

Measures of Position

Section 3.4

- Identify the position of a data value in a data set, using various measures of position such as percentiles, deciles, and quartiles

Objectives

- Are used to locate the relative position of a data value in a data set
- Can be used to compare data values from different data sets
- Can be used to compare data values within the same data set
- Can be used to help determine outliers within a data set
- Includes z-(standard) score, percentiles, quartiles, and deciles

Measures of Position (or Location or Relative Standing)

- Can be used to compare data values from different data sets by “converting” raw data to a standardized scale
- Calculation involves the mean and standard deviation of the data set
- Represents the number of standard deviations that a data value is from the mean for a specific distribution
- We will use z-scores extensively in Chapter 6

Z score or Standard Score

- Is obtained by subtracting the mean from the given data value and dividing the result by the standard deviation.
- Symbol of BOTH population and sample is **z**
- Can be positive, negative or zero
- A data point can be considered unusual if its z-score is sufficiently large or small

- Formula

- Population

$$z = \frac{x - \mu}{\sigma}$$

- Sample

$$z = \frac{x - \bar{x}}{s}$$

Z-Score

- Are position measures used in educational and health-related fields to indicate the position of an individual in a group
- Divides the data set in 100 ("per cent") equal groups
- Used to compare an individual data value with the national "norm"
- Symbolized by P_1, P_2, \dots
- Percentile rank indicates the percentage of data values that fall *below* the specified rank

Percentiles

Percentile Rank =

$$\left(\frac{\text{(number of data values below the given data point)} + 0.5}{\text{total number of values}} \right) (100\%)$$

To find the percentile rank for a given data value, x

American College Test (ACT) Scores attained by 25 members of a local high school graduating class (Data is ranked)

14	16	17	17	17
18	19	19	19	19
20	20	20	21	21
21	23	23	24	25
25	25	28	28	31

- 1) Thad scored 22 on the ACT. What is his percentile rank?
- 2) Ansley scored 20 on the ACT. What is her percentile rank?

Examples

- Step 1: Arrange data in order from lowest to highest
- Step 2: Substitute into the formula

$$c = \frac{n \cdot p}{100}$$

where n is total number of values and p is given percentile

- Step 3: Consider result from Step 2
 - If c is NOT a whole number, round up to the next whole number. Starting at the lowest value, count over to the number that corresponds to the rounded up value
 - If c is a whole number, use the value halfway between the c^{th} and $(c+1)^{\text{st}}$ value when counting up from the lowest value

Finding a Data Value Corresponding to a Given percentile

American College Test (ACT) Scores attained by 25 members of a local high school graduating class (Data is ranked)

14	16	17	17	17
18	19	19	19	19
20	20	20	21	21
21	23	23	24	25
25	25	28	28	31

To be in the 90th percentile, what would you have to score on the ACT?

Find P_{85}

Examples

- Same concept as percentiles, except the data set is divided into four groups (quarters)
- Quartile rank indicates the percentage of data values that fall *below* the specified rank
- Symbolized by Q_1 , Q_2 , Q_3
- Equivalencies with Percentiles:
 - $Q_1 = P_{25}$
 - $Q_2 = P_{50} = \text{Median of data set}$
 - $Q_3 = P_{75}$

Quartiles

Minitab calculates these for you.

- Same concept as percentiles, except divides data set into 10 groups
- Symbolized by D_1 , D_2 , D_3 , ... D_{10}
- Equivalencies with percentiles
 - $D_1 = P_{10}$ $D_2 = P_{20}$
 - $D_5 = P_{50} = Q_2 = \text{Median of Data Set}$

Deciles

- Outlier is an extremely high or an extremely low data value when compared with the rest of the data values
- A data set should be checked for "outliers" since "outliers" can influence the measures of central tendency and variation (mean and standard deviation)

Outliers

- Interquartile Range (IRQ)
 - $Q_3 - Q_1$
- Identifying Outliers (p. 149)
 - Is the data point between

$$Q_1 - 1.5IRQ \text{ and } Q_3 + 1.5IRQ$$

Identifying Outliers

American College Test (ACT) Scores attained by 25 members of a local high school graduating class (Data is ranked)

14	16	17	17	17
18	19	19	19	19
20	20	20	21	21
21	23	23	24	25
25	25	28	28	31

- 1) Emily scored 11 on the ACT. Would her score be considered an outlier?
- 2) Danielle scored 38 on the ACT. Would her score be considered an outlier?

Examples

- Data value may have resulted from a measurement or observational error
- Data value may have resulted from a recording error
- Data value may have been obtained from a subject that is not in the defined population
- Data value might be a legitimate value that occurred by chance (although the probability is extremely small)

Why Do Outliers Occur?