Modelos de Pronósticos

Claudia E. Carignano

PRONÓSTICOS

Introducción

Un aspecto esencial de la administración de cualquier organización es la planeación del futuro. En efecto, el éxito a largo plazo de una organización depende de cuán bien la gerencia anticipa el futuro y elabora las estrategias apropiadas. El buen juicio, la intuición y tener conciencia del estado de la economía pueden dar a un gerente una idea aproximada de lo que es probable que suceda en el futuro. Sin embargo, con frecuencia es difícil convertir esta intuición en un número que pueda usarse, como el volumen de ventas del siguiente trimestre o el costo de la materia prima por unidad para el año próximo.

Así, un gerente de operaciones, para establecer de manera realista los programas de producción necesita alguna estimación de las ventas futuras. Del mismo modo, una empresa que brinda servicios informáticos necesita tener una estimación de la demanda futura de sus servicios para determinar su planta laboral.

A pesar de las imprecisiones inherentes al intentar predecir el futuro, los pronósticos necesariamente guían la planeación y el establecimiento de políticas.

Los métodos de predicción pueden ser Cualitativos o Cuantitativos; dentro de los Cuantitativos existen dos importantes tipos, *métodos de Pronósticos de Series Temporales o de Extrapolación* y *métodos de Predicción Causal*.

Los métodos de Extrapolación se utilizan para pronosticar los valores futuros de una variable de interés a partir de valores anteriores observados a intervalos regulares de tiempo. Suponen que los patrones de comportamiento anteriores y las tendencias de esa variable, continuarán en los periodos futuros, por lo tanto, la información anterior se utiliza para generar los pronósticos. No toman en cuenta que ocasionó los datos anteriores, simplemente se asume que las tendencias y los patrones de comportamiento continuarán en el futuro.

Los métodos de pronósticos Causales, en cambio, se basan en el supuesto de que la variable que tratamos de pronosticar muestra una relación de causa y efecto con una o más variables. Es decir, pretenden pronosticar los valores futuros de una variable –dependiente- con ayuda de información anterior, a fin de estimar la relación entre la variable dependiente y una o más variables independientes.

La consideración primordial en la selección de un método para pronosticar es que los resultados deben facilitar el proceso de toma de decisiones a los gerentes de la organización.

Como rara vez un método funciona para todos los casos, es mejor pensar en los métodos de pronósticos como herramientas genéricas que pueden aplicarse simultáneamente, así en una situación dada es posible intentar varios métodos. La metodología que produce los pronósticos más exactos en un caso quizá no sea la mejor metodología en otra situación. Sin embargo, el método seleccionado debería producir un pronóstico que sea correcto, oportuno y comprensible para la gerencia, de manera que el pronóstico ayude a tomar mejores decisiones.

Los pronósticos Cualitativos son los que no requieren de una abierta manipulación de datos, sino que hacen uso del juicio de quien pronostica, por su naturaleza éstos suelen ser subjetivos y no utilizan modelos matemáticos. Las técnicas cualitativas se usan cuando no se tiene disponibilidad de información histórica o los datos son escasos.

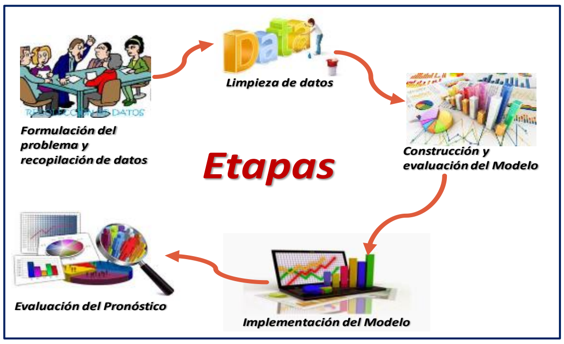
En definitiva, para realizar un pronóstico eficaz se debe establecer una hábil mezcla de pronóstico cuantitativo y buen juicio, evitando los extremos de confiar totalmente en uno o en otro.

Etapas del pronóstico

Todos los procedimientos formales para pronosticar requieren extender las experiencias del pasado hacia el futuro. Es decir, involucran la suposición de que las condiciones que generaron los datos y las relaciones pasados permanecerán en el futuro. Sin embargo, el futuro no siempre es como el pasado, cuando sí lo es, los métodos cuantitativos de elaboración de pronósticos funcionan bien, en caso contrario llegan a producirse pronósticos imprecisos. Sin embargo, generalmente es mejor tener algún pronóstico construido razonablemente, que no pronosticar.

Pasos en el proceso de pronosticar

1. Formulación del problema y recopilación de datos
2. Manipulación y limpieza de datos
3. Construcción y evaluación del modelo
4. Implementación del modelo (el pronóstico real)
5. Evaluación del pronóstico



En el ***paso 1***, la *formulación del problema* y la *recolección de datos* se tratan como un solo paso porque están íntimamente relacionadas, ya que el problema determina los datos apropiados. Si se está considerando una metodología cuantitativa para pronosticar, los datos pertinentes deben estar disponibles e igualmente espaciados.

El ***paso 2***, *manipulación y limpieza de datos*, a menudo es necesario debido a que es posible tener demasiados datos o muy pocos, o que algunos datos quizá no sean pertinentes o falten valores que se deban estimar. En muchas ocasiones ciertos datos tienen que se tienen que expresar en unidades diferentes de las originales y deben volver a procesarse. Es decir que habitualmente se requiere algún esfuerzo para obtener datos en la forma adecuada para usar ciertos procedimientos para pronosticar.

El ***paso 3***, *construcción y evaluación del modelo*, incluye ajustar los datos recolectados a un modelo de pronóstico que sea adecuado, en términos de minimizar errores en el pronóstico.

Cuanto más sencillo sea el modelo mejor será, en términos de aceptación por parte del decisor, el proceso de pronósticos. Con frecuencia, se debe alcanzar un equilibrio entre un enfoque para pronosticar complejo que ofrezca un poco más de precisión, y un enfoque sencillo que se entienda fácilmente y tenga el apoyo de quienes toman las decisiones y sea activamente usado por éstos.

El error de pronóstico (e_t), es la diferencia entre el valor observado y el pronóstico. Para medir los errores de pronóstico suele usarse la Desviación Media Absoluta (MAD) o el error cuadrático medio (EMC).

El ***paso 4***, *implementación del modelo*, es la generación del modelo real. Los datos de periodos históricos más recientes se mantienen como respaldo y más tarde se usan para verificar la exactitud del modelo.

El ***paso 5***, *evaluación del pronóstico*, implica la comparación de los valores del pronóstico con valores históricos reales. Luego de la implementación del modelo, se realizan los pronósticos para los periodos históricos más recientes, donde se conocen los valores de los datos y se analizan los errores en el pronóstico. El examen de los patrones de error a menudo lleva al analista a modificar el modelo para pronosticar.

La finalidad de un pronóstico es reducir el nivel de incertidumbre cuando se toman decisiones, esto sugiere dos reglas fundamentales:

1. El pronóstico debe ser técnicamente correcto y generar predicciones lo suficientemente precisas para satisfacer las necesidades del decisor.
2. El procedimiento de elaboración y sus resultados tienen que presentarse de manera convincente para que sean utilizados en el proceso de toma de decisiones y los resultados deben justificarse desde el punto de vista del costo-beneficio.

Métodos Cuantitativos de pronóstico

I. Series de Tiempo

Una de las partes más difícil y que más tiempo consume en la elaboración de pronósticos es la recopilación de datos válidos y confiables. Un pronóstico no puede ser tan preciso como los datos en que se basa. Sin embargo, aún el modelo de pronóstico más elaborado fallará si se aplica a datos poco confiables. Los datos deben ser fidedignos, precisos, consistentes, oportunos y relevantes; tienen que ser representativos de las circunstancias para las cuales se están usando.

El propósito es examinar esos datos y luego extrapolar o extender las relaciones identificadas a una población en general.

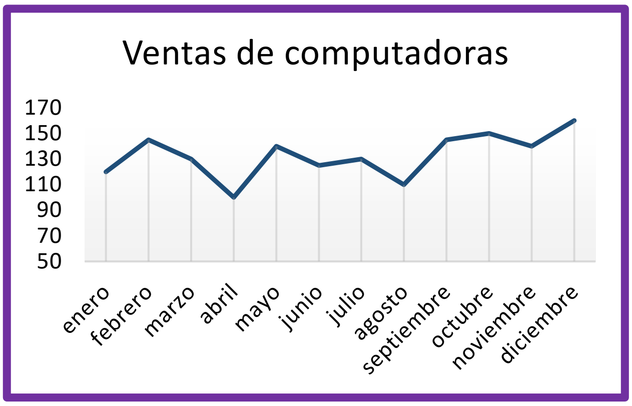
Cualquier variable integrada con datos observados, recopilados y registrados durante incrementos de tiempo fijos y sucesivos se llama serie de tiempo.

Podemos definir una ***serie de tiempo*** como *el conjunto de observaciones de una variable medida en puntos sucesivos en el tiempo o a lo largo de periodos consecutivos*.

Así si X_i es la variable aleatoria de interés en el tiempo i y si las observaciones se toman en los periodos 1 = 1, 2,….,t,… entonces los valores observados $\{X_1 = x_1 , X_2 = x_2,…. X_t = x_t\}$ son la serie de tiempo.

Por ejemplo, si nuestra variable de interés representa a las computadoras vendidas y las observaciones son mensuales, la serie de tiempo que la representa es:

Mes	Ventas
enero	120
febrero	145
marzo	130
abril	100
mayo	140
junio	125
julio	130
agosto	110
septiembre	145
octubre	150
noviembre	140
diciembre	160



En las situaciones reales no se tiene un conocimiento completo de la forma exacta del modelo que genera la serie de tiempo, por lo que se debe elegir un modelo aproximado. Con frecuencia, la elección se basa en la observación de los resultados de la serie de tiempo.

El objetivo de los métodos de serie de tiempo es descubrir un patrón en los datos históricos y luego extrapolarlo hacia el futuro; el pronóstico se basa sólo en valores pasados de la variable que tratamos de pronosticar o en errores pasados.

Componentes de una serie de tiempo

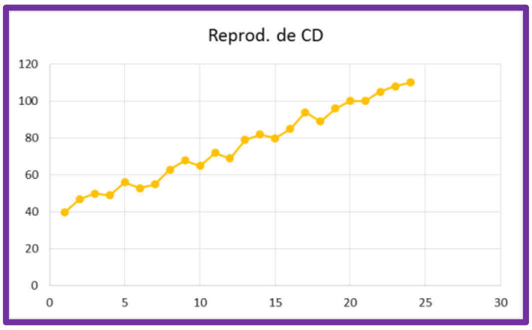
El patrón o comportamiento de los datos en una serie de tiempo tiene varios componentes. El supuesto usual es que cuatro componentes separados: tendencia, cíclico, estacional e irregular, se combinan para proporcionar valores específicos de la serie de tiempo.

I. Componente de tendencia

El cambio gradual de la serie de tiempo se conoce como tendencia. Este cambio o tendencia por lo general es el resultado de factores a largo plazo, como cambios en la población, características demográficas de la población, tecnología y preferencias de consumo.

La tendencia puede ser creciente o decreciente, lineal o no.

