

Probability/Print version

Wikibooks.org

March 24, 2011.

Contents

1	INTRODUCTION	1
1.1	TYPES OF PROBABILITY	1
1.2	ABOUT THIS BOOK	5
2	THE COUNTING PRINCIPLE	7
2.1	THE COUNTING PRINCIPLE	7
2.2	COUNTING RULES	9
3	PROBABILITY SPACES	11
3.1	CONCEPT	11
3.2	FORMAL THEORY	11
3.3	CONSEQUENCES	15
3.4	Axioms of Probability	16
4	MATHEMATICAL REVIEW	17
4.1	ELEMENTARY SET OPERATIONS	18
4.2	RULES OF SET THEORY	22
4.3	CARTESIAN PRODUCTS	23
4.4	SUMMARY OF PROBABILITY & SET THEORY	24
5	COMBINATORICS	27
5.1	THE COUNTING PRINCIPLE	27
5.2	PERMUTATIONS	29
5.3	COMBINATIONS	31
5.4	PARTITIONS	33
5.5	COMPUTING PROBABILITIES	35
6	CONDITIONAL PROBABILITY	37
6.1	DEFINITION	37
7	RANDOM VARIABLES	39
7.1	RANDOM VARIABLES: DEFINITIONS	39
8	IMPORTANT DISTRIBUTIONS	43
8.1	UNIFORM DISTRIBUTION	43
8.2	BINOMIAL DISTRIBUTION	43
8.3	THE POISSON DISTRIBUTION	44
8.4	NORMAL DISTRIBUTION	45
9	LINKS	47
10	AUTHORS	49

1 Introduction

PROBABILITY¹ theory mathematically formulates incomplete knowledge pertaining to the likelihood of an occurrence. For example, a meteorologist might say there is a 60% chance that it will rain tomorrow. This means that in 6 of every 10 times when the world is in the current state, it will rain.

A *probability* is a real number $p \in [0, 1]$. In everyday speech, the number is usually expressed as a percentage (between 0% and 100%) rather than a decimal (i.e., a probability of 0.25 is expressed as 25%). A probability of 100% means that an event is certain. In everyday speech, probability of 0% is taken to mean that the event is impossible, but (usually in where there are an infinity of possible outcomes) an event originally ascribed a probability of 0% MAY be the one that occurs. In some situations, it is CERTAIN that the event which occurs will be one that originally was ascribed zero probability (for example, in selecting a number uniformly between 0 and 1, the probability of selecting any given number is zero, but it is certain that one such number will be selected).

Another way of referring to probability of an outcome is by its *odds*: the ratio of the probability of "success" (event occurs) to the probability of "failure" (event does not occur). In the gambling world (where "odds" evolved) odds are expressed as the ratio of the STAKES risked by each participant in a wager. For instance: a bookmaker offering odds of 3/1 "against" a horse will pay a punter three times their stake (if the horse wins). In fact, the bookmaker (ignoring factors such as his potential need to "lay off" bets which are exposing him to the possibility of an unacceptable overall loss) is announcing that he thinks the horse has a 1/4 chance of winning. Using the mathematical definition of odds, "chance of winning" : "chance of not winning" = 1/4 : 3/4 = 1:3 or 1/3. So an event with a probability of 25% has odds of 33%. This disparity is even more clear where an event has a probability of 50% (e.g., the ODDS of a coin showing heads is 50%:50% = 1:1 or 1).

1.1 Types of probability

There are basically four types of probabilities, each with its limitations. None of these approaches to probability is wrong, *per se*, but some are more useful or more general than others.

In everyday speech, we express our beliefs about likelihoods of events using the same terminology as in probability theory. Often, this has nothing to do with any formal definition of probability, rather it is an intuitive idea guided by our experience, and in some cases statistics.

Consider the following examples:

¹ [HTTP://EN.WIKIBOOKS.ORG/WIKI/PROBABILITY](http://en.wikibooks.org/wiki/Probability)

- Bill says "Don't buy the avocados here; about half the time, they're rotten". Bill is expressing his belief about the probability of an event — that an avocado will be rotten — based on his personal experience.
- Lisa says "I am 95% certain the capital of Spain is Barcelona". Here, the belief Lisa is expressing is only a probability from her point of view, because only she does not know that the capital of Spain is Madrid (from our point of view, the probability is 100%). However, we can still view this as a subjective probability because it expresses a measure of uncertainty. It is as though Lisa is saying "in 95% of cases where I feel as sure as I do about this, I turn out to be right".
- Susan says "There is a lower chance of being shot in Omaha than in Detroit". Susan is expressing a belief based (presumably) on statistics.
- Dr. Smith says to Christina, "There is a 75% chance that you will live." Dr. Smith is basing this off of his research.

Probability can also be expressed in vague terms. For example, someone might say it will probably rain tomorrow. This is subjective, but implies that the speaker believes the probability is greater than 50%.

Subjective probabilities have been extensively studied, especially with regards to gambling and securities markets. While this type of probability is important, it is not the subject of this book. A good reference is "Degrees of Belief" By Steven Vick (2002).

There are two standard approaches to conceptually interpreting probabilities. The first is known as the long run (or the relative frequency approach) and the subjective belief (or confidence approach). In the Frequency Theory of Probability, probability is the limit of the relative frequency with which an event occurs in repeated trials (note that trials must be independent).

Frequentists talk about probabilities only when dealing with experiments that are random and well-defined. The probability of a random event denotes the relative frequency of occurrence of an experiment's outcome, when repeating the experiment. Frequentists consider probability to be the relative frequency "in the long run" of outcomes.

Physical probabilities, which are also called objective or frequency probabilities, are associated with random physical systems such as roulette wheels, rolling dice and radioactive atoms. In such systems, a given type of event (such as the dice yielding a six) tends to occur at a persistent rate, or 'relative frequency', in a long run of trials. Physical probabilities either explain, or are invoked to explain, these stable frequencies. Thus talk about physical probability makes sense only when dealing with well defined random experiments. The two main kinds of theory of physical probability are frequentist accounts (such as Venn) and propensity accounts.

Relative frequencies are always between 0% (the event essentially never happens) and 100% (the event essentially always happens), so in this theory as well, probabilities are between 0% and 100%. According to the Frequency Theory of Probability, what it means to say that "the probability that A occurs is p%" is that if you repeat the experiment over and over again, independently and under essentially identical conditions, the percentage of the time that A occurs will converge to p. For example, under the Frequency Theory, to say that the chance that a coin lands heads is 50% means that if you toss the coin over and over again, independently, the ratio of the number of times the coin lands heads to the total number of tosses approaches a limiting value of 50% as the number of tosses grows. Because the ratio of heads to tosses is always between 0% and 100%, when the probability exists it must be between 0% and 100%.

In the Subjective Theory of Probability, probability measures the speaker's "degree of belief" that the event will occur, on a scale of 0% (complete disbelief that the event will happen) to 100% (certainty that the event will happen). According to the Subjective Theory, what it means for me to say that "the probability that A occurs is $2/3$ " is that I believe that A will happen twice as strongly as I believe that A will not happen. The Subjective Theory is particularly useful in assigning meaning to the probability of events that in principle can occur only once. For example, how might one assign meaning to a statement like "there is a 25% chance of an earthquake on the San Andreas fault with magnitude 8 or larger before 2050?" (See Freedman and Stark, 2003, for more discussion of theories of probability and their application to earthquakes.) It is very hard to use either the Theory of Equally Likely Outcomes or the Frequency Theory to make sense of the assertion.

Bayesians, however, assign probabilities to any statement whatsoever, even when no random process is involved. Probability, for a Bayesian, is a way to represent an individual's degree of belief in a statement, given the evidence.

Evidential probability, also called Bayesian probability, can be assigned to any statement whatsoever, even when no random process is involved, as a way to represent its subjective plausibility, or the degree to which the statement is supported by the available evidence. On most accounts, evidential probabilities are considered to be degrees of belief, defined in terms of dispositions to gamble at certain odds. The four main evidential interpretations are the classical interpretation, the subjective interpretation, the epistemic or inductive interpretation, and the logical interpretation.

1.1.1 Classical theory of probability

The classical approach to probability is to count the number of *favorable outcomes*, the number of *total outcomes* (outcomes are assumed to be mutually exclusive and equiprobable), and express the probability as a ratio of these two numbers. Here, "favorable" refers not to any subjective value given to the outcomes, but is rather the classical terminology used to indicate that an outcome belongs to a given event of interest. What is meant by this will be made clear by an example, and formalized with the introduction of axiomatic probability theory.

Classical definition of probability

If the number of outcomes belonging to an event E is N_E , and the total number of outcomes is N , then the *probability* of event E is defined as $p_E = \frac{N_E}{N}$.

For example, a standard deck of cards (without jokers) has 52 cards. If we randomly draw a card from the deck, we can think of each card as a possible outcome. Therefore, there are 52 total outcomes. We can now look at various events and calculate their probabilities:

- Out of the 52 cards, there are 13 clubs. Therefore, if the event of interest is drawing a club, there are 13 favorable outcomes, and the probability of this event is $\frac{13}{52} = \frac{1}{4}$.
- There are 4 kings (one of each suit). The probability of drawing a king is $\frac{4}{52} = \frac{1}{13}$.
- What is the probability of drawing a king OR a club? This example is slightly more complicated. We cannot simply add together the number of outcomes for each event separately ($4 + 13 = 17$)

as this inadvertently counts one of the outcomes twice (the king of clubs). The correct answer is $\frac{16}{52}$ from $\frac{13}{52} + \frac{4}{52} - \frac{1}{52}$ where this is essentially $p(\text{club}) + p(\text{king}) - p(\text{king of clubs})$.

Classical probability suffers from a serious limitation. The definition of probability implicitly defines all outcomes to be equiprobable. While this might be useful for drawing cards, rolling dice, or pulling balls from urns, it offers no method for dealing with outcomes with unequal probabilities.

This limitation can even lead to mistaken statements about probabilities. An often given example goes like this:

I could be hit by a meteor tomorrow. There are two possible outcomes: I will be hit, or I will not be hit. Therefore, the probability I will be hit by a meteor tomorrow is

$$\frac{1}{2} = 50\%$$

Of course, the problem here is not with the classical theory, merely the attempted application of the theory to a situation to which it is not well adapted.

This limitation does not, however, mean that the classical theory of probability is useless. At many points in the development of the axiomatic approach to probability, classical theory is an important guiding factor.

1.1.2 Frequency of occurrence

This approach to probability is well-suited to a wide range of scientific disciplines. It is based on the idea that the underlying probability of an event can be measured by repeated trials.

Probability as a measure of frequency

Let n_A be the number of times event A occurs after n trials. We define the *probability* of event A as

$$p_A = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

It is of course impossible to conduct an infinite number of trials. However, it usually suffices to conduct a large number of trials, where the standard of large depends on the probability being measured and how accurate a measurement we need.

A note on this definition of probability: How do we know the sequence $\frac{n_A}{n}$ in the limit will converge to the same result every time, or that it will converge at all? The unfortunate answer is that we don't. To see this, consider an experiment consisting of flipping a coin an infinite number of times. We are interested in the probability of heads coming up. Imagine the result is the following sequence:

HTHHTTHHHHTTTTHHHHHHHHTTTTTTTHHHHHHHHHHHHHHHHTTTTTTTTTTTTTTTT...

with each run of k heads and k tails being followed by another run twice as long. For this example, the sequence $\frac{n_A}{n}$ oscillates between roughly $\frac{1}{3}$ and $\frac{2}{3}$ and doesn't converge.

We might expect such sequences to be unlikely, and we would be right. It will be shown later that the probability of such a run is 0, as is a sequence that converges to anything other than the underlying probability of the event. However, such examples make it clear that the limit in the definition above does not express convergence in the more familiar sense, but rather some kind of convergence in probability. The problem of formulating exactly what this means belongs to axiomatic probability theory.

1.1.3 Axiomatic probability theory

Axiomatic probability theory, although it is often frightening to beginners, is the most general approach to probability, and has been employed in tackling some of the more difficult problems in probability. We start with a set of axioms, which serve to define a **probability space**. Although these axioms may not be immediately intuitive, be assured that the development is guided by the more familiar classical probability theory.

1.2 About This Book

This book is going to discuss the topic of **mathematical probability** using CALCULUS² and ABSTRACT ALGEBRA³. Readers of this book should have a good understanding of both those topics before attempting to read and understand this book completely.

² [HTTP://EN.WIKIBOOKS.ORG/WIKI/CALCULUS](http://en.wikibooks.org/wiki/Calculus)

³ [HTTP://EN.WIKIBOOKS.ORG/WIKI/ABSTRACT%20ALGEBRA](http://en.wikibooks.org/wiki/Abstract%20Algebra)

2 The Counting Principle

2.1 The Counting Principle

Before we can delve into the properties of probability and odds, we need to understand the **Counting Principle**. We use the Counting Principle to determine how many different ways one can choose/do certain events. It is easier to define the Principle through examples:

2.1.1 Independent Events

Let's say that John is at a deli sub restaurant. There are 3 different breads, 3 different cheeses, 3 different condiments, and 3 different vegetables he can place on his sandwich, assuming that he can only place one from each category on to his sandwich. How many different ways can he set up his sandwich?

Since choosing a cheese doesn't affect the amount of choices of vegetables, condiments, or bread, these events are called **independent events**. For this problem, we will multiply 3 by 3 by 3 by 3, so 3^4 , which is 81. So there are 81 different possible combinations to form this sandwich.

Practice Problems

1) Thomas goes to a McDonalds' restaurant and chooses to create his own burger. He has 2 different kinds of cheese, 3 different breads, and 3 different sauces he can choose from, but he can only choose one of each category. How many different ways can he create this burger?

