

PROBLEM SET 1

A. FARRIS | STATISTICS 103 | UC DAVIS
WINTER 2021

Instructions: Here is our first set of practice problems. The TAs will go over the solutions to some of these problems in detail during discussion. Some problems have a subjective difficulty rating associated to it: level 1, 2 or 3, roughly corresponding to easy, moderate and hard.

Unless told otherwise, you can take “standard deviation” to refer to the *sample* standard deviation.

Practice problems

- (1) (Level 1) Suppose $\bar{x} = 11$ and define $y_i = 2x_i - 5$. Find \bar{y} .
- (2) (Level 2) Suppose $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = 1202$ and set $y_i = 2x_i - 5$. Find $\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$.
- (3) (Level 1) If last weeks average temperature was 65° Fahrenheit, what was the average temperature in Celsius?
- (4) (Level 1) Referring to FIG. 1, if there are a total of 1109 tennis balls used at the health club, how many of those are Wilson balls?
- (5) (Level 2) Referring to FIG. 1. suppose you want to combine the circles corresponding to *Penn* and *Wilson* into one category called *Commercial*. What should the radius be for this new category?
- (6) (Level 2) The graph in FIG. 2 uses relative area to represent the gender decomposition and relative size of the population in two countries A and B.

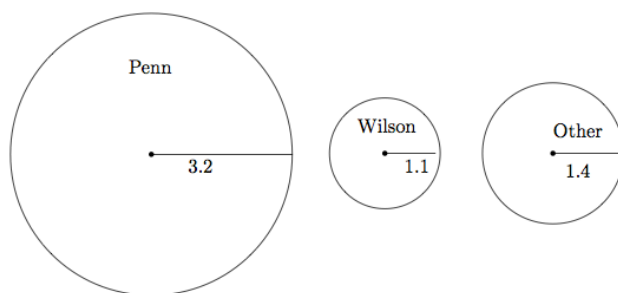


Figure 1: The following diagram uses relative area to graphically illustrate the relative frequency of the different brands of tennis balls used at a certain health club (note: the area of a circle is $\pi(\text{radius})^2$).

- (a) What is the percentage of men in country A? What is the percentage of men in country B?
- (b) Suppose country A has a population of 10 million, what is the population of country B?
- (c) If you combine the two countries, what would be the resulting percentage of women?

