4.1 Series temporales.

Definición: Una serie temporal es una variable bivariante donde una de las variables es la **variable de estudio** y la otra variable es el **tiempo**.

$$(t, X) \rightarrow X_t$$

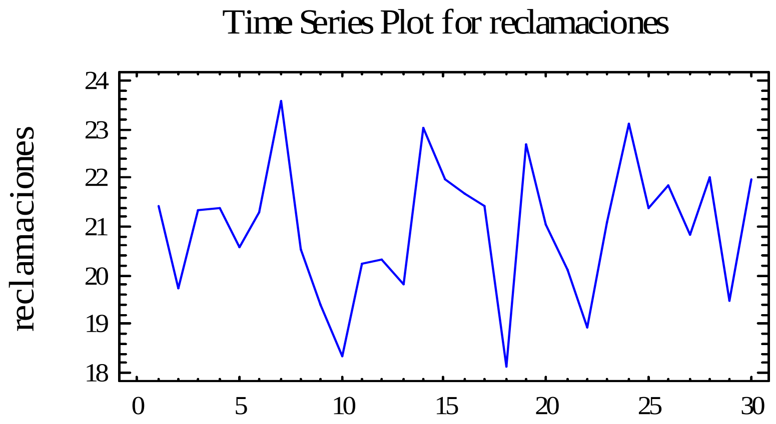
Objetivo: **Explicar la evolución** de la variable a lo largo del tiempo y **prever** sus valores futuros.

Observación: Los datos están indexados con respecto del tiempo y, por tanto, **el orden de llegada** de los datos es importante.

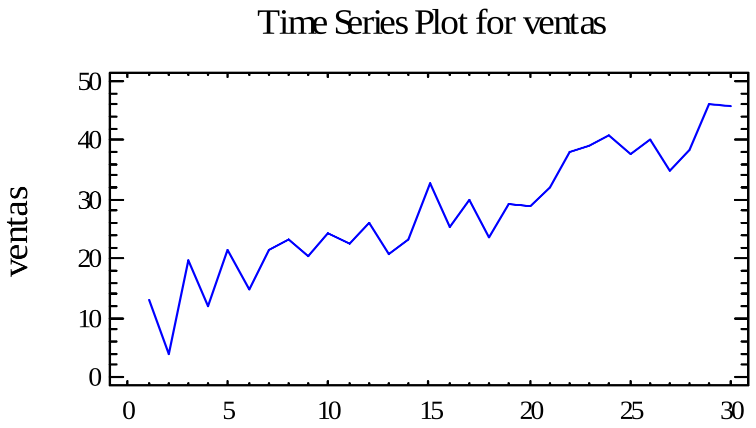
Representación gráfica: Mediante un gráfico temporal. Se trata de un diagrama de dispersión de la variable bivariante ***(t,X)*** con los puntos unidos mediante segmentos.

Cuatro ejemplos de series temporales

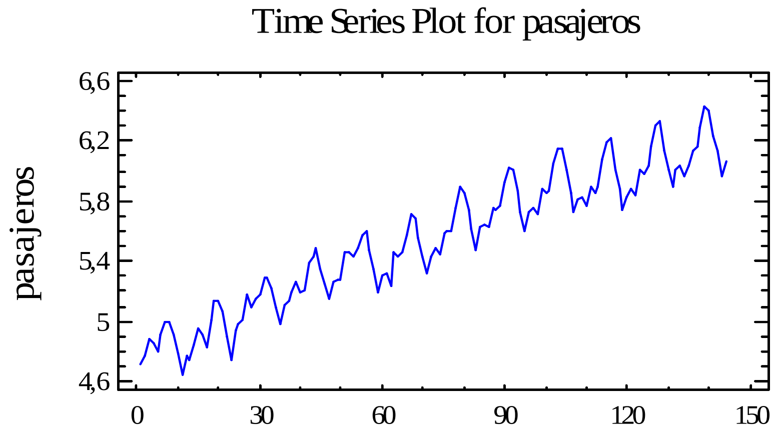
Reclamaciones semanales en un centro de asistencia al cliente.



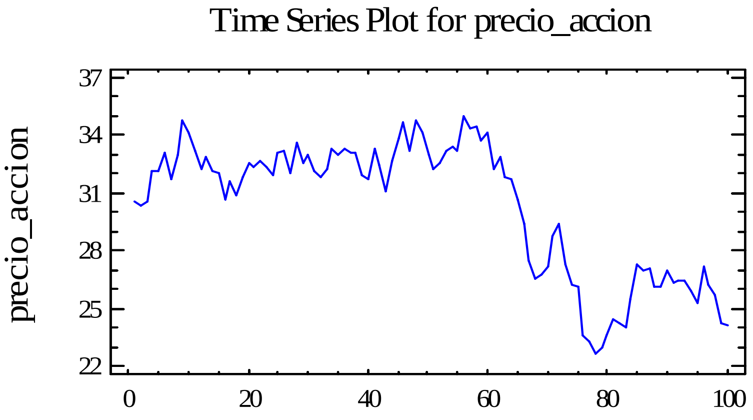
Ventas cuatrimestrales de un producto de consumo.



Pasajeros transportados mensualmente en vuelos internacionales.



Precio diario de una **acción** en la bolsa de Madrid.

