Lab 1b: Histograms

Stat 131A, Fall 2018

Learning Objectives:

- Distinguish between categorical and quantitative variables;
- Analyze the distribution of a quantitative variable using a histogram.
- Describe shape, give a general estimate of center and the overall range.

Problem 1

Identify the type of variable (qualitative or quantitative) for the list of questions from a survey applied to college students in a statistics class.

    a. Name of student.

    b. Birth date (e.g. 10/21/1995).

    c. Age (in years).

    d. Home Address (e.g. 1234 Shattuck Ave).

    e. Telephone number (e.g. 510-123-4567).

    f. Major field of study.

    g. College year-grade: freshman, sophomore, junior, senior.

    h. Score on midterm test (based on 100 possible points).

    i. Overall grade: A, B, C, D, F.

    j. Length of time—in minutes—to complete Stat 131A final test.

    k. Birth date: (e.g. 10/21/1995)

    l. Number of siblings.

Problem 2

Consider a variable with numeric values describing electronic ways of expressing personal opinions: 1 = Twitter; 2 = email; 3 = text message; 4 = Facebook; 5 = blog. Is this a quantitative or a qualitative variable?

Problem 3

For each research question, (1) identify the individuals of interest (the group or groups being studied), (2) identify the variable(s) (the characteristic that we would gather data about), and (3) determine whether each variable is categorical or quantitative.

    a. What is the average number of hours that community college students work each week?

    b. What proportion of all U.S. college students are enrolled at a community college?

    c. In community colleges, do female students have a higher GPA than male students?

    d. Are college athletes more likely than non-athletes to receive academic advising?

Problem 4

If we gathered data to answer the previous research questions, which data could be analyzed using a histogram? How do you know?

Problem 5

The following figure shows four histograms (A, B, C, D).
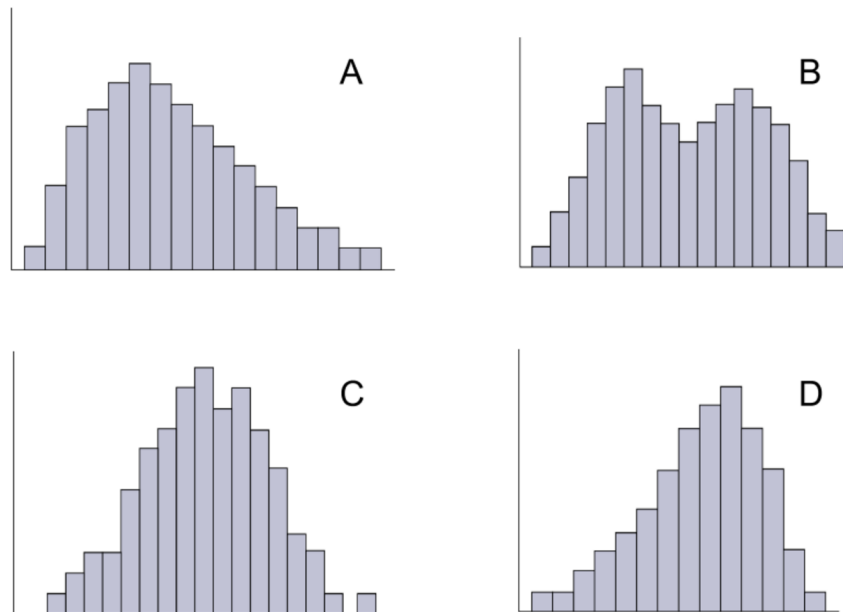
  a. Which distribution seems to be skewed right?

  b. Which distribution seems to be skewed left?

  c. Which distribution seems to be symmetrical or "bell" shaped?

  d. Which distribution seems to be bimodal?

 e. Which distribution shows a gap?

Problem 6

Match the following descriptions to the histograms 1-3.

| distribution 1 | distribution 2 | distribution 3 |
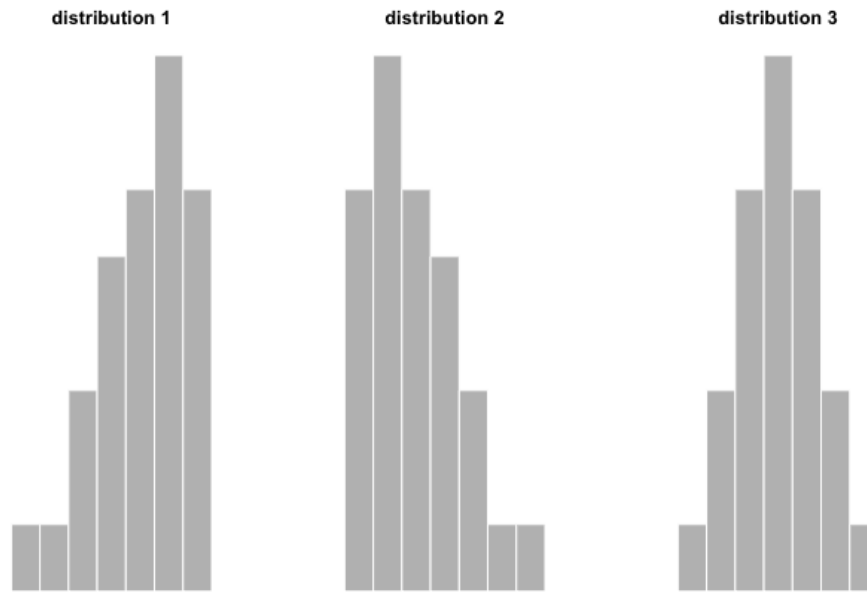
a. Scores on an easy exam for a class of students

b. Scores on a hard exam for a class of students

c. Number of siblings for a large sample of U.S. adults

d. Exact volume of soda in a one-liter bottle for a case of 24 bottles

e. Dates on the pennies I have in my car ashtray

f. Weights for a large sample of newborn babies

Problem 7

This histogram shows the distribution of exam scores for 15 students in a Biology class.

