Contents

Chapter 2

FUNDAMENTAL CONCEPTS OF PROBABILITY

2.1 Introduction

Objective of chapter:

(1) Define the meaning of probability

(2) Develop various relationships among probabilities

Basic concepts:

- **Chance (or random) experiment**
- **Outcome**

Example:

A coin tossing game:

What is the probability of the head showing up?

Past approaches to probability: [1]

1. Personal approach
2. Relative frequency approach
3. Equally likely approach

Remark: These approches are:

(a) Useful for assgining probability to an outcome of chance experiment

(b) NOT suitable for the theoretical basis of probability

$\Longrightarrow$ **Axiomatic approach** [2] [3]

[1]To be discussed in the next section.

[2]Developed by the Russian mathematician called Kolmogorov in 1933.

[3]Provides the rules, or properties that probabilities should satisfy, but NOT the means of assigning values of probability.

2.2 Approaches to Probability

Basic approach:

Rather than *defining* probability, we take *axiomatic* approach...

$\Longrightarrow$ providing the list of properties that probability must satisfy

Definitions of terminology

1. **chance experiment**
   - Not necessarily actually performed, but is often conceptual
   - called random, since the results do not obey deterministic laws of nature
   - denoted by boldface uppercase letter, e.g. $\mathbf{E}$
2. **outcome**
   - result of chance experiment
   - denoted by lowercase Greek letter, e.g. ω
3. **event**
   - collection of outcomes
   - denoted by uppercase italic letter, e.g. A
4. **sample space**
   - collection of *all possible* outcomes for a given chance experiment
   - denoted by uppercase italic S, i.e. S
5. **Venn diagram**
   - graphic representation a chance experiment
   - sample space: rectangle
   - outcome: point within rectangle
   - event: closed curve in rectangle

Example:

Suppose we perform a game of casing a die, and denote this chance experiment as **D**. The outcomes, i.e. numbers from 1 to 6 could be represented as ω_1, ω_2, ω_3, ω_4, ω_5, ω_6 repectively, and the event of an odd number showing up can be expressed as $O = \{\omega_1, \omega_3, \omega_5\}$.

Figure 2.1: Venn diagram for visualizing outcomes and events for a chance experiment.

Previous approaches to probability:

(1) Personal approach: subjective

1. Determining whether to take umbrella with you or not in the morning
2. Different from person to person
3. *NOT represent the idea of probability, but rather an educated guessing based on experience*

(2) Equally likely approach

1. A chance experiment of throwing a die
2. Assumes each outcome has equal chance of occurrence: restrictive
3. $P(A)$= (# of outcomes in A)/(# of outcomes in S)
4. *Use of probability in defining probability*: main flaw
5. Nevertheless, useful in many situations

Example 2.1

In a game of tossing a fair die, find the probability of getting an odd number using the equally likely approach.

Solution:

This is the event O in the above example. There are 6 outcomes in the sample space S, and 3 of them belong to the event O. Therefore, we have:

$$P(O) = \frac{\text{\# of outcomes in } O}{\text{\# of outcomes in } S} = \frac{3}{6} = \frac{1}{2}$$

(3) Relative frequency approach

1. A chance experiment of tossing a coin repeatedly
2. Expect half of heads and half of tails: assuming the coin and tosser are fair
3. For an arbitrary event A, the definition of $P(A)$ in relative frequency sense is:

   $$P(A) \triangleq \lim_{n \to \infty} \frac{n_A}{n}$$

   where n_A is the number of occurrence of event A in n repetition of experiment.
4. *The limit can never be reached*: main dilemma

Figure 2.2: Relative frequency for a sequence of 1000 coin tossing: (a) first 100; (b) entire 1000.

2.3 Axioms of Probability

Remark:

Due to the flaws and dilemmas of past approaches to probability as pointed out in the previous section, we adopt a different approach:

$\Longrightarrow$ instead of stating what the probability *is*, we give 3 *properties* that it must satisfy.

