Chapters 1 & 2 *

September 6, 2017

1 Introduction to Probability

The world is full of *random events* that we seek to understand.

- 'Will it rain tomorrow?'
- 'What are our chances of winning the lottery?'

The outcome of these questions **cannot** be exactly predicted in advance. But the occurrence of these phenomena is **not** completely haphazard. The outcome is uncertain, but there is a pattern to the outcome of these random phenomena.

Example: If I flip a fair coin, I do not know in advance whether the result will be a head or tail, but I can say in the long run approximately half the time the result will be heads.

Probability theory is a mathematical representation of random phenomena.

A *probability* is a numerical measurement of the uncertainty.

Applications:

- Medicine (chance of developing a genetic disease)
- Weather forecasting (chance of a major flood/earthquake in the next 10 years)

*Notes for Chapter 1 of DeGroot and Schervish adapted from Giovanni Motta's, Bodhisattva Sen's and Martin Lindquists notes for STAT W4109/W4105.

- Finance (predict stock prices)
- Computer science (analyze the performance of computer algorithms that employ randomization)
- Actuarial science (computing insurance risks and premiums)
- Provides the foundation for the study of statistics.

The mathematical theory of probability is often attributed to the French mathematicians *Pierre de Fermat* and *Blaise Pascal* (around 1650).

The theory of probability has its roots in attempts to analyze games of chance by *Gerolamo Cardano* in the sixteenth century – gambling was popular in those days. As the games became more complicated and the stakes increased, there was a need for mathematical methods for computing chance.

1.1 Interpretation of Probability

Through history there have been a variety of different definitions of probability.

Classical definition: Suppose a game has n *equally likely outcomes* of which m outcomes ($m < n$) correspond to 'winning'. Then the probability of winning is m/n.

Example: A box contains 4 blue balls and 6 red balls. If I randomly pick a ball from the box, what is the probability it is blue?

Solution: $n = 10, m = 4$, $\mathbb{P}(\text{blue ball}) = 4/10 = 2/5$.

The classical definition requires a game to be broken down into equally likely outcomes – ideal situation with very limited scope.

Relative frequency definition: Repeat a game a large number of times under the same conditions. The probability of winning is approximately equal to the number of wins in the repeated trials.

Example: Flip a coin 100,000 times. We get approximately 50,000 heads. $\mathbb{P}(\text{Heads}) = 1/2$.

Frequency approach is consistent with the classical approach.

Axiomatic Approach: Kolmogorov developed the first rigorous approach to probability (around 1930). He built up probability theory from a few fundamental axioms.

The book and this course will use these axioms to build up the theory of probability.

Modern research in probability is closely related to measure theory. This course avoids most of these issues and stresses intuition.

1.2 Sample Space and Events

One of the main objectives of statistics and probability theory is to draw conclusions about a population of objects by conducting an experiment.

Experiment: An experiment is any process, real or hypothetical, in which the possible *outcomes* can be identified ahead of time.

We must begin by identifying the possible outcomes of an experiment, or the *sample space*.

Sample space: The set, S, of all possible outcomes from a random experiment is called the *sample space*.

Example: Flip a coin. Two possible outcomes: Heads (H) or Tails (T). Thus, $S = \{H, T\}$.

Example: Roll a die. Six possible outcomes: $S = \{1, 2, 3, 4, 5, 6\}$.

Example: Roll two dice. 36 possible outcomes:

$$S = \{(i, j) : i = 1, \ldots, 6, j = 1, \ldots, 6\}.$$

Example: Lifetime of a battery. $S = \{x : 0 \leq x < \infty\}$.

Event: An *event* is a well-defined set of possible outcomes of an experiment, that is, any subset of S (including S itself).

If the outcome of the experiment is contained in an event E, then we say that E has occurred.

Example: Flip two coins. Four possible outcomes: $S = \{HH, HT, TH, TT\}$.

An event could be obtaining 'at least one H' in the two tosses. Thus,

$$E = \{HH, HT, TH\}.$$

Note E does not contain TT.

Example: Battery example. $E = \{x : x < 3\}$ (the battery lasts less than 3 hours)

1.3 Set Theory

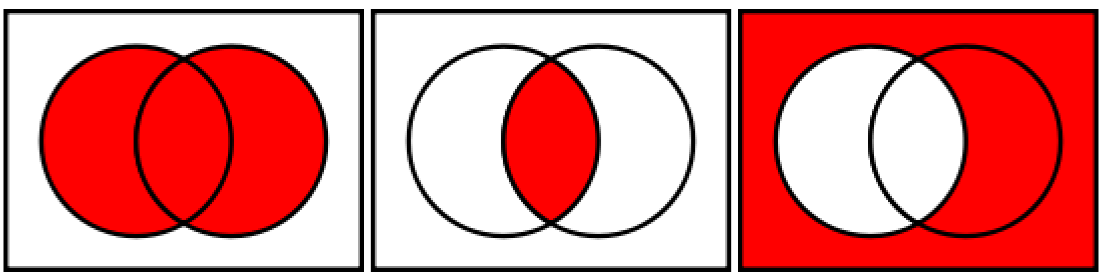
Given any two events (or sets) A and B, we have the following elementary set operations:

Union: The union of A and B, written $A \cup B$, is the set of elements that belong to *either* A *or* B *or both*;

$$A \cup B = \{x \in S : x \in A \text{ or } x \in B\}.$$

If A and B are events in an experiment with sample space S, then the union of A and B is the event that occurs if and only if A occurs or B occurs.

Venn diagram: The sample space is all the outcomes in a large rectangle. The events are represented as consisting of all the outcomes in given circles within the rectangle.



Intersection: The intersection of A and B, written $A \cap B$ or AB, is the set of elements that belong to both A and B:

$$AB = \{x \in S : x \in A \text{ and } x \in B\}.$$

If A and B are events in an experiment with sample space S, then the intersection of A and B is the event that occurs if and only if A occurs and B occurs.

Complementation: The complement of A, written A^c, is the set of all elements that are *not* in A;

$$A^c = \{x \in S : x \notin A\}.$$

If A is an event in an experiment with sample space S, then the complement of A is the event that occurs if and only if A does not occur.

Example: Select a card at random from a standard deck of cards, and note its suit: clubs ($\mathcal{C}$), diamonds ($\mathcal{D}$), hearts ($\mathcal{H}$) or spades ($\mathcal{S}$).

The sample space is $S = \{\mathcal{C}, \mathcal{D}, \mathcal{H}, \mathcal{S}\}$.

Some possible events are $A = \{\mathcal{C}, \mathcal{D}\}$, $B = \{\mathcal{D}, \mathcal{H}, \mathcal{S}\}$, $C = \{\mathcal{H}\}$.

$$\begin{aligned} A \cup B &= \{\mathcal{C}, \mathcal{D}, \mathcal{H}, \mathcal{S}\} = S \\ AB &= \{\mathcal{D}\} \\ A^c &= \{\mathcal{H}, \mathcal{S}\} \\ AC &= \emptyset \text{ (null event – event consisting of no outcomes)} \end{aligned}$$

Mutually exclusive (Disjoint) sets: If $AB = \emptyset$ then A and B are said to be *mutually exclusive* or *disjoint*.

