Instituto Nacional de Estadística y Geografía

Metodología del ajuste estacional

INEGI

INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

Catalogación en la fuente INEGI:

519.5501 Instituto Nacional de Estadística y Geografía (México).
Metodología del ajuste estacional / Instituto Nacional de Estadística y Geografía.-
- México : INEGI, c2013.

9 p.

ISBN En trámite.

1. Análisis de series de tiempo - México - Metodología. 2. Estadística matemática
- Metodología.

Si requiere más información sobre esta obra, favor de contactarnos a través de:
Centros de Información (consulte el domicilio en Internet)
Larga distancia sin costo: 01 800 111 46 34
www.inegi.org.mx
atencion.usuarios@inegi.org.mx

Síguenos en:

Introducción

El propósito del presente apartado es describir brevemente la metodología que se utiliza para realizar el ajuste estacional de los indicadores económicos de coyuntura que difunde el INEGI.

Para la generación de estas series, se utiliza el paquete estadístico X12-Arima que es distribuido y generado por el Buró de Censos de los Estados Unidos y que es ampliamente utilizado a nivel mundial.

La necesidad de realizar el ajuste estacional es debido a que las series económicas se ven afectadas por factores estacionales periódicos, que se repiten cada año, y cuyas causas pueden considerarse ajenas a la naturaleza económica de las series, como pueden ser: ciertas festividades fijas, el hecho de que algunos meses tienen más días que otros, los periodos de vacaciones escolares, el efecto del clima en las diferentes estaciones del año y las expectativas de fluctuaciones estacionales. Por ejemplo, la elevada producción de juguetes en los meses previos a la Navidad es provocada por la expectativa de ventas elevadas en diciembre.

Dichos efectos no permiten comparar adecuadamente los resultados de un determinado periodo respecto al inmediato anterior, ya que su presencia dificulta diagnosticar o describir fielmente el comportamiento de una serie económica. La desestacionalización o el ajuste estacional de series económicas, consiste en remover precisamente aquellas influencias intra-anuales periódicas.

Para poder realizar el ajuste estacional mediante el paquete estadístico X12-Arima, se requiere que la longitud de la serie sea de por lo menos cinco años de datos mensuales o trimestrales.

Componentes de una serie de tiempo

Las series de tiempo están constituidas por varios componentes que, si bien no se pueden ver de manera directa, existen y pueden separarse de la serie original. Los componentes de una serie de tiempo son: *factor estacional, tendencia, ciclo y factor irregular.*

El componente estacional corresponde a los efectos periódicos que se repiten cada año, aproximadamente en las mismas fechas y con la misma magnitud y cuyas causas pueden considerarse ajenas a la naturaleza económica de las series, como pueden ser ciertas festividades fijas, el clima, las vacaciones, etcétera.

La tendencia representa el comportamiento general del indicador, describe su evolución de largo plazo permitiendo identificar el patrón en la dirección que siguen los elementos de un entorno particular durante un periodo de tiempo; en tanto que el ciclo es una oscilación con cierta periodicidad que se caracteriza por lapsos alternos de expansión y contracción. Al realizar la descomposición mediante el paquete estadístico X12-Arima estos dos componentes de obtienen como uno solo denominado tendencia-ciclo.

Finalmente el componente irregular caracteriza los movimientos imprevisibles en cuanto a su ocurrencia, impacto y duración y que son considerados como aleatorios. Estos movimientos pueden surgir por desastres naturales, sequías, hambrunas, huelgas, etc., o ser resultado del error estadístico.

Al realizar la descomposición, la serie original se ajusta mediante técnicas econométricas, por los efectos de calendario, el factor estacional y el irregular para obtener nuevas series de tiempo que son de gran utilidad para el análisis.

