Chapter 1

Introduction

1.1 Defining Probability

The Classical Definition

The probability of an event is

$$\frac{\text{the number of ways the event may occur}}{\text{the total number of possible outcomes}}$$

provided all outcomes are equally likely.

Example 1.1.1

The probability of a fair dice landing on 3 is $1/6$ because there is one way in which the dice may land on 3 and 6 total possible outcomes of faces the dice may land on. The sample space of the experiment, $\mathbb{S}$, is $\{1, 2, 3, 4, 5, 6\}$ and the event occurs in only one of these six outcomes.

The main limitation of this definition is that it demands that the outcomes of a sample space are equally likely. This is a problem since a definition of "likelyhood" (probability) is needed to include this postulate in a definition of probability itself.

The Relative Frequency Definition

The probability of an event is the limiting proportion of times that an event occurs in a large number of repetitions of an experiment.

Example 1.1.2

The probability of a fair dice landing on 3 is $1/6$ because after a very large series of repetitions (ideally infinite) of rolling the dice, the fraction of times the face with 3 is rolled tends to $1/6$.

The main limitation of this definition is that we can never repeat a process indefinitely so we can never truly know the probability of an event from this definition. Additionally, in some cases we cannot even obtain a long series of repetitions of processes to produce an estimate due to restrictions on cost, time, etc.

The Subjective Definition

The probability of an event occurring is a measure of how sure the person making the statement is that the event will occur.

Example 1.1.3

The probability that a football team will win their next match can be predicted by experts who regard all the data of past matches and current situations to provide a subjective probability.

This definition is irrational and leads to many people having different probabilities for the same events, with no clear “right” answer. Thus, by this definition, probability is not an objective science.

Probability Model

To avoid many of the limitation of the definitions of probability, we can instead treat probability as a mathematical system defined by a set of axioms. Thus, we can ignore the numerical values of probabilities until we consider a specific application. The model is defined as follows

- A sample space of all possible outcomes of a random experiment is defined.
- A set of events, to which we may assign probabilities, is defined.
- A mechanism for assigning probabilities to events is specified.

Chapter 2

Mathematical Probability Models

2.1 Sample Spaces

A sample space, $\mathbb{S}$, is a set of distinct outcomes for an experiment or process, with the property that in a single trial, one and only one of these outcomes occurs. The outcomes that make up a sample space are called sample points or simply points.

Example 2.1.1

The sample space for a roll of a six-sided die is

$$\{a_1, a_2, a_3, a_4, a_5, a_6\} \qquad \text{where } a_i \text{ is the event the top face is } i$$

More simply we could define the sample space as

$$\{1, 2, 3, 4, 5, 6\}$$

Note that a sample space of a probability model for a process is not necessarily unique. Often times, however, we try to chose sample points that are the smallest possible or "indivisible".

Example 2.1.2

If we define E to be the event that the top face of a six-sided die is even when rolled and O to be the event the top-face is odd, then the sample space, $\mathbb{S}$, can be defined as

$$\{E, O\}$$

This is the same process as Example 2.1.1 (rolling a six-sided die), so since the sample spaces differ, clearly, sample spaces are not unique. Moreover, if we are interested in the event that a 3 is rolled, this sample space is not suitable since it groups the event in question with other events.

A sample space can be either **discrete** or **non-discrete**. If a sample space is discrete, it consists of a finite or countably infinite number "simple events". A countably infinite set is one that can be put into a one-to-one correspondence with the set of real numbers. For example, $\left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right\}$ is countably infinite whereas $\{\, x \mid x \in \mathbb{R} \,\}$ is not.

Simple Events

An event in a discrete sample space is a subset of the sample space, i.e., $A \subset \mathbb{S}$. If the event is indivisible, so as to only contain one point, we call it a simple event, otherwise it is a compound event.

Example 2.1.3

A simple event for a roll of a six-sided die is $A = \{a_1\}$ where a_i is the event the top face is i. A compound event is $E = \{a_2, a_4, a_6\}$.

