Exercise 10B

1 A discrete random variable R has the probability distribution given in this table.

r	1	5	10
$\mathrm{P}(R=r)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Find the value of

a $\mathrm{E}(R)$ **b** $\mathrm{E}(R^2)$ **c** $\mathrm{Var}(R)$ **d** standard deviation of R

2 A discrete random variable X can take only the values 0, 1, 2, 3, 4 and 5.

The probability distribution of X is given in this table.

x	0	1	2	3	4	5
$\mathrm{P}(X=x)$	a	a	a	b	b	b

Given that $\mathrm{P}(X \geq 2) = 3\mathrm{P}(X < 2)$

a find the values of a and b

b calculate $\mathrm{E}(X)$ and $\mathrm{E}(X^2)$

c find the value of $\mathrm{Var}(X)$.

3 A pack of 10 cards numbered from 1 to 6 is shuffled.

a Find the probability that the number on the bottom card is larger than the number on the top card. Justify your answer.

Let S be the random variable 'sum of the numbers on the top and bottom cards'.

b Find $\mathrm{P}(S = 4)$.

c Construct a table for the probability distribution of S.

d Find $\mathrm{E}(S)$ and $\mathrm{Var}(S)$.

4 Find the fair price to pay to enter a game in which you can win £20 with probability 0.2, win £10 with probability 0.4 or lose the amount you paid.

EXAM-STYLE QUESTION

5 A random variable T takes only integer values and has probability distribution function defined by:

$$f(t) = \mathrm{P}(T=t) = \begin{cases} kt^2 & \text{if } t = 1, 2, 3 \\ k(8-t)^2 & \text{if } t = 4, 5, 6, 7 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

a Find the value of k.

b Calculate $\mathrm{P}(T=4)$, $\mathrm{P}(T \leq 4)$ and $\mathrm{P}(T = 4 \mid T \leq 4)$.

c Find $\mathrm{E}(T)$ and $\mathrm{Var}(T)$.

d Determine the mode of T.

There are three equations here for $f(t)$. Choose the equation that corresponds to each value of t to work out $\Sigma f(t)$, then find k.

Example 6

Construct the cumulative distribution table for the larger number obtained when two dice are thrown and find the median of the distribution.

Answer

Let X be the random variable 'the larger number obtained when two dice are thrown'.

X can take the values 1, 2, 3, 4, 5 and 6.

$P(X=1)=\frac{1}{36}$ $\quad P(X=4)=\frac{7}{36}$

$P(X=2)=\frac{3}{36}$ $\quad P(X=5)=\frac{9}{36}$

$P(X=3)=\frac{5}{36}$ $\quad P(X=6)=\frac{11}{36}$

P(X = 6) is the probability of rolling either a double 6, or 6 on the first dice and a number between 1 and 5 on the second or 6 on the second dice and a number between 1 and 5 on the first.

x	1	2	3	4	5	6
$F(x)$	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{25}{36}$	1

See p 507 for more details on this.

Add together the values of the probabilities P(X = t) for all values of t less than or equal to x to obtain the CDF table.

Median of $X=\frac{4+5}{2}=4.5$

F(4) < 0.5 and F(5) > 0.5

Exercise 10C

1 A discrete random variable X has the probability distribution shown in the table.

x	5	10	15	20	25	30
$P(X=x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$

a Find $P(X \le 15)$.

b Construct the CDF table for this distribution.

c Hence find the median of the distribution.

2 An emergency call centre has five service lines that operate 24 hours every day.

A random variable L denotes the number of lines in use during any specific 5-minute period.

Data collected over a long period of time shows that L has this PDF:

n	0	1	2	3	4	5
$P(L=n)$	0.07	0.21	0.25	0.31	0.12	0.04

a What is the probability that at least three lines are in use simultaneously?

b Find $E(L)$ and $Var(L)$.

c Construct the CDF table for this distribution.

d Hence find the value of the median of the distribution.

3 A fair dice has faces numbered 1, 2, 2, 3, 3 and 3. The dice is thrown twice. S is the random variable 'sum of numbers on the uppermost face of the dice'

   a Construct tables for the probability distribution f and cumulative distribution F of S.
   b Find the mean, median and mode of S.
   c Calculate the standard deviation of S.

