Introduction to Statistics Notes 2020 − 2021

Dr. M. P. M. M. M^{c}Loughlin

based on the notes that I have created since I first taught such a course in 1987 AD
(not BC it just seems to be the case).

Math 140

Kutztown University of Pennsylvania

2021/01/27

1. Introduction

Mathematics is predicated on logic (for the rules of inference) and on Set Theory. From those foundations we build arithmetic, geometry, real analysis, abstract algebra, the theory of differential equations and difference equations, probability theory, point-set topology, general topology, numerical analysis, number theory, counting theory, combinatorics, complex variables, dynamical systems, graph theory, game theory, statistics, coding theory, cryptology, etc.

The axiomatic approach to sets, logic, analysis, etc. is one of the most impressive accomplishments of modern mathematics. Concepts which were vague or indistinct took on the property of clarity. Precise meanings replaced quasi-definitions. Adequate axioms established the foundation of modern mathematics for they provided clear, unambiguous, and understandable premises for theories which before were sound but people *didn't know why they were sound*. This is not to say that the original work of mathematicians was always without its problems; but, it is to say that mathematicians should be able to elucidate what is being assumed and then what ensues from those assumptions.

For non-mathematicians in this class, we will consider the axioms in this area of mathematics and learn how to use them to make rigourous our understanding of 'chance,' 'odds,' 'games,' and probability.

The mathematics most Americans in Secondary school or before studies includes elementary arithmetic and counting theory, geometry, analysis, algebra, trigonometry, functions, and a bit of statistics.

We are begin by focusing on the branch of mathematics called Probability Theory. It forms the basis for Statistics and without it Statistical Theory cannot be meaningful. Probability Theory is essentially a twist on Real Analysis, Logic, Set Theory, and Counting Theory.

1.1. **Review of Definitions & Notation.** We need to review some mathematical notation (symbols).

From logic we use the following:

Let us have a well-defined domain of definition.

Let P and Q be statements.

Notation 1.1. Not P is written as $\neg P$.

Notation 1.2. P and Q is written as $P \wedge Q$.

Notation 1.3. P or Q is written as $P \vee Q$.

Notation 1.4. P implies Q is written as $P \Longrightarrow Q$; $P \rightarrow Q$; or, $P \Rightarrow Q$.

Notation 1.5. P if and only if (iff) Q is written as $P \Leftrightarrow Q$ or $P \Longleftrightarrow Q$. It means $P \Longrightarrow Q \wedge Q \Longrightarrow P$.

From Set Theory we use the following:

Let U be a well-defined universe. Let A be a set and let m be an element of A.

Notation 1.6. The element m is in A is denoted as $m \in A$.

Notation 1.7. The element p is not in A is denoted as $p \notin A$.

Let U be a well-defined universe. Let A be a set and The set of every element that is not in A is called the complement of A.

Notation 1.8. The complement set of A is denoted as A^C.

Let U be a well-defined universe. Let A be a set and let B be a set where every element in A is also in B. Then the set A is called a subset of B.

Notation 1.9. The set A is a subset of B is denoted as $A \subseteq B$.

Notation 1.10. The set B is a superset of A is denoted as $A \subseteq B$.
It is also written as $B \supseteq A$.

Let U be a well-defined universe. Let A be a set and let B be a set where every element in A is also in B and moreover there is at least one element in B that is not in A. Then the set A is called a proper subset of B.

Notation 1.11. The set A is a proper subset of B is denoted as $A \subset B$.

Let U be a well-defined universe. Let C be a set and let D be a set. The union of C and D is the set of all elements that are either in C or in D.

Notation 1.12. The union of C and D is denoted as $C \cup D$.

Let U be a well-defined universe. Let C be a set and let D be a set. The intersection of C and D is the set of all elements that are in both C and D.

Notation 1.13. The intersection of C and D is denoted as $C \cap D$.

Let U be a well-defined universe. Let A be a set and let B be a set where there are no elements of A in common with B.

Notation 1.14. The set A is disjoint from B is denoted as $A \cap B = \varnothing$.

Let U be a well-defined universe. Let C be a set and let D be a set. The set of elements of C that are not in D also is called the difference set of C minus D.

Notation 1.15. The difference set of C minus D is denoted as $C - D$.

Notation 1.16. $\exists$ means 'there exists.' It is called the existential quantifier.

Example 1.17. Let U be the well-defined universe $\{1, 2, 3\}$; A be the set $\{1, 2\}$; $W = \{1\}$; and, M be the set $\{1, 3\}$. It is true that $\exists\ p$ in A that is also in M (namely 1).

When two sets have at least one point in common they are not disjoint and their intersection is non-empty. So, from the example notice $A \cap M \neq \varnothing$.

Notation 1.18. $\forall$ means 'for all;' 'every;' or, 'each,' It is called the universal quantifier.

When all the elements of one set is an element of another that is what is meant by subsethood ("$\subseteq$").
So, from the example notice
$A \subseteq M$ is false since $2 \in A \wedge 2 \notin M$. This means $\forall\ x \in A, x \in M$ is false.

Notice $W \subset A$ so it is true that $\forall\ x \in W, x \in A$ and $\exists\ k \in A, k \notin W$.

Also, some symbols from Geometry and Analysis are of assistance:
There are some special standard sets and symbols for them to denote sets that we use often. The sets under discussion are formulated from the real line (are either points of the line or are generalisations of the line. The real line exists from the first axiom of Euclidean Geometry, namely there exists a point, a line, and a plane. So, for this discussion let U be the real line.

Definition 1.19. $\mathbb{R} = \{x | x \text{ is a point on the line }\}$.

Your knowledge of high school geometry will, no doubt, be of use in making concrete these abstract ideas that follow.
One of the most basic of sets is called the **natural numbers**. It has been with us since antiquity, and we will denote it as $\mathbb{N} = \{1, 2, 3, 4, ...\}$ where the set never ends and includes all the whole or counting numbers that the student learnt in kindergarten or before.
We shall denote the set of natural numbers along with zero as the set
$\mathbb{N}^* = \{0, 1, 2, 3, 4, ...\}$.
We shall denote the set $\{1\}$ as $\mathbb{N}_1$, the set $\{1, 2\}$ as $\mathbb{N}_2$, the set $\{1,\ 2,\ 3\}$ as $\mathbb{N}_3$, and so forth so that $\mathbb{N}_k = \{1, 2, 3, 4, ..., (k-1), k\}$.
All of these are referred to as *initial segments* of the natural numbers. This definition is known as a **recursive definition** since we are inductively defining a

myriad of sets at once; the three dots signify that the enumeration of the elements continues.[1]

Continuing this train of thought we shall denote the set $\{0\}$ as $\mathbb{N}_0^*$, the set $\{0,1\}$ as $\mathbb{N}_1^*$the set $\{0,\ 1,\ 2\}$ as $\mathbb{N}_2^*$, etc. so that $\mathbb{N}_p^* = \{0,1,2,3,...,(p-1),p\}$.

Definition 1.20. Let U be the reals.
(1) $\mathbb{N} = \{1,2,3,...,(p-1),p,(p+1),\ldots\}$.
(2) $\mathbb{N}^* = \{0,1,2,3,...,(p-1),p,(p+1),\ldots\}$.
(3) $\mathbb{N}_p = \{1,2,3,...,(p-1),p\}$.
(4) $\mathbb{N}_p^* = \{0,1,2,3,...,(p-1),p\}$.

Another of the most basic of sets is called the **integers**. It has been with us quite a long time (the people of India invented the symbol of zero and were really the first to use it and negative numbers (in fact the number system that we use is of course the Hindu-Arabic number system since the Hindus created it, the Arabs adopted it and brought it west); another interesting fact is the Mayans also invented a zero independent of the Hindus).
We will denote the set of integers as $\mathbb{Z}$ such that
$\mathbb{Z} = \{0,1,-1,2,-2,3,-3,4,-4,...\}$.

Definition 1.21. Let U be the reals.
$\mathbb{Z} = \{0,1,-1,2,-2,3,-3,\ldots\}$.