For any two events A and B, if all of the outcomes in A are also in B, then we say that A is contained in B. This is written $A \subset B$.

Example: Roll a fair die. $S = \{1, 2, 3, 4, 5, 6\}$. Let $E = \{1, 3, 5\}$ and $F = \{1\}$, then $F \subset E$.

Note that if $A \subset B$ then the occurrence of A implies the occurrence of B.

Also if $A \subset B$ and $B \subset A$ then $A = B$.

Basic laws

The operations of forming unions, intersections and complements of events follow certain basic laws.

Commutative Laws: $E \cup F = F \cup E$; $EF = FE$.

Associative Laws: $(E \cup F) \cup G = E \cup (F \cup G)$; $(EF)G = E(FG)$.

Distributive Laws: $(E \cup F)G = EG \cup FG$; $(EF) \cup G = (E \cup G)(F \cup G)$.

De Morgan's Laws: $(A \cup B)^c = A^c B^c$ (*); $(AB)^c = A^c \cup B^c$.

Proof of (*): We will show that $(A \cup B)^c \subset A^cB^c$ and also $(A \cup B)^c \supset A^cB^c$. First, suppose that $x \in (A \cup B)^c$. Then,

$$\begin{aligned} x \in (A \cup B)^c &\Rightarrow x \notin A \cup B \\ &\Rightarrow x \notin A \text{ and } x \notin B \\ &\Rightarrow x \in A^c \text{ and } x \in B^c \\ &\Rightarrow x \in A^cB^c. \end{aligned}$$

Hence, $(A \cup B)^c \subset A^cB^c$. Now suppose that $x \in A^cB^c$. Then

$$\begin{aligned} x \in A^cB^c &\Rightarrow x \in A^c \text{ and } x \in B^c \\ &\Rightarrow x \notin A \text{ and } x \notin B \\ &\Rightarrow x \notin A \cup B \\ &\Rightarrow x \in (A \cup B)^c. \end{aligned}$$

Hence, $A^cB^c \subset (A \cup B)^c$.

Since $(A \cup B)^c \subset A^cB^c$ and $A^cB^c \subset (A \cup B)^c$, it must hold that $A^cB^c = (A \cup B)^c$.

[The proofs of the other laws are done in a similar manner.]

Often we are interested in unions and intersections of more than two events.

Definition:

$$\cup_{i \in I} A_i = \{x : x \in A_i \text{ for at least one } i \in I\}$$

and

$$\cap_{i \in I} A_i = \{x : x \in A_i \text{ for every } i \in I\}.$$

Generalized De Morgan's Laws:

(a) $(\cup_{i \in I} A_i)^c = \cap_{i \in I} A_i^c$

(b) $(\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c$.

1.4 Axioms of Probability

Consider an experiment whose sample space is S. For each event E of the sample space S we assume that a number $\mathbb{P}(E)$ is defined and satisfies the following three axioms.

Axiom 1: $0 \leq \mathbb{P}(E) \leq 1$.

Axiom 2: $\mathbb{P}(S) = 1$.

Axiom 3: For any sequence of *mutually exclusive* events $E_1, E_2, \ldots$,

$$\mathbb{P}(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

We refer to $\mathbb{P}(E)$ as the probability of the event E.

Proposition 1.1. $\mathbb{P}(\emptyset) = 0$.

Proposition 1.2. *For any finite sequence of mutually exclusive events* $E_1, E_2, \ldots, E_n$,

$$\mathbb{P}(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} \mathbb{P}(E_i).$$

Proposition 1.3. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.

Proof. E and E^c are mutually exclusive and $S = E \cup E^c$. Thus by Axioms 2 and 3,

$$1 = \mathbb{P}(S) = \mathbb{P}(E \cup E^c) = \mathbb{P}(E) + \mathbb{P}(E^c).$$

Hence, $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$. $\square$

Example 1: Let $E = S$. Then $E^c = \emptyset$. Thus,

$$\mathbb{P}(\emptyset) = 1 - \mathbb{P}(S) = 1 - 1 = 0.$$

Example 2: Roll a fair dice 3 times. What is the probability of at least one six, if we know that the probability of getting 0 sixes is 0.488?

$\mathbb{P}(\text{at least one six}) = 1 - \mathbb{P}(0 \text{ sixes}) = 1 - 0.488 = 0.512.$

Proposition 1.4. *If $E \subset F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$.*

Proof. Since $E \subset F$, $F = E \cup (E^c F)$.

Note that E and $E^c F$ are mutually exclusive, and Axiom 3 gives

$$\mathbb{P}(F) = \mathbb{P}(E) + \mathbb{P}(E^c F).$$

But $\mathbb{P}(E^c F) \geq 0$ by Axiom 1. Hence, $\mathbb{P}(F) \geq \mathbb{P}(E)$. $\square$

Proposition 1.5. $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(EF)$.

Proof. Note that $E \cup F = E \cup (E^c F)$, and E and $E^c F$ are mutually exclusive events. Thus,

$$\mathbb{P}(E \cup F) = \mathbb{P}(E \cup (E^c F)) = \mathbb{P}(E) + \mathbb{P}(E^c F),$$

which implies that $\mathbb{P}(E^c F) = \mathbb{P}(E \cup F) - \mathbb{P}(E)$. Now, $F = (EF) \cup (E^c F)$, where EF and $E^c F$ are mutually exclusive. Thus,

$$\mathbb{P}(F) = \mathbb{P}((EF) \cup (E^c F)) = \mathbb{P}(EF) + \mathbb{P}(E^c F),$$

which implies that $\mathbb{P}(E^c F) = \mathbb{P}(F) - \mathbb{P}(EF)$. Together, we have

$$\begin{aligned} \mathbb{P}(F) - \mathbb{P}(EF) &= \mathbb{P}(E \cup F) - \mathbb{P}(E) \\ \Rightarrow \mathbb{P}(E \cup F) &= \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(EF). \end{aligned}$$

□

Example: A store accepts either VISA or Mastercard. 50% of the stores customers have VISA, 30% have Mastercard and 10% have both.

(a) What percent of the stores customers have a credit card the store accepts?

Solution: Let A = customer has VISA; and B = customer has Mastercard. Then,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB) = 0.50 + 0.30 - 0.10 = 0.70.$$

(b) What percent of the stores customers have exactly one of the credit cards the store accepts?

Solution: $\mathbb{P}(A \cup B) - \mathbb{P}(AB) = 0.70 - 0.10 = 0.60$.

(c) What percent of the stores customers doesn't have a credit card the store accepts?

Solution: $\mathbb{P}(0 \text{ cards}) = 1 - \mathbb{P}(\text{at least one card}) = 1 - 0.70 = 0.30$.

