será un proceso $ARMA$. Así, X_t sigue un proceso $ARMA(p, q)$, y en este proceso habrá p términos autoregresivos y q términos de media móvil. Este se verá de la siguiente forma:

$$X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} + \epsilon_t$$

donde ϵ_t es un proceso de ruido blanco, y $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ son los parámetros del modelo.

Para un proceso $ARMA(p, q)$ una condición de estacionariedad es la misma que para un proceso $AR(p)$, y una condición de invertibilidad es la misma que para el proceso $MA(q)$.

El modelo $ARMA(p, q)$ se puede escribir en términos de los operadores de retardo de la siguiente manera.

Sea $c = 0$.

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t \\ \implies \phi_p(B) X_t &= \theta_q(B) \epsilon_t \end{aligned}$$

Donde:

- $\phi_p(B)$ es el polinomio autoregresivo
- $\theta_q(B)$ es el polinomio de medias móviles.

Hay que observar lo siguiente:

$$\begin{aligned} X_t &= \frac{\theta_q(B)}{\phi_p(B)} \epsilon_t &< \leftarrow MA(\infty) \\ \epsilon_t &= \frac{\phi_p(B)}{\theta_q(B)} X_t &< \leftarrow AR(\infty) \end{aligned}$$

- Los modelos $ARMA(p, q)$ siempre compartirán las características de los modelos $AR(p)$ y $MA(q)$, ya que contiene ambas estructuras a la vez.
- El modelo $ARMA(p, q)$ tiene media cero y varianza constante y finita además de que la función de autocorrelación es infinita y decrece rápidamente hacia cero.
- Un proceso $ARMA(p, q)$ es **estacionario** si y sólo si el modulo de las raíces del polinomio $\phi_p(B)$ está fuera del círculo unitario.
- Un proceso $ARMA(p, q)$ es **invertible** si y sólo si el modulo de las raíces del polinomio $\theta_q(B)$ está fuera del círculo unitario.

Ejemplo

Sea $Y_t : ARMA(2, 1)$ con $\epsilon_t \sim N(0, 1)$ tal que $Y_t = 1.5Y_{t-1} - 0.9Y_{t-2} - 0.4\epsilon_{t-1} + \epsilon_t$.

Sabemos que es invertible por que $|\theta| = |-0.4| < 1$ y es estacionario por lo siguiente:

$$\begin{aligned}\phi_2(B) &= (1 - 1.5B + 0.9B^2) \\ \implies B_1, B_2 &= \frac{1.5 \pm \sqrt{1.5^2 - 4(0.9)}}{2(0.9)} = 0.83 \pm 0.645 \\ \implies |B| &= \sqrt{0.83^2 + 0.65^2} = 1.05423 > 1\end{aligned}$$

8.1 ARMA(1, 1)

En un modelo $ARMA(1, 1)$ la serie de tiempo X_t se determina en función de su pasado hasta el primer retardo, la innovación actual y el pasado de la innovación hasta el primer retardo.

$$X_t = c + \phi_1 X_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

donde ϵ_t es un proceso de ruido blanco, y ϕ_1 y θ_1 son los parámetros del modelo. Ahora se verán las características de un proceso $ARMA(1, 1)$ estacionario.

8.1.1 Media

$$\mathbb{E}(X_t) = \mathbb{E}(c + \phi_1 X_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}) = c + \phi_1 \mathbb{E}(X_{t-1})$$

Por lo que Suponiendo estacionariedad $\mathbb{E}(X_t) = \frac{c}{1-\phi_1}$.

8.1.2 Covarianzas

Recordando que $\mathbb{E}(X_t) = c + \phi_1 \mathbb{E}(X_{t-1}) \implies X_t - \mathbb{E}(X_t) = \phi_1 (X_{t-1} - \mathbb{E}(X_{t-1})) + \epsilon_t + \theta_1 \epsilon_{t-1}$

$$\begin{aligned}\gamma_0 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))^2] \\ &= \mathbb{E}[(\phi_1 (X_{t-1} - \mathbb{E}(X_{t-1})) + \epsilon_t + \theta_1 \epsilon_{t-1})^2] \\ &= \mathbb{E}[(\phi_1 (X_{t-1} - \mathbb{E}(X_{t-1})))^2 + 2(\phi_1 (X_{t-1} - \mathbb{E}(X_{t-1}))) (\epsilon_t + \theta_1 \epsilon_{t-1}) + (\epsilon_t + \theta_1 \epsilon_{t-1})^2] \\ &= \phi_1^2 \mathbb{E}[(X_{t-1} - \mathbb{E}(X_{t-1}))^2] + 2\phi_1 \theta_1 \mathbb{E}[X_{t-1} \epsilon_{t-1}] + \mathbb{E}[(\epsilon_t + \theta_1 \epsilon_{t-1})^2] \\ &= \phi_1^2 \text{Var}(X_{t-1}) + 2\phi_1 \theta_1 \mathbb{E}[X_{t-1} \epsilon_{t-1}] + \mathbb{E}[\epsilon_t^2] + 2\theta_1 \mathbb{E}[\epsilon_t \epsilon_{t-1}] + \mathbb{E}[\theta_1^2 \epsilon_{t-1}^2] \\ &= \phi_1^2 \text{Var}(X_{t-1}) + 2\phi_1 \theta_1 \sigma^2 + \sigma^2 + \theta_1^2 \sigma^2 \\ \iff (1 - \phi_1^2) \text{Var}(X_t) &= \sigma^2 (1 + \theta_1^2 + 2\phi_1 \theta_1) \\ \implies \text{Var}(X_t) = \gamma_0 &= \frac{\sigma^2 (1 + \theta_1^2 + 2\phi_1 \theta_1)}{1 - \phi_1^2}\end{aligned}$$

$$\begin{aligned}\gamma_1 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-1} - \mathbb{E}(X_{t-1}))] \\ &= \mathbb{E}[(\phi_1 (X_{t-1} - \mathbb{E}(X_{t-1})) + \epsilon_t + \theta_1 \epsilon_{t-1})(X_{t-1} - \mathbb{E}(X_{t-1}))] \\ &= \mathbb{E}[\phi_1 (X_{t-1} - \mathbb{E}(X_{t-1}))^2] + \mathbb{E}[\epsilon_t X_{t-1}] - \mathbb{E}[\epsilon_t \mathbb{E}(X_{t-1})] + \theta_1 \mathbb{E}[\epsilon_{t-1} X_{t-1}] - \theta_1 \mathbb{E}[\epsilon_{t-1} \mathbb{E}(X_{t-1})] \\ &= \phi_1 \gamma_0 + \theta_1 \mathbb{E}[\epsilon_{t-1} X_{t-1}] \\ &= \phi_1 \gamma_0 + \theta_1 \mathbb{E}[\epsilon_{t-1} (c + \phi_1 X_{t-2} + \epsilon_{t-1} + \theta_1 \epsilon_{t-2})] \\ &= \phi_1 \gamma_0 + \theta_1 \mathbb{E}[c \epsilon_{t-1} + \epsilon_{t-1} \phi_1 X_{t-2} + \epsilon_{t-1}^2 + \theta_1 \epsilon_{t-1} \epsilon_{t-2}] \\ &= \phi_1 \gamma_0 + \theta_1 \sigma^2\end{aligned}$$

$$\begin{aligned}
\gamma_2 &= \mathbb{E}[(X_t - \mathbb{E}(X_t))(X_{t-2} - \mathbb{E}(X_{t-2}))] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1})) + \epsilon_t + \theta_1\epsilon_{t-1})(X_{t-2} - \mathbb{E}(X_{t-2}))] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-2} - \mathbb{E}(X_{t-2})) + \epsilon_t(X_{t-2} - \mathbb{E}(X_{t-2})) + \theta_1\epsilon_{t-1}(X_{t-2} - \mathbb{E}(X_{t-2}))] \\
&= \mathbb{E}[(\phi_1(X_{t-1} - \mathbb{E}(X_{t-1}))(X_{t-2} - \mathbb{E}(X_{t-2}))] \\
&\quad + \mathbb{E}[\epsilon_t X_{t-2}] - \mathbb{E}[\epsilon_t \mathbb{E}(X_{t-2})] + \theta_1 \mathbb{E}[\epsilon_{t-1} X_{t-2}] - \theta_1 \mathbb{E}[\epsilon_{t-1} \mathbb{E}(X_{t-2})] \\
&= \phi_1 \gamma_1
\end{aligned}$$