<u>**PROBABILITY AXIOMS**</u>

For a chance experiment **E**, the probability of an event A in a sample space S is defined as a number $P(A)$ which satisfies the following three axioms:

1. $P(A)$ is a non-negative number, i.e.,

$$P(A) \geq 0$$

2. The probability of S (i.e. the certain event) is unity:

$$P(S) = 1$$

3. If two events A and B are mutually exclusive (or disjoint) [4], the probability of the event $\{A \text{ or } B\}$ is the sum of their probabilities:

$$P(A \text{ or } B) = P(A) + P(B)$$

Figure 2.3: Sample space S showing two disjoint events A and B, and two events C and D which are not.

[4] Two events are called mutually exclusive if they have NO common outcomes.

NOTE:

(1) Probability axioms constitute the building blocks for a self-consistent theory of probability

(2) The axioms DO NOT relate probabilities to physical events

(3) Probabilities usually are assigned by other approaches $\ni$: relative frequency and equally likely etc..

$\Longrightarrow$ *should obey the axioms*

<u>**Verification:**</u> (Relative frequency approach)

Suppose that we perform a chance experiment **E** repeatedly n times , and observed the event A occurring n_A times out of it. Then, if n is sufficiently large enough, the probability of A in a relative frequency sense is defined as:

$$P(A) = \frac{n_A}{n}$$

1. **Axiom #1:**

$$P(A) = \frac{n_A}{n} > 0 \qquad \text{since } \; n > 0, \;\; n_A > 0$$

2. **Axiom #2:**

   Let $A = S$, then

$$P(A) = P(S) = \frac{n}{n} = 1$$

3. **Axiom #3:**

   Assume two events A and B are mutually exclusive. Then the event $\{A \text{or} B\}$ will be observed $n_A + n_B$ times, since they cannot happen simultaneously. Therefore, we have:

$$\begin{aligned} P(A \text{ or } B) &= \frac{n_A + n_B}{n} \\ &= \frac{n_A}{n} + \frac{n_B}{n} \\ &= P(A) + P(B) \end{aligned}$$

Example 2.2 *Self study*

2.4 Set Theory

Remark:

Deveolpment of various theories and relations are fundamentally based on the *set theory.*

$\Longrightarrow$ We first look at and review the set theory

Definition 2.1 Set and element*:*
Sets are collections of objects, with the objects in the set called elements.

Examples:

1. $S_1 = \{1, 3, 5, \cdots\}$ is the set of non-negative, odd integers
2. $S_2 = \{\text{head, tail}\}$ is the set of outcomes in tossing a coin

NOTE: Sets and Events:
Sets and *elements* in set theory correspond to the *events* and *outcomes* in probability theory!!!

Definition 2.2 Element of a set*:*
Indication of an element ξ_i belonging to a set A is denoted by:

$$\xi_i \in A$$

where ξ_j is not contained in a set A is represented as:

$$\xi_j \notin A$$

Definition 2.3 Subset*:*
Set B is called a subset *of a set A if every element in B is contained in A, and denoted by:*

$$B \subset A$$

Definition 2.4 Equality*:*
Two sets A and B are called equal *if:*

$$B \subset A \quad \textit{and} \quad A \subset B$$

Definition 2.5 Proper subset*:*
If $B \subset A$, but $B \neq A$, then B is called the proper subset *of A.*

Definition 2.6 Null set*:*
Set containing no element is called the null set*, and denoted by ϕ*

Definition 2.7 Universal set*:*
Set which contains all of the element of interest is called the universal set *or* universe*, and denoted by S.*

(cf) Universal set corresponds to the sample space!!!

Definition 2.8 Union*:*
The union *of two sets C and D is defined and denoted as follows:*

$$E = C \cup D \triangleq \{x \mid x \in C \ \ \textit{and/or} \ \ x \in D\}$$

Definition 2.9 Intersection*:*
The intersection *of two sets C and D is defined and denoted as follows:*

$$F = C \cap D \triangleq \{\omega \mid \omega \in C \ \ AND \ \ \omega \in D\}$$

Definition 2.10 Complement*:*
The complement *of a set A is defined and denoted as follows:*

$$A^c = \triangleq \{\omega \mid \omega \ni A \ \ but \ \ \omega \in S\}$$

(cf) Complement of a set A is also denoted as A' or $\bar{A}$.

Definition 2.11 Disjoint sets*:*
Two sets A and B are called disjoint *if:*

$$A \cap B = \phi$$

Definition 2.12 Countable set*:*
A set is called countable *if its elements can be put into one-to-one correspondence with the positive integers.*

(cf) A set which is NOT countable is called *uncountable*.