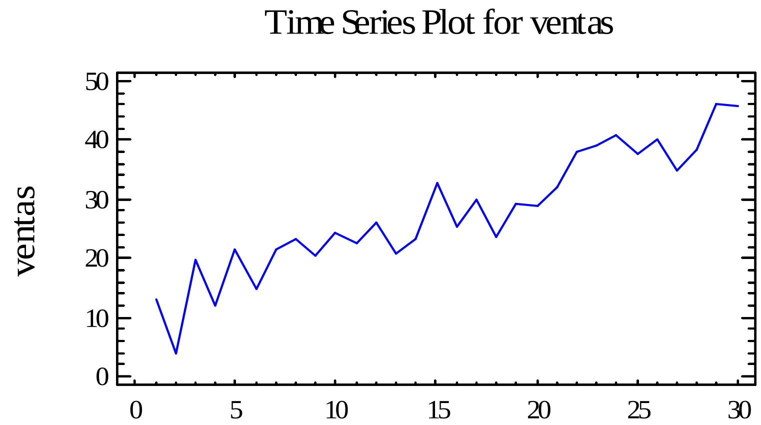


4.1.1. Clasificación de las series temporales.

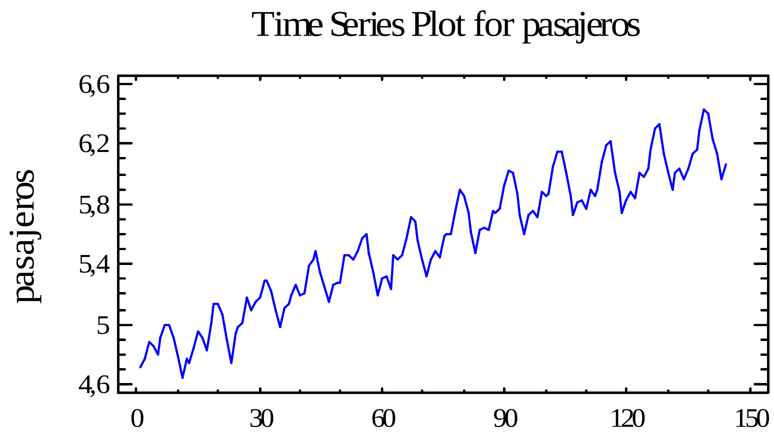
Estacionarias: la media y la variabilidad son **constantes** a lo largo del tiempo.	Oscilaciones de la serie alrededor de una media constante. La variabilidad de la serie con respecto a esta media también se mantiene constante. **Estudio de la serie**: histograma, descriptiva numérica (Tema 1).
No estacionarias: la media y/o la variabilidad **no son constantes** a lo largo del tiempo.	El cambio en la media puede marcar una **tendencia** de la serie a crecer o a decrecer. Si existe una pauta que se repite a lo largo del tiempo, diremos que la serie es **estacional** (habitual en series mensuales o diarias). **Estudio de la serie**: los métodos descriptivos vistos en el Tema 1 **no son adecuados**.

Serie estacionaria.

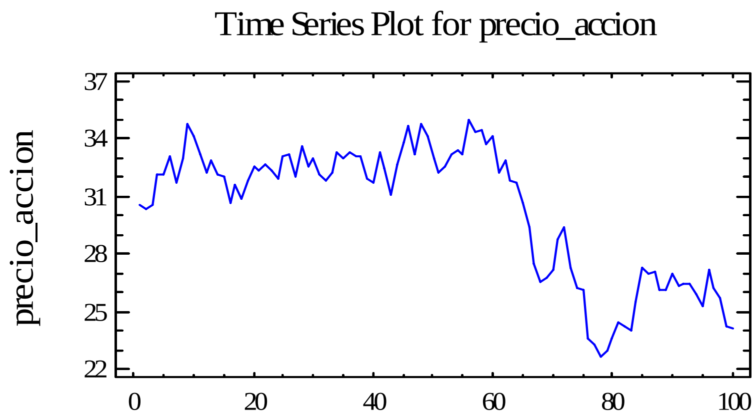
Serie no estacionaria, con tendencia creciente.



Serie no estacionaria con tendencia creciente y estacionalidad.



Serie no estacionaria con media y variabilidad estocásticas.



4.1.2 Descomposición de una serie temporal.

Expresaremos la serie temporal X_t como suma de tres componentes distintos:

$$X_t = T_t + S_t + I_t$$

- T_t es la **tendencia**: movimiento suave de la serie a largo plazo.
  a) Creciente constante
  b) Decreciente constante
  c) Variable o estocástica (cambios de dirección y/o de magnitud).

- S_t es la **estacionalidad**: comportamiento de la serie por el cual para unos períodos de tiempo los valores de la serie están sistemáticamente por encima o por debajo de la media.

- I_t es el componente **irregular**: incluye las variaciones aleatorias alrededor de los componentes anteriores.

4.1.3 Análisis de la tendencia.

Sea X_t una serie no estacionaria **sin estacionalidad**, es decir:

$$X_t = T_t + I_t$$

Hipótesis sobre la tendencia:

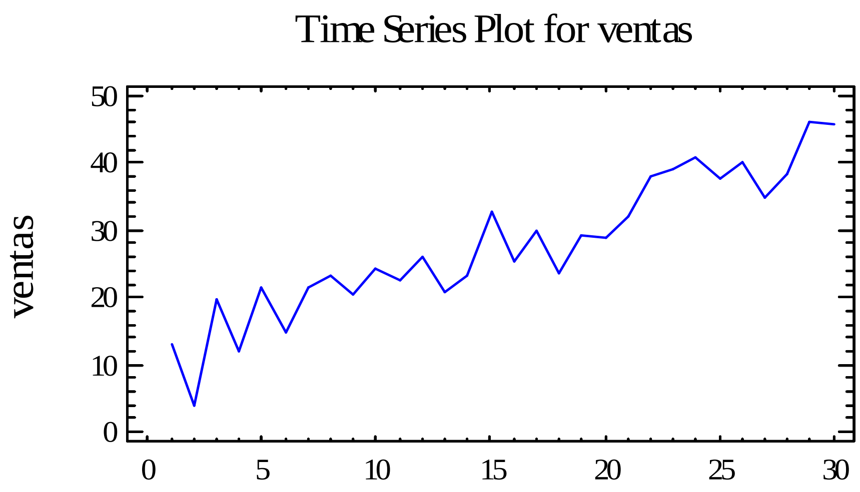
a) **Tendencia determinista**:

T_t es una función determinista del tiempo, por ejemplo:

$$T_t = a + b\,t.$$

b) **Tendencia no determinista**:

T_t es una función no determinista, desconocida, que evoluciona **suavemente** a lo largo del tiempo.



a) Tendencia determinista →Suponer un Modelo para la tendencia: $T_t = a + b\,t$

T_t es una función lineal del tiempo: ***a*** y ***b*** son los parámetros de la recta de regresión de X_t sobre ***t***. El método de los mínimos cuadrados proporciona los valores de ***a*** y ***b*** :

$$b = \frac{s_{t\,X_t}}{s_t^2}, \qquad a = \bar{x}_t - b\bar{t}$$

Componente irregular: se obtiene restando a la serie los valores calculados para la tendencia:

