Econ 83A
Statistics and Economic Analysis

Prof. Linda T.M. Bui

Spring 2005

Confidence Intervals and Hypothesis Testing

I. Introduction:

Suppose that we assume that a random sample of size n is drawn from a population that is normally distributed such that:

(1) $$X_i \sim N(\mu, \sigma^2)$$

Then, by the Central Limit Theorem, we know that the sample mean will have the following properties:

(2) $$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right),$$

Note *carefully* that the *estimator* ($\bar{X}$) is a *random variable*, although the *true* population parameter (μ) is not. But, given the Central Limit Theorem, we can construct a *confidence interval* for the estimate that we observe.

II. Constructing Confidence Intervals:

A. Known σ^2

1. Standardizing the Estimator

If the population variance is *known*, then ($\bar{X}$) can be transformed in a fashion such that the transformed random variable has a *standard* normal (denoted N(0, 1)) distribution. Specifically, if

(3a) $$W \sim \mathrm{N}(\mu, \Sigma^2),$$

then the *transformed* random variable

(3b) $$Z \equiv \frac{W - \mu}{\Sigma} \sim \mathrm{N}(0, 1).$$

For the sample mean, whose probability distribution is given by (2), the corresponding transformation is:

(4) $$(a) \qquad Z_{\bar{X}} \equiv \frac{\bar{X} - \mu}{SE(\bar{X})} \sim \mathrm{N}(0,1)$$

Some things to note:

(1) The denominator of $Z_{\bar{X}}$ is just the *Standard Error* of the estimator (as shown in (2)) using the known σ^2.
(2) It should be easy for you to show that $Z_{\bar{X}}$ has expectation zero and variance one.

(3) $Z_{\bar{X}}$ is usually called a "z-statistic".

2. Constructing a Confidence Interval

Using the transformed variable, $Z_{\bar{X}}$, we can use the *standard normal* tables at the back of WONNACOT AND WONNACOT (or any other statistics or econometrics text) to create confidence intervals around the population parameter, μ in the following way:

Step 1: Recall that since the standard normal distribution is just a probability distribution, the area under its probability distribution function is equal to one. The values given in the standard normal tables at the back of the text tell you that the *cumulative probability* that Z (a standard normal *random variable*) exceeds *z* (any given number between $\pm\infty$). So, for example, the probability that Z is greater than 1.96 is 0.025 or 2.5%, and the probability that Z is greater than 0 is 0.500, or 50% (which should not surprise you, since the standard normal distribution is symmetric and centered around 0). These numbers, together with a couple of others, which you should first *verify* in the standard normal table in your text, and should then *know*, include:

$$Prob[Z \geq 0] = 0.5000;$$

$$Prob[Z \geq 1] = 0.1587;$$

$$Prob[Z \geq 1.96] = 0.0250;$$

$$Prob[Z \geq 2.57] = 0.005.$$

Some additional things to note:

(1) Each of the numbers above is a *right* (or *upper*) tail probability. The probability that a standard normal random variable is *both* greater than some number *z and* less than -*z* -- that is, the probability that it is in *either* the right *or* left (lower) tail -- will be twice the given numbers. So the probability that an N(0, 1) random variable is > 1 *and* < -1 is about 0.3174 (or a little less than 32 percent).

(2) Since the variance of Z has been standardized to one, its standard deviation is also one. Thus, a z-value chosen in constructing a test using the N(0, 1) distribution has a natural interpretation, in terms of the probability that a standard normal random variable will fall so many "standard deviations away from zero." To use the numbers in the preceding paragraph, there is about a 32 percent probability that an N(0, 1) random variable will fall more than *one* standard deviation above *or* below zero. (You should also verify that there is a 5 percent probability that an N(0, 1) random variable will fall more than approximately *two* standard deviations above or below zero.)

