Social Studies 201
Notes for November 26, 2003

These notes contain a discussion of a test for a mean with small sample size (the t-test).

A. Test for a mean, small sample size – Section 9.3, p. 607.

When a researcher has only a small sample size available, the central limit theorem does not describe the distribution of sample means. Under certain assumptions, the t-distribution can be used to describe the distribution of sample means. Using this distribution for sample means, tests of hypothesis for a population mean μ can be conducted. The procedure for conducting an hypothesis test using the t-distribution is exactly the same as for large sample size, except that the t-distribution and t-value, rather than a normal distribution and Z-value are used. These notes provide a short review of the t-distribution, followed by examples of hypothesis tests for small sample size.

The t-distribution

The t-distribution was described in the notes for November 14 and the table of the t-distribution is contained in Appendix I, p. 911, of the text. The same distribution can be used for hypothesis tests; the conditions required for using the t-distribution and its description are:

> If a normally distributed population has a mean of μ and a standard deviation of σ, and if small random samples of size n (less than 30 cases) are drawn from this population, then the sample means $\bar{X}$ of these samples have a t-distribution with mean μ, standard deviation $s/\sqrt{n}$, and $n-1$ degrees of freedom, where s is the standard deviation obtained from the sample.

This can be stated symbollically. If

$$X \text{ is Nor } (\mu, \sigma)$$

where μ and σ are unknown, and if random samples of size n are drawn from this population,

$$\bar{X} \text{ is } t_d \left(\mu, \frac{s}{\sqrt{n}} \right).$$

where $d = n - 1$ is the degrees of freedom and $\bar{X}$ and s are the mean and standard deviation, respectively, from the sample.

Using this distribution for the sample mean $\bar{X}$, the number of standard deviations this sample mean is from the hypothesized mean of μ is

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}.$$

When conducting an hypothesis test for μ, if this t-value is in the critical region, we reject H_0 but if this t-value is not in the critical region, there is insufficient evidence to reject the null hypothesis.

Example – wages of workers after plant shutdown

In the example of November 14, confidence intervals were constructed for the wages of employees following a plant shutdown. One conclusion drawn from this example was that male wages may not have declined but female wages declined. A more direct way to examine this issue is to conduct an hypothesis test for each group.

The example was drawn from Belinda Leach and Anthony Winson, "Bringing 'Globalization' Down to Earth: Restructuring and Labour in Rural Communities" in the August, 1995 issue of the *Canadian Review of Sociology and Anthropology*. Before shutdown, mean male wages were \$13.76 per hour and mean female wages were \$11.80 per hour. After shutdown, some of the workers found new jobs and the data from small samples of such workers is contained in Table 1. Using data in this table,

1. Test whether there has been any change in the mean wage of all male workers who lost jobs as a result of the plant shutdown.
2. Test whether the wages of all female workers who lost jobs as a result of the plant shutdown have declined.
3. Comment on the findings.

1. Hypothesis test for male wages

For males, the parameter to be estimated is μ, the true mean wage for all male workers who lost jobs because of the plant shutdown. Organizing the

Table 1: Data on Hourly Wages of Workers with Jobs, After Plant Shutdown

Type of Worker	Hourly Wage in Dollars Mean	St. Dev.	Sample Size
Male	12.20	3.27	12
Female	8.11	3.53	12

answer in terms of the five steps involved in hypothesis testing (see notes of November 21), the test is as follows.

1. **Hypotheses**. Since an hypothesis test must begin with an equality for the null hypothesis, the hypothesis that makes most sense here is that there was no change in male wages, that is, that $\mu = 13.76$. The alternative suggested is that there has been a change in males wages. The null and alternative hypotheses are

$$\text{Null hypothesis} \quad H_0 : \mu = 13.76$$

$$\text{Alternative hypothesis} \quad H_1 : \mu \neq 13.76$$

2. **Test statistic**. The claim is about μ, the mean wage of male workers after the plant shutdown. The sample mean, $\bar{X}$, is the test statistic.

3. **Distribution of test statistic**. Since the sample size of $n = 12$ is small, the sample mean $\bar{X}$, has a t-distribution with mean μ and standard deviation $s/\sqrt{n}$ with $n - 1 = 12 - 1 = 11$ degrees of freedom. This assumes that the distribution of wages of all male workers who lost jobs in the shutdown is a normal distribution and the sample is a random sample of such male workers. In symbols,

$$\bar{X} \text{ is } t_d \left(\mu, \frac{s}{\sqrt{n}} \right).$$

where $d = n - 1 = 11$.

4. **Significance level**. The level of significance is not stated, so the default level of $\alpha = 0.05$ is adopted here. Since the alternative hypothesis is that $\mu \neq 13.76$, this is a two-tailed test. The represents the combined area in the two tails of the t-distribution.

5. **Critical region**. The critical region is the extreme area of the distribution, equal in area to the significance level and, in this case, located in the two tails of the distribution. Using the t-table with $n - 1 = 12 - 1 = 11$ degrees of freedom, $\alpha = 0.05$, and a two-tailed test, the t-value is $t = 2.201$. The critical region is all t-values less than -2.201 or greater than $+2.201$.

