AP Statistics Midterm Exam – 2 hours **Name** Key

Directions: Work on these sheets only. Read each question carefully and answer completely but concisely (point values are from 1 to 3 points so no written answer should be long or involved). All tables that you need are given on the last page and you might want to tear it off.

1. 20 people were asked how many movies they saw last month. The results are in the table below:

4	3	2	0	0	2	6	3	4	7
5	4	1	2	1	3	2	3	9	3

a. Find the mean and median *(2 pts)*

$\bar{x} = 3.2$ $M = 3$

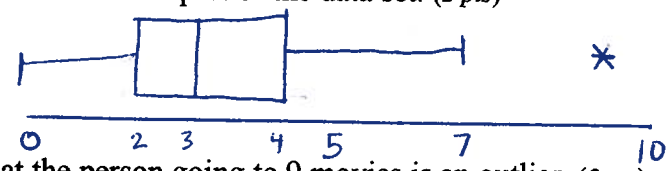
b. **Based on the values you get in a),** what can you say about the shape of the distribution? Explain briefly. *(2 pts)*

Skewed right – mean is greater than median.

c. The persons observing 5 movies would be in what percentile of the sample. *(2 pts)*

$\frac{17}{20} \approx 85^{th}$ percentile

d. Draw a modified boxplot of the data set. *(2 pts)*



e. Show that the person going to 9 movies is an outlier. *(2 pts)*

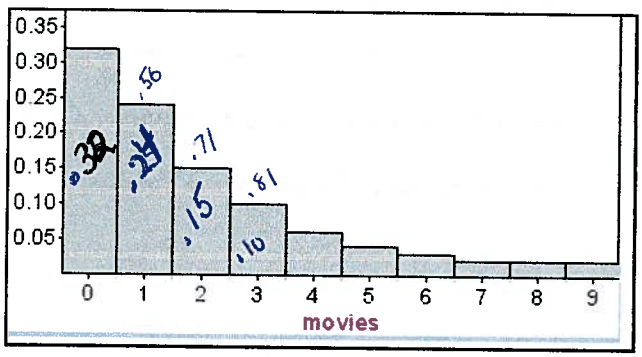
$1.5\ IQR = 1.5(4-2) = 3$

$Q_3 + 3 = 4 + 3 = 7$ over 7 is an outlier.

2. A larger sample of people were polled as to how many movies they saw in the last month. A histogram is below.

a. Circle the type of histogram *(2 pts)*

frequency diagram (relative frequency diagram)



b. What is the shape of the distribution? *(2 pts)*

skewed right

c. What is the median of the data? *(2 pts)*

1

d. What is the IQR of the data? *(2 pts)*

$Q_3 - Q_1 = 3 - 0 = 3$

3. The average cost of an ice cream cone on the Ocean City boardwalk is \$1.97 with a standard deviation of 0.32. (for these questions, work does not need to be shown. Percentage questions should be accurate to one decimal place – ex: 23.6%) $\mu = 1.97 \quad \sigma = 0.32$

a. What percent of the ice cream prices are above \$2.25? (*2 pts*)

$$Z = \frac{2.25 - 1.97}{0.32} = .875 \qquad P(Z > .875) = .1908$$

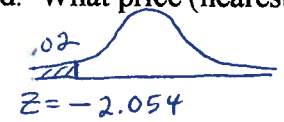
b. What percentage of the ice cream prices are below \$1.50 (*2 pts*)

$$Z = \frac{1.50 - 1.97}{0.32} = -1.46875, \quad P(Z < -1.46875) \approx .0710$$

c. What percentage of the ice cream prices are between \$1 and \$2? (*2 pts*)

