Pronóstico de series de tiempo y simulaciones en R

Clase 3: Pronóstico de series de tiempo lineales

Germán Eduardo González

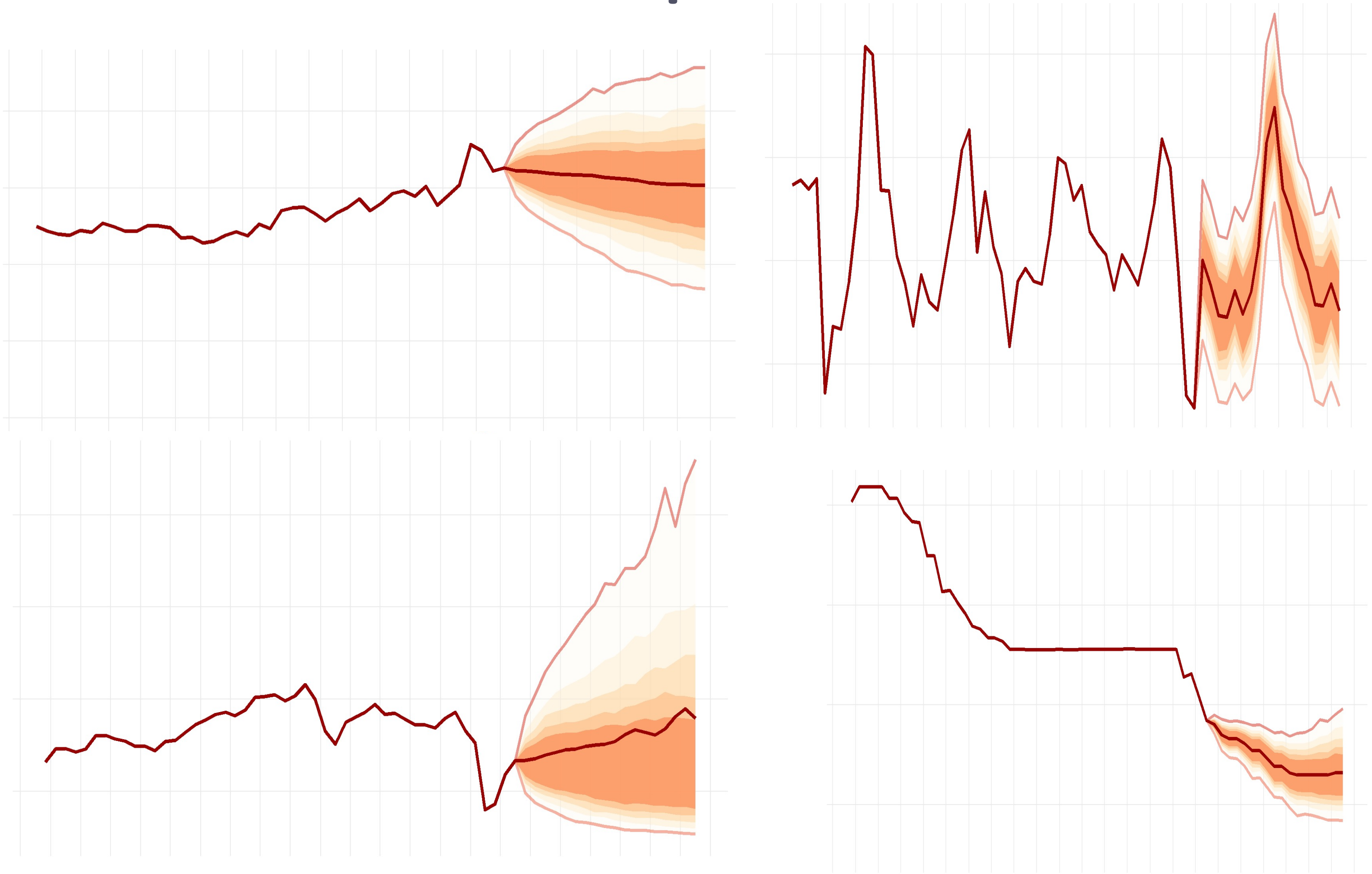
Índice

Índice

Pronóstico punto, intervalo y distribución

- Son pocas veces en la practica en la que se está interesado en un único número como pronóstico:
  - **Pronóstico punto:** Valor esperado predicho.
  - **Pronóstico intervalo:** Valores de dos percentiles de interés (por ejemplo si es un intervalo del 95% de confianza sería el valor del percentil 2.5 y 97.5)
  - **Pronostico de la distribución:** Contiene los dos anteriores y otras propiedades propias de una distribución (varianza, simetría, curtosis, etc).

Pronóstico punto, intervalo y distribución



Ilustrativamente, la información de estos 3 tipos de pronósticos se pueden resumir en un fanchart.

Pronóstico punto, intervalo y distribución

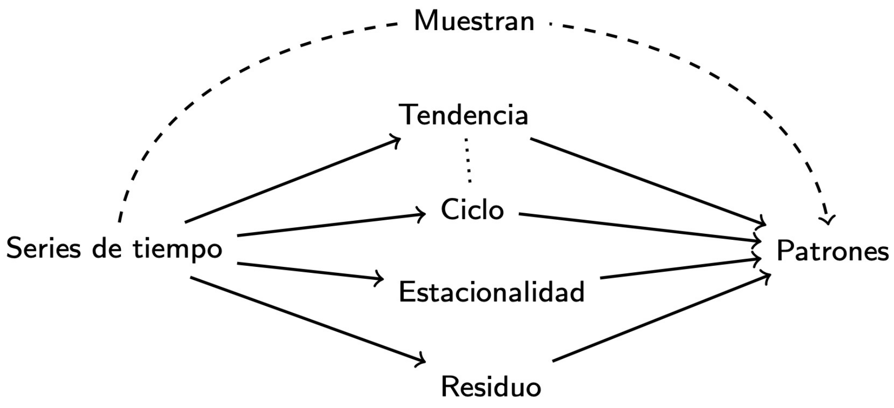
Nos concentraremos primero en el pronóstico punto:

- Entre mejor modelado esté, mejor modelado quedarán también los otros tipos de pronósticos.
- Los supuestos que se hacen para mapear correctamente este pronóstico, pueden definir directamente los otros dos (por ejemplo el caso de MCO).
- En la práctica, cuando el modelo es muy sofisticado, suele ser complicado llegar a una forma cerrada del pronóstico de intervalo y de distribución, por lo que se recurre a metodologías de simulación (Montecarlo, bootstraping, etc).

