Contenido

[Semana 3 1](#_Toc91194921)

[Series de tiempo 1](#_Toc91194922)

[Estacionariedad 2](#_Toc91194923)

[AR - Auto regressive 5](#_Toc91194924)

[MA-Medias móviles 7](#_Toc91194925)

[ARMA 8](#_Toc91194926)

[ARIMA 9](#_Toc91194927)

[ARIMA y regresión lineal 10](#_Toc91194928)

# Semana 3

## Series de tiempo

¿Qué es una serie de tiempo?

Una serie de tiempo es una serie de puntos de datos indexados en orden de tiempo.

Más comúnmente, una serie de tiempo es una secuencia de instantáneas de un proceso realizado en puntos sucesivos igualmente espaciados en el tiempo.

Por lo tanto, es una secuencia de datos de tiempo discreto.

Patrón de fondo

Descripción generada automáticamente

## Estacionariedad

¿Qué son los datos estacionarios?

"Estacionario" significa que la estructura estadística de la serie es independiente del tiempo.

¿Cómo sabemos que los datos son estacionarios?

● parcelas

● Estadísticas de resumen

● Pruebas estadísticas

¿Cómo convertir los datos de moda a los datos "estacionarios"?

Una forma de convertir los datos a los datos estacionarios es la diferencia o “medianisarlos”, a cada dato se le resta el anterior, pueden “diferenciarse” mas de una ves hasta que se encuentren centralizados

Interfaz de usuario gráfica, Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Después de diferenciarlo dos veces, estos datos ahora parecen estar "estacionarios". Puede revisarlo doble con una prueba de [Dickey-fuller aumentada](https://es.wikipedia.org/wiki/Prueba_de_Dickey-Fuller_aumentada)

¿Por qué la estacionariedad es importante en el modelado de la serie de tiempo?

La estacionaria permite la preservación de la estabilidad del modelo, es decir, un modelo cuyos parámetros y estructura se mantienen estables a lo largo del tiempo. La estacionariedad es importante porque proporciona un marco en el que se puede usar un promedio (utilizado en los procesos AR y MA) para describir adecuadamente el comportamiento de las series de tiempo.

¿Cuáles son las otras propiedades de los datos de la serie de tiempo?

La mayoría de las series horarias contienen uno o más de los siguientes componentes:

● tendencia

● estacional: Esta componente refleja comportamientos recurrentes, aunque no tienen por qué ser exactamente periódicos, con un periodo superior a un año. Muestran, habitualmente, cómo se suceden las etapas de bonanza económica con las de crisis, o al menos, desaceleración.

● cíclico: Este es el motivo por el que se consideran los periodos agrupados en otros periodos más amplios en el tratamiento de muchas series temporales. Muchas series económicas presentan oscilaciones regulares en el mismo mes de cada año, y con unas pautas que se presentan, sin repetirse exactamente, todos los años. Son las llamadas "variaciones estacionales", y se deben básicamente a causas climatológicas, vacacionales o fiscales.

● irregular (o residual): También llamado "ruido", recoge alteraciones de la serie, pequeñas en su incidencia, y sin una pauta periódica ni tendencial reconocible. Se considera que está ocasionada por múltiples factores, de pequeña entidad y diferentes ritmos temporales, que no se pueden estudiar individualmente. Esto en la teoría, porque en la práctica lo que ocurre es que la consideración de una serie como compuesta por componentes tendenciales, cíclicas y estacionales no deja de ser un modelo y, como tal, una representación aproximada e imperfecta, aunque valiosa, del mundo real. La componente irregular recogería, en consecuencia, la incapacidad del modelo para explicar a la perfección el comportamiento de la serie tempora

¿Cómo cambiamos los precios de las acciones para ser más estacionarios?

Los precios de las acciones suelen tener la tendencia (arriba o abajo). Pero en cualquier caso, tienen un cambio cambiante con el tiempo. Por lo tanto, debemos diferenciar los precios de las acciones para en ventanas diarias, mensuales o anuales para hacerlas "estacionarias"

## AR - Auto regressive

Una correlación de una variable con sí misma en diferentes períodos de tiempo en el pasado se conoce como "autocorrelación. Un proceso AR es donde se produce la autoregresión. Nuestro objetivo es encontrar el retraso del tiempo "correcto" que mejor capture el "orden" de un proceso de este AR. Este no es un procedimiento de un paso, pero es un proceso iterativo.

Aquí es cómo se ve un proceso de AR con un retraso de un período de tiempo de 1.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza baja

Observe la buena correlación para t=1, y la gran dispersión para t=6

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Aquí es cómo se ve un proceso de AR con un retraso de un período de tiempo de 2.

Texto

Descripción generada automáticamente con confianza baja

## MA-Medias móviles

Los modelos de promedio móvil (MA) no toman los valores “y” anteriores como entradas, sino que toman los términos de error anteriores. ¿Cómo obtenemos los términos de error si no predecimos?

Intentaremos predecir el siguiente valor Y basado en un promedio móvil (MA), así es como obtenemos los términos de error: ¿Qué tan equivocado se comparó nuestro promedio móvil con el valor real? Pero puede haber más de un término de error, ya que puede haber muchos promedios móviles.

