

Inferència Estadística Actividad 3

Yolanda Colom Torrens

Introducción

Para esta actividad se explorará la dinámica de las series temporales. Para ello seleccionaremos una serie estableceremos su tipo y luego aplicaremos los métodos de predicción requeridos.

Tras realizar estos pasos pasaremos a comparar el periodo muestral y extra muestral corroborando cual es más eficiente u optimo entre ambos. Al elegir cual es más eficiente se utilizará este para predecir un año futuro.

Problema 1

En este caso tenemos los métodos A y B de los cuales deberemos calcular sus diferencias, además será necesario calcular la diferencia entre las predicciones y los valores reales de YT de ambos métodos.

Para ello se utiliza el método del error absoluto medio y error cuadrático medio (MAE, MSE), ya que con estos tenemos la capacidad de determinar que método es más exacto. El método con el MAE y MSE más bajo es el que ha hecho las mejores predicciones durante el periodo muestral. Para el periodo extra muestral, se debe repetir el proceso anterior, pero solo con los años en el periodo extra muestral.

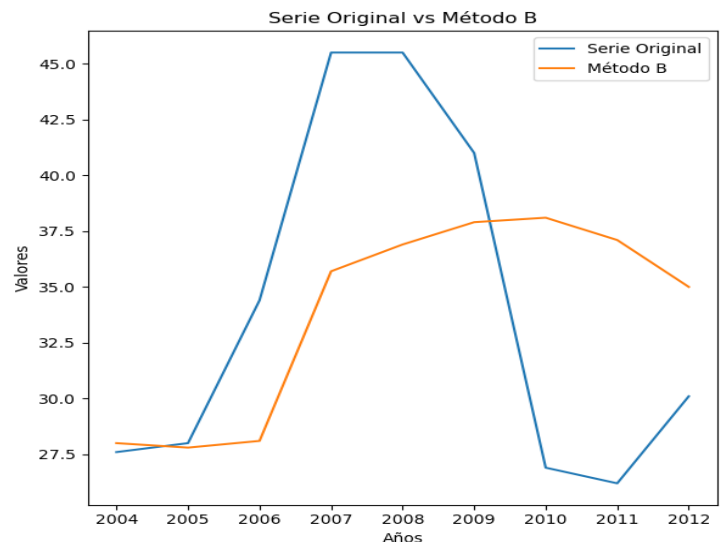
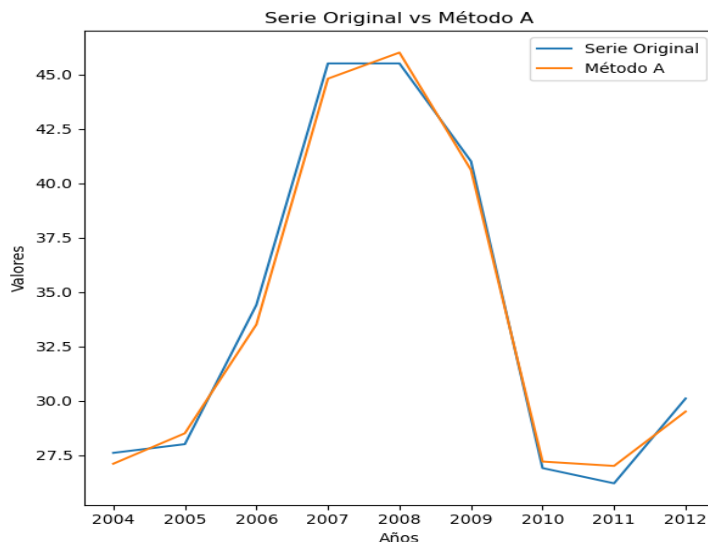
		Predicción	
	YT	Metodo A	Metodo B
2003	28	--	--
2004	27,6	27,1	28
2005	28	28,5	27,8
2006	34,4	33,5	28,1
2007	45,5	44,8	35,7
2008	45,5	46	36,9
2009	41	40,6	37,9
2010	26,9	27,2	38,1
2011	26,2	27	37,1
2012	30,1	29,5	35

Para efectos de este cálculo se utilizará los años en que ambos métodos hicieron predicciones ya que se pueden intersectar y calcular la diferencia absoluta, este resultado es mejor conocido como el error absoluto. Debido a la tabla de datos los años que utilizaremos son desde 2004 hasta 2012.

Se suman todos los errores absolutos y luego se dividen por el total de años, con esto obtendremos el MAE.

Con el cálculo del cuadrado de las diferencias entre los valores reales y las predicciones, se generan los errores cuadráticos, luego hacemos la sumatoria de estos y los dividimos por los años y obtenemos MSE.

MAE del Método A: 0.5777777777777781
MAE del Método B: 6.1555555555555555
MSE del Método A: 0.366666666666666703
MSE del Método B: 54.195555555555556



Problema 2

Para este caso a diferencia del anterior deberemos calcular las predicciones de series temporales con los métodos AES cuyos factores corresponden a 0.3 y 0.6. También se utilizará el método MM con ventanas de 2 y 3 por último, el método ingenuo y el método de media simple

Con el AES estaremos calculando el siguiente número de una serie temporal esto se basa en una ponderación entre el punto actual el cual utiliza un factor de suavizado que se mueven en un rango de 0 a 1 y la predicción que se calculó previamente entre más alto sea el factor más relevancia tiene el punto actual, de lo contrario el punto anterior es más relevante.

MM=Promedio (N ptos)

Método ingenuo = asume que el próximo punto es igual al último observado

Media simple = promedio de todos los puntos previos

```
1  import numpy as np
2  import pandas as pd
3  from statsmodels.tsa.holtwinters import SimpleExpSmoothing
4
5  # Datos de entrada
6  yt = [80.2, 78.0, 74.2, 76.5, 78.4, 72.6]
7
8  # AES(0.3) y AES(0.6)
9  model_03 = SimpleExpSmoothing(yt).fit(smoothing_level=0.3)
10 model_06 = SimpleExpSmoothing(yt).fit(smoothing_level=0.6)
11
12 pred_03 = model_03.predict(0, len(yt))
13 pred_06 = model_06.predict(0, len(yt))
14
15 # MM(2) y MM(3)
16 yt_series = pd.Series(yt)
17 pred_mm2 = yt_series.rolling(2).mean()
18 pred_mm3 = yt_series.rolling(3).mean()
19
20 # Método ingenuo
21 pred_ingenuo = [None] + yt[:-1]
22
23 # Método de la media simple
24 pred_media_simple = [np.mean(yt[:i]) for i in range(1, len(yt) + 1)]
25
26 # Función para calcular MAE
27 def calculate_mae(y_true, y_pred):
28     return np.mean(np.abs(np.array(y_true) - np.array(y_pred)))
29
30 # Calcular y mostrar el MAE para cada método
31 print("MAE AES(0.3): ", calculate_mae(yt, pred_03))
32 print("MAE AES(0.6): ", calculate_mae(yt, pred_06))
33 print("MAE MM(2): ", calculate_mae(yt[1:], pred_mm2[1:])) # Excluimos el primer valor
34 print("MAE MM(3): ", calculate_mae(yt[2:], pred_mm3[2:])) # Excluimos los primeros dos valores
35 print("MAE Método Ingenuo: ", calculate_mae(yt[1:], pred_ingenuo[1:])) # Excluimos el primer valor
36 print("MAE Método de la Media Simple: ", calculate_mae(yt[1:], pred_media_simple[1:])) # Excluimos el primer valor
```