Generalising from the integers, we have the **rational numbers**. The rationals are denoted as $\mathbb{Q}$ such that $\mathbb{Q} = \{x|x = \frac{m}{n}$ where $m \in \mathbb{Z}, n \in \mathbb{Z}, n \neq 0\}$. This statement is read as the rationals are the set consisting of elements x such that x is equal to m divided by n where m is an integer, n is an integer, and n is not zero. Thus, the symbol "|" in this context means **such that**. Indeed, as I type this opus I am getting very tired of writing the words, "such that," (I suppose it is an occupational hazard, I think we mathematicians are a lazy lot so we have invented many symbols to create a short hand to ease the amount of words necessary to communicate). A more general symbol for such that is, "$\ni$" or "϶" and will be used liberally from this point onward. Typically it is not used in the notation inside the braces for a set only as a free-standing symbol. However, it is not incorrect to use it. Therefore, it is technically correct to write:

Definition 1.22. Let U be the reals.
$\mathbb{Q} = \{x \ni x = \frac{a}{b}$ where $a \in \mathbb{Z}, b \in \mathbb{Z}, b \neq 0\}$

Generalising from the rational numbers, we have the **real numbers**. However, this is axiomatically executable, but practically most difficult to do in a basic introduction to sets. Therefore, we shall consider the set of reals from a geometric standpoint. The set of real numbers are denoted as , $\mathbb{R}$ such that $\mathbb{R} = \{x|x$ is a

[1]The three dot symbol (the ellipsis) is most problematic. Many students think that the ellipsis establishes a pattern. It does not. Consider . is not 3.14. It is not 3.141592. It is 3.14159265359 . . . (the decimal does not repeat or have a pattern); thus, the three dot symbol means on and on but not necessarily in a pattern.

point on the line}. One could also define the reals from a sequential (or decimal) perspective by defining the reals to be $\mathbb{R} = \{x|x$ is a number where x is an integer followed by a decimal and then a sequence of digits where each digit belongs to the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Yes, this is a rather cumbersome definition, but one can prove that the sequential definition and the geometric definitions are equivalent.

Definition 1.23. Let U be well defined.
$\mathbb{R} = \{x|x$ is a number where x is an integer followed by a decimal and then a sequence of digits where each digit belongs to the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}\}$.

Note, gentle reader that I skipped another standard set; that is because in a basic introduction to sets it is oft easier to 'jump' up to the reals, then return back to describe another set. That set is, of course, the irrationals. By the very nature of its name one can understand that it is composed of elements that are not rational. But, recall our discussion of the need for defining a domain of discourse or a universe. To say what something is not presupposes that everything has been specified! Putting it another way saying that the irrationals are numbers that are not rational is wrong since is a number that is not rational; but, it is also not irrational. Thus, the set of **irrational numbers** is denoted as $\mathbb{I}$ such that $\mathbb{I} = \{x : x \in \mathbb{R}$ and $x \ni \mathbb{Q}\}$. So, the irrationals are the set of all real numbers that are *not* rational. Note that this is a definition of something such that it is defined by what it is not (the complement). This definition by negation is oft quite useful; but one must understand what the first thing is (a real number) and the second thing is (a rational number) in order to understand what the third thing is (an irrational number) by way of what it isn't.

Definition 1.24. Let U be the reals.
$\mathbb{I} = \{x|x \in \mathbb{R}$ and $x \notin \mathbb{Q}\}$.

Notation 1.25. Let $U = \mathbb{R}$ be the universe. Let $a \in \mathbb{R}$ and $b \in \mathbb{R}$ so that $a < b$ The set $[a, b] = \{x|a \leq x \leq b\}$. The set is called an interval.

FIGURE 1. An Interval

Notation 1.26. Let $U = \mathbb{R}$ be the universe. Let $a \in \mathbb{R}$ and $b \in \mathbb{R}$ so that $a < b$ The set $(a, b) = \{x|a < x < b\}$. The set is called a segment.

FIGURE 2. A Segment

Notation 1.27. Let $U = \mathbb{R}$ be the universe. Let $a \in \mathbb{R}$ and $b \in \mathbb{R}$ so that $a < b$ The set $[a, b) = \{x|a \leq x < b\}$ The set is called a half-segment or half-interval.

FIGURE 3. A Half-Interval

Notation 1.28. Let $U = \mathbb{R}$ be the universe. Let $a \in \mathbb{R}$ and $b \in \mathbb{R}$ so that $a < b$ The set $(a, b] = \{x|a < x \leq b\}$The set is called a half-segment or half-interval.

FIGURE 4. Another Half-Interval

One other way to write $\mathbb{R}$ is by writing $(-\infty, \infty)$ which is segment notation. The symbols ∞ and $-\infty$ are **not numbers**, they simply represent that the line goes on ad infinitum to the left in the case of $-\infty$ and to the right in the case of ∞.

From Logic and from Set Theory that every set, A, is a subset of some well-defined universe, U. So, for probability theory we rename the universe as a **sample space** [the space from whence a sample may be chosen]; an arbitrary set is called an **event**; and an element of that set an **outcome**.

Definition 1.29. Let U be a well-defined universe. A **sample space** is a subset of U. The elements of the sample space are outcomes.

Definition 1.30. t U be a well-defined universe. Let S be a well-defined sample space. An **event** is a subset of S.

Definition 1.31. Let U be a well-defined universe. Let S be a well-defined sample space. Let A be an event. An **outcome** of A is an element of A.

What is 'tricky' in learning basic probability is you are working with sets; functions; and real numbers. So, before understanding these definitions, lemmas, theorems, corollaries, and claims think about what is precisely being said, Is it a numeric, set-theoretic, or both? One also has to be keen to read every word in a problem and understand the rigourous connotation of the word; else, mistakes can be made (horrendous mistakes perhaps).

Also, the wording changes (ever so slightly) or a term or two is added:

Definition 1.32. Let S be a well-defined sample space. When referencing sets, we say the set A is disjoint from B is denoted as $A \cap B = \varnothing$. In probability theory when referencing events, when the event A is disjoint from B we say two events are **mutually exclusive**. So, when the event A is mutually exclusive from B it is the case that $A \cap B = \varnothing$.

Logic	**Set Theory**	**Probability Theory**
Domain of definition	Universe, U	Sample space, S
	example $U = \{1, 2, 3, 4, 5\}$	example $S = \{1, 2, 3, 4, 5\}$
Statements p, q, and r	Sets A, B, and C.	Events E, F, and G
	example $A = \{1, 2\}$ $B = \{2, 3, 4\}$ $C = \{1, 4\}$	example $E = \{1, 2\}$ $F = \{2, 3, 4\}$ $G = \{1, 4\}$
Simple statements	Elements example	Outcomes example
true statements like 1 is in A	$1 \in A$ and $1 \notin B$	$\{2, 4\} \subseteq F$
false statements like 3 is in A	$1 \in B$, $\{2\} \subseteq C$	$\{2, 4\} \subseteq E$, $4 \in E$

Definition 1.33. Let S be a well-defined sample space. The events A and B are **exhaustive** when $A \cup B = S$.

2. The Axioms of Probability

[2]

Let S denote the sample space, E, E_i, F, etc. events and the notation $Pr(\cdot)$ the probability of whatever.

Axiom 2.1. (The Space Axiom) S is the space $\Rightarrow Pr(S) = 1$.

Axiom 2.2. (The Event Axiom) E is an event $\Rightarrow 0 \leq Pr(E) \leq 1$.

Axiom 2.3. (The Collection of M. E. E.[3] Axiom) Let I be an index set. The collection $\{E_i\}_{i \in I}$ being mutually exclusive $\Rightarrow \Pr(\bigcup\limits_{i \in I} E_i) = \sum\limits_{i \in I} \Pr(E_i)$.[4]

Notation 2.4. Remember summation notation $\sum_{i \in I} f(i)$ is defined as plugging in the i from the set I into the expression $f(i)$ and adding it all up.

[2]The Kolmogorov Axioms of Probability are named after the creator of the axioms: the Russian mathematician Kolmogorov. These are so named since he created them as an answer to one of the famous Hilbert 20^{th} century 10 questions. The axioms of probability are one of the shortest lists of axioms I can recall for an area of mathematics.

