STA 570 Spring 2011

Lecture 19 *Thursday, March 31*

- **Correspondence Between Significance Tests and Confidence Intervals**
- **Small Sample Inference for Means**

Correspondence Between Tests and Confidence Intervals

- Results of confidence intervals and of two-sided significance tests are consistent:
  - Whenever the hypothesized mean μ_0 is not in the confidence interval around $\bar{Y}$, then the *p*-value for testing $H_0 : \mu = \mu_0$ is smaller than 5% (significance at the 5%-level)
  - In general, a (1-alpha)-confidence interval corresponds to a test at significance level alpha
  - This is true for means as well as proportions

Example

- The mean score for all high school seniors taking a college entrance exam equals 500.
- A study is conducted to see whether a different mean applies to those students born in a foreign country.
- For a random sample of 100 of such students, the mean and standard deviation on this exam equal 508 and 100.

1. Set up hypotheses for a significance test.
2. Compute the test statistic.
3. Report the *P*-value, and interpret.
4. Can you conclude that the population mean for students born in a foreign country equals 500? Why or why not?
5. Make a decision about H_0, using alpha=0.05
6. ***Construct a 95% confidence interval for mu.***

One-Sided Significance Tests

- Example: Usually, students have an average score of 85% on the STA 570 midterm exam
- You want to prove that a certain learning method helps improve the score
- 40 students try out the new method
- Null hypothesis: $H_0 : \mu = 85\%$
- Alternative hypothesis: $H_1 : \mu > 85\%$

One-Sided Significance Tests

- Attention! For one-sided and two-sided tests, the calculation of the *p*-value is different!
- For this example, "everything *at least as extreme* as the observed value" is everything ***above*** the observed value

$$(\text{if } H_1 : \mu > \mu_0)$$

One-Sided Significance Tests

- If we want to prove that μ is smaller than a particular number μ_0 , then

$$H_0 : \mu = \mu_0 \text{ vs. } H_1 : \mu < \mu_0$$

- The P-value is obtained taking the probability of all Y-values ***less*** than the observed Y-value

One-Sided Versus Two-Sided Test

- Two-sided tests are more common
- Look for formulations like
  - "test whether the mean has ***changed***"
  - "test whether the mean has ***increased***"
  - "test whether the mean is ***the same***"
  - "test whether the mean has ***decreased***"

Significance Test for a Proportion

Assumptions

- What type of data?
  - *Qualitative*
- Which sampling method has been used?
  - *Random sampling*
- What is the sample size?
  - *n=20 if* p_0 *is between 0.25 and 0.75*
  - *In general (rule of thumb): Choose n such that*

$$n > 5/p_0 \quad \text{and} \quad n > 5/(1-p_0)$$

Significance Test for a Proportion

Hypotheses

- Null hypothesis $H_0 : p = p_0$

  where p_0 is a priori specified
- Alternative hypotheses can be one-sided or two-sided
- Again, two-sided is more common

Significance Test for a Proportion

z_{obs}

$$= \frac{\text{estimator of the parameter} - \text{null hypothesis value of the parameter}}{\text{standard error of the estimator}}$$

$$= \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

P-Value

- Calculation is exactly the same as for the test for a mean
- Find one- or two-sided tail probabilities using online tools

Example

- Let ***p*** denote the proportion of Kentuckians who think that government environmental regulations are too strict
- Test $\boldsymbol{H_0: p=0.5}$ against a two-sided alternative using data from a telephone poll of 834 people in which 26.6% said regulations were too strict

1. Calculate the test statistic
2. Find the *p*-value and interpret
3. Using alpha=0.01, can you determine whether a majority or minority think that environmental regulations are too strict, or is it plausible that $\boldsymbol{p=0.5}$?
4. Construct a 99% confidence interval. Explain the advantage of the confidence interval over the test.

Summary
Large Sample Significance Test for a Mean

	One-Sided Tests		Two-Sided Test
Null Hypothesis	$H_0 : \mu = \mu_0$		
Research Hypothesis	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Test Statistic	$z = \dfrac{\bar{Y} - \mu_0}{s / \sqrt{n}}$		
p-value	$P(Z < z_{obs})$	$P(Z > z_{obs})$	$2 \cdot P(Z > \lvert z_{obs} \rvert)$

Large Sample Significance Test for a Population Proportion

	One-Sided Tests		Two-Sided Test
Null Hypothesis	$H_0 : p = p_0$		
Research Hypothesis	$H_1 : p < p_0$	$H_1 : p > p_0$	$H_1 : p \neq p_0$
Test Statistic	$z_{obs} = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$		
p-value	$P(Z < z_{obs})$	$P(Z > z_{obs})$	$2 \cdot P(Z >\lvert z_{obs} \rvert)$

