



# **UNIVERSIDAD NACIONAL DE INGENIERÍA**

**ESCUELA CENTRAL DE POSGRADO**

**Facultad de Ingeniería Económica, Estadística y CCSS**

**Unidad de Posgrado - FIEECS**

**Curso: Forecasting for Data Science**

**Docente: Mg. Omar A. Chincaro Del C.**

# Forecasting for Data Science

1.- Introducción

2.- Componentes de una serie temporal

3.- Suavizado de un serie

4.- Series estacionarias:

Modelos Autorregresivos ( $AR(p)$ )

Modelos Medias Móviles ( $MA(q)$ )

Modelos Mixtos:  $ARMA(p,q)$

5.- Series no estacionarias

Modelos Mixtos:  $ARIMA(p,d,q)$

6.- Series estacionales: SMA, SAR, SARI, SIMA, SARIMA

7.- Series heterocedásticas: ARCH y GARCH

8.- Pronósticos

# INTRODUCCIÓN

En los últimos 30 años los modelos de series de tiempo se han convertido en metodologías poderosas para analizar conjuntos de datos, teniendo como propósito exclusivo la de realizar pronósticos. La característica principal de este conjunto de datos, es que se hayan observado puntos equidistantes a lo largo del tiempo y pueden provenir de las distintas áreas del conocimiento humano. La economía fue el área más privilegiada con el tratamiento de series de tiempo, su aplicación ha revolucionado a sus principales áreas: Macroeconomía, Microeconomía, Gestión, Comercio Exterior, Fianzas entre otras.

Actualmente el estudio de series de tiempo, abarcan desde estructuras más simples hasta las más complejas y se mezclan con otras áreas de estudio como en Análisis Bayesiano formándose los Modelos Dinámicos Bayesianos.

## Introducción

Las series temporales,  $Y_t$ , son observaciones de una variable a lo largo del tiempo. Por ejemplo, datos del paro nacional a lo largo del tiempo, ventas de una empresa durante un período de tiempo, evolución de la temperatura, etc.



# COMPONENTES DE UNA SERIE TEMPORAL

## Componentes de una serie temporal

Una serie temporal se puede descomponer en cuatro tipos de componentes:

- ▶ **Efecto estacional (EE):** se define como una conducta repetitiva a lo largo del tiempo, provocada por factores que se presentan de forma periódica (trimestralmente, semestralmente...) y que influyen en el comportamiento de la serie. La periodicidad del efecto estacional es inferior a un año. Por ejemplo, el consumo de energía presenta picos más altos en verano e invierno debido al consumo de los aires acondicionados y calefacciones.
- ▶ **Componente ciclica (CC):** algunas series muestran conductas repetitivas sin periodo fijo, debido a las fluctuaciones de la actividad social y económica. Por ejemplo, el resultado de la sucesión de las fases expansivas y recesivas de la economía. La periodicidad del efecto cíclico es superior a un año.

## Componentes de una serie temporal

- ▶ **Tendencia (T):** esta componente recoge el comportamiento de la serie a lo largo del tiempo. Es importante que la serie cuente con un gran número de observaciones para poder detectar la tendencia. El comportamiento de la tendencia puede ser lineal, exponencial, logístico... Por tanto, una serie tiene tendencia cuando hay una variación significativa en el valor medio de la serie. Mediante la tendencia se puede estudiar si la serie es **estacionaria** o **evolutiva**. Una serie es **estacionaria** si los datos fluctúan en torno a un valor constante, es decir, si no hay tendencia en la serie.



## Componentes de una serie temporal

- ▶ **Otras variaciones irregulares (VI):** fluctuaciones presentes en los datos no explicadas por la componente estacional ni por la tendencia. Por ejemplo, las huelgas, inundaciones son factores fortuitos que afectan de forma aislada y no permanente en una serie.

