Stat 185 - Problems off of past exams

Our final exam will be comprehensive and based off of past exams. Treat the following list of problems list of problems like a review sheet!

- Exam I

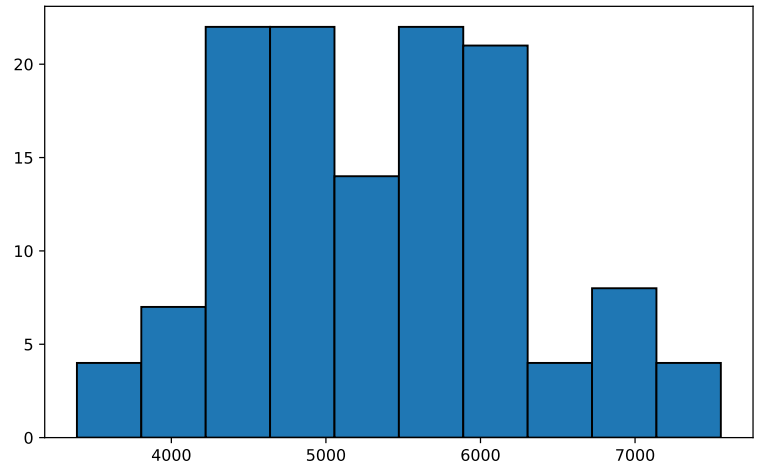

Figure 1: A histogram of total yards gained by each team for the 2016 season

5. On average, the 127 college football teams in Division I gained 5322 yards throughout the 2016 season with a standard deviation of 887 yards.
   (a) Use a normal model to estimate the number of teams that rushed more than 6000 yards.
   (b) A histogram for this data is shown in figure 1. Based on the histogram, do you see any issues with using a normal model for this estimate?

- Exam 2

3. Referring to our random sample of 50 of 3281 total students from table 1, use a normal distribution to estimate the probability that more than 28 of those 50 students are *not* from Western NC.
4. Suppose now that we are interested in the proportion of in-state UNCA students who are from outside Western NC *this* year. Since data on the entire student

Region	Western NC	Piedmont	Eastern NC	Total
Enrollment	1508	1541	232	3281

Table 1: In state UNCA Enrollment by NC Region

body is not yet available for this year, we draw a simple random sample of 64 in-state UNCA students and find that 34 of them are from outside Western NC. Use this data to write down a 95% confidence interval for the proportion of in-state UNCA students who from outside Western NC.

5. UNCA claims that 90% of it's students are from North Carolina but we suspect that it might be less than that. Suppose we draw a random sample of 50 UNCA students and find that 41 of them are in state. Let's use this to explore the claim that 90% of UNCA students are in-state vs the possibility that fewer than 90% of UNCA students are in-state.
   (a) Construct appropriate null and alternative hypotheses to explore our question.
   (b) Compute the associated p-value.
   (c) Based on a 95% level of confidence, what is the conclusion of our hypothesis test?

- Exam 3

1. In planning for an early afternoon, outside wedding ceremony on November 18, I looked up the temperature at 12:54PM for the last nineteen November 18^{th}s. I found an average temperature of 52.3 degrees with a standard deviation of 11.8 degrees. Use this data to write down a 95% confidence interval for the temperature.
2. An organic farmer sells bags of Brussels sprouts labeled as one pound each. Let's use a t-test to investigate the farmer's claim that bags each contain one pound. I purchased one bag each of 4 weeks during the Summer farmer's market, weighed them at home and recorded the following weights:

1.05	1.15	1.2	1.1

   (a) Write down the hypothesis statement for the problem.
   (b) Write down a formula showing that the mean is $\mu = 1.125$
   (c) Write down a formula showing that the standard deviation is $\sigma = 0.06454972$.
   (d) Assuming the mean $\mu_0 = 1$ is correct, compute the t-score for the observed mean.
   (e) Use a t-table to determine the outcome of the hypothesis test.
3. A recent study surveyed people's attitudes on the relative importance of economic issues. In particular, the study examined the question - what is more important:
   – Reducing the federal deficit or
   – Cutting taxes for corporations

   The results are summarized in table 1. Let p_D denote the proportion of democrats

	Democrats	Republicans
Deficit over taxes	97	66
Taxes over deficit	43	48

Table 2: Views on deficit reduction over tax reduction

who think that reducing the deficit is the more important of the two issues and let p_R denote the proportion of republicans who think that reducing the deficit is the more important of the two issues. We suspect that $p_D < p_R$ so let's use a hypothesis test to examine this question.

(a) Compute the observed proportions $\hat{p}_R$ and $\hat{p}_D$, as well as the difference

$$\hat{p} = \hat{p}_D - \hat{p}_R.$$

(b) Compute the standard error and the test statistic.

(c) Use a normal table to compute the p-value.

(d) State the conclusion of the test.

(e) Why was it OK to use a normal table?

4. A statistics professor teaching two sections of the same introductory statistics course is concerned that exam scores in one section (let's call it section A) are noticeably lower than exam scores in another section (section B). To test this hypothesis, he computed the mean and standard deviation for the combined scores on the first two exams for both classes and found the following:

   Section A: Mean=118.4, std dev = 40, number of students = 30

   Section B: Mean=133.5, std dev = 32.7, number of students = 30

   Viewing these data as samples from a random process, find

   (a) The difference between the two means.

   (b) The associated standard error and test statistic.

   (c) Compute the p-value using a normal distribution.

   (d) From the computations, can we conclude with a 95% level of confidence that section A scores lower than class B?

5. A professor using an open source introductory statistics book predicts that 60% of the students will purchase a hard copy of the book, 25% will print it out from the web, and 15% will read it online. At the end of the semester he asks his students to complete a survey where they indicate what format of the book they used. Of the 126 students, 71 said they bought a hard copy of the book, 30 said they printed it out from the web, and 25 said they read it online.

   (a) State the hypotheses for testing if the professor's predictions were inaccurate.

   (b) How many students did the professor expect to buy the book, print the book, and read the book exclusively online?

   (c) Calculate the chi-squared statistic, the degrees of freedom associated with it, and the t^*-value for a 95% level of confidence

   (d) Based on the t^*-value calculated in part (d), what is the conclusion of the hypothesis test? Interpret your conclusion in this context.