

## Discrete Random Variables

- 1 The discrete random variable  $X$  has the probability distribution shown in the table.

| $x$        | 1   | 2   | 3   | 4   |
|------------|-----|-----|-----|-----|
| $P(X = x)$ | 0.1 | 0.3 | $p$ | 0.2 |

- (a) Find the value of  $p$ . [1]  
 (b) Find  $P(X > 2)$ . [1]

- 2 The random variable  $X$  takes the values 0, 1, 2 and 3, and  $P(X = x) = k(x + 1)$  for these values, where  $k$  is a constant.  
 Find the value of  $k$  and hence find  $P(X < 2)$ . [3]

- 3 The table is claimed to show the probability distribution of a discrete random variable  $X$ .

| $x$        | 1   | 2   | 3   | 4   |
|------------|-----|-----|-----|-----|
| $P(X = x)$ | 0.3 | 0.4 | 0.2 | 0.3 |

- (a) Show that the table cannot be a probability distribution. [2]  
 (b) The first three probabilities are correct. Find the correct value of  $P(X = 4)$ . [1]

- 4 The discrete random variable  $X$  has the probability distribution shown.

| $x$        | 0     | 1     | 2     | 3     |
|------------|-------|-------|-------|-------|
| $P(X = x)$ | $1/8$ | $3/8$ | $1/4$ | $1/4$ |

- (a) Find  $P(1 \leq X < 3)$ . [2]  
 (b) Find the value of the cumulative distribution function  $F(x)$  at  $x = 1.5$ . [2]

- 5 A fair coin is tossed three times. The random variable  $X$  is the number of heads.  
 Write down the probability distribution of  $X$ . [2]

- 6 A card is chosen at random from ten cards numbered 1 to 10. The random variable  $X$  is the number on the card.

Find (a)  $P(X = 7)$ , (b)  $P(X \text{ is a multiple of } 3)$ , (c)  $P(X^2 > 40)$ . [5]

- 7 Two fair four-sided dice, each with faces numbered 1, 2, 3 and 4, are rolled. The random variable  $X$  is the positive difference between the two scores.

- (a) Draw up the probability distribution of  $X$ . [3]  
 (b) Find  $P(X \geq 2)$ . [1]

- 8 The discrete random variable  $X$  takes the values 1, 2, 3, 4 and 5. Its cumulative distribution function is shown in the table.

| $x$           | 1   | 2   | 3    | 4   | 5 |
|---------------|-----|-----|------|-----|---|
| $P(X \leq x)$ | 0.1 | 0.3 | 0.55 | 0.8 | 1 |

- (a) Find  $P(X = 3)$ . [1]  
 (b) Find  $P(X > 2)$ . [1]  
 (c) Write down the probability distribution of  $X$  as a table of  $P(X = x)$ . [2]

- 9 A fair twelve-sided die has faces numbered 1 to 12. The random variable  $X$  is the score when it is rolled.

- (a) Find  $P(3 < X \leq 8)$ . [2]  
 (b) Find  $P(X^2 > 50 \mid X \text{ is even})$ . [3]

- 10** The cumulative distribution function of the discrete random variable  $X$  is shown, where  $c$  is a constant. Given that  $P(X = 3) = 0.35$ , find the value of  $c$  and the probability distribution of  $X$ .

| $x$           | 1    | 2   | 3   | 4 |
|---------------|------|-----|-----|---|
| $P(X \leq x)$ | 0.15 | 0.4 | $c$ | 1 |

- (a) Find  $c$ . [2]  
 (b) Write down  $P(X = x)$  for  $x = 1, 2, 3, 4$ . [2]

- 11** Four cards are numbered 1, 2, 3 and 4. Two cards are drawn at random without replacement. The random variable  $X$  is the larger of the two numbers drawn. Find the probability distribution of  $X$ . [3]

- 12** The discrete random variable  $X$  has  $P(X = x) = k(6 - x)$  for  $x = 1, 2, 3, 4, 5$ , where  $k$  is a constant. Find  $k$ , the mode of  $X$  and  $P(X \text{ is even})$ . [4]

- 13** The random variable  $X$  has the distribution shown. Two independent observations of  $X$  are made. Find the probability that their sum is 4.

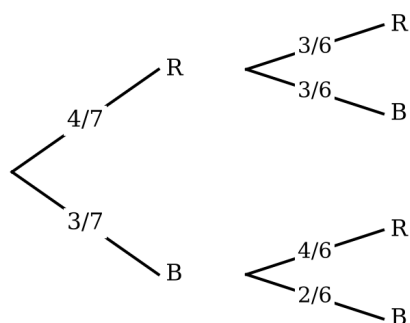
| $x$        | 1   | 2   | 3   |
|------------|-----|-----|-----|
| $P(X = x)$ | 0.5 | 0.3 | 0.2 |

[3]

- 14** The discrete random variable  $X$  has probability distribution  $P(X = x) = kx(6 - x)$  for  $x = 1, 2, 3, 4$ , where  $k$  is a constant.

- (a) Show that  $k = \frac{1}{30}$ . [2]  
 (b) Find  $P(X \geq 2 \mid X \leq 3)$ . [2]  
 (c) Two independent observations of  $X$  are made. Find the probability that their sum is 5. [3]

- 15** A bag contains 4 red balls and 3 blue balls. Two balls are taken at random from the bag, one after the other, without replacement. The random variable  $X$  is the number of red balls taken. The tree diagram shows the possible outcomes ( $R$  = red,  $B$  = blue).



- (a) Find the probability distribution of  $X$ . [4]  
 (b) Given that at least one red ball is taken, find the probability that both balls are red. [2]

- 16** A fair die is rolled until a six is obtained, but no more than three rolls are made. The random variable  $X$  is the number of rolls made.

- (a) Show that  $P(X = 2) = \frac{5}{36}$  and find  $P(X = 3)$ . [4]  
 (b) Find  $P(X \geq 2)$ . [2]

- 17** The discrete random variable  $X$  has the probability distribution shown, where  $a$  is a constant.

| $x$        | <b>1</b>   | <b>2</b> | <b>3</b>   |
|------------|------------|----------|------------|
| $P(X = x)$ | $0.35 - a$ | $2a$     | $0.65 - a$ |

- (a) State the range of possible values of  $a$ . [2]  
 (b) Two independent observations  $X_1$  and  $X_2$  of  $X$  are made. Show that  $P(X_1 + X_2 = 4) = 6a^2 - 2a + 0.455$ . [4]

- 18** The discrete random variable  $X$  has the distribution shown. Given that  $E(X) = 2.4$ , find the values of  $a$  and  $b$ .

| $x$        | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> |
|------------|----------|----------|----------|----------|
| $P(X = x)$ | 0.3      | $a$      | $b$      | 0.2      |

[5]

- 19** The discrete random variable  $X$  takes the values 1, 2, 3, 4 and its cumulative distribution function is  $F(x) = \frac{x^2 + x}{20}$ .

