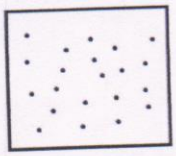
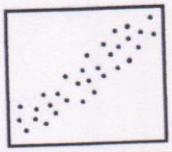
Marginal probability distribution on 2-way table.

Math 247: Test 2 (Fall, 2018)
(100 points)

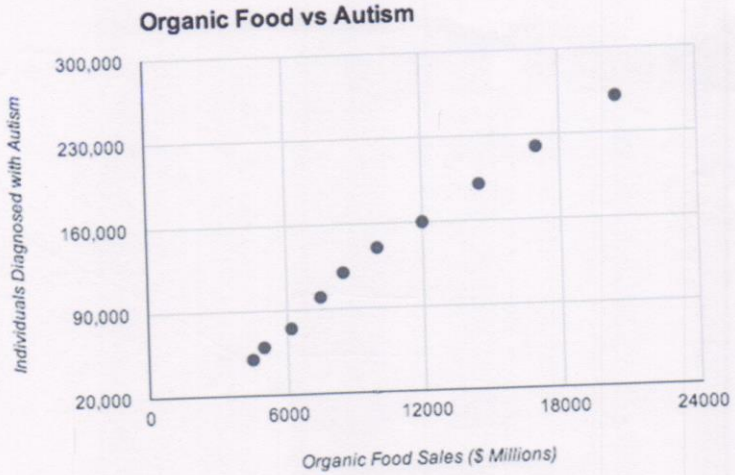
Name: KEY
Class Time: _____

1. (3 pts) Label each scatterplot with one of the following:

Linear Association, Non-linear Association, No Association

A. No association	B. linear association	C. Non-linear association
		

2. (8 pts) The following graph shows the relationship between organic food sales in the U.S. and autism rates.



3 (a) Choose the Correlation Coefficient that best describes this graph (circle your answer):

r = 0 r = -1 r = 2 r = .95 r = -.95

2 (b) Circle the correct answer: This data shows that organic food sales and autism rates. . .

(i) are strongly negatively correlated

(ii) are weakly negatively correlated

(iii) have zero correlation

(iv) are weakly positively correlated.

(v) are strongly positively correlated

3 (c) **True** or **false**: Organic food causes autism.

3. (20 pts) A sample of 6 households was monitored for one year. The household income (in thousands of dollars) and the amount of power they used (in kilowatts) is given in the table

The Correlation Coefficient and Regression Line Equation are also given.

Income ($ thousand)	31	40	23	48	195	96
Power (Kilowatts)	5	13	10	15	61	33

Correlation: Income , Power
Correlation Coefficient = 0.991

Regression Line Equation:
Power = - 0.116 + 0.3180 Income

4 (a) Which variable is the predictor: Household Income ($ thousand)

and which is the response: Power usage (KWatts)

2 (b) The r-value is .991. Judging by this, how close would the data points be to the regression line?

(Very close) Somewhat close Somewhat scattered Very scattered Can't tell

4 (c) Find the Coefficient of Determination, r-squared, and interpret it in the context of the problem.

(1) $r^2 = (.991)^2 = .982 = 98.2\%$

(3) 98.2% of the variation in power usage can be explained by (or predicted by) household income.

4 (d) What is the slope of the regression line? (1) $.3180 = \frac{.3180}{1} = \frac{\text{change in Power Use}}{\text{change in Income}}$

(3) What does the slope mean in terms of income and power? Be specific and use units.

For every increase of $1 thousand in household income, power usage increases by .3180 kilowatt hours.

4 (e) Use the regression equation to predict how much power a household would use annually if their household income was $100 thousand dollars. Income = 100

(3) Power = .116 + .3180(100) = 31.916 KW ~~usage~~ by a household with an income of $100K. — about 32 KW is the predicted power

(1) This value is (choose one) an extrapolation (a prediction (no extrapolation))

2 (f) Locate the y-intercept in the equation and interpret it in the context of the problem. Is it meaningful?

The y-intercept is −.116. This means a household with $0 income would be using negative power (which means they would be producing power with no income!) So, no, this isn't meaningful in this situation.

4. (4 pts) Use your knowledge of the world to determine whether the following pairs of events are mutually exclusive (ME) or not mutually exclusive

A person lives full time in SLO A person lives full time in Paso	ME	not ME	can't tell
A student is a business major. A student is on the basketball team.	ME	not ME	can't tell

5. (2 pts) Determine which of the following variables is continuous and which is discrete (circle the answer):

X = the number of cars not stopping at a stop sign.	DISCRETE	CONTINUOUS
X = the weight of a 2-year-old boy	DISCRETE	CONTINUOUS

6. (8 pts) (a) If you were to flip a coin 3 times, list all the possible outcomes. Use H for heads, and T for tails

2 TTT THH TTH HTH THT HHH HHT HTT

(b) Fill in the probability distribution for the number of heads obtained on three flips.

6

X = number of heads	0	1	2	3
P(X) = probability	$\frac{1}{8}$ = .125	$\frac{3}{8}$ = .375	$\frac{3}{8}$ = .375	$\frac{1}{8}$ = .125

7. (14 pts) Suppose you have a bag with 1 yellow marbles, 3 red marbles, and 6 blue mables. Find the following probabilities and express each as a fraction, a decimal, and a percent.