Example: For three events E, F and G,

$\mathbb{P}(E \cup F \cup G) = \mathbb{P}(E) + \mathbb{P}(F) + \mathbb{P}(G) - \mathbb{P}(EF) - \mathbb{P}(EG) - \mathbb{P}(FG) + \mathbb{P}(EFG).$

Exercise: A doctor knows that $\mathbb{P}(\text{bacterial infection}) = 0.7$ and $\mathbb{P}(\text{viral infection}) = 0.4$. What is $\mathbb{P}(\text{both})$ if $\mathbb{P}(\text{bacterial} \cup \text{viral}) = 1$?

1.5 Finite Sample Spaces

Suppose that there are a finite number of outcomes in $S = \{s_1, s_2, \ldots, s_n\}$.

Define $p_i = \mathbb{P}(\{s_i\})$ as the probability function. Then,

$$p_i \geq 0, \sum_{i=1}^{n} p_i = 1$$

$$\text{If } A = \{s_{i_1}, s_{i_2}, \ldots, s_{i_k}\} \text{ then } \mathbb{P}(A) = \sum_{j=1}^{k} p_{i_j}.$$

1.6 Counting Methods

Many problems in probability theory can be solved by simply, counting the number of different ways an event can occur. The mathematical theory of counting is formally known as *combinatorial analysis*.

The Sum Rule: If a first task can be done in m ways and a second task in n ways, and if these tasks cannot be done at the same time, then there are $m + n$ ways to do either task.

Example: A library has 10 textbooks dealing with statistics and 15 textbooks dealing with biology. How many textbooks can a student choose from if he is interested in learning about either statistics or biology?

Solution: 10+15 = 25.

Basic principle of counting: If two successive choices are to be made with m choices in the first stage and n choices in the second stage, then the total number of successive choices is $n \times m$. (This is sometimes referred to as the product rule.)

Proof. We can write the two choices as (i, j), if the first choice is i and the second j; $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Let us enumerate all the possible choices:

$$\begin{aligned}(1,1), (1,2), \ldots, (1,n)\\ (2,1), (2,2), \ldots, (2,n)\\ \vdots\\ (m,1), (m,2), \ldots, (m,n)\end{aligned}$$

The set of possible choices consists of m rows, each row consisting of n elements. Hence $m \times n$ total choices. $\square$

The generalized basic principle of counting: If r successive choices are to be made with exactly n_j choices at each stage, then the total number of successive choices is $n_1 \times n_2 \times \ldots \times n_r$.

Example: A license plate consists of 3 letters followed by 2 numbers. How many possible license plates are there?

Solution: $26 \times 26 \times 26 \times 10 \times 10 = 1757600$.

Example (cont.): What if repetitions of letters and numbers are not allowed?

Solution: $26 \times 25 \times 24 \times 10 \times 9 = 1404000$.

Permutations

An **ordered** arrangement of the objects in a set is called a *permutation* of the set.

Example: How many ways can we rank three people: A, B and C?

Solution: $ABC, ACB, BAC, BCA, CAB, CBA$.

Each arrangement is called a permutation. There are 6 different permutations in this case.

Factorial: For an integer $n \geq 0$, $n! = n \times (n-1) \times (n-2) \ldots 3 \times 2 \times 1$. We make the assumption that $0! = 1$.

Note that $n! = n(n-1)!$

The number of different permutations of n distinct objects is $n \times (n-1) \times \ldots \times 1 = n!$

Example: How many different letter arrangements can be made of the word COLUMBIA.

Solution: $8!$

Example: How many ways can we choose a President, Vice-President and Secretary from a group consisting of five people?

Solution: President can be chosen 5 different ways; Vice-President can be chosen 4 different ways, Secretary can be chosen 3 different ways. Therefore, total number of ways is $5 \times 4 \times 3 = 60$.

The number of different permutations of r objects taken from n distinct objects is $n \times (n-1) \times \ldots \times (n-r+1)$.

Note that $n \times (n-1) \times \ldots \times (n-r+1) = n!/(n-r)!$.

Example: How many ways can we choose a President, Vice-President and Secretary from a group consisting of five people if A will only serve if he is President?

Solution: If A serves in the committee then the total number of ways the committee can be formed is $1 \times 4 \times 3 = 12$. If A doesn't serve then the total number of ways the committee can be $4 \times 3 \times 2 = 24$.

Hence total number of ways the committee can be formed is $12 + 24 = 36$.

Finally, what is the number of permutations of a set of n objects when certain objects are indistinguishable from one another?

Example: How many permutations of the letters in the set $\{M, O, M\}$ can we make?

There are 3! different ways to arrange a three letter word, when all the letters are different. But in this case the two Ms are indistinguishable from one another.

Order the objects as if they were distinguishable. Rewrite as (M_1, O, M_2). Then, there are 3! different ways to arrange these three letters:

$$M_1OM_2, M_1M_2O, OM_1M_2, OM_2M_1, M_2OM_1, M_2M_1O.$$

But M_1OM_2 is the same as M_2OM_1 when the subscripts are removed.

We need to compensate for this. How many ways can we arrange the 2 Ms in the two slots. (2!)

Idea: Divide out the arrangements that look identical.

Therefore, the total number of distinct arrangements is $3!/2! = 3$.

In general there are $n!/(n_1! \ldots n_r!)$ different permutations of n objects, of which n_1 are alike, n_2 are alike, etc.

Example: How many different letter arrangements can be made from the letters MISSISSIPPI?

Example: 1 M, 4 I's, 4 S's, 2 P's. Therefore number of different arrangements is

$$\frac{11!}{1!4!4!2!} = 34650.$$

Circular Permutation: The number of ways to arrange n distinct objects along a fixed (i.e., cannot be picked up out of the plane and turned over) circle is $(n-1)!$ (instead of the usual factorial $n!$ since all cyclic permutations of objects are equivalent because the circle can be rotated).

For example, of the $3! = 6$ permutations of three objects, the $(3-1)! = 2$ distinct circular permutations are $\{1, 2, 3\}$ and $\{1, 3, 2\}$.

Similarly, of the $4! = 24$ permutations of four objects, the $(4-1)! = 6$ distinct circular permutations are $\{1, 2, 3, 4\}$, $\{1, 2, 4, 3\}$, $\{1, 3, 2, 4\}$, $\{1, 3, 4, 2\}$, $\{1, 4, 2, 3\}$, and $\{1, 4, 3, 2\}$.

Combinations

A permutation is an ordered arrangement of r objects from a set of n objects.

Combination: A *combination* is an **unordered** arrangement of r objects from a set of n objects.

Example: Suppose instead of picking a President, Vice-President and Secretary we would just like to pick three people out of the group of five to be on the leadership board, i.e., how many teams of three can we create from a larger group consisting of five people?

Solution: In this case order does not matter. Each combination of three people is counted 3! times. We need to compensate for this. Therefore, required number of groups $= 5!/(2!3!)$.

When counting permutations *abc* and *cab* are different. When counting combinations they are the same.

Notation: $\binom{n}{r}$; $\binom{n}{r} = \frac{n!}{(n-r)!r!}$ for $0 \le r \le n$.