[3]M. E. E. : Mutually Exclusive Events

[4]We will explain in detail this axiom in a manner that will help us understand the idea and not focus on the hideous 'chicken scratch.'

Example: $\sum_{i=1}^{6}(\frac{i^2}{i+1}) = \frac{1}{2} + \frac{4}{3} + \frac{9}{4} + \frac{16}{5} + \frac{25}{6} + \frac{36}{7}$

Notation 2.5. Also product notation exists $\prod_{j \in J} f(j)$ is defined as plugging in the j from the set J into the expression $f(j)$ multiplying the result.

Example: $\prod_{k=2}^{5}(\frac{k}{7}) = \frac{2}{7} \cdot \frac{3}{7} \cdot \frac{4}{7} \cdot \frac{5}{7} = \frac{1120}{2401}$

2.1. Some Theorems, Lemmas, or Corollaries of Probability.

Lemma 2.6. Let S be a well-defined sample space; whilst E and F are events.

(1) E is an event $\Rightarrow E^c$ is an event.
(2) E and F are events $\Rightarrow E \cap F$ is an event.
(3) E and F are events $\Rightarrow E \cup F$ is an event.
(4) E and F are events $\Rightarrow E - F$ is an event.

Corollary 2.7. (The Null Corollary) Let S be a well-defined sample space whilst E $= \varnothing$. It is the case that $Pr(E) = 0$.

Corollary 2.8. (The Complement Corollary) E is an event $\Rightarrow Pr(E^c) = 1 - Pr(E)$.

Corollary 2.9. (The Subset Corollary) E and F are events $\ni E \subseteq F \Rightarrow Pr(E) \leq Pr(F)$.

Corollary 2.10. (The Spanning Corollary) Let I be an index set. The collection $\{E_i\}_{i \in I}$ being mutually exclusive and exhaustive $\Rightarrow \text{Pr}(\bigcup_{i \in I} E_i) = \sum_{i \in I} \text{Pr}(E_i) = 1$.

Corollary 2.11. (The Same Event Corollary) E and F are events $\ni E = F \Rightarrow Pr(E) = Pr(F)$.

Theorem 2.12. (The Union Theorem) E and F are events.
$Pr(E \cup F) = Pr(E) + Pr(F) - Pr(E \cap F)$.

3. Naïve or Intuitive Idea of Probability

Informal understanding of probability pre-dated the formal theory of probability. In this section we will simply look at the ideas that first were encountered before the rigour of the axioms. Many of the concepts seem to have arisen from gaming.

3.1. **Finite Sample Space.** Recall please:

Definition 3.1. Let U be a well-defined universe. A **sample space** is a subset of U.

Definition 3.2. Let U be a well-defined universe. Let S be a well-defined sample space. An **event** is a subset of S.

Definition 3.3. Let U be a well-defined universe. Let S be a well-defined sample space. Let A be an event. An **outcome** is an element of A.

The intuitive idea of probability of an event, E, from a sample space, S, is $\dfrac{|E|}{|S|}$ where S is **finite**.[5]

Definition 3.4. Let S be a well-defined sample space. An outcome is defined as **fair** *iff* the probability that an outcome in the space is picked is equal to the probability that any other outcome in the space is picked (another way to express this is that all outcomes are *equiprobable* of being selected).

Notation 3.5. Let S be a well-defined sample space. Let E be an event. The probability of E occurring is denoted as **Pr(E)** or **P(E)**.

The result of the Kolmogorov Axioms of Probability, these definitions, corollaries, theorems, and lemmas is the probability of an event is a real number in the interval $[0, 1]$ – probability can never be less than zero or greater than 1.

When the sample space is finite one can easily compute the probability of an event by taking the number of elements of the event and dividing by the number of elements in the sample space. Such a method of calculating probabilities works for many practical elementary problems. So, let us try some of the following exercises to illustrate such.

[5]The meaning of this is $|S| < \aleph_0$ which is quite an elegant concept. A set, A is **finite** if and only if $\exists \quad g : A \longrightarrow \mathbb{N}_k \ni k \in \mathbb{N} \ni g$ is bijective or $A = \varnothing$. A set, A is **infinite** if and only if it is not finite. The cardinal number $\aleph_0$ is defined as the cardinality of $\mathbb{N}$.
Also, a subset of a finite set is also finite.

3.2. **Exercises.**

Exercise 3.6. Let $U = \mathbb{R}$ and $S = \{1, 2, 3, 4, 5, 6, 7\}$.

a) Let $A = \{1, 3, 5, 7\}$. Find $Pr(A)$.
b) Let $B = \{2, 4, 6\}$. Find $Pr(B)$.
c) Let $C = \{x|x < 3\}$. Find $Pr(C)$.
d) Let $D = \{x|2 \leq x < 3\}$. Find $Pr(D)$.
e) Let $E = \{x|2 < x < 3\}$. Find $Pr(E)$.
f) Let $F = A \cap B$. Find $Pr(F)$.
g) Let $G = A \cup B$. Find $Pr(G)$.
h) Let $H = A \cap C$. Find $Pr(H)$.

Exercise 3.7. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing exactly two heads amongst the tosses.

Exercise 3.8. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing at least two heads.

Exercise 3.9. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing of tossing no heads.

Exercise 3.10. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing at most two heads.

Exercise 3.11. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out six balls from the urn. Find the probability of choosing exactly 4 red balls.

Exercise 3.12. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out six balls from the urn. Find the probability of choosing exactly 4 red balls amongst the six balls chosen.

Exercise 3.13. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out six balls from the urn. Find the probability of choosing exactly 2 red balls, exactly 2 white balls, and exactly 2 blue balls.

Exercise 3.14. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball from the urn; then we draw another ball from the urn; then we draw another ball from the urn; then we draw another ball from the urn; then we draw another ball from the urn; and, finally we draw another ball from the urn. Find the probability of choosing exactly 2 red balls, exactly 2 white balls, and exactly 2 blue balls.

Exercise 3.15. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out two balls from the urn. Find the probability of choosing 2 blue balls.

Exercise 3.16. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball. We replace the ball shake up the urn and then draw a ball from the urn. Find the probability of choosing a blue ball first and a white ball second.

Exercise 3.17. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball. We replace the ball shake up the urn and then draw a ball from the urn. Find the probability of choosing a blue ball and a white ball.

Exercise 3.18. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball. We replace the ball shake up the urn and then draw a ball from the urn. Find the probability of choosing at least one red ball.

With dice (unless otherwise stated) we toss a pair of dice and sum up the number of dots of the up-turned faces.

Exercise 3.19. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a sum of seven or an eleven.

Exercise 3.20. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a seven or an eleven.

Exercise 3.21. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a sum of five or better.

Exercise 3.22. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a product of the upturned faces of a twelve.

Exercise 3.23. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing exactly two heads and then twice exactly no heads.

Exercise 3.24. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing two heads.

Exercise 3.25. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing at least two heads.

Exercise 3.26. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing no heads.

Exercise 3.27. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing at most two heads.

Exercise 3.28. An urn contains 5 red balls; 4 blue balls; and 2 white balls. Find the probability of picking a red ball from the urn.

Exercise 3.29. An urn contains 5 red balls; 4 blue balls; and 2 white balls. Find the probability of picking a ball that is not blue from the urn.

Exercise 3.30. An urn contains 5 red balls; 4 blue balls; and 2 white balls. Find the probability of picking a ball that is not blue from the urn.

Exercise 3.31. An urn contains 5 red balls; 4 blue balls; and 2 white balls. Find the probability of picking two balls that are both blue from the urn.

Exercise 3.32. An urn contains 5 red balls; 4 blue balls; and 2 white balls. Find the probability of picking two balls that are both blue from the urn.

Exercise 3.33. Let $S = \mathbb{N}_{15}$ Let $A = \{x|x = 2p, p \in S\}$. Find the probability of the event A.

Exercise 3.34. Let $S = \mathbb{N}_{15}$ Let $B = \{x|x \text{ is prime }\}$. Find the probability of the event B.

3.3. **Non-finite Sample Space.** An intuitive idea of probability of an event, E, from a sample space, S, is NOT $\frac{|E|}{|S|}$ because the space we have is **not** finite.

