POL 502
Lecture 5
February 24, 2005

Hypothesis Testing – Chapter 5 in Studenmund.

We know how to get an estimate of the relationship in the sample, $\hat{\beta}$.

BUT statistical inference is what we are really interested in.
The weights of 9 students or the conflicts of 10 countries are not what we want to explain.
We want to be able to make generalizable statements about the variables in the entire population.

The regression model for the entire population is:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

(We switch to Greek when talking about the model as a whole.)
We can't be sure about this equation, we only use our sample to guess what it is.

Y is the dependent variable
X is independent
β_0 & β_1 are unknown PARAMETERS we seek to estimate.
ε are the random error fluctuations.

In our sample, we estimate $\hat{\beta}_0$ & $\hat{\beta}_1$.
For different samples, we are going to get different values for $\hat{\beta}_0$ & $\hat{\beta}_1$.
We want to know how our many estimates from our many samples of $\hat{\beta}_0$ & $\hat{\beta}_1$ would be distributed around β_0 & β_1.

This is how we test our hypothesis.

When we make a hypothesis in regression, what are we hypothesizing?

About β_1: usually that it is positive or negative or just non-zero.

If there is no relationship between two variables, β_1=0.

We use our sample to generate $\hat{\beta}_1$ and then use that to tell us what β_1 might be.
If we can conclude it is positive or negative or non-zero, we reject null hypothesis of no relationship.

Our null Hyp: β_1=0

If we say $\beta_1 \neq 0$, two tailed hypothesis.

If we say β_1>0 or β_1<0, we are making an one tail hypothesis.

We use the t-distribution to test whether the observed $\widehat{\beta}_1$ differs significantly from the population parameter β_1.

Remember, the t distribution is like the normal but not quite.

with high n, the same,
with low n, fatter tails.

$$t = \frac{\widehat{\beta}_1 - \beta_1}{SE_{\widehat{\beta}}}$$

But, since we hypothesize that $\beta_1 = 0$,

$$t = \frac{\widehat{\beta}_1}{SE_{\widehat{\beta}}}$$

Degrees of freedom for t = N - K – 1

$SE_{\widehat{\beta}}$ is the estimate of the standard error of the regression coefficient.

Also, theoretically, it is the standard deviation of the Sampling Distribution of $\widehat{\beta}_1$.

95% Confidence intervals for β_1:

$$\widehat{\beta}_1 \pm t_{.025}(SE_{\widehat{\beta}_1})$$

$$\widehat{\beta}_0 \pm t_{.025}(SE_{\widehat{\beta}_0})$$

$$SE(\widehat{\beta}_1) = \sqrt{\frac{\sum e_i^2/(n-2)}{\sum (X_i - \overline{X})^2}}$$

Examine this:

The more error there is, the higher the standard error is.

The more cases there are the lower the standard error is.

The more variance there is in X, the lower the standard error is.

SE is average miss for $\widehat{\beta}$.

$$SE(\widehat{\beta_0}) = \sqrt{\sum e_i^2 / n-2} * \sqrt{\frac{\left(\frac{1}{n}+\overline{X}\right)^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}}$$

Our hypothesis is about β – slope in population.

Our hypothesis test is about $\widehat{\beta}_1$, is it significantly different from zero.

Steps in the hypothesis test are the same as they were with z tests, t tests, from last term. All we are doing here is another type of t test.
So, what were the steps we learned?
1. Advance a research hypothesis: H_A and a Null hypothesis: H_0.

Our null hypothesis and alternative hypothesis are complements – meaning that one and only one must be true.

We concentrate on the null hypothesis knowing that if we prove it is wrong, we are thereby supporting our alternative hypothesis.

Make the hypothesis directional if possible.

If you use a level of significance other than .05 – justify why you are doing so.

This value is the probability of being wrong by chance alone due to sampling error.

Since we know that it is possible to have an *unrepresentative* sample we want to take into account its possibility.

We can quantify exactly how possible it is that our results come from an unrepresentative sample.

This is what our probability level is.

It tells us the probability that we could be wrong by chance alone.