$$I_t = X_t - T_t = X_t - (a + bt)$$

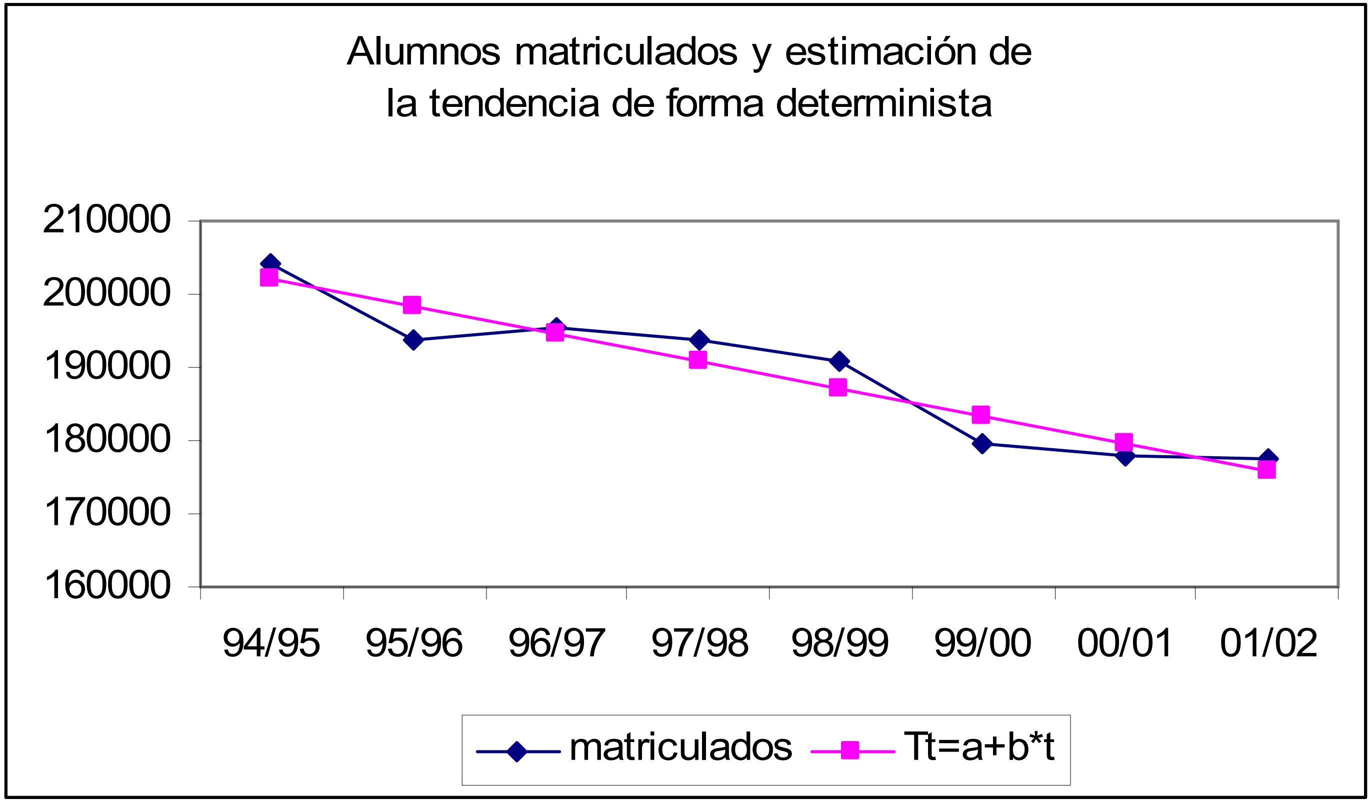
Los valores del componente irregular son los **residuos** de la recta de regresión de X_t sobre ***t***. La **desviación típica del componente irregular** es una medida de la **variabilidad** de la serie alrededor de su tendencia.

$$s_{I_t} = s_{X_t}\sqrt{1 - r^2_{t\,X_t}}$$

Ejemplo 16.

Alumnos universitarios matriculados y graduados en la Comunidad de Madrid.

CURSO	matriculados	graduados
94/95	203980	26436
95/96	193728	24622
96/97	195605	27086
97/98	193594	26362
98/99	190679	37562
99/00	179383	36183
00/01	177858	34815
01/02	177317	

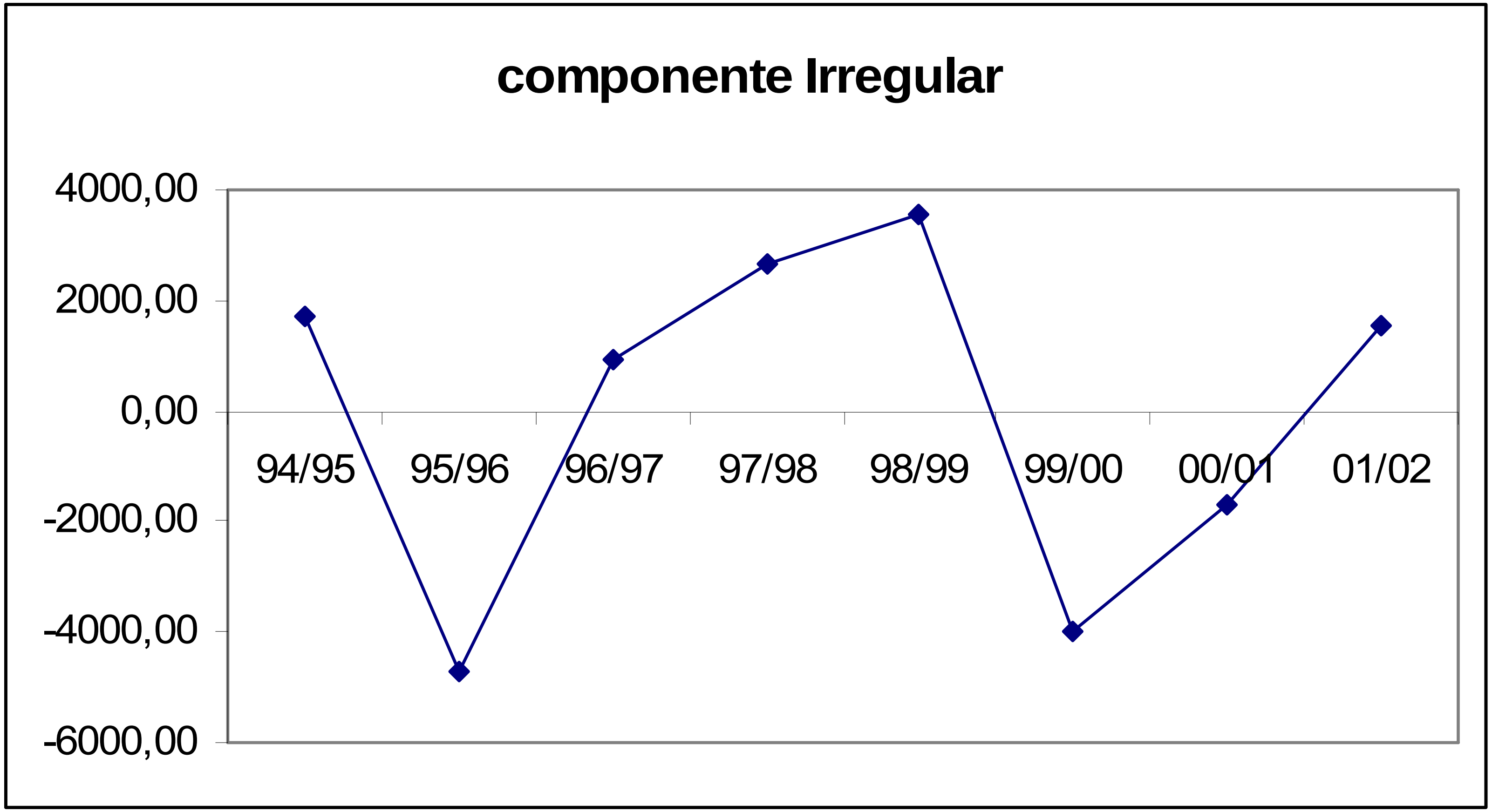


Estimación de la tendencia de forma determinista para los datos de alumnos matriculados

