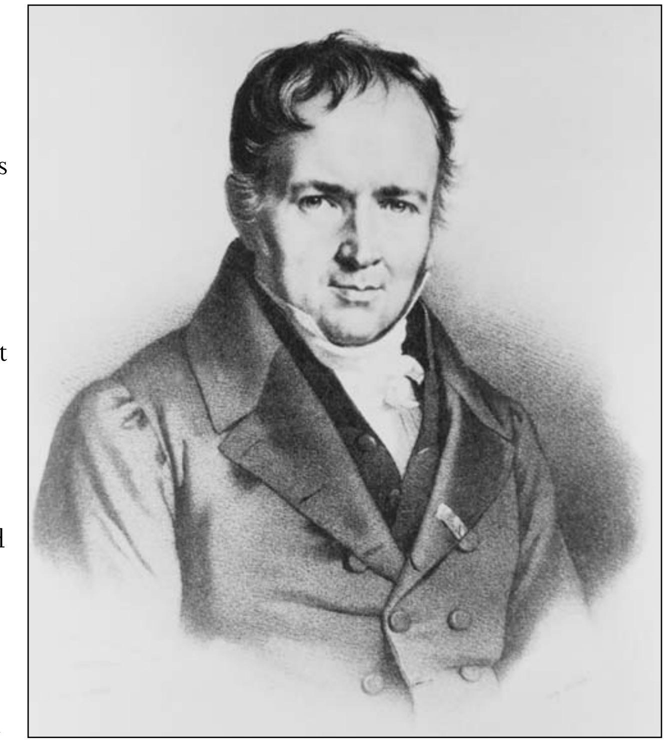
21 Discrete Probability Distributions



Siméon Denis Poisson

In this chapter we will meet the concept of a discrete probability distribution and one of these is called the Poisson distribution. This was named after Siméon-Denis Poisson who was born in Pithivier in France on 21 June 1781. His father was a great influence on him and it was he who decided that a secure future for his son would be the medical profession. However, Siméon-Denis was not suited to being a surgeon, due to his lack of interest and also his lack of coordination. In 1796 Poisson went to Fontainebleau to study at the École Centrale, where he showed a great academic talent, especially in mathematics. Following his success there, he was encouraged to sit the entrance examinations for the École Polytechnique in Paris, where he gained the highest mark despite having had much less formal education than most of the other entrants. He continued to excel at the École Polytechnique and his only weakness was the lack of coordination which made drawing mathematical diagrams virtually impossible. In his final year at the École Polytechnique he wrote a paper on the theory of equations that was of such a high quality that he was allowed to graduate without sitting the final examinations. On graduation in 1800, he became répétiteur at the École Polytechnique, which was rapidly followed by promotion to deputy professor in 1802, and professor in 1806. During this time Poisson worked on differential calculus and later that decade published papers with the Academy of Sciences, which included work on astronomy and confirmed the belief that the Earth was flattened at the poles.

Poisson was a tireless worker and was dedicated to both his research and his teaching. He played an ever increasingly important role in the organization of mathematics in France and even though he married in 1817, he still managed to take on further duties. He continued to research widely in a range of topics based on applied mathematics. In *Recherches sur la probabilité des jugements en matière criminelle et matière civile,* published in 1837, the idea of the Poisson distribution first appears. This describes the probability that a random event will occur when the event is evenly spaced, on average, over an infinite space. We will learn about this distribution in this chapter. Overall, Poisson published between 300 and 400 mathematical works and his name is attached to a wide variety of ideas, including Poisson's integral, Poisson brackets in differential equations, Poisson's ratio in elasticity, and Poisson's constant in electricity. Poisson died on 25 April 1840.

21.1 Introduction to discrete random variables

In Chapter 20 we met the idea of calculating probability given a specific situation and found probabilities using tree diagrams and sample spaces. Once we have obtained these values, we can write them in the form of a table and further work can be done with them. Also, we can sometimes find patterns that allow us to work more easily in terms of finding the initial distribution of probabilities. In this chapter we will work with discrete random variables. A **discrete random variable** has the following properties.

- It is a discrete (exact) variable.
- It can only assume certain values, $x_1, x_2, \ldots, x_n$.
- Each value has an associated probability, $P(X = x_1) = p_1$, $P(X = x_2) = p_2$ etc.
- The probabilities add up to 1, that is $\sum_{i=1}^{i=n} P(X = x_i) = 1$.
- A discrete variable is only random if the probabilities add up to 1.

A discrete random variable is normally denoted by an upper case letter, e.g. X, and the particular value it takes by a lower case letter, e.g. x.

Example

Write out the probability distribution for the number of threes obtained when two tetrahedral dice are thrown. Confirm that it is a discrete random variable.

$$P(\text{no threes}) = \frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$$

$$P(\text{1 three}) = P(3\bar{3}) + P(\bar{3}3) = \frac{1}{4} \times \frac{3}{4} + \frac{3}{4} \times \frac{1}{4} = \frac{6}{16}$$

$$P(\text{2 threes}) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

Hence the probability distribution is:

Number of threes	0	1	2
Probability	$\frac{9}{16}$	$\frac{6}{16}$	$\frac{1}{16}$

The distribution is discrete because we can only find whole-number values for the number of threes obtained. Finding a value for 2.5 threes does not make sense. It is also random because $\frac{9}{16} + \frac{6}{16} + \frac{1}{16} = \frac{16}{16} = 1$.

We can now use the notation introduced above and state that if X is the number of threes obtained when two tetrahedral dice are thrown, then X is a discrete random variable, where $P(X = 0) = \frac{9}{16}$, $P(X = 1) = \frac{6}{16}$, $P(X = 2) = \frac{1}{16}$. In a table:

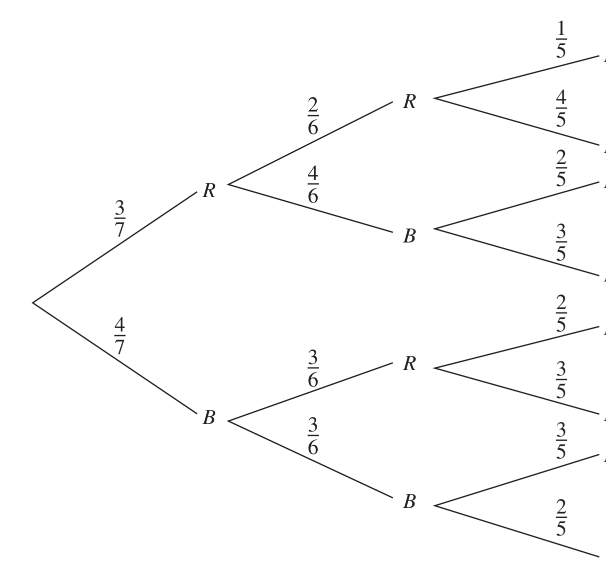
X	0	1	2
$P(X = x)$	$\frac{9}{16}$	$\frac{6}{16}$	$\frac{1}{16}$

These probability distributions can be found in three ways. Firstly they can be found from tree diagrams or probability space diagrams, secondly they can be found from a specific formula and thirdly they can be found because they follow a set pattern and hence form a special probability distribution.

Example

A bag contains 3 red balls and 4 black balls. Write down the probability distribution for R, where R is the number of red balls chosen when 3 balls are picked without replacement, and show that it is random.

The tree diagram for this is shown below.



Hence $P(R = 0) = \frac{24}{210}$

$P(R = 1) = \frac{108}{210}$

$P(R = 2) = \frac{72}{210}$

$P(R = 3) = \frac{6}{210}$

In tabular form this can be represented as:

r	0	1	2	3
$P(R = r)$	$\frac{24}{210}$	$\frac{108}{210}$	$\frac{72}{210}$	$\frac{6}{210}$

To check that it is random, we add together the probabilities.

$\frac{24}{210} + \frac{108}{210} + \frac{72}{210} + \frac{6}{210} = \frac{210}{210} = 1$

Thus we conclude that the number of red balls obtained is a random variable.

Alternatively the probabilities may be assigned using a function which is known as the probability density function (p.d.f) of X.

Example

The probability density function of a discrete random variable Y is given by $P(Y = y) = \frac{y^2}{14}$ for $y = 0, 1, 2$ and 3. Find $P(Y = y)$ for $y = 0, 1, 2$ and 3, verify that Y is a random variable and state the mode.

y	0	1	2	3
$P(Y = y)$	0	$\frac{1}{14}$	$\frac{4}{14}$	$\frac{9}{14}$

To check that it is random, we add together the probabilities.

$0 + \frac{1}{14} + \frac{4}{14} + \frac{9}{14} = \frac{14}{14} = 1$

Thus we conclude that Y is a random variable.

The mode is 3 since this is the value with the highest probability.

Example

The probability density function of a discrete random variable X is given by $P(X = x) = kx$ for $x = 9, 10, 11$ and 12. Find the value of the constant k.

In this case we are told that the variable is random and hence $\sum P(X = x) = 1$.

Therefore $9k + 10k + 11k + 12k = 1$

$$\Rightarrow k = \frac{1}{42}$$

Exercise 1

1 A discrete random variable X has this probability distribution:

x	0	1	2	3	4	5	6
$P(X = x)$	0.05	0.1	0.3	b	0.15	0.15	0.05

Find

a the value of b **b** $P(1 \le X \le 3)$ **c** $P(X < 4)$ **d** $P(1 < X \le 5)$

e the mode.

2 A discrete random variable X has this probability distribution:

x	4	5	6	7	8	9	10
$P(X = x)$	0.02	0.15	0.25	a	0.12	0.1	0.03

Find

a the value of a **b** $P(4 \le X \le 8)$ **c** $P(X < 8)$

d $P(4 < X < 8)$ **e** $P(5 < X < 7)$ **f** the mode.

3 Find the discrete probability distribution for X in the following cases and verify that the variable is random. X is defined as

a the number of tails obtained when three fair coins are tossed

b the number of black balls drawn with replacement from a bag of 4 black balls and 3 white balls, when 3 balls are picked

c the number of sixes obtained on a die when it is rolled three times

d the sum of the numbers when two dice are thrown

e the number of times David visits his local restaurant in three consecutive days, given that the probability of him visiting on any specific day is 0.2 and is an independent event.

