

Tema 3: Números Índices y Series Temporales

Departamento Matemática Aplicada

Universidad de Málaga

Curso 2015-2016

Números índices

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

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.

Ejemplo

Si en 1970 una variable estadística X vale 123 y en 1980 vale 148, el índice simple $X_{1980/1970} = \frac{X_{1980}}{X_{1970}} = \frac{148}{123} = 1.1172$, indicando que se ha incrementado en un 11.72%.

En este caso 1970 es el periodo base.

Ejemplo índice elemental

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: $C_{2000/2000} = \frac{1254}{1254} = 1$ (siempre $C_{a/a} = 1$).

$C_{2001/2000} = \frac{1345}{1254} = 1.0726$, $C_{2002/2000} = \frac{1408}{1254} = 1.1228$ y

$C_{2003/2000} = \frac{1451}{1254} = 1.1571$.

Propiedades

- ① **Identidad:** $X_{a/a} = 1$
- ② **Inversión:** $X_{a/b} = \frac{1}{X_{b/a}}$
- ③ **Cíclica:** $X_{a/b}X_{b/c}X_{c/a} = 1$
- ④ **Cíclica modificada:** $X_{a/b}X_{b/c}X_{c/d} = X_{a/d}$

Definición

*Se llaman **índices simple de cadena** cuando cada índice elemental se va calculando respecto al periodo anterior.*

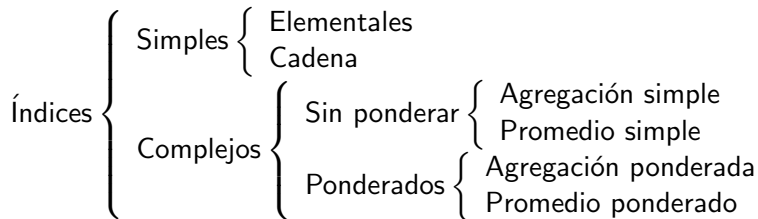
Ejemplo: Calcular los índices de cadena del ejemplo de la óptica:

$C = \{1254, 1345, 1408, 1451\}$.

$$I_{2001/2000} = \frac{1345}{1254} = 1.0726, \quad I_{2002/2001} = \frac{1408}{1345} = 1.0468,$$

$$I_{2003/2002} = \frac{1451}{1408} = 1.0305$$

Tipos de números índices



Índices de precios:



Índices complejos

Sirven para describir la evolución conjunta de varios productos.

Índices complejos sin ponderar

Definición

Media agregativa simple: Definimos el índice **media agregativa simple** (I_t) de varias variables X_i , respecto al periodo base $t = 0$ como:

$$I_t = \frac{\sum_i X_{it}}{\sum_i X_{i0}}$$

Definición

Media aritmética simple: Definimos el índice **media aritmética simple** (I_t) de n variables X_i , respecto al periodo base $t = 0$ como:

$$I_t = \frac{\sum_i \frac{X_{it}}{X_{i0}}}{n}$$

Ejemplo

Ejemplo

La óptica del problema anterior vende varios productos:

Año	2000	2001	2002	2003
Monturas	1254	1345	1408	1451
Cristales	1870	1760	1903	1964
Prismáticos	72	75	87	89

Calcular los índices media agregativa simple y media aritmética simple para el año 2003 respecto al año base 2000.

Media agregativa: $I_{Ma} = \frac{1451+1964+89}{1254+1870+72} = 1.0964 \Rightarrow 109.64 \%$

Media aritmética: $I_{Me} = \frac{\frac{1451}{1254} + \frac{1964}{1870} + \frac{89}{72}}{3} = 1.1478 \Rightarrow 114.78 \%$

Índices complejos ponderados

Usualmente unos productos tienen más importancia que otros para lo que usamos una ponderaciones w_i para cada producto.

Definición

*Definimos la **media agregativa ponderada** (I_t) de varias variables X_i , respecto al periodo base $t = 0$, como:*

$$I_t = \frac{\sum_i X_{it} w_i}{\sum_i X_{i0} w_i}$$

Definición

*Definimos la **media aritmética ponderada** (I_t) de varias variables X_i , respecto al periodo base $t = 0$, como:*

$$I_t = \frac{\sum_i \frac{X_{it}}{X_{i0}} w_i}{\sum_i w_i}$$

Ejemplo

Continuando ejemplo de la óptica, calcular los índices para 2003 ponderando las monturas con 2, los cristales con 3 y los prismáticos con 1.

