

UNIDAD IX

Técnicas de Suavización

UNIDAD IX



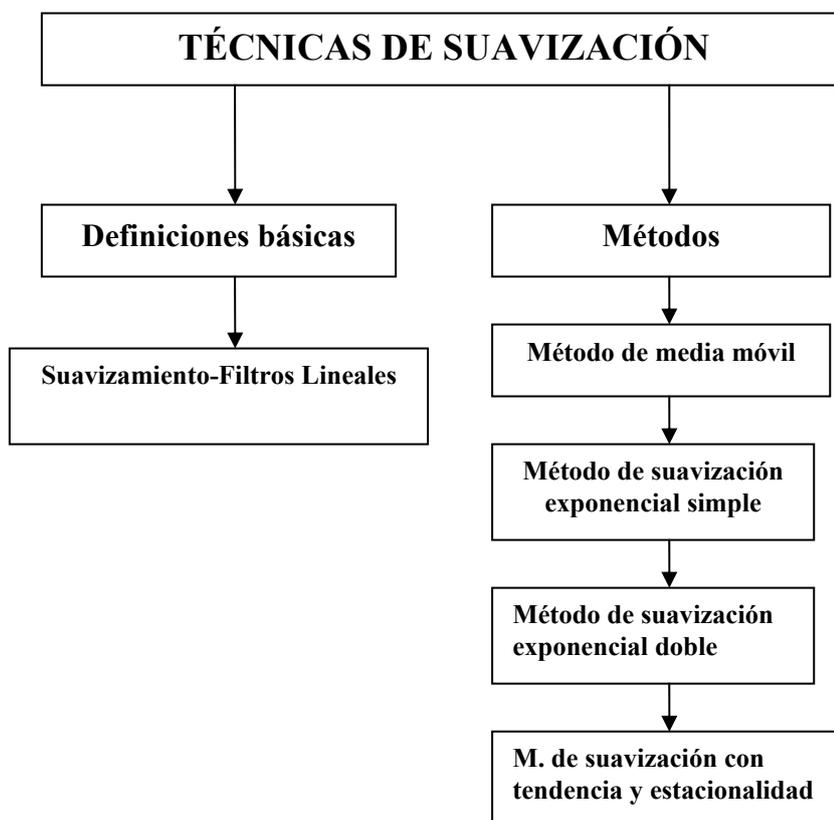
“La estadística demuestra que suele ser más fácil hacer algo bien que explicar por qué se hizo mal.”

Allen L. Webster, 1998

- **¿Cuál es el objetivo de la Técnica de suavización?**
- **¿Qué es un método de media móvil?**
- **¿Cómo se calcula el Método de suavización exponencial simple?**
- **¿Cuáles son las fórmulas del método de suavización exponencial doble (Método de Brown)?**
- **¿Cuáles son las ecuaciones que se utilizan para el Método de suavización con tendencias y estacionalidad (Método de Holt-Winters)?**

TÉCNICAS DE SUAVIZACIÓN

ESQUEMA CONCEPTUAL



COMPETENCIAS A LOGRAR

CONCEPTUAL	PROCEDIMENTAL	ACTITUDINAL
Explica las técnicas de suavización para alcanzar un buen nivel de predicción.	Aplica los métodos de suavización en las series de tiempo.	Analiza estas técnicas para una buena determinación de la dirección de la tendencia.

CONCEPTOS –CLAVE

Suavizamiento, pronóstico, exponencial, estacionalidad.

LECCIÓN 1

MÉTODO DE MEDIA MÓVIL

INTRODUCCIÓN

El objetivo de los métodos a usarse en esta unidad es “suavizar” a las fluctuaciones aleatorias causadas por el componente irregular de la serie. Estos métodos resultan apropiados para series estables, es decir, aquellas que no exhiban ningún comportamiento de tendencia, ni variaciones cíclicas ni estacionales, además es conveniente suavizar cuando existen cambios bruscos o movimientos irregulares en la serie

Son relativamente simples y generalmente alcanzan un buen nivel de predicción en períodos de tiempos cortos.

SUAVIZAMIENTO - FILTROS LINEALES

Una forma de visualizar la tendencia, es mediante el suavizamiento de la serie. La idea central es definir a partir de la serie observada un nueva serie que suavice los efectos ajenos a la tendencia (estacionalidad, efectos aleatorios), de manera que se pueda determinar la dirección de la tendencia (ver figura 9.1).

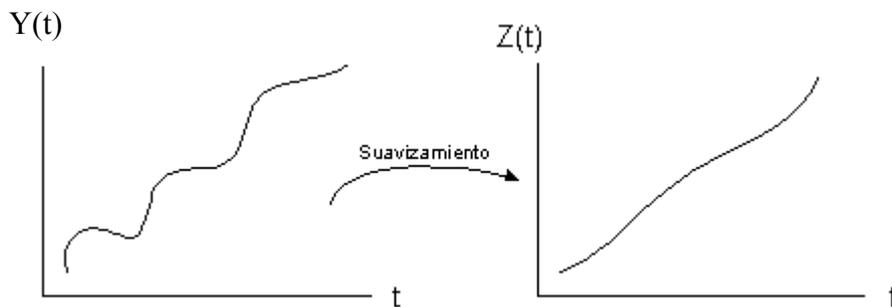


Figura 9.1

Lo que se hace es usar una expresión lineal que transforma la serie $Y(t)$ en una serie suavizada $Z(t)$: $Z(t) = F(Y(t))$, $t = 1, \dots, n$

$$Y(t) \longrightarrow \boxed{F} \longrightarrow T(t)$$

De tal modo que $F(Y(t)) = T(t)$; donde la función F se denomina Filtro Lineal, el filtro lineal más usado es la media móvil.

Método de Medias Móviles

“El objetivo es eliminar de la serie los componentes estacionales y accidentales”.

Promedios móviles simples (PMS)

Utiliza como pronóstico para el siguiente período, el promedio de los “n” valores de los datos más recientes de la serie de tiempo, matemáticamente puede expresarse como:

$$\text{Promedio Móvil} = \frac{\sum (\text{n valores de datos más recientes})}{n}$$

El término móvil indica que conforme se tenga disponible una nueva observación de la serie de tiempo, se reemplaza la observación más antigua en la ecuación y se calcula un nuevo pronóstico. Como resultado el promedio se modificará, a medida que se agreguen nuevas observaciones.

La variable **n** es una indicación de cuantos períodos habrán que tomarse para calcular el promedio, generalmente suele variar entre tres (3) a cinco (5), dependiendo de cuantos elementos tiene la serie de datos.

A continuación pasamos a ver una serie de variantes del método mencionado que se utilizan para eliminar el componente **irregular** y/o el componente **estacional**, los cuales se basan en la construcción de una serie de variables a partir de los datos originales, realizando para ello promedio de estos datos.

Casos:

1.-Media móvil anual o de orden 12 (con datos mensuales)

Consiste en sustituir el dato de cada mes por el promedio de los datos de los últimos 12 meses. Así, por ejemplo, para calcular la Media Móvil Anual de enero de 1997, calculamos el promedio de los doce meses comprendidos entre febrero de 1996 y enero de 1997, es decir:

$$M_t^{12} = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-11}}{12}$$

2.-Para una Serie Mensual con Estacionalidad Anual (s = 12)

La serie suavizada se obtiene de la forma siguiente

$$Y(k) = \frac{\frac{1}{2}Y(k-6) + Y(k-5) + \dots + Y(k+5) + \frac{1}{2}Y(k+6)}{12}, \quad 7 \leq k \leq n-6$$

3.- Para una serie trimestral, con estacionalidad anual (s = 4),

La serie suavizada está dada por

$$Y(k) = \frac{\frac{1}{2}Y(k-2) + Y(k-1) + Y(k) + Y(k+1) + \frac{1}{2}Y(k+2)}{4}, \quad 3 \leq k \leq n-2$$

A este procedimiento se le llama: filtro simétrico finito.

4.- Media Móvil de Orden “s”:

Podemos construir medias móviles de orden inferior a 12, es decir, que no sean anuales. Así, la Media Móvil de Orden 3 se construye promediando los tres últimos datos. De manera general, la Media Móvil de Orden s se construye de la siguiente forma:

$$M_t^s = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-s+1}}{s} = \frac{\sum_{i=1}^{s-1} y_{t-i}}{s}$$

Para realizar las predicciones y pronósticos de la serie utilizamos la siguiente fórmula:

$$\hat{y}_t = M_{t-1}^s.$$

La principal ventaja que presenta esta media móvil será que no retrasará tanto la evolución de la variable como lo hacía la media móvil anual, ya que al no utilizar tantas observaciones, su comportamiento se asemejará más al de la serie original.

Sin embargo, el principal inconveniente de esta media móvil viene precisamente de esa menor utilización de observaciones. Al no construirse con 12 observaciones, no se elimina completamente el componente estacional, llegando en ciertos casos a multiplicarse este efecto. Así, supongamos que trabajamos con una media móvil de orden 3 para la serie PBI, la cual tiene un componente estacional muy marcado (todos los meses de agosto son sistemáticamente más bajos que el resto de los meses del año). Al construir esta media móvil de orden 3, va a ocurrir que el dato sistemáticamente más bajo de agosto en la serie original, no sólo afecte al dato de la media móvil de agosto sino también lo hará al dato de septiembre y al dato de octubre.