Histogram of grades
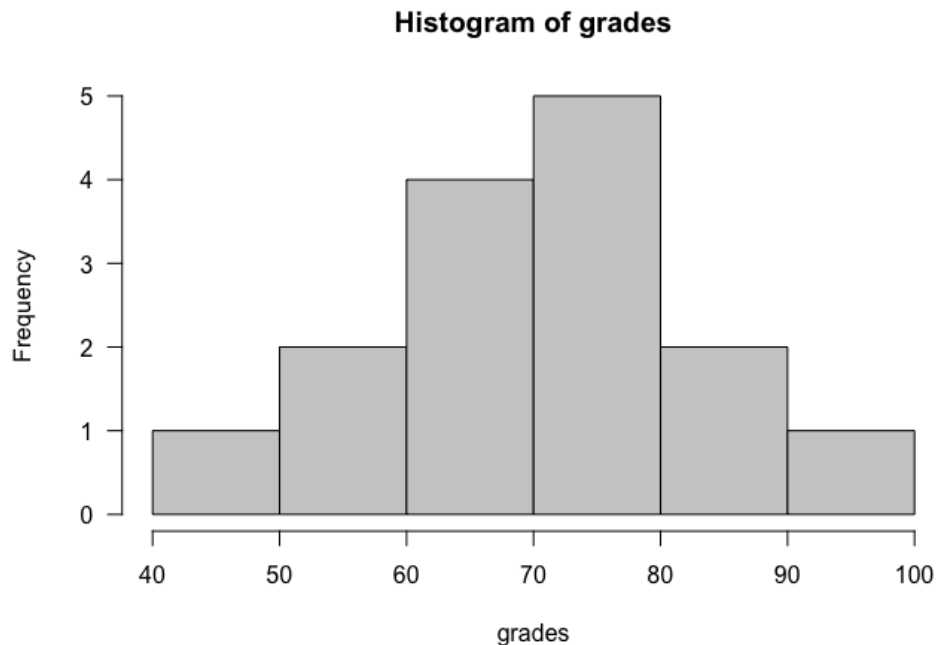
a) How would you describe the shape of this distribution of exam scores? (Use statistical vocabulary.)

b) Give an interval that describes typical grades on this exam.

c) Estimate the overall range of grades on this exam. (Range = Max – Min)

d) What percentage of the students made a D on the exam (a grade of 60-69)?

e) What percentage of the students passed the exam with a 70 or better?

f) What percentage of the students made an A ($>=90$) or a B (80-89)?

g) What percentage of the students who passed the exam made an A or a B?

h) What percentage of the students who failed the exam (grades lower than 70) made a D (a grade of 60-69%)?

Problem 8

True or False

a. A histogram allows you to visualize how values are distributed.

b. Bars in a histogram can be rearranged.

c. The area of the bars in a histogram represent the frequency of values in a class interval.

d. Class intervals must be of equal width in a histogram.

e. The height of the bars in a histogram must be equal to the frequency of the corresponding class intervals.

f. The total area of the rectangles in a histogram must be equal to 100%.

Problem 9

Below are sketches of histograms for test scores in three different classes. The scores range from 0 to 100; a passing score was 50.
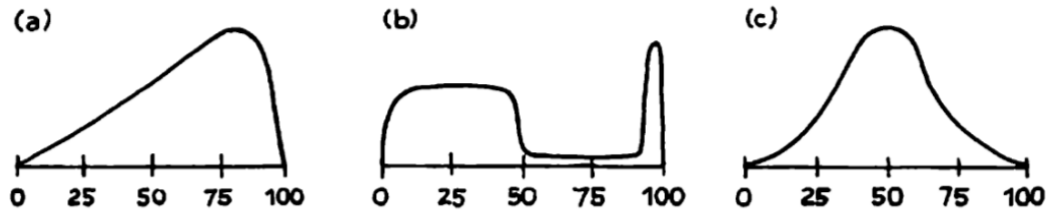
a. For each class, was the percentage who passed about 50%, well over 50%, or well under 50%?

b. One class had two quite distinct groups of students, with one group doing rather poorly on the test, and the other group doing very well. Which class was it?

c. In class (b), Were there more people with scores in the range 40-50 or 90-100?

Problem 10

In one hypothetical study, 100 people had their heights measured to the nearest eighth of an inch. A histogram for the results is shown below. Two of the following lists have this histogram. Which ones, and why?
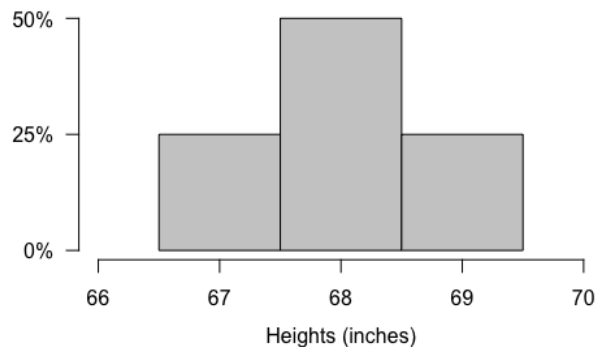
a. 30 people, 67 inches tall; 40 people, 68 inches tall; 30 people, 69 inches tall.

b. 10 people, $66\frac{3}{4}$ inches tall; 15 people, $67\frac{1}{4}$ inches tall; 50 people, 68 inches tall; 25 people, 69 inches tall.

c. 25 people, 67 inches tall; 50 people, 68 inches tall; 25 people, 69 inches tall.