Pronóstico punto

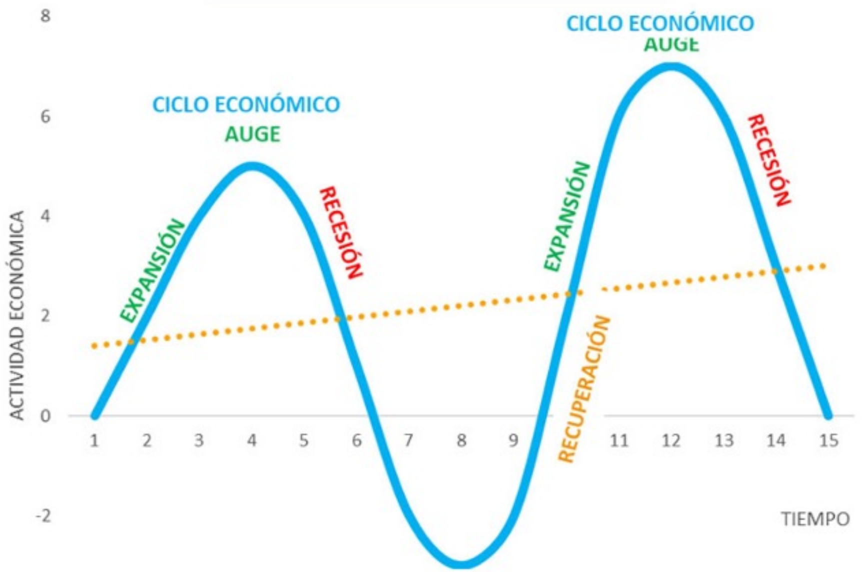
Recordando la sesión pasada, el pronóstico punto tiene 3 componentes tendencia, ciclo y estacionalidad.

- Hoy nos concentraremos en ciclo.
  - Cada componente se puede modelar por separado.
  - La **tendencia** puede modelarse de manera determinística (alguna función de t) o no determinística (diferenciando la serie) y terminar con una serie sin tendencia.
  - La **estacionalidad** se puede modelar de manera determinística (aditiva o multiplicativa) o no determinística (extensión de lo que veremos hoy, modelos SARIMA).



Ciclo

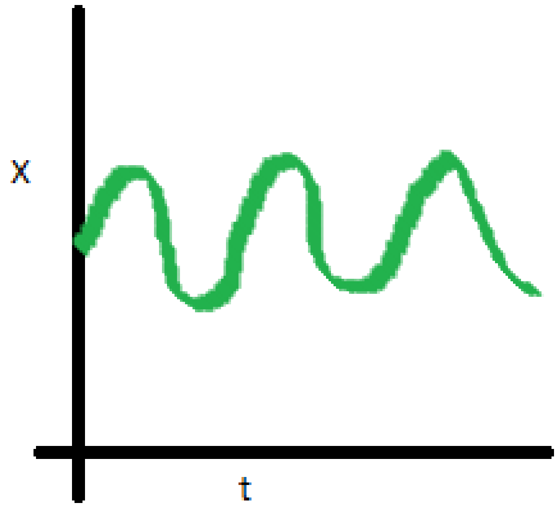
- Recapitulando de la sesión anterior, ciclo son los comportamientos alrededor de una media.
- En particular, se esperan oscilaciones alrededor de esta, pero siempre retorno a la misma.
- Implícitamente, lo que se está diciendo es que, controlando por los otros 2 componentes del pronóstico, se espera que la serie sea **estacionaria.**



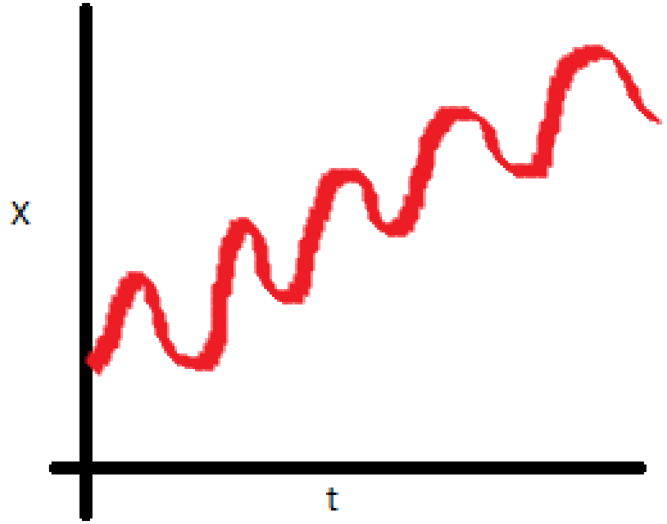
Estacionariedad

- Un proceso es estacionario si no cambia sus propiedades estadísticas con el tiempo (i.e media, varianza y covarianza).
  - La constancia en la varianza se denomina homoscedasticidad (en particular, la próxima sesión suavizaremos un poco este supuesto, al menos condicionalmente).
  - La función de la covarianza no depende del tiempo, solo debe depender de la distancia entre observaciones.
  - La media incondicional debe ser constante en el tiempo.