2) Diane is ordering pizza for her family. There are 4 different possible sizes of the pizza. Also, she has to choose one of 5 toppings to place on the pizza and one of 3 different types of cheese for the pizza. In addition, she must choose one of 3 different kinds of crust. How many different ways can she have her pizza?

3)a) How many 3-digit numbers can be formed from the digits 2, 3, 4, 5, 7 and 9?

b) How many of these numbers are less than 400?

Answers

1) $(2)(3)(3) = 18$

2) $(4)(5)(3)(3) = 180$

3) a) Since there are six available digits, the answer is $(6)(6)(6) = 216$

b) For the value of the 3-digit number to be less than 400, we only have two choices for the first digit, namely 2 or 3. After that we can choose the other two digits freely. The answer is thus $(2)(6)(6) = 72$.

2.1.2 Dependent Events

Assume that John is now working at a library. He must put 5 books on a shelf in any order. How many different ways can he order the books on the shelf? Unlike the independent events, when John puts a book on the shelf, that eliminates one book from the remaining choices of books to put next on the shelf; thus these are referred to as **dependent events**. At first, he has 5 different choices, so our first number in the multiplication problem will be 5. Now that one is missing, the number is reduced to 4. Next, it's down to 3, and so on. So, the problem will be

$$(5)(4)(3)(2)(1)$$

However, there is a symbol for this very idea. A $!$ represents the term *factorial*. So, for example, $3! = (3)(2)(1)$. Factorials are very useful in statistics and probability.

Therefore, the problem can be rewritten as $5!$, which ends up being equal to 120.

However, not all dependent event problems are that simple. Let's say that there are 10 dogs at a dog competition. How many different ways can one select the Champion AND the Runner-Up? This problem could actually be considered simpler than the last one, but it doesn't involve a factorial. So, how many different ways can the judge determine the Champion? Of course, there are 10 different dogs, so 10 different ways. Next, how many dogs are left to select the Runner-Up from? Well, you removed one, so it's down to 9. Instead of placing a factorial, though, you will only multiply 10 and 9, resulting in 90.

2.1.3 Independent Or Dependent?

To help you differentiate between the two, we will do a few more examples, but we will have to decide if it is dependent or independent before solving the problem.

Let's say that you are creating a 5-digit garage door opener code (the digits would include 0-9). If there were absolutely no restrictions, would the events be independent of each other or dependent on each other? Of course, there are no restrictions, since you could have five 4's for the code, so to solve the problem, you multiply 10 by itself 5 times, resulting in 100000.

Alternatively, suppose that the first number of the code cannot be 0, and that there can be no repeated numbers whatsoever. Obviously these events are dependent on each other, because there cannot be any repetitions. Let's look at this problem one number at a time.

The first number can be all the numbers except 0, reducing the possible amount to 9. The second number **can** be 0 this time, so the possible amount returns to 10. However, it cannot be a repeat of the previous number, so there are 9 possible choices again. After that, the numbers will reduce by one each time, due to the fact that there cannot be repeats, so the problem would look like this

$$(9)(9)(8)(7)(6) = (9)(9!)/(5!) = 27216$$

Now, just one more example. Let's say that you were choosing your schedule for school. There are 8 periods each day, and there are 7 classes to choose from. Nonetheless, you must have a lunch period during the 4th period. We can think of 4th period as non-existent because it is in a constant position and therefore does not affect the possible choices. With 7 slots and 7 options, the answer is simply $7!$.

$$(7!) = 5040$$

2.1.4 Review Of The Counting Principle

So, we use the Counting Principle to determine the different unique ways we can do something, such as a sandwich or a selection of classes. Sometimes, these events will affect each other, such as when you can't choose the same number twice for your garage door code, so they are dependent events. However, other times, one event has no effect on the next event, such as when you have different cheeses and breads to choose for your sandwich, so they are independent events. The Counting Principle is a fundamental mathematical idea and an essential part of probability.

2.2 Counting Rules

Rule 1: If any one of K mutually exclusive and exhaustive events can occur on each of N trials, there are K^N different sequences that may result from a set of such trials Example: Flip a coin three times, finding the number of possible sequences. $N=3$, $K=2$, therefore, $K^N = 2^3 = 8$

Rule 2: If K_1, K_2, \dots, K_N are the numbers of distinct events that can occur on trials $1, \dots, N$ in a series, the number of different sequences of N events that can occur is $(K_1)(K_2)\dots(K_N)$ Example: Flip a coin and roll a die, finding the number of possible sequences. Therefore, $(K_1)(K_2) = (2)(6) = 12$

Rule 3: The number of different ways that N distinct things may be arranged in order is $N! = (1)(2)(3)\dots(N-1)(N)$, where $0! = 1$. An arrangement in order is called a permutation, so that the total number of permutations of N objects is $N!$ (the symbol $N!$ is called N -factorial) Example: Arrange 10 items in order, finding the number of possible ways. Therefore, $10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3,628,800$

Rule 4: The number of ways, N , of selecting and arranging r objects from among N distinct objects is: $N!/(N-r)!$, or as seen on calculators, $[nPr]$ Example: pick 3 things from 10 items, and arrange them in order. Therefore $N=10$, $r=3$, so $10!/(10-3)! = 10!/7! = 720$

Rule 5: The total number of ways of selecting r distinct combinations of N objects, irrespective of order (ie order NOT important), is: $N!/r!(N-r)!$ or as seen on calculators, $[nC_r]$ Example: Pick 3 items from 10 in any order, where $N=10$, $r=3$. Therefore, $10!/3!(7!) = 720/6 = 120$

3 Probability Spaces

3.1 Concept

Although we came up with a basic definition of probability in the previous chapter, we will now proceed to develop a more axiomatic theory. Our theory will avoid the ambiguities of probability, and allow for a simpler mathematical formalism. We shall proceed by developing the concept of probability space, which will allow us to harness many theorems in mathematical analysis.

3.2 Formal Theory

3.2.1 Set of Outcomes

The set of all possible outcomes is called the sample space, denoted by Ω . For every problem, you must pick an appropriate sample space. This is important because we can't make any conclusions about the probability of an event if we don't know the exact size of the sample space. In a coin toss the states could be "Heads" and "Tails". For a die there could be one state for each side of the die. We have a probability function that specifies the probability of each state. Events are sets of states. In the die example an event could be rolling an even number.

3.2.2 Probability Space Definition

A *Probability Space* consists of (Ω, S, P) where Ω is a non-empty set, called the *sample space*, its elements are called the *outcomes*, $S \subset \text{Power}(\Omega)$, containing the *events*, and P is a function $S \rightarrow$

\mathbb{R} . Ω is called the sample space, and is a set of all the possible outcomes. Outcomes are all the possibilities of what can occur, where only one occurs. S is the set of events. Events are sets of outcomes, and they occur when any of their outcomes occur. For example rolling an even number might be an event, but it will consist of the outcomes 2, 4, and 6. The probability function gives a number for each event, and the probability that something will occur is 1.

E.g, when tossing a single coin Ω is $\{H, T\}$ and possible events are $\{\}, \{H\}, \{T\}$, and $\{H, T\}$. Intuitively, the probability of each of these sets is the chance that one of the events in the set will happen; $P(\{H\})$ is the chance of tossing a head, $P(\{H, T\}) = 1$ is the chance of the coin landing either heads or tails, $P\{\} = 0$ is the probability of the coin landing **neither** heads **nor** tails, etc.

3.2.4 Other Definitions

- Mutually exclusive

two or more events that can NOT occur at the same time; have no outcomes in common. Events are mutually exclusive if they cannot both occur simultaneously. Events are said to be mutually exclusive if the occurrence of any one event automatically implies the non-occurrence of the remaining $n-1$ events. Mutually exclusive events have the property in which $\Pr(A \cap B) = 0$. Also, when A and B are mutually exclusive events, $P(A \text{ or } B) = P(A) + P(B)$. In short, this implies that at most, one of the events may occur. In terms of statistics, the definition of mutually exclusive is: A property of a set of categories such that an individual or object is included in only one category. The occurrence of one event means that none of the other events can occur at the same time.

(Collectively) Exhaustive

events are said to be collectively exhaustive, which means that at least one of the events must occur. A set is jointly or exhaustive if at least one of the events must occur. Another way to describe collectively exhaustive events, is that their union must cover all the events within the entire sample space. For example, events A and B are said to be collectively exhaustive if where S is the sample space.

Test of Independence

If for two events A and B , $p(A \cap B)$ is not equal to $p(A)p(B)$, then A and B are said to be associated or dependent. If $p(A \cap B) > p(A)p(B)$, so that $p(A|B) > p(A)$ and $p(B|A) > p(B)$, then the events are said to be positively associated (answer to question #2 above). If, however, $p(A \cap B) < p(A)p(B)$, so that $p(A|B) < p(A)$ and $p(B|A) < p(B)$, one says that the events are negatively associated.

Simple Random Sampling

In a simple random sample, one person must take a random sample from a population, and not have any order in which one chooses the specific individual. In statistics, a simple random sample is a subset of individuals (a sample) chosen from a larger set (a population). Each individual is chosen randomly and entirely by chance, such that each individual has the same probability of being chosen at any stage during the sampling process, and each subset of k individuals has the same probability of being chosen for the sample as any other subset of k individuals. Simple random sampling can be done with or without replacement, though it is typically done without .e., one deliberately avoids choosing any member of the population more than once. When the sample is drawn without replacement, the same elementary event can occur more than once. When the sample is drawn with replacement, the same elementary event can occur no more than once in a given sample. Therefore, random sampling of one individual at a time means that every possible individual in the large group has an equal probability of being drawn.

Independence of random variables

If X is a real-valued random variable and a is a number then the event $\{X \leq a\}$ is the set of outcomes that correspond to X being less than or equal to a . Since these are sets of outcomes that have probabilities, it makes sense to refer to events of this sort being independent of other events of this sort.

Why are the p and q Bernoulli trial probabilities multiplied together in the binomial formula? The probability of an event can be expressed as a binomial probability if its outcomes can be broken down into two probabilities p and q , where p and q are complementary (i.e. $p + q = 1$). Binomial probability typically deals with the probability of several successive decisions, each of which has two possible outcomes. The binomial distribution is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p . Such a success/failure experiment is also called a Bernoulli experiment or Bernoulli trial. In fact, when $n = 1$, the binomial distribution is a Bernoulli distribution. Bernoulli process is a discrete-time stochastic process consisting of a sequence of independent random variables taking values over two symbols. Prosaically, a Bernoulli process is coin flipping several times, possibly with an unfair coin. A variable in such a sequence may be called a Bernoulli variable. In other words, a Bernoulli process is a sequence of independent identically distributed Bernoulli trials. Independence of Bernoulli trials implies memorylessness property: past trials do not provide any information regarding future outcomes. From any given time, future trials is also a Bernoulli process independent of the past (fresh-start property). a sequence or other collection of random variables is independent and identically distributed (i.i.d.) if each random variable has the same probability distribution as the others and all are mutually independent. Two events A and B are independent if and only if $\Pr(A \cap B) = \Pr(A)\Pr(B)$. Here $A \cap B$ is the intersection of A and B , that is, it is the event that both events A and B occur. This is called the multiplication rule for independent events.

Universal Set

all the elements for any specific discussion, and is symbolized by the symbol U .

Example: $U = \{A, E, I, O, U\}$

Intersection

the elements 2 or more sets and is denoted by the symbol, \cap .

Union

elements in two or more sets and is denoted by the symbol, \cup .

Complement

all the elements in the universal set that not the original set and is denoted by the symbol, $\bar{\cdot}$. Sometimes represented also by the symbol, \sim , meaning "not" (i.e. $p(A \cup \bar{B})$ denotes "A union NOT B").

Example:

$$U = \{1,2,3,4,5,6,7,8,9,0\}; A = \{1,2,3\}; B = \{2,3,4,5,6\}$$

$$A \cap B = \{2,3\}$$

$$A \cup B = \{1,2,3,4,5,6\}$$

$$A \cap \bar{B} = \{1,7,8,9,0\}$$

Empty or Null Set

a set that contains no elements and are denoted by the symbols $\{\}$, \emptyset . This means that the event probability in question is impossible and thus, cannot occur.

3.2.5 Further Concepts in Set Theory

Union and intersection: We can indicate the union and intersection of two sets with the symbols \cup , respectively. If two sets, then the set $A \cup B$ is the set that consists of all the elements that are either in set A or in set B . Similarly, read $A \cap B$, which is the set of all elements that are in both A and B .

Set difference. We can "subtract" one set from another using the symbol \setminus . If sets, then the set $A \setminus B$ consisting of all the elements that are members of A but are not members of B . This is a little different from the subtraction you may be used to in arithmetic.

Whole numbers: 1. Non-negative integers (aka real numbers) – $\{0, 1, 2, 3, \dots\}$ 2. Positive integers – $\{1, 2, 3, 4, \dots\}$ 3. All possible integers – $\{-3, -2, -1, 0, 1, 2, 3, \dots\}$

AND \rightarrow 2 or more things/events happen. Therefore, we MULTIPLY (product) probabilities together -- INTERSECTION OR \rightarrow 1 of the 2 events/things occur. Therefore we ADD (sum) probabilities together -- UNION

Events that are NOT mutually exclusive means we could draw them both (for example, we want to draw a 2 OR an 'M'...these events are NOT mutually exclusive, because they cannot occur together at the same time).

Events that are NOT mutually exclusive, but rather are independent, we must subtract: $1 - p(A \cap B)$ from $p(A) + p(B)$. So that, $p(A) + p(B) - p(A \cap B)$.

Probabilities for independent events often involve exponents, and probabilities for dependent events (conditional probability) often involve factorials.

3.2.6 Permutations and Combinations

Permutations

arrangement of objects without repetition where order is important. Permutations using all the objects: n objects, arranged into group size of n without repetition, and order being important – $P(n, n) = n!$

Example: Find all permutations of A, B, C.

