Solutions Project 8
21 points

1. (4 pts, Exercise 10.2)

Note that $Y \sim \text{Bin}(n = 20, p)$. Thus, for large n, $\hat{p} = \frac{Y}{n}$ is approximately normally distributed.

a. If the experimenter concludes that less than 80% of insomniacs respond to the drug when actually the drug induces sleep in 80% or more of insomniacs, a type I error has occurred.

b. $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true}) = P(y \le 12 \mid H_0 \text{ true}) = 0.032$ because $y \sim \text{Bin}(n, p{=}0.8)$ when H_0 is true (cf. `pbinom(12,20,.8)`). The normal approximation for α is $P(\hat{p} \le 12/20 \mid p = .8) = P\left(z < \frac{12/20-.8}{\sqrt{\frac{.8(1-.8)}{20}}}\right) = P(z < -2.24) = 0.0125$.

c. If the experimenter fails to find that 80% of insomniacs respond to the drug when actually the drug induces sleep in fewer than 80% of insomniacs, a type II error has occurred.

d. Denote the Type II error rate under H_a: $p = p_a$ as $\beta(p_a)$. Then $\beta(.6) = P(\text{fail to reject } H_0 \mid H_a\text{: } p = .6 \text{ is true}) = P(y > 12 \mid p = .6) = 0.416$ (cf. `1-pbinom(12,20,.6)`). The normal approximation for $\beta(.6)$ is $P(\hat{p} > 12/20 \mid p = .6) = P\left(z > \frac{.6-.6}{\sqrt{\frac{.6(1-.6)}{20}}}\right) = P(z > 0) = 0.5$.

e. $\beta(.4) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(y > 12 \mid p = .4) = 0.021$ (cf. `1-pbinom(12,20,.4)`). The normal approximation for $\beta(.4)$ is $P(\hat{p} > 12/20 \mid p = .4) = P\left(z > \frac{.6-.4}{\sqrt{\frac{.4(1-.4)}{20}}}\right) = P(z > 1.83) = .0336$.

2. (3 pts, Exercise 10.24)

Let p = proportion of overweight children and adolescents. Then the six steps of the hypothesis test are:

1. H_0: $p = .15$, H_a: $p < .15$
2. Check the assumptions: (a) The problem states that a SRS was taken. (b) For these binomial data, the sample size $n = 100$ is "large" since 100(.15) and 100(.85) are both larger than 10. (c) The sample size $n = 100$ is much smaller than 5% of the population of all children.
3. The large sample test statistic for a proportion is $z = \frac{\frac{13}{100}-.15}{\sqrt{.15(.85)/100}} = -0.56$.
4. The rejection region corresponding to $\alpha = .05$ is RR = $\{z < -1.645\}$.
5. The test statistic $z = -.56$ is not in the rejection region, so fail to reject H_0.
6. The evidence fails to suggest that less than 15% of children are overweight.

3. (6 pts, Exercise 10.23)

a.-b. Let μ_1 and μ_2 denote the true mean distances. Since we want to detect any difference between the two populations, then we will perform a two–tailed test with hypotheses H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$.

c. We have 3 choices regarding which test to perform: a 2-sample z-test; a pooled 2-sample t-test, or a Welch 2-sample t-test.

When performing a pooled *t*-test, the six steps are:
1. H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$,
2. Check the assumptions: (a) SRSs from each of the two populations cannot be confirmed. (b) Independence between the two samples of deer cannot be confirmed. (c) The sample sizes of 40 in each SRS are larger than 30, and so are considered large enough to invoke the CLT. (d) The sample sizes of 40 may NOT be less than 5% of the total number of deer that live in the small geographical locations of interest.
3. The computed large sample test statistic is t = –.954
4. The p-value is 2P(t < -.954) = 2(.1762) = .3430, calculated using a t(78) distribution and a pooled standard deviation s_p = 1055.22.
5. The p-value 0.3430 > α = .10, so fail to reject H_0.
6. The evidence fails to suggest that the mean distances traveled by the deer populations are different.

d. It could be argued that pooling is appropriate since the sample standard deviations of 1140 and 963 are close (our "rule of thumb" is within a factor of 2) to each other.

e. The un-pooled approach yields a test statistic t = –.954 (the same as with the pooled approach), but now the degrees of freedom must be calculated using Satterthwaite's approximation. Instead, I will be conservative and use df = `min(39,39)` = 39, in which case the p-value is 0.3480. The rejection region is RR = {t| |t| > 1.685}. Either way we make a decision, we fail to reject the null hypothesis.
f. The un-pooled approach is almost always preferable to the pooled approach since then we are un-encumbered with the homogeneity of variance assumption.