¿Cómo saber cuantos errores usar?

El modelo se construye iterativamente al asumir cierto valor, q, para el número de términos de error anteriores, debemos ingresarlo en nuestro pronóstico para calcular nuestra predicción para valores futuros.

Tomando un error

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Tomando 2 errores

Imagen que contiene Gráfico

Descripción generada automáticamente

## ARMA

Toma los modelos AR y los Modelos MA y los combina

Interfaz de usuario gráfica, Aplicación, Word

Descripción generada automáticamente

El Lag "P", el término significa cuántos períodos de tiempo previos que cada observación está altamente correlacionada, lo que aumenta la "P" (LAG más larga) aumentaría aún más la dependencia de los valores anteriores.

El término "Q" significa cuántos períodos de tiempo previos que estamos considerando para observar cambios de tendencias repentinos.

## ARIMA

Vayamos un paso más allá y tratemos de agregar otro término al modelo ARMA, se llama ARIMA.

¿Qué significa el "I"? "I" significa integrado: estamos combinando las técnicas AR y MA en un solo modelo integrado: ARIMA. "I" ayuda a "estacionar" los datos. ARIMA necesitara un parámetro adicional “d”,. El término "D" adicional le dice a ARMA que ahora estamos prediciendo la diferencia entre un período anterior y el nuevo período, en lugar de predecir el valor del nuevo período.

D = 1 puede causar estacionariedad para un modelo

D = 2 puede capturar movimientos exponenciales en nuestra serie

(1,0,2) implica que cada término en el modelo se correlaciona con una salida anterior (P = 1) y dos términos de error previos pueden ser útiles para predecir el siguiente valor y (Q = 2). El cero en el medio implica que los datos ya están estacionarios (D = 0).

Evaluando nuestro modelo ARIMA.

Además, hay algunas tácticas clave que podemos explorar:

* trazar nuestros residuos
  + Si no observamos un patrón en nuestros términos de error residual, podemos detener el iteración
* Prueba de caja de ljung
  + Podemos probar matemáticamente el supuesto anterior usando la prueba de la caja Ljung

Gráfico

Descripción generada automáticamente

## ARIMA y regresión lineal

¿Qué tienen en común?

* Requiere estacionariedad lineal(Normalidad en regresión)
* Correlación entre la respuesta y las variables dependientes.
* Métodos para estimar los coeficientes.
* Pruebas estadísticas para evaluar la calidad.

¿Qué hay de diferencia?

● No hay orden natural a las observaciones en la regresión lineal.

● Regresión lineal utiliza dos variables diferentes.

● Regresión lineal enfatiza una variable depende del otro.

● Hay ordenamiento secuencial a las observaciones en series de tiempo.

● ARIMA utiliza la misma variable, con valores rezagados de la respuesta.

● ARIMA no necesita hacer esta elección de variables

Terminología de la serie de tiempo: Correlación automática (ARIMA) Elija:

● los datos

● d

● p y q

● la forma de arima

● El método de estimación.

● el mejor modelo

**Los datos**

* Paso más difícil
* En series temporales a menudo tienen demasiados datos en lugar de muy poco
* Necesita elegir la frecuencia de los datos que necesita para su modelo
* Lo más importante
* Comercialización entre la puntualidad y la relevancia.
* Cómo dividir para entrenamiento y prueba.
* Relevancia a la meta de predicción

**D**

* Si D = 0, estás trabajando datos estacionarios.
  + Precios
  + Rendimiento
* Si D = 1, estás trabajando con diferencias.
  + Devoluciones y devoluciones de registro
  + Cambios de rendimiento

**P y Q**

¡Estos tienen que ser enteros!

● ... pero puedes elegir configurarlos a cero

**La forma de ARIMA**

Variaciones

* estimación fraccionada
* estacionalidad
* El modelo tradicional puede simplificar demasiado el análisis.
* Modelo más complejo puede sobrecomplicar el análisis.

**La forma de estimación**

* Máxima verosimilitud
* Método de momentos.
* Métodos no paramétricos.
* Simulación

**El mejor modelo**

* Puede ejecutar múltiples formas de modelo con diferentes parámetros de estimación
* Puede usar un software diferente
* Elija el modelo con la mejor calidad de ajuste.

Sensibilidad de la estrategia comercial.

* ¿Cuál es la media?
  + Cada una de las 6 decisiones modelo puede afectar la media.
  + Elección de datos puede cambiar significa arriba o abajo
* ¿Hasta dónde se desvía normalmente de la media y cuánto tiempo se tarda en volver?
  + El promedio a largo plazo, la desviación estándar y el valor actual le indican dónde se encuentra en términos relativos y también ...
  + Tiempo esperado para revertir a la media
* ¿Cuál es el impacto de su elección de P y Q?
  + P y Q Dan la forma del modelo.
  + Alto P significa mercado tiene memoria alta
  + Low P significa que el pasado tiene un bajo impacto en el futuro.
* ¿Cuánta potencia explicativa está en el modelo?
  + ¡Si P y Q son cero, tu modelo es un paseo aleatorio!
  + Si su modelo se ajusta mejor por AR (1), entonces hay una estructura
  + La estrategia de negociación es sensible a las estimaciones del modelo que dependen de sus elecciones