Chapter 20

Discrete Random Variables

20.1 Introduction

A **random variable** is a variable that can take any one of the values associated with the set of outcomes of an experiment.

> For example, if a die is going to be rolled three times then we might define the random variable
>
> $$X = \text{the number of sixes obtained from the three throws.}$$
>
> Similarly, if a coin is going to be tossed four times then we might define the random variable
>
> $$Y = \text{the number of heads obtained from the four tosses.}$$

Recall from Chapter 16 that a variable can be **discrete** or **continuous**.

We will focus on **discrete** random variables in this current chapter (and also in Chapters 21 and 22). In contrast, we will focus on **continuous** random variables in Chapters 23 and 24.

The **probability distribution** of a **discrete** random variable contains

- all the values the random variable may take, together with
- the corresponding probabilities of the variable taking each value.

Example 1. A coin will be tossed three times. Suppose we define the random variable

$$X = \text{the number of heads obtained from the three throws.}$$

Write down the probability distribution of X.

Solution:

x	$\Pr(X = x)$
0	$\begin{aligned}\Pr(X=0) &= \Pr(TTT) \\ &= \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} \\ &= \tfrac{1}{8}\end{aligned}$
1	$\begin{aligned}\Pr(X=1) &= \Pr(HTT \ \text{ or } \ THT \ \text{ or } \ TTH) \\ &= \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} + \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} + \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} \\ &= \tfrac{3}{8}\end{aligned}$
2	$\begin{aligned}\Pr(X=2) &= \Pr(HHT \ \text{ or } \ HTH \ \text{ or } \ THH) \\ &= \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} + \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} + \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} \\ &= \tfrac{3}{8}\end{aligned}$
3	$\begin{aligned}\Pr(X=3) &= \Pr(HHH) \\ &= \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2} \\ &= \tfrac{1}{8}\end{aligned}$

$\square$

Notice that we have $\sum \Pr(X = x) = 1$.

- This is true for *every* discrete probability distribution.

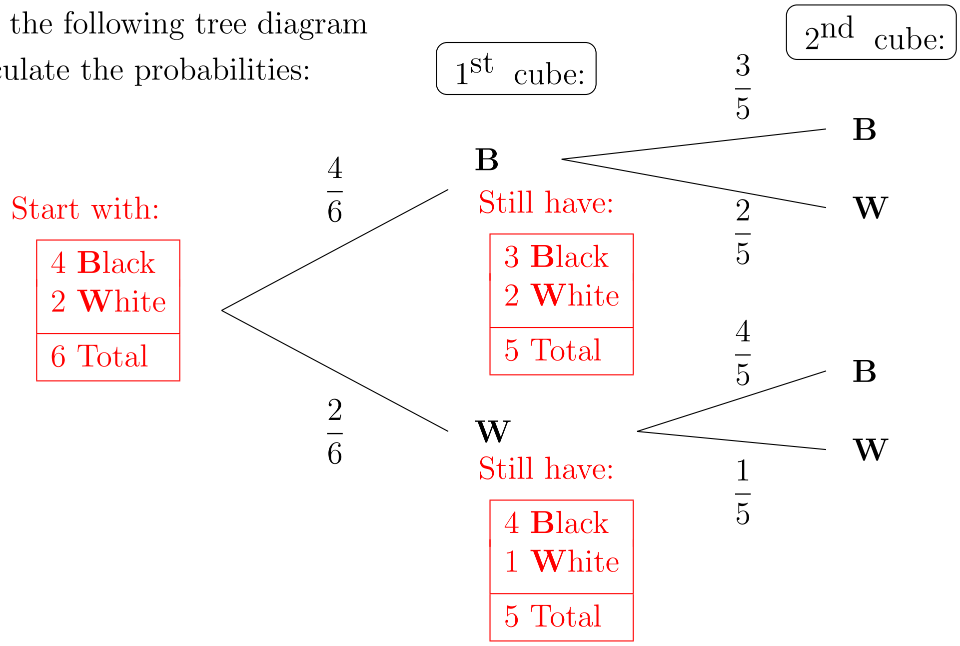
Example 2. A box contains 4 black cubes and 2 white cubes. Two cubes are going to be taken from the box at random (without replacement). Let

$$X \;=\text{ the number of white cubes selected in that sample.}$$

Find the probability distribution of X .