- (a) Show that  $F(4) = 1$  and find  $P(X = 3)$ . [3]  
 (b) Find  $P(2 < X \leq 4)$ . [2]

- 20** Two students propose models for the score  $X$  on a spinner with outcomes 1, 2, 3, 4.

- (a) Model A:  $P(X = x) = \frac{x}{6}$ . Explain why Model A is not a valid probability distribution. [2]  
 (b) Model B:  $P(X = x) = \frac{x}{10}$ . Show that Model B is valid and find  $P(X \text{ is prime})$ . [3]

- 21** The random variable  $X$  takes the values 1, 2, 3, ... and  $P(X = x) = k(\frac{1}{3})^x$ , where  $k$  is a constant.

- (a) Show that  $k = 2$ . [3]  
 (b) Find the probability that  $X$  is even. [3]  
 (c) Find the least integer  $n$  such that  $P(X > n) < 0.001$ . [3]

- 22** The discrete random variable  $X$  has the probability distribution shown in the table, where  $k$  is a constant.

| $x$        | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> |
|------------|----------|----------|----------|----------|
| $P(X = x)$ | 0.2      | $k$      | $k^2$    | 0.24     |

- (a) Show that  $k^2 + k - 0.56 = 0$ , and hence find the value of  $k$ , explaining why the other root is rejected. [4]  
 (b) Find  $P(X > 2 \mid X \geq 2)$ . [3]  
 (c) Two independent observations of  $X$  are made. Find the probability that they are equal. [2]

- 23** The discrete random variable  $X$  has  $P(X = x) = \frac{k}{x(x+1)}$  for  $x = 1, 2, 3, \dots, n$ , where  $k$  is a constant.

- (a) Express  $\frac{1}{x(x+1)}$  in partial fractions. [2]  
 (b) Hence show that  $k = \frac{n+1}{n}$ . [4]  
 (c) For  $n = 4$ , find  $P(X \geq 3)$ . [2]

- 24** The discrete random variable  $X$  has the probability distribution shown. Two independent observations  $X_1$  and  $X_2$  of  $X$  are made and  $Y = X_1 + X_2$ .

| $x$        | <b>0</b> | <b>2</b> | <b>3</b> |
|------------|----------|----------|----------|
| $P(X = x)$ | 0.4      | 0.4      | 0.2      |

- (a) Find the probability distribution of  $Y$ . [4]  
 (b) Find  $P(X_1 > X_2)$ . [2]  
 (c) Find  $P(X_1 > X_2 \mid Y \geq 3)$ . [3]

- 25** The discrete random variable  $X$  has the distribution shown. Two independent observations of  $X$  are made, and the probability that they are equal is 0.38.

| $x$        | <b>0</b> | <b>1</b> | <b>2</b>  |
|------------|----------|----------|-----------|
| $P(X = x)$ | $a$      | 0.5      | $0.5 - a$ |

- (a) Show that  $2a^2 - a + 0.12 = 0$ . [4]  
 (b) Find the possible values of  $a$  and, for the larger value, find  $E(X)$ . [4]

- 26** A fair die is rolled twice. The random variable  $X$  is defined as follows:  $X = 0$  if the two scores are the same, otherwise  $X$  is the score on the first roll.

- (a) Show that  $P(X = 0) = \frac{1}{6}$  and  $P(X = 4) = \frac{5}{36}$ . [4]  
 (b) Find  $P(X \text{ is even})$ . [3]  
 (c) Find  $P(X > 3 \mid X \neq 0)$ . [2]

- 27** The discrete random variable  $X$  has probability distribution  $P(X = x) = kx^2$  for  $x = 1, 2, \dots, n$ , where  $n$  is a positive integer and  $k$  is a constant.

- (a) Using the formula for  $\sum r^2$ , show that  $k = \frac{6}{n(n+1)(2n+1)}$ . [3]  
 (b) Given that  $P(X = n) = \frac{8}{15}$ , find the value of  $n$ . [5]  
 (c) For this value of  $n$ , find the probability that  $X$  is odd. [2]

- 28** Two fair  $n$ -sided dice, with faces numbered 1 to  $n$ , are rolled. The random variable  $X$  is the larger of the two scores ( $X = x$  if both dice show  $x$ ).

- (a) Show that  $P(X \leq x) = \frac{x^2}{n^2}$  for  $x = 1, 2, \dots, n$ , and hence that  $P(X = x) = \frac{2x-1}{n^2}$ . [4]  
 (b) For  $n = 4$ , two independent observations of  $X$  are made. Find the probability that they sum to 7. [3]  
 (c) Find the least value of  $n$  for which  $P(X = n) < 0.1$ . [5]

- 29** The discrete random variable  $X$  has the probability distribution shown. Three independent observations  $X_1$ ,  $X_2$  and  $X_3$  of  $X$  are made.

| $x$        | <b>1</b> | <b>2</b> | <b>4</b> |
|------------|----------|----------|----------|
| $P(X = x)$ | 0.5      | 0.3      | 0.2      |

- (a) Find the probability that all three observations are different. [3]  
 (b) Find the probability that  $X_1 + X_2 + X_3 = 6$ . [4]  
 (c) Find the probability that  $X_1 + X_2 + X_3 \geq 9$ , and, given that the total is at least 9, the probability that exactly two of the observations equal 4. [7]

**30**

The discrete random variable  $X$  takes the values 1, 2, 3, ... with  $P(X = x) = \left(\frac{1}{2}\right)^x$ . Two independent observations  $X_1$  and  $X_2$  of  $X$  are made and  $M$  is the smaller of the two values.