Basic Set Properties:

1. Identity property

$$A \cup S = S$$

$$A \cup \phi = A$$

$$A \cap S = A$$

$$A \cap \phi = \phi$$

2. Commutative property

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

3. Associative property

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

4. Distributive property

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

5. Complement property

$$A \cup A^c = S \quad \text{and} \quad A \cap A^c = \phi$$

6. DeMorgan's law

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

Note: Set relations can usually, and easily verified by using the Venn diagrams.

Figure 2.4: Venn diagrams for illustrating set relationships.

Example 2.3 *Self study*

Definition 2.13 Partition of S:
Subsets $A_1, A_2, \ldots, A_m$ are called the partition *of the universal set S if:*

$$A_i \cap A_j = \phi \quad (\forall\, i \textit{ and } j,\ i \neq j)$$

$$A_1 \cup A_2 \cup \cdots \cup A_m = S$$

Figure 2.5: Partition of a sample space S into disjoint events.

Definition 2.14 Cartesian product*:*
The Cartesian product [5] *of two sets A and B is a set of pairs which is denoted and defined as:*

$$C = A \times B \triangleq \{(\alpha_i, \beta_j) \mid \alpha_i \in A \;\; \textit{and} \;\; \beta_j \in B\}$$

Example 2.4

Consider tossing of two coins, where possible outcomes for each coin are sets $A = \{h_1, t_1\}$ and $B = \{h_2, t_2\}$ respectively. Determine the Cartesian product of events A and B

Solution:

From the definition of the cartesian product, we have:

$$A \times B = \{h_1h_2, h_1t_2, t_1h_2, t_2h_2\}$$

[5] The Cartesian product is sometimes called the *order space*, since the order of the pair is important!!!

2.5 Derived Probability Relationships

Remark:

Every probability relationships MUST be proved based only on the probability axioms!!!

1. **Relation # 1:**

$$P(A^c) = 1 - P(A)$$

proof:

From the complement property of sets, we have

$$A \cup A^c = S \quad \text{and} \quad A \cap A^c = \phi$$

Since A and A^c are disjoint, applying the axiom 2 and axiom 3, we get

$$1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$$

Q.E.D.

2. **Relation # 2:**

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

proof:

From the Venn diagram in figure2.6, we can easily verify the following set relations, which are expressed as the union of two *disjoint sets*:

$$A \cup B = A \cup (B \cap A^c)$$

$$B = (A \cap B) \cup (B \cap A^c)$$

Applying the axiom 3, we have:

$$P(A \cup B) = P(A) + P(B \cup A^c)$$

$$P(B) = P(A \cap B) + P(B \cap A^c)$$

Substacting above two equations proves the relation.

Q.E.D.

Figure 2.6: Venn diagrams for representing sets as a union of two disjoint sets.

Example 2.5 *Self study*

Example 2.6

Two *independent* random number generator produced numbers U and V, which are equally likely to be anywhere in the interval $[0, 1]$. Two events A and B are given as follows:

$$\begin{aligned} A &= \{0 \leq U \leq 1,\ 0.5 \leq V \leq 1\} \\ B &= \{0.5 \leq U \leq 1,\ 0 \leq V \leq 1\} \end{aligned}$$

Find $P(A \cup B)$.

Solution:

From the Venn diagram in figure2.7, it follows:

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{2}, \quad P(A \cap B) = \frac{1}{4}$$

Applying the relation #2, we have

$$P(A \cup B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}$$

Figure 2.7: Venn diagram for the probability of the union of two NOT disjoint sets.

2.6 Conditional Probabilities and Statistical Independence

Definition 2.15 Conditional probability:
The conditional probability of event A given that event B has occurred is defined as:

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

NOTE:

1. Similaly, the conditional probability of event B given that event A has occurred is:

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

2. Also, the probability of event A and B simultaneously occurring can be expressed using conditional probabilities as:

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$$

3. Conditional probability satisfies the three axioms of probability!!!

   proof: assignment

4. Conditional probability of event A given B, i.e. $P(A|B)$ corresponds to confining the sample space from S to B, (refer to fingure2.8 below)

   (cf)

   Probability of a certain event A is equivalent to the conditional probability of event A given S, i.e.

$$P(A) = P(A|S)$$

Figure 2.8: Venn diagram illustrating the conditional probability.

