Section 3.3

Definition: For whole numbers P (where $1 \leq P \leq 99$), the Pth **percentile** of a distribution is a value such that $P\%$ of the data fall at or below it and $(100 - P)\%$ of the data fall at or above it.

Definition: **Quartiles** are measures of location, denoted Q_1, Q_2, and Q_3, which divide a set of data into four groups with about 25% of the values in each group.

Definition: The **first quartile** separates the lower 25% of the ordered data list from the top 75%. It is found by locating the median of the lower half of the data list.

Definition: The **third quartile** separates the lower 75% of the ordered observations from the top 25%. It is found by locating the median of the upper half of the data list.

Exercise 1. Find the 1st quartile, median, and 3rd quartile for the following values:

10, 14, 15, 19, 20, 20, 24, 28, 29, 36, 38, 38, 44, 45, 46, 50, 51, 68, 70.

.

Class Exercise 1. Find the 1st, 2nd, and 3rd quartiles for the following values:

59, 77, 18, 10, 87, 99, 54, 46, 59, 20, 66, 55, 46, 24, 7, 84.

Definition: The **interquartile range** is the difference between the third quartile and the first quartile.

Example: In our example, the interquartile range $= 46 - 20 = 26$.

Remark: Like the median, the interquartile range is a better measure for data with extremely large or small values.

Class Exercise 2. You've just bought a new car that claims to get a highway fuel efficiency of 31 miles per gallon. Of course, your mileage will "vary". If you had to guess, would you expect the IQR of gas mileage attained by all cars like yours to be 30 mpg, 3 mpg, or 0.3 mpg? Why?

Definition: For a set of data, the **5-number summary** consists of the minimum value, the first quartile Q_1, the median (or second quartile Q_2), the third quartile Q_3, and the maximum value.

Boxplots

A **boxplot** (or **box-and-whisker diagram**) is a graph of a data set that consists of a line extending from the minimum value to the maximum value, and a box with lines drawn at the first quartile Q_1, the median,and the third quartile Q_3.

Exercise 2. For the previous set of data, find the 5-number summary and construct a boxplot:

Class Exercise 3. Here are the heights of a group of 40 American women (in inches):

52, 55, 56, 54, 50, 56, 53, 54, 50, 55, 53, 56, 51, 51, 53, 58, 58, 54, 52, 60, 53, 49, 53, 55, 54, 52, 56, 52, 56, 50, 57, 57, 48, 50, 51, 52, 53, 57, 53, 51.

Here are the heights of a group of 40 American men (in inches):

61, 61, 59, 54, 67, 57, 63, 56, 63, 55, 61, 60, 60, 56, 63, 61, 64, 55, 60, 62, 64, 57, 62, 60, 67, 59, 61, 56, 62, 66, 62, 67, 55, 59, 57, 63, 62, 60, 59, 61.

Construct boxplots for the heights of men and women on the same axis. (Hint: use a stem-and-leaf plot to sort the data.) From the boxplots, what can you conclude about the relationship between the heights of men and women?

Remark: Boxplots are great for comparing different sets of data. However, a boxplot does not portray certain features of a distribution, such as shape, as clearly as does a histogram.

Student Feedback

My teaching methods are (I hope) continually subject to improvement. If you have any comments, suggestions, or ideas, please email them to me at Sithparran.Vanniasegaram@evc.edu.

Homework

C problems

Section 3.3: 5, 7, 9

B problems

Section 3.3: 1, 11

A problems

Section 3.3: 3