Tema 3: Series Temporales

Departamento Matemática Aplicada

Universidad de Málaga

Curso 2020-2021

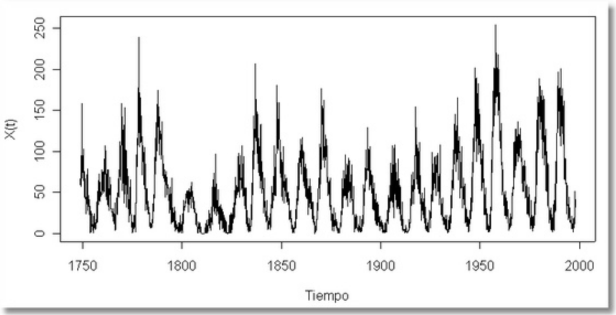
EXAMEN:

*IPC, DESESTACIONALIZACION, COMPONENTE ESTACIONAL

Series Temporales (Time series)

Una serie temporal es un conjunto de observaciones tomadas a intervalos regulares y puede considerarse como una variable bidimensional, donde una de las variables es el tiempo t (variable independiente), y la otra el fenómeno cuantitativo que se desea estudiar X (variable dependiente).

Representación gráfica: Representamos la nube de puntos X_t:



Componentes de las Series Temporales

Cada uno de los valores observados puede considerarse como la conjunción de 4 factores que lo determina:

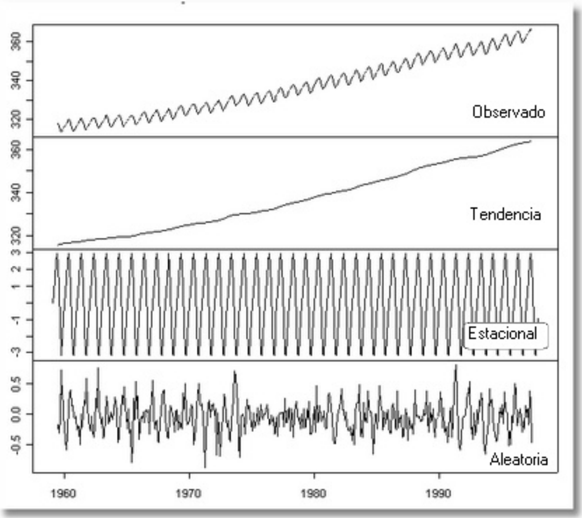
- **Tendencia secular (T):** Es la dirección dominante de la serie observada.
- **Componente estacional (E):** La mayoría de series temporales están afectadas por variaciones periódicas (semana, mes, año...).
- **Componente estacional segunda o cíclica (C):** Muchas variables presentan otros períodos de cambios adicionales. A veces de período largo (varios años).
- **Componente aleatoria (A):** Variaciones aleatorias sobre el modelo.

Hipótesis aditivas y multiplicativas

Podemos modelizar una serie temporal bajo diversos modelos

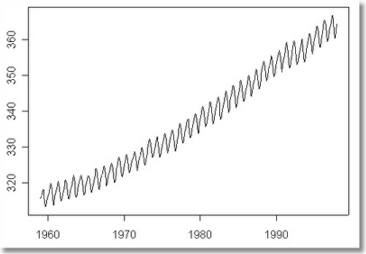
- **Hipótesis Multiplicativa:** $\mathbf{X} = \mathbf{T} \cdot \mathbf{E} \cdot \mathbf{C} \cdot \mathbf{A}$. Lo más frecuente es partir de una hipótesis de este tipo y que los cambios estacionales sean proporcionales a los cambios de tendencia.
- **Hipótesis Aditiva:** $\mathbf{X} = \mathbf{T} + \mathbf{E} + \mathbf{C} + \mathbf{A}$. A veces, los cambios estacionales no son proporcionales a los cambios de tendencia, sino que son cantidades fijas.
- **Hipótesis Mixta:** $\mathbf{X} = \mathbf{T} \cdot \mathbf{E} \cdot \mathbf{C} + \mathbf{A}$. Una variación sobre la hipótesis multiplicativa, en la que sumamos la componente aleatoria.