$\binom{n}{r}$ represent the number of ways we can choose groups of size r from a larger group of size n, when the order of selection is not relevant.

Homework: Property 1: $\binom{n}{r} = \binom{n}{n-r}$. Property 2: $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

Example: How many ways can I choose 5 cards from a deck of 52?

Solution: $\binom{52}{5} = 2598960$ (Total number of poker hands).

Example: A person has 8 friends, of whom 5 will be invited to a party.

(a) How many choices are there if 2 of the friends are feuding and will not attend together?

(b) How many choices if 2 of the friends will only attend together?

Solution:

(a) One of the feuding friends attends and the other does not: $\binom{2}{1}\binom{6}{4}$; or neither of the feuding friends attend: $\binom{6}{5}$.

Therefore, total number of choices is $2 \times 15 + 6 = 36$.

(b) The two friends attend together: $\binom{6}{3} = 20$; or the two friends do not attend: $\binom{6}{5} = 6$.

Therefore, total number of choices is $20 + 6 = 26$.

Example: A committee of 6 people is to be chosen from a group consisting of 7 men and 8 women. If the committee must consist of at least 3 women and at least 2 men, how many different committees are possible?

Solution: 3 women/3 men: $\binom{8}{3}\binom{7}{3}$; or 4 women/2 men: $\binom{8}{4}\binom{7}{2}$.

Total number of different committees is $\binom{8}{3}\binom{7}{3} + \binom{8}{4}\binom{7}{2} = 3430$.

The symbol $\binom{n}{k}$ is sometimes called a *binomial* coefficient.

1.7 The binomial theorem

If x and y are variables and n is a positive integer, then

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Example: $(x+y)^2 = \binom{2}{0}x^0y^2 + \binom{2}{1}x^1y^1 + \binom{2}{2}x^2y^0 = y^2 + 2xy + x^2$.

Example: Prove that $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

Solution: Set $x = 1$ and $y = 1$. By binomial theorem, $(1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k}$, i.e., $2^n = \sum_{k=0}^{n} \binom{n}{k}$.

Example: Suppose that a similar red balls and b similar black balls are to be arranged in a row. Number of distinguishable ways to order in a row is

$$\binom{a+b}{a} = \binom{a+b}{b}.$$

Why? Since the red balls will occupy a positions in the row, each different arrangement of the $a+b$ balls correspond to a different choice of the a (or b)

positions occupied by the red (or black) balls.

Example: Suppose that n, k are given. How many ways to split n indistinguishable objects into k distinguishable sets?

Solution: Visualize the balls in boxes, in a line: fix the outer walls, rearrange the balls and the separators.

If you x the outer walls of the first and last boxes, you can rearrange the separators and the balls using the binomial theorem.

There are n balls and $k-1$ separators (k boxes). Number of different ways to arrange the balls and separators is

$$\binom{n+k-1}{n} = \binom{n+k-1}{k-1}.$$

1.8 Multinomial Coefficients

Example: How many ways are there of assigning 10 police officers to 3 different tasks, if 5 must be on patrol, 2 in the office and 3 on reserve?

Solution: $\binom{10}{5} \times \binom{5}{2} \times \binom{3}{3} = \frac{10!}{5!2!3!}$.

In general there will $n!/(n_1!n_2!\ldots n_r!)$ ways to divide up n objects into r different categories so that there will be n_1 objects in category 1, n_2 in category 2, and so on, where $(n_1 + \ldots + n_r = n)$.

Notation: If $n_1 + n_2 + \ldots + n_r = n$, we define $\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! \ldots n_r!}$.

The symbol $\binom{n}{n_1, n_2, \ldots, n_r}$ is sometimes called a multinomial coefficient.

The multinomial theorem:

$$(x_1 + x_2 + \ldots + x_r)^n = \sum_{(n_1, \ldots, n_r): n_1 + \ldots + n_r = n} \binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}.$$

That is, the sum is over all nonnegative integer-valued vectors $(n_1, \ldots, n_r)$ such that $n_1 + \ldots + n_r = n$.

Example: 20 members of a club need to be split into 3 committees (A, B, C) of 8, 8, and 4 people, respectively. How many ways are there to split the club into these committees?

Solution: Number of ways to split is $\binom{20}{8,8,4} = \frac{20}{8!8!4!}$.

1.9 Sample Spaces having equally likely outcomes

The assignment of probabilities to sample points can be derived from the physical setup of an experiment. Consider a scenario where we have n sample points (outcomes), each *equally likely* to occur. Then each has a probability $1/n$.

Example: A fair coin, a fair die or a poker hand.

Assume an event consists of m sample points. Then the probability of the event occurring is m/n.

Example: Flip two fair coins. Then, $S = \{HH, HT, TH, TT\}$. Our sample space consists of 4 points, each of which is equally likely to occur. Thus,

$$\mathbb{P}(HH) = \frac{1}{4}.$$

Let E = "at least one head" $= \{HH, HT, TH\}$. Then, $\mathbb{P}(E) = 3/4$.

Let F = "exactly one head" $= \{HT, TH\}$. Then, $\mathbb{P}(E) = 2/4 = 1/2$.

Example: Roll two dice. The sample space $S = \{(i, j) : i = 1, \ldots, 6,\ j = 1, \ldots, 6\}$. There are 36 possible outcomes.

What is the probability that the sum of the two dice is less than or equal to 3?

Let E = "sum of the two dice are less than or equal to 3". Then,

$$E = \{(1, 1), (1, 2), (2, 1)\}$$

and $\mathbb{P}(E) = 3/36 = 1/12$.

Example: Two cards are chosen at random from a deck of 52 playing cards. What is the probability that they

(a) are both aces?

(b) are both spades?

Solution:

(a) Total number of ways to choose two cards is $\binom{52}{2}$. The total number of ways to choose two aces is $\binom{4}{2}$. Thus,

$$\mathbb{P}(\text{two aces}) = \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{1}{221}.$$

(b) Total number of ways to choose two spades is $\binom{13}{2}$. Thus,

$$\mathbb{P}(\text{two spades}) = \frac{\binom{13}{2}}{\binom{52}{2}} = \frac{1}{17}.$$

Example: Poker hands. Choose a five-card poker hand from a standard deck of 52 playing cards. What is the probability of four aces?

Solution: There are $\binom{52}{5} = 2598960$ possible hands.

Having specified that four cards are aces, there are 48 different ways of specifying the fifth card. Thus, $\mathbb{P}(\text{four aces}) = \frac{48}{2598960}$.

Exercise: (Playing Bridge) Players A, B, C, and D each get 13 cards. What is the probability $\mathbb{P}(A - 6\mathcal{H}, B - 4\mathcal{H}, C - 2\mathcal{H}, D - 1\mathcal{H})$?

Solution: $\frac{\binom{13}{6,4,2,1}\binom{39}{7,9,11,12}}{\binom{52}{13,13,13,13}}$.

1.9.1 The Birthday Problem

What is the probability that at least two people in this class share the same birthday?