Example 2.7

Find the conditional probability of obtaining two heads(event B) when flipping a coin twice given that at least one head was obtained(event A).

Solution:

The sample space is $S = \{t_1t_2, t_1h_2, h_1t_2, h_1h_2\}$ and $P(A) = \frac{3}{4}$ and $P(A \cap B) = \frac{1}{4}$ using equally likely approach of paobability. Therefore we have:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

Note that that sample space is now shrinked to event A whose has 3 outcomes among which only one outcome corresponds to event B, i.e. the probability is $\frac{1}{3}$!!!

Example 2.8

In example2.6 define additional two events C and D as follows:

$$C = \{U > V\}$$

$$D = \{0 \leq U \leq 1,\ 0 \leq V \leq 0.5\}$$

Find the conditional probabilities $P(A|C)$ and $P(A|D)$.

Solution:

From the venn diagram in figure2.9 below, we have:

Figure 2.9: Venn diagram for example 2.8.

$$P(A) = P(C) = P(D) = 0.5$$

$$P(A \cap C) = 0.125$$

$$P(D \cap C) = 0.375$$

Applying the definition of the conditional probability, we get respective probability as follows:

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{0.125}{0.5} = 0.2$$

and

$$P(D|C) = \frac{P(D \cap C)}{P(C)} = \frac{0.375}{0.5} = 0.75$$

Note: Notice that $P(A) = 0.5 > P(A|C) = 0.25$ whereas $P(D) = 0.5 < P(D|C) = 0.75$. The formers are called *a priori* probability and the latters are called *a posteriori* probability. Sometimes it could be $P(A) = P(A|B)$, and this will be the foundation of the statistical independence concept below!!!

Definition 2.16 Statistical independence*:*

Two events S and B are said to be statistically independent *if:*

$$P(A \cap B) = P(A)\, P(B)$$

Example:

Let events A, B and C defined as follows:

$$A = \{\text{the weather is rainy in the morning}\}$$

$$B = \{\text{you take the umbrella with you}\}$$

$$C = \{\text{you attend the probability and random process class}\}$$

Are events A and B statistically independent?
In the similar manner, are events A and C statistically independent, or should they be?

Answer:

Example 2.9 *Self study*

2.7 Total probability and Bayes' Theorem

Theorem 2.1 Theorem of total probability:

Let the events $A_1, A_2, \cdots, A_m$ be a partition of the sample space S, i.e. they are munually exclusive, and the union of them is S. Then, the probability of any event B can be expressed as:

$$P(B) = \sum_{i=1}^{m} P(B \cap A_i) = \sum_{i=1}^{m} P(B|A_i)P(A_i)$$

Proof:

Using the idenity property and the dsitributive property of sets, any event B can be represented as:

$$\begin{aligned} B &= B \cap S \\ &= B \cap (A_1 \cup A_2 \cup \cdots \cup A_m) \\ &= (B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_m) \end{aligned}$$

From figure2.10, notice that $B \cap A_i$ are mutually exclusive, since the events A_i's are mutually exclusive. Therefore, applying the axiom #3, we get:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots P(B \cap A_m)$$

Replacing the probability $P(B \cap A_i)$'s with the conditional prbabilities below, the theorem follows.

$$P(B \cap A_i) = P(B|A_i)P(A_i)$$

Q.E.D.

Figure 2.10: Partition of sample space S with an event B: total probability law.

Example 2.10

Box #1 contains 8 nickels and 2 quarters, whereas box #2 has 5 nickels and 20 quarters. You select a box at random and draw a coin. What is the paprobability that the coin you draw is a quarter?