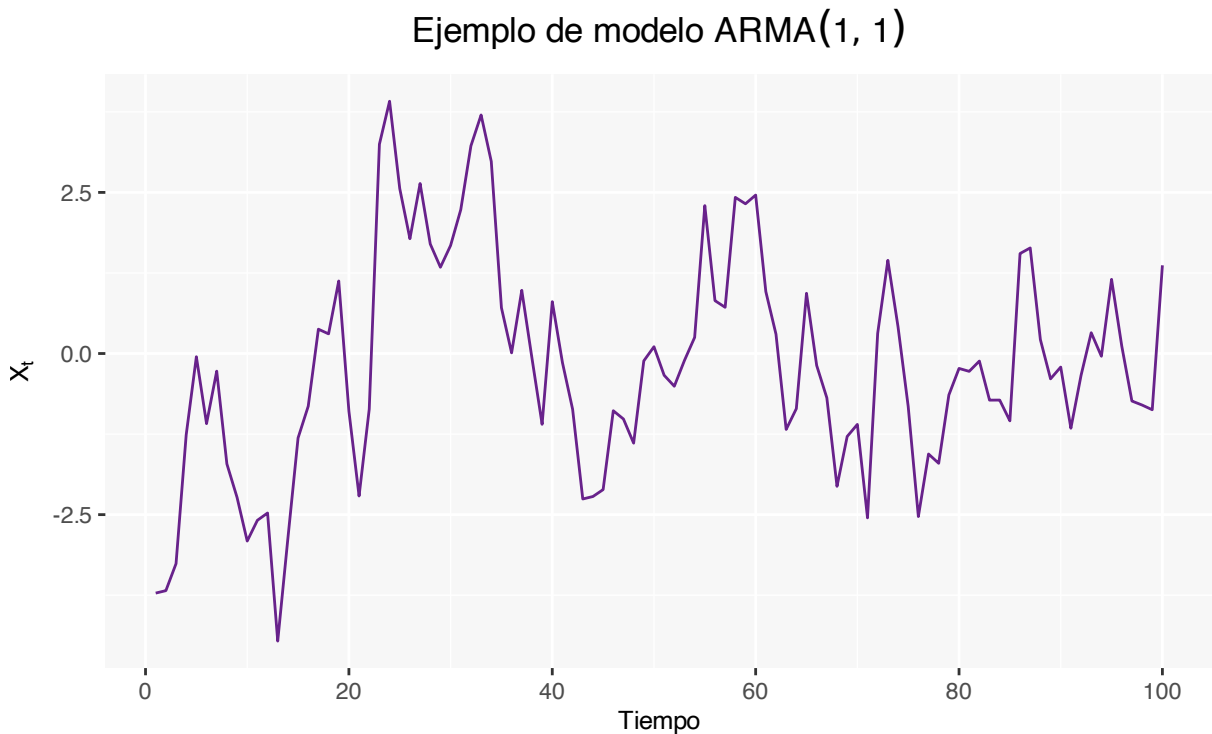
Entonces la **función de autocovarianza** de un proceso $ARMA(1,1)$ es:

$$\gamma_k = \begin{cases} \frac{\sigma^2(1+\theta_1^2+2\phi_1\theta_1)}{1-\phi_1^2} & k = 0 \\ \phi_1\gamma_0 + \theta_1\sigma^2 & k = 1 \\ \phi_1\gamma_{k-1} & k > 1 \end{cases}$$

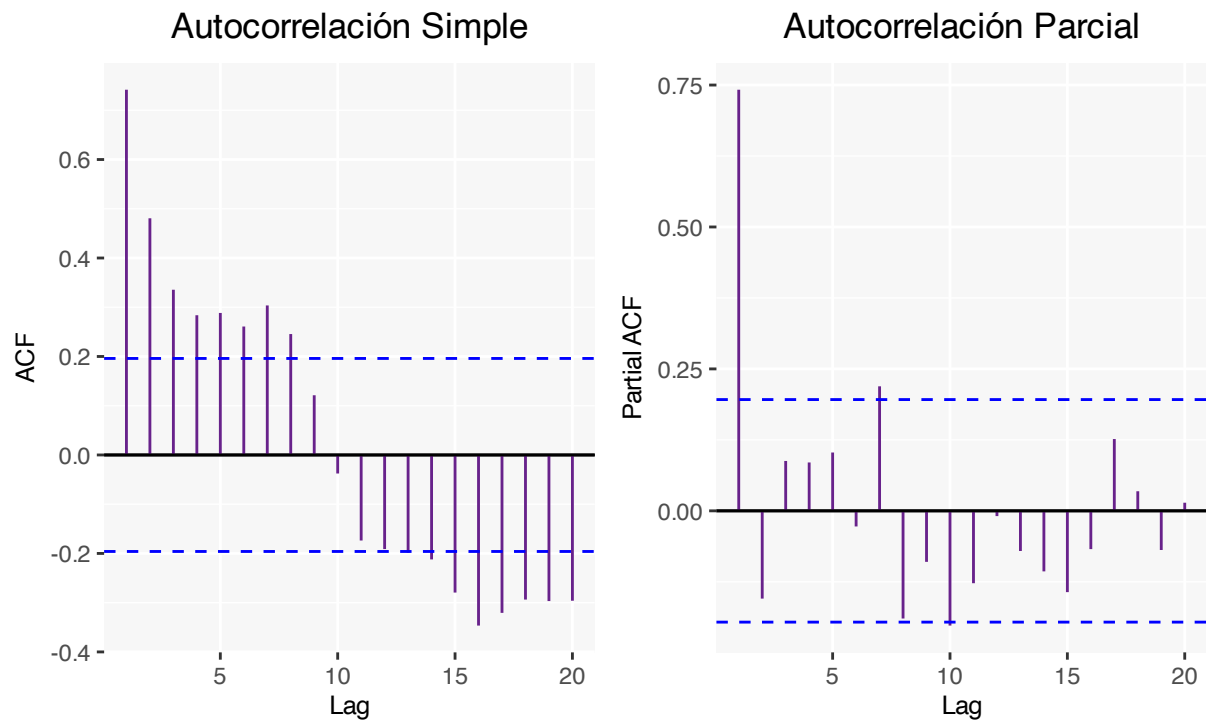
Y la **función de autocorrelación** de un proceso $ARMA(1,1)$ es:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{\phi_1\gamma_0 + \theta_1\sigma^2}{\gamma_0} = \phi_1 + \frac{\theta_1\sigma^2}{\gamma_0} & k = 1 \\ \frac{\phi_1\gamma_{k-1}}{\gamma_0} = \phi_1\rho_{k-1} & k > 1 \end{cases}$$

A continuación se muestran los resultados para un modelo $ARMA(1,1)$ de la siguiente forma $X_t = 0.65X_{t-1} + .30\epsilon_{t-1}$



Además de las gráficas de Autocorrelación simple y parcial.



La sección especializada a este tema en el blog es el apartado ARMA.

8.2 Ejercicios

- 1) Los modelos $ARMA(p, q)$ tienen componente autoregresivo y componente de medias móviles
 - a. Cierto
 - b. Falso
 - c. Sólo si p y q son mayores a cero
 - d. Sólo si p y q son distintos de cero
- 2) En un modelo definido por $X_t = 0.65X_{t-1} + \epsilon_t + 0.30\epsilon_{t-1}$ con $\epsilon_t \sim N(0, 1)$, calcular ρ_2
- 3) Suponga que el correlograma de una serie de tiempo con 100 observaciones tiene: $\hat{\rho}_1 = 0.31$, $\hat{\rho}_2 = 0.07$, $\hat{\rho}_3 = -0.05$, $\hat{\rho}_4 = 0.06$, $\hat{\rho}_5 = -0.03$, $\hat{\rho}_6 = 0.27$, $\hat{\rho}_7 = 0.08$, $\hat{\rho}_8 = 0.05$, $\hat{\rho}_9 = 0.02$, $\hat{\rho}_{10} = -0.01$. ¿Qué modelo ARMA es apropiado?

Part III

Procesos lineales no estacionarios

Chapter 9

$ARIMA(p, d, q)$: Proceso Autoregresivo Integrado y de Medias Móviles

Los modelos de series de tiempo AR , MA y $ARMA$ se basan en el supuesto de estacionariedad del proceso, es decir, la media, la varianza y las covarianzas son constantes en el tiempo. Sin embargo, muchas series de tiempo relacionadas con aplicaciones reales no son estacionarias, ya sea porque cambian de nivel en el tiempo (existe tendencia) o la varianza no es constante en el tiempo, a este tipo de procesos se les conoce como **procesos integrados**.

Para trabajar con estas series de tiempo lo que se hace es calcular las diferencias de la serie de tiempo d veces para hacerla estacionaria y posteriormente aplicar a la serie diferenciada un modelo $ARMA(p, q)$, en este caso se diría que la serie original es un proceso $ARIMA(p, d, q)$, donde p es el número de términos autoregresivos, d el número de veces que la serie debe ser diferenciada para hacerla estacionaria y q el número de términos de la media móvil.

La expresión algebraica del proceso $ARIMA(p, d, q)$ es:

$$X_t^d = c + \phi_1 X_{t-1}^d + \dots + \phi_p X_{t-p}^d - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} + \epsilon_t$$

donde X_t^d es la serie de las diferencias de orden d , ϵ_t es un proceso de ruido blanco, y $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ son los parámetros del modelo.

El modelo $ARIMA(p, d, q)$ se puede escribir en términos de los operadores de retardo de la siguiente manera:

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

Los modelos integrados se usan para reducir la **NO** estacionariedad de la serie al usar las diferencias y en la mayoría de las aplicaciones $d = 1$ y hasta $d = 2$ es suficiente para volver la serie estacionaria, pero en algunas ocasiones, aplicar transformaciones a los datos tienen mejor resultado para lograr la estacionariedad en comparación con las diferencias. Por ejemplo, en las series de tiempo económicas, la variabilidad aumenta cuando el nivel promedio del proceso aumenta. Sin embargo el porcentaje de cambio en las observaciones es relativamente independiente del nivel, entonces aplicar una transformación logarítmica puede ser más eficiente para lograr la estacionariedad de la serie.

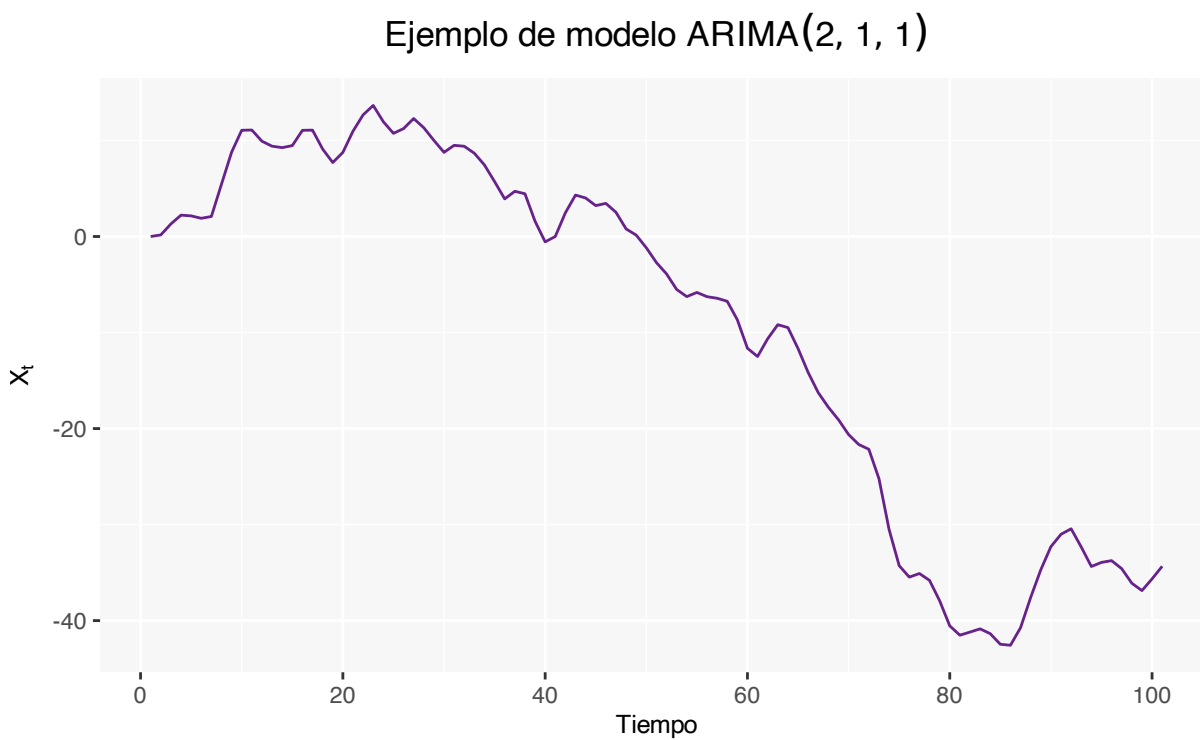
Ejemplo

Se tiene el siguiente modelo: $X_t = \mu_t + Z_t$ donde $\mu_t = \beta_0 + \beta_1 t$ y Z_t es un proceso estacionario.

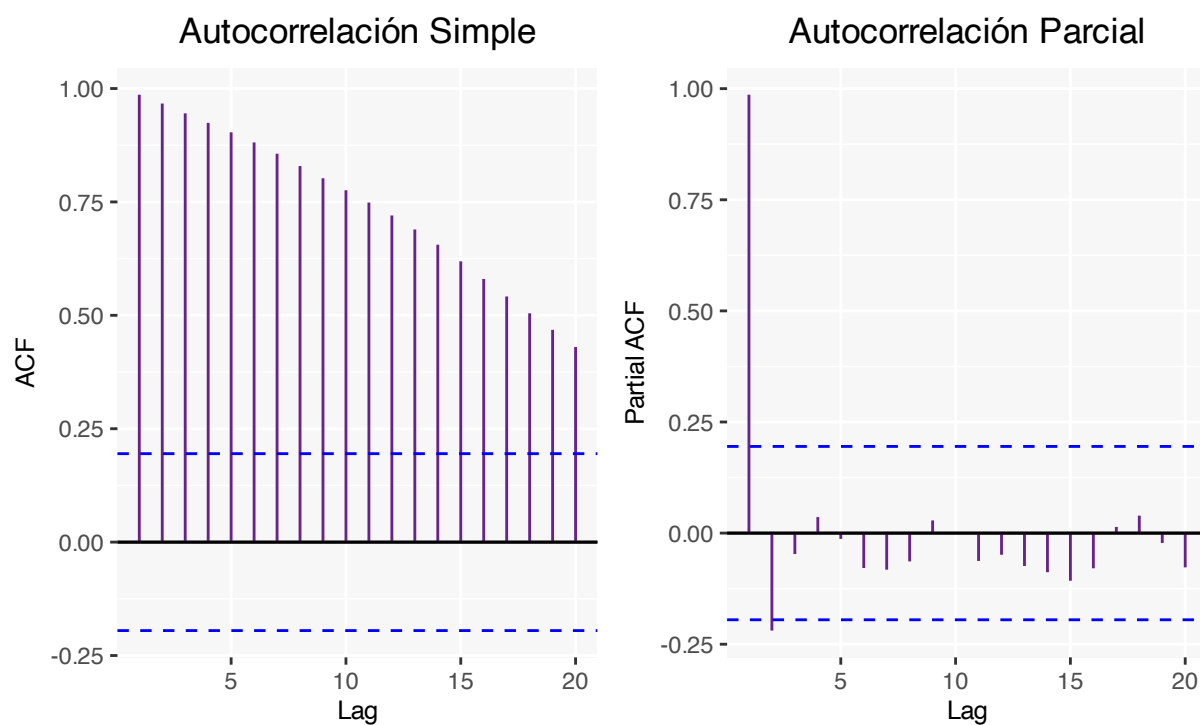
$$\begin{aligned}
 X_t - X_{t-1} &= \mu_t + Z_t - \mu_{t-1} - Z_{t-1} \\
 &= \beta_0 + \beta_1 t + Z_t - \beta_0 - \beta_1(t-1) - Z_{t-1} \\
 &= \beta_1 + Z_t - Z_{t-1}
 \end{aligned}$$

El cual ya no depende de t .

