Stat 100 Exam 2 Oct 18, 2017

STAT 100 EXAM 2 FALL 2017 Key - 18

PRINT NAME_____ _____ _____

(Last name) (First name) net ID (email)

CIRCLE SECTION: L1 (Flanagan 12:30pm) L2 (Flanagan 3:30pm) S1 (Danielle Sass) Online

(This page is worth 1 point. It is graded on writing your name and net id clearly and circling section.)

Write answers in appropriate blanks. When no blanks are provided **CIRCLE** your answers.
SHOW WORK when requested, otherwise no credit.
Do NOT use scrap paper.

Make sure you have all 7 pages including the normal table (16 problems).

For questions using the normal table, you may "round" z scores and percents to fit the closest line on the normal table and you may round percents on the table to the nearest whole number.

DO NOT WRITE BELOW THIS LINE _____

The numbers written in each blank below indicate how many points you missed on each page. The numbers printed to the right of each blank indicate how many points each page is worth.

Page 1 _____17

Page 2 _____19

Page 3 _____11

Page 4 _____19

Page 5 _____16

Page 6 _____17

Cover Page_____1

Total Score _____

There is NO CLASS tomorrow!

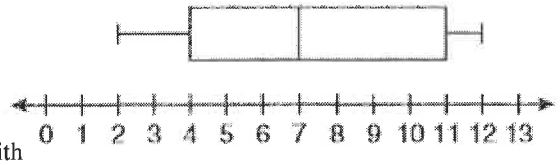
Scores will be posted on Compass by Friday morning (the drop deadline) and exams will be returned in class next week.

Online students may pick up their exam in 23 Illini Hall during office hours next week.

Page 1 - 18

Stat 100 Exam 2 Oct 18, 2017

Question 1 *(6 pts.)* Given the boxplot below, answer the following questions:



a) What is the value of Q3?
i) 2 ii) 4 iii) 7 iv) 11 v) 12

b) Which of the following is the median?
i) 2 ii) 4 iii) 7 iv) 11

c) What percent of data are less than 4? Fill in the blank with a percentage: 25 %

d) Which interval has a greater % of the data: 4-7 or 7-11?
i) They are the same ii) 4-7 iii) 7-11

e) What is the range of the data represented in the boxplot above? i) 2 ii) 7 iii) 10 iv) 12 v) 13

f) Are there any outliers in the boxplot above?
i) Yes ii) No iii) Impossible to tell

Question 2 *(5 pts.)*
Below is a distribution table for a Stat 100 exam. The right-hand column shows the % of students in each interval. To draw a box plot of the data you'd have to find the median, Q1 and Q3.

Score	%
0-45	5
45-75	20
75-85	25
85-95	25
95-100	25

a) Median= 85

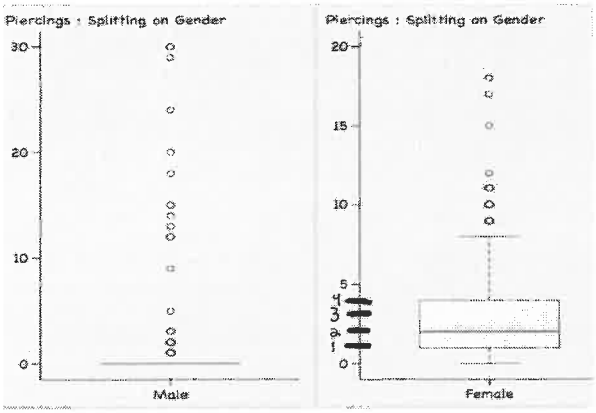
b) Q1= 75

c) Q3= 95

d) What percent of the students are low outliers?
i) 0% ii) 1% iii) 2% iv) 5% v) 25% vi) 45%

e) What percent of the students are high outliers?
i) 0% ii) 1% iii) 2% iv) 5% v) 25% vi) 45%

Question 3 *(4 pts.)* The 2 boxplots below depict the Stat 100 survey responses of 392 males and 778 females to the question: "How many piercings do you have? Fill in the 8 blanks in the table below. ***All answers are whole numbers.***



	Males	Females
Average	0.997	2.97
SD	4.178	2.613
Min	0	0
Q1	0	1
Median	0	2
Q3	0	4
Max	30	18
IQR	0	3
n	392	778