Estacionariedad



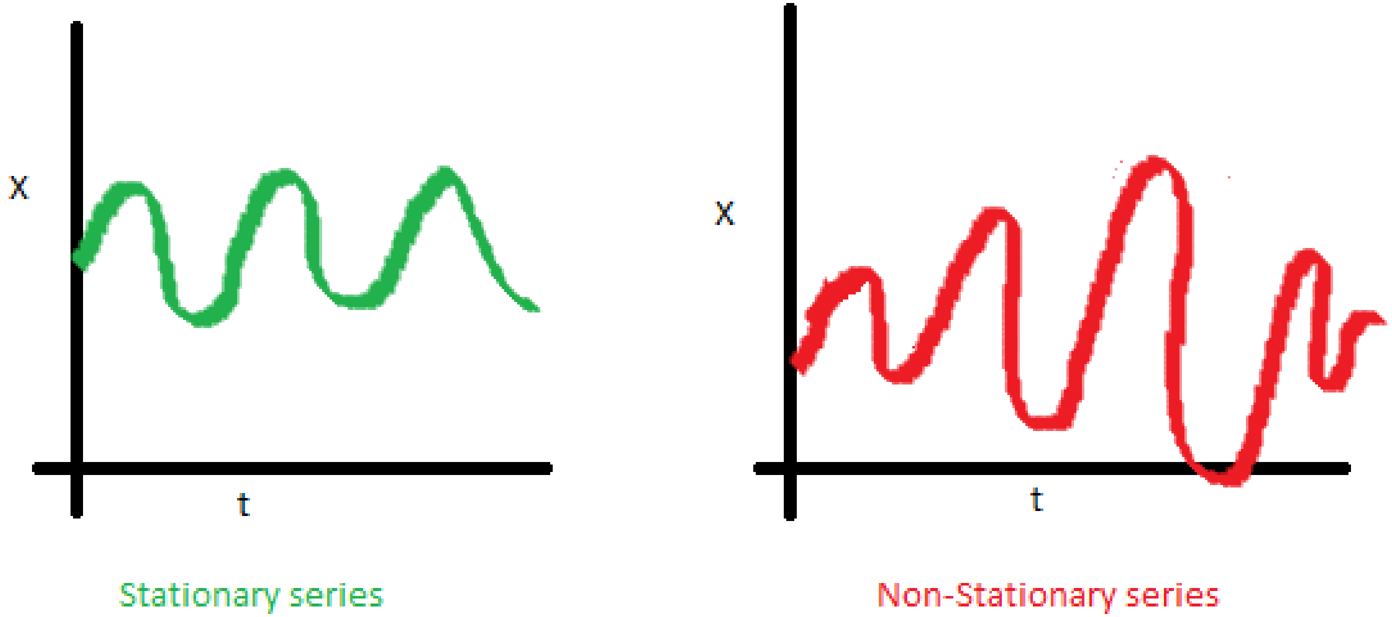
Stationary series



Non-Stationary series

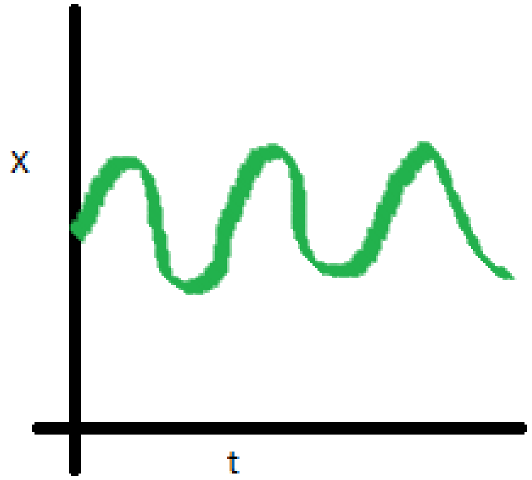
- La media de la serie de tiempo no debe ser una función del tiempo.
- La serie roja es no estacionaria pues su media aumenta en eltiempo.

Estacionariedad

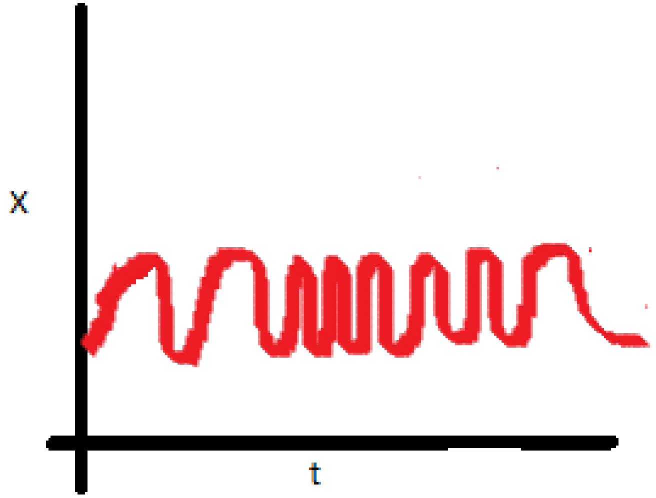


La varianza de la serie de tiempo no debe ser una función del tiempo. La serie roja es no estacionaria pues note la dispersión variable de datos a lo largo del tiempo. **Este supuesto lo suavizaremos un poco.**

Estacionariedad



Stationary series



Non-Stationary series

La covarianza del momento i y el momento i + m no debe ser una función del tiempo. La serie roja es no estacionaria pues la dispersión de sus datos disminuye con el tiempo.

¿Cómo verificar estacionariedad?

- La prueba más famosa para verificar estacionariedad es la Dickey-Fuller (DF).
- Esta prueba explota la noción de raíz unitaria (caminata aleatoria).

$$Y_t = \alpha_0 + Y_{t-1} + \epsilon_t$$

- Una serie que se comporta de esta manera **nunca será estacionaria (tiene una tendencia estocástica).**
- Entonces:

$$Y_t = \alpha_0 + \phi_1 Y_{t-1} + \epsilon_t$$
$$Y_t - Y_{t-1} = \alpha_0 + \phi_1 Y_{t-1} - Y_{t-1} + \epsilon_t$$
$$\Delta Y_t = \alpha_0 + (\phi_1 - 1) Y_{t-1} + \epsilon_t$$
$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \epsilon_t$$

¿Cómo verificar estacionariedad?

$$\Delta Y_t = \alpha_0 + (\phi_1 - 1)Y_{t-1} + \epsilon_t$$
$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \epsilon_t$$

- Lo que hace la prueba es hacer inferencia sobre γ con el fin de verificar si hay raíz unitaria.

$$H_0: \gamma = 1 \ \text{ Raíz } unitaria \ \rightarrow \text{no es estacionaria}$$
$$H_1: \gamma < 1 \text{ No raíz } unitaria \ \rightarrow \text{estacionaria}$$

Las raíces unitarias son una causa de no estacionariedad. Queremos que se rechace la hipótesis nula, y se acepte la hipótesis alternativa la serie es estacionaria. Queremos un p-valor <0.05 que indica que rechazamos la hipótesis nula de no estacionariedad.

¿Cómo verificar estacionariedad?

- Hay versiones aumentadas de esta prueba, Augmented Dickey-Fuller (ADF) donde se controla por rezagos de la diferencia para aumentar la eficiencia de esta.

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \sum_{k=1}^{K} \alpha_k \Delta Y_{t-k} + \epsilon_t$$

- **Dificultades**: existe el riesgo de tratar una serie estacionaria con cambio estructural como si presentase una raíz unitaria.
- **Solución**: La solución propuesta por Perron consistió en extender el contraste de Dickey-Fuller (ADF) introduciendo variables ficticias para recoger el efecto del cambio, considerando el punto de ruptura conocido a priori.

¿Cómo verificar estacionariedad?

$$\Delta Y_t = C + \beta t + \gamma Y_{t-1} + \epsilon_t$$

Dónde la prueba de Perron se puede escribir de 3 formas diferentes

a) $C = 0 \,\&\, \beta = 0$
b) $C \neq 0 \,\&\, \beta = 0$
c) $C \neq 0 \,\&\, \beta \neq 0$

$$H_0: \gamma = 1 \text{ Raíz } unitaria \rightarrow \text{no es estacionaria}$$
$$H_1: \gamma < 1 \text{ No raíz } unitaria \rightarrow \text{estacionaria}$$

Las raíces unitarias son una causa de no estacionariedad. Queremos que se rechace la hipótesis nula, y se acepte la hipótesis alternativa la serie es estacionaria. Queremos un p-valor <0.05 que indica que rechazamos la hipótesis nula de no estacionariedad.

Estacionariedad

- ¿Por qué nos interesa tener series estacionarias?
  - Son fáciles de predecir ya que podemos suponer que las propiedades estadísticas futuras son **iguales** o **proporcionales** a las **propiedades estadísticas actuales.**
  - La mayoría de los modelos que utilizamos en análisis de ST asumen **estacionariedad por covarianza.** Esto significa que las estadísticas descriptivas que estos modelos predicen (**medias, varianzas y correlaciones**) solo son confiables si la ST es **estacionaria** e inválidas de lo contrario.