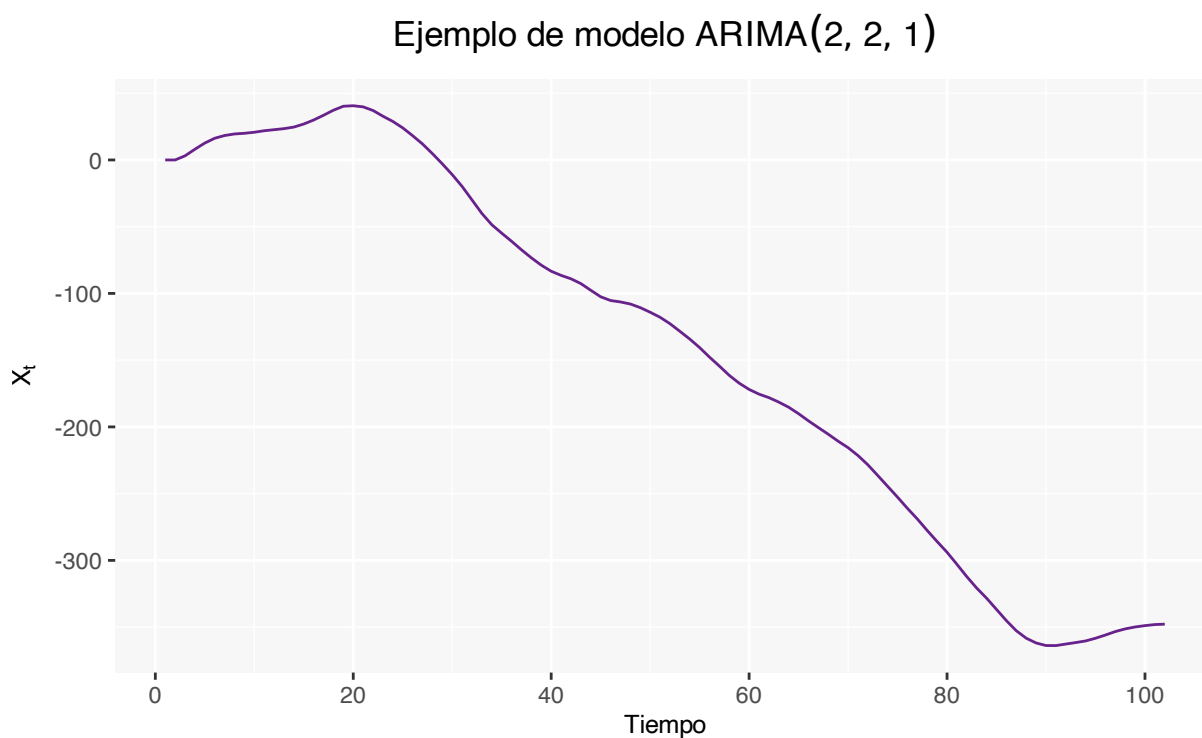
A continuación se muestran los resultados para un modelo $ARIMA(2, 1, 1)$ de la siguiente forma $X_t = 0.85X_{t-1} - 0.45X_{t-2} + .30\epsilon_{t-1}$



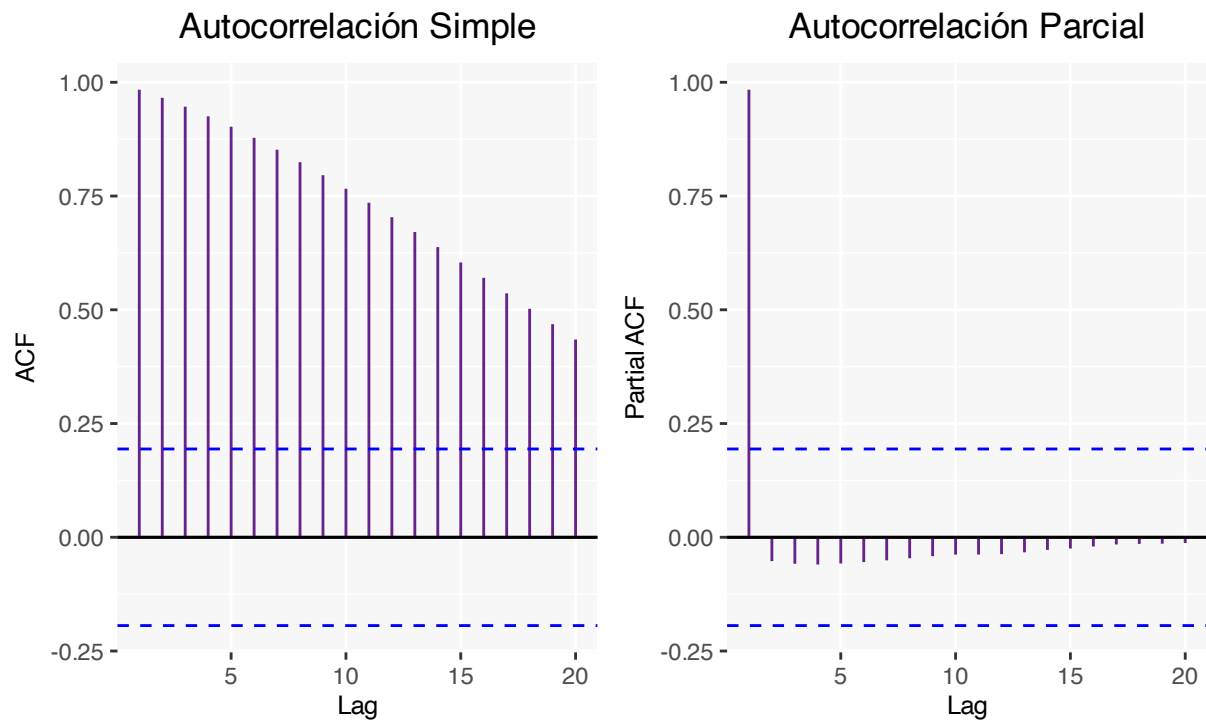
Además de las gráficas de Autocorrelación simple y parcial.



A continuación se muestran los resultados para un modelo $ARIMA(2, 2, 1)$ de la siguiente forma $X_t = 0.85X_{t-1} - 0.45X_{t-2} + .30\epsilon_{t-1}$



Además de las gráficas de Autocorrelación simple y parcial.



¿Qué se puede decir del efecto que tiene el parámetro d al comparar las gráficas de ambas series?

La sección especializada a este tema en el blog es el apartado ARIMA

9.1 Ejercicios

- 1) 57 mediciones consecutivas de una máquina están en el objeto *deere3* del paquete TSA:
 - a. Grafique la serie de tiempo. ¿Qué patrón observa?, ¿Un modelo estacionario podría ser adecuado a ese gráfico?
 - b. Usando herramientas como ACF y PACF especifique el modelo tentativo ($AR(p)$; $MA(q)$; $ARMA(p,q)$). Justifique.
- 2) Con los datos seleccionados para las prácticas
 - g. Apliqué un autoarima y comparé respecto a lo que esperaba derivado del análisis de los correlogramas realizados en la práctica 3

Chapter 10

$IMA(d, q)$: Proceso Integrado y de Media Móvil

Este proceso es un caso particular del proceso $ARIMA(p, d, q)$ cuando $p = 0$. Se calculan las diferencias de la serie de tiempo d veces para convertir la serie en estacionaria y luego se aplica un modelo $MA(q)$.

Por ejemplo:

Sea Z_1, Z_2, Z_3, Z_4, Z_5 la serie de tiempo original. Entonces la serie diferenciada de orden uno será:

$$\begin{aligned}Y_1 &= Z_2 - Z_1 \\Y_2 &= Z_3 - Z_2 \\Y_3 &= Z_4 - Z_3 \\Y_4 &= Z_5 - Z_4\end{aligned}$$

