

STAT 201: Elementary Statistics Session 13 & 14

Exam 1

Instructor: Haigang Liu

1. (16 points) Suppose you are interested in the average age of women at marriage in China, in Ming Dynasty (1368 A.D.–1644 A.D.). Within the archives, you find out marriage records between 1500 A.D. – 1600 A.D, which can be treated as a sample of all marriage records in Ming Dynasty.

In the records, the average age of women when they got married is 18.1 years. Using appropriate statistical method, you conclude that the average age of brides in Ming Dynasty is between 17.5 and 18.7. Answer the following questions.

- a) Which part of the example gives a descriptive summary of the data?
- b) Which part of the example draws an inference about the population?
- c) What is the population in your study?
- d) The average marriage age of the sample is 18.1. Is this quantity a statistic or a parameter?

2. (16 points) Identify each of the following variables by labeling them as quantitative or categorical. If you identify one variable as quantitative, also write down whether it is continuous or discrete.

- a) Number of children in family
- b) Amount of time in a soccer game
- c) Type of music (rock, jazz, classical, folk, other)
- d) Postal code in South Carolina (29201, 29202, . . . , 29210)

3. (14 points) Ancient Greek was a prosperous civilization, with a number of city-states, including Athens and Spartan. Table 1 shows the population of city-states (in ten thousand). Read the table and answer the following questions.

Narona	4.7
Nymphaion	3.3
Issa	2.6
Apollonia	2.1
Athens	12.9
Epetium	3.7
Olokenion	2.8
Spartan	5.2
Tragurion	4.4

Table 1: The population of city-states

- a) Find the mean, median and mode for these data.
- b) Which one do you think is a sensible numerical summary, mean, median or mode? Explain.

4. (20 points) Suppose we have 5 test scores from STAT 201, i.e.,

82, 84, 79, 84, 81.

a) Find the variance and standard deviation for the 5 test scores. Note that $\sum_{i=1}^5 x_i = 410$ and $\sum_{i=1}^5 (x_i - \bar{x})^2 = 18$, where \bar{x} is mean of x . You can use the formula as follows.

$$\text{Var}(X) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

b) Suppose we have another student who have received 60 points in the test. Compute the z score for that student. Is his score a potential outlier?

- c) According to empirical rule, find the interval that contains approximately 68% of all students.
- d) What assumption is required when applying the empirical rule?

5. (20 points) Researchers in a study gathered 1000 observations and plotted them in a histogram, which is displayed in Figure 1. Answer the following questions.

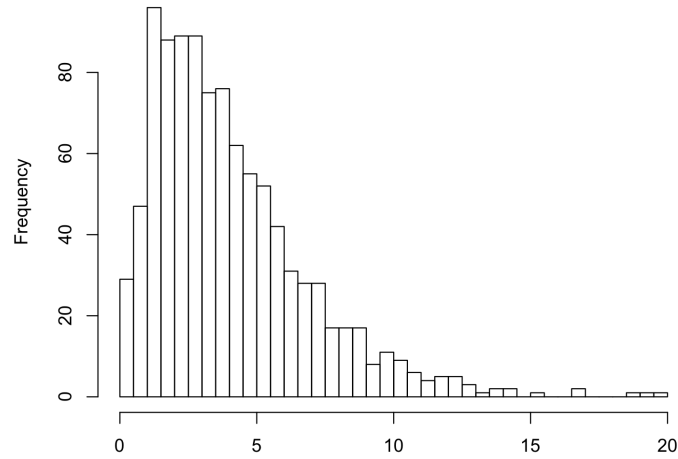


Figure 1: Histogram (Question 5)

- a) What is the difference between a histogram and a bar graph?
- b) Is the distribution in Figure 1 left skewed or right skewed?
- c) For a right-skewed distribution, which quantity is larger, mean or median? What about a left-skewed distribution?
- d) Can you come up with a real world example with a distribution skewed to right?

6. (14 points) Researchers sampled 11 observations from an unknown distribution,

55, 52, 55, 53, 49, 47, 53, 51, 57, 56, 49.

- a) Figure 2 gives a box plot for the data above. Calculate the first, the second and the third quartile in the box plot.
- b) Briefly describe the shape of distribution based on the box plot.
- c) Suppose we have another observation of 90. Should we denote this observation as outlier? Justify your judgement.

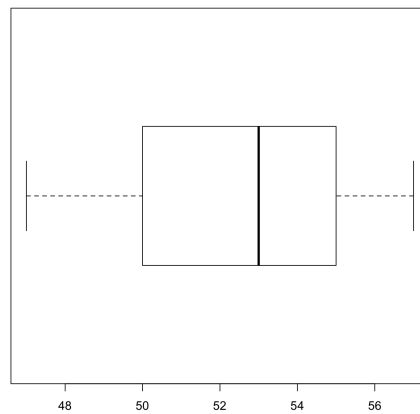


Figure 2: Box plot (Question 5)

ANSWER SHEET PAGE ONE

First Name:

Last Name:

ANSWER SHEET PAGE TWO

First Name:

Last Name:

ANSWER SHEET PAGE THREE

First Name:

Last Name:

ANSWER SHEET PAGE FOUR

First Name:

Last Name:

ANSWER SHEET PAGE FIVE

First Name:

Last Name: