**Stats Formula Sheet**

By Colin Glory

**Definitions**

**Definition 1.1**

**Definition 1.2**

**Definition 1.3**

**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 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**

The probability of pairwise exclusive events is the sum of the probabilities.

**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

**Definition 2.9**

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

**Definition 2.10**

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.

**Definition 2.11**

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

1.

2.

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

**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 numbers 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.

**Definition 3.1**

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**

The probability that Y takes on the value 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 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 for all y.

**Definition 3.4**

Let Y be a discrete random variable with the probability function p(y). Then

the expected value of , is defined to :

**Definition 3.5**

If Y is a random variable with mean , the variance of a random

variable Y is defined to be the expected value of .That is,

**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:

**Definition 3.8**

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

only if:

**Definition 3.9**

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

if and only if :

**Definition 3.10**

A random variable Y is said to have a hypergeometric probability distribution

if and only if:

**Definition 3.11**

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

**Definition 3.12**

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

**Definition 3.13**

The kth moment of a random variable Y taken about its mean, or the kth central moment of Y , is defined to be E[(Y − μ)^k ] and is denoted by .

**Definition 3.14**

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 m(t) is finite for |t| ≤ b.

**Definition 4.1**

Let Y denote any random variable. The distribution function of Y , denoted by F(y), is such that

**Definition 4.2**

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

**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.

**Definition 4.4**

Let Y denote any random variable.

**Definition 4.5**

The expected value of a continuous random variable Y is , provided that the integral exists.

**Definition 4.6**

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**

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

**Definition 5.1**

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

**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.

**Definition 5.4**

1. Let Y1 and Y2 be jointly discrete random variables with probability function p(y1, y2). Then the marginal probability functions of Y1 and Y2, respectively, are given by:
2. Let Y1 and Y2 be jointly continuous random variables with joint density function f (y1, y2). Then the marginal density functions of Y1 and Y2, respectively, are given by:

**Definition 5.5**

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:

**Definition 5.6**

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

**Definition 5.7**

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:

**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.

**Definition 5.9**

Let be a function of the discrete random variables, , which have probability function . Then the expected value of is:

**Theorems**

**Theorem 2.1**

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.

**Theorem 2.2**

**Theorem 2.3**

The number of ways of partitioning n distinct objects into k distinct groups containing n1, n2,...,nk objects, respectively, where each object appears in exactly one group and is:

**Theorem 2.4**

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

**Theorem 2.5**

Multiplicative Law of Probability:

If independent:

**Theorem 2.6**

General Addition Rule

If mutually exclusive: P(A ∩ B) = 0

**Theorem 2.7**

**Theorem 2.8**

Assume that {, ,..., } is a partition of S (see Definition 2.11) such that P() > 0, for i = 1, 2,..., k. Then for any event A:

**Theorem 2.9**

Bayes Rule, Assume that {, ,..., } is a partition of S (see Definition 2.11) such that P() > 0, for i = 1, 2,..., k. Then for any event A:

**Theorem 3.1**

For any discrete probability distribution, the following must be true:

1. , where the summation is over all values of y with nonzero probability.

**Theorem 3.2**

Let Y be a discrete random variable with probability function p(y) and g(Y)

be a real-valued function of Y. Then the expected value of g(Y ) is given by

**Theorem 3.3**

Let Y be a discrete random variable with probability function p(y) and c be a

constant. Then

**Theorem 3.4**

Let Y be a discrete random variable with probability function p(y), g(Y ) be a

function of Y, and c be a constant.

**Theorem 3.5**

Let Y be a discrete random variable with probability function p(y) and

be k functions of Y . Then

**Theorem 3.6**

Let Y be a discrete random variable with probability function p(y) and mean

**Theorem 3.7**

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

**Theorem 3.8**

If Y is a random variable with a geometric distribution

**Theorem 3.9**

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

**Theorem 3.10**

If Y is a random variable with a hypergeometric distribution

**Theorem 3.11**

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

**Theorem 3.12**

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

**Theorem 3.14**

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

**Theorem 4.1**

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

**Theorem 4.2**

Properties of a Density Function If f (y)is a density function for a continuous

random variable, then

**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]:

**Theorem 4.4**

Let g(Y) be a function of Y ; then the expected value of g(Y) is given by E [g(Y )] = 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:

.

**Theorem 4.6**

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

**Theorem 5.1**

If Y1 and Y2 are discrete random variables with joint probability function

p(y1, y2), then:

**Theorem 5.2**

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

**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 , for constants a, b, c, and d; and 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.

**Bonus**

**Immediate corollaries of the axioms**

**Permutation**

**Combination**

**Complement**

s ∈ A’ ⇐⇒ s ∈ S but s ∈/ A

A ⋂ A’ = Ø and A U A’ = S

**Distributive Laws**

**De Morgan’s Laws**

**Conditional Probability**