Estacionariedad

- Hablando particularmente de las series financieras, generalmente se puede hacer alguna **transformación** para convertirlas en estacionarias.
    - Los precios en niveles no suelen ser estacionarios, sin embargo sus **retornos** sí.
    - Las tasas de interés y los nodos de las curvas de descuento a veces no son estacionarios, pero sus **cambios** sí.
    - Las series macroeconómicas no suelen ser estacionarias (i.e la inflación, la tasa de desempleo, etc), pero **luego de controlar por estacionalidad** sí.

Índice

Función de autocorrelación serial (ACF y PACF)

- Cuando se habla de estacionariedad, una de los elementos más importantes por revisar es la correlación entre los rezagos, esta se denomina la **autocorrelación.**
- Existen dos análisis con respecto a este elemento:
  - Autocorrelación simple (ACF).
  - Autocorrelación parcial (PACF).

Autocorrelación:

En la mayoría de serie de tiempo los valores que toma una variable en un t particular, no son independientes entre sí, sino que un valor determinado depende de los valores anteriores, existen dos formas de medir esta dependencia de las variables.

1. Función de autocorrelación (ACF):

La autocorrelación mide la correlación entre dos variables separadas por k periodos.

$$\rho_j = corr(X_j, X_{j-k}) = \frac{cov(X_j, X_{j-k})}{\sqrt{V(X_j)}\sqrt{V(X_{j-k})}}$$

Si la serie es estacionaria se vería así:

$$\rho_1 = \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)} \qquad Y_t = \alpha_0 + \rho_1 Y_{t-1} + \epsilon_t$$

$$\rho_2 = \frac{Cov(Y_t, Y_{t-2})}{Var(Y_t)} \qquad Y_t = \alpha_0 + \rho_2 Y_{t-2} + \epsilon_t$$

- **No controla (no tiene en cuenta) el efecto que ya tuvo a través de rezagos más recientes.**

Autocorrelación Parcial (PACF)

Es similar a ACF pero busca eliminar el efecto de los valores intermedios.

Forma 1:

$$\pi_j = \frac{cov(X_j - \hat{X}_j, X_{j-k} - \hat{X}_{j-k})}{\sqrt{V(X_j - \hat{X}_j)}\sqrt{V(X_{j-k} - \hat{X}_{j-k})}}$$

Forma 2:

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_{k-1} x_{t-(k-1)} + u_t$$

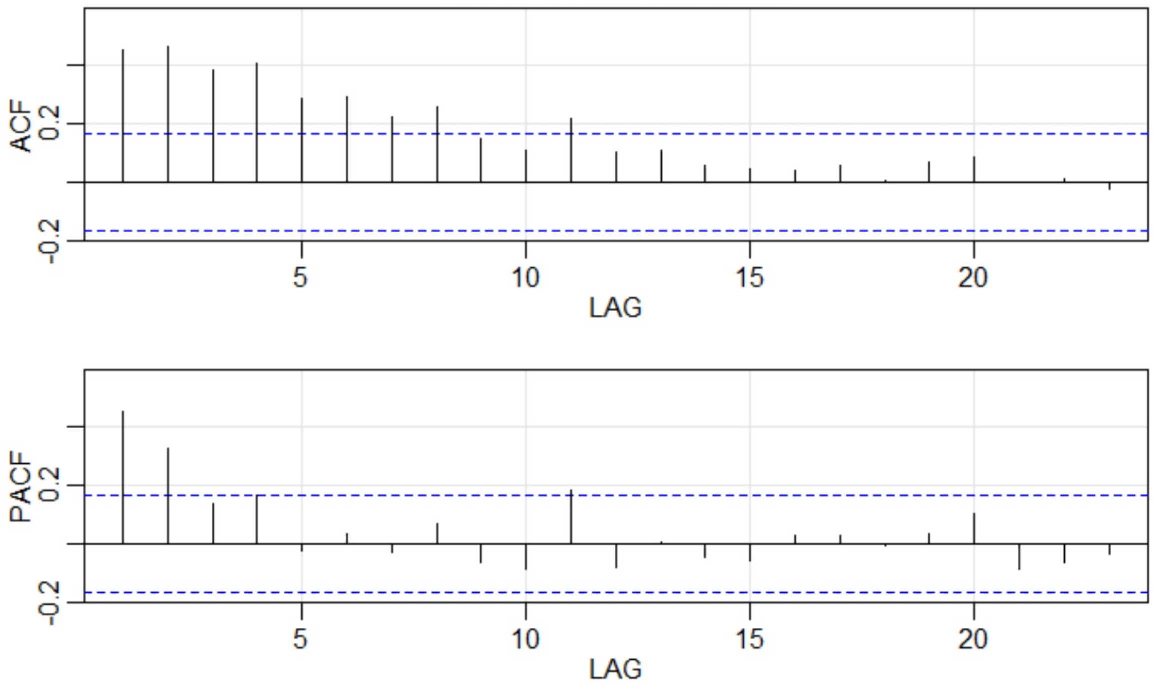
$$x_{t-k} = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_{k-1} x_{t-(k-1)} + v_t$$

$$\phi_{kk} = \frac{Cov(u_t v_t)}{\sqrt{Var(u_t)}\sqrt{Var(v_t)}}$$

El residuo significa todo lo que no es explicado por las variables dependientes. La intuición es comparar si existe correlación entre dos periodos controlando por todos los rezagos.

Ejemplo: si la tasa de interés de anteayer afectaba la de hoy **solo** por su efecto en la tasa de ayer, la autocorrelación simple lo marcaría como importante y la parcial no.

ACF y PACF



Todas las barras que terminen fuera de la banda de confianza significa que la serie temporal presenta auto correlación en el período indicado.

Cuando la ACF decrece lentamente o no entra a las bandas de contingencia el proceso tiende a ser no estacionario.

Índice

Pronóstico punto, intervalo y distribución

Función de autocorrelación serial (ACF y PACF).

Modelos AR, MA, AR-MA y AR(I)MA

Series de tiempo estacionales.