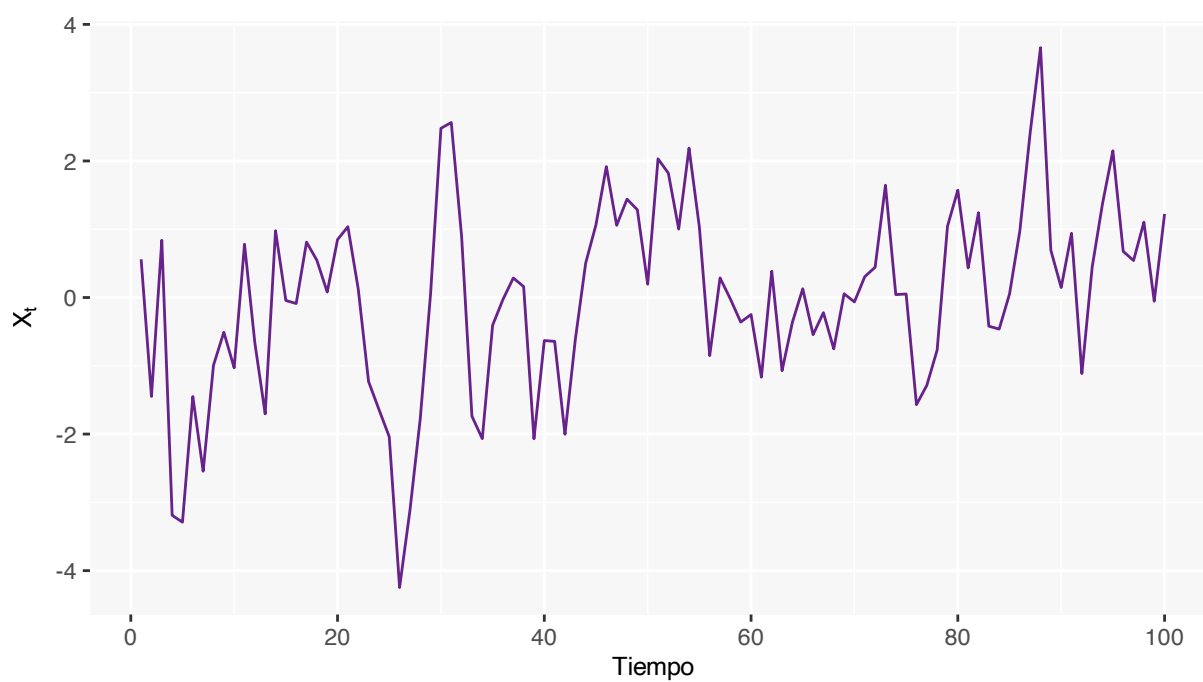
Y la serie diferenciada de orden dos será:

$$\begin{aligned}X_1 &= Y_2 - Y_1 = Z_3 - Z_2 - Z_2 + Z_1 = Z_3 - 2Z_2 + Z_1 \\X_2 &= Y_3 - Y_2 = \dots = Z_4 - 2Z_3 + Z_2 \\X_3 &= Y_4 - Y_3 = \dots = Z_5 - 2Z_4 + Z_3\end{aligned}$$

Por lo tanto al hacer diferencia de orden 2 a la serie de tiempo original Z_t , se obtiene una serie de tiempo X_1, X_2, X_3 que ya es estacionaria y a esta se le aplicaría un modelo de medias móviles $MA(q)$, por ejemplo si $q = 1$ se ajustaría un modelo $X_t = X_{t-1} + \epsilon_t$, donde ϵ_t es ruido blanco.

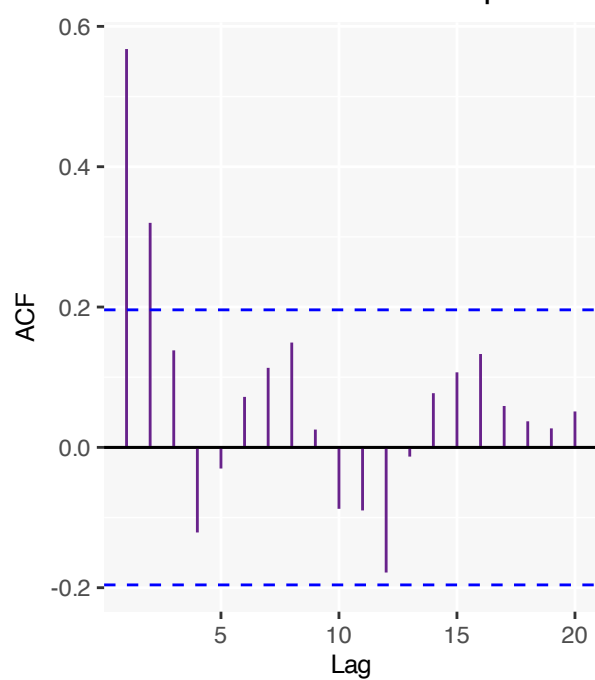
A continuación se modelan los datos de un modelo $MA(3)$ de la siguiente forma $X_t = 0.80\epsilon_{t-1} + 0.20\epsilon_{t-2} + 0.50\epsilon_{t-3}$

Ejemplo de modelo MA(3)

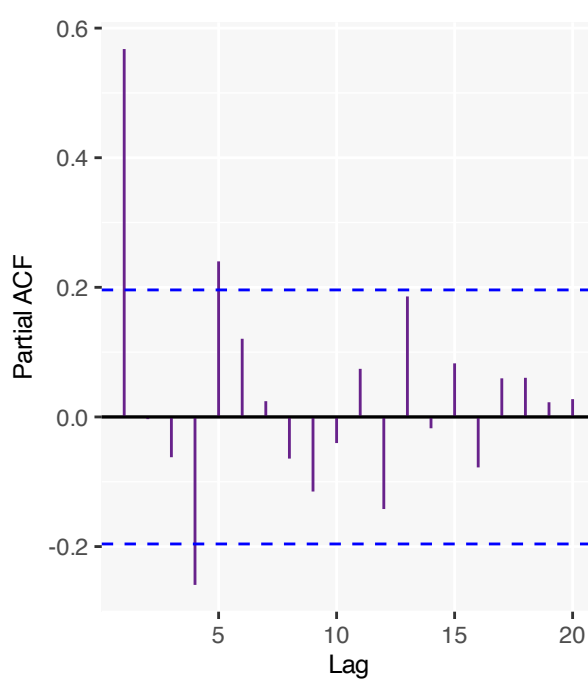


Además de las gráficas de Autocorrelación simple y parcial.

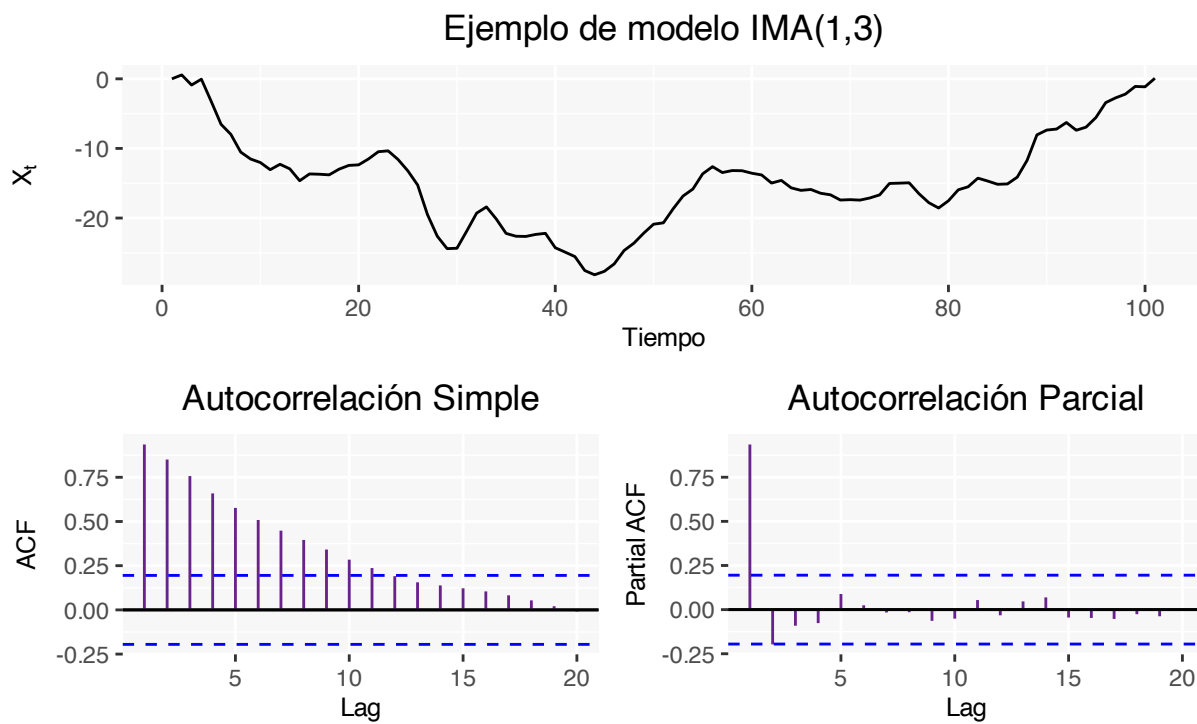
Autocorrelación Simple



Autocorrelación Parcial



Ahora veamos como queda la serie si le incluimos un proceso integrado (una diferencia) de orden 1.



Haga la prueba aumentando el orden de las diferencias, ¿cuál es el efecto en la serie de tiempo?