- (a) Show that  $P(X > x) = \left(\frac{1}{2}\right)^x$ . [3]
- (b) Show that  $P(M \geq m) = \left(\frac{1}{4}\right)^{m-1}$  for  $m \geq 1$ . [3]
- (c) Hence find  $P(M = m)$ , and show that  $M$  has a geometric distribution. [4]
- (d) Find  $P(M = 2)$  and  $P(M \text{ is even})$ . [3]

## Expectation and Variance

- 1** The discrete random variable  $X$  has the distribution shown. Find  $E(X)$ .

| $x$        | 1   | 2   | 3   | 4   |
|------------|-----|-----|-----|-----|
| $P(X = x)$ | 0.1 | 0.2 | 0.3 | 0.4 |

[2]

- 2** A spinner scores 0, 1 or 2 with probabilities 0.5, 0.3 and 0.2 respectively. The score is  $X$ . Find  $E(X)$  and  $\text{Var}(X)$ . [3]

- 3** A fair six-sided die is rolled and  $X$  is the score. Find  $E(X)$  and  $E(X^2)$ . [4]

- 4** The random variable  $X$  has  $E(X) = 3$  and  $\text{Var}(X) = 7$ . Find  $E(X^2)$ . [2]

- 5** Find the mean and variance of  $X$ , whose distribution is shown.

| $x$        | -1  | 0   | 1   |
|------------|-----|-----|-----|
| $P(X = x)$ | 0.2 | 0.5 | 0.3 |

[3]

- 6** In a game a fair coin is tossed. A player wins £4 for a head and loses £1 for a tail. Find the expected winnings. [2]

- 7** The random variable  $X$  has  $E(X) = 4$  and  $\text{Var}(X) = 9$ .
- (a) Find  $E(3X - 2)$ . [1]
- (b) Find  $\text{Var}(3X - 2)$ . [2]
- (c) Find  $E(X^2)$ . [2]

- 8** In a game, a fair six-sided die is rolled once. If a 6 is rolled the player wins £5. If a 4 or 5 is rolled the player wins £1. Otherwise the player loses £2. The player's winnings, in pounds, are  $W$  (a loss counts as negative).
- (a) Find the probability distribution of  $W$  and hence find  $E(W)$ . [3]
- (b) The game is played 120 times. Find the expected total winnings and say whether the game is fair. [2]

- 9** A raffle sells 1000 tickets at £2 each. There is 1 prize of £300, 5 prizes of £50 and 20 prizes of £5.
- (a) Find the expected value of the prize won by one ticket. [2]
- (b) Find the expected profit per ticket to the organiser. [2]

- 10** The random variable  $X$  has the distribution shown. Find  $E(X)$ ,  $\text{Var}(X)$  and the standard deviation of  $X$ .

| $x$        | 1   | 2   | 3   |
|------------|-----|-----|-----|
| $P(X = x)$ | 0.3 | 0.5 | 0.2 |

[5]

- 11** The random variable  $X$  has the distribution shown. Find  $E(X^2 - 3)$  and  $E(\frac{1}{X+1})$ .

| $x$        | 0   | 1   | 2   | 3   |
|------------|-----|-----|-----|-----|
| $P(X = x)$ | 0.2 | 0.3 | 0.4 | 0.1 |

[4]

- 12** The cumulative distribution function of the discrete random variable  $X$  is shown in the table.

| $x$           | 1   | 2   | 3   | 4   | 5 |
|---------------|-----|-----|-----|-----|---|
| $P(X \leq x)$ | 0.2 | 0.5 | 0.7 | 0.9 | 1 |

(a) Find  $E(X)$  and  $\text{Var}(X)$ .

[5]

(b) The random variable  $Y = 3X - 2$ . Find  $E(Y)$  and  $\text{Var}(Y)$ .

[3]

- 13** A box contains 4 red discs and 2 white discs. Three discs are chosen at random without replacement. The random variable  $X$  is the number of red discs chosen.

(a) Show that  $P(X = 1) = \frac{1}{5}$  and find the rest of the probability distribution of  $X$ .

[4]

(b) Find  $E(X)$  and  $\text{Var}(X)$ .

[3]

- 14** The discrete random variable  $X$  has the probability distribution shown. Its mean is  $\mu$  and its standard deviation is  $\sigma$ .

| $x$        | 1   | 2    | 3   | 4    | 5   |
|------------|-----|------|-----|------|-----|
| $P(X = x)$ | 0.1 | 0.25 | 0.3 | 0.25 | 0.1 |

(a) Find  $\mu$  and  $\sigma$ .

[4]

(b) Find  $P(\mu - \sigma \leq X \leq \mu + \sigma)$ .

[2]

(c) Find  $E(2X^2 + 1)$ .

[2]

- 15** The discrete random variable  $X$  has probability distribution  $P(X = x) = kx$  for  $x = 1, 2, 3, 4$ .

(a) Show that  $k = 0.1$ .

[1]

(b) Find  $E(X)$ .

[2]

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

[2]

(d) Find  $E(\frac{1}{X})$ .

[2]

- 16** An insurance company models the claim  $C$  (in pounds) made on a policy in a year as follows:  $P(C = 0) = 0.90$ ,  $P(C = 500) = 0.08$  and  $P(C = 2000) = 0.02$ .

(a) Find the expected claim and the standard deviation of the claim.

[5]

(b) The company charges a premium of  $\pounds p$  per policy. Find the value of  $p$  that gives an expected profit of  $\pounds 20$  per policy, and state the variance of the profit.

[3]

- 17** Two fair dice are rolled and  $X$  is the larger of the two scores. The probability distribution of  $X$  is shown.

| $x$        | 1    | 2    | 3    | 4    | 5    | 6     |
|------------|------|------|------|------|------|-------|
| $P(X = x)$ | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 |

- (a) Show that  $E(X) = \frac{161}{36}$ . [3]  
 (b) Find  $\text{Var}(X)$  to 3 significant figures. [4]

- 18** The discrete random variable  $X$  has  $P(X = x) = \frac{k}{x}$  for  $x = 1, 2, 3, 4$ . Find the value of  $k$ ,  $E(X)$  and  $\text{Var}(X)$ . [6]

- 19** The random variable  $X$  has mean 10 and variance 16. The random variable  $Y = aX + b$ , where  $a > 0$ , has mean 0 and variance 1. Find the values of  $a$  and  $b$ , and verify that  $E(Y^2) = 1$ . [5]

- 20** A player chooses one of two games. In Game A the winnings  $\pounds X$  are 3 or 5 with equal probability. In Game B the winnings  $\pounds Y$  are 0, 4 or 8 with probabilities 0.25, 0.5 and 0.25. Show that the games have the same expected winnings, and state, with a reason, which game suits a player who wants the more predictable outcome. [5]

- 21** The discrete random variable  $X$  has the distribution shown. A student calculates  $E(X) = 4.2$  and  $E(X^2) = 19.6$ , then writes:  
 "Var( $X$ ) =  $[E(X)]^2 - E(X^2) = 17.64 - 19.6 = -1.96$ , so the variance is negative."

| $x$        | 2   | 4   | 6   |
|------------|-----|-----|-----|
| $P(X = x)$ | 0.2 | 0.5 | 0.3 |

- (a) Explain the error in the student's working and say why the answer cannot be correct. [2]  
 (b) Find the correct value of  $\text{Var}(X)$ . [2]  
 (c) The random variable  $Y = 5 - 2X$ . Find  $E(Y)$  and the standard deviation of  $Y$ . [4]

- 22** A bag contains 2 red balls and 3 green balls. Two balls are drawn at random without replacement. The random variable  $X$  is the number of red balls drawn.