Solution:

We can use the following tree diagram to help calculate the probabilities:



So the probability distribution of X is:

x	$\Pr(X = x)$
0	$\Pr(X=0) = \Pr(\text{BB})$ $= \frac{4}{6} \times \frac{3}{5}$ $= \frac{2}{5}$
1	$\Pr(X=1) = \Pr(\text{BW} \quad \text{or} \quad \text{WB})$ $= \frac{4}{6} \times \frac{2}{5} + \frac{2}{6} \times \frac{4}{5}$ $= \frac{8}{15}$
2	$\Pr(X=2) = \Pr(\text{WW})$ $= \frac{2}{6} \times \frac{1}{5}$ $= \frac{1}{15}$

$\square$

20.2 Expected Value of a Random Variable

The **expected value** of a random variable X is the *long-run average* value of X. That is, it is the average value that might be expected if the experiment was repeated many times. The expected value of X is also known as the **mean** of X, and is denoted by $E(X)$ or by μ . If X is *discrete,* then the expected value (or mean) of X is defined as follows:

$$\begin{aligned} E(X) &= \mu \\ &= \sum x \Pr(X = x). \end{aligned}$$

(This is included on the Formula Sheet provided in the Maths 1 exams.)

Example 3. A coin is going to be tossed three times. Let X be the number of heads obtained from the three throws. Find the expected value of X .

Solution:

In Example 1 we obtained the following probability distribution:

x	$\Pr(X = x)$
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

$$\begin{aligned} \text{Then} \quad E(X) &= \sum x \Pr(X = x) \\ &= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} \\ &= 0 + \frac{3}{8} + \frac{6}{8} + \frac{3}{8} \\ &= \frac{12}{8} \\ &= 1.5 \end{aligned}$$

Thus, the expected value of X is 1.5 .

□

In Example 3 (on the previous page) we calculated that the expected number of heads obtained in three tosses is 1.5. This answer should seem 'sensible':

> On average we would expect half of a coin's tosses to show heads.
> Therefore, if the experiment of tossing the coin three times was repeated many times, we would expect to get an average of 1.5 heads per three tosses.

Let's explore this idea further by now supposing that the coin is tossed *one hundred* times.

Perhaps it seems immediately 'obvious' that on average we would expect to get 50 heads from one hundred tosses. However, notice that, in this case, it would be *incredibly slow and tedious* to calculate this 'obvious' answer by using the expected value formula given previously:

$$\begin{aligned} E(X) &= \sum x \Pr(X = x) \\ &= 0 \times \Pr(X = 0) \;+\; 1 \times \Pr(X = 1) \;+\; \ldots \;+\; 100 \times \Pr(X = 100) \\ &= \ldots \\ &= 50. \end{aligned}$$

Fortunately there are other versions of the expected value formula which allow us to calculate the expected value very quickly in certain circumstances. We will encounter these other versions of the expected value formula later in this chapter, and in the next chapter.

For example, in Chapter 21.1, we will learn a formula which will allow us to quickly calculate the expected number of heads from one hundred tosses as

$$\begin{aligned} E(X) &= 100 \times \frac{1}{2} \\ &= 50. \end{aligned}$$

But for now, let us consider some more examples in which we calculate the expected value by using the general formula stated previously:

$$\begin{aligned} E(X) &= \mu \\ &= \sum x \Pr(X = x). \end{aligned}$$

Example 4. In a gambling game, a player will toss a coin three times. He will win \$5 if three heads occur, \$3 if two heads occur, \$1 if one head occurs, but he will *lose* \$9 if no heads occur. What are his expected winnings (in dollars) per game?

Solution: Let X be the number of heads obtained per game, and
let Y be the winnings (in dollars) per game.

We have

x	y	$\Pr(X = x) = \Pr(Y = y)$
0	-9	$\frac{1}{8}$
1	1	$\frac{3}{8}$
2	3	$\frac{3}{8}$
3	5	$\frac{1}{8}$

(from Example 1).