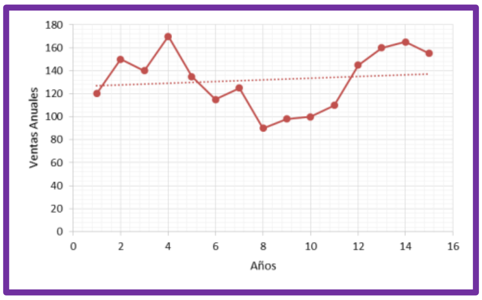
*La **tendencia** es el componente de largo plazo que representa el crecimiento o el descenso en la serie de tiempo, durante un periodo extenso.*



II. Componente cíclico

Aunque una serie de tiempo puede mostrar una tendencia durante periodos prolongados, todos los valores futuros de la serie de tiempo no caen exactamente en la línea de tendencia. De hecho, las series de tiempo con frecuencia muestran secuencias de puntos que se alternan por encima y por debajo de la línea de tendencia. Cualquier secuencia de puntos recurrente por encima y por debajo de la línea de tendencia que dura más de un año puede atribuirse al componente cíclico de las series de tiempo. Por lo general, este componente da como resultado movimientos cíclicos de múltiples años en la economía. Por ejemplo, periodos de inflación modesta seguidos por periodos de inflación rápida pueden conducir a series de tiempo que alternan por encima y por debajo de una línea de tendencia.

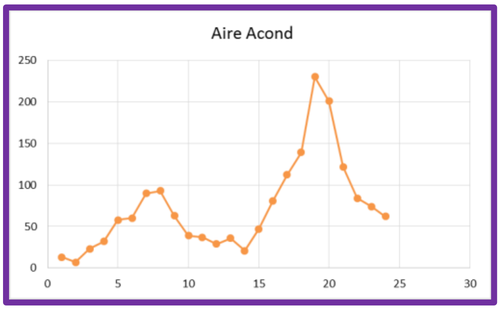
*El componente **cíclico** es la fluctuación con forma de onda alrededor de la tendencia y, por lo común, se ve afectada por las condiciones económicas generales.*



III. Componente estacional

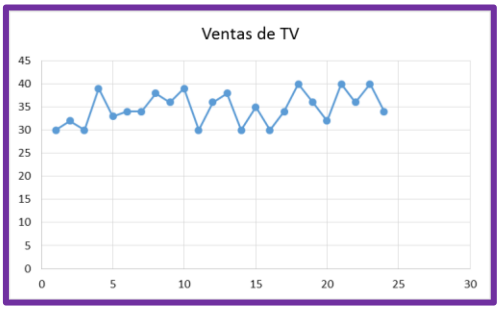
Mientras que el componente cíclico y de tendencia de una serie de tiempo, se identifican mediante el análisis de los movimientos de datos históricos durante varios años, muchas series de tiempo muestran un patrón regular durante periodos de un año. Aunque por lo general consideramos que el movimiento estacional en una serie de tiempo ocurre en un año -a semejanza de las estaciones del año-, el componente estacional también puede utilizarse para representar cualquier patrón que se repite con regularidad y tiene una duración menor a un año. En definitiva, si una serie es estacional, un patrón de comportamiento relacionado con el calendario se repite a sí mismo durante un intervalo de tiempo específico.

*El componente **estacional** es un patrón de cambio que se repite año tras año o regularmente en periodos menores a un año*



IV. Componente irregular

El componente irregular es el factor residual que incluye las desviaciones de los valores de serie de tiempo reales de aquellos esperados según los efectos del componente cíclico, de tendencia y estacional. Este componente representa la variabilidad aleatoria en las series de tiempo y es resultado de factores a corto plazo, imprevistos y no recurrentes que afectan a la serie de tiempo. Como este componente representa la variabilidad aleatoria, es impredecible; no podemos intentar prever su impacto en las series de tiempo.



Podemos identificar las componentes de una serie de tiempo a través de un análisis gráfico o realizando una autocorrelación.
Autocorrelación es la correlación que existe entre una variable, retrasada en uno o más periodos, consigo misma.
Los patrones de datos que incluyen componentes como tendencia y estacionalidad pueden estudiarse usando autocorrelaciones.
Los patrones se identifican examinando los coeficientes de autocorrelación de una variable en diferentes retrasos de tiempo.

Tiempo	Mes	x_t	x retrasada un periodo x_{t-1}	x retrasada dos periodos x_{t-2}
1	Enero	123		
2	Febrero	130	123	
3	Marzo	125	130	123
4	Abril	138	125	130
5	Mayo	145	138	125
6	Junio	142	145	138
7	Julio	141	142	145
8	Agosto	146	141	142
9	Septiembre	147	146	141
10	Octubree	157	147	146
11	Noviembre	150	157	147
12	Diciembre	160	150	157

¿Los datos muestran una tendencia?
Si una serie muestra una tendencia, hay una relación significativa entre los valores sucesivos de la serie de tiempo. Los coeficientes de autocorrelación son usualmente grandes para varios de los primeros retrasos de tiempo y luego, conforme se incrementa el número de retrasos, caen gradualmente hacia cero.
¿Los datos son estacionales?
Si una serie es estacional las observaciones de la misma posición, en diferentes periodos estacionales, tienden a estar relacionadas. Si se analizan datos trimestrales que tienen un patrón estacional, los primeros trimestres tienden a parecerse, los segundos trimestres tienden a parecerse, y así sucesivamente, y habrá un coeficiente de autocorrelación significativo en el retraso de tiempo 4.

II. Métodos de pronósticos de series temporales

Simbología

T = periodo total que abarca la serie de tiempo

t = sub-periodos en los cuales está dividido T

x_t = valor de la variable de interés que se está estudiando y para la cual se han recopilado los datos, en el momento t.

f_t = valor pronosticado o pronóstico para el momento t.

e_t = error del pronóstico para el periodo t.

Errores de pronóstico

Las diferencias que ocurren entre el valor de la variable en estudio (x_t) y su valor pronosticado (f_t) en el momento *t,* representa el error del pronóstico, es decir $e_t = x_t - f_t$, y nos da una estimación de la precisión del valor pronosticado.

Para medir la precisión del método de pronóstico adoptado, podemos usar al *Error Cuadrado Medio* (EMC), la *Desviación Media Absoluta* (MAD) o el *Error Porcentual Absoluto Medio* (MAPE).

EMC ➔ promedio del cuadrado de los errores (e_t).

$$EMC = \frac{\sum_{i=1}^{N}(x_i - f_i)^2}{N}$$

N = número de observaciones pronosticadas

- ☒ Este método penaliza los desvíos más grandes.

MAD ➔ promedio de los valores absolutos de los errores

$$MAD = \frac{\sum_{i=1}^{N}|x_i - f_i|}{N}$$

N = número de observaciones pronosticadas

- ☒ Los desvíos están expresados en las mismas unidades de la serie.

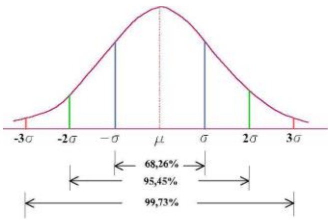
MAPE ➔ promedio del error absoluto del periodo dividido por el valor real observado en ese periodo y multiplicado por 100 para expresarlo como porcentaje.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - f_i}{x_i}\right|(100)$$

- ☒ Al ser un valor que no tiene unidades de medición es muy útil para comparar la exactitud de un método, o de más de un método, en series completamente diferentes.
- ☒ En general es de más fácil interpretación.

Precisión del pronóstico

Si el error de los pronósticos está distribuido en forma normal, podemos utilizar el MAD para estimar la desviación estándar de los errores del pronóstico con la relación: ***Se = 1,25 MAD***



- Podremos decir que un 68% de las estimaciones tendrán un error de ± 1,25(MAD).
- Y si el valor pronosticado para el siguiente periodo es f_t podremos decir que el valor de x_t, con un 0,95 de probabilidad, estarán entre f_t ± 1,25(2)(MAD).

Por ejemplo, si el valor pronosticado de ventas de computadoras para el mes de noviembre es de 135 unidades y el MAD es 16,04. Podemos decir, con un 0,95 de probabilidad, que la venta de computadoras estará entre 135 ± 1,25 **.** (2)**.**16,04 , es decir entre 95 y 175 unidades.

MÉTODO DEL PRONÓSTICO DEL ÚLTIMO VALOR

Usan el valor de la serie de tiempo observado en el tiempo *t* (x_t) como pronóstico para el periodo *t+1*. Esto es:

$$f_{t+1} = x_t$$

Desventaja: es impreciso dado que su varianza es grande debido a que se basa en una muestra de tamaño uno.

Es conveniente considerarlo solo si:

- ✓ El supuesto sobre el nivel constante de la serie es "insegura" y el proceso cambia con rapidez de manera que lo que ocurre antes del tiempo t es irrelevante o lleva a conclusiones equivocadas.
- ✓ Si la suposición de que el error aleatorio (e_t) tiene varianza constante es poco razonable y la varianza condicional en el periodo t es en realidad mucho más pequeña que en tiempos anteriores.

Se trata de un método ingenuo, sin embargo, cuando las condiciones cambian con rapidez, quizás el último valor sea el único dato relevante para pronosticar el siguiente valor.

MÉTODO DEL PROMEDIO

En lugar de usar una muestra de tamaño 1 usa todos los datos y obtienen un promedio de la serie.

$$f_{t+1} = \frac{\sum_{i=1}^{t} x_i}{t};$$ donde ***t*** representa el número de periodos transcurridos.

Esta estimación es excelente si el proceso es muy estable, es decir si se supone un nivel constante, se limita en general a procesos jóvenes.

MÉTODO DE PROMEDIOS MÓVILES

Obtiene el promedio de los datos de los últimos *n* periodos

$$f_{t+1} = \frac{\sum_{i=t-n+1}^{t} x_i}{n}$$

Se actualiza con facilidad de un periodo a otro, para lo cual se debe eliminar la primera observación y agregar la última.

- Ventaja: usa varias observaciones de datos recientes.
- Desventaja: todas las observaciones tienen el mismo peso.

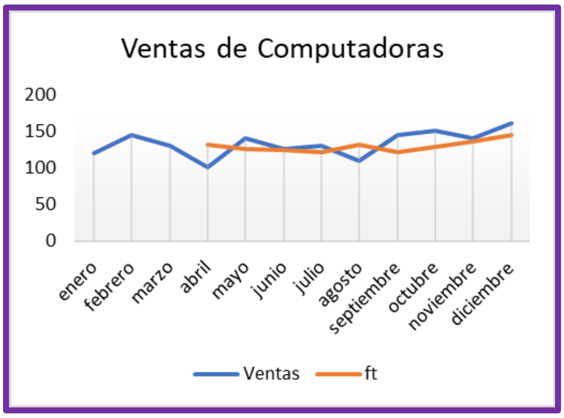
Volvamos a nuestro ejemplo de la venta de computadoras y calculemos los promedios móviles con $n = 3$

Promedio ***n = 3***

Mes	*Ventas*	f_t	*|et|*	*MAD*
enero	120			
febrero	145			
marzo	130			
abril	100	131,67	31,67	31,67
mayo	140	125,00	15,00	23,33
junio	125	123,33	1,67	16,11
julio	130	121,67	8,33	14,17
agosto	110	131,67	21,67	15,67
septiembre	145	121,67	23,33	16,94
octubre	150	128,33	21,67	17,62
noviembre	140	135,00	5,00	16,04
diciembre	160	145,00	15,00	15,93

$$f_{abril} = \frac{120+145+130}{3} = 131{,}667$$

$$f_{mayo} = \frac{145+130+100}{3} = 125$$



PROMEDIO MÓVIL PONDERADO

Si asignamos diferentes pesos o niveles de importancia a las observaciones promediadas, se obtiene un *Promedio Móvil Ponderado*. Generalmente se les da un peso mayor a las observaciones más recientes, se debe tener en cuenta que los pesos tienen que ser positivos y su suma debe ser igual a uno.

$$f_{t+1} = \sum_{i=t-n+1}^{t} p_i x_i, \qquad \text{para} \quad 0 \le p_i \le 1 \ \text{y} \ \sum_{i=t-n+1}^{t} p_i = 1$$

Continuando con el mismo ejemplo y seleccionando como pesos $p_1 = 0{,}15$; $p_2 = 0{,}30$ y $p_3 = 0{,}55$