Media agregativa ponderada:

$$I_{Maw} = \frac{1451 * 2 + 1964 * 3 + 89}{1254 * 2 + 1870 * 3 + 72} = \mathbf{1.0846} \Rightarrow 108.46 \%$$

Media aritmética ponderada:

$$I_{Mew} = \frac{\frac{1451*2}{1254} + \frac{1964*3}{1870} + \frac{89}{72}}{6} = \mathbf{1.1169} \Rightarrow 111.69 \%$$

Índices de precios

Los índices más usados son los de precios y usualmente se ponderan con las cantidades (vendidas, compradas, producidas, ...)

Definición

*Definimos el **índice de Laspeyres** (L_t) de varias variables X_i de precios p_{it} y cantidades vendidas q_{it} en el año t , respecto al periodo base $t = 0$, como:*

$$L_t = \frac{\sum_i p_{it} q_{i0}}{\sum_i p_{i0} q_{i0}}$$

El IPC (Índice de Precios al Consumo) es un índice de Laspeyres, donde los precios de multitud de productos son ponderados por las cantidades consumidas en el periodo base.

Índices de Paasche y de Fisher

Definición

Definimos la **índice de Paasche** (P_t) de varias variables X_i de precios p_{it} y cantidades vendidas q_{it} en el año t , respecto al periodo base $t = 0$, como: $P_t = \frac{\sum_i p_{it} q_{it}}{\sum_i p_{i0} q_{it}}$

Definición

Definimos el **índice ideal de Fisher** como la media geométrica del de Laspeyres y el de Paasche: $F_t = \sqrt{L_t P_t}$

Ejemplo

Ejemplo

Una empresa de ventanas compra 3 tipos de productos X =perfiles de aluminio (m.), Y =cristal (m^2) y Z =cerraduras (uds.) con los precios p_i y cantidades q_i :

Año	2006	2007	2008	2009	2010	2011
p_X	25.6	28.7	30.0	32.2	32.1	31.8
q_X	734	814	923	940	903	867
p_Y	37.5	41.3	42.5	43.2	45.1	45.5
q_Y	203	197	240	267	240	233
p_Z	25	25	27	28	28	28
q_Z	78	82	97	104	98	103

Hallar los índices de precios de Laspeyres, Paasche y Fisher de los diferentes años y periodo base 2006.

Solución

Los índices del periodo base siempre son 1: $L_{2006} = P_{2006} = F_{2006} = 1$.

Para el resto mediante **Laspeyres**:

$$L_{2007} = \frac{28.7(734) + 41.3(203) + 25(78)}{25.6(734) + 37.5(203) + 25(78)} = 1.1075,$$

$$L_{2008} = \frac{30.0(734) + 42.5(203) + 27(78)}{25.6(734) + 37.5(203) + 25(78)} = 1.1552,$$

$$L_{2009} = \frac{32.2(734)+43.2(203)+28(78)}{25.6(734)+37.5(203)+25(78)} = 1.2199, L_{2010} = 1.2309, L_{2011} = 1.2260$$

Mediante **Paasche**: $P_{2007} = \frac{28.7(814)+41.3(197)+25(82)}{25.6(814)+37.5(197)+25(82)} = 1.1081$,

$$P_{2008} = \frac{30.0(923) + 42.5(240) + 27(97)}{25.6(923) + 37.5(240) + 25(97)} = 1.1556,$$

$$P_{2009} = \frac{32.2(940)+43.2(267)+28(104)}{25.6(940)+37.5(267)+25(104)} = 1.2192, P_{2010} = 1.2311,$$

$$P_{2011} = 1.2253$$

Mediante **Fisher**: $F_{2007} = \sqrt{1.1075(1.1081)} = 1.1078$,

$$F_{2008} = \sqrt{1.1552(1.1556)} = 1.1554, F_{2009} = \sqrt{1.2199(1.2192)} = 1.2195,$$

$$F_{2010} = \sqrt{1.2309(1.2311)} = 1.2310, F_{2011} = \sqrt{1.2260(1.2253)} = 1.2257.$$

Cambio de periodo base. Renovación y empalme

- **Cambio de periodo base:** Periódicamente es necesario cambiar el periodo base a una fecha más reciente, resultando necesario comparar los nuevos índices con los anteriores.
- **Renovación y empalme:** A menudo los productos que son usados para el cálculo, cambian de importancia y aparecen otros que deben ser considerados, a eso se le llama **renovación** de un número índice. Debemos ser capaces de enlazar un número índice referido a un grupo de productos y con una ponderación dada, con otros calculados con otros productos ligeramente diferentes y con otra ponderación. A eso se le llama **empalme** de un número índice.

