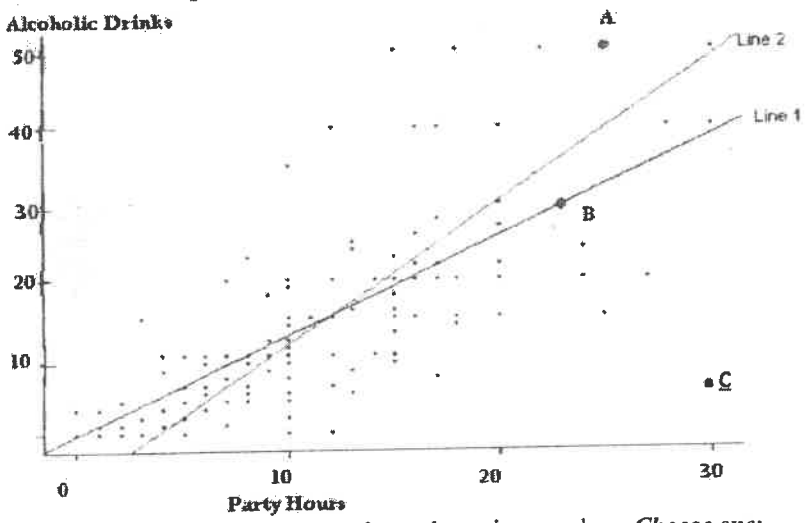
Page 1 - March 8

Statistics 100 Exam 2 March 8, 2017

Question 1 pertains to the scatter diagram below which shows party hours per week on the X axis and the number of alcoholic drinks per week on the Y axis for the 148 Greek members who answered Survey 2 this semester. (12 pts.)



a) (1 pt) The average number of party hours is around ***Choose one:*** i) 10 ii) 12 iii) 15

b) (2 pts) Which is the regression line? ***Choose one:*** i) Line 1 ii) Line 2

c) (2 pts) The correlation between partying and drinking is closest to ***Choose one:*** i) 0 ii) 0.3 iii) 0.7 iv) 1

d) (1 pt) The residual for person A is closest to ***Choose one:*** i) 0 ii) -10 iii) -20 iv) 10 v) 20

e) (1 pt) The residual for person B is closest to ***Choose one:*** i) 0 ii) -10 iii) -20 iv) 10 v) 20

f) (1 pt) If a new scatter plot was drawn with partying measured in minutes instead of hours then the correlation between party hours and drinks would.... ***Choose one:***
i) increase ii) decrease iii) stay the same

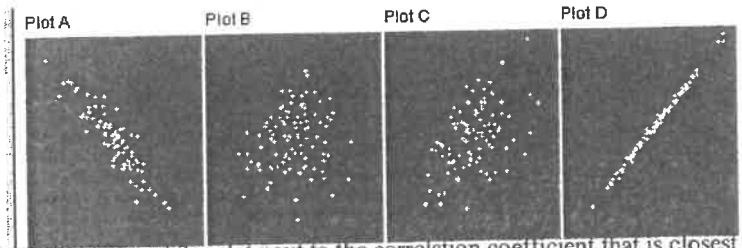
g) (1 pt) The regression equation for predicting drinks from party hours is: **Drinks = 1. 3 (Party Hours) + 0. 3**
If a student parties 9 hours per week, predict how much he drinks per week? ***Circle answer.***

drinks = 1.3(9) + 0.3 = 12

h) (2 pts) If point C was removed from the scatter plot above, the correlation coefficient would....
Choose one: i) increase ii) decrease iii) stay the same

i) (1 pt) One student is exactly average in both drinks and party hours. Which line does he fall on?
Choose one: i) Line 1 only ii) Line 2 only iii) Both iv) Neither

Question 2 pertains to the 4 scatter plots below: (4 pts)



Write the letter of the plot next to the correlation coefficient that is closest to it

r = -0.9 A r = 0.13 B r = 1.00 D r = 0.49 C

Page 2 - March 8

Question 3 (9 pts total)
Fill in the blanks of the table below and compute the correlation coefficient : (4 pts--1/2 for each blank)

