Chapter 14—Hypothesis Tests Applied to Means: Two Independent Samples

14.1 Reanalysis of Exercise 13.1 as if the observations were independent:

Males	Mean = 2.725	s = 1.165	$N_M = 91$
Females	Mean = 2.791	s = 1.080	$N_F = 91$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} = \frac{2.725 - 2.791}{\sqrt{\frac{1.165^2}{91} + \frac{1.080^2}{91}}}$$

$$= \frac{-0.066}{\sqrt{.028}} = \frac{-0.066}{0.166} = -.40$$

[$t_{.05}(180) = \pm 1.98$] Do not reject the null hypothesis.

We can conclude that we have no reason to doubt the hypothesis that males and females are equal with respect to sexual satisfaction.

> There was no need to pool the variances here because the sample sizes were equal. If we did pool them, the pooled variance would have been 1.262.

14.3 The difference between the *t* in Exercises 13.1 and 14.1 is small because the relationship between the two variables was so small.

14.5 Random assignment plays the role of assuring (as much as is possible) that there were no *systematic* differences between the subjects assigned to the two groups. Without random assignment it might be possible that those who signed up for the family therapy condition were more motivated, or had more serious problems, than those in the control group.

14.7 You can not use random assignment to homophobic categories for a study like the study of homophobia because the group assignment is the property of the participants themselves. The lack of random assignment here will not invalidate the findings.

14.9 In Exercise 14.8 it could well have been that there was much less variability in the schizophrenic group than in the normal group because the number of TATs showing positive parent-child relationship could have had a floor effect at 0.0. This did not happen, but it is important to check for it anyway.

14.11 Experimenter bias effect:

Expect Good	Mean = 18.778	s = 3.930	N = 9
Expect Poor	Mean = 17.625	s = 4.173	N = 8

$$s_p^2 = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2} = \frac{8(15.44) + 7(17.41)}{9 + 8 - 2} = 16.359$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_p^2}{N_1} + \frac{s_p^2}{N_2}}} = \frac{18.778 - 17.625}{\sqrt{\frac{16.359}{9} + \frac{16.359}{8}}}$$

$$= \frac{1.153}{\sqrt{3.863}} = \frac{1.153}{1.965} = 0.587$$

[$t_{.05}(15) = \pm 2.131$] Do not reject the null hypothesis.

We cannot conclude that our data show the experimenter bias effect.

14.13 Effect size for Ex14.11

$$d = \frac{\overline{X}_1 - \overline{X}_2}{s_p} = \frac{1.153}{\sqrt{16.359}} = \frac{1.153}{4.045} = 0.285$$

> Note that the answers to Exercises 14.11 and 14.12 are in line with the hypothesis test, in that when we rejected the null hypothesis the confidence limits did not include 0, and when we did not reject the null, they did include 0.

14.15 Comparing GPA for those with low and high ADDSC scores:

$$\overline{X}_L = 2.59 \quad s_L^2 = 0.658 \quad N_L = 75$$

$$\overline{X}_H = 1.68 \quad s_H^2 = 0.560 \quad N_H = 13$$

$$s_p^2 = \frac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1+N_2-2} = \frac{74(0.658)+12(0.560)}{75+13-2} = 0.644$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_p^2}{N_1}+\frac{s_p^2}{N_2}}} = \frac{2.59-1.68}{\sqrt{\frac{0.644}{75}+\frac{0.644}{13}}}$$

$$= \frac{0.91}{\sqrt{.058}} = \frac{0.91}{0.241} = 3.77$$

[$t_{.05}(86) = \pm 1.98$] Reject H_0 and conclude that people with high ADDSC scores in elementary school have lower grade point averages in ninth grade than people with lower scores.

> Here I pooled the variances even though the *N*s were substantially different because the variance estimates were so similar.

14.17 The answer to 14.15 tells you that ADDSC scores have significant predictability of grade point average several years later. Moreover the answer to Exercise 14.16 tells you that this difference is substantial.

> This is a nice example of a situation in which it is easy to see a test of means as a test of predictability.

14.19 Anger with a reason is just fine.

$$\overline{X}_{NoAttrib} = 3.40 \quad s_{NoAttrib}^2 = 2.0736 \quad N_{NoAttrib} = 41$$

$$\overline{X}_{Attrib} = 5.02 \quad s_{Attrib}^2 = 2.7556 \quad N_{Attrib} = 41$$

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(41-1)(2.0736^2)+(41-1)(2.7556^2)}{41+41-2} = 5.9466$$

$$t = \frac{\overline{X}_A - \overline{X}_S}{\sqrt{\frac{s_p^2}{N_1}+\frac{s_p^2}{N_2}}} = \frac{3.40-5.02}{\sqrt{\frac{5.9466}{41}+\frac{5.9466}{41}}} = \frac{-1.62}{0.538} = 3.01$$

The critical value is approximately 2.00, so we will reject the null hypothesis and conclude that when given a reason for a woman's anger, she is given more status than when no reason was given for the anger. [For R, prob = (2*(pt(q = 3.01, ncp = 0, df = 80) = .0035.]

c) We certainly would appear to have a double standard.

14.21 If the two variances are equal, they will be equal to the pooled variance.

> If students have a problem seeing this, they can take any two equal variances and unequal *N*s and try it for themselves. The answer becomes obvious when you do.

14.23 *R* on Ex14.8

```
data <-
read.table("http://www.uvm.edu/~dhowell/fundamentals9/DataFiles/Ex14-8.dat",
header = TRUE)
attach(data)
Group = factor(Group)
t.test(Number ~ Group, alternative = c("two.sided"), var.equal =
TRUE)
```

```
Two Sample t-test
data:  Number by Group
t = 2.662, df = 38, p-value = 0.01132
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.3472894 2.5527106
sample estimates:
mean in group 1 mean in group 2
          3.55           2.10
```