Números Índices y Series Temporales

Depto. Matemática Aplicada

Universidad de Málaga

Números índices

Medida estadística diseñada para poner de manifiesto los cambios producidos en una(s) variable(s), generalmente respecto al tiempo.

Definición

Llamamos **número índice elemental** *a la 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 $X_{1980/1970} = \frac{X_{1980}}{X_{1970}} = \frac{148}{123} = 1.2033$, *indicando que se ha incrementado en un $20.33\,\%$.*

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

Ejemplo: Los índices de cadena para $C = \{1254, 1345, 1408, 1451\}$ son

$$I_{\frac{2001}{2000}} = \frac{1345}{1254} = 1.0726 \quad I_{\frac{2002}{2001}} = \frac{1408}{1345} = 1.0468 \quad I_{\frac{2003}{2002}} = \frac{1451}{1408} = 1.0305$$

Tipos de números índices

$$
\text{Índices}\left\{\begin{array}{l}
\text{Simples}\left\{\begin{array}{l}\text{Elementales}\\ \text{Cadena}\end{array}\right.\\[2ex]
\text{Complejos}\left\{\begin{array}{l}
\text{Sin ponderar}\left\{\begin{array}{l}\text{Agregación simple}\\ \text{Promedio simple}\end{array}\right.\\[2ex]
\text{Ponderados}\left\{\begin{array}{l}\text{Agregación ponderada}\\ \text{Promedio ponderado}\end{array}\right.
\end{array}\right.
\end{array}\right.
$$

Índices complejos
Sin ponderar

Sirven para describir la evolución conjunta de varios productos.

Definición

La **media agregativa simple** *de varias variables* X_i*, respecto al periodo base* $t = 0$ *viene dada por:*

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

Definición

Media aritmética simple: *La* **media aritmética simple** *de* n *variables* X_i*, respecto al periodo base* $t = 0$ *se define como:*

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

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

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

Índices complejos

Ponderados

Usualmente unos productos tienen más importancia que otros, esto se ve reflejado en el uso de distintasponderaciones w_i.

Definición

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

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

Definición

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

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

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:

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

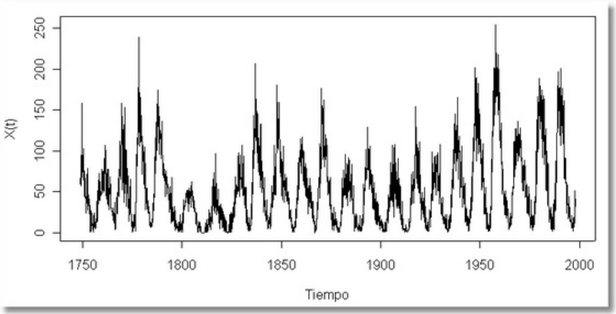
Media aritmética ponderada:

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

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 una Serie Temporal

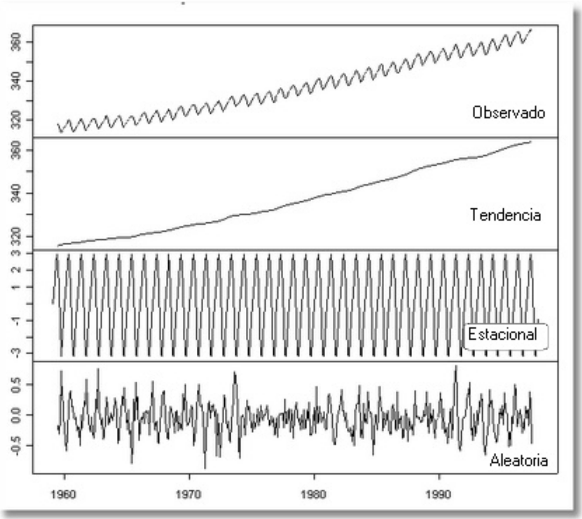
Cada uno de los valores observados puede considerarse como la conjunción de cuatro factores que lo determina:

- **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.