Step 2: To create a confidence interval we have to decide upon (*i.e.*, we *choose*) the *confidence level* (or, equivalently, the *significance level*) that we want to employ. Let γ be the *significance level* and 1 - γ be the *confidence level*, where γ is a number between 0 and 1. (If it's confusing, just think of a 5 percent significance and a 95 percent confidence level.) Then, to create a (1 - γ) percent (95 percent) confidence interval for a standard normal random variable, we need to find the z-critical values such that $\gamma/2$ percent (2.5 percent) of the area under the probability distribution function (or "probability weight") is found in each tail of the distribution. That is, we want to require that

$$Prob\left(Z < -z_{\gamma/2}\right) = \gamma/2 = Prob\left(Z > z_{\gamma/2}\right),$$

or, equivalently, that

$$Prob\left(-z_{\gamma/2} < Z < z_{\gamma/2}\right) = 1 - \gamma.$$

Thus, to test an hypothesis at the 5 percent *significance level* ($\gamma = 0.05$), equivalent to constructing a 95 percent confidence interval ($1 - \gamma = 1 - 0.05 = 0.95$), we would choose the z-critical value corresponding to $\gamma/2 = 0.025$, which as given above is 1.96, so that

$$Prob(-1.96 < Z < 1.96) = 0.95.$$

In other words, we can say with 95 percent confidence that an N(0, 1) random variable will fall within ± two standard deviations around zero.

3. Confidence Intervals for the Population Mean

From the above discussion, to create a (1 - γ) percent confidence interval for the true μ from the sample mean, we need to find the z-critical values such that:

(5a)
$$Prob\left(-z_{\gamma/2} < Z_{\bar{X}} < z_{\gamma/2}\right] = 1 - \gamma,$$

or, using (4), that

(5b)
$$Prob\left(-z_{\gamma/2} < \frac{\bar{X} - \mu}{SE(\bar{X})} < z_{\gamma/2}\right) = 1 - \gamma,$$

or, equivalently,

(5c)
$$Prob\left(\bar{X} - z_{\gamma/2} SE(\bar{X}) < \mu < \bar{X} + z_{\gamma/2} SE(\bar{X})\right) = 1 - \gamma.$$

The (1 - γ) percent confidence interval around μ can also be written: $\bar{X} \pm z_{\gamma/2} SE(\bar{X})$.

4. An Example

Suppose that you estimate from that $\bar{X} = 4.3$, and that the standard error of the estimator is 2.1 (found using the *known* variance of the population disturbance term). To create a 95% confidence interval around μ, we know that we need first to find the z-critical values such that 2.5% of the distribution lies in each of tail. Those values (which you should *already* have committed to memory) are ± 1.96. So, the 95% confidence interval around μ can be written as

$$4.3 \pm 1.96 * 2.1 = 4.3 \pm 4.12$$

$$= (0.18,\ 8.42).$$

If we had wanted to construct a *99%* confidence interval around μ, we would use the z-critical value ± 2.57, producing a confidence interval (which you should verify) of (-1.1, 9.7). Observe that if we want an interval with a *higher* confidence level, the interval is larger, increasing the probability that the random variable falls within that interval.

B. Unknown σ^2

The preceding discussion is a little unrealistic, in that the true population variance is typically *not* known. When it isn't, we are compelled to use the (unbiased) estimator for σ^2:

(6)
$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}.$$

Now, we can no longer transform $\bar{X}$ into a standard normal variable. We *can*, however, transform it into a random variable that follows the Student t-distribution. In class we showed that this can be done using the following transformation:

(7)
$$t_{\bar{X}} \equiv \frac{\bar{X} - \mu}{\hat{SE}(\bar{X})} \sim t_{n-1} \textit{ where}$$

(8)
$$\hat{SE}(\bar{X}) \equiv \sqrt{\frac{\hat{\sigma}^2}{n}}$$

More things to note:

(1) The denominator for $t_{\bar{X}}$ is just the *standard error* of the estimator using the unbiased estimator for σ^2.
(2) $t_{\bar{X}}$ is called a "t-statistic".
(3) The t-critical values depend upon the degrees of freedom.

Creating confidence intervals using the t-distribution is carried out in *exactly* the same way as with the N(0, 1) distribution, except that you must now use the t-table to find the appropriate critical values.

C. Interpretation

For each estimator, so long as we know what distribution the estimator follows, we can create the $(1 - \gamma)$ percent confidence region around the *true* parameter. The way to interpret the confidence interval is that $(1 - \gamma)$ percent of the time, the confidence interval will contain the true parameter. It is *not* the same thing to say that $(1 - \gamma)$ percent of the time the true parameter will fall in this interval. The true parameter is *not* a random variable. Recall that it is the *confidence interval* that is random.