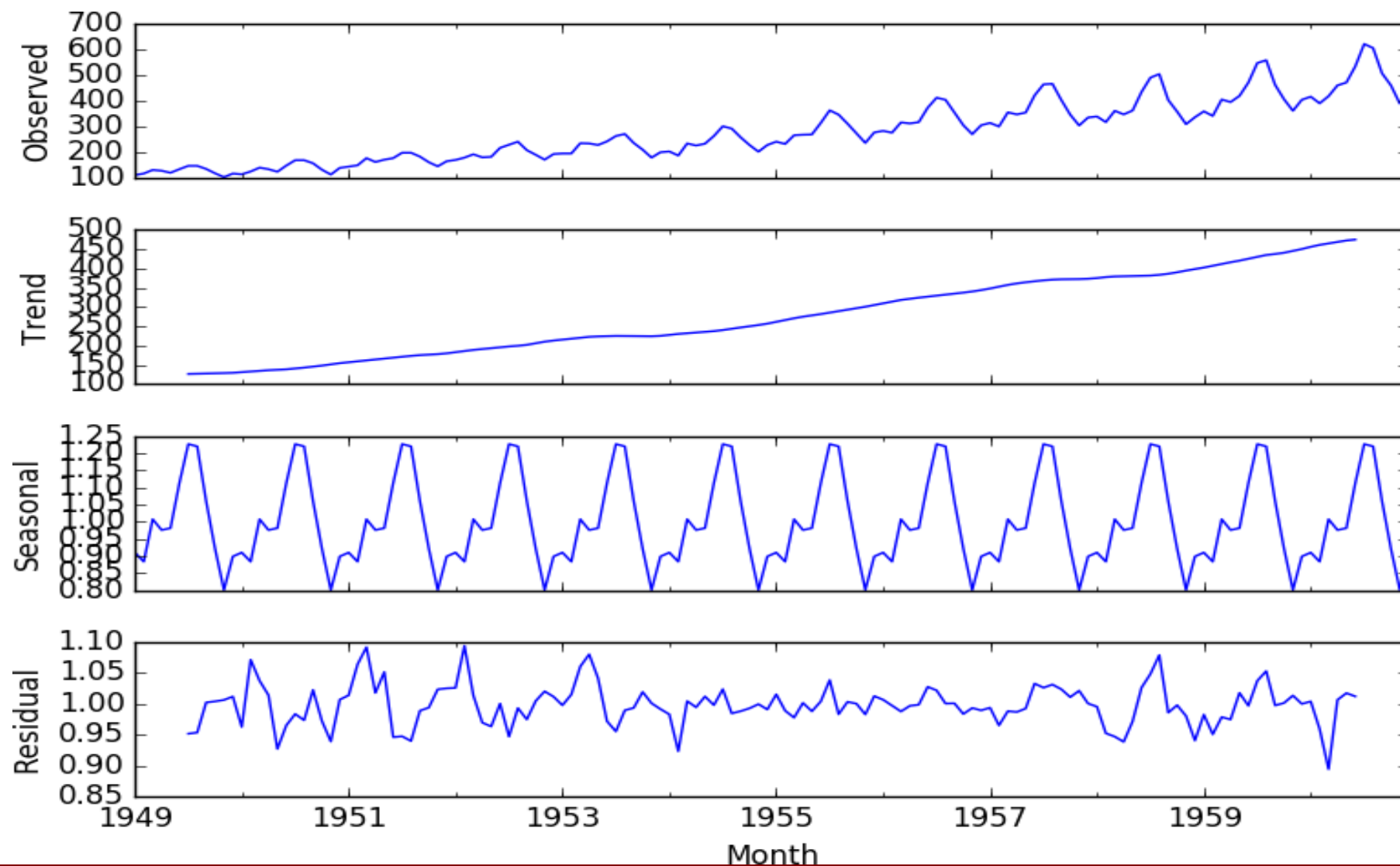
En resumen, una serie temporal se podr´a representar como

$$Y_t = f(EE_t, CC_t, T_t, VI_t), \quad t = 1, 2, \dots, n.$$

Dependiendo de la forma que tome  $f$ , los modelos de descomposición más frecuentes son:

- ▶ Modelo multiplicativo:  $Y_t = EE_t \cdot CC_t \cdot T_t \cdot VI_t$ .
- ▶ Modelo aditivo:  $Y_t = EE_t + CC_t + T_t + VI_t$ .
- ▶ Modelo mixto:  $Y_t = EE_t \cdot CC_t \cdot T_t + VI_t$ .

