

CAPITULO 4.- SERIES TEMPORALES

4.1 Introducción.

Hasta ahora todas las variables que se han estudiado tenían en común que, por lo general, nunca han estado fechadas, es decir no estaban vinculadas al tiempo en forma alguna y menos explícitamente. Se trataba de datos de corte transversal o atemporales. Sin embargo es muy frecuente, especialmente en el ámbito económico y en general en las ciencias sociales, que las observaciones de los caracteres de una población se realicen ligadas al tiempo o fechadas en instantes determinados del tiempo. Así, por ejemplo, una de los caracteres de una empresa susceptible de ser observado puede ser su volumen de ventas y podemos estar interesados en estudiar el comportamiento y evolución temporal de esa característica de la empresa. En este caso esa observación se realizará de forma repetida durante una serie de momentos del tiempo. Esa observación repetida en el tiempo da lugar a una serie temporal. En este sentido diremos que *una serie temporal, cronológica, histórica o de tiempo es una sucesión de observaciones cuantitativas de un fenómeno ordenadas en el tiempo.*

El análisis de series temporales, desde el punto de vista de su comportamiento, tanto pasado como futuro, requiere el uso de nuevas técnicas, pues las presentadas hasta el momento, aunque le son aplicables, no cubren las necesidades que surgen en el tratamiento de este tipo de datos.

Desde el momento que los valores de una serie temporal van ligados a instantes del tiempo, entonces, podemos decir que el análisis de una serie implica el manejo conjunto de dos variables, siendo una de ellas nuestra serie temporal y la otra los intervalos o instantes del tiempo sobre los cuales se han realizado las observaciones. Hay que señalar que esa observación sincronizada de la variable en el tiempo implica que los valores de la misma han de estar perfectamente ordenados, de igual modo que los intervalos del tiempo lo están.

Esas observaciones de una variable cuantitativa pueden estar referidas, como ya se ha señalado, a un instante del tiempo o a un intervalo del mismo, dando lugar a dos tipos de magnitudes. En el primer caso hablaremos de *magnitudes stocks* o *niveles*. En el segundo se habla de *flujos*. Una variable stock es el número de empleados de una empresa en un instante del tiempo, a final de cada mes, mientras que un flujo serían las ventas de esa empresa a lo largo de ese mes, por ejemplo. La diferencia

entre una y otra es que la primera no es sumable para los distintos instantes de un intervalo, pues se incurriría en duplicaciones de los valores de esa magnitud. En cambio, el segundo tipo de magnitud si es sumable o acumulable a lo largo de un periodo o intervalo de tiempo. Para este segundo tipo, los intervalos para los que se acumulan deben ser siempre de igual amplitud. Es decir, si se dan datos de ventas de una empresa, estos deberán ser siempre mensuales, trimestrales, etc, pero lo que no nunca deberá hacerse es intentar trabajar con una serie que mezcle datos semanales con mensuales o referidos a cualquier otro periodo temporal.

Este requisito lleva implícita la idea de homogeneidad. Para que el análisis de una serie temporal nos conduzca a conclusiones acertadas no basta con utilizar las técnicas apropiadas, sino que será imprescindible que esos datos sean comparables y no lo serán nunca si no son homogéneos. Si cada año cambia la metodología de observación, se cambian las definiciones, se modifica la población de referencia, etc, el resultado será una serie temporal compuesta por un conjunto de valores no comparables porque son muy heterogéneos. Esta falta de homogeneidad se pierde, de una forma natural, con el transcurso del tiempo, de manera que cuando las series son muy largas no hay garantía de que los datos iniciales y finales sean comparables. Pero esta necesidad de que las series no sean muy largas, para que sus datos no pierdan la deseable homogeneidad, entra en contradicción con el objetivo más elemental de la Estadística que es el de detectar regularidades en los fenómenos de masas.

Lo que se pretende con una serie es describir y predecir el comportamiento de un fenómeno que cambia en el tiempo. Esas variaciones que experimenta una serie temporal pueden ser de naturaleza doble. Por un lado las variaciones pueden ser evolutivas o estacionarias. Diremos que las variaciones son evolutivas cuando el valor medio de la serie cambia, no permanece fijo a lo largo del tiempo, mientras que las variaciones estacionarias son aquellas en las su valor medio no cambia, aunque sufra oscilaciones en torno a ese valor medio fijo o constante¹. Esta clasificación de las variaciones de una serie permite hablar de series evolutivas y estacionarias.

¹ Esta forma de definir una serie estacionaria es solo una aproximación al concepto de la misma. No basta con que el valor medio no cambie a lo largo del tiempo. También es preciso que la variabilidad de la misma sea constante.

4.2 Componentes de una serie temporal.

La forma más sencilla de iniciar el análisis de una serie temporal, al igual que se ha venido haciendo con datos de corte transversal, es mediante su representación gráfica. Para ello se hará uso de un sistema cartesiano en el que los valores o periodos de tiempo se llevan al eje de abscisas y los valores de la serie, y_t , se llevan al eje de ordenadas. El resultado es un diagrama de dispersión, con la particularidad de que el eje de abscisas se reserva siempre a la misma variable: el tiempo.

Mediante este tipo de representación se pueden detectar las características más sobresalientes de una serie, tales como el movimiento a largo plazo, la amplitud de las oscilaciones, la posible existencia de ciclos, los puntos de ruptura, la presencia de valores atípicos o anómalos, etc. Un ejemplo de este tipo de gráficas es el que aparece en la Figura 1, donde se ha representado la serie que recoge el paro registrado en España para un periodo de cinco años con datos mensuales. Estos datos son los que se dan en la Tabla 1.

Una vez iniciado el proceso de descripción de una serie y superado el primer paso que consiste en su representación gráfica, para poder llegar a conclusiones más definitivas respecto del comportamiento de la serie, es conveniente recurrir a otras técnicas que superen el mero análisis gráfico.

Tabla 1. Evolución del paro registrado en España (Miles de parados).

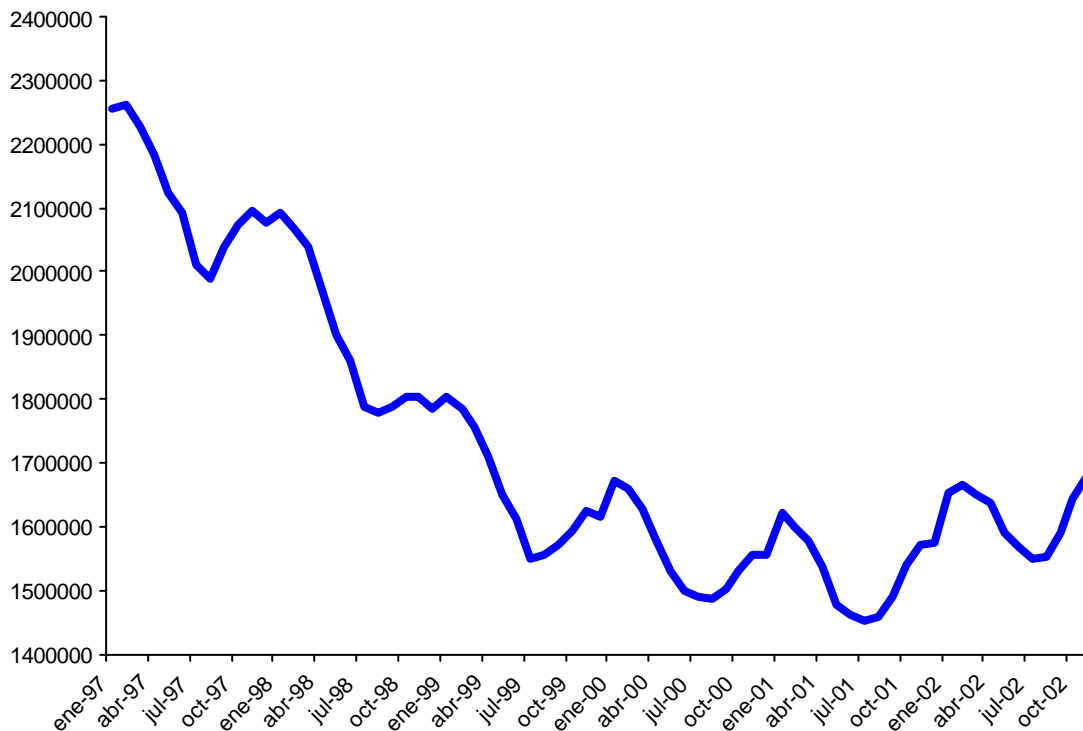
	1997	1998	1999	2000	2001	2002
Enero	2256,5	2091,3	1804,2	1670,6	1620,7	1651,7
Febrero	2262,7	2067,8	1783,9	1659,8	1598,9	1666,0
Marzo	2227,5	2039,1	1757,2	1628,5	1578,5	1649,0
Abril	2181,7	1968,0	1708,0	1578,9	1535,1	1636,3
Mayo	2123,8	1902,2	1649,1	1531,2	1478,1	1589,0
Junio	2091,9	1860,6	1612,5	1500,1	1460,6	1567,4
Julio	2009,2	1786,1	1551,0	1488,8	1451,5	1548,4
Agosto	1989,0	1777,1	1554,5	1487,6	1459,0	1552,0
Septiembre	2040,1	1788,4	1570,0	1501,4	1488,6	1590,3
Octubre	2072,9	1803,7	1591,7	1530,1	1540,0	1641,7
Noviembre	2093,9	1804,5	1623,7	1556,9	1572,8	1678,0
Diciembre	2075,7	1785,7	1613,8	1556,4	1574,8	1688,1

Fuente: Servidor web del INE

El objetivo del análisis de series temporales es doble. Por un lado se busca explicar las variaciones observadas en la serie en el pasado, tratando de determinar si responden a un determinado patrón de comportamiento. Y por otro, si se consigue definir ese patrón o modelo, se intentará predecir el comportamiento futuro de la misma.

Para alcanzar este doble objetivo se utiliza una metodología bastante consolidada, según la cual se admite que la serie temporal es una función del tiempo: $y_t = f(t)$. Bajo este esquema, la serie sería una variable dependiente y el tiempo una independiente o explicativa. Sin embargo, es necesario dejar bien claro que el tiempo, en si, no es una variable explicativa, es simplemente el “soporte” o escenario en el que se realiza o tiene lugar la serie temporal. El tiempo no sirve para explicar el comportamiento de la serie. A esta forma de abordar el estudio de una serie temporal se le conoce como enfoque clásico, frente al causal, según el cual, cualquier serie, como variable que es, puede ser explicada por otra u otras series.

Figura 1. Evolución del paro registrado en España



Desde este punto de vista, cualquier serie temporal se supone que es el resultado de cuatro componentes: **tendencia**, **variaciones estacionales**, **variaciones cíclicas** y **variaciones residuales o accidentales**. Pero esta descomposición de la serie no deja de ser un procedimiento diseñado para que el estudio de la misma resulte más fácil, pues esas componentes no siempre existen. Así cuando se trabaja con datos anuales la serie no puede presentar estacionalidad. A su vez las variaciones cíclicas son una componente ligada especialmente a las variables de tipo económico, pero que en variables de otra naturaleza puede que no esté presente. Estas componentes se definen en la forma siguiente:

1ª **Tendencia (T)**. De forma amplia podemos definir la tendencia como aquella componente que recoge el comportamiento de la serie a largo plazo. Para poder detectarla es necesario que la serie conste de un número de observaciones elevado, a lo largo de muchos años, para que se puede determinar si la serie muestra un movimiento a largo plazo que responda a una determinada ley de crecimiento, decrecimiento o estabilidad. Ese comportamiento tendencial puede responder a distintos perfiles: lineal, exponencial, parabólico, logístico, etc. Para el ejemplo del paro registrado, puede observarse que la tendencia de esa serie a lo largo de esos seis años (este periodo de tiempo no es muy largo para hablar de tendencia a largo plazo) es prácticamente una línea recta con pendiente negativa, aunque el ritmo de decrecimiento no solo se reduce al final del periodo sino que a lo largo de 2002 parece tener lugar un incipiente cambio de tendencia. O sea, que se tiene una serie que es, básicamente, decreciente para el periodo considerado. Mediante la tendencia se puede ver si la serie es estacionaria o evolutiva. Al considerar estos movimientos a largo plazo, prescindiremos de las variaciones a corto y medio plazo.

2ª **Variaciones estacionales (VE)**. Son movimientos de la serie que se repiten de forma periódica. La razón de estas variaciones se basa en causas de tipo climatológico (producción, turismo, etc.) o de ordenación del tiempo (los días de la semana condiciona el comportamiento de ciertas series). La periodicidad generalmente es el año, aunque puede ser el mes, la semana o incluso el día. En el ejemplo de la Figura 1 se observa un patrón de estacionalidad bastante bien definido: el paro registrado desciende notablemente en los meses estivales y el resto del año se mantiene en niveles más elevados, salvo en el mes de diciembre que, de forma sistemática es algo más reducido que en los anteriores y posteriores.

3ª **Variaciones cíclicas (C)** Esta componente tiene un marcado carácter económico, pues suele ser el resultado de la sucesión de las fases expansivas y recesivas de la economía. Son movimientos a plazo medio, periodos superiores al año, que se repiten de forma casi periódica, aunque no son tan regulares como las variaciones estacionales. Esta componente resulta difícil de aislar, pues ocurre, con frecuencia, que se pueden superponer ciclos de distintos periodos o amplitudes. La amplitud es el número de años que dura un ciclo completo. En nuestro ejemplo no se detecta de forma clara la presencia de ciclos, bien sea porque el periodo de tiempo estudiado sea muy corto o porque realmente no hay ciclos, aunque lo más verosímil en este caso sea la primera razón, pues el empleo responde a los ciclos de la economía.

