



Estadística Aplicada 2

Tema 4. Series de Tiempo y Pronósticos



Tema 4. Series de Tiempo

Introducción

Obtención cuantitativa de pronósticos

1. Métodos intuitivos
2. Medias móviles
3. Alisado exponencial
4. Métodos de descomposición

Modelos de
Serie de Tiempo

5. Modelos asociativos

Modelos causales
El tiempo no se
considera un
factor relevante

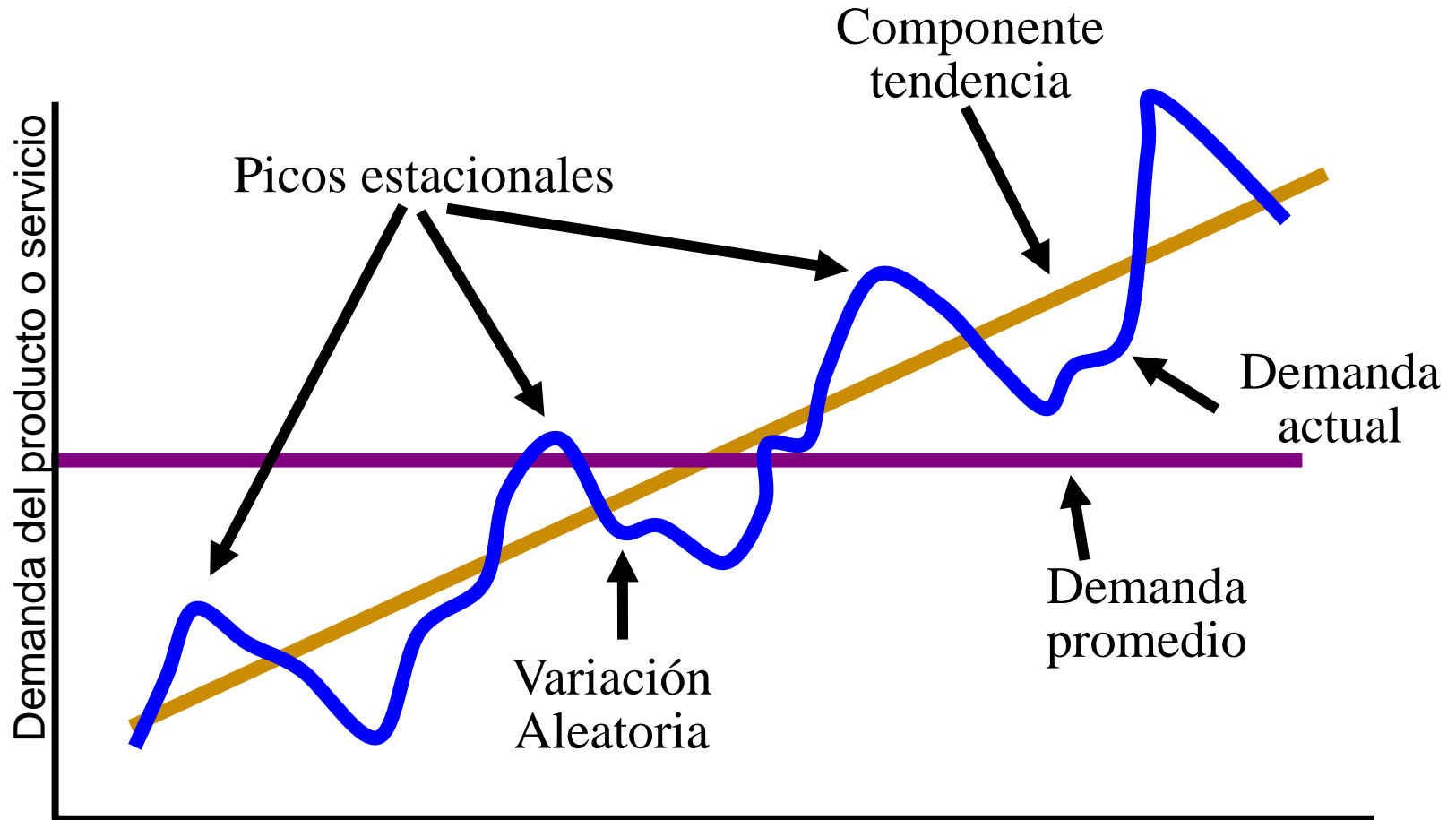
Concepto de serie de tiempo

- Secuencia de datos igualmente espaciados (semanales, mensuales, etc.)
 - Obtenida a través de la observación de la variable de respuesta en periodos de tiempo regulares
- Se pretende pronosticar valores futuros usando únicamente valores pasados
 - Se asume que otros factores (aparte del tiempo) no han cambiado ni cambian a lo largo del horizonte.

Componentes de una serie temporal

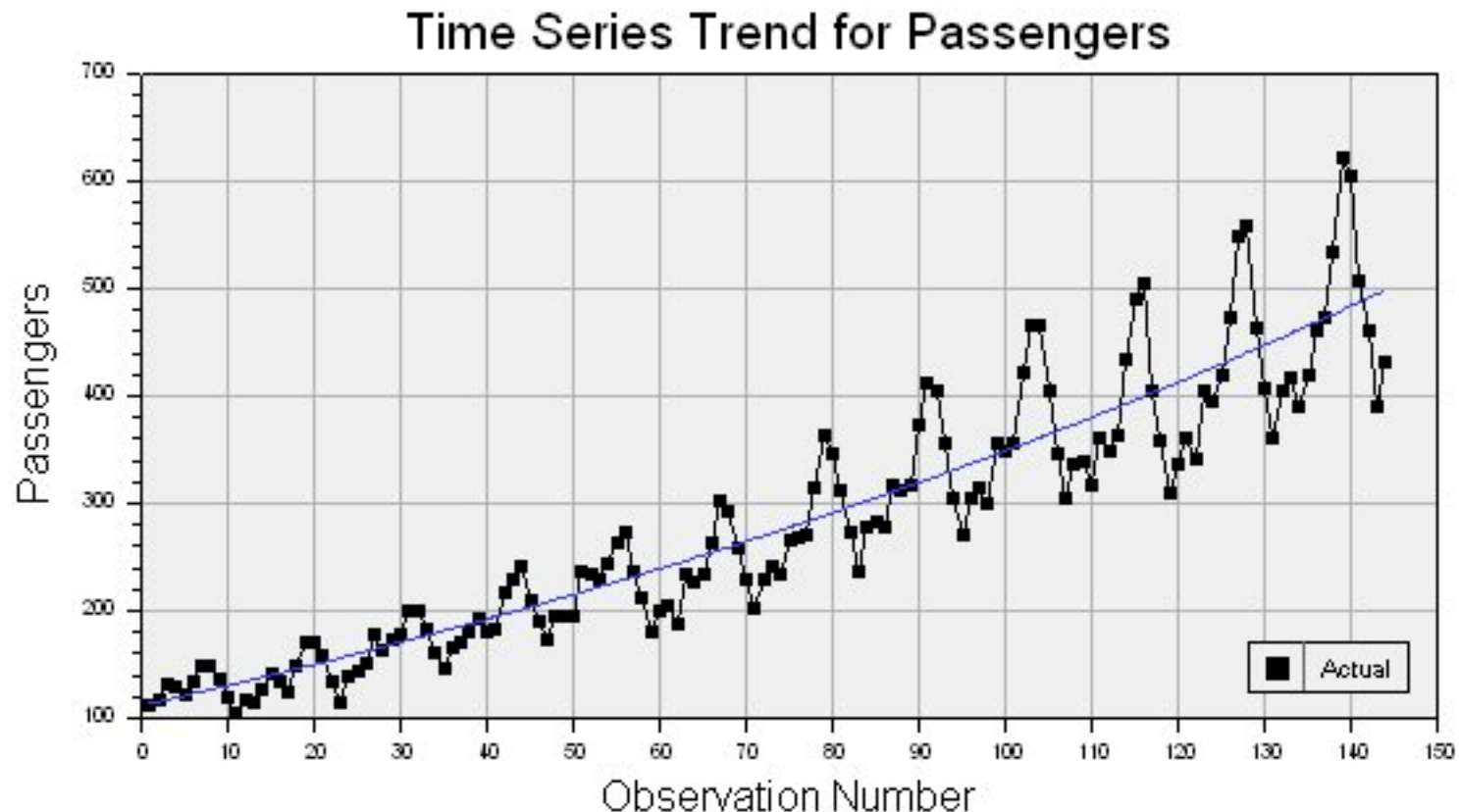
- Una serie de tiempo puede tener hasta cuatro componentes principales:
 - Tendencia
 - Estacionalidad
 - Ciclo
 - Aleatoriedad

Componentes de una serie de tiempo



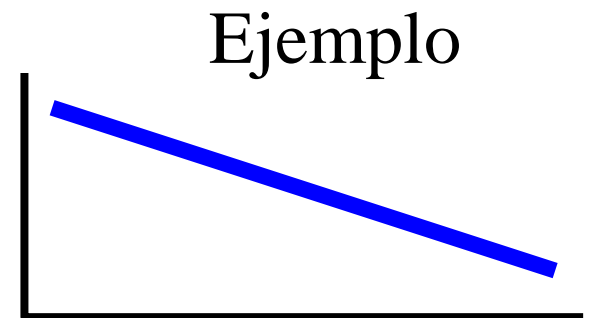
Tendencia

- Corresponde al alza que puede observarse en la serie



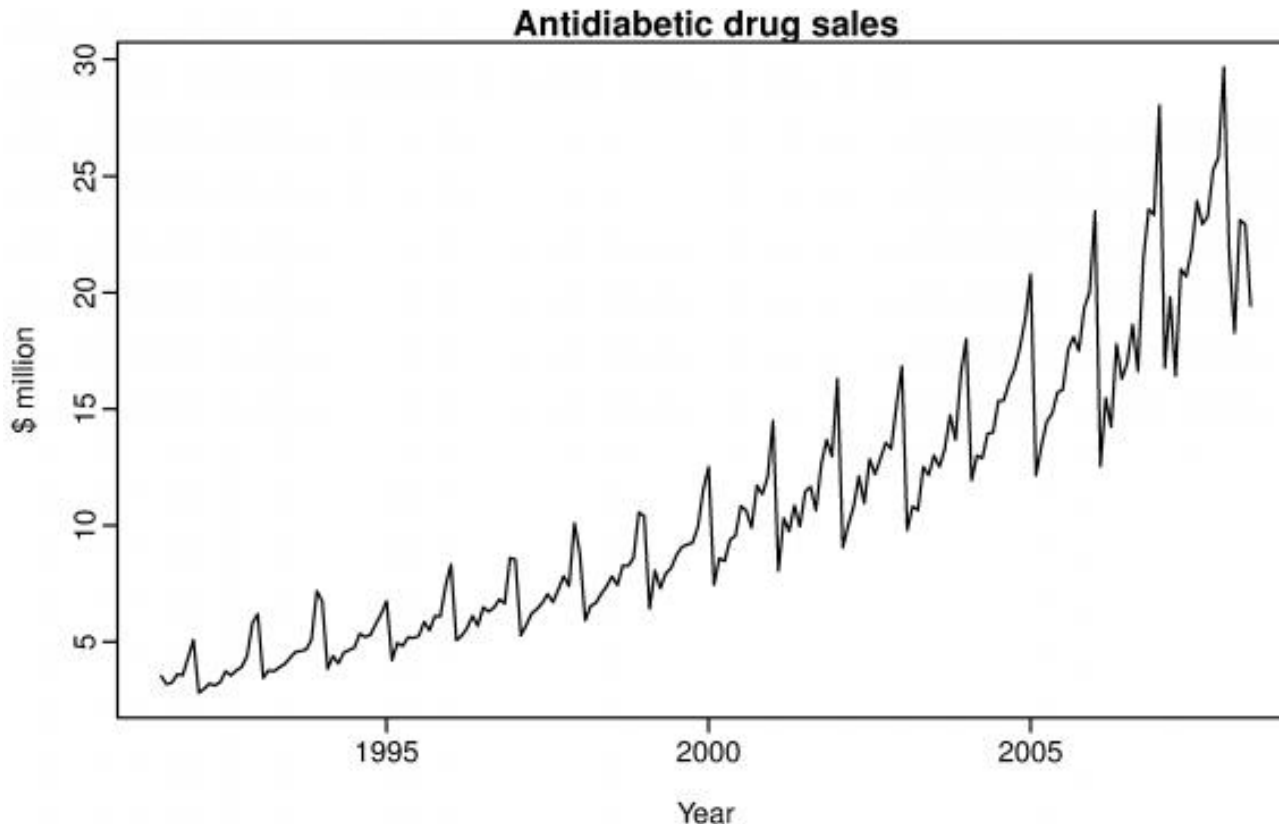
Tendencia

- Patrón que representa el incremento (o decremento) de la serie a lo largo del tiempo.
- La tendencia podría ser explicable por el ciclo económico, el ciclo de vida del producto, ...
- Representa la evolución por cambios de ingreso, tamaño de la población, cambios tecnológicos, ...
- Dura varios años



Estacionalidad

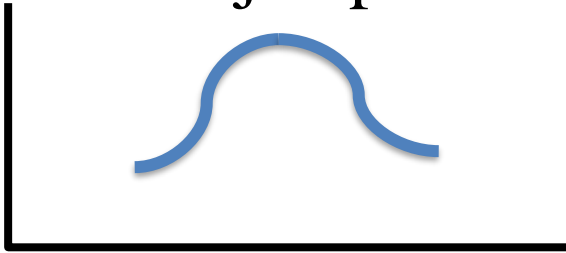
- Corresponde a las repeticiones visibles asociadas a épocas concretas del año



Estacionalidad

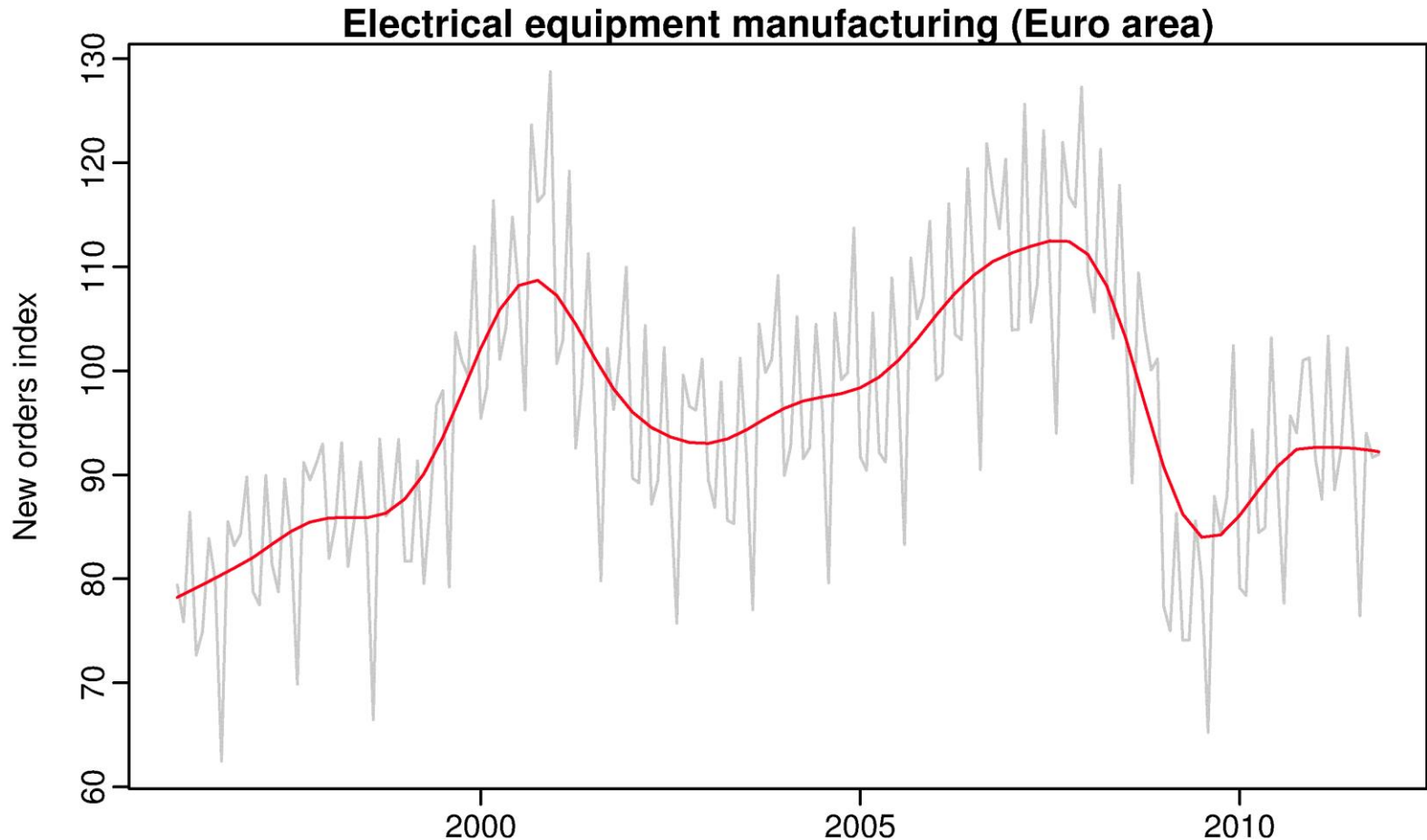
- Patrón que se repite cada periodo de tiempo (horas dentro de un día, días dentro de una semana, meses en un año...)
- Representa cambios por clima, costumbres sociales y fiestas

Ejemplo

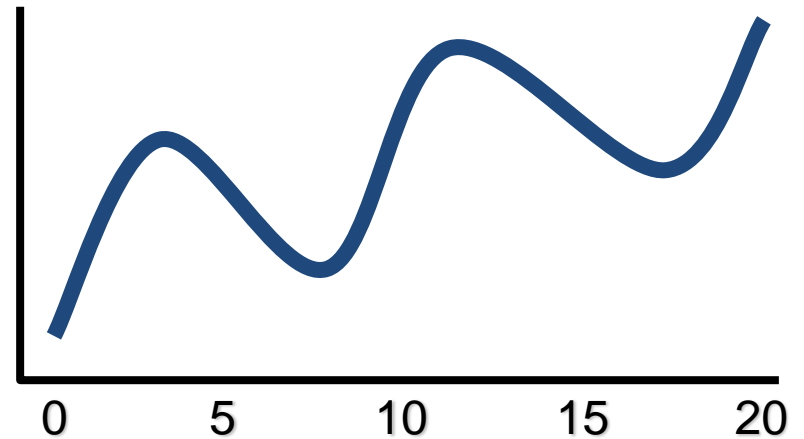


Periodo	Longitud	# periodos
Semana	Día	7
Mes	Día	28-31
Mes	Semana	4-4,5
Año	Semana	52
Año	Trimestre	4
Año	Mes	12

- Efecto del largo plazo en la serie

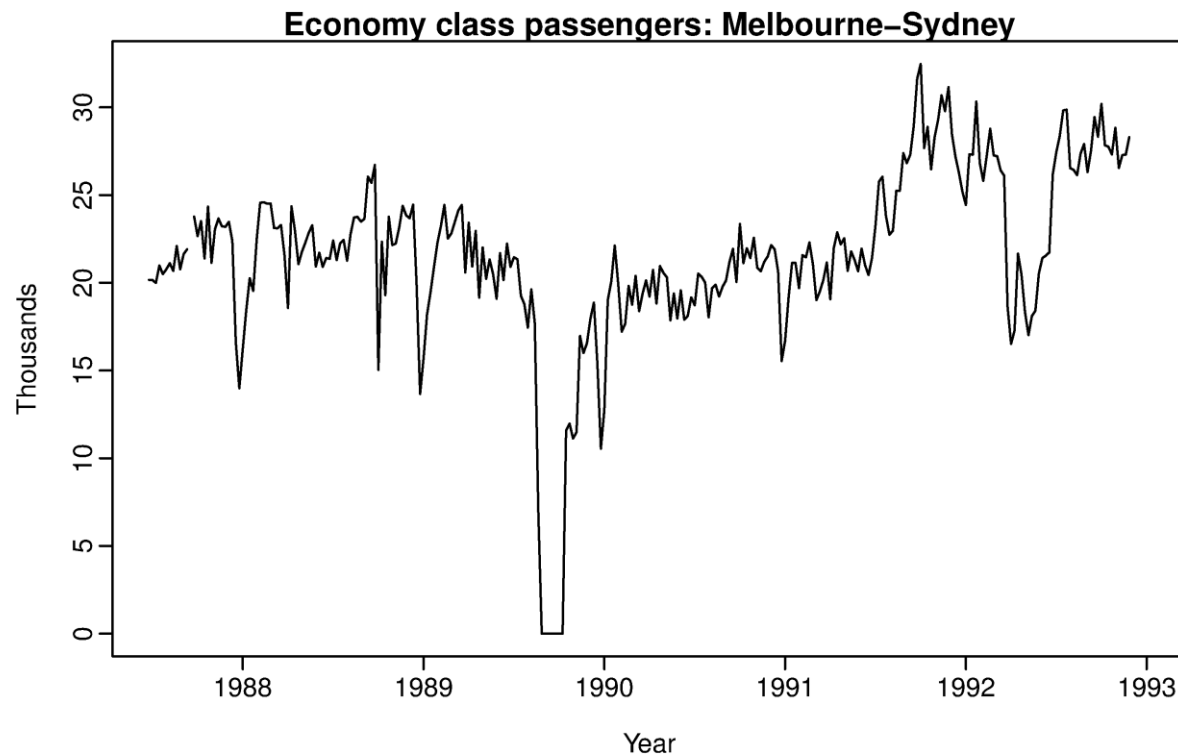


- Patrón que se repite cada cierta cantidad de años
- Representa los altibajos de las actividades económicas
- Difícil de asociar mediante series de tiempo



Aleatoriedad

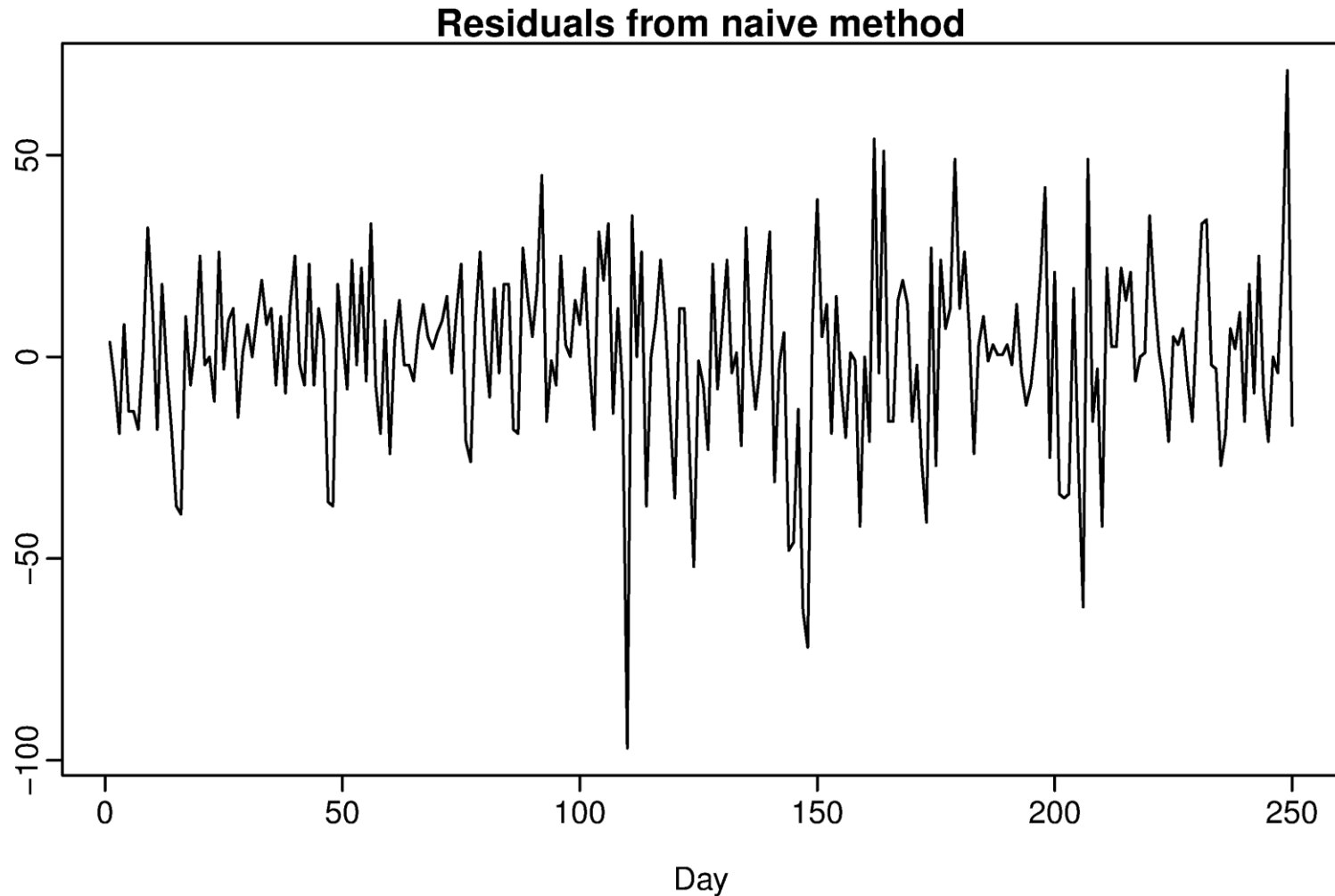
- Corresponde a los cambios no debidos a factores inexplicables (no confundir con el mínimo de este gráfico que se debe a una huelga).



Aleatoriedad

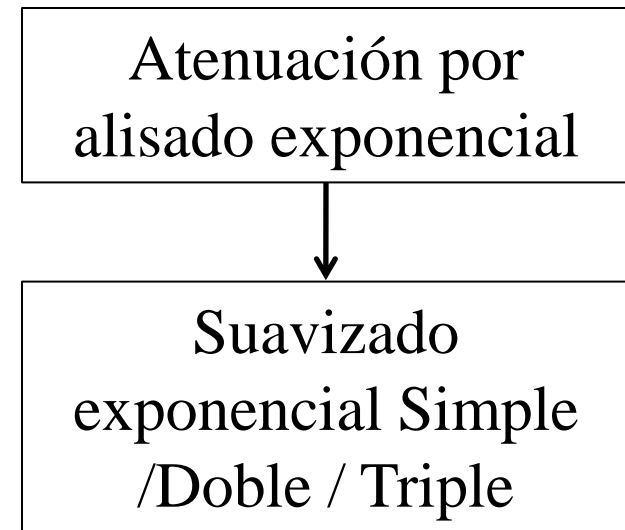
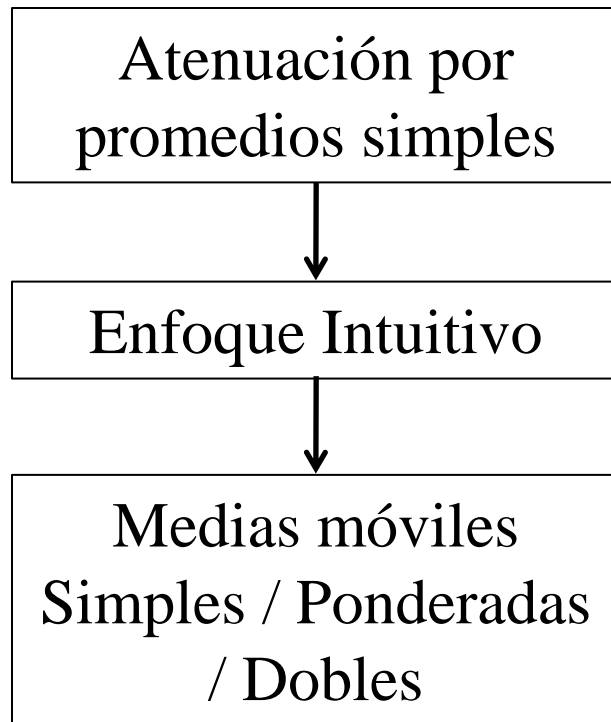
- Variaciones inexplicadas de los datos
- Se pueden deber a casualidades o a situaciones inusuales no incluidas en el modelo
- No deben presentar patrones (si presentan patrones, éstos pueden utilizarse para realizar predicciones)
- De lo anterior puede deducirse que sus efectos no son repetitivos
- Idealmente, debería obtenerse ruido blanco

- El ruido blanco no contiene información



Métodos de predicción

- Una amplia gama de métodos predictivos se basa en la atenuación de variaciones aleatorias para realización de predicciones.



