Jackson Campbell
Professor Hoy

Probability and Applied Statistics
Formula Worksheet

Definition 1.1: Mean

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

Definition 1.2: Variance

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

Definition 1.3: Standard Deviation

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

Definition 2.7: Permutations

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

Theorem 2.3: Partitioning *n* distinct objects into *k* distinct groups

$$N = \binom{n}{n_k} = \frac{n!}{n_k!}$$

Definition 2.8: Combinations

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

Definition 2.9 Conditional Probability

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

Definition 2.10: Independent events are evident if one of the following holds:

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

Theorem 2.5: Multiplicative Law of Probability

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

If independent: $P(A \cap B) = P(A)P(B)$

Theorem 2.6: Additive Law of Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If mutually exclusive: $P(A \cup B) = P(A) + P(B)$

Theorem 2.7: Compliments

$$P(A) = 1 - P(\bar{A})$$

Definition 2.11: Partitions

$S = B_1 \cup B_2 \cup \dots \dots \cup B_k$

$B_i \cap B_j = \emptyset \;\; for\ i \neq j$

$\{B_k\}$ is a partition of *S*.

Theorem 2.8: Decomposition of events

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

Theorem 2.9: Bayes' Rule

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

Definition 3.4: Expected Value

$$E(Y) = \sum_{y} yp(y)$$

Theorem 3.2: Expected value of g(Y)

$$E[g(Y)] = \sum_{all\ y}^{n} g(y)p(y)$$

By definition 3.4,

$$E[g(Y)] = \sum_{j=1}^{n} g(y_j)p(y_j)$$

Definition 3.5: Variance of Y

$$V(Y) = E[(Y - \mu)^2]$$

Theorem 3.3: Mean/Expected value of c

$$E(c) = \sum_y cp(y) = c\sum_y p(y)$$

By Theorem 3.1, $E(c) = c(1) = c$

Theorem 3.4: Mean/Expected value of product of c * random variable function

$E[cg(Y)] = cE[g(Y)]$

By Theorem 3.2, $E[cg(Y)] = \sum_y cg(y)p(y) = c\sum_y g(y)p(y) = cE[g(Y)]$

Theorem 3.5: Mean/Expected value of sum of random variable Y function

$E[g_1(Y) + g_2(Y) + \ldots\ldots + g_k(Y)] = E[g_1(Y)] + E[g_2(Y)] + \ldots\ldots + E[g_k(Y)]$

Theorem 3.6: Variance of discrete random variable

$$V(Y) = \sigma^2 = E[(Y-\mu)^2] = E(Y^2) - \mu^2$$

Definition 3.6: Binomial experiment properties:

1. Consists of a fixed number, *n,* of identical trials
2. Each trial results in one of two outcomes: success, *S,* or failure, *F.*
3. The probability of success on a single trial equates to some value *p and* remains from trial to trial. Failure is equal to *q = 1 – p*
4. The trials are independent.
5. The random variable of interest, Y, is the number of successes observed during the *n* trials.

Definition 3.7: Binomial Distribution

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

Theorem 3.7: Mean and Variance of Binomial Distribution

$$\mu = E(Y) = np \qquad \sigma^2 = V(Y) = npq$$

Definition 3.8: Geometric Distribution

A random variable Y has geometric probability distribution IFF

$$p(y) = q^{y-1}p \qquad y = 1, 2, 3, \dots\dots, etc \qquad 0 \leq p \leq 1$$

Theorem 3.8: Mean and Variance of Geometric Distribution

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

Definition 3.9: Negative Binomial Probability Distribution

A random variable Y is said to have a NBPD IFF

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r} \qquad y = r, r+1, \dots, etc \qquad 0 \leq p \leq 1$$

Theorem 3.9: Expected and Variance of NBPD

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

Definition 3.10: Hypergeometric probability distribution

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \quad y\ is\ an\ int. \quad y \leq r, \quad n - y \leq N - r$$

Theorem 3.10: Expected and Variance of Hypergeometric

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

Definition 3.11: Poisson probability distribution

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda} \qquad y = 0, 1, \dots, etc \qquad \lambda > 0.$$

Theorem 3.11: Expected and Variance of Poisson

$$\mu = E(Y) = \lambda \qquad \sigma^2 = V(Y) = \lambda$$

Theorem 3.14: Tchebysheff's Theorem

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

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

Definition 4.1: Distribution for Random Variable Y

$$F(y) = P(Y \leq y) \ for \quad -\infty < y < \infty$$

Definition 4.3: Distribution Function for Continuous Random Variable Y

$$f(y) = \frac{dF(y)}{dy} = F'(y)$$

Theorem 4.3: Density function f(y) and a < b

$$P(a \leq Y \leq b) = \int_a^b f(y)dy$$

Definition 4.5: Expected of a Continuous Random Variable Y

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy$$

Theorem 4.4: Let g(Y) be a function of Y, Expected Value of g(Y)

