ACF/PACF

Procedures ACF and PACF print and plot the sample autocorrelation and partial autocorrelation functions of a series of data.

Notation

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

x_i	*i*th observation of input series, $i = 1, \ldots, n$
r_k	*k*th lag sample autocorrelation
$\hat{\phi}_{kk}$	*k*th lag sample partial autocorrelation

Basic Statistics

The following formulas are used if no missing values are encountered. If missing values are present, see the section "Series with Missing Values" below for modification of some formulas.

Sample Autocorrelation r_k

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where $\bar{x}$ is the average of the *n* observations.

Standard Error of r_k

There are two formulas for the standard error of r_k based on different assumptions about the autocorrelation.

Under the assumption that the true MA order of the process is k-1, the approximate variance of r_k given by Bartlett (1946) is

$$\operatorname{var}(r_k) \cong \frac{1}{n}\left(1+2\sum_{l=1}^{k-1} r_l^2\right)$$

The standard error is the square root (see Box and Jenkins, 1976, p. 35). Under the assumption that the process is white noise,

$$\operatorname{var}(r_k) \cong \frac{1}{n}\left(\frac{n-k}{n+2}\right)$$

Box-Ljung Statistic

At lag k, the Box-Ljung statistic is defined by

$$Q_k = n(n+2)\sum_{l=1}^{k}\frac{r_l^2}{n-l}$$

When n is large, Q_k has a chi-square distribution with degrees of freedom k-p-q, where p and q are autoregressive and moving average orders, respectively. The significance level of Q_k is calculated from the chi-square distribution with k-p-q degrees of freedom.

Sample Partial Autocorrelation

$$\hat{\phi}_{11} = r_1$$

$$\hat{\phi}_{22} = \left(r_2 - r_1^2\right) \Big/ \left(1 - r_1^2\right)$$

$$\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad k = 2,\ldots, \quad j = 1,2,\ldots,k-1$$

$$\hat{\phi}_{kk} = \left(r_k - \sum_{j=1}^{k-1} \phi_{k-1,j} r_{k-j}\right) \Bigg/ \left(1 - \sum_{j=1}^{k-1} \phi_{k-1,j} r_j\right), \quad k = 3,\ldots$$

Standard Error of $\hat{\phi}_{kk}$

Quenouville (1949) has shown that, under the assumption that the AR(p) model is correct and $p \le k-1$,

$$\hat{\phi}_{kk} \cong N\left(0, \frac{1}{n}\right)$$

Thus

$$\operatorname{var}\left(\hat{\phi}_{kk}\right) \cong \frac{1}{n}$$

Series with Missing Values

If there are missing values in x, the following statistics are computed differently. The approach for handling missing values in ACF/PACF can be found in Cryer (1986). First, let us define

$$\bar{x} = \text{average of nonmissing } x_1, \ldots, x_n,$$

$$a_i = \begin{cases} x_i - \bar{x}, & \text{if } x_i \text{ is not missing} \\ \text{SYSMIS}, & \text{if } x_i \text{ is missing} \end{cases}$$

for $k = 0,1,2,\ldots,$ and $j = 1,\ldots,n$

$$b_j^{(k)} = \begin{cases} a_j a_{j+k}, & \text{if both are not missing} \\ \text{SYSMIS}, & \text{otherwise} \end{cases}$$

m_k = the number of nonmissing values in $b_1^{(k)},\ldots,b_{n-k}^{(k)}$.

m_0 = the number of nonmissing values in x.

Sample Autocorrelation

$$r_k = \frac{\text{sum of nonmissing } b_1^{(k)},\ldots,b_{n-k}^{(k)}}{\text{sum of nonmissing } b_1^{(0)},\ldots,b_n^{(0)}}$$

Standard Error of r_k

$$\text{se}(r_k) = \sqrt{\frac{1}{m_0}\left(1 + \sum_{l=1}^{k-1} r_l^2\right)} \quad \text{(MA assumption)}$$

$$\text{se}(r_k) = \sqrt{\frac{m_k}{(m_0 + 2)m_0}} \quad \text{(white noise)}$$

Box-Ljung Statistic

$$Q = m_0(m_0 + 2)\sum_{l=1}^{k}\frac{r_l^2}{m_l}$$

Standard Error of $\hat{\phi}_{kk}$

$$\text{se}\left(\hat{\phi}_{kk}\right) \cong \sqrt{\frac{1}{m_0}}$$

References

Bartlett, M. S. 1946. On the theoretical specification of sampling properties of autocorrelated time series. *Journal of Royal Statistical Society*, *Series B*, 8: 27.

Box, G. E. P., and Jenkins, G. M. 1976. *Time series analysis*: *Forecasting and control*. San Francisco: Holden-Day.

Cryer, J. D. 1986. *Time series analysis*. Boston: Duxbury Press.

Quenouville, M. H. 1949. Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B*, 11: 68.