FORM I FORM I FORM I

1. Here is the distribution of poor adults who live in a neighborhood in Los Angeles:

poor	Freq.	Percent	Cum.
white male	750	15.00	15.00
white female	500	10.00	25.00
non-white male	2,250	45.00	70.00
non-white female	1,500	30.00	100.00
Total	5,000	100.00	

25% → 60%

UCLA is planning a research study on the neighborhood and will need 3 residents to serve as community advisors. UCLA decides to select 3 at random. Assume independence for all parts of this problem please.

+3 a. What is the probability that they will select at least 2 male residents? (3 points)

$(.6)^3 + 3(.6 \times .6 \times .4) = 0.648$ ✓

+3 b. What is the probability that none of the three selected residents are non-white? (3 points)

all white

$.25^3 \rightarrow 0.015625$ ✓

+3 c. What is the probability that the first white male selected is the third randomly selected resident? (3 points)

$.85 \times .85 \times .15 =$

0.108375 ✓

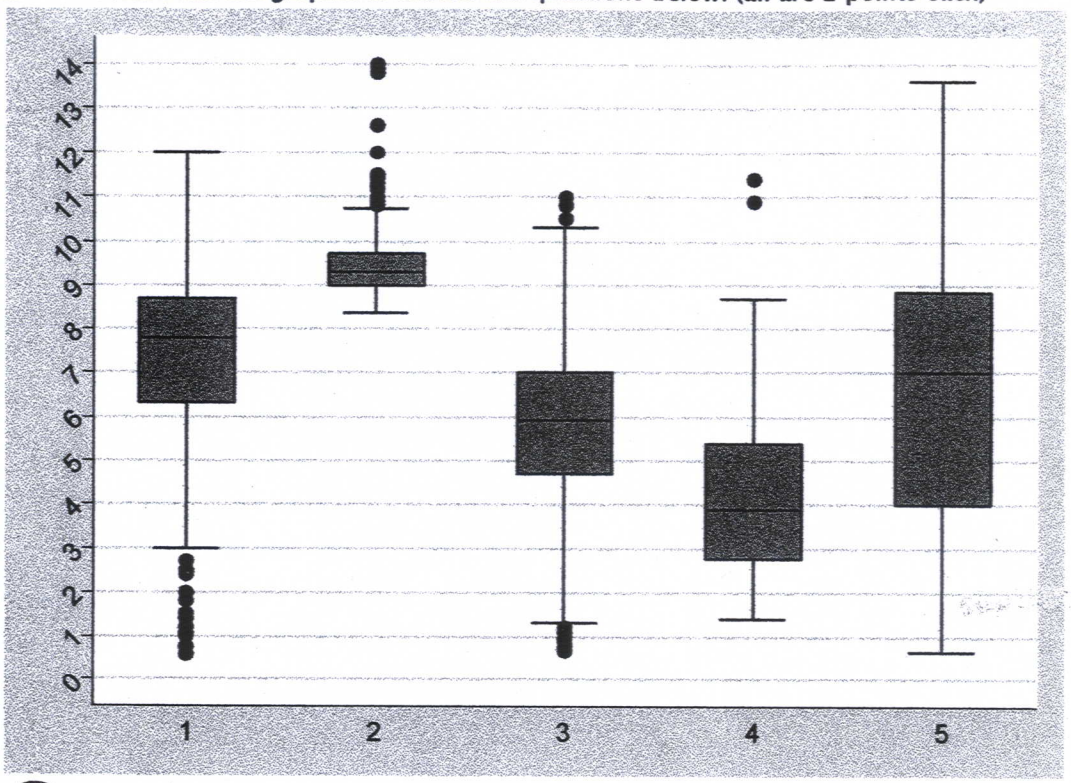
+3 d. Suppose UCLA decides it really needs 7 residents, not 3. Six residents are selected at random, none of them are white males. Is the probability of selecting a white male as the 7th resident (3 points)

A. Equal to the probability of selecting a non-white female?
B. Equal to the probability of selecting a resident from the other 3 categories?
C. One-Half the probability of selecting a non-white male?
(D.) One-Half the probability of selecting a non-white female?
E. Three times the probability of selecting a non-white male?
F. Six times the probability of selecting a white female?

$.85^6 \cdot .15$

/12

2. Please use the graphic to answer the questions below: (all are 2 points each)



a. Which box plot appears to have the strongest right skewness? (circle one)

1 2 3 4 5 can't be determined

b. Which box plot appears to have the strongest left skewness? (circle one)

1 2 3 4 5 can't be determined

c. Which box plot has the largest (or greatest) range? (circle one)

1 2 3 4 5 can't be accurately determined

d. Which box plot has the largest (or highest) median? (circle one)

1 2 3 4 5 can't be accurately determined

e. Which box plot has the largest (or highest) mean? (circle one)

1 2 3 4 5 can't be accurately determined

f. Which box plot has the largest (or highest) Q3 or 3rd Quartile? (circle one)

1 2 3 4 5 can't be accurately determined

g. Which box plot has the largest (or greatest) interquartile range? (circle one)

1 2 3 4 5 can't be accurately determined

/14

3. October is traditionally the beginning of flu season in the United States and this year there is a shortage of flu vaccine. Suppose it is known that 63% of all American adults are concerned about the shortage. In response to public concerns, the medical community has assigned grades to Americans based on their risk of death from flu: "A" for the elderly and babies; "B" for persons with chronic illness; "C" for pregnant women; "D" for those with low risk and "F" for those with no risk. The average duration of illness for all groups combined is 5 days with a standard deviation of 4 days. Only persons with grade "A" (highest risk) are asked to get shots, all others have been asked to wait.