5. Media Móvil Centrada:

Esta media móvil no se construye promediando los datos anteriores al dato original, sino que se utilizan simétricamente los datos adyacentes. Por ejemplo, para construir el dato de julio de 1996 de la media centrada de orden 3 se utiliza el promedio de los meses de junio, julio y agosto de 1996. En general, la Media Móvil Centrada de orden “s” va a estar definida por la siguiente expresión:

$$MC_t^s = \frac{y_{t+\frac{(s-1)}{2}} + \dots + y_{t+1} + y_t + y_{t-1} + \dots + y_{t-\frac{s-1}{2}}}{s}$$

Cuando el orden de la media móvil centrada es par (por ejemplo 6) se utilizan para calcularla “s+1” observaciones, ponderando cada uno de los datos extremos con un valor igual a 0.5:

$$MC_t^6 = \frac{0.5y_{t+3} + y_{t+2} + y_{t+1} + y_t + y_{t-1} + y_{t-2} + 0.5y_{t-3}}{6}$$

La principal ventaja de esta media móvil es que al utilizar los datos adyacentes, tanto anteriores como posteriores, no vamos a tener el problema de retrasar la evolución de la variable, es decir, tanto la serie original como la media móvil centrada van a tener en las mismas fechas los valores máximos y mínimos.

El principal inconveniente será que al utilizar las observaciones posteriores a cada dato original, para su construcción, no tendremos observaciones de la MC al final de la muestra disponible, siendo ésta precisamente la parte de la muestra más interesante para el análisis. Esta dificultad se resuelve si disponemos de predicciones fiables para los próximos meses. Así, por ejemplo si se dispone de datos hasta diciembre de 2002, la MC^5 para ese mes de diciembre de 2002 se formará con el promedio de las observaciones de octubre, noviembre y diciembre de 2002 y las predicciones para enero y febrero de 2003.

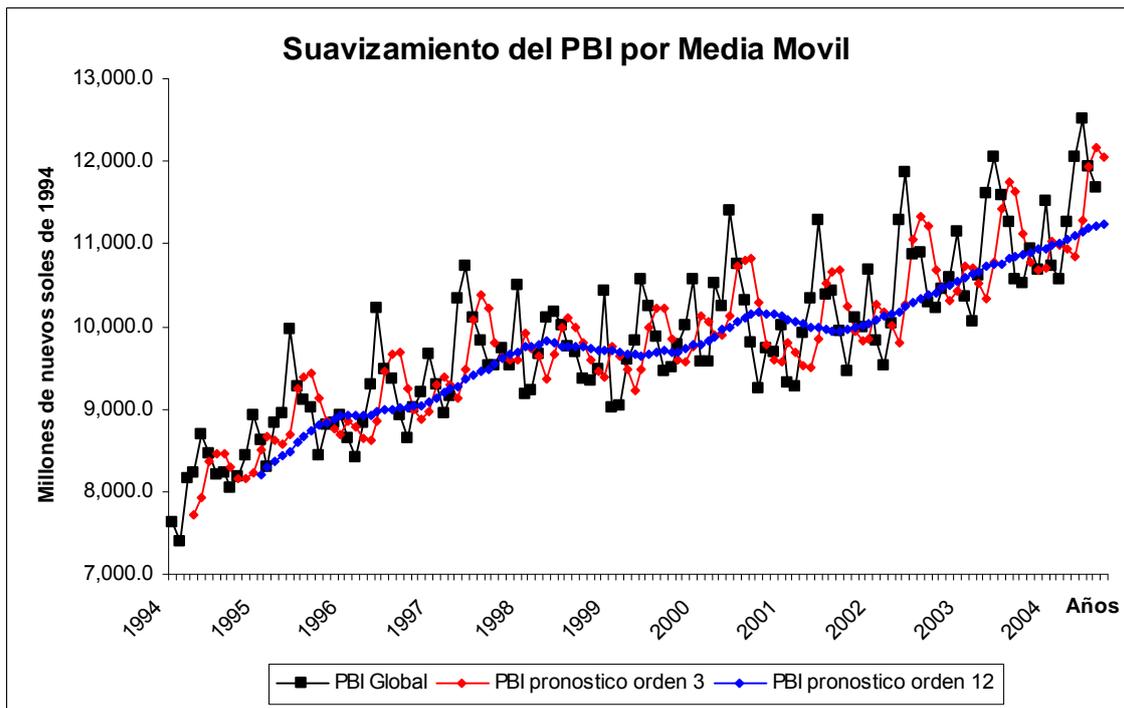
Resumiendo, cada una de las diferentes medias móviles va a tener sus ventajas como también sus desventajas y dependiendo de para que querramos utilizarlas y las características de la serie en cuestión, emplearemos a una o a otra. Por ejemplo, si la serie en estudio no tiene componente estacional muy marcado y es fácil conseguir predicciones fiables, podría utilizarse una MC^3 . En otros casos, sería más conveniente utilizar una media móvil de mayor orden (o anual) o no centrada, etc. Normalmente, estas medias móviles son interpretadas como una aproximación a la evolución de la serie temporal y, de hecho, en ocasiones se utilizan como estimación de la tendencia de la misma. Hay que destacar que esta tendencia no va a tener las limitaciones que tenían las tendencias deterministas, las cuales dependían únicamente del tiempo. Cabe mencionar que estas medias móviles tienen una naturaleza cambiante con el paso del tiempo, en función del comportamiento de la variable.

Nota: Se suaviza cuando existen muchos cambios bruscos o movimientos irregulares.

Ejemplo Ilustrativo 1:

A continuación se presentarán los datos correspondientes al PBI mensual, en millones de nuevos soles a precios de 1994, para el periodo enero 1994-julio 2004

Año	Mes	Valor del PBI	Media Movil Centrada de orden 3	Media Asimetrica de orden 3	Pronostico PBI	Media Movil Centrada de Orden 12	Media Movil Asimetrica de orden 12	Pronostico PBI
1994	Enero	7,615.2						
	Febrero	7,405.3	7,725.3					
	Marzo	8,155.6	7,932.8	7,725.3				
	Abril	8,237.6	8,364.6	7,932.8	7,725.3			
	Mayo	8,700.6	8,466.4	8,364.6	7,932.8			
	Junio	8,460.9	8,452.7	8,466.4	8,364.6			
	Julio	8,196.6	8,297.5	8,452.7	8,466.4	8,256.88		
	Agosto	8,235.1	8,154.9	8,297.5	8,452.7	8,335.71		
	Setiembre	8,033.1	8,147.5	8,154.9	8,297.5	8,400.30		
	Octubre	8,174.3	8,217.5	8,147.5	8,154.9	8,457.99		
	Noviembre	8,445.0	8,512.5	8,217.5	8,147.5	8,540.21		
	Diciembre	8,918.3	8,662.9	8,512.5	8,217.5	8,625.84	8214.79	
1995	Enero	8,625.4	8,610.3	8,662.9	8,512.5	8,697.03	8298.97	8214.79
	Febrero	8,287.1	8,578.8	8,610.3	8,662.9	8,767.74	8372.45	8298.97
	Marzo	8,823.9	8,688.3	8,578.8	8,610.3	8,817.75	8428.15	8372.45
	Abril	8,953.9	9,245.1	8,688.3	8,578.8	8,861.29	8487.83	8428.15
	Mayo	9,957.7	9,390.2	9,245.1	8,688.3	8,903.67	8592.59	8487.83
	Junio	9,259.0	9,441.3	9,390.2	9,245.1	8,919.83	8659.09	8592.59
	Julio	9,107.1	9,129.2	9,441.3	9,390.2	8,920.25	8734.97	8659.09
	Agosto	9,021.4	8,858.5	9,129.2	9,441.3	8,925.91	8800.50	8734.97
	Setiembre	8,447.0	8,757.9	8,858.5	9,129.2	8,931.86	8834.99	8800.50
	Octubre	8,805.4	8,694.5	8,757.9	8,858.5	8,946.61	8887.58	8834.99
	Noviembre	8,831.1	8,852.2	8,694.5	8,757.9	8,971.62	8919.76	8887.58
	Diciembre	8,920.0	8,794.9	8,852.2	8,694.5	8,991.65	8919.90	8919.76
...
2004	Enero	10,733.6	10,942.9	10,975.6	11,041.8	11,229.10	11010.86	10979.73
	Febrero	10,576.7	10,855.1	10,942.9	10,975.6		11054.04	11010.86
	Marzo	11,255.0	11,292.8	10,855.1	10,942.9		11106.69	11054.04
	Abril	12,046.6	11,940.5	11,292.8	10,855.1		11142.52	11106.69
	Mayo	12,519.8	12,164.2	11,940.5	11,292.8		11182.35	11142.52
	Junio	11,926.2	12,041.0	12,164.2	11,940.5		11211.51	11182.35
	Julio	11,677.0		12,041.0	12,164.2		11246.69	11211.51
2004	Agosto				12,041.0			11246.69



Como se puede observar, a medida que aumenta el orden, mayor es el efecto de suavización, pero también aumenta la pérdida de información.

EXACTITUD DEL PRONÓSTICO

Una consideración de gran importancia al seleccionar un método de pronóstico es la exactitud del mismo. Evidentemente se desea que los errores de pronóstico sean lo más pequeños posibles. Se suele tomar como medida de error general el valor de **Error Medio Cuadrado** (EMC) o Mean Square Error (MSE) de la bibliografía inglesa.

