



NAME: _____

**STAT 2120: Introduction to
Statistical Analysis
Spring 2010**

Midterm Exam, March 22, 2010

This exam is closed book and closed notes but you are allowed to use the formula sheets that are provided for you. You may also use a calculator but not your laptop computer. All cell phones, ipods, and other electronic devices must be turned off (the time will be provided for you). You are not permitted to work in groups or discuss the exam with anyone. The work you submit must be your own. No questions of any kind will be taken during the exam (unless you believe there is an error or typo).

Please fill out your identifying information on the top of the scantron form now. On the right side of your form, write and carefully bubble your UVA computing ID (the front part of your UVA email address, *i.e.*, kcs6g) in the space provided. Left justify your computing ID (*i.e.*, leave space to the right). For each question, completely fill in the label of your answer in the corresponding space on the accompanying scantron form. Each question is designed to have a single correct answer. If it seems that more than one statement is correct, then you should select the statement that best answers the question. If you calculate a numerical answer, but its exact value is not listed as a choice of answer then you should select the listed value that is closest to yours. Also, if the question asks for a probability but states that a “good approximation” is acceptable, then you should assume the context is such that any relevant approximation formula is valid (*e.g.*, the sample size is large enough for validity of the central limit theorem).

Some problems will require an evaluation one of the Excel functions `normsdist` and `normsinv`. For these you should refer to the list of selected evaluations of those functions that is provided with the exam.

Print your name clearly above, sign the honor pledge below, and write out this honor pledge on the space provided on the back of the scantron form. You have 90 minutes to complete the exam. There are thirty-three questions on the exam. Each counts the same amount toward your final score.

Honor Pledge:

I have neither given nor received unauthorized aid on this exam.

Signed: _____

1) In a questionnaire, respondents are asked to mark their self-perceived height as short, just right, or tall. Perceived height is an example of a:

- a) Categorical variable
- b) Quantitative variable
- c) Categorical or quantitative variable, depending on how it is used
- d) None of the above

2) Suppose the distribution of monthly interest paid by a population of credit-card holders is Normal with a mean of \$24 dollars and a standard deviation of \$5. Use the 68-95-99.7 rule to deduce the proportion of the population who pay between \$19 and \$34 of monthly interest.

- a) 68.0%
- b) 77.5%
- c) 81.5%
- d) 95.0%

3) The miles-per-gallon of a sample of eighteen compact cars are:

18, 21, 24, 24, 24, 26, 26, 27, 27, 27, 28, 28, 30, 30, 30, 32, 34, 39

What is the difference between the first and third quartiles of these data?

- a) 3 or less
- b) 4 or 5
- c) At least 6 but no more than 10
- d) More than 10

4) Suppose the distribution of price-earnings ratios for a population of stocks is Normal with a mean of 12.5 and a standard deviation of 2.1. What value is such that 22% of the stocks have a price-earnings ratio higher than that value?

- a) 17.8
- b) 17.1
- c) 16.7
- d) 14.1

5) When is the use of a control group in a clinical trial ethically justified?

- a) It is not clear that the treatment is more effective than a placebo.
- b) All patients in the study would be expected to recover without treatment.
- c) There is no evidence that the treatment has any serious side effects.
- d) All patients are selected by an institutional review board.

6) Which of the following is a principle of experimental design?

- a) Use replication to reduce random variability
- b) Use comparisons to cancel the effects of lurking variables
- c) Use randomization to allocate subjects among treatments
- d) All of the above

- 7) The least-squares regression line is the line that
- minimizes the sum of the squared prediction errors, where a prediction error is the horizontal distance between a point and the line
 - minimizes the sum of the absolute values of prediction errors, where a prediction error is the vertical distance between a point and the line
 - minimizes the sum of the squared prediction errors, where a prediction error is the vertical distance between a point and the line
 - minimizes the sum of the absolute values of prediction errors, where a prediction error is the horizontal distance between a point and the line
- 8) A regression analysis investigating the dependency of fuel consumption (y , in liters per 100 kilometers) on driving speed (x , in kilometers per hour) leads to the least-squares regression line $\hat{y} = 0.813 + 0.077x$. According to this line, what would be expected of fuel consumption if driving speed is increased by 25 kilometers per hour?
- It would increase by 2.738 liters per 100 km
 - It would decrease by 2.738 liters per 100 km
 - It would increase by 1.925 liters per 100 km
 - It would decrease by 1.925 liters per 100 km
- 9) A local community group is interested in the proportion of residents in a city of 100,000 who have a library card. They interviewed 500 of the city's residents, 36% of whom responded by saying that they have a library card. What is the population of interest in this problem?
- The 180 respondents (36%) who were interviewed and said they have a library card
 - About 36,000 residents (36%) of the city who have a library card
 - The 500 residents who were interviewed
 - The 100,000 residents of the city
- 10) Data on selling price and appraised value were collected for sixteen beachfront condominiums. The data on selling price have a mean of 879.4 and standard deviation of 174.7, in thousands of dollars. The data on appraised value have a mean of 732.5 and standard deviation of 163.0, in thousands of dollars. The correlation between selling price and appraised value is 0.95. What is the slope of the least-squares regression line for these data, assuming one is interested in predicting selling price from appraised value?
- 133.6
 - 1.02
 - 0.89
 - 230.1