

Make sure that this examination has 10 numbered pages

Cornell University
Final Examination
May 17, 2004

Mathematics 171
Statistical Theory and Application in the Real World

Time: 2.5 hours

Name: _____

Instructor: _____

Section: _____

Read all of the following information before starting the exam.

*You have 2.5 hours to complete this exam. Please read all instructions carefully, and check your answers. Show all work neatly and in order, and clearly indicate your final answers. Answers must be justified whenever possible in order to earn full credit. **Unless otherwise specified, no credit will be given for unsupported answers, even if your final answer is correct.** Several problems require written explanations in context. Only complete solutions written in the context specified by the problem will be awarded full points, and points will be deducted for incoherent, incorrect, and/or irrelevant statements.*

You may use standard notation; however, any new notations or abbreviations that you introduce must be clearly defined.

*Calculators are permitted; however, you must still show all your work. A copy each of Table A, Table C, and Table E from the textbook will be provided. You are also permitted to have **TWO** 8.5×11 pages of handwritten notes (each page double-sided) for your personal use. Other than these exceptions, no other aids are allowed.*

*This test has **10** numbered pages with **5** questions totalling **100** points. The number of points per question is indicated: each subpart is worth 5 points.*

Solutions will be available on the Math 171 web site after the examinations have all been graded.

DO NOT WRITE BELOW THIS LINE

Problem 1 _____ Problem 3 _____ Problem 5 _____

Problem 2 _____ Problem 4 _____

TOTAL _____

1. (*20 points*) In a well-known carnival game, a player pays \$3 to spin the *Wheel of Fortune*. If the player spins a star, then the player wins \$7. Otherwise, the player wins nothing. Note that there is a 40% chance of spinning a star. Answer the following questions by clearly defining appropriate random variables.

(a) What are the mean and the standard deviation of the player's net winnings after one spin of the *Wheel of Fortune*?

(b) If the person spins the wheel 100 times independently, what are the mean and standard deviation of the average net winnings?

(continued)

(c) What is the approximate probability that the average net winnings in (b) is negative?

(d) If the person spins the wheel over and over again, what is the limiting value of the average net winnings?

2. (15 points) A national restaurant chain conducts a simple random sample of some of its customers. The card these customers are given to fill out asks their opinions about their meal, the service, the cleanliness, and so on. One question asks the customer to rate the quality of the service as poor, below average, average, above average, or outstanding. The following data represent the results obtained for the Ithaca branch of this restaurant chain.

Poor	Below average	Average	Above average	Outstanding
7	14	33	67	19

(a) Describe (i) the population of interest, and (ii) the sample. (iii) What is the variable of interest? Is it categorical or quantitative?

(b) Let p denote the true (but unknown) proportion of customers that would rate the sampled restaurant as either above average or outstanding. Give a point estimate \hat{p} of p , and use \hat{p} to construct an approximate 99% confidence interval for p .

(c) Can the results derived from the given data be extended to *all* other branches in this restaurant chain? Explain why or why not.

3. (*20 points*) The historian Raymond Dumett of Purdue University was examining British colonial records for the Gold Coast in Africa, and he suspects that the death rate was higher among African miners than among European miners. In the year 1936, incomplete records show there were 223 deaths among 33,809 African miners and 7 deaths among 1541 European miners on the Gold Coast. Determine if there is good evidence that the proportion of African miners who died during 1936 was higher than the proportion of European miners who died during that year by answering the following questions.

(a) Define your notation, and clearly state an appropriate null hypothesis and an appropriate alternative hypothesis.

(b) Calculate the test statistic.

(continued)

(c) Using your test statistic, give the corresponding P -value.

(d) Clearly state your conclusion in words. That is, explain whether or not there is good evidence at the $\alpha = 5\%$ significance level that the proportion of African miners who died during 1936 was higher than the proportion of European miners who died that year. Your response must include the phrases *African miners*, *European miners*, *Gold Coast*, and *evidence*.

4. (25 points) How strong is the association between student scores on the Math and Verbal sections of the SAT? The Math and Verbal sections of the SAT are scored separately. Scores for each of these two sections range from 200 to 800, and are widely used by college admissions offices. A popular statistics software package gives the following as summaries and plots of the scores for a recent graduating class at Ithaca High School.

Variable	Count	Mean	Median	StdDev	Range	IntQRang
Verbal	162	596.296	610	95.5199	490	140
Math	162	612.099	630	98.1343	440	150

Dependent variable is: Math

R -squared= 46.9%

$s = 71.75$ with $162 - 2 = 160$ df

Variable	Coefficient	SE(Coeff)
Constant	209.554	34.35
Verbal	0.675075	0.0568

(i)

(ii)

(iii)

- (a) Suppose that we wish to determine if there is evidence of an association between Math and Verbal scores. Clearly state an appropriate null hypothesis and an appropriate alternative hypothesis.

(continued)

(b) Discuss the assumptions necessary for inference; specifically, comment on pictures (ii) and (iii) above.

(c) Test your hypotheses from (a) at the $\alpha = 5\%$ significance level, and state an appropriate conclusion in the context of this problem. Your response must include the phrases *Math section*, *Verbal section*, *association*, and *evidence*.

(continued)

(d) Find a 90% confidence interval for the mean SAT Math score for all students with an SAT Verbal score of 500.

(e) Find a 90% prediction interval for Jennifer's SAT Math score if you know she scored 730 on the Verbal section.

5. (20 points) Many industrial air pollutants adversely affect plants. In a study of the effect of sulfur dioxide in the air on three types of garden vegetables, 40 plants of each type were exposed to a given concentration of sulfur dioxide under controlled greenhouse conditions. After exposure, it was noted if there was severe leaf damage or not. The observed counts are as follows:

	Severe leaf damage	Not severe or no leaf damage	Total
Lettuce	32	8	40
Spinach	28	12	40
Tomato	21	19	40
Total	81	39	120

(a) What is the appropriate null hypothesis of interest here?

(b) Compute all 6 expected counts assuming your null hypothesis H_0 is true. Display your expected counts in the following table. (Show your work for the computation of at least one expected count.)

	Severe leaf damage	Not severe or no leaf damage
Lettuce		
Spinach		
Tomato		

(continued)

(c) Compute the Chi-square statistic and its associated P -value.

(d) Carry out a Chi-square test at the $\alpha = 10\%$ significance level. Based on your analysis, write a brief conclusion. Your response must include the phrases *leaf damage*, *plant type*, and *evidence*.