

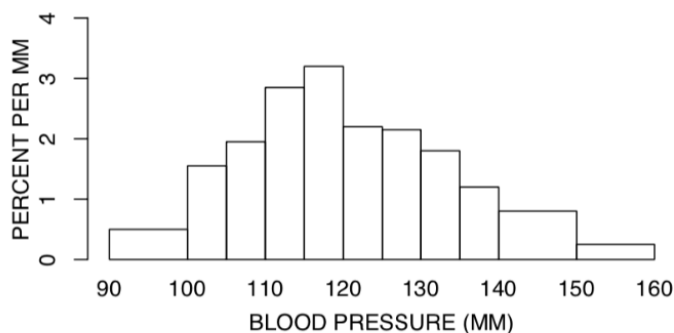
# Quiz 1 Solutions

Stats 60/160

July 23, 2020

## 1 Problem 1

The figure below is a histogram showing the distribution of blood pressure for all 14,148 women in the Drug Study (section 5). Use the histogram to answer the following questions.



- (a) Is the percentage of women with blood pressures above 130 mm around 25%, 50%, or 75%?
- (b) Is the percentage of women with blood pressures between 90 mm and 160 mm around 1%, 50%, or 99%.
- (c) In which interval are there more women: 135-140 mm or 140-150 mm?
- (d) Which interval is more crowded: 135-140 mm or 140-150 mm?
- (e) On the interval 125-130 mm, the height of the histogram is about 2.1% per mm. What percentage of the women had blood pressures in this class interval?
- (f) Which interval has more women: 97-98 mm or 102-103 mm?
- (g) Which is the most crowded millimeter of call?

(a) About 25%. The area of the histogram to the right of 130 mm is certainly less than half of whole histogram, and closer to 25%.

(b) About 99%. The area under a whole histogram is 100%, thus the area is closest to 99%.

(c) There's more women in the range of 140-150 mm since it has a larger area.

(d) The histogram is more crowded because it is taller in the interval 135-140 mm.

(e) We need to compute the area of this range, which is about  $5 \times 2.1\% = 10.5\%$ .

(f) 102-103 mm. Since the length of the two intervals are the same, comparison of the heights of the bars gives the answer.

(g) Somewhere between 115 and 120 mm, where the histogram peaks. We do not have information to tell exactly where.

## 2 Problem 2

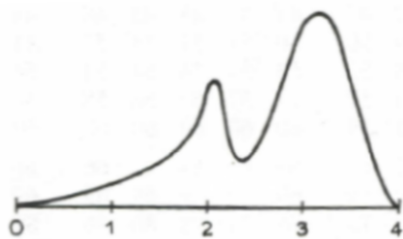
Two histograms sketched below. One shows the distribution of age at death from natural causes (heart disease, cancer, and so forth). The other shows age at death from trauma (accident, murder, suicide). Which is which, and why?



**Answer** The first is death from natural causes. Because the rate of death from natural causes typically increase when the age increases, but younger people are more likely to die from accidents, murders or suicide.

## 3 Problem 3

In a survey carried out at the University of California, Berkeley, a sample of students were interviewed and asked what their grade-point average was. A histogram of the results is shown below (GPA ranges from 0 to 4, and 2 is a bare pass.)



- (a) True or false: more students reported a GPA in the range 2.0 to 2.1 than in the range 1.5 to 1.6.
- (b) True or false: more students reported a GPA in the range 2.0 to 2.1 than in the range 2.5 to 2.6.
- (c) What accounts for the spike at 2?

(a) True. We need to compare area of the two intervals, and since the length are the same, the area depends on the height. The height at 2.05 is higher than the height at 1.55.

(b) True. Same reason as part (a).

(c) Perhaps it is the passing grade, and some students perform just enough to pass. Other explanations are also okay.

## 4 Problem 4

An investigator has a computer file showing family incomes for 1,000 subjects in a certain study. These range from \$5,800 a year to \$98,600 a year. By accident, the highest income in the file gets changed to \$986,000.

- (a) Does this affect the average? If so, by how much?
- (b) Does this affect the median? If so, by how much?

- (a) Yes. It moves the sum up by  $986000 - 98600$  and the thus average by

$$(986000 - 98600)/1000 = 887.4.$$

- (b) No, this change does not affect the median.

## 5 Problem 5

True or False, and explain briefly:

- (a) If you add 7 to each entry on a list, that adds 7 to the average.
- (b) If you add 7 to each entry on a list, that adds 7 to the SD.
- (c) If you double each entry on a list, that doubles the average.
- (d) If you double each entry on a list, that doubles the SD.
- (e) If you change the sign of each entry on a list, that changes the sign of the average.
- (f) If you change the sign of each entry on a list, that changes the sign of the SD.

You can verify these with some made-up numbers. Note you need to explain briefly to get full credit.

- (a) True. Adding 7 to each element increases the sum by  $7 \times N$  if  $N$  is the total number of elements, and thus the average by  $7 \times N/N = 7$ .
- (b) False. The SD depends on distances from the mean, and since the average moves up by 7, the deviations remain the same.
- (c) True. Because every number doubles, the sum also doubles, and thus the average doubles.
- (d) True. Because the deviations from the average doubles, the SD doubles.
- (e) True. If every number changes sign, the sum also changes sign (but it still has the same magnitude), therefore the average change the sign.
- (f) False. Because the distances from the average stays the same.