El EMC se calcula de la siguiente manera:

$$EMC = \frac{\sum((\text{Valor Real} - \text{Valor Pronosticado})^2)}{\text{Cantidad Valores Pronosticados}}$$

Ejemplo Ilustrativo 2:

Sean las ventas semanales de nafta (en miles de litros) de un establecimiento la variable que se desea pronosticar; utilizando el método de Media móvil de orden 3, obtenemos los resultados del siguiente cuadro:

Semana	Ventas y_t	Promedio móvil M_t^3	Predicciones \hat{y}_t	Error $\mu=(y_t - \hat{y}_t)$	Error Cuadrado (μ^2)
1	17				
2	21				
3	19	19			
4	23	21	19	4	16
5	18	20	21	-3	9
6	16	19	20	-4	16
7	20	18	19	1	1
8	18	18	18	0	0
9	22	20	18	4	16
10	20	20	20	0	0
11	15	19	20	-5	25
12	22	19	19	3	9
13	PRONOSTICO		19	0	10,22

Media móvil:

$$M_t^3 = (y_t + y_{t-1} + y_{t-2})/3$$

Por ejemplo:

$$M_3^3 = (19+21+17)/3 = 57/3 = 19$$

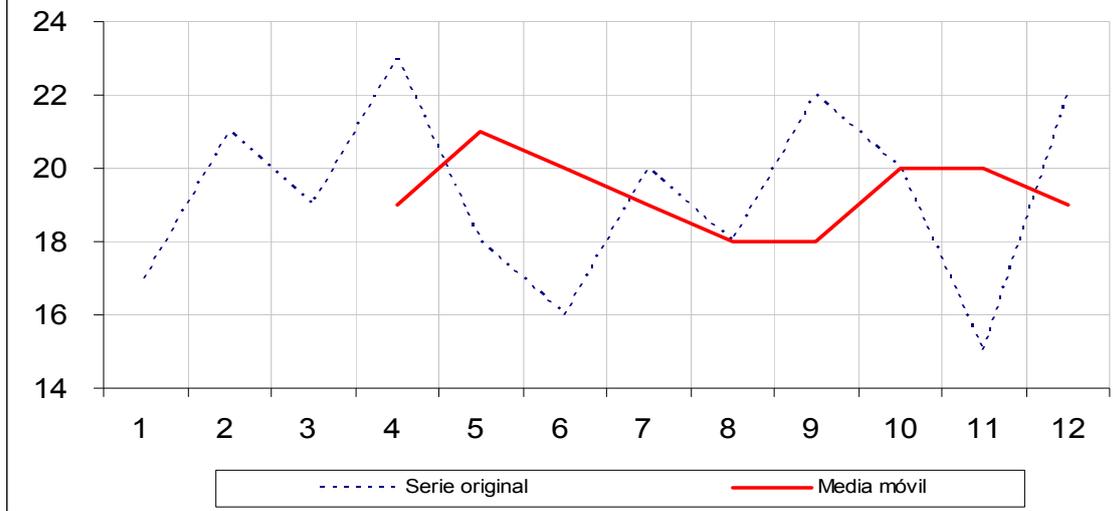
Predicciones:

En este caso se dará de la forma: $\hat{y}_t = M_{t-1}^3$, por ejemplo para el caso de semana 4 se tendrá:

$$\hat{y}_4 = M_{4-1}^3 = M_3^3 = 19$$

La semana 13 es el valor que deseamos pronosticar. En este caso 19 mil litros de nafta podrían llegar a venderse. El gráfico siguiente resume los valores reales y los predichos por el modelo de **Promedios Móviles**

**VENTAS SEMANALES DE NAFTA REALES Y SUAVIZADOS
POR EL MÉTODO DE MEDIA MÓVIL SIMPLE (s=3)**



LECCIÓN 2

MÉTODO DE SUAVIZACIÓN EXPONENCIAL SIMPLE

Esta técnica se basa en la atenuación de los valores de la serie de tiempo, obteniendo el promedio de estos de manera exponencial; es decir, los datos se ponderan dando un mayor peso a las observaciones más recientes y uno menor a las más antiguas. Al peso para ponderar la observación más reciente se le da el valor α , la observación inmediata anterior se pondera con un peso de $(1 - \alpha)$, a la siguiente observación inmediata anterior se le da un peso de ponderación de $(1 - \alpha)^2$ y así sucesivamente hasta completar el número de valores observados en la serie de tiempo a tomar en cuenta para realizar la atenuación, es decir, para calcular el promedio ponderado. La estimación o pronóstico será el valor obtenido del cálculo del promedio.

Por lo que la expresión para realizar el cálculo de la suavización exponencial simple es:

$$P_{t+1} = \alpha Y_t + \alpha(\alpha - 1)Y_{t-1} + \alpha(\alpha - 1)^2 Y_{t-2} + \dots \alpha(\alpha - 1)^{n-1} Y_{t-(n-1)}$$

Otra expresión equivalente a esta es la siguiente:

$$P_{t+1} = \alpha Y_t + (\alpha - 1)P_t$$

Donde

- Y_t : Valor de la serie en el período “t”.
- P_{t+1} : Pronóstico o predicción para el período “t+1”.
- P_t : Pronóstico o predicción en el período “t”.
- α : Factor de suavización, ($0 \leq \alpha \leq 1$)

Es decir el valor de la serie suavizada en el período “t+1” es igual a “ α ” veces el valor de la serie en el período “t”, más “ $1-\alpha$ ” veces el valor predicho en el período “t”.

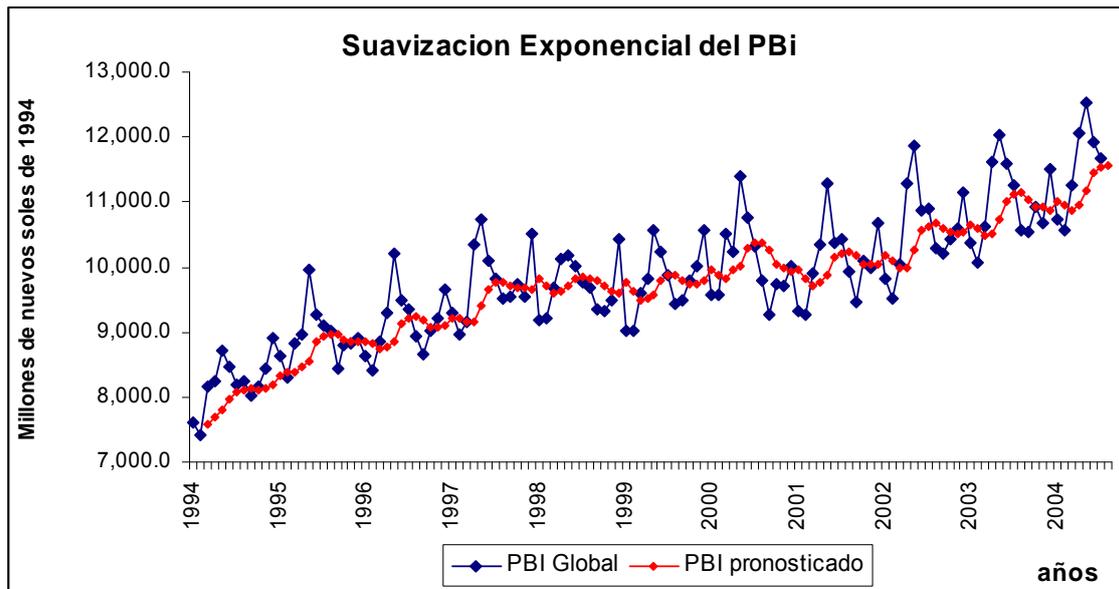
Es así que para determinar los valores de la serie suavizada se necesita un valor inicial P_0 , el cual puede ser un promedio de los datos anteriores o simplemente el primer valor de la serie.

Ejemplo Ilustrativo 1:

Continuando con la serie de tiempo del PBI Global.

La constante de suavización que se utilizara será de 0.2 y el valor inicial será la primera observación, es decir, el valor del PBI en enero de 1994.

Año	Mes	Valor del PBI	Suavizacion exponencial	Pronostico PBI
1994	Enero	7,615.2	7,615.2	
	Febrero	7,405.3	7,573.2	
	Marzo	8,155.6	7,689.7	7,573.2
	Abril	8,237.6	7,799.3	7,689.7
	Mayo	8,700.6	7,979.5	7,799.3
	Junio	8,460.9	8,075.8	7,979.5
	Julio	8,196.6	8,100.0	8,075.8
	Agosto	8,235.1	8,127.0	8,100.0
	Setiembre	8,033.1	8,108.2	8,127.0
	Octubre	8,174.3	8,121.4	8,108.2
	Noviembre	8,445.0	8,186.1	8,121.4
	Diciembre	8,918.3	8,332.6	8,186.1
1995	Enero	8,625.4	8,391.1	8,332.6
	Febrero	8,287.1	8,370.3	8,391.1
	Marzo	8,823.9	8,461.0	8,370.3
	Abril	8,953.9	8,559.6	8,461.0
	Mayo	9,957.7	8,839.2	8,559.6
	Junio	9,259.0	8,923.2	8,839.2
	Julio	9,107.1	8,960.0	8,923.2
	Agosto	9,021.4	8,972.2	8,960.0
	Setiembre	8,447.0	8,867.2	8,972.2
	Octubre	8,805.4	8,854.8	8,867.2
	Noviembre	8,831.1	8,850.1	8,854.8
	Diciembre	8,920.0	8,864.1	8,850.1
....	
2004	Enero	10,733.6	10,951.2	11,005.6
	Febrero	10,576.7	10,876.3	10,951.2
	Marzo	11,255.0	10,952.1	10,876.3
	Abril	12,046.6	11,171.0	10,952.1
	Mayo	12,519.8	11,440.7	11,171.0
	Junio	11,926.2	11,537.8	11,440.7
	Julio	11,677.0	11,565.7	11,537.8
2004	Agosto			11,565.7



Se puede observar que por método de suavización exponencial simple, suaviza de mejor manera que el método de media móvil. Empleando este método obtenemos que para agosto del 2004 el PBI Global pronosticado es de 11,565.7 millones de nuevos soles a precios de 1994.