Hipótesis Multiplicativa: $\mathbf{X} = \mathbf{T} \cdot \mathbf{E} \cdot \mathbf{C} \cdot \mathbf{A}$

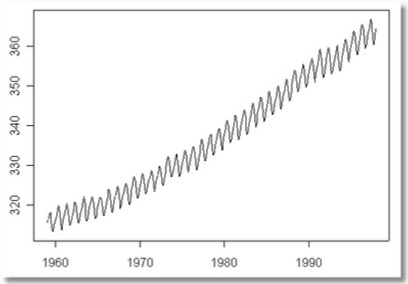
(también existe la hipótesis aditiva, por la que $X = T + E + C + A$)

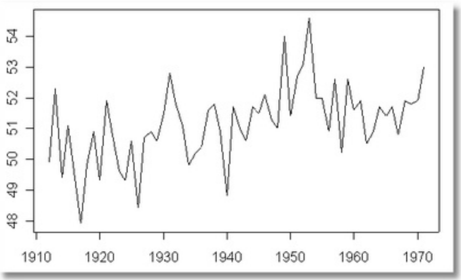
Descomposición



Tendencia

Marca la dirección (generalmente recta) dominante.



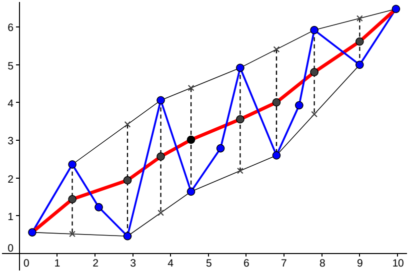


En ambas gráficas se observa una tendencia creciente, más marcada en la de la izquierda.

Determinación de la tendencia

Método gráfico

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



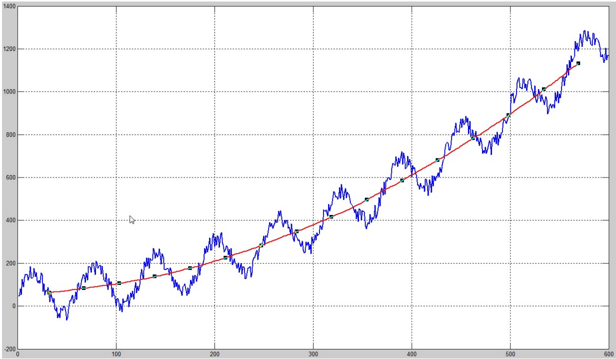
Determinación de la tendencia

Medias móviles

Definición

Llamamos **media móvil de orden k** *de una sucesión de valores* X_t*, a la nueva sucesión dada por:*

$$Y_1 = \frac{X_1 + X_2 + \ldots + X_k}{k}, \quad Y_2 = \frac{X_2 + X_3 + \ldots + X_{k+1}}{k}, \ldots$$



Determinación de la tendencia

Medias móviles

Ejemplo

Dada la serie de valores: $x = \{2.6767, 3.2050, 4.8714, 4.6275, 5.9008, 5.5200, 5.1186, 4.9507, 4.6612, 6.2572, \ldots\}$.
Hallar los primeros términos de las sucesiones de medias móviles de orden 3 y 4 asociadas.

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

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.

Determinación de la tendencia

Medias móviles

Las medias móviles suavizan los valores de forma que se evita la dependencia de ciclos y factores accidentales.
Cada media móvil se centra en el tiempo:

- 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 resulta centrada entre dos mediciones, con lo que se realizan 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

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

calcule 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	–	5.7667	5.4000	5.0667	5.3000	6.0000	–

Determinación de la tendencia

Medias móviles: ejemplo de orden par

Para orden 4, al ser par se realizan dos fases:

t	X	T_1	T_2
1959	6.0		–
1960	6.2		–
		5.55	
1961	5.1		5.45
		5.35	
1962	4.9		5.30
		5.25	
1963	5.2		5.4875
		5.725	
1964	5.8		–
1965	7.0		–

Determinación de la tendencia

Medias móviles

- 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, etc.

Determinación de la tendencia

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 con un cambio de variable:

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

Comparación

De estos 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:

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

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 índica 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).