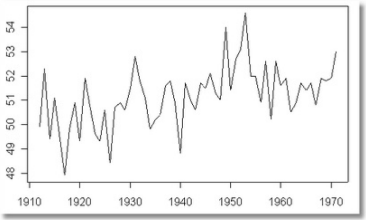
Descomposición de una serie



Tendencia

Marca la dirección dominante.





En ambas gráficas se observa una tendencia creciente, más marcada en la de la izquierda.
Para su cálculo veremos dos métodos:

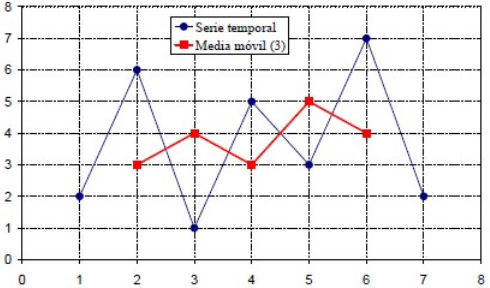
- **Medias móviles**.
- **Mínimos cuadrados** (lo veremos una vez desestacionalizada la serie).

Medias móviles

Definición

Dada una sucesión de valores X_t, llamamos **media móvil de orden k** *a la sucesión $Y_t = (\hat{X}_k)_t$ obtenida a partir de la primera:*

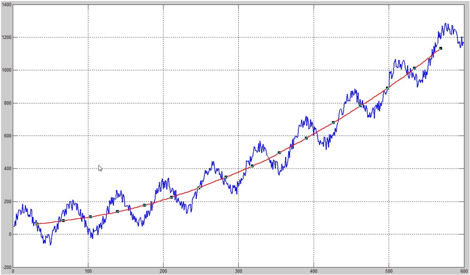
$$Y_1 = \frac{X_1+X_2+...+X_k}{k},\ Y_2 = \frac{X_2+X_3+...+X_{k+1}}{k},\ Y_3 = \frac{X_3+X_4+...+X_{k+2}}{k},\ \ldots$$



Media móvil de orden 3

Representación gráfica

Como se observa en la figura anterior la serie de medias móviles da lugar a una serie temporal más suave que la anterior. Si el periodo de la media móvil se hace coincidir exactamente con el periodo de cierta fluctuación sistemática, esta fluctuación queda eliminada en la serie resultante al aplicar la media móvil, tal y como se muestra en la figura.



Observaciones

- Si el orden es impar, la media móvil queda centrada pues su valor se asigna al dato que ocupa la posición central, es decir, si usamos medias de orden 3, la primera media móvil la consideramos en t=2, si son de orden 5 en t=3 y así consecutivamente.
- Si el orden es par, la media móvil queda descentrada pues no hay ningún valor del conjunto que ocupe la posición central. Por tanto será necesario centrarla o corregirla. Para ello, se vuelve a calcular una media móvil, esta vez de orden dos o equivalentemente multiplicar por 0.5 las muestras de los extremos, para que se sume un número par de muestras centrado.

Ejemplo

Ejemplo

Dada la serie de valores:

$$X = \{2.6767, 3.2050, 4.8714, 4.6275, 5.9008, 5.5200, 5.1186, 4.9507,\ 4.6612,\ 6.2572,\ \dots\}$$

Hallar:

- *Los 4 primeros términos de la sucesión de medias móviles de orden 3 asociada.*
- *Los 4 primeros términos de la sucesión de medias móviles de orden 4 asociada.*

Ejemplo-Solución

Solución:

- Primeros cuatro términos de la sucesión de medias móviles de orden 3:

$$y_1 = \frac{2.6767 + 3.2050 + 4.8714}{3} = 3.5844$$

$$y_2 = \frac{3.2050 + 4.8714 + 4.6275}{3} = 4.2346$$

$$y_3 = \frac{4.8714 + 4.6275 + 5.9008}{3} = 5.1332$$

$$y_4 = \frac{4.6275 + 5.9008 + 5.1186}{3} = 5.3494$$

Ejemplo-Solución-Continuación

Para hallar los primeros términos de la sucesión de medias móviles de orden 4 hay que tener en cuenta que el orden es par. Al ser una media de orden par queda descentrada y para corregirla hay que volver a calcular la media entre dos valores consecutivos o multiplicar por 0.5 los extremos, es decir:

$$y_1 = \frac{x_1 + x_2 + x_3 + x_4}{4}; \qquad y_2 = \frac{x_2 + x_3 + x_4 + x_5}{4}$$

$$z_1 = \frac{y_1 + y_2}{2} = \frac{1}{2}\left(\frac{x_1 + x_2 + x_3 + x_4 + x_2 + x_3 + x_4 + x_5}{4}\right)$$

$$z_1 = \frac{1}{2}\left(\frac{x_1 + 2x_2 + 2x_3 + 2x_4 + x_5}{4}\right)$$

$$z_1 = \frac{1}{2}\cdot\left(\frac{x_1}{4}\right) + \frac{x_2 + x_3 + x_4}{4} + \frac{1}{2}\cdot\left(\frac{x_5}{4}\right)$$

Habría que seguir el mismo procedimiento para hallar el resto de los términos.

Ejemplo-Solución-Continuación

- Por tanto, los primeros cuatro términos de la sucesión de medias móviles de orden 4 serán:

$$z_1 = \frac{0.5 \cdot 2.6767 + 3.2050 + 4.8714 + 4.6275 + 0.5 \cdot 5.9008}{4} = 4.2482$$

$$z_2 = \frac{0.5 \cdot 3.2050 + 4.8714 + 4.6275 + 5.9008 + 0.5 \cdot 5.5200}{4} = 4.9406$$

$$z_3 = \frac{0.5 \cdot 4.8714 + 4.6275 + 5.9008 + 5.5200 + 0.5 \cdot 5.1186}{4} = 5.2608$$

$$z_4 = \frac{0.5 \cdot 4.6275 + 5.9008 + 5.5200 + 5.1186 + 0.5 \cdot 4.9507}{4} = 5.3321$$

Ambas suponen una suavización de la serie inicial.

Cálculo de la tendencia mediante medias móviles

- La tendencia es una suavización de los valores que evita la dependencia de estacionalidad y factores accidentales o aleaotrios.
- Podemos utilizar medias móviles con este fin ya que con este procedimiento podemos eliminar las variaciones estacionales y las aleatorias, obteniendo por tanto la tendencia secular.

Ejemplo: Orden impar

Ejemplo

Dada la serie temporal:

t	1959	1960	1961	1962	1963	1964	1965
X	6	6.2	5.1	4.9	5.2	5.8	7

Hallar la tendencia usando medias móviles de orden 3 y orden 4.

Solución: Para k=3, el orden es impar y sumamos una cantidad entera de datos:

t	1959	1960	1961	1962	1963	1964	1965
X	6	6.2	5.1	4.9	5.2	5.8	7
$T = \hat{X}_3$	–	5.7667	5.4000	5.0667	5.3000	6.0000	–

Ejemplo: Orden par

Para orden 4, al ser par necesitamos multiplicar por 0.5 los extremos:

t	X	$T = \hat{X}_4$
1959	6.0	–
1960	6.2	–
1961	5.1	5.45
1962	4.9	5.30
1963	5.2	5.49
1964	5.8	–
1965	7.0	–

Cálculos: $\frac{0.5 \cdot 6 + 6.2 + 5.1 + 4.9 + 0.5 \cdot 5.2}{4} = 5.45,$

$\frac{0.5 \cdot 6.2 + 5.1 + 4.9 + 5.2 + 0.5 \cdot 5.8}{4} = 5.30, \ldots$

Comentarios

- Las medias móviles constituyen un potente método para suavizar la serie temporal, estimando conjuntamente tendencia y ciclos de periodo largo. De esta forma se eliminan las variaciones estacionales y las accidentales.
- Dada la dificultad para separar la componente cíclica y la tendencia, debido a que para ello necesitaríamos un número muy elevado de observaciones, llamaremos a esta componente tendencia-ciclo.
- Tienen la ventaja de no presuponer la forma de la función tendencia.

