

Statistics 2

Problems from past midterms: midterm 1

- (10 points) The men enrolled in a large physiology course had an average weight of 160 pounds and an SD of 25 pounds. Their weights followed the normal curve closely.
 - About what percent of the men weighed between 180 and 200 pounds?
 - Out of just those men—that is, the ones who weighed between 180 and 200 pounds—about what percent weighed between 180 and 190 pounds?
- (10 points) A distribution table for family income is shown below. The class intervals include the left end-point but not the right. (The table is a much simplified version of the 1991 income distribution of US families receiving a social security cheque.)

Income (in thousands of dollars)	Percentage (of families)
0 – 10	20
10 – 30	45
30 – 50	20
50 – 100	15

- Plot the histogram. Mark the horizontal and vertical scales carefully. Label the axes.
 - True or False, and explain: “The median is \$20,000.”
- (5 points) There are 300 students in a large physical education class. The instructor gives everyone in the class a reaction time test and a physical dexterity test. The dexterity test is scored out of 10 points. The averages for the class are as follows:

average reaction time: 382 milliseconds
average dexterity score: 7.4 points

The instructor compares two methods for predicting the dexterity score of someone in the class:

- Method 1: Ignore any information about the person. Just use the average 7.4 points as the prediction.
- Method 2: Take the reaction time of the person, and use the regression method to predict his or her dexterity test score.

The instructor works out the r.m.s. error for each method and finds that the two r.m.s. errors are about equal to one another.

The correlation coefficient between reaction time and dexterity score (for the students in the course) must be around:

–0.1 –0.5 –0.9 can't tell without more information

Please circle your choice and then explain it.

4. (5 points) The height and weight of a group of five men are shown in the table below:

Person	Height	Weight
A	67	158
B	69	162
C	70	160
D	71	154
E	73	166

Height is measured in inches and weight in pounds. One of these five men is going to be picked somehow or other and a statistician must estimate the weight of that man—but without being told the height. The statistician knows the numbers in the table.

- (a) Describe a good method for the statistician to use.
(b) What would be the r.m.s. error for the method you recommended in part (a).

5. (5 points) From Dear Abby, San Francisco Chronicle, October 4, 1990:

“Dear Abby. In a recent column, a writer stated that it was foolish to work for nothing as a hospital volunteer. Thanks for saying, “ The rewards are far more valuable than money.”

In 1988, a study was done by the University of Michigan Survey Research Center. They followed 2,700 people in Tecumseh, Michigan, for 10 years to determine the impact of social relationships on health. They found that people who regularly did volunteer work had dramatically longer life expectancy.

This was especially significant for men: Men who did no volunteer work were 2.5 times more likely to die during the course of the 10-year study than those who volunteered at least once a week.

Research at Yale, the University of California, John Hopkins, the National Institute of Health and Ohio State supports these findings.

Longtime Volunteer,
Mechanicsville, VA.

Dear Longtime. People who spend their time doing for others feel useful, productive and good about themselves. Volunteers, particularly those who work in hospitals, hospices and nursing homes, are too busy to dwell on their own troubles or feel depressed. Those who give-get!”

Does the evidence show that volunteer work increases life expectancy? Answer yes or no, and explain briefly.

6. (10 points) (a) What is the correlation coefficient for the data set below?

x	y
1	1
4	8
6	10
6	10
6	14
7	17

- (b) If possible, fill in the five slots below so that the correlation will be equal to the correlation for data given in part (a). If this is not possible, explain why not.

x	y
1	5
4	
6	
6	
6	
7	

7. (10 points) Pearson and Lee obtained the following results for about 1,000 families:

Father's Height: Average = 68 inches, SD = 2.7 inches
 Son's Height: Average = 69 inches, SD = 2.7 inches
 Correlation = 0.50

The scattered diagram is football shaped.

- (a) Estimate the average height of the sons whose fathers are 66 inches tall.
 (b) One father is picked at random from the 1,000 fathers in the study. He turns out to be 66 inches tall. Predict the height of his son.
 (c) Find the r.m.s. error of the regression method for predicting son's height from father's height.
8. (5 points) For this question, pretend that you made no mistake in doing part (a) and part (b) of question 7.

Check one option below.

- _____ The answer in (a) is likely to be more accurate than the answer in (b).
 _____ The answer in (a) is likely to be less accurate than the answer in (b).
 _____ The answer in (a) and answer in (b) are equally accurate.

Please explain your choice.

9. (10 points) In a large course, the correlation between midterm score and final score was 0.6. The average score on the final was 58 points and the SD was 20 points. The scatter diagram was football shaped.

Estimate the average score on the final of those who were in the 40th percentile on the midterm.