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy$$

Theorem 4.5: Let c be a constant, let g(Y), g1(Y), ... gk(Y) be functions of a Continuous Random Variable Y. Said results hold:

$$E(c) = c$$

$$E[cg(Y)] = cE[g(Y)]$$

$$E[g1(Y) + g2(Y) + \cdots + gk(Y)] = E[g1(Y)] + E[g2(Y)] + \cdots + E[gk(Y)]$$

Definition 4.6: Continuous Uniform Probability Distribution

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1}, & \theta_1 \le y \le \theta_2 \\ 0, & elsewhere \end{cases}$$

Theorem 4.6: Expected and Variance for Uniform Distribution

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \qquad \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}$$

Definition 4.8: Normal Probability Distribution

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}, \quad -\infty < y < \infty$$

Theorem 4.7: Expected and Variance of Normal Probability Distribution

$$E(Y) = \mu \qquad V(Y) = \sigma^2$$

Definition 4.9: Gamma Distribution

$$f(y) = \begin{cases} \frac{y^{\alpha-1} e^{-y/\beta}}{\beta^\alpha \Gamma(\alpha)}, & 0 \le y \le \infty \\ 0, & elsewhere \end{cases}$$

Theorem 4.8: Expected and Variance of Gamma Distribution

$$\mu = E(Y) = \alpha\beta \qquad \sigma^2 = V(Y) = \alpha\beta^2$$

Theorem 4.9: Chi-Square Random Variable with v Degrees of Freedom

$$\mu = E(Y) = v \qquad \sigma^2 = V(Y) = 2v$$

Definition 4.11: Exponential Distribution with parameter β > 0

$$f(y) = \begin{cases} \frac{1}{\beta} e^{-y/\beta}, & 0 \le y < \infty \\ 0, & elsewhere \end{cases}$$

Theorem 4.10: Expected and Variance for Exponential Distribution

$$\mu = E(Y) = \beta \qquad \sigma^2 = V(Y) = \beta^2$$

Definition 5.1: Joint Probability Function

$$p(x, y) = P(X = x, Y = y), \qquad -\infty < x < \infty, -\infty < y < \infty$$

Definition 5.2: Joint Distribution Function

$$F(x, y) = P(X \leq x, Y \leq y), \qquad -\infty < x < \infty, -\infty < y < \infty$$

Definition 5.3: Joint Distribution Function nonnegative

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t_1, t_2) dt_2 dt_1$$

Theorem 5.2.1: If Y1 and Y2 are random variables with JDF F(x, y), then

1. $F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0$
2. $F(\infty, \infty) = 1$
3. $If\ x^* \geq x\ and\ y^* \geq y, then\ F(x^*, y^*) - F(x^*, y) - F(x, y^*) + F(x, y) \geq 0$

Theorem 5.2.2: If Y1 and Y2 are jointly continuous random variables with a JDF given by f(x,y), then

$$F(x, y) \geq 0\ for\ all\ x, y$$
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

Definition 5.4:
a. Let X and Y be jointly discrete random variables with probability function p(x,y). Then the marginal probability functions of X and Y respectively are:

$$p(x) = \sum_{all\ y} p(x, y) \quad and \quad p(y) = \sum_{all\ x} p(x, y)$$

b. Let X and Y be jointly continuous random variables with joint density function f(x,y). Then the marginal density functions of X and Y respectively are:

$$f(x) = \int_{-\infty}^{\infty} f(x,y)dy \quad and \quad f(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

Definition 5.5: If X and Y are jointly discrete random variables with joint probability function p(x,y) and marginal probability functions p(x) and p(y) respectively, then the conditional discrete probability function of X given Y is:

$$p(x|y) = P(X = x \,|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x,y)}{p(y)}$$

Definition 5.6: If X and Y are jointly continuous random variables with joint density function f(x,y), then the conditional distribution function of X given Y = y is:

$$F(x|y) = P(X \leq x|Y = y)$$

Definition 5.7: Let X and Y be jointly continuous random variables with joint density f(x,y) and marginal densities f(x) and f(y) respectively. For any y such that f(y) > 0, the conditional density of X given Y = y is:

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

And for any x such that f(x) > 0, the conditional density of Y given X = x is:

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

Definition 5.8: Let X have distribution function F(x), Y have distribution function F(y), and X and Y have joint distribution function F(x, y). Then X and Y are said to be independent IFF:

$$F(x,y) = F(x)F(y)$$

For every pair of real numbers (x, y). If X and Y are not independent, they are said to be dependent.

Theorem 5.4: If X and Y are discrete random variables with joint probability function p(x, y) and marginal probability functions p(x) and p(y) respectively, then X and Y are independent IFF:

$$p(x,y) = p(x)p(y)$$

For all pairs of real numbers (x, y).

If X and Y are continuous random variables with joint density function f(x, y) and marginal density functions f(x) and f(y) respectively, then X and Y are independent IFF:

$$f(x,y) = f(x)f(y)$$

For all pairs of real numbers (x, y).

Theorem 5.5: Let X and Y have a joint density f(x, y) that is positive IFF a ≤ x ≤ b and c ≤ y ≤ d, for constants a, b, c, and d; and f(x, y) = 0 otherwise. Then X and Y are independent random variables IFF:

$$f(x,y) = g(x)h(y)$$

Where g(x) is a nonnegative function of x alone and h(y) is a nonnegative function of y alone.