Chapter 9

Comparing Two Populations

The ideas of chapter 8 are fine when you are dealing with situations where you have some known standards to which to compare some sample. However, you are often put in the position of comparing two different samples to each other with no set standard. The situation is not difficult to manage. Usually, you just want to determine whether the two samples truly represent two different populations. This can usually done by simple subtraction – literally finding the difference between the two samples. The only tricky part is determining the standard error to use to determine a confidence interval or significance test. This chapter shows you how to do this.

9.1 Vocabulary

The only new vocabulary in this chapter is the term "paired comparison" which is discussed at the end of the chapter.

9.2 Equations, specific rules, etc.

As in previous chapters, we will be forming confidence intervals and looking at test statistics for testing a null hypothesis against an alternate hypothesis. Again, the text separates these cases into questions involving frequencies or proportions and cases involving the mean of some measurement. The program should look familiar by now. The key is always some statistic which looks like a difference from the expected value divided by the standard error.

But in comparing two samples, it is the distribution of the difference in the samples that you need to consider. Assuming that the two samples are independent, their variances should add. This means that the standard deviation of the distribution under consideration (the standard error that you want) is given by the following formulas:

$$\text{For proportions} \qquad SE_{\hat{p_1}-\hat{p_2}} = \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$$

$$\text{which, when } \hat{p_1} = \hat{p_2} \text{ reduces to} \qquad SE_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}$$

$$\text{where} \qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$\text{Similarly, for the means of two samples} \qquad SE_{\bar{x_1}-\bar{x_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

These standard errors can be used in the usual ways to form confidence intervals and to compute significance tests using the t-distribution.

9.3 Specific notes

- p. 159: Think carefully about each of these questions. Can you see that they all involve the comparison of two samples? Can you see that they are all important questions?
- p. 162: This is the important leap of this chapter. It is not a difficult step, but it is important. The key is to recognize situations which fit this model, and there are many such situations.
- pp. 163-167: These pages take the general method of the chapter and apply it to specific questions about frequency or proportion.
- p. 164: The formula for the confidence interval should be familiar by now. Just remember that the critical value ($z_{\frac{\alpha}{2}}$) is usually 2 for a 95% confidence interval.
- pp. 168-171: Again, the general plan is applied to a specific situation. This time the hypotheses in question are about the average value of some characteristic of the two samples, so the difference of the means of the samples is used as the estimator. The t-distribution is used for small sample sizes, and it approaches the z-distribution as the sample size increases.
- p. 171: The pooled variance used on this page is really a special case which we are not going to cover in this program. In addition, the text has a typo near the middle of the page on the right side where it says "with $n_1 - n_2 - 2$ degrees of freedom. The correct formula in this case for degrees of freedom is $n_1 + n_2 - 2$.

- pp. 172-179: The example here makes a good point about knowing when to use different methods. There are times when paired comparisons are appropriate and times when they are not. Usually, common sense will tell you when you can use paired comparisons.

9.4 Exercises

Exercise 9.1: The one-hole test is used to test the manipulative skill of job applicants. This test requires subjects to grasp a pin, move it to a hole, insert it, and return for another pin. The score on the test is the number of pins inserted in a fixed time interval. In one study, male college students were compared with experienced female industrial workers. Here are the data for the first minute of the test. (Based on G. Salvendy, "Selection of industrial operators: the one-hole test," *International Journal of Production Research*, 13 (1973), pp. 303-321.)

Group	n	$\bar{x}$	s
Students	750	35.12	4.31
Workers	412	37.32	3.83

(a) It was expected that the experienced workers would outperform the students, at least during the first minute, before learning occurs. State the hypotheses for a statistical test of this expectation and perform the test. Give a P-value and state your conclusion.

(b) One purpose of the test was to develop performance norms for job applicants. Based on the data above, what is the range that covers the middle 95% of experienced workers? (Be careful! This is not the same as a 95% confidence interval for the mean score of the experienced workers.)

Exercise 9.2: In a study of chromosome abnormalities and criminality, data on 4124 Danish males born in Copenhagen were collected. They were classified as having criminal records or not, using the penal registers maintained in the offices of the local police chiefs. They were also classified as having the normal male XY chromosome pair or one of the abnormalities XYY or XXY. Of the 4096 men with normal chromosomes, 381 had criminal records, while 8 of the 28 men with chromosome abnormalities had criminal records. Some experts believe that chromosome abnormalities are associated with increased criminality. Do these data lend support to this belief? (Data taken from H. A. Witkin *et al.*, "Criminality in XYY and XXY men," *Science*, 193 (1976), pp. 547-555.)