Promedio n = 3 con pesos $p_1 = 0{,}15$ $p_2 = 0{,}30$ y $p_3 = 0{,}55$

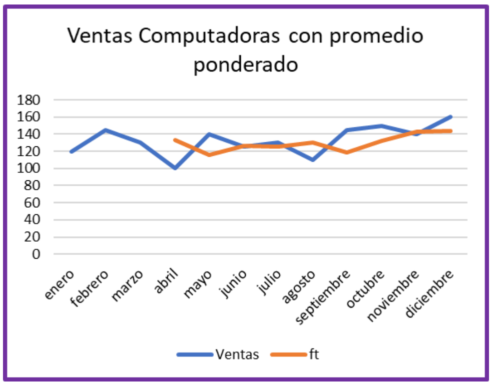
Mes	Ventas	f_t	$\lvert e_t \rvert$	MAD
enero	120			
febrero	145			
marzo	130			
abril	100	133,00	33,00	33,00
mayo	140	115,75	24,25	28,625
junio	125	126,50	1,50	19,583
julio	130	125,75	4,25	15,750
agosto	110	130,00	20,00	16,600
septiembre	145	118,25	26,75	18,292
octubre	150	132,25	17,75	18,214
noviembre	140	142,50	2,50	16,250
diciembre	160	143,75	16,25	16,250

$f_{abril} = 0{,}15(120) + 0{,}30(145) + 0{,}55(130)$

$f_{abril} = 133$

$f_{mayo} = 0{,}15(145) + 0{,}30(130) + 0{,}55(100)$

$f_{mayo} = 115{,}75$



MÉTODO DE SUAVIZADO EXPONENCIAL

Es un caso especial del método de promedios móviles ponderados en el cual sólo seleccionamos el peso (α) que le corresponde a la observación más reciente.

$f_{t+1} = \alpha x_t + (1-\alpha) f_t \qquad 0 < \alpha < 1, \qquad \alpha$ se llama constante de suavizado

Así el pronóstico es una suma ponderada de la última observación y el pronóstico anterior.

Podemos demostrar que el pronóstico de la suavización exponencial para cualquier periodo también es un promedio ponderado de todos los valores reales previos.

Por ejemplo, para una serie de tiempo que consta de cinco periodos de datos obtenemos:

Para el <u>periodo 1</u> el pronóstico es el valor observado en ese periodo

$f_1 = x_1$

Para el <u>periodo 2</u> el pronóstico es

$f_2 = \alpha x_1 + (1-\alpha) f_1 \quad = \alpha x_1 + (1-\alpha) x_1 = x_1$

Es decir que el pronóstico para el periodo 2 es igual al valor observado de la serie para el periodo 1.

Para el periodo 3

$f_3 = \alpha x_2 + (1-\alpha) f_2 \quad = \alpha x_2 + (1-\alpha) x_1$

Para el periodo 4

$$f_4 = \alpha x_3 + (1-\alpha) f_3 \quad = \alpha x_3 + (1-\alpha)\left[\alpha x_2 + (1-\alpha) x_1\right]$$
$$= \alpha x_3 + \alpha(1-\alpha) x_2 + (1-\alpha)^2 x_1$$

Para el periodo 5

$$f_5 = \alpha x_4 + (1-\alpha) f_4 \quad = \alpha x_4 + (1-\alpha)\left[\alpha x_3 + \alpha(1-\alpha) x_2 + (1-\alpha)^2 x_1 \right]$$
$$= \alpha x_4 + \alpha(1-\alpha) x_3 + \alpha(1-\alpha)^2 x_2 + (1-\alpha)^3 x_1$$

Obsérvese que el pronóstico para el periodo 5 es el promedio ponderado de los primeros cuatro valores de la serie.

También puede estimarse el pronóstico por suavización exponencial calculando el pronóstico para el periodo t+1 como el pronóstico para el periodo t corregido por el error del periodo t.

$$f_{t+1} = \alpha x_t \; + (1-\alpha) f_t = \alpha x_t \; + f_t - \alpha f_t$$
$$f_{t+1} = f_t \; + \alpha(x_t - f_t) \quad , \qquad f_{t+1} = f_t \; + \alpha e_t \qquad e_t = \text{ error de pronóstico del periodo } t$$

Observaciones:

Un valor bajo de la constante de suavizado hace que el pronóstico reaccione de manera más lenta a los cambios, por otro lado, un valor más alto es más sensible, pero podría reaccionar ante simples movimientos aleatorios.

Si la serie tiene una variabilidad aleatoria considerable, es preferible utilizar un valor pequeño para la constante de suavización. Como gran parte del error de pronóstico se debe a la variabilidad aleatoria, un α pequeño evita una reacción exagerada que ajuste los pronósticos demasiado rápido. Para una serie con poca variabilidad, valores más grandes de α permiten ajustar rápidamente los pronósticos cuando ocurren errores de pronóstico y por lo tanto el pronóstico reacciona más rápido ante condiciones cambiantes.

En general se selecciona un α entre 0,1 y 0,3 aunque es conveniente elegir el valor de α que minimiza el EMC o la MAD. Para seleccionar el valor de α se puede formular un modelo de optimización no lineal y resolverlo con el módulo Solver de Excel.

Si el valor de α que minimiza el MAD es superior a 0,5 entonces la serie presenta tendencia, estacionalidad o variación cíclica, por lo que no se recomienda suavización exponencial. En estos casos, el método *Holt* o el método de *Winter* proporcionan mejores pronósticos. El modelo de suavizado exponencial siempre estará a la zaga de todo incremento o reducción sistemática de los datos

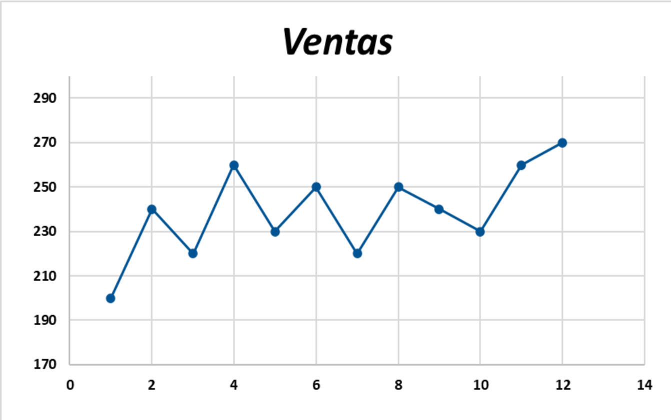
Desventajas:

- Queda retrasado cuando hay una tendencia continua.
- Es difícil elegir una constante de suavizado apropiada. Se ha sugerido que no debe exceder de 0,3 y una elección razonable aproximada es de 0,1.

Ejemplo

Los datos que se presentan a continuación reflejan las ventas de zapatillas deportivas de una fábrica situada en la localidad de Villa María, registrados durante los últimos 12 meses.
Sugiera un modelo adecuado de pronóstico para esta serie.

Mes	*Ventas*
1	200
2	240
3	220
4	260
5	230
6	250
7	220
8	250
9	240
10	230
11	260
12	270



Promedio Móvil con n = 3

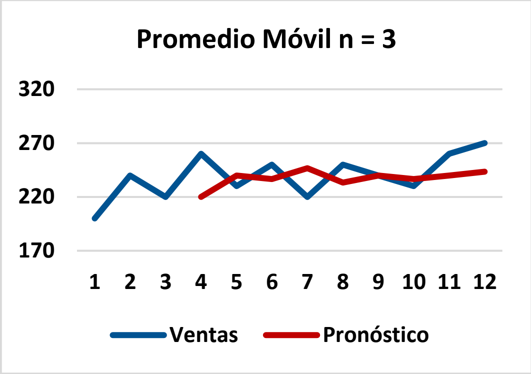
Mes	*Ventas*	*Pronóstico (f_t)*	e_t	$\|e_t\|$	*MAPE*
1	200				
2	240				
3	220				
4	260	220,00	40,0000	40,0000	0,1538
5	230	240,00	-10,0000	10,0000	0,0435
6	250	236,67	13,3333	13,3333	0,0533
7	220	246,67	-26,6667	26,6667	0,1212
8	250	233,33	16,6667	16,6667	0,0667
9	240	240,00	0,0000	0,0000	0,0000
10	230	236,67	-6,6667	6,6667	0,0290
11	260	240,00	20,0000	20,0000	0,0769
12	270	243,33	26,6667	26,6667	0,0988
		Promedio	**8,1481**	**17,7778**	**7,147%**

Ejemplo de cálculo promedio móvil y error para t = 4:

$$f_4 = \frac{200 + 240 + 220}{3} = 220 \quad ; \quad e_4 = 260 - 220 = 40$$

Ejemplo de cálculo promedio móvil y error para t = 5:

$$f_4 = \frac{240 + 220 + 260}{3} = 240 \quad ; \quad e_4 = 230 - 240 = -10$$



Pronóstico con suavizado exponencial   Alfa = 0,3

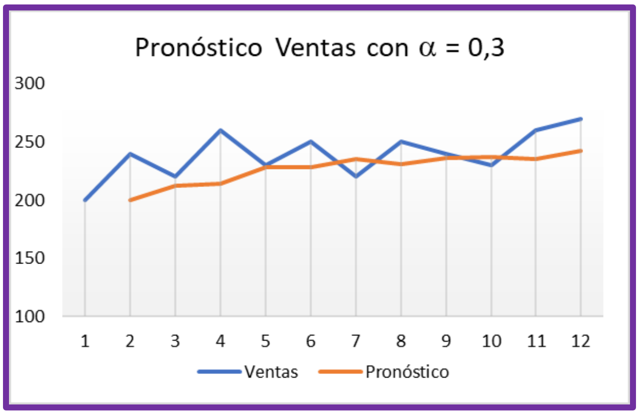
Mes	*Ventas*	*Pronóstico*	e_t	$\|e_t\|$	*MAD*	*MAPE*
1	200					
2	240	200,00	40,0000	40,0000	40,0000	0,1667
3	220	212,00	8,0000	8,0000	24,0000	0,0364
4	260	214,40	45,6000	45,6000	31,2000	0,1754
5	230	228,08	1,9200	1,9200	23,8800	0,0083
6	250	228,66	21,3440	21,3440	23,3728	0,0854
7	220	235,06	-15,0592	15,0592	21,9872	0,0685
8	250	230,54	19,4586	19,4586	21,6260	0,0778
9	240	236,38	3,6210	3,6210	19,3763	0,0151
10	230	237,47	-7,4653	7,4653	18,0520	0,0325
11	260	235,23	24,7743	24,7743	18,7240	0,0953
12	270	242,66	27,3420	27,3420	**19,5077**	0,1013
		Promedio	***15,4123***	***19,5077***		***7,8411%***