Permutations of some of the objects: n objects, group size r , order is important. $P(n, r) = n! / (n-r)!$
Example: Find all 2-letter combinations using letters A, B, C.

Distinguished permutations

if a word has n letters, k of which are unique, let $n_1, n_2, n_3, \dots, n_k$ be the frequency of each of the k letters. $n! / (n_1!)(n_2!)(n_3!)$

Combinations

arrangement of objects without repetition, where order is NOT important. A combination of n objects, arranged in groups of size r , without repetition, and order being important. $C(n, r) = n! / r!(n-r)!$

Another way to write a combination of n things, r at a time, is using the Binomial notation (Binomial Distribution), sometimes described as " n choose r ".

3.2.7 Counting Rules

Rule 1

If any one of K mutually exclusive and exhaustive events can occur on each of N trials, there are KN different sequences that may result from a set of such trials

Example: Flip a coin three times, finding the number of possible sequences. $N=3$, $K=2$, therefore, $KN = 2^3 = 8$

Rule 2

If K_1, K_2, \dots, K_N are the numbers of distinct events that can occur on trials $1, \dots, N$ in a series, the number of different sequences of N events that can occur is $(K_1)(K_2)\dots(K_N)$

Example: Flip a coin and roll a die, finding the number of possible sequences. Therefore, $(K_1)(K_2) = (2)(6) = 12$

Rule 3

The number of different ways that N distinct things may be arranged in order is $N! = (1)(2)(3)\dots(N-1)(N)$, where $0! = 1$. An arrangement in order is called a permutation, so that the total number of permutations of N objects is $N!$ (the symbol $N!$ is called N -factorial)

Example: Arrange 10 items in order, finding the number of possible ways. Therefore, $10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3628800$

Rule 4

The number of ways of selecting and arranging r objects from among N distinct objects is: $N!/(N-r)!$ [nPr]

Example: pick 3 things from 10 items, and arrange them in order. Therefore $N=10$, $r=3$, so $10!/(10-3)! = 10!/7! = 720$

Rule 5

The total number of ways of selecting r distinct combinations of N objects, irrespective of order (ie order NOT important), is: $N!/r!(N-r)!$ [nCr]

Example: Pick 3 items from 10 in any order, where $N=10$, $r=3$. Therefore, $10!/3!(7!) = 720/6 = 120$

3.3 Consequences

We can now give some basic theorems using our axiomatic probability space.

3.3.1 Theorem 1

Given a probability space (Ω, S, P) , for events $A, B \in S$:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

3.4 Axioms of Probability

A probability function has to satisfy the following three basic axioms or constraints: To each event a a measure (number) $P(a)$ which is called the probability of event a is assigned. $P(a)$ is subjected to the following three axioms:

1. $P(a) \geq 0$
2. $P(S) = 1$
3. If $a \cap b = \emptyset$, then $P(a \cup b) = P(a) + P(b)$

Corollaries:

- $P(\emptyset) = 0$
- $0 \leq P(a) \leq 1$
- If $a \cap b = \emptyset$, then $P(a \cup b) = P(a) + P(b)$
- If $b \subset a$, $P(a) = P(b) + P(a \setminus b)$

4 Mathematical Review

The review of set theory contained herein adopts a naive point of view. We assume that the meaning of a set as a collection of objects is intuitively clear. A rigorous analysis of the concept belongs to the foundations of mathematics and mathematical logic. Although we shall not initiate a study of these fields, the rules we follow in dealing with sets are derived from them. A *set* is a collection of objects, which are the *elements* of the set.

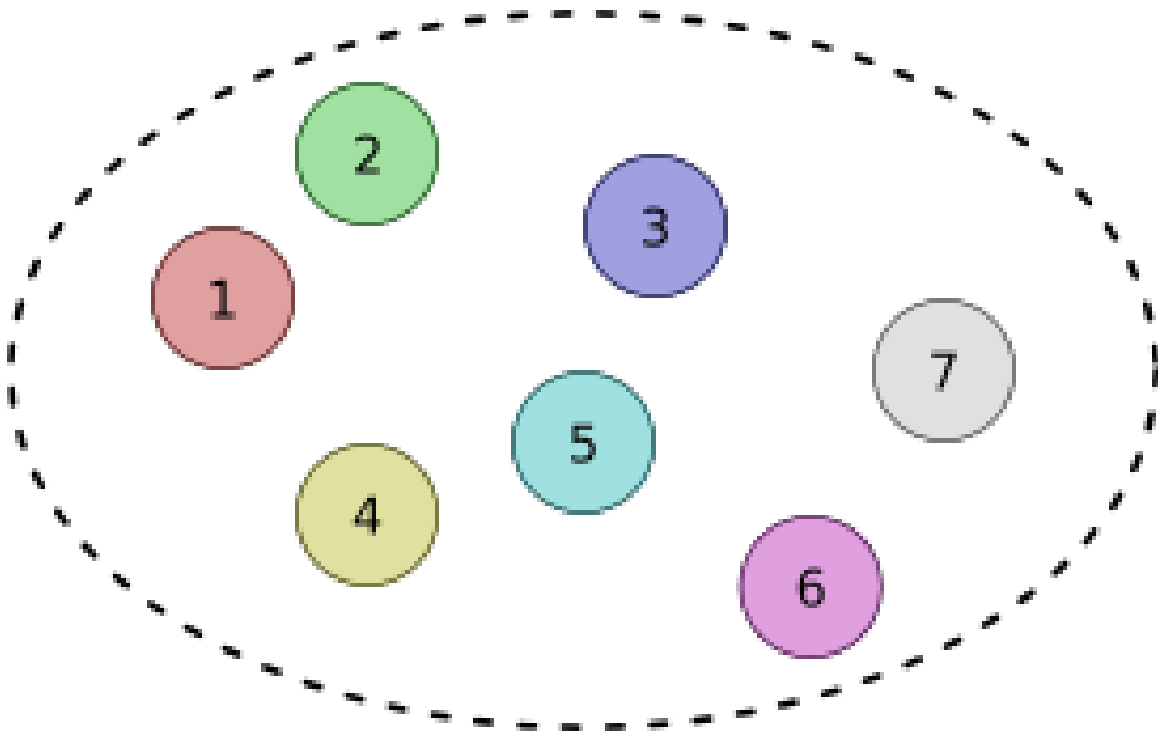


Figure 1

If an element x belongs to a set S , we express this fact by writing $x \in S$. If x does not belong to S , we write $x \notin S$. We use the equality symbol to denote *logical identity*. For instance, $x = y$ means that x and y are symbols denoting the same object. Similarly, the equation $S = T$ states that S and T are two symbols for the same set. In particular, the sets S and T contain precisely the same elements. If x and y are different objects then we write $x \neq y$. Also, we can express the fact that S and T are different sets by writing $S \neq T$.

A set S is a *subset* of T if every element of S is also contained in T . We express this relation by writing $S \subset T$. Note that this definition does not require S to be different from T . In fact, $S = T$ if and only if $S \subset T$ and $T \subset S$. If $S \subset T$ and S is different from T , then S is a *proper subset* of T and we write $S \subsetneq T$.

There are many ways to specify a set. If the set contains only a few elements, one can simply list the objects in the set;

$$S = \{x_1, x_2, x_3\}.$$

The content of a set can also be enumerated whenever S has a countable number of elements,

$$S = \{x_1, x_2, \dots\}.$$

Usually, the way to specify a set is to take some collection S of objects and some property that elements of S may or may not possess, and to form the set consisting of all elements of S having that property. For example, starting with the integers \mathbb{Z} , we can form the subset of S consisting of all even numbers

$$S = \{x \in \mathbb{Z} | x \text{ is an even number}\}.$$

More generally, we denote the set of all elements that have a certain property P by

$$S = \{x | x \text{ satisfies } P\}.$$

The braces are to be read as the words "the set of" whereas the symbol $|$ stands for the words "such that."

It is convenient to introduce two special sets. The *empty set*, denoted by \emptyset , is a set that contains no elements. The *universal set* is the collection of all objects of interest in a particular context, and it is denoted by Ω . Once a universal set Ω is specified, we need only consider sets that are subsets of Ω . In the context of probability, Ω is often called the *sample space*.

The *complement* of a set S , with respect to the universal set Ω , is the collection of all objects in Ω that do not belong to S ,

$$S^c = \{x \in \Omega | x \notin S\}.$$

We note that $\Omega^c = \emptyset$.

4.1 Elementary Set Operations

Probability theory makes extensive use of elementary set operations. Below, we review the ideas of set theory, and establish the basic terminology and notation. Consider two sets, S and T .

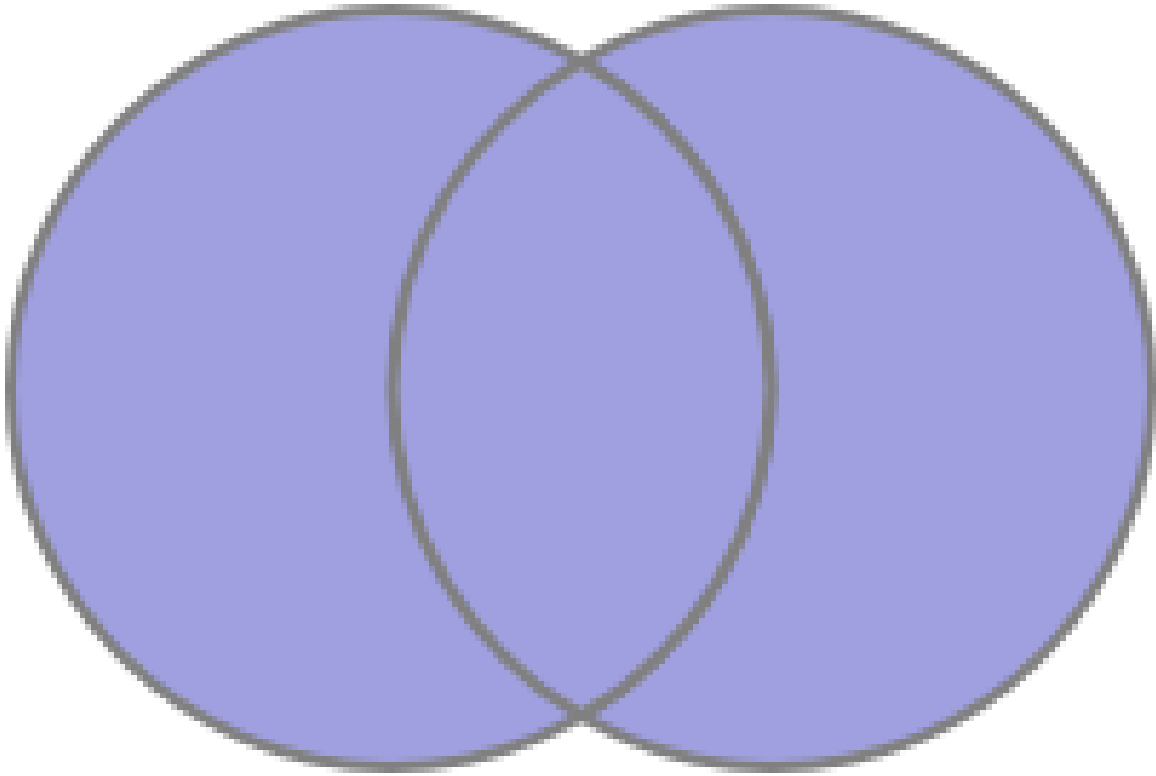


Figure 2

The *union* of sets S and T is the collection of all elements that belong to S or T (or both), and it is denoted by $S \cup T$. Formally, we define the union of these two sets by

$$S \cup T = \{x \mid x \in S \text{ or } x \in T\}.$$

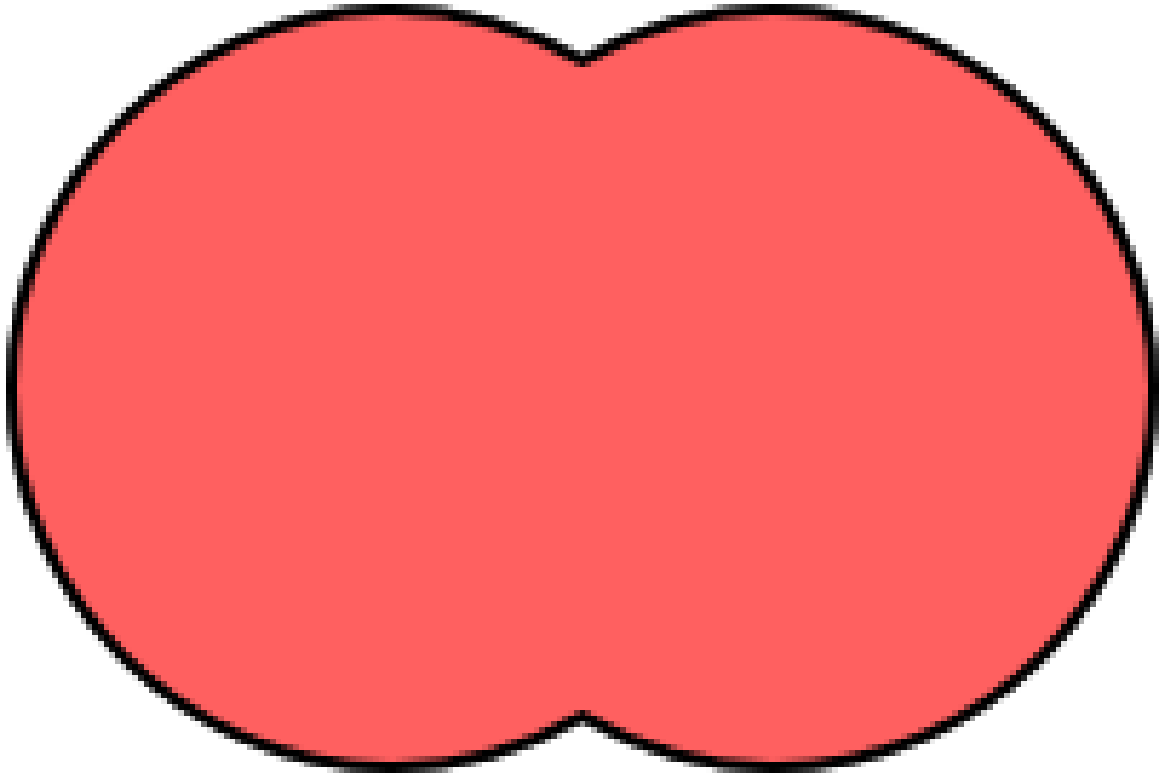


Figure 3

The *intersection* of sets S and T is the collection of all elements that belong to S and T . It is denoted by $S \cap T$, and it can be expressed mathematically as