Descomposición de la serie

Para llevar a cabo el proceso de descomposición de una serie de tiempo, se puede utilizar un método basado en filtros de promedios móviles. El paquete X12-Arima utiliza este método mediante el proceso x11 que tiene implementado. Éste consiste en la aplicación consecutiva e iterativa de promedios móviles de diferente orden y distinto tipo (simples o compuestos). Estos serán simples o compuestos dependiendo de la etapa en la descomposición de la serie. Supongamos como ejemplo un promedio móvil simple de orden 3 (n=3) sobre una serie de 9 observaciones:

Serie original	Promedio móvil
1	-
3	(1+3+6)/3= 3.33
6	(3+6+8)/3= 5.67
8	(6+8+24)/3= 12.67
24	(8+24+13)/3= 15.00
13	(24+13+4)/3= 13.67
4	(13+4+16)/3= 11.00
16	(4+16+21)/3= 13.67
21	-

Los resultados se obtuvieron al aplicar un promedio móvil de orden 3 de la siguiente forma: para obtener el primer valor se calcula un promedio de los primeros 3 datos que corresponden al dato del periodo de referencia, al dato anterior y al dato posterior, es decir:

$$(1+3+6) / 3 = 3.33$$

Para obtener el segundo valor, el promedio se recorre un sitio y se vuelven a promediar tres elementos:

$$(3+6+8) / 3 = 5.67$$

Y así sucesivamente hasta obtener los siete datos de la nueva serie.

Los promedios móviles compuestos se obtienen al aplicar dos promedios simples consecutivos; por ejemplo, el promedio móvil de orden 3x5 es la aplicación de un promedio móvil simple de orden 3 y a la serie que se obtiene se le aplica un promedio móvil simple de orden 5.

El proceso x11 también utiliza otros filtros que son promedios ponderados simétricos, es decir, el primer dato y el último dato considerados tiene el mismo peso o ponderación, el segundo y el penúltimo también tendrán el mismo peso y así respectivamente.

La aplicación de filtros en una serie ocasiona que se pierdan valores al inicio y al final; esto se debe a que se promedian "*n*" observaciones alrededor de cada uno de los datos de la serie, de tal modo que si una serie tiene "*N*" datos se obtiene una nueva serie de valores promediados con $N-(n-1)$ datos.

Dado que el mayor interés de la descomposición es analizar los últimos datos de la serie, es necesario obtener estimaciones sobre el comportamiento futuro de la serie original, para que después de la aplicación de los filtros se conserve la longitud inicial de la misma. Las estimaciones futuras se realizan mediante un modelo autorregresivo, integrado de promedios móviles (ARIMA).

- **Modelo**

Para efectuar la descomposición con x11, es necesario primero ajustar la serie original por efectos de calendario y, posteriormente, obtener la serie estimada; para esto, es necesario realizar un análisis estadístico de la serie a fin de determinar el modelo más adecuado para ella, que incluye los siguientes aspectos:

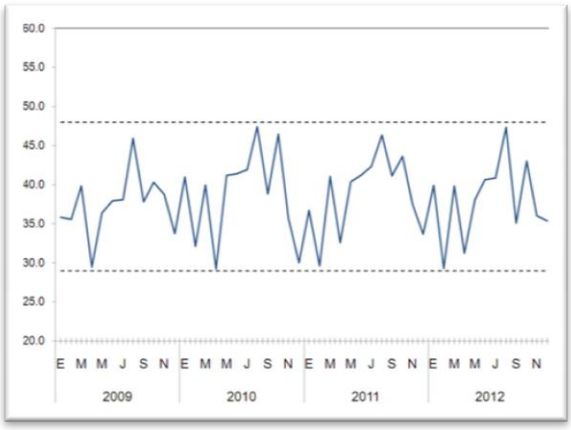
Transformación.
Factores de calendario.
Modelo ARIMA (autorregresivos, integrados y de promedios móviles).
Outliers (valores atípicos).