Métodos de predicción

- Además de los factores de atenuación la literatura recoge otros elementos a predecir:
- Estudiar la autocorrelación de las observaciones. Para ello, se debe estudiar un modelo con un componente autoregresivos (errores correlacionados)
- Estudiar los componentes de la serie en partes individuales. En este caso, aplicar un método especial a cada elemento de la serie.



Tema 4. Series de Tiempo

Medias móviles

Enfoque intuitivo

- Asume que no existen cambios en la demanda
- En tal caso la mejor predicción para un periodo es el valor de demanda del periodo anterior.
 - Ejemplo: Las ventas de mayo fueron iguales a 840 unidades, por tanto podemos esperar unas ventas en junio de 840 unidades.
- En algunos casos, es un enfoque rápido y eficiente.

Enfoque intuitivo

- Matemáticamente:

$$F_{t+1} = X_t$$

Dónde:

- F_{t+1} : Pronóstico (*forecast*) de demanda para el periodo $t+1$
- t : Subíndice que identifica a los periodos.
- X_t : Demanda del periodo t .

Medias móviles

- El método de medias móviles se basa en la aplicación de medias aritméticas entre las observaciones.
- En una versión simple sólo se recomienda utilizar cuando la tendencia es casi inexistente.
- Tiene también usos en la descomposición de series para detectar tendencias, eliminar estacionalidad y reducir el efecto de la aleatoriedad.
- En inglés se conoce por sus siglas MA (Moving Average)

Medias móviles

- En su versión más simple, se puede describir matemáticamente como:

$$F_{t+1} = MA_t = \frac{1}{n} \sum_{i=0}^{n-1} X_{t-i} \quad , t > n$$

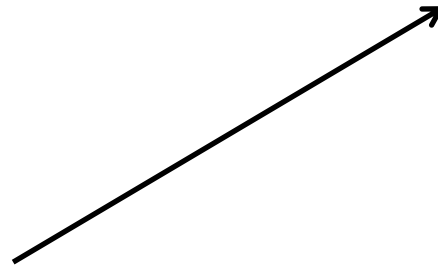
Dónde:

- F_{t+1} : Pronóstico (*forecast*) de demanda para el periodo $t+1$
- t : Subíndice que identifica a los periodos.
- X_t : Demanda del periodo t .

Relación con enfoque intuitivo

- Evidentemente el enfoque intuitivo comentado anteriormente es un caso especial de la media móvil con una única observación promediada.
- Usando la terminología estándar: $MA(1)$

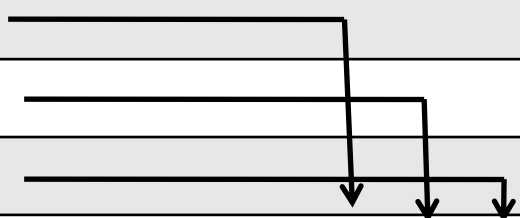
Número de periodos usados en
la determinación de la media



Ejemplo 1.

- Se desea realizar un pronóstico de ventas para las ventas de enero usando una técnica de medias móviles que tenga en cuenta las ventas de los últimos 3 meses.
- A continuación se muestran los datos y los cálculos.

Solución

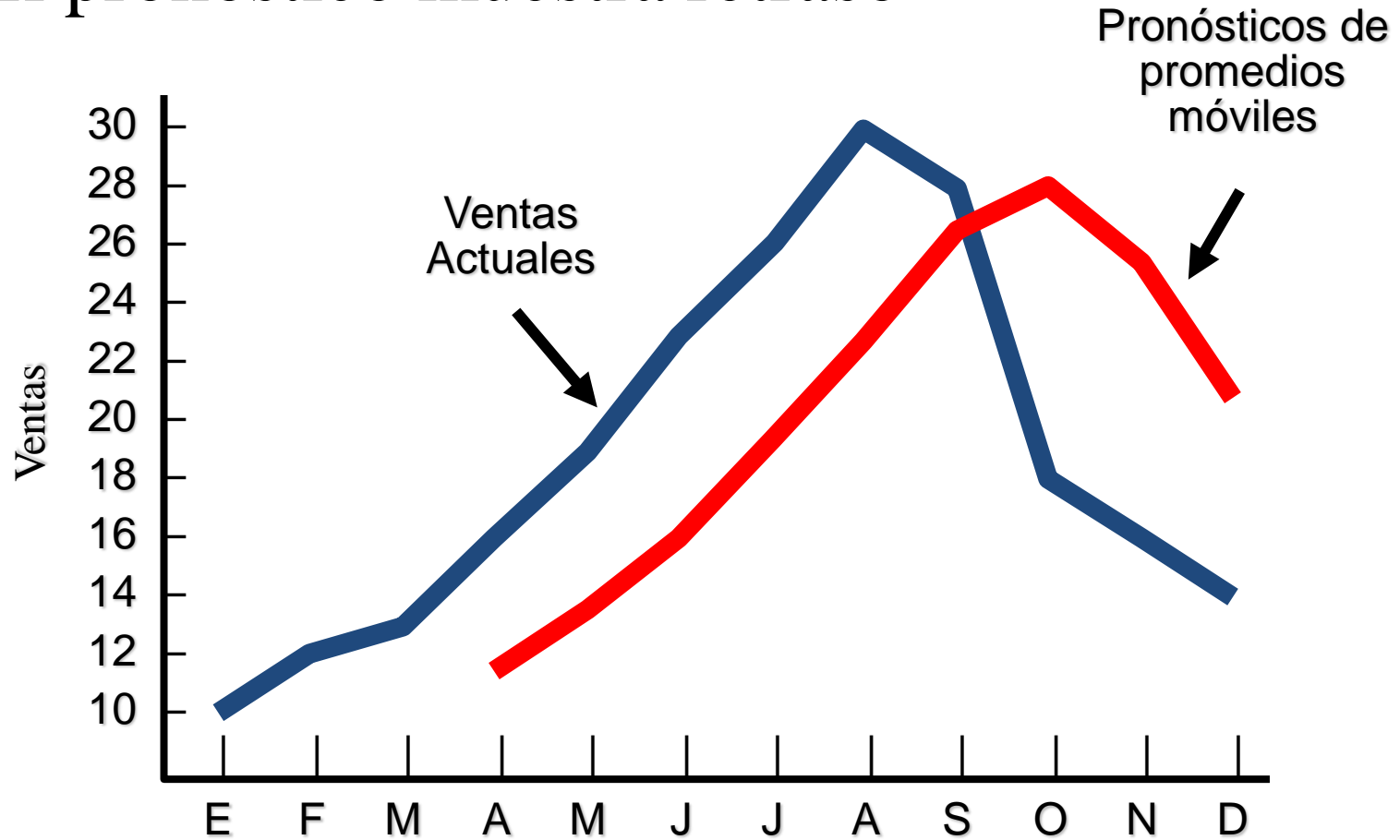
Mes	Ventas	Media Móvil
Enero	10	
Febrero	12	
Marzo	13	
Abril	16	$(10+12+13)/3 = 11 \frac{2}{3}$
Mayo	19	$(12+13+16)/3 = 13 \frac{2}{3}$
Junio	23	$(13+16+19)/3 = 16$
Julio	26	$(16+19+23) / 3 = 19 \frac{1}{3}$
Agosto	30	$(19+23+26) / 3 = 22 \frac{2}{3}$
Septiembre	28	$(23+26+30) / 3 = 26 \frac{1}{3}$
Octubre	18	$(26+30+28) / 3 = 28$
Noviembre	16	$(30+28+18) / 3 = 25 \frac{1}{3}$
Diciembre	14	$(28+18+16) / 3 = 20 \frac{2}{3}$
Enero	¿?	$(18+16+14) / 3 = 16$

Solución

Mes	Ventas	Media Móvil
Enero	10	
Febrero	12	
Marzo	13	
Abril	16	$(10+12+13)/3 = 11 \frac{2}{3}$
Mayo	19	$(12+13+16)/3 = 13 \frac{2}{3}$
Junio	23	$(13+16+19)/3 = 16$
Julio	26	$(16+19+23) / 3 = 19 \frac{1}{3}$
Agosto	30	$(19+23+26) / 3 = 22 \frac{2}{3}$
Septiembre	28	$(23+26+30) / 3 = 26 \frac{1}{3}$
Octubre	18	$(26+30+28) / 3 = 28$
Noviembre	16	$(30+28+18) / 3 = 25 \frac{1}{3}$
Diciembre	14	$(28+18+16) / 3 = 20 \frac{2}{3}$
Enero	¿?	$(18+16+14) / 3 = 16$

Ejemplo 1

- El pronóstico muestra retraso



Ejemplo 2

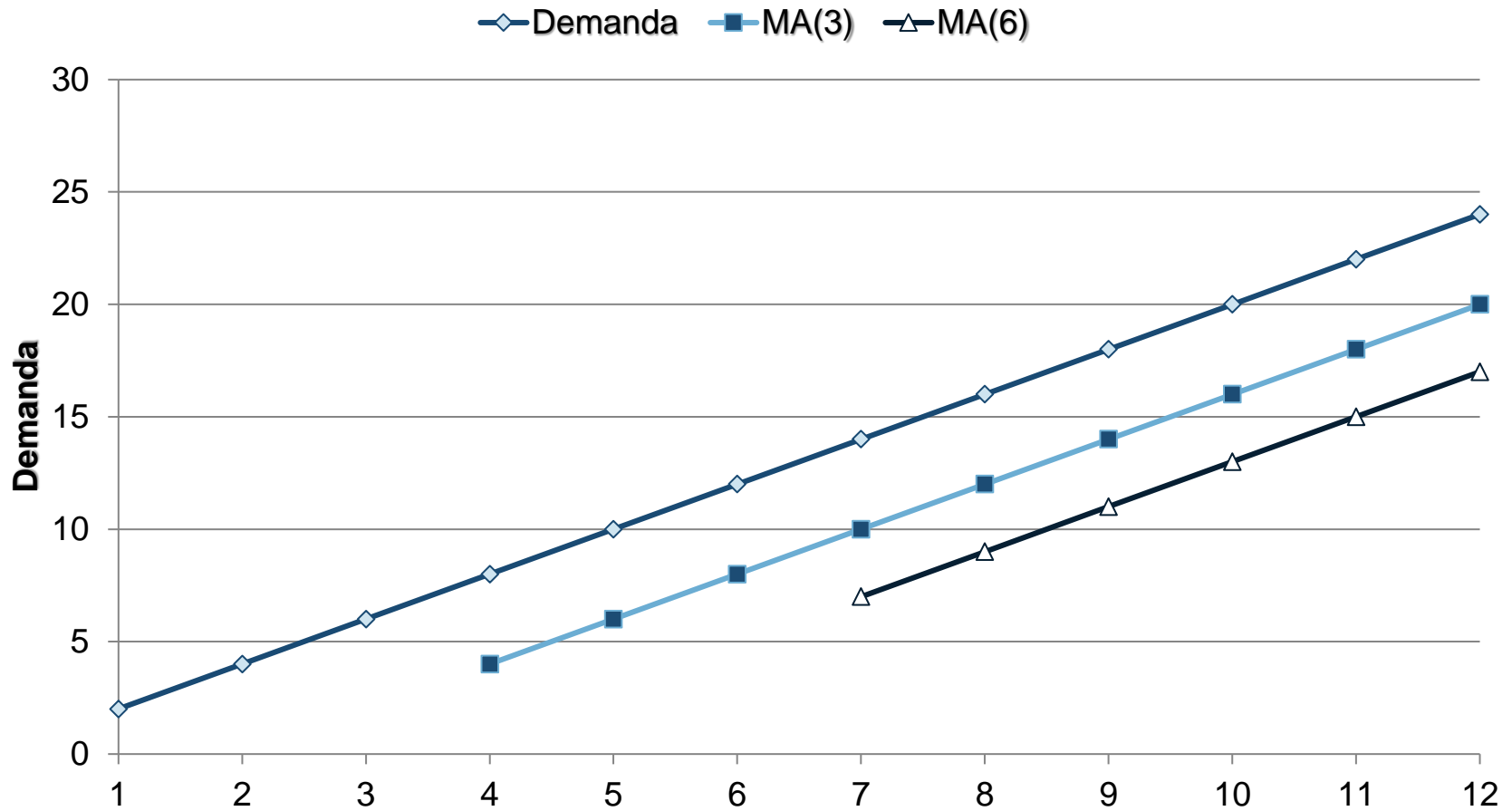
- El proceso analizado muestra una tendencia definida. Por ejemplo, las ventas de los últimos meses han sido: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24.
- Analicemos las previsiones de:
 - MA(3): Medias móviles calculadas usando los últimos 3 periodos.
 - MA(6): Medias móviles calculadas usando los últimos 6 periodos.

Solución

- Resultados:

Periodo	Demanda	MA(3)	MA(6)
1	2		
2	4		
3	6		
4	8	4	
5	10	6	
6	12	8	
7	14	10	7
8	16	12	9
9	18	14	11
10	20	16	13
11	22	18	15
12	24	20	17

- Gráficamente

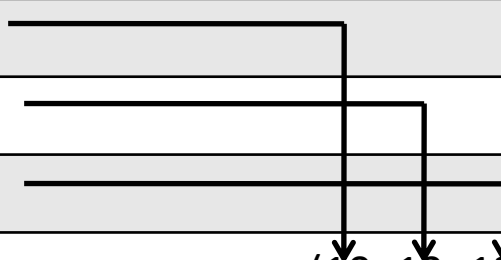


Solución

- En el gráfico puede observarse que las medias móviles están retrasadas en el tiempo. Cuanto mayor es el número de observaciones utilizadas, mayor es el retraso observado.
- Esto indica que el método no es el adecuado para hacer predicciones de series con tendencia.
- Como contra partida, se obtienen previsiones robustas ante efectos aleatorios

Análisis crítico del modelo simple

- Parte de los problemas del método se deben al siguiente motivo:

Mes	Ventas	Media Móvil
Enero	10	
Febrero	12	
Marzo	13	
Abril	16	$(10+12+13)/3 = 11 \frac{2}{3}$

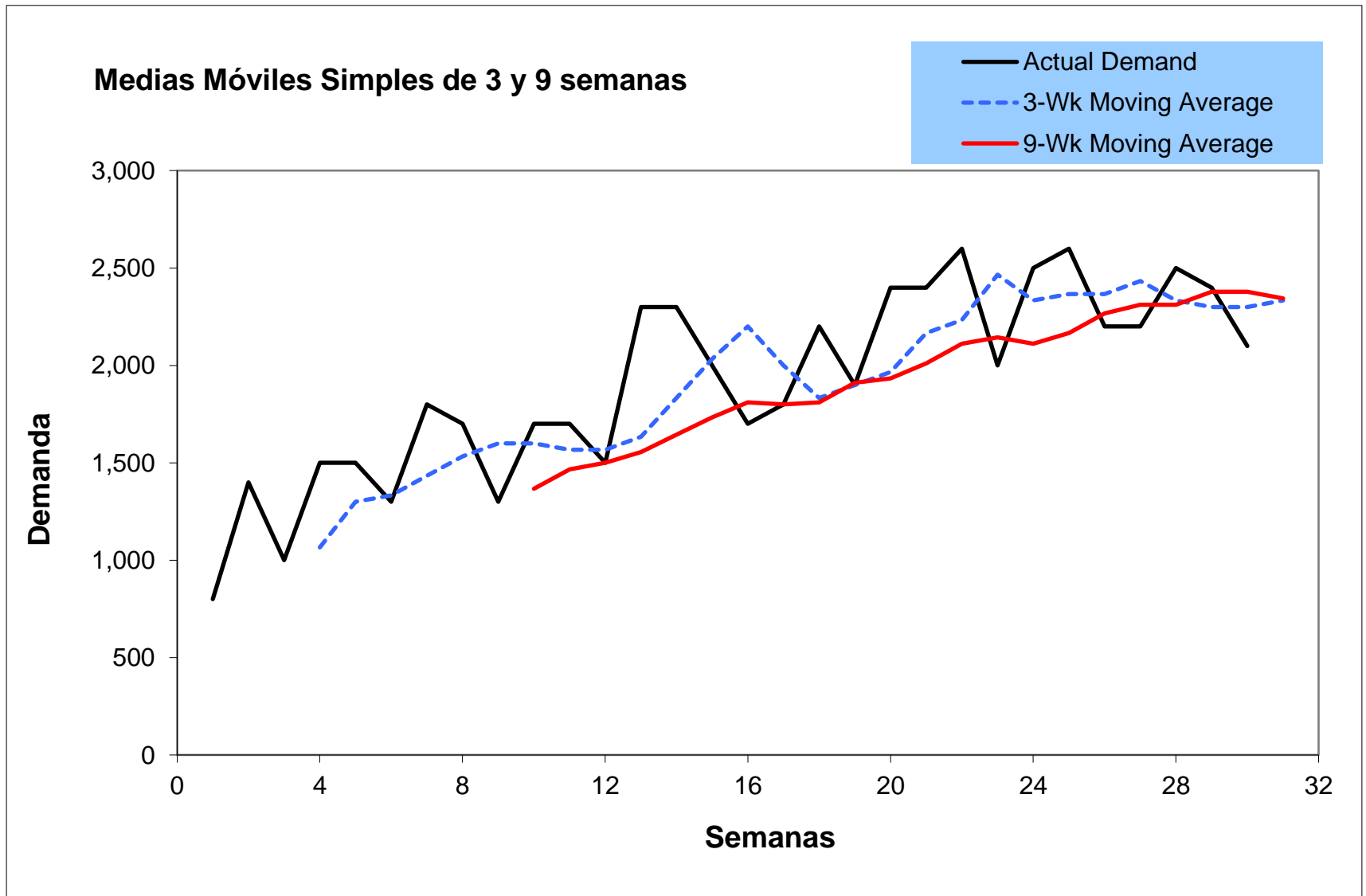
- Estamos proyectando ventas de Abril contando observaciones de Enero/Febrero/Marzo. Si existe una tendencia el promedio debería equivaler al mes de febrero no al de Abril.

Ejemplo 3

- Considere las siguientes 30 demandas semanales. Utilice promedios móviles simples basados en 3 y 9 semanas para predecir la demanda en la semana 31. Realice un gráfico comparativo entre las demandas y los pronósticos e indique lo que observa.

Solución

Semana	Demanda	Promedios Móviles Simples	
		3 Semanas	9 Semanas
1	800		
2	1.400		
3	1.000		
4	1.500	1.067	
5	1.500	1.300	
6	1.300	1.333	
7	1.800	1.433	
8	1.700	1.533	
9	1.300	1.600	
10	1.700	1.600	1.367
11	1.700	1.567	1.467
12	1.500	1.567	1.500
13	2.300	1.633	1.556
14	2.300	1.833	1.644
15	2.000	2.033	1.733
16	1.700	2.200	1.811
17	1.800	2.000	1.800
18	2.200	1.833	1.811
19	1.900	1.900	1.911
20	2.400	1.967	1.933
21	2.400	2.167	2.011
22	2.600	2.233	2.111
23	2.000	2.467	2.144
24	2.500	2.333	2.111
25	2.600	2.367	2.167
26	2.200	2.367	2.267
27	2.200	2.433	2.311
28	2.500	2.333	2.311
29	2.400	2.300	2.378
30	2.100	2.300	2.378
31	----	2.333	2.344



Análisis de la solución

- Se observa que la demanda tiene una tendencia creciente hasta estabilizarse al final de la serie.
- MA(3) produce una serie que sigue mejor los cambios de demanda (siempre con un retraso).
- MA(9) produce una serie más estable (permite incluso identificar la tendencia de la serie).

Se prefiere un método con periodo largo si se considera que la serie es estable y uno con periodo corto si se opina lo contrario.

Medias móviles ponderadas

- Una alternativa para hacer frente a la problemática de la tendencia es el uso de series ponderadas.
- El método usa pesos diferentes para las observaciones, dando mayor importancia a las observaciones recientes.
- Estas ponderaciones también pueden reflejar consideraciones de estacionalidad.
- Existen diversos métodos para determinar la ponderación a usar (experiencia/intuición/optimización).

Medias móviles ponderadas

- Matemáticamente:

$$F_{t+1} = WMA_t = \sum_{i=0}^{n-1} w_i X_{t-i} \quad ; t > n; \sum_{i=0}^{n-1} w_i = 1$$

Dónde:

- n : Número de periodos usados
- w_i : Ponderación del i -ésimo periodo.

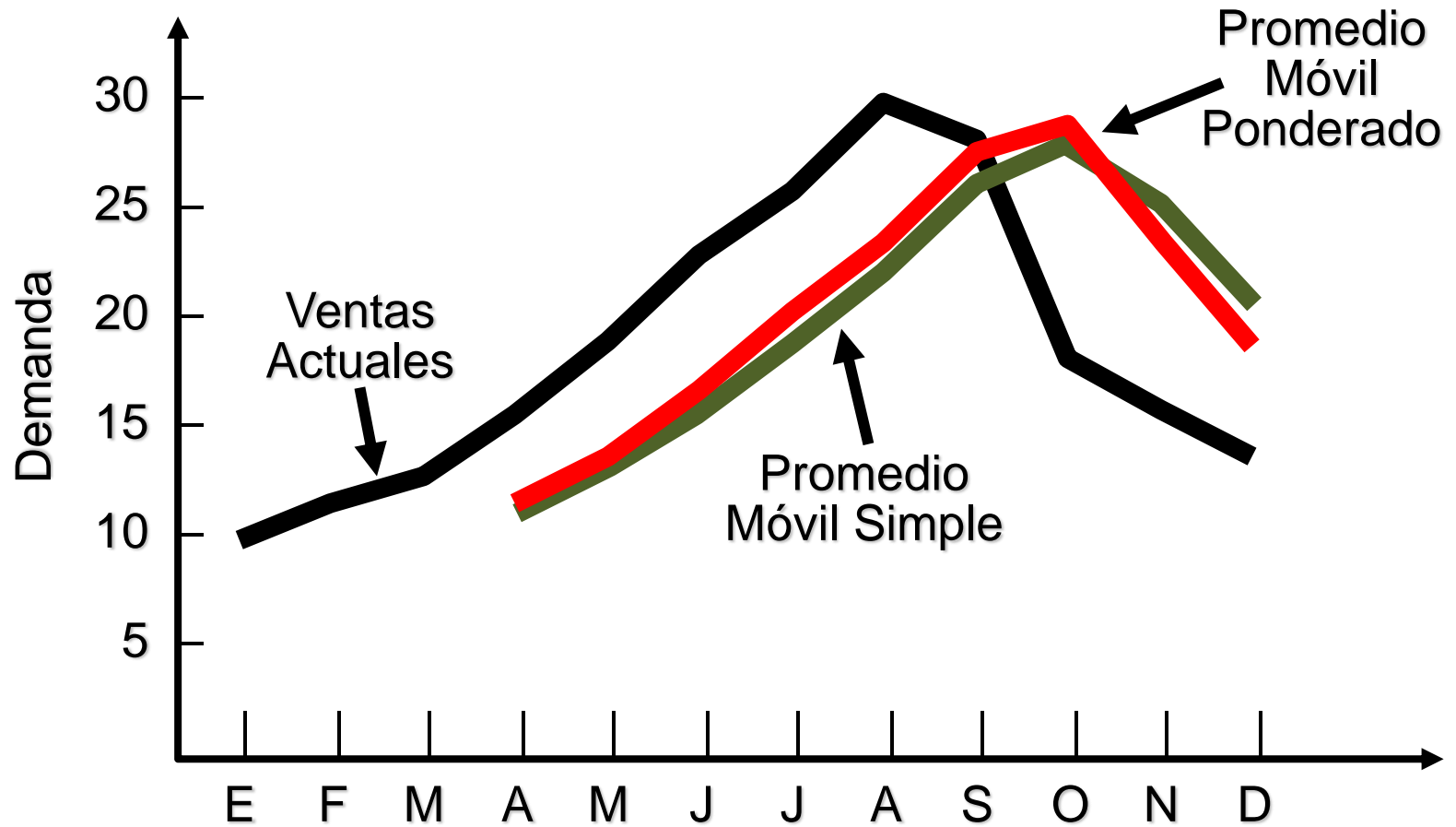
Ejemplo 4

- Consideremos de nuevo los datos del ejemplo 1.
- Aplicaremos la siguiente ponderación:
 - $3/6$ al periodo más reciente
 - $2/6$ al segundo periodo
 - $1/6$ al tercer periodo

Solución

Mes	Ventas	Media Móvil
Enero	10	
Febrero	12	
Marzo	13	
Abril	16	$(1*10+2*12+3*13)/6 = 12 \frac{1}{6}$
Mayo	19	$(1*12+2*13+3*16)/6 = 14 \frac{1}{3}$
Junio	23	$(1*13+2*16+3*19)/3 = 17$
Julio	26	$(1*16+2*19+3*23) / 3 = 20 \frac{1}{2}$

Solución



Medias móviles en Minitab

- Minitab provee de dos opciones para el cálculo de medias móviles:
 - Básica: En este caso se calcula la media móvil tal como se muestra en las diapositivas
 - Centradas: En este caso se rectifica el “lag” observado en el proceso (rectifica el origen de las medias)
 - Si la constante es impar: Se asigna el valor previsto al periodo medio de la serie (p.ej. Si $MA(3)$ la primera media móvil correspondería la observación 2.
 - Si es par: Se promedia los dos valores de media móvil y se asignan al periodo superior. (p.ej. Si $MA(4)$ se determinan primer y segundo valor y se asignan al observación 3.

Conclusiones

- Entre los múltiples problemas de las medias móviles:
 - Incrementar el número de periodos considerados suaviza el pronóstico pero lo hace menos sensibles a cambios.
 - No refleja tendencias (pero curiosamente las detecta).
- A pesar de los defectos indicados es un método ampliamente utilizado que conforma la base del método de alisado exponencial.



Tema 4. Series de Tiempo

Alisado exponencial
simple

Suavizado exponencial simple

- Se trata de un método de media móvil ponderada.
 - Las ponderaciones declinan exponencialmente
 - Por tanto, los datos más recientes son más importantes
 - El método reacciona mejor ante cambios
- Utiliza una constante de suavización (α)
 - Su rango está comprendido entre 0 y 1
 - Puede optimizarse u obtenerse subjetivamente
 - Si $\alpha=2/(n+1)$ se obtiene el método de medias móviles simple.

Suavizado exponencial simple

- Matemáticamente:

$$F_{t+1} = a \times y_t + a \times (1 - a) y_{t-1} + a \times (1 - a)^2 y_{t-2} + a \times (1 - a)^3 y_{t-3} + \dots$$

Observación	$\alpha=0,2$	$\alpha=0,4$	$\alpha=0,6$	$\alpha=0,8$
y_t	0,2	0,4	0,6	0,8
y_{t-1}	0,16	0,24	0,24	0,16
y_{t-2}	0,128	0,114	0,096	0,032
y_{t-3}	0,1024	0,0864	0,0384	0,0064
y_{t-4}	0,08192	0,05184	0,01536	0,00128
y_{t-5}	0,06554	0,03110	0,00614	0,000256

Suavizado exponencial simple

- La fórmula anterior no es práctica, por lo que se acostumbra organizar los operandos como sigue:

$$F_{t+1} = a \times y_t + (1 - a) F_t$$

- Para verificar la validez de esta fórmula apliquemos sustitución, empezando por el periodo 2 ($t+1=2$) y consideremos que la previsión para el periodo 1 es igual a l_1 .

$$F_2 = a \times y_1 + (1 - a) l_1$$

$$F_3 = a \times y_2 + (1 - a) F_2, \dots$$

Suavizado exponencial simple

- Sustituyendo las ecuaciones anteriores en la ecuación original:

$$F_3 = a \times y_2 + (1 - a) [a \times y_1 + (1 - a) l_1] =$$
$$a \times y_2 + a(1 - a) y_1 + (1 - a)^2 l_1$$

$$F_4 = a \times y_3 + (1 - a) F_3 =$$
$$a \times y_3 + (1 - a) [a \times y_2 + a(1 - a) y_1 + (1 - a)^2 l_1] =$$
$$a \times y_3 + a(1 - a) y_2 + a(1 - a)^2 y_1 + (1 - a)^3 l_1$$

Suavizado exponencial simple

- Por tanto:

$$F_{t+1} = a \times y_t + a \times (1 - a) y_{t-1} + \\ a \times (1 - a)^2 y_{t-2} + a \times (1 - a)^3 y_{t-3} + \dots + (1 - a)^{t-1} l_1$$

- Que es la fórmula presentada originalmente excepto: $(1 - a)^{t-1} l_1$ que tiende a 0 porque $0 < a < 1$

$$F_{t+1} = a \times y_t + (1 - a) F_t$$

Efecto del parámetro α

- Un valor de $\alpha=0$ significa que:

$$F_{t+1} = 0 \times y_t + (1 - 0) F_t$$

– La predicción se basa sólo en históricos antiguos

- Un valor de $\alpha=1$ significa que:

$$F_{t+1} = 1 \times y_t + (1 - 1) F_t$$

– La predicción se basa sólo en la información más reciente

- Un valor de $\alpha=0,2$ significa que:

$$F_{t+1} = 0,2 \times y_t + (1 - 0,2) F_t$$

Ejemplo 5

- En Enero, un vendedor de automóviles predijo que la demanda para Febrero sería de 142 coches. La demanda real en Febrero fue de 153 automóviles. Usando la constante de suavizado que eligió gerencia ($\alpha = 0,2$) el vendedor quiere pronosticar la demanda para Marzo usando el modelo de suavizado exponencial simple.

