Adam Gincel
MA331
I pledge my honor that I have abided by the Stevens Honor System.

Homework 3

Pages 369-370: 6.17, 6.27, 6.28;
Pages 391-393: 6.58, 6.59, 6.71, 6.73;
Page 401: 6.99;
Page 412: 6.120;
Pages 442: 7.22, 7.23.

Answers are in **bold**.

6.17: ***The state of stress in the United States.*** Since 2007, the American Psychological Association has supported an annual nationwide survey to examine stress across the US. A total of 340 Millennials (18-33 years old) were asked to indicate their average stress level (10 point scale) during the month. The mean score was 5.4. Assume that the population standard deviation is 2.3.

a) Give the margin of error and find the 95% confidence interval for this sample.
M = z * (2.3) / sqrt(340)
Z* = 1.96
M = 0.244
Interval = [5.4 - 0.244, 5.4 + 0.244] = [5.166, 5.644]

b) Repeat these calculations for a 99% confidence interval. How do these results compare with those in part a?
M = z* (2.3) / sqrt(340)
Z* = 2.576
M = 0.321
Interval = [5.4 - 0.321, 5.4 + 0.321] = [5.079, 5.721]

6.27: ***Average hours per week listening to the radio.*** The Student Monitor surveys 1200 undergrads from 4-year colleges and universities throughout the US semiannually to understand trends among college students. Recently the Student Monitor reported that the average amount of time listening to the radio per week was 11.5 hours. Of the 1200 students surveyed 83% said they listened to the radio, so this collection of listening times has around 204 (17% * 1200) zeros. Assume the standard deviation is 8.3 hours.

a) Give a 95% confidence interval for the mean time spent per week listening to the radio.
M = Z*(8.3) / sqrt(1200)
z* = 1.96
M = (1.96)(8.3)/sqrt(1200) = 0.470

Interval = [11.5 - 0.47, 11.5 + 0.47] = [11.03, 11.97]

b) Is it true that 95% of the 1200 students reported weekly times that lie in the interval you found in part a? Explain your answer.
204 of the people sampled answered they do not listen to the radio at all; this means at least 17% of the sample population lies outside of our 95% confidence interval. This is because our confidence interval is based on the average time, not individual times.

c) It appears that the population distribution has many zeroes and is skewed to the right. Explain why the confidence interval based on the Normal distribution should nevertheless be a good approximation.
This approximation should still be fine because we have a large sample size. The normal distribution is best with larger sample sizes.

6.28: Refer to the previous exercise.

a) Give the mean and standard deviation in minutes.
   Mean: 11.5 hrs * 60 = 690 minutes
   Standard Deviation: 8.3 * 60 = 498 minutes
b) Calculate the 95% confidence interval in minutes from your answer to part a.
   M = (1.96 * 498)/sqrt(1200)
   M = 28.177 minutes
   M = [690 - 28.177, 690 + 28.177] = [661.823, 718.177]
c) Explain how you could have directly calculated this interval from the 95% interval that you calculated in the previous exercise.
   Because our answers for the 95% confidence interval in problem 6.27 were in hours, we could have simply multiplied both values in our original confidence interval by 60 to get their value in minutes.

6.58: ***Computing the p-value.*** A test of the null hypothesis H_0: u = u_0 gives the test statistic z = 1.77

a) What is the P-value if the alternative is H_a: u > u_0?
   P-value = 0.962
b) What is the P-value if the alternative hs H_a: u < u_0?
   P-Value = 0.0384
c) What is the P-value if the alternative is H_a: u != u_0?
   P-Value = 0.0767

6.59: ***More on computing the p-value.*** A test of the null hypothesis H_0: u = u_0 gives the test statistic z = -1.69

a) What is the P-value if the alternative is H_a: u > u_0?
   P-value = 0.9545
b) What is the P-value if the alternative hs H_a: u < u_0?
   P-Value = 0.0455
c) What is the P-value if the alternative is H_a: u != u_0?

P-Value = 0.091028

6.71: ***Attitudes Toward School.*** The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures the motivation, attitude toward school, and study habits of students. Scores range from 0 to 200. The mean score for US college students is about 115, and the standard deviation is about 30. A teacher who suspects that older students have better attitudes toward school gives the SSHA to 25 students who are at least 30 years old. Their mean score is 127.8.

a) Assuming that standard deviation is 30 for the population of older students, carry out a test of

H_0: u = 115

H_a: u > 115

Report the P-value of your test and state your conclusion clearly.

