Math 107 (110 points) Test 1 Pledged Key

Part I **Fill in the blanks: (2 points each, 30 points)**

1. According to Empirical Rule, about 68% of the data falls within one standard deviation of the mean.

2. In stratified sampling, we divide the population into groups according to some characteristic that is important to the study, then sample from each group.

3. What levels of measurement of data is most appropriate for the following:

   1) heights of rose bushes ratio

   2) years in which the bushes were planted ordinal/interval

   3) colors of the rose bushes nominal

4. Likert scales are ordinal (level of measurement) and need nonparametric statistics.

5. If the mean is the appropriate measure of central tendency for a particular set of data, then the best measure of variability is the standard deviation (variance)
   If the median is the appropriate measure of central tendency for a particular set of data, then the best measure of variability is the interquartile range

6. Jerome Cardan wrote Liber de Ludo Aleae or The Book on Games of Chance.

7. In 1936, Literary Digest predicted Alf Landon to be president when Roosevelt won by a landslide.

8. There are 50,400 many ways to arrange the letters in the word STATISTICS. $\frac{10!}{3!3!2!}$

9. The sample space lists or gives all the outcomes of the experiment.

10. Two events are independent if the occurrence of one does not affect the occurrence of the other.
    Two events are mutually exclusive if they cannot occur at the same time. (or disjoint)

11. A coin is tossed 5 times. The probability of getting at least one Head is .96875 OR 31/32

$1 - P(\text{no heads})$

$1 - (.5)^5$

$1 - (\frac{1}{2})^5 = \frac{31}{32}$

Part II Descriptive (30 points) Pearson's index for skewness: $I = \frac{3(\bar{x} - Q_2)}{s}$

Given the following scores of a test:

~~28~~	71	71	73	74	75	[77	80	82	83	84
86	86	86	87	88	90	90	92]	95	98	100

1. Determine if there is an outlier. If so, discard it. Clearly explain your reasoning.

$\bar{X} = 81.6363$
$s = 14.5881$
$Q_3 - Q_1 = 90 - 75 = 15$

$\bar{X} \pm 3s = 81.6363 \pm 3(14.5881)$
$(37.872, 125.4006) \rightarrow$ 28 is an outlier

$Q_1 - 1.5(IQR) = 75 - 22.5$
$= 52.5 \rightarrow$ 28 is an outlier

New: $n = 21$
$\bar{X} = 84.1905$
$s = 8.5301$
$Q_3 - Q_1 = 90 - 76$ $Q_2 = 86$

$Q_3 + 1.5(IQR) = 90 + 22.5 = 112.5$

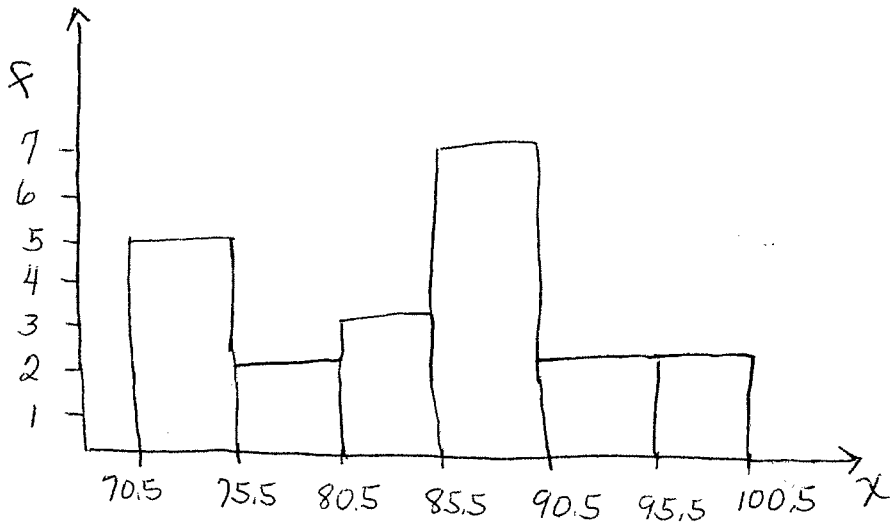
Take out 28

If there were outliers, do not use those data values in the following problems:

2. Draw a frequency histogram using 6 categories that is consistent with the rules for making histograms.

boundaries	f
100.5 - 95.5	2
95.5 - 90.5	2
90.5 - 85.5	7
85.5 - 80.5	3
80.5 - 75.5	2
75.5 - 70.5	5

