

Math 110
Final Exam Sample
April 30, 2010

Do all your work in the blue book. Be sure to show your calculations and provide explanations as indicated. You may use your formula sheet (half of one-side of one $8\frac{1}{2} \times 11$ piece of paper), your calculator, and the z , t , and χ^2 tables that accompany the test.

1. (30 pts.) Explain the following statistical concepts:
 - (a) Residual plot.
 - (b) Bootstrap method.
 - (c) Probability histogram.
 - (d) Regression fallacy.

2. (30 pts.) Short Answer:
 - (a) When drawing without replacement from a box, why is it necessary to use a correction factor when calculating the standard error (SE)?
 - (b) (Hypothetical.) Two Holy Cross students, Gannon and Oglethorpe, have to design a study of student cell-phone use for class. Gannon insists on using a simple random sample but Oglethorpe counters that a simple random sample is too simple and a more sophisticated technique is required. Why would Oglethorpe say a simple random sample is too simple when surveying Holy Cross students?
 - (c) (Hypothetical.) A town near Worcester is the location of a Superfund* toxic waste site. A doctor in the town observes an unusually high number of cases of autism in his town and calculates that for a town that size, there is a 1 in 15 chance of having so many cases. Is this reason to conclude that there is a connection between the Superfund site and the number of cases of autism? Why or why not?

*From Wikipedia: Superfund is the common name for the United States environmental policy officially known as the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA, 42 U.S.C. 9601-9675), enacted by the United States Congress on December 11, 1980 in response to the Love Canal disaster and the environmental contamination at the Valley of the Drums. The Superfund law was created to protect people, families, communities and others from heavily contaminated toxic waste sites that have been abandoned.
 - (d) (Hypothetical.) A major university health service agrees to participate in a trial of new medication for acne. It contacts all of the students who are currently on prescription medication for acne and offers them the opportunity to try the new medication under the supervision of a health service doctor. Of the 145 students contacted, 53 agree to take the new medication. The health service uses these 53 students for the treatment group and the remaining 92 students as controls. Is this a well-designed study? If so, why? If not, why not?

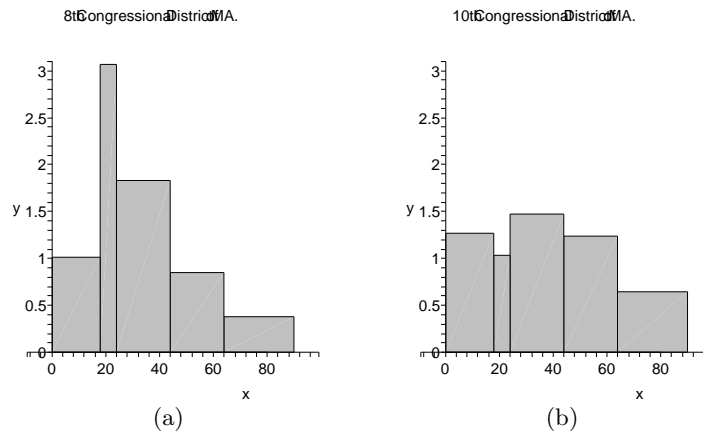


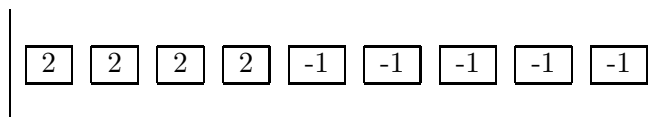
Figure 1: (a) Boston and Cambridge. (b) Cape Cod, Nantucket, and Martha's Vineyard.

3. (20 pts.) The above histograms are for the population by age in the 8th and 10th Congressional Districts of Massachusetts for the 106th Congress. There were 620,372 residents of the 8th District and 663,508 residents of the 10th District. (Data from the US Census 2000.) The 8th District contains most of Boston and all of Cambridge. The 10th District contains Cape Cod, Nantucket, and the Martha's Vineyard.
 - (a) Based on the histograms, which district has a higher median age. Explain your answer.
 - (b) Based on the histograms, how do the population distributions in the 8th District and 10th Districts compare?
 - (c) Is this difference to be expected? Why or why not?

4. (20 pts.) (Hypothetical) The dean at a nearby engineering school wants to analyze the grades in a two course introduction to engineering sequence that must be taken by all first year students. The course grade is based on a raw total of 500 points in each of the courses. In the fall semester the average point total is 385 with an SD of 30 and in the spring semester, the average point total is 350 with an SD of 25. The correlation coefficient for the data is $r = .6$.
 - (a) Find the formula for the regression line for this data.
 - (b) If a student is chosen at random what would you estimate their score to be in the second course?
 - (c) If a student scores 400 in the first semester, what would you estimate his or her score to be in the second semester?
 - (d) If a student scores in the 70th percentile first semester, what would you estimate for his or her percentile in the second semester?

5. (20 pts.) Answer the following questions about rolling a fair die six times:
 - (a) What are the chances of rolling a 1 or a 2 on the first three rolls and *no* 1 or 2 on the last three rolls?
 - (b) What are the chances of rolling 1 or a 2 exactly 3 times out of six rolls?
 - (c) What are the chances of not rolling a 1 or 2 at all?

6. (20 pts.) Sixty draws with replacement are made from the following box:



- (a) What are chances that the sum of the draws will be between 15 and 18 inclusive?
- (b) What are chances that number of 2s drawn will be 30 or more? (*Hint*: What is the box model for this question?)
7. (20 pts.) (Hypothetical) After a number of complaints, the librarians in Dinand decide to conduct a poll to determine whether study rooms in the library should be kept open 24 hours a day during study period and exam period. They conducted a simple random sample of 78 library users. Of the 78, 43 said study rooms should remain open 24 hours a day and 35 said it should not. Before changing policy the librarians put a plus-minus number on the result to make sure the favorable response is not due to the chance in sampling.
- (a) What box model should the librarians use?
- (b) Compute a 95% confidence interval for this poll.
- (c) Is it likely that a majority of all students favor keeping the library open based on the results of this poll? Explain.
8. (20 pts.) A large university mathematics department has a policy that students in a calculus class should work on average at least 7 hours per week on calculus outside of class. If the average in an instructor's class falls below this number, the instructor is told that he or she must assign more homework. Since monitoring every student's hours is time consuming and expensive, the department takes sample numbers from each class. If the average falls below 7 hours per week, the department runs a test of significance and if the difference in the sample average is statistically significant, the homework policy is applied. In one instructor's section, the out-of-class hours worked in a six person sample were

5, 3, 4, 9, 3, 7

- (a) What test of significance should be used to determine if this instructor will need to assign more homework? Why?
- (b) What are the null and alternative hypotheses for this test?
- (c) Carry out your test of significance.
- (d) Will the department require the instructor to assign more homework?

9. (20 pts.) The following question and responses appear in the SDA archive:

Question: Do you ever read a horoscope or your personal astrology report?

The valid responses to this question by sex among 18-22 year old men and women are as follows:

	Male	Female	Row Total
Yes	31	64	95
No	27	11	38
Col Total	58	75	133

Statistically, we would like to know whether the difference in percentages of responses for men and women is “real” or due to the chance error in sampling.

- Formulate a null and alternative hypothesis for this question.
- What significance test should you use to determine the answer?
- Carry out your test of significance. What do you get for a test statistic and observed significance level?
- Is the difference real or not?