Homework 3

Psychology 310

Instructions. Answer the following questions. Show your R code, your input, and your output. Feel free to email me for hints if you get stumped.

1. Supppose you observe a set of $n = 122$ independent observations, and calculate the sample mean to be $\overline{X} = 105$, and the sample standard deviation to be $s = 14.5$. You wish to test the statistical null hypothesis that

   $$H_0 : \quad \mu = 100$$

   against the alternative that $\mu \neq 100$.

   (a) Is this a one-sided or two-sided test?

   (b) What is the value of the t-statistic, and what is its p-value?

   (c) Is the null hypothesis rejected at the 0.01 significance level?

   (d) What are the endpoints of a 0.95 confidence interval for μ?

2. You are running an opinion poll regarding a zoning initiative in your home town. You sample 188 registered voters at random, and 119 say that they are in favor of the initiative. Suppose the population proportion of voters supporting the initiative is symbolized p, and the statistical null hypothesis is

   $$H_0 : \quad p \leq .50$$

   (a) Is this a two-tailed or one-tailed hypothesis?

   (b) Can you reject the statistical null hypothesis at the 0.05 level?

   (c) What is the p-value?

   (d) State, in your own words, the meaning of the statistical null hypothesis in practical terms (i.e., regarding the voting initiative).

   (e) Construct a 0.95 confidence interval for p.

3. Use the same data as in question 1 above. Consider the statistical null hypothesis that

   $$H_0 : \quad \sigma = 12$$

   (a) Is this a one-sided or two-sided hypothesis?

(b) Test this hypothesis at the 0.05 significance level. What is the p-value? Is the null hypothesis rejected?

(c) What are the endpoints of a 0.95 confidence interval for σ?

4. Construct some artificial data using the following code:

```
> set.seed(20131006)
> x <- rnorm(50)
> y <- rnorm(50) + x
> w <- rnorm(50) + x
```

(a) Fit a linear regression model for predicting y from just x.

(b) Is the slope of the regression line significantly different from zero?

(c) Construct a "rough" confidence interval for the slope of the regression line using the method described in class

$$\hat{\beta}_1 \pm 2 * \hat{\sigma}_{\hat{\beta}_1}$$

that is, the estimate plus or minus 2 standard errors.

(d) Make sure the MASS library is installed and loaded in your version of R. Then construct the normal-theory confidence intervals with the command `confint(fit)`, where `fit` is the name of your fit object.

(e) Add w as a predictor to your model. With w and x in the model, what is the squared multiple correlation?

(f) Are the regression coefficients for x and w both significantly different from zero?

(g) Construct confidence intervals for the regression coefficients for x and w using any method you choose.

(h) Using the verb|anova| command, construct a hierarchical regression test of the null hypothesis that adding w produces no significant improvement in the linear regression for predicting y. What is the p-value?

(i) Do you reject the null hypothesis of no improvement at the 0.01 level?

(j) Now, construct a linear regression model in which y is predicted solely from w. Is the regression slope significant?

(k) Next, construct a hierarchical linear regression test of the hypothesis that adding x produces no improvement of the regression model over and above a model with just w. Is the null hypothesis rejected?

(l) So you had a significant w, you added an x. Now, examine the two t-tests for x and w in the full model. Are they both significant? Does the result of these tests contradict the previous results? Discuss.

5. Suppose that exam grades in a large course are approximately bivariate normally distributed, with a correlation $\rho_{X,Y}$ of 0.65 between scores on the two exams. If a person scores 2 standard deviations above the mean on the first exam, what is the probability that they will score *at least one standard deviation* above the mean on the second exam?