Example 3.35. Let $U = \mathbb{R}$ and $S = \mathbb{N}$ whilst $E = \{x|x = 2 \cdot a, a \in \mathbb{N}\}$.
$Pr(E) = \frac{|E|}{|S|}$ fails; the concept of 'infinity divided by infinity.'

What is intuitively clear is the probability of E is $\frac{1}{2}$ since half the naturals are odd and half are even. Therefore, when the sample space is not finite we have to approach solutions in a different way.

Example 3.36. Let $U = S = \mathbb{R}$ and $E = \mathbb{N}$. $Pr(E) = \frac{|E|}{|S|}$ fails.

What is intuitively clear is the probability of E is 0 since so few of the reals are natural. So, notice the sample space is not finite - again problems. We can compute some probabilities when the sample space is not finite using methods which are familiar to you all and are rigourous, clear, and understandable. For a one-dimensional sample space we use **length**. In order to understand what is being requested; I advise draw the graph. For a two-dimensional sample space we use **area**. In order to understand what is being requested; I also advise draw the graph. For a three-dimensional sample space we use **volume** (which we will not do in this class - drawing the graph is a bother).

A key is to realise that an outcome is **in** an event (an element belongs to a set) and the probability of the event is a real **number** restricted to the interval of zero to one, inclusive.

Example 3.37. Suppose $U = \mathbb{R}$ and $S = [0, 10]$. Let $E = [0, 5]$. Let $F = [0, 5)$. Let $H = (0, 4)$. Let $J = \{x|x \in S \wedge \exists \quad m \in \mathbb{N} \ni x = 2 \cdot m\}$.

$$Pr(J) = \frac{\ell(J)}{\ell(S)} = \frac{0}{10}$$

$$Pr(F) = \frac{\ell(F)}{\ell(S)} = \frac{5}{10}$$

$$Pr(E) = \frac{\ell(E)}{\ell(S)} = \frac{5}{10}$$

$$Pr(H) = \frac{\ell(H)}{\ell(S)} = \frac{4}{10}$$

$$Pr(E \cap F^c) = \frac{\ell(E \cap F^c)}{\ell(S)} = \frac{0}{10}$$

Example 3.38. Suppose $U = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ and let $S = \{(x, y)| \ \ x^2 + y^2 \leq 4\}$.

Notice the sample space (S) is a disc with center $(0, 0)$ and a radius of 2 (most of you please recall $S = \{(x, y)| \ \ x^2 + y^2 = 4\}$ is a circle of radius 2 centred at the origin. Let $A = \{(x, y)|x > 0, y > 0\}$. Note the probability of A is $\frac{1}{4}$ since it is a quarter of the disc (draw the disc and notice a quarter of it is shaded).

Let $B = \{(x, y)|x \geq 0, y \geq 0\}$. Note the probability of B is $\frac{1}{4}$ since it is still just a quarter of the disc (draw the disc and notice a quarter of it is shaded).

Let $C = \{(x, y)|x \geq 0, y > 0\}$. Note the probability of C is $\frac{1}{4}$ since it is still just a quarter of the disc (draw the disc and notice a quarter of it is shaded).

Let $E = \{(x, y)|x > 0\}$. Note the probability of E is $\frac{1}{2}$ since it is half of the disc (draw the disc and notice half of it is shaded).

Let $F = \{(x, y)|y \geq 0\}$. Note the probability of F is $\frac{1}{2}$ since it is half of the disc (draw the disc and notice half of it is shaded).

Let $G = \{(x, y)|x = 0\}$. Note the probability of G is 0 since it is a line (having zero area) of the disc (draw the disc and notice no shading).

This illustrates that the event, G is non-empty but it has probability 0. Let $H = \{x|x = 0\}$. Note the probability of G is 0 since it is a line (having zero area) of the disc (draw the disc and notice no shading).

This illustrates that the event, G is non-empty but it has probability 0.

Let $J = \{(x, y)| \ \ x^2 + y^2 > 16\}$ Note the probability of J is 0 since it is actually $\varnothing$ (draw the disc and notice J does not intersect S). What should (I hope) become clear is that any time one works with $\varnothing$ the probability of $\varnothing$ is zero.
Let $M = \{(x, y)| \ \ x^2 + y^2 = 4\}$. Note the probability of M is 0.
Such is also a consequence of the definitions and axioms of probability; recall:

Definition 3.39. Let U be a well-defined universe. A **sample space** is a subset of U.

Definition 3.40. Let U be a well-defined universe. Let S be a well-defined sample space. An **event** is a subset of S.

Remember (and notice in all your answers to exercises before this and after how the solutions do not violate the axioms):

Axiom 3.41. (The Space Axiom) S is the space $\Rightarrow Pr(S) = 1$.

Axiom 3.42. (The Event Axiom) E is an event $\Rightarrow 0 \leq Pr(E) \leq 1$.

Axiom 3.43. (The Collection of M. E. E.[6] Axiom) Let I be an index set. The collection $\{E_i\}_{i \in I}$ being mutually exclusive $\Rightarrow \Pr(\bigcup_{i \in I} E_i) = \sum_{i \in I} \Pr(E_i)$.

Corollary 3.44. (The Complement Corollary) E is an event $\Rightarrow Pr(E^c) = 1 - Pr(E)$.

Notice by the space axiom, the probability of the space is 1. $S^C = \varnothing$. So, $Pr(\varnothing) = 1 - Pr(S)$. Since $Pr(S) = 1$ we get $Pr(\varnothing) = 1 - 1 = 0$.

3.4. **Exercises.**

Exercise 3.45. Let $U = \mathbb{R}$ and $S = [0, 8]$. Let $E = [1, 3]$. Let $F = [0, 3)$. Let $H = (1, 3)$. Let $J = \{x | x \in S \wedge \exists \quad m \in \mathbb{N} \ni x = 2 \cdot m\}$.

A. Find $Pr(E)$. B. Let $E = [1, 3]$ C. Find $Pr(E^c)$. D. Find $Pr(F)$.

E. Find $Pr(H)$. F. Find $Pr(J)$. G. Let $G = E - F$ Find $Pr(G)$.

Exercise 3.46. Let $U = \mathbb{R}$ and $S = [1, \pi)$. Let $E = [2, 3]$. Let $F = (2, 3)$. Let $G = (3, 3)$. Let $H = (e, \pi)$. Let $J = (1, \frac{\pi}{2})$. Let $K = (1, \frac{\pi - 1}{2}]$. Let $M = [3, 3]$.

A. Find $Pr(E)$. B. Find $Pr(E^c)$. C. Find $Pr(F)$. D. Find $Pr(G)$.

E. Find $Pr(H)$. F. Find $Pr(J)$. G. Find $Pr(M)$. H. Find $Pr(K)$.

Exercise 3.47. Let $U = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ and let $S = \{(x, y) | \ (x - 1)^2 + (y + 2)^2 = 9\}$.

Think to yourself what is the sample space geometrically.
A. Let $K = \{(1, 1)\}$. Find $Pr(K)$. B. Let $J = \{(1, -2)\}$. Find $Pr(J)$.

C. Let C be the set of all points in S such that $x > 0$. Find $Pr(C)$.

D. Let D be the set of all points in S such that $x \geqslant 0 \wedge y \geqslant 0$. Find $Pr(D)$.

E. Let E be the set of all points in S such that $(x - 1)^2 + (y + 2)^2 = 9$. Find $Pr(E)$.

F. Let $F = \left\{(x, y) | \ (x - 1)^2 + (y - 4)^2 < \frac{1}{2}\right\} \bigcap S$. Find $Pr(F)$.

G. Let $G = \{(x, y) | \ y = -2x\} \bigcap S$. Find $Pr(G)$.

H. Let $J = \{(x, y) | \ y = -2x\} \bigcap S$. Find $Pr(H)$.

[6] M. E. E. : Mutually Exclusive Events

J. Let $J = \{(x, y)|\ \ y \leq -2x\} \bigcap S$. Find $Pr(J)$.

Exercise 3.48. Let $U = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ and let $S = \{(x, y)|\ \ (x-1)^2 + (y+2)^2 \leqslant 9\}$. Think to yourself what is the sample space geometrically.