Average of X = 1, SD of X = 1 **Average of Y = 2, SD of Y = 2**

X	Y	Zx	Zy	Zx * Zy
0	2	-1	0	0
1	0	0	-1	0
3	6	2	2	4
0	0	-1	-1	1
1	2	0	0	0
1	2	0	0	0

a) (1 pt) The correlation coefficient (r) = 5/6 or 0.83 (round to 2 decimals)

b) (1 pt) If -1 is added to all the y values the correlation coefficient would ..
i) stay the same ii) change sign iii) impossible to tell

c) (1 pt) If all the original y values are doubled the correlation coefficient would....
i) stay the same ii) change sign iii) impossible to tell

d) (1 pt) If all the original y values are multiplied by negative 2 the correlation coefficient would ...
i) stay the same ii) change sign iii) double and change sign

e) (1 pt) If the first and second values of x (the 0 and 1) were switched, Would the correlation coefficient stay the same?
i) Yes ii) No

Question 4 (5 pts)
For each of the following pairs of variables, check the box under the column heading that best describes its correlation.

	Correlation	Exactly -1	Between -1 and 0	About 0	Between 0 and 1	Exactly +1
a)	Temperature in Fahrenheit and Temperature in Centigrade (F = 9/5 C =32)	☐	☐	☐	☐	☑
b)	Shoe Size and Reading Level among all elementary school children	☐	☐	☐	☑	☐
c)	Shoe Size and Reading Level among 6th grade children only	☐	☐	☑	☐	☐
d)	Hours Spent Exercising each Day and Body Fat percentage	☐	☑	☐	☐	☐
e)	The number of heads and the number of tails in 100 tosses of a coin.	☑	☐	☐	☐	☐

Question 5 pertains to drawing at random from a bin containing the twelve boxes below. (4 pts.)
One of the boxes labeled "1", two of the boxes labeled "2", and one of the boxes labeled "3" have prizes inside them. The rest are empty. **For parts a-d, *please circle your answer choice.***

a) What is the chance of drawing 3 boxes ***without replacement*** and getting first a #1, then a #2 , and last a #3 box?
i) 1/12 * 2/11 * 3/10 **ii)** 1/6* 1/4 * 1/2 **iii)** 6/12* 4/12 * 2/12 **iv)** 6/12* 4/11 * 2/10 **v)** 6/12* 5/11 * 4/10

b) What is the chance of drawing 2 boxes ***with replacement*** and getting a prize in the first box but not in the second?
i) 1/6 * 6/12 **ii)** 4/12* 8/12 **iii)** 1/6 + 2/4 + 1/2 **iv)** 1/6* 2/4 * 1/2 *8/12 **v)** 4/12 + 8/12

c) Draw one box at random. What is the chance that it is ***either*** odd-numbered ***or*** has a prize in it?
i) 8/12 + 4/12 ii) 8/12 + 4/12 - 2/12 iii) 8/12*4/12 iv) 8/12*4/12

d) Draw 4 boxes at random ***without replacement***. What is the chance that ***at least one*** of the boxes has a prize?
i) 4/12 * 3/11 * 2/10 *1/9 **ii)** 1- 8/12 * 7/11 * 6/10 *5/9 **iii)** 1- 4/12 * 3/11 * 2/10 *1/9 iv) 1- $(8/12)^4$

Question 6 (19 pts.)
The heights and weights of the 386 female students who filled out survey 1 this semester yielded the following 5 (rounded) summary statistics:

	Average	*SD*
Height	*65"*	*2.5"*
Weight	*135 lbs.*	*22 lbs.*

Correlation: r = 0.5

a) Student A and Student B are both exactly 1 SD above average in height.

i) Student A falls right on the SD line, how much does she weigh? 157 lbs. (1 pt)

ii) Student B falls right on the regression line. How much does she weigh? 146 lbs (1 pt)

b) (3 pts) One of the women who filled out the survey is 60" tall. What's your best estimate for how much she weighs in pounds? (Solve by the 3-step process.) Show work.

i) Convert her height to Z score: Z= -2 ii) Z score for weight = -1 iii) Her weight = 113 lbs.
cont error

c) (1 pt) There's about a 68% chance that your regression estimation in (b iii), give or take about _____ lbs. is correct.
Choose one.
i) 19 ii) 22 iii) 2.5 iv) 2.2 v) 1.9 vi) 0

d) (1 pt) Suppose you don't know the survey respondent's height, what is your best guess for her weight? 135 lbs.