Solución

- Previsión de Febrero (F_t) = 142
- Demanda (y_t) = 153
- t: febrero
- $\alpha = 0,2$

$$\begin{aligned} F_{t+1} &= 0,2 \times y_t + (1 - 0,2) F_t = \\ &0,2 \times 153 + (1 - 0,2) \times 142 = \\ &30,6 + 113,6 = 144,2 \gg 144 \end{aligned}$$

Ejemplo 6

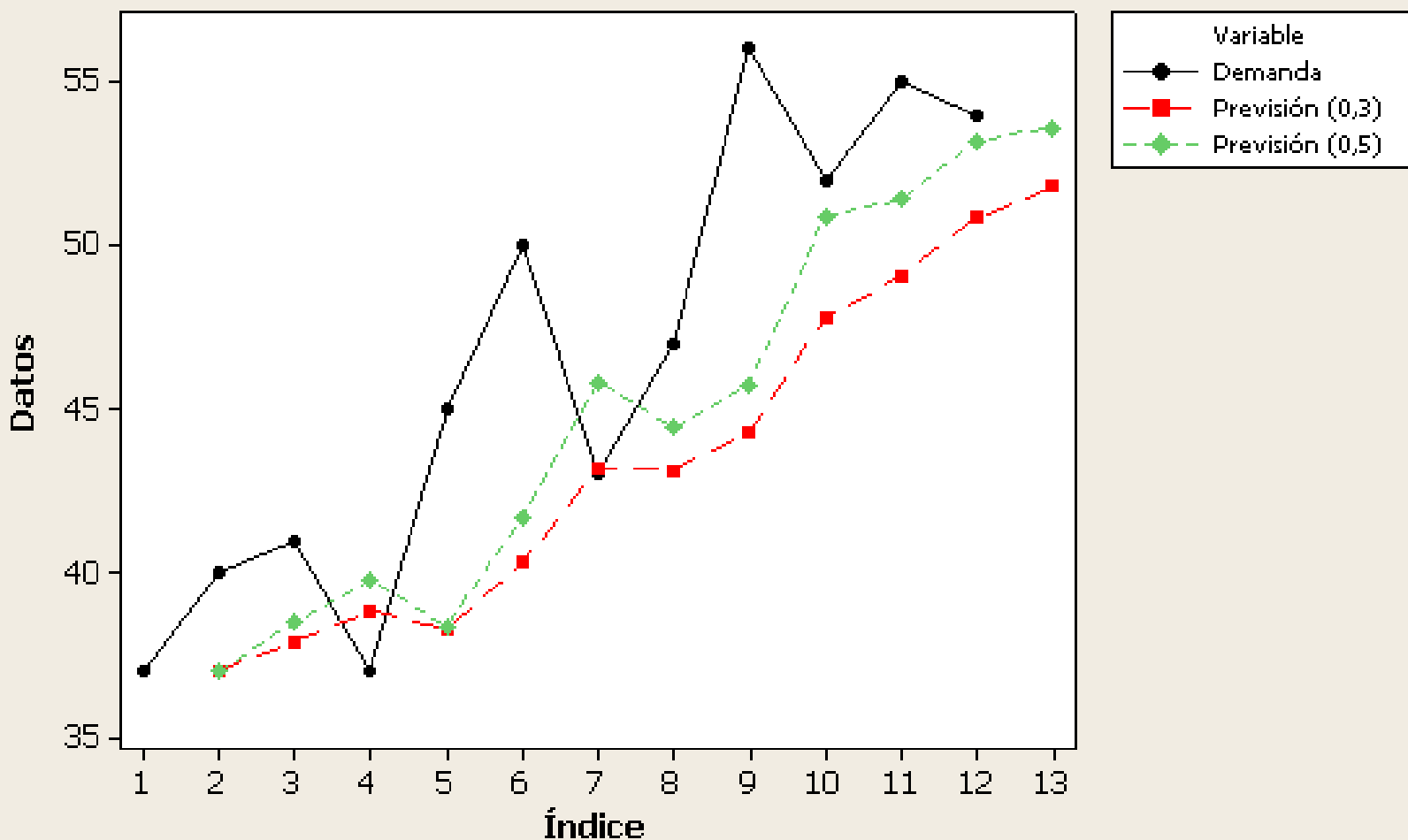
- Hitek Computer Services es una empresa que repara ordenadores.
- Hitek ha tenido un incremento paulatino de la demanda en el último tiempo. Es por esto que necesitan realizar una previsión de su demanda para estudiar su política de aprovisionamientos
- Se dispone de la demanda del último año, y se decide utilizar una constante de suavización de 0,3 y 0,5 para realizar los pronósticos.

Solución

Periodo	Mes	Demanda	Previsión ($\alpha=0,3$)	Previsión ($\alpha=0,5$)
1	ENE	37	-	-
2	FEB	40	37	37
3	MAR	41	37,9	38,5
4	ABR	37	38,83	39,75
5	MAY	45	38,28	38,37
6	JUN	50	40,29	41,68
7	JUL	43	43,20	45,84
8	AGO	47	43,14	44,42
9	SEP	56	44,30	45,71
10	OCT	52	47,81	50,85
11	NOV	55	49,06	51,42
12	DIC	54	50,84	53,21
13	ENE	-	51,79	53,61

Representación gráfica

Gráfica de series de tiempo de Demanda; Previsión (0,3); Previsión (0,5)



Determinación de parámetros

- El modelo de alisado exponencial necesita dos parámetros para su correcto funcionamiento:
 - Una previsión inicial (valor de previsión del primer periodo) equivalente a l_1 en la demostración de la fórmula de suavizado exponencial.
 - Un constante de suavizado.

Determinación de previsión inicial

- Es habitual utilizar la primera observación como previsión inicial.
- Otro método (utilizado por Minitab) consiste en promediar las k primeras observaciones
- Existen técnicas más complejas que estudian la serie en orden inverso para determinar el mejor valor inicial.
- En cualquier caso, éste parámetro es poco importante si la serie tiene muchas observaciones.

Determinación constante de suavizado

- En cuanto la constante de suavizado, es habitual seleccionar manualmente la constante.
- Los valores de α que generalmente son escogidos, varían entre 0,1 y 0,3.
- α toma valores altos cuando hay (o se esperan) cambios en la variable observada.
- Se eligen valores bajos de α cuando la variable observada es estable (el pronóstico reaccionará lentamente a los cambios).
- Es posible optimizar el valor de los parámetros (ver Minitab).

Determinación constante de suavizado

- Una técnica para obtener el mejor parámetro posible consiste en la resolución de un modelo más complejo (conocido como ARIMA) que obtiene el conjunto de parámetros que minimiza una medida de error.
- Otra técnica sería resolver un modelo no lineal con dos parámetros que minimice alguna de las medidas de error planteadas.

Suavizado simple en Minitab

- Minitab provee de dos opciones:
 - Determinación óptima:
 - Resuelve un modelo ARIMA(0,1,1) que se verá posteriormente.
 - El valor de α se iguala al coeficiente de media móvil del método ARIMA.
 - El valor de l_1 se determina usando el valor previsto por el modelo ARIMA para el segundo periodo siguiendo:

$$l_1 = [F_2 - a \times y_1] / (1 - a) \quad \text{Backcasting}$$

- Parámetro especificado:
 - En este caso se necesita conocer l_1 que se determina promediando las seis primeras observaciones (si hay menos de 6 observaciones, promedia todas las observaciones disponibles).

Errores

- Como en un modelo lineal, una aproximación posible consistiría en medir un error de previsión calculando la diferencia entre la previsión y el valor real:

$$e_t = F_{t+1} - y_t$$

- Se intentaría reducir el error total del modelo, que sería equivalente a la minimización de una de las siguientes tres funciones que presentan objetivos diferentes.

Medidas de error

- Desviación absoluta media:

$$MAD = \frac{\sum_t |e_t|}{n}$$

- Error medio cuadrado

$$MSE = \frac{\sum_t e_t^2}{n}$$

- Error porcentual absoluto medio:

$$MAPE = \frac{\sum_t 100 \times |e_t| / y_t}{n}$$

Ejemplo 7

- Durante los últimos 8 trimestres, en el puerto de Baltimore se han descargado de los barcos grandes cantidades de grano.
- El administrador de operaciones del puerto quiere probar el uso del alisado exponencial simple para predecir el tonelaje descargado.
- El pronóstico de grano descargado durante el primer trimestre fue de 175 toneladas.
- Compare constantes $\alpha = 0,1$ y $0,5$. A través de las medidas de error, decida con cuál se queda.

Trimestre	Observación	Previsión		Previsión	
		n	Error Absoluto	n	Error Absoluto
		0,1	0	0,1	0
1	180	175	5	175	5
2	168	175,5	7,5	177,5	9,5
3	159	174,75	15,75	172,75	13,75
4	175	173,18	1,82	165,88	9,12
5	190	173,36	16,64	170,44	19,56
6	205	175,02	29,98	180,22	24,78
7	180	178,02	1,98	192,61	12,61
8	182	178,22	3,78	186,3	4,3

Solución

$$MAD = \frac{\sum_t |e_t|}{n} \left. \begin{array}{l} \bullet \alpha=0,1 \rightarrow 10,31 \\ \bullet \alpha=0,5 \rightarrow 12,33 \end{array} \right\}$$

$$MSE = \frac{\sum_t e_t^2}{n} \left. \begin{array}{l} \bullet \alpha=0,1 \rightarrow 190,8 \\ \bullet \alpha=0,5 \rightarrow 201,5 \end{array} \right\}$$

$$MAPE = \frac{\sum_t 100 \times |e_t| / y_t}{n} \left. \begin{array}{l} \bullet \alpha=0,1 \rightarrow 5,59\% \\ \bullet \alpha=0,5 \rightarrow 6,85\% \end{array} \right\}$$

- Aunque es frecuente usar los métodos anteriores para comparar diferentes métodos de predicción, la comparación directa usando todas las observaciones puede llevar a sobre ajustar los parámetros a la serie.
- Este problema lleva a que en normalmente se divida la serie en dos partes, una primera que se utiliza para “entrenar” y determinar los parámetros obtenidos por diversos métodos y otra que se utiliza para comparar la eficiencia de los métodos.



Tema 4. Series de Tiempo

Alisado exponencial
doble

Introducción

- Los métodos de medias móviles y de suavizado siempre mostrarán un retraso entre las observaciones y las predicciones en caso de existir una tendencia.
- Si se percibe tendencia será necesario el uso de un modelo que la incorpore para ajustar las predicciones al fenómeno esperado
- Un método basado en el alisado exponencial que permite tener en cuenta tendencias es el método de Holt, que usa dos constantes de alisado.

Método de Holt

- El método determina mediante una ecuación la previsión del nivel de la serie y mediante otra ecuación determina una previsión de tendencia.
- Ambos niveles se comportan mediante funciones de suavizado exponencial.

$$\left. \begin{aligned} F_t &= aX_t + (1 - a)(F_{t-1} + T_{t-1}) \\ T_t &= b(F_t - F_{t-1}) + (1 - b)T_{t-1} \end{aligned} \right\} H_{t+1} = F_t + m \times T_t$$

Nomenclatura

$$\left. \begin{aligned} F_t &= \alpha X_t + (1 - \alpha)(F_{t-1} + T_{t-1}) \\ T_t &= \beta(F_t - F_{t-1}) + (1 - \beta)T_{t-1} \end{aligned} \right\} H_{t+1} = F_t + m \times T_t$$

- F_t : valor suavizado para el nivel en el periodo t
- α : constante de alisado para el nivel de la serie ($0 \leq \alpha \leq 1$)
- x_t : demanda real en el periodo t
- T_t : factor de tendencia de la serie en el periodo t
- β : constante de alisado para la tendencia de la serie ($0 \leq \beta \leq 1$)
- m : número de periodos hasta el que se desea predecir
- H_t : predicción del método para el periodo t .

Ejemplo 8

- Un fabricante desea determinar la demanda de un equipo industrial.
- Una revisión de las ventas históricas, indica que hay presente una tendencia.
- Los datos se presentan a continuación.
- Utilice el método de Holt para predecir la demanda del mes 10, con $\alpha = 0,2$ y $\beta = 0,4$.
- Considere que el pronóstico para el mes 1 fue de 11 unidades y que la tendencia para el mismo mes fue de 2 unidades.

		Nivel Tendencia Pronóstico		
Mes	Demanda	Ft	Tt	Ht
1	12	11	2	13
2	17			
3	20			
4	19			
5	24			
6	21			
7	31			
8	28			
9	36			
10				

Paso 1. Determinar F_2 :
$$F_2 = \alpha X_1 + (1 - \alpha)(F_1 + T_1) =$$
$$F_2 = 0,2 \times 12 + 0,8 \times (11 + 2) =$$
$$12,8 \text{ unidades.}$$

Mes	Demanda	Nivel Tendencia Pronóstico		
		Ft	Tt	Ht
1	12	11	2	13
2	17	12,8		
3	20			
4	19			
5	24			
6	21			
7	31			
8	28			
9	36			
10				

Paso 2. Tendencia T_2 :

$$T_2 = \beta(F_2 - F_1) + (1 - \beta) T_1 =$$
$$T_2 = 0,4 \times (12,8 - 11) + 0,6 \times 2 =$$

1,92 unidades.

Mes	Demanda	Nivel Tendencia Pronóstico		
		Ft	Tt	Ht
1	12	11	2	13
2	17	12,8	1,92	
3	20			
4	19			
5	24			
6	21			
7	31			
8	28			
9	36			
10				

Paso 3. Pronóstico H_2 :

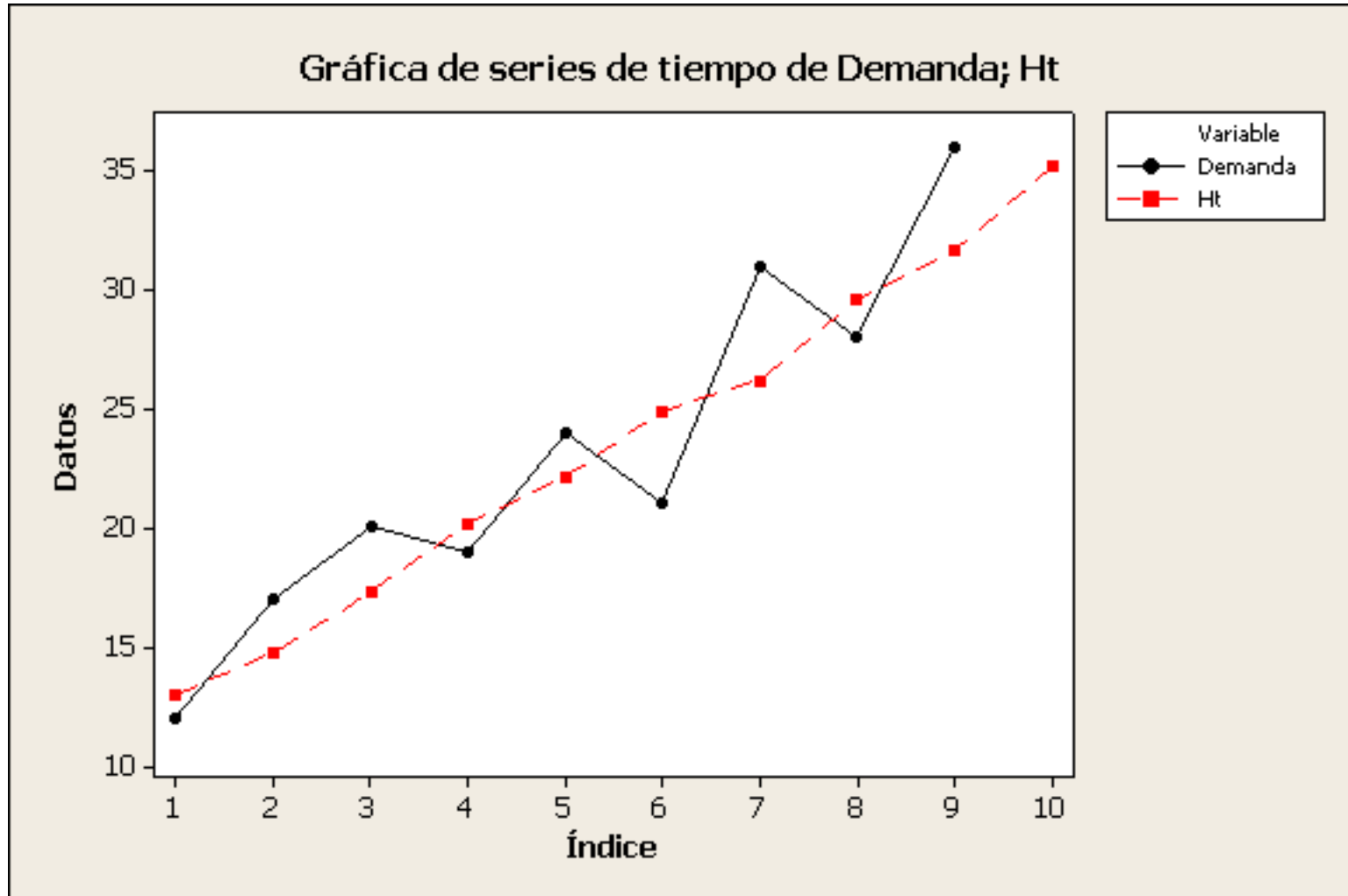
$$H_2 = F_2 + T_2 =$$

$$H_2 = 12,8 + 1,92 =$$

14,72 unidades.

Solución

		Nivel	Tendencia	Pronóstico
Mes	Demanda	Ft	Tt	Ht
1	12	11	2	13
2	17	12,8	1,92	14,72
3	20	15,18	2,1	17,28
4	19	17,82	2,32	20,14
5	24	19,91	2,23	22,14
6	21	22,51	2,38	24,89
7	31	24,11	2,07	26,18
8	28	27,14	2,45	29,59
9	36	29,28	2,32	31,6
10		32,48	2,68	35,16



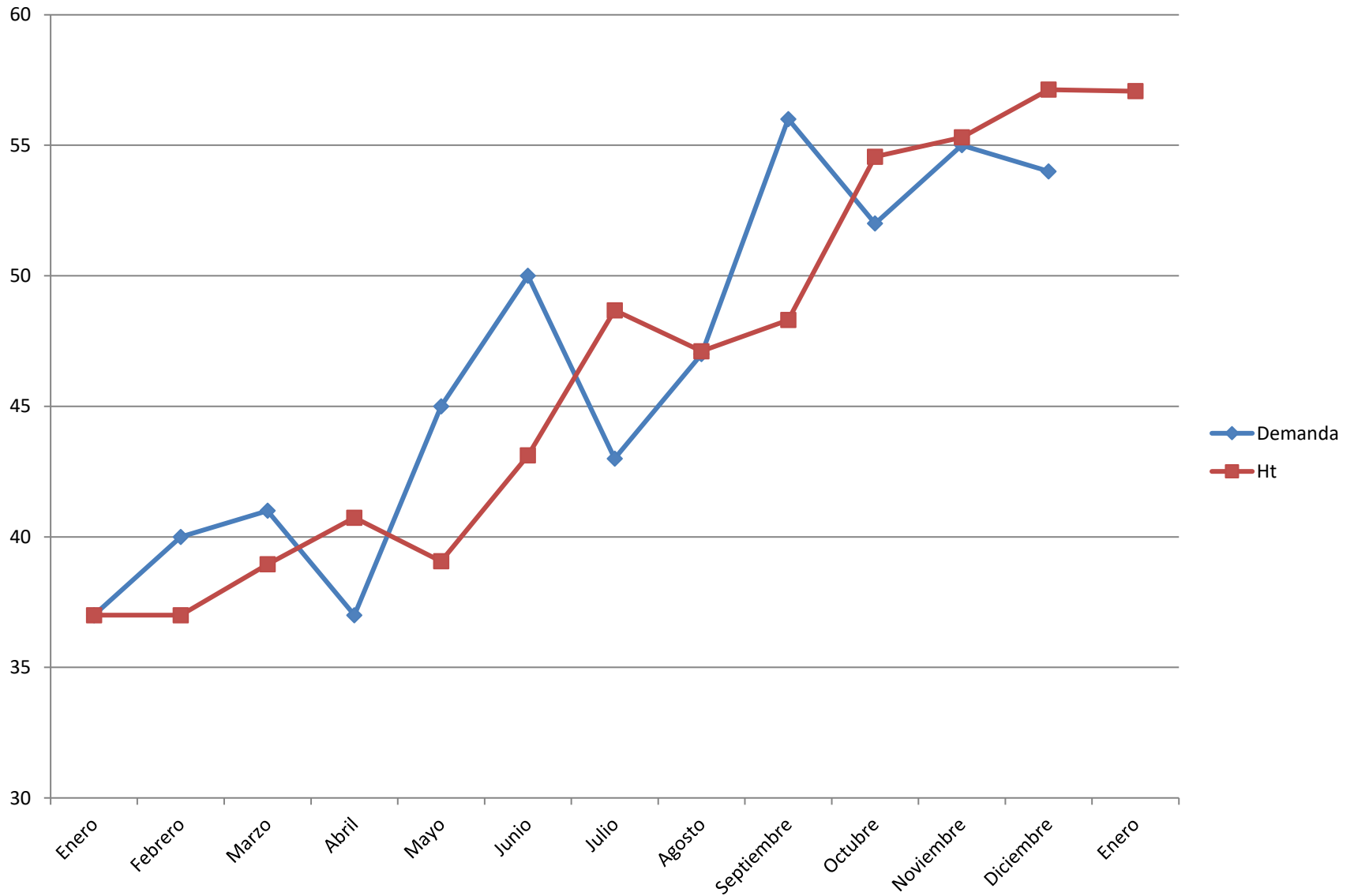
Ejemplo 9

- Realice un ajuste con tendencia para los datos del ejemplo 6.
- Use como constantes de suavizado $\alpha=0,5$ y $\beta=0,3$
- Estime la demanda del próximo mes de Enero según el modelo planteado.

Solución

	Demanda	Nivel	Tendencia	Pronóstico
Mes	X_t	F_t	T_t	H_t
Enero	37	37,00		37,00
Febrero	40	37,00	0,00	37,00
Marzo	41	38,50	0,45	38,95
Abril	37	39,98	0,76	40,73
Mayo	45	38,87	0,20	39,06
Junio	50	42,03	1,09	43,12
Julio	43	46,56	2,12	48,68
Agosto	47	45,84	1,27	47,11
Septiembre	56	47,05	1,25	48,31
Octubre	52	52,15	2,41	54,56
Noviembre	55	53,28	2,02	55,30
Diciembre	54	55,15	1,98	57,13
Enero		55,56	1,51	57,07

Solución



Suavizado doble en Minitab

- Minitab provee de dos opciones:
 - Determinación óptima:
 - Resuelve un modelo $ARIMA(0,2,2)$ que se verá posteriormente.
 - El valor de α se iguala al coeficiente de media móvil del método ARIMA.
 - Los valores de l_1 y β se determina usando las fórmulas en sentido contrario (técnica conocida como backcasting y usada en el Suavizado simple) con las predicciones obtenidas por el método ARIMA.
 - Parámetros especificados:
 - Se resuelve una regresión lineal con las observaciones y se asigna el valor previsto por la regresión para la observación 1 como la predicción inicial.



Tema 4. Series de Tiempo

Monitoreo y control de las predicciones

Control de pronósticos

- En una aplicación real es necesario controlar las desviaciones entre los valores reales y los obtenidos mediante predicciones.
- El objetivo es detectar cuándo el modelo deja de ser representativo de la situación real
- Entre las múltiples causas que pueden generar estas desviaciones cabe destacar:
 - Cambios de tendencia
 - Patrones originados por clima/fiestas/...
 - Promociones internas

Señal de rastreo

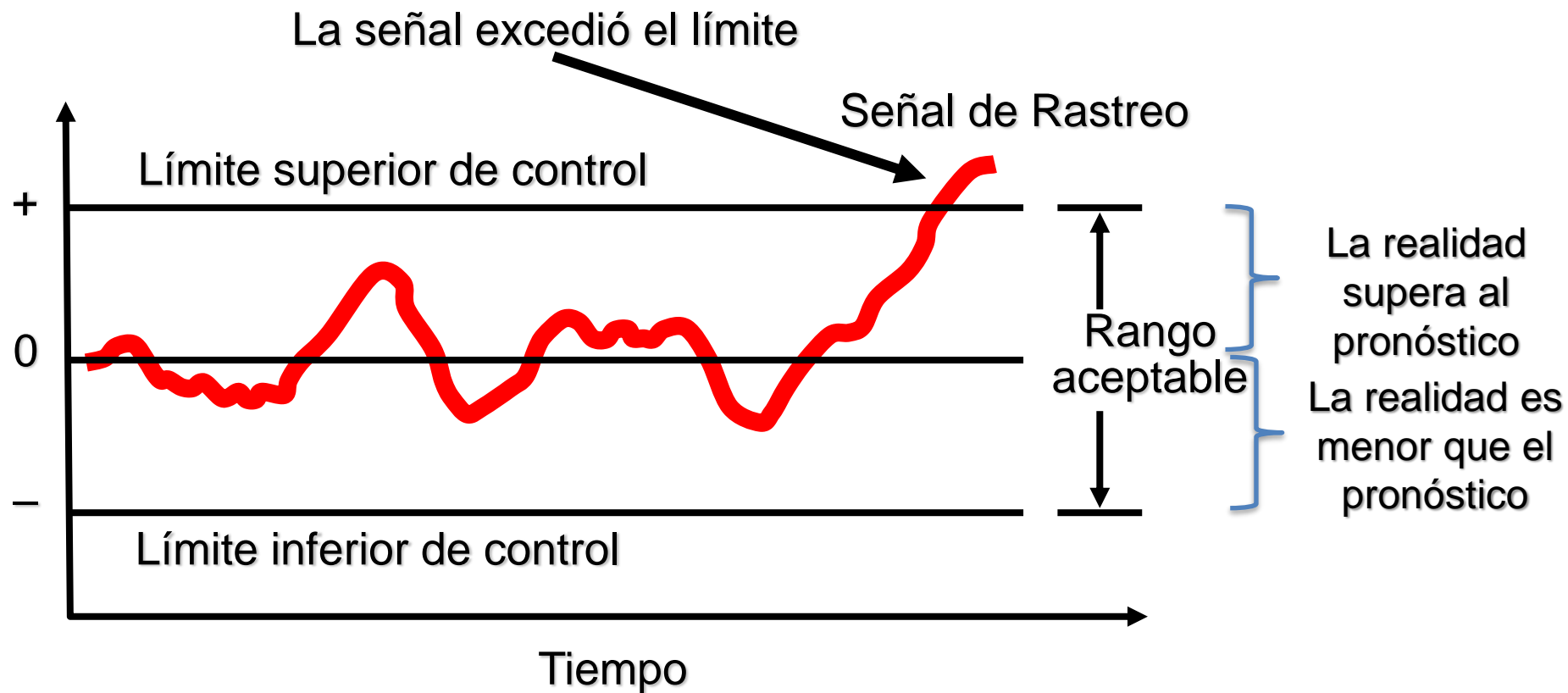
- La señal de rastreo es un método que permite medir si el pronóstico se ajusta a los cambios observados
- Se calcula como el cociente de la suma de errores de pronóstico (RSFE) y la desviación absoluta media (MAD)
 - Una buena señal de rastreo tiene valores bajos
 - Si el pronóstico está constantemente apreciando o despreciando el valor observado, se dice que el error del método sufre de sesgo.

Señal de rastreo

- Se puede entender que el valor corresponde al número de desviaciones absolutas que el valor está por encima (o por debajo) de los datos reales
- Se calcula (actualiza) para cada periodo
- El valor obtenido debe compararse con unos límites prefijados de control
- Se considera que el pronóstico está controlado si se encuentra entre estos límites de control.

$$Rastreo = \frac{\sum errores}{MAD} = \frac{\sum_t e_t}{\sum_t |e_t| / n}$$

- El numerador debería tender a 0 ya que considera suma de valores sin importar signo.
- El denominador debería ser constante una vez alcanzado un estacionario.