Solution:

Let events B_1, B_2, and Q be respectively defined as follows:

$$B_1 = \{\text{select box } \#1\}$$

$$B_2 = \{\text{select box } \#2\}$$

$$Q = \{\text{draw a quarter}\}$$

Then, we have:

$$P(B_1) = P(B_2) = \frac{1}{2} \quad \text{and} \quad P(Q|B_1) = \frac{1}{5} \quad \text{and} \quad P(Q|B_2) = \frac{4}{5}$$

Note that $Q \cap B_1$ and $Q \cap B_2$ are disjoint events, and applying the total probability theorem, the probability of drawing a quarter is then:

$$\begin{aligned} P(Q) &= P(Q|B_1)P(B_1) + P(Q|B_2)P(B_2) \\ &= \frac{1}{5} \cdot \frac{1}{2} + \frac{4}{5} \cdot \frac{1}{2} = \frac{1}{2} \end{aligned}$$

Exercise: repeat above example with $P(B_1) = p$ and $P(B_2) = 1 - p$, and plot the probability $P(Q)$ as a function of p.

Theorem 2.2 Bayes' Theorem*:*

The relationship below is known as *Bayes' theorem*:

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

Proof: *assignment (easy!)*

Example 2.11

Certain class is given an examination. The probability of a student preparing for the exam is 0.7, and of the students who prepared, the probability of passing the exam is 0.9, while if a student does not study for the exam, his or her chance of passing the exam is 0.05.

Given that a student did not pass the exam, what is the probability that he or she studied?

Solution:

Let W and V be the events that a student studied and that a student pass the exam respectively. Then, the probability we want to compute is as follows:

$$P(W|V^c) = \frac{P(V^c|W)P(W)}{P(V^c)}$$

where we are given that $P(W) = 0.7$, $P(V^c|W) = 1 - P(V|W) = 1 - 0.9 = 0.1$, and the denominator follows from the total probability theorem as:

$$\begin{aligned} P(V^c) &= P(V^c|W)P(W) + P(V^c|W^c)P(W^c) \\ &= (0.1)(0.7) + (0.95)(0.3) = 0.355 \end{aligned}$$

Therefore, we have:

$$P(W|V^c) = \frac{0.07}{0.355} = 0.197$$

Theorem 2.3 Genelaized Bayes' Theorem*:*

The generalization of the Bayes' theorem to the partition of a sample space into m disjoint events $A_!, A_2, \ldots, A_m$ is in the following form:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{m} P(B|A_j)P(A_j)}, \quad 1 \le i \le m$$

Proof: *In the Bayes' theorem, replace $P(B)$ with the total probability theorem...*

Example 2.12

In the process of testing defects for certain manufacured devices, let events F and D be that the device fails the test and that the device is identified to be defective. Given the following probabilities:

$$P(D^c) = 0.95$$

$$P(F|D) = 0.98$$

$$P(F|D^c) = 0.03$$

what are the probabilities $P(D|F)$ and $P(D^c|F^c)$?

Solution:

From the given probabilities, we get:

$$P(D) = 1 - P(D^c) = 1 - 0.95 = 0.05$$

$$P(F^c|D) = 1 - P(F|D) = 1 - 0.98 = 0.02$$

$$P(F^c|D^c) = 1 - P(F|D^c) = 1 - 0.03 = 0.97$$

Applying the total prob. theorem,

$$\begin{aligned} P(F) &= P(F|D)P(D) + P(F|D^c)P(D^c) \\ &= (0.98)(0.05) + (0.03)(0.95) = 0.078 \end{aligned}$$

Now, from the Bayes's theorem,

$$\begin{aligned} P(D|F) &= \frac{P(F|D)P(D)}{P(F)} \\ &= \frac{(0.98)(0.05)}{0.078} = 0.632 \end{aligned}$$

and

$$\begin{aligned} P(D^c|F^c) &= \frac{P(F^c|D^c)P(D^c)}{P(F^c)} \\ &= \frac{(0.97)(0.95)}{0.922} = 0.999 \end{aligned}$$

(cf) Note that the probability that the device fails the test even if it is NOT defective [6] is:

$$P(F|D^c) = 1 - P(F^c|D^c) = 1 - 0.97 = 0.03$$

Example 2.13 Monte Hall problem: *self study*

[6] This is called the type I error, or the probability of false alarm, while $P(F^c|D)$ is called the type II error, or the probability of miss.

2.8 Counting Techniques

Remark: In equally likely approach for probabilty assignment:

$\Longrightarrow$ need to count the # of favorable outcomes corresponding to the event is *necessary*

1. Multiplication principle

(e.g.) Suppose you have 3 suits, 2 ties, and 2 shoes. How many different selections of clothing do you have?

solution: Obviously from the tree diagram below, you have $3 \times 2 \times 2 \times = 12$ different selections to choose.