Assumptions:

- 365 days in a year (no leap year).
- Each day of the year has an equal chance of being somebody's birthday (birthdays are evenly distributed over the year).

Let E = no two people share a birthday. Then E^c = at least two people share a birthday.

Use Proposition 1.3 to get

$\mathbb{P}(\text{at least two people share birthday}) = 1 - \mathbb{P}(\text{no two people share birthday})$.

What is the probability that no two people will share a birthday?

Solution: The first person can have any birthday. The second person's birthday has to be different. There are 364 different days to choose from, so the chance that two people have different birthdays is 364/365. That leaves 363 birthdays out of 365 open for the third person.

To find the probability that both the second person and the third person will have different birthdays, we have to multiply:

Table 1: Common birthday probability

n	$\mathbb{P}(E^c)$
10	0.117
15	0.253
20	0.411
23	0.507
50	0.970

$$\frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} = 0.9918.$$

If we want to know the probability that four people will all have different birthdays, we multiply again:

$$\frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} = 0.9836.$$

We can keep on going the same way as long as we want.

A formula for the probability that n people have different birthdays is

$$\frac{365-1}{365} \times \frac{365-2}{365} \times \frac{365-3}{365} \times \ldots \times \frac{365-n+1}{365} = \frac{365!}{(365-n)!} \times 365^{-n}.$$

Let $E =$ no two people share a birthday. Then,

$$\begin{aligned} \mathbb{P}(\text{two people share birthday}) = \mathbb{P}(E^c) &= 1 - \mathbb{P}(\text{no two people share birthday}) \\ &= 1 - \frac{365!}{(365-n)! \times 365^n}. \end{aligned}$$

Example: How large must a class be to make the probability of finding two people with the same birthday at least 50%?

Solution: It turns out that the smallest class where the chance of finding two people with the same birthday is more than 50% is a class of 23 people. (The probability is about 50.73%.)

1.9.2 Matching Problem

Suppose that each of N men at a party throws his hat into the center of the room. The hats are first mixed up, and then each man randomly selects a hat. What is the probability that none of the men select their own hat?

Version 2: You have n letters and n envelopes. Randomly put the letters into the envelopes. What is the probability that at least one letter will match its intended envelope?

Calculate the probability of at least one man selecting his own hat and use Proposition 1.3.

Let E_i be the event that the ith man selects his own hat, for $i = 1, 2, \ldots, N$.

Then,

$$\mathbb{P}(\text{at least one man selecting his own hat}) = \mathbb{P}(E_1 \cup E_2 \cup \ldots \cup E_N).$$

We use the following result: For any set of n events $A_1, A_2, \ldots, A_n$,

$$\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i A_j) + \sum_{i<j<k} \mathbb{P}(A_i A_j A_k) - \ldots$$

The outcome of this experiment can be seen as a vector of N elements, where the ith element is the number of the hat drawn by the ith man. For example $(1, 2, 3, \ldots, N)$ means that every man chooses his own hat.

The sample space S is the set of all permutations of the set $\{1, 2, \ldots, N\}$. Hence S has $N!$ elements. There are $N!$ possible outcomes.

Thus, $\mathbb{P}(E_i) = \frac{(N-1)!}{N!} = \frac{1}{N}$ (permute everyone else if just E_i is in the right place); for each $i = 1, 2, \ldots, N$.

Also, $\mathbb{P}(E_i E_j) = \frac{(N-2)!}{N!} = \frac{1}{N(N-1)}$ for $i \neq j$. Similarly, $\mathbb{P}(E_{i_1} E_{i_2} \ldots E_{i_k}) = \frac{(N-k)!}{N!}$.

Thus,

$$\begin{aligned}\mathbb{P}(E_1 \cup E_2 \cup \ldots \cup E_N) &= N\frac{1}{N} - \binom{N}{2}\frac{(N-2)!}{N!} + \binom{N}{3}\frac{(N-3)!}{N!} - \ldots \\ &\quad + (-1)^{N+1}\binom{N}{N}\frac{(N-N)!}{N!}.\end{aligned}$$

The kth term is $\binom{N}{k}\frac{(N-k)!}{N!} = \frac{1}{k!}$. Thus the required probability is

$$1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots + (-1)^{N+1}\frac{1}{N!}.$$

Recall: Taylor series for $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$

For $x = -1$, $e^{-1} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \ldots$

Therefore, when N is large, the probability converges to $1 - e^{-1} \approx 0.63$.

2 Conditional probabilities

Sometimes if we are given additional information we can reduce our sample space.

If we know an event F occurs, then the probability of another event E occurring may change.

Example: It snows tomorrow, given it is snowing today.

If event F occurs what is the probability that event E occurs?

This probability is called the *conditional probability* of E given F, written $\mathbb{P}(E|F)$. $E|F$ symbolizes the event E given the event F.

Example: Flip two coins. Then $S = \{HH, HT, TH, TT\}$, then $\mathbb{P}(HH) = 1/4$.

Let F = at least one H. If we know F occurs, then our new sample space is $S' = \{HH, HT, TH\}$. Thus, $\mathbb{P}(HH|F) = 1/3$.

The probability of getting two H is greater if we already know F.

Let G = the first flip is H; $S'' = \{HH, HT\}$. Then $\mathbb{P}(HH|G) = 1/2$. Thus, the probability of two heads is greater yet if we already know G.

Definition: If $\mathbb{P}(B) > 0$, then

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}.$$

[Draw Venn diagram.]

What if A and B are disjoint? If A and B are disjoint then $\mathbb{P}(AB) = 0$. Hence, $\mathbb{P}(A|B) = 0$. If B occurs, then A cannot occur.

Example: What is the probability of rolling an even number with a single die, given the die roll is 3 or less?

Solution: Here $S = \{1, 2, 3, 4, 5, 6\}$. Let A = roll an even number = $\{2, 4, 6\}$, B = roll a 3 or less = $\{1, 2, 3\}$. Then $AB = \{2\}$ and

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{2\})}{\mathbb{P}(\{1, 2, 3\})} = \frac{1}{3}.$$

Example: What is the probability the die roll is 3 or less, given the die roll is an even number?

Solution: $\mathbb{P}(B|A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\{2\})}{\mathbb{P}(\{2,4,6\})} = \frac{1}{3}$.

Example: A family has two children. What is the probability that both are boys, given that at least one is a boy?

Solution: Let (b, g) denote the event that the older child is a boy and the younger child a girl. Then

$$S = \{(b, b), (b, g), (g, b), (g, g)\}.$$

Assume that all outcomes are equally likely.

Let E = both children are boys, and F = at least one of them is a boy. Then,

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(EF)}{\mathbb{P}(F)} = \frac{\mathbb{P}(\{(b, b)\})}{\mathbb{P}(\{(b, b), (b, g), (g, b)\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

What is the probability that both are boys given that the younger is a boy?