Cálculo del pronóstico para t = 4

$f_4 = 0,3(\ 220) + 0,7(\ 212) = 214,40$

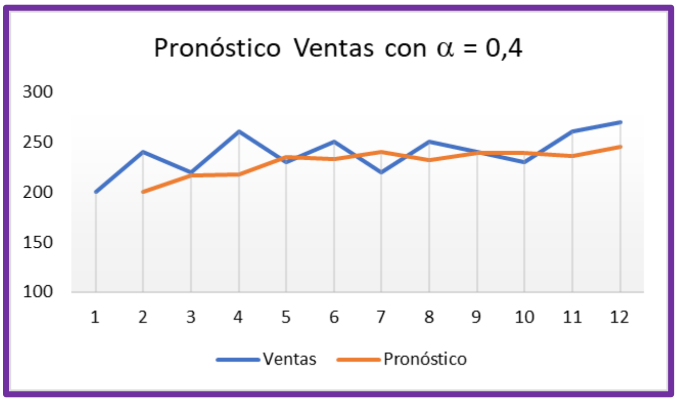
Cálculo del pronóstico para t = 5

$f_5 = 0,3(\ 260) + 0,7(\ 214,4) = 228,08$



Alfa = 0,4

Mes	Ventas	Pronóstico	e_t	$\lvert e_t \rvert$	MAD	MAPE
1	200					
2	240	200,00	40,0000	40,0000	40,0000	0,1667
3	220	216,00	4,0000	4,0000	22,0000	0,0182
4	260	217,60	42,4000	42,4000	28,8000	0,1631
5	230	234,56	-4,5600	4,5600	22,7400	0,0198
6	250	232,74	17,2640	17,2640	21,6448	0,0691
7	220	239,64	-19,6416	19,6416	21,3109	0,0893
8	250	231,78	18,2150	18,2150	20,8687	0,0729
9	240	239,07	0,9290	0,9290	18,3762	0,0039
10	230	239,44	-9,4426	9,4426	17,3836	0,0411
11	260	235,67	24,3344	24,3344	18,0787	0,0936
12	270	245,40	24,6007	24,6007	**18,6716**	0,0911
		Promedio	**12,5544**	**18,6716**		**7,5323%**



PRONÓSTICOS EN SERIES DE TIEMPO CON TENDENCIA Y/O ESTACIONALIDAD

Pronóstico en presencia de tendencia

Cuando la serie de datos de la variable de interés presenta una tendencia lineal, podemos utilizar el método conocido como *Proyección de Tendencia*. Este método ajusta una recta de tendencia a una serie de observaciones históricas y luego la utiliza para obtener pronósticos a mediano y largo plazo. Existen varias ecuaciones de tendencia que se pueden desarrollar de acuerdo con el comportamiento de los datos, como por ejemplo lineal, exponencial o cuadrática; sin embargo, solo analizaremos series de datos con tendencia lineal.

Una recta de tendencia es una ecuación de regresión con el tiempo como variable independiente.

La forma de la ecuación de la recta de tendencia lineal es:

$\hat{y} = b_0 + b_1 x$, donde

$\hat{y} =$ valor pronosticado

$b_0 =$ ordenada al origen

$b_1 =$ pendiente de la recta

$x =$ perido (es decir $x = 1,2,3.....n$)

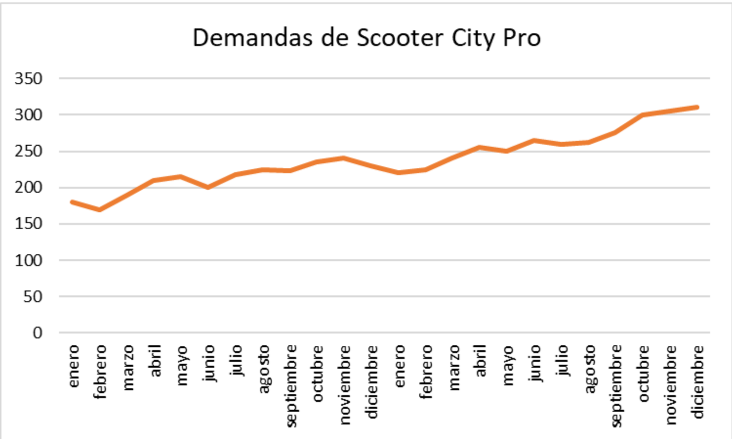
Se puede utilizar el método de regresión de mínimos cuadrados para encontrar los coeficientes que minimizan la suma de los cuadrados de los errores y de esta manera se minimiza el error cuadrático medio (ECM).

$$b_1 = \frac{\sum x_i y_i - \left(\frac{\sum x_i \sum y_i}{n}\right)}{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}} \qquad b_0 = \overline{y} - b_1 \overline{x}$$

Ejemplo:

MotorcycleCba, se dedica a la venta de motos y scooter eléctricos. En este momento necesita pronosticar la demanda para los próximos tres meses del scooter City Pro. En la tabla que se encuentra a continuación se presenta la información sobre la demanda mensual de los dos últimos años.

	Demanda en unidades	
	2021	**2022**
enero	180	220
febrero	170	225
marzo	190	240
abril	210	255
mayo	215	250
junio	200	265
julio	218	260
agosto	225	262
septiembre	223	275
octubre	235	300
noviembre	240	305
diciembre	230	310



Para estimar la recta de tendencia, el primer paso es ordenar los datos colocando un año a continuación del otro y reemplazando el nombre del mes por el número del periodo observado. Con los datos ya ordenados podemos aplicar las fórmulas para calcular los parámetros de la ecuación de la recta de tendencia.

x	y	x^2	xy
1	180	1	180
2	170	4	340
3	190	9	570
4	210	16	840
5	215	25	1075
6	200	36	1200
7	218	49	1526
8	225	64	1800
9	223	81	2007
10	235	100	2350
11	240	121	2640
12	230	144	2760
13	220	169	2860
14	225	196	3150
15	240	225	3600
16	255	256	4080
17	250	289	4250
18	265	324	4770
19	260	361	4940
20	262	400	5240
21	275	441	5775
22	300	484	6600
23	305	529	7015
24	310	576	7440
300	**5703**	**4900**	**77008**

$\bar{x}$ $\bar{y}$

12,5   237,625   ***n* =** 24

$$\frac{\sum x_i \sum y_i}{n} \quad 71287{,}5 \qquad \frac{\left(\sum x_i\right)^2}{n} \quad 3750$$

$b_1 =$ 4,97435

$b_0 =$ 175,4457

La pendiente de la recta de tendencia $b_1 = 4{,}97435$ implica que durante los 24 meses pasados las ventas de scooter City Pro han presentado un crecimiento promedio de 4,97435 unidades mensuales.

Con esta ecuación podemos estimar las demandas de los tres próximos meses reemplazando a x por los periodos 25, 26 y 27. Así las estimaciones de demanda serán:

Valor pronosticado para Enero $\quad y_{25} = 175{,}4457+4{,}97435*25 = 299{,}80445$

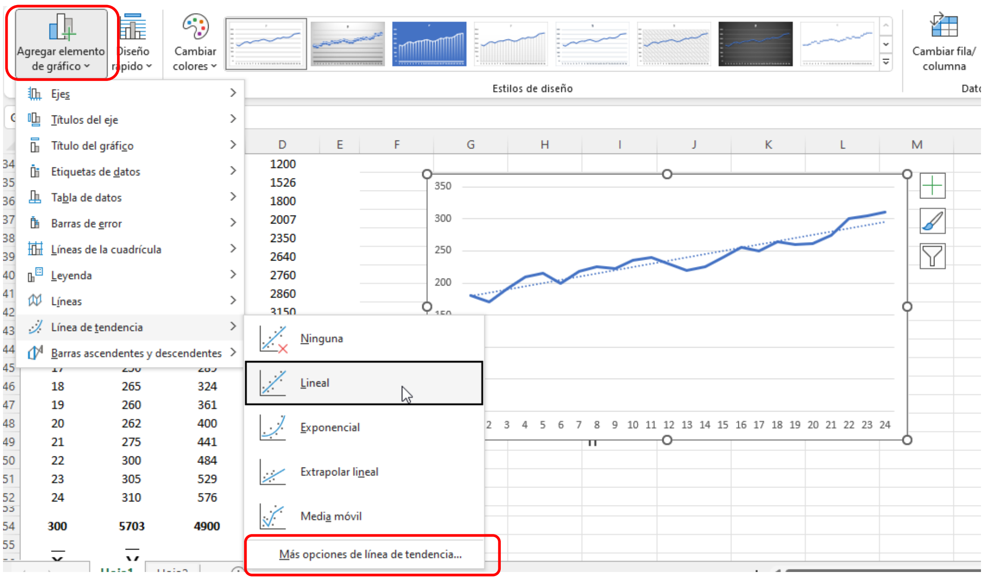
Valor pronosticado para Febrero y_{26} = 175,4457+4,97435*26 = 304,7788

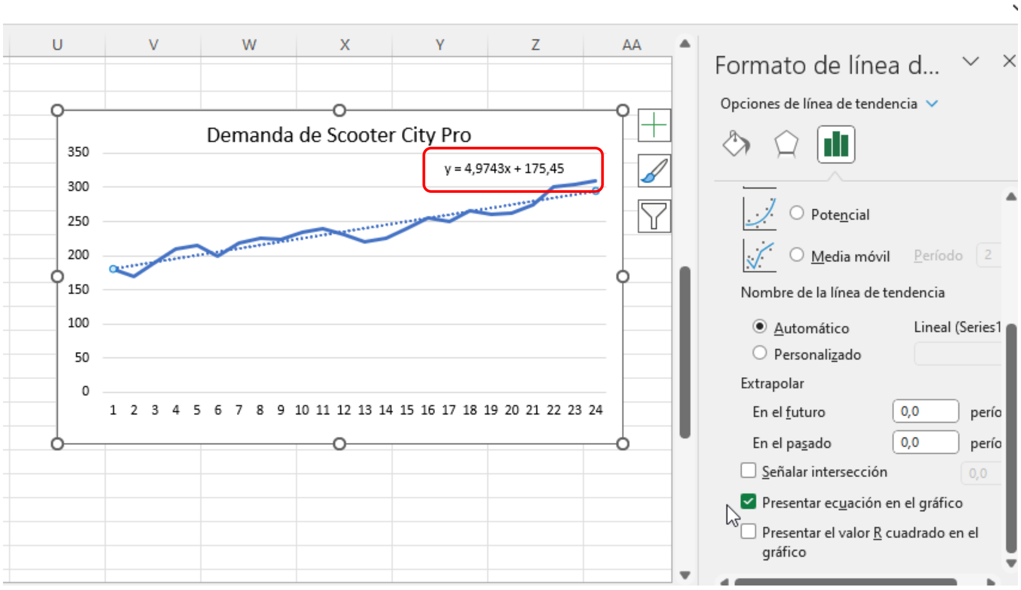
Valor pronosticado para Marzo y_{27} = 175,4457+4,97435*27 = 309,75315

Ejemplo usando Excel:

En una hoja de cálculo cargamos los datos, ya preparados con los meses como periodos. Primero vamos a "Insertar gráfico" y seleccionando el conjunto de datos insertamos un gráfico de dispersión, luego en "Agregar elementos al gráfico", solicitamos que incluya la recta de tendencia lineal y en "Más opciones de tendencia" le pedimos que inserte la ecuación de la recta.

x	y
1	180
2	170
3	190
4	210
5	215
6	200
7	218
8	225
9	223
10	235
11	240
12	230
13	220
14	225
15	240
16	255
17	250
18	265
19	260
20	262
21	275
22	300
23	305
24	310





Observa que la ecuación de la recta coincide con la que calculamos anteriormente.

PRONÓSTICO EN PRESENCIA DE TENDENCIA Y ESTACIONALIDAD

Índice estacional

La demanda de ciertos productos o servicios suele presentar picos durante algunos periodos de tiempo, por ejemplo, la demanda de calefactores suele ser mayor en los períodos fríos, en tanto la demanda de aire acondicionados lo es en los meses cálidos. Analizar los datos en términos de meses o trimestres facilita la detección de los patrones estacionales. Es frecuente utilizar un índice estacional en los modelos de pronósticos multiplicativos, para realizar un ajuste en la predicción, cuando existe una componente estacional.

La eliminación del efecto estacional de una serie se conoce como *desestacionalización*; eliminar este efecto, además de ser necesario para las comparaciones entre periodos, nos ayuda a identificar la componente de tendencia.

Un índice estacional indica la comparación de una estación dada (como mes o trimestre) y una estación promedio. Cuando no hay una tendencia, el índice se determina dividiendo el valor promedio para una estación específica entre el promedio de todos los datos. Así, un índice de 1 significa que la estación es promedio.