e) (1 pt) There's about a 68% chance that this guess, give or take _____ lbs. is correct.
Choose one:
i) 19 ii) 22 iii) 2.5 iv) 2.2 v) 1.9 vi) 0

f) (2 pts) One student is exactly average in height and weighs 145 lbs. What is her residual when predicting weight from height?
residual = actual - predicted
= 145 - 135 = 10
10 lbs.

g) (1 pt) What is the average of all the residuals when estimating weight from height? 0 (No work is necessary.)

h) (2 pts) What is the SD of all the residuals when estimating weight from height? ***Show work. Circle answer. Round to 2 decimal places.***
$\sqrt{1-0.5^2} \times 22 = 19.05$

i) (2 pts) What is the slope of the regression equation for ***predicting weight*** from height? ***Show work. Circle answer.***
$r \times \frac{SD_y}{SD_x} = 0.5 \times \frac{22}{2.5} = 4.4$

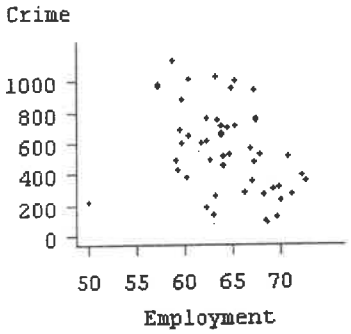
j) (2 pts) The women in our class who are 63" weigh about 126 lbs. on the average. Can you conclude that the women in our class who weigh about 126 lbs. are 63" tall on the average?
Choose one:
i) Yes
ii) No, they'd be taller than 63" on the average.
iii) No, they'd be shorter than 63" on the average.

k) (2 pts) The regression equation for predicting height from weight is : **Height = .06 * (Weight) + b**
Find the intercept (b). ***Show work, write answer in blank below.***
$65 = 0.06(135) + b$
$b = 56.9$
b = 56.9 inches

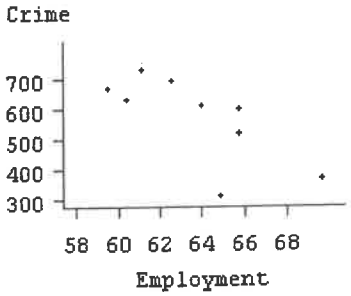
Page 4 - March 8

Question 7 (2 pts.)
Below are 2 scatter plots depicting the relationship between employment rate and crime rate by state and by geographical region. The graph on the left has 50 points: one for each ***individual*** state's employment and crime rate. The graph on the right has the same information condensed into 9 points; one for each geographical region in the country. In other words, each point in the division plot represents the average employment rate and the average crime rate of the states in that region.

By State



By Division



How do the correlation coefficients of the 2 plots compare. *Choose one:*

a) The 2 plots are based on the same data so they must have the same correlation coefficients.
b) This is an example of ecological correlations with the State Plot having a stronger correlation (higher in absolute value) than the Division Plot.
c) This is an example of ecological correlations with the Division Plot having a stronger correlation (higher in absolute value) than the State Plot.

Question 8 (7 pts.)
a) (1 pt) If x is always exactly half of y then the correlation between x and y is
Choose one:
i) 0.5 ii) -0.5 iii) 1.0 iv) -1.0 v) 0 vi) Not enough information is given to determine.

b) (1 pt) The regression line is the same as the SD line when
Choose one:
i) The correlation is 0
ii) The correlation is perfect (1 or -1)
iii) The average and SD of both variables are the same

c) (1 pt) The regression line is a horizontal line through the average of Y when…
Choose one:
i) The correlation is 0
ii) The correlation is perfect (1 or -1)
iii) The average and SD of both variables are the same

d) (2 pts) True or False?
The slope of the regression line for predicting Y from X is the same as r (the correlation coefficient) whenever…