Let G = the younger is a boy. Then,

$$\mathbb{P}(E|G) = \frac{\mathbb{P}(EG)}{\mathbb{P}(G)} = \frac{\mathbb{P}(\{(b, b)\})}{\mathbb{P}(\{(g, b), (b, b)\})} = \frac{1/4}{1/2} = \frac{1}{2}.$$

Example: At a fast food restaurant, 90% of the customers order a hamburger. If 72% of the customers order a hamburger and french fries, what is the probability that a customer who orders a hamburger will also order french fries?

Solution: Let A = customer orders a hamburger, and B = customer orders french fries. Then

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)} = \frac{0.72}{0.9} = 0.8.$$

Conditional probabilities can be used to calculate the probability of the intersection of two events.

If $\mathbb{P}(E) > 0$, and $\mathbb{P}(F) > 0$, then we can rewrite $\mathbb{P}(E|F) = \frac{\mathbb{P}(EF)}{\mathbb{P}(F)}$ to yield

$$\mathbb{P}(EF) = \mathbb{P}(E|F)\mathbb{P}(F) = \mathbb{P}(F|E)\mathbb{P}(E).$$

Example: A box contains 8 blue balls and 4 red balls. We draw two balls from the box without replacement. What is the probability that both are red?

Solution: Let E = first ball red, and let F = second ball red. Then,

$$\mathbb{P}(\text{both balls are red}) = \mathbb{P}(EF) = \mathbb{P}(E)\mathbb{P}(F|E).$$

Note that $\mathbb{P}(E) = 4/12$ and $\mathbb{P}(F|E) = 3/11$. Thus,

$$\mathbb{P}(EF) = \mathbb{P}(E)\mathbb{P}(F|E) = \frac{4}{12} \times \frac{3}{11} = \frac{1}{11}.$$

Another way to solve this problem (using the methods learnt previously):

$$\mathbb{P}(\text{two red}) = \frac{\binom{4}{2}}{\binom{12}{2}} = \frac{4!/(2!2!)}{12!/(10!2!)} = \frac{1}{11}.$$

Multiplication Rule:

$$\mathbb{P}(E_1 E_2 \ldots E_n) = \mathbb{P}(E_1)\ \mathbb{P}(E_2|E_1)\ \mathbb{P}(E_3|E_1 E_2) \cdots \mathbb{P}(E_n|E_1 E_2 \ldots E_{n-1}).$$

Proof.

$$\begin{aligned} & \mathbb{P}(E_1)\ \mathbb{P}(E_2|E_1)\ \mathbb{P}(E_3|E_1 E_2) \cdots \mathbb{P}(E_n|E_1 E_2 \ldots E_{n-1}) \\ = \ & \mathbb{P}(E_1) \frac{\mathbb{P}(E_1 E_2)}{\mathbb{P}(E_1)} \frac{\mathbb{P}(E_1 E_2 E_3)}{\mathbb{P}(E_1 E_2)} \cdots \frac{\mathbb{P}(E_1 E_2 \ldots E_n)}{\mathbb{P}(E_1 E_2 \ldots E_{n-1})} \\ = \ & \mathbb{P}(E_1 E_2 \ldots E_n) \end{aligned}$$

□

Example: A box contains five red balls and five green balls. Four balls are sampled without replacement. What is the probability of drawing four red balls?

Solution: Let $E_i = i$'th ball is red; for $i = 1, 2, 3, 4$. Then,

$$\mathbb{P}(E_1 E_2 E_3 E_4) = \mathbb{P}(E_1)\ \mathbb{P}(E_2|E_1)\ \mathbb{P}(E_3|E_1 E_2)\mathbb{P}(E_4|E_1 E_2 E_3).$$

Here, $\mathbb{P}(E_1) = 5/10 = 1/2$, $\mathbb{P}(E_2|E_1) = 4/9$, $\mathbb{P}(E_3|E_1 E_2) = 3/8$, and $\mathbb{P}(E_4|E_1 E_2 E_3) = 2/7$. Thus,

$$\mathbb{P}(E_1 E_2 E_3 E_4) = \frac{1}{2} \times \frac{4}{9} \times \frac{3}{8} \times \frac{2}{7} = \frac{24}{1008} = \frac{1}{42}.$$

Solve this problem using the method of counting learnt previously!

It is very important to note that conditional probabilities are probabilities in their own right. By this we mean that conditional probability satisfy all the properties of ordinary probabilities:

(1) $0 \leq \mathbb{P}(E|F) \leq 1$

(2) $\mathbb{P}(S|F) = 1$

(3) If E_i, $i = 1, 2, \ldots$, are mutually exclusive events, then

$$\mathbb{P}(\cup_{i=1}^{\infty} E_i | F) = \sum_{i=1}^{\infty} \mathbb{P}(E_i | F).$$

[Easy to show.]

Thus, $\mathbb{P}(\cdot | F)$ is till a probability, but the sample space has been redefined.

All the propositions we derived for regular probabilities also hold for conditional probabilities, e.g., $\mathbb{P}(A^c | B) = 1 - \mathbb{P}(A | B)$.

2.1 Bayes' Formula

Let E and F be events. We can express E as $E = (EF) \cup (EF^c)$. Therefore, by the axioms of probability,

$$\mathbb{P}(E) = \mathbb{P}(EF) + \mathbb{P}(EF^c).$$

But

$$\mathbb{P}(EF) = \mathbb{P}(E|F)\mathbb{P}(F) \qquad \text{and} \qquad \mathbb{P}(EF^c) = \mathbb{P}(E|F^c)\mathbb{P}(F^c).$$

So, $\mathbb{P}(E) = \mathbb{P}(E|F)\mathbb{P}(F) + \mathbb{P}(E|F^c)\mathbb{P}(F^c)$. Another way of writing this is

$$\mathbb{P}(E) = \mathbb{P}(E|F)\mathbb{P}(F) + \mathbb{P}(E|F^c)(1 - \mathbb{P}(F)).$$

This formula enables us to calculate the probability of an event by conditioning upon whether or not a second event occurs.

Very useful!

Example: Box 1 contains two white balls and three blue balls, while Box 2 contains three white and four blue balls. A ball is drawn at random from Box 1 and put into Box 2, and then a ball is picked at random from Box 2 and examined. What is the probability it is blue?

Solution: Let $A =$ the ball chosen from Box 1 is blue, and let $B =$ ball chosen from Box 2 is blue. Then,

$$\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)$$

where $\mathbb{P}(B|A) = 5/8$, $\mathbb{P}(A) = 3/5$, $\mathbb{P}(B|A^c) = 4/8$, $\mathbb{P}(A^c) = 2/5$,

$$\mathbb{P}(B) = \frac{5}{8} \times \frac{3}{5} + \frac{4}{8} \times \frac{2}{5} = \frac{23}{40}.$$

An important consequence is to note is that (**Bayes' theorem**),

$$\mathbb{P}(F|E) = \frac{\mathbb{P}(EF)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|F)\mathbb{P}(F)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|F)\mathbb{P}(F)}{\mathbb{P}(E|F)\mathbb{P}(F) + \mathbb{P}(E|F^c)\mathbb{P}(F^c)}.$$

Example: A lab test is 95% correct in detecting a certain disease when the disease is actually present. However the test also yields a false positive for 1% of the healthy people tested. If 0.5% of the population has the disease, what is the probability that somebody has the disease, given that he tests positive.