t	x_t	t^2	$x_t^{\ 2}$	$t\cdot x_t$	T_t	I_t
1	203980	1	41607840400	203980	202250,17	1729,83
2	193728	4	37530537984	387456	198469,55	-4741,55
3	195605	9	38261316025	586815	194688,93	916,07
4	193594	16	37478636836	774376	190908,31	2685,69
5	190679	25	36358481041	953395	187127,69	3551,31
6	179383	36	32178260689	1076298	183347,07	-3964,07
7	177858	49	31633468164	1245006	179566,45	-1708,45
8	177317	64	31441318489	1418536	175785,83	1531,17
36	1512144	204	2,8649E+11	6645862		

medias	4,5	189018	25,5	35811232454	830732,75
varianzas	5,25	83428129,5			
desv. tip.	2,29	9133,90			
covarianza	-19848,25			**correlación**	**-0,948389**
				R^2	**0,899442**

T_t=a+b·t	**a**	**b**
	206030,8	-3780,61905



$$s_{I_t} = s_{X_t}\sqrt{1-r^2_{t\,X_t}} = 9133.90\sqrt{1-0.899442} = 2896{,}44$$

b) Tendencia no determinista → Suponer que la tendencia sigue una función que evoluciona lentamente a lo largo del tiempo.

b1) Método de las medias móviles: La tendencia en cada punto se aproxima por la media del valor de la serie en ese período y los valores inmediatamente anteriores e inmediatamente posteriores.

Para *cada instante **t***, se calcula:

$$T_t \cong m_t = \frac{x_{t-1} + x_t + x_{t+1}}{3} \rightarrow \text{media móvil de orden } 3$$

$$T_t \cong m_t = \frac{x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2}}{5} \rightarrow \text{media móvil de orden } 5$$

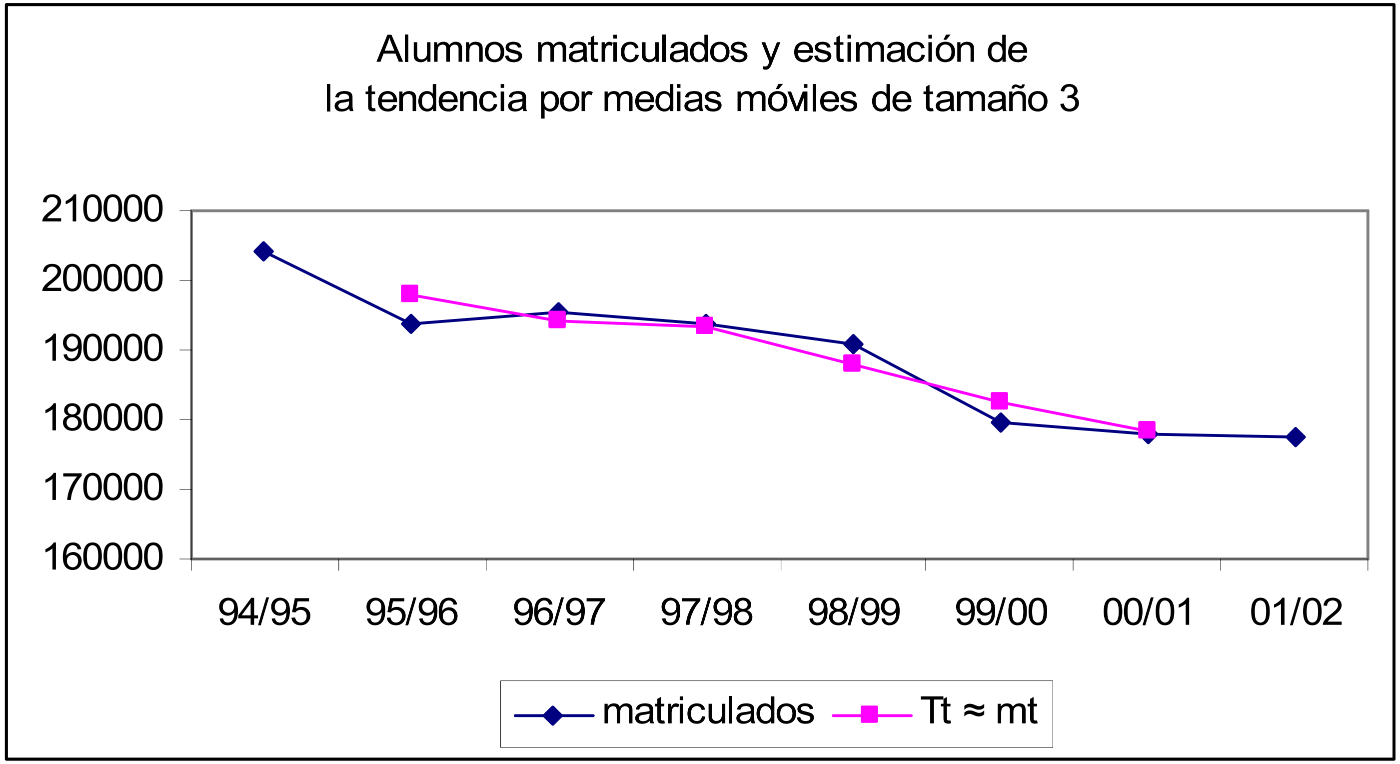
Componente irregular: se obtiene restando a la serie los valores calculados para la tendencia:

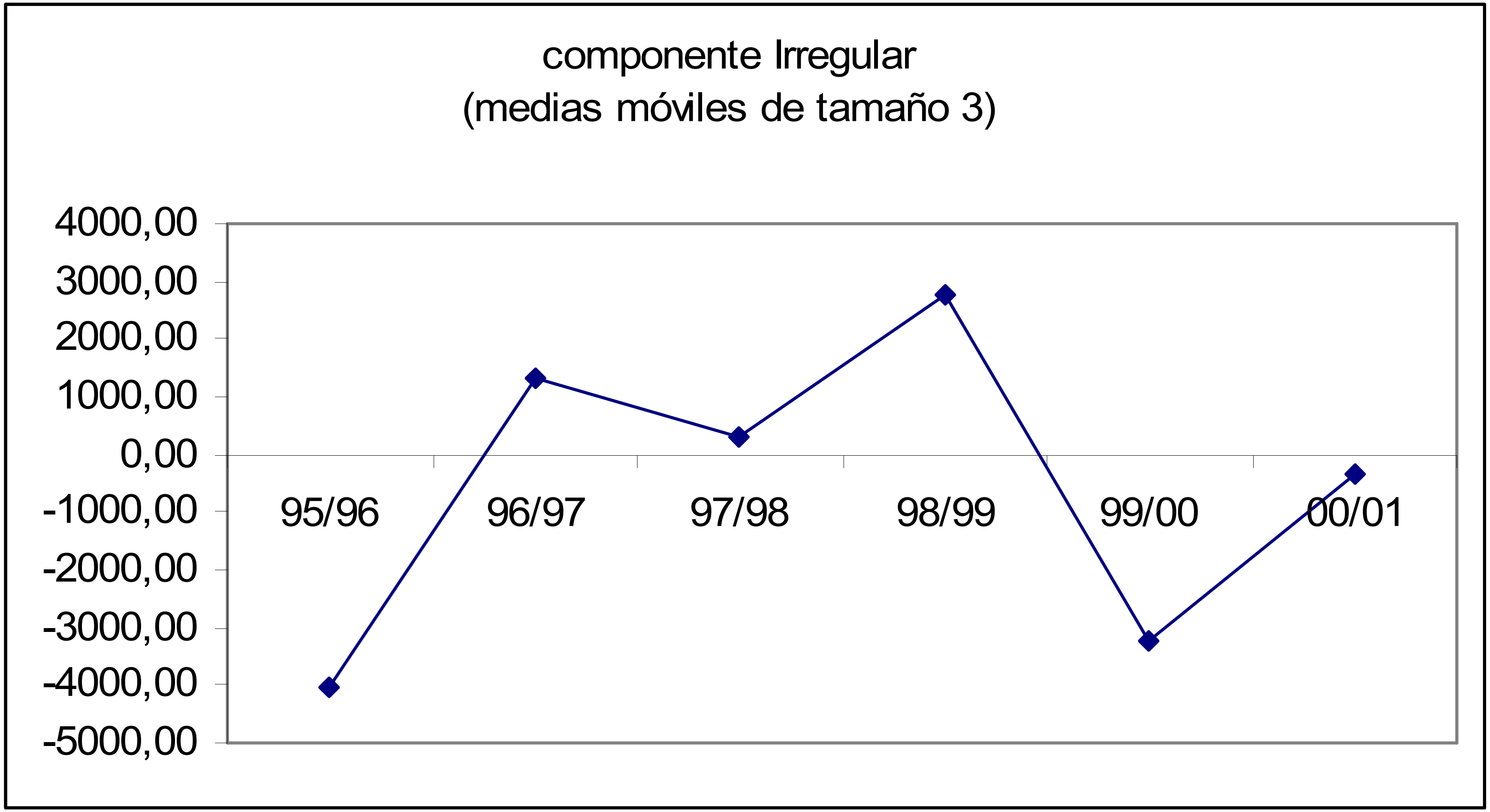
$$I_t = X_t - T_t = X_t - m_t$$

Ejemplo 17.

Estimación de la tendencia mediante medias móviles de tamaño 3 para los datos del ejemplo 16.

matriculados	$T_t \approx m_t$	I_t
203980		
193728	197771,00	-4043,00
195605	194309,00	1296,00
193594	193292,67	301,33
190679	187885,33	2793,67
179383	182640,00	-3257,00
177858	178186,00	-328,00
177317		




componente Irregular
(medias móviles de tamaño 3)
4000,00
3000,00
2000,00
1000,00
0,00
-1000,00
-2000,00
-3000,00
-4000,00
-5000,00
95/96
96/97
97/98
98/99
99/00
00/01

b2) Diferenciación de la serie.

Es el método más general para calcular la tendencia y el que más se utiliza en la actualidad.

Supone que la tendencia evoluciona **lentamente** a lo largo del tiempo, de manera que la tendencia en el instante ***t*** debe ser próxima a la tendencia en el instante ***t-1***.