$$S \cap T = \{x | x \in S \text{ and } x \in T\}.$$

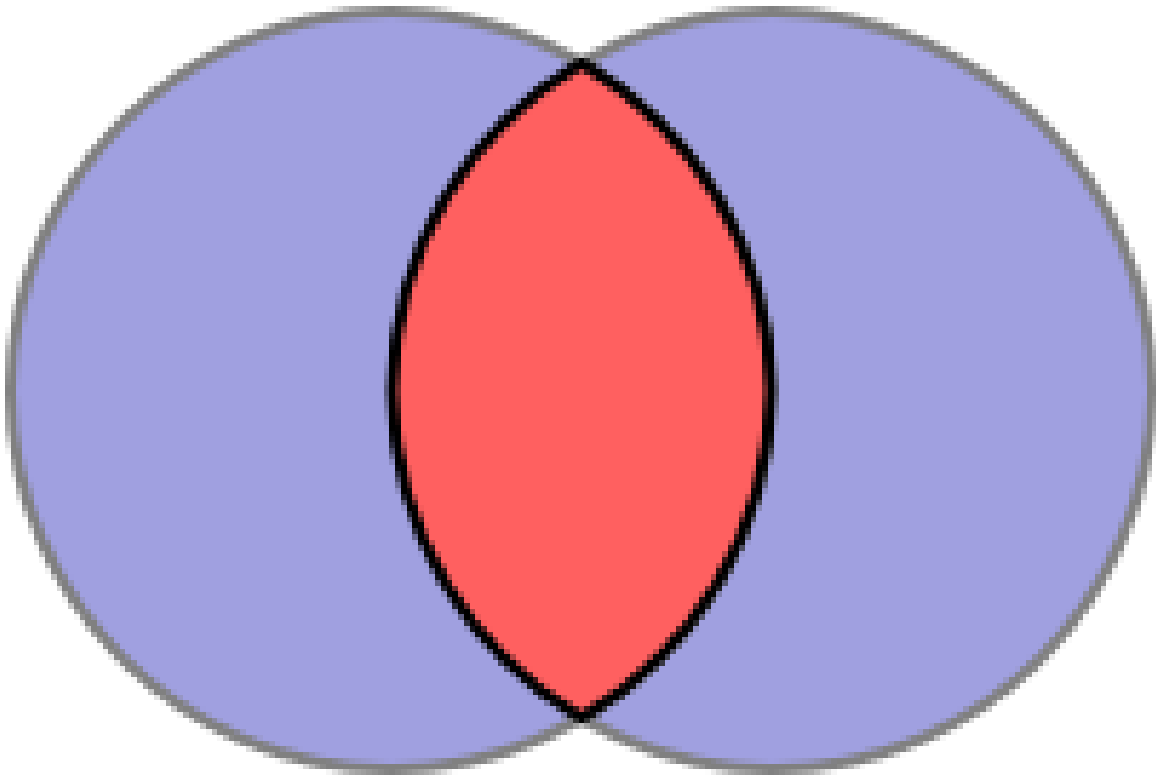


Figure 4

When S and T have no elements in common, we write $S \cap T = \emptyset$. We also express this fact by saying that S and T are *disjoint*. More generally, a collection of sets is said to be disjoint if no two sets have a common element. A collection of sets is said to form a *partition* of S if the sets in the collection are disjoint and their union is S .

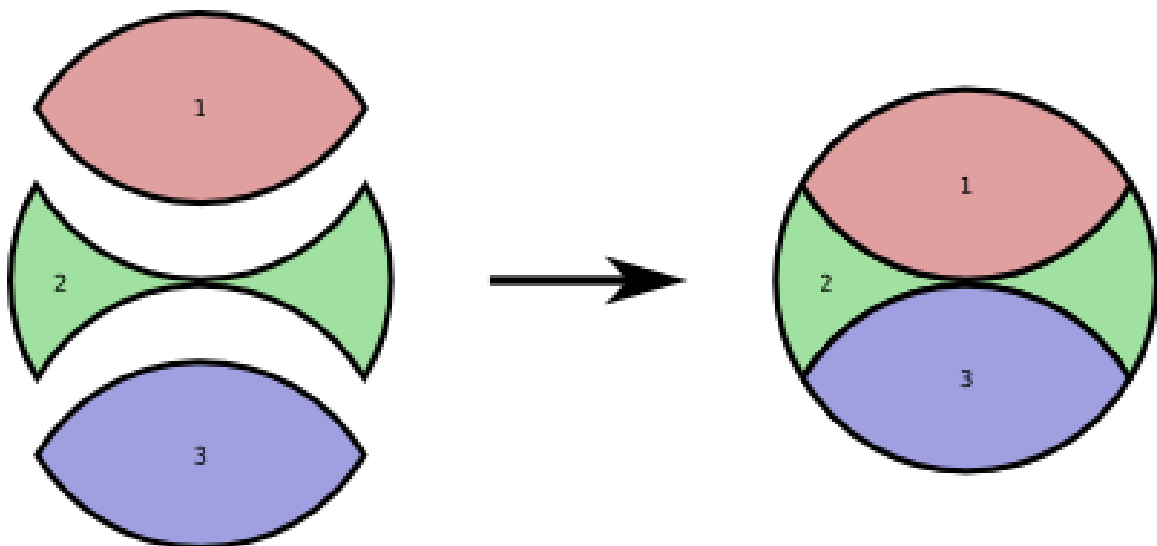


Figure 5

The *difference* of two sets, denoted by $S - T$, is defined as the set consisting of those elements of S that are not in T ,

$$S - T = \{x \mid x \in S \text{ and } x \notin T\}.$$

This set is sometimes called the complement of T relative to S , or the complement of T in S .

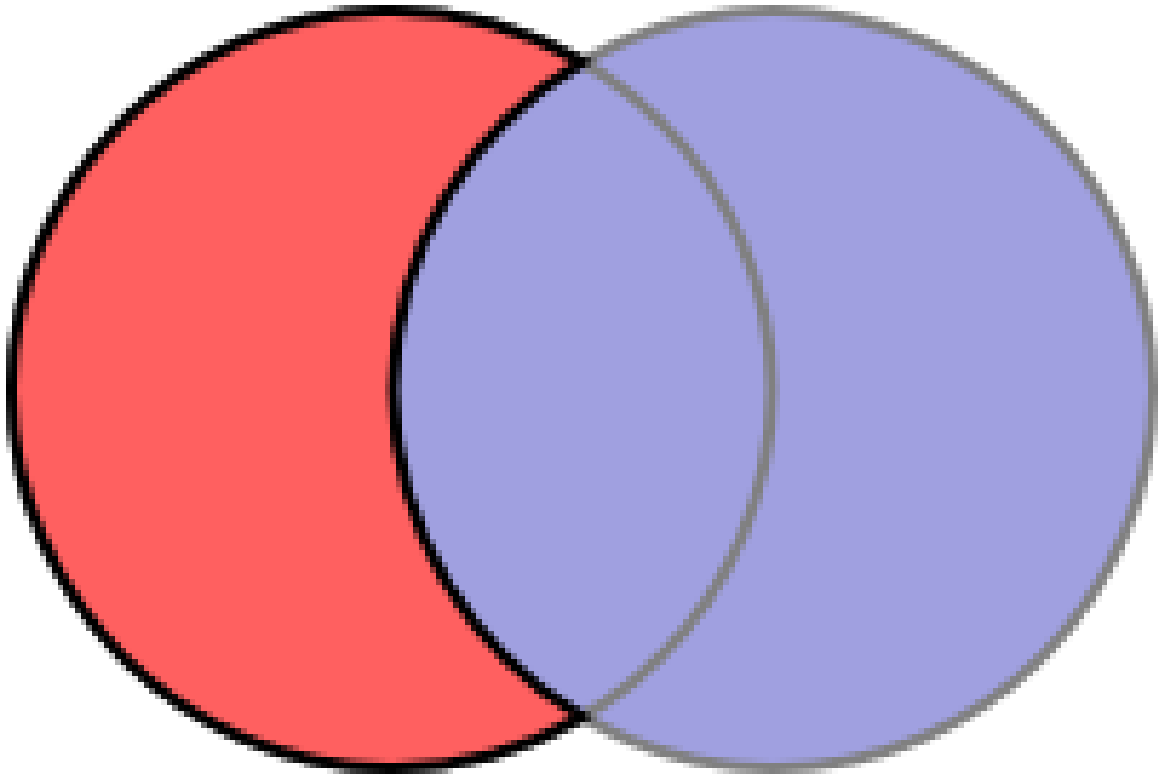


Figure 6

We have already looked at the definition of the union and the intersection of two sets. We can also form the union or the intersection of arbitrarily many sets. This is defined in the obvious way,

$$\bigcup_{\alpha \in I} S_{\alpha} = \{x \mid x \in S_{\alpha} \text{ for some } \alpha \in I\}$$

$$\bigcap_{\alpha \in I} S_{\alpha} = \{x \mid x \in S_{\alpha} \text{ for all } \alpha \in I\}.$$

The index set I can be finite or even infinite.

4.2 Rules of Set Theory

Given a collection of sets, it is possible to form new ones by applying elementary set operations to them. As in algebra, one uses parentheses to indicate precedence. For instance, $R \cup (S \cap T)$ denotes the union of two sets R and $S \cap T$, while $(R \cup S) \cap T$ represents the intersection of two sets $R \cup S$ and T . The sets thus formed are quite different.

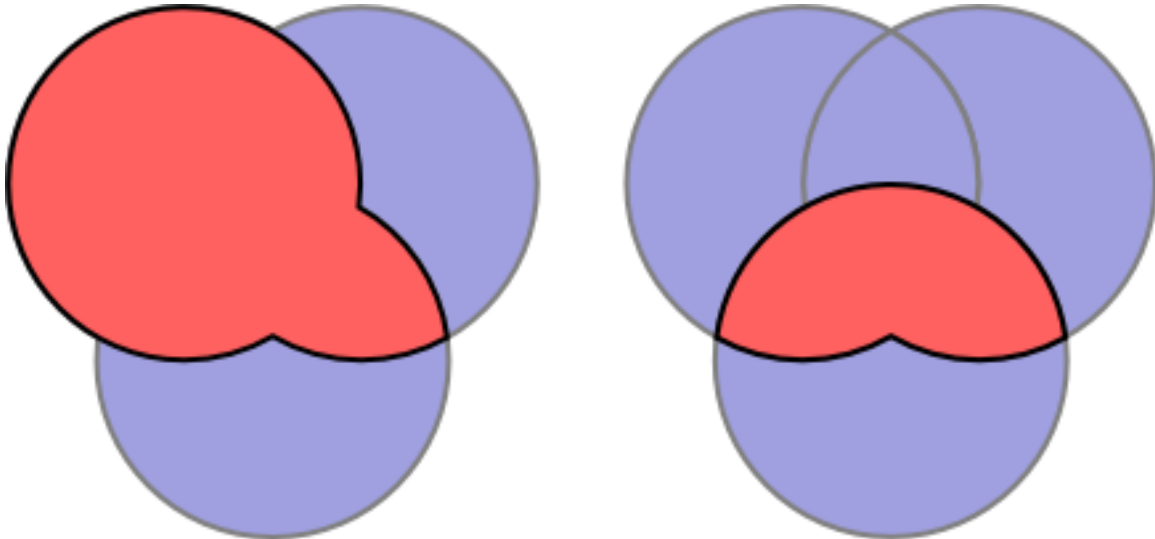


Figure 7

Sometimes different combinations of operations lead to the same set. For instance, we have the two distributive laws

$$R \cap (S \cup T) = (R \cap S) \cup (R \cap T)$$

$$R \cup (S \cap T) = (R \cup S) \cap (R \cup T).$$

Two particularly useful equivalent combinations of operations are given by *De Morgan's laws*, which state that

$$R - (S \cup T) = (R - S) \cap (R - T)$$

$$R - (S \cap T) = (R - S) \cup (R - T).$$

These two laws can be generalized to

$$(\cup_{\alpha \in I} S_{\alpha})^c = \cap_{\alpha \in I} S_{\alpha}^c$$

$$(\cap_{\alpha \in I} S_{\alpha})^c = \cup_{\alpha \in I} S_{\alpha}^c$$

when multiple sets are involved. To establish the first equality, suppose that x belongs to $(\cup_{\alpha \in I} S_{\alpha})^c$. Then x is not contained in $\cup_{\alpha \in I} S_{\alpha}$. That is, x is not an element of S_{α} for any $\alpha \in I$. This implies that x belongs to S_{α}^c for all $\alpha \in I$, and therefore $x \in \cap_{\alpha \in I} S_{\alpha}^c$. We have shown that $(\cup_{\alpha \in I} S_{\alpha})^c \subset \cap_{\alpha \in I} S_{\alpha}^c$. The converse inclusion is obtained by reversing the above argument. The second law can be obtained in a similar fashion.