Then
$$\begin{aligned} E(Y) &= \sum y \Pr(Y = y) \\ &= -9 \times \frac{1}{8} + 1 \times \frac{3}{8} + 3 \times \frac{3}{8} + 5 \times \frac{1}{8} \\ &= -\frac{9}{8} + \frac{3}{8} + \frac{9}{8} + \frac{5}{8} \\ &= \frac{8}{8} \\ &= 1. \end{aligned}$$

So the expected winnings per game are \$1.

□

We already know that $E(X) = \sum x \Pr(X = x)$.

More generally, we have $\boldsymbol{E(f(X)) = \sum f(x) \Pr(X = x)}$.

Thus, for example, we have $E(X^2) = \sum x^2 \Pr(X = x)$.

Example 5. Consider the random variable X with the following probability distribution:

x	$\Pr(X = x)$
-0.3	0.05
-0.1	0.1
0.1	0.35
0.4	0.5

Find $E(X)$ and $E(X^2)$.

Solution:

$$
\begin{aligned}
\text{We have } \quad E(X) &= \sum x \Pr(X = x) \\
&= -0.3 \times 0.05 \ + \ -0.1 \times 0.1 \ + \ 0.1 \times 0.35 \ + \ 0.4 \times 0.5 \\
&= -\ 0.015 \ - \ 0.01 \ + \ 0.035 \ + \ 0.2 \\
&= 0.21\,,
\end{aligned}
$$

$$
\begin{aligned}
\text{and} \quad E(X^2) &= \sum x^2 \Pr(X = x) \\
&= (-0.3)^2 \times 0.05 \ + \ (-0.1)^2 \times 0.1 \ + \ 0.1^2 \times 0.35 \ + 0.4^2 \times 0.5 \\
&= 0.09 \times 0.05 \ + \ 0.01 \times 0.1 \ + \ 0.01 \times 0.35 \ + \ 0.16 \times 0.5 \\
&= 0.0045 \ + \ 0.001 \ + \ 0.0035 \ + \ 0.08 \\
&= 0.089\,.
\end{aligned}
$$

$\square$

Notice that, for this example, we have $E(X^2) \neq (E(X))^2$.

However, it can be shown that for *any* random variables X and Y, and for *any* real numbers a and b , we will *always* have

$$E(aX + bY) \ = \ a\,E(X) \ + \ b\,E(Y).$$

Because of this property, the expected value is called a **linear** function.

Note that when $Y = b$ is a *constant* random variable (that is, a random variable with one outcome equal to the number b), this formula simplifes to

$$E(aX + b) \ = \ a\,E(X) \ + \ b.$$

20.3 Variance and Standard Deviation (of a Random Variable)

In Chapter 16.3, we learnt about the variance and standard deviation *of a sample.* Now we are going to learn about the variance and standard deviation *of a random variable.*

The **variance** of a random variable X is defined by σ^2 or $\mathrm{Var}(X)$, and is defined to be

$$\boldsymbol{E\left((X-\mu)^2\right).}$$

If X is *discrete,* then the variance of X is defined as follows:

$$\begin{aligned}\sigma^2 = \mathrm{Var}(X) &= E\left((X-\mu)^2\right)\\ &= \sum (x-\mu)^2 \, \mathrm{Pr}(X=x)\,.\end{aligned}$$

By expanding the brackets and rearranging the terms in this formula, we can obtain the following *alternative* (but completely equivalent) definition:

$$\begin{aligned}\sigma^2 = \mathrm{Var}(X) &= E(X^2) - \mu^2\\ &= \left(\sum x^2 \, \mathrm{Pr}(X=x)\right) - \mu^2.\end{aligned}$$

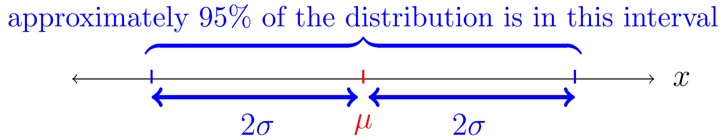
(Both of these two alternative (and equivalent) definitions are included on the Formula Sheet provided in the Maths 1 exams.)

- Variance is *not* a linear function. Instead, if a and b are real numbers, then we always have

$$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)\,.$$

The **standard deviation** of X is just σ, where $\sigma = \sqrt{\mathrm{Var}(X)}$.