4ª **Variaciones accidentales (R)**. Esta componente no responde a ningún patrón de comportamiento, sino que es el resultado de factores fortuitos o aleatorios que inciden de forma aislada y no permanente en una serie. Estos factores pueden ser de índole muy diversa tales, como inundaciones, huelgas y otras similares.

La interacción de estas cuatro componentes genera la serie temporal. La forma en que se combinen puede ser muy variada, pero tradicionalmente se ha optado por dos modelos distintos. El aditivo y el multiplicativo, aunque en algunas ocasiones se mezclan ambos. Según el modelo que se adopte, la serie temporal será:

$$y_t = T_t + VE_t + C_t + R_t \quad (4.1)$$

en el caso del modelo aditivo, y

$$y_t = (T_t)(VE_t)(C_t)(R_t) \quad (4.2)$$

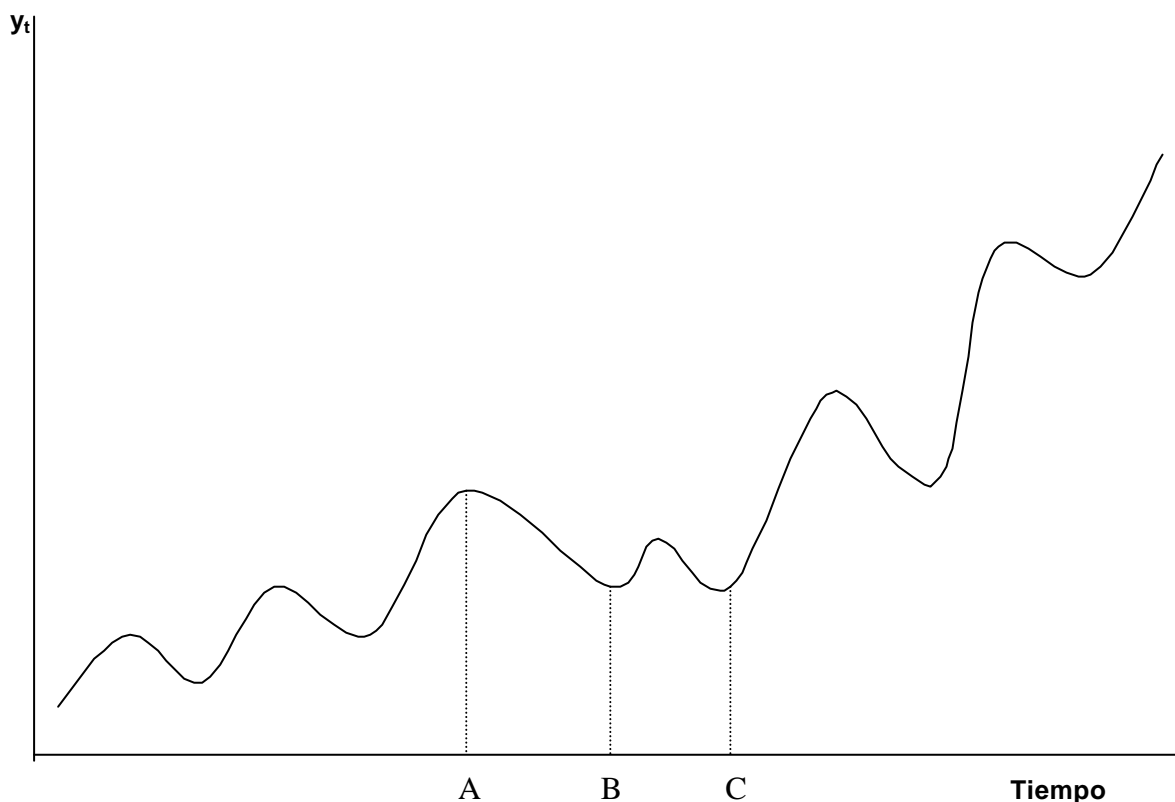
si se admite un esquema multiplicativo.

Frente a este tratamiento clásico de las series temporales, también se puede optar por otro enfoque de tipo causal, donde las variaciones de una serie podrían explicarse mediante las de otro conjunto de series temporales.

4.3 Obtención de la tendencia.

Para aislar esta primera componente de una serie se pueden utilizar distintos métodos alternativos. Pero con independencia del que se utilice, debe quedar bien claro que, tal y como se ha definido la misma, el periodo de información necesario debe ser lo suficientemente largo para evitar identificar como tendencia otros movimientos distintos de la serie.

Figura 2.



Para entender esta idea se puede hacer uso de la Figura 2. En la misma se ha representado una serie ficticia que muestra un perfil creciente, por lo que se puede concluir que la tendencia de la misma, como movimiento a largo plazo, no es decreciente, sino más bien lo contrario. Sin embargo, si nos hubiéramos limitado al periodo de tiempo que va de los puntos A al B, la conclusión sería justo la contraria de la indicada antes. Además, como dentro de ese periodo de tiempo no se ha observado ninguna perturbación de la “tendencia”, salvo que se tuviera información extra, no habría motivos para dudar de una tendencia lineal decreciente. Al mismo resultado se habría llegado si el periodo de observación hubiera sido desde A hasta C, solo que

ahora se diría que ha habido un ciclo por medio. Pero, tanto en un caso como en otro, las conclusiones serían poco afortunadas, pues se habría detectado de forma correcta la tendencia pero circunscrita a un periodo de tiempo muy corto, lo que entra en contradicción con la definición misma de tendencia, que se asocia con periodos de tiempo largos.

El problema es que el concepto de largo plazo va íntimamente ligado a la naturaleza de la variable, por lo que la longitud de esos periodos no siempre es comparable. Algo similar a lo descrito en este párrafo le ocurre a la serie recogida en la Figura 1, pues la tendencia que muestra el paro registrado en el periodo considerado es, en realidad, la manifestación de la fase expansiva del ciclo económico experimentado por la economía española a lo largo de la segunda mitad de los años noventa.

A continuación vamos a comentar algunos de los métodos más habituales en la determinación de la tendencia.

4.3.1 Análisis gráfico.

Se trata del método más simple para la obtención de la tendencia. Pero en su sencillez está su debilidad, pues al no hacer uso de ningún procedimiento analítico que garantice, tanto la objetividad del resultado como la posibilidad de que dos analistas distintos lleguen al mismo resultado, este procedimiento es impreciso y no garantiza fiabilidad alguna. Todo depende del conocimiento que de esa serie tenga el investigador que la esté trabajando. A grandes rasgos, este método pasa por la representación gráfica de la serie para posteriormente, bien a mano alzada o por cualquier otro procedimiento de características similares, obtener la tendencia. Lo que si es cierto es que, mediante esta vía sencilla, se podrá tener una idea preliminar de cual es la tendencia. Además la representación grafica de la serie es un paso previo aconsejable en el análisis de la misma.

4.3.2 Medias móviles.

Con este método lo que se hace es “suavizar” la serie promediando los valores de la misma para periodos de tiempo fijos pero que se desplazan a lo largo de todo el horizonte de la serie. El resultado de este proceso mecánico es la eliminación de los movimientos a corto y medio plazo así como las irregularidades debidas a factores no controlables ni predecibles. Es decir, a la serie se le quitan tres de sus componentes y se le deja solo la cuarta, la tendencia. La idea que subyace detrás de este método es

que la media de cualquier conjunto de valores sirve para eliminar la dispersión o variabilidad de la serie motivada por factores coyunturales o esporádicos.

Este método de suavizado consiste, como se ha indicado, en promediar la serie. Estos promedios serán las medias aritméticas de un conjunto k de valores consecutivos, con el requisito de que k sea inferior al total de observaciones. El procedimiento específico sería el siguiente. Supóngase que k es un entero impar. Entonces las sucesivas medias se obtendrían de forma siguiente:

$$y_t^* = \frac{\sum_{i=-\frac{k-1}{2}}^{\frac{k-1}{2}} y_{t+i}}{k} = \frac{y_{t-\frac{k-1}{2}} + y_{t-\frac{k-1}{2}+1} + y_{t-\frac{k-1}{2}+2} + \dots + y_t + \dots + y_{t+\frac{k-1}{2}-2} + y_{t+\frac{k-1}{2}-1} + y_{t+\frac{k-1}{2}}}{k} \quad (4.3)$$

A la media y_t^* se le llama centrada porque al ser impar el número de sumandos con el que se ha obtenido, la media resultante se le hace corresponder con la observación del momento t , que es el valor central de la suma. Según esta expresión de cálculo, la primera media que se puede calcular es la correspondiente al grupo de valores cuya observación central se corresponde al instante $t = \frac{k-1}{2}$, siendo en este caso la primera observación y_0 . Una vez obtenida esa media, la siguiente se calcula para los valores que tienen por observación central la del periodo $t+1$, y así sucesivamente. Esta forma de obtener medias eliminando la primera observación del grupo y añadiendo la siguiente es lo que le da el adjetivo de “móvil” a las mismas.

Par fijar un poco las ideas supongamos que $k=5$. En tal caso, las sucesivas medias móviles vendrían dadas por:

$$y_2^* = \frac{y_0 + y_1 + y_2 + y_3 + y_4}{5} \quad y_3^* = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5} \quad y_4^* = \frac{y_2 + y_3 + y_4 + y_5 + y_6}{5}$$

Ahora bien, si k fuera par, entonces la media de esos k valores no se correspondería con ninguno de los observados de la serie original, sino con el punto medio de los dos

centrales. Pero ese instante no es observable ($t = \frac{k-1}{2}$ no sería un entero), por lo que las medias calculadas de esta forma habría que promediarlas de dos en dos y de forma sucesiva para que el resultado si fuera una serie de valores (medias) centrados, es decir, que se correspondan con valores para periodos o instantes de tiempo observados².

Esta serie no centrada se obtendría mediante la expresión:

$$y_{t-0,5}^* = \frac{\sum_{i=\frac{k}{2}-1}^{\frac{k}{2}} y_{t+i}}{k} \quad \left(t = \frac{k}{2}, \frac{k}{2} + 1, \dots, n - \frac{k}{2} \right) \quad (4.4)$$

Otra cuestión importante a la hora de calcular las medias móviles es determinar cuantas observaciones deben tomarse en cada caso. Si k es muy grande entonces el proceso de suavizado puede llegar a ser tan fuerte que se pierda más información de la deseada. Piense en la situación extrema de que k fuera igual al total de observaciones. En ese caso solo habría una media, por lo que el suavizamiento de la serie sería máximo, tanto que no habría ni tendencia ni componente alguna. Por esta razón k no debe ser demasiado grande, pues se podría incurrir en un suavizado excesivo. En general, cuanto mayor es k , menor será el número de términos de la serie suavizada resultante (se pierden observaciones al principio y al final de la serie). En conclusión, si se toma un grupo de observaciones muy alto se incurre en el peligro de perder información por dos vías: a) la serie se suaviza más de lo necesario, ocultando ciertos movimientos tedenciales; b) el número de términos de la nueva serie se reduce considerablemente, y perder datos nunca es bueno.

² Este proceso de tomar medias móviles en dos ocasiones par centrar la serie puede reducirse a una media móvil ponderada, donde el número de sumando es $k+1$ y todos los valores se ponderan doble, salvo el primero y el último cuya ponderación es la unidad, y esa suma se divide entre $2k$. Así si $k = 4$, las

dos primeras medias descentradas serían: $\frac{y_0 + y_1 + y_2 + y_3}{4}$ y $\frac{y_1 + y_2 + y_3 + y_4}{4}$. A partir de estas medias, el valor centrado sería: $\frac{\frac{y_0 + y_1 + y_2 + y_3}{4} + \frac{y_1 + y_2 + y_3 + y_4}{4}}{2} = \frac{y_0 + 2y_1 + 2y_2 + 2y_3 + y_4}{8}$.

Por el contrario, si k es muy pequeño entonces no se conseguirán eliminar todas las perturbaciones ajenas a la tendencia. De forma similar a como se razonó antes, si $k=1$, entonces la serie original y la suavizada coinciden, con lo cual no se ha conseguido nada.

En algunos casos, ese valor de k es fácil de determinar. Así ocurre cuando la serie muestra un patrón de comportamiento que se repite de manera sistemática cada k periodos de tiempo. Tal sería el caso de la estacionalidad. Si se trabaja con datos mensuales y la serie está sometida a un esquema de estacionalidad que se repite todos los años, entonces la forma de suavizar esa serie y eliminar la componente estacional sería tomar una media móvil de doce meses ($k = 12$). A la serie resultante se le habrían eliminado dos componentes: la estacionalidad y las variaciones accidentales. Pero al ser k par, la serie resultante no estaría centrada, por lo que habría que volver a tomar medias móviles con $k = 2$.

Una vez que de la serie original se han eliminado esas dos componentes cabría preguntarse cómo proceder con las variaciones cíclicas. En este caso la elección de k es más difícil, pues los ciclos no son movimientos de la serie que se repitan con una periodicidad fija, como ocurre con la estacionalidad. En estos casos, si esa periodicidad no puede determinarse de forma clara y sin que perjudique notoriamente a los resultados, la mejor forma de proceder es trabajar con lo que ha dado en denominarse componente ciclo-tendencia.