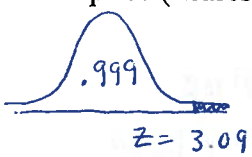
$$P(1 < X < 2) = P(-3.03 < Z < .09) \approx .5361$$

d. What price (nearest cent) would an ice cream cone have to have to be in the bottom 2% of prices? (*2 pts*)



$$-2.054 = \frac{x - 1.97}{.32} \Rightarrow x \approx \$1.31$$

e. What price (nearest cent) would an ice cream cone have to have to be in the 99.9th percentile? (*2 pts*)



$$3.09 = \frac{x - 1.97}{.32} \Rightarrow x \approx \$2.96$$

4. Jay, James, and Joe each drive different types of cars. The price they paid for their cars are in the table below along with the mean and standard deviation of the type of car they drive. Show necessary work to determine who paid the most and least amount for their cars relative to their type. Assume that prices for each type is normally distributed. (*7 pts*)

	Jay	**James**	**Joe**
Type	Compact	SUV	Luxury
Price	\$24,857	\$37,925	\$55,290
μ	\$21,335	\$32,853	\$48,936
σ	\$2,455	\$4,065	\$5,654

$$Z_{JAY} = \frac{24,857 - 21,335}{2455} = 1.44$$

$$Z_{JAMES} = 1.25$$

$$Z_{JOE} = 1.12$$

Jay paid the most and Joe paid the least (relative to the type of car).

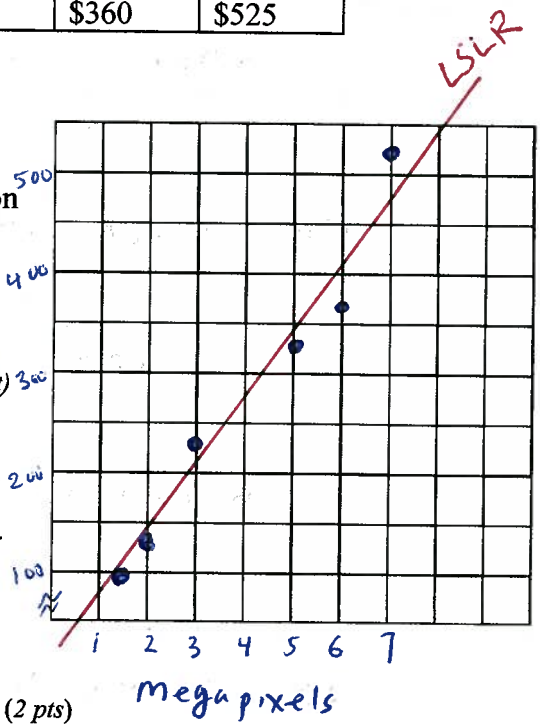
5. Prices of digital cameras are believed to be related to the number of megapixels the camera has. Following is a chart of the average camera price for that number of megapixels.

Megapixels	1.5	2	3	5	6	7
Avg price	$99	$129	$236	$329	$360	$525

a. Draw the scatterplot of the data on the graph to the right. (*3 pts*)

b. Find the least-squares regression line for this data and draw it on the graph to the right (*2 pts*)

$\hat{y} = 69.13x - 2.62$



c. Explain the meaning of the slope in the LSRL for this data. (*2 pt*)

Each megapixel raises the price by about $69.13

d. Compute the residual (nearest dollar) for a 5 megapixel camera. (*2 pts*)

Residual $= y - \hat{y}$

$= 329 - 343 \approx \$-14$

e. Predict the price for a 4 megapixel camera to the nearest dollar. (*2 pts*)

$274

f. Find the values of r and explain its significance in the context of the problem. (*2 pts*)

$r \approx .98$. There is a very strong positive correlation between megapixel and camera price.

g. Find the value of r^2 and explain its significance in the context of the problem. (*2 pts*)

$r^2 \approx .95$. 95% of the variation in camera price is due to the linear relationship between megapixels and price.

h. Every least-squares regression line must pass through a certain point. Find this point for this data. (*2 pts*)