The standard deviation of a random variable measures the *spread* of its distribution. In fact, for many random variables, approximately 95% of the distribution lies within two standard deviations of the mean, as shown here:



We can re-state this idea in symbols as shown next.

If X is a random variable with mean μ and standard deviation σ, then we often have

$$\boxed{\Pr\Big(\mu - 2\sigma \le X \le \mu + 2\sigma\Big) \approx 0.95}$$

This tells us that just two numbers, μ and σ, summarise most of the values of X. Indeed, most researchers use precisely these two numbers to summarise many different types of data.

Example 6. Find the mean and the variance of the following probability distribution:

x	$\Pr(X = x)$
0	0.5
1	0.3
2	0.2

Solution:

$$\begin{aligned} \text{We have} \quad E(X) &= \sum x \Pr(X = x) \\ &= 0 \times 0.5 + 1 \times 0.3 + 2 \times 0.2 \\ &= 0.7\,. \end{aligned}$$

We can find the variance using *either* of the formulae given on the Formula Sheet. That is, we can find the variance by calculating

$$\begin{aligned} \operatorname{Var}(X) &= E\left((X-\mu)^2\right) \\ &= \sum (x-\mu)^2 \Pr(X = x) \\ &= (0-0.7)^2 \times 0.5 + (1-0.7)^2 \times 0.3 + (2-0.7)^2 \times 0.2 \\ &= 0.245 + 0.027 + 0.338 \\ &= 0.61 \end{aligned}$$

or, alternatively, by calculating

$$\begin{aligned} \operatorname{Var}(X) &= E(X^2) - \mu^2 \\ &= \Big(\sum x^2 \Pr(X = x)\Big) - \mu^2 \\ &= 0^2 \times 0.5 + 1^2 \times 0.3 + 2^2 \times 0.2 - 0.7^2 \\ &= 0 + 0.3 + 0.8 - 0.49 \\ &= 0.61\,. \end{aligned}$$

$\square$

Example 7. The probability distribution of the number of misprints, M, per page in a textbook is

m	$\Pr(M = m)$
0	0.35
1	0.45
2	0.15
3	0.05

Find the mean and the standard deviation of the number of misprints per page.

Solution: The mean is given by

$$\begin{aligned}\mu &= \sum m \Pr(M = m) \\ &= 0 \times 0.35 \ + \ 1 \times 0.45 \ + \ 2 \times 0.15 \ + 3 \times 0.05 \\ &= 0.9\,.\end{aligned}$$

Thus the mean number of misprints per page is 0.9.

Before we find the *standard deviation*, we will calculate the *variance*.
One way of finding the *variance* is to calculate

$$\begin{aligned}\sigma^2 &= E(M^2) \ - \ \mu^2 \\ &= \sum m^2 \Pr(M = m) \ - \ 0.9^2 \\ &= 0^2 \times 0.35 \ + \ 1^2 \times 0.45 \ + \ 2^2 \times 0.15 \ + 3^2 \times 0.05 \ - \ 0.81 \\ &= 0.69\,.\end{aligned}$$

Then the *standard deviation* is given by

$$\begin{aligned}\sigma &= \sqrt{variance} \\ &= \sqrt{0.69} \\ &= \frac{\sqrt{69}}{10}\,.\end{aligned}$$

That is, the standard deviation of the number of misprints per page is $\dfrac{\sqrt{69}}{10}$. $\square$

Exercises for Sections 20.1 – 20.3

1. Consider the probability distribution given below:

x	$\Pr(X = x)$
1	0.2
2	0.1
3	0.5
4	0.2

   (a) Find $E(X)$.
   (b) Find $E(3X)$.
   (c) Find $E(X^2)$.
   (d) Find $\text{Var}(X)$.

2. Consider the probability distribution given below:

x	$\Pr(X = x)$
0	0.5
1	0.3
2	0.2

   (a) Find $E(X)$.
   (b) Find $E(X^2)$.
   (c) Find $E(\sqrt{X+1}\,)$.
   Give this answer to 3 decimal places.
   (d) Find $E(2X)$.

3. A coin will be tossed three times. Let X be the number of tails obtained.
   (a) Write down the probability distribution of X .
   (b) Find $E(X)$.
   (c) Find $\text{Var}(X)$.
   (d) Find $E(X^2 - 2)$.
   (e) Find $E(X^3)$.