   The critical region and the associated conclusions that can be made are as follows:

$$\text{Region of rejection of } H_0 : \; t < -2.201 \text{ or } t > +2.201$$

$$\text{Area of nonrejection of } H_0 : \; -2.201 \leq t \leq +2.201$$

6. **Conclusion**. In order to determine whether the sample mean $\bar{X}$ is within the critical region or not, it is necessary to determine the distance $\bar{X}$ is from the hypothesized mean μ. This can be determined by obtaining the t-value associated with the sample mean – that is, how many standard deviations $\bar{X} = 12.20$ is from the hypothesized mean of $\mu = 13.76$.

$$\begin{aligned} t &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ &= \frac{12.20 - 13.76}{3.27/\sqrt{12}} \\ &= \frac{-1.56}{3.27/3.464} \\ &= \frac{-1.56}{0.944} \\ &= -1.653 > -2.201 \text{ and } < +2.201. \end{aligned}$$

   That is, the sample mean is 1.654 standard deviations below the hypothesized mean of $\mu = 13.76$. This t-value is greater than 2.201 standard deviations below the hypothesized mean so is not distant enough from the hypothesized mean to be in the critical region. As a result, the null hypothesis cannot be rejected. This conclusion is made at the 0.05 level of significance, using a two-tailed test.

2. Test for female wages

For female workers, the parameter to be estimated is μ, the true mean wage for all female workers who lost jobs because of the plant shutdown. Organizing the answer in terms of the six steps involved in hypothesis testing, the answer is as follows.

1. **Hypotheses**. Since an hypothesis test must begin with an equality for the null hypothesis, the null hypothesis is no change in female wages, that is, $\mu = 11.80$. Since the question asks whether there has been a decline in female wages, the alternative hypothesis is that $\mu < 11.80$. The null and alternative hypotheses are

$$\text{Null hypothesis} \quad H_0 : \mu = 11.80$$

$$\text{Alternative hypothesis} \quad H_1 : \mu < 11.80$$

2. **Test statistic**. The claim is about μ, the mean wage of female workers after the plant shutdown. The sample mean, $\bar{X}$, is the test statistic.

3. **Distribution of test statistic**. Since the sample size of $n = 12$ is small, the sample mean $\bar{X}$, has a t-distribution with mean μ and standard deviation $s/\sqrt{n}$ with $n-1 = 12-1 = 11$ degrees of freedom. This assumes that the distribution of wages of all female workers who lost jobs in the shutdown is a normal distribution and the sample is a random sample of such female workers. In symbols,

$$\bar{X} \text{ is } t_d\left(\mu, \frac{s}{\sqrt{n}}\right).$$

where $d = n - 1 = 11$.

4. **Significance level**. The level of significance is not stated, so the default level of $\alpha = 0.05$ is adopted. Since the alternative hypothesis is that $\mu < 11.80$, this is a one-tailed test.

5. **Critical region**. In this case of a one-tailed or one-directional alternative hypothesis, the critical region is the extreme area of $\alpha = 0.05$ in the left tail of the t-distribution. Using the t-table with $n-1 = 12-1 = 11$ degrees of freedom, $\alpha = 0.05$, and a one-tailed test, the t-value is $t = -1.796$. The critical region is all t-values less than -1.796.

   The critical region and the associated conclusions that can be made are as follows:

$$\text{Region of rejection of } H_0 : \ t < -1.796$$

$$\text{Area of nonrejection of } H_0 : \ t \geq -1.796$$

6. **Conclusion**. In order to determine whether the sample mean $\bar{X}$ is within the critical region or not, it is necessary to determine the distance $\bar{X}$ is from the hypothesized mean μ. This can be determined by obtaining the t-value associated with the sample mean – that is, how many standard deviations $\bar{X} = 12.20$ is from the hypothesized mean of $\mu = 11.80$.

$$\begin{aligned} t &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ &= \frac{8.11 - 11.80}{3.53/\sqrt{12}} \\ &= \frac{-3.69}{3.53/3.464} \\ &= \frac{-3.69}{1.019} \\ &= -3.621 < -1.796. \end{aligned}$$

   That is, the sample mean is 3.621 standard deviations below the hypothesized mean of $\mu = 11.80$ and is in the left 0.05 of the t-distribution, below $t = -1.796$. The conclusion is to reject the null hypothesis of no change in female wages and, at the 0.05 level of significance, conclude that female wages have declined.

3. Comment

The sample means in Table 1 for males and females are both lower than comparable hourly wages prior to the plant shutdown. As a result, there is some evidence that mean hourly wages for both males and females are lower than previously. But the sample size is only 12, a small sample size, and the standard deviation is between \$3 and \$4. For the males sampled, the mean wage after plant shutdown is approximately \$1.50 lower than the previous hourly wage of \$13.76. The t-test in part 1 of this question demonstrates that the decline of \$1.50 is insufficient to conclude that the mean wage of all male workers has changed. This conclusion is made at the 0.05 level of significance.

In the case of female workers, the decline was greater, from \$11.80 to the \$8.11 reported for the sample of twelve female workers. This is a decline of over \$3.50 and the t-test of part 2 of the question provides strong evidence that the mean hourly wage of all female workers has declined. This conclusion is made at the 0.05 level of significance.

There is the possibility of type I error in the latter conclusion. That is, it is possible that female wages have not declined, but that this sample of female workers is a sample that includes a great number of lower paid female workers. But if this is a random sample of all employed female workers who were previously employed at the plant, the probability is less than 0.05 the conclusion to reject H_0 is in error. In fact, the t-value is -3.621, a lot less than the critical t-value of -1.796, and well into the critical region. As a result, a researcher can be quite confident that the conclusion is correct, always remembering that there is a small chance of type I error.

In the case of male workers, the sample mean in Table 1 is not enough less than the previous mean wage to conclude that the mean wage of all males has changed. There is the possibility that male wages have changed, so there is likely type II error – failing to reject the hypothesis that wages have not changed, when in fact they have. But if wages have changed, they may not have changed all that much, at least that is the evidence presented. The type II error does not seem to be all that serious; if male wages have declined, they do not appear to have declined nearly as much as female wages.

Next day: Test of a proportion, large sample size.