Este índice se puede calcular de varias maneras, por ejemplo, utilizando los promedios móviles centrados (PMC), los pasos para obtener los índices estacionales basados en los PMC pueden sintetizarse en:

1. Calcular el PMC para cada observación (cuando sea posible).

2. Calcular la Razón Estacional (RE) realizando el cociente entre cada observación y el PMC (observación/PMC para esa observación).

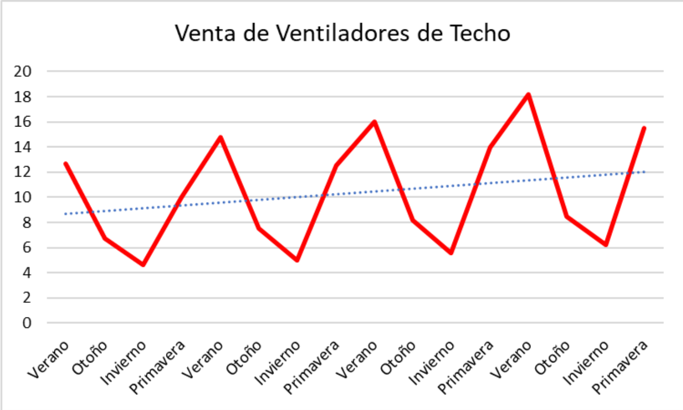
3. Promediar las razones estacionales para obtener los índices estacionales (IE).

4. Si la suma de los IE no es igual al número de estaciones (n), multiplicar cada índice por el cociente entre el número de estaciones y la suma de los IE ($n/\sum IE$).

Ejemplo:

La serie de datos de la tabla muestra las ventas de ventiladores de techo de los últimos 4 años agrupadas por trimestre.

		2019	**2020**	**2021**	**2022**
1	Verano	12,7	13,6	14,1	15
2	Otoño	6,7	6,5	6,9	7
3	Invierno	4,6	4,6	5	5,5
4	Primavera	10	9,8	10,4	10,8



Vamos a utilizar el método de Promedios Móviles Centrados para calcular los índices estacionales.

1. El primer paso es calcular el PMC para cada observación, para lo esto acomodamos los datos en una tabla.

Año	Trimestre	Ventas (miles $)	Promedio Móvil de 4 trimestres	Promedio Móvil Centrado
2019	1	12,7		
	2	6,7	8,50	
	3	4,6	9,03	8,76
	4	10	9,23	9,13
2020	1	14,8	9,33	9,28
	2	7,5	9,95	9,64
	3	5	10,25	10,10
	4	12,5	10,43	10,34
2021	1	16	10,58	10,50
	2	8,2	10,95	10,76
	3	5,6	11,50	11,23
	4	14	11,58	11,54
2022	1	18,2	11,73	11,65
	2	8,5	12,10	11,91
	3	6,2		
	4	15,5		

8,50 = (12,7+6,7+4,6+10)/4

9,03 = (6,7+4,6+10+14,8)/4

8,76 = (8,50+9,03)/2

9,13 = (9,03+9,23)/2

2. Calcular la Razón Estacional (RE) realizando el cociente entre cada observación y el PMC (observación/PMC para esa observación).

Año	Trimestre	Ventas (miles $)	Promedio Móvil Centrado	Razón Estacional
2019	1	12,7		
	2	6,7		
	3	4,6	8,76	0,525
	4	10	9,13	1,096
2020	1	14,8	9,28	1,596
	2	7,5	9,64	0,778
	3	5	10,10	0,495
	4	12,5	10,34	1,209
2021	1	16	10,50	1,524
	2	8,2	10,76	0,762
	3	5,6	11,23	0,499
	4	14	11,54	1,213
2022	1	18,2	11,65	1,562
	2	8,5	11,91	0,714
	3	6,2		
	4	15,5		

0,525 = 4,6/8,76

1,096 = 10/9,13

3. Promediar las razones estacionales para obtener los índices estacionales (IE)
Para hacer los cálculos en este paso, acomodamos las RE en una tabla como la original.

		2019	2020	2021	2022	IE
1	Verano		1,596	1,524	1,562	1,561
2	Otoño		0,778	0,762	0,714	0,751
3	Invierno	0,525	0,492	0,499		0,505
4	Primavera	1,096	1,209	1,213		1,173

4. Si la suma de los IE no es igual al número de estaciones (n), multiplicar cada índice por el cociente entre el número de estaciones y la suma de los IE (n/∑IE).

		IE	IE corregido
1	Verano	1,561	**1,565**
2	Otoño	0,751	**0,753**
3	Invierno	0,505	**0,507**
4	Primavera	1,173	**1,176**
	Suma	3,99	4
Factor de corrección		1,0025	

Otro método para calcular los índices estacionales

El índice estacional se calcula como el promedio de los índices estacionales de cada año considerado, es decir, calcular un índice estacional para cada año y luego promediar los índices de cada estación para eliminar la componente irregular.

Los pasos para obtener los índices estacionales basados en este método, sintetizarse en:

1. Calcular el promedio anual.

2. Dividir cada observación estacional por el promedio anual.

3. Promediar las observaciones de cada estación.

Ejemplo:

1. Calcular el promedio anual.

		2019	2020	2021	2022
1	Verano	12,7	13,6	14,1	15
2	Otoño	6,7	6,5	6,9	7
3	Invierno	4,6	4,6	5	5,5
4	Primavera	10	9,8	10,4	10,8
	Promedio	8,5	8,625	9,1	9,575

2. Dividir cada observación estacional por el promedio anual.

Índice est./año		2019	2020	2021	2022
1	Verano	1,4941	1,5768	1,5495	1,5666
2	Otoño	0,7882	0,7536	0,7582	0,7311
3	Invierno	0,5412	0,5333	0,5495	0,5744
4	Primavera	1,1765	1,1362	1,1429	1,1279

3. Promediar las observaciones de cada estación.

Índice est./año		2019	2020	2021	2022	*IE*
1	Verano	1,4941	1,5768	1,5495	1,5666	1,5467
2	Otoño	0,7882	0,7536	0,7582	0,7311	0,7578
3	Invierno	0,5412	0,5333	0,5495	0,5744	0,5496
4	Primavera	1,1765	1,1362	1,1429	1,1279	1,1459
						4,0000

Observación: los índices calculados por ambos métodos tienen pequeñas diferencias de decimales.

MÉTODO DE DESCOMPOSICIÓN DEL PRONÓSTICO CON COMPONENTES DE TENDENCIA Y ESTACIONAL

El primer paso para desestacionalizar una serie de tiempo es calcular los índices estacionales para cada estación. Luego, se elimina la estacionalidad de los datos dividiendo cada número entre su índice. A continuación, se estima la recta de tendencia usando los datos sin estacionalidad.

Con esta ecuación calculamos el pronóstico basado en la tendencia, y luego se lo ajusta con el índice estacional multiplicando cada valor por el índice estacional correspondiente.

En resumen, los pasos de este método son:

1. Calcular los índices estacionales usando alguno de los métodos vistos.
2. Eliminar la estacionalidad de los datos dividiendo cada número entre su índice estacional.
3. Encontrar la ecuación de la recta de tendencia empleando los datos sin estacionalidad.
4. Pronosticar para periodos futuros con la recta de tendencia.
5. Multiplicar el pronóstico de la recta de tendencia por el índice estacional correspondiente.

Ejemplo:

1. Calcular los índices estacionales usando alguno de los métodos vistos.

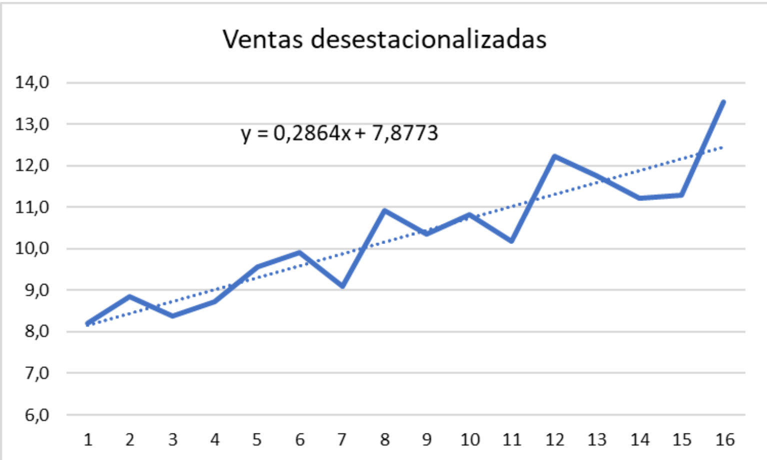
Vamos a desarrollar el ejemplo con los índices estacionales calculados por el segundo método.

3. Eliminar la estacionalidad de los datos dividiendo cada número entre su índice estacional. Para esto acomodamos los datos en otra tabla

Año	Trimestre	Ventas (miles $)	IE	Ventas desestacionalizadas
2019	1	12,7	1,5467	8,2
	2	6,7	0,7578	8,8
	3	4,6	0,5496	8,4
	4	10	1,1459	8,7
2020	1	14,8	1,5467	9,6
	2	7,5	0,7578	9,9
	3	5	0,5496	9,1
	4	12,5	1,1459	10,9
2021	1	16	1,5467	10,3
	2	8,2	0,7578	10,8
	3	5,6	0,5496	10,2
	4	14	1,1459	12,2
2022	1	18,2	1,5467	11,8
	2	8,5	0,7578	11,2
	3	6,2	0,5496	11,3
	4	15,5	1,1459	13,5

3. Encontrar la ecuación de la recta de tendencia empleando los datos sin estacionalidad.

A este paso lo realizamos usando Excel



4. Pronosticar para periodos futuros con la recta de tendencia.

Pronosticamos para los próximos 4 trimestres usando la ecuación de la recta de tendencia y así obtenemos la proyección de ventas, pero sin el efecto de la estacionalidad. Observa que en la ecuación de la recta se debe reemplazar el número del trimestre por el número del periodo a pronosticar.

			Ventas desestacionalizadas
2023	1	17	12,7461
	2	18	13,0325
	3	19	13,3189
	4	20	13,6053

Pronóstico para el primer trimestre de 2023 considerando solo el efecto de la tendencia, se calcula como:

12, 7461 = 0,2864 (17) + 7,8773

5. Multiplicar el pronóstico de la recta de tendencia por el índice estacional correspondiente.

			Ventas desestacionalizadas	IE	Ventas pronosticadas
2023	1	17	12,7461	1,5467	19,71
	2	18	13,0325	0,7578	9,88
	3	19	13,3189	0,5496	7,32
	4	20	13,6053	1,1459	15,59

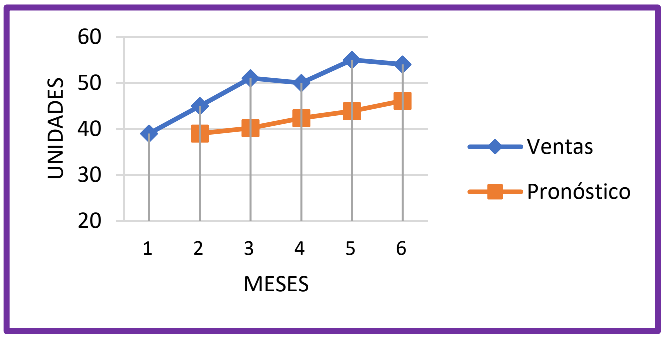
Multiplicando el pronóstico de ventas desestacionalizadas por el índice estacional correspondiente obtenemos el pronóstico con los efectos de tendencia, estacionalidad e irregular.

