# Ryan Minneo Formula Sheets Project 1 and 2

# Table of Contents

[Project 1:](#bookmark=id.digl42fhs9ub)

[Project 2:](#bookmark=id.qsqww7v3kqf1)

[Chapter 3:](#_heading=h.p10yqvf2kxaw)

[Definitions](#_heading=h.r8iezs2cdjhf)

[Theorems](#_heading=h.lxom4oiqdi34)

[Chapter 4:](#_heading=h.v60i6jtmfxvi)

[Definitions](#_heading=h.adui39g249gf)

[Theorems](#_heading=h.3k77s3ezwozw)

[Chapter 5:](#_heading=h.qnfaenxukp4o)

[Definitions](#_heading=h.jnrty5mw8a)

[Theorems](#_heading=h.14a06ff6hssj)

# 

Project 1:

Definition 1.1:

𝑦=1𝑛𝑡=1𝑛𝑦𝑡

𝑠2=1𝑛−1𝑖 = 1𝑛𝑦𝑖−𝑦2

Definition 1.3: 𝑠 = 𝑠2

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

Theorem 2.1: With m elements a1, a2, …, am and n elements b1, b2,.., bn, it is possible to form mn = m x n pairs containing one element from each group.

Theorem 2.2: 𝑃𝑟𝑛=𝑛𝑛−1𝑛−2𝑛−𝑟+1 = 𝑛!𝑛−𝑟!

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 𝑖 = 1𝑘𝑛𝑖 = 𝑛

, is

𝑁 = ├𝑛1𝑛2𝑛𝑘𝑛=𝑛!𝑛1! 𝑛2! … 𝑛𝑘!

Theorem 2.4: The number of unordered subsets of size r chosen (without replacement) from n available objects is

(n r) = 𝐶𝑟𝑛=𝑃𝑟𝑛𝑟!=𝑛!𝑟!𝑛−𝑟!

Theorem 2.5: The Multiplicative Law of Probability The probability of the intersection of two events A and B is 𝑃𝐴𝐵 = 𝑃𝐴𝑃𝐵𝐴 = 𝑃𝐵𝑃𝐴𝐵

If A and B are independent, then 𝑃𝐴𝐵=𝑃𝐴𝑃𝐵

Theorem 2.6: The Additive Law of Probability The probability of the union of two events A and B is

𝑃𝐴𝐵 = 𝑃𝐴+𝑃𝐵−𝑃𝐴𝐵

If A and B are mutually exclusive events, 𝑃𝐴𝐵 = 0

and

𝑃𝐴𝐵 = 𝑃𝐴 + 𝑃𝐵

Theorem 2.7: If A is an event, then

𝑃𝐴 = 1−𝑃𝐴

Theorem 2.8: Assume that {B1, B2, … , Bk} is a partition of S such that P(Bi) > 0, for I = 1,2,…,k. Then for any event A

𝑃𝐴=𝑖=1𝑘𝑃𝐴𝐵𝑖𝑃𝐵𝑖.

Theorem 2.9: Bayes’ Rule Assume that {B1,B2,…,Bk} is a partition of S such that P(Bi) > 0, for I = 1,2,…,k. Then

𝑃𝐵𝑗𝐴=𝑃𝐴𝐵𝑗𝑃𝐵𝑗𝑖=1𝑘𝑃𝐴𝐵𝑖𝑃𝐵𝑖

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

𝑃𝐴1  𝐴2  𝐴3 =𝑖 = 1𝑃𝐴𝑖

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 𝑃𝑟𝑛

.

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 𝐶𝑟𝑛

or (n r). PS I physically cannot for the life of me figure out how to put it properly for that (n r) notation.

Definition 2.9: The conditional probability of an event A, given that an event B has occurred is equal to

𝑃𝐴  𝐵= 𝑃𝐴  𝐵𝑃𝐵

,

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

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

𝑆 = 𝐵1  𝐵2  …  𝐵𝑘.

𝐵𝑖  𝐵𝑗 = ∅,

for i

j

Then the collection of sets {B1, B2, … , Bk} 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 (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.

Theorem 3.1: For any discrete probability distribution, the following must be true:

0≤𝑝𝑦 ≤1

for all y.