Selección de la transformación

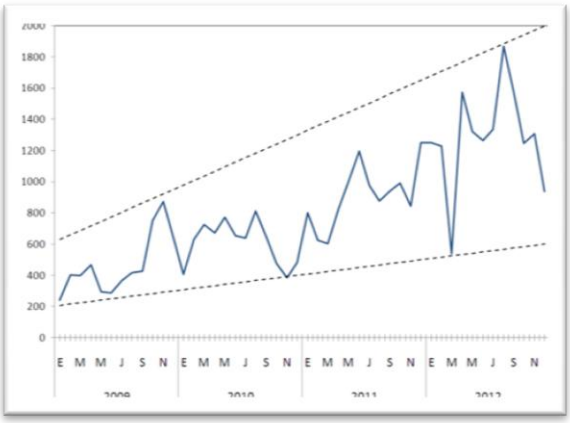
Para que una serie sea analizada estadísticamente primero se evalúa la variabilidad que presenta, si ésta no es constante en el tiempo es necesario aplicar la transformación logaritmo, obteniendo así que la variabilidad se haga constante, para poder determinar el resto de los parámetros del modelo.

INEGI. Metodología del ajuste estacional. 2013.

Variabilidad constante



Variabilidad no constante



Si la serie requiere la transformación logaritmo, se utiliza un modelo multiplicativo en la descomposición; es decir, para obtener la serie original los componentes de la serie se multiplican.

*Serie Original = Tendencia-Ciclo * Estacional * Irregular*

Si no requiere transformación, el modelo será aditivo, donde los componentes se suman.

Serie Original = Tendencia-Ciclo + Estacional + Irregular

Identificación de los factores de calendario

Una vez determinado si se requiere transformación o no, se identifica la presencia de los factores de calendario, que se presentan debido a que no todos los meses o trimestres tienen el mismo número de días o de fines de semana. Estos efectos pueden medirse y separarse de otros aspectos que afecten a la serie.

En particular los factores que ajustan son: el efecto por frecuencia de días de la semana y el efecto de Semana Santa. Algunas series contienen ambos efectos; otras, sólo uno de ellos; y otras más, no se ven afectadas por ninguno de los dos.

La magnitud de los efectos depende de cada serie y se calcula a partir de un modelo de regresión sobre la serie original del indicador económico en cuestión.

1. Efecto por la frecuencia de los días de la semana

El efecto por la frecuencia de los días de la semana considera la frecuencia que tiene cada día (lunes, martes,…) en un mes, ya que algunas series económicas pueden estar fuertemente influenciadas por la composición diaria del mes. Un sábado de más o de menos, en un determinado mes, puede hacer variar de manera significativa los indicadores. Aquí también se identifica el impacto de un día de más en el mes de febrero para los años bisiestos.

Existen varias formas de medir este efecto; sin embargo, los ajustes utilizados son:

a) La instrucción *Td* calcula una ponderación para cada día de la semana, además de considerar si el mes de febrero corresponde a un año bisiesto o no.
b) La instrucción *Tdnolpyear* calcula una ponderación para cada día de la semana, sin considerar años bisiestos.
c) La instrucción *Td1coef* calcula una ponderación igual de lunes a viernes y otra para el fin de semana considerando los años bisiestos.
d) La instrucción *Td1nolpyear* calcula una ponderación igual de lunes a viernes y otra para el fin de semana, sin considerar años bisiestos.
e) La instrucción L*pyear* únicamente considera los años bisiestos.
f) La instrucción Lom considera la longitud del periodo.