10. (5 points) The San Francisco Chronicle of Friday, May 8, 1992, reported one study of the effects of vitamin C on heart disease. The lead investigator was James E. Enstrom, an epidemiologist at UCLA School of Public Health. Here are some quotations from the Chronicle article:

“A 10- year health study has found evidence linking vitamin C to decreased mortality rates-especially from heart disease-among American men, according to a team of researchers at the University of California at Los Angeles.”

“Debate has clouded the question of the benefits of vitamin C for many years, and Linus Pauling, the Nobel laureate chemist, persists in claiming that it can prevent and fight both cancer and the common cold.”

“The report published yesterday makes no such claim. But after carefully examining the nutrition health record of more than 11,000 American adults who were surveyed between 1971 and 1974 and followed up for 10 additional years, the researchers found strong arguments in support of the vitamin’s benefits ... Entstrom was able to divide the 11,000 adults into those taking less than 50 milligrams of vitamin C daily and those whose intake was far higher-about 300-400 milligrams a day, half of it from food and half from pills.”

“Men in the second group showed a 42 percent lower overall death rate during the 10 years and a 45 percent lower death rate from heart disease than men who took lesser amounts of the vitamin. That added up to an increased life expectancy of up to six years, according to Enstrom.”

Something happened in the Clofibrate trial which is relevant to Enstrom’s finding. What is it and how is it relevant?

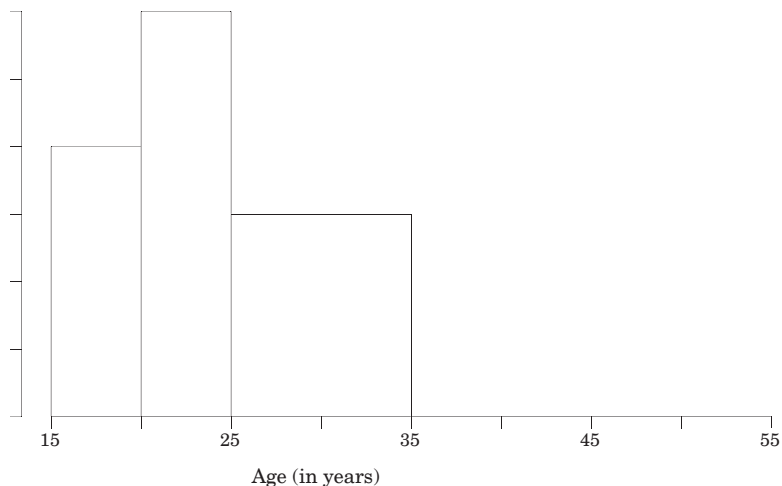
11. (10 points) The summary statistics for height in the HANES study are shown below:

Men: Average = 69.0 inches, SD = 3.0 inches
Women: Average = 63.5 inches, SD = 2.5 inches

There are about 6,000 men and 6,600 women in the study.

- (a) About what percent of the men heights that are between the two averages?
(b) About what percent of the women have heights that are between the two averages?

12. (10 points) (a) At a large city college system, the 80th percentile of age of the students is 35 years and the 20th percentile is 20 years. No one is under 15 or over 55. The histogram is shown below, but the block over the class interval from 35 to 55 is missing. Draw in the missing block. State your reasoning.



- (b) What is the median age of the students in the system?

13. (10 points) (a) What is the correlation coefficient for the data set below?

x	y
1	20
3	17
3	20
4	3
4	19
5	21
8	40

- (b) If possible, fill in the four blanks below so that the correlation coefficient is equal to zero. If this is not possible, explain why not.

x	y
1	3
1	_____
2	_____
2	_____
3	_____
3	4

14. (5 points) An aerobic study involves 1,040 men, age 18 to 24. The histogram of systolic blood pressure for these men followed the normal curve closely. The average is 120 mm, and 600 of the men have blood pressures between 110 and 130 mm. How many men have blood pressure between 115 and 125 mm?

15. (5 points) As part of a longitudinal study of personality change, 2500 normal subjects were given a psychological test twice; once at age 21, and the second time at age 35. At both ages the average score was 50, and the SD was 10. There was a strong positive association between the two scores.

For each of the 2500 subjects, the investigators used the regression method to predict score at age 35 from score at age 21. The predicted score turned out to be within 10 points of the actual score for _____ of the 2500 subjects.

Check the best option below to fill in the blank. Explain your choice.

- _____ more than 68%
- _____ around 68%
- _____ less than 68%

16. (10 points) A large study of human growth obtained the following results for males:

Height at age 4: Average = 41 inches, SD = 1.5 inches
Height at age 18: Average = 70 inches, SD = 2.5 inches
Correlation = 0.75

The scatter diagram was football shaped.

- (a) Predict the height at eighteen for someone in the study who was 38 inches tall at age four.
- (b) Of those in the study who were 38 inches tall at age four, what percent had added 30 inches or more to their height by the time they were eighteen?