If we set our level at .05 (which is standard) and find a relationship exists we are still admitting that there is a 5% chance that although we found a relationship in the sample, the relationship might not exist in the population.

The lower the probability level that we set the lower the chance that we are wrong by chance alone.

Thus, at the .01 level there is only a 1% chance that the relationship we find in the sample isn't true of the population as a whole and at the .001 level there is only a 0.1% chance.

Why don't we always try for the highest level of probability?

Because there is a trade-off: The lower our probability level the harder it is to say that we have found something. But the harder a goal we set for ourselves, the better the value of attaining it is.

For example: We might be able to say that we accept our hypothesis at the .05 level – thereby allowing the 5% chance that we are wrong in doing so. But the relationship in our sample would have to be stronger to be able to accept the hypothesis at the .01 level.

2. Choose the test that is appropriate based upon what data you have available.

E.g. If data are arrayed in tabular form and we are looking at 2 variables, we want to know if the population is the same as the sample, Chi-square test is appropriate.

If we can do a regression, then t is appropriate.

3. Calculate the test statistic.

This will depend on which statistic we are using.

4. Locate the appropriate critical value.

To get the critical value you will need to know:

A. The statistic you are interested in
B. The degrees of freedom (based on the size of your sample and sometimes the number of independent variables)
C. The desired probability level (usually .05).
D. Whether your hypothesis is one tailed (directional – positive or negative) or two tailed (non-directional).

5. Compare the test statistics you calculated with the appropriate critical value.

6. If your test statistic is less than the critical value you "Fail to reject the null hypothesis." You cannot say that the results are statistically significant.

7. If your test statistic > the critical value you "Reject the null hypothesis and provisionally accept the alternative hypothesis."

8. Assign a level of probability to your findings.
E.g. "Fail to reject the null hypothesis at the .05 level."

“Reject the null hypothesis (.05 level or whichever level you choose) and provisionally accept the alternative hypothesis.”

That is, you find your hypothesis to be statistically significant at the .05 level.

Another way of saying this is that there is a 5% chance that you have made a Type I error.
Which is?

Finding a relationship where none truly exists. Falsely rejecting the null.

A Type II error?
Failing to find a true relationship. Falsely accepting the null.

Again, we can think about hypothesis testing and significance tests in terms of a smoke detector.

Think about how they can go off if you make your toast too dark.

This is a Type I error – an alarm without a fire.

To avoid this, you take out the batteries or put the detector in a closet.

Increases chances of a Type II error – a fire without an alarm.

Decreasing the chances of having a fire without an alarm increases the chances of a false alarm..

	NO FIRE	FIRE
NO ALARM	No error	Type II
ALARM	Type I	No error

TRUE STATE

	H_0	H_A
Accept H_0	No error	Type II
Reject H_0	Type I	No error

For a significance test:

$Pr(\text{Rejecting } H_0 \mid H_0) = Pr(\text{Type I error} \mid H_0) = \alpha$

1- α measures our confidence that any alarm bells we hear are genuine.
High confidence means rarely setting off false alarms.

Example

What should be the relationship between voting in the House and greenhouse emissions?
Hypothesis?
Null hypothesis?

Choose t test

Probability level .05

Open the State Environment data set.

Graph the two variables.

Run the regression.

What is N?
Why?

What are each of the following?

$\widehat{\beta}_1, \widehat{\beta}_0, SE, t.$

95% confidence interval.

What are the outlying cases?

One case can move the regression line.

Try without the case.

Lets construct confidence intervals for our estimates.

$$\widehat{\beta}_1 \pm t_{.025}(SE_{\widehat{\beta}_1})$$

$$\widehat{\beta}_0 \pm t_{.025}(SE_{\widehat{\beta}_0})$$

Use the t-distribution and n-2 degrees of freedom.

If we don't want to think of units of measurement of our independent variable, we can standardize our coefficients.

Beta Coefficient or Beta weight.

$$\hat{\beta}_1^* = \hat{\beta}_1 (\frac{SD_X}{SD_Y})$$

We can do this after our estimation or we can do it before by converting all of our data to z-scores.

For a bivariate regression, Beta Coefficient is the same as a correlation.

Homework
Studenmund exercises 1,2,3,6,11.