USP 634
Spring 2017
Assignment 2: Confidence Intervals, Hypothesis testing (one- and two-sample tests)

1. State law requires every county in the state to average a minimum expenditure of \$5,000 per student across all public schools in the county. A random survey of 200 schools in Multnomah County found an average expenditure of \$4,782, with a standard deviation of \$600. Can you inform the state's Secretary of Education at a 95 percent confidence level that Multnomah County is meeting the standard? Show how you arrive at the conclusion. (15 pts)

This question can be translated, "Does the 95 percent confidence interval around our sample mean include \$5000?" If it does not, then you can say with 95 percent confidence that Multnomah County does not meet the standard. The equation for the confidence interval is $\bar{X} \pm Z_{\alpha/2}(\frac{s}{\sqrt{n}})$. For 95% confidence level $Z_{\alpha/2}$ is 1.96. The 95% confidence interval around the mean is:

$$\$4{,}782 \pm 1.96\left(\frac{600}{\sqrt{200}}\right) = \$4{,}782 \pm 83 = [\$4{,}699, \$4{,}865]$$

Since the standard (\$5,000) is not contained in the 95% confidence interval, we can inform the Secretary of Education that Multnomah County is not meeting the standard.

Alternatively, you can conduct hypothesis testing of whether the mean expenditure per student is greater than or equal to \$5,000:
H_0: $\mu \geq 5{,}000$
H_1: $\mu < 5{,}000$

Conduct a one tailed *t*-test (or a z test as we have more than 100 observations) at 95% significance level ($\alpha = 0.05$)

Test statistics: tstar = $\frac{\bar{X}-5000}{\sigma/\sqrt{n}} = \frac{4782-5000}{600/\sqrt{200}} = -5.12$
p = pt(tstar, n-1, lower.tail=T) = 3.59e-07 < 0.05, we can reject the Null hypothesis H_0 that Multnomah is meeting the standard.

2. You have been asked to manage a regional study on attitudes toward regional growth management. You've been given a budget of \$75,000 to conduct a regional mail survey of households. Your Board wants to know the proportion of residents in favor of an urban growth boundary within an accuracy of plus or minus two percent and at a confidence level of 90 percent. (Recall that when you don't know the true population proportion, to be conservative you should assume a maximum standard error). What is the minimum number of households you need to randomly sample? Can you conduct the survey within budget given an estimated cost of \$25 for administering each survey? (15 pts)

For this problem we need to use the formula for the confidence interval for proportions (which is different from the confidence interval for means because the standard deviation is calculated differently). Recall that we make a conservative assumption that the underlying population parameter is 0.5, which in this case might be close to the true value. This is a conservative assumption because it produces the largest possible estimate of the confidence interval. We solve for the sample size "N" that will yield a confidence interval of plus or minus 0.02. Here's how to do this:

$$P_s \pm Z_{\alpha/2}\sqrt{\frac{P_u(1-P_u)}{n}} = P_s \pm 0.02$$

$$Z_{\alpha/2}\sqrt{\frac{P_u(1-P_u)}{n}} = 0.02$$

$$1.645\sqrt{\frac{0.5(1-0.5)}{n}} = 0.02$$

$$\frac{0.5(1-0.5)}{N} = (0.02/1.645)^2$$

$$N = \frac{0.5(1-0.5)}{(0.02/1.645)^2} = 1691$$

Since there is $75,000 available, up to 3,000 surveys can be paid for which comfortably exceeds the target number of 1,691 (unless of course you accounted for non-response rates, which we also gave credit for). You may also get 1681 or 1701 (or 1702) due to precision.

3. A random sample of adults living in 60 traditional, mixed-use, pedestrian-friendly neighborhoods and adults living in 55 postwar, auto-oriented neighborhoods, all with comparable household income levels, revealed the following:

- Traditional neighborhoods averaged 15.8 daily vehicle miles traveled (VMT) per adult household member, with a standard deviation of 5.3 VMT.
- Auto-oriented neighborhoods averaged 18.3 VMT per adult household member and a standard deviation of 7.5 VMT.

Create 90% confidence intervals for VMT for each type of neighborhoods. Do the intervals overlap with each other? Using the five-step hypothesis testing process, test at α = .10 level the hypothesis of New Urbanists that people living in traditional neighborhoods have lower automobile usage. (20 pts)

Confidence Intervals

$$CI = \bar{X} \pm Z_{\alpha/2}(\sigma/\sqrt{n})$$

$$Z_{.10/2} = 1.645$$

VMT confidence intervals for traditional neighborhood residents:
CI = 15.8 +/- 1.645 * (5.3/sqrt(60)) = 15.8 +/- 1.135 = [14.674–16.926]
VMT confidence intervals for traditional neighborhood residents:
CI = 18.3 +/- 1.645 * (7.5/sqrt(55)) = 18.3 +/- 1.679 = [16.636–19.964]
These two confidence intervals overlap with each other.