A. Let $K = \{(2, 0)\}$. Find $Pr(K)$. B. Let $J = \{(-1, 2)\}$. Find $Pr(J)$.

C. Let C be the set of all points of such that $x > 0$. Find $Pr(C)$.

D. Let D be the set of all points such that $x \geqslant 0 \wedge y \geqslant 0$. Find $Pr(D)$.

E. Let E be the set of all points such that $(x-1)^2 + (y+2)^2 = 9$. Find $Pr(E)$.

4. Understanding Basic Theoretical Claims

Note: Some of these claims are true; others are false. Can you figure out which are true or not?
Perhaps create an example or two if the ideas are 'difficult' (usually the ideas are not difficult – the idea plus translating the 'chicken scratch' (mathematical notation) is the hard part, I opine).

Claim 4.1. Let S be a well-defined sample space (w.d.s.s.) whilst $E = \varnothing$. So, $Pr(E) = 0$.

Claim 4.2. Let S be a well-defined sample space (w.d.s.s.) whilst $Pr(E) = 0$. It is the case that $E = \varnothing$.

Claim 4.3. Let S be a w.d.s.s. whilst E is an event. It is the case that $Pr(E^c) = 1 - Pr(E)$.

Claim 4.4. Let S be a w.d.s.s. whilst E is an event. $Pr(E) > \frac{1}{2} \Longrightarrow Pr(E^c) < \frac{1}{2}$

Claim 4.5. Let S be a w.d.s.s. whilst E is an event. $Pr(E) > \frac{1}{2} \Longrightarrow Pr(E^c) \leqslant \frac{1}{2}$

Claim 4.6. Let S be a w.d.s.s. whilst E is an event. $Pr(E) \geqslant \frac{1}{2} \Longrightarrow Pr(E^c) < \frac{1}{2}$

Claim 4.7. Let S be a w.d.s.s. whilst E, F are events such that $E \subseteq F$. It is the case that $Pr(E) \leqslant Pr(F)$

Claim 4.8. Let S be a w.d.s.s. whilst E, F are events $\ni E \subset F$. It is the case that $Pr(E) < Pr(F)$.

Claim 4.9. Let S be a w.d.s.s. whilst E, F are events $\ni E = F$. It is the case that $Pr(E) = Pr(F)$.

Claim 4.10. Let S be a w.d.s.s. whilst E, F are events $\ni Pr(E) = Pr(F)$. It is the case that $E = F$.

Claim 4.11. Let S be a w.d.s.s. whilst E, F are events. Let $G = E - F$. It is the case that $Pr(G) = Pr(E) - Pr(F)$.

Claim 4.12. Let S be a w.d.s.s. whilst E, F are events. It is the case that $Pr(E \cup F) = Pr(E) + Pr(F) - Pr(E \cap F)$

Claim 4.13. There exists a w.d.s.s. S and events E, F such that $Pr(E \cup F) > 1$

Claim 4.14. $\exists$ a w.d.s.s. S and event E such that $Pr(E) \in \mathbb{I}$

Claim 4.15. $\exists$ a w.d.s.s. S and event E such that $E \neq S$ but $Pr(E) = Pr(S)$

Claim 4.16. $\exists$ a w.d.s.s. S and event E such that $Pr(E) > Pr(S)$

Of the claims on the previous page:
The first is true: Let S be a well-defined sample space (w.d.s.s.) whilst $E = \varnothing$. So, $Pr(E) = 0$.
The second is false: Let S be a well-defined sample space (w.d.s.s.) whilst $Pr(E) = 0$. It is the case that $E = \varnothing$.
The third is true: Let S be a w.d.s.s. whilst E is an event. It is the case that $Pr(E^c) = 1 - Pr(E)$.
and we could go on.

Two other claims that most people believe are true; but, are actually false is the eighth claim and the eleventh claim:
Claim 8 Let S be a w.d.s.s. whilst E, F are events $\ni E \subset F$.
It is the case that $Pr(E) < Pr(F)$.
Claim 11 Let S be a w.d.s.s. whilst E, F are events $\ni Pr(E) = Pr(F)$.
It is the case that $E = F$.

4.1. **Exercises.**

Exercise 4.17. For these exercises let $U = \mathbb{R}$ and $S = \mathbb{N}_{12}$ along with

$A = \{1, 3, 5, 8\}$	$B = \{1, 3, 5, 7, 9, 11\}$	$C = \{1, 2, 3, 4\}$
$D = \{1, 2, 3, 5, 7, 11\}$	$E = \{8, 9, 10, 11, 12\}$	$F = \{6, 7, 8\}$
$G = \{1, 2, 3, 4, 9, 10, 11, 12\}$	$H = \{2, 4, 6, 8, 10, 12\}$	$J = \{2, 3\}$

1. Find $Pr(E)$	2. Find $Pr(F)$	3. Find $Pr(E \cap F)$
4. Find $Pr(D)$	5. Find $Pr(D - G)$	6. Find $Pr(E \cup F)$
7. Find $Pr(H)$	8. Find $Pr(H^C)$	9. Find $Pr(J \cup F)$
10. Find $Pr(F \cap G \cup B)$	11. Find $Pr(G \cap H^C)$	12. Find $Pr(J \cup J^C)$
13. Find $Pr(A \cap B \cap C)$	14. Find $Pr(G \cup D \cup H)$	15. Find $Pr(A \cap J)$
16. Find $Pr(F \cap C)$	17. Find $Pr(G) + Pr(H)$	18. Find $Pr(G \cup H)$
19. Find $Pr(G \cap C)$	20. Find $Pr(J \cap E)$	21. Find $Pr(J \cup E)$
22. Find $Pr((A \cap B) \cup C)$	23. Find $Pr(A \cap (B \cup C))$	24. Find $Pr(A \cap B \cup C)$

Exercise 4.18. For these exercises let there be an urn with 5 red, 6 blue, and 4 white balls.

1. Draw a ball from the urn. Find the probability the ball is red.
2. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is not red.
3. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue or white.
4. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue or green.
5. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue or not blue.

6. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue and red.
7. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue and not blue.
8. Put the ball back in the urn. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue, or red, or white.

If one did NOT put the ball back in urn notice how the questions change.

Exercise 4.19. For these exercises let there be an urn with 5 red, 6 blue, and 4 white balls.

1. Draw a ball from the urn. Find the probability the ball is red.
2. Assume a red ball was drawn in part 1. Shake up the urn. Draw a ball from the urn. Find the probability the ball is not red.
3. Assume a ball that was not red was drawn in part 2. Shake up the urn. Draw a ball from the urn. Find the probability the ball is blue or white.

Notice the difference in the question from the previous exercise?

Nonetheless, note that all the exercises illustrate how the axioms of probability hold.

4.2. **More Exercises.**

Exercise 4.20. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing exactly two heads amongst the tosses.

Exercise 4.21. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing at least two heads.

Exercise 4.22. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing no heads.

Exercise 4.23. A fair coin is tossed 4 times and the sequence of heads and tails is observed. Find the probability of tossing at most two heads.

Exercise 4.24. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out two balls from the urn. Find the probability of choosing exactly 2 red balls.

Exercise 4.25. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out two balls from the urn. Find the probability of choosing a red ball and a blue amongst the 2 balls chosen.

Exercise 4.26. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out two balls from the urn. Find the probability of choosing at least one red ball amongst the 2 balls chosen.

Exercise 4.27. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out two balls from the urn. Find the probability of choosing a blue ball and a white ball.

Exercise 4.28. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball and then draw another ball from the urn. Find the probability of choosing a blue ball and a white ball.

Exercise 4.29. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball. We replace the ball shake up the urn and then draw a ball from the urn. Find the probability of choosing a blue ball first and a white ball second.

Exercise 4.30. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw a ball. We replace the ball shake up the urn and then draw a ball from the urn. Find the probability of choosing a blue ball and a white ball.

Exercise 4.31. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out three balls from the urn. Find the probability of choosing a blue ball and two white balls.