4 The probability distribution of a discrete random variable X is given by

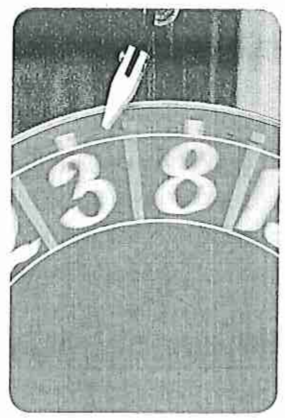
   $f(x) = kx$ where $x = 1, 2, \ldots, n$ and k is a parameter.

   a Show that $k = \frac{2}{n^2 + n}$ b Hence find $E(X)$

5 The probability distribution of a discrete random variable X is given by $f(x) = 3^{a-x}$ where $x = 1, 2, 3, \ldots$ and a is a parameter.

   a Show that $a = \log_3 2$
   b Hence find an expression for the cumulative distribution of S.

10.2 Binomial distribution



In a TV game show the final round consists of a Wheel of Fortune game. Each contestant that reaches this round has the chance to double the amount won in previous rounds. The Wheel of Fortune is a roulette wheel with 10 slots, one of which is gold. The contestant spins the wheel and wins if the ball lands in the gold slot. The producer of the show wants to know roughly how much will be paid in prize money so he needs a probability model that allows him to estimate the number of times the contestants will win in the last stage during the 50 planned sessions of the contest.

This Wheel of Fortune game is a typical example of a real-world situation that you can model mathematically.

→ The characteristic features of this type of problem are:

- There are a fixed number of trials n.
- The trials are independent and are done under exactly the same conditions.
- Each trial has exactly two possible outcomes: **success** and **failure**.
- For each trial the probability of success p is constant.
- The probability of failure is denoted by $q = 1 - p$ which is also constant.

Jacob Bernoulli (1654–1705), a member of the famous Swiss family of mathematicians, was the first person to study problems like this extensively.

Probability models make interesting topics for the Extented Essay.

Example 10

Let $X \sim B(4, 0.25)$

Find **a** The mode of X **b** The median of X.

Answers

x	0	1	2	3	4
$P(X = x)$	0.316	0.421	0.210	0.046...	0.003
$P(X \le x)$	0.316	0.738	0.949	0.996	1

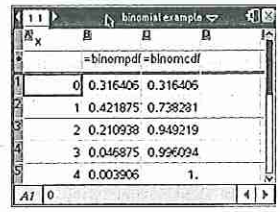
a mode of $X = 1$ as this is the maximum PDF value.

b $P(X \le 0) < 0.5$ and $P(X \le 1) > 0.5$

median of $X = \frac{0+1}{2} = 0.5$

PDF is maximum when $x = 1$

GDC help on CD: *Alternative demonstrations for the TI-84 Plus and Casio fx-9860GII GDCs are on the CD.*



Exercise 10D

1 A fair coin is tossed ten times. Find the probability of getting:

a exactly four heads **b** at least six heads

c not more than five heads.

2 A coin is biased so that the probability of obtaining head is 0.6. Find the probability of getting:

a exactly 2 heads if the coin is tossed 5 times

b at least 3 heads if the coin is tossed 7 times

c more heads than tails if the coin is tossed 9 times.

EXAM-STYLE QUESTIONS

3 In the mass production of light bulbs the probability that one bulb is defective is 1%. Bulbs are selected at random and put in packs of eight.

a If a pack is selected at random, what is the probability that it will contain:

i at least one defective bulb

ii not more than two defective bulbs?

b Given that a pack selected at random contains at least one defective bulb, what is the probability that it contains exactly two defective bulbs?

What assumption do you need to make to answer question 3?

4 Let $X \sim B(6, 0.35)$. Find

a The mode a of X

b The median b of X

c $P(X < 2a \mid X > b)$

5 Let $X \sim B(n, 0.4)$

a Construct three tables for the binomial CDF of X when $n = 2$, 5 and 10

b Given that $P(X \le 10) > 0.5$, find the largest possible value of n.