En el esquema presentado hasta el momento, para la obtención de la tendencia mediante medias móviles, se ha trabajado con el supuesto de que los datos tenían una periodicidad inferior al año (semestres, cuatrimestres, trimestres, meses, etc.) lo que implicaba aceptar la posibilidad de que esa serie presentara estacionalidad. Ahora bien, si los datos fueran anuales entonces la estacionalidad quedaría descartada, pues las únicas componentes de la serie serían la tendencia, los ciclos y las variaciones accidentales. Esta nueva situación nos lleva a que sea poco verosímil que la serie presente un esquema repetitivo a lo largo del tiempo tan estable como presentaban las variaciones estacionales. Ante estas circunstancias se hace difícil saber cuál debiera ser el número adecuado de observaciones que debieran tomarse para calcular las medias móviles. La forma de salir de esta situación incómoda es obtener medias móviles de tres o cinco datos (número impar y pequeño) para de esa forma eliminar la componente accidental. Una vez que se ha procedido de esta forma, la serie suavizada resultante contiene una mezcla de ciclo-tendencia. Si la componente cíclica

fuera regular con periodos definidos y fijos, entonces la tendencia se obtendría aplicando una media móvil con un k igual a la longitud del ciclo. Pero es poco probable que los ciclos tengan ese comportamiento tan sistemático, por lo que quizás la mejor solución sea, como se indicó en el párrafo anterior, no manipular más los datos y trabajar con esa mixtura de componentes ciclo-tendencia.

Este método de obtención de la tendencia presenta, frente a su sencillez, algunos inconvenientes que deben ser señalados. Al igual que en el método gráfico, también aquí se introduce un cierto grado de subjetividad, pues la elección del número de observaciones a promediar queda a la elección del analista y, salvo que sea muy claro cual debe ser ese número (caso de la estacionalidad), esa decisión no siempre es la acertada, por lo que los valores de la componente tendencia variaran según quién los calcule. Por otro lado, esta forma de obtener la tendencia no permite alcanzar el objetivo de la predicción en el análisis de las series temporales, pues la tendencia obtenida mediante medias móviles no permite que se proyecte hacia el futuro.

Ejemplo 1. *Obtégase la tendencia de la serie de la Tabla 1 mediante medias móviles.*

En este caso, dado que los datos son mensuales y la serie muestra una clara componente estacional que se repite todos los años, el periodo de la media móvil debe ser de doce datos (doce meses). Pero al ser par el valor de k se deben tomar medias móviles en dos ocasiones. Primero con $k=12$ y después con $k = 2$, para de esta forma obtener una serie centrada, que será la tendencia, pues, como puede observarse, la serie original, para el conjunto de años considerado, no muestra componente cíclica clara. Los resultados de estas operaciones son los que aparecen en las Tablas 2 y 3. Adicionalmente, en la Figura 2, se ha vuelto a representar la serie original y la suavizada que recoge la tendencia. Esta última tiene menos observaciones que la primera. En este caso se ha perdido doce datos, seis al inicio y seis al final.

La forma en la que se han obtenido esos datos es la siguiente:

$$y_{97JUL}^* = \frac{\frac{y_{96ENE} + y_{96FEB} + \dots + y_{97NOV} + y_{97DIC}}{12} + \frac{y_{96FEB} + y_{96MAR} + \dots + y_{97DIC} + y_{97ENE}}{12}}{2} =$$

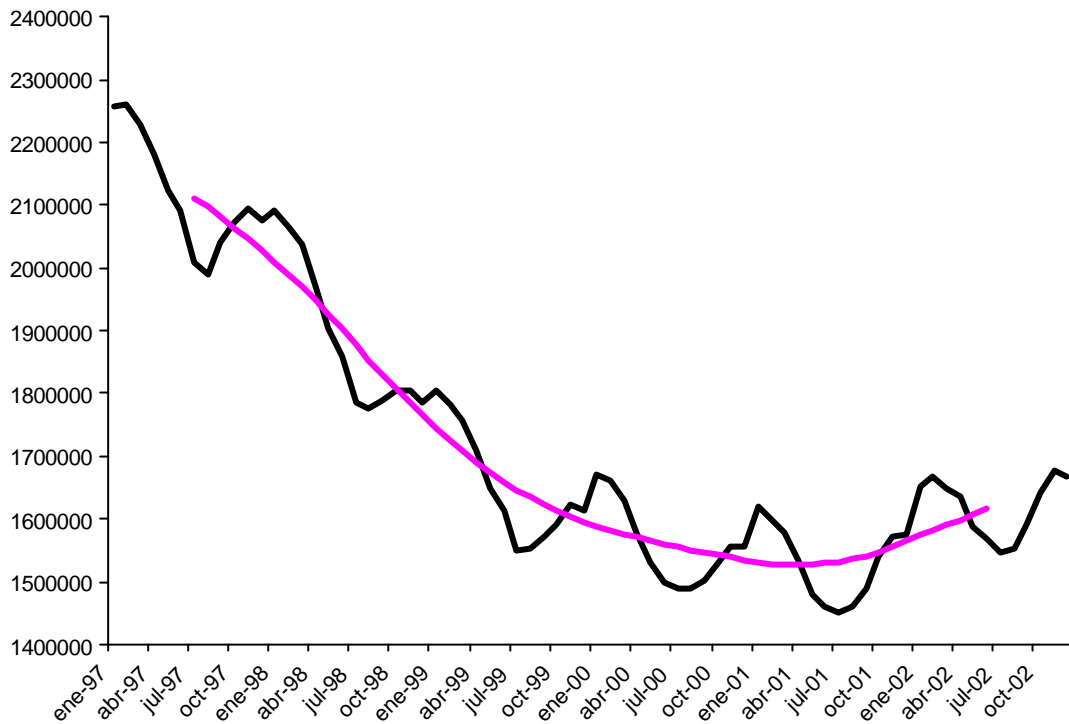
$$= \frac{\frac{2256,5 + 2262,7 + \dots + 2093,9 + 2075,7}{12} + \frac{2262,7 + 2227,5 + \dots + 2075,7 + 2091,3}{12}}{2} = \frac{2118,7 + 2105,0}{2} = 2111,9$$

Tabla 2. **Medias móviles de doce meses no centradas de la serie del paro registrado en España.**

	1997	1998	1999	2000	2001	2002
Enero	„	2017,5	1755,0	1589,5	1532,8	1570,5
Febrero	„	1998,9	1735,4	1584,3	1529,7	1578,6
Marzo	„	1981,2	1716,9	1578,7	1527,3	1586,3
Abril	„	1960,3	1698,7	1573,0	1526,2	1594,8
Mayo	„	1937,8	1681,0	1567,9	1527,0	1603,3
Junio	„	1913,7	1665,9	1562,3	1528,3	1612,1
Julio	2118,7	1889,5	1651,6	1557,5	1529,9	1619,8
Agosto	2105,0	1865,6	1640,5	1553,4	1532,5	„
Septiembre	2088,7	1842,0	1630,1	1548,3	1538,1	„
Octubre	2073,0	1818,5	1619,4	1544,1	1543,9	„
Noviembre	2055,2	1796,8	1608,7	1540,5	1552,4	„
Diciembre	2036,8	1775,7	1598,8	1536,1	1561,6	„

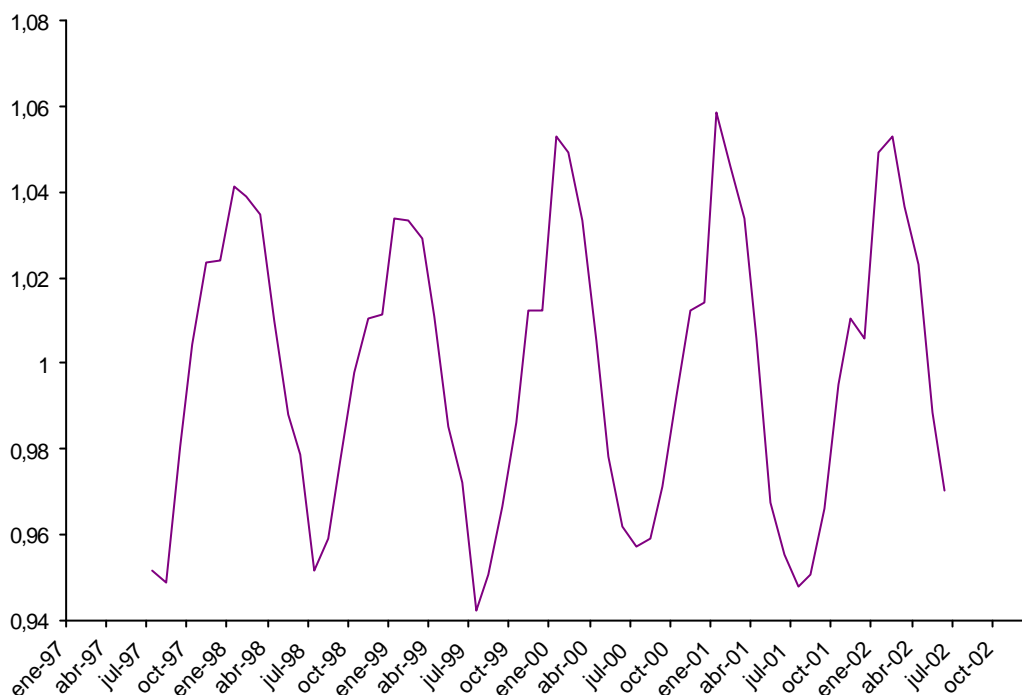
Tabla 3. **Tendencia del paro registrado en España obtenida por medias móviles. (Medias móviles centradas)**

	1997	1998	1999	2000	2001	2002
Enero	„	2008,2	1745,2	1586,9	1531,2	1574,6
Febrero	„	1990,1	1726,2	1581,5	1528,5	1582,5
Marzo	„	1970,7	1707,8	1575,9	1526,7	1590,6
Abril	„	1949,0	1689,9	1570,4	1526,6	1599,1
Mayo	„	1925,8	1673,5	1565,1	1527,7	1607,7
Junio	„	1901,6	1658,8	1559,9	1529,1	1615,9
Julio	2111,9	1877,6	1646,1	1555,5	1531,2	„
Agosto	2096,8	1853,8	1635,3	1550,8	1535,3	„
Septiembre	2080,9	1830,2	1624,8	1546,2	1541,0	„
Octubre	2064,1	1807,6	1614,0	1542,3	1548,2	„
Noviembre	2046,0	1786,3	1603,7	1538,3	1557,0	„
Diciembre	2027,1	1765,4	1594,1	1534,4	1566,1	„

Figura 3. Tendencia del paro registrado obtenida por medias móviles.

Una vez que se ha obtenido la componente ciclo-tendencia y si se admite un esquema multiplicativo de generación de la serie, entonces el cociente entre estas dos componentes nos daría la estacionalidad y las variaciones accidentales. Los resultados de este cociente aparecen en la Tabla 7 y su representación gráfica en la Figura 4.

Figura 4. Serie de paro registrado corregida de la componente ciclo-tendencia.



4.3.3 Método analítico.

Con este procedimiento de obtención de la tendencia, lo que se pretende es seleccionar una función matemática que modelice de forma adecuada el comportamiento a largo plazo de la serie temporal objeto de estudio. Se trata, por tanto, de ajustar los datos observados a esa función, donde la variable a explicar o dependiente es la propia serie temporal y la independiente o explicativa es, ahora, el tiempo. Ni que decir tiene que el tiempo, como ya se ha indicado, no explica nada, sino que es un mero soporte en el que se mueve la serie. El procedimiento de ajuste que puede utilizarse no es único, aunque el más utilizado es el de los mínimos cuadrados.

En consecuencia, se trata de seleccionar aquella función del tiempo que minimice la suma de los cuadrados de los errores. Pero aunque este es el criterio último, como paso previo para seleccionar esa función se puede recurrir a su representación gráfica, la cual nos informará de manera aproximada sobre la función que se debe ajustar. Otra alternativa es hacer uso del posible conocimiento de la naturaleza de esa serie o de lo que la teoría establezca.

Las funciones más utilizadas para modelizar series económicas son las siguientes:

- *Lineal.*

Se trata de un modelo en el que la serie temporal se hace depender linealmente del tiempo y que responde a aquellas magnitudes que presentan unas variaciones constantes en periodos sucesivos. La forma general del mismo es:

$$y_t = y_t^* + e_t = a + bt + e_t \quad (4.5)$$

donde: y_t , la serie original, se descompone en la tendencia y_t^* y, (bajo un supuesto aditivo), las otras componentes que aparecen de forma residual y conjunta bajo e_t . En este modelo b es la variación media entre periodos y t el tiempo cronológico.

- *Polinomial.*

La función Polinomial es en realidad una familia de funciones que se diferencian unas de otras en el grado del polinomio con el que se trabaje. La más común de todas es la parabólica. En este caso las variaciones de la serie no son constantes, ni en términos absolutos ni relativos. Para este modelo la serie temporal se expresa como:

$$y_t = y_t^* + e_t = a + bt + ct^2 + e_t \quad (4.6)$$

- *Exponencial*

Esta función surge cuando la serie cambia a razón de una tasa constante. Para este tipo de series, su tendencia viene dada por:

$$y_t^* = AB^t = y_0^* e^{bt} = e^{a+bt} \quad (4.7)$$

donde b es la tasa de variación instantánea de la serie temporal que es constante, y_0^* es el valor de la tendencia en punto $t = 0$, t es, de nuevo, el tiempo cronológico y e es la base de los logaritmos naturales.

El ajuste por mínimos cuadrados de cualquiera de estos modelos no plantea ningún problema especial, pues los mismos son lineales o fácilmente linealizables y, en estos casos, la aplicación del método de ajuste de los mínimos cuadrados es inmediato.