4.3 Cartesian Products

There is yet another way to create new sets from existing ones. It involves the notion of an *ordered pair* of objects. Given sets S and T , the *cartesian product* $S \times T$ is the set of all ordered pairs (x, y) for which x is an element of S and y is an element of T ,

$$S \times T = \{(x, y) | x \in S \text{ and } y \in T\}.$$

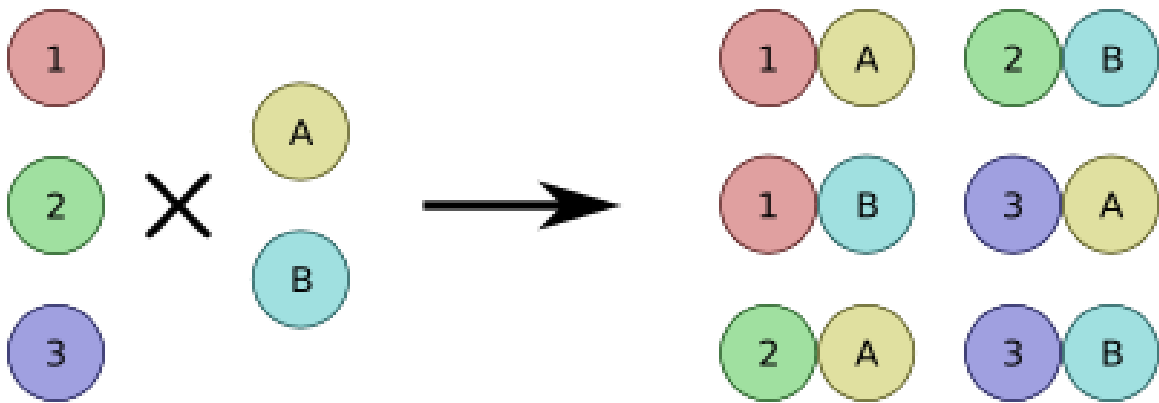


Figure 8

4.4 Summary of Probability & Set Theory

- $A = P(A)$ [0, 1]
- NOT A = $P(A') = 1 - P(A)$
- A OR B = $P(A \cup B) = P(A) + P(B) - P(AB) = P(A) + P(B)$ [*if and only if A and B are mutually exclusive]
- A AND B = $P(AB) = P(A/B) * P(B) = P(A) * P(B)$ [* if and only if A and B are independent]
- A GIVEN B = $P(A|B) = P(AB)/P(B)$ [*conditional]

Basic Set Theory

UNION: combined area of set A and B, known as $A \cup B$. The set of all items which are members of either A or B

- Union of A and B are added together
- Some basic properties of unions:
- $A \cup B = B \cup A$
- $A \cup (B \cap C) = (A \cup B) \cap C$
- $A \subset (A \cup B)$
- $A \cup A = A$
- $A \cup \emptyset = A$, where \emptyset = null, empty set
- $A \subset B$, if and only if $A \cup B = B$

INTERSECTION: area where both A and B overlap, known as AB . It represents which members the two sets A and B have in common

- If $AB = \emptyset$, then A and B are said to be **DISJOINT**.
- Some basic properties of intersections:
- $AB = BA$
- $A(B \cap C) = (AB) \cap C$
- $AB \subset A$
- $AA = A$

- $A \cap 0 = 0$
- $A \subset B$, if and only if $AB = A$
- UNIVERSAL SET: space of all things possible, which contains ALL of the elements or elementary events.
- U/A is called the absolute complement of A

Complement (set): 2 sets can be subtracted. The relative complement (set theoretic difference of B and A). Denoted by B/A (or $B - A$) is the set of all elements which are members of B , but not members of A

Some basic properties of complements (\bar{A} , or A'):

- $A \cup A' = U$
- $AA' = 0$
- $(A')' = A$
- $A/A = 0$
- $U' = 0$, and $0 = U$
- $A/B = 'AB'$

Summary

- Intersection (AB) --> AND – both events occur together at the same time
- Union ($A \cup B$) --> OR – everything about both events, A and B
- Complement (\bar{A}) --> NOT A – everything else except A (or the event in question)
- $A \cup \bar{A} = S$ (sample space)
- $A \bar{A} = 0$ (impossible event)

Union and Intersection are:

Commutative:

- $A \cup B = B \cup A$
- $AB = BA$

Associative:

- $A \cup (B \cap C) = (A \cup B) \cap C$
- $A(B \cap C) = (AB)C$

Distributive:

- $A \cup (BC) = (A \cup B)(A \cup C)$
- $A(B \cup C) = (AB) \cup (AC)$

5 Combinatorics

Often, in experiments with finite sample spaces, the outcomes are equiprobable. In such cases, the probability of an event amounts to the number of outcomes comprising this event divided by the total number of outcomes in the sample space. While counting outcomes may appear straightforward, it is in many circumstances a daunting task. For example, consider the number of distinct subsets of the integers $\{1, 2, \dots, n\}$ that do not contain two consecutive integers. This number is equal to

$$\frac{\phi^{n+2} - (1-\phi)^{n+2}}{\sqrt{5}},$$

where $\phi = (1 + \sqrt{5})/2$ is the *golden ratio*. It can also be obtained recursively through the *Fibonacci* recurrence relation.

Calculating the number of ways that certain patterns can be formed is part of the field of *combinatorics*. In this section, we introduce useful counting techniques that can be applied to situations pertinent to probability theory.

5.1 The Counting Principle

The Fundamental Rule of Counting If a set of choices or trials, $T_1, T_2, T_3, \dots, T_k$, could result, respectively, in $n_1, n_2, n_3, \dots, n_k$ possible outcomes, the entire set of k choices or trials has $n_1 \times n_2 \times n_3 \times \dots \times n_k$ possible outcomes. (The numbers $n_1, n_2, n_3, \dots, n_k$ cannot depend on which outcomes actually occur.)

By the Fundamental Rule of Counting, the total number of possible sequences of choices is $5 \times 4 \times 3 \times 2 \times 1 = 120$ sequences. Each sequence is called a permutation of the five items. A permutation of items is an ordering of the items. More generally, by the Fundamental Rule of Counting, in ordering n things, there are n choices for the first, $(n-1)$ choices for the second, etc., so the total number of ways of ordering a total of n things is $n \times (n-1) \times (n-2) \times \dots \times 1$. This product, $n \times (n-1) \times (n-2) \times \dots \times 1$, is written $n!$, which is pronounced "n factorial." By convention, $0! = 1$

The counting principle is the guiding rule for computing the number of elements in a cartesian product as well.

$$S \times T = \{(x, y) | x \in S \wedge y \in T\}.$$

The number of elements in the cartesian product $S \times T$ is equal to mn . Note that the total number of outcomes does not depend on the order in which the experiments are realized.

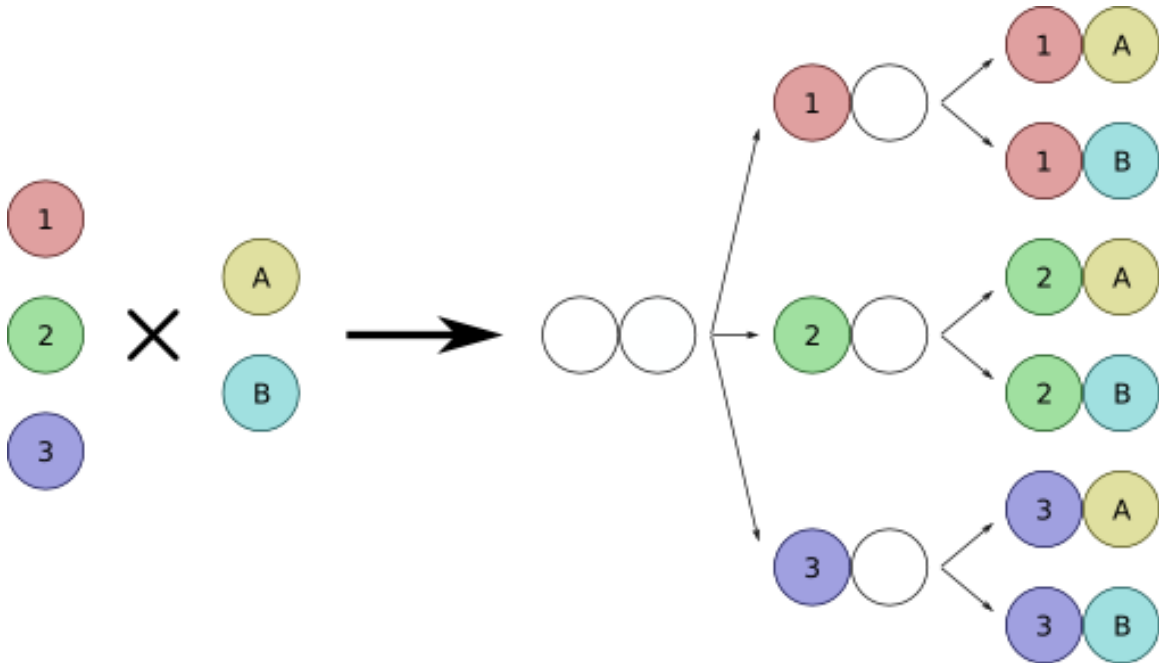


Figure 9

Example: Consider an experiment consisting of flipping a coin and rolling a die. There are two possibilities for the coin, heads or tails, and the die has six sides. The total number of outcomes for this experiment is $2 \times 6 = 12$. That is, there are twelve outcomes for the roll of a die followed by the toss of a coin: 1H, 1T, 2H, 2T, 3H, 3T, 4H, 4T, 5H, 5T, 6H, 6T.

The counting principle can be extended to compute the number of elements in the cartesian product of multiple sets. Consider the finite sets S_1, S_2, \dots, S_r and their cartesian product

$$S_1 \times S_2 \times \dots \times S_r = \{(s_1, s_2, \dots, s_r) | s_p \in S_p\}.$$

If we denote the cardinality of S_p by $n_p = |S_p|$, then the number of distinct ordered r-tuples of the form (s_1, s_2, \dots, s_r) is $n = n_1 n_2 \dots n_r$.

Example - Sampling with Replacement and Ordering: An urn contains n balls numbered 1 through n . A ball is drawn from the urn, and its number is recorded on an ordered list. The ball is then replaced in the urn. This procedure is repeated k times. We wish to compute the number of possible sequences that results from this experiment. There are k drawings and n possibilities per drawing. Using the counting principle, we conclude that the number of distinct sequences is n^k .

Example: The power set of S , denoted by 2^S , is the set of all subsets of S . In set theory, 2^S represents the set of all functions from S to $\{0, 1\}$. By identifying a function in 2^S with the corresponding preimage of one, we obtain a bijection between 2^S and the subsets of S . In particular, each function in 2^S is the characteristic function of a subset of S .

Suppose that S is finite with $n = |S|$ elements. For every element of S , a characteristic function in 2^S is either zero or one. There are therefore 2^n distinct characteristic functions from S to $\{0, 1\}$. Hence, the number of distinct subsets of S is given by 2^n .

Formulae The formula for counting combinations has special cases that are worth memorizing: $nC0 = nCn = 1$ (There is only one way to pick no thing and only one way to pick all n things.) $nC1 = nCn-1 = n$ (there are n ways to pick one thing or to leave one thing out) $nCk = nCn-k$ (There are the same number of ways of picking k of n things as there are of leaving out k of n things)

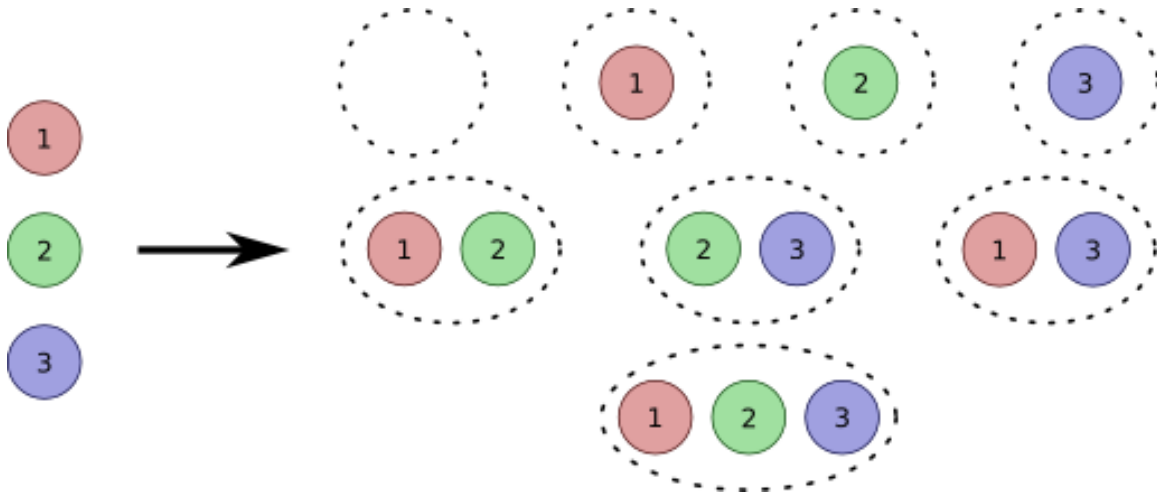


Figure 10

5.2 Permutations

Permutation: an arrangement of objects without repetition where order is important. Permutations using all the objects: n objects, arranged into group size of n without repetition, and order being important. $P(n, n) = n!$ Example: Find all permutations of A, B, C

Permutations of some of the objects: n objects, group size r , order is important. $P(n, r) = N!/(n-r)!$ Example: Find all 2-letter combinations using letters A, B, C **Distinguished permutation:** if a word has n letters, k of which are unique, let $n_1, n_2, n_3, \dots, n_k$ be the frequency of each of the k letters: $N!/(n_1!)(n_2!)(n_3!)$.