6 If $X \sim B(n, 0.3)$ and $P(X > 3) > 0.7$, find the least possible value of n.

b

x	$P(X=x)$	$x\,P(X=x)$	$x^2\,P(X=x)$
0	0.0041	0	0
1	0.0412	0.0412	0.0412
2	0.1646	03292	0.6284
3	03292	0.9877	2.9630
4	0.3292	1.3169	5.2675
5	0.1317	0.6584	3.2922
Total	1	3.3333	12.2222

Use your GDC to obtain the table.

$E(X) = 3.33$ (3 sf)

$E(X^2) = 12.22\ldots$

$\mathrm{Var}(X) = 12.22\ldots - (3.33)^2$

$= 1.11$ (3 sf)

$E(X)$ is the total of the third column.

$E(X^2)$ is the total of the fourth column.

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

As this question says 'hence' you must use the table to calculate the parameters of the distribution. Here the formulae for E(X) and Var(X) can only be used to check the answers.

You can check your answers using the formulae

$E(X) = np$ and $\mathrm{Var}(X) = npq$

$$E(X) = 5 \times \frac{2}{3} = \frac{10}{3} = 3.33\ (3\text{ sf})$$

$$\mathrm{Var}(X) = 5 \times \frac{2}{3} \times \frac{1}{3} = \frac{10}{9} = 1.11\ (3\text{ sf})$$

Exercise 10E

1 Given that $X \sim B(8, 0.4)$, find:

a $P(X = 5)$ **b** $P(X \le 5)$ **c** $P(X < 5)$

d the mean of X **e** the variance of X.

Only round at the final values. Early rounding can give you incorrect answers.

2 Given that $Y \sim B(7, 0.3)$, find:

a $P(Y = 1) + P(Y = 2)$ **b** $P(Y \le 2)$

c $P(Y \ge 2)$ **d** the median of Y.

3 Given that $T \sim B\left(5, \frac{1}{2}\right)$

a Show that $P(T = 5) = \frac{1}{32}$.

b Construct a table for the probability distribution function of T. Hence state the mode of T.

c Construct a table for the distribution function F of T.

d Write down the value of the median of T.

4 The probability that it rains on any given day in June of any given year in Drycity is 0.02.

a What is the probability that it rains on exactly three days in June in a given year?

b What is the probability that it does not rain on the first five days in June in a given year?

c Find the expected number of rainy days in Drycity in June.

EXAM-STYLE QUESTION

5 A random variable R follows a binomial distribution $B(n, p)$ with mean 2 and variance 1.5. Find the values of n and p.

EXAM-STYLE QUESTIONS

6 In a multiple choice test there are 20 questions. For each question there is a choice of four answers and only one of these is correct.

a If a student guesses each answer find the probability that he gets

i none correct **ii** more than ten correct

iii not more than five correct.

b Calculate the mean and standard deviation of the number of correct answers.

c Suppose that five students guess the answers to the test. What is the probability that at least two of them get more than ten answers correct?

7 In a large city 18 % of the people are left-handed. If a random sample of ten people from this city is selected:

a find the probability that exactly two of them will be left-handed

b find the probability that at least one person in the sample is left-handed

c find the most likely number of left-handed people in the sample. If another sample of 25 people from the city is selected,

d find the expected number of left-handed people in the sample.

e find the variance of the number of left-handed people in this sample. If a sample of size n is to be selected randomly,

f Find the minimum value of n for which the probability that it contains at least two left-handed people will be greater than 95%.

g In the same city the percentage of left-handed women is 16% and the percentage of left-handed men is 22%. A random sample of five women and five men is selected from the population of the city. What is the probability that the sample contains at least one left-handed woman and one left-handed man?

8 On a TV news channel, the evening news starts at the same time every day. The probability that Mr Li gets home from work in time to watch the news is 0.3

a Calculate the probability that, in a particular week of five working days, he gets home in time to watch the news:

i on exactly four days **ii** on at least three days.

b What is the probability that Mr Li gets home in time to watch the news on three consecutive days?

9 Let $X \sim B(9, p)$

a Draw bar graphs to represent the probability distribution of X when $p = 0.1, 0.3, 0.5, 0.7$ and 0.9.

b Compare the graphs obtained in part **a** and comment on their symmetries.

c For each value of p in part **a**, find the mean, median and mode of the distribution and comment on their values in relation to the symmetries (or asymmetries) of the graph.

