Formula Sheet for CSCI-3327-001 Probability and Applied Stats

Formulas taken from Wackerly, D., Mendenhall, W., & Scheaffer, R. (2008). *Mathematical Statistics with Applications* (7th ed.). Thomson Brooks/Cole.

Definition 1.1 - Mean

$$\mu = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

"The *mean* of a sample of n measured responses $y_1, y_2, \ldots, y_n$ … The corresponding population mean is denoted μ" (Wackerly et al., 2008, p. 9).

Definition 1.2 - Variance

$$\sigma^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

"The *variance* of a sample of measurements $y_1, y_2, \ldots, y_n$ is the sum of the square of the differences between the measurements and their mean, divided by n-1… The corresponding population variance is denoted by the symbol σ^2" (Wackerly et al., 2008, p. 10).

Definition 1.3 - Standard Deviation

$$\sigma = s = \sqrt{s^2}$$

"The *standard deviation* of a sample of measurements is the positive square root of the variance… The corresponding *population* standard deviations is denoted by $\sigma = \sqrt{\sigma^2}$" (Wackerly et al., 2008, p. 10).

Definition 2.7 - Permutation

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

"An ordered arrangement of r distinct objects is called a *permutation*. The number of ways of ordering n distinct objects taken r at a time will be designated by the symbol P_r^n" (Wackerly et al., 2008, p. 43).

Definition 2.8 - Combination

$$\binom{n}{r} = C_r^n = \frac{n!}{r!\,(n-r)!}$$

"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 the n objects. This number will be denoted by C_r^n or $\binom{n}{r}$" (Wackerly et al., 2008, p. 46).

Definition 2.9 - Conditional

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

"The *conditional probability of an event A*, given that an event B has occurred… provided $P(B) > 0$. [The symbol $P(A|B)$ is read 'probability of A given B.']" (Wackerly et al., 2008, p. 52).

Definition 2.10 - Two Events are Independent

$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$
$$P(A \cap B) = P(A)\,P(B)$$

"Two events A and B are said to be *independent* if any one of the following holds… Otherwise, the events are said to be *dependent*" (Wackerly et al., 2008, p. 53).

Theorem of Total Probability
From Professor during Final Topic of Ch. 2 Lecture on Sept. 27, 2023

$$P(A) = \sum_{i=1}^{k} P(A|B_i)\,P(B_i)$$

Can be found in Theorem 2.8 on p. 70 – (Wackerly et al., 2008, p. 70).

Theorem 2.9 - Bayes' Rule | Bayes' Theorem

$$P(B_j|A) = \frac{P(A|B_j)\,P(B_j)}{\sum_{i=1}^{k} P(A|B_i)\,P(B_i)}$$

"Assume that $\{B_1, B_2, \ldots, B_k\}$ is a partition of S… such that $P(B_i) > 0$, for $i = 1, 2, \ldots, k$" (Wackerly et al., 2008, p. 71). – Basically: Bayes' theorem is when B **HAS NOT** been observed <u>given</u> that A **HAS** been observed.

Definition 3.1 - Random Variable

"A *random variable* is a real-valued function for which the domain is a sample space…random variable Y is said to be *discrete* if it can assume only a finite or countably infinite number of discrete values" (Wackerly et al., 2008, p. 76, 87).

Definition 3.3 & Theorem 3.1 - Probability Distribution and Two Requirements

"The *probability distribution* for a discrete variable Y can be represented by a formula, a table, or a graph that provides $p(y) = P(Y = y)$ for all y" (Wackerly et al., 2008, p. 88).

"For any discrete probability distribution, the following must be true:
   1. $0 \leq p(y) \leq 1$ for all y.

2. $\Sigma_y\, p(y) = 1$, where the summation is over all values of y with nonzero probability"

(Wackerly et al., 2008, p. 89).

Definition 3.6 - Binomial Expression Requirements
From Professor during Binomial Distribution lecture on Oct. 4, 2023

Rules:
1. Fixed # of identical trials.
2. Results in TWO outcomes, successes and failures.
3. Probability of failure is equal to $q = (1 - p)$ where q is the probability of failure.
4. Trials are INDEPENDENT.
5. RV is Y = # of successes observed during the n trials

Can be found in Definition 2.6 on p. 101 – (Wackerly et al., 2008, p. 101).

Definition 3.7 - Binomial Distribution

$$p(y) = \binom{n}{y} p^y q^{n-y}, \; y = 0, 1, 2, \ldots, n \; and \; 0 \le p \le 1$$

"A random variable Y is said to have a *binomial distribution* based on n trials with success probability p..." (Wackerly et al., 2008, p. 103).

Theorem 3.7 - Binomial Expected Value and Expected Variance

$$\mu = E(Y) = n\,p \; and \; \sigma^2 = V(Y) = n\,p\,q$$

"Let Y be a binomial random variable based on n trials and success probability p..." (Wackerly et al., 2008, p. 107).

3.5 - Geometric Experiment

"...the experiment consists of a series of trials that concludes with the first success. Consequently, the experiment could end with the first trial if a success is observed on the very first trial, or the experiment could go on infinitely" (Wackerly et al., 2008, p. 114).

Definition 3.8 - Geometric Distribution

$$p(y) = q^{y-1} p, \quad y = 1, 2, 3, \ldots, \quad 0 \le p \le 1$$

"A random variable Y is said to have a *geometric probability distribution*..." (Wackerly et al., 2008, p. 115).

Theorem 3.8 - Geometric Expected Value and Expected Variance

$$\mu = E(Y) = \frac{1}{p} \; and \; \sigma^2 = V(Y) = \frac{1-p}{p^2}$$

(Wackerly et al., 2008, p. 116).

Extra Formulas
From Professor during Geometric Distribution lecture on Oct. 6, 2023

Use the following when a success occurs…
1. **ON** or **BEFORE** the nth trial

$$p(x \leq n) = 1 - (1 - p)^n$$

2. **BEFORE** the nth trial

$$p(x < n) = 1 - (1 - p)^{n-1}$$

3. **ON** or **AFTER** the nth trial

$$p(x \geq n) = (1 - p)^{n-1}$$

4. **AFTER** the nth trial

$$p(x > n) = (1 - p)^n$$

Definition 3.10 - Hypergeometric Probability Distribution

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

"A random variable Y is said to have a *hypergeometric probability distribution*… where y is an integer $0, 1, 2, . . ., n$, subject to the restrictions $y \leq r$ and $n - y \leq N - r$" (Wackerly et al., 2008, p. 126).

Theorem 3.10 - Hypergeometric Expected Value and Expected Variance

$$\mu = E(Y) = \frac{nr}{N} \ and \ \sigma^2 = V(Y) = n(\frac{r}{N})(\frac{N-r}{N})(\frac{N-n}{N-1})$$

(Wackerly et al., 2008, p. 127).