Errores

- Idealmente los errores deberían comportarse según una ley normal de media 0 y desviación estándar:

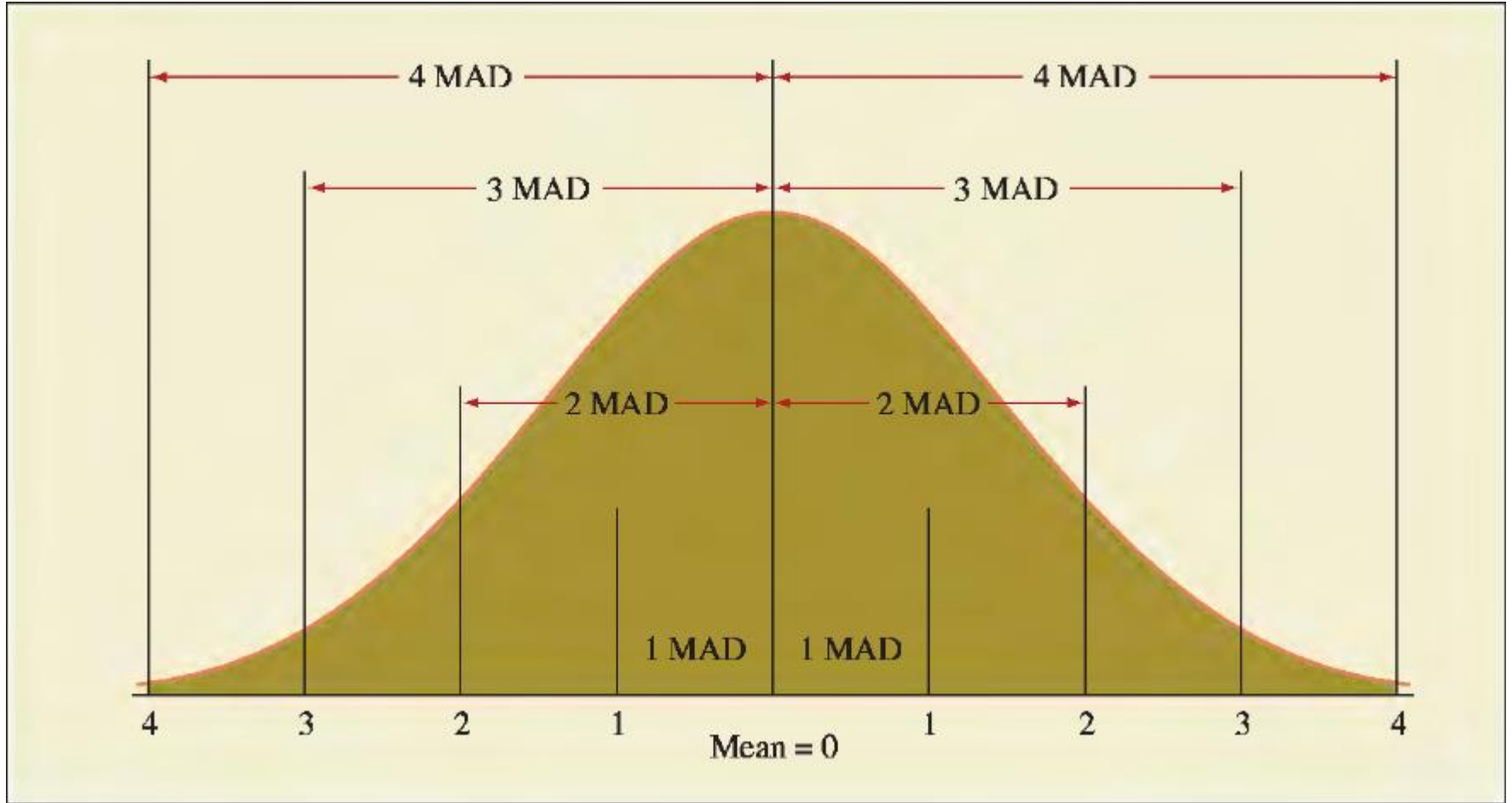
$$S = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n - 1}}, \quad e = X_i - F_i$$

- Además, existe la siguiente relación entre σ y MAD:

$$S = \left(\sqrt{p/2} \right) \times MAD \gg 1,25 \times MAD; \quad MAD = 0,798S \gg 0,8S$$

Número de MAD	Número de desviaciones estándar	% de puntos dentro de los límites de control
± 1	0,798	57,084
± 2	1,596	88,946
± 3	2,394	98,334
± 4	3,192	99,856

Curva de Bell asociada



Ejemplo 10

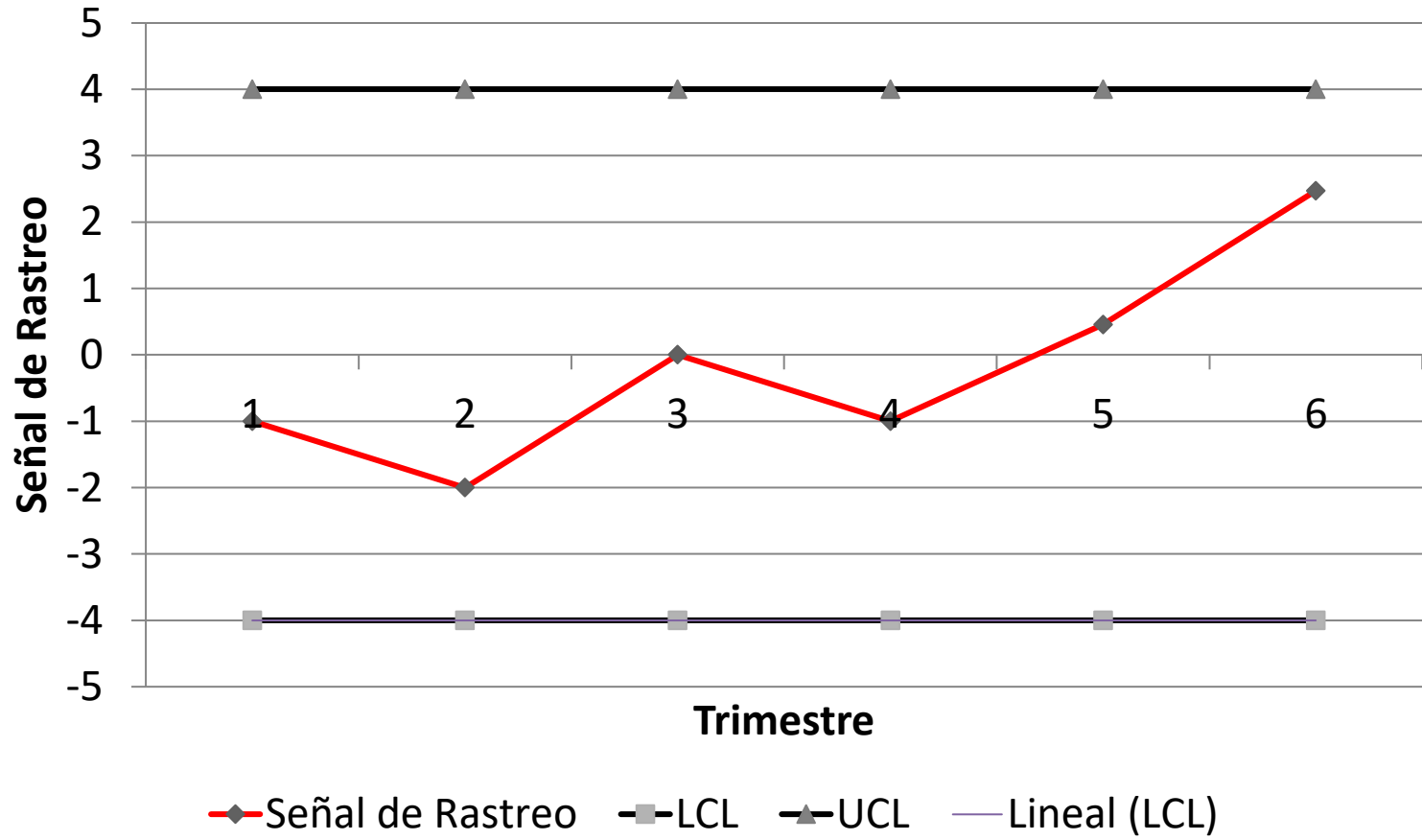
- Una empresa desea evaluar el desempeño de su método de pronóstico.
- Desarrolle una señal de rastreo que verifique si éste se haya dentro de los parámetros razonables, que se han definido como $\pm 4\text{MAD}$.
- Datos:

Trimestre	Demanda	Pronóstico
1	90	100
2	95	100
3	115	100
4	100	110
5	125	110
6	140	110

Trimestre	Demand a	Pronóstico o	Error	RSFE	MAD	Señal
1	90	100	-10	-10	10	-1
2	95	100	-5	-15	7,5	-2
3	115	100	15	0	10	0
4	100	110	-10	-10	10	-1
5	125	110	15	5	11	0,45
6	140	110	30	35	14,17	2,47

- El método está dentro de los límites fijados

Monitoreo de los Pronósticos





Tema 4. Series de Tiempo

Alisado exponencial
triple

Suavizado Exponencial Triple

- Intentan dar explicación a fenómenos de estacionalidad en los datos.
- Este factor puede considerarse aditivo o multiplicativo (en la práctica tiene más sentido lo segundo)
- El método de Winters (basado en el método de Holt) plantea una aproximación al problema basada en la técnica de suavizado. El método utiliza 3 constantes y 4 ecuaciones.

Método de Winters

$$\left. \begin{aligned} F_t &= a \frac{X_t}{S_{t-p}} + (1-a)(F_{t-1} + T_{t-1}) \\ T_t &= b(F_t - F_{t-1}) + (1-b)T_{t-1} \\ S_t &= g \frac{X_t}{F_t} + (1-g)S_{t-p} \end{aligned} \right\} W_{t+m} = (F_t + m \times T_t) S_{t+m-p}$$

Donde:

F_t : Nivel de la serie sin estacionalidad ni tendencia.

T_t : Factor de tendencia de la serie para el periodo t .

S_t : Factor de estacionalidad para el periodo t .

Método de Winters

X_t : Demanda del periodo t .

α, β, γ : Constantes de suavización. (entre 0 y 1).

p : Número de periodos en que aparece el efecto de estacionalidad.

m : Número de periodos a futuro a predecir

W_{t+m} : Predicción ofrecida por el método de Winters.

Explicación del método de previsión

- La fórmula de la previsión funciona de la siguiente manera:

$$W_{t+m} = \left(F_t + m \times T_t \right) S_{t+m-p}$$

- El primer factor (entre paréntesis) determina, por una parte, la previsión sin tendencia (F_t) y le suma el efecto de la tendencia m veces ($m \times T_t$) para predecir el nivel a m meses vista.
- Posteriormente ese valor se multiplica por el índice de estacionalidad del periodo que pertoque ($t+m-p$ indica el último periodo asociado al periodo $t+m$).

Procedimiento

- El método necesita un mínimo de dos ciclos ($2p$) de estacionalidad completos. Por ejemplo si los datos se presentan por meses y se cree que la estacionalidad es anual se necesitará un mínimo de 24 datos.
- Por comodidad, se considera que el periodo actual corresponde al periodo 0 y que el subíndice de los periodos pasados son negativos (X_{-2p+1}, \dots, X_{-1} para los $2p$ periodos citados anteriormente)

Método de cálculo

- Primero calcule una estimación de la media de cada estación de datos:

$$V_1 = \frac{1}{p} \mathring{a}_{i=-2p+1}^{-p} X_i \quad V_2 = \frac{1}{p} \mathring{a}_{i=-p+1}^0 X_i$$

- La estimación del factor de tendencia (pendiente) es igual a:

$$T_0 = \frac{V_2 - V_1}{p}$$

- Si $s > 2$ estaciones, calcule V_1, \dots, V_s para cada estación y defina el factor de tendencia como:

$$T_0 = \frac{V_s - V_1}{(s - 1)p}$$

Método de cálculo

- Para estimar el valor de la serie en el tiempo $t=0$ use:

$$F_0 = V_2 + T_0 \frac{p-1}{2}$$

- Los factores estacionales iniciales se calculan para cada periodo del que disponemos de datos y posteriormente se promedian para obtener un conjunto inicial de factores estacionales.

$$S_t = \frac{X_t}{V_i - T_0 \frac{p+1}{2} - j}$$

Para $-2p+1 \leq t \leq 0$, donde i es el subíndice de estación y j es el periodo de la estación

Método de cálculo

- Promedie los factores estacionales:

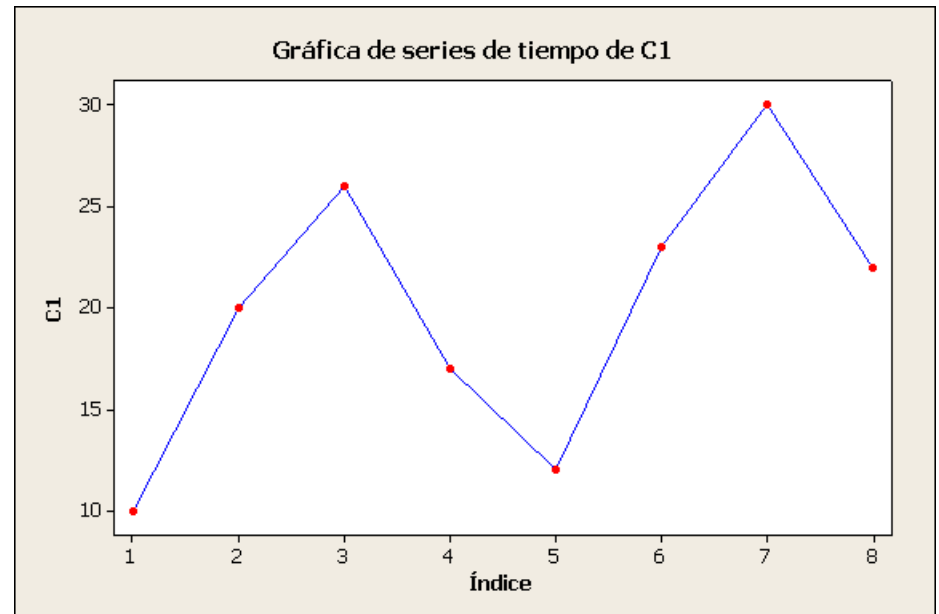
$$S_{-p+1} = \frac{S_{-2p+1} + S_{-p+1}}{2}, \dots, S_0 = \frac{S_{-p} + S_0}{2}$$

- Normalice los factores estacionales:

$$S_j = \frac{S_j}{\sum_{i=0}^{-p+1} S_i} \times p \quad \text{Para } -p+1 \leq j \leq 0$$

Ejemplo 11

- Supongamos que el historial de demanda de un producto para los últimos ocho meses es 10, 20, 26, 17, 12, 23, 30, 22.
- El gráfico de la serie muestra la existencia de dos estaciones completas (anuales) y una tendencia creciente.



Ejemplo 11

- Determinar las demandas del año próximo
- Determine las demandas del segundo, tercer y cuarto trimestre si la demanda del primer trimestre ha sido igual a 16 y que las constantes de suavización son $\alpha=0,2$; $\beta=0,1$ y $\gamma=0,1$.

Solución

- Disponemos del mínimo de dos años de datos.
- Inicializamos los parámetros de tendencia y nivel mediante:

$$V_1 = \frac{1}{p} \dot{a}_{i=-2p+1}^{-p} X_i = \frac{1}{4} (10 + 20 + 26 + 17) = 18,25$$

$$V_2 = \frac{1}{p} \dot{a}_{i=-p+1}^0 X_i = \frac{1}{4} (12 + 23 + 30 + 22) = 21,75$$

$$T_0 = \frac{V_2 - V_1}{p} = \frac{21,75 - 18,25}{4} = 0,875$$

$$F_0 = V_2 + T_0 \frac{p-1}{2} = 21,75 + 0,875 \frac{4-1}{2} = 23,0625$$

Solución

- Calculemos valores iniciales de estacionalidad

$$S_{-7} = \frac{X_{-7}}{V_1 - T_0 \left(\frac{4+1}{2} - 1 \right)} = \frac{10}{18,25 - 0,875 \left(\frac{4+1}{2} - 1 \right)} = 0,5904$$

$$S_{-6} = \frac{X_{-6}}{V_1 - T_0 \left(\frac{4+1}{2} - 1 \right)} = \frac{20}{18,25 - 0,875 \left(\frac{4+1}{2} - 1 \right)} = 1,1228$$

- De forma similar:

$$S_{-5} = 1,3913; S_{-4} = 0,8690; S_{-3} = 0,5872;$$

$$S_{-2} = 1,0792; S_{-1} = 1,3521; S_0 = 0,9539;$$

Solución

- Promediamos los factores de estacionalidad:

$$S_{-3} = (0,5904 + 0,5872) / 2 = 0,5888$$

$$S_{-2} = (1,1228 + 1,0792) / 2 = 1,1001$$

$$S_{-1} = (1,3913 + 1,3521) / 2 = 1,3717$$

$$S_0 = (0,8690 + 0,9539) / 2 = 0,9115$$

- Y rectificamos:

$$S_{-3} = 0,5888 \times 4 / (S_{-3} + S_{-2} + S_{-1} + S_0) = 0,5928$$

$$S_{-2} = 1,1001 \times 4 / (S_{-3} + S_{-2} + S_{-1} + S_0) = 1,1085$$

$$S_{-1} = 1,3717 \times 4 / (S_{-3} + S_{-2} + S_{-1} + S_0) = 1,3810$$

$$S_0 = 0,9115 \times 4 / (S_{-3} + S_{-2} + S_{-1} + S_0) = 0,9177$$

Solución

- Ya se disponen de todos los parámetros y por tanto puede contestarse la primera pregunta:

Determinar las demandas del año próximo

$$W_{t+m} = (F_t + m \times T_t) S_{t+m-p}$$

$$W_1 = (F_0 + 1 \times T_0) S_{-3} = (23,0625 + 1 \times 0,875) \times 0,5928 = 14,1899$$

$$W_2 = (F_0 + 2 \times T_0) S_{-2} = (23,0625 + 2 \times 0,875) \times 1,1085 = 27,5044$$

$$W_3 = (F_0 + 3 \times T_0) S_{-1} = (23,0625 + 3 \times 0,875) \times 1,3810 = 35,4756$$

$$W_4 = (F_0 + 4 \times T_0) S_{-0} = (23,0625 + 4 \times 0,875) \times 0,9177 = 24,3757$$

Solución

- Para la segunda pregunta, primero actualizamos las componentes del pronóstico:

$$F_1 = a \frac{X_1}{S_{-3}} + (1 - a)(F_0 + T_0) =$$

$$0,2 \frac{16}{0,5928} + 0,8(23,0625 + 0,875) = 24,5482$$

$$T_1 = b(F_1 - F_0) + (1 - b)T_0 =$$

$$0,1(24,5482 - 23,0625) + 0,9 \times 0,875 = 0,9361$$

$$S_1 = g \frac{X_1}{F_1} + (1 - g)S_{-3} =$$

$$0,1 \frac{16}{24,5482} + 0,9 \times 0,5928 = 0,5987$$

Solución

- El resto de factores estacionales son:

$$S_{-2} = 1,1085; \quad S_{-1} = 1,3810; \quad S_0 = 0,9177$$

- Rectificamos para que sumen 4:

$$S_{-2} = \frac{1,1085}{(1,1085 + 1,3810 + 0,9177 + 0,5987)} \cdot 4 = 1,1069$$

$$S_{-1} = \frac{1,3810}{(1,1085 + 1,3810 + 0,9177 + 0,5987)} \cdot 4 = 1,3790$$

$$S_0 = \frac{0,9177}{(1,1085 + 1,3810 + 0,9177 + 0,5987)} \cdot 4 = 0,9163$$

$$S_1 = \frac{0,5987}{(1,1085 + 1,3810 + 0,9177 + 0,5987)} \cdot 4 = 0,5978$$

Solución

- Ya podemos realizar un pronóstico de los trimestres 2, 3 y 4:

$$W_{t+m} = (F_t + m \times T_t) S_{t+m-p}$$

$$W_2 = (F_1 + 1 \times T_1) S_{-2} = (24,5482 + 1 \times 0,9361) \times 1,1069 = 28,2074$$

$$W_3 = (F_1 + 2 \times T_1) S_{-1} = (24,5482 + 2 \times 0,9361) \times 1,3790 = 36,4340$$

$$W_4 = (F_1 + 3 \times T_0) S_0 = (24,5482 + 3 \times 0,9361) \times 0,9163 = 25,0673$$

Suavizado triple en Minitab

- Se conoce como el modelo de Winters
- Permite modelos lineales y aditivos
- Todos los coeficientes se obtienen resolviendo un modelo de regresión que determine pendiente y punto de origen.
- Los coeficientes de estacionalidad iniciales se obtienen usando variables dummy en la regresión que determinen la estacionalidad.



Tema 4. Series de Tiempo

Métodos de descomposición

Introducción

- Un método de descomposición es aquél que intenta explicar la demanda a través de su descomposición en sus componentes:
 - Tendencia
 - Estacionalidad
 - Ciclo
- Poseen diversas ventajas a nivel práctico:
 - Su representación es sencilla
 - Son fáciles de comprender
 - Son fáciles de integrar en un proceso de toma de decisiones

Método de descomposición

- La ecuación general del método es la siguiente:

$$X_t = f(S_t, T_t, C_t, I_t)$$

- S_t : Componente de estacionalidad.
- T_t : Componente de tendencia.
- C_t : Componente de ciclo.
- I_t : Componente irregular.

Modelos para la demanda

- La demanda, representada por una serie de tiempo, puede ser descrita a través de diferentes modelos. Los más populares son el aditivo y el multiplicativo.

- El modelo aditivo viene dado por:

$$X_t = S_t + T_t + C_t + I_t$$

- El modelo multiplicativo viene dado por:

$$X_t = S_t \times T_t \times C_t \times I_t$$

- El segundo modelo puede linearizarse.

Modelo Multiplicativo

- Centraremos la atención en el modelo multiplicativo: $X_t = S_t \times T_t \times C_t \times I_t$
- Puede entenderse en términos de cambios porcentuales (recordar interpretación logarítmica de las variables en una regresión)
- En caso de disponer de una tendencia, es posible eliminar de la demanda mediante: X_t / T_t
- Así, si $X_t / T_t = 1,05$, puede entenderse que la observación está un 5% por encima que lo esperable por la tendencia.

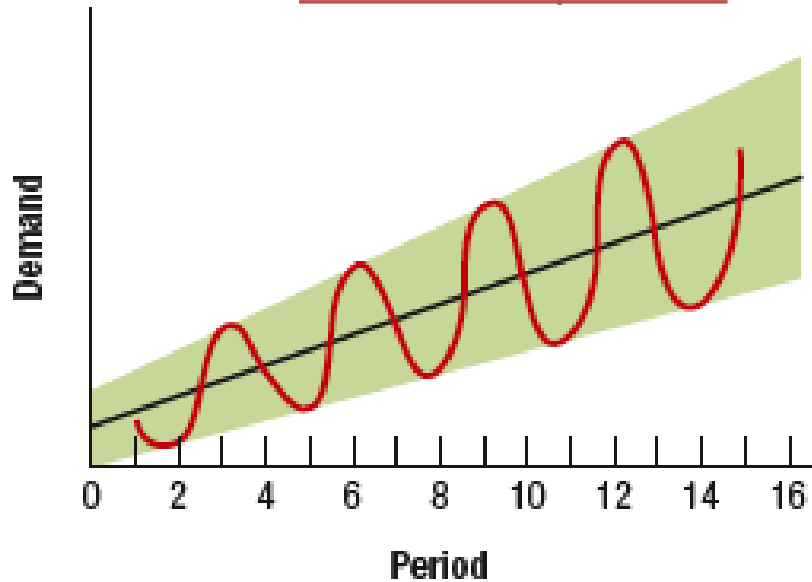
Modelo aditivo

- El modelo aditivo es: $X_t = S_t + T_t + C_t + I_t$
- Debe entenderse en unidades absolutas, no porcentuales
- Para eliminar la tendencia es suficiente con restarla de la observación: $X_t - T_t$
- Así, si $X_t - T_t = 320$, puede entenderse que se venden 320 unidades más de las que se esperaría por la tendencia.

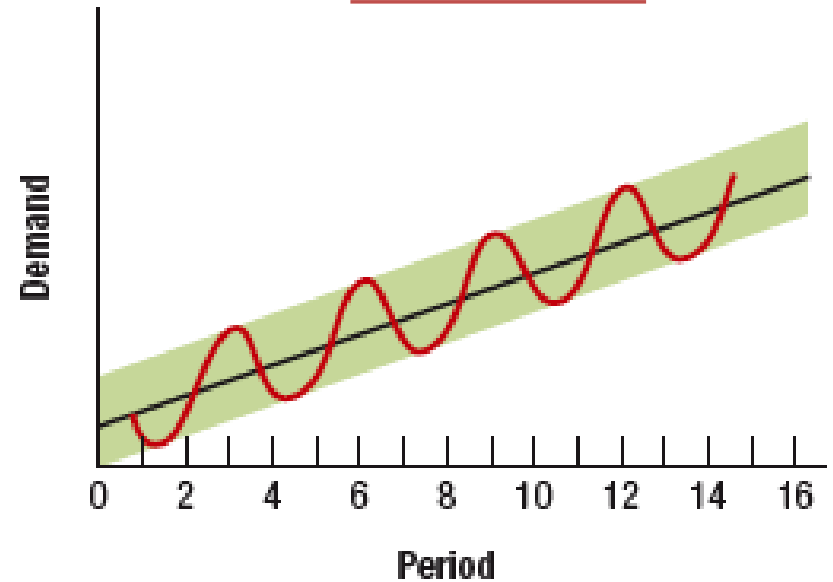
Efecto de la estacionalidad

- A continuación pueden verse los patrones que tendrán las series con estacionalidad aditiva o multiplicativa

Patrón Multiplicativo



Patrón Aditivo



Cálculo de estacionalidad

- La estacionalidad representa el patrón causado por la estación en la serie.
- Para aislarlo, se utilizará una media móvil simple corregida (esto, es “centrada”)
- El número de periodos a promediar dependerá de la estructura de las observaciones:
- Estos promedios eliminan variaciones estacionales y compensan irregularidades (aleatoriedad)
- Matemáticamente $MA = T \times C \rightarrow X/MA = S \times I$

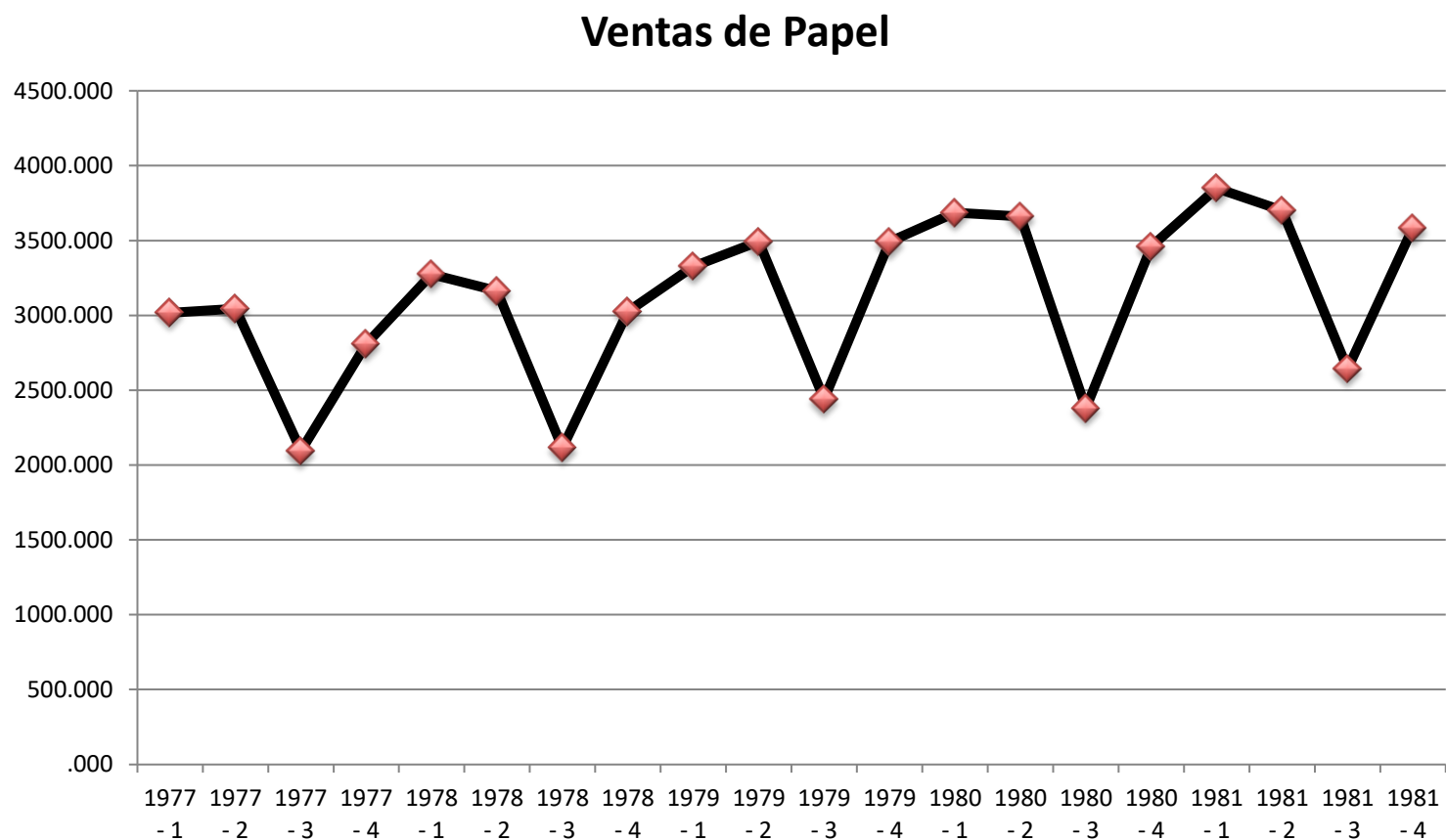
Ejemplo 12

- Utilice los datos de la venta de papel en Francia entre 1977 y 1981 para calcular los factores de estacionalidad.