4. A die is loaded so that it has the following probability distribution:

x	$\Pr(X = x)$
1	k
2	$2k$
3	$3k$
4	$4k$
5	$5k$
6	$6k$

   (a) Find the value of k .
   (b) If an even number is thrown, what is the probability that it is a 4?
   (c) If a number greater than 4 is thrown, what is the probability that it is a 6?

5. A random variable X is defined only for the values $X = 1,\ 2,\ 3,\ 4$. The probabilities are given by $\Pr(X = x) = kx$, where k is a constant.

   (a) Find the value of k .
   (b) Hence find the mean and variance of X .

20.4 Hypergeometric Distribution

In these examples, we will be choosing samples *without replacement.* There will be two possible outcomes for each item chosen. Those two outcomes are often referred to by the two words "success" and "failure", and we will be interested in the number of "successes" chosen in the sample. Because we are sampling *without replacement*,

the probability of a "success" can alter

(and similarly the probability of a "failure" can alter).

- *Please note that the two words "success" and "failure" are* <u>*not*</u> *intended to imply any judgement of whether those particular outcomes are "good" or "bad".*

Example 8. A group of people contains 3 males and 6 females. A random sample of 5 people will be selected, without replacement. Find the probability that exactly 1 male will be chosen.

Solution:

$$
\begin{aligned}
&\Pr(\text{exactly 1 male will be chosen}) \\
&= \frac{\text{number of ways of choosing 1 male and 4 females}}{\text{total number of ways of choosing 5 people}} \\
&= \frac{{}^{3}\mathrm{C}_{1} \times {}^{6}\mathrm{C}_{4}}{{}^{9}\mathrm{C}_{5}} \\
&= \frac{3 \times 15}{126} \\
&= \frac{5}{14}\,.
\end{aligned}
$$

So the probability that exactly one male will be chosen is $\dfrac{5}{14}$.

□

Let us define

$$X = \text{the number of males selected in the sample of 5 people.}$$

This variable X is an example of a **hypergeometric** variable.

A **hypergeometric distribution** has the following features:

- There are n *dependent* trials.
- Each trial has two possible outcomes (called "success" and "failure").
- The probability of success in each trial *alters*.

In the previous example, we found $\Pr(X = 1)$ where

$$X = \text{the number of males selected in the sample of 5 people.}$$

Similarly, we could find $\Pr(X = 0)$ and $\Pr(X = 2)$ and $\Pr(X = 3)$, which results in the following probability distribution for this random variable X :

x	$\Pr(X = x)$
0	$\dfrac{{}^3C_0 \times {}^6C_5}{{}^9C_5} = \dfrac{1 \times 6}{126} = \dfrac{1}{21}$
1	$\dfrac{{}^3C_1 \times {}^6C_4}{{}^9C_5} = \dfrac{3 \times 15}{126} = \dfrac{5}{14}$
2	$\dfrac{{}^3C_2 \times {}^6C_3}{{}^9C_5} = \dfrac{3 \times 20}{126} = \dfrac{10}{21}$
3	$\dfrac{{}^3C_3 \times {}^6C_2}{{}^9C_5} = \dfrac{1 \times 15}{126} = \dfrac{5}{42}$

Suppose we have a population of size N, of which

- D items have a particular property, and so
- $N - D$ items do not have that property.

$$\text{Overall group:} \quad \left\{ \begin{array}{c} D \text{ items have property} \qquad \text{and} \qquad N - D \text{ items do not have property.} \\ N \text{ items in total} \end{array} \right.$$

Let X be the number of items with the particular property obtained in a sample of n items.

$$\text{Sample:} \quad \left\{ \begin{array}{c} x \text{ items have property} \qquad \text{and} \qquad n - x \text{ items do not have property.} \\ n \text{ items in sample} \end{array} \right.$$

Then
$$\boxed{\Pr(X = x) \;=\; \frac{{}^{D}\mathrm{C}_{x} \times {}^{N-D}\mathrm{C}_{n-x}}{{}^{N}\mathrm{C}_{n}} \,.}$$

This formula is given on the Formula Sheet provided in the Maths 1 exams.