Ejemplo Ilustrativo 2:

Siguiendo con el mismo ejemplo de las ventas semanales de nafta (en miles de litros); utilizando el método de Suavización exponencial simple con $\alpha=0.8$, obtenemos los resultados del siguiente cuadro:

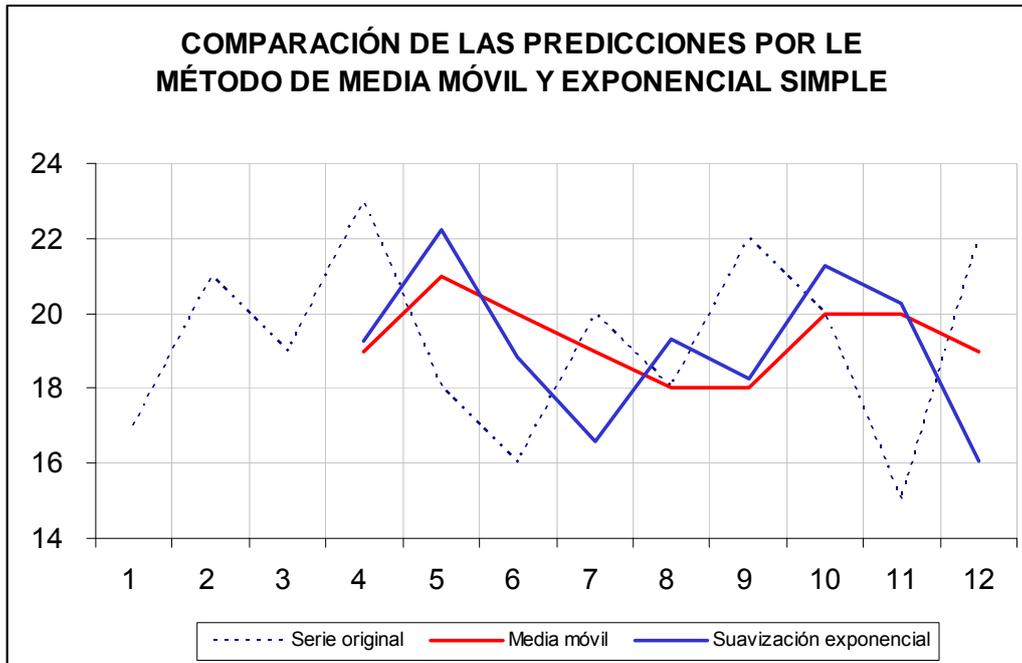
Semana	Ventas y_t	Suavización Exponencial $\alpha = 0,80$	Error $\mu=(y_t - \hat{y}_t)$	Error Cuadrado (μ^2)
1	17	17		
2	21	17		
3	19	20,20		
4	23	19,24	3,76	14,14
5	18	22,25	-4,25	18,05
6	16	18,85	-2,85	8,12
7	20	16,57	3,43	11,77
8	18	19,31	-1,31	1,73
9	22	18,26	3,74	13,97
10	20	21,25	-1,25	1,57
11	15	20,25	-5,25	27,57
12	22	16,05	5,95	35,40
13	PRONOSTICO	20.81	1,96	14,70

Predicciones:

$$P_3 = 0.8 \times Y_2 + (1 - 0.8) \times P_2$$

$$P_3 = 0.8 \times 21 + 0.2 \times 17 = 20.2$$

En la siguiente gráfica observamos las predicciones halladas por los métodos de media móvil simple y suavización exponencial.



NOTA: Para utilizar el método de suavización exponencial se debe elegir un coeficiente α adecuado de tal forma que el EMC sea el menor posible. En nuestro ejemplo podemos observar que la media móvil suaviza mejor los datos, asimismo presenta un menor error medio cuadrático (EMC).

LECCIÓN 3

MÉTODO DE SUAVIZACIÓN EXPONENCIAL DOBLE (Método de Brown)

En este método se calcula primero una suavización exponencial simple para cada valor de la serie y luego se vuelve a calcular otra suavización exponencial sobre los datos resultantes de la primera. Para ello se usan las siguientes formulas:

Suavización Exponencial Simple

$$P_t = \alpha Y_t + (1-\alpha)P_{t-1}$$

Suavización Exponencial Doble

$$Y'_t = \alpha P_t + (1-\alpha)Y'_{t-1}$$

Donde:

P_t : Valor atenuado según modelo de suavización exponencial simple en el tiempo t .

P_{t-1} : Valor atenuado según modelo de suavización exponencial simple en el tiempo $t-1$.

Y'_t : Valor pronosticado sobre segunda suavización exponencial en el tiempo t

Y'_{t-1} : Valor pronosticado sobre segunda suavización exponencial, en el tiempo $t - 1$

Y_t : Valor experimental de la serie de datos de tiempo

α : Constante de suavización exponencial

Para pronosticar hacia el futuro, se usa una interpolación lineal que contempla el componente de tendencia (segunda suavización exponencial) del siguiente tipo:

$$\hat{Y}_{t+j} = a_t + j \times (b_t)$$

Siendo:

$$a_t = 2 \times P_t - Y'_t$$

$$b_t = \left(\frac{\alpha}{1-\alpha} \right) \times (P_t - Y'_t)$$

Donde:

\hat{Y}_{t+j} : Valor pronosticado agregando tendencia lineal para el período $t+j$

a_t : Ordenada de origen para modelo lineal en el tiempo t

- b_t : Pendiente de tendencia lineal en el tiempo t
 j : Cantidad de períodos a pronosticar ($j = 1,2,3,\dots$)

La constante empírica α es la única variable en este modelo que debe ser determinada de manera experimental sobre los valores disponibles de la serie de datos.

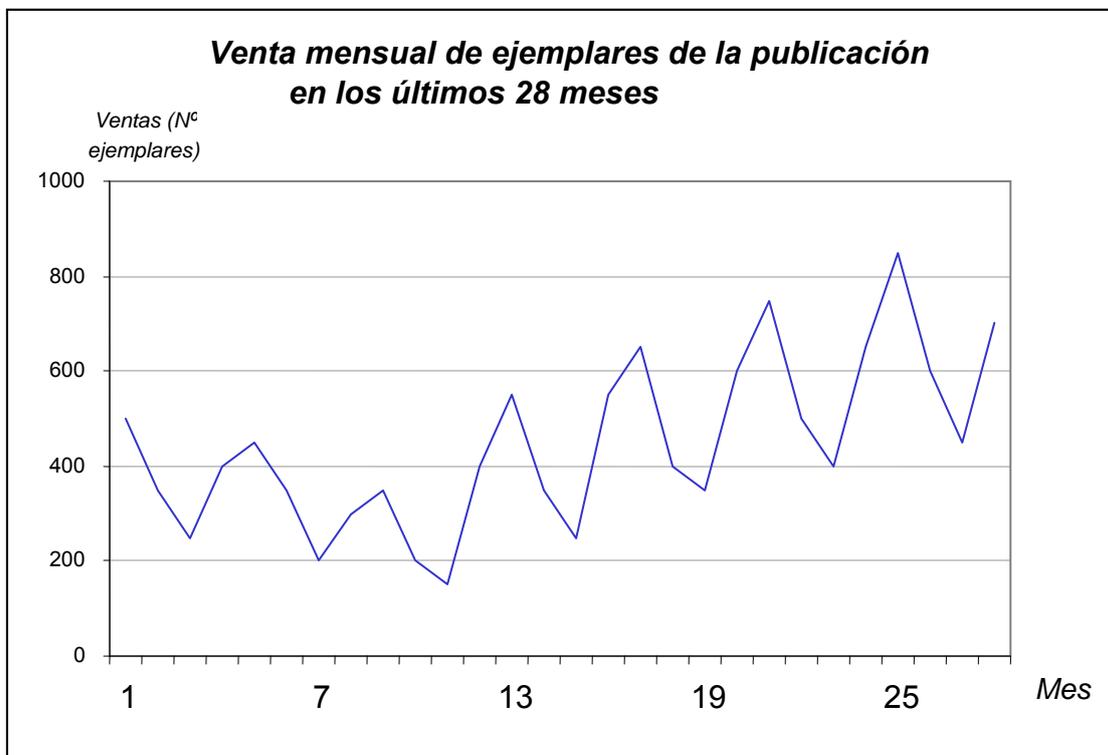
Ejemplo Ilustrativo

Aplicaremos este modelo al siguiente conjunto de datos reales que representan las ventas de una determinada publicación, durante los últimos 28 meses.

