Project 8 Solutions

Statistics 401: Fall 2006

1. Regarding the study of middle aged Finnish men (see the abstract for *Coffee Drinking Is Dose-Dependently Related to the Risk of Acute Coronary Events in Middle-Aged Men* from the September 2004 Journal of Nutritional Epidemiology:

   (a) Since a can of Miller High Life admits that 12 fluid ounces is equal to 355 ml of the champagne of beers, then 950ml of coffee is equal to $950\left(\frac{12}{355}\right) = 32.11$ fluid ounces of coffee, which is equal to a little over 4 cups of coffee each day.

   I drink about 1 cup of tea each day, so I am not a heavy coffee drinker.

   (b) A 99% confidence interval for the mean difference between non-coffee drinker and heavy drinker activity levels using an un-pooled procedure is given by:

   $$2.84 - 15.21 \pm t_{.995,\ df=189}\sqrt{\frac{4.58^2}{77} + \frac{7.87^2}{351}} \approx (-14.1, -10.6)$$

   Note that $t_{.995,\ df=189} = 2.602092$. Also, $df = \frac{(V_1+V_2)^2}{\frac{V_1^2}{76}+\frac{V_2^2}{350}} \approx 189$ where $V_1 = \frac{4.58^2}{77}$ and $V_2 = \frac{7.87^2}{351}$.

   (c) The sample sizes of $n_1 = 77$ and $n_2 = 351$ are both larger than 30, and so the Central Limit Theorem applies: both $\overline{X}_1$ and $\overline{X}_2$ are normal, which implies that $\overline{X}_1 - \overline{X}_2$ is normal.

   (d) I am 99% confident that heavy coffee drinkers smoke between 10.6 and and 14.1 more pack years than non-coffee drinkers on average.

   (e) This was an observational study, not an experiment. A "coffee treatment" was not randomly assigned to the individuals in the study. Thus, cause-and-effect statements can not be made on the basis of this study alone.

   (f) The assertion in the abstract that "heavy coffee consumption increases the short-term risk of acute myocardial infarction or coronary death," solely based on this study, seems dubious since this is an observational study. This is highlighted by the fact that the abstract points out that the "evidence remains equivocal" due to conflicting results from many previous observational studies.

2. Problem 11.4 on page 471:

   (a) A pooled test is appropriate since the sample standard deviations are so close $s_1 = s_2 = .6$, and so we will assume that $\sigma_1 = \sigma_2$.

   (b) i. **Hypotheses:** $H_0 : \mu_{\text{ginkgo}} = \mu_{\text{placebo}}$ versus $H_a : \mu_{\text{ginkgo}} > \mu_{\text{placebo}}$

   ii. **Assumptions:** We have samples of size $n_1 = 104$ and $n_2 = 115$ and both are larger than 30, so we can assume that $\overline{X}_1 - \overline{X}_2$ is normally distributed.

   iii. **Test statistic:** $t = \frac{5.6-5.5}{.6\sqrt{\frac{1}{104}+\frac{1}{115}}} \approx 1.2317$ since $s_p = \sqrt{\frac{103s_1^2+114s_2^2}{104+115-2}} = .6$.

iv. **p-value:** Using a $t(217)$ (since $df = 104 + 115 - 2 = 217$), the p-value $= P(T > 1.2317) \approx 0.1098$.

v. **Decision:** Since the p-value $> .05$, we fail to reject H_0.

vi. **Conclusion:** The evidence fails to suggest that subjects taking ginkgo perform higher on average on the Wechsler Memory Scale than others taking a placebo.

3. Problem 11.24 on page 475-6 of your textbook.

   (a) i. **Hypotheses:** $H_0 : \mu_{\text{adv}} = \mu_{\text{inter}}$ versus $H_a : \mu_{\text{adv}} > \mu_{\text{inter}}$

   ii. **Assumptions:** Since we have samples of size $n_1 = 6$ and $n_2 = 8$ and both are smaller than 15, then we must assume that the force distributions for both advanced and intermediate players are normal.

   iii. **Test statistic:** $t = \frac{40.3-21.4}{\sqrt{\frac{11.3^2}{6}+\frac{8.3^2}{8}}} \approx 3.4568$

   iv. **p-value:** Using a $t(8)$ since $df = \frac{(V_1+V_2)^2}{\frac{V_1^2}{6}+\frac{V_2^2}{8}} \approx 8.838$ where $V_1 = \frac{11.3^2}{6}$ and $V_2 = \frac{8.3^2}{8}$, the p-value $= P(T > 3.4568) \approx 0.0037$.

   v. **Decision:** Since the p-value $< .05$, we reject H_0 in favor of H_a.

   vi. **Conclusion:** The evidence suggests that the mean force after impact is greater for advanced tennis players than it is for intermediate tennis players.