4 Write down the discrete probability distributions given the following probability density functions:

a $P(X = x) = \frac{x^2}{55}$ for $0 \le x \le 5, x \in \mathbb{N}$

b $P(X = x) = \frac{x}{21}$ for $x = 1, 2, 3, 4, 5, 6$

c $P(Y = y) = \frac{y - 1}{30}$ for $y = 7, 8, 9, 10$

d $P(S = s) = \frac{s - 3}{42}$ for $s = 12, 13, 14, 15$

5 Find the value of k in each of the probability density functions shown below, such that the variable is random. In each case write out the probability distribution.

a $P(X = x) = k(x - 1)$ for $x = 3, 4, 5$

b $P(X = x) = k(x^2 - 1)$ for $x = 4, 5, 6$

c $P(Y = y) = ky^3$ for $y = 1, 2, 3, 4, 5$

d $P(B = b) = \frac{b + 2}{k}$ for $b = 3, 4, 5, 6, 7$

6 A man has six blue shirts and three grey shirts that he wears to work. Once a shirt is worn, it cannot be worn again in that week. If X is the discrete random variable "the number of blue shirts worn in the first three days of the week", find

a the probability distribution for X

b the probability that he wears at least one blue shirt during the first three days.

7 In a game a player throws three unbiased tetrahedral dice. If X is the discrete random variable "the number of fours obtained", find

a the probability distribution for X

b $P(X \ge 2)$.

8 Five women and four men are going on holiday. They are travelling by car and the first car holds four people including the driver. If Y is the discrete random variable "the number of women travelling in the first car", write down

a the probability distribution for Y

b the probability that there is at least one woman in the first car.

21.2 Expectation and variance

The expectation, E(X)

In a statistical experiment:

- A practical approach results in a frequency distribution and a mean value.
- A theoretical approach results in a probability distribution and an expected value.

The **expected value** is what we would expect the mean to be if a large number of terms were averaged.

The expected value is found by multiplying each score by its corresponding probability and summing.

$$\mathrm{E}(X) = \sum_{\text{all } x} x \cdot \mathrm{P}(X = x)$$

If the probability distribution is symmetrical about a mid-value, then E(X) will be this mid-value.

Example

The probability distribution of a discrete random variable X is as shown in the table.

x	1	2	3	4	5
$\mathrm{P}(X = x)$	0.2	0.4	a	0.1	0.05

a Find the value of a.
b Find E(X).

a Since it is random

$0.2 + 0.4 + a + 0.1 + 0.05 = 1$

$\Rightarrow a = 0.25$

b $\mathrm{E}(X) = 1 \times 0.2 + 2 \times 0.4 + 3 \times 0.25 + 4 \times 0.1 + 5 \times 0.05 = 2.4$

Example

The probability distribution of a discrete random variable Y is shown below.

y	5	6	7	8	9
$\mathrm{P}(Y = y)$	0.05	0.2	b	0.2	0.05

a Find the value of b.
b Find E(Y).

a Since the variable is random $0.05 + 0.2 + b + 0.2 + 0.05 = 1$

$\Rightarrow b = 0.5$

b In this case we could use the formula $\mathrm{E}(Y) = \sum_{\text{all } y} y \cdot \mathrm{P}(Y = y)$ to find the expectation, but because the distribution is symmetrical about $y = 7$ we can state immediately that $\mathrm{E}(Y) = 7$.

Example

A discrete random variable has probability density function $\mathrm{P}(X = x) = \frac{kx^3}{2}$ for $x = 1, 2, 3$ and 4.

a Find the value of k.
b Find E(X).

a $\frac{k}{2} + 4k + \frac{27k}{2} + 32k = 1$

$\Rightarrow k = \frac{1}{50}$

b $\mathrm{E}(X) = 1 \times \frac{1}{100} + 2 \times \frac{8}{100} + 3 \times \frac{27}{100} + 4 \times \frac{64}{100}$

$\Rightarrow \mathrm{E}(X) = 3.54$

Example

A discrete random variable X can only take the values 1, 2 and 3. If $\mathrm{P}(X = 1) = 0.15$ and $\mathrm{E}(X) = 2.4$, find the probability distribution for X.

The probability distribution for X is shown below:

x	1	2	3
$\mathrm{P}(X = x)$	0.15	p	q

Since the variable is random $0.15 + p + q = 1$

$\Rightarrow p + q = 0.85$

If $\mathrm{E}(X) = 2.4$ then $0.15 + 2p + 3q = 2.4$

$\Rightarrow 2p + 3q = 2.25$

Solving these two equations simultaneously gives $p = 0.3$ and $q = 0.55$.
Hence the probability distribution function for X is:

x	1	2	3
$\mathrm{P}(X = x)$	0.15	0.3	0.55

Example

Alan and Bob play a game in which each throws an unbiased die. The table below shows the amount in cents that Alan receives from Bob for each possible outcome of the game. For example, if both players throw a number greater than 3, Alan receives 50 cents from Bob while if both throw a number less than or equal to 3, Alan pays Bob 60 cents.

		B	
		≤ 3	**> 3**
A	**≤ 3**	−60	x
	> 3	40	50

Find

a **i** the expected value of Alan's gain in one game in terms of x

ii the value of x which makes the game fair to both players

iii the expected value of Alan's gain in 20 games if $x = 40$.

b Alan now discovers that the dice are biased and that the dice are three times more likely to show a number greater than 3 than a number less than or equal to 3. How much would Alan expect to win if $x = 30$?

a **i** On throwing a die, if X is the number thrown, then $P(X \le 3) = P(X > 3) = \frac{1}{2}$.

The probability of each combination of results for Alan and Bob is

$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Hence the probability distribution table is shown below where the discrete random variable X is Alan's gain.

x	-60	x	40	50
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

$$\Rightarrow E(X) = \frac{1}{4} \times -60 + \frac{1}{4}x + \frac{1}{4} \times 40 + \frac{1}{4} \times 50 = \frac{30 + x}{4}$$

ii If the game is fair, then neither player should gain or lose anything and hence $E(X) = 0$

$$\Rightarrow \frac{30 + x}{4} = 0$$

$$\Rightarrow x = -30$$

iii His expected gain in one game when $x = 40$ is $\frac{30 + 40}{4} = 17.5$ cents.

Hence his expected gain in 20 games is $20 \times 17.5 = 350$ cents.

b In this case the probability of die showing a number greater than 3 is $\frac{3}{4}$ and the probability of it showing a number less than or equal to 3 is $\frac{1}{4}$.

Hence the probability distribution table is now:

x	-60	30	40	50
$P(X = x)$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{9}{16}$

$$\Rightarrow E(X) = \frac{1}{16} \times -60 + \frac{3}{16} \times 30 + \frac{3}{16} \times 40 + \frac{9}{16} \times 50 = 37.5 \text{ cents}$$

The expectation of any function f(x)

If $E(X) = \sum_{\text{all } x} x \cdot P(X = x)$, then $E(X^2) = \sum_{\text{all } x} x^2 \cdot P(X = x)$, $E(X^3) = \sum_{\text{all } x} x^3 \cdot P(X = x)$ etc.

In general, $E(f(x)) = \sum_{\text{all } x} f(x) \cdot P(X = x)$.

Example

For the probability distribution shown below, find:

x	0	1	2	3	4	5
$P(X = x)$	0.08	0.1	0.2	0.4	0.15	0.07

a $E(X)$ **b** $E(X^2)$ **c** $E(2X)$ **d** $E(2X - 1)$

a $E(X) = 0 \times 0.08 + 1 \times 0.1 + 2 \times 0.2 + 3 \times 0.4 + 4 \times 0.15 + 5 \times 0.07 = 2.65$

b In this case the probability distribution is shown below:

x^2	0	1	4	9	16	25
$P(X = x^2)$	0.08	0.1	0.2	0.4	0.15	0.07

$$\Rightarrow E(X^2) = 0 \times 0.08 + 1 \times 0.1 + 4 \times 0.2 + 9 \times 0.4 + 16 \times 0.15 + 25 \times 0.07 = 8.65$$

c The probability distribution for this is:

$2x$	0	2	4	6	8	10
$P(X = 2x)$	0.08	0.1	0.2	0.4	0.15	0.07

$$\Rightarrow E(2X) = 0 \times 0.08 + 2 \times 0.1 + 4 \times 0.2 + 6 \times 0.4 + 8 \times 0.15 + 10 \times 0.07 = 5.3$$

d The probability distribution for this is:

$2x - 1$	-1	1	3	5	7	9
$P(X = 2x - 1)$	0.08	0.1	0.2	0.4	0.15	0.07

$$\Rightarrow E(2X - 1) = -1 \times 0.08 + 1 \times 0.1 + 3 \times 0.2 + 5 \times 0.4 + 7 \times 0.15 + 9 \times 0.07 = 4.3$$

This idea becomes important when we need to find the variance.

The variance, Var(X)

From Chapter 19, we know that for a frequency distribution with mean $\bar{x}$, the variance is given by

$$s^2 = \frac{\sum f(x-\bar{x})^2}{\sum f} \text{ or } s^2 = \frac{\sum fx^2}{\sum f} - \bar{x}^2$$

Using the first formula we can see that the variance is the mean of the squares of the deviations from the mean. If we now take a theoretical approach using a probability distribution from a discrete random variable, where we define $E(X) = \mu$ and apply the same idea, we find

$\text{Var}(X) = E(X - \mu)^2$.

However, we do not normally use this form and the alternative form we usually use is shown below.

$$\begin{aligned}\text{Var}(X) &= E(X-\mu)^2\\ &= E[X^2 - 2\mu X + \mu^2]\\ &= E(X^2) - 2\mu E(X) + \mu^2\\ &= E(X^2) - 2\mu^2 + \mu^2\\ &= E(X^2) - \mu^2\end{aligned}$$

$$\text{Var}(X) = E(X^2) - E^2(X)$$

The variance can never be negative. If it is, then a mistake has been made in the calculation.

Example

For the probability distribution shown below for a discrete random variable X, find:

x	-2	-1	0	1	2
$P(X = x)$	0.1	0.25	0.3	0.25	0.1

a $E(X)$ **b** $E(X^2)$ **c** $\text{Var}(X)$

a $E(X) = 0.3$ since the distribution is symmetrical.

b $E(X^2) = 4\times0.1 + 1\times0.25 + 0\times0.3 + 1\times0.25 + 4\times0.1 = 1.3$

c $\text{Var}(X) = E(X^2) - E^2(X)$

$= 1.75 - 0.3^2 = 1.21$

Example

A cubical die and a tetrahedral die are thrown together.

a If X is the discrete random variable "total scored", write down the probability distribution for X.

b Find $E(X)$.

c Find $\text{Var}(X)$.

