

**22S:25 TA REVIEW PROBLEMS – EXAM #1 – FALL 2008 / SPRING / FALL 2009**

- 1) A statistician collects the following data and would like it organized into a histogram. Prepare a histogram for this data using a left endpoint convention.

Class Interval	Frequency
0-1	2
1-3	4
3-4	10
4-7	6
7-12	5

In which interval does the median value fall?

- 2) Find the standard deviation of the list: 2,4,8,16,32,64
- 3) If a normal population has an average of 50 and a standard deviation of 10, then what percentage of the population falls between 35 and 60? 73 is at the \_\_\_\_ percentile?
- 4) Find the new average, range, standard deviation, and median for a linear transformation of (new value) =  $4(\text{old value}) + 2$  if the old average = 6, the old range = 14, the old SD = 2, and the old median = 7.
- Also, find the average and SD of the list 1,3,5,10,11 and the average and SD of the transformation (new value) =  $2(\text{old value}) + 4$ .
- 5) If a normal population has an average of 30 and a standard deviation of 6, what percentage of the population is either less than 23 or greater than 35.
- 6) Find the equation of the line which passes through the point (3, 4) and has a slope 2.
- 7) Five students took a quiz. For each individual, the total time spent on the quiz and the number of questions answered correctly were recorded:

	Average number of minutes (x)	Total number of questions (y)
Amy	22	16
Bob	24	10
Carol	25	12
Donna	26	4
Eric	28	8

Find the correlation coefficient,  $r$ , and the equation of the SD line.

8) You are given 3 sets of x values and corresponding y values and you note that the last y value in each set is unknown. Determine whether or not you are given enough data to find the correlation coefficient for each set.

(a)

x	y
2	2
4	3
8	5
14	8
16	?

No additional information is given for (a).

(b)

x	y
1	2
4	11
5	14
7	19
10	?

For (b), average  $y = 15$ .

(c)

x	y
2	4
4	?
8	1
10	0
16	?

For (c), average  $y = 5$

9) You are given the information that the correlation coefficients of set (1) and (2) are  $-0.8$  and  $0.87$  respectively. Find the correlation coefficients of sets (a) to (d).

\*\*\*Hint: no computation is necessary.

(1)  $r = -0.8$

x	y
2	32
4	16
5	20
6	24
8	8

(2)  $r = 0.87$

x	y
0	20
3	24
2	26
6	28
4	32

Questions:

(a)

x	y
2	8
4	4
5	5
6	6
8	2

(b)

x	y
0	0
3	4
2	6
6	8
4	12

(c)

x	y
16	2
8	4
10	5
12	6
4	8

(d)

x	y
12	20
15	24
14	26
18	28
16	32

10) Determine whether each of the following statements is true or false.

- (a) If the correlation between salary and number of years of service is 0.8, then the correlation between number of years of service and salary is also 0.8
- (b) In a certain group, it is known that each man marries a woman who is exactly 5% shorter than himself. The correlation coefficient between heights of husbands and wives is negative.
- (c) To find the slope of the SD line, you should divide the SD of y by the SD of x and then multiply it by the correlation coefficient r.

11) The heights and weights of 200 men were recorded. The following summary statistics resulted:

Height            Average = 69"            SD = 3"

Weight            Average = 160 lb            SD = 20 lb

$r = 0.58$

Estimate the average weight of

- (i) A man having height 76"  
(ii) A group of men having height 76"  
(iii) A man whose height is unknown

12) The 50 students in a typing class were each given a 4 page article to type for a test. The correlation coefficient between the time taken to type the article and the number of typing mistakes was (-0.56). Assume both time taken and number of mistakes to be normally distributed.

If a student's time was at the 46<sup>th</sup> percentile, the estimate for the mistake percentile would be \_\_\_\_.

- (a) 46 (b) 55 (c) 52 (d) 49

If the time taken was unknown, the estimate for the mistake percentile would be \_\_\_\_.

- (a) 56 (b) 50 (c) 46 (d) can't say

13) In a certain High School 1200 students took the SAT. The following summary statistics resulted:

Math-SAT Ave= 530 SD=85

Verbal-SAT Ave=550 SD=80

$r = 0.65$

Given that a student's score on the Math-SAT was at the 75<sup>th</sup> percentile, estimate the student's **score** (not percentile) on the Verbal-SAT. You may assume that both Math-SAT and Verbal-SAT scores are normally distributed.

14) 250 students, each from a different household, participated in a survey. The size of the student's household (number of people) and rent per student in the household were recorded.

Household size Ave = 4 SD = 1

Rent per student Ave = \$500 SD = \$100  $r = -0.8$

- (a) Predict the per student rent for students in a household of size 2.  
(b) Estimate the size of the household for students who pay \$660 per month.

15) For 700 union workers, a record was kept of the number of years of education completed and annual income. The following summary statistics were computed:

	Average	SD	
Education completed	14	1	
(in no. of years)			$r = 0.72$
Annual Income	40	7	
(in thousands of dollars)			

- (i) Find the equation of the regression line used to predict annual income from education completed.  
(ii) For the group of workers whose education completed is at the 23<sup>rd</sup> percentile, estimate the percentile for annual income. (Assume that years of education completed and annual income are normally distributed).