Para el tipo de ajuste *Td*, el cual es el más utilizado, el factor en el caso mensual se obtiene mediante la siguiente expresión que asigna una ponderación distinta a cada día de la semana:

$$M_t = \frac{N_t}{N_t^*} \exp \sum_{i=1}^{6} \beta_i \left(D_{it} - D_{7t} \right)$$

Donde

M_t = factor de ajuste por la frecuencia de los días de la semana

i = 1 corresponde a lunes, i = 2 martes,... y i = 7 domingo

D_{it} = número de días j en el mes t

N_t = Número de días del mes t

N_t* = la media, sobre 4 años, de la longitud del mes (30, 31 ó 28.25) [1]

β_i = parámetros de las ponderaciones diarias

2. Efecto por Semana Santa

La Semana Santa produce un cambio de nivel de actividad en numerosos sectores debido a que esta festividad no es fija, pues en ocasiones ocurre en marzo y otras en abril; a este efecto se le conoce como Efecto de la Semana Santa. Cada indicador se ve afectado por un cierto número de días previos a que ocurra la Semana Santa, que puede ser desde un día hasta quince, los cuales se contabilizan desde el Domingo Santo hacía atrás. El número de días cambia dependiendo del indicador que se trate por lo que se busca el óptimo para cada caso de acuerdo al estadístico para cada uno de los números de días.

Este factor se obtiene mediante la siguiente expresión:

$$E_{ij} = \hat{a} + \hat{b} X_{ij}(w)$$

[1] En el caso de los meses con 30 o 31 días el promedio es igual a los días del mes. Para Febrero la suma es 28+28+28+29 por el año bisiesto por lo que el promedio es igual a 28.25.

INEGI. Metodología del ajuste estacional. 2013.

E_{ij} = factor de ajuste por la Semana Santa

a y b = parámetros estimados de la regresión

w = número de días previos a la Semana Santa, en que se ve afectada la serie en estudio

w_{ij} = el número de días de los w días antes de la Semana Santa que caen en el periodo j

$X_{ij}(w) = w_{ij} / w - \underline{Z}_{.j}(w)$

$\underline{Z}_{.j}(w)$ = la media w_{ij} / w

Modelo ARIMA

Una vez determinados los efectos de calendario se busca el modelo ARIMA (p,d,q)(P,D,Q) más adecuado a cada serie. Estos modelos son aquellos en los que una variable y su evolución, presente y futura, se explican utilizando exclusivamente el comportamiento histórico de la variable en cuestión. Esto es, las estimaciones futuras vienen explicadas por los datos del pasado y no por otras variables.

Los modelos ARIMA se forman por una parte regular (p,d,q) y una parte estacional (P,D,Q) las cuales se conforman por parámetros autorregresivos (p o P), un orden de integración (d o D) y parámetros de promedios móviles (q o Q). Es muy importante encontrar una especificación del orden de los parámetros[2], a fin de que el modelo ARIMA represente de manera apropiada a la serie.

Una forma de evaluar que el modelo sea adecuado es que cumpla con los supuestos estadísticos. Algunos de estos supuestos involucran los residuales del modelo, que miden la discrepancia entre los valores observados y los valores estimados por el modelo, por lo que se calculan mediante la diferencia de valores observados y los estimados.

Los supuestos estadísticos que se deben cumplir son:

* Los residuales tienen media cero.
* Los residuales tienen varianza constante.
* Los residuales son variables independientes.
* Los residuales tienen una distribución normal.
* Dentro de los residuales no existen observaciones aberrantes.
* El modelo es considerado parsimonioso.
* El modelo es admisible (condiciones de estacionariedad e invertibilidad).
* El modelo es estable.

[2] El orden de los parámetros hace referencia al valor más alto de los parámetros (autorregresivos o promedios móviles) o diferencias tanto regulares como estacionales que conforman el modelo ARIMA para cada serie.