In a game, a player pays an entry fee of  $\pounds c$ , draws two balls and is paid  $\pounds 3$  for each red ball drawn.

- (a) Show that  $P(X = 0) = 0.3$ ,  $P(X = 1) = 0.6$  and  $P(X = 2) = 0.1$ . [3]  
 (b) Find the value of  $c$  for which the game is fair. [3]  
 (c) With the entry fee set at the fair value, find the variance of the player's profit. [3]

- 23**  $X$  is a discrete random variable with mean  $\mu$  and  $Y = aX + b$ , where  $a$  and  $b$  are constants.

- (a) Write down  $E(Y)$  in terms of  $a$ ,  $b$  and  $\mu$ . [1]  
 (b) Using  $\text{Var}(Y) = E[(Y - E(Y))^2] = \sum (y - E(Y))^2 P(Y = y)$ , prove that  $\text{Var}(Y) = a^2 \text{Var}(X)$ . [5]  
 (c) Hence explain why the standard deviation of  $Y$  is  $|a|$  times that of  $X$ , and find the standard deviation of  $5 - 3X$  when  $\text{Var}(X) = 4$ . [3]

- 24** A fair coin is tossed until a head appears, with a maximum of four tosses.  $X$  is the number of tosses.

- (a) Show that  $P(X = 3) = \frac{1}{8}$  and  $P(X = 4) = \frac{1}{8}$ . [4]  
 (b) Find  $E(X)$  and  $\text{Var}(X)$ . [4]  
 (c) A player pays  $\pounds 1$  to play and receives  $\pounds (2X - 1)$  pounds back. Find the expected profit. [3]

- 25** The random variable  $X$  has the distribution shown.  $X_1$  and  $X_2$  are independent observations of  $X$ .

| $x$        | <b>1</b> | <b>2</b> | <b>4</b> |
|------------|----------|----------|----------|
| $P(X = x)$ | 0.25     | 0.5      | 0.25     |

- (a) Find  $E(X)$  and  $\text{Var}(X)$ . [4]  
 (b) Find  $E(X_1 X_2)$  and  $E(|X_1 - X_2|)$ . [4]  
 (c) Find  $\text{Var}(X_1 + X_2)$  and  $\text{Var}(X_1 - X_2)$ , and comment. [4]

- 26** The random variable  $X$  takes the values 1, 2 and 3 with  $P(X = 1) = a$ ,  $P(X = 2) = b$  and  $P(X = 3) = c$ . Given that  $E(X) = 2.1$  and  $\text{Var}(X) = 0.49$ , find  $a$ ,  $b$  and  $c$ .

| $x$        | <b>1</b> | <b>2</b> | <b>3</b> |
|------------|----------|----------|----------|
| $P(X = x)$ | $a$      | $b$      | $c$      |

[7]

- 27** The random variable  $X$  takes each of the values 1, 2, 3, ...,  $n$  with equal probability  $\frac{1}{n}$ .

- (a) Show that  $E(X) = \frac{n+1}{2}$ . [2]  
 (b) Show that  $\text{Var}(X) = \frac{n^2 - 1}{12}$ . [4]  
 (c) Given that  $\text{Var}(X) = 8.25$ , find  $n$ . [2]  
 (d) For this value of  $n$ , find the mean and variance of  $Y = 2X - 7$ . [2]

- 28** The discrete random variable  $X$  has probability distribution

$P(X = x) = \frac{x}{24}$  for  $x = 1, 2, 3, 4$ ,  $P(X = 6) = \frac{7}{12}$ , and  $P(X = x) = 0$  otherwise.

- (a) Verify that this is a valid probability distribution and find  $P(X \leq 4)$ . [3]  
 (b) Show that  $E\left(\frac{1}{X}\right) = \frac{19}{72}$ . [3]  
 (c) Find  $\text{Var}\left(\frac{1}{X}\right)$ . [5]  
 (d) The random variable  $Y = \frac{X+2}{X}$ . Find  $E(Y)$  and  $\text{Var}(Y)$ . [4]

- 29** The random variable  $X$  takes the values 0, 1, 2, ...,  $n$  with  $P(X = x) = \frac{1}{2^n} C(n, x)$ , where  $C(n, x) =$

$$\frac{n!}{x!(n-x)!}.$$

- (a) Use  $x C(n, x) = n C(n-1, x-1)$  and the binomial expansion of  $(1+1)^{n-1}$  to show that  $E(X) = \frac{n}{2}$ . [4]  
 (b) Given that  $x(x-1)C(n, x) = n(n-1)C(n-2, x-2)$ , show that  $E[X(X-1)] = \frac{n(n-1)}{4}$ . [3]  
 (c) Hence show that  $\text{Var}(X) = \frac{n}{4}$ . [4]  
 (d) For  $n = 12$ , find the mean and standard deviation of  $3X - 5$ . [4]

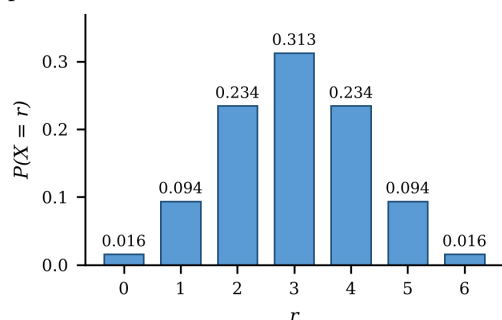
- 30** In a game a fair die is rolled. The player may either accept the score  $\pounds x$  as the prize, or reject it and roll again, in which case the prize is the score on the second roll (which must be accepted).
- (a) Find the expected prize when a player rejects the first roll, and explain why a player who wishes to maximise the expected prize should reject a first score of 1, 2 or 3 but accept 4, 5 or 6. [4]
- (b) Show that with this strategy the expected prize is  $\pounds 17/4$  (that is,  $\pounds 4.25$ ). [4]
- (c) Find the variance of the prize with this strategy. [5]
- (d) The organiser charges  $\pounds 4$  to play. Find the expected profit per game for the player and comment on whether the game is fair. [3]