A game is now played with the two dice. Anna has the cubical die and Beth has the tetrahedral die. They each gain points according to the following rules:

- If the number on both dice is greater than 3, then Beth gets 6 points.
- If the tetrahedral die shows 3 and the cubical die less than or equal to 3, then Beth gets 4 points.
- If the tetrahedral die shows 4 and the cubical die less than or equal to 3, then Beth gets 2 points.
- If the tetrahedral die shows a number less than 3 and the cubical die shows a 3, then Anna gets 5 points.
- If the tetrahedral die shows a number less than 3 and the cubical die shows a 1 or a 2, then Anna gets 3 points.
- If the tetrahedral die shows 3 and the cubical die greater than 3, then Anna gets 2 points.
- If the tetrahedral die shows a number less than 3 and the cubical die shows a number greater than 3, then Anna gets 1 point.

d Write out the probability distribution for Y, "the number of points gained by Anna".

e Calculate $E(Y)$ and $\text{Var}(Y)$.

f The game is now to be made fair by changing the number of points Anna gets when the tetrahedral die shows a number less than 3 and the cubical die shows a 1 or a 2. What is this number of points to the nearest whole number?

a A probability space diagram is the easiest way to show the possible outcomes.

6	7	8	9	10
5	6	7	8	9
4	5	6	7	8
3	4	5	6	7
2	3	4	5	6
1	2	3	4	5
	1	2	3	4

Hence the probability distribution for X is:

x	2	3	4	5	6	7	8	9	10
$P(X = x)$	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{3}{24}$	$\frac{4}{24}$	$\frac{4}{24}$	$\frac{4}{24}$	$\frac{3}{24}$	$\frac{2}{24}$	$\frac{1}{24}$

b Since the probability distribution is symmetrical, $E(X) = 6$.

c $E(X^2) = 4\times\frac{1}{24} + 9\times\frac{2}{24} + 16\times\frac{3}{24} + 25\times\frac{4}{24} + 36\times\frac{4}{24}$

$+ 49\times\frac{4}{24} + 64\times\frac{3}{24} + 81\times\frac{2}{24} + 100\times\frac{1}{24}$

$= \frac{964}{24} = \frac{241}{6}$

$\text{Var}(X) = E(X^2) - E^2(X)$

$= \frac{241}{6} - 6^2 = \frac{25}{6}$

d By considering the possibility space diagram again the probability distribution is:

y	−6	−4	−2	5	3	2	1
$P(Y = y)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$

e $E(Y) = -6\times\frac{1}{8} - 4\times\frac{1}{8} - 2\times\frac{1}{8} + 5\times\frac{1}{12} + 3\times\frac{1}{6} + 2\times\frac{1}{8} + 1\times\frac{1}{4} = -\frac{1}{12}$

$E(Y^2) = 36\times\frac{1}{8} + 16\times\frac{1}{8} + 4\times\frac{1}{8} + 25\times\frac{1}{12} + 9\times\frac{1}{6} + 4\times\frac{1}{8} + 1\times\frac{1}{4} = \frac{34}{3}$

$\text{Var}(Y) = E(Y^2) - E^2(Y)$

$= \frac{34}{3} - \frac{1}{144} = \frac{1631}{144}$

f Let the number of points Anna gains when the tetrahedral die shows a number less than 3 and the cubical die shows a 1 or a 2 be y.

In this case the probability distribution is now:

y	−6	−4	−2	5	y	2	1
$P(Y = y)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{4}$

$E(Y) = -6\times\frac{1}{8} - 4\times\frac{1}{8} - 2\times\frac{1}{8} + 5\times\frac{1}{12} + y\times\frac{1}{6} + 2\times\frac{1}{8} + 1\times\frac{1}{4} = -\frac{1}{12}$

Since the game is now fair, $E(Y) = 0$

$\Rightarrow -\frac{7}{12} + \frac{y}{6} = 0$

$\Rightarrow y = 3.5$

Exercise 2

1 Find the value of b and $E(X)$ in these distributions.

a

x	3	4	5	6
$P(X = x)$	0.1	b	0.4	0.2

b

x	0	1	2	3	4
$P(X = x)$	0.05	1.5	0.6	b	0.05

c

x	−3	0	3	6	9
$P(X = x)$	0.15	0.2	b	0.22	0.15

d

x	−4	1	3	5	6
$P(X = x)$	0.1	0.25	0.3	0.25	b

2 If three unbiased cubical dice are thrown, what is the expected number of threes that will occur?

3 A discrete random variable has a probability distribution function given by $f(x) = \frac{cx^2}{12}$ for $x = 1, 2, 3, 4, 5, 6$.

a Find the value of c. **b** Find $E(X)$.

4 Caroline and Lisa play a game that involves each of them tossing a fair coin. The rules are as follows:

- If Caroline and Lisa both get heads, Lisa gains 6 points.
- If Caroline and Lisa both get tails, Caroline gains 6 points.
- If Caroline gets a tail and Lisa gets a head, Lisa gains 3 points.
- If Caroline gets a head and Lisa gets a tail, Caroline gains x points.

a If X is the discrete random variable "Caroline's gain", find $E(X)$ in terms of x.

b What value of x makes the game fair?

5 A discrete random variable X can only take the values -1 and 1. If $E(X) = 0.4$, find the probability distribution for X.

6 A discrete random variable X can only take the values -1, 1 and 3. If $P(X = 1) = 0.25$ and $E(X) = 1.9$, find the probability distribution for X.

7 A discrete random variable Y can only take the values 0, 2, 4 and 6. If $P(Y \le 4) = 0.6$, $P(Y \le 2) = 0.5$, $P(Y = 2) = P(Y = 4)$ and $E(Y) = 2.4$, find the probability distribution for Y.

8 A five-a-side soccer team is to be chosen from four boys and five girls. If the team members are chosen at random, what is the expected number of girls on the team?

9 In a chemistry examination each question is a multiple choice with four possible answers. Given that Kevin randomly guesses the answers to the first four questions, how many of the first four questions can he expect to get right?

10 A discrete random variable X has probability distribution:

x	0	1	2	3	4
$P(X = x)$	0.1	0.2	0.35	0.25	0.1

Find:

a $E(X)$ **b** $E(X^2)$ **c** $E(2X - 1)$ **d** $E(3X + 2)$

11 A discrete random variable X has probability distribution:

x	−2	0	2	4	6
$P(X = x)$	0.05	0.15	0.25	0.35	0.2

Find:

a $E(X)$ **b** $E(X^2)$ **c** $E(2X + 1)$ **d** $E(3X - 1)$

12 Find $\text{Var}(X)$ for each of these probability distributions.

a

x	0	1	2	3	4
$P(X = x)$	0.2	0.2	0.3	0.15	0.15

b

x	1	3	5	7	9
$P(X = x)$	0.05	0.2	0.2	0.3	0.25

c

x	−4	−2	0	2	4
$P(X = x)$	0.1	0.3	p	0.2	0.15

d

x	−2	−1	0	1	2
$P(X = x)$	0.03	0.2	p	0.35	0.1

13 If X is the sum of the numbers shown when two unbiased dice are thrown, find:
a $E(X)$ **b** $E(X^2)$ **c** $Var(X)$

14 Three members of a school committee are to be chosen from three boys and four girls. If Y is the random variable "number of boys chosen", find:
a $E(Y)$ **b** $E(Y^2)$ **c** $Var(Y)$

15 A discrete random variable X has probability distribution:

x	0	1	2	3	4
$P(X = x)$	k	0.2	$2k$	0.3	$4k$

a Find the value of k. **b** Calculate $E(X)$. **c** Calculate $Var(X)$.

16 A teacher randomly selects 4 students from a class of 15 to attend a careers talk. In the class there are 7 girls and 8 boys. If Y is the number of girls selected and each selection is independent of the others, find
a the probability distribution for Y **b** $E(Y)$ **c** $Var(Y)$

17 A discrete random variable X takes the values $x = 1, 3, 5$, with probabilities $\frac{1}{7}, \frac{5}{14}$ and k respectively. Find
a k
b the mean of X
c the standard deviation of X.

18 One of the following expressions can be used as a probability density function for a discrete random variable X. Identify which one and calculate its mean and standard deviation.

a $f(x) = \frac{x^2 + 1}{35}, x = 0, 1, 2, 3, 4$

b $g(x) = \frac{x - 1}{7}, x = 0, 1, 2, 4, 5$

19 Jim has been writing letters. He has written four letters and has four envelopes addressed. Unfortunately he drops the letters on the floor and he has no way of distinguishing which letters go in which envelopes so he puts each letter in each envelope randomly. Let X be the number of letters in their correct envelopes.
a State the values which X can take.
b Find the probabilities for these values of X.
c Calculate the mean and variance for X.

20 A box contains ten numbered discs. Three of the discs have the number 5 on them, four of the discs have the number 6 on them, and three of the discs have the number 7 on them. Two discs are drawn without replacement and the score is the sum of the numbers shown on the discs. This is denoted by X.
a Write down the values that X can take.
b Find the probabilities of these values of X.
c Calculate the expectation and variance of X.
d Two children, Ahmed and Belinda, do this. Find the probability that Ahmed gains a higher score than Belinda.