Every LSLR passes through $(\bar{x}, \bar{y})$. This point is $(4.08\overline{3}, 279.\overline{6})$

i. Is there evidence that the least-squares line is a good model for this data? Explain. (*2 pts*)

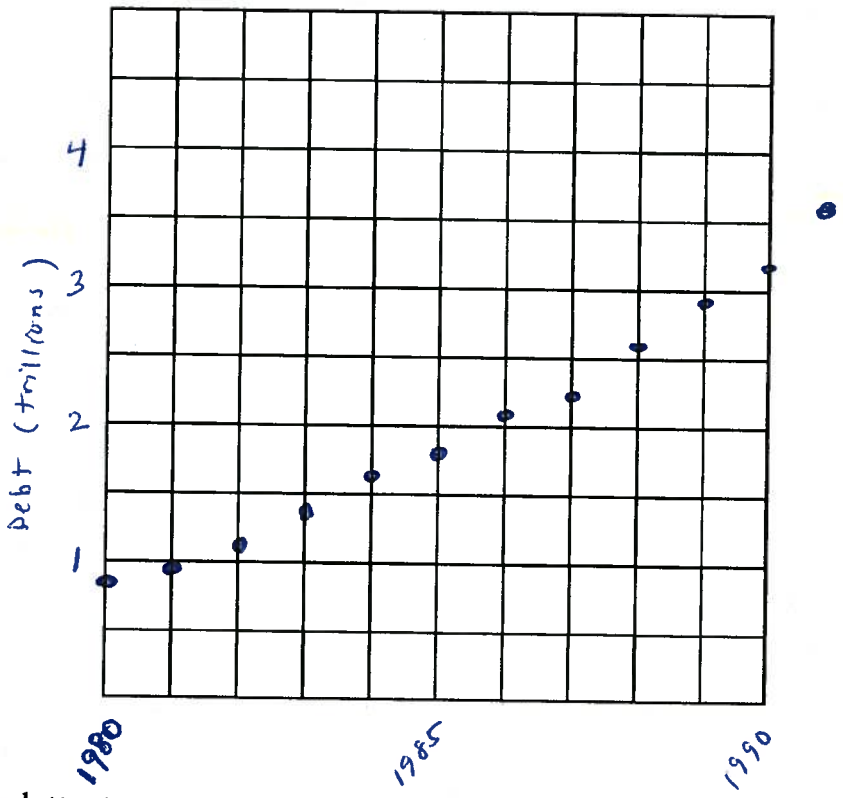
YES, a residual plot shows a random scatter of points with no discernable pattern

j. Predict the average price for a 10-megapixel camera and your confidence in the truth of this value. (*2 pts*)

$689. I'm not very confident since $x=10$ is outside the range of data.

6. The following table shows the federal debt for a short period of time from 1980 through 1991.

Year	Federal debt (in trillions)
1980	0.909
1981	0.994
1982	1.1
1983	1.4
1984	1.6
1985	1.8
1986	2.1
1987	2.3
1988	2.6
1989	2.9
1990	3.2
1991	3.6



a. Construct a scatterplot on the grid provided. (*3 pts*)

b. Determine if the data is linear or exponential and explain how you reached your conclusion. (*2 pts*)

Exponential – when successive y-values are divided, the quotient is always around 1.1.

c. Perform exponential regression and generate an equation that predicts the federal debt based on the year. (*3 pts*) Let x = years after 1980

$$\hat{Y} = .9153 \times 1.1373^{x}$$

d. Find the residual for this model for the year 1990. (*2 pts*)

$$\text{Residual} = y - \hat{y}$$
$$= 3.2 - 3.31 \approx -.11 \text{ trillion}$$

e. Use this model to predict the national debt for the year 2000. (*2 pts*)

$\approx$ 12 trillion

7. In 1912, the *Titanic* struck an iceberg and sank. Some passengers got off the ship in lifeboats but because of the lack of lifeboats, many died. Following is a chart of the breakdown of the *Titanic's* passengers in terms of class and whether they died or survived the tragedy.