Comentarios

- Tienen el inconveniente de que se pierden valores extremos. Cuánto mayor sea el orden, más valores se pierden aunque se suaviza más, por lo que habrá que mantener un cierto equilibrio.
- Si conocemos el periodo de los factores estacionales conviene tomar como medias móviles dicho valor, esto es, si las medidas se corresponden con meses tomar k=12, si con trimestres tomar k=4 (4 medidas en el año), etc.

Estacionalidad

Una vez estimada la tendencia-ciclo, pasamos a estudiar los factores estacionales mediante la eliminación previa de las otras componentes. Para ello determinaremos unas medidas de la variación estacional: los índices de variación estacional.
Lo veremos siguiendo un ejemplo:

X	Primavera	Verano	Otoño	Invierno
1980	178.2	153.2	185.9	163.6
1981	196.3	156.9	**197.9**	166.4
1982	197.3	159.7	202.6	175.6
1983	209.5	169.5	202.4	179.8
1984	200.0	168.6	216.1	178.5

Lo primero será eliminar el factor 'Tendencia-Ciclo' de la serie original X.

Cálculo de la tendencia

Usamos medias móviles orden 4 para calcular la tendencia-ciclo:

$T \cdot C$	Primavera	Verano	Otoño	Invierno
1980	–	–	172.4875	175.2125
1981	177.1750	179.0250	**179.5000**	179.9750
1982	180.9125	182.6500	185.3250	188.0750
1983	189.2750	189.7750	189.1125	187.8125
1984	189.4125	190.9625	–	–

Al suavizar mediante medias móviles, realmente se calcula el conjunto de efectos de la tendencia y ciclos largos $T \cdot C$, que no se manifiestan para orden pequeño y se elimina la estacionalidad y la accidentalidad .

Para eliminar sus efectos, dividimos la tabla X por $T \cdot C$:

$$\frac{X}{T \cdot C} = \frac{T \cdot E \cdot C \cdot A}{T \cdot C} = E \cdot A$$

Así, quedan solo las componentes estacional y aleatoria.

Eliminación de la tendencia

Si nuestro modelo fuera aditivo, lo haríamos restando la componente tendencial en lugar de dividiendo.

$E \cdot A = X/T \cdot C$	Primavera	Verano	Otoño	Invierno
1980	–	–	1.0778	0.9337
1981	1.1079	0.8764	**1.1025**	0.9246
1982	1.0906	0.8743	1.0932	0.9337
1983	1.1069	0.8932	1.0703	0.9573
1984	1.0559	0.8829	–	–

Por ejemplo, el valor 1.1025 obtenido para Otoño de 1981 resulta de dividir 197.9 (valor para Otoño de 1981 en X) por 179.5 (valor para esa fecha en T). Igual para el resto.

Cálculo índices de estacionalidad

Para eliminar la componente accidental, se calcula el índice de cada estación, haciendo la media de los valores correspondientes en la tabla $X/T \cdot C$:

$$I_P = \frac{1.1079 + 1.0906 + 1.1069 + 1.0559}{4} = 1.0903 \quad \Rightarrow 109.03\,\%$$

$$I_V = \frac{0.8764 + 0.8743 + 0.8932 + 0.8829}{4} = 0.8817 \quad \Rightarrow 88.17\,\%$$

$$I_O = \frac{1.0778 + 1.1025 + 1.0932 + 1.0703}{4} = 1.0859 \quad \Rightarrow 108.59\,\%$$

$$I_I = \frac{0.9337 + 0.9246 + 0.9337 + 0.9573}{4} = 0.9373 \quad \Rightarrow 93.73\,\%$$

Los valores así obtenidos representan la variación estacional. El índice de primavera 109.03 % se interpreta como que en esa estación el valor es un 9.03 % superior a la media anual, mientras que el de verano índica una disminución del 11.83 %.