Criterios de error: RSME, MAE, MAPE

AR(p)

- Una aproximación natural al problema de predicción cuando se tiene una serie estacionaria, es suponer que los valores pasados tienen información de lo que debería ser la variable hoy.
- Es decir, si ayer la variable Y tuvo un valor alto, lo más probable es que hoy también lo tenga.
- Es más, el valor hace p días la variable Y debe tener algo de información del valor de hoy
- Terminando entonces con una especificación de este tipo:

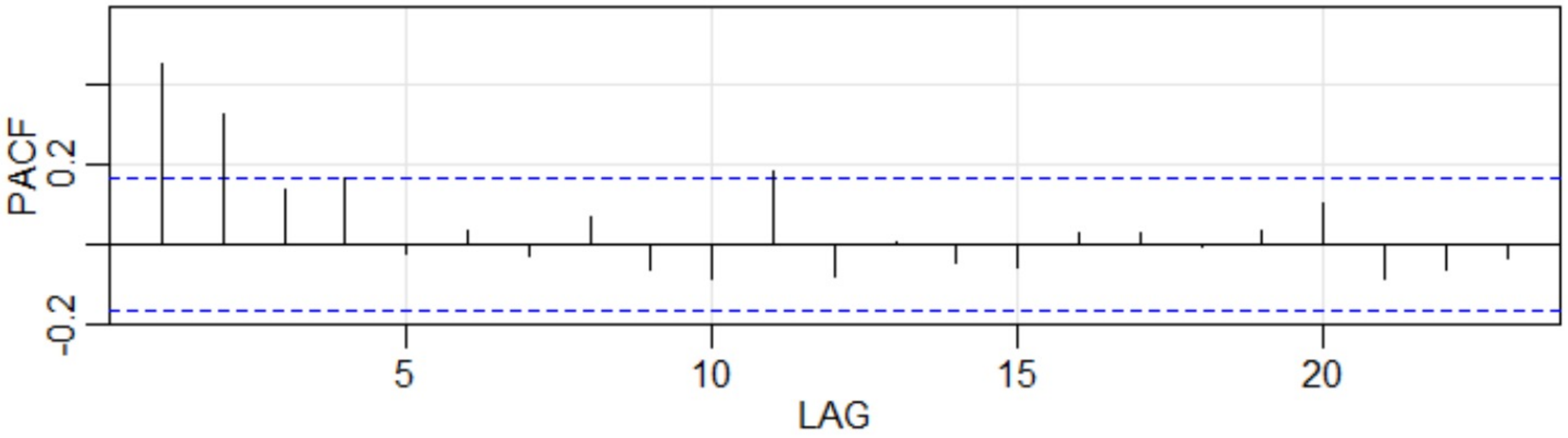
$$Y_t = \alpha_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t$$

- Donde ϵ_t se supone ruido blanco: $\epsilon_t \sim N(0, \sigma^2)$

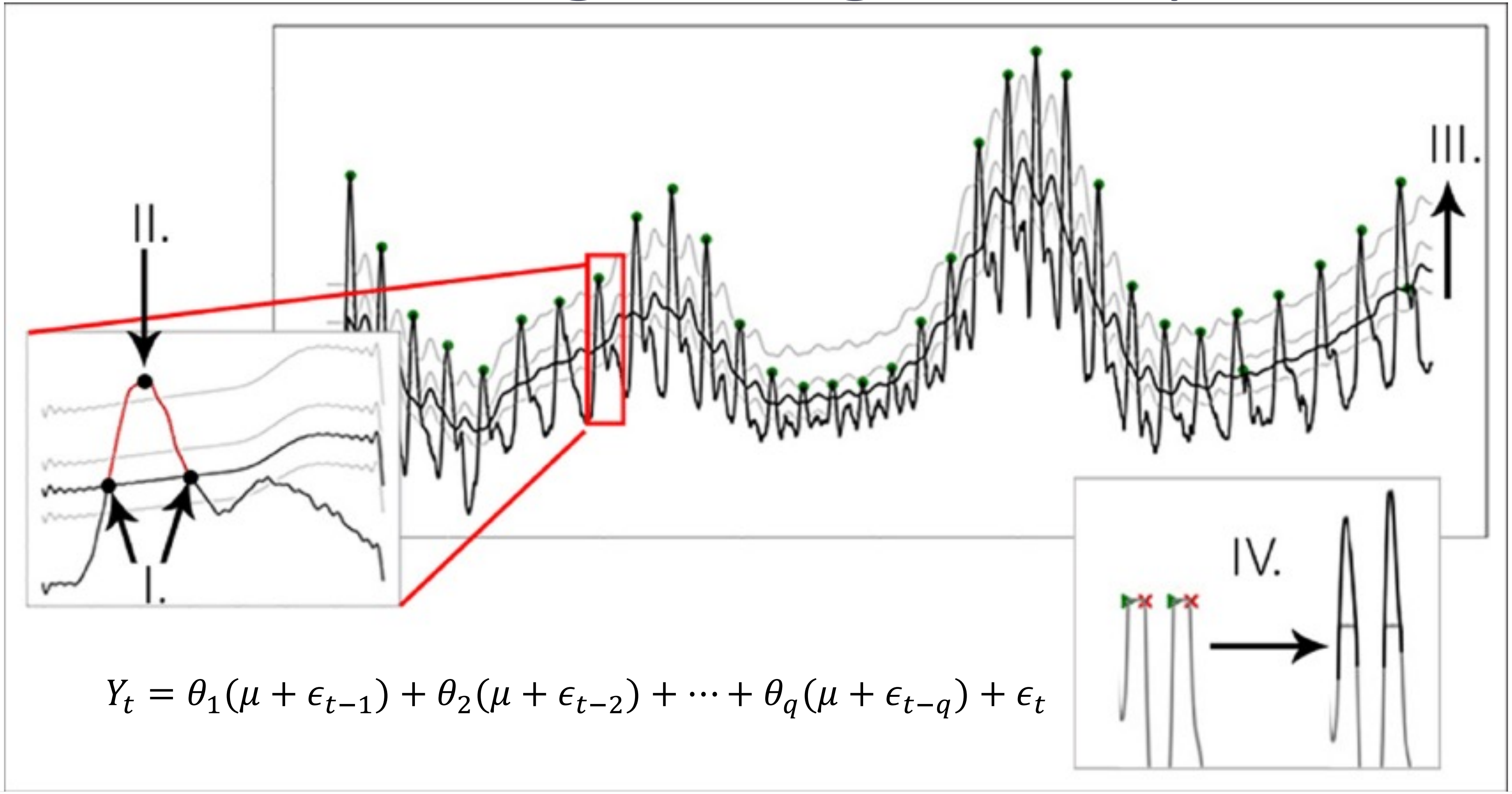
AR(p)

$$Y_t = \alpha_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t$$

- Note que esta es la misma especificación que tiene por detrás un PACF.
- En ese orden de ideas, este gráfico es informativo de cuántos rezagos debería considerar en una especificación AR(p)



Moving Average- MA(q)



$$Y_t = \theta_1(\mu + \epsilon_{t-1}) + \theta_2(\mu + \epsilon_{t-2}) + \cdots + \theta_q(\mu + \epsilon_{t-q}) + \epsilon_t$$

Moving Average MA(q)

- Si se tiene un proceso estacionario, la media incondicional no cambia entre periodos. (Constante)
- En ese orden de ideas, un promedio ponderado de los últimos q rezagos equivale a tomar el promedio incondicional (μ) y sumarle la desviación del rezago con respecto a este (θ_j).