(An example)



3. Is this distribution significantly skewed? Clearly explain.

$$I = \frac{3(\bar{x} - Q_2)}{s}$$

$$I = \frac{3(84.1905 - 86)}{8.5301} = -.6364$$

not significantly skewed

I is NOT less than -1.00

4. What percentage of data is within one standard deviation from the mean?

$\bar{x} \pm s = 84.1905 \pm 8.5301$

75.66 to 92.72

$\frac{13}{21} = .619$ or about 62%

Part III Counting and Probability (50 points)

1. (10 points) How many 4-digit even numbers can be formed with digits {1,2,3,4,5,6,7} if the digits may not be repeated?

$n = 7$

$6 \cdot 5 \cdot 4 \cdot 3$

$\uparrow$ $\{2,4,6\}$

$\boxed{360}$

2. (15 points) There is a urn with 5 red balls, 7 blue balls and 9 green balls. You select three balls. 5R 7B 9G = 21

a) What is the probability that you get one of each color?

$$\frac{{}_5C_1 \cdot {}_7C_1 \cdot {}_9C_1}{{}_{21}C_3} = \frac{5 \cdot 7 \cdot 9}{1330} = \frac{315}{1330} \doteq \boxed{.237}$$

b) What is the probability that all three are of the same color?

$$\frac{{}_5C_3 \cdot 1 + {}_7C_3 \cdot 1 + {}_9C_3 \cdot 1}{{}_{21}C_3} = \frac{10 + 35 + 84}{1330}$$

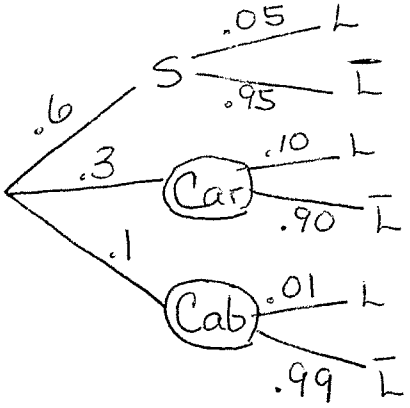
$$\doteq \boxed{.097}$$

c) What is the probability that you get two red balls and one blue?

$$\frac{{}_5C_2 \cdot {}_7C_1 \cdot {}_9C_0}{{}_{21}C_3} = \frac{10 \cdot 7 \cdot 1}{1330} \doteq \boxed{.0526}$$

3. (15 points) A person takes the subway, drives his car or takes a cab with probabilities 0.6, 0.3 and 0.1 respectively. When he takes the subway, he is late 5% of the time. When he takes his car, he is late 10% of the time. When he takes a cab, he is late 1% of the time.

a) Make a tree diagram.



b) Given that the person is late, what is the probability that he took his car?

$$P(\text{Car} \mid L) = \frac{P(\text{Car and late})}{P(\text{late})}$$

$$\frac{(.3)(.10)}{(.6)(.05)+(.3)(.10)+(.1)(.01)} = \frac{.03}{.03+.03+.001}$$

$$\doteq \boxed{.492}$$

c) Given that the person is on time, what is the probability that he took a cab?

$$P(\text{Cab} \mid \bar{L}) = \frac{P(\text{cab and not late})}{P(\text{not late})}$$

$$\frac{(.1)(.99)}{(.6)(.95)+(.3)(.90)+(.1)(.99)} = \frac{.099}{.57+.27+.099}$$

$$\doteq \boxed{.105}$$

4. (10 points) What is the minimum number of people in a room so that the probability of having at least two people with the same birthday is at least 1/4?

$\frac{1}{4} = 1 - P(\text{no two with same b'day})$

OR $P(\text{no two with same b'day}) = \frac{3}{4} = .75$

$$\frac{365 \cdot 364 \cdot 363 \cdots 351}{365^{15}} \doteq .747$$

so, n is $\boxed{15}$

for n people

OR $P(\text{at least two with the same B'day})$

$= 1 - P(\text{no two with same B'day})$

$= 1 - \frac{365P_n}{365^n} \geq \frac{1}{4}$

smallest $n = 15$

$1 - \frac{365P_{15}}{365^{15}} = 0.2529$

$1 - \frac{365P_{14}}{365^{14}} = 0.2231$