Since we are working with relatively small samples, you can also use *t* statistic to construct confidence intervals:

$$CI = \bar{X} \pm t_{\alpha/2,df}(\sigma/\sqrt{n})$$

$$t_{.10/2,59} = 1.671 \text{ and } t_{.10/2,54} = 1.674$$

VMT confidence intervals for traditional neighborhood residents:

CI = 15.8 +/- 1.671 * (5.3/sqrt(60)) = 15.8 +/- 1.153 = [14.657, 16.943]
VMT confidence intervals for auto-oriented neighborhood residents:
CI = 18.3 +/- 1.674 * (7.5/sqrt(55)) = 18.3 +/- 1.708 = [16.607, 19.993]
These two confidence intervals are slightly wider and have larger overlap with each other than when using z statistic, as the CIs based on t-statistc are more conservative.

Hypothesis Testing
Step 1: Set hypothesis
H_o: mu.t = mu.auto;
H_A: mu.t < mu.auto

Step 2: Calculate point estimate
n.t = 60
n.auto = 55
Xbar.t = 15.8;
Xbar.auto = 18.3
sd.t = 5.3
sd.auto = 7.5

Step 3: Check conditions
a. The sample size of 60 traditional neighborhoods is less than 10% of the total traditional neighborhoods, so the samples can be assumed to be independent.
b. The sample size is more than 30, so the sampling distribution is nearly normal.

Step 4: Conduct hypothesis test under assumption of Ho is true via CLT method.
```
> tstar = (Xbar.t – Xbar.auto)/sqrt( sd.t^2/n.t + sd.auto^2 / n.auto)
> df=min(n.t-1, n.auto-1)
> p=pt(-abs(tstar), df)
[1] 0.021
```

Step 5: Make a decision
If the traditional neighborhoods have equal auto usage, there is a 2.1% chance of obtain random samples of 60 traditional neighborhoods yield a sample mean higher than 55 auto-oriented neighborhood. So, it is unlikely to observe higher auto usage data for traditional neighborhood when the null hypothesis is true, and hence we should reject Ho. The data provide convincing evidence that the average VMT in traditional neighborhoods is less than auto-oriented neighborhoods, and our data support New Urbanists' claim that people living in traditional neighborhoods have lower automobile usage.

Note that here that even though the CIs do overlap, the hypothesis testing indicates a statistically significant difference between group means. In the case of single sample hypothesis testing of mean, the conclusion of whether the CI contains μ_0 should always be consistent with the hypothesis testing of $\bar{X} = \mu_0$. However, in the case of two group means, it is more complicated. If the CIs for the two groups do not overlap, then the group means are statistically different in hypothesis testing at the same significance level. However, if they overlap, it is still possible for the hypothesis testing to be significant. See https://www.cscu.cornell.edu/news/statnews/stnews73.pdf and https://www.cscu.cornell.edu/news/statnews/Stnews73insert.pdf for explanation why this happens.

4. Housing values were compared between residences with a view of the Bay and otherwise comparable residences (e.g., amenities, neighborhood quality) though without a view. The analysis sought to measure the imputed value of a Bay view on home sales prices. Two factors were also controlled for in the analysis: age of home and distance to downtown San Francisco. Thus, randomly selected homes that just sold were matched on these two factors, yielding the data shown in the excel spreadsheet **view.xls**. Test at $\alpha = .05$ level that homes with a view enjoy significant value premiums, as reflected by differences in housing values. (10 pts)

The sample size here is pretty small so a t-test of differences in means is to be used here rather than a Z test. We are using a paired *t*-test to control for age of home and distance to downtown.

Set the null hypothesis to no home value differences between properties with view and without view:
H0: Homes with a view have the same average value as homes without a view.
Ha: Homes with a view have higher values than homes without.

A paired *t*-test is essentially testing the mean of differences between each pair. The null hypothesis is equivalent that mean difference in population is 0 (μ=0), and the alternative hypothesis is that mean difference in population is greater than 0 (μ>0). Thus we conduct a **one-tailed *t*-test**.

We first calculate the difference in per sqft housing value for each pair, and then the mean and standard deviation of the differences: $\bar{X} = 20.68, s = 19.362,$ and n=25.

The test statistic is:

$$\text{tstar} = \frac{\bar{X}-0}{s/\sqrt{n}} = \frac{20.68}{19.362/\sqrt{25}} = 5.34$$

Since p = pt(tstar, n-1, lower.tail=F) = 8.80e-06 < .05 confidence level, we are well within range of rejecting the null hypothesis and conclude that homes with a view cost more than those without.

In R, you can do a paired t-test with t.test() function, passing “paired=T” argument:

```
t.test(view$Bay.View, view$No.View, alternative='greater',
conf.level=.95, paired=TRUE)

      Paired t-test
data:  view$Bay.View and view$No.View
t = 5.3, df = 24, p-value = 8.8e-06
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
  14 Inf
sample estimates:
mean of the differences
                     21
```

Notice that the mean in differences is the same as difference in mean between two paired samples. The difference between an unpaired independent *t*-test and a paired *t*-test is in the standard deviation for the test statistic and degrees of freedom. The standard deviation for paired *t* test is smaller, with lower degrees of freedom. The statistically significance level (p-value) from these two tests can be very different.