Part IV

Otras consideraciones

Chapter 11

Predictores Lineales

Sea X_t un proceso estacionario. Se desea predecir X_{n+1} en una función lineal de lo observado hasta el tiempo n es decir con X_1, X_2, \dots, X_n . Entonces lo que buscamos es encontrar los coeficientes a_i con $i = 1, \dots, n$ tales que

$$\hat{X}_{n+1} = a_1 X_n + a_2 X_{n-1} + \dots + a_n X_1$$

Utilizando las varianzas y covarianzas entre los elementos de la serie de tiempo se obtiene que una estimación de las a_i 's son las que den solución al siguiente sistema:

$$\begin{bmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) & \text{Cov}(X_1, X_3) & \dots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) & \text{Cov}(X_2, X_3) & \dots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \text{Cov}(X_n, X_3) & \dots & \text{Cov}(X_n, X_n) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \text{Cov}(X_{n+1}, X_n) \\ \text{Cov}(X_{n+1}, X_{n-1}) \\ \vdots \\ \text{Cov}(X_{n+1}, X_1) \end{bmatrix}$$

$$\Gamma \underline{a} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} = \underline{\gamma}$$

En general si deseamos predecir X_{n+h} dado X_1, X_2, \dots, X_n , el predictor lineal será

$$\hat{X}_{n+h} = a_1 X_n + a_2 X_{n-1} + \dots + a_n X_1$$

En donde el mejor predictor lineal serán los a_1, a_2, \dots, a_n tales que

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \gamma_h \\ \gamma_{h+1} \\ \vdots \\ \gamma_{h+n} \end{bmatrix}$$

$$\Gamma \underline{a} = \underline{\gamma}$$

Ejemplo ¿Cuál es el predictor lineal de X_3 dado X_1, X_2 ?

$$\begin{bmatrix} \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$$

Como $\rho_k = \frac{\gamma_k}{\gamma_0}$

$$\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \implies \begin{aligned} a_1 + \rho_1 a_2 &= \rho_1 \\ a_1 \rho_1 + a_2 &= \rho_2 \end{aligned}$$

Y habría que resolver el sistema de ecuaciones para encontrar los valores a_1 y a_2 .

En general: $\hat{X}_{n+h} = a_1 X_n + a_2 X_{n-1} + \dots + a_n X_1$.

Para este tema, se tiene preparado el material del tema Evaluación y pruebas de hipótesis.

11.1 Ejercicios

- 1) Sean X_1, X_2, X_4, X_5 obs. de un modelo $MA(2)$ de la forma $X_t = Z_t + 0.5Z_{t-1} + 0.25Z_{t-2}$. Encontrar el predictor lineal para X_3 en términos de X_1 y X_2 . $Z_t \sim N(0, 1)$.
- 2) Sean X_1, X_2, X_3, X_4 obs. de un modelo $MA(2)$ de la forma $X_t = Z_t - 0.25Z_{t-2}$. Encontrar el predictor lineal para X_5 en términos de X_3 y X_4 . $Z_t \sim N(0, 1)$.
- 3) Sean X_1, X_2 obs. de un modelo $AR(1)$ de la forma $X_t = 0.35X_{t-1} + Z_t$. Encontrar el predictor lineal para X_3 en términos de X_1 y X_2 . $Z_t \sim N(0, 1)$

Chapter 12

Construcción de modelos para series de tiempo univariadas

Para construir un modelo *ARIMA* que ajuste a una serie tiempo dada, se debe seguir un proceso iterativo de tres etapas. Primero identificar un modelo *ARIMA*(p, d, q) tentativo, segundo, estimar los parámetros desconocidos del modelo. Tercero, mediante el análisis de residuales verificar si el modelo propuesto es el adecuado.

- **Identificación:** Utilizando los datos ordenados cronológicamente haciendo uso de gráficos (correlograma, diagrama de dispersión, otros) se seleccionan los modelos *ARIMA*(p, d, q) que valga la pena investigar. En esta etapa es posible identificar más de un modelo candidato que describa la serie. Observando las gráficas del ACF y PACF de la serie transformada podemos hacernos una idea del modelo que describe nuestra serie, o al menos de cuáles son los primeros candidatos que debemos probar.
- **Estimación:** Considerando el modelo o modelos apropiados seleccionados en el paso anterior, se procede a realizar inferencia sobre los parámetros del modelo. Algunos paquetes permiten la selección del método de estimación (verosimilitud, momentos, mínimos cuadrados) que mejor se ajuste a las especificaciones del problema.
- **Verificación:** Si el modelo es el adecuado, es decir los valores de p y q han sido correctamente especificados, entonces el modelo deberá ajustar bien a los datos y los residuales (la diferencia entre lo observado y lo estimado con el modelo) deberán comportarse como ruido blanco y no tener ninguna estructura. Esto significa que hay que probar que los residuales son estacionarios y no relacionados.

Para verificar la no autocorrelación se puede usar la prueba de Ljung-Box vista en el capítulo 4, pero también la prueba Box-Pierce que es una versión simplificada de la prueba Ljung-Box sirve para evaluar si la serie de tiempo está autocorrelacionada.

Sea $n = \text{longitud}(X_t)$, $\rho_i = \text{autocorrelación de } X \text{ en el retraso } i$, entonces el estadístico de la prueba de Ljung-Box es $n * (n + 2) * (\frac{\rho_1^2}{n-1} + \frac{\rho_2^2}{n-2} + \dots + \frac{\rho_k^2}{n-k})$. Y el estadístico de la prueba Box-Pierce es $n * (\rho_1^2 + \rho_2^2 + \dots + \rho_k^2)$. Bajo la hipótesis nula definida como que no existe autocorrelación, las estadísticas de prueba tienen una distribución Chi-cuadrado con grados de libertad los rezagos(k).

Asimismo, la prueba Durbin-Watson también puede utilizarse para probar autocorrelación en las series de tiempo.

La prueba aumentada de Dickey Fuller (prueba ADF) es una prueba estadística comúnmente utilizada en series de tiempo para probar la estacionariedad de la serie. Recordemos que para el ajuste de modelos tipo ARIMA, el primer paso es determinar el número de diferenciaciones necesarias para que la serie sea estacionaria. Por lo tanto, pruebas como la Dickey Fuller Aumentada (ADF), la KPSS y la Phillips-Perron se utilizan para determinar la estacionariedad de una serie en los modelos autoregresivos. La prueba ADF es fundamentalmente una prueba de hipótesis en donde se plantea una hipótesis nula y una alternativa y a partir del estadístico de prueba y el valor p que se puede inferir si una serie de tiempo es estacionaria o no.

Si se ajustan varios modelos candidatos $ARIMA(p, d, q)$, un buen modelo será aquel que tenga los residuales semejantes al de un ruido blanco, además que tenga los valores del AIC (Criterio de Información de Akaike) y BIC (Criterio de Información Bayesiana) menores con relación al resto de los modelos candidatos.

A manera de guía de aplicación (la cual regularmente es cíclica) se recomienda revisar o repasar los temas para el ajuste de modelos univariados de series de tiempo:

- Autocorrelación
- Transformaciones
- ARIMA
- Evaluación y pruebas de hipótesis
- Predicción

12.1 Ejercicios

- 1) Con los datos seleccionados para las prácticas
 - h. Definir el modelo de serie de tiempo que mejor ajusta a los datos. ¿Tiene alguna limitación?
 - i. Realizar las predicciones con base en el modelo propuesto. ¿Cuáles son las consideraciones que habría que tener? En este ejercicio pondrás en práctica todos tus conocimientos adquiridos además de tu creatividad para generar resultados interesantes.

Part V

Procesos no lineales

Chapter 13

Introducción

Los procesos vistos hasta ahora son lineales, pero muchas series temporales económicas, y especialmente series financieras, no cumplen con el supuesto de varianza constante, y no sólo eso, sino que estos cambios tienden a estar correlacionados serialmente, en el sentido de que cambios de gran magnitud en el valor de la serie son seguidos por grandes cambios (periodos de mucha volatilidad) mientras que a cambios pequeños en el valor de la serie les siguen cambios pequeños (periodos de poca volatilidad).

También puede ocurrir que por la propia naturaleza del fenómeno económico que se está analizando se requiera conocer no solo aspectos de su nivel medio sino que también nos interese su varianza (o volatilidad). Usar un proceso ARIMA en una serie con estas características se observará que los residuos recogerán todo el efecto de la variabilidad de la serie, ya que los modelos lineales no consideran la variabilidad dentro del modelo.

Existen diferentes tipos de modelos no lineales, entre ellos los más populares son los ARCH y GARCH para predecir la volatilidad.