Exercise 4.32. Suppose there is an urn that contains 5 red, 4 white, and 11 blue balls. We draw out three balls from the urn. Find the probability of choosing two blue balls and a red ball.

With dice (unless otherwise stated) we toss a pair of dice and sum up the number of dots of the up-turned faces.

Exercise 4.33. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a sum of seven or an eleven.

Exercise 4.34. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a seven or an eleven.

Exercise 4.35. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a sum of five or better.

Exercise 4.36. Suppose we toss a pair of fair six sided dice. We view the up-turned faces of the dice. Find the probability of tossing a product of the upturned faces of a twelve.

With cards (unless otherwise stated) we deal from a fair deck of 52 standard bridge cards with the standard four suits and 13 ranks from 2 to ace. 'Face' cards are jacks, queens, kings, or aces (even though aces do not have a face aces are 'special.') 'Number' cards are 2 through 10.

Exercise 4.37. Suppose a card is drawn from a fair deck of 52 standard bridge cards. What is the probability the card is a face card?

Exercise 4.38. Suppose a card is drawn from a fair deck of 52 standard bridge cards. What is the probability the card is a heart?

Exercise 4.39. Suppose a card is drawn from a fair deck of 52 standard bridge cards. What is the probability the card is a heart or a face card?

5. Understanding Probability: Graphs

Drawing graphs to visualise the sample space and events is important; so, let us look at this in detail.
Let $U = \mathbb{R}$, the line.

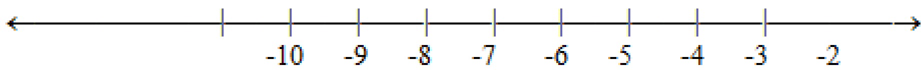


Figure 5. Line

Let us say $S = [0, 5], E = [1, 5], F = [2, 4), G = \{1, 4\}$.

For S notice $\ell(S) = 5$

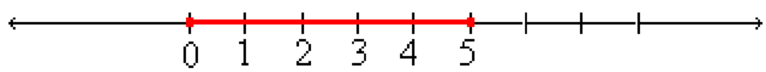


Figure 6. The graph of S is red.

For E notice $\ell(E) = 4$

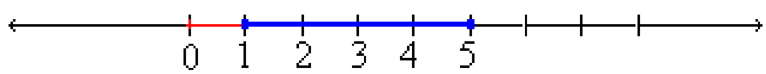


Figure 7. The graph of E is blue.

For F notice $\ell(E) = 2$

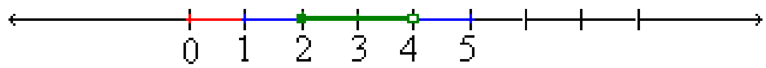


FIGURE 8. The graph of F is green.

For G notice $\ell(G) = 0$ (just two points both of length 0 and $0 + 0 = 0$).

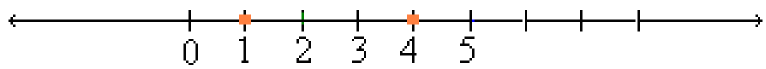
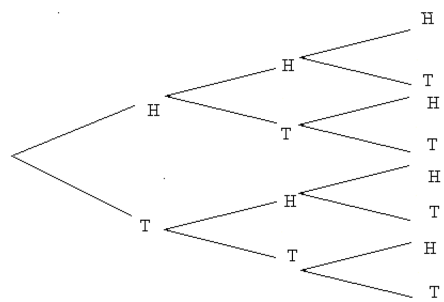


FIGURE 9. The graph of G is orange.

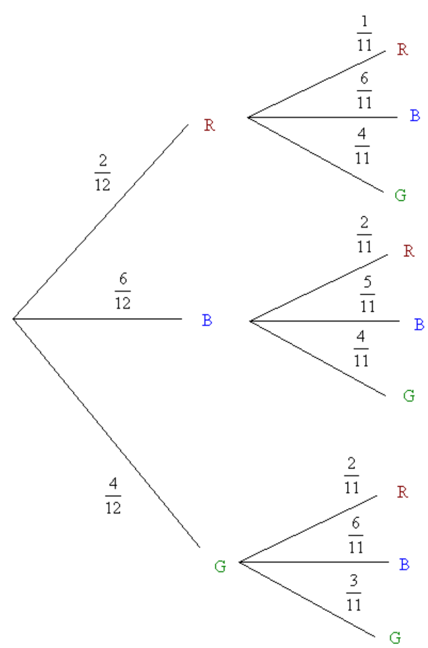
So, now the probabilities: $\Pr(E) = \dfrac{4}{5}$, $\Pr(F) = \dfrac{2}{5}$, $Pr(G) = 0$, etc.
Indeed, $\Pr(S) = \dfrac{5}{5} = 1$.
Moreover notice the $Pr(G) = 0$ and $G \neq \varnothing$; so we have an interesting fact that if the event is $\varnothing$ then $Pr(\varnothing) = 0$; but, if the probability of an event is zero that does not imply the event is $\varnothing$.

In drawing a tree diagramme it is easy if there are not many events like flipping a coin 3 times, it is easy. We are then going to draw a tree to illustrate the experiment. It shows the temporal sequencing of the experiment. But if there are a



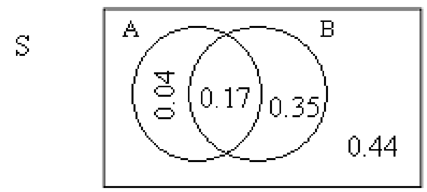
lot of things but there aren't too many stages of the experiment, then a tree is best drawn with the events and probabilities on the branches:

Let us say there is an urn with 2 red, 6 blue, ∧ 4 green balls. We draw a ball then another ball from the urn. We are then going to draw a tree to illustrate the experiment.

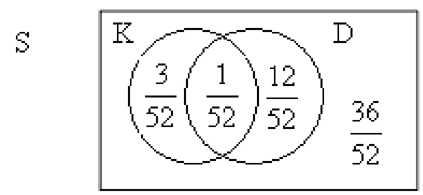


However, if the experiment is not in stages but the descriptors are multiple descriptions of things (***at the same time***) then it is best to draw a Venn diagramme with the probabilities.

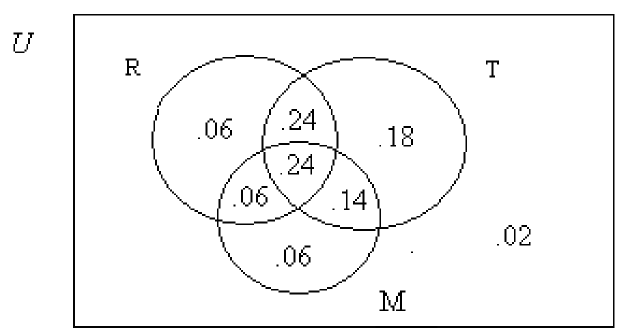
S is a well-defined sample space, $Pr(A) = 0.21, Pr(B) = 0.52$, and $Pr(A \cap B) = 0.17$ We get:



Pick a card from a standard 52 deck of cards. Finding probability of picking a king, but not a spade: Let K be the event draw a king and D be the event draw a spade. We get:



Or even three events. Let S be a well-defined sample space, $Pr(R) = 0.60$, $Pr(T) = 0.80, Pr(M) = 0.50, Pr(R \cap T) = 0.48, Pr(R \cap M) = 0.30, Pr(M \cap T) = 0.38$, and $Pr(R \cap T \cap M) = 0.24$. We get:



6. Binomial Random Variable

1. The **Bernoulli** random variable (or sometimes we say trial) is a probabilistic (or stochastic) experiment that can have one of two outcomes, success ($X = 1$) or failure ($X = 0$) in which the probability of success is p. The parametre is p. We call the variable X a Bernoulli (discrete) random variable. $p \in (0,1)$[7] and $X \in \{0,1\}$

$$Ber(x,p) = Pr(X = x) = \begin{cases} p, & x = 1 \\ 1-p & x = 0 \\ 0 & else \end{cases}$$