Solution: Let $A =$ person has the disease, and $B =$ person tests positive. Then

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

Here, $\mathbb{P}(B|A) = 0.95$, $\mathbb{P}(A) = 0.005$,

$\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c) = 0.95 \times 0.005 + 0.01 \times 0.995 = 0.0147.$

Also, $\mathbb{P}(A|B) = \frac{0.95\times 0.005}{0.0147} = 0.323.$

Question: What if we test twice, i.e., what is the probability that somebody has the disease, given that he tests positive twice?

Solution: Let $B' =$ person tests positive twice. Then, $\mathbb{P}(B'|A) = \mathbb{P}(B|A)\mathbb{P}(B|A) = 0.95^2$, and

$$\begin{aligned}\mathbb{P}(B') &= \mathbb{P}(B'|A)\mathbb{P}(A) + \mathbb{P}(B'|A^c)\mathbb{P}(A^c)\\ &= 0.95 \times 0.95 \times 0.005 + 0.01 \times 0.01 \times 0.995 = 0.0046.\end{aligned}$$

Thus,

$$\mathbb{P}(A|B') = \frac{\mathbb{P}(B'|A)\mathbb{P}(A)}{\mathbb{P}(B')} = \frac{0.95 \times 0.95 \times 0.005}{0.0046} = 0.98.$$

Example: In answering a multiple choice question the student either knows the answer or guesses. Let p be the probability the student knows the answer and $(1-p)$ the probability he guesses. Assume that a student who guesses gets the question correct with probability $1/m$, where m are the number of choices. What is the probability that the student knew the answer to the question, given that he gets the correct answer?

$P(B|\bar{A}) = \frac{1}{m}$

Solution: Let $A =$ student knows answer, and $B =$ student gets answer correct. ... Show that $\mathbb{P}(A|B) = mp/(1-p+mp)$.

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

$P(B|A)\cdot P(A)$
$= p \cdot 1$

$P(B|\bar{A})\cdot P(\bar{A}) = \frac{1}{m}\cdot(1-p)$

$P(AB) = \frac{P(B|A)\cdot P(A)}{P(B|\bar{A})\cdot P(\bar{A}) + P(B|A)\cdot P(A)}$

$$p(A|B) = \frac{p}{\frac{1}{m}(1-p)+p} = \frac{p}{\frac{(1-p)}{m}+\frac{mp}{m}} = \frac{mp}{1-p+mp}.$$

2.2 Law of Total Probability

This is the denominator in the Bayes' ratio. We have so far seen that,

$$\mathbb{P}(E) = \mathbb{P}(E|F)\mathbb{P}(F) + \mathbb{P}(E|F^c)\mathbb{P}(F^c).$$

In general, if $F_1, F_2, \ldots, F_n$ are mutually exclusive events such that $\cup_{i=1}^n F_i = S$ then

$$\mathbb{P}(E) = \sum_{i=1}^n \mathbb{P}(EF_i) = \sum_{i=1}^n \mathbb{P}(E|F_i)\mathbb{P}(F_i).$$

2.3 Bayes' Formula

A general version of this equation is given by Bayes' Formula:

Suppose that $F_1, F_2, \ldots, F_n$ are mutually exclusive events, such that $S = \cup_{i=1}^n F_i$; then

$$\mathbb{P}(F_i|E) = \frac{\mathbb{P}(E|F_i)\mathbb{P}(F_i)}{\sum_{j=1}^n \mathbb{P}(E|F_j)\mathbb{P}(F_j)}.$$

Example: There are three machines (A, B and C) at a factory that are used to make a certain product. Machine A makes 25% of the products, B makes 35% and C 40%. Of the products that A makes, 5% are defective, while for B 4% are defective and for C 2%. The products from the different machines are mixed up and sent to the customer.

(a) What is the probability that a customer receives a defective product?

(b) What is the probability that a randomly chosen product was made by A, given the fact that it is defective?

Solution: Let E = product defective, and F_1 = product comes from A, F_2 = product comes from B, F_3 = product comes from C.

(a)

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E|F_1)\mathbb{P}(F_1) + \mathbb{P}(E|F_2)\mathbb{P}(F_2) + \mathbb{P}(E|F_3)\mathbb{P}(F_3) \\ &= 0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.40 \\ &= 0.0345.\end{aligned}$$

(b) $\mathbb{P}(F_1|E) = \frac{0.25\times 0.5}{0.0345} = 0.36.$

Assume that $F_i, i = 1, \ldots, n$, are competing hypothesis. Then the Bayes' formula gives us a way to compute the conditional probabilities of these hypotheses when additional evidence E becomes available.

If E has occurred, what is the probability that F_i occurred as well. Bayes' formula gives us a way to update our personal probability.

In the context of Bayes' formula, $\mathbb{P}(F_i)$ is the prior probability of F_i and $\mathbb{P}(F_i|E)$ is the posterior probability of F_i. 先验

后验

Example: A plane is missing and it is assumed that it has crashed in any of three possible regions with equal probability. Assume that the probability of finding the plane in Region 1, if in fact the plane is located in there, is 0.8. What is the probability the plane is in the i-th region given that the search of Region 1 is unsuccessful?

Solution: Let R_1 = plane is in region 1, R_2 = plane is in region 2, R_3 = plane is in region 3, and let E = search of region 1 unsuccessful.

Note that $\mathbb{P}(R_1|E) = \frac{\mathbb{P}(E|R_1)\mathbb{P}(R_1)}{\sum_{i=1}^{3} \mathbb{P}(E|R_i)\mathbb{P}(R_i)}$.

Now, $\mathbb{P}(R_i) = 1/3$, for $i = 1, 2, 3$; $\mathbb{P}(E|R_1) = 0.2 = 1/5$, and $\mathbb{P}(E|R_i) = 1$ for $i = 2, 3$.

Then,

$$\mathbb{P}(R_1|E) = \frac{\frac{1}{5} \times \frac{1}{3}}{\frac{1}{5} \times \frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \frac{\frac{1}{15}}{\frac{11}{15}} = \frac{1}{11}.$$

Similarly, for $i = 2, 3$,

$$\mathbb{P}(R_i|E) = \frac{1 \times \frac{1}{3}}{\frac{1}{5} \times \frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \frac{\frac{1}{3}}{\frac{11}{15}} = \frac{5}{11}.$$

2.4 Independent Events

In most of the examples we have studied so far the occurrence of an event A changes the probability of an event B occurring.

In some cases the occurrence of a particular event, B, has no effect on the probability of another event A.

In this situation we can say, $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Since, $\mathbb{P}(A|B) = \mathbb{P}(AB)/\mathbb{P}(B)$ we have that $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$.

Definition: Two events A and B, are statistically *independent* if

$$\mathbb{P}(AB) = \mathbb{P}(A)\ \mathbb{P}(B).$$

If two events are not independent they are said to be dependent.