Año	Trimestre	Ventas de Papel
1977	1	3.017,60
	2	3.043,54
	3	2.094,35
	4	2.809,84
1978	1	3.274,80
	2	3.163,28
	3	2.114,31
	4	3.024,57
1979	1	3.327,48
	2	3.493,48
	3	2.439,93
	4	3.490,79

Año	Trimestre	Ventas de Papel
1980	1	3.685,08
	2	3.661,32
	3	2.378,43
	4	3.459,55
1981	1	3.849,63
	2	3.701,18
	3	2.642,38
	4	3.585,52

- Serie:



Solución

- En el gráfico anterior podemos apreciar que hay estacionalidad.
- Para aislar la estacionalidad, calculamos MA(4). Luego como deben ser “centrados”, notamos que por ejemplo el centro de los períodos 1, 2, 3 y 4 es 2,5, así como el centro de 2, 3, 4, 5 es 3,5. para evitar la problemática de los “medios” promediamos MA(4) de dos periodos consecutivos, creándose promedios móviles dobles quedan en el periodo que corresponde: 3, 4, 5,...

Solución

- La opción anterior (dobles promedios) es una de las alternativas disponibles.
- Una vez calculados los promedios móviles dobles, la demanda observada en el respectivo periodo se divide por el promedio doble respectivo, obteniendo los índices iniciales de estacionalidad.
- La siguiente diapositiva muestra los resultados del proceso hasta aquí realizado

Año	Trimestre	Periodo	Ventas de Papel	Promedio Movable	Movable Doble	Estacionalidad
1977	1	1	3017,6			
	2	2	3043,54			
	3	3	2094,35	2741,33	2773,48	75,51%
	4	4	2809,84	2805,63	2820,60	99,62%
1978	1	5	3274,8	2835,57	2838,06	115,39%
	2	6	3163,28	2840,56	2867,40	110,32%
	3	7	2114,31	2894,24	2900,83	72,89%
	4	8	3024,57	2907,41	2948,69	102,57%
1979	1	9	3327,48	2989,96	3030,66	109,79%
	2	10	3493,48	3071,37	3129,64	111,63%
	3	11	2439,93	3187,92	3232,62	75,48%
	4	12	3490,79	3277,32	3298,30	105,84%
1980	1	13	3685,08	3319,28	3311,59	111,28%
	2	14	3661,32	3303,91	3300,00	110,95%
	3	15	2378,43	3296,10	3316,66	71,71%
	4	16	3459,55	3337,23	3342,22	103,51%
1981	1	17	3849,63	3347,20	3380,19	113,89%
	2	18	3701,18	3413,19	3428,93	107,94%
	3	19	2642,38	3444,68		
	4	20	3585,52			

Solución

- Con los coeficientes, se agrupan y se corrigen para que la suma de coeficientes sea igual a la cantidad de periodos (en este caso 4).
- En caso que la estacionalidad se considerara aditiva se corregirían para que la suma de coeficientes fuera igual a 0.
- Ambas operaciones requieren de un cambio de escala de valores. En el caso multiplicativo el reparto se realiza multiplicando cada coeficiente por el número de periodos y dividiéndolo por la suma original.

Solución

- Recordar que en este caso la suma de los índices debiera ser 4.

Índices Estacionales Sin Ajuste					
Año/Trimestre	1	2	3	4	
1977			0,7551	0,9962	
1978	1,1539	1,1032	0,7289	1,0257	
1979	1,0979	1,1163	0,7548	1,0584	
1980	1,1128	1,1095	0,7171	1,0351	
1981	1,1389	1,0794			
Suma	4,5035	4,4083	2,9559	4,1154	15,9831
Promedio	1,1259	1,1021	0,7390	1,0288	3,9958

Factor de Ajuste					1,0010572
Ajustados:	1,1271	1,1032	0,7398	1,0299	

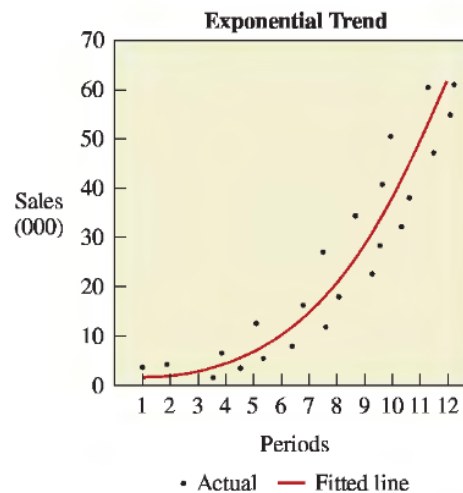
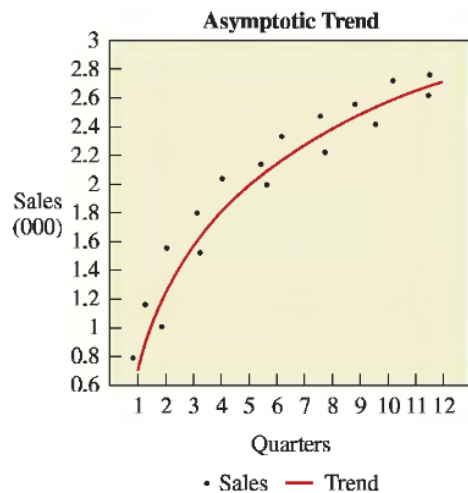
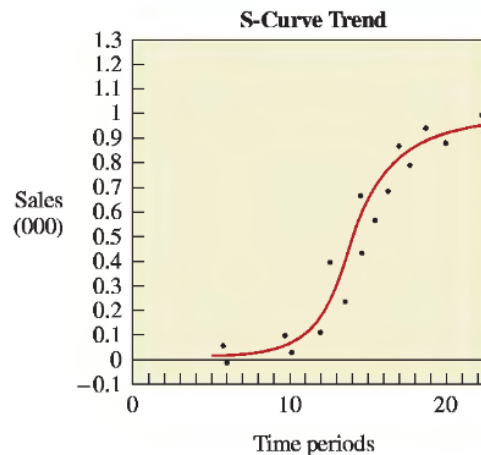
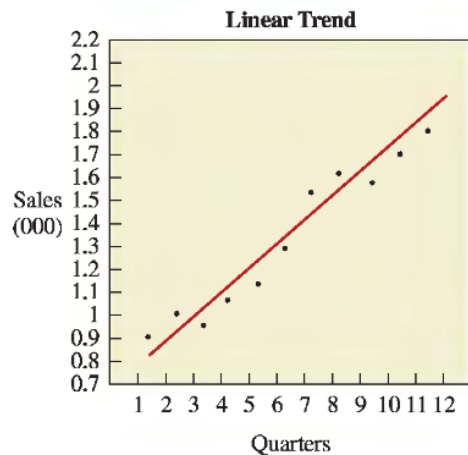
Solución

- Los índices que aparecen en la última fila (promedio) de la tabla anterior son los factores estacionales que se usarán más adelante, es decir: $S_1 = 1,1271$, $S_2 = 1,1032$, $S_3 = 0,7398$ y $S_4 = 1,0299$ para los trimestres 1, 2, 3 y 4 respectivamente.

Tendencia

- La tendencia es un movimiento de largo plazo en una serie de tiempo.
- Es la dirección básica (tendencia ascendente o descendente) y el grado de cambio en una serie de tiempo, cuando se han asignado los demás componentes.
- En ocasiones se puede describir a través de una línea recta o bien de una curva suavizada.
- Existen varios tipos de tendencia: lineal, cuadrática, exponencial, asintótica (logarítmica), curva S, etc.

Tipos de tendencia



Tendencia lineal

- Los datos deben mostrar una tendencia creciente o decreciente pero relativamente lineal.
- El modelo a utilizar es: $T_t = a + bt + \varepsilon_t$
- La demanda observada es representada por T_t .
- El error se considera que se distribuye según una ley normal.

Tendencia lineal

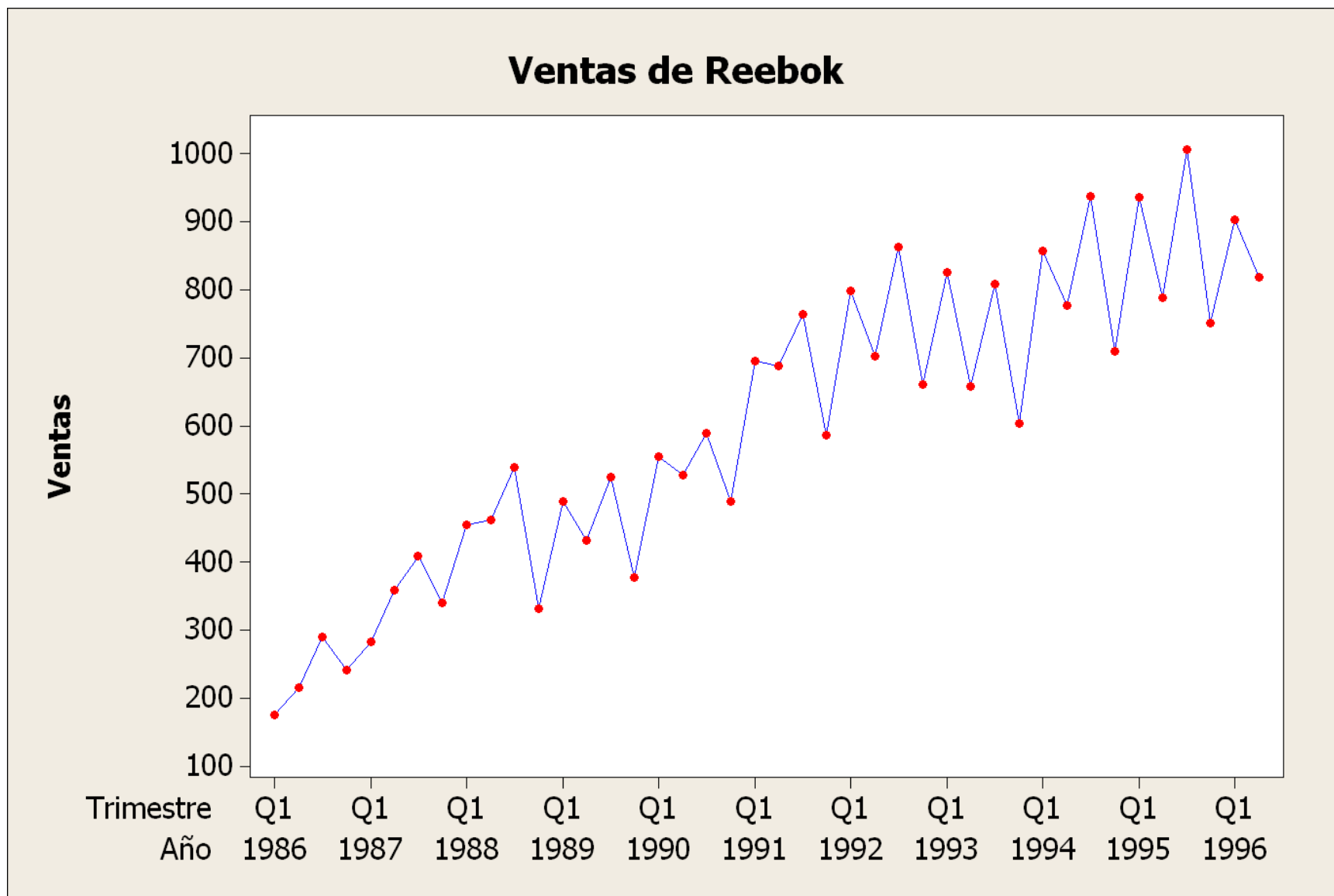
- El intercepto (a) de la ecuación representa el nivel de demanda inicial en los datos analizados.
- La pendiente (b) representa el cambio esperado de la demanda desde un periodo al siguiente.
- Los parámetros del modelo y la adecuación del modelo se realiza mediante un análisis de regresión.

Ejemplo 13

- La tabla siguiente muestra las ventas trimestrales realizadas por Reebok desde el primer trimestre de 1986 hasta el segundo trimestre de 1996.
- Las ventas se han incrementado desde US\$ 174,52 millones en el primer trimestre de 1986 hasta US\$ 817,57 millones en el segundo trimestre de 1996.
- ¿El correcto un ajuste lineal?
- ¿Los residuos son aleatorios?

Periodo	Trimestre	Ventas
1	T1-86	174,52
2	T2-86	215,57
3	T3-86	289,06
4	T4-86	240,25
5	T1-87	281,77
6	T2-87	358,79
7	T3-87	408,42
8	T4-87	340,22
9	T1-88	453,98
10	T2-88	462,03
11	T3-88	539,36
12	T4-88	330,57
13	T1-89	488,49
14	T2-89	431,48
15	T3-89	524,51
16	T4-89	377,61
17	T1-90	554,40
18	T2-90	527,23
19	T3-90	588,89
20	T4-90	488,72
21	T1-91	695,27

Periodo	Trimestre	Ventas
22	T2-91	688,30
23	T3-91	764,29
24	T4-91	586,61
25	T1-92	797,37
26	T2-92	701,98
27	T3-92	863,21
28	T4-92	660,06
29	T1-93	825,22
30	T2-93	657,61
31	T3-93	808,49
32	T4-93	602,57
33	T1-94	857,37
34	T2-94	776,75
35	T3-94	937,15
36	T4-94	709,15
37	T1-95	935,48
38	T2-95	788,69
39	T3-95	1.005,98
40	T4-95	751,30
41	T1-96	902,92
42	T2-96	817,57



Solución

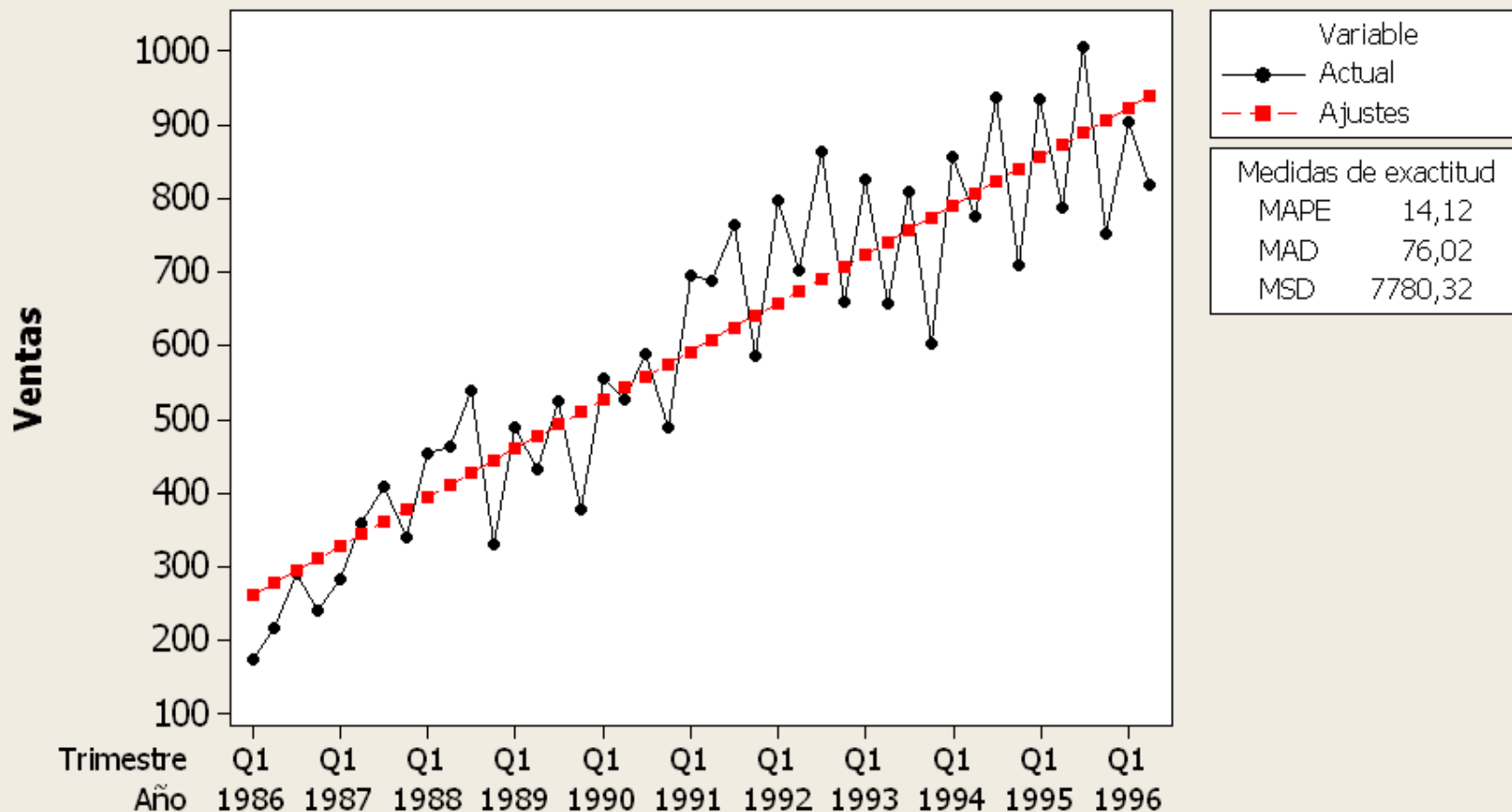
- El gráfico anterior muestra con claridad la tendencia creciente en las ventas de Reebok. A simple vista, parece apropiado realizar un ajuste de tendencia lineal a los datos.
- Lo que primero debe hacerse es codificar los trimestres desde 1 a 42, y de esta forma t se transforma en la “variable independiente”.
- Una vez hecho eso, Minitab entrega los siguientes resultados:

Solución

Gráfica de análisis de tendencia de Ventas

Modelo de tendencia lineal

$$Y_t = 244,8 + 16,5 * t$$



Análisis de regresión: Ventas vs. Periodo

La ecuación de regresión es

$$\text{Ventas} = 245 + 16,5 \text{ Periodo}$$

Predictor	Coef	Coef. de EE	T	P
Constante	244,82	28,40	8,62	0,000
Periodo	16,530	1,151	14,37	0,000

S = 90,3844 R-cuad. = 83,8% R-cuad. (ajustado) = 83,4%

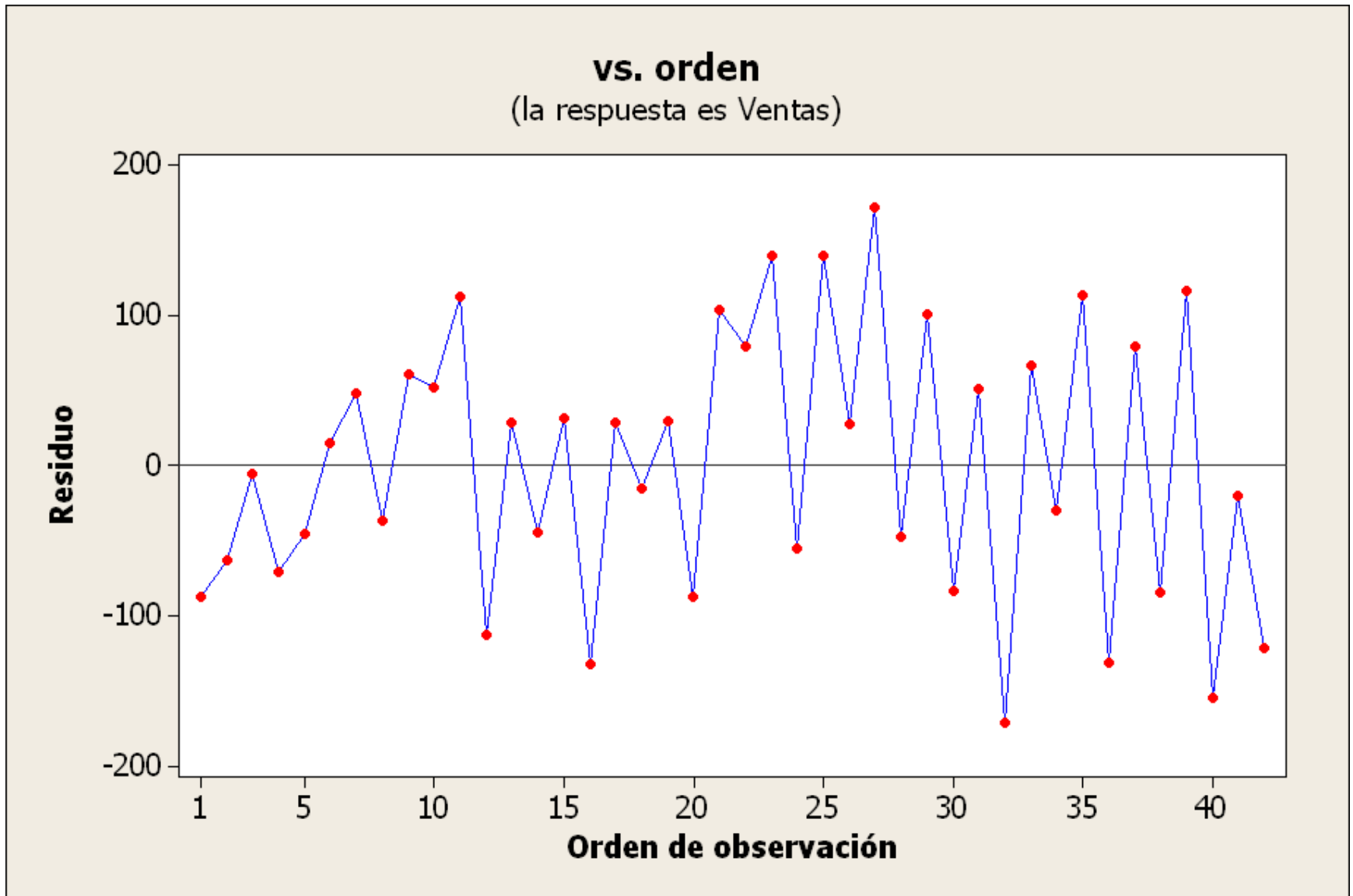
Análisis de varianza

Fuente	GL	SC	MC	F	P
Regresión	1	1686122	1686122	206,40	0,000
Error residual	40	326773	8169		
Total	41	2012895			

Análisis de la solución

- Del gráfico anterior, notamos que el ajuste de tendencia lineal parece bastante aceptable.
- Además, cuantitativamente podemos ratificar lo anterior puesto que $R^2 = 83,8\%$ y el test F de la regresión es altamente significativo.
- Las medidas de error reportadas por Minitab son $MAPE = 14,12$, $MAD = 76,02$ y $MSE = 7780,32$.
- Por tanto: las ventas aumentan, en promedio, 16,53 millones de US\$ por trimestre.

Autocorrelación



- Del gráfico de los residuos observamos lo siguiente:
 - Existe cierta regularidad en los residuos.
 - Ellos se comportan en forma de zigzag más que de forma aleatoria.
 - Probablemente hay algún patrón estacional en los datos de ventas, que podría determinarse con un método de pronóstico más sofisticado.
- El test de Durbin-Watson indicaría que la solución muestra autocorrelación negativa, por lo que los residuos no se comportan correctamente.

Ejemplo 14

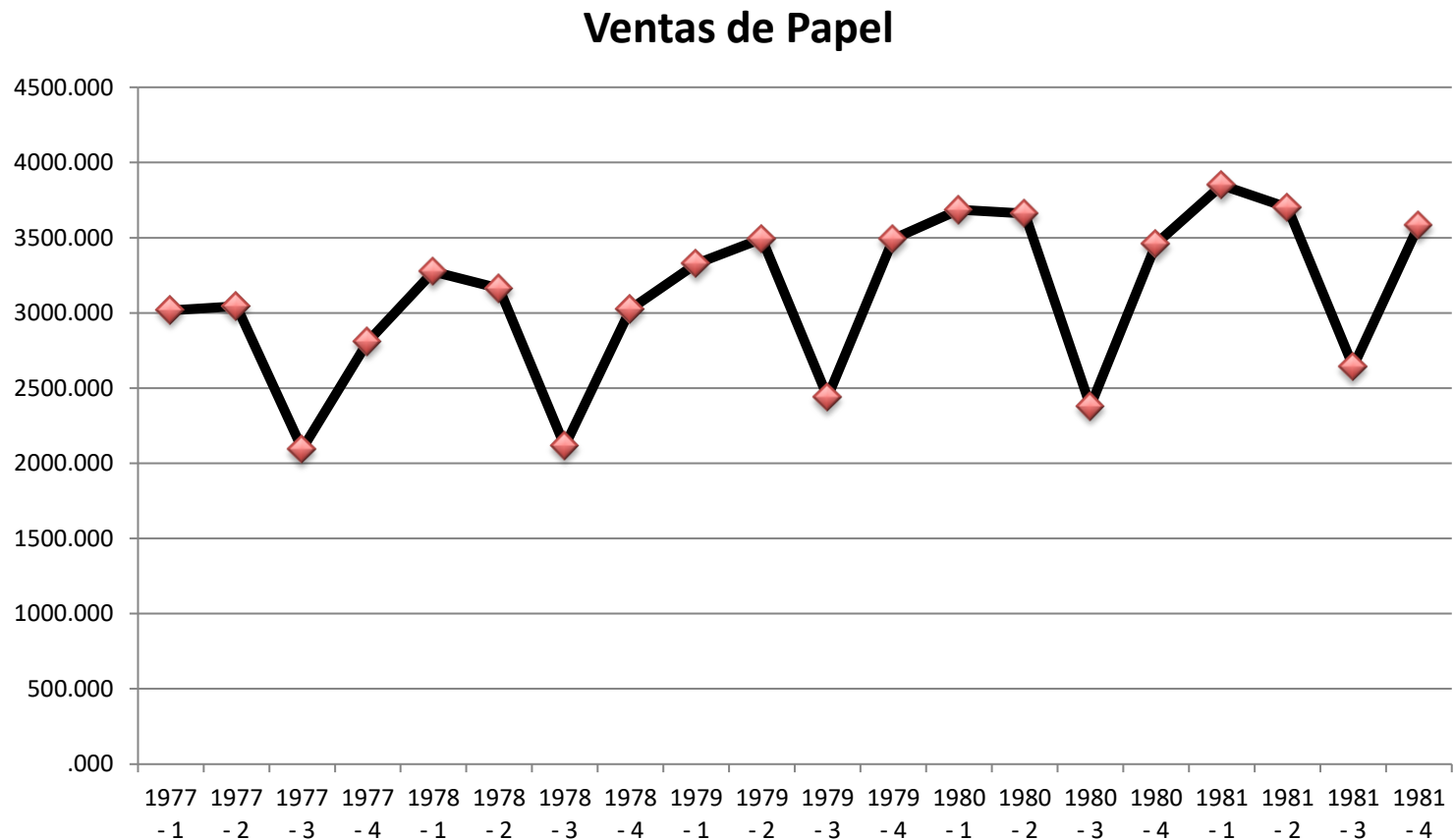
- Utilice los datos de la venta de papel en Francia entre 1977 y 1981 (ejercicio 12) para calcular la tendencia en ellas.