III. Hypothesis Testing:

A. The Confidence Interval Approach

The first step in hypothesis testing is to define the *null* hypothesis, denoted as H_0. The null hypothesis is usually tested against an *alternative* hypothesis, denoted as H_1. The null hypothesis is constructed in such a way so that it is desirable to reject (statistically) the hypothesis. The reason that hypothesis testing is set up this way is because we can never "accept" a hypothesis as being true -- we don't know what actually is the "truth". Therefore, we can only reject a hypothesis.

For each of the estimated coefficients found using ordinary least squares, we know that we can construct confidence intervals. We can also do hypothesis testing on each of the random variables. That is, we can test whether or not the true value of the parameter equals a particular number (remember, we don't actually know what that value is). For example, we can test the null hypothesis that $\mu = \mu_0$ against the appropriate alternative hypotheses. Hypothesis testing may be done by constructing confidence intervals (as described above) around the population parameters and asking whether or not the observed (estimated) coefficient is consistent with the null hypothesis. For example, consider the following hypothesis:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

(9)

Now we need to determine whether or not the estimated μ, $(\bar{X})$ is consistent with the null hypothesis. To do this, we can first construct a confidence interval around μ with a confidence level of $(1 - \gamma)$ percent. We know that with $(1 - \gamma)$ percent confidence that the true μ will be contained in this interval. Consequently, if μ_0 does *not* fall in this confidence interval, we can say that with $(1 - \gamma)$ percent confidence we may reject the null hypothesis that $\mu = \mu_0$. If μ_0 does fall within the confidence interval then we *cannot reject* the null hypothesis.

B. The Test-of-Significance Approach

An equivalent method for doing hypothesis testing is based upon the "test-of-significance" approach. This procedure relies upon the creation of a "test statistic" and the sampling distribution of the statistic under the null hypothesis. The test statistic is created using the *observed* data. Recall that with unknown σ^2 we can transform the distribution of our estimator for μ into a t-distribution by using (7). If the *true* value of μ is specified under the null hypothesis then the value for $t_{\bar{X}}$ can be calculated from the sample data. Therefore, $t_{\bar{X}}$ can serve as a test statistic. We know what distribution $t_{\bar{X}}$ follows under the null hypothesis and it can be calculated using the sample data. Since we know the distribution that the test statistic follows, we can create a confidence interval around the test statistic. To see this, consider the following hypothesis test:

(10)
$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Under the null hypothesis that $\mu = \mu_0$, we can create the following test statistic:

(11)
$$t_{\bar{X}} \equiv \frac{\bar{X} - \mu_0}{\hat{SE}(\bar{X})} \sim t_{n-1}$$

and since $t_{\bar{X}}$ follows a t-distribution under the null hypothesis, we can create the $(1 - \gamma)$ percent confidence interval around $t_{\bar{X}}$:

(12a)
$$Prob\left(-t_{\gamma/2} < \frac{\bar{X} - \mu_0}{SE(\bar{X})} < t_{\gamma/2} \right) = 1 - \gamma,$$

which we can re-arrange in the following manner:

(12b)
$$Prob\left(\mu_0 - t_{\gamma/2} SE(\bar{X}) < \bar{X} < \mu_0 + t_{\gamma/2} SE(\bar{X}) \right) = 1 - \gamma.$$

This confidence interval gives the interval in which $\bar{X}$ will fall with probability $(1 - \gamma)$ *given* $\mu = \mu_0$. That is, *given* the hypothesized value of the true μ ($\mu = \mu_0$), with $(1 - \gamma)$ percent confidence the estimated value should fall within this confidence interval.

Notice how this relates to hypothesis testing using the confidence interval method. The confidence interval method creates confidence intervals around the *true* but unknown value of the parameter and tests whether or not with some reasonable level of certainty the true value falls within the confidence region. The test-statistic method constructs confidence intervals around the *estimated* value of the parameter and tests whether or not with some reasonable level of certainty the estimate falls within the (hypothesized) confidence region.