   (b) See the Appendix for R-code and R-output.

4. Do problem 11.32 on pages 485-6 of your textbook.

   (a) Since the measurements can be paired for each hospital, define the variable d =(Inpatient Ratio) - (Outpatient Ratio). Thus, $\mu_d = \mu_{\text{in}} - \mu_{\text{out}}$. A one sample t-test on these differences is a matched pairs t-test:

   i. **Hypotheses:** $H_0 : \mu_{\text{in}} = \mu_{\text{out}}$ versus $H_a : \mu_{\text{in}} > \mu_{\text{out}}$ which is equivalent to testing $H_0 : \mu_d = 0$ versus $H_a : \mu_d > 0$

   ii. **Assumptions:** Since we have paired samples of size 6, a very small sample size, then we must assume that the differences d between Inpatient ratios and Outpatient ratios are normal.

   iii. **Test statistic:** The mean of the differences is $\bar{d} = 18.8\bar{3}$ and $s_d = 5.6716$. Thus, the test statistic is $t = \frac{18.8-0}{\frac{5.6716}{\sqrt{6}}} \approx 8.1339$.

   iv. **p-value:** p-value $= P(Z > 8.1339) \approx 0$.

   v. **Decision:** Since the p-value is tiny, we reject H_0 in favor of H_a.

   vi. **Conclusion:** The preponderance of evidence suggests that the mean cost-to-charge ratio for Inpatient care is larger than the mean cost-to-charge ratio for Outpatient care.

   (b) See the Appendix for R-code and R-output.

5. Regarding the AP article *Search-and-rescue dogs not sickened by post-9/11 work, scientists say*:

   (a) , (b), (c)

i. **Hypotheses:** $H_0 : \pi_1 = \pi_2$ versus $H_a : \pi_1 > \pi_2$ where π_1 is the mortality rate of search-and-rescue dogs who worked at Ground Zero in the days and weeks after 9/11/2001, and π_2 is the mortality rate of search-and-rescue dogs who did not work at Ground Zero.

ii. **Assumptions:** To check the sample size, first compute the pooled proportion $p = \frac{41}{97+55} \approx 0.2697$. Now, we can see that $97p \approx 26 > 10$ and $97(1-p) \approx 71 > 10$; $55p \approx 15 > 10$ and $55(1-p) \approx 40 > 10$. Thus, p_1 and p_2 each have an approximate normal distribution, and so $p_1 - p_2$ is approximately normal.

iii. **Test statistic:** $t = \frac{\frac{29}{97} - \frac{12}{55}}{\sqrt{\frac{.2697(1-.2697)}{97} + \frac{.2697(1-.2697)}{55}}} \approx 1.0784$

iv. p**-value:** $P(Z > 1.0784) \approx 0.1404$.

v. **Decision:** Since the p-value $> .05$, we fail to reject H_0.

vi. **Conclusion:** The evidence fails to suggest that the proportion of search-and-rescue dogs who worked at Ground Zero in the days and weeks after 9/11/2001 have a higher mortality rate than the search-and-rescue dogs who did not work at Ground Zero.

(d) Yes, this conclusion agrees with the conclusion in the article that "the difference was not considered statistically significant."

(e) See the Appendix for R-code and R-output.

6. Regarding the AP article *Bad News For Male Mountain Bikers* available through the STAT401 web site:

(a) We have that $p_1 = \frac{49.5}{55} = .9$ and $p_2 = \frac{6.5}{25} = .26$. Thus, a 99% CI for the difference of proportions $\pi_1 - \pi_2$ is $.9 - .26 \ \pm \ 2.576\sqrt{\frac{.9(1-.9)}{55} + \frac{.26(1-.26)}{35}} \approx (0.4224, 0.8576)$.

(b) To check that $p_1 - p_2$ is approximately normal, first check that n_1 is large enough: $n_1 p_1 = 55(.9) = 49.5 > 10$. But $n_1(1 - p_1) = 55(.1) = 5.5 < 10$. Thus we can not be assured that p_1 is normal, and so $p_1 - p_2$ is not assured to be approximately normal. Furthermore, $n_2 p_2 = 25(.26) = 6.5 < 10$ shows that p_2 is not assured to be normal either. The assumptions are not satisfied.

(c) Since the sample sizes n_1 and n_2 are not large enough, we can not use the confidence interval for inference. However, our point estimate for the difference between the percentage of bikers with low sperm count and the percentage of non-bikers with low sperm count is $p_1 - p_2 = .9 - .26 = .64$. Although we can not use a CI to claim that this is a statistically significant difference, this large difference is troubling. One ought to measure more male bikers and more non-bikers!

(d) See the Appendix for R-code and R-output.