$$Y_t = \theta_1(\mu + \epsilon_{t-1}) + \theta_2(\mu + \epsilon_{t-2}) + \cdots + \theta_q(\mu + \epsilon_{t-q}) + \epsilon_t$$

Factorizando

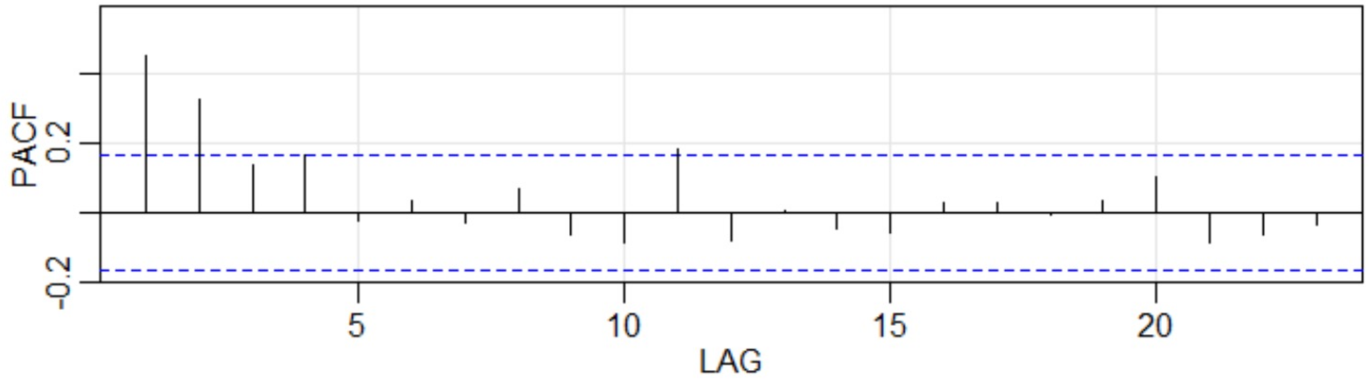
$$Y_t = \alpha_0 + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q} + \epsilon_t$$

Dónde $\alpha_0 = \mu \sum_{i=1}^{q} \theta_i$

MA(q)

$$Y_t = \alpha_0 + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q} + \epsilon_t$$

- Note que esta especificación está condicionando por la desviación del periodo pasado **pero no por su valor** Y_{t-q} particularmente. (Recordar que ϵ_t es ruido blanco).
- En este orden de ideas, intuitivamente se está verificando una **correlación no condicionada** en el valor del periodo anterior, tal como lo hace el ACF.
- Este gráfico entonces es informativo de cuántos rezagos q tomar.

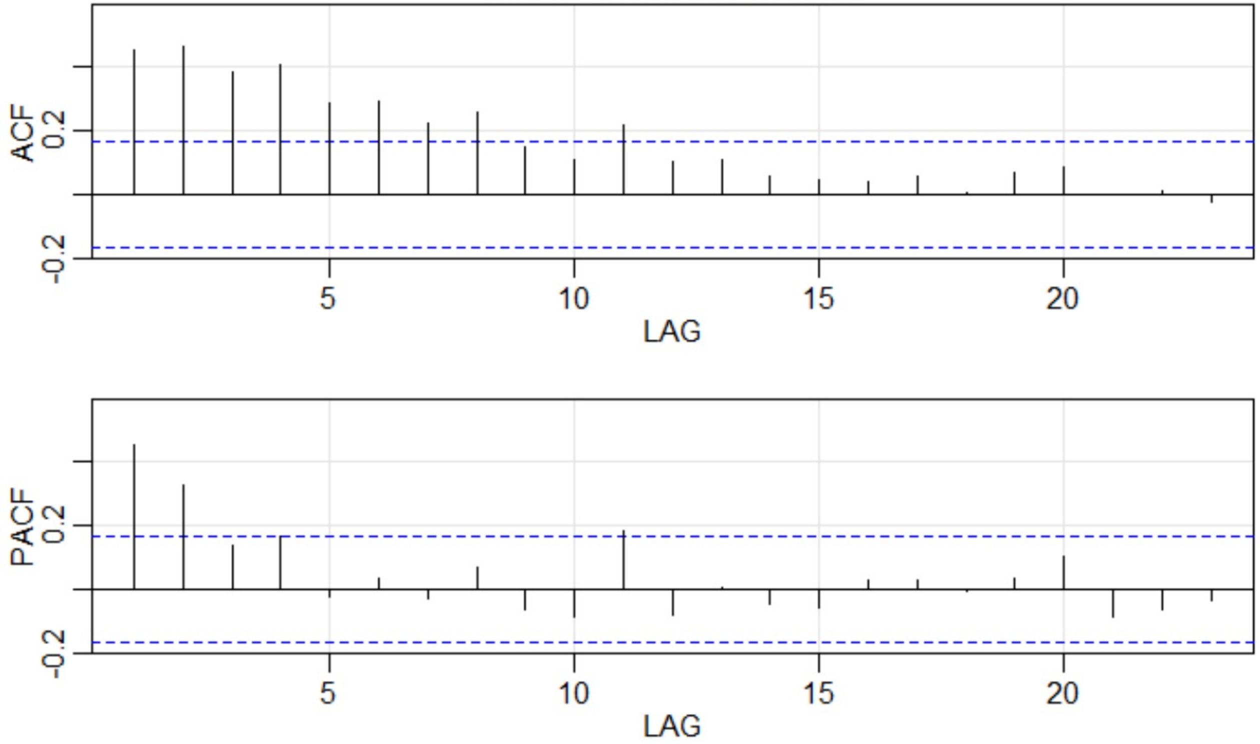


ARMA(p,q)

- Con el fin de aprovechar ambos procesos aparecen los modelos ARMA(p,q).
- Estos modelos controlan tanto por los rezagos de la variable como por las desviaciones de la misma.

$$Y_t = \alpha_0 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t$$

- Para escoger los rezagos de cada componente, se puede utilizar el ACF y PACF simultáneamente.