10 Consider two independent random variables X and Y such that $X \sim \mathrm{B}(2, a)$ and $Y \sim \mathrm{B}(2, b)$.
Let W be the random variable that represents the product of each value of X with each value of Y.

a Construct a table showing the probability distribution of W.

b Hence find an expression for $\mathrm{E}(W)$ in terms of a and b.

Extension material on CD: *Worksheet 10*

10.3 Poisson distribution

The Poisson distribution is named after **Siméon-Denis Poisson** (1781–1840) who first discovered this model as an approximation to the binomial distribution when the number of trials n gets larger and larger while the probability of success p gets smaller and smaller.

Poisson was an extremely hard-working mathematician. His major work on probability was a book with over 400 pages where just one was dedicated to the derivation of the Poisson distribution.

→ A discrete random variable X follows a Poisson distribution when it models situations that satisfy these conditions:

- The occurrence of an event at a particular point in space or in time is **independent** of what happens elsewhere.
- The probability of an event occurring within a **small fixed interval** (or in a small region of space) is constant.
- There is no chance that two events will occur at precisely the same moment or at the same place.

In theory, the Poisson distribution has no upper value, and in this way it differs from the binomial distribution.

→ If an event is randomly scattered in time or space, the discrete random variable X that models the number of its occurrences in a given interval follows a **Poisson distribution** with parameter m, $X \sim \mathrm{Po}(m)$

The probability distribution function (PDF) of X is

$$f(x) = \mathrm{P}(X = x) = \frac{e^{-m} m^x}{x!} \text{ where } x = 0, 1, 2, \ldots.$$

The Poisson cumulative distribution function (CDF) is

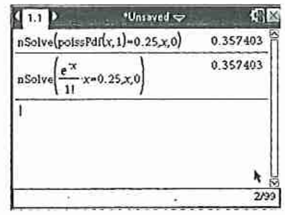
$$F(x) = \mathrm{P}(X \le x) = \sum_{t=0}^{x} \frac{e^{-m} m^t}{t!}$$

Example 16

If $X \sim \text{Po}(\lambda)$ find the value of λ, correct to 4 decimal places, given that $\text{P}(X = 1) = 0.25$

Answer	
$\text{P}(X = 1) = 0.25$	$\text{P}(X = x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$
$\dfrac{e^{-1}\lambda^1}{1!} = 0.25$	
$l = 0.3574$ (4 dp)	*Use GDC numerical solver.*

GDC help on CD: *Alternative demonstrations for the TI-84 Plus and Casio fx-9860GII GDCs are on the CD.*



Exercise 10G

1 The number of accidents in a randomly chosen week at Safe School can be modeled by a Poisson distribution with parameter 0.7. Find the probability that, in a randomly chosen week, there are

a exactly two accidents

b at least two accidents.

2 The number of bacteria per millilitre of a certain liquid follows a Poisson distribution with parameter 3. Find the probability that in a millilitre of the liquid there will be

a at least 4 bacteria

b not more than 2 bacteria.

3 $X \sim \text{Po}(m)$. Find the value of m if $\text{P}(X = 1) = 0.1$

4 $X \sim \text{Po}(m)$. Find the value of m if $\text{P}(X \le 1) = 0.5$

Parameters of the Poisson distribution

The variable X in a Poisson distribution is a discrete variable, and therefore you can use the formulae

$$\text{E}(X) = \sum_{x=0}^{\infty} x\,\text{P}(X = x) \text{ and } \text{Var}(X) = \text{E}(X^2) - (\text{E}(X))^2$$

to calculate the mean and variance. The variable X can take an infinite number of values: 0, 1, 2, 3, 4,... so instead of a finite sum for E(X) and Var(X) you get infinite series. As the study of infinite series is beyond the scope of this course, you cannot directly deduce formulae for these parameters.

This example shows you how to estimate the values of the mean and variance of a Poisson variable using your GDC.

The exact calculation of the expected value and variance of Poisson variables requires knowledge of techniques that are part of the *Calculus* option.

Example 19

On Sunday mornings, cars arrive at a petrol station at an average rate of 30 per hour. Assuming that the number of cars arriving at the petrol station follows a Poisson distribution, find the probability that:

a in a half-hour period 12 cars arrive

b no cars arrive during a particular 5-minute interval

c more than 5 cars arrive in a 15-minute interval.