SUAVIZADO EXPONENCIAL CON TENDENCIA: *MÉTODO DE HOLT*

Veamos el siguiente ejemplo

α = 0,20

Mes	*Ventas*	*Pronóstico*
1	39	
2	45	39
3	51	40
4	50	42
5	55	44
6	54	46



Si utilizamos como método de pronóstico un suavizado exponencial con un α = 0,2 se ve claramente que el pronóstico siempre queda retrasado con respecto a los valores reales. Al analizar la serie de datos, podemos darnos cuenta de que las ventas presentan un crecimiento sostenido, esto es una tendencia lineal ascendente. En estos casos, para un mejor ajuste, deberíamos utilizar un método que tenga en cuenta esta componente.

Siempre que existe una tendencia continuada, la suavización exponencial permanecerá rezagada en el tiempo con respecto a los valores reales.

El método de suavizado exponencial simple, supone que el nivel de las series de tiempo varía ocasionalmente, y solo necesita una estimación del nivel actual.

En algunas situaciones, los datos observados tienen una tendencia clara y contienen información que permite anticipar movimientos futuros ascendentes. Cuando éste es el caso, se necesita una función que estime la tendencia lineal del pronóstico. Es decir que se requiere una estimación del nivel de base, así como de la pendiente actual.

Supongamos que estamos en el periodo *t* = 4 y luego de observar x_4 suponemos que *la línea o nivel de base* (L_t) es L_4 = 48 y que la *tendencia* (T_t) o *pendiente de la línea de tendencia* es T_4 = 5, es decir que el nivel de base se incrementa en cinco unidades por periodo, entonces estimamos que el valor pronosticado para el periodo *t* = 6 será de 58 unidades

$$f_5 = 48 + (2)5 = 58 \quad \text{unidades}$$

Si f_{t+k} el pronóstico para x_{t+k}, realizado al final del periodo **t**, entonces

$$f_{t+k} = L_t + kT_t$$

Si el pronóstico es para el periodo inmediato siguiente entonces k = 1

Un método apropiado para este tipo de series con tendencia lineal y sin estacionalidad es el método de Holt, el que explícitamente permite incluir una tendencia lineal ascendente o descendente. Al final del periodo *t* este método genera una estimación del nivel de base L_t y de la tendencia T_t de la serie.

Luego de observar x_t podemos actualizar la estimación del nivel de base y la tendencia utilizando suavizado exponencial.

Actualización de la línea de base ***L_t***

El nivel de base actual L_t se calcula tomando un promedio ponderado por los pesos α y (1-α) de dos estimaciones de nivel: una estimación está dada por la observación actual x_t y la otra estimación está dada por la suma de la tendencia previa T_{t-1} y el nivel previamente actualizado L_{t-1}, en ausencia de tendencia no es necesario el término T_{t-1}. Observe que ($L_{t-1} + T_{t-1}$) en definitiva es f_{t-1}

$$L_t = \alpha x_t + (1 - \alpha)\ (L_{t-1} + T_{t-1})$$

como $f_t = L_{t-1} + T_{t-1}$

reemplazamos a f_t $L_t = \alpha x_t + (1 - \alpha) f_t$

Actualización de tendencia ***T**_t*

Para la actualización de la estimación de la tendencia también se utiliza suavizado exponencial, con una constante de suavizado β.

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\ (T_{t-1})$$

Esta ecuación indica que la tendencia actual es un promedio ponderado por los pesos β y ($1-\beta$), de dos tendencias estimadas, una está dada por el cambio en el nivel de base entre los periodos *t-1* y *t* es decir ($L_t - L_{t-1}$) y la otra es la tendencia calculada en el periodo *t-1.*

Para elegir las constantes de suavizado α y β, pueden ensayarse varias combinaciones de ellas y optar por la que minimice el MAD. Si estos valores son mayores a 0,5 debería usarse otro método de pronóstico ya que esto estaría indicando la presencia de estacionalidad o comportamiento cíclico.

Pesos grandes tienen como resultado cambios más rápidos en el componente; pesos pequeños tienen como resultado cambios menos rápidos. Por lo tanto, cuanto mayores son los pesos, los valores de suavizado seguirán más a los datos; cuanto menores son los pesos, los valores de suavizado seguirán más a los valores de suavizado previos.

Para aplicar el método, necesitamos una estimación inicial de la base (L_0) y una estimación inicial de la tendencia (T_0).

Un enfoque consiste en fijar la primera estimación del nivel de base suavizado igual a la primera observación y la tendencia igual a cero. Un segundo enfoque es usar el promedio de las primeras cinco o seis observaciones como el valor suavizado inicial de L. Luego, se estima la tendencia usando la pendiente de una línea ajustada a estas cinco o seis observaciones.

También se puede usar como estimación de L_0 a la observación del último periodo y T_0 como el incremento promedio por periodo de la serie.

Utilicemos Holt con los datos del ejemplo

Si suponemos $L_0 = 35$ y $T_0 = 2$

$\alpha = 0,4$ $\beta = 0,1$

Mes	*Ventas*	f_t	L_t	T_t	f_{t+1}	$\lvert e_t \rvert$	*MAD*
0			35	2	37		
1	39	37	37,8	2,08	39,88	2	2
2	45	39,88	41,93	2,28	44,21	5,12	3,56
3	51	44,21	46,93	2,56	49,48	6,79	4,64
4	50	49,48	49,69	2,58	52,27	0,52	3,61
5	55	52,27	53,36	2,69	56,05	2,73	3,43
6	54	56,05	55,23	2,60	57,83	2,05	**3,20**

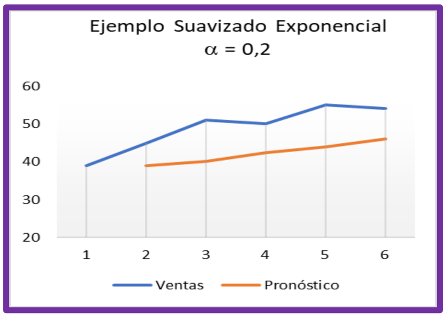
Veamos los cálculos de L_t y T_t para los periodos 1 y 2

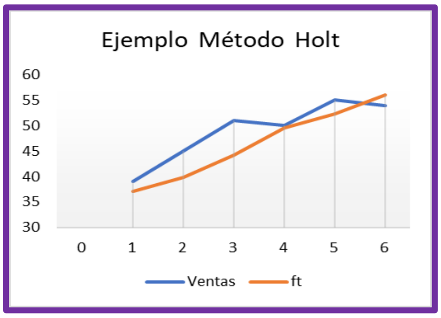
$L_1 = 0,4(39) + (1 - 0,4)\ (35 + 2)$ =0,4(39)+0,6(37)=37,8

$T_1 = 0,1(37,8 - 35) + (1 - 0,1)\ (2)$ =0,1(2,8)+0,9(2)=2,08

$L_2 = 0,4(45) + (1 - 0,4)\ (37,8 + 2,08)$ =0,4(45)+0,6(39,88)=41,93

$T_2 = 0,1(41,93 - 37,8) + (1 - 0,1)\ (2,08)$ =0,1(4,13)+0,9(2,08)=2,28





El método de Holt proporcionará buenos pronósticos de una serie, cuando la componente de tendencia sea lineal. Existe una versión multiplicativa del método de Holt que se puede utilizar para generar buenas previsiones para una serie de la forma $x_t = ab^t\ \varepsilon_t$.
Dónde:

a ➔ representa el nivel de base

b ➔ Actualiza el nivel de base en cada periodo para un crecimiento expresado en porcentaje, es decir que si estimamos que el crecimiento del nivel de base de la serie será del 5% por período entonces *b* = 1,05

ε_t ➔ factor aleatorio de error con media igual a 1.

Ejemplo de la fábrica de zapatillas con el Método de Holt

Usamos este método con la misma serie del ejemplo de la fábrica de zapatillas. Para comenzar necesitamos una estimación para la base (L_0), una estimación para la tendencia (T_0) y definir las constantes de suavizado α y β. Vamos a suponer que en el último mes del año anterior se vendieron 200 unidades y que el promedio de los incrementos es de 2,3 unidades y seleccionamos un α = 0,3 y β = 0,1

			ALFA	***BETA***					
			0,3	0,1					
Mes	***Ventas***	f_t	L_t	T_t	f_{t+1}	e_t	$\lvert e_t \rvert$	***MAD***	***MAPE***
0			200	2,30	202,30				
1	200	202,30	201,61	2,23	203,84	-2,30	2,30	2,3	0,01
2	240	203,84	214,69	3,32	218,00	36,16	36,16	19,23	0,15
3	220	218,00	218,60	3,38	221,98	2,00	2,00	13,48	0,01
4	260	221,98	233,39	4,52	237,90	38,02	38,02	19,62	0,15
5	230	237,90	235,53	4,28	239,81	-7,90	7,90	17,28	0,03
6	250	239,81	242,87	4,58	247,45	10,19	10,19	16,09	0,04
7	220	247,45	239,22	3,76	242,98	-27,45	27,45	17,72	0,12
8	250	242,98	245,08	3,97	249,06	7,02	7,02	16,38	0,03

9	240	249,06	246,34	3,70	250,04	-9,06	9,06	15,57	0,04
10	230	250,04	244,03	3,10	247,13	-20,04	20,04	16,01	0,09
11	260	247,13	250,99	3,49	254,47	12,87	12,87	15,73	0,05
12	270	254,47	259,13	3,95	263,08	15,53	15,53	***15,71***	0,06
					Promedio	**4,59**	**15,71**		**0,06**
									6%

Veamos el cálculo del pronóstico para t = 1 y t = 2

$$L_1 = 0,3(200) + 0,7(200 + 2,3) = 201,61$$

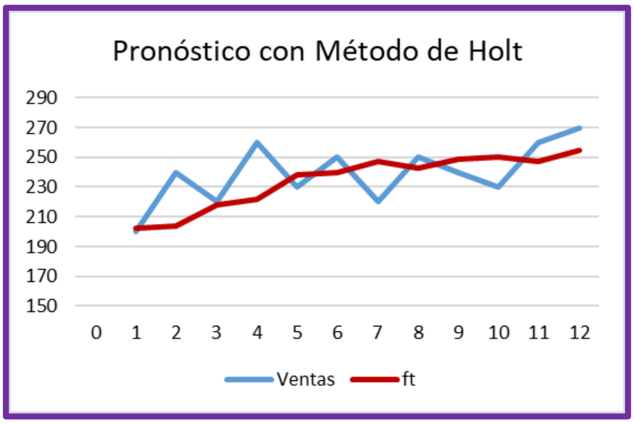
$$T_1 = 0,1(201,61 - 200) + 0,9(2,3) = 2,23$$

$$f_2 = 201,61 + 2,23 = 203,84$$

$$L_2 = 0,3(240) + 0,7(201,61 + 2,23) = 214,69$$

$$T_2 = 0,1(214,69 - 201,61) + 0,9(2,23) = 3,32$$

$$f_3 = 214,69 + 3,32 = 218,01$$

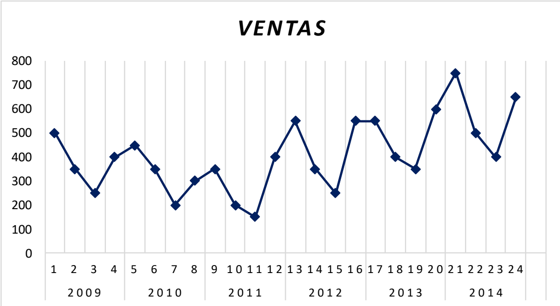


Analizando los errores de pronóstico, podemos decir que para esta serie el método que mejor ajusta es el de suavizado exponencial con tendencia o método de Holt.
Podemos también encontrar los valores de α y β que minimicen el MAD a través de una optimización no lineal.

Suavizado exponencial con estacionalidad: *Método de Winter*

La tabla siguiente indica la venta de automóviles en miles de unidades, una revisión de los datos muestra que las ventas son más altas durante el primero y el cuarto trimestres y más bajas durante el tercer trimestre de manera sistemática. Esto podría indicar la existencia de un patrón estacional.