## Binomial Distribution

- 1** The random variable  $X \sim B(10, 0.3)$ . Find  $P(X = 3)$ . [2]
- 2** A fair coin is tossed 8 times. Find the probability of obtaining at most 2 heads. [2]
- 3** The random variable  $X \sim B(6, 0.4)$ . Find  $P(X \leq 2)$ . [2]
- 4** The random variable  $X \sim B(20, 0.1)$ . Find  $P(X \geq 1)$ . [2]
- 5** The random variable  $X \sim B(30, 0.2)$ . Write down  $E(X)$  and  $\text{Var}(X)$ . [2]
- 6** A fair die is rolled 12 times and  $X$  is the number of sixes obtained.
- (a) State the distribution of  $X$ , giving the values of its parameters. [2]
- (b) Find  $P(X = 3)$ . [2]
- 7** A multiple choice test has 12 questions, each with four possible answers of which exactly one is correct. A student guesses every answer. The number of correct answers is  $X$ .
- (a) State two assumptions needed to model  $X$  by a binomial distribution. [2]
- (b) Find  $P(X \geq 5)$ . [3]
- 8** The random variable  $X \sim B(n, p)$  has mean 6 and variance 4.2.
- (a) Find  $n$  and  $p$ . [4]
- (b) Find  $P(X = 6)$ . [1]
- 9**  $X \sim B(15, 0.35)$ . Find  $P(4 \leq X \leq 8)$ . [3]
- 10** Each seed in a packet germinates with probability 0.85, independently. A gardener sows 12 seeds.
- (a) Find the probability that at least 10 seeds germinate. [2]
- (b) Find the probability that all 12 germinate. [1]
- 11** A pack of 52 playing cards is shuffled. Five cards are dealt without replacement, and  $X$  is the number of hearts. Explain why  $X$  cannot be modelled by a binomial distribution. [2]
- 12** The random variable  $X \sim B(10, 0.2)$ .
- (a) Find  $P(X < 2)$ . [2]
- (b) Two independent observations of  $X$  are made. Find the probability that exactly one of them equals 2. [3]
- 13** The random variable  $X \sim B(18, p)$  has variance 4.
- (a) Show that  $9p^2 - 9p + 2 = 0$  and hence find the two possible values of  $p$ . [4]
- (b) Given also that the mean of  $X$  is greater than 9, find  $P(X = 12)$ . [3]

- 14** A factory produces components, 8% of which are defective independently of each other. A random sample of 15 components is taken.
- Find the probability that the sample contains no defective components. [2]
  - Find the probability that the sample contains at most 2 defective components. [2]
  - Find the smallest sample size  $n$  for which the probability of at least one defective component exceeds 0.95. [3]

- 15** The bar chart shows the probability distribution of  $X \sim B(n, p)$ , with the probabilities given to 3 decimal places.



- Write down the values of  $n$  and  $p$ , giving reasons. [3]
  - Find  $P(X \geq 5)$ . [2]
  - Find the variance of  $X$ . [1]
- 16** A multiple choice test has 20 questions, each with 5 possible answers of which exactly one is correct. A candidate guesses every answer at random. A correct answer scores 4 marks and an incorrect answer scores  $-1$  mark. The number of correct answers is  $X$  and the candidate's total mark is  $M$ .
- State the distribution of  $X$ , with its parameters. [1]
  - Show that  $E(M) = 0$ . [3]
  - Find  $\text{Var}(M)$ . [3]
- 17**  $X \sim B(25, 0.3)$ . Find the largest integer  $c$  such that  $P(X \leq c) \leq 0.05$ , and write down the value of  $P(X \leq c)$ .  
Show your trial values. [4]
- 18**  $X \sim B(8, 0.25)$ . Find  $P(X \geq 3 | X \geq 1)$ . [4]
- 19** Eggs are packed in boxes of 10. Each egg is cracked with probability 0.08, independently. Find the probability that, in a delivery of 5 boxes, exactly 2 boxes contain no cracked eggs. [5]
- 20**  $X \sim B(5, p)$  and  $P(X = 0) = 0.07776$ . Find the value of  $p$  and hence  $P(X = 3)$ . [5]
- 21** The random variable  $X \sim B(12, p)$ . It is given that  $P(X = 1) = 3 P(X = 0)$ .
- Show that  $p = 0.2$ . [4]
  - Find  $P(X \geq 2)$ . [3]
  - Using calculator values of  $P(X = r)$ , find the most likely value (the mode) of  $X$ . [2]
- 22** A quiz has 10 questions, each with 5 possible answers of which exactly one is correct. Sam knows the correct answers to 4 of the questions and guesses the other 6 independently. The number of correct guesses is  $G$ , and Sam's total score is the number of correct answers.  
Sam needs a score of at least 7 to pass.
- Find the probability that Sam passes. [3]
  - Given that Sam passes, find the probability that Sam scores exactly 7. [3]
  - Give one reason why the binomial model for  $G$  might not be realistic. [1]

- 23** A doctor's surgery finds that 20% of the patients who make an appointment do not attend. The doctor has time to see at most 20 patients in a day.
- On a day when 20 appointments are made, find the probability that all the patients attend, and the probability that more than 3 patients do not attend. [4]
  - The doctor books 21 appointments. Find the probability that every patient who attends can be seen. [3]
  - The doctor books 25 appointments. Find the probability that at least one patient who attends cannot be seen. [3]

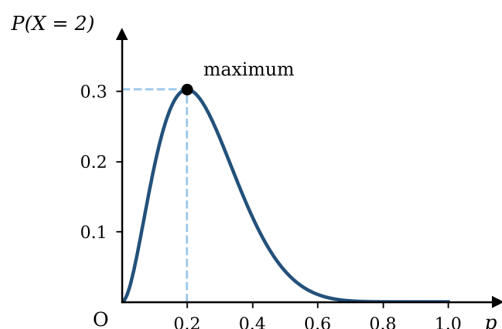
- 24**  $X \sim B(8, 0.3)$  and  $Y \sim B(12, 0.3)$  are independent.
- Find  $P(X = 1 \text{ and } Y = 1)$ . [2]
  - By considering all the ways in which  $X + Y = 2$  can occur, find  $P(X + Y = 2)$ . [3]
  - Explain why  $P(X + Y = 2)$  can be found directly using the distribution  $B(20, 0.3)$ , and verify this. [3]

- 25** A machine produces components, each defective with probability 0.03 independently. A sample of  $n$  components is taken.
- Show that the probability of at most one defective is  $0.97^{n-1}(0.97 + 0.03n)$ . [3]
  - Find the largest value of  $n$  for which this probability exceeds 0.9. [4]

- 26**  $X \sim B(6, p)$ . Given that  $X = 2$ , show that the probability that the successes occurred on the first two trials is  $\frac{1}{15}$ , whatever the value of  $p$ . [5]

- 27** The random variable  $X \sim B(n, p)$ , and  $q = 1 - p$ .
- Show that  $\frac{P(X = r + 1)}{P(X = r)} = \frac{(n - r)p}{(r + 1)q}$ . [3]
  - For  $n = 20$  and  $p = 0.35$ , use part (a) to find the most likely value of  $X$ . [4]
  - For  $p = 0.35$ , find all values of  $n$  for which the mode of  $X$  is 11. [4]

- 28** The random variable  $X \sim B(n, p)$ , where  $n \geq 3$  is fixed and  $p$  can vary. The graph shows  $P(X = 2)$  against  $p$  for  $n = 10$ .



