**Formula Sheet**

## **Chapter 1**

### **Definition 1.1 Mean**

The mean of a sample of n measured responses y1, y2,..., yn is given by

The corresponding population mean is denoted µ

### **Definition 1.2 Variance**

The variance of a sample of measurements y1, y2,..., yn is the sum of the square of the differences between the measurements and their mean, divided by n − 1. Symbolically, the sample variance is

The corresponding population variance is denoted by the symbol σ2

### **Definition 1.3 Standard Deviation**

The standard deviation of a sample of measurements is the positive square root

of the variance; that is,

The corresponding population standard deviation is denoted by σ = √ σ2.

## **Chapter 2**

### Definition 2.1 Experiment

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

### Definition 2.2 Simple Event

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 Sample Space

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 Discrete Sample Space

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

### Definition 2.5 Event

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

### Definition 2.6 Probability

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:

Axiom 2:

Axiom 3: If A1, A2, A3, . . . form a sequence of pairwise mutually exclusive events in S (that is, if ) then

**Sample-point Method**

The sample-point method is outlined in Section 2.4. The following steps are used to find the probability of an event:

1. Define the experiment and clearly determine how to describe one simple event.

2. List the simple events associated with the experiment and test each to make certain that it cannot be decomposed. This defines the sample space S.

3. Assign reasonable probabilities to the sample points in S, making certain that P(Ei) ≥ 0 and P(Ei) = 1.

4. Define the event of interest, A, as a specific collection of sample points. (A sample point is in A if A occurs when the sample point occurs. Test all sample points in S to identify those in A.)

5. Find P(A) by summing the probabilities of the sample points in A.

### Theorem 2.1 mn = m x n

With m elements a1, a2,..., am and n elements b1, b2,..., bn, it is possible to form mn = m × n pairs containing one element from each group.

**Proof** Verification of the theorem can be seen by observing the rectangular table in Figure 2.9. There is one square in the table for each ai, bj pair and hence a total of m × n squares.

### Definition 2.7 Permutation

An ordered arrangement for 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 Pn r .

### Definition 2.8 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 the *n* objects. This number will be denoted by or ()

### Definition 2.9 Conditional Probability of an Event

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

provided P(B) > 0. [The symbol P(A|B) is read “probability of A given B.”]

### Definition 2.10 Independent Cases

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.

### Theorem 2.5 The Multiplicative Law of Probability

The Multiplicative Law of Probability The probability of the intersection of two events A and B is

If A and B are independent, then

**Proof** The multiplicative law follows directly from Definition 2.9, the definition of conditional probability

### Theorem 2.6 The Additive Law of Probability

The probability of the union of two events A and B is

If A and B are mutually exclusive events, P(A ∩ B) = 0 and

**Proof** The proof of the additive law can be followed by inspecting the Venn diagram in Figure 2.10. Notice that A ∪ B = A ∪ (A ∩ B), where A and (A ∩ B) are mutually exclusive events. Further, B = (A ∩ B)∪(A ∩ B), where (A ∩ B) and (A ∩ B) are mutually exclusive events. Then, by Axiom 3, P(A ∪ B) = P(A) + P(A ∩ B) and P(B) = P(A ∩ B) + P(A ∩ B). The equality given on the right implies that P(A ∩ B) = P(B) − P(A ∩ B). Substituting this expression for P(A ∩ B) into the expression for P(A ∪ B) given in the left-hand equation of the preceding pair, we obtain the desired result:

### Theorem 2.7 Mutual Exclusive Events

If A is an event, then

**Proof** Observe that S = A ∪ A. Because A and A are mutually exclusive events, it follows that P(S) = P(A) + P(A). Therefore, P(A) + P(A) = 1 and the result follows.

### Definition 2.11 Partition

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

1. S = B1 ∪ B2 ∪···∪ Bk .

2. Bi ∩ Bj = ∅, for i = j.

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

### Theorem 2.9 Bayes’ Rule

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 Random Variable

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

### Definition 2.13 Random Sample

Let N and n represent the numbers of elements in the population and sample, respectively. If the sampling is conducted in such a way that each of the N n 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 Discrete Values

A random variable Y is said to be discrete if it can assume only a finite or countably infinite1 number of distinct values.

### Definition 3.2 Sum of the Probabilities of all Sample Points

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 will sometimes denote P(Y = y) by p(y).