2. The **Binomial** random variable or variate) is the number (x) of successes in n-independent finitely many Bernoulli trials where the probability of success at each trial is p. The parametres are p and n (the number of trials).We call the variable X a binomial random variable. $p \in (0,1)$[8] and $n \in \mathbb{N}$. $x \in \mathbb{N}_n^*$ meaning, of course, $x \in \{0,1,2,...,(n-1),n\}$

$$Bin(x,p,n) = Pr(X = x) = \begin{cases} \begin{pmatrix} n \\ x \end{pmatrix} \cdot p^x \cdot (1-p)^{(n-x)} & x = 0,1,2,...,n \\ 0 & else \end{cases}$$

Example 6.1. Let $X \sim Bin(x, \frac{1}{5}, 3)$. The function to compute the probabilities is

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$
$$f(x) = \frac{3!}{x! \cdot (3-x)!} \cdot (\frac{1}{5})^x \cdot (\frac{4}{5})^{(3-x)} \ni x = 0,1,2,3$$

So,

$$f(0) = \frac{3!}{0! \cdot 3!} \cdot (\frac{1}{5})^0 \cdot (\frac{4}{5})^3 = 1 \cdot 1 \cdot \frac{64}{125} = \frac{64}{125}$$
$$f(1) = \frac{3!}{1! \cdot 2!} \cdot (\frac{1}{5})^1 \cdot (\frac{4}{5})^2 = 3 \cdot \frac{1}{5} \cdot \frac{16}{25} = \frac{48}{125}$$
$$f(2) = \frac{3!}{2! \cdot 1!} \cdot (\frac{1}{5})^2 \cdot (\frac{4}{5})^1 = 3 \cdot \frac{1}{25} \cdot \frac{4}{5} = \frac{12}{125}$$
$$f(3) = \frac{3!}{3! \cdot 0!} \cdot (\frac{1}{5})^3 \cdot (\frac{4}{5})^0 = 1 \cdot \frac{1}{125} \cdot 1 = \frac{1}{125}$$
$$f(x) = 0 \qquad \forall x \notin \{0,1,2,3\}$$

[7] It is really $p \in [0,1]$ but the trivial cases are not interesting.

[8] Once again it is really $p \in [0,1]$ but the trivial cases are not interesting.

The graph of which is: So, we will see that it is more useful to draw what we call a

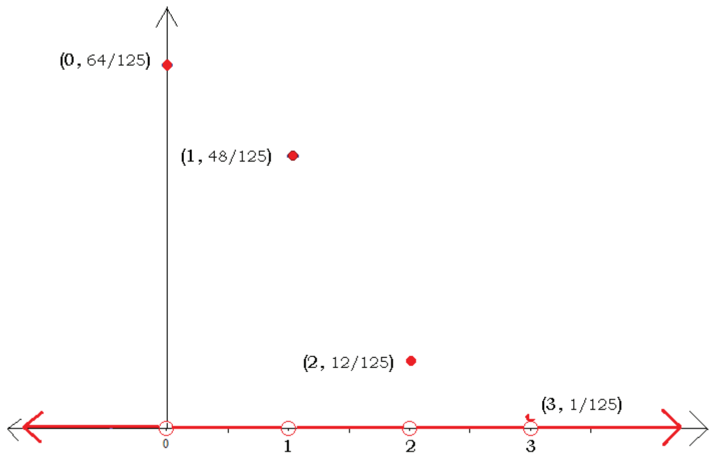


FIGURE 10. Graph of Binomial Variable with Parameters 3 and one-fifth

histogramme (more on that later)

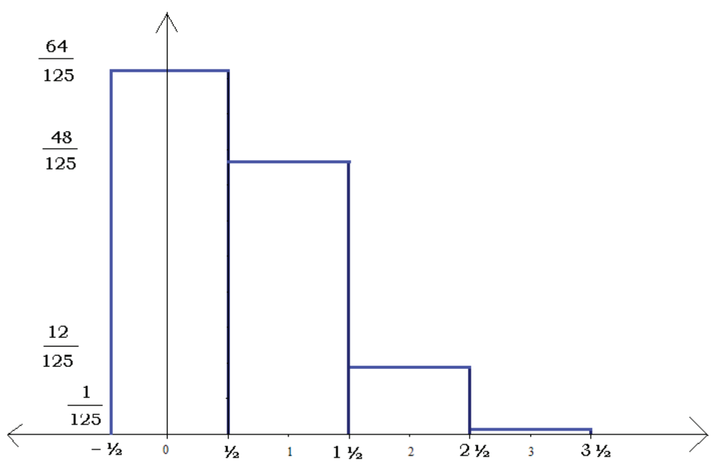


FIGURE 11. Histogramme of Binomial Variable with Parameters 3 and one-fifth

Another graph that will be important is the cumulative distribution graph.

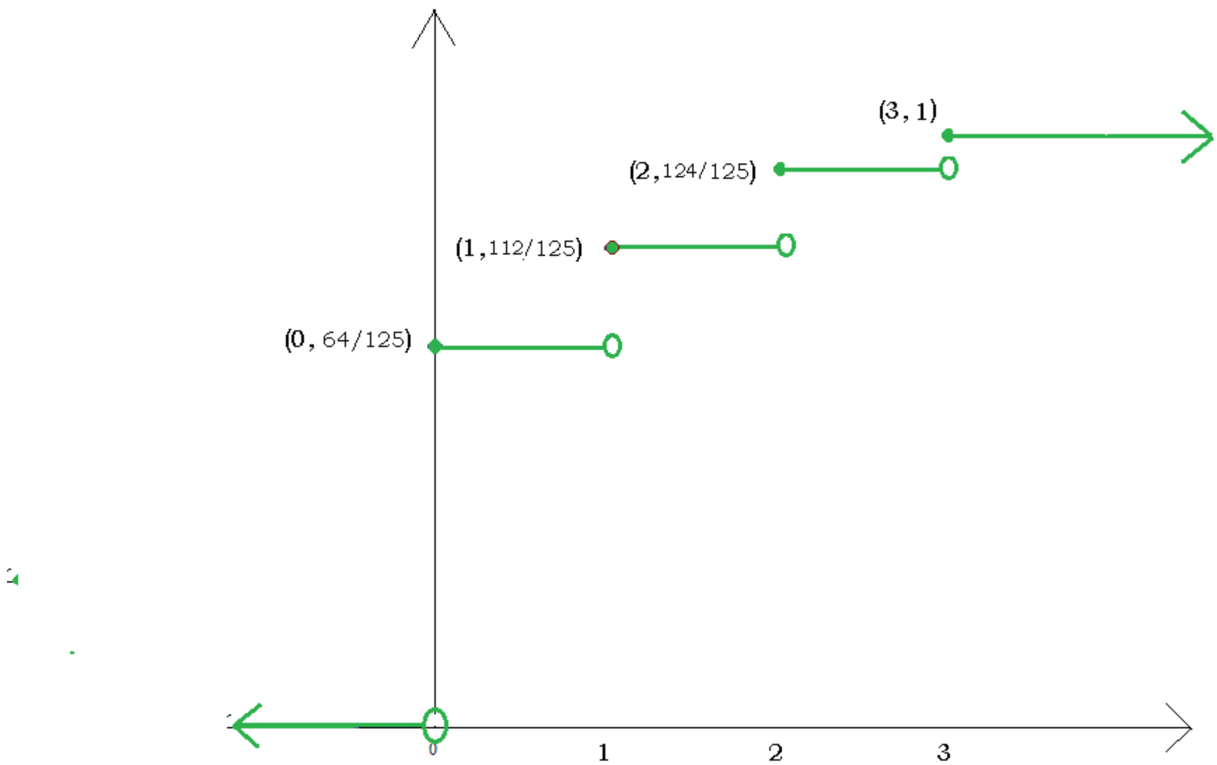


Figure 12. Graph of Binomial Variable Cumulative Distribution (for 3 and one-fifth)

6.1. **Exercises.**

Exercise 6.2. Let $X \sim Bin(x, \frac{1}{2}, 4)$.

Find

A. $Pr(X = 0)$ B. $Pr(X = 1)$ C. $Pr(X = 2)$ D. $Pr(X = 3)$ E. $Pr(X = 4)$

F. $Pr(X = 5)$ G. $Pr(X = \pi)$ H. $Pr(X \leq 3)$ I. $Pr(X > 4)$ J. $Pr(X < 4)$

K. $Pr(X \leq 4)$ L. $Pr(|X - 2| \leq 1)$

Exercise 6.3. Let $X \sim Bin(x, \frac{1}{3}, 4)$.