Para cada instante de tiempo ***t*** se construye la serie

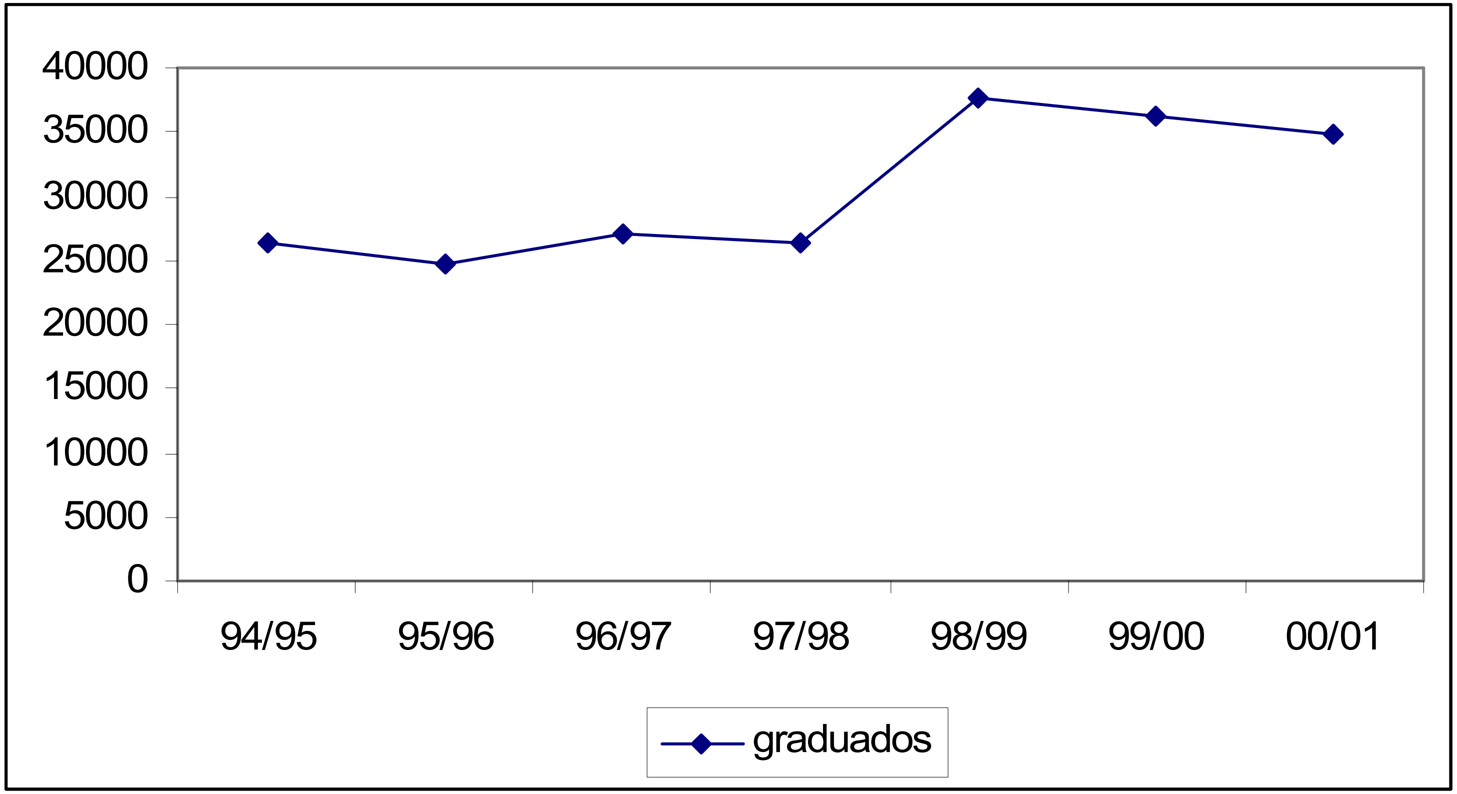
$$y_t = x_t - x_{t-1}$$

que se denomina **serie diferenciada** y que se supone libre de tendencia. La **serie diferenciada equivale al componente irregular**.

Ejemplo 18.

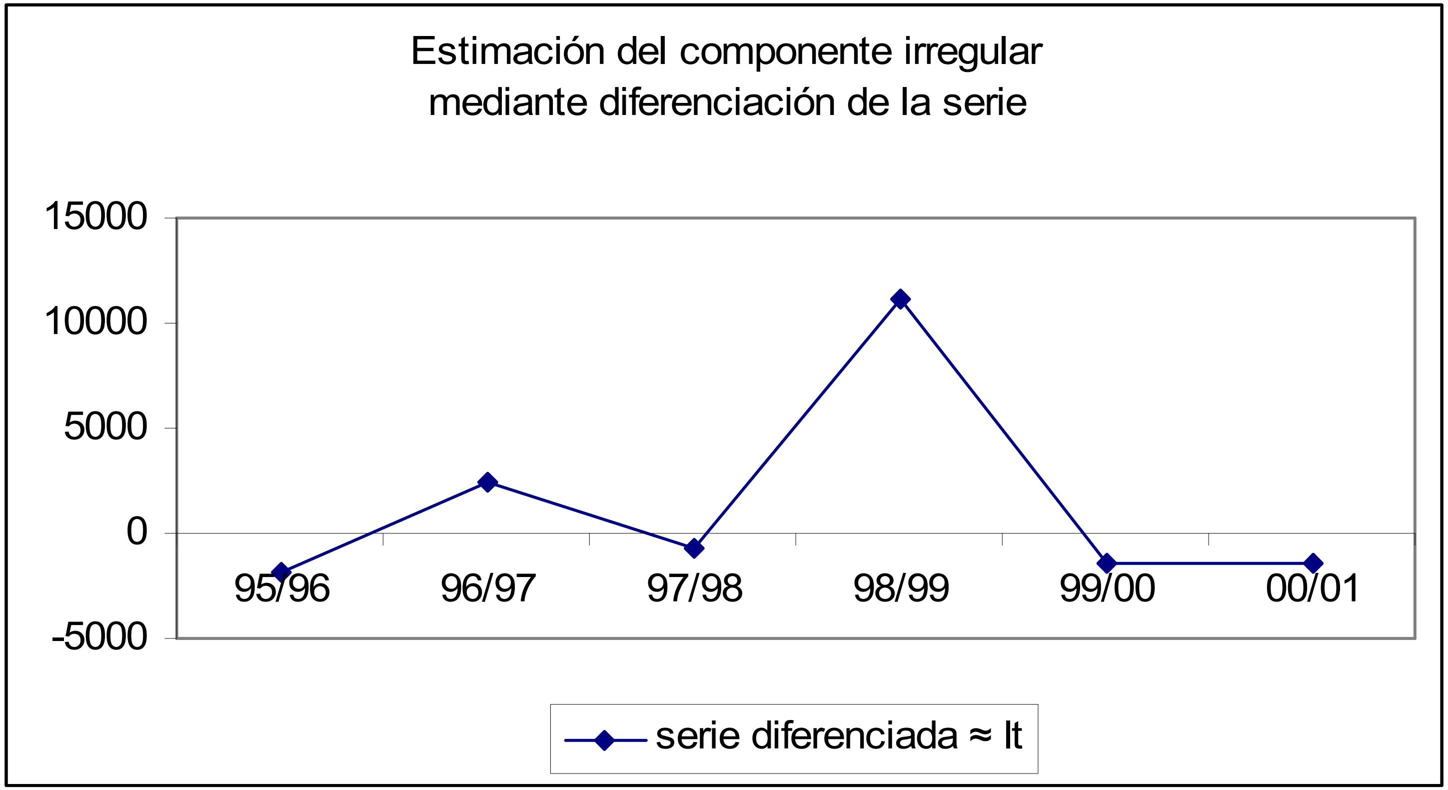
Estimación del componente irregular de la serie “graduados” del ejemplo 16.

CURSO	matriculados	graduados
94/95	203980	26436
95/96	193728	24622
96/97	195605	27086
97/98	193594	26362
98/99	190679	37562
99/00	179383	36183
00/01	177858	34815
01/02	177317	



Diferenciación de la serie "graduados"

serie	serie diferenciada ≈ I_t
26436	
24622	-1814
27086	2464
26362	-724
37562	11200
36183	-1379
34815	-1368



4.1.4 Análisis de la estacionalidad.

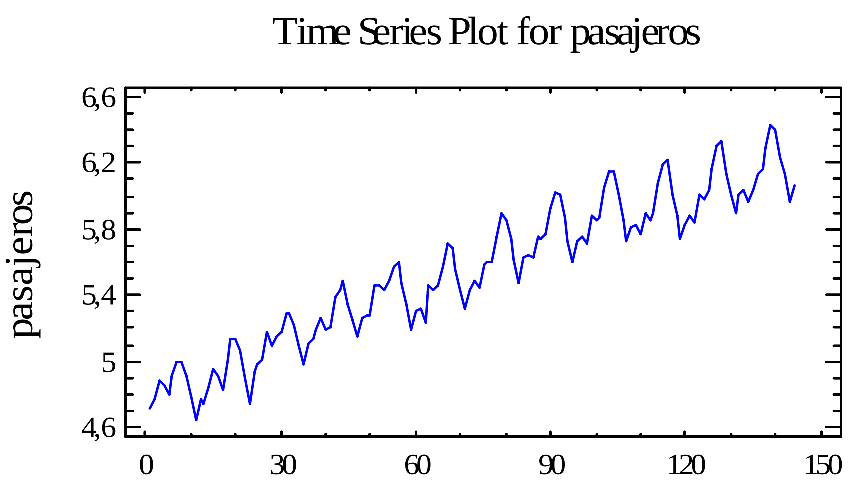
Cuando la serie observada es mensual o semanal (y, en general, siempre que resulte de recoger datos igualmente espaciados dentro del año) es plausible encontrar un **efecto estacional**.