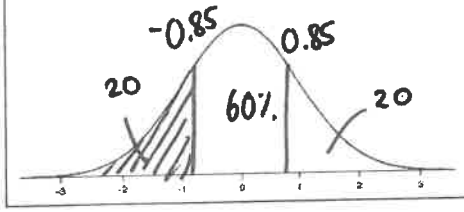
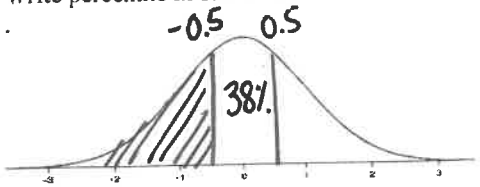
i) the SD of the X's and Y's are the same. *Circle one:* True False
ii) r =1 or -1. *Circle one:* True False
iii) the X's and Y's are in standard units, *Circle one:* True False
iv) the averages of the X's and the Y's are the same. *Circle one:* True False

e) (1 pt) If the RMSE = 0 the correlation coefficient must be …..
Choose one:
i) 0 **ii)** 1 **iii)** -1 **iv)** either 1 or -1 **v)** the same as the slope of the regression line **v)** SDy

f) (1 pt) If r=0 what must the RMSE be … ***Choose one:***
i) 0 **ii)** 1 **iii)** -1 **iv)** either 1 or -1 **v)** the same as the slope of the regression line **v)** SDy

Question 9 (8 pts total)
Suppose scores on a physical fitness test and an IQ test both follow the normal curve but have different correlation coefficients in different schools.

a) (4 pts) At School A the correlation between physical fitness scores and IQ scores = 0.6
If a student is in the 20th percentile in physical fitness, estimate his IQ percentile at School A.

Fitness Percentile = 20th percentile	Fitness Z	r=0.6	IQ Z	IQ Percentile = 31st (1 pt) or 30.8
Mark 20th percentile on curve. What middle area on the table should you look up to find the Z score? 60 % (1 pt.) 	Z = -0.85 (1 pt) -½ for no ⊖	r=0.6	Z = -0.5 or -0.51 (1pt.)	Mark the Z score on the graph below. Write percentile in blank above. 

b) (2 pts) At School B the correlation between physical fitness scores and IQ scores = -1
If a student is in the 20th percentile in physical fitness, his estimated IQ percentile = 80 th percentile at School B.
(No work is necessary.)

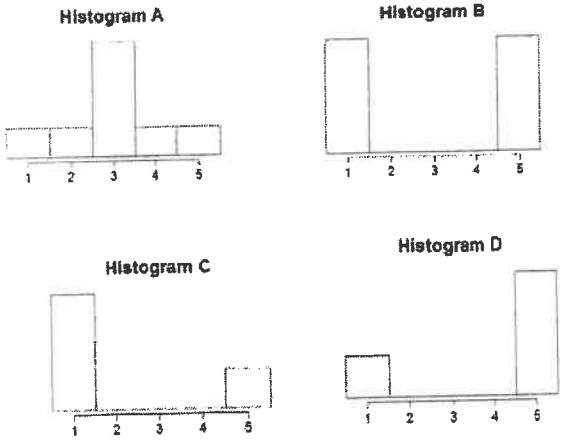
c) (2 pts) At School C the correlation between physical fitness scores and IQ scores = 0
If a student is in the 20th percentile in physical fitness, his estimated IQ percentile = 50 th percentile at School C.
(No work is necessary.)

Question 10 *(7 pts.)*
Below is a distribution table for US income. The right-hand column shows the % of population in each interval. To draw a box plot of the data you'd have to find the median, Q1 and Q3.

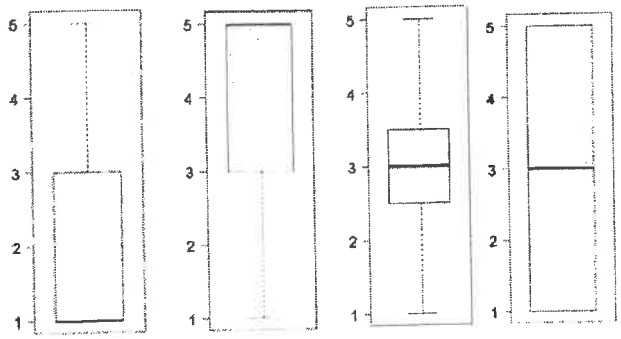
Income	%
$0-$7,500	12
$7,500-$16,000	13
$16,000-$30,000	25
$30,000-$56,000	25
$56,000-$75,000	10
$75,000-$116,000	10
> $116,000	5

a) **Median= $** 30,000
b) **Q1 = $** 16,000
c) **Q3= $** 56,000
d) **The middle 50% of the incomes lie between $** 16,000 **and $** 56,000 . cont error from Q1 + Q3
e) **What percent of the population are high outliers?**
i) 0% ii) 1% iii) 2% (iv) 5% v) 10% vi) 12%
f) **What percent of the population are low outliers?**
(i) 0% ii) 1% iii) 2% iv) 5% v) 10% vi) 12%