In practice, there is no need to actually estimated (12b) directly. We can compute $t_{\bar{X}}$ directly given the null hypothesis and see whether or not it lies between the t-critical values. If the t-statistic is larger than the t-critical value, we should reject the null hypothesis:

(13) $$Prob\left(-t_{\gamma/2} < t_x < t_{\gamma/2}\right) = 1 - \gamma,$$

that is, if $t_{\bar{X}}$ falls outside the confidence interval bounded above and below in (18) by the t-critical values, with (1 - γ) percent confidence we can reject the null hypothesis that $\mu = \mu_0$.

C. One-Tail Tests:

So far, we have discussed the construction of two-tailed confidence intervals and hypothesis testing where the alternative hypothesis has been one of "negation". The one-sided (1 - γ) percent confidence interval may be set up by finding the *single* critical value such that either:

(14a) $$Prob\left[t \geq t_\gamma\right] = 1 - \gamma;$$

where (1 - γ) percent of the distribution lies to the right of the critical value or

(14b) $$Prob\left[t \leq t_\gamma\right] = 1 - \gamma;$$

where (1 - γ) percent of the distribution lies to the left of the critical value.

Examples of a One-Sided Tests for μ:

A. Suppose that we wanted to test the following hypothesis:

(15) $$H_0: \mu \geq \mu_0$$
$$H_1: \mu < \mu_0$$

Now we must create a one-sided test. With a (1 - γ) confidence level (or γ level of significance) we want to find the t-critical value such that γ percent of the distribution falls in the *left* tail of the distribution. Note that the t-critical value depends upon the *alternative* hypothesis and not the null hypothesis. The alternative says that μ is *less than* a particular number (μ_0) -- that is, it falls to the left of μ_0. We only wish to reject the null hypothesis if we have a test statistic that falls to the left of the critical value γ percent of the time. If γ is 5% and n = 1000 we can use the t-critical value of -1.645. Therefore, we will reject the null hypothesis if:

$$t_{\bar{X}(n-1)} = \frac{\bar{X} - \mu_0}{\hat{SE}(\bar{X})} < -1.65$$

(16)

The rejection criteria for the one-sided test is then:

Case A:

$H_0 : \mu \leq \mu_0$

$H_1 : \mu > \mu_0$

Reject the null hypothesis if the test statistic > +ve critical value with α in the tail.

Case B:

$H_0 : \mu \geq \mu_0$

$H_1 : \mu < \mu_0$

Reject the null hypothesis if the test statistic < -ve critical value with α in the tail.

B. Sample Question:

NOTE: You should feel comfortable doing these types of calculations without having to consult a "formula".

Suppose that $\overline{X} = 0.5091$, the estimated standard error of $\overline{X} = 0.0357$ and there are 8 degrees of freedom. Calculate the following:

(a) Create the 95% confidence interval around μ.

(b) Test the hypothesis that $\mu = 0$ at a 5% significance level against the alternative hypothesis that $\mu \neq 0$.

(c) Test the hypothesis that $\mu = 0.3$ at a 5% significance level against the alternative hypothesis that $\mu \neq 0.3$.

(d) Test the hypothesis that $\mu \leq 0.3$ at a 5% significance level against the alternative hypothesis that $\mu > 0.3$.

Solution:

a. The t-critical values for 8 degrees of freedom and $\gamma/2 = 2.5\%$ are 2.306 and -2.306. The 95% confidence interval around μ is given by:

$0.5091 \pm 2.306*(0.0357) = [0.4268, 0.5914]$.

b. To test the null hypothesis that $\mu = 0$ at a 5% significance I just need to notice that 0 doesn't fall within my 95% confidence interval calculated in (a) so I know that I can reject the null hypothesis.

c. To test the null hypothesis that $\mu = 0.3$, I can create the t-statistic:

$t_{\overline{X}} = (0.5091 - 0.3)/0.0357 = 5.86 > 2.306$, therefore I can reject the null hypothesis that $\mu = 0.3$ with 5% significance.

d. To construct the one-sided test of the hypothesis that $\mu \leq 0.3$ against the alternative that $\mu > 0.3$ I need to find the t-critical value such that 5% of the distribution lies to the *right* of the distribution. The t-critical value with 8 degrees of freedom is 1.86. The t-statistic is as above (5.86) which is greater than the t-critical value of 1.86, therefore, I reject the null hypothesis.