	Died	Survived	
First Class	117	183	300
Second Class	163	112	275
Third Class	526	186	712
	806	481	1287

a. Find the marginal distribution of passenger survival in terms of percents. (*2 pts*)

Died	Survived
62.6%	37.4%

b. Find the conditional distribution of survival based on class in terms of percents. (*3 pts*)

	Percent survived
First Class	61%
Second Class	40.7%
Third Class	26.1%

c. The White Star Line (who owed the *Titanic*) said that all classes were treated equally in terms of survival. Do you agree with that or not? Explain why using a statistic. (*2 pts*)

61% of the First Class survived as opposed to only 26% of Third Class so I disagree with the White Star Line.

8. A medical study of heart surgery a drug called a beta-blocker will reduce the pulse rate of a patient during surgery. The pulse rate will be measured at a specific point during the operation. The investigators will use as subjects 10 patients during heart surgery. You have a list of these patients in alphabetical order.

a. Write a description in paragraph form or diagram form for a randomized experimental design for this study. (*3 pts*)

Randomly choose 5 patients to receive the drug and 5 will receive a placebo. Double-blind the experiment so neither the patients nor the medical staff know who's getting the drug. Compare the pulse rates of the experimental group and the control group.

b. Give one lurking variable that may exist in your design. (*1 pt*)

The pre-existing pulse rates of the patients?

9. A suburban school has 4 classes, freshman, sophomores, juniors, and seniors. The administration wants to compare the quality and taste of the current company that provides pizza at lunch (A) with another company (B). It is interested in which pizza the students like the most, A or B. The population of the school is 1,000 students with each class having roughly the same number or students. A stratified sample of 40 students will be chosen to taste the pizza (plain pizza only) with each class having equal representation.

a. Decide exactly how you will label the students in order to select your sample. Since you are going to be using a table of random numbers, think this through first. *(2 pts)*

001 – 250 Frosh
251 – 500 Soph
501 – 750 JR
751 – 000 SR

OR
000 – 249 FROSH
250 – 499 SOPH
500 – 749 JR
750 – 999 SR

ANSWERS WILL VARY

b. Use Table B, beginning at line 130, to select the first 6 students in the sample and write in the parentheses whether the student is a freshman (F), Sophomore (Sh), Junior (J), or Senior (Sr)

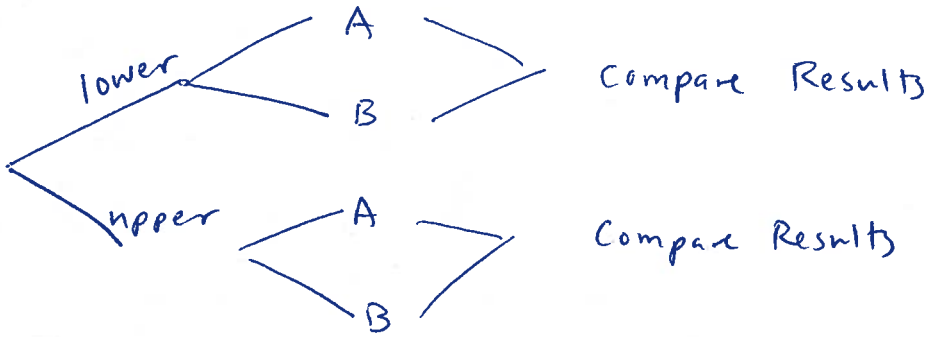
690 (J), 516 (J), 481 (Sh), 787 (Sr), 174 (F), 095 (F)

c. Diagram how the following experiments would be set up.

i) completely randomized *(3 pts)*

Randomly choose 20 students to get Pizza A and the other group gets Pizza B. The students rate the Pizza (1–10?). We compare the ratings between A and B

ii) block design with 2 blocks (lower grades 9 and 10) and upper grades (11 and 12) *(3 pts)*



iii) matched pairs (students acting as their own control) *(3 pts)*

Each student gets either A or B first (random coin flip), rates the pizza (1–10?) then gets the other pizza and rates it (1–10)

d) A decision is made to go with the matched pairs design. Assuming that the experimenters have taken reasonable precautions, describe a lurking variable that could present itself in trying to identify what pizza (A or B) the students prefer. *(1 pt)*

They may prefer the pizza that they have 2nd since it's fresh in their mind which is why we should flip a coin to determine order.