Notice that this hypergeometric random variable X is *discrete*, with values such as $0, 1, 2, \ldots$. Notice also that we must have $X \leq n$ and also $X \leq D$.

For the probability distribution given on the previous page, we *could* use the general formula for expected value from Section 20.2 to calculate

$$\begin{aligned} E(X) &= \sum x \Pr(X = x) \\ &= 0 \times \frac{1}{21} \;+\; 1 \times \frac{5}{14} \;+\; 2 \times \frac{10}{21} \;+\; 3 \times \frac{5}{42} \\ &= \frac{5}{3}\,. \end{aligned}$$

However, whenever we have a hypergeometric distribution, we do not need to use the general formula $E(X) = \sum x \Pr(X = x)$ from Section 20.2 to find the expected value. Instead we can use the following much easier formula.

If X has the *hypergeometric* distribution, then $\boxed{E(X) \;=\; n \, \dfrac{D}{N}\,.}$

This formula is also on the Formula Sheet provided in the Maths 1 exams.

Furthermore, if X has the *hypergeometric* distribution, then

$$\boxed{\mathrm{Var}(X) \;=\; n\,\frac{D}{N}\left(1-\frac{D}{N}\right)\,\frac{N-n}{N-1}\;.}$$

This formula is also on the Formula Sheet provided in the Maths 1 exams.

Example 9. Find the expected value and the variance for the probability distribution given on page 13.

Solution:

Let X be the number of males chosen in the sample of 5 people.

We have $N = 9,\ D = 3$ and $n = 5$.

$$\begin{aligned} \text{Therefore we have}\quad E(X) &= n\,\frac{D}{N} \\ &= 5 \times \frac{3}{9} \\ &= \frac{5}{3}\,. \end{aligned}$$

That is, the expected value is $\dfrac{5}{3}$.

$$\begin{aligned} \text{Also we have}\quad \mathrm{Var}(X) &= n\,\frac{D}{N}\left(1-\frac{D}{N}\right)\frac{N-n}{N-1} \\ &= 5 \times \frac{3}{9}\left(1-\frac{3}{9}\right)\frac{4}{8} \\ &= \frac{5}{3} \times \frac{6}{9} \times \frac{1}{2} \\ &= \frac{5}{9}\,. \end{aligned}$$

That is, the variance is $\dfrac{5}{9}$.

$\square$

Example 10. Five cards will be chosen from a pack of 52. Find the probability of choosing exactly 2 spades. Give your answer to 4 decimal places.

(Recall that the contents of a pack of cards are summarized on the last page of Chapter 18.)

Solution:

$$\text{Overall pack: } \begin{cases} 13 \text{ spades} \qquad \text{and} \qquad 39 \text{ non-spades} \\ \qquad\qquad 52 \text{ cards in total} \end{cases}$$

$$\text{Sample: } \begin{cases} 2 \text{ spades} \qquad \text{and} \qquad 3 \text{ non-spades} \\ \qquad\qquad 5 \text{ chosen cards} \end{cases}$$

Let X be the number of spades obtained among the 5 chosen cards.

Then $\Pr(X = 2)$

$$= \frac{\text{number of ways of choosing 2 spades and 3 non–spades}}{\text{total number of ways of choosing 5 cards}}$$

$$= \frac{{}^{13}\mathrm{C}_2 \times {}^{39}\mathrm{C}_3}{{}^{52}\mathrm{C}_5}$$

$$= \frac{78 \times 9139}{2\,598\,960}$$

$$= 0.2743 \ \ (4 \text{ d.p.})$$

So the probability of choosing exactly 2 spades in the sample is 0.2743 (4 d.p.).

□

Example 11. In a class of 12 students, 7 of the students have understood a particular topic whereas the other 5 students have **not** understood that topic. The teacher does not know how many students have understood the topic, so she adopts the following plan:

> Three students will be selected at random, and each of those three students will be asked a question on the topic.
>
> - If at least 2 of these students answer correctly, the teacher will proceed to the next topic.
> - If, however, fewer than 2 of the students answer correctly then the teacher will select another 3 students (out of the 9 students who have **not** already been asked), and ask them a question.
>   - If each of the three students from the second sample answers correctly then the teacher will proceed to the next topic.
>   - However, if *fewer* than three of the students from the second sample answers correctly then the teacher will re–teach the old topic.