21 Pushkar buys a large box of fireworks. The probability of there being X fireworks that fail is shown in the table below.

x	0	1	2	3	≥ 4
$P(X = x)$	$9k$	$3k$	k	k	0

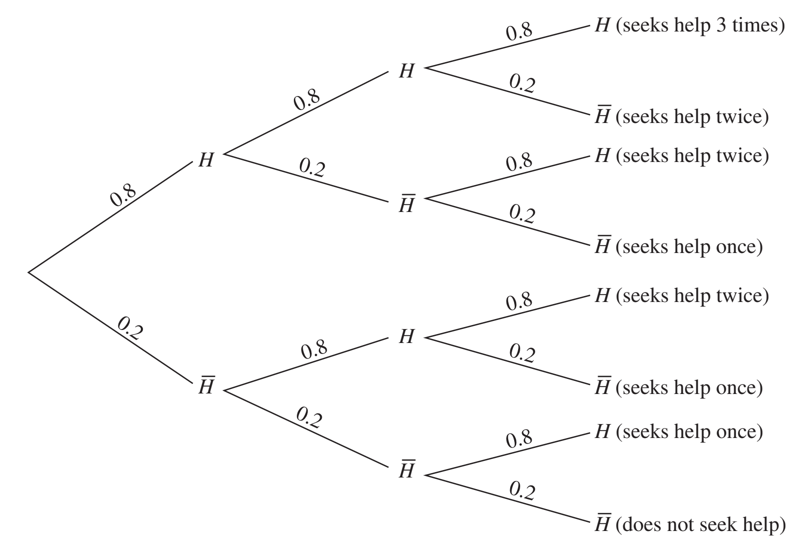
a Find the value of k.
b Find $E(X)$ and $Var(X)$.
c His friend, Priya, also buys a box. They put their fireworks together and the total number of fireworks that fail, Y, is determined. What values can Y take?
d Write down the distribution for Y.
e Find the expectation and variance of Y.

21.3 Binomial distribution

This is a distribution that deals with events that either occur or do not occur, so there are two complementary outcomes. We are usually told the number of times an event occurs and we are given the probability of the event happening or not happening.

Consider the three pieces of Mathematics Higher Level homework done by Jay. The probability of him seeking help from his teacher is 0.8.

The tree diagram to represent this is shown below.



If X is the number of times he seeks help, then from the tree diagram:

$P(X = 0) = 0.2 \times 0.2 \times 0.2 = 0.008$

By using the different branches of the tree diagram we can calculate the values for $x = 1, 2, 3$.

Without using the tree diagram we can see that the probability of him never seeking help is $0.2 \times 0.2 \times 0.2$ and this can happen in ${}^3C_0$ ways, giving $P(X = 0) = 0.008$.

The probability of him seeking help once is $0.8 \times 0.2 \times 0.2$ and this can happen in ${}^3C_1$ ways, giving $P(X = 1) = 0.096$.

The probability of him seeking help twice is $0.8 \times 0.8 \times 0.2$ and this can happen in ${}^3C_2$ ways, giving $P(X = 2) = 0.384$.

The probability of him seeking help three times is $0.8 \times 0.8 \times 0.8$ and this can happen in ${}^3C_3$ ways, giving $P(X = 3) = 0.512$.

Without using the tree diagram we can see that for 20 homeworks, say, the probability of him seeking help once would be ${}^{20}C_1 \times 0.8 \times 0.2^{19}$.

If we were asked to do this calculation using a tree diagram it would be very time consuming!

Generalizing this leads to a formula for a binomial distribution.

If a random variable X follows a **binomial distribution** we say $X \sim \text{Bin}(n, p)$ where n = number of times an event occurs and p = probability of success.
The probability of failure $= q = 1 - p$.
n and p are called the parameters of the distribution.

If $X \sim \text{Bin}(n, p)$ then $P(X = x) = {}^nC_x p^x q^{n-x}$.

Example

If $X \sim \text{Bin}\left(7, \frac{1}{4}\right)$, find:

a $P(X = 6)$

b $P(X \leq 2)$

a In this case $n = 7$, $p = \frac{1}{4}$ and $q = 1 - \frac{1}{4} = \frac{3}{4}$.

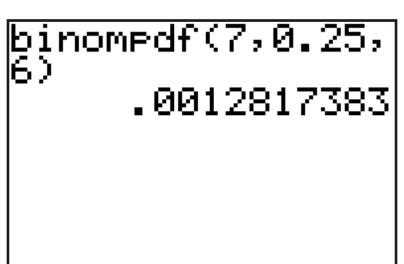
Hence $P(X = 6) = {}^7C_6 \left(\frac{1}{4}\right)^6\left(\frac{3}{4}\right)^1 = 0.00128$

b $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

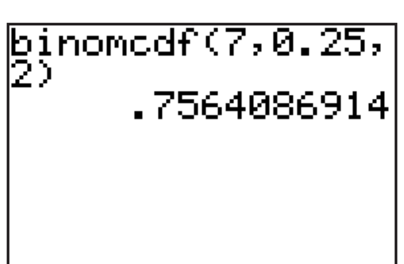
$$= {}^7C_0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^7 + {}^7C_1\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^6 + {}^7C_2\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^5 = 0.756$$

It is usual to do these calculations on a calculator. The screen shots for these are shown below.

a



b



So how do we recognize a binomial distribution? For a distribution to be binomial there must be an event that happens a finite number of times and the probability of that event happening must not change and must be independent of what happened before. Hence if we have 8 red balls and 6 black balls in a bag, and we draw 7 balls from the bag one after the other with replacement, X = "the number of red balls drawn" follows a binomial distribution. Here the number of events is 7 and the probability of success (drawing a red ball) is constant. If the problem were changed to the balls not being replaced, then the probability of drawing a red ball would no longer be constant and the distribution would no longer follow a binomial distribution.

Example

Market research is carried out at a supermarket, looking at customers buying cans of soup. If a customer buys one can of soup, the probability that it is tomato soup is 0.75. If ten shoppers buy one can of soup each, what is the probability that

a exactly three buy tomato soup

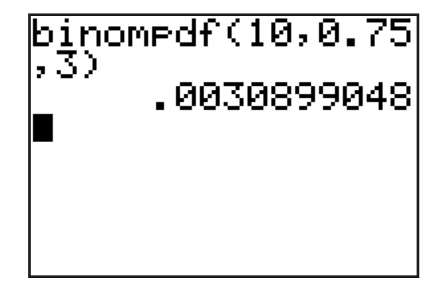
b less than six buy tomato soup

c more than four buy tomato soup?

The distribution for this is $X \sim \text{Bin}(10, 0.75)$.

a $P(X = 3) = {}^{10}C_3\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^7 = 0.00309$

Or from the calculator:

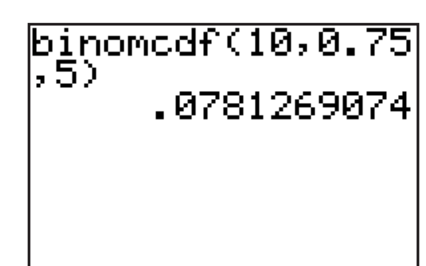


b $P(X < 6) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$$= {}^{10}C_0\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)^{10} + {}^{10}C_1\left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right)^9 + {}^{10}C_2\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^8 + {}^{10}C_3\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^7 + {}^{10}C_4\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right)^6 + {}^{10}C_5\left(\frac{3}{4}\right)^5\left(\frac{1}{4}\right)^5$$

$= 0.0781$

Or from the calculator:



c In a binomial distribution, the sum of the probabilities is one and hence it is sometimes easier to subtract the answer from one.

In this case $P(X > 4) = 1 - P(X \le 4)$

$$= 1 - \{P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)\}$$

$$= 1 - \left\{{}^{10}C_0\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)^{10} + {}^{10}C_1\left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right)^9 + {}^{10}C_2\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^8 + {}^{10}C_3\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^7 + {}^{10}C_4\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right)^6\right\}$$

$$= 1 - 0.0197\ldots = 0.980$$

On the calculator we also subtract the answer from 1.

```
binomcdf(10,0.75
,4)
       .0197277069
1-Ans
       .9802722931
```

In questions involving discrete distributions, ensure you read the question. If a question asks for more than 2, this is different from asking for at least 2. This also affects what is inputted into the calculator.

Example

Scientists have stated that in a certain town it is equally likely that a woman will give birth to a boy or a girl. In a family of seven children, what is the probability that there will be at least one girl?

"At least" problems, i.e. finding $P(X \ge x)$, can be dealt with in two ways. Depending on the number, we can either calculate the answer directly or we can work out $P(X < x)$, and subtract the answer from 1.

In this case, $X \sim \text{Bin}(7, 0.5)$ and we want $P(X \ge 1)$.

$$P(X \ge 1) = 1 - P(X = 0)$$

$$= 1 - {}^7C_7(0.5)^7(0.5)^0 = 0.992$$

Or from the calculator:

```
binompdf(7,0.5,0
)
         .0078125
1-Ans
         .9921875
```

Example

The probability of rain on any particular day in June is 0.45. In any given week in June, what is the most likely number of days of rain?

If we are asked to find the most likely value, then we should work through all the probabilities and then state the value with the highest probability. In this case the calculator is very helpful.

If we let X be the random variable "the number of rainy days in a week in June", then the distribution is $X \sim \text{Bin}(7, 0.45)$.

From the calculator, the results are:

x	$P(X = x)$
0	0.0152 …
1	0.0871 …
2	0.214 …
3	0.291 …
4	0.238 …
5	0.117 …
6	0.0319 …
7	0.00373 …

Hence we can state that the most likely number of days is 3.

If asked to do a question of this sort, it is not usual to write out the whole table. It is enough to write down the highest value and one either side and state the conclusion from there. This is because in the binomial distribution the probabilities increase to a highest value and then decrease again and hence once we have found where the highest value occurs we know it will not increase beyond this value elsewhere.

Expectation and variance of a binomial distribution

If $X \sim \text{Bin}(n, p)$
$E(X) = np$
$\text{Var}(X) = npq$

The proofs for these are shown below, but they will not be asked for in examination questions.

Proof that $E(X) = np$

Let $X \sim \text{Bin}(n, p)$

Hence $P(X = x) = {}^nC_x p^x q^{n-x}$

Therefore the probability distribution for this is:

x	0	1	2	…	n
$P(X = x)$	q^n	$nq^{n-1}p$	$\frac{n(n-1)}{2!}q^{n-2}p^2$		p^n

Now $E(X) = \sum_{\text{all } x} x \cdot P(X = x)$

$$= 0 \cdot q^n + 1 \cdot nq^{n-1}p + 2 \cdot \frac{n(n-1)}{2!}q^{n-2}p^2 + \ldots + n \cdot p^n$$

$$= np[q^{n-1} + (n-1)q^{n-2}p + \ldots + p^{n-1}]$$

$$= np(q + p)^{n-1}$$

Since $q + p = 1$, $E(X) = np$.