Año	Trimestre	Ventas de Papel
1977	1	3.017,60
	2	3.043,54
	3	2.094,35
	4	2.809,84
1978	1	3.274,80
	2	3.163,28
	3	2.114,31
	4	3.024,57
1979	1	3.327,48
	2	3.493,48
	3	2.439,93
	4	3.490,79

Año	Trimestre	Ventas de Papel
1980	1	3.685,08
	2	3.661,32
	3	2.378,43
	4	3.459,55
1981	1	3.849,63
	2	3.701,18
	3	2.642,38
	4	3.585,52

Recordatorio

- La serie se comporta como:

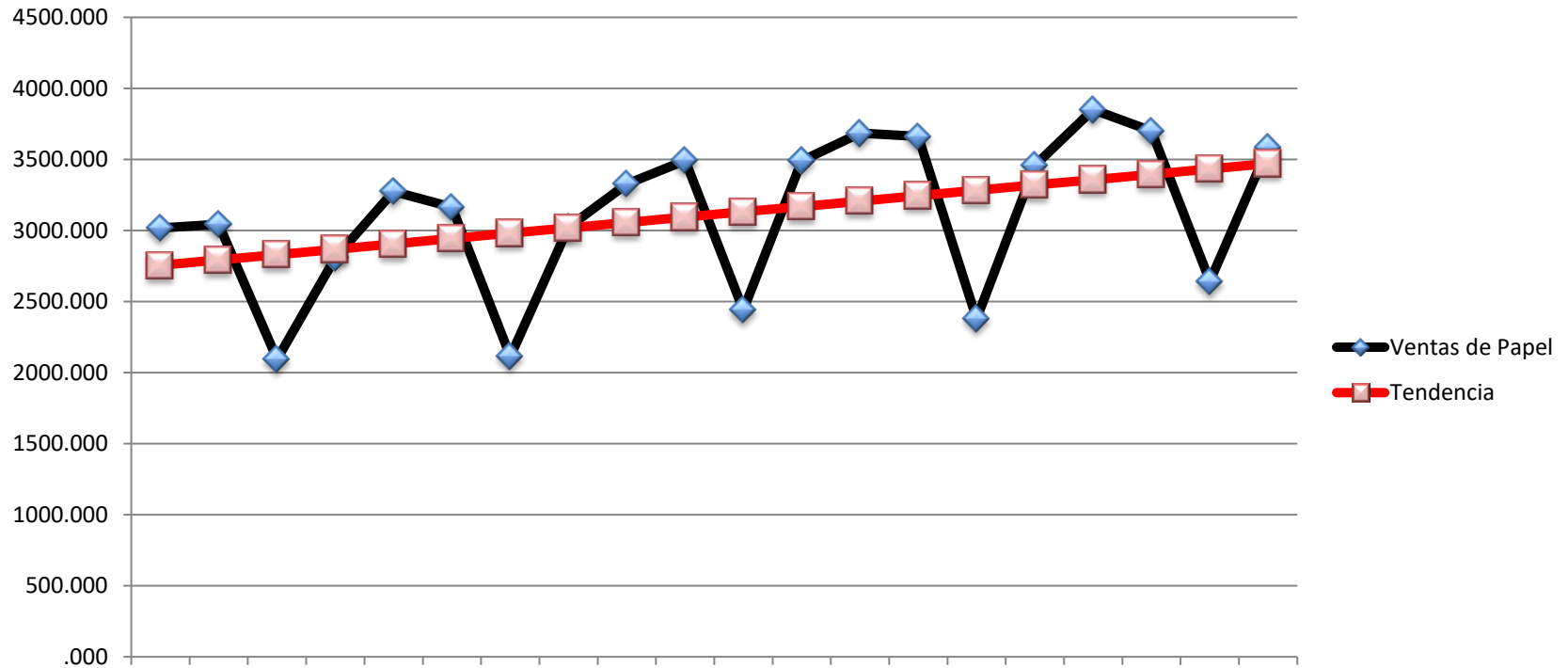


Solución

- En el gráfico anterior apreciamos que existe una tendencia creciente en las ventas de papel, pero ésta parece completamente lineal y no exponencial.
- Así, el modelo que utilizaremos es $T_t = a + b \times t$
- Los valores de intercepto y pendiente son: $a = 2.717,1129$, $b = 37,6895$.
- A continuación se muestra la tendencia calculada sobreimpuesta en la serie.

- La serie se comporta como:

Tendencia de la serie

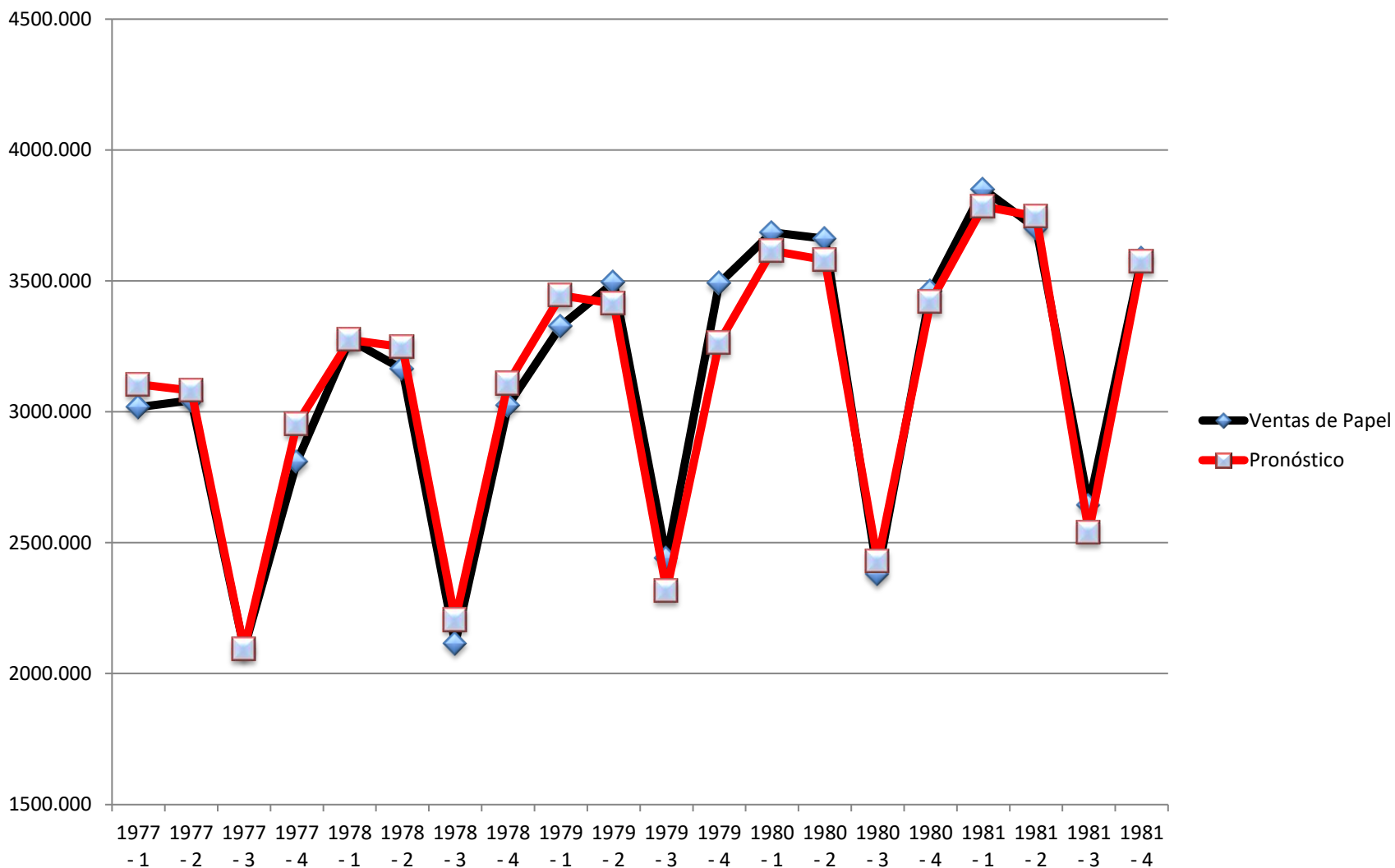


Recapitulación ejercicios 12 y 14

- Teniendo ya calculados los índices de estacionalidad y la tendencia podríamos hacer pronósticos sobre las ventas de papel en Francia.
- La expresión matemática para realizar los pronósticos de las ventas es $F_t = T_t \times S_t = (a + b \times t) \times S_t$
- Como medidas de desempeño se tiene que: $MSE = 8.568,13$, $MAD = 77,03$ y $MAPE = 2,54\%$.

Resultados conjuntos

Pronósticos de las ventas con Tendencia y Estacionalidad



Caso tendencia exponencial

- En contraste a la tendencia lineal, una tendencia exponencial es apropiada cuando la serie de tiempo cambia en un porcentaje constante (en lugar de una cantidad constante) cada periodo.
- El modelo a utilizar es: $T_t = ae^{bt}\varepsilon_t$
- La demanda observada es representada por T_t .
- La constante b (expresada como porcentaje) es aproximadamente el porcentaje de cambio por periodo. Una mejor estimación de este porcentaje es $e^b - 1$.

Tendencia exponencial

- Por ejemplo, si $b = 0,05$, la serie de tiempo de la demanda aumenta aproximadamente en un 5% por periodo. Por otra parte, si $b = -0,05$, la serie de tiempo de la demanda disminuye aproximadamente en un 5% por periodo.
- Para poder hacer la estimación de las constantes a y b , se deben aplicar logaritmos al modelo:

$$T_t = ae^{bt} e_t$$

$$\ln(T_t) = \ln(a) + bt + \ln(e_t)$$

$$\ln(T_t) = a' + bt + e'_t$$

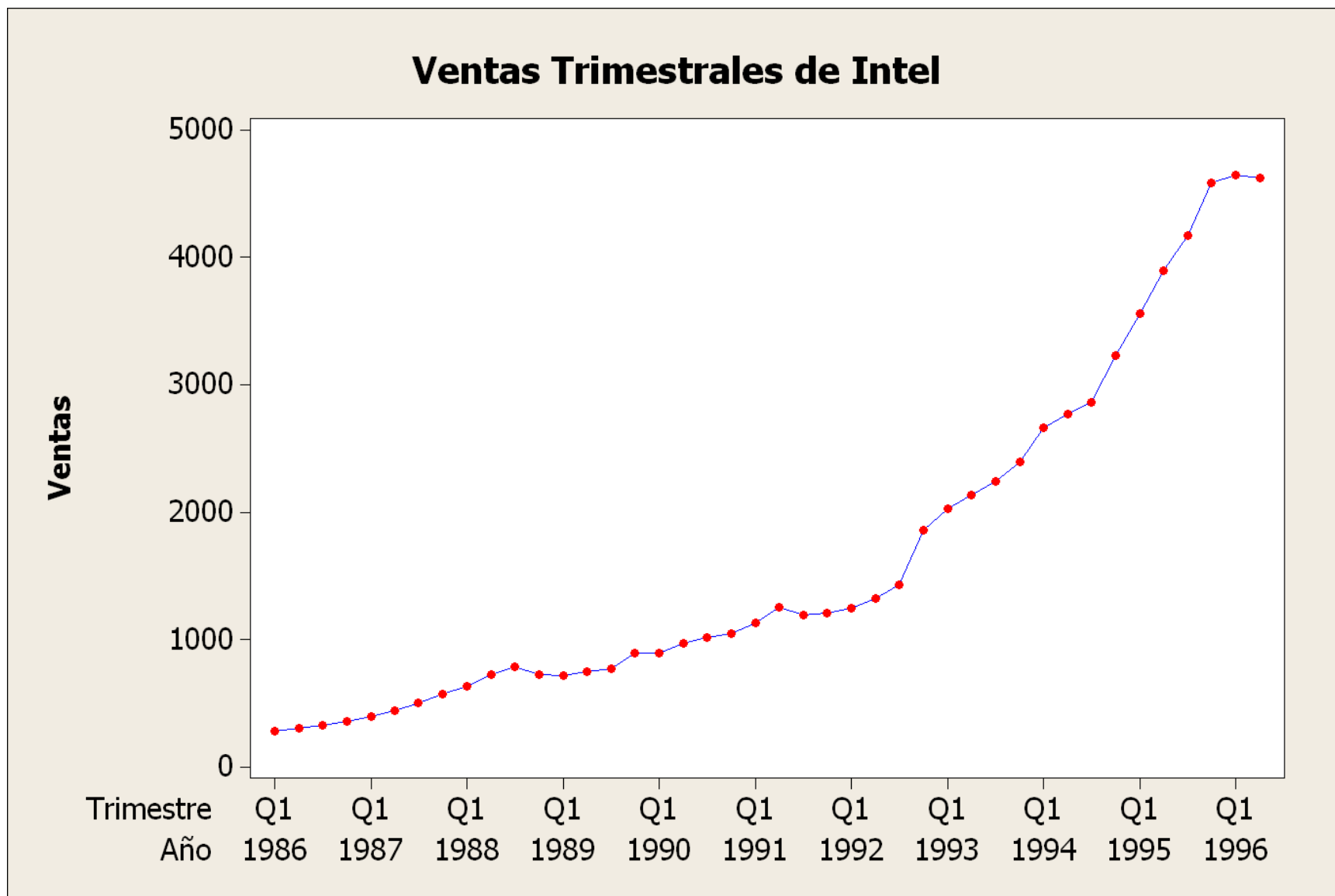
Ejemplo 15

- A continuación se presentan las ventas trimestrales para la compañía productora de chips Intel, desde comienzos de 1986 hasta el segundo trimestre de 1996.
- Cada valor de ventas está expresado en millones de dólares.
- Compruebe que un ajuste de tendencia exponencial se ajusta bastante bien a estos datos de ventas.
- Posteriormente, estime la relación e interprétela.

Datos

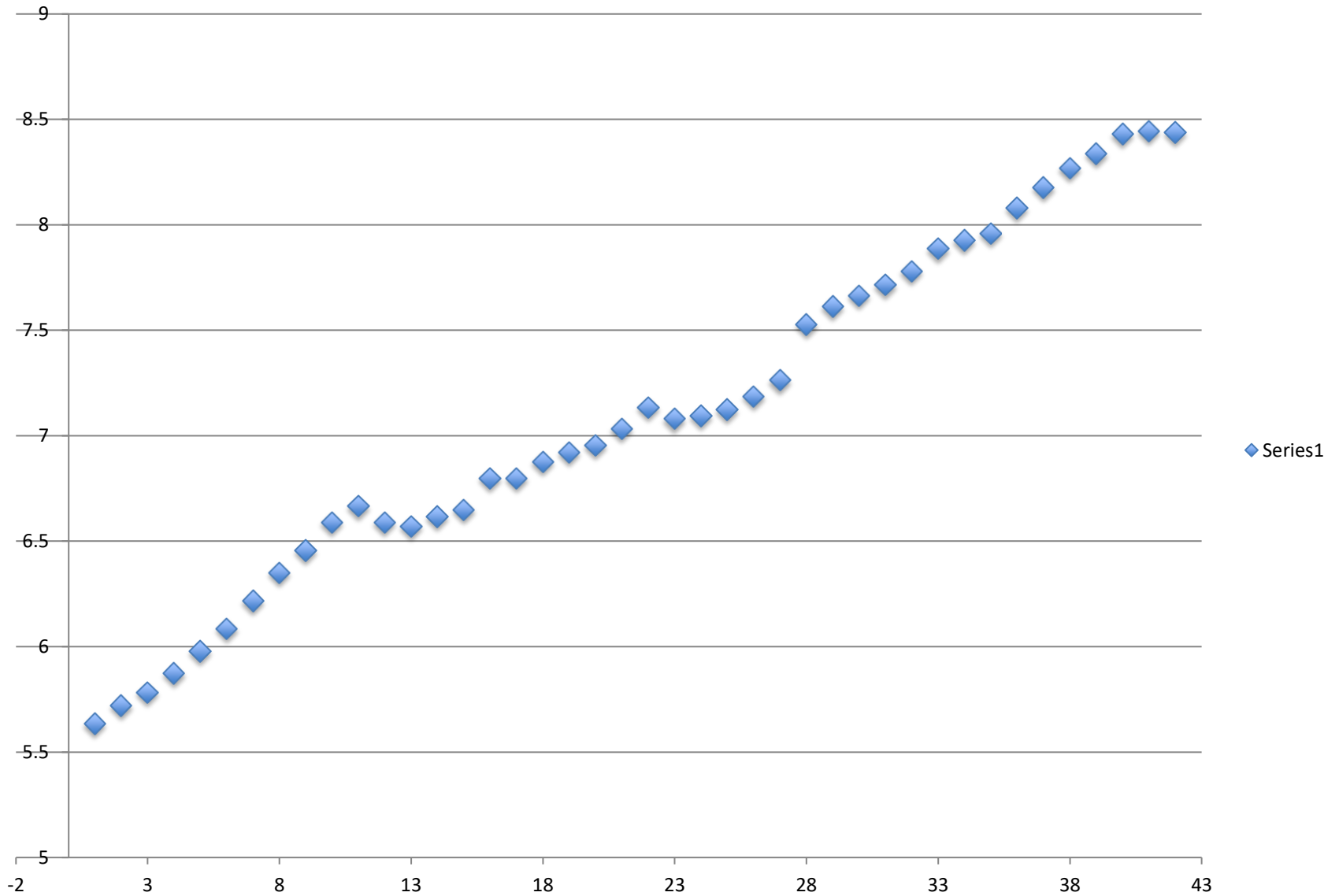
Periodo	Trimestre	Ventas
1	T1-86	280,05
2	T2-86	305,18
3	T3-86	324,14
4	T4-86	355,64
5	T1-87	394,53
6	T2-87	438,96
7	T3-87	501,13
8	T4-87	572,49
9	T1-88	635,81
10	T2-88	726,68
11	T3-88	784,94
12	T4-88	727,34
13	T1-89	713,08
14	T2-89	747,34
15	T3-89	771,44
16	T4-89	894,97
17	T1-90	894,46
18	T2-90	968,30
19	T3-90	1.012,44
20	T4-90	1.046,07
21	T1-91	1.132,78

Periodo	Trimestre	Ventas
22	T2-91	1.252,69
23	T3-91	1.187,70
24	T4-91	1.205,45
25	T1-92	1.241,29
26	T2-92	1.319,69
27	T3-92	1.426,17
28	T4-92	1.856,83
29	T1-93	2.023,56
30	T2-93	2.129,56
31	T3-93	2.239,85
32	T4-93	2.389,03
33	T1-94	2.660,00
34	T2-94	2.770,00
35	T3-94	2.863,00
36	T4-94	3.228,00
37	T1-95	3.557,00
38	T2-95	3.894,00
39	T3-95	4.171,00
40	T4-95	4.580,00
41	T1-96	4.644,00
42	T2-96	4.621,00

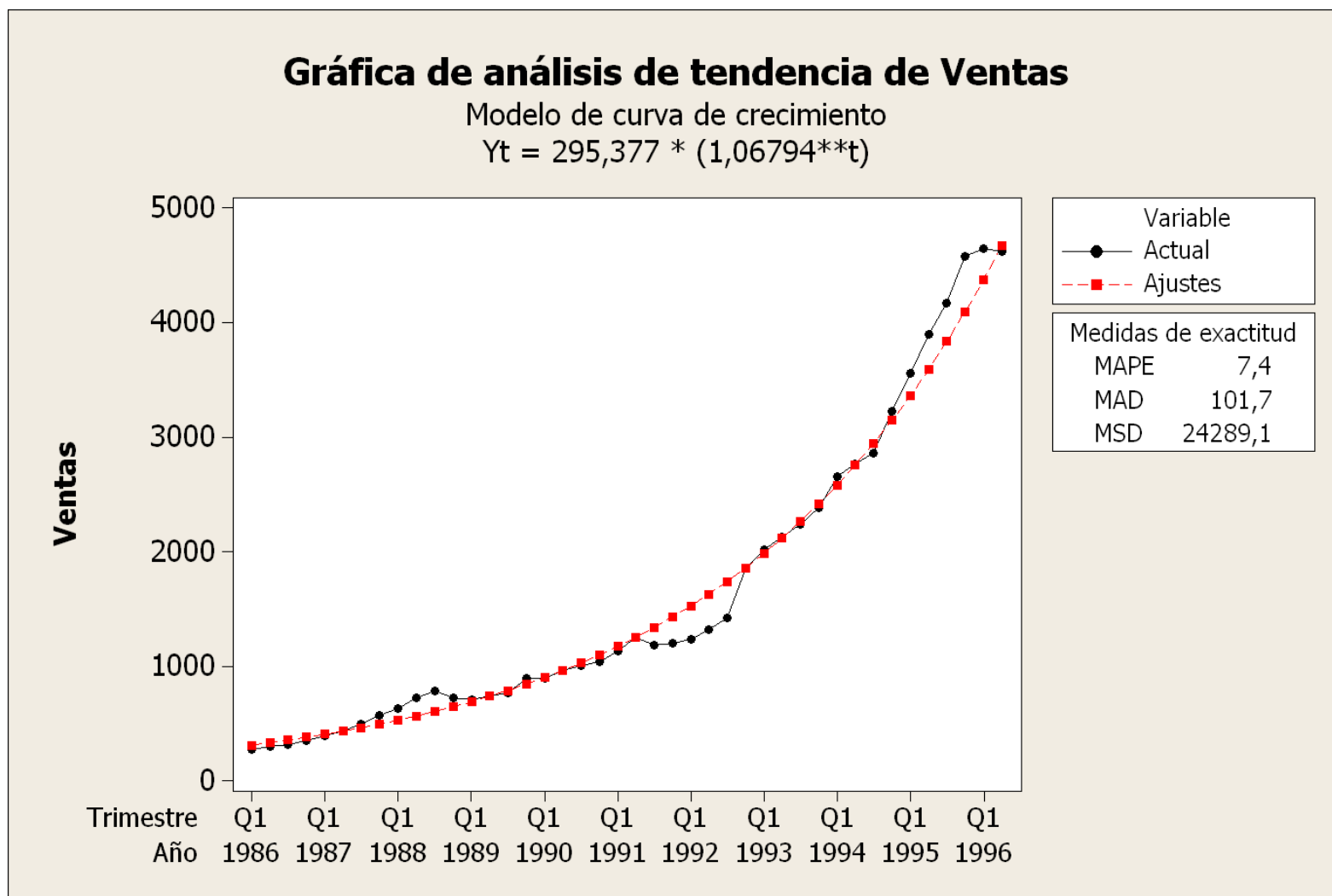


Solución

- El gráfico muestra que las ventas se incrementan de forma exponencial.
- Así, una curva suavizada adecuada a la serie debería mostrar una tendencia exponencial.
- Una forma de corroborar que nuestra intuición es correcta, es tratar de "arreglar" los datos, aplicándole el logaritmo natural a las ventas y graficar esta nueva serie.
- Si esta nueva serie de tiempo muestra una tendencia lineal, entonces existe una tendencia exponencial.



- Obtenemos modelo exponencial:



Análisis de regresión: $\ln(\text{Ventas})$ vs. Periodo

La ecuación de regresión es

$$\ln(\text{Ventas}) = 5,69 + 0,0657 \text{ Periodo}$$

Predictor	Coef	Coef. de EE	T	P
Constante	5,68825	0,03330	170,80	0,000
Periodo	0,065734	0,001349	48,72	0,000

$S = 0,105993$ $R\text{-cuad.} = 98,3\%$ $R\text{-cuad. (ajustado)} = 98,3\%$

Análisis de varianza

Fuente	GL	SC	MC	F	P
Regresión	1	26,662	26,662	2373,28	0,000
Error residual	40	0,449	0,011		
Total	41	27,112			

Solución

- De los resultados anteriores obtenemos que $a' = \ln(a) = 5,68825$, y por consiguiente $a = e^{5,68825} = 295,3763$. Además notamos que $b = 0,065734$.
- Así, nuestro modelo para predecir las ventas trimestrales de Intel es: $T_t = 295,3763e^{0,065734t}$.
- Pero como $e^{0,065734} = 1,06794$, entonces podemos reescribir el modelo como $T_t = 295,3763(1,06794)^t$.

Interpretación de los coeficientes

- Los coeficientes de la regresión se podrían interpretar como sigue:
- Las ventas de Intel en un comienzo (primer trimestre de 1986) fueron de aproximadamente US\$ 295,38 millones.
- Para cada trimestre posterior éstas mostraron una tendencia creciente que las aumentaba en aproximadamente un 6,8% ($e^b - 1$)
-

Determinación de ciclo

- Los ciclos representan patrones en los datos de largo plazo, marcados generalmente por ciclo económicos o comerciales
- Recordemos que $MA = T \times C$
- Por lo cual, la componente cíclica (índice de ciclo), puede ser calculado como: $C = (T \times C) / T = MA / T$
- Así, los índice de ciclo se calculan dividiendo los promedios móviles dobles (calculados para la estacionalidad) por la tendencia del periodo respectivo.

Ejemplo 16 (continuación ejemplo 12)

- Utilice los datos de la venta de papel en Francia entre 1977 y 1981 para calcular índice de ciclo y luego realice los pronósticos utilizando el método de descomposición.

Año	Trimestre	Ventas de Papel
1977	1	3.017,60
	2	3.043,54
	3	2.094,35
	4	2.809,84
1978	1	3.274,80
	2	3.163,28
	3	2.114,31
	4	3.024,57
1979	1	3.327,48
	2	3.493,48
	3	2.439,93
	4	3.490,79

Año	Trimestre	Ventas de Papel
1980	1	3.685,08
	2	3.661,32
	3	2.378,43
	4	3.459,55
1981	1	3.849,63
	2	3.701,18
	3	2.642,38
	4	3.585,52

Solución

- Como ya se calcularon los promedios móviles dobles, los índices estacionales y la tendencia, podemos calcular los índices de ciclo.
- El cálculo que hacemos consiste simplemente en calcular el cociente entre los promedios móviles dobles y la tendencia, en los periodos donde existan cálculos de los promedios móviles dobles.
- En la siguiente diapositiva se muestra la tabla con los cálculos.

