STAT 804

Lecture 16 Notes

Distribution theory for sample autocovariances

The simplest statistic to consider is

$$\frac{1}{T}\sum X_t X_{t+k}$$

where the sum extends over those t for which the data are available. If the series has mean 0 then the expected value of this statistic is simply

$$\frac{T-k}{T}C_X(k)$$

which differs negligibly for T large compared to k from $C_X(k)$. To compute the variance we begin with the second moment which is

$$\frac{1}{T^2}\sum_s\sum_t \mathrm{E}(X_s X_{s+k} X_t X_{t+k})$$

The expectations in question involve the fourth order product moments of X and depend on the distribution of the X's and not just on C_X. However, for the interesting case of white noise, we can compute the expected value. For $k > 0$ you may assume that $s < t$ or $s = t$ since the $s > t$ cases can be figured out by swapping s and t in the $s < t$ case. For $s < t$ the variable X_s is independent of all 3 of X_{s+k}, X_t and X_{t+k}. Thus the expectation factors into something containing the factor $\mathrm{E}(X_s) = 0$. For $s = t$, we get $\mathrm{E}(X_s^2)\mathrm{E}(X_{s+k}^2) = \sigma^4$. and so the second moment is

$$\frac{T-k}{T^2}\sigma^4$$

This is also the variance since, for $k > 0$ and for white noise, $C_X(k) = 0$.

For $k = 0$ and $s < t$ or $s > t$ the expectation is simply σ^4 while for $s = t$ we get $\mathrm{E}(X_t^4) \equiv \mu_4$. Thus the variance of the sample variance (when the mean is known to be 0) is

$$\frac{T-1}{T}\sigma^4 + \mu_4/T - \sigma^4 = (\mu_4 - \sigma^4)/T\,.$$

For the normal distribution the fourth moment μ_4 is given simply by $3\sigma^4$.

Having computed the variance it is usual to look at the large sample distribution theory. For $k = 0$ the usual central limit theorem applies to $\sum X_t^2/T$ (in the case of white noise) to prove that

$$\sqrt{T}(\hat{C}_X(0) - \sigma^2)/\sqrt{\mu_4} \to N(0,1)\,.$$

The presence of μ_4 in the formula shows that the approximation is quite sensitive to the assumption of normality.

For $k > 0$ the theorem needed is called the m-dependent central limit theorem; it shows that

$$\sqrt{T}\hat{C}_X(k)/\sigma^2 \to N(0,1)\,.$$

In each of these cases the assertion is simply that the statistic in question divided by its standard deviation has an approximate normal distribution.

The sample autocorrelation at lag k is

$$\hat{C}_X(k)/\hat{C}_X(0)\,.$$

For $k > 0$ we can apply Slutsky's theorem to conclude that

$$\sqrt{T}\hat{C}_X(k)/\hat{C}_X(0) \to N(0,1)\,.$$

This justifies drawing lines at $\pm 2/\sqrt{T}$ to carry out a 95% test of the hypothesis that the X series is white noise based on the kth sample autocorrelation.

It is possible to verify that subtraction of $\bar{X}$ from the observations before computing the sample covariances does not change the large sample approximations, although it does affect the exact formulas for moments.

When the X series is actually not white noise the situation is more complicated. Consider as an example the model

$$X_t = \phi X_{t-1} + \epsilon_t$$

with ϵ being white noise. Taking

$$\hat{C}_X(k) = \frac{1}{T} X_t X_{t+k}$$

we find that

$$T^2 \mathrm{E}(\hat{C}_X(k)^2) = \sum_s \sum_t \sum_{u_1} \sum_{u_2} \sum_{v_1} \sum_{v_2} \phi^{u_1+u_2+v_1+v_2} \mathrm{E}(\epsilon_{s-u_1}\epsilon_{s+k-u_2}\epsilon_{t-v_1}\epsilon_{t+k-v_2})$$