Find the probability that the teacher proceeds to the next topic.
Write the answer to 4 decimal places.

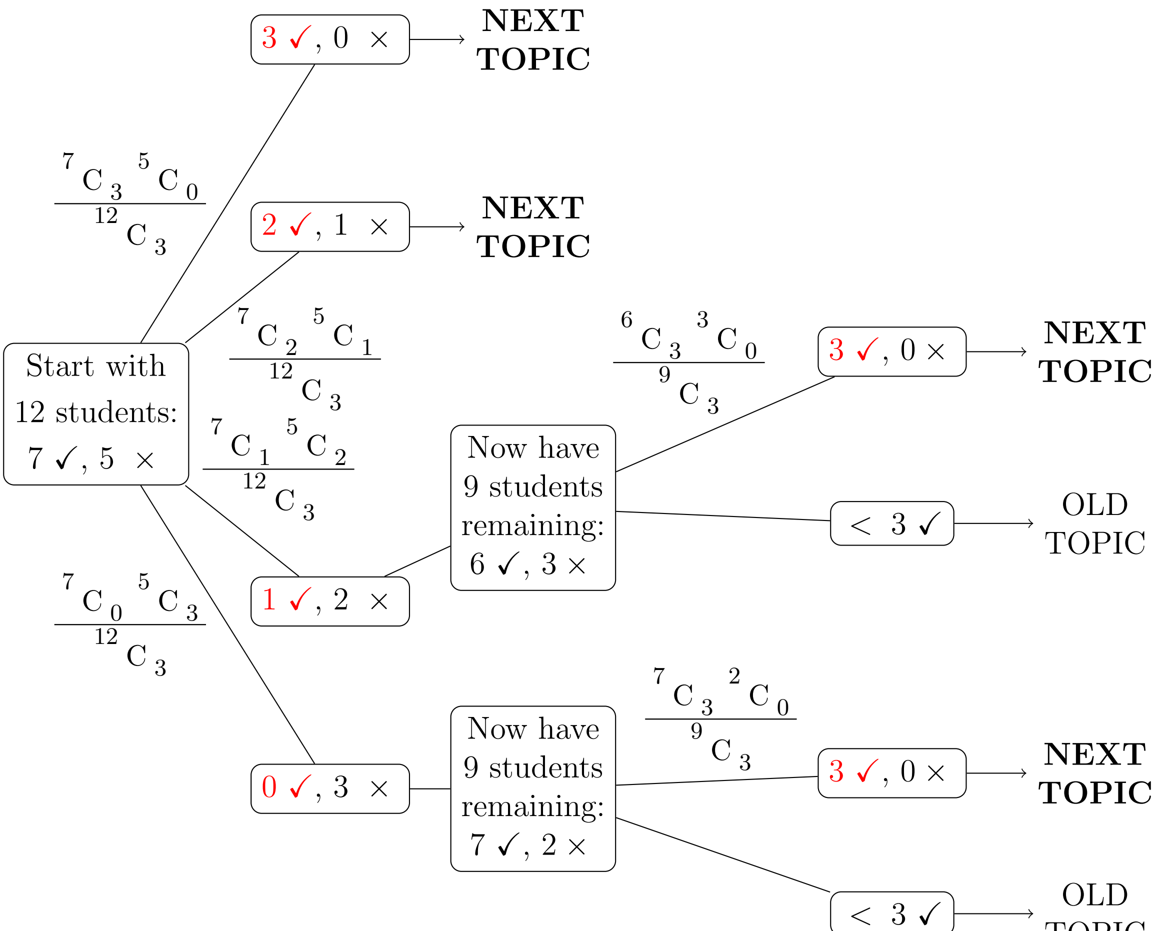
Solution:

A tree diagram (shown on the next page) is very helpful for summarizing the information for this example.

Note: The symbols used in that tree diagram represent the following meanings:

$\checkmark$ = Understands the topic (and so answers the question correctly).