Así, un Índice de Precios al Consumo en 1900 debería dar una alta ponderación al carbón, pienso para animales, etc., sin embargo, el de 2010 deberá dársele a la gasolina e incluir productos como teléfono, móvil, ordenador, . . . , que no figuraban en el primero.

Deflacción

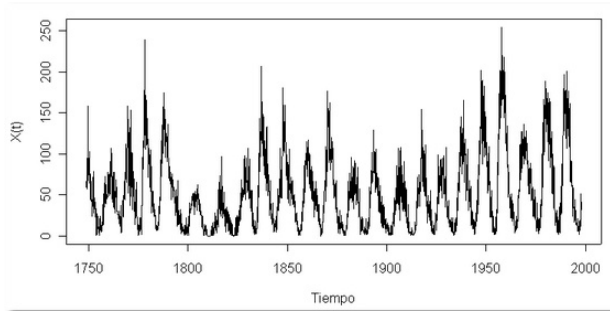
Al considerar el índice del valor de un producto a lo largo del tiempo, resulta conveniente tener en cuenta el IPC, distinguiendo entre el 'valor nominal' y el 'valor real' del producto.

Así, la serie de valores nominales vendrá expresada en la moneda del año en curso, pero si dividimos por el IPC vendrá expresada en moneda constante del año base. A este proceso se le llama **deflacción**.

Series Temporales o Cronológicas

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 , y la otra el fenómeno cuantitativo que se desea estudiar X .

Representación gráfica: Representamos la nube de puntos X_t :



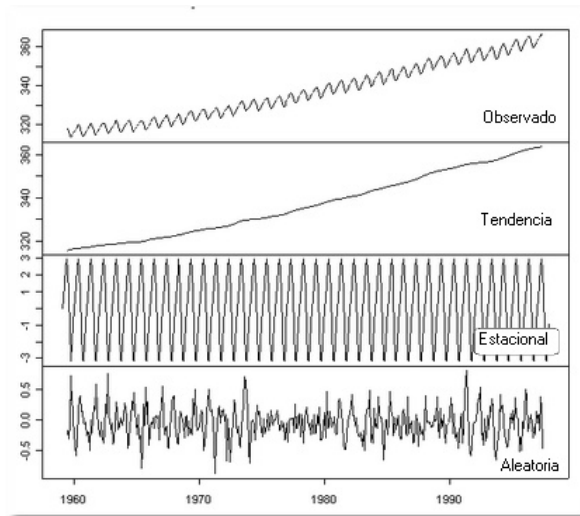
Componentes de las S.T.

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

- **Tendencia: (T)** Es la dirección dominante de la serie observada.
- **Variaciones estacionales: (E)** Muchas variables vienen afectadas por variaciones periódicas de periodo corto (semana, mes, año).
- **Variaciones cíclicas: (C)** Muchas variables económicas presentan variaciones periódicas de periodo largo. Para observarlas se necesita tener un gran número de observaciones.
- **Variaciones accidentales: (A)** Son debidas a fenómenos aleatorios.

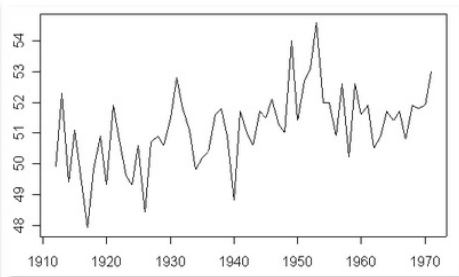
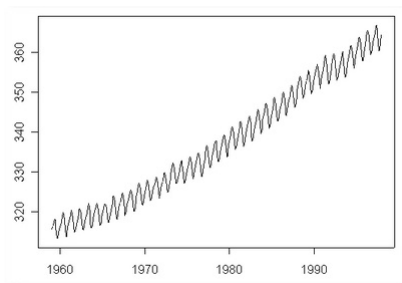
Hipotesis Multiplicativa: $X = T \cdot E \cdot C \cdot A$
(aunque existe $X = T + E + C + A$)

Descomposición



Tendencia

Marca la dirección (generalmente recta) dominante.



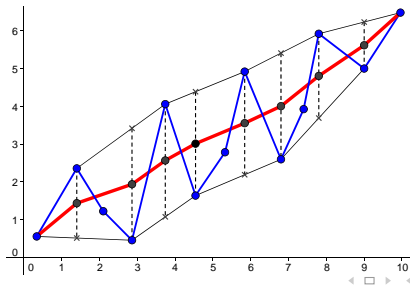
En ambas gráficas se observa una tendencia creciente, más marcada en la de la izquierda.