Sample Size: Test for a Mean

	One-Sided Tests		Two-Sided Test
Null Hypothesis	$H_0 : \mu = \mu_0$		
Research Hypothesis	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Sample Size	$\left(\sigma \dfrac{z_{1-\beta} + z_{1-\alpha}}{\mu_0 - \mu_1} \right)^2$		$\left(\sigma \dfrac{z_{1-\beta} + z_{1-\alpha/2}}{\mu_0 - \mu_1} \right)^2$

z_p = pth quantile of standard normal

Sample Size: Test for a Proportion

	One-Sided Tests		Two-Sided Test
Null Hypothesis	$H_0 : p = p_0$		
Research Hypothesis	$H_1 : p < p_0$	$H_1 : p > p_0$	$H_1 : p \neq p_0$
Sample Size	$\left(\dfrac{z_{1-\beta} s_1 + z_{1-\alpha} s_0}{p_0 - p_1} \right)^2$		$\left(\dfrac{z_{1-\beta} s_1 + z_{1-\alpha/2} s_0}{p_0 - p_1} \right)^2$

z_p = pth quantile of standard normal

Multiple Choice Question I

- A 95% confidence interval for mu is (96,110). Which of the following statements about significance tests for the same data are correct?
  a) In testing the null hypothesis mu=100 (two-sided), P>0.05
  b) In testing the null hypothesis mu=100 (two-sided), P<0.05
  c) In testing the null hypothesis mu=x (two-sided), P>0.05 if x is any of the numbers inside the confidence interval
  d) In testing the null hypothesis mu=x (two-sided), P<0.05 if x is any of the numbers outside the confidence interval

Multiple Choice Question II

- The P-value for testing the null hypothesis mu=100 (two-sided) is P=.001. This indicates
  a) There is strong evidence that mu = 100
  b) There is strong evidence that mu does not equal 100
  c) There is strong evidence that mu > 100
  d) There is strong evidence that mu < 100
  e) If mu were equal to 100, it would be unusual to obtain data such as those observed

Multiple Choice Question II

- The P-value for testing the null hypothesis mu=100 (two-sided) is P=.001. Suppose that in addition you know that the z score of the test statistic was z=3.29. Then

a) There is strong evidence that mu = 100

b) There is strong evidence that mu > 100

c) There is strong evidence that mu < 100

Small Sample Confidence Interval for a Mean

- What if we want to make inference about the population mean, but our sample size is not big enough to meet the minimal sample size requirement n>25 to apply the central limit theorem?
- Confidence intervals are constructed in the same way as before, but now we are using ***t*-values instead of *z*-values**
- For a random sample ***from a normal distribution***, a 95% confidence interval for mu is

$$\bar{Y} \pm t_{0.025} \frac{s}{\sqrt{n}}$$

- where $t_{0.025}$ is a t-score (instead of z-score) from a site like *http://stattrek.com/Tables/T.aspx*
- degrees of freedom are *df=n-1*

Small Sample Hypothesis Test for a Mean

- Assumptions
  - Quantitative variable, random sampling, **population distribution is normal**, ***any* sample size**
- Hypotheses
  - Same as in the large sample test for the mean

$$H_0 : \mu = \mu_0 \text{ vs. } H_1 : \mu \neq \mu_0$$

$$\text{or} \quad H_0 : \mu = \mu_0 \text{ vs. } H_1 : \mu > \mu_0$$

$$\text{or} \quad H_0 : \mu = \mu_0 \text{ vs. } H_1 : \mu < \mu_0$$

Small Sample Hypothesis Test for a Mean

- Test statistic
  - Exactly the same as for the large sample test

$$t_{obs} = \frac{\bar{Y} - \mu_0}{s / \sqrt{n}}$$

- *p*-Value
  - Same as for the large sample test (one-or two-sided), but using an online tool for the *t* distribution
- Conclusion
  - Report p-value and make formal decision

Small Sample Hypothesis Test for a Mean: Example

- A study was conducted of the effects of a special class designed to improve children/s verbal skills
- Each child took a verbal skills test twice, both before and after a three-week period in the class
- Y=2^{nd} exam score – 1^{st} exam score
- If the population mean for Y, E(Y)=mu equals 0, the class has no effect
- Test the null hypothesis of no effect against the alternative hypothesis that the effect is positive
- Sample (*n=4*): 3,7,3,3

Normality Assumption

- An assumption for the *t*-test is that the population distribution is normal
- In practice, it is impossible to be 100% sure if the population distribution is normal
- It is useful to look at histogram or stem-and-leaf plot (or normal probability plot) to check whether the normality assumption is reasonable