½ point for each blank

- ½ if they switch males + females

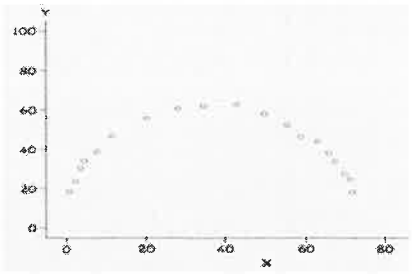
Cont error Q3-Q1

Question 4 *(2 pts.)* The scatter plot on the right shows a definite pattern. Does that mean the correlation must be close to 1 in absolute value? ***Choose one:***

i) Yes, there is a clear pattern so r must be close to 1 in absolute value.

ii) No, r measures how closely points hug a line, not a curve. In this plot r is close to 0 since about half the points slope up and half slope down.

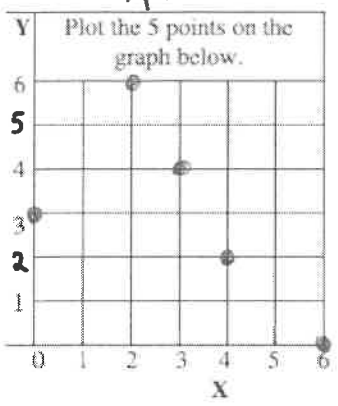
iii) No, since about 50% of the points slope up and 50% slope down, r would be about 0.5



Page 2 -18

Question 5 *(7 pts.)* Compute the correlation coefficient (r) between X and Y by filling in the table below. Then check that your answer makes sense by graphing the points in the box provided. ***The average of X and Y is 3 and the SD of X and Y is 2.**

1 pt



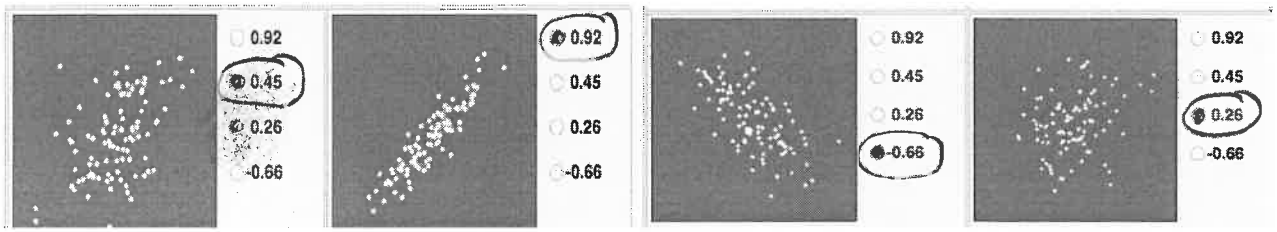
X	Y	X in Standard Units (Zx)	Y in Standard Units (Zy)	Products
0	3	-1.5	0	0
2	6	-0.5	1.5	-0.75
3	4	0	0.5	0
4	2	0.5	-0.5	-0.25
6	0	1.5	-1.5	-2.25

1 pt for each row

$r = \frac{0 + (-0.75) + 0 + (-.25) + (-2.25)}{5}$ What is r? -0.65 *(1 pt)*

Question 6 *(7 pts.)*

a) (4 pts.) Match the scatter plots to their correlation coefficients by circling the correct r next to each plot.



b) *(3 pts.)* How would the correlation coefficients for the 4 plots above change if you did the following?

For each circle all that are true.

i) Switch X and Y for each plot.
a) increase in absolute value b) decrease in absolute value c) stay the same d) change sign

ii) Change X and Y to Z scores in the original plots.
a) increase in absolute value b) decrease in absolute value c) stay the same d) change sign

iii) Multiplied all the X values by -2 in the original plots.
a) increase in absolute value b) decrease in absolute value c) stay the same d) change sign

Question 7 *(5 pts)* For each of the following pairs of variables, check the box under the column heading that best describes its correlation among typical STAT 100 students: (**Hint:** Every column should have exactly one box checked.)

