Lecture 9

Sampling Distributions of Statistics
Textbook Sections: 2.4

Sample Estimates

Recall that given observations $x_1, x_2, \cdots, x_n$, we can compute estimates of μ, $\gamma(h)$ and $\rho(h)$ as

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t,$$

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_t - \bar{x})(x_{t+h} - \bar{x}),$$
$$\hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0),$$

where $h = 0, 1, \cdots, n-1$.

Sample Mean

The expectation and variance of $\bar{X}$ are

$$
\begin{aligned}
E(\bar{X}) &= \mu, \\
Var(\bar{X}) &= \tau_n^2/n, \text{ where} \\
\tau_n^2 &= \sum_{h=-(n-1)}^{n-1} (1 - |h|/n)\gamma(h) = \gamma(0) + 2\sum_{h=1}^{n-1}(1 - h/n)\gamma(h).
\end{aligned}
$$

If $\gamma(h)$ is negligible for large h, then it can be shown that

$$\tau_n^2 \approx \gamma(0) + 2\sum_{h=1}^{\infty}\gamma(h) = \gamma(0)[1 + 2\sum_{h=1}^{\infty}\rho(h)].$$

Though the last expression looks formidable, it can have simple expressions in some cases (Examples 1 and 2 below).

Now let's consider the distribution of $\bar{X}$.

1. If $\{X_t\}$ are normally distributed, then $\bar{X} \sim N(\mu, \tau_n^2/n)$.
2. For many processes, $\bar{X}$ is still approximately normally distributed, especially when n is large. In particular, this is true for $ARMA(p, q)$ processes, which we will use as models.

A consequence of this result is that we can construct approximate confidence intervals for μ. For instance if we have an estimate $\hat{\tau}_n$ of τ_n, then an approximate 95% confidence interval for μ is given by $\bar{X} \pm 1.96\hat{\tau}_n/\sqrt{n}$.

We can estimate τ_n using estimates of the autocovariances as

$$\hat{\tau}_n^2 = \hat{\gamma}(0) + 2\sum_{h=1}^{L}(1 - h/n)\hat{\gamma}(h) = \hat{\gamma}(0)[1 + 2\sum_{h=1}^{L}(1 - h/n)\hat{\rho}(h)],$$

where L is some (large) integer. For example, we can use $L = \sqrt{n}$ or the integer part of $\sqrt{n}$. Sometimes the value of L can be guessed by looking at the ACF plot. Choose L so that all the sample autocorrelations of order $L+1$ or higher seem to be negligible. [Note that we have used the relation $\hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0)$, and hence $\hat{\gamma}(h) = \hat{\gamma}(0)\hat{\rho}(h)$.]

Special Cases

In some cases it may be possible to find simple expressions for τ_n if there is some reasonably simple model for the sequence $\{X_t\}$.

Example 1.
If $\{X_t\}$ can be described by a moving average model, then estimation of τ_n is particularly simple. If we have an MA(1) model, then all the autocovarianes (and autocorrelations) of order 2 or higher are zero. For an $MA(2)$ model, all the autocovarianes (and autocorrelations) of order 3 or higher are zero. Similarly, for an $MA(3)$ model, all the autocovarianes (and autocorrelations) of order 4 or higher are zero.

$$\begin{aligned}
MA(1): \tau_n^2 &= \gamma(0) + 2(1-1/n)\gamma(1),\\
\hat{\tau}_n^2 &= \hat{\gamma}(0) + 2(1-1/n)\hat{\gamma}(1)\\
MA(2): \tau_n^2 &= \gamma(0) + 2(1-1/n)\gamma(1) + 2(1-2/n)\gamma(2),\\
\hat{\tau}_n^2 &= \hat{\gamma}(0) + 2(1-1/n)\hat{\gamma}(1) + 2(1-2/n)\hat{\gamma}(2),\\
MA(3): \tau_n^2 &= \gamma(0) + 2(1-1/n)\gamma(1) + 2(1-2/n)\gamma(2) + 2(1-3/n)\gamma(3),\\
\hat{\tau}_n^2 &= \hat{\gamma}(0) + 2(1-1/n)\hat{\gamma}(1) + 2(1-2/n)\hat{\gamma}(2) + 2(1-3/n)\hat{\gamma}(3).
\end{aligned}$$