2.2 Assigning Probabilities

Let $\mathbb{S} = \{a_1, a_2, a_3, \ldots\}$ be a discrete sample space. We assign probabilities, $P(a_i)$, for $i = 1, 2, 3, \ldots$ to each sample point a_i such that the following two conditions hold

- $0 \leq P(a_i) \leq 1$
- $\sum\limits_{\text{all } i} P(a_i) = 1$

The set of probabilities $\{\, P(a_i) \mid i = 1, 2, 3, \ldots \}$ is called a **probability distribution** on $\mathbb{S}$.

Note that P is a function with the sample space as its domain.

The second condition, that the sum of the probabilities of all sample points is 1, relates to the property that for a given experiment one simple event in the sample space must occur. Every experiment or process always has an outcome thus the probability of any outcome being achieved must be 1.

Compound Events

The probability of an event A is the sum of the probability of all the simple events that make up A.

$$P(A) = \sum_{a \in A} P(a)$$

Example 2.2.1

In the previous example we saw that $E = \{a_2, a_4, a_6\}$ is a compound event. Thus, the probability of the compound event E is

$$P(E) = P(a_2) + P(a_4) + P(a_6)$$

Note that the probability model that we defined does not specify what actual numbers to assign to the simple events of a process. It only defines the properties that guarantee mathematical consistency. Thus, if we assigned $P(a_2)$ to be 0.9, our model would still be mathematically consistent but would not be consistent with the frequencies we obtain in multiple repetitions of the experiment.

In actual practice we try to define probabilities that are approximately consistent with the frequencies of the events in multiple repetition of the process.

Chapter 3

Counting Techniques

3.1 Counting Arguments

If we have a sample space, $\mathbb{S}$, of some experiment that has a **uniform distribution** (all sample points are equally likely), then we can calculate the probability of a compound event A as the number of sample points in A divided by the total number of sample points.

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

where k is the number of sample points in A and n is the total number of sample points in the sample space.

Addition Rule

Consider we can perform process 1 in p ways and process 2 in q ways. Suppose we want to do process 1 **or** process 2 **but not both**, then there are $p + q$ ways to do so.

Example 3.1.1

Suppose a keyboard only has 26 letters and 20 special characters (!%#$), there are 46 ways in which a typist may type a **single** character. (Process 1: typing a letter. Process 2: typing a special character).

Multiplication Rule

Again, consider we can perform process 1 in p ways and process 2 in q ways. Suppose we want to do process 1 **and** process 2, then there are $p \times q$ ways to do so. This is because **for each way** of doing process 1 we can do process 2 in q ways.

Example 3.1.2

Suppose the same typist with the same keyboard wants to type a single letter **and** a single special character. The typist can do so in 520 ways, since there are 26 ways to select the letter and **for each** possible letter selection there are 20 possible special character selections.

Try to associate **OR** and **AND** with **addition** and **multiplication** respectively in your mind.

Often times, **OR**'s and **AND**'s are not explicit or obvious so you must re-word your problem to identify implicit **OR**'s and **AND**'s.

Example 3.1.3

A young boy gets to pick 2 toys from a store for his birthday. How many ways can he pick 2 toys if the store contains 12 toys? He may pick the same toy multiple times and picks the toys at random.

We can re-word this problem as follows: A young boy selects one of 12 toys **and** again, selects one of 12 toys. Thus there are $12 \times 12 = 144$ ways in which he can select 2 toys. Furthermore, we have that since selections are random, each selection is equally likely. So the probability that the boy selects any pair of toys is $1/144$.

In this case the boy was allowed to select the same toy more than once. This is often referred to as **with replacement**. The addition and multiplication rules are generally sufficient to find probability of processes with replacement but if processes occur without replacement solutions become more complex and other techniques are often used.

The phrase **at random** or **uniformly**, indicates that each point in the sample space is equally likely.

Example 3.1.4

Consider a farmer with 500 different seeds. How many ways can he select 3 seeds randomly to plant?

We can re-word this problem to become: A farmer selects one seed from 500 **and** then selects one seed of 499 **and** then one seed of 498. So there are $500 \times 499 \times 498$ ways to do so.

Now, how many ways can he select 5 and 50 seeds randomly?

He can select 5 seeds in $500 \times 499 \times 498 \times 497 \times 496$ ways and 50 seeds in $500 \times \cdots \times 451$ ways.

Generally, if there are n ways of doing a process and it is done k times **without replacement**, that is you can only do the process a specific way once, there are $n \times \cdots \times (n - k + 1)$ ways to do it.