Mathematics 376 – Probability and Statistics II
Lab Project 4 – Hypothesis Testing
April 2, 2012

Background

We have now discussed various tests of hypotheses for means of normal distributions and proportions. In this lab we will work through several realistic examples to illustrate the thinking process statisticians would use to select appropriate tests, and interpret the results.

Lab Questions

A) An office furniture manufacturer has developed a new glue application process for assembling tables. To compare the new process with the old one currently in use, random samples with $n = 30$ are selected from inventories produced with the two processes. Each table is subjected to "destructive testing" in which the force (in pounds) needed to break the glue in the table was measured. Let X be the force in pounds needed to break one of the new tables, and Y be the force in pounds needed to break one of the old tables. The goal is to determine whether the new gluing process has significantly increased the strength of the tables. The data collected was as follows:

$$X : \begin{array}{cccccccccc} 1250 & 1210 & 990 & 1310 & 1320 & 1200 & 1290 & 1360 & 1200 & 1150 \\ 1120 & 1360 & 1310 & 1110 & 1320 & 980 & 950 & 1430 & 1100 & 1080 \\ 960 & 1050 & 1310 & 1240 & 1420 & 1170 & 1470 & 1060 & 1230 & 1300 \end{array}$$

and

$$Y : \begin{array}{cccccccccc} 1180 & 1360 & 1310 & 1190 & 920 & 1060 & 1440 & 1010 & 1000 & 950 \\ 1310 & 980 & 1310 & 1030 & 960 & 800 & 1280 & 1080 & 900 & 1030 \\ 930 & 1050 & 1010 & 1310 & 940 & 860 & 1450 & 1070 & 840 & 1100 \end{array}$$

1) Construct parallel "box-and-whisker plots" for these data sets as in Lab 1 from the fall, and make an informal conjecture about whether or not the new process (the X data) has increased the strength of the tables, compared with the Y data.
2) Describe an appropriate test of the null hypothesis $H_0 : \mu_X = \mu_Y$ versus $H_a : \mu_X > \mu_Y$. Say what your assumptions about the data are, what your test statistic is, what the rejection region will be, and so forth.
3) Carry out your test at the $\alpha = .01$ level of significance. Give a clear and concise statement of the conclusion you draw from your test.
4) What is the attained significance level of your test (the p-value)? What does this say?

B) Let X be the lengths of male spiders of a particular large species and let Y be the lengths of female spiders of the same species (both in mm). A random sample of $n_X = 9$ X values were taken:

$$20.4, 21.7, 21.9, 21.4, 21.1, 23.6, 18.9, 22.6, 21.3$$

Similarly, a random sample of $n_Y = 13$ observations of Y were made:

$$20.5, 20.4, 20.3, 21.1, 21.2, 20.9, 21.0.21.3, 20.9, 20.0, 20.4, 20.8, 20.3$$

Is there a statistically demonstrable difference in the length distributions of two sexes, though? If you think about discussion of small-sample confidence intervals for a difference of means from earlier in the class, you can see that there would be a corresponding hypothesis test. However, there were assumptions that we needed to check to apply the confidence interval formulas based on the t-distribution.

1) What were they?
2) There was one assumption about the overall distributions of the populations the sample values were coming from. Is there any reason to suspect that assumption is not satisfied for this data? Check using *two* features of `R` that we saw earlier this semester. Note: You should be able to interpret the output from `shapiro.test` more precisely now. Can you guess what the null and alternative hypotheses are for the Shapiro-Wilk test? What does the p-value tell you? (*Comment*: The Shapiro-Wilk test is really a test designed to detect *non-normality* of distribution in a data set. Not being able to reject the null hypothesis in this test does not mean the data definitely comes from a normal population. But if we are not able to reject the null hypothesis, the test at least indicates there is no strong evidence to suspect that the population is not normally distributed, and that is about as well as we can do for samples this small.)
3) There was also a second assumption about population variances. This will be tested using a hypothesis test based on an F-statistic as described in §5.9 of the text (pages 530-537). We have not discussed this explicitly in class, but the idea is very similar to what we did before for confidence intervals for the ratio of two variances. So you should be able to look up what you need and follow the book's discussion. Test the null hypothesis $H_0 : \sigma_X^2 = \sigma_Y^2$ against the alternative hypothesis $H_a : \sigma_X^2 \neq \sigma_Y^2$. You may select the significance level α. Interpret your results and give the attained significance level (p-value).
4) Then, if there is *no demonstrable difference* in the variances, test for equality of the means using the t-test from §10.8 in the text. Use the pooled estimator S_p for the common variance, and test statistic

$$t = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}}\ .$$

For the alternative hypothesis $H_a : \mu_X \neq \mu_Y$ take the rejection region

$$RR = \{t \mid |t| \geq t_{\alpha/2}(n_X + n_Y - 2)\}.$$

5) If there *is a demonstrable difference* in the variances, the basic t-test is not all that reliable in some cases. With small sample sizes, many experienced statisticians would use a different method due to Welch. Use the test statistic:

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}},$$

but set up the rejection region for the test using a t-distribution with r degrees of freedom where

$$r = \left[\frac{(S_X^2/n_X + S_Y^2/n_Y)^2}{(S_X^2/n_X)^2/(n_X - 1) + (S_Y^2/n_Y)^2/(n_Y - 1)} \right] \tag{1}$$

($[z]$ = greatest integer less than or equal to z). In this case, to test the null hypothesis $H_0 : \mu_X = \mu_Y$ against the alternative hypothesis $H_a : \mu_X \neq \mu_Y$, you would compute t as above and reject H_0 if $|t| \geq t_{\alpha/2}(r)$ where r is the number of degrees of freedom given in (1).

(*) Note: You will need to select which of the two alternatives 4 and 5 you will use here, how you will set up the rejection region for your test, the test statistic, etc. based on the results of your test from part 1. Explain your choice. Also clearly state the conclusion you draw from the test.

Assignment

Write-ups due in class *Wednesday, April 4.*