Chapter 4

Probability

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.

4.1 Definitions

We often use the letter P to represent a probability. For example, P(Rain) would be the probability that it rains.

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 midday 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 event is a set of outcomes. For example "the shoe size of the next customer is less than 9" is an event. It is made up 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

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

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 an event with probability one is a *certain* event.

The collection of all possible outcomes, that is the *sample space*, 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 or failure) = 1.

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

Two events are said to be *independent* if the occurrence 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).

4.2 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.

4.2.1 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

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

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 occurring. Hence the $P(1) = \frac{1}{6}$, $P(2) = \frac{1}{6}$,

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.

4.2.2 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 cannot 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.

4.2.3 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.

Example 4.1

A fast–food chain with 700 outlets describes the geographic location of its restaurants with the following table:

		Region				
		NE	SE	SW	NW	
Population	Under 10,000	35	42	21	70	
	10,000 - 100,000	70	105	84	35	
	Over 100,000	175	28	35	0	

A health and safety organisation selects a restaurant at random for a hygiene inspection.

- (a) Which of the three approaches to probability would you use to find:
 - (i) P(NE restaurant chosen),
 - (ii) P(SW and city with a population < 100,000)?
- (b) Now use this approach to find the probabilities given above.



On the probability scale, how likely do you think it is that Newcastle United will finish in a top–six position this season? Which approach to probability would you use to estimate this?

4.3 Laws of Probability

4.3.1 Multiplication Law

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

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

and this is known as the *multiplication law* of probability. For example, 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, and we shall consider next term.

4.3.2 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.

(Example 4.2)

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(Morn.) = 0.5, P(Aft.) = 0.65 and P(Morn. and Aft.) = 0.3. Therefore

$$P(\text{at least one paper}) = P(\text{Morning or Afternoon})$$

= $P(\text{Morning}) + P(\text{Afternoon}) - P(\text{Morning and Afternoon})$
= $0.5 + 0.65 - 0.3$
= 0.85 .

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 $P(E_1 \text{ and } E_2) = 0$.

4.4 Tree Diagrams

Tree diagrams or probability trees are simple clear ways of presenting probabilistic information. Let us first consider a simple example in which a die is rolled twice. Suppose we are interested in the probability that we score a six on both rolls. This probability can be calculated as

$$\begin{array}{rcl} P(\mbox{ Six and Six}) &=& P(\mbox{ Six on 1st throw}) \times P(\mbox{ Six on 2nd throw}) \\ &=& \frac{1}{6} \times \frac{1}{6} \\ &=& \frac{1}{36}. \end{array}$$

This example can be represented as a tree diagram in which experiments are represented by circles (called *nodes*) and the outcomes of the experiments as *branches*:



Example 4.3

A machine is used to produce components. Each time it produces a component there is a chance that the component will be defective. When the machine is working correctly the probability that a component is defective is 0.05. Sometimes, though, the machine requires adjustment and, when this is the case, the probability that a component is defective is 0.2. At the time in question there is a probability of 0.1 that the machine requires adjustment. Components produced by the machine are tested and either accepted or rejected. A component which is not defective is accepted with probability 0.97 and (falsely) rejected with probability 0.03. A defective component is (falsely) accepted with probability 0.15 and rejected with probability 0.85.

Display this information on a tree diagram.



4.5 The binomial distribution

Suppose we are interested in the number of sixes we get from 3 rolls of a die. Each roll of the dice is an experiment or trial which gives a "six" (success, or s) or "not a six" (failure, or f). The probability of a success is p = P(six) = 1/6. We have n = 3 independent experiments or trials (rolls of the dice). Let X be the number of sixes obtained. We can now obtain the full *probability distribution* of X; a probability distribution is a list of all the possible outcomes for X with along with their associated probabilities.

For example, suppose we want to work out the probability of obtaining three sixes (three "successes" — i.e. sss — or P(X = 3)). Since the rolls of the die can be considered independent, we get:

$$P(sss) = P(s) \times P(s) \times P(s) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^3$$

That one's easy! What about the probability that we get two sixes — i.e. P(X = 2)? This one's a bit more tricky, because that means we need two s's and one f — i.e. two sixes and one "not six" — but the "not six" could appear on the first roll, or the second roll, or the third! Thinking about it, there are actually eight possible outcomes for the three rolls of the die.



So, for P(X = 2), we could have:

$$P(fss) = \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \frac{5}{6}$$

or we could have:

$$P(sfs) = \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} = \left(\frac{1}{6}\right)^2 \times \frac{5}{6},$$

or even:

$$P(ssf) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} = \left(\frac{1}{6}\right)^2 \times \frac{5}{6},$$

Can you see that we therefore get:

$$P(X=2) = 3 \times \left(\frac{1}{6}\right)^2 \times \frac{5}{6}.$$

Which takes the form:

P(X = 2) = Number of ways to get two sixes $\times P(2 \text{ sixes}) \times P(1 \text{ "not six"}).$

Using the same argument as above we can calculate the other probabilities:

$$P(X = 0) = \left(\frac{5}{6}\right)^3 = 0.579$$

$$P(X = 1) = 3 \times \left(\frac{1}{6}\right) \times \left(\frac{5}{6}\right)^2 = 0.347$$

$$P(X = 2) = 3 \times \left(\frac{1}{6}\right)^2 \times \frac{5}{6} = 0.069$$

$$P(X = 3) = \left(\frac{1}{6}\right)^3 = 0.005,$$

and so the full probability distribution for X is:

x	0	1	2	3	
P(X=x)	0.579	0.347	0.069	0.005	

This probability distribution shows that most of the time we would get either 0 or 1 sixes and, for example, 3 sixes would be quite rare. You might like to try rolling a die three times and observing how many sixes you obtain. Repeating this a large number of times will give you relative frequencies which should be close to these "theoretical" probabilities.

Now this is a bit long–winded ... and that was just for three rolls of the die! Imagine what it would be like to calculate for 100 rolls of the die!

We would like a more concise way of working these probabilities out without having to list all the possible outcomes as we did above. You should see from the tree diagram that we can construct a general formula, taking the form:

P(X = r) = # ways to get r successes out of n trials $\times P(r \text{ successes}) \times P(n - r \text{ failures})$

4.5.1 The binomial distribution: notation

In many surveys and experiments data is collected in the form of counts. For example, the number of people in the survey who bought a CD in the past month, the number of people who said they would vote Labour at the next election, the number of defective items in a sample taken from a production line, and so on. All these variables have common features:

- Each person/item has only two possible (exclusive) responses (Yes/No, Defective/Not defective etc)
 - this is referred to as a *trial* which results in a *success* or *failure*.
- 2. The survey/experiment takes the form of a random sample the responses are independent.
- 3. The probability of a success in each trial is p (in which case the probability of a failure is 1 p).
- 4. We are interested in the random variable X, the total number of successes out of n trials.

If these conditions are met, then X has a binomial distribution with index n and probability p. We write this as $X \sim Bin(n, p)$, which reads as "X has a binomial distribution with index n and probability p". Here, n and p are known as the parameters of the binomial distribution.

Returning to the dice-rolling experiment above: we were interested in the number of sixes obtained from three rolls of a six-sided die. Treating each roll of the die as a *trial*, with a six representing a *success* and "not a six" representing a *failure*, we can see that we have n = 3 independent trials, each with probability of success p = 1/6. Thus if X represents the number of sixes on the three rolls, we have that X has a binomial distribution with parameters n = 3 and p = 1/6, that is

$$X \sim Bin(3, 1/6).$$

4.5.2 Probability calculations

Recall our derived formula for working out probabilities in the die example:

P(X = r) = # ways to get r successes out of n trials $\times P(r \text{ successes}) \times P(n - r \text{ failures})$. We can write this more succinctly as

$$P(X = r) = {^{n}C_{r} \times p^{r} \times (1 - p)^{n-r}}$$

The binomial coefficient ${}^{n}C_{r}$ works out how many ways we can choose r objects out of n, and so is commonly read as "n choose r". There is a formula for this, given in terms of the factorial function:

$${}^{n}\mathbf{C}_{r} = \frac{n!}{r!(n-r)!}$$

where A! = "A factorial" = $A \times (A - 1) \times (A - 2) \times (A - 3) \times \cdots$; however, this is more easily obtained using the *nCr* button on the calculator!

The expression for P(X = 2) in the die rolling example can be calculated using this formula. In this example n = 3, r = 2, the probability of success (i.e. rolling a six) is $p = \frac{1}{6}$, therefore the probability of failure (i.e. not rolling a six) is $1 - p = \frac{5}{6}$. So,

$$P(X=2) = {}^{3}C_{2} \times \left(\frac{1}{6}\right)^{2} \times \left(1-\frac{1}{6}\right)^{3-2}$$
$$= 3 \times \left(\frac{1}{6}\right)^{2} \times \frac{5}{6}$$
$$= 0.069.$$

which matches up exactly with what we calculated before (as it should)! You should try this formula for r = 0, 1 and 3 to make sure you can recover the other values in the table on page 112.

4.5.3 Mean and variance

If we have the probability distribution for X rather than the raw observations, we denote the mean for X not by \bar{x} but by E[X] (which reads as "the expectation of X"), and the variance by Var(X).

If X is a random variable with Binomial Bin(n, p) distribution, then its mean and variance are

$$E[X] = n \times p$$
 $Var(X) = n \times p \times (1-p).$

For example, in the die-rolling experiment,

$$E[X] = 3 \times \frac{1}{6} = 0.5$$

and

$$Var(X) = 3 \times \frac{1}{6} \times \frac{5}{6} = 0.417$$
 and so
 $SD(X) = \sqrt{0.417} = 0.645.$

Example 4.4

A salesperson has a 40% chance of making a sale on a customer visit and she arranges 6 visits in a day. Suppose that the salesperson's visits result in sales *independently*.

- (a) What are the probabilities of her making 0, 1, 2, 3, 4, 5 and 6 sales?
- (b) Find the salesperson's expected number of sales. What is the standard deviation?



4.6 The Poisson distribution

The Poisson distribution is a very important discrete probability distribution which arises in many different contexts. Typically, Poisson random quantities are used in place of binomial random quantities in situations where n is large, p is small, and both np and n(1-p) exceed 5. In general, it is used to model data which are counts of (random) events in a certain area or time interval, without a known fixed upper limit but with a known *rate* of occurrence.

For example, consider the number of calls made in a 1 minute interval to an Internet service provider (ISP). The ISP has thousands of subscribers, but each one will call with a very small probability. If the ISP knows that on average 5 calls will be made in the interval, the actual number of calls will be a Poisson random variable, with mean 5.

4.6.1 Notation and probability calculations

If X is a random variable with a Poisson distribution with parameter λ (Greek lower case "lambda") then the probability it takes different values is

$$P(X = r) = \frac{\lambda^r e^{-\lambda}}{r!}, \quad r = 0, 1, 2, \dots$$

We write this as $X \sim Po(\lambda)$. The parameter λ has a very simple interpretation as the rate at which events occur.

4.6.2 Mean and variance

The distribution has mean and variance

$$E(X) = \lambda, \qquad Var(X) = \lambda.$$

Thus, when approximating binomial probabilities by Poisson probabilities, we match the means of the distributions: $\lambda = np$.

Returning to the ISP example, suppose we want to know the probabilities of different numbers of calls made to the ISP. Let X be the number of calls made in a minute. Then $X \sim Po(5)$ and, for example, the probability of receiving 4 calls is

$$P(X=4) = \frac{5^4 e^{-5}}{4!} = 0.1755.$$

We can use the formula for Poisson probabilities to calculate the probability of all possible outcomes:

	Probability	Cumulative Probability
r	P(X=r)	$P(X \le r)$
0	0.0067	0.0067
1	0.0337	0.0404
2	0.0843	0.1247
3	0.1403	0.2650
4	0.1755	0.4405
5	0.1755	0.6160
6	0.1462	0.7622
7	0.1044	0.8666
8	0.0653	0.9319
:	:	÷
sum	1.000000	

Therefore the probability of receiving between 2 and 8 calls is

$$P(2 \le X \le 8) = P(X \le 8) - P(X \le 1) = 0.9319 - 0.0404 = 0.8915$$

and so is very likely. Probability calculations such as this enable ISPs to calibrate the likely demand for their service and hence the resources they need to provide the service. Using such a model we can also account for "extreme" situations. For example, suppose that, for this ISP, we observed the following number of calls per minute over a five minute period: 6, 3, 5, 4, 6. Using simple frequentist reasoning, we would have

$$P(7 \text{ calls made}) = \frac{0}{5} = 0,$$

i.e. we will *never* observe seven calls in any one minute period! However, using the Poisson model, we have

$$P(X = 7) = 0.1044,$$

which is probably more realistic. You can see from the previous table of probabilities that, although when r is large the associated probabilities are small, at least they are accounted for and are, more realistically, non-zero.

Example 4.5

Pronto Pizzeria have recently opened a new restaurant in Newcastle. They have noticed that, during any one hour period on a Saturday night, they usually see about eight parties of customers arrive.

- (a) What probability distribution might be appropriate for X, the number of parties that arrive between 7pm and 8pm this Saturday evening?
- (b) What is the probability that, between 7pm and 8pm this Saturday, at least three parties of customers arrive at the restaurant?
- (c) The management are considering a Half Price Happy Hour Menu between 6pm and 7pm every Saturday night. Why might the model you proposed in part (a) no longer be valid?

4.7 Further examples

(Example 4.6)

Consider the following two scenarios:

Scenario 1

RyanJet currently leads all European airlines in on–time arrivals, with 82.5% of it's flights arriving on time or earlier. Of the next 6 scheduled *RyanJet* flights, X is the number that will land on time.

Scenario 2

Customer bookings on the RyanJet website occur at a rate of 20 per hour. Due to a technical problem, the RyanJet website will be down for maintenance between 17:30 and 18:00 this evening. We are interested in Y, the number of bookings we can expect RyanJet to lose, because of this website down-time.

1. Would you use the binomial distribution or the Poisson distribution to model the scenarios outlined above? In each case explain why, and state fully the most appropriate distribution, including the values of the associated parameters.



2. Find the expectation, and standard deviation, of the number of on-time flights of the next 6 scheduled *RyanJet* flights. Also find the expected number of lost *RyanJet* bookings between 17:30 and 18:00, with the associated standard deviation.



3. Find the probability that, of *RyanJet*'s next six flights, exactly 4 arrive on time.



4. Find the probability that *RyanJet* will lose

- (a) exactly 8 bookings;
- (b) more than 2 bookings,

during their website down–time.



5. It is decided to move the website down–time to 03:00–03:30. Why might the model you proposed in question 1 no longer be appropriate?



Example 4.7

When a customer places an order with *Staples Online Office Supplies*, a computerised accounting information system automatically checks to see if the customer has exceeded his or her credit limit. Let X be the number of customers who have exceeded their credit limit.

In a random sample of n orders, we can expect 1 customer to have exceeded their credit limit, with a variance of 0.95.

1. Would the binomial or the Poisson distributions be most appropriate to model X?



2. Find the parameters of the model you proposed in question 1, and hence find the probability that at least 1 customer, of the next n orders placed, exceeds their credit limit.



4.8 Chapter 4 practice questions

1. The following data refer to a class of 18 students. Suppose that we will choose one student at random from this class.

Student		Height	Weight	Shoe	Student		Height	Weight	Shoe
Number	\mathbf{Sex}	(m)	(kg)	Size	Number	\mathbf{Sex}	(m)	(kg)	Size
1	М	1.91	70	11.0	10	М	1.78	76	8.5
2	F	1.73	89	6.5	11	Μ	1.88	64	9.0
3	М	1.73	73	7.0	12	Μ	1.88	83	9.0
4	М	1.63	54	8.0	13	Μ	1.70	55	8.0
5	\mathbf{F}	1.73	58	6.5	14	Μ	1.76	57	8.0
6	М	1.70	60	8.0	15	Μ	1.78	60	8.0
7	М	1.82	76	10.0	16	\mathbf{F}	1.52	45	3.5
8	М	1.67	54	7.5	17	Μ	1.80	67	7.5
9	\mathbf{F}	1.55	47	4.0	18	Μ	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.
- 2. 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.
- 3. Do you think the following pairs of events are independent or dependent? Why?
 - (a) E: An individual has a high IQF: An individual is accepted for a University place
 - (b) A: A student plays table tennisB: A student is good at maths
 - (c) E_1 : An individual has a large outstanding credit card debt E_2 : An individual is allowed to extend his bank overdraft

4.8. CHAPTER 4 PRACTICE QUESTIONS

- 4. A company has installed a new computer system and some employees are having difficulty logging on to the system. They have been given training and the problems which arose during training were recorded and their probabilities calculated as follows:
 - An employee has a probability of 0.9 of logging on successfully on the first attempt.
 - If the employee logs in successfully then the employee will also be successful on each later attempt with probability 0.9.
 - If the employee tries to log in and is not successful then the employee loses confidence and the probability of a successful log-in on later occasions drops to 0.5.

Use a tree diagram to find the following probabilities:

- (a) An employee successfully logs on in each of the first three attempts.
- (b) An employee fails in the first attempt but is successful in the next two attempts.
- (c) An employee logs on successfully only once in three attempts.
- (d) An employee does not manage to log on successfully in three attempts.
- 5. Which of the following random variables could be modelled with a binomial distribution and which could be modelled with a Poisson distribution? In each case state the value(s) of the parameter(s) of the distribution.
 - (a) A salesperson has a 30% chance of making a sale on a customer visit. She arranges 10 visits in a day. Let X be the number of sales she makes in a day.
 - (b) Calls to the British Passport Office in Durham occur at a rate of 7 per hour on average. Let Y be the number of calls at the passport office in a 1 hour period.
 - (c) History suggests that 10% of eggs from a family–run farm are bad. Let Z be the number of bad eggs in a box of a dozen (i.e. 12) eggs.
- 6. An operator at a call centre has 20 calls to make in an hour. History suggests that they will be answered 60% of the time. Let X be the number of answered calls in an hour.
 - (a) What probability distribution does X have?
 - (b) What is the mean and standard deviation of X?
 - (c) Calculate the probability of getting a response exactly 9 times.
 - (d) Calculate the probability of getting less than 2 responses.

- 7. Customers join a queue at a bank at a rate of 10 per minute. Let X be the number of customers joining the queue in one minute.
 - (a) What probability distribution does X have?
 - (b) What is the mean and standard deviation of X?
 - (c) Calculate the probability that there are 12 arrivals to the queue in one minute.
 - (d) Calculate the probability there are no more than 2 arrivals to the queue in a minute.
- 8. Recently published figures show that 89% of orders taken at McDonald's Drive–Thru windows are filled correctly. Alicia Baker, Robyn Andrew and Josh Cook each visit a McDonald's Drive–Thru (independently of each other). Let X be the number of correctly–filled orders.
 - (a) Using standard notation, write down a suitable probability model for X.
 - (b) Find E[X] and SD(X).
 - (c) Find the probability that at least two of the orders are filled correctly.
- 9. Between 1896 when the Dow Jones Index was created and 2009, the index rose in 64% of the years (data taken from M. Hulbert, "What the Past Can't Tell Investors", New York Times, January 3rd, 2010). Based on this information, and assuming a binomial distribution, estimate the probability that the stock market will rise
 - (a) In exactly four of the next five years;
 - (b) In none of the next five years.
 - (c) What assumption of the binomial distribution might not be valid here?
- 10. On average, 28 bookings per week are made for holidays in *Thomas Cook*'s "Exotic and Faraway" brochure. Let X be the number of holidays booked from this brochure today.
 - (a) Using standard notation, write down a suitable probability model for X.
 - (b) What is the expected number, and standard deviation, of the number of holidays booked from this brochure today?
 - (c) Find the probability that exactly 3 holidays are booked from this brochure today.
 - (d) Find the probability that no holidays are booked from this brochure today.
 - (e) Use your answer to part (d) to find the probability that at least one holiday from this brochure is booked today.

4.8. CHAPTER 4 PRACTICE QUESTIONS

- (f) The management at *Thomas Cook* are interested in the number of days Y, over the period of a week, in which no holidays are booked from their "Exotic and Faraway" brochure. Write down a suitable proability model for Y, and hence find the probability that no holidays are sold on fewer than two days, i.e. P(Y < 2).
- 11. Employees at the *Sky Northeast* call centre earn commission for each customer they sign up to Sky's "triple bundle", which includes Sky TV, broadband and telephone.

For n outbound sales calls to be made by one employee this afternoon, the expected number of successful sign-ups is 2.2, with a variance of 1.76.

- (a) Assuming a binomial distribution, find the number of outbound calls that will be made (n) and the probability of successfully signing a customer up to the "triple bundle" (p).
- (b) Find the probability that this particular employee will get at least 9 successful sign-ups to the "triple bundle".
- 12. In 2012, the airline Cathay Pacific "mishandled" 3.89 bags per 1,000 passengers. What is the probability that, in the next 1,000 passengers, Cathay Pacific will have
 - (a) no mishandled bags;
 - (b) at least one mishandled bags;
 - (c) exactly two mishandled bags?
 - (d) What is the expected number and standard deviation of mishandled bags?
 - (e) Find the probability that the airline will have exactly one mishandled bag on a plane of 200 passengers.
 - (f) Find the probability that the airline will have exactly one mishandled bag on both it's flight from Hong Kong to Shanghai (100 passengers) and it's flight from London to Hong Kong (200 passengers), assuming bags are misplaced independently on these two flights.