Find

A. $Pr(X = 0)$ B. $Pr(X = 1)$ C. $Pr(X = 2)$ D. $Pr(X = 3)$

E. $Pr(X = 4)$ F. $Pr(X = 5)$ G. $Pr(X = \frac{4}{3})$ H. $Pr(X \leq 3)$

I. $Pr(X > 4)$ J. $Pr(X < 4)$ K. $Pr(X \leq 4)$ L. $Pr(|X - \frac{4}{3}| \leq 1)$

Exercise 6.4. Let $X \sim Bin(x, \frac{1}{2}, 3)$.

Find

A. $Pr(X = 0)$ B. $Pr(X = 1)$ C. $Pr(X = 2)$ D. $Pr(X = 3)$

E. $Pr(X = 1.5)$ F. $Pr(X < 3)$ G. $Pr(X \leq 3)$ H. $Pr(|X - 1.5| \leq 1)$

Exercise 6.5. Let $X \sim Bin(x, \frac{1}{3}, 3)$.

Find

A. $Pr(X = 0)$ B. $Pr(X = 1)$ C. $Pr(X = 2)$ D. $Pr(X = 3)$

E. $Pr(X < 2)$ F. $Pr(X \leq 2)$ G. $Pr(X \leq 3)$ H. $Pr(X > 3)$

7. Normal Random Variable

3. The **Normal or Gaussian**[9] random variable is a probabilistic (or stochastic) random variable that is the **most** important of all the random variables (or so it would seem). The normal variable is a probabilistic (or stochastic) experiment that can have any outcome on $(-\infty, \infty)$. The parameters are μ and σ (or μ and σ^2 and it **must** always be clearly delineated whether the parameters are the mean and variance or the mean and standard deviation).

Thus, it is defined by its mean and standard deviation. Its applications are many and its uses quite important. A substantial number of empirical studies have indicated that the normal function provides an adequate representation of, or at least a decent approximation to, the distributions of many physical, mental, economic, biological, and social variables. (for example: meteorological data (temperature and rainfall), measurements of living organisms (height or weight of humans, populations, etc.), scores on aptitude tests, physical measurements of manufactured parts, instrumental errors, deviations from social norms, and above all else as we take means of means the limiting distribution is normal.).
$x \in \mathbb{R}$, $\mu \in \mathbb{R}$, and $\sigma \in \mathbb{R} \ni \sigma \in (0, \infty)$.[10]

$$Nor(x, \mu, \sigma) = \frac{e^{\frac{-(x-\mu)^2}{2\cdot\sigma^2}}}{\sqrt{2\cdot\pi\cdot\sigma^2}} \quad \forall x \in \mathbb{R}$$
$$Nor(x, \mu, \sigma) = \frac{1}{\sqrt{2\cdot\pi\cdot\sigma^2}\cdot e^{\frac{(x-\mu)^2}{2\cdot\sigma^2}}} \quad \forall x \in \mathbb{R}$$

A very special case of the normal family is the **standard normal** function:

$$Nor(z, 0, 1) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-z^2}{2}} \quad \forall z \in \mathbb{R}$$

Note: We will use the notation $Z \sim Nor(z, 0, 1)$ for the standard normal.
The c. d. f. of the standard normal function[11] also called the unidimensional normal ogive function is:

$$\Phi(Z \leq z) = CDFN(z, 0, 1) = \int\limits_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\cdot e^{t^2/2}} dt \quad \forall z \in \mathbb{R}$$

[9] Gaussian after K. F. Gauss, ‘the Prince of Mathematics,’ who laid much of the ground work for Modern Probability & Statistics (1809 and later). The Normal curve family of functions was first considered by DeMoivre (1733). Modern Probability & Statistics, however, was formalised by R. A. Fisher & K. Pearson.

[10] We need , $\sigma \neq 0$ for if it were zero it would mean there is no variation and all values are the mean which implies the integral across all reals is not 1 (contradicting the Kolmogorov axioms).

[11] Note it is a non-integrable [closed form] function.

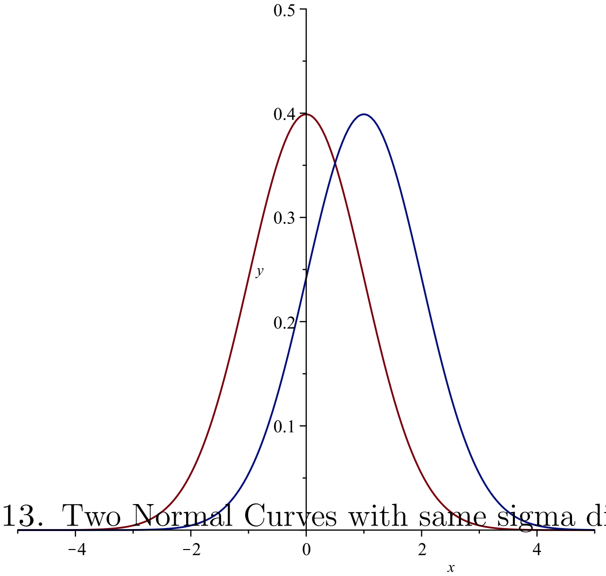


FIGURE 13. Two Normal Curves with same sigma different mus.

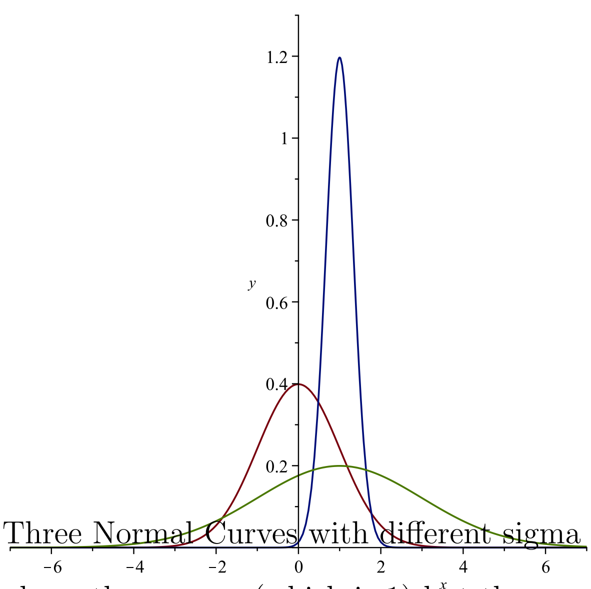


FIGURE 14. Three Normal Curves with different sigma or different mus.

The green and blue have the same μ (which is 1) but the green one has a larger σ than the blue. The red curve has a μ less than the green or blue (which is zero) and a σ that is less than the green σ but greater than the blue σ.

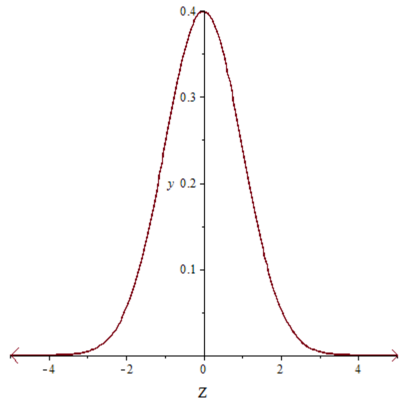


FIGURE 15. Standard Normal Curve

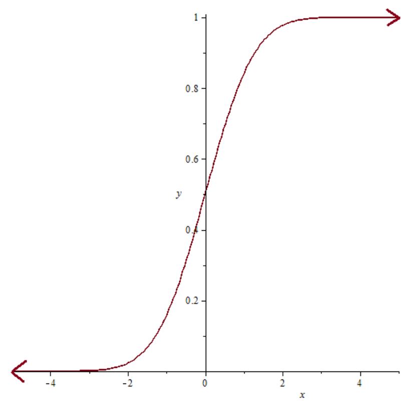


FIGURE 16. Cumulative Standard Normal Distribution Curve