	Correlation		Exactly -1	Between -1 and 0	About 0	Between 0 and 1	Exactly +1
a)	Height in centimeters	Height in inches	☐	☐	☐	☐	☒
b)	Height in inches	Weight in pounds	☐	☐	☐	☒	☐
c)	Hours awake (per 24)	Hours asleep (per 24)	☒	☐	☐	☐	☐
d)	GPA	Height in inches	☐	☐	☒	☐	☐
e)	Distance from home	# times you visit home	☐	☒	☐	☐	☐

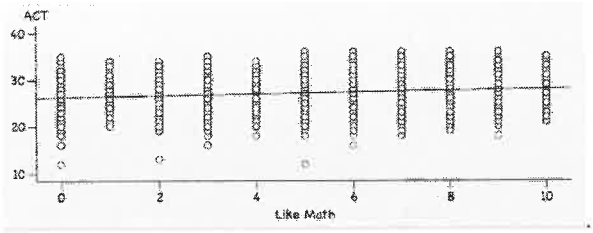
Page 3 - 18

Stat 100 Exam 2 Oct 18, 2017

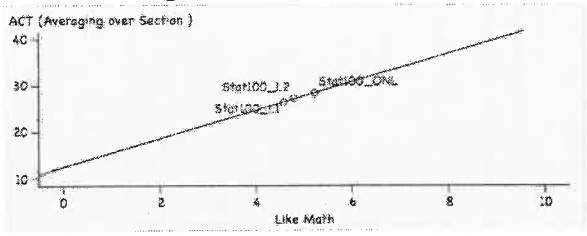
Question 8 *(6 pts.)*

The 2 plots below depict your survey responses to the 2 questions: On a scale of 0-10, how much do you like math? and What's your ACT score?. Plot A shows the individual responses of the 1314 students who filled out the survey and Plot B shows the average responses for each section (L1, L2, and Online)

Plot A- individuals



Plot B- averages



a) *(1 pt)* One of the regression equations below corresponds to Plot A and one to Plot B. Which corresponds to Plot A?
Circle one: (i) ACT= 26.22 + 0.888(Like Math) ii) ACT= 12.47 + 3.044 (Like Math) iii) Not enough info

b) *(1 pt)* One of the correlation coefficients below corresponds to Plot A and one to Plot B. Which corresponds to Plot B?
Circle one: i) 0.1328 (ii) 0.9956 iii) Not enough info

c) *(1 pt)* One of the RMSE's below corresponds to Plot A and one to Plot B. Which corresponds to Plot A?
Circle one: i) 0.07685 (ii) 3.974 iii) Not enough info

For d-f,* make *regression estimates (predictions) using the correct equation. Use the equations given in part a. Fill in the first blank with the prediction and the second with the correct RMSE. Use the RMSEs from part c. ½ point for each blank ***(Round predictions to 2 decimal places.)***

d) *(1 pt)* Doug hates math (Like math = 0), we estimate his ACT = 26.22, give or take about 3.974.

cont error?

e) *(1 pt)* Alma loves math (Like math =10) and is in L1, we estimate her ACT = 35.1, give or take about 3.974.

f) *(1 pt.)* 521 students in the L1 section responded to the survey. Their average Like Math rating= 4.3.
We estimate their average ACT = 25.56 give or take about 0.07685 accept rounded answer

Question 9 *(5 pts)* pertains to drug screening for bus drivers. Bus drivers are given random drug tests. If they test positive for drugs, they fail the test and face losing their jobs. Suppose only **1%** of drivers who get tested for drugs are really using them. If a driver is using drugs, then 95% of the time they will correctly fail the test, but 10% of the clean drivers will incorrectly fail the test.
Fill in the 6 blanks in the following table for 10,000 drivers who get tested. *(3 pts.)*

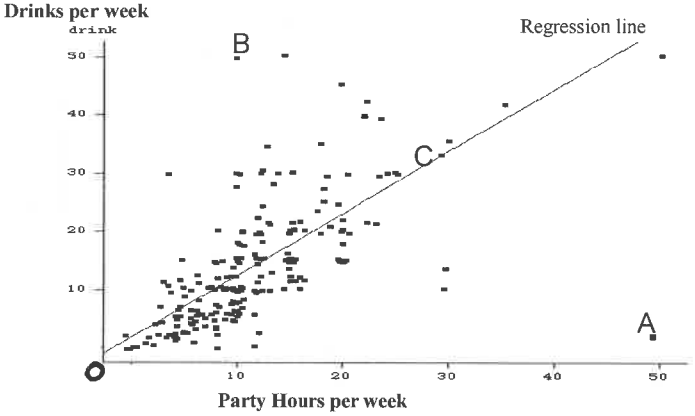
	Fail Test	**Pass Test**	**Total**
Drug Users	95	5	**(Hint: Fill in this blank first. Use underlined info above.)** 100
Clean Drivers	990	8910	9900
Total	**1085**	**8915**	**10,000**

a) *(1 pt.)* If a driver fails the drug test what's chance the he's really clean? 990/1085 (leave answer as a fraction)

b) *(1 pt.)* If a driver passes the test what is the chance he is really using drugs? 5/8915 (leave answer as a fraction)

cont error

Page 4-18



Question 10 *(7 pts.)*
The scatter plot to the right shows the responses of 215 fraternity and sorority members in Stat 100 to the survey questions: "How much do you drink?" and How much do you party?" per week.
Here are the 5 summary stats: ***r = 0.7***