Proof that $\text{Var}(X) = npq$

$\text{Var}(X) = \text{E}(X^2) - \text{E}^2(X)$

Now $\text{E}(X^2) = \sum_{\text{all } x} x^2 \cdot \text{P}(X = x)$

$$= 0 \cdot q^n + 1 \cdot nq^{n-1}p + 4 \cdot \frac{n(n-1)}{2!}q^{n-2}p^2 + 9 \cdot \frac{n(n-1)(n-2)}{3!}q^{n-3}p^3 + \ldots + n^2 \cdot p^n$$

$$= np\left[q^{n-1} + 2(n-1)q^{n-2}p + \frac{3(n-1)(n-2)}{2!}q^{n-3}p^2 + \ldots + np^{n-1}\right]$$

This can be split into two series:

$$= np\left\{\left[q^{n-1} + (n-1)q^{n-2}p + \frac{(n-1)(n-2)}{2!}q^{n-3}p^2 + \ldots + p^{n-1}\right] + \left[(n-1)q^{n-2}p + \frac{2(n-1)(n-2)}{2!}q^{n-3}p^2 + \ldots + (n-1)p^{n-1}\right]\right\}$$

$$= np(q+p)^{n-1} + np\left[(n-1)q^{n-2}p + \frac{2(n-1)(n-2)}{2!}q^{n-3}p^2 + \ldots + (n-1)p^{n-1}\right]$$

Since the first series is the same as the one in the proof of $\text{E}(X)$.

$$= np\{1 + (n-1)p[q^{n-2} + (n-2)q^{n-3}p + \cdots + p^{n-2}]\}$$

Since $p + q = 1$.

$$= np\{1 + (n-1)p(q+p)^{n-2}\}$$

$$= np\{1 + (n-1)p\}$$

Again since $p + q = 1$.

Hence $\text{Var}(X) = np\{1 + (n-1)p\} - (np)^2$

$$= np + n^2p^2 - np^2 - n^2p^2$$

$$= np(1-p) = npq$$

Example

X is a random variable such that $X \sim \text{Bin}(n, p)$. Given that $\text{E}(X) = 3.6$ and $p = 0.4$, find n and the standard deviation of X.

Since $p = 0.4$ then $q = 1 - 0.4 = 0.6$.

Using the formula $\text{E}(X) = np$

$\Rightarrow 3.6 = 0.4n$

$\Rightarrow n = 9$

$\text{Var}(X) = npq = 9 \times 0.4 \times 0.6 = 2.16$

Hence the standard deviation is $\sqrt{2.16} = 1.47$.

21 Discrete Probability Distributions

Example

In a class mathematics test, the probability of a girl passing the test is 0.62 and the probability of a boy passing the test is 0.65. The class contains 15 boys and 17 girls.

a What is the expected number of boys to pass?
b What is the most likely number of girls to pass?
c What is the probability that more than eight boys fail?

If X is the random variable "the number of boys who pass" and Y is the random variable "the number of girls who pass", then $X \sim \text{Bin}(15, 0.65)$ and $Y \sim \text{Bin}(17, 0.62)$.

a $\text{E}(X) = np = 15 \times 0.65 = 9.75$

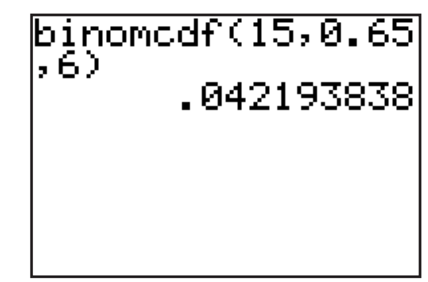
b From the calculator the results are:

y	10	11	12
$\text{P}(Y = y)$	0.186 …	0.193 …	0.158 …

Hence we can state that the most likely number of girls passing is 11.

The expectation is the theoretical equivalent of the mean, whereas the most likely is the equivalent of the mode.

c The probability of more than eight boys failing is the same as the probability of no more than six boys passing, hence we require $\text{P}(X \leq 6)$.



Therefore the probability that more than eight boys fail is 0.0422.

Example

Annabel always takes a puzzle book on holiday with her and she attempts a puzzle every day. The probability of her successfully solving a puzzle is 0.7. She goes on holiday for four weeks.

a Find the expected value and the standard deviation of the number of successfully solved puzzles in a given week.
b Find the probability that she successfully solves at least four puzzles in a given week.
c She successfully solves a puzzle on the first day of the holiday. What is the probability that she successfully solves at least another three during the rest of that week?
d Find the probability that she successfully solves four or less puzzles in only one of the four weeks of her holiday.

Let X be the random variable "the number of puzzles successfully completed by Annabel". Hence $X \sim \text{Bin}(7, 0.7)$.

a $E(X) = np = 7 \times 0.7 = 4.9$

$\text{Var}(X) = npq = 7 \times 0.7 \times 0.3 = 1.47$

Hence standard deviation $= \sqrt{1.47} = 1.21$

b $P(X \geq 4) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]$

$$= 1 - \left\{ {}^7C_0(0.7)^0(0.3)^7 + {}^7C_1(0.7)^1(0.3)^6 + {}^7C_2(0.7)^2(0.3)^5 + {}^7C_3(0.7)^3(0.3)^4 \right\}$$

$= 1 - 0.126\ldots = 0.874$

```
binomcdf(7,0.7,3
)
            .126036
1-Ans
            .873964
```

c This changes the distribution and we now want $P(Y \geq 3)$ where $Y \sim \text{Bin}(6, 0.7)$.

$P(Y \geq 3) = 1 - P(Y \leq 2)$

$$= 1 - \left\{ {}^6C_0(0.7)^0(0.3)^6 + {}^6C_1(0.7)^1(0.3)^5 + {}^6C_2(0.7)^2(0.3)^4 \right\}$$

$= 1 - 0.704\ldots = 0.930$

```
binomcdf(6,0.7,2
)
             .07047
1-Ans
             .92953
```

d We first calculate the probability that she successfully completes four or less in a week, $P(X \leq 4)$.

$P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= {}^7C_0(0.7)^0(0.3)^7 + {}^7C_1(0.7)^1(0.3)^6 + {}^7C_2(0.7)^2(0.3)^5 + {}^7C_3(0.7)^3(0.3)^4 + {}^7C_4(0.7)^4(0.3)^3$$

$= 0.353$

```
binomcdf(7,0.7,4
)
           .3529305
```

We now want $P(A = 1)$ where $A \sim \text{Bin}(4, 0.353)$.

$P(A = 1) = {}^4C_1(0.353)^1(0.647)^3 = 0.382$

Exercise 3

1 If $X \sim \text{Bin}(7, 0.35)$, find:
a $P(X = 3)$ **b** $P(X \leq 2)$ **c** $P(X > 4)$

2 If $X \sim \text{Bin}(10, 0.4)$, find:
a $P(X = 5)$ **b** $P(X \geq 3)$ **c** $P(X \leq 5)$

3 If $X \sim \text{Bin}(8, 0.25)$, find:
a $P(X = 3)$ **b** $P(X \geq 5)$ **c** $P(X \leq 4)$ **d** $P(X = 0 \text{ or } 1)$

4 A biased coin is tossed ten times. On each toss, the probability that it will land on a head is 0.65. Find the probability that it will land on a head at least six times.

5 Given that $X \sim \text{Bin}(6, 0.4)$, find
a $E(X)$ **b** $\text{Var}(X)$ **c** the most likely value for X.

6 In a bag of ten discs, three of them are numbered 5 and seven of them are numbered 6. A disc is drawn at random, the number noted, and then it is replaced. This happens eight times. Find
a the expected number of 5's
b the variance of the number of 7's drawn
c the most likely number of 5's drawn.

7 A random variable Y follows a binomial distribution with mean 1.75 and variance 1.3125.
a Find the values of n, p and q.
b What is the probability that Y is less than 2?
c Find the most likely value(s) of Y.

8 An advert claims that 80% of dog owners, prefer Supafood dog food. In a sample of 15 dog owners, find the probability that
a exactly seven buy Supafood
b more than eight buy Supafood
c ten or more buy Supafood.

9 The probability that it will snow on any given day in January in New York is given as 0.45. In any given week in January, find the probability that it will snow on
a exactly one day **b** more than two days
c at least three days **d** no more than four days.

10 A student in a mathematics class has a probability of 0.68 of gaining full marks in a test. She takes nine tests in a year. What is the probability that she will
a never gain full marks
b gain full marks three times in a year
c gain full marks in more than half the tests
d gain full marks at least eight times?

11 Alice plays a game that involves kicking a small ball at a target. The probability that she hits the target is 0.72. She kicks the ball eight times.
a Find the probability that she hits the target exactly five times.
b Find the probability that she hits the target for the first time on her fourth kick.

12 In a school, 19% of students fail the IB Diploma. Find the probability that in a class of 15 students
a exactly two will fail **b** less than five will fail
c at least eight will pass.

13 A factory makes light bulbs that it distributes to stores in boxes of 20. The probability of a light bulb being defective is 0.05.

a Find the probability that there are exactly three defective bulbs in a box of light bulbs.

b Find the probability that there are more than four defective light bulbs in a box.

c If a certain store buys 25 boxes, what is the probability that at least two of them have more than four defective light bulbs?

The quality control department in the company decides that if a randomly selected box has no defective light bulbs in it, then all bulbs made that day will pass and if it has two or more defective light bulbs in it, then all light bulbs made that day will be scrapped. If it has one defective light bulb in it, then another box will be tested, and if that has no defective light bulbs in it, all light bulbs made that day will pass. Otherwise all light bulbs made that day will be scrapped.

d What is the probability that the first box fails but the second box passes?

e What is the probability that all light bulbs made that day will be scrapped?

14 A multiple choice test in biology consists of 40 questions, each with four possible answers, only one of which is correct. A student chooses the answers to the questions at random.

a What is the expected number of correct answers?

b What is the standard deviation of the number of correct answers?

c What is the probability that the student gains more than the expected number of correct answers?

15 In a chemistry class a particular experiment is performed with a probability of success p. The outcomes of successive experiments are independent.

a Find the value of p if probability of gaining three successes in six experiments is the same as gaining four successes in seven experiments.

b If p is now given as 0.25, find the number of times the experiment must be performed in order that the probability of gaining at least one success is greater than 0.99.

16 The probability of the London to Glasgow train being delayed on a weekday is $\frac{1}{15}$. Assuming that the delays occur independently, find

a the probability that the train experiences exactly three delays in a five-day week

b the most likely number of delays in a five-day week

c the expected number of delays in a five-day week

d the number of days such that there is a 20% probability of the train having been delayed at least once

e the probability of being delayed at least twice in a five-day week

f the probability of being delayed at least twice in each of two weeks out of a four-week period (assume each week has five days in it).