X_t una serie no estacionaria **con estacionalidad:** $X_t = T_t + S_t + I_t$

Ejemplo 19: Consideremos la serie de pasajeros de avión. Se trata de una serie de datos mensuales recogidos a lo largo de 6 años. (Por tanto, el período estacional es 12).

Para calcular el **componente estacional** disponemos los datos en una tabla de doble entrada donde los valores de la serie se clasifican por año y por mes.

	90	91	92	93	94	95
enero	5,49	5,65	5,75	5,83	5,89	6,03
febrero	5,45	5,62	5,71	5,76	5,83	5,97
marzo	5,59	5,76	5,87	5,89	6,01	6,04
abril	5,59	5,75	5,85	5,85	5,98	6,13
mayo	5,60	5,76	5,87	5,89	6,04	6,16
junio	5,75	5,92	6,05	6,08	6,16	6,28
julio	5,90	6,02	6,14	6,20	6,31	6,43
agosto	5,85	6,00	6,15	6,22	6,33	6,41
septiembre	5,74	5,87	6,00	6,00	6,14	6,23
octubre	5,61	5,72	5,85	5,88	6,01	6,13
noviembre	5,47	5,60	5,72	5,74	5,89	5,97
diciembre	5,63	5,72	5,82	5,82	6,00	6,07



Se calculan las medias por filas y columnas y se obtienen los **coeficientes estacionales** como la diferencia entre la media de cada mes y la media global.

Los coeficientes estacionales representan cómo se sitúa la media de cada mes respecto de la media global.

	90	**91**	**92**	**93**	**94**	**95**	**media mensual**	**coeficiente de estacionalidad**
enero	5,49	5,65	5,75	5,83	5,89	6,03	5,77	**-0,14**
febrero	5,45	5,62	5,71	5,76	5,83	5,97	5,72	**-0,18**
marzo	5,59	5,76	5,87	5,89	6,01	6,04	5,86	**-0,05**
abril	5,59	5,75	5,85	5,85	5,98	6,13	5,86	**-0,05**
mayo	5,60	5,76	5,87	5,89	6,04	6,16	5,89	**-0,02**
junio	5,75	5,92	6,05	6,08	6,16	6,28	6,04	**0,13**
julio	5,90	6,02	6,14	6,20	6,31	6,43	6,17	**0,26**
agosto	5,85	6,00	6,15	6,22	6,33	6,41	6,16	**0,25**
septiembre	5,74	5,87	6,00	6,00	6,14	6,23	6,00	**0,09**
octubre	5,61	5,72	5,85	5,88	6,01	6,13	5,87	**-0,04**
noviembre	5,47	5,60	5,72	5,74	5,89	5,97	5,73	**-0,18**
diciembre	5,63	5,72	5,82	5,82	6,00	6,07	5,84	**-0,07**
media anual	5,64	5,78	5,90	5,93	6,05	6,15	**5,91**	

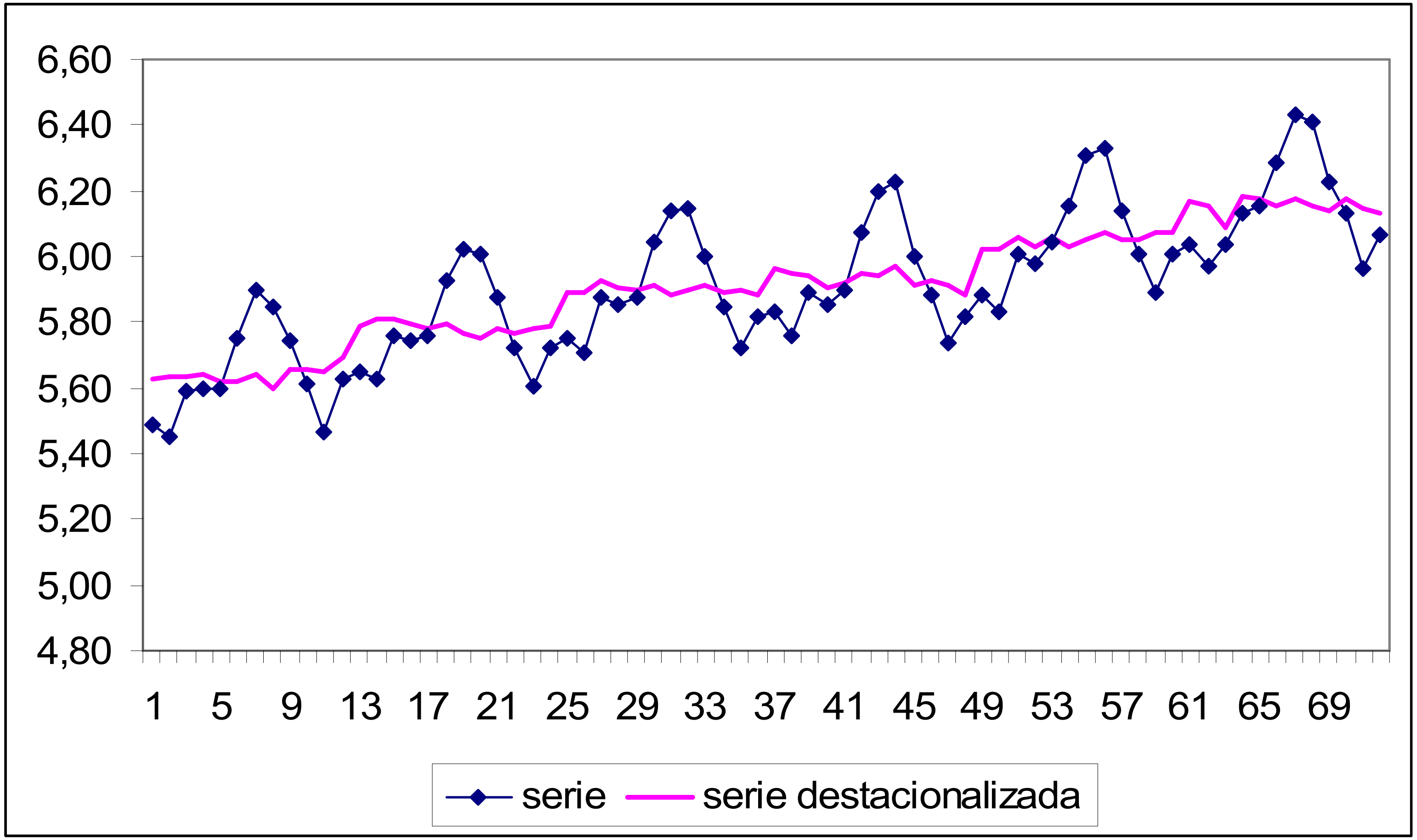
Desestacionalización de la serie (eliminación de la componente estacional): consiste en restar a cada valor de la serie el coeficiente estacional de su correspondiente mes. Se obtiene así la **serie desetacionalizada**.