Trimestre	***2009***	***2010***	***2011***	***2012***	***2013***	***2014***
1	500	450	350	550	550	750
2	350	350	200	350	400	500
3	250	200	150	250	350	400
4	400	300	400	550	600	650



El método de *Winters* podría representar mejor los datos y reducir el error de pronóstico. Este método es una extensión del método de *Holt* e introduce una ecuación adicional para estimar la estacionalidad (S_t).

Al final del periodo t la predicción $f_{t,k}$ para el periodo t+k se obtiene con

$$f_{t+k} = \left(L_t + kT_t\right) s_{t+k-c}$$ donde c es el número de periodos en el patrón estacional

En la versión multiplicativa del método de *Winters*, la estimación de la estacionalidad está dada por un índice estacional (S_t) y se calcula como:

$$s_t = \gamma \; \frac{x_t}{L_t} \; + (1 - \gamma) \; s_{t-c}$$

Es decir que el componente estacional actual S_t surge del producto entre una constante de suavizado (γ) y un estimado del índice estacional dado por x_t / L_t , al que se le suma (1-γ) veces el componente estacional previo S_{t-c}.

El valor observado x_t se divide entre el nivel actual estimado L_t, para crear un índice que pueda usarse de forma multiplicativa para ajustar un pronóstico tomando en cuenta los picos y valles estacionales.

Las ecuaciones del método son:

$$L_t = \alpha \left(\frac{x_t}{s_{t-c}} \right) + (1 - \alpha) \; (L_{t-1} + T_{t-1}) \quad \}$$

Actualiza la estimación de la base

Observación desestacionalizada

$$T_t = \beta (L_t - L_{t-1}) + (1 - \beta) \; (T_{t-1}) \quad \}$$

Actualiza la tendencia

$$s_t = \gamma \left(\frac{x_t}{L_t} \right) + (1 - \gamma) \left(s_{t-c} \right) \quad \}$$

Actualiza la estacionalidad

Estimación más reciente de la estacionalidad del periodo t

Estimación de la estacionalidad del periodo t obtenida a partir del periodo actual

El pronóstico para k periodos futuros, $f_{t,k}$, se obtiene con:

$$f_{t+k} = (L_t + kT_t) s_{t+k-c}$$

Producto de la estimación para el periodo (t+k) por la estimación de la estacionalidad del periodo (t+k)

Donde:

c ➔ número de periodos en la duración del patrón estacional o longitud de la estacionalidad. Por ejemplo, si es trimestral c = 4, si es mensual c = 12

s_t ➔ factor multiplicativo estacional para el periodo t, obtenido después de observar x_t.

L_t ➔ nivel de base

T_t ➔ tendencia

Al igual que en la suavización exponencial lineal de Holt, los pesos α, β y γ pueden seleccionarse subjetivamente o generarse al minimizar una medida de error de pronóstico. También se puede emplear un algoritmo de optimización para obtener las constantes óptimas de suavización, por ejemplo, podría utilizar el módulo Solver de Excel.
Para iniciar el algoritmo, se deben establecer los valores iniciales para las series suavizadas (nivel de base) L_t, la tendencia T_t y los índices estacionales S_t.
Existen varios métodos para la inicialización de la estimación del nivel de base, tendencia y estacionalidad.

Índice estacional

Una forma de estimar el factor o índice estacionales inicial consiste en dividir el promedio de las observaciones del periodo estacional entre el promedio general de las observaciones.

Tendencia

La estimación inicial de la tendencia (T_0) la calculamos como la diferencia del promedio de ventas de los dos últimos años, dividido por el número de periodos estacionales, esto es:

$$T_0 = \frac{\bar{V}_{-1} - \bar{V}_{-2}}{n} \quad \text{dónde}$$

$\bar{V}$: ventas promedio del año

n: número de periodos estacionales

Nivel inicial de base

En cuanto al cálculo del nivel inicial de base (L_0), existen en la bibliografía diferentes propuestas, nosotros lo calcularemos como:

$$L_0 = \bar{V}_{-1} + \frac{n-1}{2}\,\mathrm{T}_0 \quad \text{dónde}$$

$\bar{V}_{-1}$:ventas promedio del año anterior

n: número de periodos estacionales

<u>Por ejemplo</u>:

- Estacionalidad trimestral

$$L_0 = \overline{V}_{-1} + \frac{4-1}{2}\,T_0 = \overline{V}_{-1} + 1{,}5\,T_0$$

$\overline{V}_{-1}$:ventas promedio del año anterior

n: número de periodos estacionales, en este caso el 4 trimestres

- Estacionalidad bimestral

$$L_0 = \overline{V}_{-1} + \frac{6-1}{2}\,T_0 = \overline{V}_{-1} + 2{,}5\,T_0$$

$\overline{V}_{-1}$:ventas promedio del año anterior

n: número de periodos estacionales, en este caso el 6 bimestres

- Estacionalidad mensual

$$L_0 = \overline{V}_{-1} + \frac{12-1}{2}\,T_0 = \overline{V}_{-1} + 5{,}5\,T_0$$

$\overline{V}_{-1}$:ventas promedio del año anterior

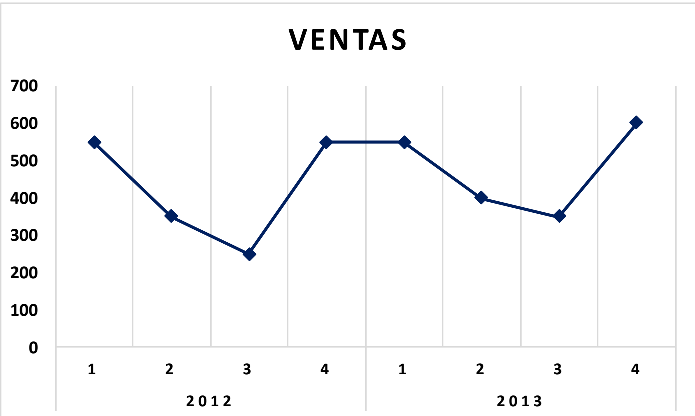
n: número de periodos estacionales, en este caso 12 meses

Otro método consiste en igualar la primera estimación de la serie suavizada (nivel) a la primera observación. Luego, la tendencia se iguala a cero y cada índice estacional se fija en uno (1).

Ejemplo con Método de Winter

Volvamos a nuestro ejemplo de la venta de autos. De todos los datos vamos a usar los de los años 2012 y 2013 para estimar los parámetros iniciales y calcularemos los pronósticos de ventas para los años 2014 y 2015.

Año	*t*	*Ventas*
2012	1	550
	2	350
	3	250
	4	550
2013	1	550
	2	400
	3	350
	4	600



Para iniciar el proceso necesitamos los índices estacionales y una estimación para el nivel de base (L_0) y la tendencia (T_0).

1. **Cálculo del índice estacional**

El índice estacional trimestral se calcula como el promedio de los índices estacionales de cada año considerado. En nuestro caso el índice estacional del trimestre 1 del año 2012 es igual al cociente entre las ventas del primer trimestre y el promedio de ventas anuales, si las ventas del trimestre están por encima del promedio anual el índice será mayor a uno de lo contrario el índice será menor a uno.

550/425 = 1,2941

(1,2941 + 1,1579)/2 = 1,2260

Trimestre	2012	2013	Índice Trim 1	Índice Trim 2	Índice Trimestral	
1	550	550	1,2941	1,1579	**1,2260**	S_{-3}
2	350	400	0,8235	0,8421	**0,8328**	S_{-2}
3	250	350	0,5882	0,7368	**0,6625**	S_{-1}
4	550	600	1,2941	1,2632	**1,2786**	S_0
Promedios	**425**	**475**				

2. **Cálculo de la Tendencia inicial**

Como T_0 vamos a considerar la diferencia entre el promedio de ventas de los dos años divido 4 (número de trimestres del año).

$T_0 = (475 – 425)/4 = 12,5$

3. **Cálculo del Nivel de Base inicial**

Nivel de base inicial L_0 lo calculamos como el promedio de ventas del año 2013 más la estimación de la tendencia trimestral multiplicada por 1,5 ya que las observaciones promedio se dan en mitad de cada trimestre.

$$\boxed{L_0 = \overline{V}_{-1} + \frac{4-1}{2}\,T_0 = \overline{V}_{-1} + 1{,}5\,T_0} \qquad L_0 = 475 + 12{,}5\,(1{,}5) = 493{,}75$$

Para las constantes de suavizado usaremos los valores: $\alpha = 0{,}2$ $\beta = 0{,}1$ y $\gamma = 0{,}7$

				Alfa	Beta	Gama				
				0,1	0,5	0,9	1,2260			
Año	**Trimestre**	**t**	x_t	f_t	L_t	T_t	S_t	f_{t+1}	e_t	$\lvert e_t \rvert$
		0			493,75	12,50	1,226	620,67		
2014	1	1	750	620,67	516,80	17,77	1,429	445,20	129,33	129,33

$$f_{0+1} = f_1 = \left(L_0 + T_0\right) s_{-3} \qquad f_1 = (493{,}75 + 12{,}50)\ 1{,}226 = 620{,}67$$

$$L_1 = \alpha\ \frac{x_1}{s_{1-4}} + (1-\alpha)\ (L_{1-1} + T_{1-1}) = \alpha\ \frac{x_1}{s_{-3}} + (1-\alpha)\ (L_0 + T_0) = 0{,}1\left(\frac{750}{1{,}226}\right) + 0{,}9\left(493{,}75 + 12{,}5\right) = 516{,}799$$

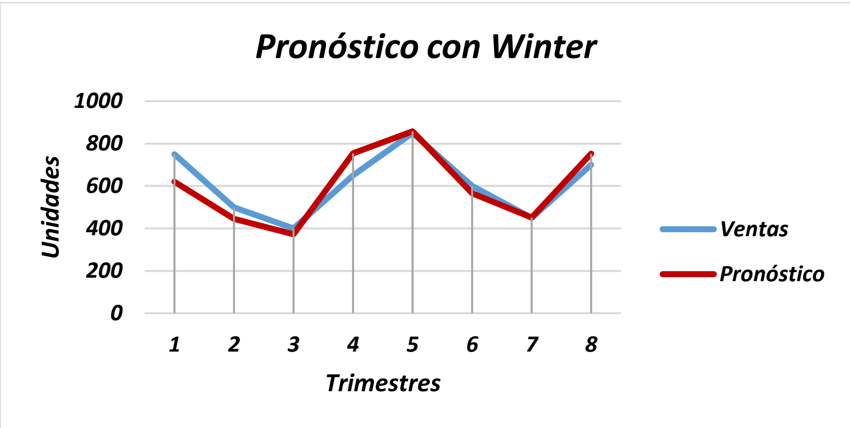
$$T_1 = \beta(L_1 - L_0) + (1-\beta)\ (T_0) = 0{,}5\ (516{,}80 - 493{,}75{,}75) + 0{,}5\ (12{,}5) = 17{,}77$$

$$s_1 = \gamma \; \frac{x_1}{L_1} \; + (1-\gamma)\; s_{1-4} \; = 0,9 \left(\frac{750}{516,80} \right) + 0,1\;(1,226) = 1,429$$

Utilizando las fórmulas complete los valores faltantes en la tabla:

				Alfa	***Beta***	***Gama***				
				0,2	0,1	0,9				
Año	***Trimestre***	***t***	***x_t***	***f_t***	***L_t***	***T_t***	***S_t***	***f_{t+1}***	***e_t***	***$\|e_t\|$***
		0			493,75	12,50	1,226	620,67		
2014	1	1	750	620,67	516,80	17,77	1,429	445,20	129,33	129,33
	2	2	500	445,20						
	3	3	400		566,37	23,14	0,702	753,77	27,51	27,51
	4	4	650	753,77	581,39	19,08	1,134	857,91	-103,77	103,77
2015	1	5	850	857,91	599,92	18,81	1,418	566,04	-7,91	7,91
	2	6	600	566,04	622,44	20,66	0,959	451,38	33,96	33,96
	3	7	450	451,38	642,91	20,56	0,700			
	4	8	700							
2016	1	9								
									MAD	

El gráfico muestra el ajuste entre los valores reales de venta y los pronosticados.



