MAS187/AEF258

University of Newcastle upon Tyne

2005-6

Contents

Chapter 5

Introduction to Probability

5.1 Introduction

Probability is the language we use to model uncertainty. We all intuitively understand that few things in life are certain. There is usually an element of uncertainty or randomness around outcomes of our choices. In business this uncertainty can make all the difference between a good investment and a poor one. Hence an understanding of probability and how we might incorporate this into our decision making processes is important. In this chapter, we look at the logical basis for how we might express a probability and some basic rules that probabilities should follow. In the next chapter, we look at how we can use probabilities to aid decision making.

5.1.1 Definitions

We often use the letter P to represent a probability. For example, $P(Rain)$ would be the probability of the event of it raining.

Experiment An experiment is an activity where we do not know for certain what will happen but we will observe what happens. For example:

- We will ask someone whether or not they have used our product.
- We will observe the temperature at mid day tomorrow.
- We will toss a coin and observe whether it shows "heads" or "tails".

Outcome An outcome, or *elementary event*, is one of the possible things that can happen. For example, suppose that we are interested in the (UK) shoe size of the next customer to come into a shoe shop. Possible outcomes include "eight", "twelve", "nine and a half" and so on. In any experiment, one and only one outcome occurs.

Sample space The sample space is the set of all possible outcomes. For example it could be the set of all shoe sizes.

Event An even is a set of outcomes. For example "the shoe size of the next customer is less than 9" is an event. It is made of of all of the outcomes where the shoe size is less than 9. Of course an event might contain just one outcome.

Probabilities are usually expressed in terms of fractions or decimal numbers or percentages. Therefore we could express the probability of it raining today as

$$P(Rain) = \frac{1}{20} = 0.05 = 5\%.$$

All probabilities are measured on a scale ranging from zero to one. The probabilities of most events lie strictly between zero and one as an event with probability zero is an impossible event and one with probability one is a certain event.

The collection of all possible outcomes, that is the sample sapce, has a probability of 1. For example, if an event consists of only two outcomes *success* or *failure* then the probability of either a *success* or a *failure* is 1. That is $P(success \text{ or } failure) = 1$.

Two events are said to be *mutually exclusive* if both can not occur simultaneously. In the example above, the outcomes *success* and a *failure* are mutually exclusive.

Two events are said to be *independent* if the occurence of one does not affect the probability of the second occurring. For example, if you toss a coin and look out of the window, it would be reasonable to suppose that the events "get heads" and "it is raining" would be independent. However, not all events are independent. For example, if you go into the Students' Union Building and pick a student at random, then the events "the student is female" and "the student is studying engineering" are not independent since there is a greater proportion of male students on engineering courses than on other courses at the University (and this probably applies to those students found in the Union).

5.2 How do we measure Probability?

There are three main ways in which we can measure probability. All three obey the basic rules described above. Different people argue in favour of the different views of probability and some will argue that each kind has its uses depending on the circumstances.

5.2.1 Classical

If all possible outcomes are "equally likely" then we can adopt the *classical* approach to measuring probability. For example if we tossed a fair coin, there are only two possible outcomes, a head or a tail both of which are equally likely and hence

$$P(Head) = \frac{1}{2} \quad \text{and} \quad P(Tail) = \frac{1}{2}.$$

The underlying idea behind this view of probability is *symmetry*. In this example, there is no reason to think that the outcome *Head* and the outcome *Tail* have different probabilities and so they should have the same probability. Since there are two outcomes and one of them must occur, both outcomes must have probability 1/2.

Another commonly used example is rolling dice. There are six possible outcomes (1,2,3,4,5,6) when a die is rolled and each of them should have an equal chance of occuring. Hence the $P(1) = \frac{1}{6}$, $P(2) = \frac{1}{6}, \ldots$.

Other calculations can be made such as $P(\text{Even Number}) = \frac{3}{6} = \frac{1}{2}$. This follows from the formula

$$P(\text{Event}) = \frac{\text{Total number of outcomes in which event occurs}}{\text{Total number of possible outcomes}}.$$

Note that this formula only works when all possible events are equally likely – not a practical assumption for most real life situations.

5.2.2 Frequentist

When the outcomes of an experiment are not equally likely, we can conduct experiments to give us some idea of how likely the different outcomes are. For example, suppose we were interested in measuring the probability of producing a defective item in a manufacturing process. This probability could be measured by monitoring the process over a reasonably long period of time and calculating the proportion of defective items. What constitutes a reasonably long period of time is, of course, a difficult question to answer. In a more simple case, if we did not believe that a coin was fair, we could toss the coin a large number of times and see how often we obtained a head. In both cases we perform the same experiment a large number of times and observe the outcome. This is the basis of the frequentist view. By conducting experiments the probability of an event can easily be estimated using the following formula:

$$P(\text{Event}) = \frac{\text{Number of times an event occurs}}{\text{Total number of times experiment done}}.$$

The larger the experiment, the closer this probability is to the "true" probability. The frequentist view of probability regards probability as the long run relative frequency (or proportion). So, in the defects example, the "true" probability of getting a defective item is the proportion obtained in a very large experiment (strictly an *infinitely* long sequence of trials).

In the frequentist view, probability is a property of nature and, since, in practice, we can not conduct infinite sequences of trials, in many cases we never really know the "true" values of probabilities. We also have to be able to imagine a long sequence of "identical" trials. This does not seem to be appropriate for "one-off" experiments like the launch of a new product. For these reasons (and others) some people prefer the *subjective* or *Bayesian* view of probability.

5.2.3 Subjective/Bayesian

We are probably all intuitively familiar with this method of assigning probabilities. When we board an aeroplane, we judge the probability of it crashing to be sufficiently small that we are happy to undertake the journey. Similarly, the odds given by bookmakers on a horse race reflect people's beliefs about which horse will win. This probability does not fit within the frequentist definition as the race cannot be run a large number of times.

One potential difficulty with using subjective probabilities is that it *is* subjective. So the probabilities which two people assign to the same event can be different. This becomes important if these probabilities are to be used in decision making. For example, if you were deciding whether to launch a new product and two people had very different ideas about how likely success or failure of this product was, then the decision to go ahead could be controversial. If both individuals assessed the probability of success to be 0.8 then the decision to go ahead could easily be based on this belief. However, if one said 0.8 and the other 0.3, then the decision is not straightforward. We would need a way to reconcile these different positions.

Subjective probability is still subject to the same rules as the other forms of probability, namely that all probabilities should be positive and that the probability of all outcomes should sum to one. Therefore, if you assess $P(Success) = 0.8$ then you should also assess $P(Failure) = 0.2$.

5.3 Laws of Probability

5.3.1 Multiplication Law

The probability of two *independent* events E_1 and E_2 both occurring can be written as

$$P(E_1 \text{ and } E_2) = P(E_1) \times P(E_2).$$

For example, if the probability of throwing a six followed by another six on two rolls of a die is calculated as follows. The outcomes of the two rolls of the die are independent. Let E_1 denote a six on the first roll and E_2 a six on the second roll. Then

$$P(\text{two sixes}) = P(E_1 \text{ and } E_2) = P(E_1) \times P(E_2) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}.$$

This method of calculating probabilities extends to when there are many *independent* events

$$P(E_1 \text{ and } E_2 \text{ and } \cdots \text{ and } E_n) = P(E_1) \times P(E_2) \times \cdots \times P(E_n).$$

(There is a more complicated rule for multiplying probabilities when the events are not independent).

5.3.2 Addition Law

The multiplication law is concerned with the probability of two or more independent events occurring. The *addition law* describes the probability of any of two or more events occurring. The addition law for two events E_1 and E_2 is

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2).$$

This describes the probability of *either* event E_1 *or* event E_2 happening.

Consider the following information: 50 percent of families in a certain city subscribe to the morning newspaper, 65 percent subscribe to the afternoon newspaper, and 30 percent of the families subscribe to both newspapers. What proportion of families subscribe to at least one newspaper?

We are told $P(\text{Morning}) = 0.5$, $P(\text{Afternoon}) = 0.65$ and $P(\text{Morning and Afternoon}) = 0.3$. Therefore

$$\begin{aligned} P(\text{at least one paper}) &= P(\text{Morning or Afternoon}) \\ &= P(\text{Morning}) + P(\text{Afternoon}) - P(\text{Morning and Afternoon}) \\ &= 0.5 + 0.65 - 0.3 \\ &= 0.85. \end{aligned}$$

So 85% of of the city subscribe to at least one of the newspapers.

A more basic version of the rule works where events are mutually exclusive: if events E_1 and E_2 are mutually exclusive then

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2).$$

This simplification occurs because when two events are mutually exclusive they cannot happen together and so $P(E_1 \text{ and } E_2) = 0$.

These two laws are the basis of more complicated problem solving we will see later.

5.3.3 Example

A building has three rooms. Each room has two separate electric lights. There are thus six electric lights altogether. After a certain time there is a probability of 0.1 that a given light will have failed and all light are independent of all other lights. Find the probability that, after this time, there is at least one room in which both lights have failed.

Solution

For a given light, the probability that it has failed is 0.1.

For a given room, the probability that *both* lights have failed is

$$0.1 \times 0.1 = 0.01.$$

For a given room, the probability that it is not true that both lights have failed, that is the probability that at least one of the two lights is working, is

$$1 - 0.01 = 0.99.$$

The probability that at least one light is working in every one of the three rooms (that is, in Room A *and* in Room B *and* in Room C) is

$$0.99 \times 0.99 \times 0.99 = 0.99^3 = 0.970299.$$

The probability that there is at least one room in which both lights have failed (that is the probability that it is not true that there is at least one light working in every room) is

$$1 - 0.970299 = 0.029701$$

or just under 3%.

N.B. We also can obtain this answer by extending the addition law to cover three events. Let $A,\ B,\ C$ be the events "both lights have failed in Room A," " both lights have failed in Room B," "both lights have failed in Room C." We can show that

$$\begin{aligned} P(A \text{ or } B \text{ or } C) &= P(A) + P(B) + P(C) - P(A \text{ and } B) - P(A \text{ and } B) - P(B \text{ and } C) \\ &\quad + P(A \text{ and } B \text{ and } C) \end{aligned}$$

where "A or Bor C" means "at least one of $A,\ B,\ C$" and "A and $B$and C" means "all three of $A,\ B,\ C$". So, the required probability is

$$\begin{aligned} P(A \text{ or } B\text{or } C) &= 0.01 + 0.01 + 0.01 - (0.01 \times 0.01) - (0.01 \times 0.01) - (0.01 \times 0.01) \\ &\quad + (0.01 \times 0.01 \times 0.01) \\ &= 3 \times 0.01 - 3 \times 0.0001 + 0.000001 \\ &= 0.03 - 0.0003 + 0.000001 = 0.029701. \end{aligned}$$

5.4 Exercises 5

1. A company manufactures a device which contains three components A, B and C. The device fails if any of these components fail and the company offers to its customers a full money-back warranty if the product fails within one year. The company has assessed the probabilities of each of the components lasting at least a year as 0.98, 0.99 and 0.95 for A, B and C respectively. The three components within a single device are considered to be independent. Consider a single device chosen at random. Calculate the probability that
   (a) all three components will last for at least a year;
   (b) the device will be returned for a refund.
2. The following data refer to a class of 18 students. Suppose that we will choose one student at random from this class.

Student Number	Sex	Height (m)	Weight (kg)	Shoe Size	Student Number	Sex	Height (m)	Weight (kg)	Shoe Size
1	M	1.91	70	11.0	10	M	1.78	76	8.5
2	F	1.73	89	6.5	11	M	1.88	64	9.0
3	M	1.73	73	7.0	12	M	1.88	83	9.0
4	M	1.63	54	8.0	13	M	1.70	55	8.0
5	F	1.73	58	6.5	14	M	1.76	57	8.0
6	M	1.70	60	8.0	15	M	1.78	60	8.0
7	M	1.82	76	10.0	16	F	1.52	45	3.5
8	M	1.67	54	7.5	17	M	1.80	67	7.5
9	F	1.55	47	4.0	18	M	1.92	83	12.0

Find the probabilities for the following events.

(a) The student is female.

(b) The student's weight is greater than 70kg.,

(c) The student's weight is greater than 70kg. and the student's shoe-size is greater than 8,

(d) The student's weight is greater than 70kg. or the student's shoe-size is greater than 8.