- Show that  $P(X = 2)$  is greatest when  $p = \frac{2}{n}$ . [5]
- Find the maximum value of  $P(X = 2)$  when  $n = 10$ . [2]
- By considering the expansions of  $(q + p)^n$  and  $(q - p)^n$ , where  $q = 1 - p$ , show that the probability that  $X$  is even is  $\frac{1}{2}[1 + (1 - 2p)^n]$ . Hence find this probability for  $n = 10$  and  $p = 0.3$ . [4]

- 29** A batch of components is accepted if a random sample of 20 contains at most one defective. Each component is defective with probability  $p$ , independently.
- (a) Find the probability that the batch is accepted when  $p = 0.02$  and when  $p = 0.10$ . [4]
  - (b) Show that  $P(\text{accept}) = (1 - p)^{19}(1 + 19p)$ , and explain why this decreases as  $p$  increases. [3]
  - (c) Use a numerical method to find the value of  $p$  for which the probability of acceptance is 0.5, giving your answer to 2 significant figures. [4]
  - (d) A supplier claims  $p = 0.02$ . Using your answers, comment on the risk to the supplier of a good batch being rejected. [2]
- 
- 30** A machine is in one of two states. In state N (normal, which has probability 0.9) each item produced is defective with probability 0.02. In state F (faulty, probability 0.1) each item is defective with probability 0.2. Items are defective independently. A sample of 10 items is taken and  $D$  is the number of defective items.
- (a) Find  $P(D \leq 1)$  when the machine is in state N, and when it is in state F. [3]
  - (b) Find the overall probability that  $D \leq 1$ . [2]
  - (c) Given that  $D \leq 1$ , find the probability that the machine is faulty. [3]
  - (d) Given that  $D \geq 2$ , find the probability that the machine is faulty, and comment on what a result of  $D \geq 2$  suggests. [3]

## Poisson Distribution

- 1** The random variable  $X \sim \text{Po}(3)$ . Find  $P(X = 2)$ . [2]
- 
- 2** The random variable  $X \sim \text{Po}(2.5)$ . Find  $P(X \leq 1)$ . [2]
- 
- 3** The random variable  $X \sim \text{Po}(2)$ .
- (a) Show that  $P(X = 3) = \frac{4}{3}e^{-2}$ . [2]
  - (b) Find  $P(X \leq 1)$ , giving your answer in terms of  $e$ . [2]
- 
- 4**  $X \sim \text{Po}(1.2)$ . Find  $P(X \geq 1)$ . [2]
- 
- 5**  $X \sim \text{Po}(7)$ . Write down  $E(X)$ ,  $\text{Var}(X)$  and the standard deviation of  $X$ . [2]
- 
- 6** Emails arrive at random at an average rate of 6 per hour. Find the probability that no email arrives in a 20 minute period. [2]
- 
- 7** The number of telephone calls received by a helpline in one hour is modelled by a Poisson distribution with mean 4.5.
- (a) State two conditions needed for a Poisson model to be suitable. [2]
  - (b) Find the probability of exactly 6 calls in one hour. [2]
- 
- 8** Cars arrive at a toll booth at an average rate of 12 per hour, at random and independently.
- (a) State the distribution of the number of cars arriving in a 10-minute period. [1]
  - (b) Find the probability that at least 3 cars arrive in a 10-minute period. [3]
- 
- 9**  $X \sim \text{Po}(5.5)$ . Find  $P(3 \leq X \leq 7)$ . [3]
- 
- 10** Orders for a particular product arrive at random at an average rate of 3.5 per day. A shop starts a day with 5 of the product in stock.
- (a) Find the probability that exactly 2 orders arrive in the first half of the day. [3]
  - (b) Find the probability that the stock is not enough to meet the orders arriving during the whole day. [3]

- 11** Buses on a route run to a timetable and arrive every 10 minutes. A student says the number of buses arriving at a stop in an hour follows a Poisson distribution.  
Give a reason why this is unlikely to be a good model. [2]

- 12**  $X \sim \text{Po}(\lambda)$  and  $P(X = 0) = 0.1$ . Find the exact value of  $\lambda$  and hence  $P(X = 2)$ . [4]

- 13** The random variable  $X \sim \text{Po}(\lambda)$  and  $P(X = 5) = P(X = 6)$ .  
Find the value of  $\lambda$  and hence find  $P(3 < X \leq 8)$ . [5]

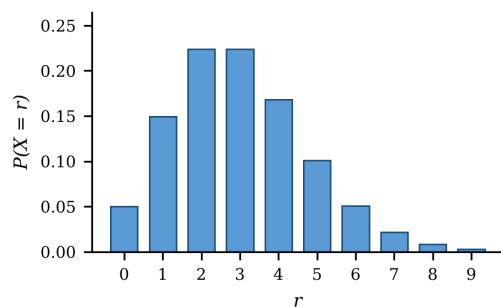
- 14** Flaws occur at random in two fabrics, A and B, at the average rates shown in the table. The numbers of flaws in non-overlapping areas are independent.

A sheet is made from  $2 \text{ m}^2$  of fabric A joined to  $1 \text{ m}^2$  of fabric B.

| Fabric | Mean flaws per $\text{m}^2$ |
|--------|-----------------------------|
| A      | 0.8                         |
| B      | 1.1                         |

- (a) State the distribution of the total number of flaws in the sheet. [2]  
 (b) Find the probability that the sheet has at least 4 flaws. [2]  
 (c) Given that the sheet has exactly 3 flaws, find the probability that exactly 2 of them are in the fabric A part. [3]

- 15** The bar chart shows the distribution of a random variable  $X \sim \text{Po}(\lambda)$ . The bar at  $X = 0$  has height 0.05, correct to 2 decimal places.



- (a) Use the height of the bar at  $X = 0$  to show that  $\lambda$  is approximately 3. [2]  
 (b) Taking  $\lambda = 3$ , find  $P(X \geq 6)$ . [3]  
 (c) Show that  $P(X = 2) = P(X = 3)$  when  $\lambda = 3$ , and say how this shows in the bar chart. [2]

- 16** Faults in a network occur at random at an average rate of 0.8 per day. Find the smallest whole number of days  $d$  for which the probability of at least one fault in  $d$  days exceeds 0.95. [5]

- 17** Light tubes in an office block fail at random at an average rate of 1.5 per working day. The caretaker checks the stock on Monday morning. The block is open for 5 working days each week.  
Find the smallest number of spare tubes the caretaker must hold so that the probability of running out during the week is less than 1%. [5]

- 18**  $X \sim B(200, 0.015)$ .  
 (a) State, with reasons, a suitable Poisson approximation for  $X$ . [2]  
 (b) Use the approximation to estimate  $P(X \leq 2)$ , and compare it with the exact binomial value. [3]