Consider the integer set $S = \{1, 2, \dots, n\}$. A *permutation* of S is an ordered arrangement of its elements, a list without repetitions. The number of permutations of S can be computed as follows. The number of distinct possibilities for the first item is n . The number of possibilities for the second item is $n-1$, namely all the integers in S except the first element in the list. Similarly, the number of distinct possibilities for the m -th item is $n - m + 1$. This pattern continues until all the elements in S are listed. Summarizing, we find that the total number of permutations of S is n factorial, $n! = n (n-1) \dots 1$.

Example: We wish to compute the number of permutations of $S = \{1, 2, 3\}$. Since the set S contains three elements, it has $3! = 6$ permutations. They can be listed as 123, 132, 213, 231, 312, 321.

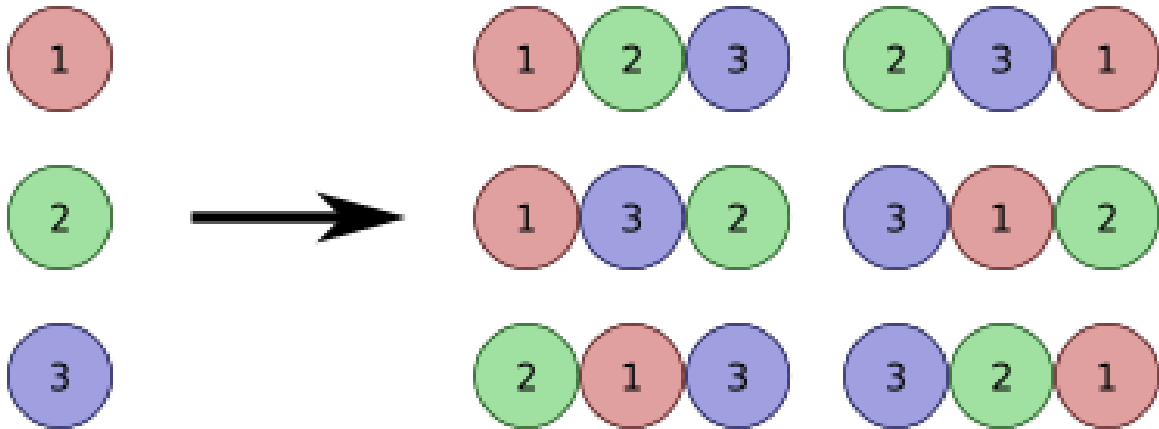


Figure 11

5.2.1 Stirling's Formula

The number $n!$ grows very rapidly as a function of n . A good approximation for $n!$ when n is large is given by Stirling's formula,

$$n! \sim n^n e^{-n} \sqrt{2\pi n}$$

as $n \rightarrow \infty$. The notation $a(n) \sim b(n)$ signifies that the ratio $a(n)/b(n)$ approaches one as n tends to infinity.

5.2.2 k-Permutations

Suppose that we list only k elements out of the set $S = \{1, 2, \dots, n\}$, where $k \leq n$. We wish to count the number of distinct k -permutations of S . Following our previous argument, we can choose one of n elements to be the first item listed, one of $n-1$ elements for the second item, and so on. The procedure terminates when k items have been listed. The number of possible sequences is

$$\frac{n!}{(n-k)!} = n(n-1) \cdots (n-k+1).$$

Example: A recently formed music group has four original songs they can play. They are asked to perform two songs at a music festival. We wish to compute the number of song arrangements the group can offer in concert. Abstractly, this is equivalent to computing the number of 2-permutations of four songs. Thus, the number of distinct arrangements is $4!/2! = 12$.

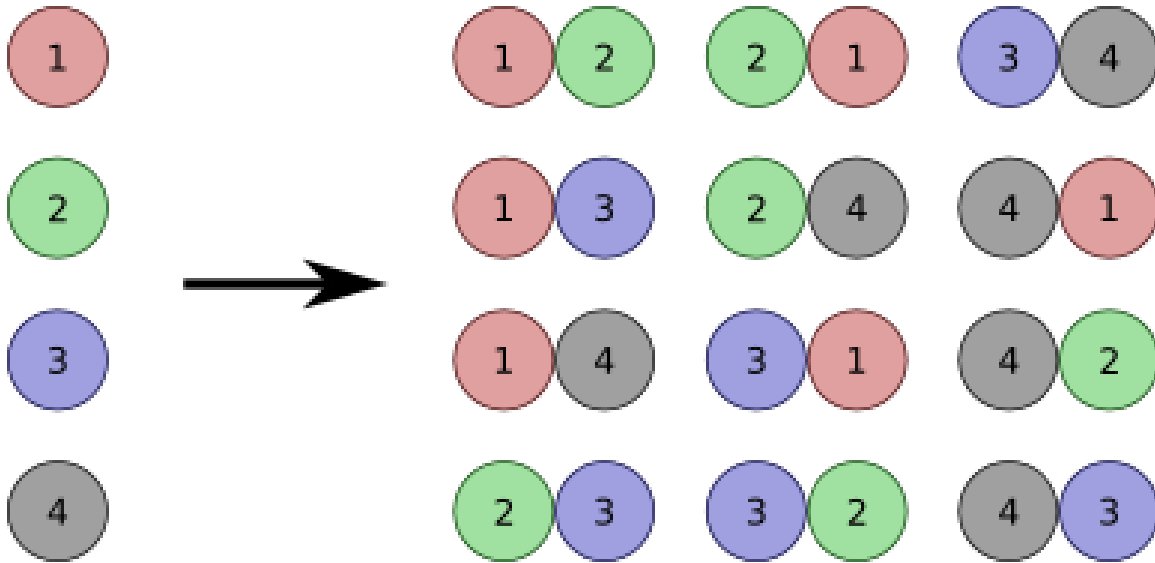


Figure 12

Example - Sampling without Replacement, with Ordering: An urn contains n balls numbered one through n . A ball is picked from the urn, and its number is recorded on an ordered list. The ball is not replaced in the urn. This procedure is repeated until k balls are selected from the urn, where $k \leq n$. We wish to compute the number of possible sequences that results from this experiment. The number of possibilities is equivalent to the number of k -permutations of n elements, which is given by $n!/(n-k)!$.

5.3 Combinations

Combinations: arrangement of objects without repetition, where order is NOT important. A combination of n objects, arranged in groups of size r , without repetition, and order being important. $C(n, r) = n!/r!(n-r)!$ Another way to write a combination of n things, r at a time, is using the Binomial notation (**Binomial Distribution**). Combinations of this type are referred to verbally as "n choose r".

Consider the integer set $S = \{1, 2, \dots, n\}$. A *combination* is a subset of S . We emphasize that a combination differs from a permutation in that elements in a combination have no specific ordering. The 2-element subsets of $S = \{1, 2, 3, 4\}$ are $\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}$, whereas the 2-permutations of S are more numerous with $(1, 2), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4), (3, 1), (3, 2), (3, 4), (4, 1), (4, 2), (4, 3)$. There are therefore fewer 2-element subsets of S than 2-permutations of S .

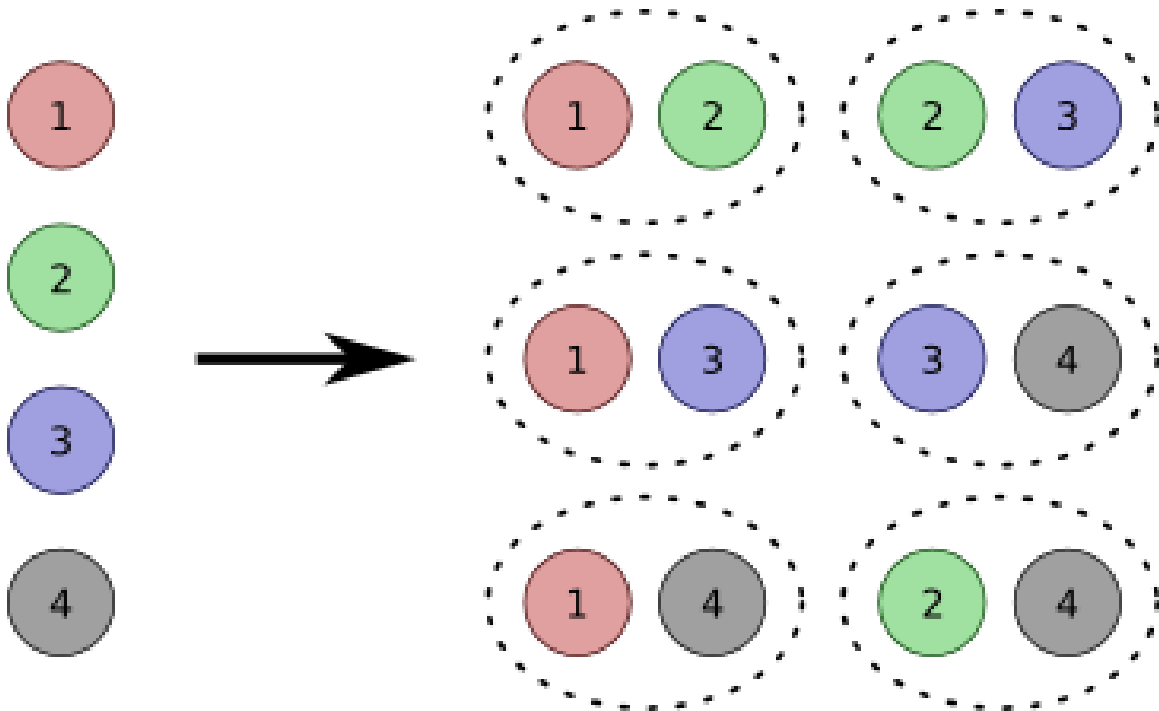


Figure 13

We can compute the number of k -element combinations of $S = \{1, 2, \dots, n\}$ as follows. First, we note that a k -permutation can be formed by first selecting k objects from S and then ordering them. There are $k!$ distinct ways of ordering k objects. The number of k -permutations must then equal the number of k -element combinations multiplied by $k!$. Since the total number of k -permutations of S is $n! / (n-k)!$, the number of k -element combinations is found to be

$$\frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k!}.$$

This expression is called a *binomial coefficient*. Observe that selecting a k -element subset of S is equivalent to choosing the $n-k$ elements that belong to its complement.

Example - Sampling without Replacement or Ordering: An urn contains n balls numbered one through n . A ball is drawn from the urn and placed in a jar. This process is repeated until the jar contains k balls, where $k \leq n$. We wish to compute the number of distinct combinations the jar can hold after the completion of this experiment. Because there is no ordering in the jar, this amounts to counting the number of k -element subsets of a given n -element set, which is given by

$$\frac{n!}{k!(n-k)!}.$$

Again, let $S = \{1, 2, \dots, n\}$. Since a combination is also a subset and the number of k -element combinations of S is $\frac{n!}{k!(n-k)!}$, the sum of the binomial coefficients over all values of k must be equal to the number of elements in the power set of S . Thus, we get

$$\sum_{k=0}^n \frac{n!}{k!(n-k)!} = 2^n.$$

5.4 Partitions

Abstractly, a combination is equivalent to partitioning a set into two subsets, one containing k objects and the other containing the $n-k$ remaining objects. In general, the set $S = \{1, 2, \dots, n\}$ can be partitioned into r subsets. Let n_1, n_2, \dots, n_r be nonnegative integers such that

$$\sum_{p=1}^r n_p = n.$$

Consider the following iterative algorithm that leads to a partition of S . First, we select a subset of n_1 elements from S . Having chosen the first subset, we select a second subset containing n_2 elements from the remaining $n - n_1$ elements. We continue this procedure by successively choosing subsets of n_p elements from the remaining $n - n_1 - \dots - n_{p-1}$ elements, until no element remains. This algorithm yields a partition of S into r subsets, with the p -th subset containing exactly n_p elements.

We wish to count the number of such partitions. We know that there are $\frac{n!}{n_1!(n-n_1)!}$ ways to form the first subset. Examining our algorithm, we see that there are

$$\frac{(n - n_1 - \dots - n_{p-1})!}{n_p!(n - n_1 - \dots - n_p)!}$$

ways to form the p -th subset. Using the counting principle, the total number of partitions is then given by

$$\frac{n!}{n_1!n_2!\dots n_r!}.$$

This expression is called a *multinomial coefficient*.

Example: A die is rolled nine times. We wish to compute the number of outcomes for which every odd number appears three times. The number of distinct sequences in which one, three, and five each appear three times is equal to the number of partitions of $\{1, 2, \dots, 9\}$ into three subsets of size three, namely

$$\frac{9!}{3!3!3!} = 1680.$$

In the above analysis, we assume that the cardinality of each subset is fixed. Suppose instead that we are interested in counting the number of ways to pick the cardinality of the subsets that form the partition. Specifically, we wish to compute the number of ways integers n_1, n_2, \dots, n_r can be selected such that every integer is nonnegative and

$$\sum_{p=1}^r n_p = n.$$

We can visualize the number of possible assignments as follows. Picture n balls spaced out on a straight line and consider $r-1$ vertical markers, each of which can be put between two consecutive balls, before the first ball, or after the last ball. For instance, if there are five balls and two markers then one possible assignment is illustrated below.