Además de las tres funciones señaladas antes también hay otras que estarían indicadas para aquellos casos en los que la serie presenta un techo en su crecimiento, es decir, que no es admisible un crecimiento indefinido. Ese tipo de variables se pueden modelizar con alguna de las siguientes funciones en forma de S:

- *Logarítmica recíproca:*

$$y_t^* = e^{a - \frac{b}{t}} \quad (4.8)$$

Si en esta función se toman logaritmos se tiene que:

$$\ln y_t^* = a - b \left(\frac{1}{t} \right) \quad (4.9)$$

que es una función lineal a la que es fácil aplicar mínimos cuadrados. Esta función se caracteriza porque su tasa de crecimiento frente a cambios unitarios en t es inversamente proporcional al cuadrado de t .

$$\frac{1}{y_t^*} \frac{dy_t^*}{dt} = \frac{b}{t^2} \quad (4.10)$$

lo que conlleva que la tasa de crecimiento no es constante.

La gráfica de esta función es del tipo representado en la Figura 4.

- *Logística:*

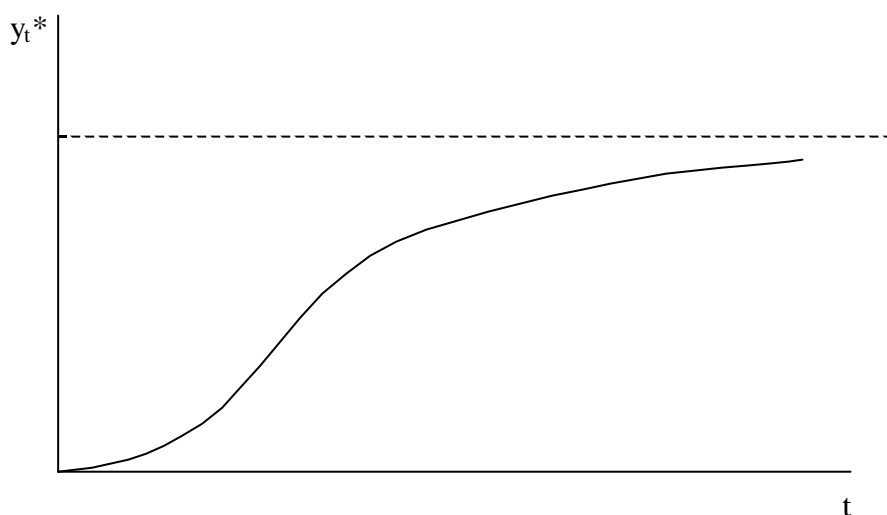
$$y_y^* = \frac{c}{1 + ae^{-bt}} \quad (4.11)$$

El perfil de esta curva es similar al de la anterior y que se ha representado en la Figura 5. Las diferencias estriban en que los puntos de inflexión de ambas son

distintos, pues mientras que en la primera el mismo es $y_t^* = 0,135e^a$, en la segunda ese punto se da en $y_t^* = c/2$. Es decir, en la primera el punto de inflexión ocurre antes que en la segunda y, por lo tanto, la caída en la tasa de crecimiento tiene lugar antes.

Al igual que el modelo logarítmico recíproco, la función logística presenta también una tasa de crecimiento variable. El gran problema de este modelo es que no es fácil linealizarlo, por lo que su ajuste es difícil con las técnicas que se manejan en este manual.

Figura 5. Función en forma de S.



Ni que decir tiene que el repertorio de funciones para modelizar la tendencia de una serie no se agota con las expuestas aquí, pero sí que puede decirse de ellas que son las más frecuentes.

La obtención de la tendencia mediante el método analítico presenta dos ventajas frente a los otros procedimientos descritos antes. La primera es que se tiene una medida de la bondad del ajuste, lo que permite determinar hasta qué punto la función seleccionada recoge de manera correcta la tendencia de la serie. Ahora la adecuación de la tendencia obtenida no se deja a la mera inspección gráfica, como ocurría con los dos procedimientos anteriores. Por otro lado, al obtener una función explícita, que relaciona los datos observados con el tiempo, se cubre mejor el segundo de los objetivos del análisis de series temporales, que como se señaló era la predicción.

Solo queda una cuestión de orden práctico que no se ha abordado hasta ahora. Es la forma de tratar a las series cuando los datos tienen periodicidad inferior al año y presentan estacionalidad. En estos casos, si el ajuste se realiza directamente con los datos observados (mensuales, trimestrales, etc.), la bondad del mismo será inferior a la que se obtendría si se trabajara con valores anuales (medias anuales), pues en este caso la serie presenta menos fluctuaciones y los errores del ajuste serán más pequeños. Pero ese criterio basado en la bondad del ajuste no es del todo trascendente en este contexto, pues la estacionalidad no tiene porque interferir en la tendencia, por lo que esta será la misma con independencia de que se obtenga con datos anuales o con periodicidad inferior al año.

Ejemplo 2 *Obtener la tendencia correspondiente a la serie de paro registrado dada en la Tabla 1 mediante el método analítico.*

A la vista de la representación gráfica de esta serie en la Figura 1, parece poco razonable ajustar una función lineal, pues el comportamiento del paro registrado a lo largo de esos seis años no puede decirse que fuera lineal. Quizás fura más apropiado ensayar un modelo parabólico, el cual podrá recoger la manifiesta no linealidad de los últimos años.

Inicialmente se trabajará con los datos mensuales. Otra decisión que habrá que tomarse es fijar el origen del tiempo cronológico. Como se sabe, el origen del tiempo se ha fijado siempre de forma arbitraria (el mundo occidental, el mundo árabe, el mundo hebreo, etc, tienen un calendario con un origen diferente). Esa libertad de elección permite que, desde el punto de vista estadístico, ese origen se fije en el punto que convenga más en cada momento, siempre que se indique donde se ha hecho $t=0$. En consecuencia, para este ejemplo, se situará ese origen en el primer mes. Es decir, se realizará un cambio de origen en la variable independiente, y este cambio ya se sabe como afecta a los distintos coeficientes del análisis de regresión. Así pues, $t=0$ se corresponde con el mes de enero de 1997. A partir de ese primer valor, la variable t experimentará variaciones unitarias, es decir, $t=1$ para febrero de 1997, $t=2$ para marzo de 1997, y así sucesivamente.

Una vez que se ha establecido ese cambio se procederá al ajuste del modelo:

$$y_t = y_t^* + e_t = a + bt + ct^2 + e_t$$

Para este modelo no lineal, el sistema de ecuaciones normales al que se llega, tras la aplicación de mínimos cuadrados, es:

$$\begin{aligned}\sum_t y_t &= Na + b \sum_t t \\ \sum_t y_t t &= a \sum_t t + b \sum_t t^2 \\ \sum_t y_t t^2 &= a \sum_t t^2 + b \sum_t t^3 + c \sum_t t^4\end{aligned}$$

La solución de este sistema de ecuaciones con los datos de la tabla 1 nos lleva a los siguientes resultados:

$$y_t = y_t^* + e_t = 2287203 - 29773,4t + 294,065t^2 + e_t$$

A su vez la bondad del ajuste medida por el coeficiente de determinación es:

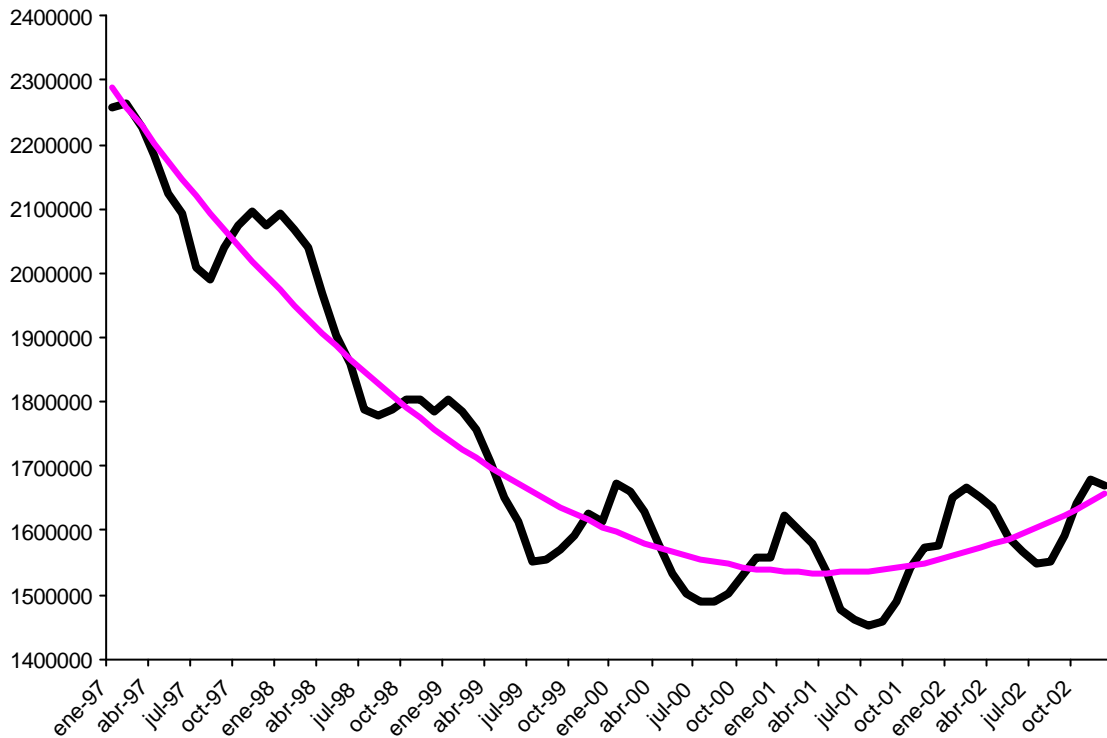
$$R^2 = 0,933$$

Todos estos resultados, así como la información previa, se pueden expresar de forma resumida en los siguientes términos:

$$\begin{aligned}Y_t^* &= 2287203 - 29773,4t + 294,065t^2 & R^2 &= 0,933 \\ &(t = 0 \text{ en enero de } 1997; \text{ unidad temporal, } 1 \text{ mes})\end{aligned}$$

En la Figura 6 se han representado la serie original del paro registrado así como la tendencia ajustada³.

³ Hay que señalar que aunque los resultados del ajuste son bastante buenos a tenor del valor del coeficiente de determinación, sin embargo, desde un punto de vista formal, es aconsejable trabajar con datos anuales, pues bajo el enfoque analítico de obtención de la tendencia, los residuos del ajuste por mínimos cuadrados recogen tres de las cuatro componentes (todas menos la tendencia), lo que hace que esos residuos no tengan el comportamiento aleatorio deseable en los mismos, pues en ellos queda recogida al menos una componente sistemática, la estacionalidad, siempre que la serie tenga esa componente. Pero en nuestro ejemplo no es del todo aconsejable trabajar con datos anuales pues la información disponible es escasa (solo seis años). Sería preferible eliminar en primer lugar la componente estacional, que es la que podría causar problemas, y, posteriormente, obtener la tendencia de la serie desestacionalizada.

Figura 6. Paro registrado y tendencia parabólica ajustada por mínimos cuadrados

Ejemplo 3. En la Tabla 4 se recoge la evolución, desde 1971, del Consumo Final para toda España, expresado en pesetas constantes de ese año. A partir de esos datos obtenga la tendencia de esa serie.

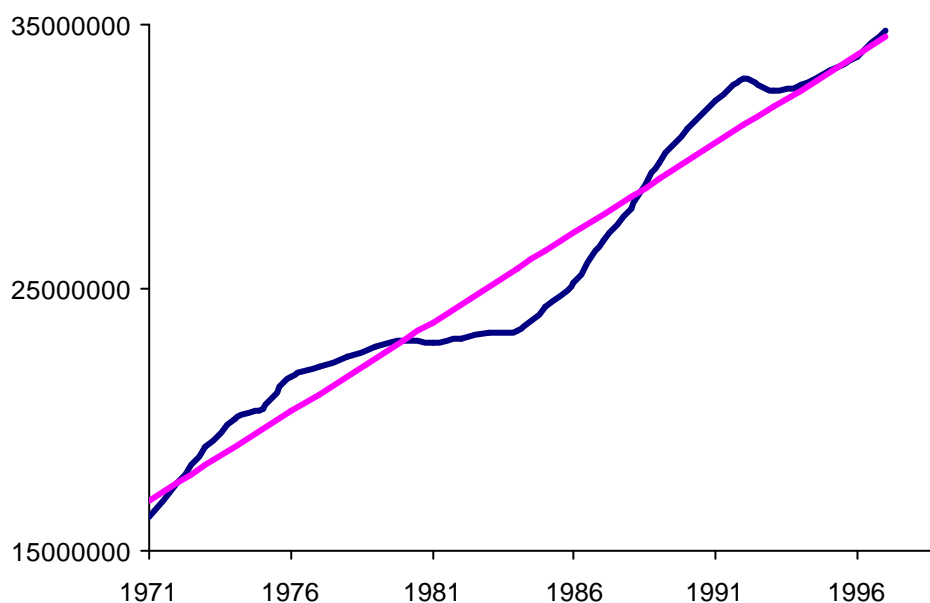
Tabla 4. Consumo Final (10⁶ pesetas de 1971)

1971	16309153	1980	23034875	1989	29779333
1972	17593135	1981	22919130	1990	31036353
1973	18932295	1982	23099009	1991	32100077
1974	20003509	1983	23308968	1992	32929885
1975	20457068	1984	23370031	1993	32514354
1976	21640271	1985	24279204	1994	32723353
1977	22039094	1986	25177951	1995	33258278
1978	22382034	1987	26781757	1996	33843015
1979	22771211	1988	28051856	1997	34770733

Fuente: Web INE.