Normality Assumption

- Good news: The *t*-test is relatively ***robust*** against violations of the assumption that the population distribution is normal
- Unless the population distribution is highly skewed, the *p*-values und confidence intervals are fairly accurate

- However: The random sampling assumption must never be violated, otherwise the test results are completely invalid

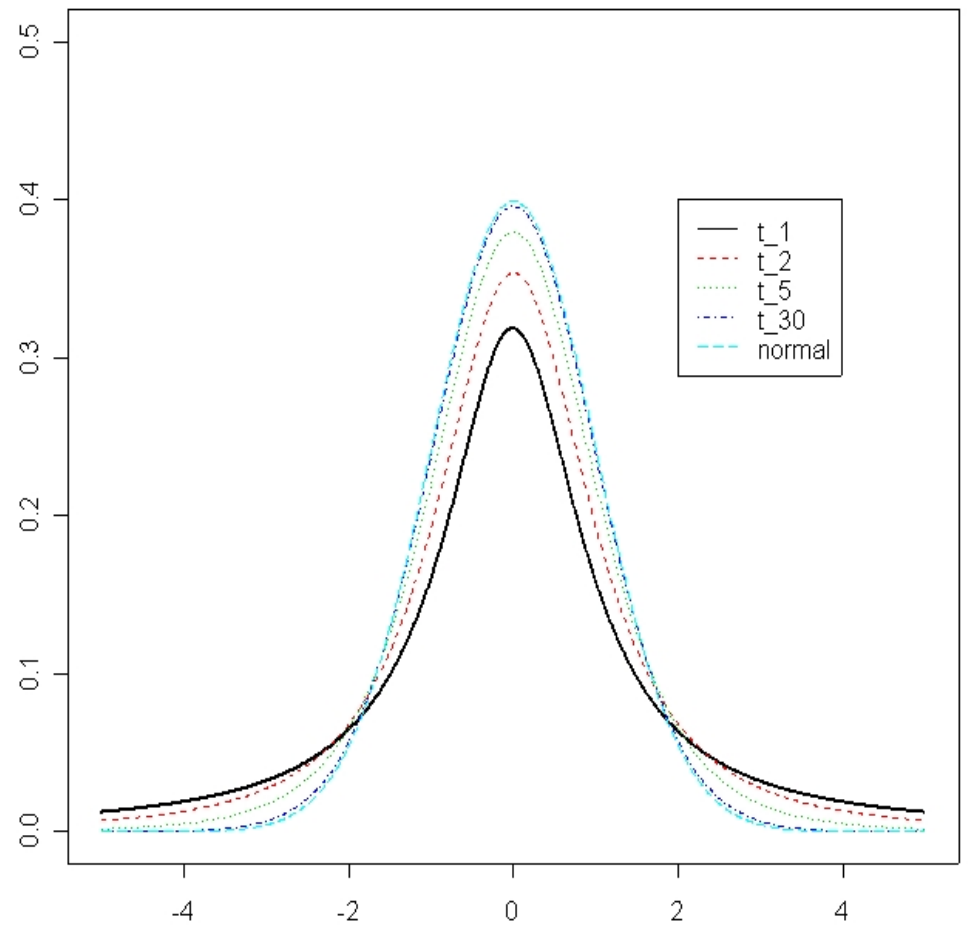
Summary: Small Sample… Significance Test for a Mean

(Assumption: Population distribution is normal)

	One-Sided Tests		Two-Sided Test
Null Hypothesis	$H_0 : \mu = \mu_0$		
Research Hypothesis	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Test Statistic	$t_{obs} = \dfrac{\overline{Y} - \mu_0}{s / \sqrt{n}}$, degrees of freedom $= n - 1$		
p-value	$P(T_{n-1} < t_{obs})$	$P(T_{n-1} > t_{obs})$	$2 \cdot P(T_{n-1} > \lvert t_{obs} \rvert)$

t-Distributions

- The t-distributions are bell-shaped and symmetric around 0
- The smaller the degrees of freedom, the more spread out is the distribution
- *t*-distributions look almost like a normal distribution
- In fact, the limit of the t-distributions is a normal distribution when *n* gets larger



Statistical Methods for One Sample Summary I

- Testing the Mean
  - Large sample size (*30 or more)*:
    Use the large sample test for the mean
    (z-Scores, normal distribution)
  - Small sample size:
    Check whether the data is not very skewed
    Use the t test for the mean
    (t-Scores, t distribution)

Statistical Methods for One Sample Summary II

- Testing the Proportion
  - Large sample size (*np>5, n(1-p)>5)*:

    Use the large sample test for the proportion

    (*z*-Scores, normal distribution)
  - Small sample size:

    Binomial distribution

    (not covered in class)