Chapter 6: The Normal Distribution

6-4: Checking for Normality

Normally distributed variables

- Over the last few days, we have solved several problems involving normally distributed variables

- We were able to use the standard normal (z) distribution to solve these problems

- However, what if the problem doesn't tell you the variable is normally distributed?

Checking for Normality

- Histogram

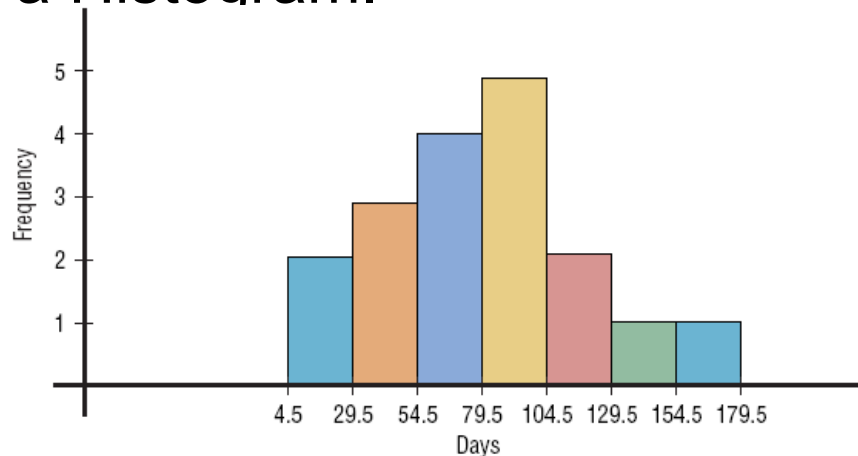- Pearson's Index PI of Skewness

- Outliers

Example 1: Technology Inventories

A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74

81 88 91 97 98 113 118 151 158

Method 1: Construct a Histogram.

Class	Frequency
5–29	2
30–54	3
55–79	4
80–104	5
105–129	2
130–154	1
155–179	1

The histogram is approximately bell-shaped.

Example 1: Technology Inventories

Method 2: Check for Skewness.

$$\bar{X} = 79.5, MD = 77.5, s = 40.5$$

$$\text{PI} = \frac{3(\bar{X} - MD)}{s} = \frac{3(79.5 - 77.5)}{40.5} = 0.148$$

The PI is not greater than 1 or less than -1, so it can be concluded that the distribution is not significantly skewed.

Method 3: Check for Outliers.

Five-Number Summary: 5 - 45 - 77.5 - 98 - 158

Q1 – 1.5(IQR) = 45 – 1.5(53) = -34.5

Q3 + 1.5(IQR) = 98 + 1.5(53) = 177.5

No data below -34.5 or above 177.5, so no outliers.

Example 1: Technology Inventories

A survey of 18 high-technology firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5 29 34 44 45 63 68 74 74

81 88 91 97 98 113 118 151 158

Conclusion:

•The histogram is approximately bell-shaped.

•The data are not significantly skewed.

•There are no outliers.

Thus, it can be concluded that the distribution is approximately normally distributed.

Example 2: Baseball Hall of Fame

The data shown consist of the number of games played each year in the career of Bill Mazeroski. Determine if the data are approximately Normally distributed.

81	148	152	135	151	152
159	142	34	162	130	162
163	143	67	112	70	

Method 1: Construct a Histogram.

The histogram shows the frequency distribution is somewhat negatively skewed.

Example 2: Baseball Hall of Fame

Method 2: Check for Skewness.

$$PI = \frac{3(\overline{X} - Median)}{s} = \frac{3(127.24 - 143)}{39.87} = -1.19$$

The PI is less than -1, so it can be concluded that the distribution is significantly skewed left.

Method 3: Check for Outliers.

Five-Number Summary: Q1=96.5, Q3=155.5, IQR=59

Q1 – 1.5(IQR) = 96.5 – 1.5(59) = 8

Q3 + 1.5(IQR) = 155.5 + 1.5(59) = 244

No data below 8 or above 244, so no outliers.

Example 2: Baseball Hall of Fame

The data shown consist of the number of games played each year in the career of Bill Mazeroski. Determine if the data are approximately Normally distributed.

81 148 152 135 151   152

159 142   34   162 130   162

163 143   67   112   70

Conclusion:

• The histogram is somewhat negatively skewed.

• The data is significantly skewed left.

• There are no outliers.

Thus, it can be concluded that the distribution is <u>NOT</u> approximately normally distributed.

Homework

p. 316: 33, 39-42