Ejemplo-Continuación

La media de los índices de variación debe ser uno o 100 si se mide en porcentajes. En el ejemplo se tiene que:
$I_P + I_V + I_O + I_I = 399.53\,\%$ y deberían sumar $400\,\%$, por lo que decimos que son índices "sin corregir".

Para corregirlos debemos calcular:

$$I'_P = \frac{400}{399.53} \cdot I_P = 109.1606, \quad I'_V = \frac{400}{399.53} \cdot I_V = 88.2746,$$

$$I'_O = \frac{400}{399.53} \cdot I_O = 108.7217, \quad I'_I = \frac{400}{399.53} \cdot I_I = 93.8432$$

Finalmente, I'_P, I'_V, I'_O, I'_I son los índices estacionales corregidos, normalizados o ajustados.

Desestacionalización

La componente estacional tiene interés por sí misma, pues nos permite conocer la evolución a corto plazo de la serie temporal. Pero además, es interesante llegar al conocimiento de la serie temporal una vez eliminadas las variaciones estacionales, y este proceso se denomina desestacionalización. Ésto se consigue dividiendo cada valor en ella (serie inicial) por el índice correspondiente:

X/E	Primavera	Verano	Otoño	Invierno
1980	163.2458	173.5494	170.9871	174.3334
1981	179.8268	177.7409	182.0244	177.3171
1982	180.7429	180.9128	186.3474	187.1207
1983	191.9191	192.0145	186.1634	191.5963
1984	183.2163	190.9950	198.7644	190.2110

Tendencia real

Una de las aplicaciones de la desestacionalización es el cálculo de la tendencia real. La eliminación de la componente estacional se utiliza para recalcular la tendencia y obtener una mejor aproximación de la trayectoria real de la serie.

Ejemplo-Tendencia real

Para calcular la tendencia real del ejemplo, recalculamos la tendencia a los datos desestacionalizados de la tabla:

X/E	Primavera	Verano	Otoño	Invierno
1980	163.2458	173.5494	170.9871	174.3334
1981	179.8268	177.7409	182.0244	177.3171
1982	180.7429	180.9128	186.3474	187.1207
1983	191.9191	192.0145	186.1634	191.5963
1984	183.2163	190.9950	198.7644	190.2110

mediante el método de los mínimos cuadrados. Para ello, será necesario asignar un número a cada periodo de tiempo (t). Comenzaremos asignando un 1 a la primavera del año 1980, un 2 al verano del año 1980, hasta asignar un 20 al invierno del año 1984.

Ejemplo- Tendencia real

La recta de regresión que se ajusta a los datos de la tabla anterior es

$$X = 1.30027 \cdot t + 169.2985$$

y determina la tendencia (T) que mostramos en la siguiente tabla:

T	Primavera	Verano	Otoño	Invierno
1980	170.599	171.899	173.199	174.5
1981	175.8	177.1	178.4	179.701
1982	181.001	182.301	183.602	184.902
1983	186.202	187.502	188.803	190.103
1984	191.403	192.703	194.004	195.304

Ejemplo-Tendencia real

Obsérvese que estos valores de la tendencia real difieren de los valores de la tendencia obtenidos anteriormente, ya que en este caso no están afectados por la componente estacional.

Para calcular las variaciones cíclicas y aleatorias podríamos tomar los datos desestacionalizados $X/E = T \cdot C \cdot A$ y lo dividimos por la tendencia corregida o real para obtener una nueva tabla con las componentes C y A.

Componente aleatoria

Podemos aislar los factores accidentales eliminando los restantes, para ello podemos dividir la tabla $(X/T \cdot C)$ por los índices de estacionalidad correspondientes, quedando solo los factores accidentales.

