Math 222: Exam 1
Fall – 2019
10/03/2019
80 Minutes

Name: *Caleb M^{c}Whorter — Solutions*

Write your name on the appropriate line on the exam cover sheet. This exam contains 10 pages (including this cover page) and 5 questions. Check that you have every page of the exam. Answer the questions in the spaces provided on the question sheets. Be sure to answer every part of each question and show all your work. If you run out of room for an answer, continue on the back of the page — being sure to indicate the problem number.

Question	Points	Score
1	20	
2	20	
3	20	
4	20	
5	20	
Total:	100	

1. uPhone advertises that their new '*XIV*' phone has download speeds of over 20 Mbps (megabytes per second). A consumer research agency suspects that the speeds are less than what is claimed in the advertisement. A random sample of download speeds using different times and locations was taken. The statistics are found below:

	n	$\overline{x}$	s
DLSpeed	30	18.4	3.7

(a) (6 points) Construct a 96% confidence interval for the mean download speeds for this phone. [You do *not need* to interpret the result.]

We use a 1-sample t*-procedure, see part (e). We have* $n = 30$, $\overline{x} = 18.4$, $s = 3.7$, *and dof 29. For a 96% confidence interval with degrees of freedom 29, we have* $t^* = 2.150$. *Then we have*

$$\overline{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 18.4 \pm 2.150 \frac{3.7}{\sqrt{30}} = 18.4 \pm 1.45$$

Therefore, the 96% confidence interval is $(16.95, 19.85)$. *Therefore, we are 96% that the average download speed of this phone is between 18.4 Mbps and 19.85 Mbps.*

(b) (2 points) Write an appropriate null and alternative hypothesis for the given problem.

$$\begin{cases} H_0 : \mu = 20 \textit{ Mbps} \\ H_a : \mu < 20 \textit{ Mbps} \end{cases}$$

(c) (4 points) If the consumer research agency uses a significance level of 5%, what is the critical value for the test you wrote in (b)?

Given the alternative hypothesis in (b), we need a critical value corresponding to 5% on the left. Because the t*-distribution is symmetric, we can find the* t*-statistic leaving 5% on the right, which using degrees of freedom 29, is 1.699. Therefore, the critical value is* -1.699.

(d) (6 points) Given the sample data collected by the research agency, use the critical value in (c) to determine whether or not you reject the H_0 you gave in (b). State your conclusions in the context of the problem.

We have

$$t = \frac{18.4 - 20}{3.7/\sqrt{30}} = \frac{-1.6}{0.6755} = -2.369 \overset{\text{dof } 29}{\rightsquigarrow} 0.01 < p < 0.02$$

The test statistic is less than the critical value, i.e. $-2.369 < -1.699$ *(equivalently,* $p < \alpha$*). Therefore, we reject the null hypothesis. There is sufficient evidence to suggest that the average download speed of this phone is less than 20 Mbps.*

(e) (2 points) Given n, what must we assume about the sample for the previous parts to be accurate?

Because $n = 30$ *and* $15 < n < 40$*, we need the sample to have no skewness or outliers. Making this assumption, the previous computations are valid.*

2. A particular survey course tends to be taken by either Freshmen or Seniors. The department is trying to determine if the course should only be available to those in their last two years. The department will make the decision based off course letter grades, only allowing students in their last two years to take the course if there is a statistically significant difference between Freshman and Senior grades. They take a random sample of students and find the following data:

	Total Number of Students	Number Receiving an 'A'
Freshman	40	10
Seniors	60	24

(a) (10 points) Using a significance level of 10%, conduct an appropriate hypothesis test to determine if the course should be limited to Juniors and Seniors. You must provide the null and alternative hypotheses, the test statistic, the p-value, and a conclusion stated in the context of the problem. Be sure to justify that the test is appropriate.

Notice a 2-sample p-procedure is appropriate because each sample has at least 5 successes/failures. We have the following data:

$$X_F = 10 \qquad n_F = 40 \qquad \hat{p}_F = \frac{10}{40} = 0.25$$
$$X_S = 24 \qquad n_S = 60 \qquad \hat{p}_S = \frac{24}{60} = 0.40$$

We test the hypothesis

$$\begin{cases} H_0 : p_F = p_S \\ H_a : p_F \neq p_S \end{cases}$$

We have $\hat{p} = \dfrac{10+24}{40+60} = \dfrac{34}{100} = 0.34$. *Then*

$$SE_{D_{\hat{p}}} = \sqrt{0.34(1-0.34)\left(\frac{1}{40}+\frac{1}{60}\right)} = \sqrt{0.00935} = 0.0967.$$

Then we have test statistic

$$z = \frac{0.40-0.25}{0.0967} = \frac{0.15}{0.0967} = 1.55 \rightsquigarrow 0.9394$$

Then $p = 2(1-0.9394) = 2(0.0606) = 0.1212 > \alpha$. *Therefore, we fail to reject the null hypothesis. The data is consistent with the percentage of Freshmen and Seniors receiving an 'A' are the same. Based on this data, no restrictions to the course need to be made.*

(b) (10 points) Find a 90% confidence interval for the difference in the *percentages* of Freshman and Seniors receiving an 'A' in the course. Interpret your answer in the context of the problem.