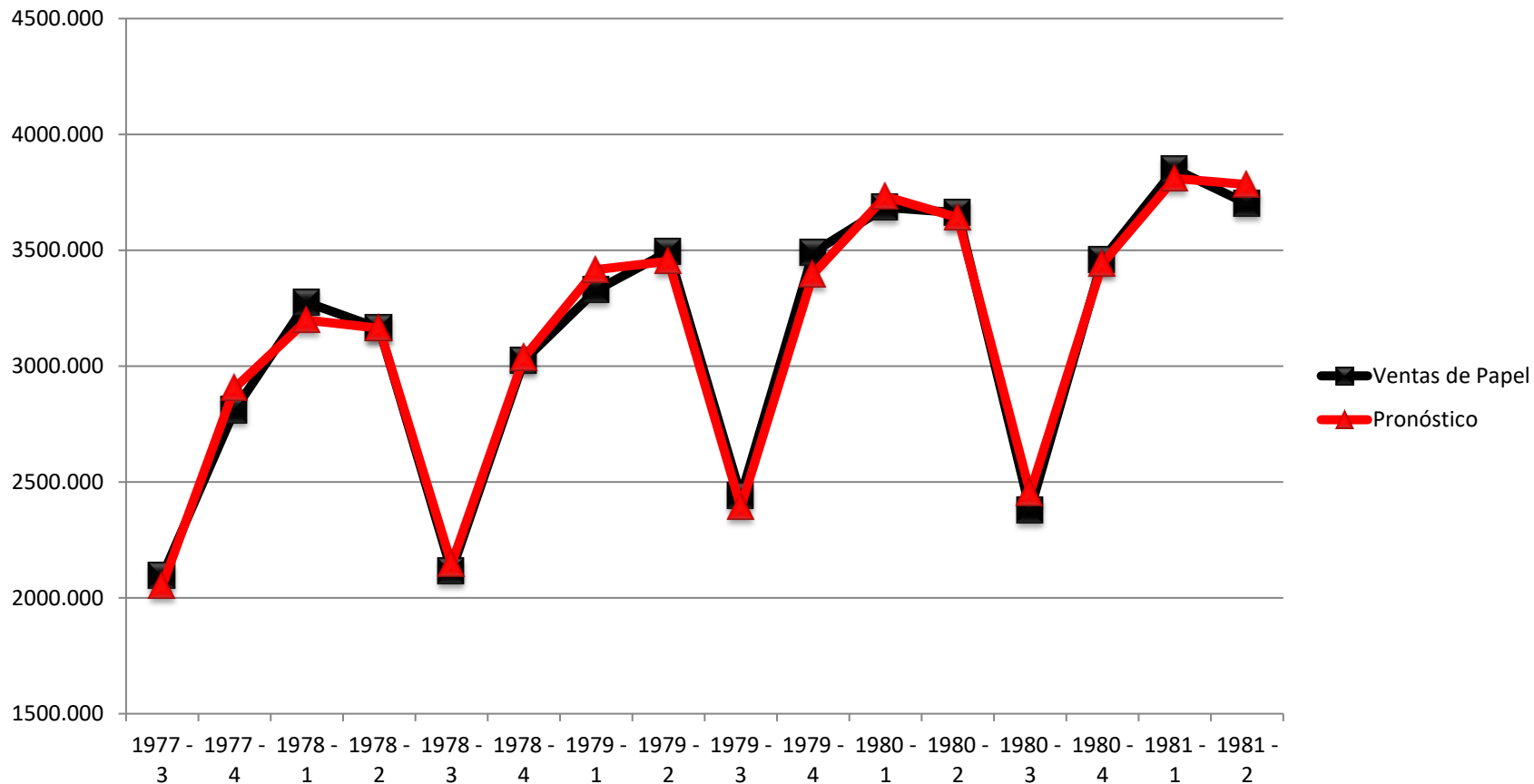
Año-Trimestre	Periodo	Ventas de Papel	Promedio Móvil	Móvil Doble	Estacionalidad	Tendencia	Ciclo	Pronóstico
1977 - 1	1	3017,60			112,71%	2754,80		
1977 - 2	2	3043,54			110,32%	2792,49		
1977 - 3	3	2094,35	2741,33	2773,48	73,98%	2830,18	98,00%	2051,70
1977 - 4	4	2809,84	2805,63	2820,60	102,99%	2867,87	98,35%	2905,03
1978 - 1	5	3274,80	2835,57	2838,06	112,71%	2905,56	97,68%	3198,67
1978 - 2	6	3163,28	2840,56	2867,40	110,32%	2943,25	97,42%	3163,45
1978 - 3	7	2114,31	2894,24	2900,83	73,98%	2980,94	97,31%	2145,90
1978 - 4	8	3024,57	2907,41	2948,69	102,99%	3018,63	97,68%	3036,95
1979 - 1	9	3327,48	2989,96	3030,66	112,71%	3056,32	99,16%	3415,74
1979 - 2	10	3493,48	3071,37	3129,64	110,32%	3094,01	101,15%	3452,77
1979 - 3	11	2439,93	3187,92	3232,62	73,98%	3131,70	103,22%	2391,35
1979 - 4	12	3490,79	3277,32	3298,30	102,99%	3169,39	104,07%	3397,03
1980 - 1	13	3685,08	3319,28	3311,59	112,71%	3207,08	103,26%	3732,37
1980 - 2	14	3661,32	3303,91	3300,00	110,32%	3244,77	101,70%	3640,72
1980 - 3	15	2378,43	3296,10	3316,66	73,98%	3282,46	101,04%	2453,52
1980 - 4	16	3459,55	3337,23	3342,22	102,99%	3320,15	100,66%	3442,26
1981 - 1	17	3849,63	3347,20	3380,19	112,71%	3357,83	100,67%	3809,68
1981 - 2	18	3701,18	3413,19	3428,93	110,32%	3395,52	100,98%	3782,96
1981 - 3	19	2642,38	3444,68		73,98%	3433,21		
1981 - 4	20	3585,52			102,99%	3470,90		

Resultados

- Con los índices de ciclo ya calculados, estamos en condiciones de calcular pronósticos “más precisos”.
- Los pronósticos los calculamos según el modelo multiplicativo: $F_t = T_t \times S_t \times C_t = (a + b \times t) \times S_t \times C_t$
- En la siguiente diapositiva se muestran los pronósticos junto con las ventas de papel observadas.
- $MSE = 3.456,75$, $MAD = 50,71$ y $MAPE = 1,68\%$.

Comparación gráfica

Ajuste del Modelo



Descomposición en Minitab

- Permite modelos lineales y aditivos
- El proceso de obtención de los coeficientes es el siguiente:
 1. Se elimina la tendencia de los datos mediante la aplicación de medias móviles centradas (tal como se explicó en el apartado de media móvil)
 2. Se resta la serie sin tendencia de la serie original para obtener los coeficientes de estacionalidad de cada observación.

Descomposición en Minitab

3. Partiendo de los valores obtenidos en el paso 2 y para cada periodo, se determina la mediana que conforman los índices de estacionalidad.
4. Se ajustan los índices de estacionalidad del paso 3 para rectificarnos (multiplicativos igual a 1, aditivos igual a 0)
5. Usando estos índices se desestacionaliza la serie (se divide/resta la serie por el coeficiente que pertoque).
6. La tendencia se obtiene mediante regresión lineal de la serie del punto 5.



Tema 4. Series de Tiempo

Modelos ARIMA

Introducción

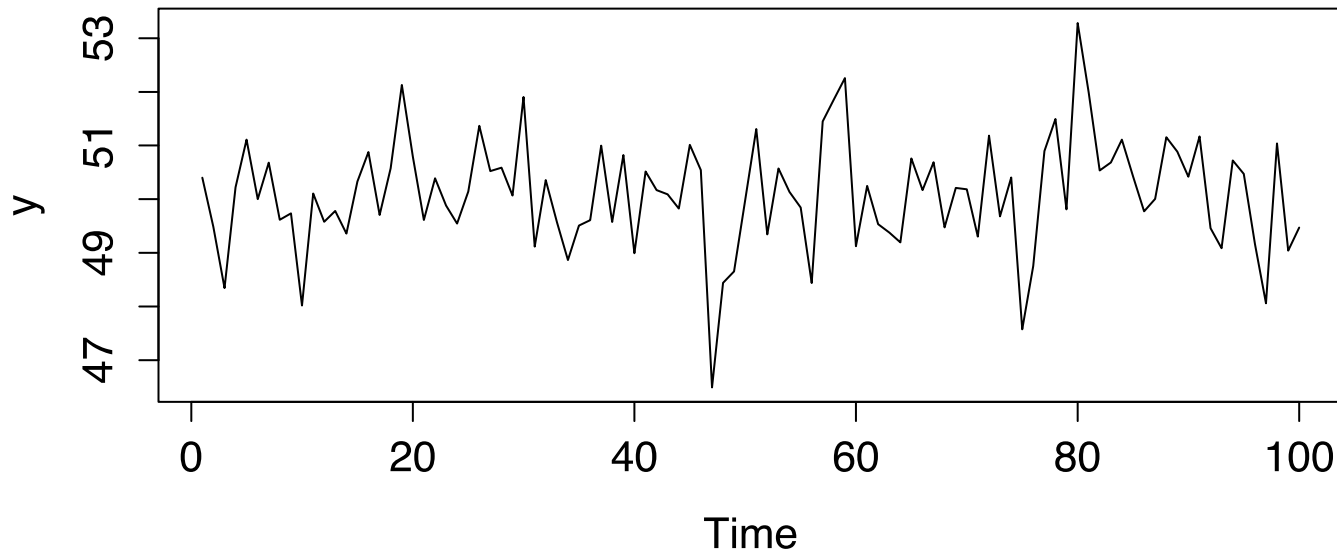
- Los modelos ARIMA son un método estadístico que pretende estudiar patrones en la serie para predecirla hacia el futuro.
- En ocasiones se les conoce como el método de Box-Jenkins.
- La notación clásica es: $\text{ARIMA}(p,d,q)$ donde:
 - **p**: valor que controla el número de componentes autoregresivas (AR) del modelo.
 - **d**: valor que controla el número de integraciones (I) del modelo.
 - **q**: valor que controla el número de componentes de media móvil (MA) del modelo

Introducción

- Son modelos orientados a estudiar procesos estacionarios.
- Se entiende por serie estacionaria una serie cuya función de probabilidad no varia en el tiempo (entre otras cosas, que la media y la varianza no cambian).
- El componente tendencia se elimina a través de integraciones (concepto explicado posteriormente)
- El componente estacionalidad requiere un modelo ARIMA con estacionalidad (SARIMA)

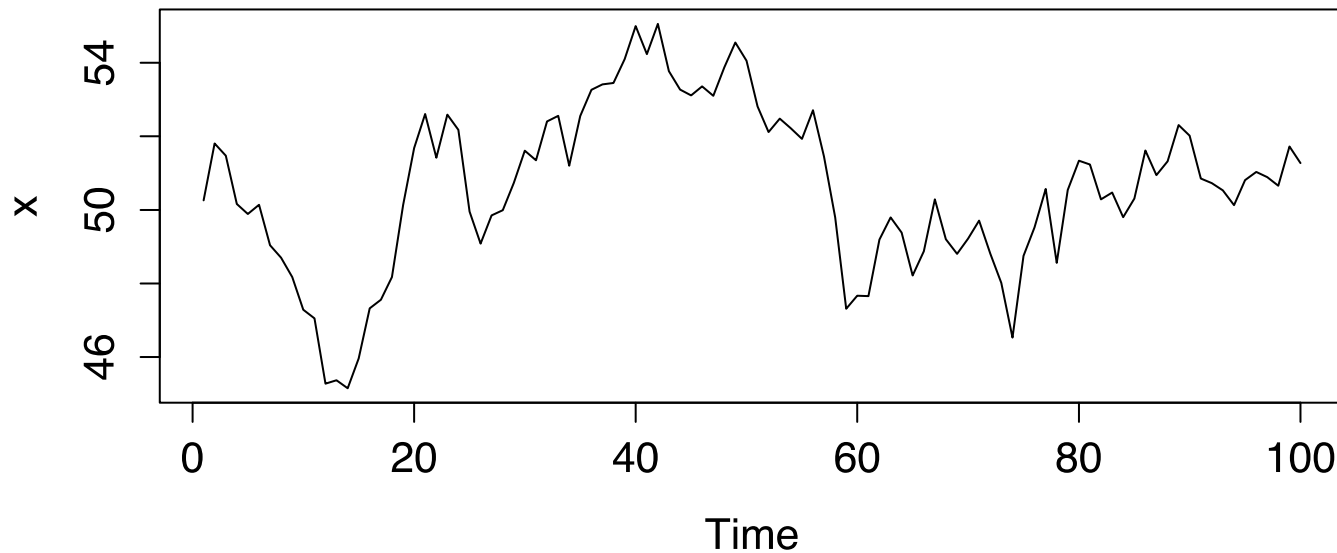
Ejemplos de series

- El objetivo de todo método de series temporales es dejar un residuo sin tendencia ni autocorrelación.



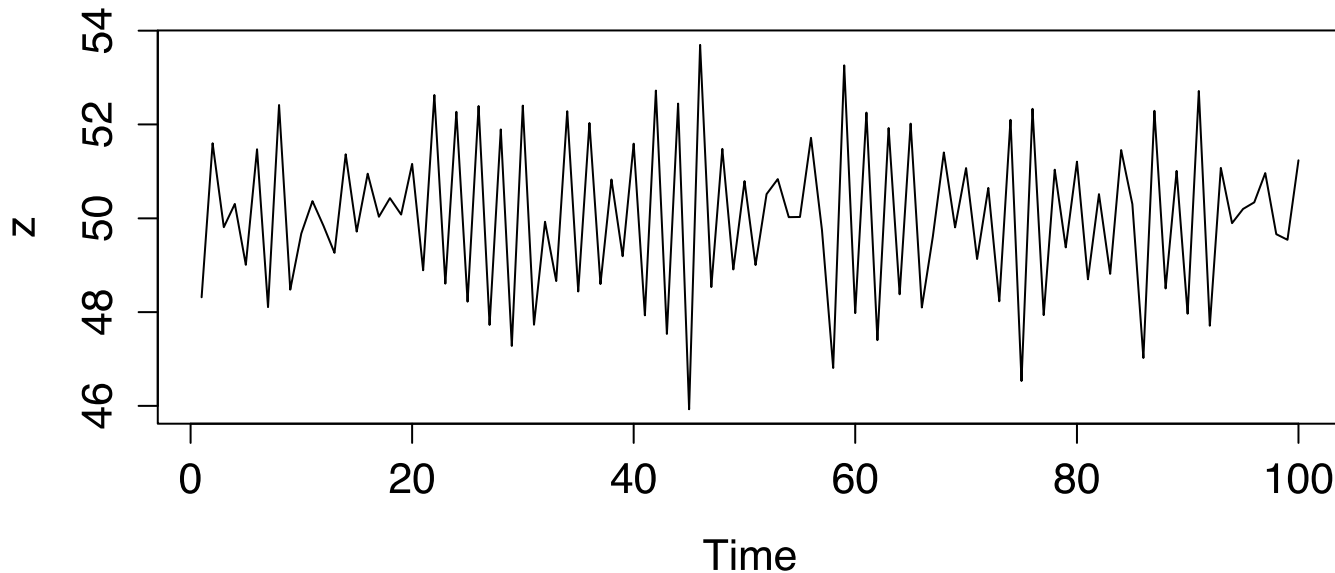
Ejemplos de series

- Autocorrelación positiva: el ejemplo muestra una tendencia a que observaciones con crecimiento sean seguidos por observaciones con crecimiento



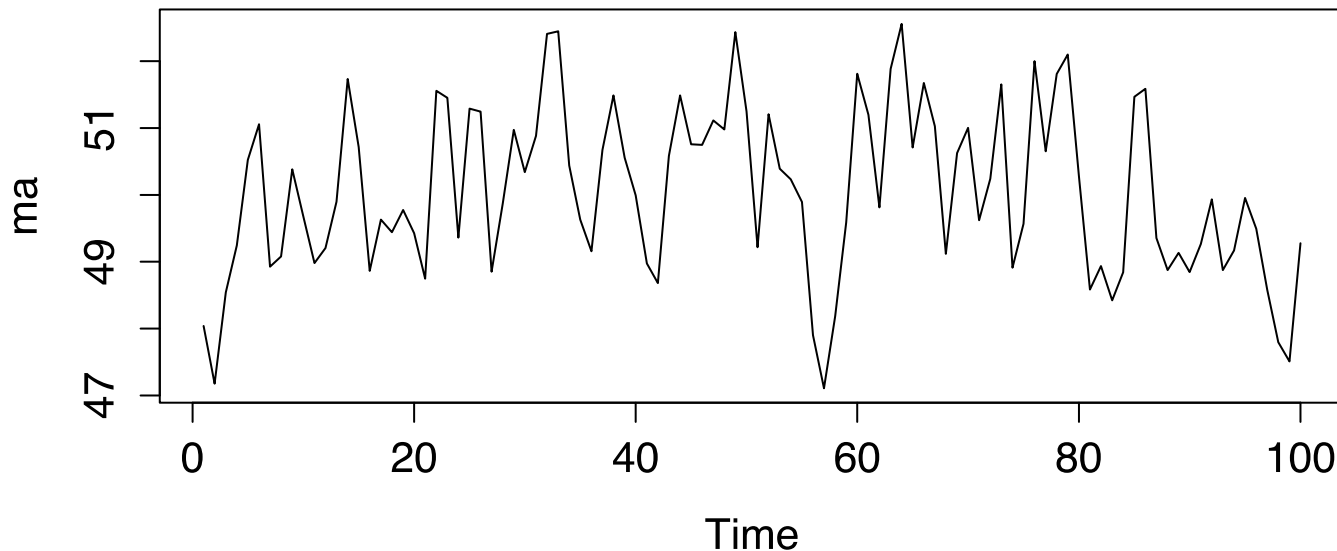
Ejemplos de series

- Autocorrelación negativa: el ejemplo muestra una tendencia a que observaciones grandes se alternen con observaciones pequeñas



Ejemplo de series

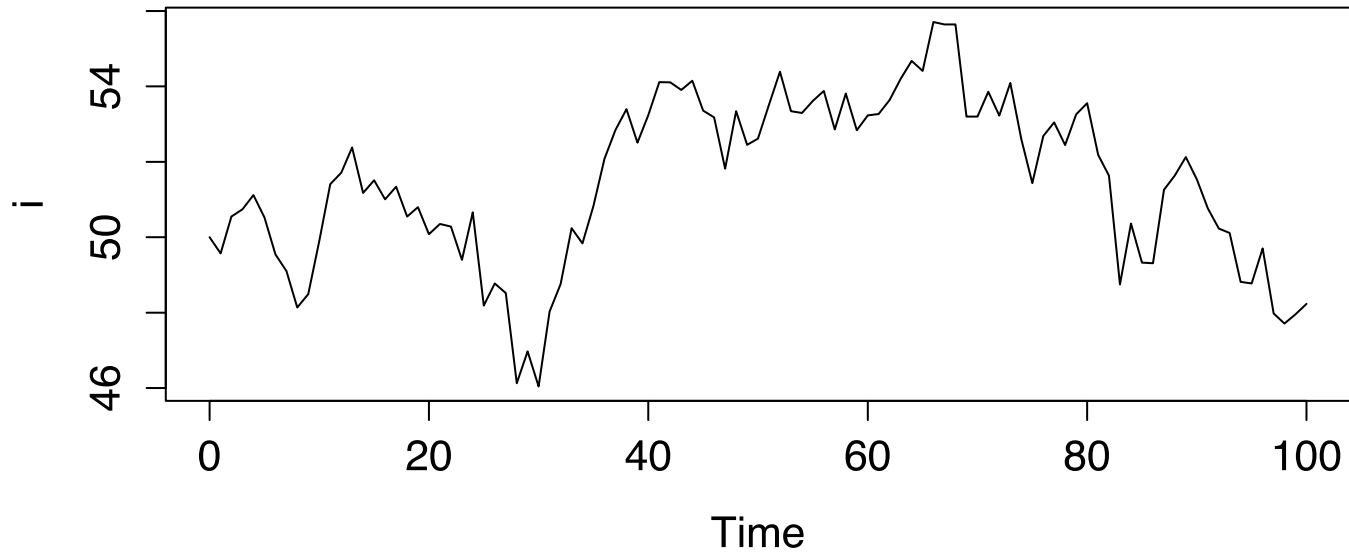
- Los casos anteriores son extremos. ¿Qué sucede con series como ésta?



- Respuesta: Sigue un comportamiento MA(1)

Ejemplo de series

- ¿O ésta?



- Respuesta: responde a una serie $I(1)$

Métodos ARIMA

- En la presente descripción no entraremos en detalles sobre el método de cálculo y nos basaremos en las herramientas que brinda MINITAB para su resolución.
- Realizaremos algunas simplificaciones aceptadas por el método para las cuales se conocen métodos más eficientes (que no consideraremos en la asignatura)

Técnicas de trabajo

- Se utilizarán únicamente:
 - Gráficos de la serie
 - Gráficos de residuo de la serie
 - Gráfico de autocorrelación
 - Gráfico de autocorrelación parcial.
- El esquema del procedimiento es:
 - Identificación y selección del modelo
 - Estimación de parámetros
 - Verificación del modelo.

Paso 1. Identificación del modelo

- Utilizando el método de descomposición propuesto con anterioridad, eliminamos la estacionalidad y (en caso que así se desee) la tendencia.
- En caso que la tendencia no se elimine mediante la descomposición, si existe tendencia debería eliminarse diferenciando la serie (determinación del coeficiente del parámetro I). Es incluso probable que aunque se elimine la tendencia sea conveniente determinar si debe diferenciarse la serie.

Determinar I(d)

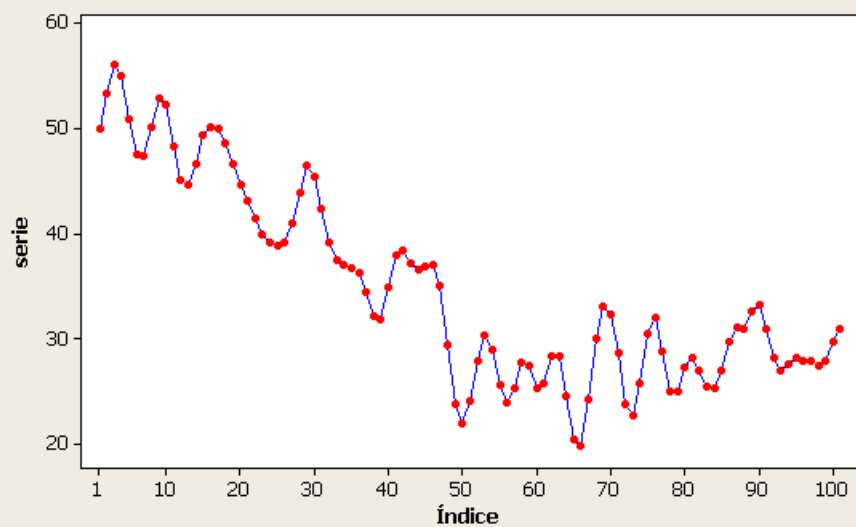
- Utilizarse la opción “diferencias” de minitab y representar el gráfico resultante.

$$I(1) X'_t = X_t - X_{t-1}$$

$$I(2) X''_t = X'_t - X'_{t-1}$$

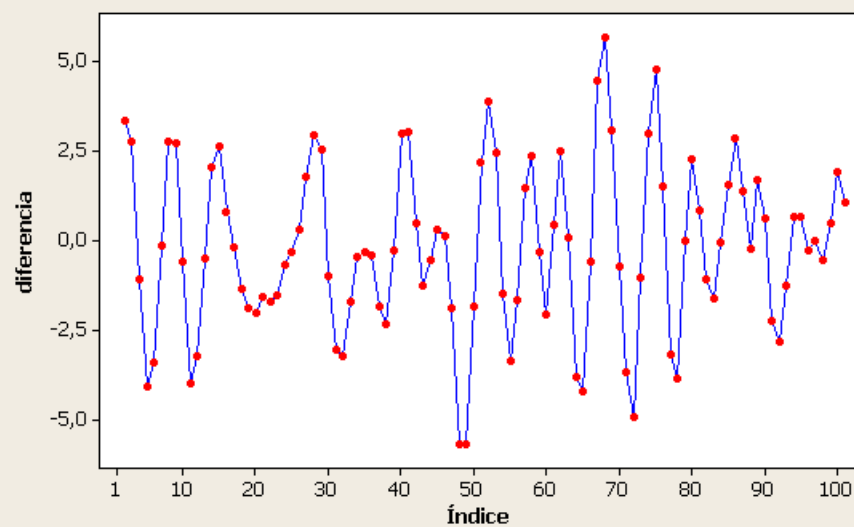
- Raramente se necesita más que una diferencia para estabilizar la serie

Gráfica de series de tiempo de serie



Serie Original

Gráfica de series de tiempo de diferencia



Serie diferenciada

Determinar parámetros AR(p) y MA(q)

- Se recomienda usar el gráfico de autocorrelación y el gráfico de autocorrelación parcial.
- Dependiendo de la combinación de parámetros, nos encontraremos con diferentes formas de ambos gráficos.
- En este caso sólo se pueden dar indicaciones y se recomienda verificar la validez de las decisiones adoptadas una vez se ha resuelto el modelo

Gráfico de Autocorrelación

- Muestra el nivel de autocorrelación entre la serie y la misma serie desfasada n posiciones.
- Los rangos significativos aparecen en rojo (por encima significativo)

Normalmente
denominado
ACF

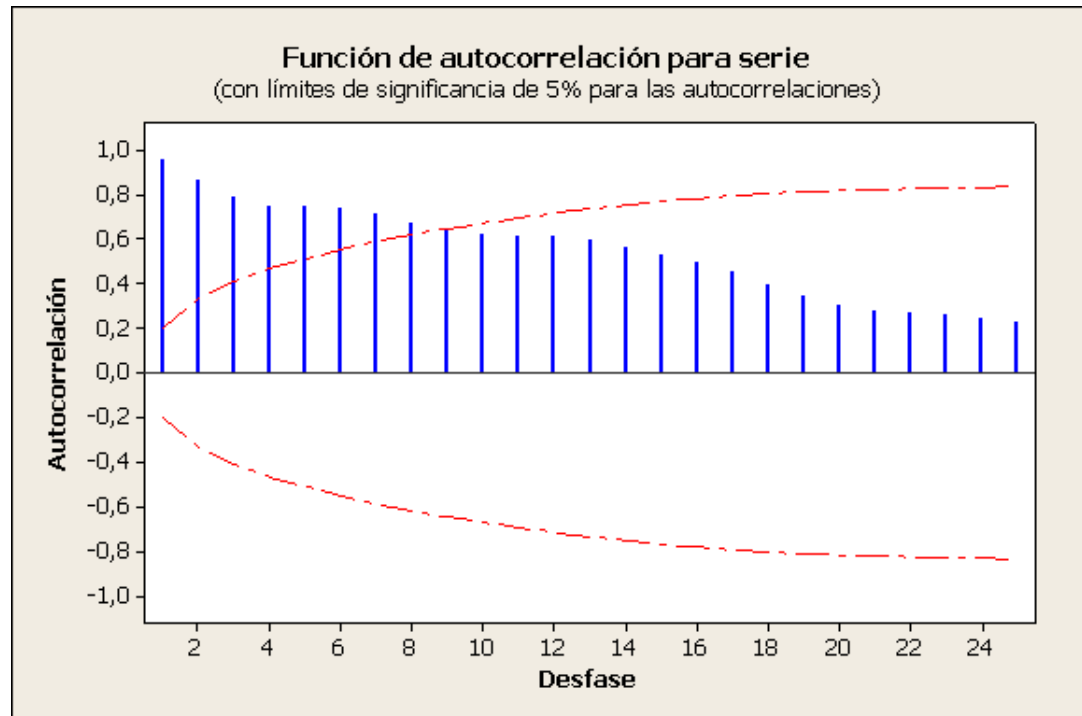
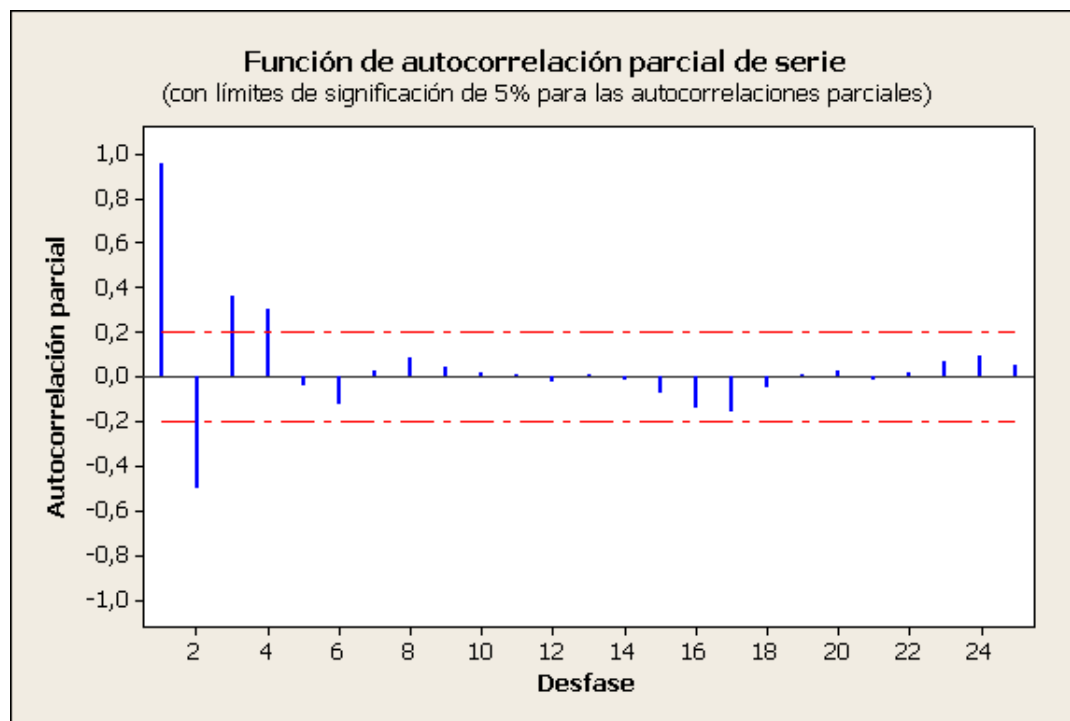


Gráfico de Autocorrelación Parcial

- Muestra la autocorrelación de la serie tras eliminar las autocorrelaciones con las series desfasadas hasta $n-1$ periodos, con la serie desfasada n periodos.

