RMV

Missing values in a time series are estimated.

Notation

The following notation is used throughout this chapter unless otherwise stated:

$X = (X_1, \ldots, X_n)$	Original series
$\hat{X}_i$	Estimate for spans
p	Number of spans
k	The number of consecutive missing values
X_i to X_{i+k-1}	Set of consecutive missing values

Methods for Estimating Missing Values

Linear Interpolation (*LINT(X)*)

$$\hat{X}_{i+l} = \begin{cases} X_{i-1} + \frac{l+1}{k+1}\left(X_{i+k} - X_{i-1}\right) & l = 0, \ldots, k-1 \\ \text{SYSMIS} & i = 1 \text{ or } i + k - 1 = n \end{cases}$$

If $k = 1$ (that is, only one consecutive missing observation), then

$$\hat{X}_i = \begin{cases} \frac{1}{2}\left(X_{i-1} + X_{i+1}\right) & i = 2, \ldots, n-1 \\ \text{SYSMIS} & i = 1 \text{ or } i = n \end{cases}$$

Mean of *p* Nearest Preceding and *p* Subsequent Values *(MEAN (X,p))*

If the number of nonmissing observations in $\left(X_1,\ldots,X_{i-1}\right)$ or $\left(X_{i+k},\ldots,X_n\right)$ is less than p, then set $\hat{X}_{i+l} = \text{SYSMIS}$; otherwise, set $\hat{X}_{i+l} =$ average of p nonmissing observations preceding X_i and p nonmissing observations following X_{i+k-1}.

Median of *p* Nearest Preceding and *p* Subsequent Values *(MEDIAN (X,p))*

If the number of nonmissing observations in $\left(X_1,\ldots,X_{i-1}\right)$ or $\left(X_{i+k},\ldots,X_n\right)$ is less than p, then set $\hat{X}_{i+l} = \text{SYSMIS}$; otherwise, set $\hat{X}_{i+l} =$ median of p nonmissing observations preceding X_i and p nonmissing observations following X_{i+k-1}.

Series Mean *(SMEAN (X))*

$\hat{X}_{i+l} =$ average of all nonmissing observations in the series.

Linear Trend *(TREND(X))*

(1) Use all the nonmissing observations in the series to fit the regression line of the form

$$\hat{X}_t = a + bt$$

The least squares estimates are

$$b = \frac{\sum\left(X_t - \overline{X}\right)\left(t - \bar{t}\right)}{\sum\left(t - \bar{t}\right)^2}$$

$$a = \overline{X} - b\bar{t}$$

(2) Apply the regression equation to replace the missing values

$$\hat{X}_{i+l} = a + b(i + l)$$