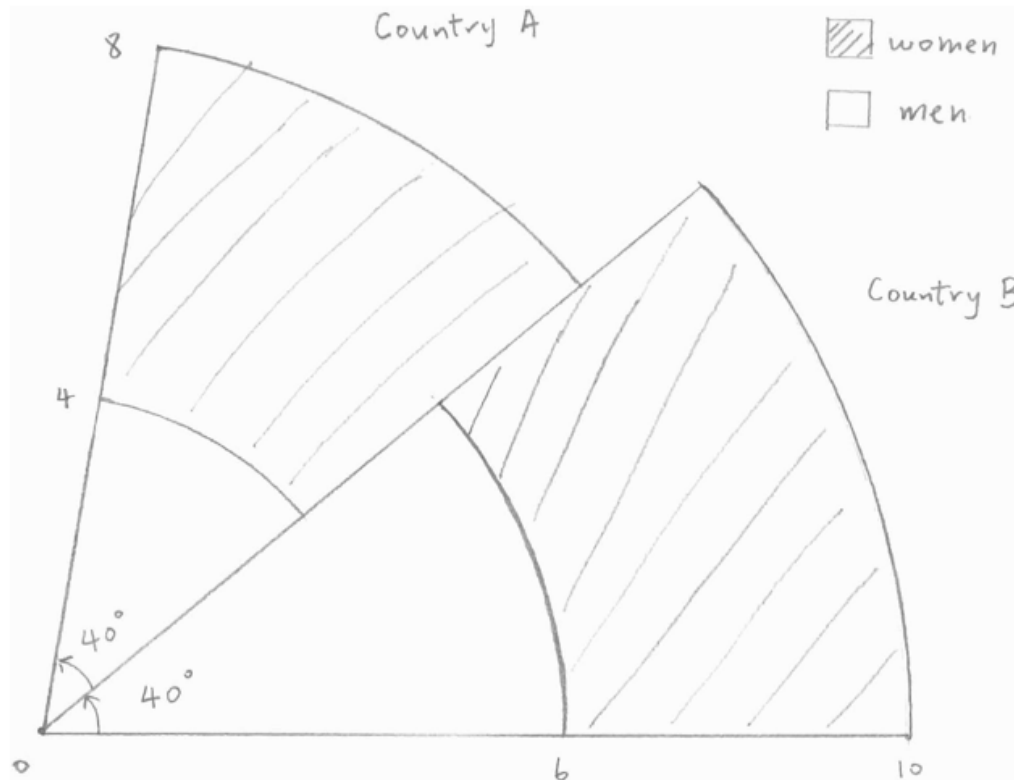


Figure 2: Using area to represent data

- (7) (Level 2) An honor list of 280 students categorizes each according to area of study: engineering, liberal arts, sciences. The diagram in FIG. 3 uses relative area to express class percentage. Compute the frequencies of each category in the list of students.
- (8) (Level 3) The diagram shown in FIG. 4 uses relative area to graphically illustrate the relative frequency of 145 drinks left in a vending machine after a week on campus. Suppose that after drawing the diagram, the vendor finds two cokes stuck in the machine. How should the radii be adjusted?

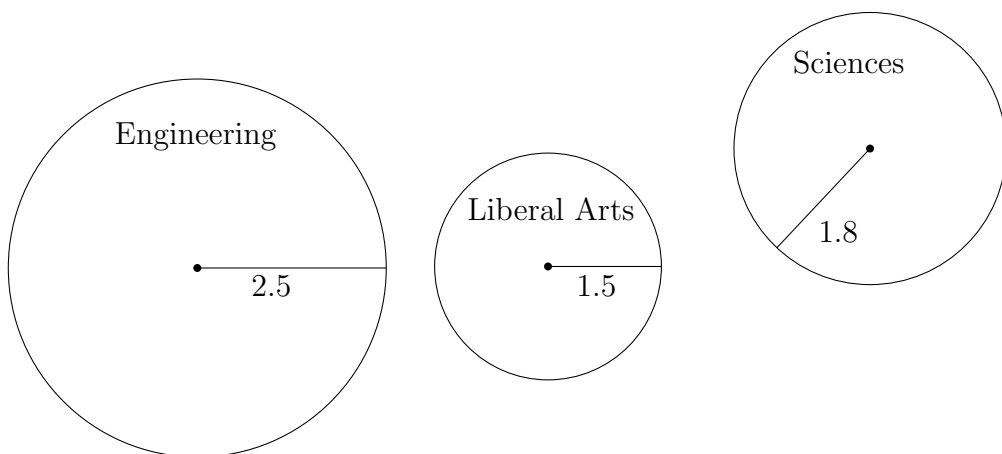


Figure 3: Relative area represents relative frequency.

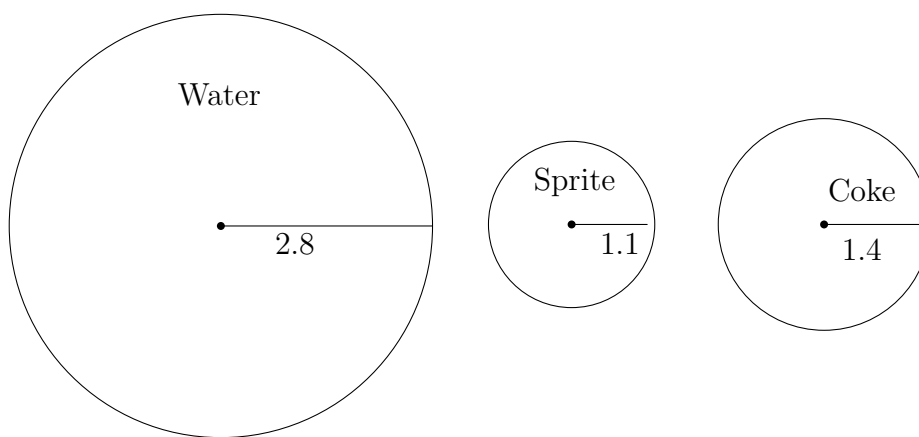


Figure 4: Relative area represents relative frequency.

