



□ filial de isa

CITese 004161-1 XM

MEDELLIN, ABR-23-2014 06:11 PM

ORIGEN : 6026 VIA : 4,2

# **PROCEDIMIENTO PARA DETERMINAR VALORES ATÍPICOS DE MODELO DE ESTIMACIÓN LBC**

**Abril 2014**

## PROCEDIMIENTO PARA DETERMINAR VALORES ATÍPICOS DE MODELO DE ESTIMACIÓN LBC

CITESE 004161-1 XM

MEDELLIN, ABR-23-2014 06:11 PM

ORIGEN : 6026 VIA : 4,2

### Captura y depuración de datos.

#### Captura:

La LBC tendrá una frecuencia diaria. La metodología se realizará sobre una ventana de tiempo de 105 días, correspondientes a las últimas quince semanas.

La verificación de los 105 días se hace teniendo en cuenta la mejor información disponible correspondiente a la última liquidación publicada por el ASIC.

#### Transformación de los valores atípicos e iguales a cero:

*Se identificará cada valor atípico o cero*, junto con el subíndice *i* correspondiente al día de la semana en que fue observado, de manera ascendente desde el dato más antiguo en fecha de operación al más reciente.

El procedimiento contiene los siguientes pasos:

1. Se identificarán los valores cero, iniciando de la fecha más antigua a la más reciente, y se reemplazarán por el promedio de los cinco días anteriores que tengan el mismo subíndice siempre y cuando corresponda con valores distintos de cero o ajustados con este procedimiento.

En el caso de no encontrarse la totalidad de los datos para los cinco días anteriores, se calculará el promedio con los días anteriores disponibles que tenga el mismo subíndice siempre y cuando correspondan con valores distintos de cero o ajustados con este procedimiento.

Un dato igual a cero será descartado de la serie si para el cálculo del promedio no hay datos anteriores distintos de cero o ajustados con este procedimiento.

2. Una vez ajustados los valores iguales a cero conforme a lo establecido en el paso 1, se realizará la identificación de los datos atípicos mediante el concepto de cuartil de un conjunto de datos, en el que se hace uso de la mediana, la cual es resistente a la presencia de datos atípicos.

Es de anotar que si el conjunto inicial de datos contiene ceros, estos habrán sido reemplazados en el paso 1, no obstante en el paso 2 se evalúa la totalidad de elementos del nuevo conjunto de datos (los datos originales diferentes de cero y los ajustados en el paso 1).

La identificación de datos atípicos mediante la utilización del concepto cuartil, así:

- a. Se ordena el conjunto de datos de forma ascendente para cada tipo de día *i*.



■ filial de isa

- b. Se identifican los cuartiles  $Q_1$ ,  $Q_2$  y  $Q_3$ , que determinan los valores correspondientes al 25 %, 50% y al 75% de los datos. Es de anotar que el cuartil  $Q_2$  coincide con la mediana del conjunto de datos.
- c. Se halla el rango intercuartil – IQR que es igual a  $Q_3$  menos  $Q_1$ .
- d. Los datos atípicos corresponden a aquellos  $X_k$  que:

$$X_k < Q_1 - 1.5 * IQR \text{ ó } X_k > Q_3 + 1.5 * IQR$$

Dónde:

$X_k$ : Dato k del conjunto

CITese 004161-1 XM

MEDELLIN, ABR-23-2014 06:11 PM

ORIGEN : 6026 VIA : 4,2

$Q_1$ : Cuartil 1

$Q_3$ : Cuartil 3

IQR: Rango intercuartil

En la imagen 1, se presenta la dispersión de datos mediante la utilización de un diagrama de cajas.

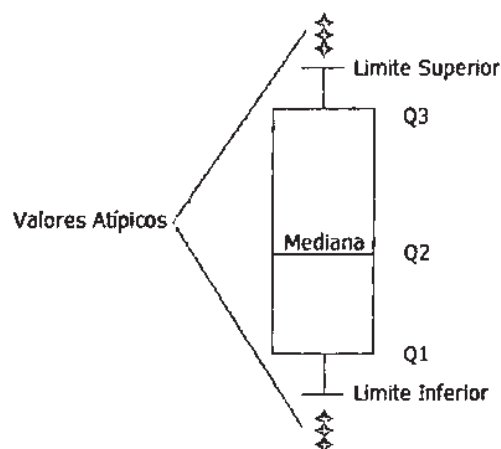


Imagen 1: Diagrama de cajas

Iniciando de la fecha más antigua a la más reciente, los valores atípicos, se transformarán por el promedio de los cinco días anteriores que tengan el mismo subíndice siempre y cuando corresponda con valores típicos o ajustados con este procedimiento.



En el caso de no encontrarse la totalidad de los datos para los cinco días anteriores, se calculará el promedio con los días anteriores disponibles que tenga el mismo subíndice siempre y cuando correspondan con valores típicos o ajustados.

Un dato atípico será descartado de la serie si para el cálculo del promedio no hay datos anteriores con valores típicos o transformados con este procedimiento.

Si alguno de los datos transformados en el paso 1 resulta ser atípico, en el paso 2 se transformará haciendo uso de lo mencionado en este paso.

CITese 004161-1 XM

MEDELLIN, ABR-23-2014 06:11 PM

ORIGEN : 6026 VIA : 4,2

