Stat 21 Homework 2 - Problem 5 Rubric

Problem 5

When a professional statistician has information to share with colleagues, they will submit an article to one of several Statistics journals for publication. This can be a lengthy process; typically the article must be circulated for "peer review" and then edited before being accepted for publication. There the article must wait in line with other articles before appearing in print. In the Winter 1998 issue of Change magazine, Eric Bradlow and Howard Wainer reported on this delay for several journals between 1990 and 1994. For 288 articles published in the journal *The American Statistician*, the mean length of time between initial submission and publication was 21 months with a standard deviation of 12 months. For 209 articles published in the journal *Applied Statistics*, the mean time to publication was 31 months with a standard deviation of 12 months.

(a) Create and interpret a 90% confidence interval for the difference in mean delay between the two journals. The critical value is 1.65.

(b) What are the assumptions needed for part (a)?

(c) State the null and alternative hypotheses for a two-sided hypothesis test that there is a difference in the publication time between the journals. Without doing any calculations in R or by hand, what are the results of this hypothesis test (at an $\alpha = 0.10$ significance level)?

Solution Problem 5: In order to get all 2 points for this problem, students must include the following in their responses.

(a) Students must show their work (if they didn't use R that's ok but they still have to show their steps) and get the correct answer (within 0.1 units). Their interpretation must be logically coherent and they must get the order correct (as to which journal has the longer delay). They must also state the units of the confidence interval (months).

```
estimate <- (21-31)
crit_val <- 1.65
SE <- sqrt( 12^2/288 + 12^2/209 )
LB <- estimate - crit_val*SE
UB <- estimate + crit_val*SE
c(LB,UB)
```

```
## [1] -11.799177  -8.200823
```

Interpretation: We are 90% confident that the true difference in mean delay between these two journals is between -11.8 and -8.2 months. That is, the delay for the American Statistician is, on average, between 8.2 and 11.8 months shorter than the delay for Applied Statistics.

(b) To get full credit the student must identify all of the following as necessary assumptions (it is ok if they provided reasonable caveats to any of the following as long as they identified what the assumption would need to be): - The publication delay of each of these journals is independent of one another. (So just because one of the journals has a delay does not have any bearing on whether or not the other journal will have a delay.) Under normal circumstances this seems reasonable unless there is a worldwide catastrophe that is drastically changing the academic publishing environment of the US... - It isn't unreasonable to think that the delay in months for any given publication follows a symmetric, unimodal histogram. Some publications will always take longer than others and some will be quicker than others. - The data is not a SRS from each population and was collected over a period of four consecutive years, however it does comprise less than 10% of the

entire population of all articles published by each of the journals. Therefore it's not unreasonable to consider this observational study to be a representative sample of the population (at least in the 80s and 90s perhaps).

(c) Let μ_1 be the average delay for articles published in The American Statistician, and let μ_2 be the average delay for articles published in Applied Statistics. The null and alternative could be written as

$$H_0 : \mu_1 - \mu_2 = 0, \quad H_A : \mu_1 - \mu_2 \neq 0$$

or

$$H_0 : \mu_2 - \mu_1 = 0, \quad H_A : \mu_2 - \mu_1 \neq 0$$

it just must be consistent with the order of their estimate in part (a). In either case, we would reject the null in a two-sided, $\alpha = 0.1$ hypothesis test because the confidence interval in part (a) does not contain zero.