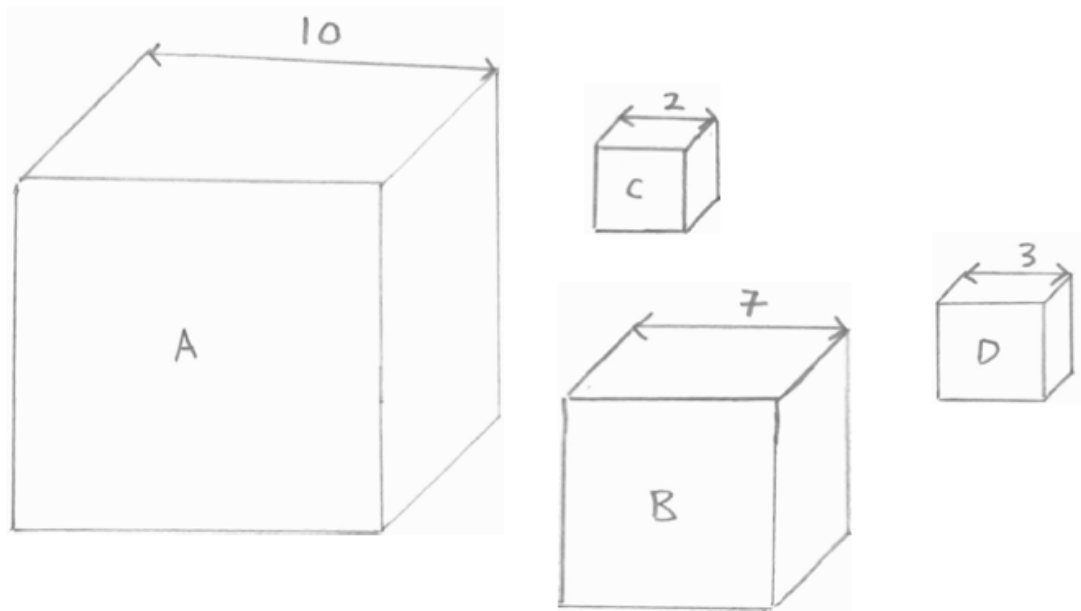


Figure 5: Using volume to represent data.

- (9) (Level 2) Referring to the same diagram referred to in question 8, suppose you want to combine the circles corresponding to *coke* and *sprite* into one category: *soda*. What should the new radius be?
- (10) (Level 2) Suppose the cubes in FIG.5 represent a qualitative data set with four classes: A, B, C, D.
- Construct a relative frequency table for the four classes.
 - Suppose you want to combine C and D into one class. What is the edge length of the new cube?
- (11) (Level 2) Suppose a bank's currency reserve contains 78% ones, 10% fives, 7% tens, 4% twenties and 1% one hundred dollar bills. If you were going to graphically represent these percentages using scaled versions of the corresponding bills, what heights would you use?
- (12) The histogram in FIG. 6 shows the distribution of the number of cigarettes smoked per day by male smokers in 1971. What percentage of this group smokes more than 20 cigarettes per day? Hint: for this histogram, relative area of a block equals relative frequency of the corresponding class interval.
- (13) Referring to the histogram shown in problem 12, suppose you want to combine the middle two blocks into one block with base over the interval from 10 to 40.

What height would you make this new block so that relative area matches relative frequency of the class intervals.

- (14) In a psychological study, 150 children were asked which shape they preferred between a circle, an equilateral triangle, and a square. Their responses were reflected in FIG. 7 below. About how many children preferred either a square or a triangle? (note: the *area of an equilateral triangle* = $\frac{\sqrt{3}}{4}(\text{side})^2$ and *area of a circle* = $\pi(\text{radius})^2$)
- (15) (Level 1) Given $n = 10$, $\sum_{i=1}^{10} x_i^2 = 100$, $\bar{x} = 1$, compute the sample standard deviation s .
- (16) Continuing with the data from problem (15), suppose that x_5 is changed from 2 to 4. What is \bar{x} after this change?
- (17) Find \bar{x} for a list of numbers which is composed of 200 ones and 40 zeros.
- (18) (Level 2) Find s^2 for a list of numbers which is composed of 200 ones and 40 zeros.
- (19) Find the mean and mode of the following list of numbers:

2, 3, 4, 5, 2, 1, 4, 2, 6, 3, 6, 7, 3, 2, 2, 5.

- (20) If the mean of a data set is 1 and we add 3 to each data observation, then multiply by 2, what is the mean of the new data?
- (21) (Level 1) Top speeds for a sample of five new automobiles are listed below:

200, 170, 125, 175, 180.

Calculate the sample standard deviation of the above list.

- (22) (Level 2) Given a sample of size $n = 10$ with $\sum_{i=1}^{10} x_i^2 = 100$ and $\sum_{i=1}^{10} x_i = 5$, if one more data point 9 is added to this sample, compute the new sample standard deviation s .