	Average	*SD*
Party hrs per week	***11.5 hours***	***7 hours***
Drinks per week	***13 drinks***	***10 drinks***

a) *(1 pt.)* The regression equation for predicting drinking from partying is: Drinks = 1*(Party hours) + 1.5 ***(Fill in blank with y-int.)*** Compute the y-intercept by using the information provided in the summary stats table above. (Don't just read the y-intercept off the graph) **Show *work.***
$13 = 1(11.5) + b \quad b = 1.5$

b) *(1 pt.)* How would you interpret what the slope in the regression equation above means? ***Choose one:***
(i) For each additional hour that students party, they consume one more drink on the average.
ii) Everyone drinks exactly 1 drink for each hour that they spend partying.
iii) Once you've had a lot to drink it's impossible for anyone to determine what it means.

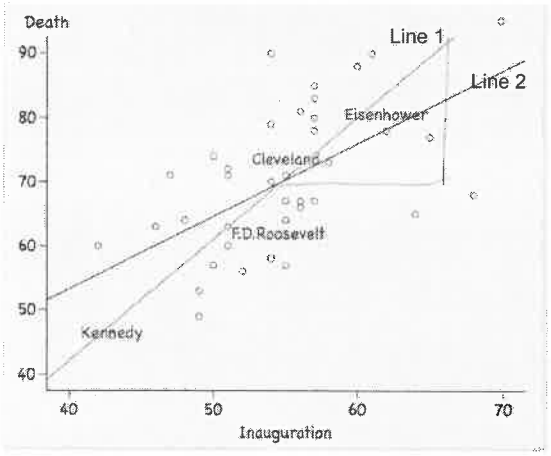
c) *(1 pt.)* How would you interpret what the Y-intercept in the regression equation above means? ***Choose one:***
i) It indicates the average number of drinks for people who party less than average.
(ii) It's the regression estimate for the number of drinks people who party 0 hours have.
iii) It is the exact number of drinks everyone who doesn't party has.

d) *(3 pts.)* Look at Students A, B and C. One has a positive prediction error, one has a negative prediction error and one has 0 error.

i) The prediction error for A is	***Circle one:***	(Negative)	Positive	0
ii) The prediction error for B is	***Circle one:***	Negative	(Positive)	0
iii) The prediction error for C is	***Circle one:***	Negative	Positive	(0)

e) *(1 pt.)* If student A was removed, the correlation coefficient would….
Circle one: (**i)** increase **ii)** decrease **iii)** stay the same.

Question 11 *(12 points)* The scatter plot below shows the age at inauguration and the age at death of 38 US Presidents.
The average age at Inauguration= 55 years with SD= 6 years and the average age at Death = 70 years with SD = 12 years.



2 pts each

a) Which line is the SD line? *Circle one:*
(i) Line 1 ii) Line 2

b) Which line is the point of averages on? *Circle one:*
i) Regression Line Only ii) SD Line Only (iii) Both iv) Neither

c) The correlation between age at Inauguration and age at Death is closest to
Circle one: i) 0 ii) 0.3 (iii) 0.6 iv) 0.9 v) 1

d) The residual for Eisenhower is closest to _____years
Circle one: (i) 0 ii) -6 iii) 6

e) FDR was inaugurated at age 52 and falls exactly on Line 1, how old was he when he died? 64 *Show work. Hint: what is his z-score and which line is he on?*
$\frac{52-55}{6} = -0.5 \quad 70 + (-0.5)(12)$

f) Cleveland falls on both lines, how old was he when he died? 70

Stat 100 Exam 2 Oct 18, 2017

Question 12 *(8 pts.)* Below are the 5 summary statistics for scores on Exam 1 and Exam 2 in a large class.

