Intro Statistics — Final Exam — Spring 2005

Show all work clearly for partial credit. In particular, write down any difficult computations before you perform them, so that, if you make a mistake, you have at least displayed your understanding of the process.

1. (18 pts) Suppose the Admissions Office at Colgate claims that its policy for maintaining a geographic mix of students has remained unchanged over the past four years. To test whether this is the case, two (fictitious) intrepid Math 102 students did a random survey of Colgate students. Their findings are below:

	NE	SE	MW	W	Total
Class 0f '05	70	20	35	15	140
Class 0f '06	81	22	17	30	150
Class 0f '07	58	15	23	27	123
Class 0f '08	65	17	20	13	115
Totals	274	74	95	85	528

- a) Create a table of expected frequencies for a χ^2 -test to verify the claim of the Admissions Office based on this sample. (You do not need to calculate the χ^2 -statistic.)
- b) Suppose the χ^2 -statistic for part a) is 20.3. Find the P-value for this test and state the conclusion of the test.
- c) Suppose a new Director of Admissions recruited the Class of '08. Suppose the university's policy on geographic mix is to have 55% of each class from the Northeast and the remaining portion equally distributed across the other regions. Create the expected frequencies for a χ^2 -test measuring the success of the new director in meeting this goal for the Class of '08.
- 2. (15 pts) In a city with 100,000 TV sets, the Nielsen rating company installs their meters on 900 sets.
 - a) At 8:17 pm, 180 of the metered sets are tuned to CBS. What is a 95% confidence interval for the percentage of TV sets in the city that are tuned to CBS at that time?
 - b) On NBC at 8:17 pm is a show that people (mostly children) like to watch in groups. Suppose that of the sets in the city tuned to NBC, the average number of people watching a given set is 3.64, with an SD of 1.2. If 100 of the metered sets are tuned to NBC, what is the probability that the average audience at these 100 sets is 3.40 or more?
- 3. (5 pts) According to an article in the November 28, 1990 edition of the San Francisco Chronicle, 74 percent of the freshman class at Berkeley scored over 500 on the verbal section of the SAT. If the verbal SAT scores for the entire class have an SD of 80 points and follow a normal distribution, what is the average?
- 4. (10 pts) An experiment is performed to see whether calculators help students do word problems. The subjects are a group of 500 thirteen-year-olds in a certain school district. All the subjects work a given math problem. Three hundred students are chosen at random and allowed to use calculators; the others do the problem with pencil and paper. In the calculator group, 180 of the students get the right answer; in the pencil-and-paper group, 116 do. Can this difference be explained by chance variation? (Check this statistically and state your conclusion.)

5. (10 pts) Demand an Experienced Surgeon?

The more experienced a doctor is, the better. As obvious as that sounds, there are still too many people out there who never ask their surgeons for a history of their work. The importance of knowing is illustrated by this study.

Peter Starek, a surgeon at University of North Carolina, reviewed 460 heart valve replacement operations and found that only 4 percent of the patients of the three most senior surgeons died. But one junior surgeon lost almost a third of his patients. That surgeon was said to be technically the best in the group. Starek says something was obviously lacking—perhaps the kind of good judgment that grows out of experience...

The last sentence of the paragraph contains the claim that the junior surgeon was obviously lacking something. Is that claim justified by Dr. Starek's evidence? Discuss briefly.

- 6. (15 pts) As computers come off the assembly line they are tested before being shipped to retailers. There is a 11% chance of a computer having a bad memory chip, a 3% chance of having a bad keyboard and a 6% chance of a bad central processor. The different faults listed are independent of each other.
 - a) For the first ten computers, what is the chance that two of them have a bad memory chip?
 - b) What is the chance of a computer having all three parts fail the tests?
 - c) What is the chance that a computer has at least one of the three parts being faulty?
 - d) In a day of testing 300 computers, what is the chance of finding 25 or fewer faulty memory chips?
- 7. (17 pts) For women age 25-29 in New York, the average income is \$23,600 with an SD of \$11,500. The average years of education is 13 years with an SD of 3.3 years. The correlation coefficient between income and education level for this group is r = 0.4.
 - a) Predict the income of a woman with 14 years of education.
 - b) This prediction is likely to be off by how much; or can this be determined from the information given?
 - c) Predict the education level of a woman with income \$18,100.
 - d) Suppose the residual plot shows that data points on the left and right sides lie above the regression line and middle points lie below it. In a brief paragraph, describe what difficulty this introduces to the measurement of correlation between education and income and how one might go about alleviating that difficulty?
- 8. (10 pts) Recall the article by Megan Beckett in the San Diego Tribune on January 2, 2003 entitled "Monitor after-school programs carefully". Describe how this article calls for methods from the hypothesis testing section of this course and from the first unit (on experimental design). Why are both concepts needed?