Two events are independent if the occurrence of one does not change the probability of the other occurring.

Example: If A and B are mutually exclusive are they independent? they must be dependent

Example: Flip two coins and let $E = H$ on first toss, and $F = H$ on second toss. Then

$$\mathbb{P}(EF) = \mathbb{P}(\{HH\}) = \frac{1}{4}, \qquad \mathbb{P}(E) = \frac{1}{2}, \qquad \mathbb{P}(F) = \frac{1}{2}.$$

As $\mathbb{P}(E)\mathbb{P}(F) = \mathbb{P}(EF)$, E and F are independent.

Theorem 2.1. *If E and F are independent, so are*

(a) *E and F^c;*

(b) *E^c and F;*

(c) *E^c and F^c.*

c)

$P(E^cF^c) = (1-p(E))(1-p(F))$

$= 1-p(F) - p(E) + p(F)\cdot p(E)$

$= 1-p(F) - p(E) + p(E)\cdot p(F)$

$(F \cup E) - (F \cap E) +$
$(F^c \cap E^c) = \Omega$

Proof. (a) As E and F are independent, $\mathbb{P}(EF) = \mathbb{P}(E)\mathbb{P}(F)$.

Now, as $E = EF \cup EF^c$, and EF and EF^c are mutually exclusive, so $\mathbb{P}(E) = \mathbb{P}(EF) + \mathbb{P}(EF^c)$. Thus,

$$\mathbb{P}(EF^c) = \mathbb{P}(E) - \mathbb{P}(EF) = \mathbb{P}(E) - \mathbb{P}(E)\mathbb{P}(F) = \mathbb{P}(E)(1 - \mathbb{P}(F)) = \mathbb{P}(E)\mathbb{P}(F^c).$$

b) $F = EF \cup E^cF$

$P(E^cF) = p(F) - p(EF) = p(F) - p(E)\cdot p(F)$

$= p(F)(1-p(E)) = p(F)\cdot p(E^c)$ ※

Hence the result is proved.

Homework: Show (b) and (c). $\square$

Definition: The three events E, F and G are said to be *independent* if

$$\mathbb{P}(EFG) = \mathbb{P}(E)\mathbb{P}(F)\mathbb{P}(G), \qquad \mathbb{P}(EF) = \mathbb{P}(E)\mathbb{P}(F), \qquad \mathbb{P}(EG) = \mathbb{P}(E)\mathbb{P}(G),$$

$$\text{and } \mathbb{P}(FG) = \mathbb{P}(F)\mathbb{P}(G).$$

Example: Roll two dice. Let A = first die is a 4, B = second die is a 3, C = sum of two dice is $7 = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$.

Then, $\mathbb{P}(A) = 1/6$, $\mathbb{P}(B) = 1/6$, $\mathbb{P}(C) = 1/6$ and $\mathbb{P}(A)\mathbb{P}(B) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$. Now $\mathbb{P}(AB) = \mathbb{P}(\{(4,3)\}) = \frac{1}{36}$. Thus, $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$.

Using similar arguments, we can see that $\mathbb{P}(AC) = \mathbb{P}(A)\mathbb{P}(C)$ and $\mathbb{P}(BC) = \mathbb{P}(B)\mathbb{P}(C)$ (Verify for yourself).

$P(AC) = \{(4,3)\} = \frac{1}{36}$

$P(A) = \frac{1}{6}$ $\quad P(C) = (\frac{1}{6})$

But, $\mathbb{P}(ABC) = \mathbb{P}(\{(4,3)\}) = \frac{1}{36}$. But $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{256}$. So A, B, C are **NOT** independent.

Definition: The events $E_1, E_2, \ldots, E_n$ are *independent* if, for every subset $E_{i_1}, E_{i_2}, \ldots, E_{i_r}$, $2 \leq r \leq n$, of these events, we have

$$\mathbb{P}(E_{i_1} E_{i_2} \ldots E_{i_r}) = \mathbb{P}(E_{i_1})\ \mathbb{P}(E_{i_2}) \cdots\ \mathbb{P}(E_{i_r}).$$

2.4.1 Independent Trials

Sometimes an experiment consists of performing a sequence of sub-experiments. If each sub-experiment is identical, then the sub-experiments are called trials.

If the trials are independent they are called *independent trials*.

Example: Roll a dice five times. What is the probability of rolling a five exactly 2 times?

Solution: Let E_i = rolling a five in the i-th trial, $i = 1, 2, \ldots, 5$. Then $\mathbb{P}(E_i) = 1/6$, and $\mathbb{P}(E^c) = 5/6$. Thus,

$$\mathbb{P}(E_1^c E_2 E_3 E_4^c E_5^c) = \frac{5}{6}\,\frac{1}{6}\,\frac{1}{6}\,\frac{5}{6}\,\frac{5}{6} = \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3.$$

There are $\binom{5}{2}$ different ways to get exactly 2 fives in five rolls. Thus,

$$\mathbb{P}(\text{Rolling exactly 2 fives in 5 tries}) = \binom{5}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3.$$

Example: What is the probability of at least one 6 in 5 tries?

Solution: Note that,

$$\mathbb{P}(1 \text{ six in 5 tries}) = \binom{5}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^4$$

$$\mathbb{P}(2 \text{ sixes in 5 tries}) = \binom{5}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3.$$

We can add up the probabilities to get the result.

Alternatively, the required probability is

$$1 - \mathbb{P}(0 \text{ sixes in 5 tries}) = 1 - \binom{5}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^5 = 1 - \left(\frac{5}{6}\right)^5 = 0.598.$$

In general, if we have a series of trials where in each trial the event E consists of a success and the event E^c consists of a failure, where $\mathbb{P}(E) = p$ and $\mathbb{P}(E^c) = (1-p)$, then the probability of exactly k successes in the n trials is:

$$\mathbb{P}(\text{exactly } k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Theorem 2.2. *If the events $E_1, E_2, \ldots, E_n$ are independent and $\mathbb{P}(E_i) = p_i$, the probability that at least one of the events occur is*

$$1 - (1-p_1)(1-p_2)\cdots(1-p_n).$$

Corollary 2.3. *If the events E_i are independent and each of them occur with a probability p, the probability that at least one of the events occur is $1-(1-p)^n$.*

Example: A person takes a certain risk at 1000 different opportunities, each time independent of the last. An accident occurs each time with probability 1/1000. What is the probability that an accident will take place at least one of the 1000 opportunities?

Solution: $1 - \left(1 - \frac{1}{1000}\right)^{1000} = 0.63 \approx 1 - e^{-1}$.

Example: A series of independent trials are performed which result in success with probability p and failure with probability $1-p$. What is the probability that exactly m failures occur before the n'th success, $m \geq 0$?

Solution: In order for exactly m failures to occur before n successes, it is equivalent to that there are exactly $n-1$ successes in the first $m+n-1$ trials. Hence, the required probability is

$$\binom{m+n-1}{m} p^n (1-p)^m.$$