Figure 2.11: Possible clothing selections.

Example 2.14

How many different combinations of security code exist for a three thumbwheel briefcase, where each thumbwhhel have numbers from 0 to 9?

Solution:

10 ways we can choose for the first number, and for each first number there are 10 ways we can choose for the second number ets....., i.e. $10 \times 10 \times 10 \times = 1000$ different combinations.

2. Permutations

(e.g.) How many different ways are there to draw 5 hearts *in a given order* from a 52 card deck?

solution: The first heart can be any one of 13 possibilities, and the second heart can be any of remaining 12 hearts, and so on..., for a total of:

$$13 \times 12 \times 11 \times 10 \times 9 = \frac{13!}{(13-5)!} = \frac{13!}{8!}$$

Definition 2.17 Permutation*:*
A permutation is the number of ways that we can line up *a set of objects.*

NOTE:

1. If there are n objects, we can permute them in $n!$ ways. [7]
2. The number of permutations of n objects taken m at a time is:

$$P_m^n = \frac{n!}{(n-m)!}$$

Example 2.15

Find the number of different ways of making password using five alphabets, where the same letter cannot be repeated.

Solution:

Permutation of size 5 from a 26 objects set:

$$P_5^{26} = \frac{26!}{(26-5)!} = 26 \times 15 \times 24 \times 23 \times 22 = 7,893,600$$

[7] Imagine it as filling slots or boxes, where once a slot is filled with an pbject, it can never be used again.

3. Combinations

(e.g.) How many different ways are there to draw 5 hearts from a 52 card deck where the *order does not count*?

solution: Among the permutation P_5^{13}, there are $5 \times 4 \times 3 \times 2 \times 1 = 5!$ containing the same cards in a different order. Since we are not interested in order, we divide P_5^{13} by 5!, i.e.

$$\frac{13 \times 12 \times 11 \times 10 \times 9}{5 \times 4 \times 3 \times 2 \times 1} = \frac{13!}{(13-5)!5!} = \frac{13!}{8!5!}$$

Definition 2.18 Combination*:*
A combination is a permutation in which the order is of no consequences.

NOTE:

1. If there are n objects, there are only one way of selecting n objects out of n, i.e.:
$$\frac{n!}{n!} = 1$$

2. The number of combinations of n objects taken m at a time is:
$$C_m^n = \begin{pmatrix} n \\ m \end{pmatrix} = \frac{n!}{(n-m)!m!}$$
and it is called the *binomial coefficient of algebra*.

3. Useful properties of binomial coefficient:
$$\begin{pmatrix} n \\ k \end{pmatrix} = \begin{pmatrix} n \\ n-k \end{pmatrix}, \quad \begin{pmatrix} n \\ k \end{pmatrix} = \begin{pmatrix} n-1 \\ k \end{pmatrix} + \begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$$

4. Let $P(\text{head}) = p$ and $P(\text{tail}) = 1 - p$ for a coin. If we flip the coin n times in a row, the probability of getting k heads out of n flips is the given by [8]:
$$\begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}$$

[8] This is called the *binomial probability distribution* and will be discussed in Chapter 3.

Example 2.16

Find the probability of flush in a 5 cards poker game.

Solution:

There exits 4 possible suits, each of which having the number of possible 5 card hands given by:

$$\begin{pmatrix} 13 \\ 5 \end{pmatrix} = \frac{13!}{5! \times 8!}$$

where the total number of possible 5 card hands from 52 card deck is

$$\begin{pmatrix} 52 \\ 5 \end{pmatrix} = \frac{52!}{5! \times 47!}$$

Using the equally likely definition of the probability, we have:

$$\begin{aligned} P(\text{flush}) &= \frac{4 \times \frac{13!}{5! \times 8!}}{\frac{52!}{5! \times 47!}} = \frac{4 \times 13! \times 5! \times 47!}{5! \times 8! \times 52!} \\ &= \frac{4 \times 13 \times 12 \times 11 \times 10 \times 9}{52 \times 51 \times 50 \times 49 \times 48} = 0.001981 \end{aligned}$$

Example 2.17 *Self study*

Example 2.18 *Self study*