Señal de Control

Es útil desarrollar una medida que nos alerte respecto a un cambio en el patrón básico de una serie de tiempo. Un sistema de control contiene un rango de desviaciones permitidas del pronóstico con respecto a los valores reales.

Un procedimiento para dar seguimiento a una técnica de pronósticos es determinar el rango en el que deben estar los errores de pronóstico. Si la técnica utilizada es razonablemente exacta, el error de pronóstico debe tener una distribución aproximadamente normal alrededor de una media igual a cero. Con esta condición, hay un 95% de probabilidad de que la observación real caiga dentro de aproximadamente dos desviaciones estándar del pronóstico. Podemos utilizar el MAD para estimar la desviación estándar de los errores del pronóstico con la relación *Se = 1,25 MAD*

Si usamos el MAD, como control podremos decir que las unidades vendidas, con un 0,95 de probabilidad, estarán entre el valor pronosticado ± (2)1,25(MAD). Los errores de pronóstico que caigan dentro de estos límites indican que no hay motivo de alarma. Los errores, particularmente una secuencia de errores, fuera de los límites sugieren un cambio.

*Una **señal de control** implica el cálculo de una medida de errores de pronóstico en el tiempo y el establecimiento de límites, de manera que, cuando los errores acumulados caen fuera de esos límites, el encargado de efectuar el pronóstico recibe una señal de alerta.*

III. Métodos Causales

Como ya lo anticipamos, los métodos de pronósticos Causales se basan en el supuesto de que la variable que tratamos de pronosticar muestra una relación de causa y efecto con una o más variables. Estos métodos procuran estimar la relación entre la variable dependiente a pronosticar y una o más variables independientes y se utilizan, en general, cuando hay pocos datos históricos o no los hay.

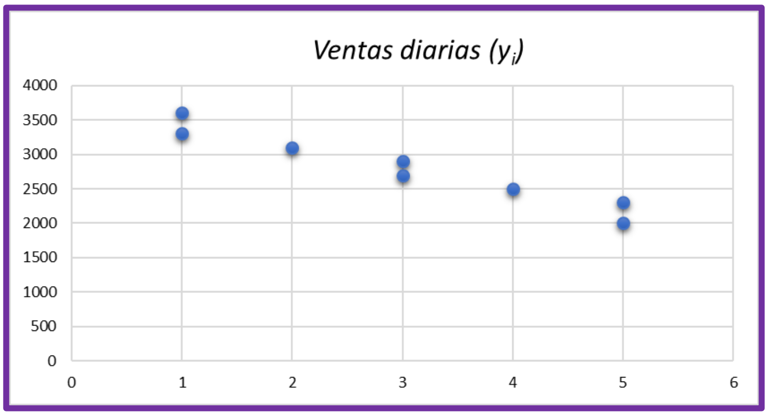
Se puede emplear como método de elaboración de pronósticos causal el *análisis de regresión* para desarrollar una ecuación matemática que relacione la variable que queremos pronosticar con otras variables que se supone la influyen o la explican.

Cuando el análisis de regresión involucra una variable independiente y una variable dependiente y la relación entre ellas se aproxima por medio de una recta se llama *regresión lineal simple.* Si el análisis incluye dos o más variables independientes se conoce como *análisis de regresión múltiple.*

Ejemplo con análisis de regresión simple

Gonzalo está por lanzar una franquicia de medialunas y reunió datos sobre la relación entre el volumen de venta diaria medido en $, y el número de panaderías que ofrecen productos similares en un radio de 700 mts. Para realizar el estudio reunión los datos siguientes:

N° de panaderías en de 700 mts	Ventas diarias en ($)
1	3600
1	3300
2	3100
3	2900
3	2700
4	2500
5	2300
5	2000



Con estos datos podemos estimar, a través del método de mínimos cuadrados, la relación que existe entre las panaderías cercanas y el volumen de ventas y luego utilizarlo para predecir el volumen de ventas del local donde se pretenda montar una franquicia, de acuerdo con la cantidad de panaderías situadas en un radio de 700 mts.

Recordemos que la ecuación de la recta de regresión estimada es:

$$\boxed{\hat{y} = b_0 + b_1 x}$$

Donde:

y = valor estimado de la variable dependiente, en nuestro caso son las ventas diarias.

b_0 = ordenada al origen de la ecuación de regresión estimada

b_1 = pendiente de la ecuación de regresión estimada

x = valor de la variable independiente, para nuestro caso representa el número de panaderías situadas en un radio 700mts.

Utilizamos datos muestrales y las ecuaciones siguientes para calcular la ordenada al origen b_0 y la pendiente b_1:

$$b_1 = \frac{\sum x_i y_i - \left(\frac{\sum x_i \sum y_i}{n} \right)}{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}} \qquad b_0 = \bar{y} - b_1 \bar{x}$$

Donde:

x_i = valor de la variable independiente para la i-ésima observación

y_i = valor de la variable dependiente para la i-ésima observación

$\bar{x}$ = valor medio de la variable independiente

$\bar{y}$ = valor medio de la variable dependiente

n = número de observaciones

n	N° de panaderías (x_i)	Ventas diarias (y_i)	$x_i y_i$	x_i^2
1	1	3600	3600	1
2	1	3300	3300	1
3	2	3100	6200	4
4	3	2900	8700	9
5	3	2700	8100	9
6	4	2500	10000	16
7	5	2300	11500	25
8	5	2000	10000	25
Totales	**24**	**22400**	**61400**	**90**

$\sum x_i = 24 \quad \sum y_i = 22400 \quad \sum x_i y_i = 61400 \quad \sum x_i^2 = 90 \quad \left(\sum x_i\right)^2 = 576$

$\sum x_i \sum y_i = 537.600 \quad \bar{y} = 2.800 \quad \bar{x} = 3 \quad n = 8$

Con esto, calculamos los valores de b_0 y b_1 y escribimos la ecuación de la recta de regresión:

$$\hat{y} = 3766,667 + (-322,222)x$$

Si por ejemplo estamos analizando una posible zona para la localización de un nuevo local y en un radio de 700 mts. identificamos 4 panaderías, podemos estimar las ventas diarias en:

$$y = 3766,667 + (-322,222) * 4 = \$2477,77$$

Métodos Cualitativos de Pronóstico

Los modelos cuantitativos requieren datos históricos sobre la variable que se quiere predecir, los que muchas veces no se encuentran disponibles. Además, estos procedimientos suponen que las condiciones del pasado se mantendrán en el futuro.

Esto tiene como consecuencia que un cambio significativo en las condiciones del entorno que afectan la serie de tiempo puede hacer cuestionable el uso de los datos históricos en la predicción de valores futuros de las series de tiempo.

Los métodos de elaboración de pronósticos cualitativos se consideran adecuados cuando se espera que el patrón de comportamiento histórico de las series de tiempo no continúe en el futuro.

El uso de técnicas de pronóstico cualitativo aumenta cuando los datos históricos son escasos o cuando se cree que éstos son parcialmente irrelevantes. En casos extremos, tal vez el analista piense que los datos históricos no son directamente relevantes para el proceso de pronóstico.

Método Delphi

Cuando se reúne a un conjunto de expertos en una sala y se les solicita un análisis acerca del futuro, la dinámica de grupo puede distorsionar el proceso de discusión y dar como resultado un consenso que no es pensado cuidadosamente por todos los participantes.

En respuesta a esta situación, el método Delphi intenta elaborar pronósticos por medio de un "consenso de grupo".

La aplicación del método Delphi comienza solicitándoles a los integrantes de un panel de expertos, que están separados físicamente y no se conocen entre sí, que respondan una serie de preguntas.

Luego, las opiniones de cada experto se resumen y cada miembro del grupo es informado de las posiciones de sus colegas anónimos. Esta retroalimentación contiene las opiniones y los juicios de todos los miembros del grupo y no sólo las de los más elocuentes.

A continuación, se le solicita a cada uno de los participantes que reconsideren y analicen su respuesta previa a la luz de la información proporcionada por el grupo. Este proceso continúa hasta que el coordinador considera que se ha llegado a cierto grado de consenso.

El objetivo del método Delphi es obtener una pluralidad de opiniones con las cuales coincida la mayoría de los expertos.

Cualquier procedimiento Delphi tiene cuatro características fundamentales: anonimato, iteración, retroalimentación controlada y agregación de respuestas del grupo.

La ventaja del método es que a los expertos notables se les puede pedir que consideren cuidadosamente el tema de interés y que contesten reflexivamente a los puntos de vista de otros sin la interferencia de la dinámica de grupo.

Juicio experto

El juicio experto es un método de elaboración de pronósticos que se recomienda, en general, cuando no es probable que las condiciones pasadas se mantengan en el futuro.

En este caso cada experto considera la información que cree influirá en la variable de interés que se quiere predecir, luego combinan sus conclusiones en un pronóstico. No se utiliza un modelo formal, y no es probable que dos expertos consideren la misma información de la misma manera.

Aun cuando no se utiliza un modelo cuantitativo formal, el juicio experto proporciona buenos pronósticos en muchas situaciones.

Redacción de escenarios

Este procedimiento cualitativo consiste en la elaboración de escenarios conceptuales del futuro con base en un conjunto de supuestos o hipótesis bien definidas. Estos conjuntos de hipótesis proporcionan escenarios distintos. La nueva tecnología, los cambios demográficos y el cambio en las demandas del consumidor son, entre otros, los factores que se consideran y entrelazan.
Se establece cuál es el escenario más probable, junto con uno o más escenarios menos probables, pero posibles.
Al proceso de describir el escenario más probable le sigue una etapa de debate, algunas veces realizado por un grupo diferente del que desarrolló los escenarios

Elección de una técnica adecuada

Algunas de las preguntas que deben considerarse, antes de decidir sobre la técnica más adecuada para la elaboración del pronóstico de un problema específico pueden ser:

- ¿Por qué es necesario un pronóstico?
- ¿Quién usará el pronóstico?
- ¿Cuáles son las características de los datos disponibles?
- ¿Qué periodo se va a pronosticar?
- ¿Cuáles son los requerimientos mínimos de datos?
- ¿Qué precisión se desea?
- ¿Cuánto costará el pronóstico?

Para elegir acertadamente la técnica de pronóstico, el analista debe hacer lo siguiente:

- Definir la naturaleza del problema que se va a pronosticar.
- Explicar la naturaleza de los datos en investigación.
- Describir las capacidades y limitaciones de las técnicas de elaboración de pronósticos potencialmente útiles.
- Desarrollar algún criterio predeterminado, con el cual se pueda tomar la decisión.

Un factor importante que influye en la selección de la técnica de elaboración del pronóstico es la identificación y comprensión de patrones históricos en los datos. Si se pueden reconocer patrones de tendencia, cíclicos o estacionales, entonces se deben seleccionar las técnicas que sean capaces de extrapolar efectivamente tales patrones.

Es importante observar que, en muchas situaciones prácticas, más de un método o modelo para pronosticar puede dar predicciones aceptables y casi sin diferencias. De hecho, se recomienda intentar varias técnicas razonables. A menudo se deben considerar varios criterios tales como, simplicidad en la implementación, costo, condiciones ambientales externas, etcétera, para seleccionar un modelo adecuado.