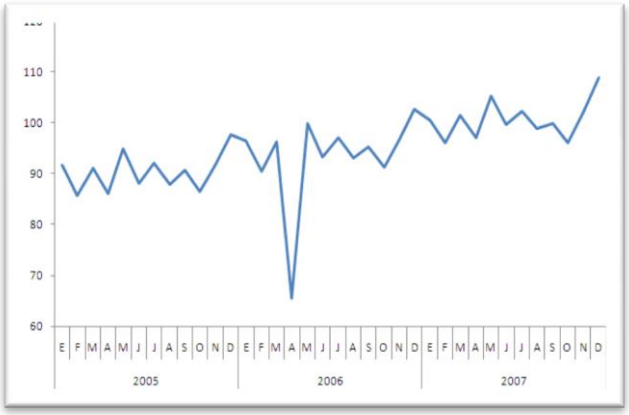
INEGI.Metodología del ajuste estacional. 2013.

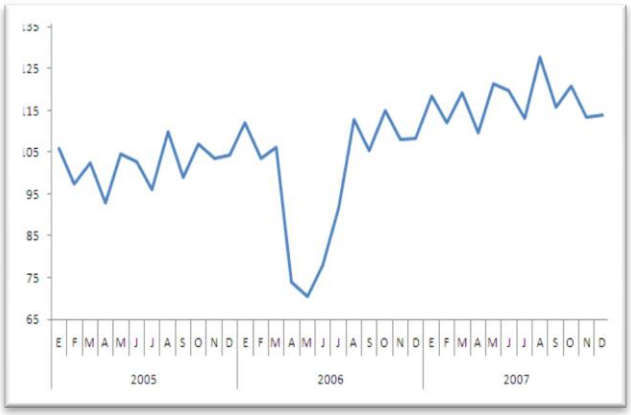
Identificación de *Outliers*

A la par de esta búsqueda se determinan los *outliers* adecuados para la serie dependiendo del modelo. Estos pueden ser del tipo: a*ditivo, cambio temporal* o *cambio de nivel*.

Aditivo

Se trata de un *outlier Aditivo* cuando hay sólo una observación inusual o atípica, fuera del comportamiento observado en la serie. Posteriormente, en el siguiente periodo la serie regresa al nivel anterior y registra el mismo comportamiento que venía presentando. Un ejemplo de esto se da por huelgas cortas, que sólo afectan un periodo.



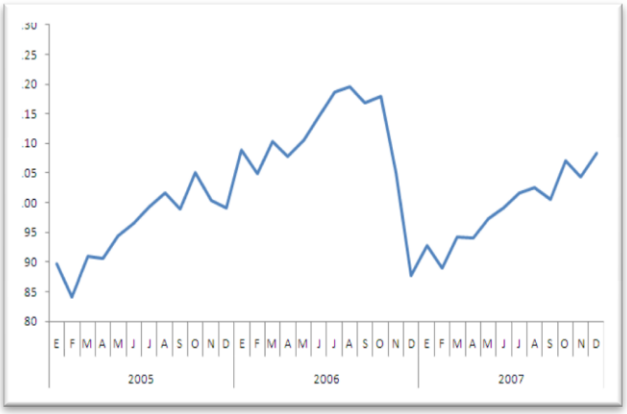


Cambio temporal

Cuando se registra una caída (o aumento) abrupto en el nivel que venía presentando la serie y, después de varios periodos regresa al nivel que se tenía anteriormente se trata de un *outlier* de *Cambio Temporal*. Esto se puede presentar por paros técnicos programados durante cierto tiempo, una vez que se terminan los paros la producción continua de forma habitual.

Cambio de nivel

Este tipo de *outlier* se observa cuando el nivel de la serie cambia de repente, es decir, cuando hay un incremento o disminución de todas las observaciones de un cierto punto en adelante en el tiempo. Después de este punto el comportamiento de la serie se mantiene sobre el nuevo nivel. El cambio de nivel se puede presentar debido a una crisis, la producción cae y empieza un nuevo comportamiento.



Para establecer el tipo de *outlier* que afecta a una serie, se analiza el comportamiento de la misma, dado que los *outliers* obedecen a factores externos a la naturaleza misma de la serie de tiempo. Un shock externo puede afectar de distinta forma a cada serie.