### Definition 3.3 Probability Distribution

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 Expected Value

Let Y be a discrete random variable with the probability function p(y). Then the expected value of Y , E(Y ), is defined to be2

### Definition 3.5 Random Variable w/Mean

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. That is,

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

### Definition 3.6 Binomial Experiment

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 Binomial Distribution

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

, and

### Definition 3.8 Geometric Probability

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

### Theorem 3.8 Random Variable with a Geometric Distribution

If Y is a random variable with a geometric distribution

and

### Definition 3.9 Negative Binomial Probability Distribution

A random variable Y is said to have a negative binomial probability distribution if and only if

,

### Theorem 3.9 Random Variable with a Negative Binomial Distribution

If Y is a random variable with a negative binomial distribution,

and

### Definition 3.10 Hypergeometric Probability DistributionA 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 Random Variable with a Hypergeometric Distribution

If Y is a random variable with a hypergeometric distribution,

and .

### Definition 3.11 Poisson Probability Distribution

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

, .

### Theorem 3.11 Random variable Possessing a Poisson Distribution

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

and

### Definition 3.12 Kth Moment of a Random Variable Taken About the Origin

The kth moment of a random variable Y taken about the origin is defined to be and is denoted by .

### Definition 3.13 Kth Moment of a Random Variable Taken About the Mean

The kth moment of a random variable Y taken about its mean, or the kth central moment of Y , is defined to be and is denoted by .

### Definition 3.14 Moment-Generating Function for a Random Variable

The moment-generating function m(t) for a random variable Y is defined to be . We say that a moment-generating function for Y exists if there exists a positive constant b such that is finite for .

### Theorem 3.12

If m(t) exists, then for any positive integer k,

In other words, if you find the kth derivative of m(t) with respect to t and then set t = 0, the result will be .

### Definition 3.15 Probability-Generating Function

Let Y be an integer-valued random variable for which , where i = 0, 1, 2, . . . . The probability-generating function P(t) for Y is defined to be

for all values of t such that P(t) is finite.

### Definition 3.16 Factorial Moment

The kth factorial moment for a random variable Y is defined to be

,

where k is a positive integer

### Theorem 3.13

If P(t) is the probability-generating function for an integer-valued random variable, Y, then the kth factorial moment of Y is given by

.

### Theorem 3.14 Tchebysheff’s Theorem

Tchebysheff’s Theorem Let Y be a random variable with mean µ and finite variance σ2. Then, for any constant k > 0,

or .

## **Chapter 4**

### Definition 4.1 Distribution Function

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

Properties of a Distribution Function1 If F(y)is a distribution function, then

1. .
2. .
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 Continuous

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

### Definition 4.3 Probability Density Function

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

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

1. for all
2. .

### Definition 4.4 pTh Quantile

Let Y denote any random variable. If 0 < p < 1, the pth quantile of Y , denoted by , is the smallest value such that P(Y ≤ ) = F(φp) ≥ p. If Y is continuous, φp is the smallest value such that F( ) = P(Y ≤ ) = p. Some prefer to call the 100pth percentile of Y .

### Theorem 4.3

If the random variable Y has density function f (y) and a < b, then the probability that Y falls in the interval [a, b] is

.

### Definition 4.5 Expected Value of a Continuous Random Variable

The expected value of a continuous random variable Y is

,

provided that the integral exists.

### Theorem 4.4

M 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.

### Theorem 4.5

Let c be a constant and let be functions of a continuous random variable Y . Then the following results hold:

1. .
2. .
3. .

### Definition 4.6 Continuous Uniform Probability Distribution

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

### Definition 4.7 Parameters

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

### Theorem 4.6

If and Y is a random variable uniformly distributed on the interval (, then

and .

### Definition 4.8 Normal Probability Distribution

A random variable Y is said to have a normal probability distribution if and only if, for and, the density function of Y is

, .

### Theorem 4.7

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

and

### Definition 4.9 Gamma Distribution

A random variable Y is said to have a gamma distribution with parameters α > 0 and β > 0 if and only if the density function of Y is

where

### Theorem 4.8

If Y has a gamma distribution with parameters α and β, then

and

### Definition 4.10 Chi-Square Distribution

0 Let ν be a positive integer. A random variable Y is said to have a chi-square distribution with ν degrees of freedom if and only if Y is a gamma-distributed random variable with parameters α = ν/2 and β = 2.

