Formula Sheet

# Variance

Variance is the measure of the spread between numbers in a data set. The variance of a sample of measurements is the sum of the square of the difference between the measurements and their mean, divided by n-1.

# Standard Deviation

Standard deviation measures the dispersion of a data set relative to its mean. The Standard Deviation of a sample of measurements is the positive square root of the variance.

# Permutations

A permutation is an ordered arrangement of *r* distinct objects. The following formula calculates the number of ways to order *n* distinct objects taken *r* at a time:

# Combinations

The number of combinations of *n* objects taken *r* at a time is the number of subsets, each of size *r,* that can be formed from *n* objects can be calculated with the following formula:

# Conditional Probability

The conditional probability of an event *A*, given that an event *B* has occurred is equal to :

Two Events *A* and *B* are said to be independent if any one of the following holds:

# Bayes’ Rule

Assume that {B1, B2, … , Bn} is a partition of S such that P(Bj) > 0, for I = 1,2,…k. Then:

# Binomial Distribution

Binomial distribution is used to model the number of successes in a sample of size n drawn with replacement from population of Size N. The binomial distribution of a random variable *Y* can be calculated by:

# Geometric Distribution

Geometric distribution is the probability distribution of the number of fails before the first success. The geometric distribution of a random variable *Y*  can be can be calculated by:

# Hyper Geometric Distribution

Hypergeometric distribution is a discrete probability distribution that describes the probability of successes in draws, without replacement. The hypergeometric distribution of a random variable *Y* can be calculated by:

# Poisson Distribution

A Poisson distribution is a statistical distribution showing the likely number of times that an event will occur within a specified period of time. The Poisson distribution of a random variable *Y*  can be calculated by:

# Tchebysheff’s Theorem

Tchebysheff’s theorem estimates the minimum proportion of numerical data that occur within a certain number of standard deviations from the mean.