Una vez que se ha detectado para cada serie el tipo de transformación, los efectos de calendario, el modelo ARIMA y los *outliers* adecuados, se define el modelo con estas especificaciones para dar paso a la descomposición x11.

Para obtener la serie desestacionalizada, se obtiene primero la serie original corregida por efectos de calendario y *outliers*. De esta manera el punto de partida para la descomposición es la *serie original corregida por efectos del calendario*.

$$\begin{matrix} Serie\ original\ corregida\ por \\ efectos\ del\ calendario \end{matrix} = \frac{Serie\ original}{factores\ del\ calendario}$$

Los efectos de calendario están formados por los efectos de frecuencia de días y el efecto de la Semana Santa, por lo que el factor de calendario resulta de multiplicar ambos factores. Si sólo uno de ellos está presente, únicamente se ajusta por ese efecto; si no hay presencia de efectos de calendario, la serie original corregida por efectos de calendario y la serie original son la misma.

A su vez, esta serie es ajustada por *outliers*. De esta serie se obtiene una primera estimación de la tendencia-ciclo utilizando promedios móviles. Posteriormente, se determina un cociente entre la serie original corregida por efectos de calendario y *outliers* y la tendencia-ciclo estimada que es la primera estimación del componente estacional-irregular (SI por sus siglas en inglés).

$$SI_{inicial} = \frac{Serie\ original\ corregida\ por\ efectos\ de\ calendario\ y\ outliers}{tendencia\text{-}ciclo\ inicial}$$

A continuación, se obtienen dos series mediante la aplicación de filtros de promedios móviles al cociente SI inicial, uno compuesto y el otro simple; se divide el primero entre el segundo para obtener el factor estacional inicial.

La estimación inicial de la serie ajustada por estacionalidad (SA por sus siglas en inglés) se realiza dividiendo la serie original corregida por efectos de calendario y *outliers* entre el factor estacional inicial.

$$SA = \frac{Serie\ original\ corregida\ por\ efectos\ de\ calendario\ y\ outliers}{factor\ estacional\ inicial}$$

A la estimación inicial de la serie ajustada por estacionalidad se le aplica un filtro de promedios ponderados simétricos para obtener la tendencia-ciclo intermedia, el orden depende de la variación de la estimación inicial de la serie ajustada por estacionalidad.

Después se obtiene el componente estacional-irregular final.

$$SI_{final} = \frac{Serie\ original\ corregida\ por\ efectos\ de\ calendario\ y\ outliers}{tendencia\text{-}ciclo\ intermedia}$$

A continuación se obtienen dos series mediante la aplicación de filtros de promedios móviles sobre el cociente SI final, uno compuesto y el otro simple; se divide el primero entre el segundo para obtener el factor estacional final.

De esta forma, la *serie desestacionalizada* se obtiene mediante el cociente de la serie original corregida por efectos del calendario entre el factor estacional final:

$$Serie\ desestacionalizada = \frac{Serie\ original\ corregida\ por\ efectos\ del\ calendario}{Factor\ estacional\ final}$$

Para obtener la tendencia-ciclo final, se le aplica a la serie desestacionalizada un filtro de promedios ponderados simétricos, el orden depende de la variación de la serie desestacionalizada, por lo que puede ser diferente al utilizado para la tendencia-ciclo intermedia.

Actualizaciones

Una vez definidas las especificaciones del modelo, se realiza el ajuste estacional, posteriormente, conforme se van recibiendo las nuevas observaciones, se incluyen, y con ellas se reestiman los parámetros del modelo ARIMA, sin modificar las especificaciones, actualizando de esta forma la serie original corregida por efectos del calendario, la serie desestacionalizada y la tendencia-ciclo.

Las recomendaciones internacionales proponen una revisión anual del modelo para cada serie. Para los indicadores económicos de coyuntura que publica el INEGI, la revisión del modelo se hace por lo general cuando hay revisiones en los datos de la fuente original.