Como paso previo a la obtención de la tendencia es aconsejable visualizar la representación gráfica de esta serie, pues ello facilita la elección del modelo que mejor se ajuste a los datos de la Tabla 4. Por tal razón en la Figura 7 se ha representado la serie del consumo final. A tenor de esta representación puede pensarse en un modelo lineal para la tendencia de esta serie. Pero aunque el modelo propuesto sea acertado en este caso, de lo que no cabe duda es que los valores anuales del consumo no están perfectamente alineados. Presentas desviaciones que se alejan de la línea de tendencia. Esas desviaciones pueden deberse tanto a la componente cíclica como las variaciones residuales. En cambio, por tratarse de una serie con datos anuales, no se verá afectada por la componente estacional.

Figura 7. Evolución del Consumo Final y tendencia lineal ajustada



En el proceso de ajuste de la tendencia a una línea recta vamos a fijar el origen del tiempo en 1971 y trabajar en términos de miles de millones de pesetas. Esto nos lleva a que:

$$\begin{aligned} \sum t &= 351 & \sum y_t &= 695105,9 & \sum t^2 &= 351 & \sum ty_t &= 10149170,3 \\ \sum y_t^2 &= 18685901587,2 \end{aligned}$$

A partir de esta información se tiene que:

$$y_t^* = 16912,97 + 679,36t \quad R^2 = 0,956$$

($t = 0$ en 1971; unidad temporal , 1 año)

Ejemplo 4. Según la Encuesta General de Medios, el número de usuarios de internet, desde 1996 a 2001, ha evolucionado como muestra la Tabla 5.

Tabla 5. Usuarios de internet en España. (Miles)

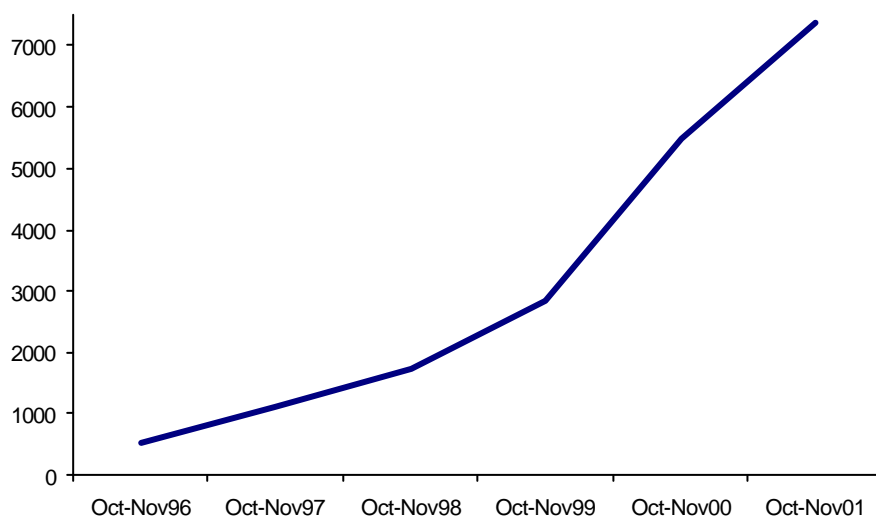
Oct-Nov96	526
Oct-Nov97	1110
Oct-Nov98	1733
Oct-Nov99	2830
Oct-Nov00	5486
Oct-Nov01	7388

Fuente: EGM

Obtener la tenencia de esta serie.

Como en los ejemplos anteriores, lo primero que debe hacerse es obtener la representación gráfica de la serie. La misma es la que aparece en la Figura 8.

Figura 8. Número de usuarios de internet en España (Miles)



A la vista de la Figura 8 no puede afirmarse que el crecimiento medio anual del número de usuarios de internet a lo largo del periodo considerado haya sido constante, por lo que no sería muy acertado modelizar la tendencia mediante una línea recta. Esta figura parece mostrar más un crecimiento exponencial. No obstante se ajustarán los dos modelos y en función de los resultados finales se optará por el mejor para representar la tendencia de esa serie.

Los cálculos necesarios para el ajuste de los dos modelos propuestos son los que aparecen en la tabla siguiente.

t	y_t	$\ln y_t$	t^2	y_t^2	ty_t	$(\ln y_t)^2$	$t \ln y_t$
0	526	6,27	0	276676	0	39,25	0
1	1110	7,01	1	1232100	1110	49,17	7,01
2	1733	7,46	4	3003289	3466	55,62	14,92
3	2830	7,95	9	8008900	8490	63,17	23,84
4	5486	8,61	16	30096196	21944	74,13	34,44
5	7388	8,91	25	54582544	36940	79,35	44,54
15	19073	46,20	55	97199705	71950	360,69	124,75

A partir de estos datos, los resultados del ajuste lineal son los siguientes:

$$y_t^* = -287,95 + 1386,71t \quad R^2 = 0,92$$

($t = 0$ en 1996; unidad temporal, 1 año)

Aunque el ajuste no es malo, sin embargo nos dice que el crecimiento medio anual ha sido de casi 1400 miles de usuarios, cuando los datos de la Tabla 5 muestran que ese crecimiento medio constante no es muy verosímil. Por esa razón se va a aprobar el ajuste con el modelo exponencial especificado en (4.7). La versión linealizada del mismo viene dada por:

$$\ln y_t^* = a + bt$$

El ajuste por mínimos cuadrados de este modelo es el siguiente:

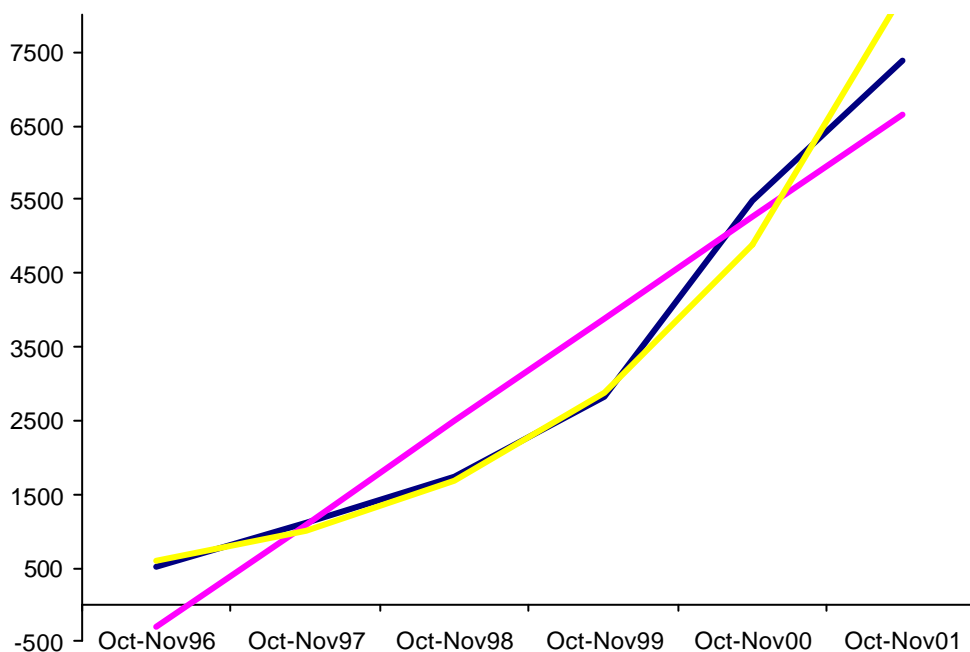
$$\ln y_t^* = 6,38 + 0,528t \quad R^2 = 0,9896$$

($t = 0$ en 1996; unidad temporal, 1 año)

Como puede observarse, estos resultados son mejores que los anteriores, pues la bondad del ajuste es superior y, además, la Figura 9 corrobora esta conclusión.

En cualquier caso hay que señalar que el periodo temporal con el que se ha trabajado es pequeño, por lo que no se puede pensar que esa tendencia observada de 1996 a 2001 se vaya a mantener de manera indefinida. Para este tipo series, la tendencia a largo plazo que mejor se ajusta es la de tipo logístico, donde se combina un crecimiento inicial de tipo exponencial con otro posterior más lento y que tiende a estancarse.

**Figura 9. Numero de usuarios de internet en España.
Series original y tendencias lineal y exponencial**



4.3.4 Alisado exponencial.

Estos procedimientos de obtención de la tendencia son muy parecidos a los de medias móviles. En ambos casos la tendencia es el resultado de promediar los valores de la serie, bien de forma ponderada o sin ponderar. En el caso de las medias móviles, los promedios son sin ponderar, salvo cuando las medias se obtienen para un número par de valores, pues entonces hay que promediar dos veces, siendo el resultado final una

media ponderada, como ya se ha visto. A estas técnicas se les podría denominar como de alisado proporcional.

Sin embargo, en los procedimientos de alisado exponencial, siempre se utilizan ponderaciones y, además, los valores suavizados o alisados que se obtienen son una combinación lineal de todas las observaciones de la serie, pero con la particularidad de que la ponderación decrece conforme nos alejamos del origen. Estos procedimientos están especialmente diseñados para la predicción.

A continuación se expondrá la idea básica del alisado exponencial, el cual está especialmente indicado para el tratamiento de series no estacionales y con una tendencia no definida, en el sentido de que ni es creciente ni decreciente.

Este procedimiento para el suavizado de una serie se basa en suponer que el nivel o valor medio de la serie en el instante t (y_t^*) se puede obtener de la forma siguiente:

$$\begin{aligned}
 y_t^* &= \alpha y_t + (1-\alpha)y_{t-1}^* = \\
 &= \alpha y_t + (1-\alpha)[\alpha y_{t-1} + (1-\alpha)y_{t-2}^*] = \\
 &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + (1-\alpha)^2 y_{t-2}^* = \\
 &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + (1-\alpha)^2 [\alpha y_{t-2} + (1-\alpha)y_{t-3}^*] = \\
 &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + (1-\alpha)^3 y_{t-3}^* = \\
 &= \dots\dots\dots = \\
 &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \dots\dots\dots + \alpha(1-\alpha)^{t-1} y_1 + (1-\alpha)^t y_0^* = \\
 &= y_{t-1}^* + \alpha(y_t - y_{t-1}^*) \qquad (0 < \alpha < 1) \qquad (4.12)
 \end{aligned}$$

Según expresión, el nivel medio en un instante t cualquiera es una combinación lineal (media ponderada) del nivel del periodo anterior y del valor observado en ese periodo, pero que a su vez, ese nivel es, también, una media ponderada de todas las observaciones de la serie (la suma de todas las ponderaciones vale la unidad). Todas las ponderaciones son una función de α , a la que se le conoce como constante de suavizado. La elección del valor de esta constante dependerá de que la serie sea más o menos estable. Así, si la misma es muy estable el valor de α deberá estar próximo a la unidad, mientras que si la serie presenta gran volatilidad, entonces es aconsejable que α sea pequeño para evitar darle demasiado peso al último dato observado. En cualquier caso, la selección del valor de α implica introducir una cierta carga de subjetividad en el análisis de la serie, lo que no deja de ser un serio inconveniente.

La única cuestión que queda pendiente para que este procedimiento de alisado sea aplicable es la forma de obtener y_0^* . Este valor se fija como: $y_0^* = y_0$.

Ejemplo 5 *Obtener la serie suavizada del número de empresas que presentaron suspensión de pagos en España mediante un alisado exponencial simple.*

En la Tabla 6 se recoge la serie original y tres alisadas según distintos valores de a . Las mismas se han representado en la Figura 10. Se ha optado por trabajar con más de un valor para la constante de suavizado. Ello permite apreciar el efecto que tiene esa constante sobre la serie alisada. Así, cuando $a=0,9$, entonces la serie original y la suavizada son prácticamente la misma. En este caso, la serie se suaviza muy poco. Ese valor de a estaría indicado cuando la serie cambia muy poco de un periodo de tiempo a otro. Cuando $a=0,1$ la serie resultante si que elimina todas las crestas que presentaban los datos originales. Pero este suavizado puede resultar excesivo en algunos casos, como es el de nuestro ejemplo, pues lo que se persigue con el alisado exponencial de una serie es eliminar de la misma solo la componente residual, mientras que en nuestro ejemplo, con $a=0,1$, se ha eliminado algo más que eso. Casi se ha eliminado también la componente cíclica, la única que aparece de forma clara en la serie original. La tercera es una solución intermedia entre las dos anteriores ($a=0,5$).

Tabla 6. Evolución del número de suspensiones de pagos de empresas en España. Series original y alisadas.

Años	Suspensión de pagos	Serie alisada ($a=0,9$)	Serie alisada ($a=0,1$)	Serie alisada ($a=0,5$)
1983	841	841	841	841
1984	814	817	838	828
1985	459	495	800	643
1986	231	257	743	437
1987	188	195	688	313
1988	154	158	635	233
1989	167	166	588	200
1990	351	333	564	276
1991	798	751	587	537
1992	1135	1097	642	836
1993	1446	1411	723	1141
1994	969	1013	747	1055
1995	650	686	738	852
1996	649	653	729	751

1997	479	496	704	615
1998	348	363	668	481
1999	290	297	630	386
2000	223	230	590	304

Fuente: Servidor web del INE y elaboración propia.

