

CLINICAL EPIDEMIOLOGY AND POPULATION HEALTH

Key Points – Descriptive Statistics

Measures of Central Tendency

Mean: $\Sigma X / n$

where Σ (Greek letter sigma) means to add, X represents the individual observations, and n is the number of observations. Commonly referred to as \bar{X} (called x-bar) for a sample mean or μ for a known population mean.

Median: the middle observation (i.e. half the observations are smaller and half are larger).

Arrange the observations from smallest to largest (or vice versa). The median is the middle value for an odd number of observations; it is defined as the mean of the two middle values for an even number of observations.

Mode: value that occurs most frequently

Measures of Spread

Range: the difference between the largest and the smallest observation

Interquartile Range (IQR): the difference between the 75th and 25th percentiles (i.e., the upper and lower quartiles). IQR is often presented with the median (i.e., 50th percentile)

Standard Deviation: measure of the average spread of data about their mean

The standard deviation is the square root of the **variance**.

Commonly referred as SD, sd or s (also σ , when referring to a known population standard deviation).

$$SD = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}}$$

If data are normally distributed (i.e., normal distribution):

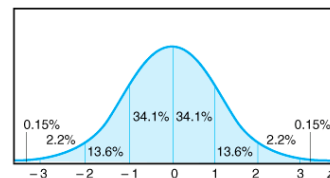
67% of the observations lie between the mean \pm 1 SD

95% of the observations lie between the mean \pm 2 SD

99.7% of the observations lie between the mean \pm 3 SD

Z-score: The Z-score is the number of standard deviations from the mean an observation is. You can transform any normal distribution into the standard normal distribution (aka Z distribution) with mean=0 and SD=1 using the following formula:

$$Z = \frac{X - \mu}{SD}$$



Source: Dawson B, Trapp RG: *Basic & Clinical Biostatistics*, 4th Edition: <http://www.accessmedicine.com>
Copyright © The McGraw-Hill Companies, Inc. All rights reserved.

Standard normal (z) distribution.

Where X is the observation, μ is the population mean and SD is the standard deviation. A Z-score of 0 is the mean. A Z-score of 1 means that the observation is 1 SD away from the mean, Z-score of 2 means the observation is 2 SD away from the mean, etc. The same percentages of the distribution as noted above in SD also apply to the Z-score.