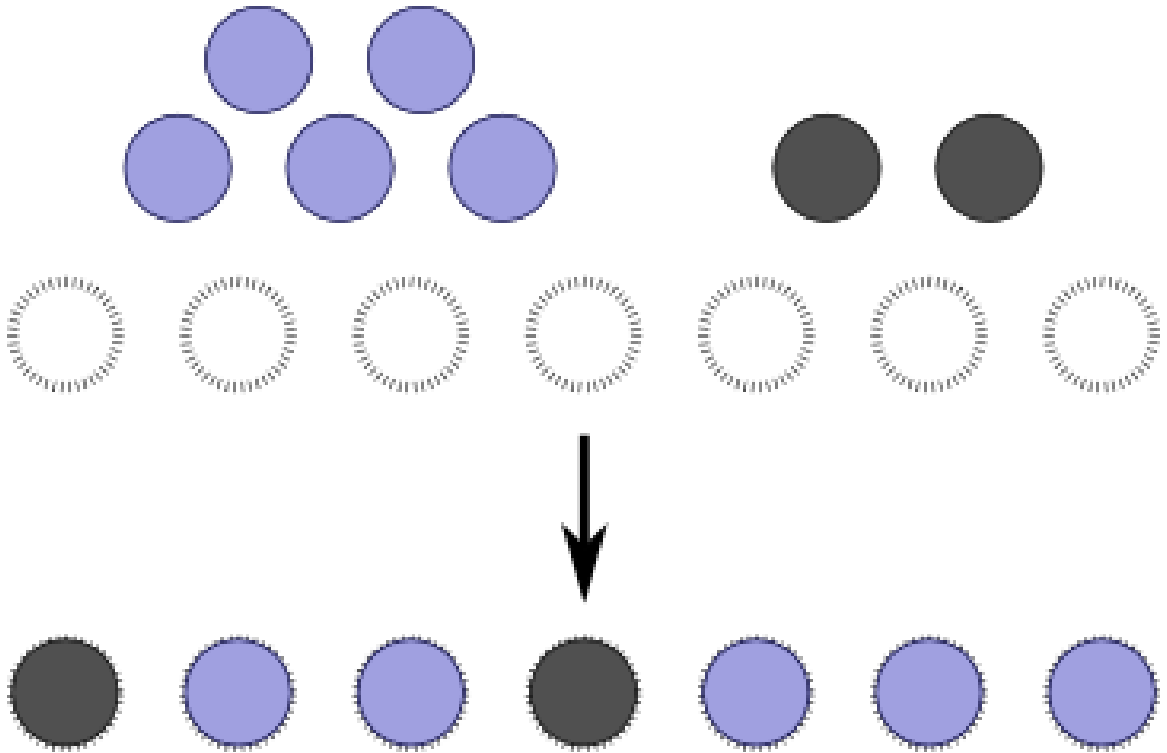


Figure 14

The number of objects in the first subset corresponds to the number balls before the first marker. Similarly, the number of objects in the p -th subset is equal to the number of balls between the p -th marker and the preceding one. Finally, the number of objects in the last subset is simply the number of balls after the last marker. In the figure, the integer assignment is

$$(n_1, n_2, n_3) = (0, 2, 3).$$

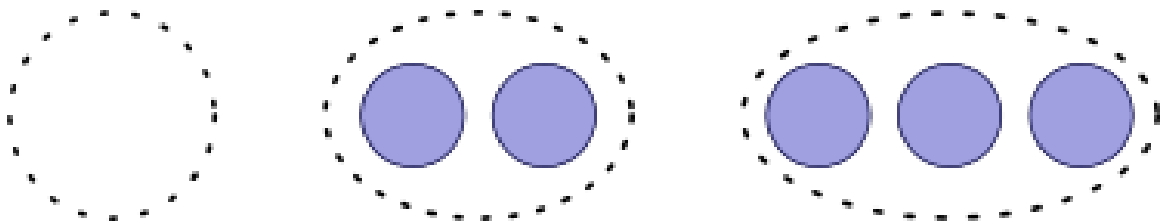


Figure 15

Two consecutive markers imply that the corresponding subset is empty. There is a natural bijection between an integer assignment and the graphical representation depicted above. To count the number of possible integer assignments, it then suffices to compute the number of ways to position the markers and the balls. In particular, there are $n + r - 1$ positions, n balls, and $r - 1$ markers. The number of ways to assign the markers is equal to the number of n -combinations of $n + r - 1$ elements,

$$\frac{(n+r-1)!}{n!(r-1)!}.$$

Example - Sampling with Replacement, without Ordering: An urn contains r balls numbered one through r . A ball is drawn from the urn and its number is recorded. The ball is then returned to the urn. This procedure is repeated a total of n times. The outcome of this experiment is a table that contains the number of times each ball has come in sight. We are interested in computing the number of distinct outcomes. This is equivalent to counting the ways a set with n elements can be partitioned into r subsets. The number of possible outcomes is therefore given by

$$\frac{(n+r-1)!}{n!(r-1)!}.$$

5.5 Computing Probabilities

In this section, we present two straightforward applications of combinatorics to computing the probability of winning the lottery.

Example - Pick 3 Texas Lottery: The Texas Lottery game *Pick 3* is easy to play. A player must pick three numbers from zero to nine, and choose how to play them: exact order, or any order. The Pick 3 balls are drawn using three air-driven machines. These machines use compressed air to mix and select each ball.

The probability of winning while playing the exact order is

$$\frac{1}{10} \frac{1}{10} \frac{1}{10} = \frac{1}{1000}.$$

The probability of winning while playing any order depends on the numbers selected. If three distinct numbers are selected then the probability of winning is $3/500$. If a number is repeated twice, the probability of winning is $3/1000$. While, if the same number is selected three times, the probability of winning becomes $1/1000$.

Example - Mega Millions Texas Lottery: To play the Mega Millions game, a player must select five numbers from 1 to 56 in the upper white play area of the play board, and one Mega Ball number from 1 to 46 in the lower yellow play area of the play board. All drawing equipment is stored in a secured on-site storage room. Only authorized drawings department personnel have keys to this door. Upon entry of the secured room to begin the drawing process, a lottery drawing specialist examines the security seal to determine if any unauthorized access has occurred. For each drawing, the Lotto Texas balls are mixed by four acrylic mixing paddles rotating clockwise. High speed is used for mixing and low speed for ball selection. As each ball is selected, it rolls down a chute into an official number display area. We wish to compute the probability of winning the Mega Millions Grand Prize, which require the correct selection of the five white balls plus the gold Mega ball. The probability of winning the Mega Millions Grand Prize is

$$\frac{51!5!}{56!} \frac{1}{46} = \frac{1}{175711536}.$$

6 Conditional Probability

In some situations we need a new kind of probability.

6.0.1 Examples

We shall throw two dice: the probability of getting double six is $1/36$. We have thrown the dice, but they are still under the cup: the probability of double six still is considered to be $1/36$. Now we have someone else look under the cup at the result, and they tell us that at least one of the dice is a 6. What will the probability be of double six? Although we ask for the "probability" of double six, it's not the original probability we're talking about. It's a new kind of probability of the event "double six" under the condition that we know that another event has occurred, i.e. the event that one of the dice shows six. We call it "conditional probability". It indicates how likely the event "double six" is amongst all events in which (at least) one of the dice shows six. In 11 of the 36 possibilities at least one of the dice shows six. Thus, the probability of one of the dice showing six is $11/36$ and the probability of double six is $1/36$. Hence, the conditional probability of double six, given one die shows six is $(1/36)/(11/36) = 1/11$. We write this conditional probability as:

$$P(\text{"double six"} \mid \text{"at least one six"}),$$

using the same letter P as for the original probability, and separating the event of which the conditional probability is indicated from the event that states the condition by a vertical line ("|").

As another example we pick at random an inhabitant of the UK. The probability of the chosen person to be a woman (W) is $1/2$, it's nothing else than the fraction women amongst the British. What if we are informed that the chosen person is working in a hospital (H)? It now comes down to the fraction of women amongst hospital staff, let's say 0.7. So we have:

$$P(W) = 0.5$$

and

$$P(W|H) = 0.7.$$

From these examples we learn:

6.1 Definition

The conditional probability of an event A, given the (occurrence of) the event B, is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

provided $P(B) > 0$. In the case where $P(B) = 0$, the conditional probability of A given B is meaningless, but for completeness we define $P(A|B) = 0$ for $P(B) = 0$

It follows from this formula that $P(A|B)P(B) = P(A \text{ and } B)$.

Independent events Intuitively, we would like two events, A and B, to be independent if $P(A|B) = P(A)$, that is, if A is as likely to occur whether or not B has occurred. In this case, the occurrence of A is independent of the occurrence of B.

However, $P(A|B) = P(A) \iff P(A \text{ and } B)/P(B) = P(A) \iff P(A \text{ and } B) = P(A)P(B)$.

Therefore, A is independent from B \iff B is independent from A $\iff P(A \text{ and } B) = P(A)P(B)$.

7 Random Variables

7.1 Random Variables: Definitions

Formally, a *random variable* on a probability space (Ω, Σ, P) is a measurable real function X defined on Ω (the set of possible outcomes)

$$X : \Omega \rightarrow \mathbb{R}$$

where the property of measurability means that for all real x the set

$$\{X \leq x\} := \{\omega \in \Omega \mid X(\omega) \leq x\} \in \Sigma$$

, i.e. is an event in the probability space.

7.1.1 Discrete variables

If X can take a finite or countable number of different values, then we say that X is a *discrete random variable* and we define the *mass function* of X , $p(x_i) = P(X = x_i)$, which has the following properties:

- $p(x_i) \geq 0$
- $\sum_i p(x_i) = 1$

Any function which satisfies these properties can be a mass function.

Variables

We need some way to talk about the objects of interest. In set theory, these objects will be sets; in number theory, they will be integers; in functional analysis, they will be functions. For these objects, we will use lower-case letters: a, b, c , etc. If we need more than 26 of them, we'll use subscripts.

Random Variable

an unknown value that may change everytime it is inspected. Thus, a random variable can be thought of as a function mapping the sample space of a random process to the real numbers. A random variable has either a associated probability distribution (discrete random variable) or a probability density function (continuous random variable).

Random Variable "X"

formally defined as a measurable function (probability space over the real numbers).

Discrete variable

takes on one of a set of specific values, each with some probability greater than zero (0). It is a finite or countable set whose probability is equal to 1.0.

Continuous variable

can be realized with any of a range of values (ie a real number, between negative infinity and positive infinity) that have a probability greater than zero (0) of occurring. $\Pr(X=x)=0$ for all X in \mathbb{R} . Non-zero probability is said to be finite or countably infinite.

7.1.2 Continuous variables

If X can take an uncountable number of values, and X is such that for all (measurable) A :

$$P(X \in A) = \int_A f(x) dx$$

we say that X is a *continuous variable*. The function f is called the (*probability*) *density* of X . It satisfies:

- $f(x) \geq 0 \forall x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f(x) dx = 1$

7.1.3 Cumulative Distribution Function

The (*cumulative*) *distribution function* (c.d.f.) of the r.v. X , F_X is defined for any real number x as:

$$F_X(x) = P(X \leq x) = \begin{cases} \sum_{i: x_i \leq x} p(x_i), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^x f(y) dy, & \text{if } X \text{ is continuous} \end{cases}$$

The distribution function has a number of properties, including:

- $\lim_{x \rightarrow -\infty} F(x) = 0$ and $\lim_{x \rightarrow \infty} F(x) = 1$
- if $x < y$, then $F(x) \leq F(y)$ -- that is, $F(x)$ is a non-decreasing function.
- F is right-continuous, meaning that $F(x+h)$ approaches $F(x)$ as h approaches zero from the right.

7.1.4 Independent variables

1

¹ [HTTP://EN.WIKIBOOKS.ORG/WIKI/CATEGORY%3A](http://en.wikibooks.org/wiki/Category%3A)

8 Important Distributions

8.1 Uniform Distribution

The **uniform distribution** is a model for "no preference". For example if we roll a fair dice numbered from 1 to 6, each side has equal probability of facing up, and the probability of each number would be 1/6. For the continuous analogon of such a distribution, there is an equal probability of picking any range of a given size. For example for a uniform distribution from 0 to 10, the probability of picking a real number between 1 and 2 is 1/10, the same as for picking a real number between 5.43 and 6.43.

This simple distribution forms the basis of much of the study of probability. It mirrors the type of activity observed from rolling a fair dice, or flipping a coin. This intuitively relates to many standard problems. A number of other distributions derive from this distribution. For instance, while any one roll of a dice has a uniform distribution, summing up the totals of rolling a dice lots of time, or taking their average, does not have a uniform distribution, but approximates a Gaussian distribution, which we will discuss later.

The uniform distribution on the interval [a,b] is a continuous distribution with probability density f given by:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a < x < b, \\ 0 & \text{for } x < a \text{ or } x > b, \\ \text{see below} & \text{for } x = a \text{ or } x = b. \end{cases}$$

The value of the density at a or b can be 0, $\frac{1}{b-a}$, or any other finite value without changing finite probabilities. For a more complete discussion of this point and for more information on the uniform distribution, see the UNIFORM DISTRIBUTION (CONTINUOUS)¹ page at Wikipedia.

A random variable X with this distribution is called a *random number* between a and b .

8.2 Binomial Distribution

Many random experiments are of the type where one counts the number of successes in a series of a fixed number of independently repeated trials which may result in either success or failure. The distribution of the number of successes is a **binomial distribution**. It is a discrete probability

¹ [HTTP://EN.WIKIPEDIA.ORG/WIKI/UNIFORM%20DISTRIBUTION%20%28CONTINUOUS%29](http://en.wikipedia.org/wiki/Uniform%20distribution%20%28continuous%29)

distribution with two parameters, traditionally indicated by n , the number of trials, and p , the probability of success.

A well known example of such an experiment is the repeated tossing of a coin and counting the number of times "heads" comes up.

One says that the random variable X follows the binomial distribution with parameters n and p , and writes $X \sim B(n, p)$ if the probability of getting exactly k successes is given by the probability mass function:

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \dots, n$ and where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

is the binomial coefficient.

The formula is easily derived from the fact that from n trials k should be successful, having probability p^k and $n - k$ failures, with probability $(1-p)^{n-k}$. This gives a specific outcome a probability of $p^k (1-p)^{n-k}$. An outcome with exactly k successes, however, may occur in a number of specific orders, i.e. one has to choose k of the n outcomes to be successful. This leads to the binomial coefficient as factor to be multiplied.