Our Z value is 2.13
Our P-value is ~0.017
As such, the null hypothesis holds.

b) Your test in part A required two important assumptions in addition to the assumption that the value of the standard deviation is known. What are they? Which of these assumptions is most important to the validity of your conclusions in part a?

We assume this is a normal distribution, though that isn't very important as there are no huge outliers. We also assume this is a simple random sample, and isn't weighted in favor or against a certain conclusion.

6.73: ***Are the measurements similar?*** Refer to Exercise 6.30 (Page 371). In addition to the computer's calculations of miles per gallon, the driver also recorded the piles per gallon by dividing the miles driven by the number of gallons at each fill-up. The following data are the differences between the computer's and the driver's calculations for that random sample of 20 records. The driver wants to determine if these calculations are different. Assume that the standard deviation of a difference is 3.0.

5.0	6.5	-0.6	1.7	3.7	4.5	8.0	2.2	4.9	3.0
4.4	0.1	3.0	1.1	1.1	5.0	2.1	3.7	-0.6	-4.2

----

a) State the appropriate H_0 and H_a to test this suspicion.

H_0: Mean = 0 mpg
H_a: Mean != 0 mpg

b) Carry out the test. Give the P-value, and then interpret the result in plain language.

Z = 4.07

This gives us a P-value of 0.000047, which is incredibly small. This supports the null hypothesis, which suggests the measurements are basically the same.

6.99: ***Practical Significance and Sample Size.*** Every user of statistics should understand the distinction between statistical significance and practical importance. A sufficiently large sample will declare very small effects stastically significant. Consider the study of elite female Canadian athletes in 6.72. Female athletes were consuming an average of 2403.7 kcal/d with a standard deviation of 880 kcal/d. Suppose that a nutritionist is brought in to implement a new health program for these athletes. This program should increase mean caloric intake but not change the standard deviation. Given the standard deviation and how calorie deficient these athletes are, a change in the mean of 50 kcal/d to 2453.7 is of little importance. However, with a large enough sample, this change can be significant. To see this, *calculate the P-value* for the test of each of the following situations:

A) A sample of 100 athletes, their average caloric intake is 2453.7
Z = 50/(880/sqrt(100)) = 0.569
P-value = 0.2843
B) A sample of 500 athletes, their average caloric intake is 2453.7
Z = 50/(880/sqrt(500)) = 1.27
P-value= 0.1020
C) A sample of 2500 athletes, their average caloric intake is 2453.7
Z = 50/(880/sqrt(2500)) = 2.841
P-value = 0.0023

6.120: ***Choose the appropriate distribution.*** You must decide which of two discrete distributions a random variable X has. We will call the distributions P_0 and P_1 Here are the probabilities they assign the the values x of X:

X	0	1	2	3	4	5	6
p_0	0.1	0.1	0.2	0.1	0.1	0.1	0.3
p_1	0.2	0.2	0.2	0.1	0.1	0.1	0.1

We have a single observation on X and wish to test
H_0: p_0 is correct
H_a: p_1 is correct

a) Find the probability of a Type I error; that is, the probability that you reject H_0 When p_0 is the correct distribution.
Probability of a Type I error = 5% (Assuming a 95% confidence interval)
b) Find the probability of a type II error.
P_{type2} = 1 - power = 1 - 0.73 = 0.27 = 27%

7.22: ***A one-sample t test.*** The one-sample t statistic for testing
H_0: u = 8
H_a: u > 8
From a sample of n = 16 observations has the value t = 2.15.

a) What are the degrees of freedom for this statistic?
Degrees of Freedom = 16 - 1 = 15

b) Give the two critical values t* from Table D that bracket t.
   1.753 < t < 2.131
c) Between what two values does the P-value of the test fall.
   .025021 < P < 0.050004
d) Is the value t = 2.15 significant at the 5% level? Is it significant at the 1% level?
   The upper bound of t* (2.131) yields a P value that is significant at a 5% level. However, it is not significant at 1% level.
e) If you have software available, find the exact P-value.
   P-value = 0.0483

7.23: ***Another one-sample t test.*** The one-sample t statistic for testing

H_0: u = 40

H_a: u != 40

From a sample of n = 27 observations has the value t = 2.01.

a) What are the degrees of freedom for t?'
   Degree of Freedom = 27 - 1 = 26
b) Locate the two critical values t* from Table D that bracket t.
   1.706 < t < 2.056
c) Between what two values does the P-value of the test fall?
   0.049951 < P < 0.099928
d) Is the value t 2.01 statistically significant at the 5% level? At the 1% level?
   T = 2.01 is not significant at either 5% or 1%
e) If you have software available, find the exact P-value.
   P-value = 0.0549