We have the following data:

$$X_F = 10 \qquad n_F = 40 \qquad \hat{p}_F = \frac{10}{40} = 0.25$$
$$X_S = 24 \qquad n_S = 60 \qquad \hat{p}_S = \frac{24}{60} = 0.40$$

Then

$$SE_D = \sqrt{\frac{0.25(1-0.25)}{40} + \frac{0.40(1-0.40)}{60}} = \sqrt{0.00869} = 0.09322$$

For a 90% confidence interval, we have $z^* = 1.645$. *Then we have*

$$(0.40 - 0.25) \pm 1.645(0.09322) = 0.15 \pm 0.1533$$

which gives $(-0.003, 0.303)$. *As percentages, this is* $(-0.3, 30.3)$. *Therefore, we are 90% certain that, on average, Seniors receive an 'A' in the course between 0.3% less to 30.3% more often than Freshmen.*

3. Education researchers are trying to determine if a course based on universal design (UD) principles increases students' performance in a course. They administer two independent Statistics courses, one based on UD principles (consisting of 45 students) and one not (consisting of 32 students). The UD course had a class average of 87.6 with standard deviation 3.3. The non-UD course had a class average of 82.3 with standard deviation 7.1.

   (a) (5 points) Is a pooled 2-sample t-test appropriate? Explain. Then regardless of whether a pooled test is appropriate, calculate s_p.

   A pooled 2-sample t-test is not appropriate because $7.1/3.3 = 2.15 > 2$. If we were to pool, we would have

$$s_p^2 = \frac{(45-1)3.3^2 + (32-1)7.1^2}{45+32-2} = \frac{2041.87}{75} = 27.2249$$

   so that $s_p = \sqrt{27.2249} = 5.22$.

   (b) (5 points) Using an appropriate method, construct a 90% confidence interval for the difference in mean grade between the two courses. [You do not need to interpret the result.]

   We have SE $= \sqrt{\frac{3.3^2}{45} + \frac{7.1^2}{32}} = \sqrt{1.81731} = 1.3481$. We have degrees of freedom dof $= \min\{45-1, 32-1\} = 31$ (we use dof 30), so that for a 90% confidence interval, we have $t^* = 1.697$. Then

$$(87.6 - 82.3) \pm 1.697 \cdot 1.3481 = 5.3 \pm 2.29$$

   which gives interval $(3.01, 7.59)$. Therefore, we are 90% certain that, on average, students in a UD course (as a whole) score between 3.01 and 7.59 more than students in a non-UD course.

(c) (5 points) Use the confidence interval in (b) to test (at the 10% significance level) the null hypothesis $H_0 : \mu_{UD} = \mu_N$ against the alternative $H_a : \mu_{UD} \neq \mu_N$. State your conclusions in the context of the problem.

Because we have a two-sided test, we can use the confidence interval (corresponding to α) to test the hypothesis. Observe that a difference of 0 is not in the 90% confidence interval; therefore, we reject the null hypothesis. There is sufficient evidence to suggest that the average course grade between the UD course and non-UD course are different. [Computing directly, we have test statistic $t = 3.93 \rightsquigarrow p < 0.0.0005$ so that $2p < 0.001$.]

(d) (5 points) What if the researchers believed that students in a UD course should score 5 points higher on average than students that were not in a UD course. Is the data consistent with this hypothesis? Answer this by writing down an appropriate null and alternative hypotheses, calculating an appropriate test statistic and p-value, and state your conclusions at the 5% significance level in the context of the problem.

We test the hypothesis

$$\begin{cases} H_0 : \mu_{UD} = \mu_N + 5 \\ H_a : \mu_{UD} \neq \mu_N + 5 \end{cases}$$

i.e., $H_0 : \mu_{UD} - \mu_N = 0$ against $H_a : \mu_{UD} - \mu_N = 5$. Then the test statistic is

$$t = \frac{(87.6 - 82.3) - 5}{1.348} = \frac{5.3 - 5}{1.348} = \frac{0.3}{1.348} = 0.22 \rightsquigarrow 0.5871$$

Then the p-value is $p = 1 - 0.5871 = 0.4129$. Therefore, we fail to reject the null hypothesis. The data is consistent with UD course students scoring, on average, 5 points higher than non-UD course students.