Chapter 14

ARCH(q): Proceso Autoregresivo con Heterocedasticidad Condicional

El nombre proviene del inglés, “Autoregressive conditional heteroscedasticity” (ARCH). Método propuesto por Engles en 1982, que determina un patrón de comportamiento estadístico para la varianza. Considera que la información pasada de una variable y su volatilidad son factores que explican su comportamiento presente y, por tanto, podrá ser extrapolado a futuro.

La expresión algebraica básica del proceso *ARCH*(r) es:

$$X_t = \sigma_t \epsilon_t$$

Donde ϵ_t (proceso de ruido blanco formado por variables aleatorias normales independientes de media cero y varianza uno) y σ_t (factor denominado volatilidad) son procesos estacionarios independientes entre sí. La condición de independencia entre ϵ_t y σ_t , garantiza que la serie X_t tenga media marginal igual a cero:

$$E(X_t) = E(\sigma_t \epsilon_t) = E(\sigma_t)E(\epsilon_t) = 0$$

La media condicional también es nula.

$$E(X_t | X_{t-1}) = E(\sigma_t \epsilon_t | X_{t-1}) = E(\sigma_t | X_{t-1})E(\epsilon_t | X_{t-1}) = E(\sigma_t | X_{t-1})E(\epsilon_t) = 0$$

La varianza marginal de X_t es constante (σ^2) y se calcula como:

$$E(X_t^2) = E(\sigma_t^2 \epsilon_t^2) = E(\sigma_t^2)E(\epsilon_t^2) = \sigma^2 \times 1 = \sigma^2$$

Sin embargo, la varianza condicionada no es constante:

$$E(X_t^2 | X_{t-1}) = E(\sigma_t^2 \epsilon_t^2 | X_{t-1}) = E(\sigma_t^2 | X_{t-1})E(\epsilon_t^2 | X_{t-1}) = E(\sigma_t^2 | X_{t-1})E(\epsilon_t^2) = \sigma_t^2 \times 1 = \sigma_t^2$$

Por tanto, σ_t^2 , representa la varianza condicionada de la serie en cada instante, que va variando con cierta estructura estacionaria.

La condición de independencia entre σ_t y ϵ_t , además de garantizar que la serie X_t tenga media marginal igual a cero, nos garantiza que la serie carezca de autocorrelación y forme un proceso de ruido blanco. Sin embargo, la serie X_t no es de variables independientes.

El caso más simple de este proceso es un modelo *ARCH*(1) (la varianza condicional depende de un retardo de la serie) y se define como:

Un proceso estacionario X_t sigue un modelo *ARCH*(1) si y sólo si

$$X_t = \sigma_t \epsilon_t$$

donde ϵ_t es ruido blanco y

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

Observemos que la varianza condicional (σ_t^2) tiene una estructura similar a un AR(1), y por tanto solo depende del último valor observado. Por tanto, si el valor de X_t^2 es alto, la varianza σ_t^2 de la siguiente observación condicionada a este valor será también alta. Esto producirá correlación entre los cuadrados de la serie, provocando rachas de valores de magnitud relativamente elevada o con mayor varianza. Pero como la media marginal y la media condicionada valen cero, aunque la varianza condicionada sea alta, siempre es posible que aparezca un valor pequeño de X_t^2 , que disminuirá la varianza condicionada de la observación siguiente y facilitará que la siguiente observación sea pequeña en valor absoluto. De manera que la serie puede presentar rachas de valores altos, pero globalmente será estacionaria.

El modelo anterior puede generalizarse permitiendo una dependencia de la varianza condicional con q retardos. De manera que el modelo será $ARCH(q)$, el cual se define como:

Un proceso estacionario X_t sigue un modelo $ARCH(q)$ si y sólo si

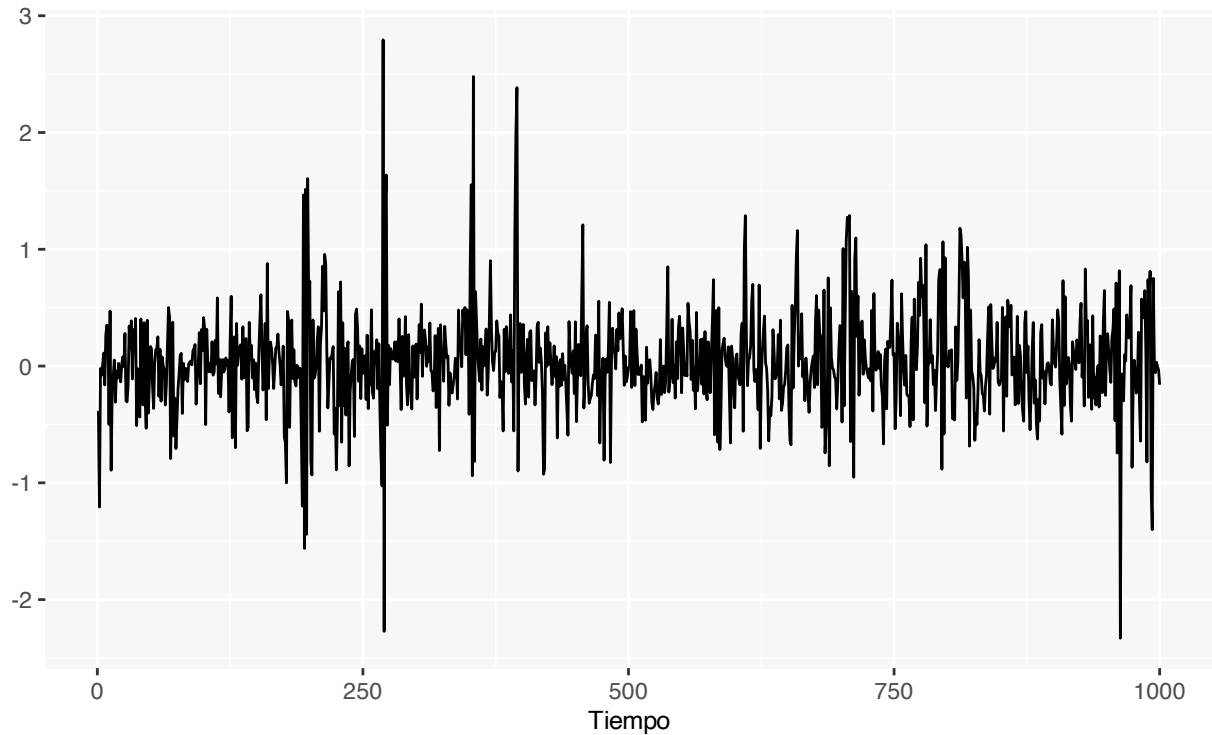
$$X_t = \sigma_t \epsilon_t$$

donde ϵ_t es ruido blanco y

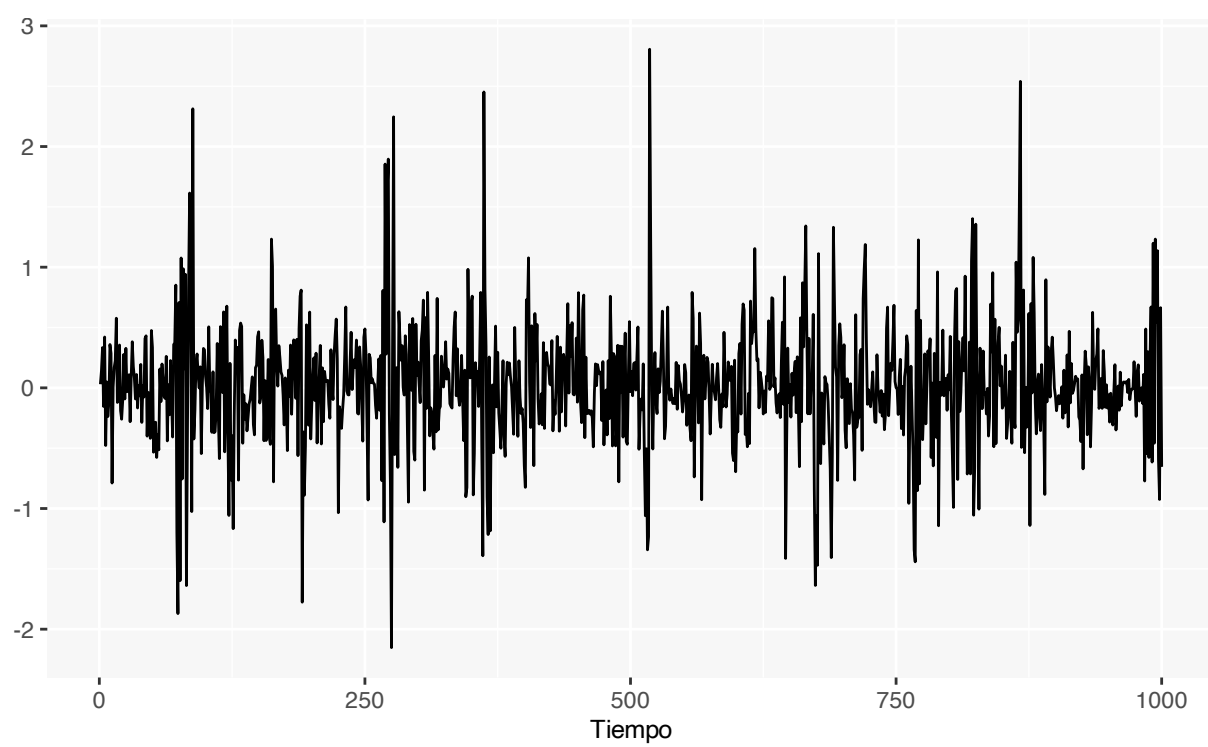
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2 + \dots + \alpha_q X_{t-q}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2$$