Answers

a In a half-hour period, the number of cars arriving at the petrol station can be modeled by a Poisson random variable $Y \sim \text{Po}(15)$
$\text{P}(Y = 12) = 0.0829$ (3 sf)

Mean is $\frac{1}{2} \times 30 = 15$

Use $\text{P}(Y = 12) = \frac{e^{-15} \times (15)^{12}}{12!}$

b In a 5-minute interval, the number of cars arriving is modeled by $Y \sim \text{Po}(2.5)$
$\text{P}(Y = 0) = 0.0821$ (3 sf)

The mean is $5 \times \frac{30}{60} = 2.5$

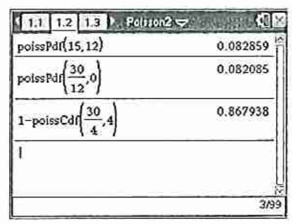
Use $\text{P}(Y = 0) = \frac{e^{-2.5} \times (2.5)^{0}}{0!}$

c In a 15-minute interval, the number of cars arriving is modeled by $Y \sim \text{Po}(7.5)$
$\text{P}(Y > 5) = 1 - \text{P}(Y \le 4)$
$= 0.868$ (3 sf)

Use $\text{P}(Y \le 4)$
$= \text{P}(Y \le 0) + \text{P}(Y = 1) + \text{P}(Y = 2) + \text{P}(Y = 3) + \text{P}(Y = 4)$

Use built-in functions on your GDC to calculate these probabilities.

GDC help on CD: *Alternative demonstrations for the TI-84 Plus and Casio fx-9860GII GDCs are on the CD.*



Exercise 10H

1 The mean number of bacteria per millilitre of a given liquid is 3.5 Find the probability that

a in 2 ml of the liquid there will be fewer than 7 bacteria

b in 0.5 ml of the liquid there will be at least 2 bacteria.

2 The mean number of flaws per square metre of fabric produced on a machine is 0.01. If flaws occur randomly and their number is modeled by a Poisson variable find the probability that

a in a randomly chosen 100 square metres of fabric there will be exactly two flaws

b in 25 square metres of chosen randomly fabric there will be at least one flaw.

3 The number of phone calls received by a school between 8:00 and 9:00 on any weekday is modeled by a Poisson distribution. If the mean number of phone calls per hour is 12, calculate

a the expected number of phone calls received between 8:00 and 8:15 in a given day

b the probability that more than five calls are received between 8:00 and 8:10 in a given day.

4 $X \sim \text{Po}(3.5)$

a Calculate

i $P(X = 3)$ **ii** $P(X > 3)$ **iii** $P(X < 5 \mid X > 3)$

b Write down the values of $E(X)$ and $\text{Var}(X)$

c Hence find the value of $E(X^2)$.

EXAM-STYLE QUESTION

5 The random variable X is Poisson distributed with mean m and satisfies

$$P(X = 0) + P(X = 1) - P(X = 4) = 0$$

a Find the value of m correct to four decimal places

b Hence calculate $P(2 \le X \le 4)$.

6 Let X be a random variable with a Poisson distribution, such that $P(X > 3) = 0.555$

Find $P(X < 3)$

EXAM-STYLE QUESTION

7 The random variable P has a Poisson distribution with mean $\lambda > 0$.

Let p be the probability that P takes the value 0, 1 or 2.

a Write down an expression for p in terms of λ.

b Show that $p = p(\lambda)$ is a decreasing function.

c Sketch the graph of p for $0 < \lambda \le 6$, showing clearly concavities and any points of inflexion.

10.4 Continuous random variables

In the previous section you studied **discrete random variables** whose density functions are suitable models for situations where the outcomes have distinct values, usually integers. In many other situations, when we you need to model the behavior of variables like height, weight, mass and time, you use **continuous random variables**.

→ A variable X is **random** when its value is the result of a random experiment and it is **continuous** when it is not possible to list all of its values but only the range of values it can take.

The values of a continuous random variable form an uncountable set. What is the meaning of uncountable? Is the set of rational numbers uncountable? What is the difference between countable and uncountable?