- 19**  $X \sim \text{Po}(\lambda)$ .  
 (a) Show that  $\frac{P(X = r + 1)}{P(X = r)} = \frac{\lambda}{r + 1}$ . [2]  
 (b) Hence find the mode of  $X$  when  $\lambda = 4.3$ , and show that when  $\lambda = 4$  there are two modes. [5]

- 20** The number of calls to a helpline was recorded on each of 100 days. The results are shown in the table.

| Calls $r$ | 0 | 1  | 2  | 3  | 4  | 5 | 6 |
|-----------|---|----|----|----|----|---|---|
| Days      | 9 | 22 | 27 | 21 | 13 | 6 | 2 |

- (a) Calculate the mean and variance of the number of calls per day. [4]  
 (b) Explain why these results support a Poisson model, and use it to find the expected number of days with exactly 2 calls. [3]

- 21** The random variable  $X \sim \text{Po}(\lambda)$ , and  $P(X = 3) = 1.5 P(X = 1)$ .

- (a) Show that  $\lambda = 3$ . [3]  
 (b) Find  $P(X \geq 4)$ . [2]  
 (c) Five independent observations of  $X$  are made. Find the probability that exactly 2 of them are at least 4. [3]

- 22** The number of typing errors on a page of a long document is modelled by a Poisson distribution with mean 0.6.

- (a) Find the probability that a page contains at least one error. [2]  
 (b) Find the probability that a section of 5 pages contains at most 2 errors. [3]  
 (c) Find the smallest number of pages  $n$  for which the probability of no errors at all is less than 0.01. [3]

- 23** Vehicles pass a point at random at an average rate of 1.5 per minute. Over 10 separate one-minute intervals, find the probability that exactly 3 of the intervals have no vehicle. [5]

- 24** The random variable  $X \sim \text{Po}(4)$ . The random variable  $Y = aX + b$ , where  $a$  and  $b$  are positive constants, has  $E(Y) = \text{Var}(Y) = 16$ .

- (a) Find the values of  $a$  and  $b$ . [4]  
 (b) Although  $E(Y) = \text{Var}(Y)$ , show that  $Y$  cannot have a Poisson distribution. [3]

- 25** A call centre receives calls at an average rate of 8 per hour in busy periods and 3 per hour in quiet periods, independently and at random. On a certain day there are 2 busy hours and 3 quiet hours.

- (a) State the distribution of the total number of calls  $T$  during these 5 hours, giving your reasoning. [3]  
 (b) Find  $P(T \leq 20)$  and  $P(T > 30)$ . [3]  
 (c) Given that  $T = 22$ , find the probability that exactly 14 of the calls were in the busy hours. [3]

- 26** Faults arise in a machine at random at an average rate of 0.7 per month. Each fault costs £50 to repair. The number of faults in a month is  $F$  and the monthly cost is  $C = 50F$  pounds.

- (a) Find  $E(C)$  and the standard deviation of  $C$ . [4]  
 (b) Find the probability that the cost in a month exceeds £100. [3]  
 (c) Find the probability that the cost over a 6-month period is less than £200. [3]

- 27**  $X \sim \text{Po}(\lambda)$  and  $Y \sim \text{Po}(\mu)$  are independent random variables.

- (a) Show that  $P(X = r \mid X + Y = n) = C(n, r) \left(\frac{\lambda}{\lambda + \mu}\right)^r \left(\frac{\mu}{\lambda + \mu}\right)^{n-r}$ , so that, given  $X + Y = n$ ,  $X$  has a binomial distribution. [5]  
 (b) For  $\lambda = 2$  and  $\mu = 3$ , find  $P(X = 2 \mid X + Y = 5)$ . [3]  
 (c) A random variable  $W \sim B(150, 0.02)$ . Use a suitable Poisson approximation to estimate  $P(W \leq 2)$ , and explain why the approximation is reasonable. [3]

- 28** Calls arrive at a switchboard as a Poisson process at an average rate of 3 per hour, as shown by the marks on the time line (time in minutes).



- Find the probability that no calls arrive in a given 20-minute period. [2]
- Let  $T$  be the waiting time, in hours, from a given moment until the next call. Show that  $P(T > t) = e^{-3t}$ . [2]
- Hence find the probability that the next call arrives between 10 and 30 minutes from now. [2]
- Given that exactly 2 calls arrive in a particular hour, find the probability that both arrived in its first 20 minutes. [4]

- 29**  $X \sim \text{Po}(\lambda)$ , so that  $P(X = r) = e^{-\lambda} \frac{\lambda^r}{r!}$  for  $r = 0, 1, 2, \dots$

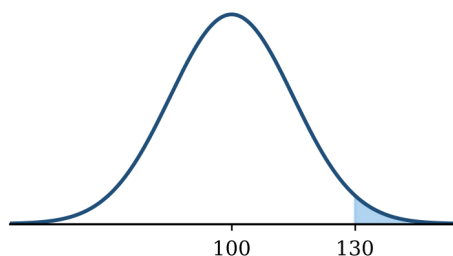
- Using the series  $e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots$ , show that the probabilities sum to 1. [2]
- Show that  $E(X) = \lambda$ . [3]
- Show that  $E[X(X - 1)] = \lambda^2$ , and hence that  $\text{Var}(X) = \lambda$ . [5]

- 30** Cars pass a checkpoint as a Poisson process at an average rate of 20 per hour. Each car, independently, is a van with probability 0.3. Let  $N$  be the number of cars in an hour and  $V$  the number of vans.

- Show that  $P(V = v) = \sum_{n=v}^{\infty} e^{-20} \frac{20^n}{n!} C(n, v)(0.3)^v(0.7)^{n-v}$ . [3]
- Simplify this to show that  $V \sim \text{Po}(6)$ . [5]
- Find the probability that at least 4 vans pass in an hour. [2]
- Given that exactly 5 cars passed in a 15 minute period, find the probability that none was a van. [2]

## Normal Distribution

- The random variable  $X \sim N(50, 4^2)$ . Find  $P(X < 54)$ . [2]
- The IQ scores of a large group are modelled by  $X \sim N(100, 15^2)$ . The shaded region in the diagram represents the proportion of the group with a score above 130. Find this proportion.