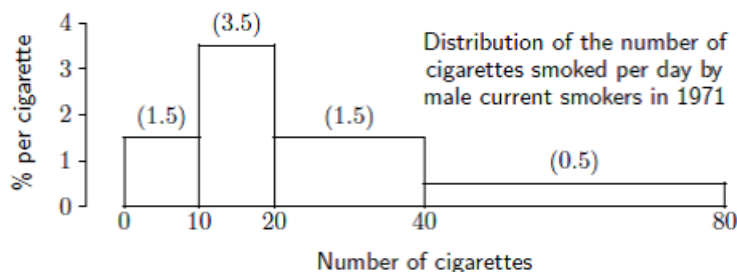


Figure 6: Smokers

- (23) (Level 1) The following table lists the frequency and relative frequency for grades on an exam. Find the missing relative frequency.

Grades on Test	Frequency	Relative Frequency
A	6	0.24
B	?	?
C	9	0.36
D	2	0.08
F	1	0.04

- (24) FIG. 8 shows a histogram for a set of numbers. What proportion of the numbers in the list are smaller than 4.5?

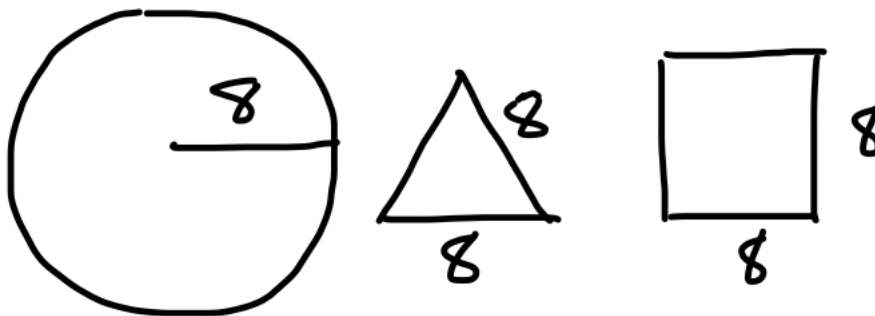


Figure 7: Shapes

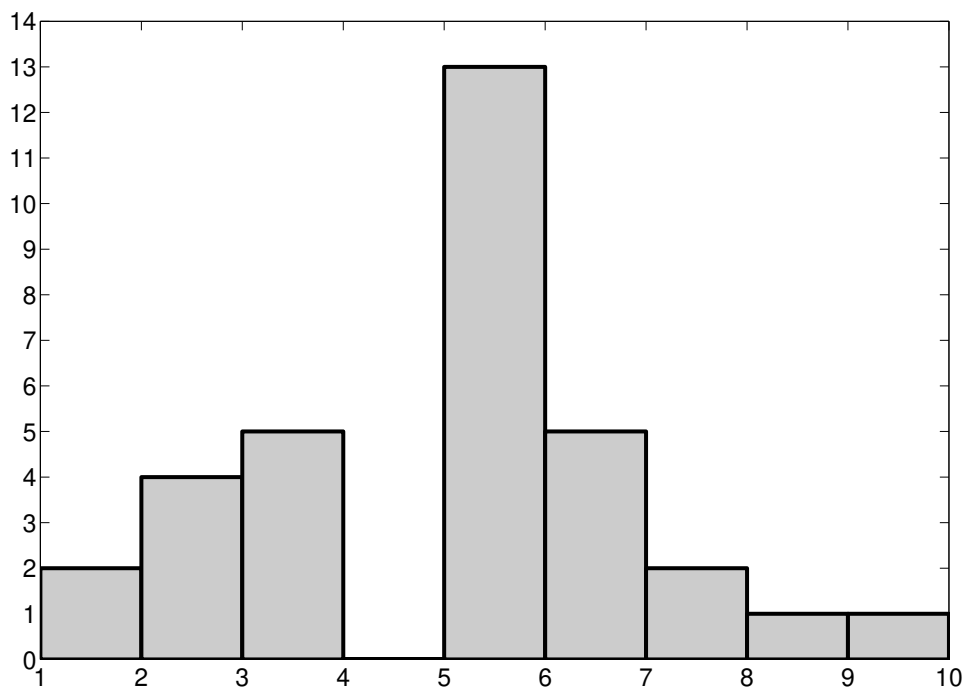


Figure 8: Histogram