A large pharmaceutical firm has developed an experimental vaccine designed to prevent the elderly from becoming ill from the flu and needs to test its effectiveness. To do so, a simple random sample of 1,200 grade "A" Americans was selected. After interviewing the 1,200, it was discovered that 400 were babies. The elderly were randomized into two groups: 300 controls and 500 treatment. The controls were given a shot that was nothing more than vitamins, the treatment group was given the experimental vaccine that looked exactly like the vitamins. The nurses who administered the shots could not distinguish the vitamins from the vaccine. After 90 days, 30% of the controls had become ill and 20% of the treatment group had become ill. The illness had an average duration of 7 days for the controls and duration was not normally distributed. The controls had an average age of 74.2 years and the treatment group had an average age of 76.8 years.

A. (2 points) The population of greatest interest to the pharmaceutical firm is:

The elderly +2

B. (2 points) The sample of interest to the pharmaceutical firm is:

800 elderly people +2

C. (3 points) The parameter of greatest interest to the pharmaceutical firm is:

proportion of elderly who avoid becoming sick from this vaccine. +3

D. (3 points) The statistic of greatest interest to the pharmaceutical firm is:

difference between proportion of sick in control group to proportion of sick in treatment group. +3

E. (2 points) The variable "assigned grade" is an example of a:

(a) ordinal variable
(b) quantitative discrete variable
(c) quantitative continuous variable
(d) categorical discrete variable
(e) categorical variable

+2

F. (2 points) This is an example of a

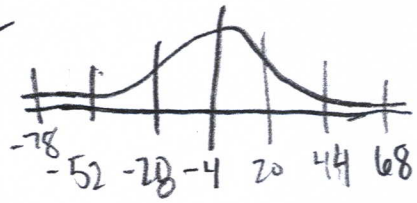
(a) a Randomized Experiment without Controls, but it is blind
(b) a Randomized Experiment without Controls, but it is double-blind
(c) a Randomized Controlled Experiment
(d) an Observational Study with historical controls
(e) an Observational Study utilizing multi-stage cluster sampling techniques
(f) an Observational Study that uses a random probability method for sample selection.

+2

4. A study performed by the Virginia Department of Education on high school students who take test preparation courses for the SAT showed an average -4% change in the SAT scores. The study went on to note that the standard deviation for test takers was 24% and the changes were normally distributed. One student had a change of -82% and was the worst in the population, another student had a change of +119%, the best. Outliers are possible in normal distributions. Both parents and test preparation centers were angered by the study.

A. A student had a -25% change after taking a test preparation course, what percentage of all test preparation course takers performed better than this student? (5 points)

+5



$z = \frac{-25+4}{24} = -0.875 \rightarrow 0.1922$

$1 - 0.1922 = .8078$

80.78%

B. Given the information above, are Q1, Q3, and the interquartile range for change calculable? If yes, please write the word yes in the space below, calculate them, and please show your work. If no, please write the word "no" in the space below and explain why this cannot be done with the information provided. If some are calculable but some are not, write "some" in the space below and calculate the ones than can be calculated (10 points)

+10

YES

Q_3 $.75 \rightarrow$ $.67 = \frac{x+4}{24} \rightarrow 16.08 = x+4 \rightarrow x = 12.08$

Q_1 $.25 \rightarrow$ $-.67 = \frac{x+4}{24} \rightarrow -16.08 = x+4 \rightarrow x = -20.08$

IQR $= 12.08 - (-20.08) = 32.16$

✓

15

FORM I FORM I FORM I

C. Ignore all of the information on the previous page to answer this question. A different study was sponsored by Punkstown Test Preparation Center. This study found out the following: only 4.85% (that's .0485) of all students who take test preparation courses had score changes worse than (lower) than -8% while 20.05% (that's .2005) of students who take the courses had score changes greater (higher) than +10%. They neglected to provide a mean and a standard deviation, but they did mention that the score changes were normally distributed.

Given the information above, is it possible to figure out the mean score change and standard deviation of the score change? If yes, please write "YES" and calculate the values in the space it below. If no, please write "no" and please explain why this cannot be done with the information that has been provided. (10 points)

t9

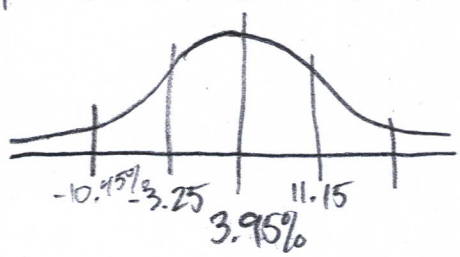
Yes ok

4.85% → $x < -8\%$

20.05% → $x > 10\%$

0.0485 → -1.66σ

$1 - 0.2005 = 0.7995 \rightarrow 0.84\sigma$



$-1.66 = \frac{y-M}{\sigma}$

$-1.66 = \frac{-8-7.2}{\sigma}$ $\quad -1.66\sigma = -15.2$

$\sigma = 9.156$

$0.84 + 1.66 = 2.5$

$10\% + 8\% = 18\%$

$-1.66 = \frac{-8-x}{7.2}$

$-11.952 = -8 - x$

$x = 3.952$

$0.84 = \frac{10-x}{7.2}$

$6.048 = 10 - x$

$x = 3.952$

$\frac{18}{2.5} = 7.2$

$.84 = \frac{10-7.2}{\sigma}$

$.84\sigma = 2.8$

$\sigma = 3.3$

SD = 7.2

mean = 3.952

10