4. (1 pt, Exercise 10.50, sample size calculation)

$$n = \frac{(z_{1-\alpha} + z_{1-\beta})^2 \sigma^2}{(\mu_a - \mu_0)^2} = \frac{(1.645 + 2.325)^2 .11^2}{.01^2} = 1908.03.$$

Thus, the recommended sample size is 1909.

5. (2 pts) When testing H_0: μ=0.6 vs. H_a: μ<0.6 at a significance level of α = 0.10, the RR = {$\bar{x}$ < k} where $k = 0.6 - 1.28\frac{0.11}{\sqrt{120}} = 0.587$. When H_a: μ<0.6 is true at μ=0.59, then $\beta = P(\bar{x} > k|\mu=0.59) = P(z > \frac{0.587-0.59}{0.11/\sqrt{120}}) = 1-.3897 = 0.6174$.

6. (5 pts) Regarding the AP article *Bad News For Male Mountain Bikers*:

(a) and (b)

i. **Hypotheses:** H_0: $p_1=p_2$ versus H_a: $p_1>p_2$, The parameter p_1 is proportion of male ``extreme" mountain bikers who have abnormalities, and p_2 is the proportion of male non-bikers who have abnormalities.

ii. **Assumptions:** 1. Since these are human subjects, it is unlikely that SRS's were taken. 2. The test also requires independence between the two samples of mountain bikers and non-mountain bikers. 3. The 5% rule is satisfied since the populations of interest are potentially infinite. 4. To check the sample size, first observe that $\hat{p}_1$=.9=49.5/55 and $\hat{p}_2$=.26=9.1/35. Now, checking $n_i \hat{p}_i$ and $n_i(1-\hat{p}_i)$, we see that we do not have enough abnormalities in either the control group or mountain biking groups to be assured of the normal approximation for $\hat{p}_2$. Nonetheless, we will proceed as if $\hat{p}_1$ and $\hat{p}_2$ each have an approximate normal distribution, and so we assume that $\hat{p}_1 - \hat{p}_2$ is approximately normal.

iii. **Test statistic:** Assuming that H_0 is true, we pool the two samples to get an estimate for $p=p_1=p_2$, which is $\hat{p}_c$ = (49.5 + 9.1)/(55 + 35) = 0.65. The test statistic is calculated using $\hat{p}_c$ as an estimate for $p=p_1=p_2$, $t = \dfrac{.9-.26}{\sqrt{\dfrac{.65(1-.65)}{55}+\dfrac{.65(1-.65)}{35}}}$ = 6.21.

iv. ***p*-value**: $P(Z>6.21) = 2.65 \times 10^{-10}$.

v. **Decision:** Since the *p*-value < .05, reject H_0.

vi. **Conclusion:** The evidence suggests that a higher proportion of males who are ``extreme" mountain bikers (ride 3,000 miles yearly - or an average of more than two hours a day, six days a week) have low sperm counts and scrotal abnormalities when compared to other males.

(c) The assertion

> Frequent mountain biking may reduce fertility in men, according to a small Austrian study ... The research suggests frequent jolts and vibration caused by biking over rough terrain may cause abnormalities, including small scars within the scrotum and impaired sperm production.

is tenuous since this is an observational study, and only well designed experiments can provide evidence for a cause-and-effect conclusion.

(d) Here is the R-code:

```
> # PROBLEM 1
> # Test Statistic
```

```
> (.9-.26)/(sqrt(.6511*(1-.6511)/55 + .6511*(1-.6511)/35))
[1] 6.210117

# p-value
> 1-pnorm(6.21)
[1] 2.649230e-10

>prop.test(c(.9*55,.26*35),c(55,35),conf.level=.99,alternative="
greater",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 = 2.646e-10
alternative hypothesis: greater
99 percent confidence interval:
 0.4435165 1.0000000
sample estimates:
prop 1 prop 2
  0.90   0.26

# Making sure that R agrees with hand calculation
> sqrt(38.5661)
[1] 6.210161
```