The expectation is 0 unless either all 4 indices on the ϵ's are the same or the indices come in two pairs of equal values. The first case requires $u_1 = u_2 - k$ and $v_1 = v_2 - k$ and then $s - u_1 = t - v_1$. The second case requires one of three pairs of equalities: $s - u_1 = t - v_1$ and $s - u_2 = t - v_2$ or $s - u_1 = t + k - v_2$ and $s + k - u_2 = t - v_1$ or $s - u_1 = s + k - u_2$ and $t - v_1 = t - +k - v_2$ along with the restriction that the four indices not all be equal. The actual moment is then μ_4 when all four indices are equal and σ^4 when there are two pairs. It is now possible to do the sum using geometric series identities and compute the variance of $\hat{C}_X(k)$. It is not particularly enlightening to finish the calculation in detail.

There are versions of the central limit theorem called mixing central limit theorems which can be used for ARMA(p,q) processes in order to conclude that

$$\sqrt{T}(\hat{C}_X(k) - C_X(k))/\sqrt{\mathrm{Var}(\hat{C}_X(k))}$$

has asymptotically a standard normal distribution and that the same is true when the standard deviation in the denominator is replaced by an estimate. To get from this to

distribution theory for the sample autocorrelation is easiest when the true autocorrelation is 0.

The general tactic is the δ method or Taylor expansion. In this case for each sample size T you have two estimates, say N_T and D_T of two parameters. You want distribution theory for the ratio $R_T = N_T/D_T$. The idea is to write $R_T = f(N_T, D_T)$ where $f(x,y) = x/y$ and then make use of the fact that N_T and D_T are close to the parameters they are estimates of. In our case N_T is the sample autocovariance at lag k which is close to the true autocovariance $C_X(k)$ while the denominator D_T is the sample autocovariance at lag 0, a consistent estimator of $C_X(0)$.

Write

$$\begin{aligned} f(N_T, D_T) &= f(C_X(k), C_X(0)) \\ &\quad +(N_T - C_X(k))D_1 f(C_X(k), C_X(0)) \\ &\quad +(D_T - C_X(0))D_2 f(C_X(k), C_X(0)) + \text{remainder} \end{aligned}$$

If we can use a central limit theorem to conclude that

$$(\sqrt{T}(N_T - C_X(k)), \sqrt{T}(D_T - C_X(0)))$$

has an approximately bivariate normal distribution and if we can neglect the remainder term then

$$\sqrt{T}(f(N_T, D_T) - f(C_X(k), C_X(0))) = \sqrt{T}(\hat{\rho}(k) - \rho(k))$$

has approximately a normal distribution. The notation here is that D_j denotes differentiation with respect to the jth argument of f. For $f(x,y) = x/y$ we have $D_1 f = 1/y$ and $D_2 f = -x/y^2$. When $C_X(k) = 0$ the term involving $D_2 f$ vanishes and we simply get the assertion that

$$\sqrt{T}(\hat{\rho}(k) - \rho(k))$$

has the same asymptotic normal distribution as $\hat{C}_X(k)/C_X(0)$.

Similar ideas can be used for the estimated sample partial ACF.

Portmanteau tests

In order to test the hypothesis that a series is white noise using the distribution theory just given, you have to produce a single statistic to base youre test on. Rather than pick a single value of k the suggestion has been made to consider a sum of squares or a weighted sum of squares of the $\hat{\rho}(k)$.

A typical statistic is

$$T \sum_{k=1}^{K} \hat{\rho}^2(k)$$

which, for white noise, has approximately a χ^2_K distribution. (This fact relies on an extension of the previous computations to conclude that

$$\sqrt{T}(\hat{\rho}(1), \ldots, \hat{\rho}(K))$$

has approximately a standard multivariate distribution. This, in turn, relies on computation of the covariance between $\hat{C}(j)$ and $\hat{C}(k)$.)

When the parameters in an ARMA(p, q) have been estimated by maximum likelihood the degrees of freedom must be adjusted to $K - p - q$. The resulting test is the Box-Pierce test; a refined version which takes better account of finite sample properties is the Box-Pierce-Ljung test. S-Plus plots the P-values from these tests for 1 through 10 degrees of freedom as part of the output of *arima.diag*.