serie desestacionalizada

	90	**91**	**92**	**93**	**94**	**95**
enero	5,63	5,79	5,89	5,97	6,02	6,17
febrero	5,64	5,81	5,89	5,95	6,02	6,15
marzo	5,64	5,81	5,92	5,94	6,06	6,09
abril	5,64	5,80	5,90	5,90	6,03	6,18
mayo	5,62	5,78	**5,89**	5,92	6,06	6,18
junio	5,62	5,79	5,91	5,95	6,03	6,15
julio	5,64	5,77	5,88	5,94	6,05	6,18
agosto	5,60	5,75	5,90	5,97	6,08	6,16
septiembre	5,65	5,78	5,91	**5,91**	6,05	6,14
octubre	5,65	5,76	5,89	5,92	6,05	6,17
noviembre	5,65	5,78	5,90	5,91	6,07	6,14
diciembre	5,69	5,79	5,88	5,89	6,07	6,13

5.87-(**-0.02**)

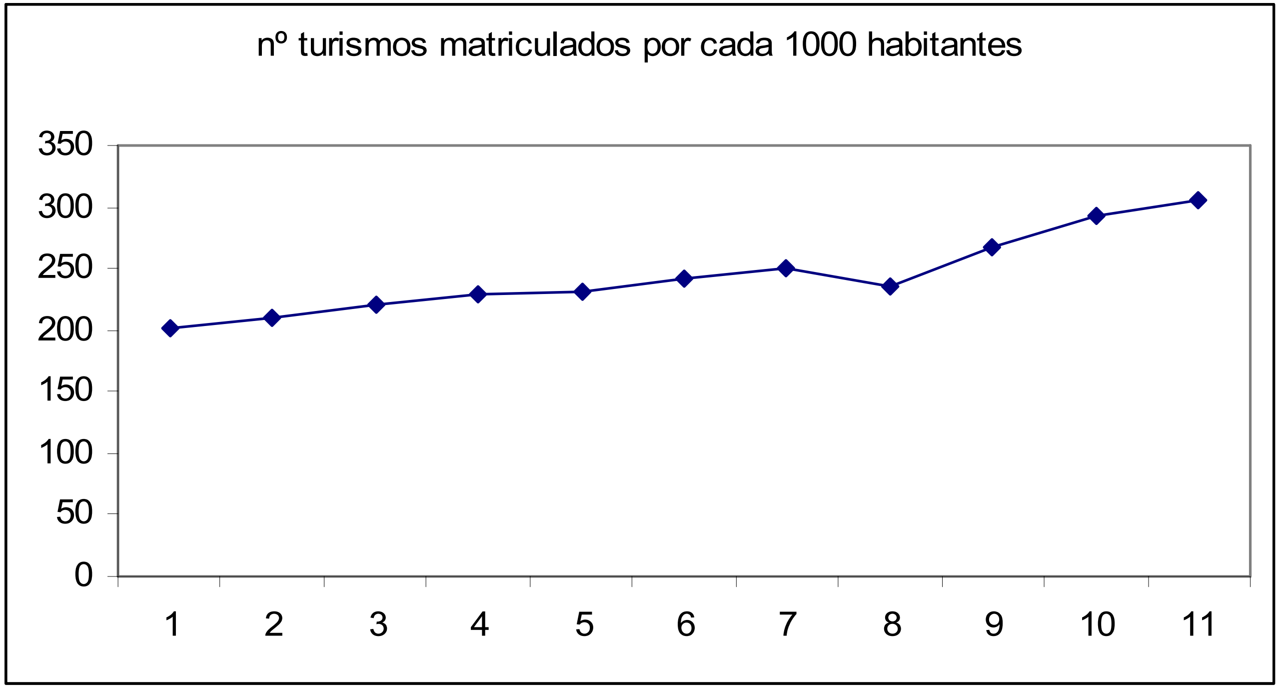
6.00-**0.09**



Ahora mediante alguno de los métodos del apartado 4.1.3 se estimaría la tendencia de la serie desestacionalizada. Se realizarían las predicciones y se estacionalizarían estos valores predichos para obtener los futuros valores de la serie.

Ejemplo 20.
Número de turismos matriculados por cada 1000 habitantes.

año (t)	nº turismos (x_t)
1	201,9
2	210,4
3	220
4	228,3
5	231,5
6	240,9
7	249,3
8	236,2
9	267,7
10	292,9
11	304,5



t	x_t	t^2	x_t^2	$t \cdot x_t$	$T_t=a+bt$	I_t
1	201,9	1	40763,61	201,9	197,61	4,29
2	210,4	4	44268,16	420,8	206,88	3,52
3	220	9	48400,00	660	216,15	3,85
4	228,3	16	52120,89	913,2	225,42	2,88
5	231,5	25	53592,25	1157,5	234,69	-3,19
6	240,9	36	58032,81	1445,4	243,96	-3,06
7	249,3	49	62150,49	1745,1	253,23	-3,93
8	236,2	64	55790,44	1889,6	262,50	-26,30
9	267,7	81	71663,29	2409,3	271,77	-4,07
10	292,9	100	85790,41	2929	281,04	11,86
11	304,5	121	92720,25	3349,5	290,31	14,19
66	2683,6	506	665292,6	17121,3		

medias	6,00	243,96	46,00	60481,15	1556,48
varianzas	10,00	962,89			
covarianza	92,7				

Predicciones	
año	nº turismos
12	299,58
13	308,85

tendencia $T_t=a+bt$

a	b
188,3436	9,2700