Average Exam 1= 85 with SD= 5, Average Exam 2 = 77 with SD =10, and r = 0.8

a) *(3pts)* A student scored 75 on Exam 1, estimate his score on Exam 2 by filling in the 3 blanks in table below. ***Don't round answers.***

Exam 1 score = 75	**Z for Exam 1 =** -2 Show work. $\frac{75-85}{5} = -2$	**r =0.8**	**Z for Exam 2=** -1.6	**Exam 2 score =** 61 Show work. $77 + (-1.6)(10)$

cont error

b) *(1 pt.)* There's about a 68% chance that this prediction, give or take _____ points is correct.

Circle one: (i) 6 ii) 8 iii) 10 iv) 16 v) 20

c) *(2 pts.)* Suppose you don't know what a student got on Exam 1.

i) What would be your best prediction for what he got on Exam 2? 77 . **Fill in the blank with a number.**

ii) There's about a 68% chance that this prediction, give or take _____points is correct.

Circle one: a) 4 b) 8 (c) 10 d) 16 e) 20

d) *(1 pt.)* The **regression equation** for predicting Exam 2 scores from Exam 1 scores is … *Circle one:*

i) Exam 2 = 1.2 (Exam 1) - 27 **ii)** Exam 2 = 0.3 (Exam 1) + 49.5 **(iii)** Exam 2 = 1.6 (Exam 1) - 59 **iv)** none of these

e) *(1 pt.)* One student scored 85 on Exam 1 and 77 on Exam 2. What is his residual? **(No work is necessary.)**

Circle one: i) 10 ii) -10 iii) 8 iv) -8 (v) 0

Question 13 *(8 pts.)*

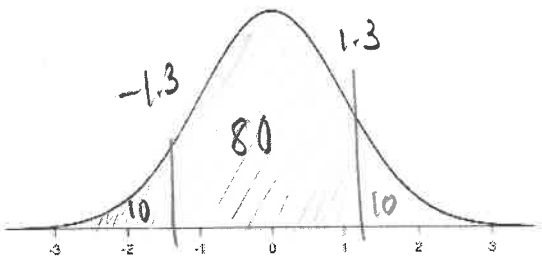
Suppose math skills and athletic skills follow the normal curve but have different correlations in different countries.

a) Consider 5 countries where the correlation coefficients between math and athletic skills are as given in the table below. If someone is in the **90th** percentile in math, estimate the person's athletic percentile for each country. (5 pts.)

Math Percentile	**r**	**Athletic Percentile**
90th	***1***	*Choose One:* 5th 10th 40th 50th 60th **(90th)** 95th
90th	***- 1***	*Choose One:* 5th **(10th)** 40th 50th 60th 90th 95th
90th	***0***	*Choose One:* 5th 10th 40th **(50th)** 60th 90th 95th
90th	***0.2***	*Choose One:* 5th 10th 40th 50th **(60th)** 90th 95th
90th	***- 0.2***	*Choose One:* 5th 10th **(40th)** 50th 60th 90th 95th

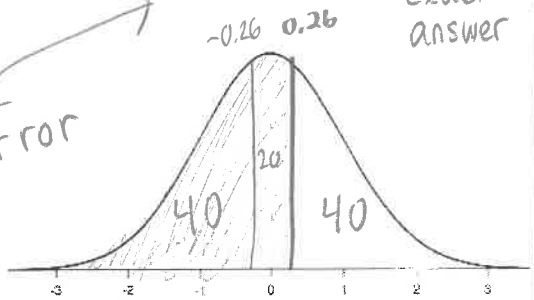
b) Check your work above for people in the ***90th* percentile** in math **where r = 0.2,** what percentile would you estimate for their athletic skills? ***Solve by filling in the table below.*** (You may round areas and z-scores to fit the nearest line on the table.) *(3 pts.)*

Math Percentile	Math Z	r	Athletic Z	Athletic Percentile
90th percentile	Z = 1.3	r = 0.2	Z = 0.26	Athletic Percentile = 60 th



cont error

accept exact answer



Page 6-18

Question 14 *(9 pts.)* The table shows Stat 100 male and female opinions on whether climate change is due to man-made or natural causes or whether it's not happening at all. ***Suppose you randomly draw from these students.***