Question 11 *(4 pts.)* Which histograms correspond to which box plots? *Write the correct letter in each blank to match the box plots with their corresponding histograms.*



i) C ii) D iii) A iv) B



Question 12 (5 pts total) pertains to a roadside drunk driving test. Suppose only **20%** of those who get stopped and tested for drunk driving are really drunk. **90%** of the drunk drivers will correctly fail the test, but **25%** of the sober drivers will also fail the test (due to nervousness or other reasons). ***Fill in the 8 blanks in the following table*** for a typical sample of 100 drivers who get stopped and tested.

	Fails Test	**Passes Test**	**Total**
Drunk Drivers	18	2	20
Sober Drivers	20	60	80
Total	38	62	**100**

a) (1 pt) Given that a driver fails the test what's the chance he or she was really sober? 20/38 (leave answer as a fraction)

b) (1 pt) Given that a driver passes the test what's the chance he or she was really drunk? 2/62 (leave answer as a fraction)

Question 13 pertains to tossing a fair coin. Circle your answer choice. (3 pts.)

a) What is the chance of tossing a coin 5 times and getting all tails?
$1 - (1/2)^5$ $(1/2)^5$ $(1/2)^5 + (1/2)^5$ $(1/2)^5 + (1/2)^5 - (1/2)^{10}$

b) What is the chance of getting this particular sequence: THTHT?
$1- (1/2)^5$ $(1/2)^5$ $(1/2)^5 + (1/2)^5$ $(1/2)^5 + (1/2)^5 - (1/2)^{10}$

c) What is the chance of getting either all heads or no heads on 5 tosses?
$1- (1/2)^5$ $(1/2)^5$ $(1/2)^5 + (1/2)^5$ $(1/2)^5 + (1/2)^5 - (1/2)^{10}$

Question 14 pertains to a well-shuffled deck of 52 cards. (A deck of cards has 4 suits: clubs, diamonds, hearts and spades. There are 13 cards in each suit: 2 through 10, jack , queen, king, ace.) (4 pts.)

a) Draw 2 cards **with replacement**. What is the chance that the first is an ace and the second is a king?
i) 4/52 + 4/52 ii) 13/52 * 12/51 iii) 16/52 iv) 13/52 * 13/52 v) 4/52* 4/52 vi) 4/52*3/51 vii) 17/52

b) Draw 2 cards **without replacement**. What is the chance that the first is a club and the second is a club?
i) 4/52 + 4/52 ii) 13/52 * 12/51 iii) 16/52 iv) 13/52 * 13/52 v) 4/52* 4/52 vi) 4/52*3/51 vii) 17/52

c) Draw ***one*** card at random. What's the chance that it's ***either*** an ace or a queen?
i) 4/52 + 4/52 ii) 13/52 * 12/51 iii) 16/52 iv) 13/52 * 13/52 v) 4/52* 4/52 vi) 4/52*3/51 vii) 17/52

d) Draw ***one*** card at random. What's the chance that it's ***either*** an ace or a club?
i) 4/52 + 4/52 ii) 13/52 * 12/51 iii) 16/52 iv) 13/52 * 13/52 v) 4/52* 4/52 vi) 4/52*3/51 vii) 17/52

Question 15 pertains to rolling fair dice. (6 pts.)

a) Two dice are rolled. What is the chance that the sum of the spots is 6? 5/36 (2,4 3,3 1,5 4,2 5,1)

b) Two dice are rolled. What is the chance of getting a total of 10 or more spots? 6/36 (5,5 6,4 6,5 4,6 5,6 6,6)

c) One die is rolled 3 times. What is the chance of getting no 2's?
i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1- (5/6)^3$ iv) $1- (1/6)^3$ v) 3/6

d) One die is rolled 3 times. What is the chance of getting at least one 2?
i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1- (5/6)^3$ iv) $1- (1/6)^3$ v) 3/6

e) Three dice are rolled. What is the chance of getting all 2's?
i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1- (5/6)^3$ iv) $1- (1/6)^3$ v) 3/6

f) Three dice are rolled. What is the chance that not all of them are 2's?
i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1- (5/6)^3$ iv) $1- (1/6)^3$ v) 3/6