𝑦𝑝𝑦 = 1

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

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

𝐸𝑔𝑌 = 𝑎𝑙𝑙 𝑦 𝑔𝑦𝑝𝑦

Theorem 3.3: Let Y be a discrete random variable with probability function p(y), g(Y) be a function of Y, and c be a constant. Then

𝐸𝑐𝑔𝑌 = 𝑐𝐸𝑔𝑌

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. Then E[cg(Y)] = cE[g(Y)].

Theorem 3.5: Let Y be a discrete random variable with probablity function p(y) and g1(Y), g2(Y), … , gk(Y) be k functions of Y. Then

𝐸𝑔1𝑌 + 𝑔2𝑌+…+𝑔𝑘𝑌 = 𝐸𝑔1𝑌+𝐸𝑔2𝑌+…+𝐸𝑔𝑘𝑌.

Theorem 3.6: Let Y be a discrete random variable with probability function p(y) and mean E(Y) = Mu; then

𝑉𝑌 = 𝜎2=𝐸𝑌−𝜇2=𝐸𝑌2−𝜇2

Theorem 3.7: Let Y be a binomial random variable based on n trials and success probability p. Then

𝜇=𝐸𝑌 = 𝑛𝑝

and 𝜎2=𝑉𝑌 = 𝑛𝑝𝑞.

Theorem 3.8: If Y is a random variable with a geometric distribution,

𝜇=𝐸𝑌=1𝑝

and 𝜎2=𝑉𝑌 = 1−𝑝𝑝2

Definition 3.1: A random variable Y is said to be discrete if it can assume only a finite or countably infinite 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 will 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: 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

𝐸𝑌 = 𝑦𝑦𝑝𝑦

Definition 3.5: 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,

𝑉𝑌 = 𝐸𝑌−𝜇2.

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

Definition 3.6: A binomial experiment possesses the following properties:

The experiment consists of a fixed number, n, of identical trials.

Each trial results in one of two outcomes: success, S, or failure, F.

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).

The trials are independent.

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

𝑝𝑦 = 𝑛 𝑦𝑝𝑦𝑞𝑛−𝑦,     𝑦 = 0, 1, 2, … , 𝑛 𝑎𝑛𝑑 0 ≤ 𝑝 ≤ 1.

Definition 3.8: A random variable Y is said to have a geometric probability distribution if and only if

𝑝𝑦 = 𝑞𝑦−1𝑝,     𝑦 =1, 2, 3, … ,  0 ≤ 𝑝 ≤ 1.

Project 2

# Chapter 3:

## Definitions:

### 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

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

### 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 E(Y k) and is denoted by µk.

### 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 µk.

### Definition 3.14:

The moment-generating function m(t) for a random variable Y is defined to be m(t) = E(e^tY). 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 3.15:

Let Y be an integer-valued random variable for which P(Y = i) = pi, 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:

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

where k is a positive integer.

## Theorems:

### 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

and

By definition,

Notice that the first term in this sum is equal to 0 (when y = 0), and, hence,

### 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 µk.

### 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:

Second favorite theorem. I like the name.

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

# Chapter 4:

## Definitions:

### 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 < .

### 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.

### Definition 4.4:

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

### 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 parametersof 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

### Definition 4.9:

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

### Definition 4.10:

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.

### Definition 4.11:

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

### Definition 4.12:

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

, where

### Definition 4.13:

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

The kth moment about the mean, or the kth central moment, is given by

### Definition 4.14:

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.

### 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

## Theorems:

### Theorem 4.1:

Properties of a Distribution Function1 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).]

### Theorem 4.2:

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] is

### 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.

### Theorem 4.5:

Let c be a constant and let g(Y), g1(Y), g2(Y), …, gk (Y) be functions of a continuous random variable Y. Then the following results hold:

1. E(c) = c.

2. E[cg(Y)] = cE[g(Y)].

3. E[g1(Y)+g2(Y)+· · ·+gk (Y)] = E[g1(Y)]+E[g2(Y)]+· · ·+E[gk (Y)].

### Theorem 4.6:

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

### Theorem 4.7:

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

### Theorem 4.8:

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

### Theorem 4.9:

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

### Theorem 4.10:

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

### Theorem 4.11:

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

### 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:

My favorite theorem. I like the name. ***Tchebysheff’s Theorem***

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

# Chapter 5:

## Definitions:

### 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 variablesY1 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:

### 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

provided that

### 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 Y2 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

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.

## Theorems:

### 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.

### 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.

### 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 ≤ ≤ b and c ≤ ≤ 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.