Name ____________________________ Feb. 15, 2007

Math 145 - Elementary Statistics Long Exam I

**Instructions:** *Please write your solutions neatly and using the appropriate notations. Explanations should be written in complete sentences. Encircle your final answers.*

**1.** The following stem-and-leaf plot summarizes the exam scores of a sample of 32 statistics students.

| | |
|---|---|
| 4 | 0 |
| 5 | 8 |
| 6 | 379 |
| 7 | 2445788 |
| 8 | 0013346777899 |
| 9 | 0111268 |

note that $\sum_{i=1}^{32} X_i = 2,578$ and $\sum_{i=1}^{32} X_i^2 = 212,152$.

**a.** Determine the following:

i. Maximum value. 98 [2 pts.]

ii. Sample mean. 80.56 [3 pts.]

iii. Standard deviation. 12 [5 pts.] note : Variance = $\dfrac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$

iv. Median. 83 [2 pts.]

v. The value of first quartile ($Q_1$). 74.5 [2 pts.]

vi. The value of third quartile ($Q_3$). 89 [2 pts.]

vii. The value of IQR. 14.5 [2 pts.]

viii. The lower limit (for the boxplot). 52.75 [2 pts.]

ix. The upper limit (for the boxplot). 110.75 [2 pts.]

**b.** Draw a (horizontal) **modified** boxplot for this data. Are there potential outliers? [5 pts.]
*Please include all relevant values.*

**c.** Construct the frequency table for this data using $40 < 50$, $50 < 60$, ..., $90 < 100$, for the classes (Class $40 < 50$ includes 40 but not 50). Then draw a histogram for this table. [6 pts.]

**d.** What can you say about the skewness of the distribution? *skewed to the left* [2 pts.]

**2.** Suppose that in a sample of 500 college students, 350 drink alcohol, 200 smoke, and 150 drink and smoke. If a student is chosen at random from this sample, let $D$ represent the event of selecting someone who drinks alcohol, and let $S$ represent the event of selecting someone who smokes.

**a.** Draw a Venn diagram that includes these 2 events and determine the number of students that fall in each separate regions. [6 pts.]

**b.** What is the probability that the randomly selected student

i. drinks alcohol but does not smoke? $\frac{200}{500}$ or $\frac{2}{5}$ [4 pts.]

ii. is engaged in at least one of these two habits? $\frac{350}{500} + \frac{200}{500} - \frac{150}{500} = \frac{400}{500}$ or $\frac{4}{5}$ [4 pts.]

iii. smokes given that he/she drinks? $\frac{150}{350}$ [4 pts.]

**c.** Are events $D$ and $S$ mutually exclusive (disjoint)? Explain. [4 pts.]
*No, because there are students who drink and smoke. The intersection of the two events is not empty.*

**d.** Are events $D$ and $S$ independent? Explain. [4 pts.]
*No. Because* $P(D \cap S) = \frac{150}{500} \neq P(D) * P(S) = \frac{350}{500} * \frac{200}{500} = \frac{140}{500}$

**3.** From a lot of 12 missiles, 5 are selected at random and fired. If a lot contains 3 defective missiles that will not fire, what is the probability that

**a.** all 5 will fire? $\frac{\binom{9}{5}}{\binom{12}{5}} = \frac{126}{792} \approx 0.159$ [6 pts.]

**b.** at least 2 will not fire? $\frac{\binom{3}{2}\binom{9}{3}+\binom{3}{3}\binom{9}{2}}{\binom{12}{5}}$ [8 pts.]

**4.** A company uses three different assembly lines $-A_1, A_2, A_3-$ to manufacture a particular component. Of those manufactured by $A_1$, 5% need rework to remedy a defect, whereas 8% of $A_2$'s components need rework and 10% of $A_3$'s need rework. Suppose that 60% of all components are produced by line $A_1$, 30% are produced by line $A_2$, and 10% come from line $A_3$. If a component is randomly selected,

**a.** construct a tree diagram, showing the different possibilities. [5 pts.]

**b.** what is the probability that it needs rework? (.6)(.05)+(.3)(.08)+(.1)(.1)=.064 [5 pts.]

**c.** what is the probability it came from line $A_3$ given that it requires rework? $(.1)(.1)/.064 = .15625$ [5 pts.]

**5.** Give two (2) examples of Categorical and Quantitative variables, then list 2 examples of actual data that you might observe from these variables. [6 pts.]

**a.** Categorical variables

i. *Blood Type* (1)A (2)AB

ii. __________ (1)__________ (2)__________

**b.** Quantitative variables

i. *Height* (1)68 inches (2)175 cm

ii. __________ (1)__________ (2)__________

**6. Essay:** Answer the following questions with at most 3 sentences. [2 pts. each]

**a.** Why do we usually have to work with a sample when we are really interested with the population?
*Because populations are usually too large to study.*

**b.** Why is it important that we work with a *representative* sample?
*So we can generalize the conclusion we obtained from the sample to the whole population.*

**c.** Discuss the main difference between $\mu$ and $\bar{X}$.
*$\mu$ is the mean of the whole population, which is constant at a certain time. $\bar{X}$ is the mean of a sample, whose value changes from sample to sample.*