ARIMA(p,d,q)

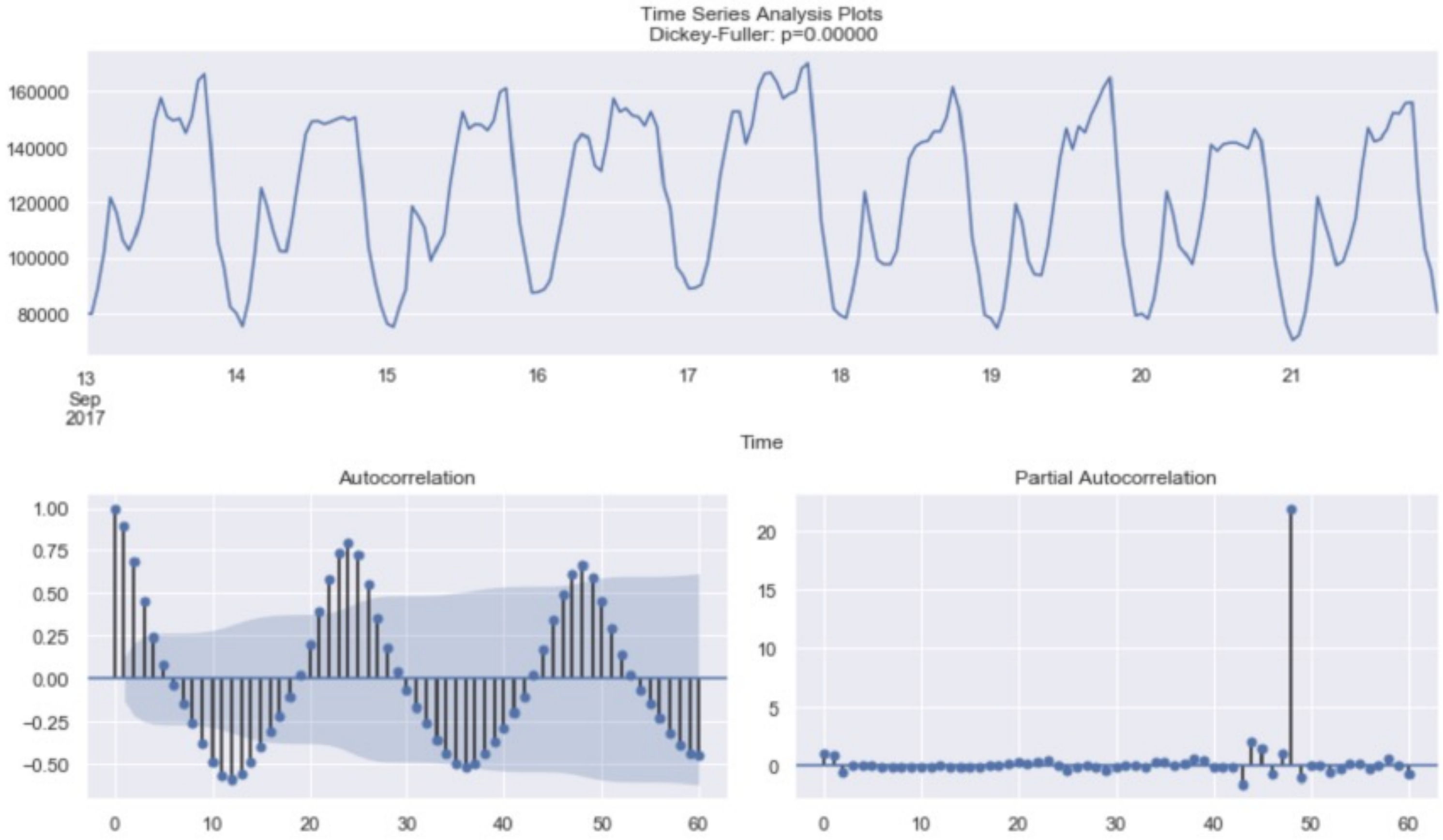
- Como se mencionó antes, no siempre las series de interés son estacionarias.
- Sin embargo, existen transformaciones de las mismas que sí lo son.
- En particular, se puede demostrar que una serie no estacionaria se convierte en estacionaria después de cierta cantidad de **diferenciaciones.**
- d hace referencia a la cantidad de veces que hay que diferenciar la serie original para convertirla en estacionaria.
- Se dice que una serie es **integrada** de orden d ($\mathrm{I}(d)$) cuando necesitó de esa cantidad de diferencias para que sea estacionaria. Una serie estacionaria de entrada es $\mathrm{I}(0)$.

Índice

$SARIMA(p,d,q)\ (P,D,Q)m$

- (p, d, q) hacen referencia a los componentes del ARIMA que ya estudiamos.
- $(P, D, Q)_m$ hacen referencia a los componentes estacionales.
  - D hace referencia a la cantidad de diferencias estacionales m que se tuvieron que hacer para convertir la serie en estacionaria.
  - En el ejemplo de la inflación, si el efecto a mitad de año daña la estacionariedad, necesitaría D=1 y m=6.
  - P,Q hace referencia a los componentes AR y MA estacionales significativos.
  - Si aun después de hacer las diferencias estacionales el comportamiento de la inflación hace 6 meses sigue siendo significativo en la predicción, debería esperar un P=1 y/o Q=1.