### Theorem 4.9

If Y is a chi-square random variable with ν degrees of freedom, then

and

### Definition 4.11 Exponential Distribution with Parameters

A random variable Y is said to have an exponential distribution with parameter β > 0 if and only if the density function of Y is

### Theorem 4.10

If Y is an exponential random variable with parameter β, then

and

### Definition 4.12 Beta Probability Distribution

A random variable Y is said to have a beta probability distribution with parameters α > 0 and β > 0 if and only if the density function of Y is

### Theorem 4.11

If Y is a beta-distributed random variable with parameters α > 0 and β > 0, thenl

and .

### Definition 4.13

If Y is a continuous random variable, then the kth moment about the origin is given by

### Definition 4.14 Moment-Generating Function

If Y is a continuous random variable, then the moment-generating function of Y is given by

The moment-generating function is said to exist if there exists a constant b > 0 such that m(t) is finite for | t | ≤ b.

### Theorem 4.12

Let Y be a random variable with density function f (y) and g(Y) be a function of Y. Then the moment-generating function for g(Y) is

### Theorem 4.13 Tchebysheff’s

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

### Definition 4.15

Let Y have the mixed distribution function

and suppose that X1 is a discrete random variable with distribution function F1(y) and that X2 is a continuous random variable with distribution function F2(y). Let g(Y) denote a function of Y. Then

## **Chapter 5**

### Definition 5.1 Join Probability Function

Let Y1 and Y2 be discrete random variables. The joint (or bivariate) probability function for Y1 and Y2 is given by

### Theorem 5.1

If Y1 and Y2 are discrete random variables with joint probability function p(y1, y2), then

1. for all y1, y2.
2. where the sum is over all values (y1, y2) that are assigned nonzero probabilities.

### Definition 5.2

For any random variables Y1 and Y2, the joint (bivariate) distribution function F(y1, y2) is

### Definition 5.3

Let Y1 and Y2 be continuous random variables with joint distribution function F(y1, y2). If there exists a nonnegative function f(y1, y2), such that

for all then Y1 and Y2 are said to be jointly continuous random variables. The function f(y1, y2) is called the joint probability density function.

### Theorem 5.2

If Y1 and Y2 are random variables with joint distribution function F(y1, y2), then

1. F() = F() = F() = 0.
2. F() = 1.
3. If y∗ 1 ≥ y1 and y∗ 2 ≥ y2, then F(y∗ 1, y∗ 2 ) − F(y∗ 1, y2) − F(y1, y∗ 2 ) + F(y1, y2) ≥ 0.

### Definition 5.4

### Definition 5.5 Conditional Discrete Probability Function of Y

If Y1 and Y2 are jointly discrete random variables with joint probability function p(y1, y2) and marginal probability functions p1(y1) and p2(y2), respectively, then the conditional discrete probability function of Y1 given Y2 is

provided that

### Definition 5.6 Conditional Distribution Function of Y

If Y1 and Y2 are jointly continuous random variables with joint density function f(y1, y2), then the conditional distribution function of Y2 given Y2 = y2 is

### Definition 5.7 Jointly Continuous Random Variables with Joint Density

Let Y1 and Y2 be jointly continuous random variables with joint density f(y1, y2) and marginal densities f1(y1) and f2(y2), respectively. For any y2 such that f2(y2) > 0, the conditional density of Y1 given Y2 = y2 is given by

and, for any y1 such that f1(y1) > 0, the conditional density of Y2 given Y1 = y1 is given by

### Definition 5.8

Let Y1 have distribution function F1(y1), Y2 have distribution function F2(y2), and Y1 and Y2 have joint distribution function F(y1, y2). Then Y1 and Y2 are said to be independent if and only if

for every pair of real numbers (y1, y2). If Y1 and Y2 are not independent, they are said to be dependent.

### Theorem 5.4

If Y1 and Y2 are discrete random variables with joint probability function p(y1, y2) and marginal probability functions p1(y1) and p2(y2), respectively, then Y1 and Y2 are independent if and only if

for all pairs of real numbers (y1, y2).

### Theorem 5.5

Let y1 and y2 have a joint density f(y1, y2) that is positive if and only if a ≤ y1 ≤ b and c ≤ y2 ≤ d, for constants a, b, c, and d; and f(y1, y2) = 0 otherwise. Then Y1 and Y2 are independent random variables if and only if

where g(y1) is a nonnegative function of y1 alone and h(y2) is a nonnegative function of y2 alone.