Example 2.
If $\{X_t\}$ is $AR(1)$, i.e.,

$$X_t - \mu = \phi(X_{t-1} - \mu) + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is white noise with variance σ^2, then we know that

$$\gamma(0) = Var(X_t) = (1-\phi^2)^{-1}\sigma^2, \rho(h) = \phi^h.$$

In such a case

$$\tau_n^2 \approx \gamma(0)[1 + 2\sum_{h=1}^{\infty}\rho(h)] = (1-\phi^2)^{-1}\sigma^2[1 + 2\sum_{h=1}^{\infty}\phi^h] = \sigma^2/(1-\phi)^2.$$

So if we have estimates of ϕ and σ, then an estimate of τ_n^2 is given by

$$\hat{\tau}_n^2 = \hat{\sigma}^2/(1-\hat{\phi})^2.$$

[Technical details: Note that

$$\sum_{h=1}^{\infty}\phi^h = \phi + \phi^2 + \cdots = \phi(1 + \phi + \phi^2 + \cdots) = \phi(1-\phi)^{-1}.$$

Hence using some algebra we can find that $\tau_n^2 \approx (1-\phi^2)^{-1}\sigma^2[1 + 2\phi(1-\phi)^{-1}] = \sigma^2/(1-\phi)^2$.]

Remark:
Typically, models are not known in advance. For this reason, it is often reasonable to estimate τ_n^2 using the formula with L above.

Sample ACF

It is known that, for large n, the distribution of $\hat{\rho}(h)$, $h \geq 1$, is approximately normal with mean $\rho(h)$ and variance w_{hh}/n, where

$$w_{hh} = \sum_{k=1}^{\infty} [\rho(k+h) + \rho(k-h) - 2\rho(h)\rho(k)]^2.$$

So if we have an estimate $\hat{w}_{hh}$ of w_{hh}, then an approximate 95% confidence interval of $\rho(h)$ is

$$\hat{\rho}(h) \pm 1.96 s(\hat{\rho}(h)), \text{ where } s(\hat{\rho}(h))^2 = \hat{w}_{hh}/n.$$

We have already discussed that in order to check if $\rho(h) = 0$, we look at the sample ACF plot to find out if $\hat{\rho}(h)$ is inside the interval $\pm 1.96/\sqrt{n}$. If $\hat{\rho}(h)$ is inside this interval, then it indicates that $\rho(h)$ may be close to zero.

In general obtaining an estimate of w_{hh} using the estimated autocorrelations requires a bit of care. However, for a few simple cases given below, it is possible to find nice expressions for w_{hh} and these espressions can be used to obtain estimate for w_{hh}.

Example 3.
Suppose that the sequence $\{X_t\}$ is i.i.d. with mean μ and variance σ^2. In this case, $\rho(h) = 0$ for $h = 1, 2, \ldots$. The quantity $w_{hh} = 1$ for any integer h. As a matter of fact, $\hat{\rho}(1), \hat{\rho}(2), \ldots$ are all independent and each is approximately normally distributed with mean zero and variance $1/n$. This result is behind the Portmanteu and Box-Ljung tests.

Example 4.
Suppose that the sequence $\{X_t\}$ has mean μ and follows an $MA(1)$ sequence,

$$X_t - \mu = \varepsilon_t + \theta\varepsilon_{t-1},$$

where $\{\varepsilon_t\}$ is white noise with variance σ^2. Then $\rho(1) = \theta/(1+\theta^2)$ and $\rho(2)$, $\rho(3), \rho(4), \ldots$ are all zero. In this case, simple mathematical calculations will show

$$w_{hh} = \begin{cases} 1 - 3\rho(1)^2 + 4\rho(1)^4 & h = 1 \\ 1 + 2\rho(1)^2 & h = 2, 3, \ldots. \end{cases}$$

So in this case, $\hat{\rho}(1)$ is approximately normally distributed with mean $\rho(1)$ and variance w_{11}/n. However, for any $h = 2, 3, \ldots$, $\hat{\rho}(h)$ is approximately normally distributed with **mean zero** and variance $(1 + 2\rho(1)^2)/n$.

Example 5.
If the sequence $\{X_t\}$ with mean μ follows an $AR(1)$ model, i.e., $X_t - \mu = \phi(X_{t-1} - \mu) + \varepsilon_t$, where $\{\varepsilon_t\}$ is white noise with variance σ^2, then $\gamma(h) = \gamma(0)\phi^h$ and $\rho(h) = \phi^h$, $h = 0, 1, \ldots$. In this case, $\hat{\rho}(h)$ is approximately normally distributed with mean $\rho(h) = \phi^h$ and variance

$$w_{hh} = (1 - \phi^{2h})(1 + \phi^2)(1 - \phi^2)^{-1} - 2h\phi^{2h}.$$

The expression for w_{hh} is obtained by using the expressions for $\{\rho(j)\}$ in the AR(1) case.