10 marbles total

(a) If you choose one marble,

2 a. what is the probability it will be red? $P(red) = \frac{3}{10} = .3 = 30\%$

2 b. What is the probability it won't be red? $P(red^C) = 1 - .3 = .7 = 70\%$

2 c. What is the probability it will be yellow? $P(yellow) = \frac{1}{10} = .1 = 10\%$

2 d. What is the probability it will be red or yellow? $P(Red \text{ OR } Yellow) = P(Red) + P(Yellow) = .30 + .10 = .40 = 40\%$

3 (b) If you choose two marbles **with replacement**, what is the probability both will be red?

$P(Red \text{ and } Red) = \frac{3}{10} \cdot \frac{3}{10} = .09 = 9\%$

3 (c) If you choose two marbles **without replacement**, what is the probability both will be red?

$P(Red \text{ and } red) = \frac{3}{10} \cdot \frac{2}{9} = .067 = 6.7\%$

1 point change numbers

8. (18 points) In Montreal, Canada, an experiment was done with parents of children who were thought to have a high risk of committing crimes when they became teenagers. Some of the families were randomly assigned to receive parental training, and the others were not. The results are summarized in the Two-Way Table below.

	Parental Training	No Parental Training	
Arrested (by age 15)	6	38	44
Not Arrested (by age 15)	37	86	123
	43	124	167

2 (a) What is the research question?

Can parental training cause (experiment!) at-risk youth to stay out of trouble?

2 (b) What are the variables?

Indep. Parent Training status

Dep. Arrest Status

3 (c) What percentage of the entire group were not arrested by age 15?

$\frac{123}{167} = .737 = 73.7\%$ were not arrested.

3 (d) What percentage were not arrested, given that their parents had training?

$\frac{37}{43} = .860 = 86\%$

3 (e) What percentage were not arrested, given that their parents did not have training?

$\frac{86}{124} = .694 = 69.4\%$

4 (f) Are the variables you described in part (b) associated or independent? Explain, and include the percentages you found in the cxplanation.

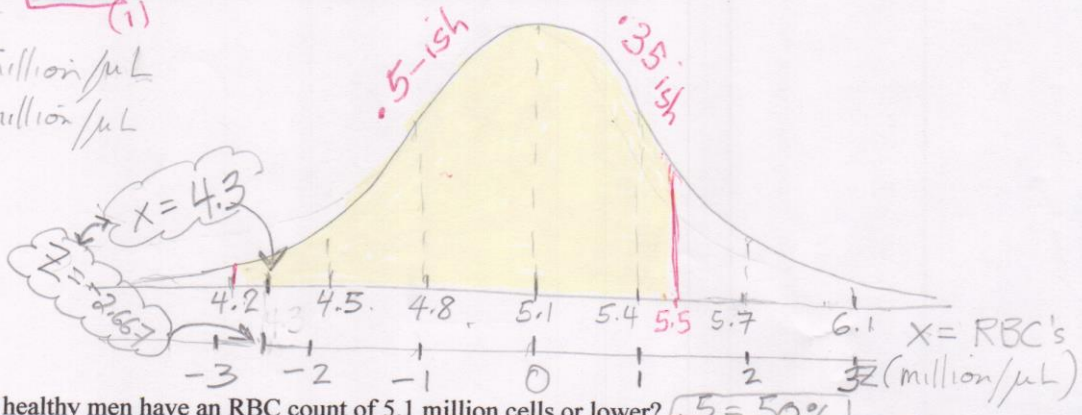
2, explain 2.

Parental training IS associated with arrest status, for just this group, because having training or not resulted in different percentages, i.e overall 73.7% of these kids stayed out of trouble, but 86%, a big increase, of the kids with trained parents stayed out of trouble, while only 69.4% of the kids with untrained parents stayed out of trouble.

9. (20 pts) Blood Cell Count. The distribution of red blood cell (RBC) counts in healthy men is approximately normal, with a mean of 5.1 million cells per microliter and a standard deviation of 0.3 million cells per microliter.

4 (a) Sketch a normal curve showing the distribution of RBC's . Clearly label the x-axis and fill in all relevant tick marks. Also sketch the z-axis below the x-axis and fill in all relevant tick marks.

parking lot
$\mu = 5.1$ million/μL
$\sigma = 0.3$ million/μL



2 (b) What percentage of healthy men have an RBC count of 5.1 million cells or lower? .5 = 50%

2 (c) What would be an unusually high RBC count? Over 5.7 million cells

2 (d) What would be an unusually low RBC count? Under 4.5 million cells

unusual
$z > 2$ or $z < -2$
wildly unusual
$z > 3$ OR $z < -3$

4 (e) Shade the region that represents the percentage of healthy men with RBC counts between 4.2 and 5.5 million cells per microliter.

Which of the following is the best estimate of this percentage? (Circle the best answer)

(i) 50% (ii) 68% (iii) 84% a bit low (iv) 16%

5 (f) Find the z-score for an RBC count of 4.3 million cells and interpret it in the context of the problem.

Plot both 4.3 and the z-score on the graph you made in part (a).

$X = 4.3$
$Z = \frac{X - \mu}{\sigma}$

$Z = \frac{4.3 - 5.1}{.3} = -2.667$

(1) interpret

An RBC count of 4.3 million cells is extremely low since it is 2.667 SD's from the mean. I'd guess this person is anemic!

10. (3 pts) Suppose a man has an RBC count of has an RBC count with a z-score of .723. What can you deduce from this (just regarding the RBC's and what would you tell him?

(a) The man has slightly more RBC's than average. "Your RBC count is fine."

(b) The man has slightly fewer RBC's than average "Your RBC count is fine."

(c) The man has many, many more more RBC's than average. "I'm concerned about your RBC count."

(d) The man has a lot fewer RBC's than average. "I'm concerned about your RBC count."

(e) Can't tell from this information.