Probability and Applied Statistics Equation Sheet

Ivan Wang

October 30, 2023

Chapter 1

Mean:

$$\text{Mean}(X) = \frac{\text{sum of all data points}}{\text{number of data points}}$$

Where X represents the data set.

Median: The middle value in a data set when arranged in ascending or descending order.

Mode: The data point(s) that appear most frequently in a set.

Standard Deviation (σ):

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \text{mean})^2}$$

Where n is the number of data points, and x_i is the i^{th} data point.

Variance (σ^2):

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \text{mean})^2$$

Where n is the number of data points, and x_i is the i^{th} data point.

Chapter 2

Factorial ($n!$):

$$n! = n(n-1)(n-2)\ldots 3 \times 2 \times 1$$

Represents the product of all positive integers less than or equal to n.

Permutations (nPr):

$$nPr = \frac{n!}{(n-r)!}$$

Where n is the total number of items and r is the number of items to choose.

Combinations (nCr):

$$nCr = \frac{n!}{r!(n-r)!}$$

Where n is the total number of items, and r is the number of items chosen at a time.

Conditional Probability ($P(A|B)$):

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Where A and B are events.

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Calculates the probability of A given that B has occurred, using prior knowledge.

Determining Independence: Events A and B are considered independent if the occurrence of one does not affect the occurrence of the other. Events A and B are independent if $P(A \cap B) = P(A)P(B)$.

Chapter 3

Binomial:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Where n is the number of trials, k is the number of successes, p is the probability of success on a given trial, and $1-p$ is the probability of failure.

Geometric:

$$P(X = k) = p(1-p)^{k-1}$$

Where k is the number of trials needed for the first success, and p is the probability of success on a given trial.

Hyper-geometric:

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Where N is the total number of items, K is the number of successes possible, n is the number of draws, and k is the number of observed successes.

Negative Binomial:

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

Where k is the number of failures until the r^{th} success, and p is the probability of success on a given trial.

Poisson:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Where λ is the average number of successes within a given time frame, and k is the actual number of successes.

Chebyshev's Inequality:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$
$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

Where μ is the mean, σ is the standard deviation, and k is a positive constant. For a given probability P, to find k:

$$k \geq \sqrt{\frac{1}{P}}$$