17 It is known that 14% of a large batch of light bulbs is defective. From this batch of light bulbs, 15 are selected at random.

a Write down the distribution and state its mean and variance.

b Calculate the most likely number of defective light bulbs.

c What is the probability of exactly three defective light bulbs?

d What is the probability of at least four defective light bulbs?

e If six batches of 15 light bulbs are selected randomly, what is the probability that at least three of them have at least four defective light bulbs?

18 In the game scissors, paper, rock, a girl never chooses paper, and is twice as likely to choose scissors as rock. She plays the game eight times.

a Write down the distribution for X, the number of times she chooses rock.

b Find $P(X = 1)$.

c Find $E(X)$.

d Find the probability that X is at least one.

19 On a statistics course at a certain university, students complete 12 quizzes. The probability that a student passes a quiz is $\frac{2}{3}$.

a What is the expected number of quizzes a student will pass?

b What is the probability that the student will pass more than half the quizzes?

c What is the most likely number of quizzes that the student will pass?

d At the end of the course, the student takes an examination. The probability of passing the examination is $\frac{n}{55}$, given that n is the number of quizzes passed. What is the probability that the student passes four quizzes and passes the examination?

21.4 Poisson distribution

Consider an observer counting the number of cars passing a specific point on a road during 100 time intervals of 30 seconds. He finds that in these 100 time intervals a total of 550 cars pass.

Now if we assume from the beginning that 550 cars will pass in these time intervals, that a car passing is independent of another car passing, and that it is equally likely that they will pass in any of the time intervals, then the probability that a car passes in any specific time interval is $\frac{1}{100}$. The probability that a second car arrives in this time interval is also $\frac{1}{100}$ as the events are independent, and so on. Hence the number of cars passing this point in this time period follows a binomial distribution $X \sim \text{Bin}\left(550, \frac{1}{100}\right)$.

Unfortunately, this is not really the case as we do not know exactly how many cars will pass in any interval. What we do know from experience is the mean number of cars that will pass. Also, as n gets larger, p must become smaller. That is, the more cars we observe, the less likely it is that a specific car will pass in a given interval. Hence the distribution we want is one where n increases as p decreases and where the mean np stays constant. This is called a Poisson distribution and occurs when an event is evenly spaced, on average, over an infinite space.

If a random variable X follows a **Poisson distribution**, we say $X \sim \text{Po}(\lambda)$ where λ is the parameter of the distribution and is equal to the mean of the distribution.

$$\text{If } X \sim \text{Po}(\lambda) \text{ then } P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}.$$

Example

If $X \sim \text{Po}(2)$, find:

a $P(X = 3)$

b $P(x \leq 4)$

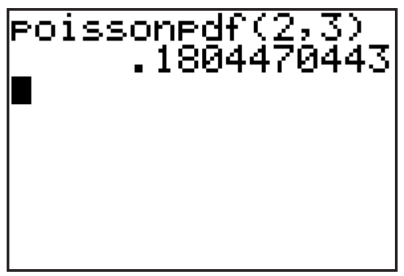
a $P(X = 3) = \dfrac{e^{-2}2^3}{3!} = 0.180$

b $P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

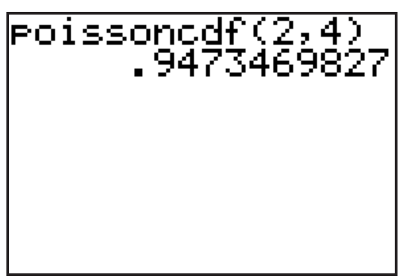
$$= \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!} + \frac{e^{-2}2^4}{4!} = 0.947$$

As with the binomial distribution, it is usual to do these calculations on the calculator.

a



b



To recognize a Poisson distribution we normally have an event that is randomly scattered in time or space and has a mean number of occurrences in a given interval of time or space.

Unlike the binomial distribution, X can take any positive integer value up to infinity and hence if we want $P(X \geq x)$ we must always subtract the answer from 1. As x becomes very large, the probability becomes very small.

Example

The mean number of zebra per square kilometre in a game park is found to be 800. Given that the number of zebra follows a Poisson distribution, find the probability that in one square kilometre of game park there are

a 750 zebra

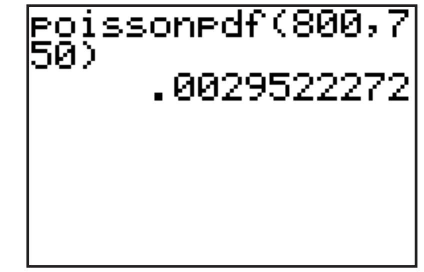
b less than 780 zebra

c more than 820 zebra.

Let X be the number of zebra in one square kilometre.
Hence $X \sim \text{Po}(800)$.

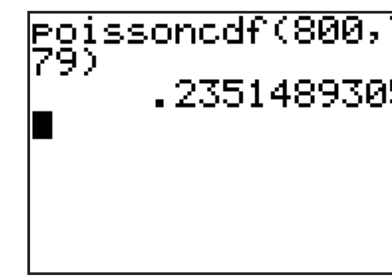
a We require $P(X = 750) = e^{-800} \cdot \dfrac{800^{750}}{750!}$.

Because of the numbers involved, we have to use the Poisson function on a calculator.



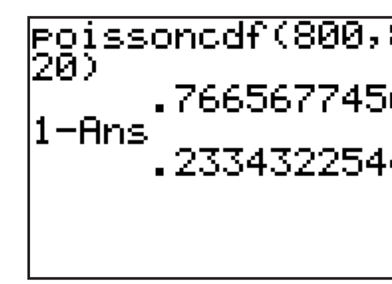
$P(X = 750) = 0.00295$

b In this case we have to use a calculator. We want less than 780, which is the same as less than or equal to 779.



$P(X < 780) = 0.235$

c We calculate $P(X > 820)$ using $1 - P(X \leq 820)$ on a calculator.



$P(X > 820) = 0.233$

With a Poisson distribution we are sometimes given the mean over a certain interval. We can sometimes assume that this can then be recalculated for a different interval.

Example

The mean number of telephone calls arriving at a company's reception is five per minute and follows a Poisson distribution. Find the probability that there are

a exactly six phone calls in a given minute
b more than three phone calls in a given minute
c more than 20 phone calls in a given 5-minute period
d less than ten phone calls in a 3-minute period.

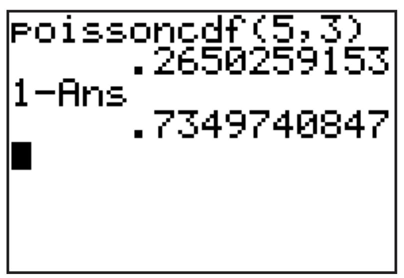
Let X be the "number of telephone calls in a minute". Hence $X \sim \text{Po}(5)$.

a $\text{P}(X = 6) = \dfrac{e^{-5}5^6}{6!} = 0.146$

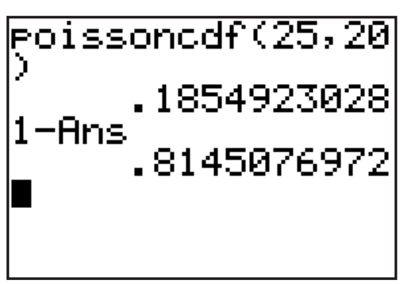
b $\text{P}(X > 3) = 1 - [\text{P}(X = 0) + \text{P}(X = 1) + \text{P}(X = 2) + P(X = 3)]$

$$= 1 - \left\{\frac{e^{-5}5^0}{0!} + \frac{e^{-5}5^1}{1!} + \frac{e^{-5}5^2}{2!} + \frac{e^{-5}5^3}{3!}\right\} = 0.735$$

On a calculator:

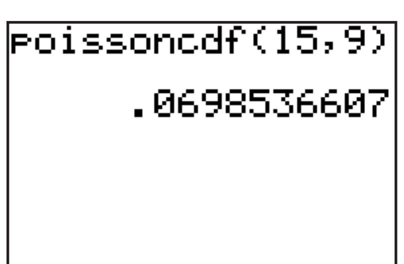


c If there are five calls in a minute period, then in a 5-minute period there are, on average, 25 calls. Hence if Y is "the number of telephone calls in a 5-minute period", then $Y \sim \text{Po}(25)$. We require $\text{P}(Y > 20)$. Because of the numbers involved we need to solve this on a calculator.



$\text{P}(Y > 20) = 0.815$

d If A is "the number of telephone calls in a 3-minute period", then $A \sim \text{Po}(15)$. We require $\text{P}(Y < 10)$. Because of the numbers involved, again we solve this on a calculator.



$\text{P}(Y < 10) = 0.0699$

Example

Passengers arrive at the check-in desk of an airport at an average rate of seven per minute.
Assuming that the passengers arriving at the check-in desk follow a Poisson distribution, find

a the probability that exactly five passengers will arrive in a given minute
b the most likely number of passengers to arrive in a given minute
c the probability of at least three passengers arriving in a given minute
d the probability of more than 30 passengers arriving in a given 5-minute period.

If X is "the number of passengers checking-in in a minute", then $X \sim \text{Po}(7)$.

a $\text{P}(X = 5) = \dfrac{e^{-7}7^5}{5!} = 0.128$

b As with the binomial distribution, we find the probabilities on a calculator and look for the highest. This time we select a range of values around the mean. As before, it is only necessary to write down the ones either side as the distribution rises to a maximum probability and then decreases again. Written as a table, the results are:

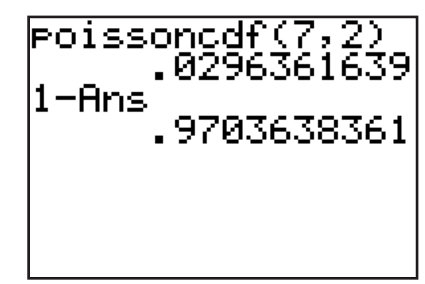
x	5	6	7	8
$\text{P}(X = x)$	0.127…	0.149…	0.149…	0.130…

Since there are two identical probabilities in this case, the most likely value is either 6 or 7.

