**Probability & Applied Statistics**

**Chapters 1-3**: Formulas, Definitions & Theorems

**Chapter 1**

**Definition 1.1**

**Mean:**

The *mean* of a sample of *n* measured responses

Sample mean (

Population mean (

**Definition 1.2**

**Variance:**

The *variance* of a sample of measurements

The *variance* of a population of measurements

Sample variance (

Population variance (

**Definition 1.3**

**Standard Deviation:**

The *standard deviation* of a sample of measurements

The *standard deviation* of a population of measurements

Sample standard deviation (

Population standard deviation (

**Empirical Rule**

For a distribution of measurements that is approximately normal (bell shaped),

it follows that the interval with end points

μ ± σ contains approximately 68% of the measurements.

μ ± 2σ contains approximately 95% of the measurements.

μ ± 3σ contains almost all of the measurements

**Chapter 2**

**Definition 2.1**

An *experiment* is the process by which an observation is made.

**Definition 2.2**

A *simple event* is an event that cannot be decomposed. Each simple event corresponds to one and only one *sample point*. The letter E with a subscript will be used to denote a simple event or the corresponding sample point.

**Definition 2.3**

The *sample space* associated with an experiment is the set consisting of all possible sample points. A sample space will be denoted by S.

**Definition 2.4**

A *discrete sample space* is the one that contains either a finite or a countable number of distinct sample points.

**Definition 2.5**

An *event* in a discrete sample space S is a collection of sample points- that is, any subset of S.

**Definition 2.6**

Suppose S is a sample space associated with an experiment. To every event A in S (A is a subset of S), we assign a number, P(A), called the *probability* of A, so that the following axioms hold:

Axiom 1: P(A) ≥ 0.

Axiom 2: P(S) = 1.

Axiom 3: If A1, A2, A3,... form a sequence of pairwise mutually

exclusive events in S (that is, Ai ∩ Aj = ∅ if i ≠ j), then

**Definition 2.7**

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 .

**Definition 2.8**

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 or .

The number of unordered subsets of size r chosen (without replacement) from *n* available objects is

**Definition 2.9**

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

provided P(B) > 0.

**Definition 2.10**

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

P(A|B) = P(A)

P(B|A) = P(B)

P(A∩B) = P(A)P(B)

Otherwise, the events are said to be *dependent*.

**Definition 2.11**

For some positive integer k, let the sets B1, B2, …, Bk be such that

1. S = B1 U B2 U … U Bk

2. Bi ∩ Bj = Ø, for i ≠ j

Then the collection of sets {B1, B2, …, Bk} is said to be a *partition* of S.

**Bayes’ Rule**

Assume that {B1, B2, …, Bk} is a partition of S (see Definition 2.11) such that P(Bi) > 0, for i = 1, 2,..., k. Then

**Definition 2.12**

A *random variable* is a real-valued function for which the domain is a sample space.

**Definition 2.13**

Let *N* and *n* represent the number of elements in the population and sample, respectively. If the sampling is conducted in such a way that each of the samples has an equal probability of being selected, the sampling is said to be random, and the result is said to be a *random sample*.

**Chapter 3**

**Definition 3.1**

A random variable Y is said to be *discrete* if it can assume only a finite or countably finite number of distinct values.

**Definition 3.2**

The probability that Y takes on the value *y*, P (Y = y), is defined as the *sum of the probabilities of all sample points in S* that are assigned the value *y*. We sometimes denote P (Y = y) by p(y).

**Definition 3.3**

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.

**Definition 3.4**

**Mean**

Let Y be a discrete random variable with the probability function p(y). Then the *expected value* of Y, E(Y), is defined to be

**Definition 3.5**

**Variance & Standard Deviation**

If Y is a random variable with mean E(Y) = μ, the variance of a random variable Y is defined to be the expected value of (Y – μ)2.

The *standard deviation* of Y is the positive square root of V(Y).

**Definition 3.6**

A *binomial experiment* possesses the following properties:

1. The experiment 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 is equal to some value *p* and remains the same from trial to trial. The probability of a failure is equal to q = (1 - p).

4. The trials are independent.

5. The random variable of interest is Y, the number of successes observed during the *n* trials.

**Definition 3.7**

A random variable Y is said to have a *binomial distribution* based on *n* trials with success probability *p* if and only if

**Theorem 3.7**

Let Y be a binomial random variable based on *n* trials and success probability p.

and

**Definition 3.8**

A random variable Y is said to have a *geometric probability distribution* if and only if

**Theorem 3.8**

If Y is a random variable with a geometric distribution,

and

**Definition 3.10**

A random variable Y is said to have a *hypergeometric probability distribution* if and only if

where y is an integer 0, 1, 2, … , n, subject to the restrictions y ≤ r and n – y ≤ N – r.

**Theorem 3.10**

If Y is a random variable with a hypergeometric distribution,

and

**Definition 3.11**

A random variable Y is said to have a *Poisson probability distribution* if and only if

, y = 0, 1, 2, …,

**Theorem 3.11**

If Y is a random variable possessing a Poisson distribution with parameter λ, then

**Theorem 3.14**

**Tchebysheff’s Theorem**

Let Y be a random variable with mean μ and finite variance σ2. For any constant k > 0,

**Chapter 4**

**Definition 4.1**

Let Y denote any random variable. The *distribution function* of Y, denoted by F(y), is such that F(y) = P (Y ≤ y) for −∞ < y < ∞.

**Theorem 4.1**

**Properties of a Distribution Function:** If F(y)is a distribution function, then

1. F (−∞) = lim y → −∞ F(y) = 0.

2. F (∞) ≡ lim y → ∞ F(y) = 1.

3. F(y) is a nondecreasing function of y. [If y1 and y2 are any values such that y1 < y2, then F(y1) ≤ F(y2).]

**Definition 4.2**

A random variable *Y* with distribution function F(y) is said to be *continuous* if F(y) is continuous, for −∞ < y < ∞.

**Definition 4.3**

Let *F*(*y*) be the distribution function for a continuous random variable Y. Then *f* (*y*), given by

wherever the derivative exists, is called the *probability density function* for the random variable Y.

**Theorem 4.2**

**Properties of a Density Function:** If f (*y*)is a density function for a continuous random variable, then

1. f (y) ≥ 0 for all y, −∞ < y < ∞.

2.

**Definition 4.5**

The expected value of a continuous random variable *Y* is

provided that the integral exists.

**Theorem 4.4**

Let g (Y ) be a function of Y ; then the expected value of g(Y ) is given by

provided that the integral exists

**Definition 4.6**

If θ1 < θ2, a random variable *Y* is said to have a continuous *uniform probability* *distribution* on the interval (θ1, θ2) if and only if the density function of *Y* is

**Definition 4.7**

The constants that determine the specific form of a density function are called *parameters* of the density function.

**Definition 4.8**

A random variable *Y* is said to have a *normal probability distribution* if and only if, for σ > 0 and −∞ < µ < ∞, the density function of *Y* is

**Theorem 4.7**

If *Y* is a normally distributed random variable with parameters µ and σ, then