10. 4 fair coins are tossed. Generate a sample space *(3 pts)*

16 outcomes

1 HHHH

4 HHHT, HHTH, HTHH, THHH

6 HHTT, HTHT, HTTH, TTHH, THTH, THHT

4 TTTH, TTHT, THTT, HTTT

1 TTTT

Find the following probabilities

a. There are exactly 2 heads *(1 pt)*

$\frac{6}{16} = \frac{3}{8}$ OR .375

b. There are at least 2 heads *(1 pt)*

$\frac{11}{16}$

c. There are no heads *(1 pt)*

$\frac{1}{16}$

d. There is at least one head *(1 pt)*

$\frac{15}{16}$

11. Two fair 4-sided dice (with the numbers 1-2-4-6 on them) are rolled and added. Generate a sample space. *(3 pts)*

	1	2	4	6
1	2	3	5	7
2	3	4	6	8
4	5	6	8	10
6	7	8	10	12

Find the following probabilities

a. The sum is even *(1 pt)*

$\frac{10}{16} = \frac{5}{8}$

b. The sum is greater than 4 *(1 pt)*

$\frac{12}{16} = \frac{3}{4}$

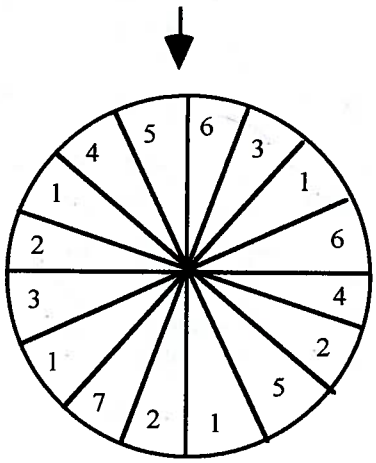
c. If the sum is even, it is greater than 4 *(1 pt)*

$\frac{8}{10} = \frac{4}{5}$

d. If the sum is greater than 4, it is even *(1 pt)*

$\frac{8}{12} = \frac{2}{3}$

12. Suppose you spin a "wheel of fortune" with 16 equally spaced slots as shown below.



a. The wheel is spun once and the number that comes under the arrow (X) is observed. Generate a probability distribution table below. **Answers should be in fraction form.** *(2 pts)*

X	1	2	3	4	5	6	7
$P(X)$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$

b. Let A be the event: The number that comes out is divisible by 2. List the outcomes in A and find $P(A)$. *(2 pts)*

$P(A) = P(2) + P(4) + P(6)$

$= \frac{3}{16} + \frac{1}{8} + \frac{1}{8} = \frac{7}{16}$

c. Let B be the event: The number that comes out is divisible by 3. List the outcomes in B and find $P(B)$. *(2 pts)*

$P(B) = P(3) + P(6)$

$= \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$

d. Are the events A and B disjoint? Explain. *(2 pts)*

No, you can have an outcome that is a multiple of 2 and 3, namely a 6.

$P(X=6) = \frac{1}{8}$.

e. *Show work* using the formal definition of independence to determine if events A and B are independent. *(2 pts)*

$P(A) \cdot P(B) \stackrel{?}{=} P(A \cap B) = P(X=6)$

$\frac{7}{16} \cdot \frac{1}{4} \stackrel{?}{=} \frac{1}{8}$

$\frac{7}{64} \neq \frac{1}{8}$ ∴ Not Independent