Stat 139: Homework 5

Due: Friday, October 10th, before noon

- **Incorporate the relevant R output in the HW write-up**. Choose the output wisely, no need to print dozens of pages. No more than **two pages** of output should be displayed for each problem and the relevant parts should be **highlighted**.
- **A hard copy of your solutions is due at noon on Friday, in your TF's drop-box outside the Science Center 300 suite**. The solutions should not be submitted by email unless arrangements are made with a TF prior to the deadline (or unless late, see below).
- **Your code should be attached to the end of your write-up.** Use comments to indicate the code associated with each question.
- See syllabus for details on homework and collaboration policies: **acknowledge your collaborators**; your lowest homework score will be dropped; solutions submitted electronically within 24 hours after deadline will be graded with a penalty; solutions more than 24 hours late will receive no credit.

Related Reading: Ramsey and Schafer, Chapters 4 & 5.
R code: Chapter 4, Chapter 5.
Supplementary: Moore, McCabe Introduction to the practice of statistics: Chapter 6, Section on "Power".

(Not Graded) Complete the following Conceptual Exercises in Chapter 4 (2nd or 3rd edition of R&S): 2, 5, 6, 8, 10, 12, and in Chapter 5: 1, 2, 4. No need to include your answers in the write-up.

1. (15 points) Suppose that Y_i, $i = 1, ..., 100$, are drawn independently from a Normal distribution with mean μ and variance 1. You are using a one-sample z-test to assess the following hypotheses:

$$H_0 : \mu = 0$$
$$H_A : \mu < 0$$

   (a) Assume that H_0 is rejected if the left-sided p-value is less than $\alpha = 0.05$. What is the Type I error? What value of $\bar{Y}$ is the cutoff for rejecting versus not rejecting H_0? Call that cutoff c_α.

   (b) Consider a specific version of the alternative hypothesis: $H_A : \mu = -1$. Assuming that this alternative hypothesis is true, what is the probability that $\bar{Y}$ is less than c_α? Specify the Type II error and the power.

   (c) More generally, assuming that $H_A : \mu = \mu_A$ is true, the probability that $\bar{Y}$ is less than c_α is equal to the probability that a standard normal is less than some value Q. Express Q as a function of c_α and μ_A.

   (d) How does the power of the test change as μ_A decreases? Briefly explain the intuition. How does the power of the test change as α increases? Briefly explain why there is a tradeoff between Type I and Type II error.

2. (10 points) In lecture, we used simulations to explore the validity and power of various tests in the context of sampling from two independent populations. Here, you will illustrate the same principles in the context of pairs sampled from a population of pairs. We'll focus on left-sided p-values, but the principles you'll illustrate are also true for right- or two-sided p-values.

Your comments on what each simulation shows are at least as important as the simulations themselves.

(a) First, you'll see that the paired t-test is valid when its assumptions are met. Draw 100 independent standard normals: these are the paired differences for a sample of 100 pairs. Perform a paired t-test, and record the left-sided p-value. Note that a paired t-test is just a one sample t-test for the paired differences. Repeat this process 10,000 times and report the proportion of left-sided p-values that are less than 0.05. Comment in one sentence.

(b) Repeat the steps in 2a, but report the proportion of left-sided p-values from the *sign test* that are less than 0.05. You can use the normal approximation for the reference distribution of the sign test, so that the R-function `pnorm` produces the p-value. Comment in one sentence.

(c) Rerun your code from 2a and 2b, but draw the sets of 100 paired differences from a normal distribution with mean -0.2 and variance 1. Report the proportions of left-sided p-values that are less than 0.05 for the paired t-test and for the sign test. In one sentence, comment on the power of the sign test versus the power of the paired t-test when the t-test assumptions hold.

(d) Last, rerun 2a and 2b, but draw the paired differences from an exponential distribution, shifted to have mean 0 (`rexp(N)-1` will generate N such draws in R). Report the proportion of left-sided p-values that are less than 0.05.

   i. Does the paired t-test appear to be robust to the normality assumption?

   ii. The sign test does not assume normality. Why would the rejection rate be different from 0.05 when the differences follow an exponential distribution? (It may help to look at the shape of this distribution: use `hist(rexp(1000)-1)`).

3. (10 points) Chapter 4, Exercises 20, 21, and 22 (either edition). All of these refer to Exercise 18 in Chapter 3, which has been scanned and posted online. Exercise 20 should be performed by hand; there is no need to use continuity correction. Both #21 and #22 should be performed in R. For #22, explain the mechanics of obtaining the confidence interval, based on what you learned from Section 4.2.4. (Note that the same procedure may be used to obtain confidence intervals in a permutation test.)

4. (15 points) Chapter 5, Exercise 24: **Vegetarians and Zinc: An Observational Study**, in the **2nd ed.** (note that it has been removed from the 3rd. edition, the scanned problem and the data set are posted online). There is no need to transform the data in this problem. In addition to the question posed in the exercise, answer the following question:

   Is there evidence that means are different for all three groups? Include the ANOVA table in your write-up (see Display 5.10 for an example). You might want to wait until the lecture on Tuesday, 10/7, before tackling this question.

   You may use R to find sample statistics as well as quantiles of the corresponding t- and F-distributions. However, the rest of the calculations in this problem should be performed by hand.

   Clearly state your hypotheses, defining all the necessary notation. Specify and check the assumptions. Word your conclusions in terms of the original question of interest and comment on the scope of inference.