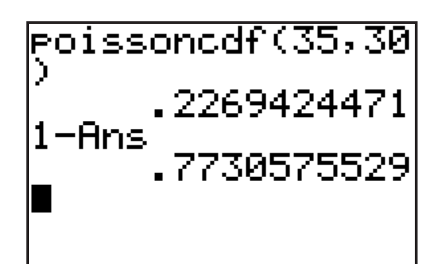
c $\text{P}(X \geq 3) = 1 - \text{P}(X \leq 2)$

$$= 1 - \left\{\frac{e^{-7}7^0}{0!} + \frac{e^{-7}7^1}{1!} + \frac{e^{-7}5^2}{2!}\right\} = 0.970$$

On the calculator:



d If seven people check-in in a minute period, then on average 35 people will check-in in a 5-minute period. Hence if Y is "the number of people checking-in in a 5-minute period", then $Y \sim \text{Po}(35)$. We require $\text{P}(Y > 30)$. Because of the numbers involved we need to solve this on a calculator.



$\text{P}(Y > 30) = 0.773$

Expectation and variance of a Poisson distribution

If $X \sim \text{Po}(\lambda)$
$\text{E}(X) = \lambda$
$\text{Var}(X) = \lambda$

The proofs for these are shown below, but they will not be asked for in examination questions.

Proof that $\text{E}(x) = \lambda$

The probability distribution for $X \sim \text{Po}(\lambda)$ is:

x	0	1	2	3	…
$\text{P}(X = x)$	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2!}e^{-\lambda}$	$\frac{\lambda^3}{3!}e^{-\lambda}$	

Now $\text{E}(X) = \sum_{\text{all } x} x \cdot \text{P}(X = x)$

$$= 0 \cdot e^{-\lambda} + 1 \cdot \lambda e^{-\lambda} + 2 \cdot \frac{\lambda^2}{2!}e^{-\lambda} + 3 \cdot \frac{\lambda^3}{3!}e^{-\lambda} + \cdots$$

$$= \lambda e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} \cdots\right)$$

The series in the bracket has a sum of e^{λ} (the proof of this is beyond the scope of this curriculum).

Hence $\text{E}(X) = \lambda$.

Proof that $\text{Var}(X) = \lambda$

$\text{Var}(X) = \text{E}(X^2) - \text{E}^2(X)$

Now $\text{E}(X^2) = \sum_{\text{all } x} x^2 \cdot \text{P}(X = x)$

$$= 0 \cdot e^{-\lambda} + 1 \cdot \lambda e^{-\lambda} + 4 \cdot \frac{\lambda^2}{2!}e^{-\lambda} + 9 \cdot \frac{\lambda^3}{3!}e^{-\lambda} + 16 \cdot \frac{\lambda^4}{4!}e^{-\lambda} + \cdots$$

$$= \lambda e^{-\lambda}\left(1 + 2\lambda + \frac{3\lambda^2}{2!} + \frac{4\lambda^3}{3!} + \cdots\right)$$

We now split this into two series.

$$= \lambda e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots + \lambda + \frac{2\lambda^2}{2!} + \frac{3\lambda^3}{3!} + \cdots\right)$$

The first of these series is the same as in the proof for $\text{E}(X)$ and has a sum of e^{λ}.

$$= \lambda e^{-\lambda}\left\{e^{\lambda} \cdots + \lambda\left(1 + \lambda + \frac{\lambda^2}{2!} + \cdots\right)\right\}$$

$$= \lambda e^{\lambda}(e^{\lambda} + \lambda e^{\lambda})$$

$$= \lambda + \lambda^2$$

Hence $\text{Var}(X) = \text{E}(X^2) - \text{E}^2(X)$

$$= \lambda + \lambda^2 - \lambda^2$$

$$= \lambda$$

Example

In a given Poisson distribution it is found that $\text{P}(X \geq 1) = 0.25$. Find the variance of the distribution.

Let the distribution be $X \sim \text{Po}(m)$.

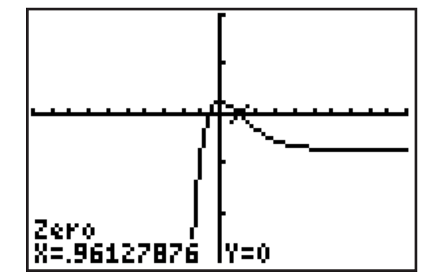
If $\text{P}(X \geq 1) = 0.25$ then

$1 - \text{P}(X = 0) - \text{P}(X = 1) = 0.25$

$\Rightarrow 1 - e^{-m} - me^{-m} = 0.25$

$\Rightarrow e^{-m} + me^{-m} - 0.75 = 0$

This can be solved on a calculator.



Since m is not negative, $m = 0.961$ and this is also $\text{Var}(X)$.

Example

In a fireworks factory, the number of defective fireworks follows a Poisson distribution with an average of three defective fireworks in any given box.

a Find the probability that there are exactly three defective fireworks in a given box.

b Find the most likely number of defective fireworks in a box.

c Find the probability that there are more than five defective fireworks in a box.

d Find the probability that in a sample of 15 boxes, at least three boxes have more than five defective fireworks in them.

Let X be the number of defective fireworks.

Hence $X \sim \text{Po}(3)$.

a $\text{P}(X = 3) = \frac{e^{-3}3^3}{3!} = 0.224$

b We select a range of values around the mean to find the most likely value. In this case we choose 1, 2, 3, 4, 5 and use a calculator.

Written as a table, the results are:

x	1	2	3	4	5
$\text{P}(X = x)$	0.149…	0.224…	0.224…	0.168…	0.100…

Since there are two identical probabilities in this case, the most likely value is either 2 or 3.

c $\text{P}(X > 5) = 1 - [\text{P}(X = 0) + \text{P}(X = 1) + \text{P}(X = 2) + \text{P}(X = 3)$
$+ \text{P}(X = 4) + \text{P}(X = 5)]$

$$= 1 - \left\{\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!} + \frac{e^{-3}3^5}{5!}\right\}$$

$$= 0.0839$$

```
poissoncdf(3,5)
        .9160820581
1-Ans
        .0839179419
```

d This is an example of where the question now becomes a binomial distribution Y, which is the number of boxes with more than five defective fireworks in them. Hence $Y \sim \text{Bin}(15, 0.0839)$.

$$P(Y \geq 3) = 1 - [P(Y = 0) + P(Y = 1) + P(Y = 2)]$$

$$= 1 - \left\{{}^{15}C_0(0.0839)^0(0.916)^{15} + {}^{15}C_1(0.0839)^1(0.916)^{14} + {}^{15}C_2(0.839)^2(0.916)^{13}\right\}$$

$$= 1 - 0.874\ldots = 0.126$$

```
binomcdf(15,0.08
39,2)
        .8742221686
1-Ans
        .1257778314
```

Exercise 4

1 If $X \sim \text{Po}(3)$, find:
a $P(X = 2)$ **b** $P(X \leq 2)$ **c** $P(X > 3)$ **d** $E(X)$

2 If $X \sim \text{Po}(6)$, find:
a $P(X = 4)$ **b** $P(X \leq 3)$ **c** $P(X > 5)$ **d** $E(X)$

3 If $X \sim \text{Po}(10)$, find:
a $P(X = 9)$ **b** $P(X \leq 8)$ **c** $P(X \geq 6)$ **d** $\text{Var}(X)$

4 If $X \sim \text{Po}(m)$ and $E(X^2) = 4.5$, find:
a m **b** $P(X = 4)$ **c** $P(X \leq 3)$

5 If $X \sim \text{Po}(\lambda)$ and $P(X \leq 2) = 0.55$, find:
a $E(X)$ **b** $P(X = 3)$ **c** $P(X < 4)$ **d** $P(X > 5)$

6 If $X \sim \text{Po}(m)$ and $P(X > 1) = 0.75$, find
a $\text{Var}(X)$
b $P(X = 5)$
c the probability that X is greater than 4
d the probability that X is at least 3
e the probability that X is less than or equal to 5.

7 If $X \sim \text{Po}(n)$ and $E(X^2) = 6.5$, find
a n
b the probability that X equals 4
c the probability that X is greater than 5
d the probability that X is at least 3.

8 A Poisson distribution is such that $X \sim \text{Po}(n)$.
a Given that $P(X = 5) = P(X = 3) + P(X = 4)$, find the value of n.
b Find the probability that X is at least 2.

9 On a given road, during a specific period in the morning, the number of drivers who break the speed limit, X, follows a Poisson distribution with mean m. It is calculated that $P(X = 1)$ is twice $P(X = 2)$. Find
a the value of m **b** $P(X \leq 3)$.

10 At a given road junction, the occurrence of an accident happening on a given day follows a Poisson distribution with mean 0.1. Find the probability of
a no accidents on a given day
b at least two accidents on a given day
c exactly three accidents on a given day.

11 Alexander is typing out a mathematics examination paper. On average he makes 3.6 mistakes per examination paper. His colleague, Roy, makes 3.2 mistakes per examination paper, on average. Given that the number of mistakes made by each author follows a Poisson distribution, calculate the probability that
a Alexander makes at least two mistakes
b Alexander makes exactly four mistakes
c Roy makes exactly three mistakes
d Alexander makes exactly four mistakes and Roy makes exactly three mistakes.

12 A machine produces carpets and occasionally minor faults are produced. The number of faults in a square metre of carpet follows a Poisson distribution with mean 2.7. Calculate
a the probability of there being exactly five faults in a square metre of carpet
b the probability of there being at least two faults in a square metre of carpet
c the most likely number of faults in a square metre of carpet
d the probability of less than six faults in 3 m^2 of carpet
e the probability of more than five faults in 2 m^2 of carpet.

13 At a local airport the number of planes that arrive between 10.00 and 12.00 in the morning is 6, on average. Given that these arrivals follow a Poisson distribution, find the probability that
a only one plane lands between 10.00 and 12.00 next Saturday morning
b either three or four planes will land next Monday between 10.00 and 12.00.

14 X is the number of Annie dolls sold by a shop per day. X has a Poisson distribution with mean 4.
a Find the probability that no Annie dolls are sold on a particular Monday.
b Find the probability that more than five are sold on a particular Saturday.
c Find the probability that more than 20 are sold in a particular week, assuming the shop is open seven days a week.

d If each Annie doll sells for 20 euros, find the mean and variance of the sales for a particular day.

Y is the number of Bobby dolls sold by the same shop per day. Y has a Poisson distribution with mean 6.

e Find the probability that the shop sells at least four Bobby dolls on a particular Tuesday.

f Find the probability that on a certain day, the shop sells three Annie dolls and four Bobby dolls.

