Math 663-101, Fall 2008

Final Exam

Name:____________

Student ID:_____________

Dec, 16

Must show all work to get full credit!

I pledge I have not violated the NJIT Honor Code______________________

1. a. – d. The following displays the results of a standardized neuropsychological test administered to a sample of high school seniors. Test measures analytic skills and is scored on a scale of 0 to 100, with higher scores indicative of stronger analytic skills. Note that 1|5 = 51 for males and also 6|3 = 63 for females.

Males		Females
1	1	
8 8	2	
4 5 6	3	2 3 3 3 4 4 6 7 8
3 5 6 9	4	4 5 5 6 7 8 8 9 9
1 1 2 3	5	3 3 4 4 5 5 6 7
	6	1 1 1 2 2 3 3 4 4 4 5 5 5 7 7 8 9
3 4 5 5 5 7 8 9	7	2 2 3 3 3 4 5 6 6 7 8
4 5 6 6 6 8 9 9 9	8	2 2 3 6 7
1 2 2 2 3 3 3 3 4 4 4 4 4 4	9	1 3 4
0 0	10	

Compute:

a) The median scores for males.
(2 points)

b) The first, second and third quartiles for **females** and then the interquartile range.
(9 points)

c) Are there any outliers among the **female** scores using quartiles? Why or Why not?
(4 points)

d) Based on your answer to **c)** would the mean or median provide a better estimate for central location of female scores? Again, would the standard deviation or the interquartile range divided by two provide a better estimate for spread of female scores? Give reason for your answers.
(6 points)

2. a. - d. Data show that 20% of all patients who make appointment at a primary-care clinic never show up. If in a given clinic session, 20 patients make appointments,
    a. What is the probability that at most 2 patients do not show up?
    (6 points)

    b. What is the probability that all patients show up?
    (3 points)

    c. How many patients would be expected to not show up?
    (2 points)
    d. How many patients would be expected to show up?
    (2 points)

3. a. – d. We wish to design a study to evaluate patients' satisfaction with the medical care received at a community health center. Before developing a sampling plan, we will attempt to understand the patients in the population according to their insurance status (which might influence their satisfaction with medical care) and gender (investigators have shown that men tend to report more satisfaction with medical care than women). Some descriptive information is available on each patient who has been seen in the community health center within the last 3 years. Assume that patients who have been seen in the last 3 years constitute our population of interest.

	Insurance Status		
Gender	Private Health Insurance	Medicaid	No Insurance
Female	250	450	200
Male	120	680	150

a. What proportion of the population has no insurance?
(3 points)

b. What proportion of the population is female and has no insurance?
(3 points)

c. Among the patients with no insurance, what proportion is female?
(5 points)

d. Are the events "No Insurance" and "Female" independent? Justify your answer.
(7 points)

4. a. – b. Suppose we measure length of hospital stay (in days) for patients undergoing a particular surgical procedure. The mean is 3 days with a standard deviation of 3.2 days.
    a. Does it appear that the length-of-stay data follow a normal distribution? Justify your answer briefly.
    (3 points)
    b. If a sample of 35 patients is selected what is the probability that their mean length of stay is less than 2.5 days?
    (6 points)

5. a. – c. The following data were collected from a random sample of nine asthmatic children enrolled in a research study and reflect the number of days each child missed school during the past 3 months: 6 12 14 3 2 4 7 8 10.
    a. Test the hypothesis that the true mean number of days asthmatic children missed school during the 3-month period is 5 days. Compute the p-value of the test. Use alpha = 0.01.
    (6 points)

    b. Construct a 95 % confidence interval for the mean number of days asthmatic children miss school during a 3-month period. What assumption did you make if any to compute this confidence interval?
    (6 points)

    c. Using the given data, how many children would be required to ensure that margin of error in the estimate of the true mean, is within 1 day with 99% confidence?
    (6 points)

6. a. – d. Data were collected from a random sample of eight patients currently undergoing treatment for hypertension. Each subject reported the average number of cigarettes smoked per day and each was assigned a numerical value reflecting their risk of cardiovascular disease (CVD). The risk assessments were computed by physicians and based on blood pressure, cholesterol level, and exercise status. The risk assessments range form 0 to 100, with higher values indicating increased risk.

Number of Cigarettes:	0	2	6	8	12	0	2	20
Risk of CVD	12	20	50	60	77	5	15	90

a. Compute the correlation between number of cigarettes and risk of CVD. (6 points)

b. Based on this sample, is there evidence that the true correlation between number of cigarettes and risk of CVD is significantly different from zero? Use alpha = 0.01.
(6 points)

c. Compute the equation of the line that best fits the data to predict risk for CVD from the number of cigarettes.

(6 points)

d. What percentage of the variability in risk of CVD is explained by Number of Cigarettes smoked?

(3 points)