A continuación se detalla la forma en que se han obtenido esas series. Como puede apreciarse, el primer valor es idéntico para las tres e igual al de la serie original. Es decir:

$$y_{83}^* = y_{83} = 841$$

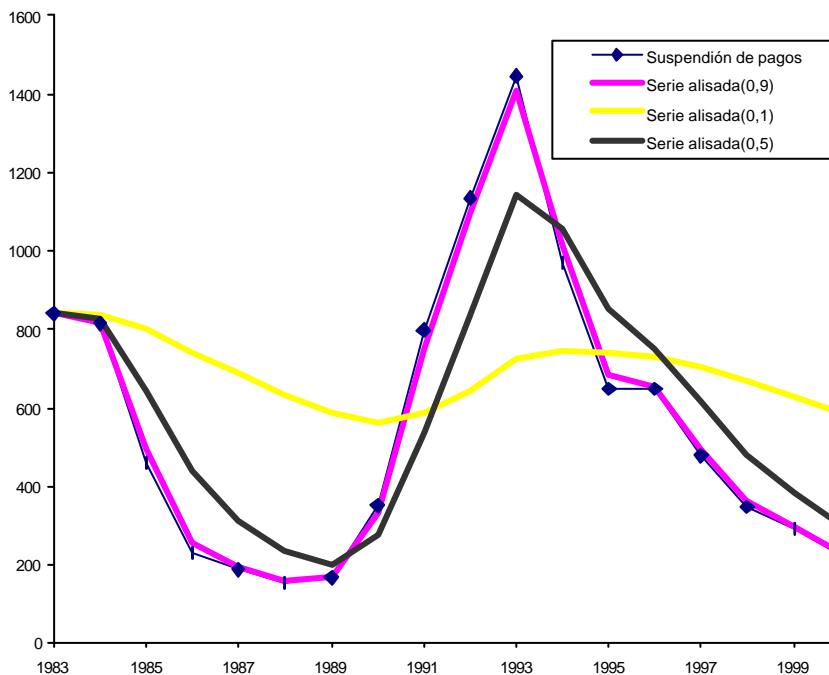
A partir de ese valor inicial, los demás se han calculado según se indicó más arriba:

$$y_{84}^* = (0,9)(y_{84}) + (1-0,9)(y_{83}^*) = (0,9)(814) + (0,1)(841) = 817$$

$$y_{85}^* = (0,9)(y_{85}) + (1-0,9)(y_{84}^*) = (0,9)(459) + (0,1)(817) = 495$$

y así para los demás años y los otros valores de a .

Figura 10. Evolución del número de suspensiones de pagos en España. Series original y alisadas exponencialmente



Cuando la serie que se pretende suavizar tiene una tendencia definida y es estacional, entonces el método de alisado exponencial simple que se acaba de exponer se sustituye por otros procedimientos que, basándose en él, permiten esas nuevas situaciones. Entre ellos está el de Holt-Winters.

4.4 Componente estacional.

Previamente ya se ha definido la componente estacional como aquellos movimientos de la serie que se repiten de forma periódica, siendo la periodicidad inferior al año. Estos movimientos de la serie, que se repiten de forma sistemática, dificultan la posibilidad de hacer comparaciones entre los valores sucesivos de una serie, pues el nivel medio de la misma se ve alterado por la estacionalidad. Para aclarar esta idea haremos uso de la variable paro registrado que se aparece en la Tabla 1. Para un año cualquiera de esta serie se puede observar como a partir de septiembre la misma empieza a tomar valores que son sistemáticamente superiores a los de meses anteriores. Este comportamiento, si se analizara para un solo año, podría hacer pensar que al serie ha cambiado de tendencia y que los niveles medios de la misma han empezado a crecer. Sin embargo ya se ha podido comprobar que ese no es el caso, pues la tendencia mostraba un perfil de continuo decrecimiento, que solo cambia a partir de la primavera del año 2001. Ese comportamiento anómalo de la tendencia, dentro de cada año, es el efecto de la estacionalidad, que hace que, para determinados meses (u otros periodos de tiempo inferiores al año), se observen movimientos ajenos a la tendencia (causados por motivos económicos, climatológicos, etc) que impiden una correcta comparación de los valores de esa serie en esos meses.

Para evitar esas distorsiones en los valores medios se recurre a lo que se conoce como **desestacionalización** de la serie o **corrección estacional**. Para realizar esta operación es necesario aislar en primer lugar la componente estacional, lo que posibilitará su posterior eliminación.

En los distintos métodos elementales que pueden utilizarse para obtener la componente estacional, siempre hay un paso previo que consiste en eliminar la tendencia, obtenida mediante alguno de los procedimientos señalados con anterioridad, medias móviles o regresión, principalmente.

Se centrará la atención en el que hace uso de la tendencia obtenida por media móviles. A este procedimiento se le conoce como **razón a las medias móviles**. La exposición de este procedimiento se realizará con el apoyo de un ejemplo. En concreto se usaran los datos de la Tabla 1, es decir, la serie de paro registrado. El primer paso a dar es suavizar la serie para de esa forma obtener la tendencia o, más bien, la componente tendencia-ciclo. Los resultados se recogen en la Tabla 3. A continuación, y admitiendo que la serie sigue un esquema multiplicativo (si fuera aditivo se procedería de una forma similar), se procede a dividir la serie original entre el resultado que se obtiene mediante aplicación de medias móviles. El resultado de ese cociente:

$$\frac{y_i}{\text{Media móvil}} = \frac{TxCxR}{TxR} = VExR \quad (4.13)$$

sería una componente mixta que recoge la estacionalidad y las variaciones residuales y se le conoce como razón a las medias móviles.

Tabla 7. Razón a las medias móviles. (VExR)

	1997	1998	1999	2000	2001	2002
Enero	,,	104,14	103,38	105,27	105,84	104,90
Febrero	,,	103,91	103,35	104,95	104,61	105,28
Marzo	,,	103,47	102,89	103,34	103,39	103,68
Abril	,,	100,97	101,07	100,54	100,56	102,33
Mayo	,,	98,77	98,54	97,83	96,76	98,84
Junio	,,	97,84	97,21	96,17	95,52	97,00
Julio	95,14	95,12	94,22	95,71	94,79	,,
Agosto	94,86	95,86	95,06	95,92	95,03	,,
Septiembre	98,04	97,72	96,63	97,10	96,60	,,
Octubre	100,42	99,78	98,62	99,21	99,47	,,
Noviembre	102,34	101,02	101,24	101,21	101,02	,,
Diciembre	102,39	101,15	101,23	101,43	100,56	,,

Como se ha señalado, el resultado (4.13) no es solo la componente estacional. También aparecen en el mismo las variaciones residuales. Si estas últimas no fueran muy importantes, entonces el contenido de la Tabla 7 nos daría directamente la estacionalidad de la serie estudiada en este ejemplo. A esos cocientes se les conoce como **índices específicos de variación estacional**. Tal es el caso del ejemplo que se viene desarrollando. En este ejemplo, además, puede observarse como el esquema

de estacionalidad cambia poco de año en año. Se trata de una estacionalidad estable, por lo que es preferible resumir todos esos índices específicos en uno general que recoja la estacionalidad de la serie. La manera más fácil de llegar al mismo es obtener un promedio de los valores de cada mes. Además, al proceder de esta forma lo que se consigue también es eliminar los posibles efectos de R, aunque estos sean poco importantes. El resultado de esta operación se recoge en la Tabla 8.

**Tabla 8. Índices generales de variación estacional (IGVE)
del paro registrado en España**

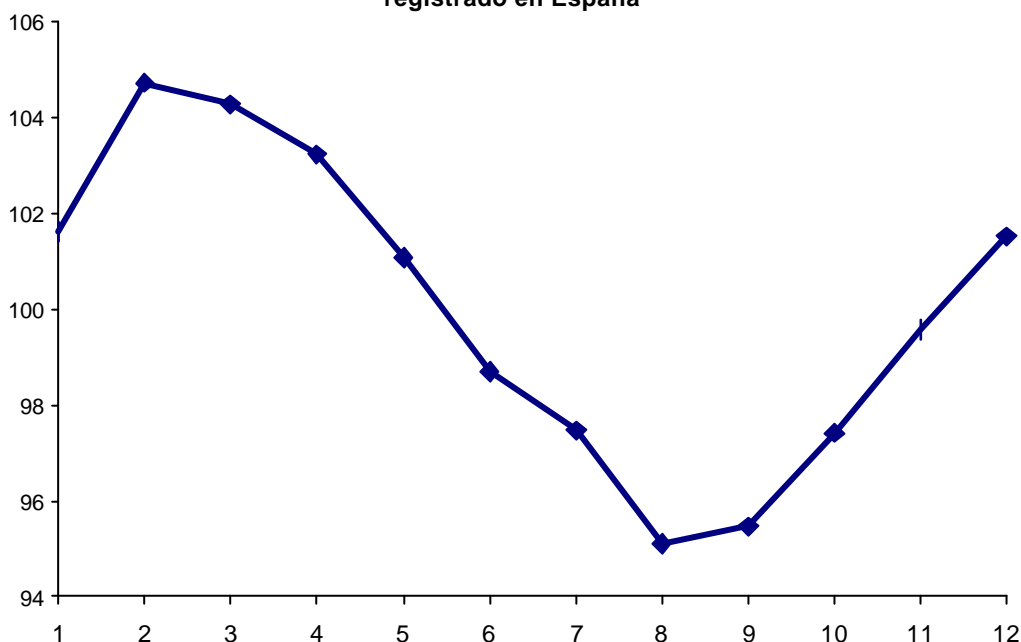
	Media	IGVE
Enero	104,71	104,86
Febrero	104,42	104,57
Marzo	103,35	103,50
Abril	101,09	101,24
Mayo	98,15	98,29
Junio	96,75	96,89
Julio	95,00	95,14
Agosto	95,35	95,49
Septiembre	97,22	97,36
Octubre	99,50	99,65
Noviembre	101,37	101,51
Diciembre	101,35	101,50

En esta tabla aparecen dos columnas, una encabezada con el nombre de media y otra con el de IGVE. La primera es el resultado de obtener la media, para cada mes, de los índices de la Tabla 7. Pero, en teoría, la media de todas esas medias debiera ser cien y, en este ejemplo, por centésimas, no lo es. Esto nos lleva a corregir estos índices generales para forzarlos a que cumplan el requisito de que su media sea cien, como ocurre en la segunda columna. La interpretación de estos índices, y para el caso de la serie de paro registrado, es que, debido a los efectos estacionales de la economía española, durante los meses de mayo a octubre el paro registrado se sitúa por debajo de la media anual, siendo durante julio y agosto cuando menos paro registrado se recoge en las Oficinas del INEM, mientras que, por el contrario, es en los meses de Enero y Febrero cuando la situación de demanda de empleo insatisfecha es mayor. Si los IGVE fueran todos igual a la unidad, entonces habría que concluir que la serie no presenta estacionalidad.

Si la estacionalidad fuera evolutiva habría que quedarse con los índices específicos. Si, además, las variaciones residuales ejercieran un efecto significativo en las razones a las medias móviles, entonces habría que eliminar esa componente. En estos casos,

como no tiene sentido calcular la media para cada mes, lo que suele hacerse es dividir y_t en tantas series independientes como periodos inferiores al años se estén considerando. Así si se trabaja con datos mensuales, el número de estas nuevas series sería doce, siendo la longitud de cada una de ellas el número de años de y_t . A continuación se suaviza cada una de esas series mediante medias móviles, con lo que se consigue eliminar la componente residual. El periodo de estas medias será de 3 a 5 años, todo ello en función de la longitud de la serie original, pues cuanto mayor es el periodo más garantía hay de haber eliminado R, pero también se pierde más información. Estos datos perdidos hay que extrapolarlos de alguna forma si lo que se persigue es desestacionalizar la serie completa.

Figura 11. Índices generales de variación estacional (IGVE) del paro registrado en España

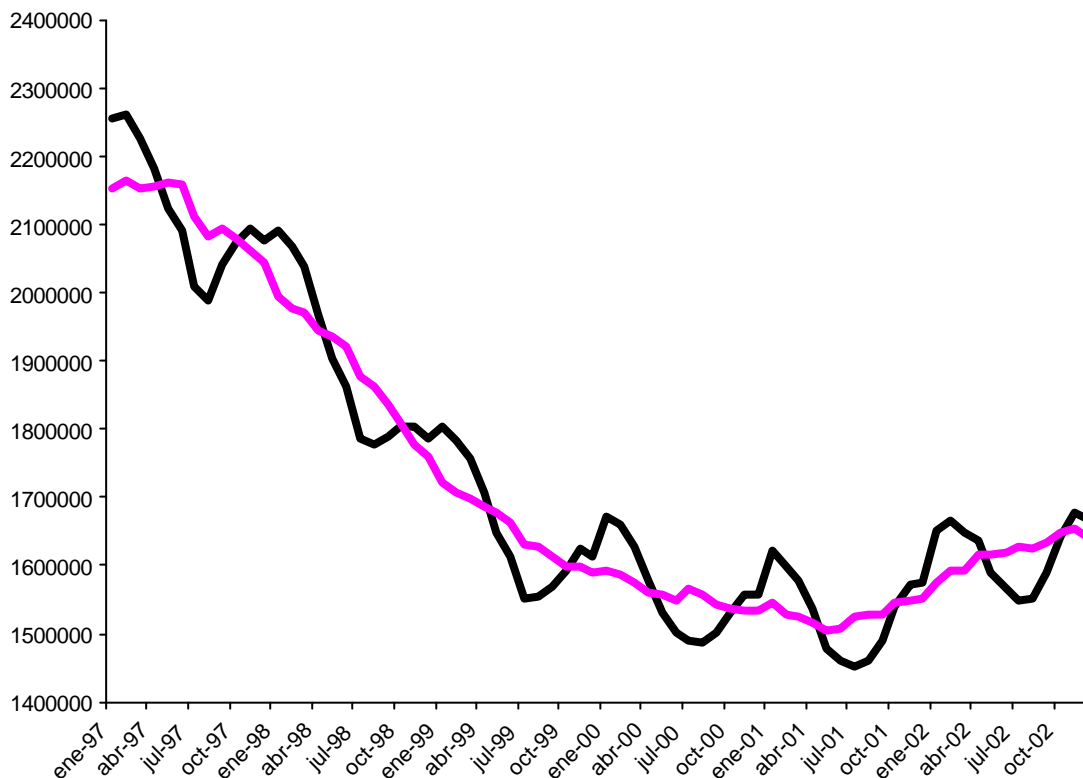


Una vez que se han obtenido los IGVE, lo que ayuda, junto con el análisis de la tendencia, al estudio descriptivo de la serie, el siguiente paso sería eliminar la estacionalidad de nuestra serie, es decir, habría que desestacionalizarla. Esto se consigue de la forma siguiente:

$$\frac{y_t}{IGVE} = \frac{TxExCxR}{VE} = TxCxR \quad (4.14)$$

Para el ejemplo del paro registrado la serie desestacionalizada aparece en la Tabla 8 a la vez que en la Figura 12.

