Lecture 21: Chapter 6

Section 6C: The Normal Distribution

Definition

The most commonly studied distribution is the normal distribution.

Definition 1. The **normal distribution** is a symmetric, bell-shaped distribution with a single peak. Its peak corresponds to the mean, median, and mode of the distribution.

Example 1. Most human related variables are given by normal distributions: IQ scores, height, weight, etc.

A data set that satisfies the following is **likely** to have a normal distribution:

1. Most data values are clustered near the mean, giving the distribution a well-defined single peak.

2. Data values are spread evenly around the mean, making the distribution symmetric.

3. Larger deviations from the mean become increasingly rare, producing the tapering tails of the distribution.

4. Individual data values result from a combination of many different factors,

such as genetic and environmental factors.

Standard Deviation

The 68-95-99.7 rule is as follows:

i) About 68%, of just over two-thirds, of the data points fall within 1 standard deviation of the mean.

ii) About 95% of the data points fall within 2 standard deviations of the mean.

iii) About 99.7% of the data points fall within 3 standard deviations of the mean.

Example 2. If the mean distribution of IQ scores is 100 and the standard deviation is 16, what fraction of people have IQ's between 84 and 116? What fraction have IQ's above 132?

Computing Standard Scores

Suppose we want to know what fraction of people in the above example have IQ's above 105. We can not do this from the 68-95-99.7 rule. We need the following:

Definition 2. The number of standard deviations a data value lies above or below the mean is called its **standard score** defined by

$$z = \text{standard score} = \frac{\text{data value } - \text{mean}}{\text{standard deviation}}$$

Example 3. Given the mean of the IQ test to be 100, and the standard deviation to be 16, compute the standard score for an IQ of 105.

Definition 3. The n**th percentile** of a distribution the unique value in the distribution such that n percent of all other values are less than it.

Standard Scores are merely representation of percentiles as seen on page 408.

Example 4. Lets say we are given a distribution with mean 58.34, and standard deviation 5.31. Compute the percentile for the values $x_1 = 54.24$, $x_2 = 72.51$, $x_3 = 59.11$.