$\times$ = Does not understand the topic (and so answers the question incorrectly).



$$
\begin{aligned}
&\Pr(\text{teacher proceeds to the next topic})\\
&= \Pr(\text{3 of the original sample answer correctly})\\
&\quad + \Pr(\text{2 of the original sample answer correctly})\\
&\quad + \Pr(\text{1 of the original sample answers correctly, } \textbf{and}\\
&\qquad\qquad \text{all 3 of the second sample answer correctly})\\
&\quad + \Pr(\text{0 of the original sample answers correctly, } \textbf{and}\\
&\qquad\qquad \text{all 3 of the second sample answer correctly})\\
&= \frac{{}^7C_3\,{}^5C_0}{{}^{12}C_3} + \frac{{}^7C_2\,{}^5C_1}{{}^{12}C_3} + \frac{{}^7C_1\,{}^5C_2}{{}^{12}C_3}\,\frac{{}^6C_3\,{}^3C_0}{{}^{9}C_3} + \frac{{}^7C_0\,{}^5C_3}{{}^{12}C_3}\,\frac{{}^7C_3\,{}^2C_0}{{}^{9}C_3}\\
&= \frac{35}{220} + \frac{105}{220} + \frac{70}{220}\times\frac{20}{84} + \frac{10}{220}\times\frac{35}{84}\\
&= \frac{193}{264} = 0.7311 \;\text{(4 d.p.)}
\end{aligned}
$$

Therefore the probability that the teacher proceeds to the next topic is 0.7311 (4 d.p.).

□

Exercises for Section 20.4

1. A mathematics class consists of 15 boys and 10 girls. A committee of 5 is chosen at random from this class. What is the probability that the committee has a majority of boys? Give your answer to 4 decimal places.

2. A box contains 8 items of which 2 are defective. A man selects 3 items from the box (without replacement).
   (a) State the type of probability distribution of the number of defective items he has drawn.
   (b) Find the expected number of defective items he has drawn.
   (c) Find the variance of the distribution.

3. A random sample of 3 items is selected, without replacement, from a batch of 10 items which contains 4 defectives. Find the probability that there is at most one defective item in the sample.

4. Packages, each containing 10 articles, are subjected to the following sampling plan. Three articles, selected at random from a package, are inspected, and the package is accepted if none of the three articles are defective. Consider packages which contain one defective article. Find the probability that such a package will be accepted.

5. A class consists of 12 boys and 8 girls. What is the probability that at least 2 of the top 3 students are girls? Give your answer to 4 decimal places.

6. The prizes in a competition are 100 books, 20 of which have been signed by the author. A particular school wins 10 prizes.
   (a) What is the expected number of signed books in the 10 prizes won by the school?
   (b) What is the probability that at most 2 of the 10 prizes won by the school have been signed? Give your answer to 4 decimal places.

20.5 Answers for the Chapter 20 Exercises

20.1 – 20.3

1. (a) 2.7 (b) 8.1 (c) 8.3 (d) 1.01
2. (a) 0.7 (b) 1.1 (c) 1.271 (3 d.p.) (d) 1.4
3. (a)

x	$\Pr(X = x)$
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

   (b) $E(X) = 1.5$

   (c) $\mathrm{Var}(X) = 0.75$

   (d) $E(X^2 - 2) = 1$

   (e) $E(X^3) = 6.75$

4. (a) We must have $k = \dfrac{1}{21}$ so that $\sum \Pr(X = x) = 1$.

   (b) The required probability is $\dfrac{1}{3}$.

   (c) The required probability is $\dfrac{6}{11}$.

5. (a) $k = \dfrac{1}{10}$ (b) $E(X) = 3$ and $\mathrm{Var}(X) = 1$.

20.4

1. The probability of having a majority of boys is 0.6988 (4 d.p.).
2. (a) The distribution is hypergeometric, with $N = 8,\ D = 2$ and $n = 3$.

   (b) The expected number is $E(X) \ = \ n\,\dfrac{D}{N} \ = \ \dfrac{3}{4}$.

   (c) The variance is $\mathrm{Var}(X) \ = \ \dfrac{45}{112}$.

3. The probability that there is at most one defective item in the sample is $\dfrac{2}{3}$.
4. The probability that such a package will be accepted is 0.7 .
5. The probability that at least 2 of the top 3 students are girls is 0.3439 (4 d.p.)
6. (a) The expected number of signed books in the 10 prizes is 2.

   (b) The probability that at most 2 of the school's prizes have been signed is 0.6812 (4 d.p.).