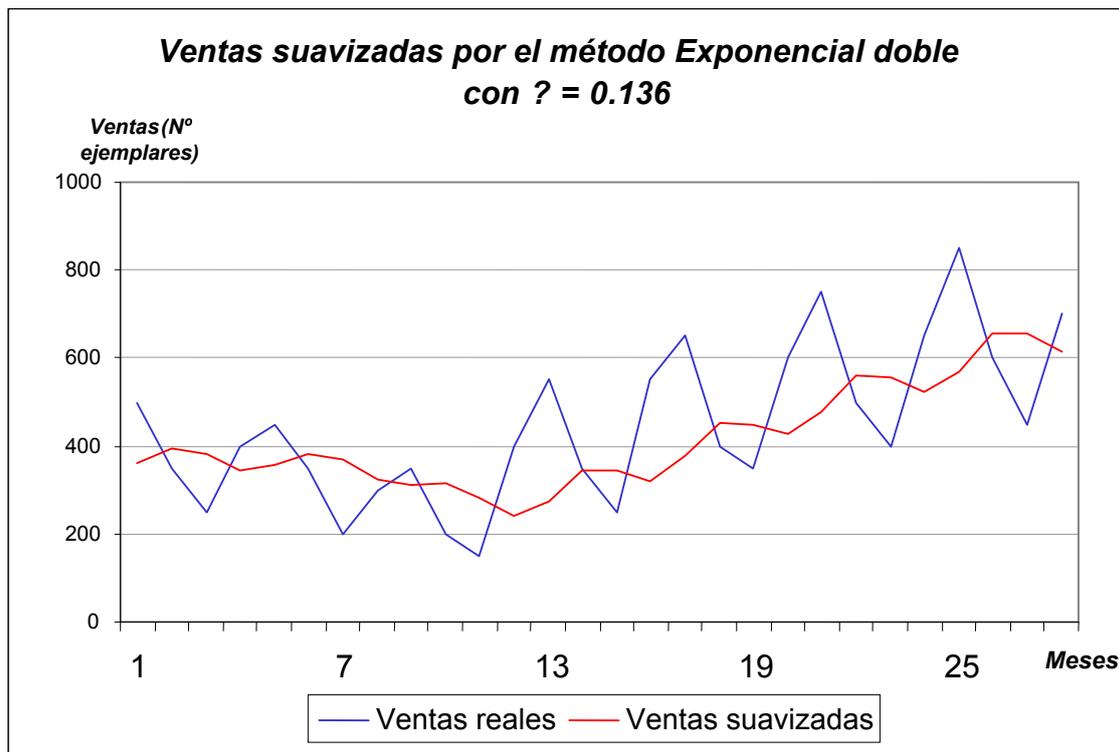
Número de ejemplares vendidos de una publicación durante los últimos 28 meses

Mes	Ventas	Mes	Ventas	Mes	Ventas	Mes	Ventas
1	500	8	300	15	250	22	500
2	350	9	350	16	550	23	400
3	250	10	200	17	650	24	650
4	400	11	150	18	400	25	850
5	450	12	400	19	350	26	600
6	350	13	550	20	600	27	450
7	200	14	350	21	750	28	700

La gráfica de la serie se presenta en el siguiente gráfico:



Con ayuda del Software Eviews hallamos la serie suavizada, utilizando un valor de $\alpha = 0.136$, los resultados se pueden apreciar en el siguiente gráfico.



LECCIÓN 4

MÉTODO DE SUAVIZACIÓN CON TENDENCIA Y ESTACIONALIDAD (Método de Holt - Winters)

ATENUACIÓN EXPONENCIAL AJUSTADO A TENDENCIA (MÉTODO DE HOLT)

Es un método sofisticado de extensión de la suavización exponencial, descrita anteriormente. A diferencia del método de suavización exponencial, el método de Holt Winters también permite el estudio de la tendencia de la serie a través de pronósticos a mediano y largo plazo

Este modelo utiliza dos constantes (α y β) para realizar los pronósticos. Estas constantes deben determinarse experimentalmente para señalar los valores reales de la serie de tiempo. Las Ecuaciones que se Utilizan son

$$\text{Nivel de la serie:} \quad E_t = \alpha Y_t + (1-\alpha)(E_{t-1} + T_{t-1})$$

$$\text{Nivel de la tendencia:} \quad T_t = \beta (E_t - E_{t-1}) + (1-\beta) T_{t-1}$$

De las ecuaciones anteriores las variables tienen el siguiente significado:

E_t	:	Estimación Atenuada de la serie en el tiempo t
E_{t-1}	:	Estimación Atenuada de la serie en el tiempo t -1
T_t	:	Valor pronosticado de Tendencia en el tiempo t
T_{t-1}	:	Valor pronosticado de Tendencia en el tiempo t -1
Y_t	:	Valor observado de la serie en el tiempo t
α	:	Constante de suavización de la serie
β	:	Constante de suavización para Tendencia

El modelo de Holt – Winters se utiliza cuando existe la presencia de una tendencia en la serie de tiempo. La elección de las constantes de suavización α y β afecta al valor de los resultados. Un valor pequeño de “ α ” da mayor peso a los valores más retrazados y un mayor valor en dicha constante da mayor peso a los niveles más recientes. Igualmente un valor pequeño de “ β ” da mayor peso a las tendencias más retrazadas en la serie y un menor valor de la constante da mayor peso a las tendencias de la serie más recientes.

Para pronosticar períodos futuros, se define la siguiente fórmula:

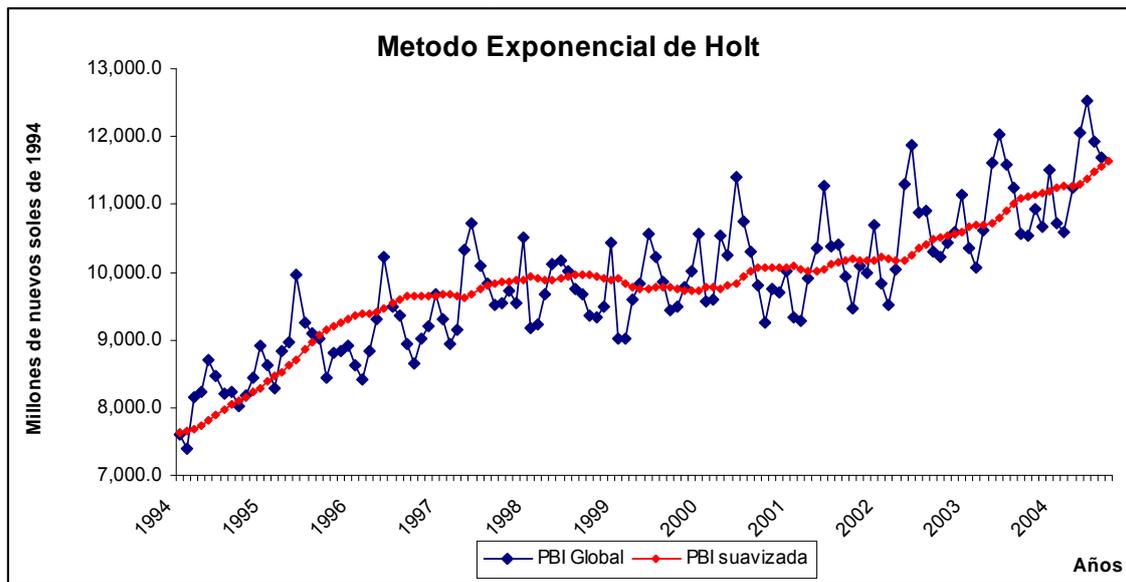
$$Y_{t+j} = E_t + j(T_t)$$

En la ecuación anterior las variables tienen el siguiente significado:

- Y_{t+j} : Valor de pronóstico de la serie para periodo futuro (t+j)
- E_t : Valor pronosticado de suavización exponencial atenuado
- T_t : Valor pronosticado de tendencia
- j : Cantidad de períodos futuros a pronosticar (j = 1,2,3,...)

Ejemplo Ilustrativo 1:

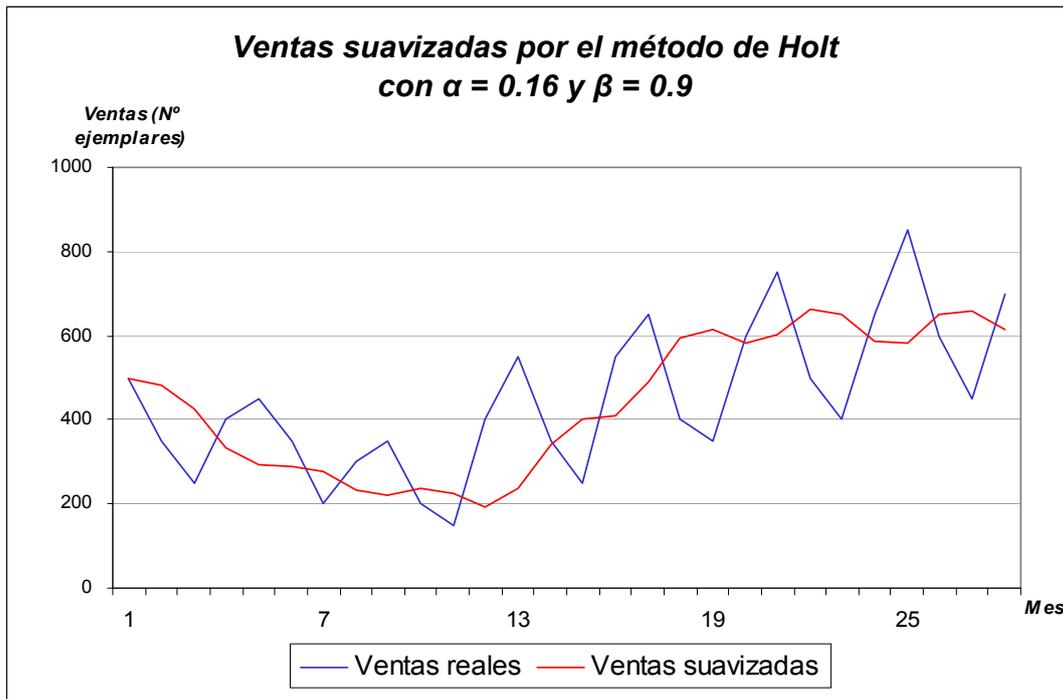
Continuando con el análisis del PBI, ahora suavizamos por medio del método de suavización de Holt, empleando los siguientes parámetros: 0.05 y 0.2



Podemos observar que por medio del método de Holt, logramos un mejor efecto en la suavización de la serie del PBI.

