

Problem 2. (30 pts) We wish to examine the reported rate of cell phone usage while driving among drivers aged 18-22 (88%). We perform our own study, in which we take a random sample of 17 such drivers, finding that 14 would not hesitate to use a cell phone while driving.

a. (6 pts) Define the sample proportion of those who would use a cell phone while driving. What would be our best point estimate for this rate?

b. (6 pts) Describe the distribution of the sample proportion p , assuming the reported rate is correct.

c. (6 pts) Would your answer change if we had done a study of 170 drivers, and 140 had said that they would not hesitate to use a cell phone while driving?

d. (6 pts) Let's assume that we used 170 drivers in our study, and that the stated rate is correct. Find the probability that 140 or fewer of these drivers use cell phones.

- e. (6 pts) Let's assume that we used 170 drivers in our study, and that the stated rate is correct. What's the chance that 160 or more drivers in a sample will use cell phones while driving?

Problem 3. (26 pts) Let's suppose that we're interested in finding out how many acres of U.S. farmland were lost to development between 1982 and 1992, per hour.

Harper's Index estimated that the number was 50 (8/1998). If their estimate was obtained from a sample of 2000 randomly chosen hours, and the standard deviation of the distribution of "acres per hour" is 5 acres,

- a. (6 pts) give a 95% estimate for the mean number of acres per hour.

- b. (4 pts) Explain what we mean by "a 95% estimate for the mean number of acres per hour".

c. (6 pts) give a 97% estimate for the mean.

d. (10 pts) Let's suppose that we want to provide an estimate for this number to within plus-or-minus one single acre. What should the sample size be to assure that we achieve this with 95% confidence?

Problem 4. (22 pts) Suppose that prevailing wisdom conjectures that all minority groups have the same (reduced) income relative to the white population. Suppose further that we suspect that a Mexican-American child will have a better income ratio than an African-American child.

Let's think of the parameter of interest as the average difference in income between the two groups.

According to Harper's Index (9/2006), the average income of an African-American today, expressed as a percentage of the average income of a white American, is 74%. The average income of a Mexican-American, expressed in the same terms, is 71%.

a. (4 pts) State the null hypothesis.

b. (4 pts) State our alternative hypothesis.

c. (6 pts) Suppose that we commissioned Harper's Index to perform a study to allow us to test our hypothesis. Based on the evidence presented by Harper's Index, what would our next step be?

d. (4 pts) Suppose that we conclude that there is a difference, when in fact there is none. How would you describe the error we made?

e. (4 pts) Suppose that we conclude that there is no difference, when in fact there is one. How would you describe the error we made?

Problem 5. (24 pts) Figures from the College Board, a nonprofit organization of colleges and universities, suggested that books and supplies cost the average student at a public four-year school \$817 (during the 2003-2004 year). Assume that the standard deviation was \$119

The school CUCU, which has been focused on using more internet-based, public domain texts, sampled 120 students in 2004 and found that the average cost for their students was \$623.

- a. (12 pts) Do the data provide sufficient evidence to conclude that CUCU is saving their students money on books? Use a .01 level of significance.

- b. (6 pts) If we had loosened up and used a .10 level of significance, would our answer have changed?

- c. (6 pts) Would a cost of \$778 at CUCU have produced a smaller or larger p-value than the one calculated in Part a.?

Problem 6. (16 pts) A manufacturer of flashlights wants to assure that their product is of high quality. They begin by sampling their product, to get a baseline of defectives. They randomly sample 500 lights from all their production lines, and find that 23 are defective.

a. (10 pts) Find a 99% confidence interval for the proportion of all lights produced by the company that are defective.

b. (6 pts) If a much larger sample of lights were selected, and another 99% confidence interval were calculated, how would it compare to the the interval found above?

Problem 7. (16 pts)

- a. (4 pts) How is the test for normality different in the case of the sample mean \bar{x} versus the sample proportion p ?

- b. (4 pts) Why is only one parameter needed to characterize the distribution of the sample proportion p , but two in the case of \bar{x} ?

- c. (4 pts) In the news (New York Times, Tue Mar 31, 2009): an innocent man serves time for nearly 14 years for a killing he did not commit. This is an error of our criminal justice system. Describe the error in hypothesis testing terms.

- d. (4 pts) Describe the trade-off in our judicial system between the errors that can arise.