Figura 12. Serie desestacionalizada del paro registrado.



Como puede observarse en la Figura 12, la serie desestacionalizada tiene un perfil muy parecido al de la tendencia obtenida por medias móviles, con la única diferencia que ésta última era más suave. La razón de esa diferencia radica en que en la desestacionalizada contiene la tendencia, los ciclos y las variaciones residuales, mientras que en la serie obtenida por medias móviles solo se recogen la tendencia y los ciclos.

Una vez que a la serie se le ha eliminado la componente estacional, los valores de cualquier mes si son comparables en lo que a niveles medios se refiere, pues ahora cada mes ha perdido la especificidad que el confería la estacionalidad de la serie. Se podría afirmar que los meses se han homogeneizado o estandarizado para que sean comparables entre si.

Tabla 8. **Serie de paro registrado en España desestacionalizada.**

	1997	1998	1999	2000	2001	2002
Enero	2151,9	1994,4	1720,6	1593,1	1545,6	1575,2
Febrero	2163,8	1977,4	1705,9	1587,2	1529,0	1593,2
Marzo	2152,1	1970,1	1697,7	1573,4	1525,0	1593,2
Abril	2154,9	1943,9	1687,1	1559,5	1516,3	1616,2
Mayo	2160,7	1935,2	1677,8	1557,8	1503,8	1616,6
Junio	2159,1	1920,4	1664,3	1548,3	1507,5	1617,7
Julio	2111,9	1877,3	1630,2	1564,9	1525,6	1627,6
Agosto	2083,0	1861,1	1627,9	1557,9	1528,0	1625,4
Septiembre	2095,4	1836,9	1612,6	1542,2	1528,9	1633,4
Octubre	2080,2	1810,1	1597,3	1535,6	1545,5	1647,5
Noviembre	2062,6	1777,6	1599,4	1533,6	1549,4	1652,9
Diciembre	2045,0	1759,3	1589,9	1533,4	1551,5	1643,4

Para finalizar este apartado hay que señalar que el procedimiento desarrollado para aislar la componente estacional no es el único. En lugar de trabajar con la tendencia obtenida por medias móviles se podría haber trabajado con la tendencia deducida mediante mínimos cuadrados. Pero en este caso lo más probable es que en el ajuste se recoja solo la tendencia, por lo que al dividir la serie original entre esta, el resultado contenga, además de las variaciones estacionales, los ciclos y las variaciones residuales. Pero como la periodicidad de los ciclos es poco regular, su eliminación puede influir negativamente en la obtención de la componente estacional. Por esta razón puede afirmarse que es preferible trabajar con la tendencia deducida por media móviles. De hecho, en esta idea se basa uno de los métodos de desestacionalización más ampliamente utilizado, como es el X-9 y su posterior desarrollo el X-11, elaborado por el Bureau of the Census de Estados Unidos y que aparece en la mayoría de los paquetes estadísticos.

4.5 Variaciones cíclicas y residuales.

La obtención de la tercera componente sistemática de una serie temporal es, sin lugar a dudas, la que entraña más problemas. Esto es así por tres razones como mínimo. En primer lugar no siempre existe. En segundo lugar hace falta mucha información, series muy largas, para que pueda detectarse la presencia de esta componente. En tercer lugar, es la menos sistemática de las tres, pues los ciclos, cuando existen, no siempre tienen la misma longitud y, además, se puede dar el caso de que se superpongan más de un ciclo de distintas longitudes de onda. Todo ello hace que sea

muy frecuente el tratamiento de las series en el que se prescinde del estudio separado de los ciclos y, en su lugar, se trabaja con la componente mixta ciclo-tendencia, a la que también se le conoce como extraestacional.

Pero para no dejar olvidada esta componente vamos a seguir con el enfoque multiplicativo usado hasta el momento e indicar los pasos a seguir para su obtención. Algunos de esos pasos ya se han dado en los apartados anteriores, por lo que ahora solo se hará alusión a los mismos.

En primer lugar habría que obtener la componente estacional para desestacionalizar la serie. Ello nos llevaría a:

$$\frac{y_t}{IGVE} = \frac{TxVExCxR}{VE} = TxCxR \quad (4.15)$$

En segundo lugar hay que obtener la tendencia para eliminarla de la serie:

$$\frac{TxCxR}{T} = CxR \quad (4.16)$$

Por último habría que suavizar, mediante medias móviles o por cualquier otro procedimiento de alisado, esa serie resultante, lo que daría, finalmente, la componente cíclica buscada.

La última componente son las variaciones residuales. La misma tiene poco interés y se obtendría como cociente entre (4.16) y las variaciones cíclicas. La utilidad de esta última componente se basa en poder verificar si satisface ciertos supuestos o hipótesis, como el que sea realmente aleatoria. Pero esto no está al alcance del material que se está manejando en este manual.

4.6 Medición de las variaciones de una serie.

En los apartados anteriores se ha descrito un esquema metodológico que permite analizar una serie temporal a partir del estudio de un conjunto de componentes

definidas previamente. Que duda cabe que esta es una forma de abordar el problema. Pero desde luego que no es la única. Otra posible forma de estudiar el comportamiento de una serie es apoyándose en otras, es decir, tratando de explicar sus variaciones como consecuencia de las variaciones de otra u otras series. En este sentido, y siguiendo con el ejemplo del paro registrado que se ha utilizado en el apartado anterior, se podría haber planteado que el paro es el resultado de la producción, de los salarios y de la oferta de mano de obra medida por la población activa. Este tipo de planteamiento nos llevaría a buscar una función que ligue esas variables para después cuantificarla mediante el análisis de la regresión. Estaríamos frente a una forma distinta de abordar el estudio de las variaciones de una serie. Pero tampoco es procedimiento agota todas las posibilidades de análisis.

En este epígrafe se presentará otra forma de analizar el comportamiento de las variaciones de una serie temporal. Ahora lo que se buscará será simplemente cuantificar los cambios que experimenta la serie al pasar de un periodo de tiempo a otro. Esos periodos pueden ser consecutivos (mes e a mes, trimestre a trimestre, año a año, etc.) o puede que estén separados entre si (un determinado mes-trimestre de un año con respecto a ese mismo mes-trimestre del año anterior). Si la serie temporal la seguimos representando por y_t , entonces la forma más simple de cuantificar la variación de la misma es:

$$\Delta y_t = y_t - y_{t-1} \quad (4.17)$$

Esta relación lo que nos dice, mediante el signo de la misma, es si la serie, para ese periodo considerado, está creciendo o decreciendo, según que aquel sea positivo o negativo, respectivamente. Pero aparte de esa información poco más nos dice, pues esa diferencia o variación observada en la serie viene expresada en las mismas unidades de medida que la propia serie, por lo que bastaría con realizar un cambio de escala para que podamos modificar la magnitud de esa variación de manera arbitraria. Pero las limitaciones de esta forma de medir las variaciones de una serie no se limitan a la ya mencionada. Puede darse también el caso de cambios de igual magnitud para dos variables distintas pero que no sean comparables si los niveles de esas series son muy distintos o si las mismas vienen expresadas en unidades medida distintas. Así por ejemplo no es comparable el crecimiento de diez euros en la renta de una familia que tiene un nivel de ingresos de mil euros que ese mismo crecimiento para otra familia con unos ingresos de diez mil euros.

Antes de pasar a estudiar las posibles maneras de solucionar las limitaciones que presenta esta forma tan simple de medir las variaciones de una serie, puede resultar de interés detenerse un momento e indagar un poco en el significado de esa diferencia. La misma, por si sola, es otra serie que para el caso de datos anuales y tendencia lineal recogería las componentes cíclica y residual. Se trataría de una serie filtrada de tendencia. Pero si la periodicidad fuera inferior a la anual, por ejemplo mensual, entonces, y como ya se ha señalado, las diferencias podría tomarse respecto al mes anterior o al mismo mes del año anterior. En el segundo caso el resultado es otra serie filtrada de estacionalidad y tendencia y que recoge, de nuevo, las componentes cíclica y residual. Esto es lo que se ha hecho con la serie del paro registrado y cuyo resultado aparece en la Figura 13⁴.

Esta serie resultante muestra que desde mediados de 1999 empezó a tener lugar no un cambio de tendencia, como se ha indicado antes, sino más bien una fase distinta del ciclo, en la que el decrecimiento de la misma es cada vez menor. Ese decrecimiento se fue anulando de manera que para octubre de 2001 las variaciones interanuales eran ya positivas. Vemos como esta forma de tratar la serie puede complementar el instrumental analítico presentado en los epígrafes anteriores.

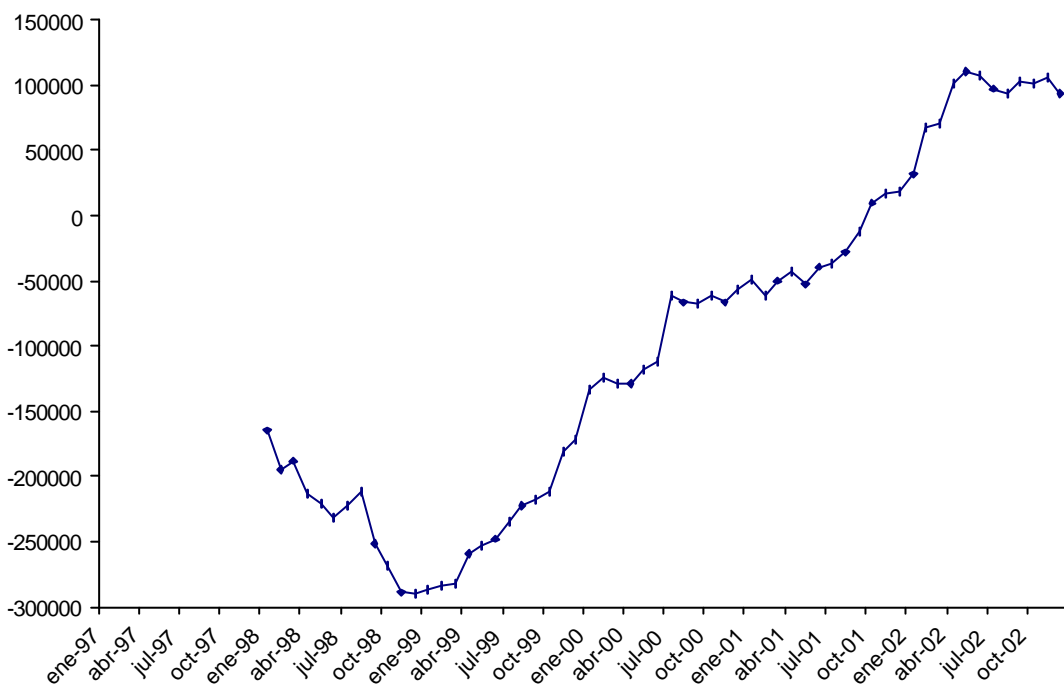
Pero la cuantificación de las variaciones de una serie en términos absolutos ya hemos visto que presenta ciertas limitaciones que es conveniente corregir. La forma más simple de eliminar esos problemas asociados a las unidades de medida de la serie es trabajar en términos relativos, es decir, medir las variaciones de forma adimensional. Esto se consigue cuando se trabaja con **tasas de variación**. Estas se obtienen comparando la variación intertemporal de la serie con respecto al valor inicial de la misma. De acuerdo con esta definición, las tasas vendrían dadas por:

$$T(h, n) = T_h^n = \frac{\Delta y_t}{y_{t-h}} = \frac{y_t - y_{t-h}}{y_{t-h}} = \frac{y_t}{y_{t-h}} - 1 \quad (4.18)$$

⁴ Hay que indicar que esta forma de eliminar la componente tendencia de una serie solo da buenos resultados si la misma es lineal, pues si las variaciones fueran de tipo exponencial, entonces, la diferenciación sucesiva de la serie no garantiza que se elimine completamente la tendencia.

donde h indica el número de periodos que hay entre los valores comparados y n es el número de observaciones que se utilizan para el cálculo de la tasa. Las tasas se pueden expresar en tantos por uno, aunque lo más habitual es que se multipliquen por cien, o cualquier otra potencia de diez, en cuyo caso se hablaría de porcentajes o lo que corresponda. Tanto por adimensionalidad como por hecho de que se está comparando variaciones absolutas con niveles de la variable, hace que esas tasas sean comparables con las obtenidas para otro tipo de variables o con la misma variable para periodos de tiempo muy alejados entre si.