Ejemplo Ilustrativo 2

Aplicaremos este modelo al ejemplo de la sección anterior sobre las ventas de una publicación, para lo cual se obtienen los valores óptimos ($\alpha= 0.1600$ y $\beta =0.9002$). Los resultados se pueden apreciar en el siguiente gráfico.



SUAVIZACIÓN EXPONENCIAL CON COMPONENTES ESTACIONALES

Se entiende como variación estacional, aquella distorsión que se produce en la serie de datos debido a que hay un patrón de comportamiento que parece repetirse año tras año (o transcurridos una cantidad de períodos), ejemplos de este tipo lo constituyen las ventas de helado los cuales se incrementan en la temporada de verano, otro podrían ser las ventas de artículos de pirotecnia las cuales se incrementan durante el último mes del año, las ventas de ropa para deportes invernales que se incremente en el segundo trimestre del año, la demanda de pasajes que se incrementa en el mes de julio hacia destinos invernales, etc.

Los modelos aplicables en estos casos, deben contener un componente de corrección debido a la tendencia que pudiese llegar a presentarse y otra corrección debido a la estacionalidad que se presenta en la serie de datos.

En el ejemplo anterior, de las ventas de publicaciones podemos observar un comportamiento similar cada cuatro meses, analizaremos esta serie utilizando la componente estacional.

Existen varios métodos que pueden ser aplicados, pero a manera de ejemplo describiremos solamente los siguientes, que resultan ser los más aplicados:

- Atenuación Exponencial Ajustada con estimación de Tendencia y Variación Estacional (**método de Holt-Winters Multiplicativo**)
- Atenuación Exponencial Ajustada con estimación de Tendencia y Variación Estacional (**método de Holt-Winters Aditivo**)

Ambos métodos son una extensión del método de Holt presentado anteriormente, con la particularidad de que agregaremos una complejidad adicional al incorporar índices de estacionalidad como una manera de corregir la serie de tiempos evitando estos “picos” que distorsionan el análisis de los pronósticos y la predicción en consecuencia.

Método de Holt-Winters Multiplicativo

Se basa en el cálculo de cuatro componentes:

- (1) Ajuste Exponencial de la Serie de Datos

$$E_t = \alpha \left(\frac{Y_t}{S_{t-L}} \right) + (1-\alpha)(E_{t-1} + T_{t-1})$$

- (2) Estimación de Tendencias

$$T_t = \beta (E_t - E_{t-1}) + (1-\beta)T_{t-1}$$

- (3) Estimación de Estacionalidad

$$S_t = \delta \left(\frac{Y_t}{E_t} \right) + (1-\delta)S_{t-L}$$

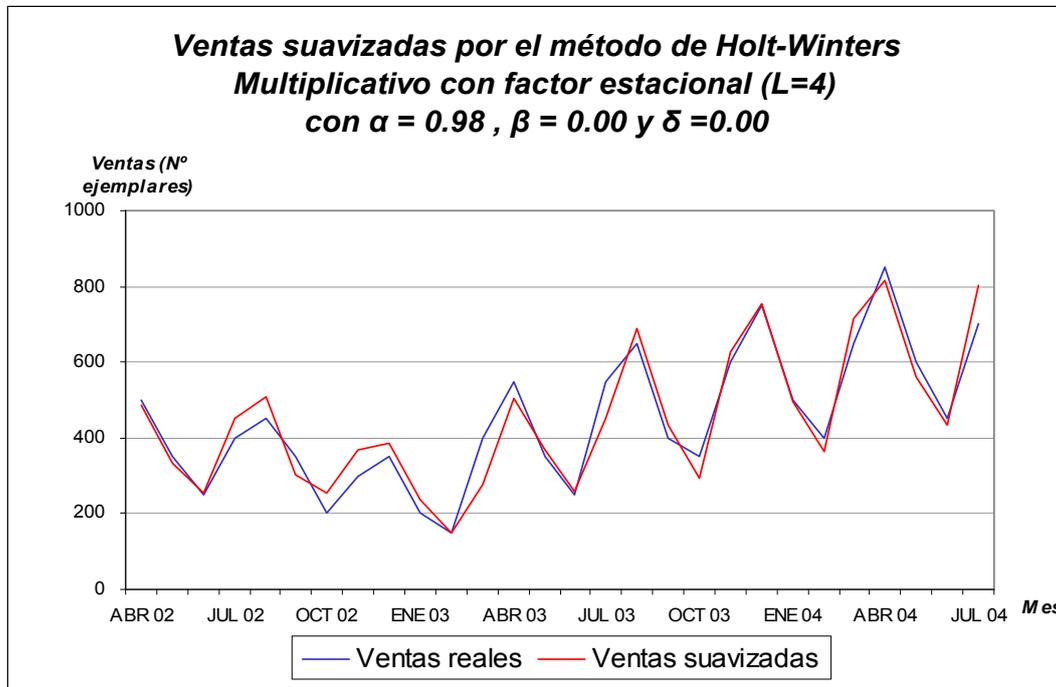
- (4) Pronóstico de Períodos Futuros

$$\hat{Y}_{t+j} = (E_t + j(T_t)) S_{t-L+j}$$

Las variables en las ecuaciones tienen el siguiente significado:

E_t	:	Estimación suavizada para el período t
T_t	:	Estimación de Tendencia para período t
S_{t-L}	:	Índice de Estacionalidad calculado para el período t-L
S_t	:	Índice de estacionalidad para el período t
Y_t	:	Valor real de la serie de tiempo para el período t
L	:	Longitud o duración de la estacionalidad, en nuestro caso L = 4 (el ciclo se repite cada cuatro meses)
j	:	Cantidad de períodos a Pronosticar hacia delante
α	:	Constante de suavización exponencial simple de la serie de datos
β	:	Constante de suavización exponencial de tendencia
δ	:	Constante de corrección de estacionalidad

Los valores de α , β y δ , deben ser calculados de manera experimental sobre el conjunto de datos disponibles de la serie de tiempo. En nuestro caso, el software Eviews estima los valores óptimos de α , β y δ , resultando **0.98**, **0.00** y **0.00** respectivamente. El modelo final se puede apreciar en la siguiente gráfica:



Método de Holt-Winters Aditivo

Se basa en el cálculo de cuatro componentes:

- (5) Ajuste Exponencial de la Serie de Datos

$$E_t = \alpha(Y_t - S_{t-L}) + (1-\alpha)(E_{t-1} + T_{t-1})$$

- (6) Estimación de Tendencias

$$T_t = \beta(E_t - E_{t-1}) + (1-\beta)T_{t-1}$$

- (7) Estimación de Estacionalidad

$$S_t = \delta(Y_t - E_t) + (1-\delta)S_{t-L}$$

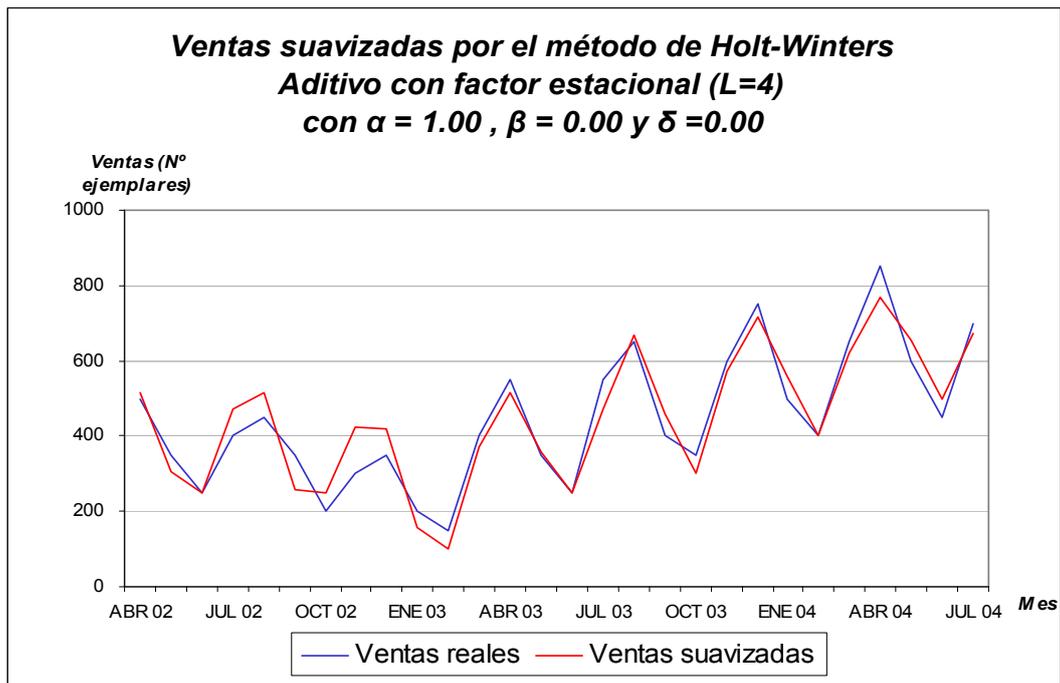
- (8) Pronóstico de Períodos Futuros

$$\hat{Y}_{t+j} = E_t + jT_t + S_{t-L+j}$$

Las variables en las ecuaciones tienen el siguiente significado:

- E_t : Es la estimación exponencialmente suavizada para el período t cualquiera
- Y_t : Valor real de la serie de tiempo para el período t cualquiera
- T_t : Estimación de Tendencia para período t
- S_t : Índice de estacionalidad para el período t
- S_{t-L} : Índice de Estacionalidad en el período t-L
- L : Longitud o duración de la estacionalidad, en nuestro caso $L = 4$
- j : Cantidad de períodos a Pronosticar hacia adelante
- α : Constante de suavización exponencial simple
- β : Constante de suavización exponencial de tendencia
- δ : Constante de corrección de estacionalidad

Los valores de α , β y δ , deben ser calculados de manera experimental sobre el conjunto de datos de la serie de tiempos disponibles. En nuestro caso, los valores de α , β y δ , óptimos resultan: **1.00**, **0.00** y **0.00** respectivamente. Basado en estos valores, obtenemos el siguiente resultado:



Para evaluar cuál de los modelos se ajusta mejor a los datos se procede a calcular el Error Medio Cuadrático (EMC) para cada caso, el resultado se resume en el siguiente cuadro:

Método	Error Medio Cuadrático (EMC)
Holt (sin estacionalidad)	23977.1
Holt – Winters (Multiplicativo)	2571.2
Holt – Winters (Aditivo)	2856.0

Como se puede observar el método de Holt-Winters Multiplicativo presenta un menor EMC por lo cual provee un mejor ajuste a los datos, por lo que será elegido para el cálculo de pronósticos.

Ejercicio de autoconocimiento

¿Por qué debo aplicar una técnica de suavización?

CRITERIOS	SI	NO	NO SÉ
1. Porque el objetivo de suavizar las fluctuaciones es apropiado para las series estables.			
2. Porque me permite determinar la dirección de la tendencia.			
3. Porque su uso sirve para modelar series estacionarias.			
4. Para visualizar mejor la tendencia.			
5. Para eliminar las fluctuaciones de la serie que no permiten observar su verdadero comportamiento.			
6. Para hacer un buen pronóstico de períodos futuros.			
7. Porque se puede aplicar en diferentes casos, como en las ventas de helado en verano.			
8. Para manejar mejor cualquier serie estacionaria o no.			
9. Porque son un medio efectivo de para realizar predicciones.			
10. Porque permite hacer un estudio de la tendencia a través de pronósticos.			

CALIFICACION

Puntuar con un punto cada respuesta “SI”.

Si obtienes de de 1 - 3 puntos tienes pocas expectativas de hacer una buena aplicación de las técnicas de suavización.

Si tienes entre 4 – 7, tienes buenas expectativas de hacer una buena aplicación de las técnicas de suavización.

Y si tienes entre 8 – 10, denotas excelentes expectativas de hacer una buena aplicación de las técnicas de suavización.

RESUMEN

Las **técnicas más usadas** para realizar la suavización son: el método de medias móviles, suavización exponencial simple, suavización exponencial doble (método de Brown) y el método de suavización con tendencia y estacionalidad (método de Holt-Winters).

Para determinar que **método elegido** es el más conveniente será necesario considerar la exactitud del mismo, lo que significa que tendrá que considerarse como adecuado aquel que sea capaz de originar el menor error de pronóstico, para tal fin se toma como medida referencial al Error Medio Cuadrático (EMC).

El **objetivo del uso** de estas técnicas es eliminar las fluctuaciones que estén presentes en la serie las cuales no permiten observar su verdadero comportamiento, esto se logra a partir de ajustar los datos históricos de a una nueva serie en la cual se pueda determinar el verdadero comportamiento de ésta.

Resumen de fórmulas

Método de Medias Móviles:

- Promedios móviles simples (PMS)

$$\text{Promedio Móvil} = \frac{\sum (\text{n valores de datos más recientes})}{n}$$

- Media móvil anual

$$M_t^{12} = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-11}}{12}$$

- Media Móvil de Orden "s"

$$M_t^s = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-s+1}}{s} = \frac{\sum_{i=1}^{s-1} y_{t-i}}{s}$$

Pronósticos: $\hat{y}_t = M_{t-1}^s$

- Media Móvil Centrada:

$$MC_t^s = \frac{y_{t+\frac{(s-1)}{2}} + \dots + y_{t+1} + y_t + y_{t-1} + \dots + y_{t-\frac{(s-1)}{2}}}{s}$$

Error Medio Cuadrado (EMC)

$$EMC = \frac{\sum((\text{Valor Real} - \text{Valor Pronosticado})^2)}{\text{Cantidad Valores Pronosticados}}$$

- **Método de Suavización Exponencial Simple**

$$P_{t+1} = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \dots + \alpha(1-\alpha)^{n-1} X_{t-(n-1)}$$

$$P_{t+1} = \alpha X_t + (1-\alpha)P_t$$

- **Método de Suavización Exponencial Doble ((Método de Brown))**

Suavización Exponencial Simple

$$P_t = \alpha Y_t + (1-\alpha)P_{t-1}$$

Suavización Exponencial Doble

$$Y'_t = \alpha P_t + (1-\alpha)Y'_{t-1}$$

Pronóstico: $\hat{Y}_{t+j} = a_t + j \times (b_t)$

Siendo: $a_t = 2 \times P_t - Y'_t$, $b_t = \left(\frac{\alpha}{1-\alpha}\right) \times (P_t - Y'_t)$

- **Método de Suavización con Tendencia y Estacionalidad (método de Holt-Winters)**

Atenuación Exponencial Ajustado A Tendencia (Método De Holt)

Nivel de la serie: $E_t = \alpha Y_t + (1-\alpha)(E_{t-1} + T_{t-1})$

Nivel de la tendencia: $T_t = \beta (E_t - E_{t-1}) + (1-\beta) T_{t-1}$

Pronóstico: $Y_{t+j} = E_t + j(T_t)$

- Suavización Exponencial Con Componentes Estacionales

Método de Holt-Winters Multiplicativo:

Ajuste Exponencial de la Serie de Datos

$$E_t = \alpha \left(\frac{Y_t}{S_{t-L}}\right) + (1-\alpha)(E_{t-1} + T_{t-1})$$

Estimación de Tendencias

$$T_t = \beta (E_t - E_{t-1}) + (1-\beta)T_{t-1}$$

Estimación de Estacionalidad

$$S_t = \delta \left(\frac{Y_t}{E_t} \right) + (1 - \delta) S_{t-L}$$

Pronóstico de Períodos Futuros

$$\hat{Y}_{t+j} = (E_t + j(T_t)) S_{t-L+j}$$

Método de Holt-Winters Aditivo

Ajuste Exponencial de la Serie de Datos

$$E_t = \alpha (Y_t - S_{t-L}) + (1 - \alpha) (E_{t-1} + T_{t-1})$$

Estimación de Tendencias

$$T_t = \beta (E_t - E_{t-1}) + (1 - \beta) T_{t-1}$$

Estimación de Estacionalidad

$$S_t = \delta (Y_t - E_t) + (1 - \delta) S_{t-L}$$

Pronóstico de Períodos Futuros

$$\hat{Y}_{t+j} = E_t + j T_t + S_{t-L+j}$$

EXPLORACION ON LINE

1. Pronósticos obtenidos mediante la **suavización** exponencial simple

<http://www.prodigyweb.net.mx/lmfs/ingenierias/pronosticos/pronosticos.html>

2. Definiciones de Suavización Exponencial y Ejercicios Desarrollados

<http://www2.adm.salvador.edu.ar/master/deusto/Metodos/Apuntes/Pronosticos/Teoria%20General%20de%20Pronosticos.doc>.

3. Aplicación de la media móvil.

http://html.rincondelvago.com/administracion-de-la-produccion_4.html

LECTURA

TÉCNICAS DE SUAVIZACIÓN

Una de las principales aplicaciones del estudio de las series temporales está relacionado a la predicción de los valores futuros, como se explicó en el capítulo anterior, siendo la mejor manera de explicar el comportamiento de una variable es el estudio de su tendencia a largo plazo, el problema que trae consigo esta afirmación es que al realizar el estudio se observará que en la serie contiene una considerable cantidad de fluctuaciones aleatorias y cambios estacionales al corto plazo lo que hará difícil encontrar cierto comportamiento particular, por ejemplo el caso de la presencia de una tendencia.