15 A school office receives, on average, 15 calls every 10 minutes. Assuming this follows a Poisson distribution, find the probability that the office receives

a exactly nine calls in a 10-minute period

b at least seven calls in a 10-minute period

c exactly two calls in a 3-minute period

d more than four calls in a 5-minute period

e more than four calls in three consecutive 5-minute periods.

16 The misprints in the answers of a mathematics textbook are distributed following a Poisson distribution. If a book of 700 pages contains exactly 500 misprints, find

a **i** the probability that a particular page has exactly one misprint

ii the mean and variance of the number of misprints in a 30-page chapter

iii the most likely number of misprints in a 30-page chapter.

b If Chapters 12, 13 and 14 each have 40 pages, what is the probability that exactly one of them will have exactly 50 misprints?

17 A garage sells Super Run car tyres. The monthly demand for these tyres has a Poisson distribution with mean 4.

a Find the probability that they sell exactly three tyres in a given month.

b Find the probability that they sell no more than five tyres in a month.

A month consists of 22 days when the garage is open.

c What is the probability that exactly one tyre is bought on a given day?

d What is the probability that at least one tyre is bought on a given day?

e How many tyres should the garage have at the beginning of the month in order that the probability that they run out is less than 0.05?

18 Between 09.00 and 09.30 on a Sunday morning, 15 children and 35 adults enter the local zoo, on average. Find the probability that on a given Sunday between 09.00 and 09.30

a exactly ten children enter the zoo

b at least 30 adults enter the zoo

c exactly 14 children and 28 adults enter the zoo

d exactly 25 adults and 5 children enter the zoo.

Review exercise

All questions in this exercise will require a calculator.

1 The volumes (V) of four bottles of drink are 1 litre, 2 litres, 3 litres and 4 litres. The probability that a child selects a bottle of drink of volume V is cV.

a Find the value of c.

b Find E(X) where X is the volume of the selected drink.

c Find Var(X).

2 The random variable X follows a Poisson distribution. Given that $P(X \leq 1) = 0.2$, find:

a the mean of the distribution

b $P(X \leq 2)$ [IB Nov 06 P1 Q7]

3 The probability that a boy in a class has his birthday on a Monday or a Tuesday during a school year is $\frac{1}{4}$. There are 15 boys in the class.

a What is the probability that exactly three of them have birthdays on a Monday or a Tuesday?

b What is the most likely number of boys to have a birthday on a Monday or a Tuesday?

c In a particular year group, there are 70 boys. The probability of one of these boys having a birthday on a Monday or a Tuesday is also $\frac{1}{4}$. What is the expected number of boys having a birthday on a Monday or Tuesday?

4 In a game a player rolls a ball down a chute. The ball can land in one of six slots which are numbered 2, 4, 6, 8, 10 and x. The probability that it lands in a slot is the number of the slot divided by 50.

a If this is a random variable, calculate the value of x.

b Find E(X).

c Find Var(X).

5 The number of car accidents occurring per day on a highway follows a Poisson distribution with mean 1.5.

a Find the probability that more than two accidents will occur on a given Monday.

b Given that at least one accident occurs on another day, find the probability that more than two accidents occur on that day. [IB May 06 P1 Q16]

6 The most popular newspaper according to a recent survey is the Daily Enquirer, which claims that 65% of people read the newspaper on a certain bus route. Consider the people sitting in the first ten seats of a bus.

a What is the probability that exactly eight people will be reading the Daily Enquirer?

b What is the probability that more than four people will be reading the Daily Enquirer?

c What is the most likely number of people to be reading the Daily Enquirer?

d What is the expected number of people to be reading the Daily Enquirer?

e On a certain bus route, there are ten buses between the hours of 09.00 and 10.00. What is the probability that on exactly four of these buses at least six people in the first ten seats are reading the Daily Enquirer?

7 The discrete random variable X has the following probability distribution.

$$P(X = x) = \frac{k}{x},\ x = 1, 2, 3, 4$$
$$= 0 \text{ otherwise}$$

Calculate:

a the value of the constant k

b E(X) [IB May 04 P1 Q13]

8 An office worker, Alan, knows that the number of packages delivered in a day to his office follows a Poisson distribution with mean 5.

a On the first Monday in June, what is the probability that the courier company delivers four packages?

b On another day, Alan sees the courier van draw up to the office and hence knows that he will receive a delivery. What is the probability that he will receive three packages on that day?

9 The number of cats found in a particular locality follows a Poisson distribution with mean 4.1.

a Find the probability that the number of cats found will be exactly 5.

b What is the most likely number of cats to be found in the locality?

c A researcher checks half the area. What is the probability that he will find exactly two cats?

d Another area is found to have exactly the same Poisson distribution. What is the probability of finding four cats in the first area and more than three in the second?

10 The probability of the 16:55 train being delayed on a weekday is $\frac{1}{10}$.

Assume that delays occur independently.

a What is the probability, correct to three decimal places, that a traveller experiences 2 delays in a given 5-day week?

b How many delays must a commuter travel before having a 90% probability of having been delayed at least once? [IB Nov 90 P1 Q20]

11 Two children, Alan and Belle, each throw two fair cubical dice simultaneously. The score for each child is the sum of the two numbers shown on their respective dice.

a **i** Calculate the probability that Alan obtains a score of 9.

ii Calculate the probability that Alan and Belle both obtain a score of 9.

b **i** Calculate the probability that Alan and Belle obtain the same score.

ii Deduce the probability that Alan's score exceeds Belle's score.

c Let X denote the largest number shown on the four dice.

i Show that $P(X \le x) = \left(\frac{x}{6}\right)^4$ for $x = 1, 2, \dots 6$.

ii Copy and complete the following probability distribution table.

x	1	2	3	4	5	6
$\mathbf{P(X = x)}$	$\frac{1}{1296}$	$\frac{5}{1296}$				$\frac{671}{1296}$

iii Calculate $E(X)$. [IB May 02 P2 Q4]

12 The probability of finding the letter z on a page in a book is 0.05.

a In the first ten pages of a book, what is the probability that exactly three pages contain the letter z?

b In the first five pages of the book, what is the probability that at least two pages contain the letter z?

c What is the most likely number of pages to contain the letter z in a chapter of 20 pages?

d What would be the expected number of pages containing the letter z in a book of 200 pages?

e Given that the first page of a book does not contain the letter z, what is the probability that it occurs on more than two of the following five pages?

13 A biased die with four faces is used in a game. A player pays 10 counters to roll the die. The table below shows the possible scores on the die, the probability of each score and the number of counters the player receives for each score.

Score	1	2	3	4
Probability	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{10}$
Number of counters player receives	4	5	15	n

Find the value of n in order for the player to get an expected return of 9 counters per roll. [IB May 99 P1 Q17]

14 The number of accidents in factory A in a week follows a Poisson distribution X, where $\text{Var}(X) = 2.8$.

a Find the probability that there are exactly three accidents in a week.

b Find the probability that there is at least one accident in a week.

c Find the probability of more than 15 accidents in a four-week period.

d Find the probability that during the first two weeks of a year, the factory will have no accidents.

e In a neighbouring factory B, the probability of one accident in a week is the same as the probability of two accidents in a week in factory A. Assuming that this follows a Poisson distribution with mean n, find the value of n.

f What is the probability that in the first week of July, factory A has no accidents and factory B has one accident?

g Given that in the first week of September factory A has two accidents, what is the probability that in the same week factory B has no more than two accidents?

15 **a** Give the definition of the conditional probability that an event A occurs given that an event B (with $P(B) > 0$) is known to have occurred.

b If A_1 and A_2 are mutually exclusive events, express $P(A_1 \cup A_2)$ in terms of $P(A_1)$ and $P(A_2)$.

c State the multiplication rule for two independent events E_1 and E_2.

d Give the conditions that are required for a random variable to have a binomial distribution.

e A freight train is pulled by four locomotives. The probability that any locomotive works is θ and the working of a locomotive is independent of the other locomotives.

i Write down an expression for the probability that k of the four locomotives are working.

ii Write down the mean and variance of the number of locomotives working.

iii In order that the train may move, at least two of the locomotives must be working. Write down an expression, in terms of θ, for P, the probability that the train can move. (Simplification of this expression is not required.)

iv Calculate P for the cases when $\theta = 0.5$ and when $\theta = 0.9$.

v If the train is moving, obtain a general expression for the conditional probability that j locomotives are working. (Again, simplification of the expression is not required.) Verify that the sum of the possible conditional probabilities is unity.

vi Evaluate the above conditional probability when $j = 2$, for the cases when $\theta = 0.5$ and when $\theta = 0.9$.

vii For $\theta = 0.5$, calculate the probability that at least one of the three trains is able to move, assuming that they all have four locomotives and that different trains work independently. [IB May 94 P2 Q15]

16 In a game, a player pays 10 euros to flip six biased coins, which are twice as likely to show heads as tails. Depending on the number of heads he obtains, he receives a sum of money. This is shown in the table below:

Number of heads	0	1	2	3	4	5	6
Amount received in euros	30	25	15	12	18	25	40

a Calculate the probability distribution for this.

b Find the player's expected gain in one game.

c What is the variance?

d What would be his expected gain, to the nearest euro, in 15 games?

17 The table below shows the probability distribution for a random variable X. Find α and $E(X)$. [IB May 93 P1 Q20]

x	1	2	3	4
$P(X = x)$	2α	$4\alpha^2$	$2\alpha^2 + 3\alpha$	$\alpha^2 + \alpha$

18 An unbiased coin is tossed n times and X is the number of heads obtained. Write down an expression for the probability that $X = r$.
State the mean and standard deviation of X.
Two players, A and B, take part in the game. A has three coins and B has two coins. They each toss their coins and count the number of heads which they obtain.

a If A obtains more heads than B, she wins 5 cents from B. If B obtains more heads than A, she wins 10 cents from A. If they obtain an equal number of heads then B wins 1 cent from A. Show that, in a series of 100 such games, the expectation of A's winnings is approximately 31 cents.

b On another occasion they decide that the winner shall be the player obtaining the greater number of heads. If they obtain an equal number of heads, they toss the coins again, until a definite result is achieved. Calculate the probability that

i no result has been achieved after two tosses

ii A wins the game. [IB Nov 89 P2 Q8]