	Man-Made	Natural Causes	Not Happening	Totals
Male	324	54	14	392
Female	667	87	24	778
Totals	991	141	38	1170

a) What is the chance of getting a male?
i) 991/1170 ii) 38/1170 iii) 324/1170 iv) 392/1170 v) 14/1170 vi) 430/1170 vii) 416/1170

b) What is the chance of getting someone who believes climate change is not happening?
i) 991/1170 ii) 38/1170 iii) 324/1170 iv) 392/1170 v) 14/1170 vi) 430/1170 vii) 416/1170

c) What is the chance of drawing once and getting ***either*** a male or someone who believes it's not happening?
i) 991/1170 ii) 38/1170 iii) 324/1170 iv) 392/1170 v) 14/1170 vi) 430/1170 vii) 416/1170

d) What is the chance of drawing once and getting a male who believes it's not happening?
i) 991/1170 ii) 38/1170 iii)324/1170 iv) 392/1170 v) 14/1170 vi) 430/1170 vii) 416/1170

e) What is the chance you'll get a female if you draw only from those who believe it's not happening?
i) 24/1170 ii) 24/778 iii) 24/38 iv) 778/38 v) 38/778 vi) 778/1170

f) What is the chance you'll get someone who believes it's not happening if you draw only from the females?
i) 24/1170 ii) 24/778 iii) 24/38 iv) 778/38 v) 38/778 vi) 778/1170

g) Draw 3 students ***without*** replacement. What is the chance that ***all*** 3 students are female?
i) $(778/1170)^3$ **ii)** $1-(778/1170)^3$ **iii)** 1-(778/1170*777/1169*776/1168) **iv)** 778/1170*777/1169*776/1168 **v)** $(392/1170)^3$

h) Draw 3 students ***without*** replacement. What's the chance that **not all** 3 students are female?
i) $(778/1170)^3$ **ii)** $1-(392/1170)^3$ **iii)** 1-(778/1170*777/1169*776/1168) **iv)** 778/1170*777/1169*776/1168 **v)** $(392/1170)^3$

i) Draw 3 students ***with*** replacement. What is the chance that at least one student is female?
i) $(778/1170)^3$ **ii)** $1-(392/1170)^3$ **iii)** 1-(778/1170*777/1169*776/1168) **iv)** 778/1170*777/1169*776/1168 **v)** $(392/1170)^3$

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

5,5 4,6 6,4

a) Two dice are rolled. What is the chance that the sum of the spots is 10?
i) 2/36 ii) 3/36 iii) 4/36 iv) 5/36 v) 1/6*1/6 vi) 7/36

5,5 4,6 5,6 6,4 6,5

b) Two dice are rolled. What is the chance that the sum of the spots is 10 or 11?
i) 2/36 ii) 3/36 iii) 4/36 iv) 5/36 v) 1/6*1/6 vi) 7/36

2,2 3,1 1,3

c) Two dice are rolled what is the chance that either the sum of the spots is 4 or that both dice show the same number of spots?
i) 5/36 ii) 6/36 iii) 7/36 iv) 8/36 v) 9/36 vi) 10/36 vii) 11/36

d) What is the chance of rolling a die *7 times and getting at least one "4"*?
i) $(5/6)^7$ ii) $(1/6)^7$ iii) $1-(5/6)^7$ iv) $1-(1/6)^7$ v) 7*(5/6)

e) What is the chance of rolling a die 7 times and getting no "4"s?
i) $(5/6)^7$ ii) $(1/6)^7$ iii) $1-(5/6)^7$ iv) $1-(1/6)^7$ v) 7*(5/6)

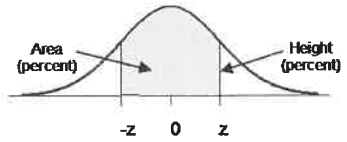
f) What is the chance of rolling a die 7 times and getting all "4"s?
i) $(5/6)^7$ ii) $(1/6)^7$ iii) $1-(5/6)^7$ iv) $1-(1/6)^7$ v) 7*(1/6)

Question 16 pertains to tossing fair coins. *(2 pts.)*

a) What is the chance of tossing a fair coin 3 times and getting all tails?
i) ½ * ½* ½ ii) ½ + ½ + ½ iii) 1- (½ * ½* ½) iv) 3/6

b) What is the chance of tossing a fair coin 3 times and getting this particular sequence: THH?
i) ½ * ½* ½ ii) ½ + ½ + ½ iii) 1- (½ * ½* ½) iv) 3/6

STANDARD NORMAL TABLE



Standard Units

z	Area	z	Area	z	Area
0.00	0.00	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991