A	Primavera	Verano	Otoño	Invierno
1980	–	–	0.9913	0.9950
1981	1.0149	0.9928	1.0141	0.9853
1982	0.9991	0.9904	1.0055	0.9950
1983	1.0140	1.0118	**0.9844**	1.0201
1984	0.9673	1.0002	–	–

los resultados, si son aleatorios deberían corresponderse con una distribución normal de media uno.
Un valor como el 0.9844 de la tabla se interpreta como que en otoño de 1983, los factores accidentales hicieron que el valor de X disminuyese en un 1.56 % (100-98.44).

Predicción. Autocorrelación

Predicción: Una vez descompuesta la serie en sus componentes, podemos estimar mejor un valor futuro de la variable X, mediante el producto de la tendencia por el factor de estacionalidad. Para este cometido resulta conveniente hallar la tendencia por el método de mínimos cuadrados.

Autocorrelación: Un método (laborioso) para buscar ciclos en una serie temporal es hallar el coeficiente de correlación lineal entre la serie X y la propia serie desplazada T lugares. Si obtenemos $r \approx 1$ habremos encontrado un ciclo.

Ejemplo

Ejemplo: Buscar si tiene un ciclo de periodo 3 la serie temporal:

$X_t = \{1.2, 1.4, 2.0, 1.23, 1.37, 2.07, 1.25, 1.47, 2.09\}$

t	1	2	3	4	5	6	7	8	9	10	11	12
X_t	1.2	1.4	2.0	1.23	1.37	2.07	1.25	1.47	2.09	–	–	–
X_{t+3}	–	–	–	1.2	1.4	2.0	1.23	1.37	2.07	1.25	1.47	2, 09

Usando solo los 6 valores comunes, obtenemos $r = 0.9937 \approx 1$ que indica que es un ciclo de la serie temporal.

Números índices

Definición

Llamamos **número índice elemental** *al porcentaje de variación (porcentual o en tanto por 1) de la variable estudiada respecto al valor obtenido para el periodo base de referencia.*

Observación: Es una medida estadística diseñada para poner de manifiesto los cambios producidos en una variable o grupo de ellas, generalmente respecto al tiempo.

Propiedades:

- **Identidad:** $X_{a/a} = 1$, **Inversión:** $X_{a/b} = \dfrac{1}{X_{b/a}}$
- **Cíclica:** $X_{a/b} \cdot X_{b/c} \cdot X_{c/a} = 1$
- **Cíclica modificada:** $X_{a/b} \cdot X_{b/c} \cdot X_{c/d} = X_{a/d}$

Ejemplos

Ejemplo

Si en 1970 una variable estadística X vale 123 y en 1980 vale 148, el índice simple:

$$X_{1980/1970} = \frac{X_{1980}}{X_{1970}} = \frac{148}{123} = 1.2033$$

Indica que se ha incrementado en un 20.33 % siendo 1970 *el periodo base.*

Ejemplo

La cantidad de monturas de gafas vendidas por una óptica, viene expresada en la tabla:

Año	2000	2001	2002	2003
Cantidad	1254	1345	1408	1451

Si tomamos como periodo base el año 2000, calcular todos los índices elementales posibles.

Solución:

$$C_{2000/2000} = \frac{1254}{1254} = 1, \quad C_{2001/2000} = \frac{1345}{1254} = 1.0726$$

$$C_{2002/2000} = \frac{1408}{1254} = 1.1228 \quad C_{2003/2000} = \frac{1451}{1254} = 1.1571$$

Ejemplo

En España, igual que en todos los países de la OCDE se calcula un Índice de Precios al Consumo a partir de los precios de una cesta estándar que se calcula estadísticamente como una media de lo que compra una familia española.

En el siguiente enlace podemos ver la serie actual de índices normalizados a partir de la media de 2012:

Serie del IPC posterior a 2002