For more information see the [BINOMIAL DISTRIBUTION²](#) article at Wikipedia.

8.3 The Poisson Distribution

The **Poisson distribution** with parameter $\lambda > 0$ is a discrete probability distribution with mass function given by:

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

where k is a nonnegative integer.

The Poisson distribution is sometimes called: "the distribution of rare events", meaning it indicates the distribution of the number of events occurring in a fixed time interval, or fixed spatial area, where the occurrence of an event is "rare" in the sense that no two events may happen together.

The random variable X indicating this number of occurrences is called a *Poisson random variable*. The Poisson random variable is of fundamental importance in communication and queuing theory.

² [HTTP://EN.WIKIPEDIA.ORG/WIKI/BINOMIAL%20DISTRIBUTION](http://en.wikipedia.org/wiki/Binomial%20distribution)

It is possible to obtain the Poisson distribution as the limit of a sequence of binomial distributions.

Let λ be fixed and consider the binomial distributions with parameters n and λ/n :

$$p_n(k) = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n!}{n^k(n-k)!} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

In the limit, as n approaches infinity, the binomial probabilities tend to a Poisson probability with parameter λ :

$$\lim_{n \rightarrow \infty} p_n(k) = \lim_{n \rightarrow \infty} \frac{n!}{n^k(n-k)!} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

This result shows that the Poisson PMF can be used to approximate the PMF of a binomial distribution. Suppose that Y is a binomial random variable with parameters n and p . If n is large and p is small, the probability that Y equals k is not easily calculated, but may be approximated by a Poisson probability with parameter np :

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{(np)^k}{k!} e^{-np},$$

The above mentioned limiting process may be illustrated with the following example. Count the yearly number X of car accidents on an important roundabout. Let us assume on the average 8 accidents happen yearly. To calculate the distribution, we consider the 12 months and count a success when an accident has happened that month and else a failure. The number of successes will constitute a binomial random variable X_{12} with parameters $n=12$ and $p=8/12$. As the average number is 8, it is still likely some months will meet more than one accident, so we consider weeks and count the number of weeks with an accident as successful. This number X_{52} may be considered binomial distributed with parameters $n=52$ and $p=8/52$. While there may still be weeks with two accidents, we turn to days. The number X_{365} of days with success may be considered binomial distributed with parameters $n=365$ and $p=8/365$. The limiting distribution by considering hours, minutes, seconds, etc. gives the Poisson distribution of X with parameter 8. We also note an important condition on the occurrence of the events (accidents): we only count them correctly if in the end we may separate them in disjoint time intervals. So it is not allowed that two accidents happen the same moment.

8.4 Normal Distribution

The normal or Gaussian distribution is a thing of beauty, appearing in many places in nature. This probably is a result of the normal distribution resulting from the law of large numbers, by which a sum of many random variables (with finite variance) becomes a normally distributed random variable according to the central limit theorem.

Also known as the bell curve, the normal distribution has been applied to many social situations, but it should be noted that its applicability is generally related to how well or how poorly the

situation satisfies the property mentioned above, whereby many finitely varying, random inputs result in a normal output.

The formula for the density f of a normal distribution with mean μ and standard deviation σ is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

A rather thorough article in Wikipedia could be summarized to extend the usefulness of this book: [NORMAL DISTRIBUTION FROM WIKIPEDIA](http://en.wikipedia.org/wiki/Normal_distribution)³.

4

³ [HTTP://EN.WIKIPEDIA.ORG/WIKI/NORMAL%20DISTRIBUTION](http://en.wikipedia.org/wiki/Normal_distribution)

⁴ [HTTP://EN.WIKIBOOKS.ORG/WIKI/CATEGORY%3A](http://en.wikibooks.org/wiki/Category:3A)

9 Links

Hyperlinks to external resources on probability:

- [PROBABILITY AND STATISTICS EBOOK](#)¹
- [GRINSTEAD, SNELL: INTRODUCTION TO PROBABILITY](#)², 520 pages PDF, licensed under GNU FDL

[CATEGORY:PROBABILITY](#)³

¹ [HTTP://WIKI.STAT.UCLA.EDU/SOCR/INDEX.PHP/EBOOK](http://wiki.stat.ucla.edu/socr/index.php/Ebook)

² [HTTP://WWW.DARTMOUTH.EDU/~{}CHANCE/TEACHING_AIDS/BOOKS_ARTICLES/PROBABILITY_BOOK/BOOK.HTML](http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html)

³ [HTTP://EN.WIKIBOOKS.ORG/WIKI/CATEGORY%3APROBABILITY](http://en.wikibooks.org/wiki/Category%3AProbability)

10 Authors

Edits	User
6	ABIGOR ¹
18	ADRIGNOLA ²
1	ANGELAS ³
7	ANTONW ⁴
2	ARYAZAND ⁵
2	CARANDOL ⁶
8	CATASTROPHICTOAD ⁷
24	CHMBRLND ⁸
4	CHRISTOPHER PARHAM ⁹
1	CHUMJOELY ¹⁰
17	CRONIAN ¹¹
9	CYBORGTOSSE ¹²
3	DAN POLANSKY ¹³
1	DAVID SCHWEIN ¹⁴
1	DAVIDCARY ¹⁵
1	DIRK HÜNNIGER ¹⁶
1	EAGLE ¹⁷
2	EIBWEN ¹⁸
5	FELIX LABRECQUE ¹⁹
1	GJEFFERY ²⁰
1	JAMES11 ²¹

1 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:ABIGOR](http://en.wikibooks.org/w/index.php?title=User:ABIGOR)
2 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:ADRIGNOLA](http://en.wikibooks.org/w/index.php?title=User:ADRIGNOLA)
3 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:ANGELAS](http://en.wikibooks.org/w/index.php?title=User:ANGELAS)
4 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:ANTONW](http://en.wikibooks.org/w/index.php?title=User:ANTONW)
5 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:ARYAZAND](http://en.wikibooks.org/w/index.php?title=User:ARYAZAND)
6 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CARANDOL](http://en.wikibooks.org/w/index.php?title=User:CARANDOL)
7 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CATASTROPHICTOAD](http://en.wikibooks.org/w/index.php?title=User:CATASTROPHICTOAD)
8 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CHMBRLND](http://en.wikibooks.org/w/index.php?title=User:CHMBRLND)
9 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CHRISTOPHER_PARHAM](http://en.wikibooks.org/w/index.php?title=User:CHRISTOPHER_PARHAM)
10 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CHUMJOELY](http://en.wikibooks.org/w/index.php?title=User:CHUMJOELY)
11 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CRONIAN](http://en.wikibooks.org/w/index.php?title=User:CRONIAN)
12 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:CYBORGTOSSE](http://en.wikibooks.org/w/index.php?title=User:CYBORGTOSSE)
13 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:DAN_POLANSKY](http://en.wikibooks.org/w/index.php?title=User:DAN_POLANSKY)
14 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:DAVID_SCHWEIN](http://en.wikibooks.org/w/index.php?title=User:DAVID_SCHWEIN)
15 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:DAVIDCARY](http://en.wikibooks.org/w/index.php?title=User:DAVIDCARY)
16 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:DIRK_H%C3%BCNNIGER](http://en.wikibooks.org/w/index.php?title=User:DIRK_H%C3%BCNNIGER)
17 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:EAGLE](http://en.wikibooks.org/w/index.php?title=User:EAGLE)
18 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:EIBWEN](http://en.wikibooks.org/w/index.php?title=User:EIBWEN)
19 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:FELIX_LABRECQUE](http://en.wikibooks.org/w/index.php?title=User:FELIX_LABRECQUE)
20 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:GJEFFERY](http://en.wikibooks.org/w/index.php?title=User:GJEFFERY)
21 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=USER:JAMES11](http://en.wikibooks.org/w/index.php?title=User:JAMES11)

- 7 JGUK²²
- 2 JOMEGAT²³
- 1 KEEPEROFDAKEYS²⁴
- 1 LEO3421²⁵
- 3 MYT²⁶
- 1 MARKYPARKY²⁷
- 7 MIKE'S BOT ACCOUNT²⁸
- 21 MIKE.LIFEGUARD²⁹
- 5 NE8ZDYH³⁰
- 17 NIJDAM³¹
- 1 NORMAN WEINER³²
- 5 QUITEUNUSUAL³³
- 2 RECENT RUNES³⁴
- 4 RS2³⁵
- 7 TANNERSF³⁶
- 1 TRISTANREID³⁷
- 2 TWINKIE EATER91³⁸
- 2 VAN DER HOORN³⁹
- 2 WHITEKNIGHT⁴⁰

-
- 22 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:JGUK](http://en.wikibooks.org/w/index.php?title=User:JGUK)
 - 23 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:JOMEGAT](http://en.wikibooks.org/w/index.php?title=User:JOMEGAT)
 - 24 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:KEEPEROFDAKEYS](http://en.wikibooks.org/w/index.php?title=User:KEEPEROFDAKEYS)
 - 25 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:LEO3421](http://en.wikibooks.org/w/index.php?title=User:LEO3421)
 - 26 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:MYT](http://en.wikibooks.org/w/index.php?title=User:MYT)
 - 27 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:MARKYPARKY](http://en.wikibooks.org/w/index.php?title=User:MARKYPARKY)
 - 28 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:MIKE%27s_BOT_ACCOUNT](http://en.wikibooks.org/w/index.php?title=User:MIKE%27s_BOT_ACCOUNT)
 - 29 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:MIKE.LIFEGUARD](http://en.wikibooks.org/w/index.php?title=User:MIKE.LIFEGUARD)
 - 30 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:NE8ZDYH](http://en.wikibooks.org/w/index.php?title=User:NE8ZDYH)
 - 31 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:NIJDAM](http://en.wikibooks.org/w/index.php?title=User:NIJDAM)
 - 32 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:NORMAN_WEINER](http://en.wikibooks.org/w/index.php?title=User:NORMAN_WEINER)
 - 33 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:QUITEUNUSUAL](http://en.wikibooks.org/w/index.php?title=User:QUITEUNUSUAL)
 - 34 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:RECENT_RUNES](http://en.wikibooks.org/w/index.php?title=User:RECENT_RUNES)
 - 35 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:Rs2](http://en.wikibooks.org/w/index.php?title=User:Rs2)
 - 36 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:TANNERSF](http://en.wikibooks.org/w/index.php?title=User:TANNERSF)
 - 37 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:TRISTANREID](http://en.wikibooks.org/w/index.php?title=User:TRISTANREID)
 - 38 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:TWINKIE_EATER91](http://en.wikibooks.org/w/index.php?title=User:TWINKIE_EATER91)
 - 39 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:VAN_DER_HOORN](http://en.wikibooks.org/w/index.php?title=User:VAN_DER_HOORN)
 - 40 [HTTP://EN.WIKIBOOKS.ORG/W/INDEX.PHP?TITLE=User:WHITEKNIGHT](http://en.wikibooks.org/w/index.php?title=User:WHITEKNIGHT)

List of Figures

- GFDL: Gnu Free Documentation License. <http://www.gnu.org/licenses/fdl.html>
- cc-by-sa-3.0: Creative Commons Attribution ShareAlike 3.0 License. <http://creativecommons.org/licenses/by-sa/3.0/>
- cc-by-sa-2.5: Creative Commons Attribution ShareAlike 2.5 License. <http://creativecommons.org/licenses/by-sa/2.5/>
- cc-by-sa-2.0: Creative Commons Attribution ShareAlike 2.0 License. <http://creativecommons.org/licenses/by-sa/2.0/>
- cc-by-sa-1.0: Creative Commons Attribution ShareAlike 1.0 License. <http://creativecommons.org/licenses/by-sa/1.0/>
- cc-by-2.0: Creative Commons Attribution 2.0 License. <http://creativecommons.org/licenses/by/2.0/>
- cc-by-2.0: Creative Commons Attribution 2.0 License. <http://creativecommons.org/licenses/by/2.0/deed.en>
- cc-by-2.5: Creative Commons Attribution 2.5 License. <http://creativecommons.org/licenses/by/2.5/deed.en>
- cc-by-3.0: Creative Commons Attribution 3.0 License. <http://creativecommons.org/licenses/by/3.0/deed.en>
- GPL: GNU General Public License. <http://www.gnu.org/licenses/gpl-2.0.txt>
- PD: This image is in the public domain.
- ATTR: The copyright holder of this file allows anyone to use it for any purpose, provided that the copyright holder is properly attributed. Redistribution, derivative work, commercial use, and all other use is permitted.
- EURO: This is the common (reverse) face of a euro coin. The copyright on the design of the common face of the euro coins belongs to the European Commission. Authorised is reproduction in a format without relief (drawings, paintings, films) provided they are not detrimental to the image of the euro.
- LFK: Lizenz Freie Kunst. <http://artlibre.org/licence/lal/de>
- CFR: Copyright free use.
- EPL: Eclipse Public License. <http://www.eclipse.org/org/documents/epl-v10.php>

List of Figures

1		GFDL
2		GFDL
3	Original uploader was CHMBRLND ⁴¹ at EN.WIKIBOOKS ⁴²	GFDL
4		GFDL
5		GFDL
6		GFDL
7		GFDL
8		GFDL
9		GFDL
10		GFDL
11		GFDL
12		GFDL
13		GFDL
14		GFDL
15		GFDL

⁴¹ [HTTP://EN.WIKIBOOKS.ORG/WIKI/%3Awikibooks%3Aen%3Auser%3Achmbrlnd](http://en.wikibooks.org/wiki/%3Awikibooks%3Aen%3Auser%3Achmbrlnd)

⁴² [HTTP://EN.WIKIBOOKS.ORG](http://en.wikibooks.org)