Figura 13. Serie filtrada de tendencia y estacionalidad



A partir de (4.18) se deduce que:

$$y_t = (T(h, n) + 1)y_{t-h} \quad (4.19)$$

En función de los valores de h y n , las tasas más habituales que suelen calcularse son las siguientes:

a)

$$T_1^1 = \left(\frac{y_t}{y_{t-1}} - 1 \right) \times 100 \quad (4.20)$$

Esta expresión se utilizaría para calcular las tasas de variación para series con datos anuales. También podría utilizarse para series cuya periodicidad sea inferior al año (mensual, trimestral, etc.). Pero en estos casos, si la serie presenta estacionalidad, no es aconsejable hacer uso de la misma para calcular las tasas de variación, pues ese factor distorsiona el valor de las mismas.

b)

$$T_{12}^1 = \left(\frac{y_t}{y_{t-12}} - 1 \right) \times 100 \quad (4.21)$$

Para datos mensuales, esto es una tasa de variación anual en un mes concreto y que se puede obtener en meses sucesivos (tasa interanual). La misma puede utilizarse para datos con y sin estacionalidad, pues los periodos que se comparan son homogéneos, el mismo mes de cada año. Es aplicable tanto a variables flujos (producción, ventas, etc.) como stock (paro registrado, empleo de una empresa, etc).

c)

$$T_6^1 = \left(\frac{y_t}{y_{t-6}} - 1 \right) \times 100 \quad (4.22)$$

Si los datos fueran mensuales se trataría de tasas semestrales en un mes concreto y que se pueden obtener en meses sucesivos. No debe utilizarse cuando la serie presente estacionalidad, pues los meses que se comparan no son homogéneos (diciembre-junio, noviembre-mayo, etc). Como la anterior, es aplicable tanto a variables flujo como de naturaleza stock.

d)

$$T_4^1 = \left(\frac{y_t}{y_{t-4}} - 1 \right) \times 100 \quad (4.23)$$

Para datos mensuales da tasas de variación cuatrimestral en un mes concreto y se pueden obtener en meses sucesivos. A este tipo de tasas le son aplicables los comentarios hechos a las semestrales en lo referido a la estacionalidad. Si la serie fuera trimestral, esta expresión nos daría las tasas de variación anual en un trimestre concreto y además ahora si que sería aplicable cuando hay estacionalidad (primer trimestre de t-primer trimestre de t-1, etc.).

e)

$$T_3^1 = \left(\frac{y_t}{y_{t-3}} - 1 \right) \times 100 \quad (4.24)$$

Para datos mensuales da tasas de variación trimestral en un mes concreto y se pueden obtener en meses sucesivos. A este tipo de tasas le son aplicables los comentarios hechos a las semestrales y cuatrimestrales en lo referido a la estacionalidad. Si la serie fuera cuatrimestral, esta expresión nos daría las tasas de variación anual en un cuatrimestre concreto y además ahora si que sería aplicable cuando hay estacionalidad (primer cuatrimestre de t-primer cuatrimestre de t-1, etc.)

f)

$$T_{12}^{12} = \left(\frac{y_t + y_{t-1} + \dots + y_{t-11}}{y_{t-12} + y_{t-13} + \dots + y_{t-23}} - 1 \right) \times 100 \quad (4.25)$$

Cuando los datos son mensuales, ahora se estaría obteniendo tasas de variación anual. Esta expresión solo es aplicable a variables de tipo flujo con o sin estacionalidad.

g)

$$T_3^3 = \left(\frac{y_t + y_{t-1} + y_{t-2}}{y_{t-3} + y_{t-4} + y_{t-5}} - 1 \right) \times 100 \quad (4.26)$$

Si los datos son mensuales, con esta expresión se obtienen tasas de variación trimestral para variables de tipo flujo. No es muy aconsejable utilizarla si la serie presenta estacionalidad.

h)

$$T_1^{12} = \left(\frac{y_t + y_{t-1} + \dots + y_{t-11}}{y_{t-1} + y_{t-2} + \dots + y_{t-12}} - 1 \right) \times 100 \quad (4.27)$$

Esta expresión, con datos mensuales, lo que da son tasas mensuales basadas en medias móviles anuales.

i)

$$T_1^3 = \left(\frac{y_t + y_{t-1} + y_{t-2}}{y_{t-1} + y_{t-2} + y_{t-3}} - 1 \right) \times 100 \quad (4.28)$$

Esta expresión, con datos mensuales, lo que da son tasas mensuales basadas en medias móviles trimestrales.

Para finalizar este epígrafe hay que hacer referencia a dos cuestiones que tienen un gran interés desde el punto de vista práctico. Se trata, en primer lugar, de la conversión (elevación) de tasas con periodicidad inferior al año a tasas anuales y, en segundo lugar, de la obtención de las tasas medias mensuales, o de otro periodo, en un año.

Por lo que respecta al primer punto, si admitimos inicialmente que se trabaja con tasas mensuales, entonces la tasa mensual elevada a anual sería aquella que representara el mismo crecimiento absoluto que la anual. Bajo este supuesto se tiene que, cuando se trabaja con tasas mensuales (que representaremos por T_i , $i = 1, 2, \dots, 12$), el valor de la serie en el último mes vendrá dado por:

$$y_{12} = y_0(1+T_1)(1+T_2)\dots(1+T_{12}) \quad (4.29)$$

pues:

$$\begin{aligned} y_1 &= y_0(1+T_1) \\ y_2 &= y_1(1+T_2) \\ &\vdots \\ y_{12} &= y_{11}(1+T_{12}) \end{aligned}$$

Por otro lado, según la tasa anual (T_a) se tiene que:

$$y_{12} = y_0(1+T_a) \quad (4.30)$$

Si ahora se iguala (4.29) y (4.30) el resultado es que:

$$T_a = (1+T_1)(1+T_2)\dots(1+T_{12}) - 1 \quad (4.31)$$

Si además se supone que las tasas mensuales son constantes e iguales a T_m , entonces:

$$T_a = (1+T_m)^{12} - 1 \quad (4.32)$$

En general esa tasa anualizada vendría dada por:

$$T_a = \left(1+T_{12/h}\right)^{12/h} - 1 \quad (4.33)$$

La segunda cuestión planteada era la de obtener una tasa media mensual a partir de las doce tasas. Esa tasa, que la representaremos por TM , se aquella que satisfaga la relación:

$$(1+TM)^{12} = [(1+T_1)(1+T_2)\dots(1+T_{12})] \quad (4.34)$$

De donde se tiene que:

$$TM = \sqrt[12]{[(1+T_1)(1+T_2)\dots(1+T_{12})]} - 1 \quad (4.35)$$

Es decir, la tasa media mensual no es la media aritmética de las tasas mensuales, sino la media geométrica de los factores de variación unitaria de cada mes $(1+T_i)$ menos la unidad.

Ejemplo 6 La fabricación mensual de vehículos (turismos) en España durante los años 1998, 1999 y 2000 fue la que se recoge en la tabla siguiente:

	1998	1999	2000
Enero	182630	191620	186070
Febrero	191620	214600	225570
Marzo	213540	230990	247000
Abril	196140	228600	195030
Mayo	205900	221610	252360
Junio	221080	220660	226210
Julio	200220	211590	217540
Agosto	60190	40070	60030
Septiembre	207430	212930	202780
Octubre	217990	198690	214570
Noviembre	210580	161350	238970
Diciembre	164530	211950	156810
Total	2271850	2344660	2422940

Fuente: Página web del INE

Analizar la variación de esta serie en términos de:

- Tasas mensuales de variación
- Tasas anuales de variación
- Tasas anuales mes a mes

a) Para obtener las tasas mensuales de variación habría que aplicar (4.20), lo que da como resultado:

	1998	1999	2000
Enero	,,	16,47	-12,21
Febrero	4,92	11,99	21,23
Marzo	11,44	7,64	9,50
Abril	-8,15	-1,03	-21,04
Mayo	4,98	-3,06	29,40
Junio	7,37	-0,43	-10,36
Julio	-9,44	-4,11	-3,83
Agosto	-69,94	-81,06	-72,41

Septiembre	244,63	431,40	237,80
Octubre	5,09	-6,69	5,81
Noviembre	-3,40	-18,79	11,37
Diciembre	-21,87	31,36	-34,38

Donde la tasa para febrero de 1998 es el resultado de:

$$4,92 = \left(\frac{191620}{182630} - 1 \right) \times 100$$

y así para los demás meses.

Como puede apreciarse, estas tasas se ven afectadas por todas las componentes de la serie. Así, los valores de agosto y septiembre son un claro exponente de la estacionalidad de la serie, mientras que la componente errática o residual se manifiesta en distintos meses de esta serie.

b) En este caso, como la serie es un flujo, para obtener las tasas de variación anual lo primero que debe hacerse es sumar la producción de los doce meses, lo que nos daría la producción anual para cada uno de los tres años considerados y, a partir de esos datos, se obtendrían las tasas correspondientes. Las mismas son el resultado de aplicar, de nuevo, la expresión (4.20):

$$3,20 = \left(\frac{2344660}{2271850} - 1 \right) \times 100 \qquad 3,34 = \left(\frac{2422940}{2344660} - 1 \right) \times 100$$

Es decir, que la producción de 1999 fue superior a la de 1998 en un 3,20% y la de 2000 superó a la de 1999 en un 3,34%.

A partir de las tasas de variación mensual del apartado anterior, también se puede calcular la tasa mensual elevada a anual. Para ello habría que obtener la tasa media mensual para cada año aplicando (4.35):

$$TM = \sqrt[12]{[(1+T_1)(1+T_2)\dots(1+T_{12})]} - 1 = \sqrt[12]{\left[\left(\frac{\text{Enero}_t}{\text{Diciembre}_{t-1}}\right)\left(\frac{\text{Febrero}_t}{\text{Enero}_t}\right)\dots\left(\frac{\text{Diciembre}_t}{\text{Noviembre}_t}\right)\right]} - 1 =$$

$$= \sqrt[12]{\left[\left(\frac{\text{Diciembre}_t}{\text{Diciembre}_{t-1}}\right)\right]} - 1$$

$$\text{pues } (1+T) = \left[1 + \left(\frac{y_t}{y_{t-1}} - 1\right)\right] = \left(\frac{y_t}{y_{t-1}}\right)$$

Todo esto lleva a que esa tasa, para 1999 sea:

$$T_m(1999) = \sqrt[12]{\left[\left(\frac{\text{Diciembre}_t}{\text{Diciembre}_{t-1}}\right)\right]} - 1 = \sqrt[12]{\left[\left(\frac{211950}{164530}\right)\right]} - 1 = 0,0213$$

y para 2000

$$T_m(2000) = \sqrt[12]{\left[\left(\frac{\text{Diciembre}_t}{\text{Diciembre}_{t-1}}\right)\right]} - 1 = \sqrt[12]{\left[\left(\frac{156810}{211950}\right)\right]} - 1 = -0,0248$$

Una vez que han obtenido esas tasas medias mensuales, para elevarlas a anuales hay hacer uso de (4.32). El resultado de ello es:

$$T_a(1999) = (1 + 0.0213)^{12} - 1 = 0,2882$$

$$T_a(2000) = (1 - 0.0248)^{12} - 1 = -0,26$$

Estos resultados, si se comparan con los obtenidos antes en este mismo apartado, merecen ser comentados. En primer lugar hay que señalar que las tasas medias mensuales elevadas a anuales no son otra cosa que las interanuales que se dan en el siguiente apartado. En segundo lugar, estas tasas se obtienen a partir de la información que suministran solo dos meses, sin tener en cuenta, para nada, lo ocurrido en los demás, de forma que si en esos periodos la componente errática (por ejemplo, una huelga) es decisiva, entonces condiciona la tasa de todo el año, cuando eso no tiene razón de ser.

Esto lleva a que este procedimiento de obtener las tasas anuales deba usarse con cautela y, desde luego, no es aconsejable cuando se trabaja con variables flujo. Estarían más indicadas en los casos de magnitudes de tipo stock, donde un mes cualquiera puede ser “representativo” de los demás, pues como se sabe, el valor del año no es la suma de los valores de todos los meses. También se pueden utilizar como un método de proyectar la variación de todo un año, cuando solo se conoce lo ocurrido en los primeros meses y se asume que para los demás la tendencia en la variación será similar a la observada en esos periodos iniciales.

Para finalizar, hay que indicar que las tasas próximas al 3%, obtenidas en primer lugar, reflejan de forma más fidedigna lo ocurrido en esos dos años que lo que sugieren las mensuales elevadas a anuales.

c) *Las tasas anuales mes a mes (tasas interanuales) se obtienen por la aplicación de (4.21) dando como resultado:*

	1998	1999	2000
Enero	„	4,92	-2,90
Febrero	„	11,99	5,11
Marzo	„	8,17	6,93
Abril	„	16,55	-14,69
Mayo	„	7,63	13,88
Junio	„	-0,19	2,52
Julio	„	5,68	2,81
Agosto	„	-33,43	49,81
Septiembre	„	2,65	-4,77
Octubre	„	-8,85	7,99
Noviembre	„	-23,38	48,11
Diciembre	„	28,82	-26,02

En este caso se pierden los datos de 1998 al no disponerse de los correspondientes a 1997, pues las tasas se obtienen comparando el valor de un mes para un año con el del mismo mes el año anterior. Es decir, siempre se perderá un año.

Estas tasas no reflejan la estacionalidad de la serie, siempre que la misma sea estable. En cambio, si que recogen las otras tres componentes. Es como si se trabajara con una serie desestacionalizada.