Determinación de la tendencia

Veremos 3 métodos:

- ① **Gráfico.**
- ② **Medias móviles.**
- ③ **Mínimos cuadrados.**

1) Método gráfico: Unimos los picos superiores por un lado y los inferiores por otro, luego calculamos la media. Con ello, conseguimos un suavizado.

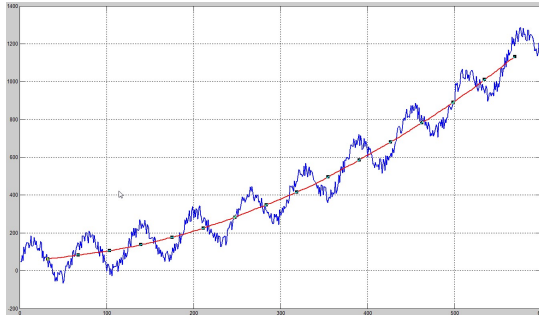


Medias móviles

Definición

Dada una sucesión de valores X_t , llamamos **media móvil de orden k** a la sucesión $Y = \hat{X}_3$ obtenida a partir de la primera:

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



Suavización mediante el uso de medias móviles de orden el periodo.

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 las sucesiones de medias móviles de orden 3 y 4 asociadas.

Solución: Para orden 3: $y_1 = \frac{2.6767+3.2050+4.8714}{3} = 3.5844$,
 $y_2 = \frac{3.2050+4.8714+4.6275}{3} = 4.2346$, $y_3 = \frac{4.8714+4.6275+5.9008}{3} = 5.1332$,
 $y_4 = \frac{4.6275+5.9008+5.1186}{3} = 5.3494$

Para el orden 4: $z_1 = \frac{2.6767+3.2050+4.8714+4.6275}{4} = 3.8452$,
 $z_2 = \frac{3.2050+4.8714+4.6275+5.9008}{4} = 4.6512$,
 $z_3 = \frac{4.8714+4.6275+5.9008+5.5200}{4} = 5.2299$,
 $z_4 = \frac{4.6275+5.9008+5.5200+5.1186}{4} = 5.2917$

Ambas suponen una suavización de la serie inicial.

Tendencia mediante medias móviles

2) Tendencia mediante medias móviles: En realidad la tendencia es una suavización de los valores que evita la dependencia de ciclos y factores accidentales. Así, podemos usar las medias móviles con este fin.

Lo que hacemos es centrarlas en el tiempo, esto es, si usamos medias de orden 3, la primera media móvil la consideramos en $t = 2$, si son de orden 5 en $t=3$, etc.

Si el orden es par tendremos que centrarlas en dos pasos. Por ejemplo, la primera media móvil para orden 4 resulta centrada en 2.5, la segunda en 3.5, etc. Posteriormente, sacamos la media entre ambas para centrarla en $t=3$.

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.

Para $k=3$, el orden es impar y resultan centradas en el tiempo:

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 dos fases:

t	X	\hat{X}_4	$T = \hat{\hat{X}}_4$
1959	6.0	5.55	—
1960	6.2		—
1961	5.1	5.35	5.45
1962	4.9	5.25	5.30
1963	5.2	5.725	5.4875
1964	5.8	5.725	—
1965	7.0		—

Calculos: $\frac{6+6.2+5.1+4.9}{4} = 5.55$, $\frac{6.2+5.1+4.9+5.2}{4} = 5.35, \dots$
 $\frac{5.55+5.35}{2} = 5.45$, $\frac{5.35+5.25}{2} = 5.30, \dots$

Comentarios

Las medias móviles constituyen un potente método para suavizar la serie temporal, adaptándose además a factores cíclicos de periodo grande y sin presuponer la forma de la función tendencia.

Tienen el inconveniente de que se pierden valores extremos. Cuánto mayor sea el orden, más valores se pierden aunque se suaviza más.

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.

Tendencia mediante mínimos cuadrados

3) Tendencia por el método de mínimos cuadrados: Consiste en ajustar a los datos una recta, parábola, etc.

Debido a que la variable t toma valores consecutivos los cálculos pueden simplificarse mucho si hacemos el cambio:

- Si el número de valores de t es impar: $t' = t - \bar{t}$
- Si el número de valores de t es par: $t' = 2(t - \bar{t})$

que produce $\sum_i t'_i = 0$ y $\sum_i t'^3_i = 0$ por lo que para ajustar una recta $x = a + bt'$ calculamos: $a = \frac{\sum_i x_i}{N}$ y $b = \frac{\sum_i t'_i x_i}{\sum_i t'^2_i}$.