4. Red tide is a discoloration of seawater caused by blooms of a toxic red dinoflagellates (a type of plankton). When weather and tide factors cause these blooms, shellfish in the area develop dangerous levels of paralysis-inducing toxins. In Massachusetts, the Division of Marine Fisheries (DMF) tracks the levels of toxin in shellfish. If the mean level of toxin exceeds 800 μg of toxin per kilogram of clam meat, clam harvesting is banned. [The standard deviation of these toxins is known to be 51 μg.] The DMF uses a sample of 122 shellfish to test the following hypothesis:

$$\begin{cases} H_0 : \mu = 800 \\ H_a : \mu > 800 \end{cases}$$

The DMF decides to use a test that reject H_0 if $\overline{x} > 812.0\,\mu$g.

(a) (6 points) Find the probability of a Type I error for this test.

Note that $n \geq 30$ so the CLT applies. We know that $P(\text{Type I error}) = P(\text{reject } H_0 \mid H_0 \text{ true})$. But assuming H_0 is true, the probability that we reject is

$$z = \frac{812 - 800}{51/\sqrt{122}} = \frac{12}{4.6173} = 2.60 \rightsquigarrow 0.9953$$

Therefore, $P(\text{Type I error}) = \alpha = 1 - 0.9953 = 0.0047 \approx 0.005$, i.e. 0.5%.

(b) (2 points) What is the significance level for this test?

The significance level is equal to the probability of making a Type I error, which we found in (a). Therefore, we have $\alpha = P(\text{Type I error}) = 0.005$, i.e. 0.5%.

(c) (2 points) What is the critical value for this test?

The critical value is the test statistic at which one begins to reject H_0. But this is exactly the z-value corresponding to the sample mean of 812.0 μg, which we found in part (a). Therefore, the critical value is 2.60.

(d) (6 points) Is the test sufficiently sensitive to detect an increase of 5 μg of toxin per kilogram of clam meat? Answer this question by calculating the power of the test against the alternative $\mu = 805$ and interpreting the answer.

We know that Power $= P($*reject* $H_0 \mid H_0$ *false*$)$. *An increase of 5* μ*g would mean the true population mean would be 805* μ*g. We reject when* $\overline{x} > 812.0\,\mu$*g. Then*

$$z = \frac{812 - 805}{51/\sqrt{122}} = \frac{7}{4.6173} = 1.52 \rightsquigarrow 0.9357.$$

Therefore, Power $= 1 - 0.9357 = 0.0643$. *This means there is only a 6.43% chance of detecting an increase of 5* μ*g of toxin per kilogram of clam meat. Therefore, the test is not sufficiently sensitive to detect this change.*

(e) (2 points) How can the DMF increase the power for this test to detect a change of 5 μg per kilogram of clam meat? Assume the DMF cannot change anything about the distribution of toxin levels, and that they do not want to change their significance level for the test.

The ways of increasing the power of a statistical test are to decrease σ *(not an option here by the first comment), increase the significance level (not an option by the second comment), choose* μ_{alt} *farther from* μ_0 *(not possible because we are considering an increase of exactly 5* μ*g), or increase the sample size. Therefore, the only option is to increase the sample size.*

(f) (2 points) Calculate the probability of a Type II error for this test if $\mu = 805$.

We know $P($*Type II error*$) = 1 -$ *Power. But we calculated this in (d). Therefore,* $P($*Type II error*$) = 1 - 0.0643 = 0.9357$. *Alternatively,* $P($*Type II error*$) = P($*fail to reject* $H_0 \mid H_0$ *false*$)$. *Given we fail to reject when* $\overline{x} < 812.0\,\mu$*g and if* $\mu = 805$*, we know from (d) that this happens with probability* 0.9357.

5. (20 points) Mark each of the following statements as True (T) or False (F).

(a) ___T___ t-procedures are robust.

(b) ___F___ The margin of error accounts for all possible sources of error in a confidence interval.

(c) ___F___ All other things equal, increasing the significance level will increase the probability of making a Type II error.

(d) ___T___ The larger (in absolute value) the test statistic is in a hypothesis test the more unusual the sample.

(e) ___T___ The power of a statistical test measures the ability to detect if the null hypothesis is false.

(f) ___F___ Failing to reject $H_0 : \mu = 52.1$ means that $\mu = 52.1$.

(g) ___F___ All other things equal, increasing the confidence level decreases the size of a confidence interval.

(h) ___F___ Two-sample t-tests are robust against the samples not being a SRS.

(i) ___T___ A Type I error is rejecting a true null hypothesis.

(j) ___T___ Assuming no sample bias and other sources of error, it is possible to compute the sample size required to estimate a population proportion p within a given margin of error with no prior information available about p.