# Chapter Six

The Normal Distribution

**Summary**

• The normal distribution is central to the theory of inferential statistics. It also provides a model for many empirical distributions that approximate normality.

• In all normal or nearly normal curves, we find a constant proportion of the area under the curve lying between the mean and any given distance from the mean when measured in standard deviation units.

• The standard normal distribution is a normal distribution represented in standard scores, or Z scores, with mean = 0 and standard deviation = 1. Z scores express the number of standard deviations that a given score is above or below the mean. The proportions corresponding to any Z score or its fraction are organized into a special table called the standard normal table.

**Outline**

* The Normal Distribution
* A theoretical distribution, also known as the normal curve
* A theoretical distribution is similar to an empirical distribution in that it can be organized into frequency distributions, displayed using graphs, and described by its central tendency and variation using measures such as the mean and the standard deviation
  + - A theoretical distribution is based on theory rather than on real data
* Properties of the Normal Distribution
  + - One of the most striking characteristics of the normal distribution is its perfect symmetry
    - The normal curve is a theoretical ideal, and real-life distributions never match this model perfectly
    - Researchers study many variables that closely resemble this theoretical model
    - It is easier to understand the properties of a normal curve if we think in terms of a real distribution that is near normal
    - In all normal or nearly normal curves we find a constant proportion of the area under the curve lying between the mean and any given distance from the mean when measured in standard deviation units
    - As long as a distribution is normal and we know the mean and the standard deviation, we can determine the relative frequency (proportion or percentage) of cases that fall between any score and the mean
* Standard Z Scores
  + - A standard (Z) score is the number of standard deviations that a given raw score (or the observed score) is above or below the mean
    - A raw score can be transformed into a Z score to find how many standard deviations it is above or below the mean
      * To transform a raw score into a Z score, we divide the difference between the score and the mean by the standard deviation
    - A Z score can be converted to a raw score to find the score associated with a particular distance from the mean when this distance is expressed in standard deviation units
      * To transform a Z score into a raw score, multiply the Z score by the standard deviation and add this product to the mean
* The Standard Normal Distribution
  + - Standard normal distribution A normal distribution represented in standard (Z) scores, with mean = 0 and standard deviation = 1
* The Standard Normal Table
  + - A standard normal table is a table showing the area (as a proportion, which can be translated into a percentage) under the standard normal curve corresponding to any Z score or its fraction
    - The structure of the standard normal table consists of three columns representing positive Z scores, the area between the mean and the Z score, and the proportion of the area beyond the Z score
    - Use the standard table to find the area between the mean and a specified positive Z score, the area between the mean and a specified negative Z score, and the area above a positive Z score or below a negative Z score
* The Variance and the Standard Deviation
  + - Changes in the elderly population
    - The variance and the standard deviation are two closely related measures of variation that increase or decrease based on how closely the scores cluster around the mean
    - Both measure variability in interval-ratio variables
    - Calculating the deviation from the mean
    - Calculating the variance and the standard deviation
* Considerations for Choosing a Measure of Variation
  + - As in choosing a measure of central tendency, one of the most basic considerations in choosing a measure of variability is the variable’s level of measurement
    - With nominal variables, your choice is restricted to the IQV as a measure of variability
    - The choice of measure of variation for ordinal variables is more problematic. One compromise is to use the IQR (reporting Q1 and Q3) alongside the median, interpreting the IQR as the range of rank-ordered values that includes the middle 50% of the observations
    - For interval-ratio variables, you can choose the variance (or standard deviation), the range, or the IQR. However, if a distribution is extremely skewed so that the mean is no longer representative of the central tendency in the distribution, the range and the IQR can be used
* Reading the Research Literature
  + - Differences in college aspirations and expectations among Latino adolescents