## Componentes de una serie temporal



# SUAVIZADO DE UNA SERIE TEMPORAL

# MEDIAS MOVILES

# Métodos de suavizados de la serie

## Medias Móviles (MA)

Un proceso que se encuentra dentro de todos los métodos de descomposición es el **suavizado de datos**. El suavizado de datos es una transformación de los datos para que la serie resultante contenga menos fluctuaciones que la original. El método más popular es el de medias móviles.

Este método consiste en promediar la serie. Se promedian los valores de la serie para periodos de tiempo fijos que se encuentran a lo largo de toda la serie.

Índice	Datos	q=3	q=5	q=4
1	28			
2	20	26,00		
3	30	26,00	25,40	25,63
4	28	26,33	24,40	25,13
5	21	24,00	24,00	24,00
6	23	20,67	23,00	22,13
7	18	22,00	22,20	22,13
8	25	22,33	24,00	23,38
9	24	26,33	24,60	25,25
10	30	26,67	26,60	26,63
11	26	28,00	25,60	26,50
12	28	24,67	24,20	24,38
13	20	21,67	22,20	22,00
14	17	19,00	21,00	20,25
15	20	19,00	19,20	19,13
16	20	19,67	18,60	19,00
17	19	18,67	19,40	19,13
18	17	19,00	21,00	20,25
19	21	22,00		
20	28			

# EXPONENCIAL

## Métodos de suavizados de la serie – Suavización exponencial

Los métodos de suavizado exponencial describen la serie temporal basándose en el estudio de las fluctuaciones que sufren las misma, sin crear un modelo matemático a partir de los datos de la serie. Aunque no se basan en ningún modelo, implícitamente sí que dependen de algún modelo sencillo debido a la naturaleza de los datos.

El objetivo que persiguen es eliminar o “suavizar” las fluctuaciones o movimientos aleatorios en los datos, aprovechando el comportamiento de la serie para predecir valores futuros.

Las predicciones que ofrece este método son medias de valores pasados, ponderando los datos según un conjunto desigual de pesos. Por ello, el nombre del método “suavizado exponencial” es porque, en general, los pesos decaen de forma exponencial desde el dato más reciente a las antiguo.

## Métodos de suavizados de la serie – Suavización exponencial

Los métodos de suavizado exponencial son:

- **Suavizado exponencial simple.** Se estima 1 parámetro. Adecuado para predecir valores futuros de series que no presentan ni componente estacional ni tendencia.
- **Suavizado exponencial doble:** método lineal de Holt. Se estiman 2 parámetros. Es adecuado para predecir valores futuros de series que presentan tendencia lineal.
- **Suavizado exponencial triple:** método de Holt-Winters. Se estiman 3 parámetros. Es adecuado para predecir valores futuros de series que presentan componente estacional.



## Métodos de suavizados de la serie – Suavización exponencial simple

Es un método iterativo que ofrece una serie suavizada  $\hat{Y}_t$  de la serie original :  $Y_t$

$$\hat{Y}_{t+1} = \alpha Y_{t+1} + (1 - \alpha) \hat{Y}_t \quad t = 1 \dots n$$

donde  $\alpha$  (parámetro de suavizado) es la constante con valores comprendidos entre 0 y 1.

Cuanto más cercano a 1 sea su valor, ....

Para obtener los pesos que se asignan a cada valor pasado de la serie, se debe de aplicar de forma repetitiva la ecuación anterior.

Por ello, la predicción en  $t+1$  es una media ponderada de todos los valores pasados.

## Métodos de suavizados de la serie – Suavización exponencial doble

Tendencia	Estacionalidad	Suavización Exponencial
No	No	Simple Un Parámetro (alfa)
Si	No	Doble (Brown, Holt) Dos Parámetros (alfa, beta)
Si	Si	Holt Winters Tres parámetros (alfa, beta, gamma)