Ejemplo

Ejemplo

Ajustar una recta de tendencia a los datos:

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

$N = 7$, y $\bar{t} = 1962 \Rightarrow$ hacemos el cambio $t' = t - 1962$ quedando:

t_i	X_i	$t'_i = t_i - 1962$	t'^2_i	$t'_i X_i$	$T = X_i^*$	$ e_i $	e^2_i	x^2_i
1959	6	-3	9	-18	5.4964	0.5036	0.2536	36.00
1960	6.2	-2	4	-12.4	5.5786	0.6214	0.3862	38.44
1961	5.1	-1	1	-5.1	5.6607	0.5607	0.3144	26.01
1962	4.9	0	0	0	5.7429	0.8429	0.7104	24.01
1963	5.2	1	1	5.2	5.8250	0.6250	0.3906	27.04
1964	5.8	2	4	11.6	5.9071	0.1071	0.0115	33.64
1965	7	3	9	21	5.9893	1.0107	1.0215	49.00
	40.2		28	2.3			3.0882	234.14

$$a = \frac{40.2}{7} = 5.7429, b = \frac{2.3}{28} = 0.0821,$$

$$\Rightarrow y = 5.7429 + 0.0786t' = 5.7429 + 0.0821(t - 1962), V_r = \frac{3.0882}{7} = 0.4412,$$

$$V_y = \frac{237.14}{7} - \left(\frac{40.2}{7}\right)^2 = 0.4682, r = 0.2401 \text{ Mal ajuste}$$

Comparación entre métodos.

De los 3 métodos, el gráfico es más simple, pero tiene el problema de que puede diferir para diversos usuarios y no se dispone de una expresión analítica.

El método de las medias móviles tiene la ventaja de no presuponer la forma de los datos y estimar conjuntamente tendencia y ciclos de periodo largo, pero no permite la predicción y pierde los extremos.

El método de los mínimos cuadrados presupone una forma a la distribución de datos y, sin embargo, permite predecir y dar una medida de la bondad del ajuste.

Estacionalidad

Una vez estimada la tendencia por alguno de los métodos anteriores, pasamos a estudiar los factores estacionales, mediante el cálculo de los índices de variación estacional.

Método de la media móvil en porcentajes: 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' de la serie original X.

Cálculo de la tendencia

Usamos medias móviles orden 4 para calcular la tendencia:

\hat{T}_4	Prim.	Ver.	Otoño	Inv.
1980		170.225	174.750	175.675
1981	178.675	179.375	179.625	180.325
1982	181.500	183.800	186.850	189.300
1983	189.250	190.300	187.925	187.700
1984	191.125	190.800		

y medias móviles orden 2 para centrarlas en el tiempo:

T	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	—	—

Eliminación de la tendencia

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. para eliminar sus efectos, dividimos la tabla X por la T Así, $X = T \cdot E \cdot C \cdot A$ y entonces $\frac{X}{T} = \frac{T \cdot E \cdot C \cdot A}{T \cdot C} = E \cdot A$, es decir quedan, solo los factores estacionales y accidentales:

$\frac{X}{T}$	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 sacar el índice de cada estación, se saca la media de los valores correspondientes en la tabla X/T :

$$\begin{aligned} I_P &= \frac{1.1079+1.0906+1.1069+1.0559}{4} = 1.0903 \Rightarrow 109.03\% \\ I_V &= \frac{0.8764+0.8743+0.8932+0.8829}{4} = 0.8817 \Rightarrow 88.17\% \\ I_O &= \frac{1.0778+1.1025+1.0932+1.0703}{4} = 1.0859 \Rightarrow 108.59\% \\ I_I &= \frac{0.9337+0.9246+0.9337+0.9573}{4} = 0.9373 \Rightarrow 93.73\% \end{aligned}$$

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 indica una disminución del 11.83 %.

De todas formas, estos índices se llaman 'sin corregir' pues deben sumar 400 % y, en realidad suman 399.53 por lo que debemos calcular

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

$$I'_O = \frac{400}{399.53} I'_0 = 108.7217, I'_I = \frac{400}{399.53} I_I = 93.8432$$

que son los índices estacionales corregidos o ajustados.

Desestacionalización

Resulta útil eliminar los factores estacionales de la serie original X , esto 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

Factores accidentales

Podemos aislar los factores accidentales eliminando los restantes, para ello podemos dividir la tabla $\frac{X}{T}$ 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 1.

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 1.54 % (100-98.44).

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: 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.