Para remediar tal situación se han propuesto un cierto tipo de técnicas como el cálculo de las medias de los datos para ciertos periodos (promedios móviles). Lográndose con el uso de estas técnicas eliminar las fluctuaciones de la serie que no permiten observar su verdadero comportamiento, esto es a partir del ajuste de los datos históricos de una nueva serie.

De todos los métodos de pronóstico por series de tiempo es la suavización exponencial el que más se adecua a la predicción a corto plazo, debido a las posibilidades de automatización del proceso de cálculo, operar con costos relativamente bajos y de dar una mayor importancia a los datos de demandas recientes cuando se realiza el pronóstico.

Existen una variedad de métodos, los cuales se emplean principalmente en las áreas de predicción en los negocios y la economía, como por ejemplo para detectar el comportamiento de las ventas de un producto que ha aparecido en el mercado hace veinticuatro meses y se desea determinar cuál es la época del año en que debe incidir con mayor fuerza (para el uso de campañas de publicidad) con el fin de levantar sus ingresos o si sus ventas presentan un comportamiento mantenido (comportamiento cíclico).

Otro ejemplo lo constituye la selección del método de suavización exponencial para proyectar la demanda de productos en las empresas hoteleras, el cual es usado debido a la existencia de un elevado número de artículos así como poseen una demanda independiente aleatoria.

El autor

ACTIVIDADES

1. Dado los números 10, 6, 2, 5, 2, 6, 9,7, 3, determinar el promedio móvil de orden 2.
2. La tabla siguiente muestra el total de ventas mensuales de una empresa productora de conservas en el Perú durante los últimos 6 años.
 - a. Construir un promedio móvil de orden 12
 - b. Construir un promedio móvil centrado de orden 12
 - c. Representar gráficamente los resultados obtenidos junto con los datos originales y comparar los resultado

	1998	1999	2000	2001	2002	2003
Enero	174132	171695	150419	167654	215773	212658
Febreo	161943	142800	138585	165798	177886	232616
Marzo	146195	156978	164806	149714	222408	249744
Abril	141146	138780	141578	175504	215793	231813
Mayo	147081	154502	136450	156909	203316	216286
Junio	168127	155887	136144	183660	209376	220976
Julio	144123	161107	144638	177009	203557	212788
Agosto	169350	162323	118136	204359	240105	195399
Septiembre	163300	175434	133284	194735	225476	207013
Octubre	186654	172045	158266	199345	242739	187727
Noviembre	182168	190468	167523	218171	217358	188220
Diciembre	181453	164692	167508	213791	210469	159060

3. Considere las observaciones de la tabla siguiente las cuales corresponden a las ventas de un determinado producto. Calcule haciendo uso del método de Holt Winters sus predicciones correspondientes, para lo cual se considere los valores de $\alpha = 0.1600$ y $\beta = 0.9002$.

	1990	1991	1992	1993	1994	1995	1996
Enero	299.53	274.01	278.92	276.85	310.87	362.34	355.34
Febrero	286.04	258.24	263.25	262.53	287.53	329.38	343.17
Marzo	295.52	271.92	281.32	296.28	333.20	356.09	352.04
Abril	276.49	293.04	272.24	290.02	332.65	358.20	364.69
Mayo	307.33	313.58	296.85	309.90	363.05	404.46	414.46
Junio	300.94	298.54	288.45	310.98	356.85	380.43	399.64
Julio	301.97	308.95	286.65	319.38	351.80	374.86	395.91
Agosto	245.94	273.84	258.45	288.26	318.12	341.56	348.26
Septiembre	204.29	249.39	237.95	262.63	293.49	304.81	306.35
Octubre	219.19	256.90	249.31	262.60	307.12	318.78	318.43
Noviembre	240.22	267.29	259.60	283.52	324.63	327.14	346.89
Diciembre	266.30	268.80	279.73	301.79	333.66	328.93	356.96

4. Considere una suavización exponencial simple para una serie que corresponde a las importaciones a precios constantes. Calcular las predicciones con sus respectivos errores de predicción.

121.65	253.39	582.23	867.72	568.72
164.35	308.14	608.55	1005.82	682.80
182.27	347.58	606.88	1028.04	776.15
197.76	384.38	681.57	723.37	786.01
173.64	410.56	866.66	591.83	974.18
222.25	479.94	841.91	540.60	1216.16
261.00	543.97	735.87	650.72	1222.97
286.89	603.19	738.45	747.35	1347.59
253.21	548.36	558.46	680.38	1350.55
211.87	544.95	666.22	508.36	

5. los siguientes datos corresponden al IPC al consumidor de lima metropolitana promedio mensual de 1990-2000. Calcule las predicciones para los meses de junio a diciembre del 2000.

MES	1990	1992	1993	1994	1995	1996	1997	1998	1999	2000
Enero	0.29	43.15	68.48	92.78	105.51	117.31	130.21	139.21	146.27	151.80
Febrero	0.38	45.19	70.49	94.46	106.71	119.11	130.33	140.93	146.73	152.53
Marzo	0.51	48.56	73.48	96.66	108.17	120.75	131.99	142.79	147.63	153.36
Abril	0.69	50.10	76.73	98.15	109.24	121.80	132.50	143.66	148.50	154.14
Mayo	0.92	51.82	79.06	98.86	110.15	122.69	133.50	144.52	149.20	154.17
Junio	1.31	53.68	80.49	99.98	111.04	123.26	134.95	145.28	149.47	154.27
Julio	2.14	55.55	82.70	100.87	111.67	124.95	136.07	146.19	149.86	
Agosto	10.66	57.12	84.79	102.42	112.83	126.11	136.38	146.58	150.12	
Septiembre	12.12	58.61	86.17	102.95	113.27	126.52	136.77	145.79	150.81	
Octubre	13.29	60.74	87.47	103.24	113.85	127.45	136.98	145.30	150.63	
Noviembre	14.08	62.89	88.87	104.50	115.26	128.04	137.09	145.35	151.04	
Diciembre	17.42	65.32	91.10	105.12	115.87	129.59	137.96	146.25	151.70	

FUENTE: INSTITUTO NACIONAL DE ESTADISTICA E INFORMATICA - Dirección Técnica de Indicadores Económicos.

AUTOEVALUACIÓN

Encierra en un círculo la letra que contenga la alternativa correcta.

1. Porque se debería aplicar una suavización:
 - a. Para aplanar eliminar los picos y valores extremos de una serie.
 - b. Cuando no se tiene certeza acerca del comportamiento de una serie.
 - c. Por la presencia del error aleatorio.
 - d. Cuando se quiere estimar el valor promedio de la serie a largo plazo.
2. Establezca la verdad o falsedad de cada una de las siguientes proposiciones:
 - a. Las técnicas de suavización son usadas para eliminar las fluctuaciones aleatorias y cambios estacionales a corto plazo.
 - b. Con el uso de la media móvil puede lograr “aplanar” los datos y producir un movimiento de la serie donde no aparezcan tantos picos como inicialmente tiene la serie.
 - c. Para formar una media móvil vasta con tener información pasada y considerar un orden “s”.
 - d. Con el uso de una técnica de suavización siempre se logrará obtener un comportamiento más estable “más claro” de la serie.
3. Identificar cada una de las siguientes técnicas de suavización a partir de las ecuaciones siguientes :

a.
$$M_t^s = \frac{y_t + y_{t-1} + y_{t-2} + \dots + y_{t-s+1}}{s}$$

b.
$$P_t = \alpha Y_t + (1-\alpha)P_{t-1}$$

c.
$$Y'_t = \alpha P_t + (1-\alpha)Y'_{t-1}$$

d. Siendo

Nivel de la serie:
$$E_t = \alpha Y_t + (1-\alpha)(E_{t-1} + T_{t-1})$$

Nivel de la tendencia:
$$T_t = \beta (E_t - E_{t-1}) + (1-\beta) T_{t-1}$$

- I. Ninguna
- II. Método de Brown
- III. Método de Holt- Winters aditivo
- IV. Media Móvil anual

4. Establezca la verdad o falsedad acerca del método de HoltWinters

- a. A diferencia con el método de suavización exponencial, el método de Holt Winters permite el estudio de la tendencia de la serie a través de pronósticos a mediano y largo plazo
 - b. Se utiliza cuando existe la certeza de la presencia de una tendencia en la serie.
 - c. El modelo de Holt-Winters es una ampliación perfeccionada del método de suavización exponencial.
 - d. En comparación con los métodos antes mencionados este da mayor peso a los valores más retrazados y un mayor valor en dicha constante da mayor peso a los niveles más recientes
5. Considere una serie del IPC que se ha obtenido para Lima Metropolitana entre enero de 1990 y diciembre de 1999, si se hace uso del método de suavización exponencial simple para valores de α en el intervalo de $[0.8; 0.99]$. ¿Cuál es el valor de éste según el cual se obtiene el mejor ajuste al método planteado?

	a	b	c	D	e	f	g
α	0.8	0.9	0.95	0.96	0.97	0.98	0.99
EMC	512.195	486.856	477.802	476.269	474.828	473.477	472.218

RESPUESTAS DE CONTROL

1. b, 2. VVFF, 3. a.IV; b.I; c.II; d. III, 4. VVVV, 5. g