Ejemplos

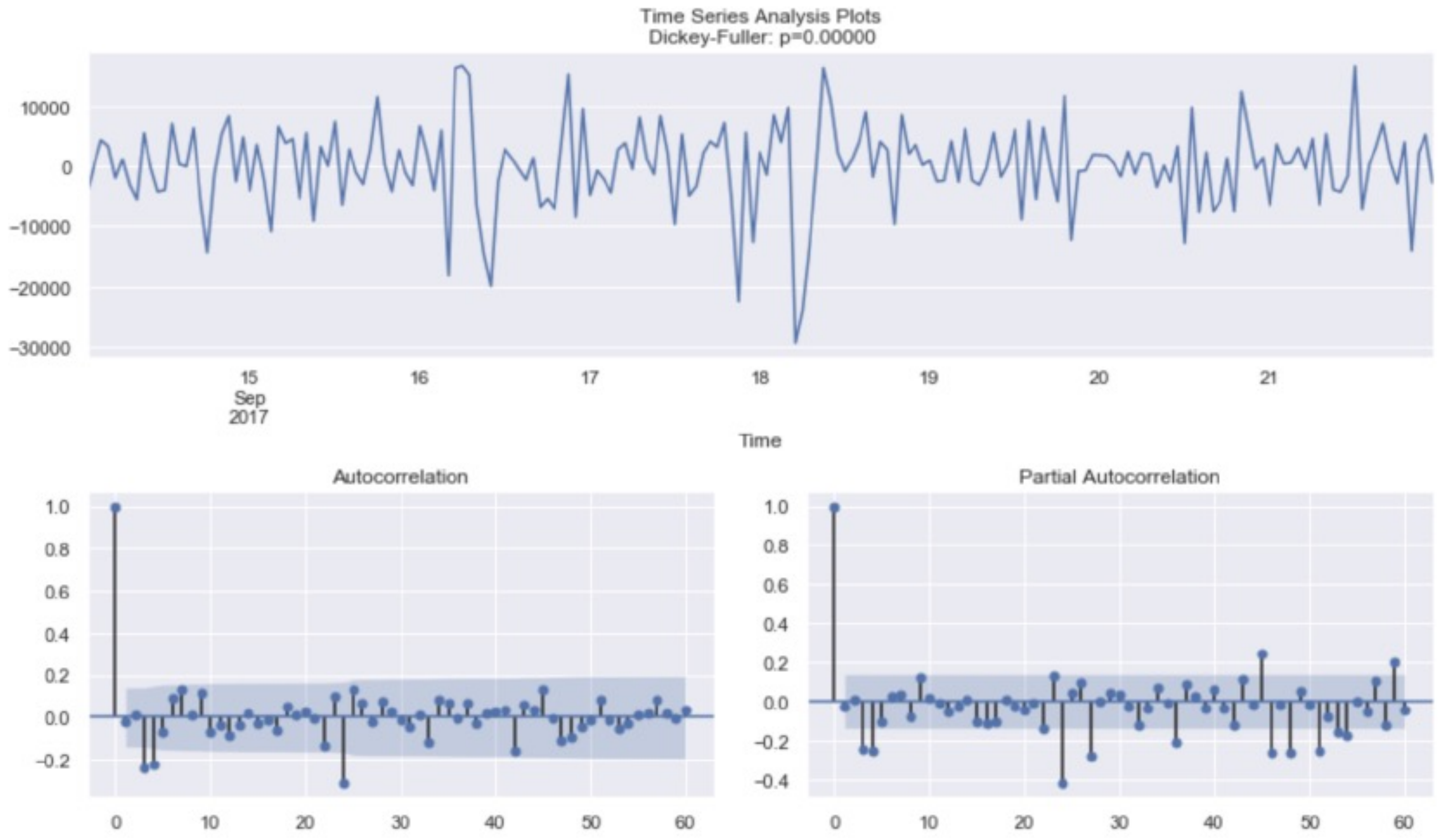


Time Series Analysis Plots
Dickey-Fuller: p=0.00000
160000
140000
120000
100000
80000
13 Sep 2017
14
15
16
17
18
19
20
21
Time
Autocorrelation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
0
10
20
30
40
50
60
Partial Autocorrelation
20
15
10
5
0
0
10
20
30
40
50
60

Ejemplos

Time Series Analysis Plots
Dickey-Fuller: p=0.04742
40000
20000
0
-20000
-40000
-60000
14
Sep
2017
15
16
17
18
19
20
21
Time
Autocorrelation
Partial Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
10
20
30
40
50
60

Ejemplos


Time Series Analysis Plots
Dickey-Fuller: p=0.00000
10000
0
−10000
−20000
−30000
15
Sep
2017
16
17
18
19
20
21
Time
Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0
10
20
30
40
50
60
Partial Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0
10
20
30
40
50
60

Ejemplos

- Los ejemplos son para motivar el hecho de que en la práctica no es obvio encontrar la mejor especificación ARIMA ni SARIMA para una serie particular, ni siquiera con la ayuda de los ACF y los PACF.
- Sin embargo, la revolución tecnológica permite probar muchas especificaciones en muy corto tiempo.
- En particular una de las funciones más famosas para esto es auto.arima en R.
- Esta permite encontrar la mejor especificación bajo el criterio AIC, automáticamente.

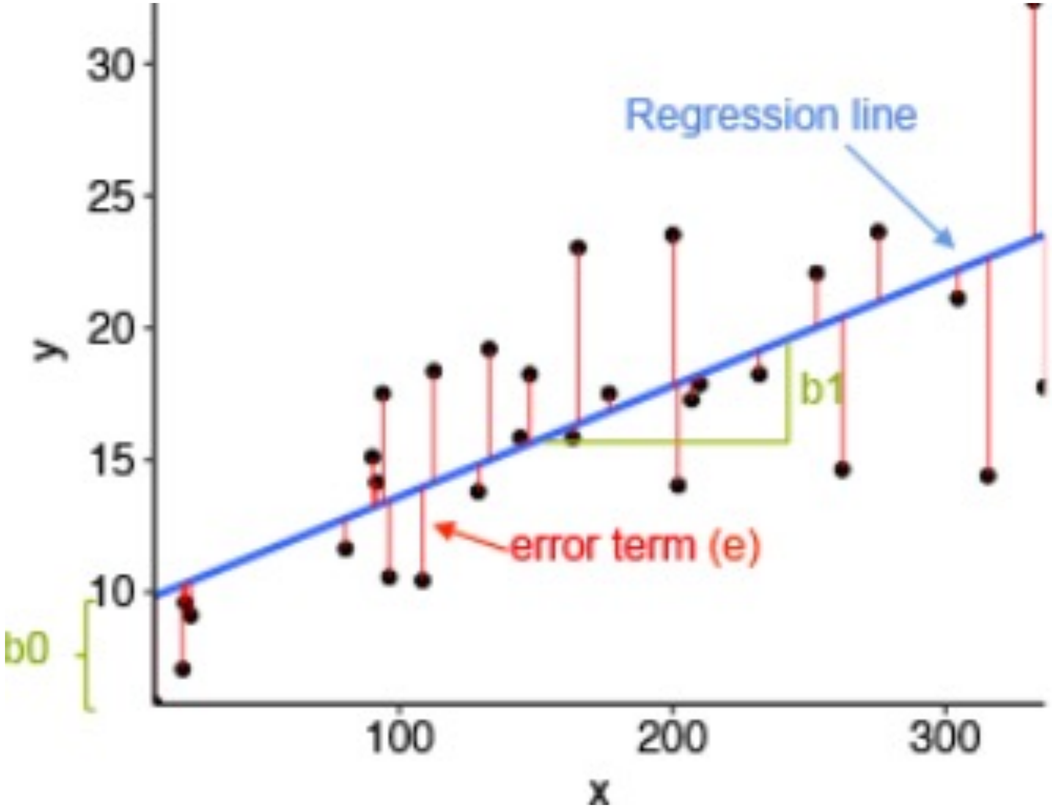
Universidad de los Andes

Índice

Métricas de desempeño

Error cuadrático medio (MSE): la métrica más utilizada que otorga una penalización más alta a los errores grandes y viceversa, $[0, \infty)$.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$



Métricas de desempeño

- **Raíz del error cuadrático medio (RMSE):**

$$\mathrm{RMSE} = \sqrt{MSE}$$

RMSE se puede interpretar como la desviación estándar del error de medición. (Raíz para tener unidades en mismo valor)

Métricas de desempeño

- **Error absoluto medio (MAE):** Métrica fácilmente interpretable dado que tiene las mismas unidades que la variable en cuestión. Muestra La diferencia absoluta entre la variable real y la estimada ajustando por el número de datos.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\, y_i - \hat{y}_i|$$

- **Error porcentual absoluto medio (MAPE):** Igual que el MAE pero en porcentaje con respecto a la variable de interés.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}$$

Se expresa la exactitud como un porcentaje del error. Por ejemplo, si el MAPE es 5, en promedio, el pronóstico está errado en un 5%.

Gracias

Universidad de los Andes

quantil

matemáticas aplicadas