1 Appendix

```
> # PROBLEM 1b

# Get df
> s1=4.58;s2=7.87;n1=77;n2=351
> v1=s1^2/n1;v2=s2^2/n2
> floor((v1+v2)^2/(v1^2/(n1-1) + v2^2/(n2-1)))
[1] 189

# Get critical value
> qt(.995,df=189)
[1] 2.602092

> 2.84-15.21+c(-1,1)*qt(.995,df=189)*sqrt(4.58^2/77 + 7.87^2/351)
[1] -14.11336 -10.62664




> # PROBLEM 2
> # test statistic
> (5.6-5.5)/sqrt(.6^2/104 + .6^2/115)
[1] 1.231664

> # Getting the p-value
> df = 104 + 115 -2
[1] 217
> 1-pt(1.231664,df=217)    # upper tail probability
[1] 0.1097038




> # PROBLEM 3a
> #test stat
> (40.3-21.4)/sqrt(11.3^2/6 + 8.3^2/8)
[1] 3.456827

> # degrees of freedom
> n1=6;x1=40.3;s1=11.3
> n2=8;x2=21.4;s2=8.3
> v1=s1^2/n1;v2=s2^2/n2
> (v1+v2)^2/(v1^2/(n1-1) + v2^2/(n2-1))
[1] 8.832066

> # p-value
> 1-pt(3.456827,df=8.832066)
[1] 0.00370183
```

```
> # PROBLEM 3b
> adv=c(44.7,26.31,55.75,28.54,46.99,39.46)
> inter=c(15.58,19.16,24.13,10.56,32.88,21.47,14.32,33.09)
> t.test(adv,inter,conf.level=0.95,alternative="greater")

        Welch Two Sample t-test

data:  adv and inter
t = 3.4571, df = 8.838, p-value = 0.003697
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 8.854022      Inf
sample estimates:
mean of x mean of y
 40.29167  21.39875



> # PROBLEM 4
> inpatient=c(68,100,71,74,100,83)
> outpatient=c(54,75,53,56,74,71)
> diffpatient=inpatient-outpatient
> rbind(inpatient,outpatient,diffpatient)
            [,1] [,2] [,3] [,4] [,5] [,6]
inpatient     68  100   71   74  100   83
outpatient    54   75   53   56   74   71
diffpatient   14   25   18   18   26   12

> # Test statistic
> mean(diffpatient)
[1] 18.83333
> sd(diffpatient)
[1] 5.671567
> 18.83333/(5.671567/sqrt(6))
[1] 8.133916

> # p-value
> 1-pnorm(8.133916)
[1] 0

> # PROBLEM 4b
> t.test(inpatient,outpatient,conf.level=0.95,alternative="greater",paired=T)

        Paired t-test
```

```
data:  inpatient and outpatient
t = 8.1339, df = 5, p-value = 0.000228
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 14.16768      Inf
sample estimates:
mean of the differences
               18.83333
```

```
> # PROBLEM 5
> # Checking Assumptions
> p=41/(97+55)
> p
[1] 0.2697368
> 97*p
[1] 26.16447
> 55*p
[1] 14.83553
> 97*(1-p)
[1] 70.83553
> 55*(1-p)
[1] 40.16447

> # Test statistic
> (29/97-12/55)/sqrt(.2697*(1-.2697)/97 + .2697*(1-.2697)/55)
[1] 1.078443

> # p-value
> 1-pnorm(1.078443)
[1] 0.1404181

> # PROBLEM 5b
> prop.test(c(29,12),c(97,55),conf.level=.95,alternative="greater",correct=F)

        2-sample test for equality of proportions without continuity
        correction

data:  c(29, 12) out of c(97, 55)
X-squared = 1.1629, df = 1, p-value = 0.1404
alternative hypothesis: greater
95 percent confidence interval:
 -0.03853113  1.00000000
sample estimates:
   prop 1    prop 2
```

```
0.2989691 0.2181818



> # PROBLEM 6

> # PROBLEM 6a
> (.9 - .26) + c(-1,1)*2.576*sqrt(.9*.1/55 + .26*.74/35)
[1] 0.4224308 0.8575692

> # PROBLEM 6b
> 55*.9;55*.1
[1] 49.5
[1] 5.5
> 25*.26
[1] 6.5

> # PROBLEM 6d
> prop.test(c(.9*55,.26*35),c(55,35),conf.level=.99,alternative="two.sided",correct=F)

        2-sample test for equality of proportions without continuity correction

data:  c(0.9 * 55, 0.26 * 35) out of c(55, 35)
X-squared = 38.5661, df = 1, p-value = 5.293e-10
alternative hypothesis: two.sided
99 percent confidence interval:
 0.4224452 0.8575548
sample estimates:
prop 1 prop 2
  0.90   0.26
```