Donde $\alpha_0 > 0$ y $\alpha_i \geq 0, i = 1, \dots, q$. Para garantizar que el proceso σ_t^2 sea estacionario se requiere que $\sum_{i=1}^q \alpha_i < 1$.

A continuación se muestran las simulaciones de un modelo $ARCH(1)$ con la varianza modelada de la siguiente forma: $\sigma_t^2 = 0.05 + 0.8X_{t-1}^2$



A continuación se muestran las simulaciones de un modelo $ARCH(3)$ con la varianza modelada de la siguiente forma: $\sigma_t^2 = 0.05 + 0.6X_{t-1}^2 + 0.2X_{t-2}^2 + 0.1X_{t-3}^2$



Chapter 15

GARCH(p, q): Proceso Autoregresivo Generalizado con Heterocedasticidad Condicional

En el modelo *ARCH*(1) el predictor al tiempo $t + 1$ de la varianza depende solo del último valor de σ_t . Sin embargo, en la práctica se desea tener mayor precisión en la predicción, para mejorarla se podría incluir todos los valores pasados σ_t con menor peso para volatilidades más distantes. Una propuesta para este problema la desarrolló Bollerslev(1986), donde introduce p retrasos de la varianza condicional al modelo, entonces p hace referencia al orden del modelo *GARCH*.

Entonces un proceso estacionario X_t sigue un modelo *GARCH*(p, q) si

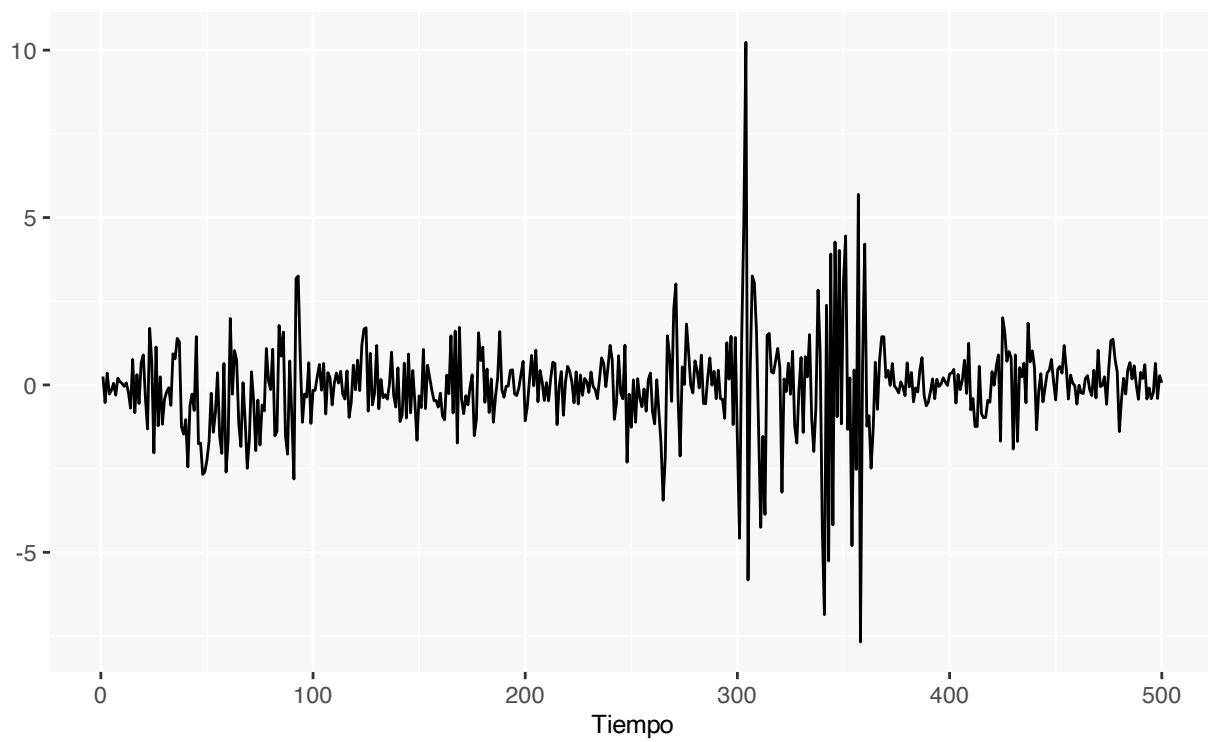
$$X_t = \sigma_t \epsilon_t$$

donde ϵ_t es ruido blanco y

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

Donde $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q$ y $\beta_j \geq 0, j = 1, \dots, p$. Para garantizar que la varianza sea positiva y existan los momentos de orden superior se requiere que $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$.

A continuación se muestran las simulaciones de un modelo *GARCH*(1, 1) con la varianza modelada de la siguiente forma: $\sigma_t^2 = 0.05 + 0.4X_{t-1} + 0.55\sigma_{t-1}^2$



Part VI

Tareas

Tarea 1

Documento ejecutivo sobre la serie de tiempo seleccionado a analizar. Debe incluir como mínimo:

- 1) Introducción sobre las características y/o el tipo de datos seleccionados y definición del problema a resolver con el análisis
- 2) Gráfico de los datos con interpretación
- 3) Hallazgos de la descomposición de la serie
- 4) Hallazgos de transformaciones aplicadas a la serie

Tarea 2

Resolver los siguientes ejercicios. Estos provienen de las secciones de ejercicios de los capítulos 6, 7, 8 y 11 del libro.

6.3.1 Encuentre la función de covarianza γ_k para el modelo $AR(1)$ en la sección 6.1, definido como $X_t = 0.35X_{t-1} + \epsilon_t$ donde $\epsilon_t \sim N(0, 1)$

7.3.1 Sea $Z_t \sim N(0, 1)$. Calcular la función de autocovarianza γ_k para el modelo $X_t = 0.3 + 0.1Z_t - 0.2Z_{t-1}$

7.3.2 Muestre que los procesos $MA(1)$ siguientes a. $Z_t = a_t + \theta a_{t-1}$ b. $Y_t = a_t + \frac{1}{\theta} a_{t-1}$ donde $0 < |\theta| < 1$ y $a_t \sim N(0, 1)$, tienen las mismas funciones de autocorrelación

8.2.2 En un modelo definido por $X_t = 0.65X_{t-1} + \epsilon_t + 0.30\epsilon_{t-1}$ con $\epsilon_t \sim N(0, 1)$, calcular ρ_2

11.1.2 Sean X_1, X_2, X_3, X_4 obs. de un modelo $MA(2)$ de la forma $X_t = Z_t - 0.25Z_{t-2}$. Encontrar el predictor lineal para X_5 en términos de X_3 y X_4 . $Z_t \sim N(0, 1)$.

11.1.3 Sean X_1, X_2 obs. de un modelo $AR(1)$ de la forma $X_t = 0.35X_{t-1} + Z_t$. Encontrar el predictor lineal para X_3 en términos de X_1 y X_2 . $Z_t \sim N(0, 1)$

Tarea 3

Documento ejecutivo sobre la serie de tiempo seleccionado a analizar. Al documento de la tarea 1 agregarle como mínimo:

- 5) Análisis de los correlogramas
- 6) Análisis e interpretación del ajuste de un ARIMA
- 7) Predicciones
- 8) Conclusión ligada con el problema que se planeó resolver

Bibliografía

Bibliography

Cryer, J. D. and Chan, K.-S. (2008). *Time series analysis: with applications in R*. Springer Science & Business Media.