Normalmente
denominado
PACF

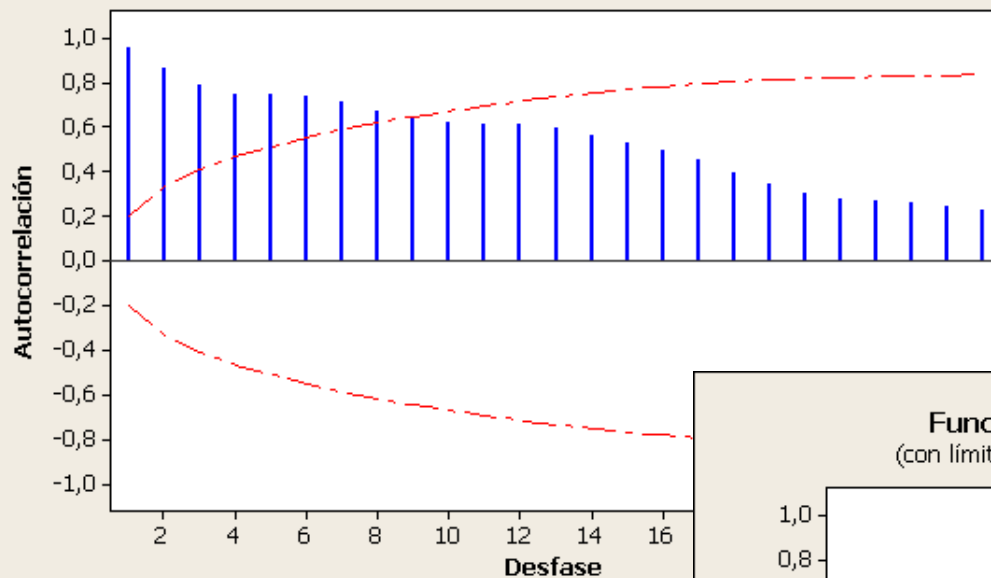


Formas habituales

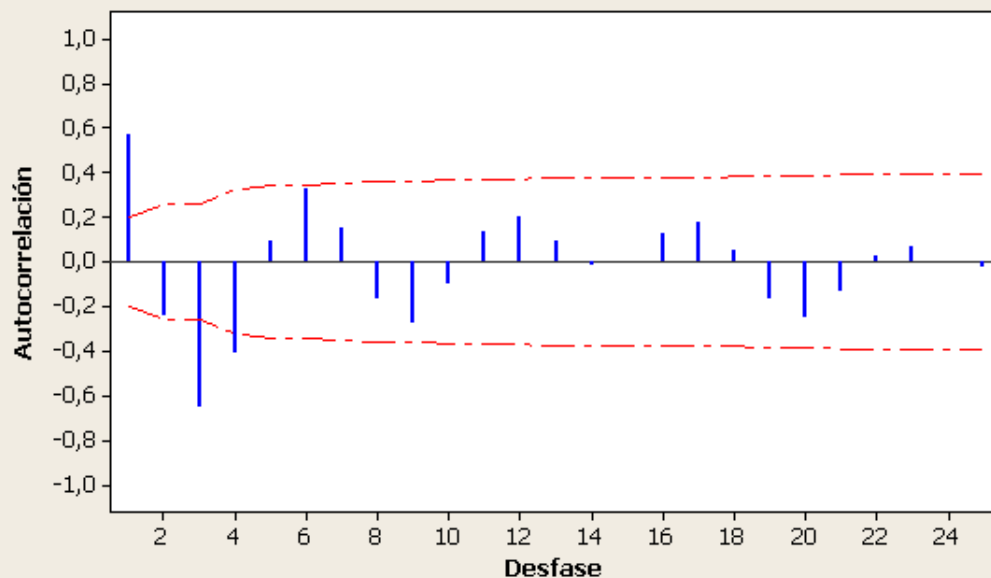
- Si la función decrece a 0 exponencialmente: AR (usar PACF para determinar el valor de p que es el valor en que la función se corta bruscamente)
- Alternancia entre valores positivos y negativos que decrece a 0 exponencialmente: AR con autoregresión negativa
- Uno o más picos, el resto insignificantes. Modelo MA con q igual al número de picos.
- La función decae tras unos cuantos picos. Corresponde a un modelo ARMA.
- Cercanos a 0: No hay componentes ARMA
- Valores altos en periodos fijos: Efecto de estacionalidad
- No decae a 0: Serie no estacionaria

ACF (sin y tras diferenciar)

Función de autocorrelación para serie
(con límites de significancia de 5% para las autocorrelaciones)



Función de autocorrelación para diferencia
(con límites de significancia de 5% para las autocorrelaciones)



Paso 2. Obtención de parámetros

- Una vez se ha escogido un modelo, éste debe ser evaluado.
- Nota: En caso de error en la elección del modelo, el paso 3 debería identificar el error.
- Escoger un modelo corresponde a determinar un valor para los parámetros (p,d,q) que configuran la fórmula a usar en la predicción.

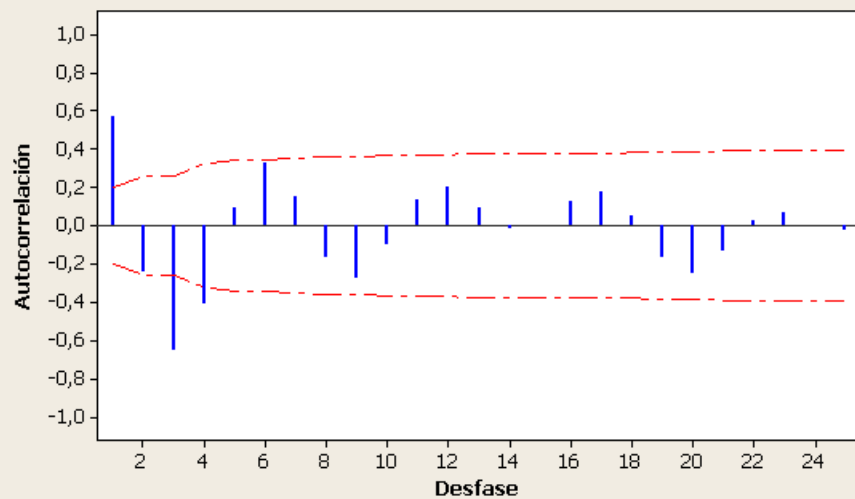
Significado de los parámetros

- Una forma de entender el efecto de los tres parámetros es el siguiente:
 - p : indica el número de periodos cuyos residuos (desviaciones respecto a lo normal) siguen influyendo en la determinación del valor actual.
 - d : indica el número de diferenciaciones necesarias para obtener una serie estacionaria.
 - q : indica el número de periodos anteriores cuyas valores promedio siguen influyendo en la determinación del valor actual.

Selección de modelo

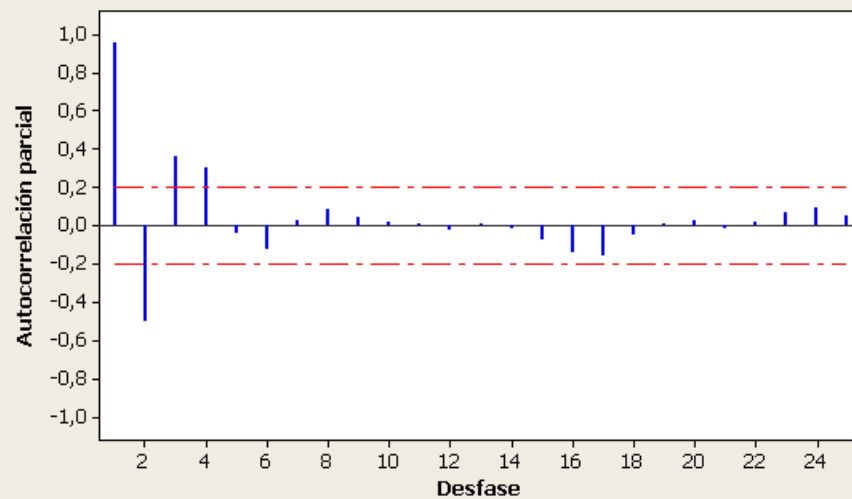
- Ante la duda, escoger el modelo más simple.
- En este caso, optaría por un modelo (1,1,1) aunque el ACF pudiera entenderse que no correspondería a un caso AR(1)

Función de autocorrelación para diferencia
(con límites de significancia de 5% para las autocorrelaciones)



ACF

Función de autocorrelación parcial de serie
(con límites de significación de 5% para las autocorrelaciones parciales)



PACF

Resultados

Estimados finales de los parámetros

Tipo		Coef	SE	Coef	T	P
AR	1	0,4636	0,0951	4,87	0,000	
MA	1	-0,8494	0,0555	-15,32	0,000	
Constante		-0,0778	0,2486	-0,31	0,755	

Diferenciación: 1 Diferencia regular

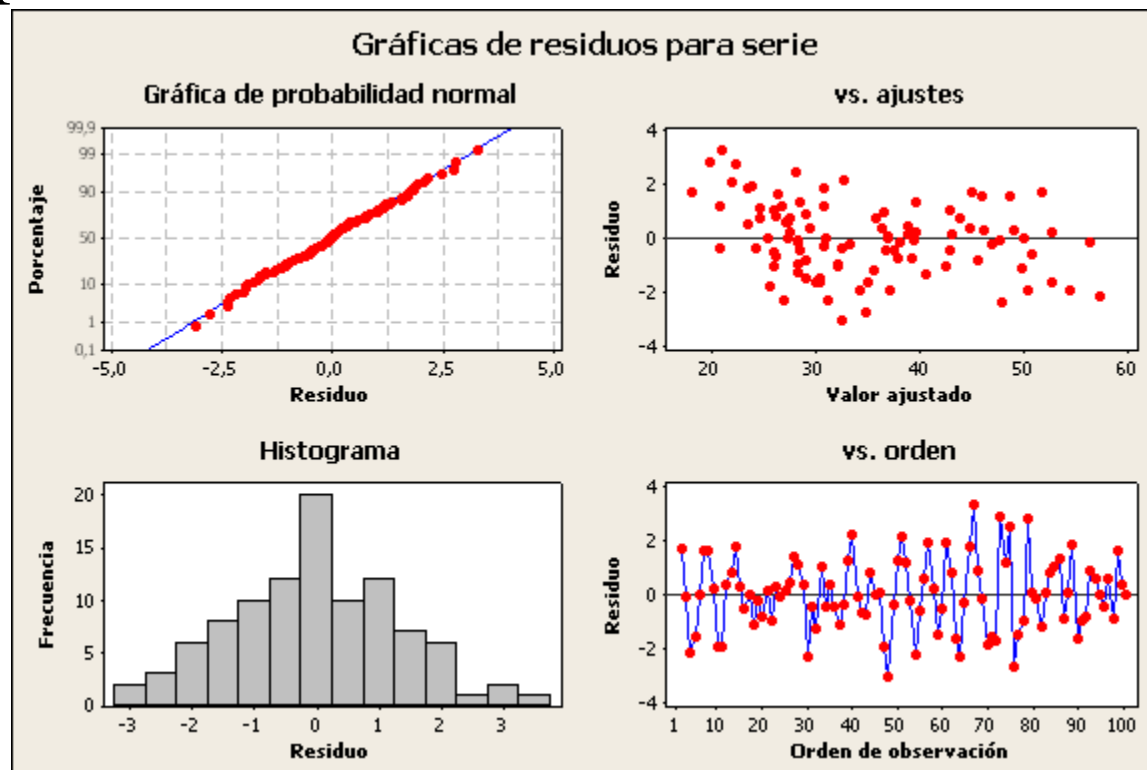
Residuos:

$$SC = 175,108$$

$$MC = 1,805 \quad GL = 97$$

Paso 3. Verificación del modelo

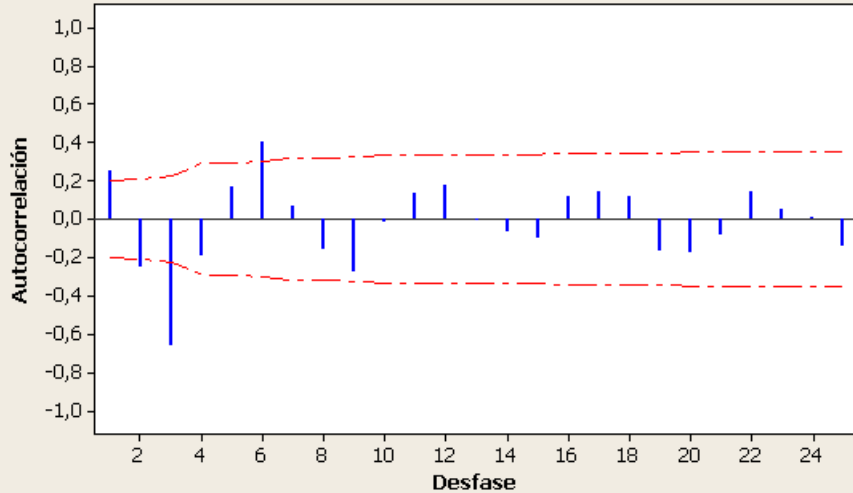
- Aparte de los p-valores, es necesario graficar los residuos y verificar si realmente no muestran patrones



ACF y PACF de los residuos

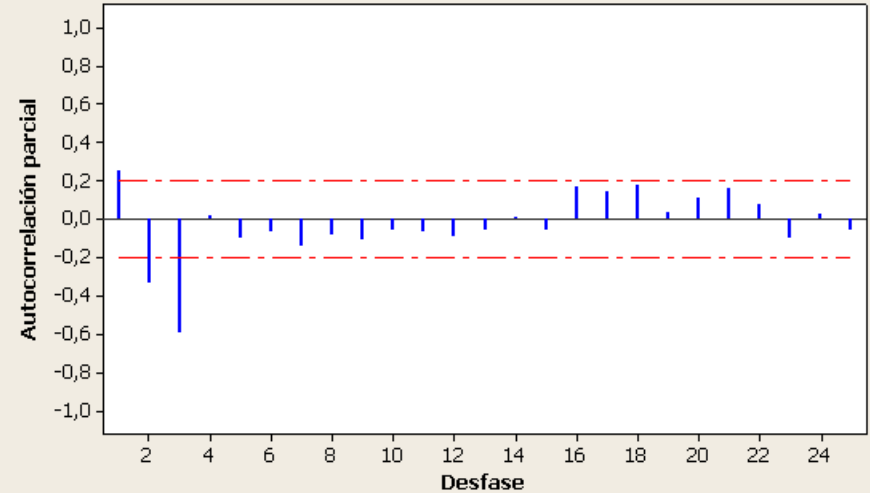
ACF de residuos para serie

(con límites de significancia de 5% para las autocorrelaciones)



PACF de residuos de serie

(con límites de significación de 5% para las autocorrelaciones parciales)



- La serie sigue mostrando patrones en los residuos (picos significativos en 2, 3 y 6 para ACF y significativos para 2 y 3 en PACF).
- Esto indicaría que faltan incluir parámetros

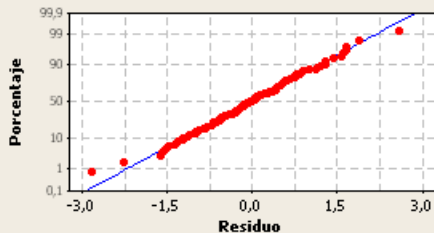
Modelos alternativos

- Se recomienda estudiar siempre el modelo más simple antes del más complejo. Por ejemplo un modelo $AR(3)$ incluye al modelo $AR(2)$ y al $AR(1)$, por tanto y aunque parezca que debiera usarse un $AR(3)$ se recomienda probar primero un $AR(2)$.
- Por tanto, en este caso, observando los gráficos se optaría por un modelo $(2,1,1)$ o un modelo $(3,1,1)$. Primero se muestra el más simple.

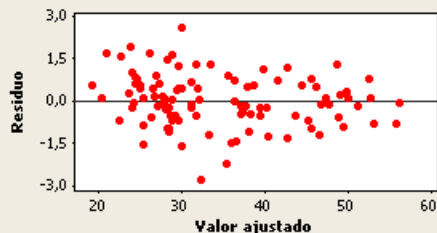
Modelo (2,1,1)

Gráficas de residuos para serie

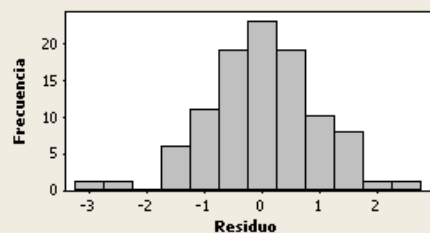
Gráfica de probabilidad normal



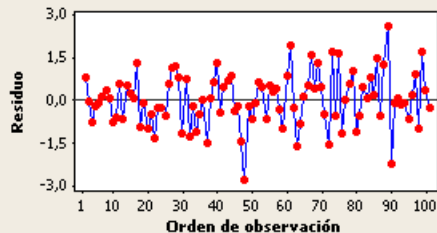
vs. ajustes



Histograma

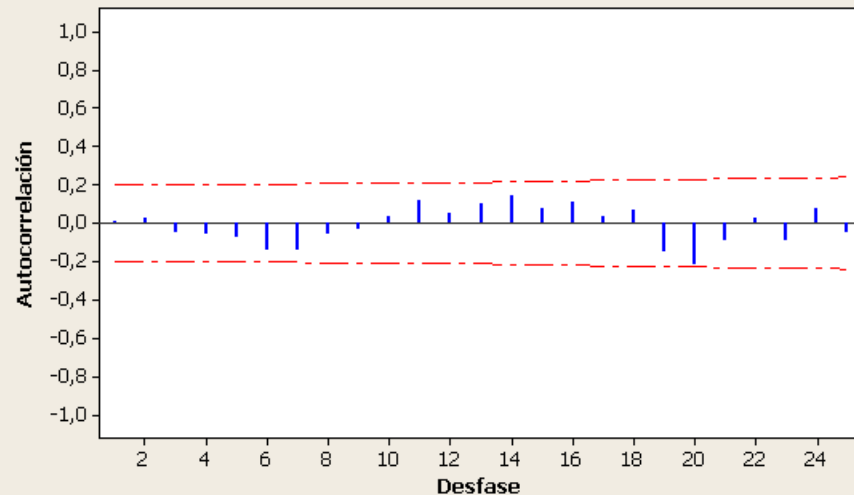


vs. orden



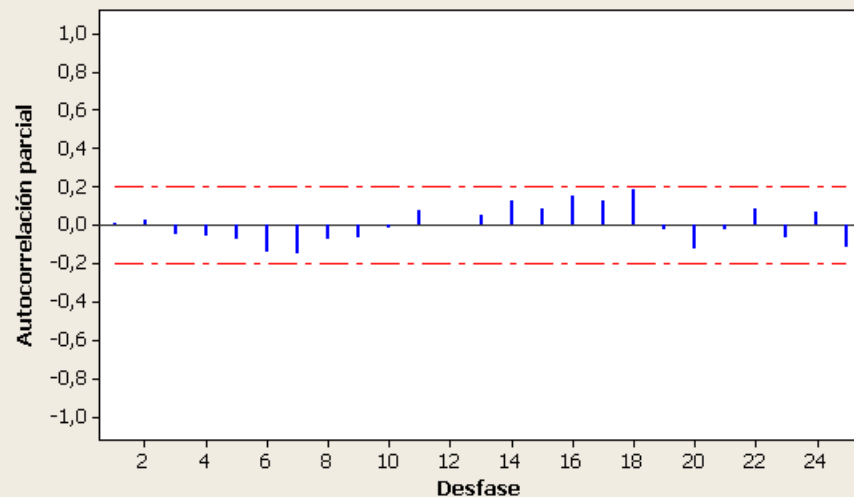
ACF de residuos para serie

(con límites de significancia de 5% para las autocorrelaciones)



PACF de residuos de serie

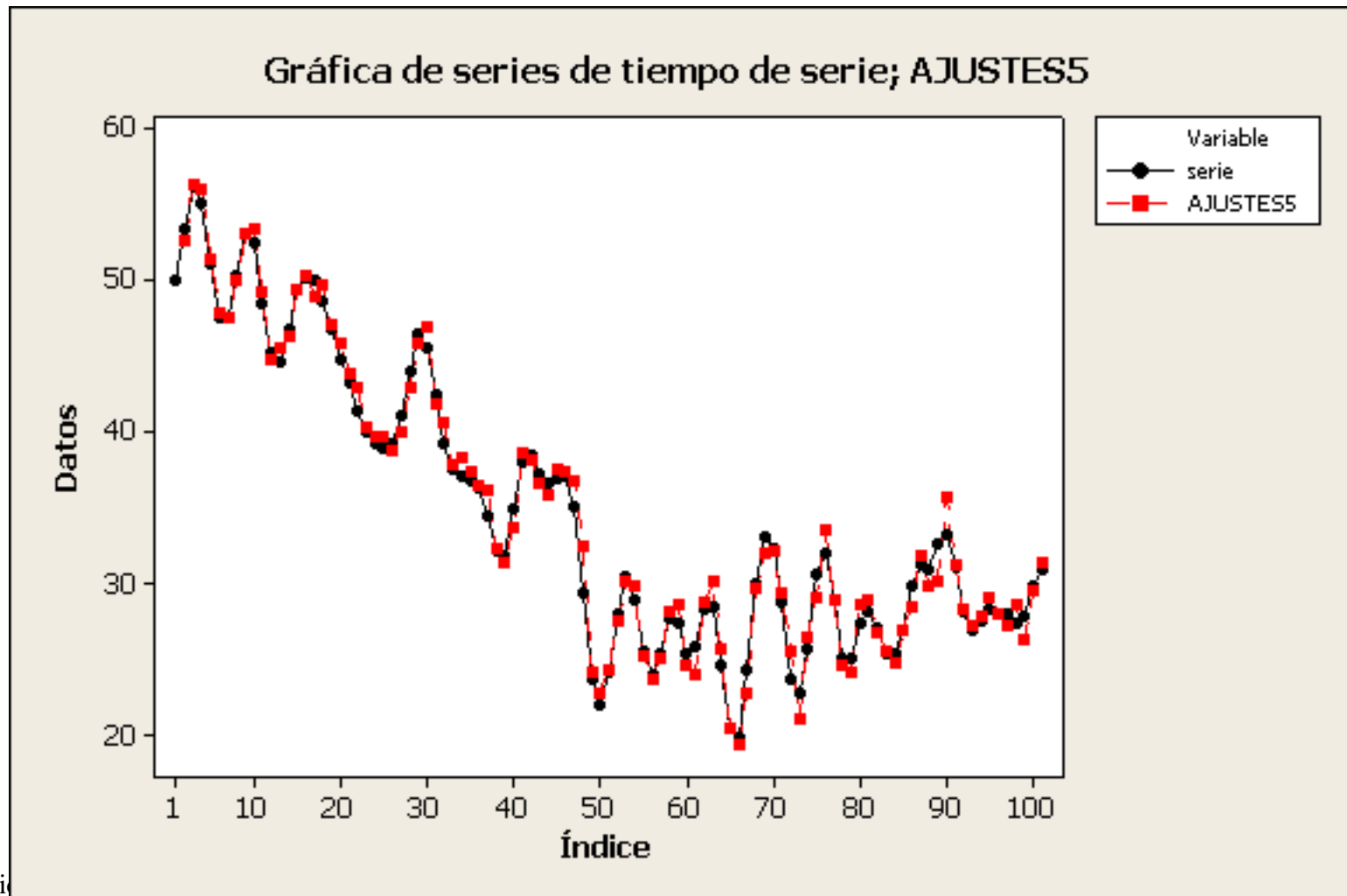
(con límites de significación de 5% para las autocorrelaciones parciales)



Tipo		Coef	SE Coef	T	P
AR	1	0,9435	0,0689	13,69	0,000
AR	2	-0,7893	0,0672	-11,74	0,000
MA	1	-0,5069	0,0959	-5,29	0,000

Gráficamente

- Para obtener la serie ajustada, debería usarse la opción de almacenamiento de los ajustes



Fórmula matemática detrás de ARIMA

- Por ahora únicamente sabemos que el modelo cumple un $ARIMA(2,1,1)$ pero qué fórmula está estudiando?
- Con el objeto de no complicar demasiado la explicación hablaremos de cada elemento de forma independiente y posteriormente de forma conjunta.

Ejemplos

- Un modelo estacionario autoregresivo puro sería equivalente a un modelo lineal con:

$$Y_t = b_0 + b_1 Y_{t-1} + \dots + b_p Y_{t-p} + e_t$$

cabe destacar que no corresponde a una regresión lineal clásica porque suponemos autocorrelación entre los predictores

- Un modelo estacionario de media móvil puro sería equivalente a un modelo lineal con:

$$Y_t = b_0 + b_1 e_{t-1} + \dots + b_q e_{t-q} + e_t$$

- Recordemos que diferenciar es: $Y'_t = Y_{t-1} - Y_{t-2}$

Ejemplos

- Un modelo ARIMA(1,0,0) es: $Y_t = c + f_1 Y_{t-1} + e_t$
- Un modelo ARIMA(0,0,1) es: $Y_t = c - q_1 Y_{t-1} + e_t$

Es importante fijarse que el coeficiente tiene por convenio valor negativo

- Un modelo ARIMA(1,0,1) es:

$$Y_t = c + f_1 Y_{t-1} + e_t - q_1 e_{t-1}$$

- Usando el concepto de backshift (B equivale a $t-1$, B^2 a $t-2$, etc...):

$$Y_t = c + f_1 B Y_t + e_t - q_1 B e_t$$

$$(1 - f_1) Y_t = c + (1 - q_1) e_t$$

Ejemplos

- Un modelo ARIMA(1,1,1):

$$(1 - f_1 B)(1 - B)Y_t = c + (1 - q_1 B)e_t$$

AR1

I1

MA1

$$(1 - f_1 B)(Y_t - Y_{t-1}) = c + (e_t - q_1 e_{t-1})$$

$$Y_t - Y_{t-1} - f_1 Y_{t-1} + f_1 Y_{t-2} = c + e_t - q_1 e_{t-1}$$

$$Y_t = c + Y_{t-1} + f_1 Y_{t-1} - f_1 Y_{t-2} + e_t - q_1 e_{t-1}$$

- Un modelo ARIMA(2,1,1):

$$(1 - f_1 B - f_2 B^2)(1 - B)Y_t = c + (1 - q_1 B - q_2 B^2)e_t$$

ARIMA en Minitab

- La fórmula general es:

$$Y_t = \underbrace{-(D^d Y_t - Y_t)}_I + \underbrace{f_0 + \sum_{i=1}^p f_i D^d Y_{t-i}}_{AR} - \underbrace{\sum_{i=1}^q q_i e_{t-i}}_{MA} + e_t$$

- Minitab intenta obtener el mejor juego de parámetros mediante la minimización de la suma de errores cuadrados.
- Para determinar previsiones de los valores iniciales se utiliza la técnica de “backcasting” comentada en el alisado exponencial simple y doble.



Tema 4. Series de Tiempo

Modelos asociativos

Métodos asociativos

- Son métodos cuantitativos
- Usan datos históricos, donde el tiempo no es una variable relevante
- Buscar asociar la demanda con otras variables
- Los modelos se eligen a través del comportamiento de las relaciones mostrada en gráficos de dispersión, o bien por experiencia del analista
- Los parámetros se estiman a través de una Regresión

Ejemplo 17

- Pharmex es una cadena de farmacias que funciona alrededor del país.
- Para determinar la eficacia de su publicidad y otras actividades promocionales, la compañía ha recogido datos de 50 regiones metropolitanas seleccionadas al azar.
- En cada región se han comparado sus gastos promocionales y ventas con los del competidor principal en la región.

Ejemplo 17

- Se dispone de dos variables:
 - Promote: Gastos promocionales de Pharmex como porcentaje de los del competidor principal
 - Sales: Ventas de Pharmex como porcentaje de las del competidor principal
- La compañía espera que haya una relación positiva entre las dos variables, de modo que las regiones con mayor gasto relativo , tenga mas ventas relativas. ¿Es cierta la afirmación?

Region	Promote	Sales
1	77	85
2	110	103
3	110	102
4	93	109
5	90	85
6	95	103
7	100	110
8	85	86
9	96	92
10	83	87
11	88	99
12	94	101
13	104	109
14	89	81
15	95	107
16	94	95
17	96	109

Region	Promote	Sales
18	92	113
19	93	84
20	98	98
21	103	112
22	95	96
23	103	93
24	89	97
25	97	92
26	97	105
27	99	102
28	101	105
29	113	118
30	86	82
31	100	101
32	96	97
33	105	101
34	101	104

Region	Promote	Sales
35	89	91
36	90	99
37	102	112
38	106	104
39	116	111
40	117	119
41	100	94
42	96	81
43	88	92
44	109	108
45	109	103
46	116	113
47	98	104
48	100	98
49	95	108
50	96	87

Sales