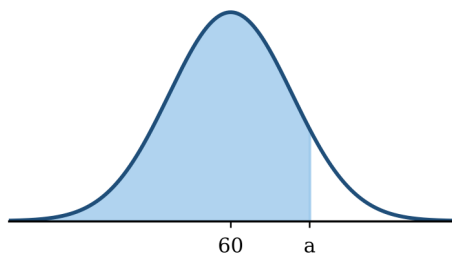
- $X \sim N(30, 3^2)$ .
  - Find  $P(X > 33)$ . [2]
  - Find  $P(27 < X < 36)$ . [2]
- $Z \sim N(0, 1)$ . Find  $P(-1 < Z < 2)$ . [2]

- 5 The masses of apples are normally distributed with mean 150 g and standard deviation 8 g. An apple is weighed to the nearest gram. Find the probability that its recorded mass is 160 g. [3]

- 6  $X \sim N(40, 5^2)$ . Find the value  $a$  such that  $P(X < a) = 0.8$ . [2]

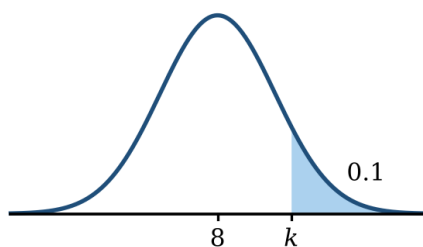
- 7 The heights of adult men in a town are normally distributed with mean 172 cm and standard deviation 8 cm.
- (a) Find the probability that a randomly chosen man has height between 165 cm and 180 cm. [3]
- (b) In a sample of 500 men, estimate how many are taller than 185 cm. [2]

- 8 The random variable  $X \sim N(60, 10^2)$ .



- (a) The shaded area in the diagram is 0.9. Find the value of  $a$ . [2]
- (b) Find the symmetrical interval about the mean which contains 95% of the distribution. [3]

- 9 The random variable  $X \sim N(8, 2^2)$ . The shaded area in the diagram is 0.1. Find the value of  $k$ .

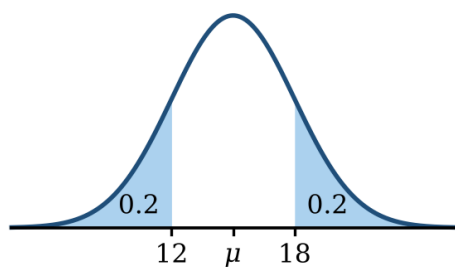


[3]

- 10 The mass of a bag of flour is normally distributed with mean 1.5 kg and standard deviation 0.02 kg. A bag is underweight if its mass is less than 1.46 kg.
- (a) Find the probability that a bag is underweight. [2]
- (b) In a delivery of 2000 bags, estimate the number that are underweight. [2]

- 11 Marks in an examination are modelled by  $N(70, 12^2)$ . Find the lower and upper quartiles of the marks. [3]

- 12  $X$  is normally distributed with mean 15. It is given that  $P(X < 12) = 0.2$ , and that  $P(X > 18) = 0.2$ , as shown.

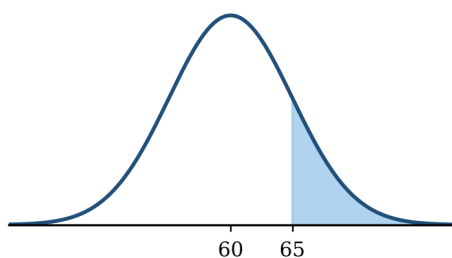


- (a) Find  $P(12 < X < 18)$ . [2]
- (b) Find the standard deviation of  $X$ . [3]

**13**  $X \sim N(\mu, 6^2)$  and  $P(X > 40) = 0.3$ . Find the value of  $\mu$ . [3]

**14** The mass  $X$  grams of a certain type of sweet is normally distributed with mean  $\mu$  and standard deviation  $\sigma$ . It is known that 10% of the sweets have mass less than 22 g and 5% have mass greater than 35 g.  
Find  $\mu$  and  $\sigma$ . [6]

**15** The mass of an egg, in grams, is normally distributed with mean 60 and standard deviation 5. An egg is graded as large if its mass exceeds 65 g.  
The diagram shows the region that corresponds to a large egg.



(a) Find the probability that an egg chosen at random is large. [2]

(b) A box contains 10 eggs chosen at random. Find the probability that at least 2 are large. [4]

**16** The random variable  $Y \sim N(\mu, \sigma^2)$ . It is given that  $P(Y < 21) = P(Y > 33) = 0.1587$ . [4]

(a) Find the values of  $\mu$  and  $\sigma$ . [2]

(b) Find  $P(24 < Y < 35)$ . [2]

**17** The random variable  $X \sim N(\mu, \sigma^2)$ . It is given that  $P(X > \mu + a) = 0.15$ , where  $a$  is a positive constant. [1]

(a) Write down  $P(X < \mu - a)$ . [2]

(b) Find  $P(\mu - a < X < \mu + a)$ . [3]

(c) Find  $P(X < \mu + a \mid X > \mu - a)$ . [3]

**18**  $X \sim \text{Po}(36)$ .

(a) Explain why  $X$  may be approximated by a normal distribution, and state its mean and variance. [2]

(b) Use the approximation, with a continuity correction, to estimate  $P(X \leq 30)$ , and compare with the exact value. [4]

**19** Heights of adult women are normally distributed with mean  $\mu$  cm and standard deviation  $\sigma$  cm. Find the probability that a woman is more than  $\sigma$  cm above the mean, given that she is taller than the mean. [4]

**20** A researcher models the time, in hours, a patient waits in a clinic by  $T \sim N(8, 4^2)$ .  
Show that this model assigns a probability of about 2.3% to a negative waiting time, and explain what this suggests about the model. [4]

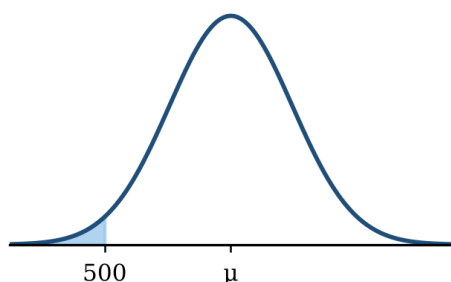
**21** The random variable  $X \sim B(80, 0.4)$ .

(a) Explain why  $X$  may be approximated by a normal distribution, and state the parameters of this normal distribution. [3]

(b) Use this approximation, with a continuity correction, to estimate  $P(X \geq 38)$ . [3]

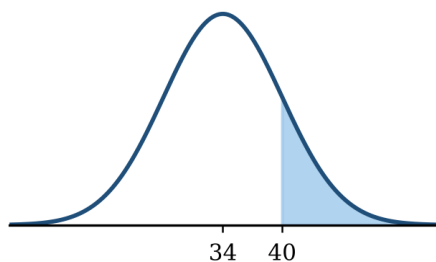
(c) Estimate  $P(30 \leq X \leq 36)$  by the normal approximation and compare with the exact binomial value 0.5617. [2]

- 22** A machine fills bottles with a liquid. The volume of liquid, in ml, in a bottle is normally distributed with mean  $\mu$  and standard deviation 0.8. The label on each bottle states 500 ml.



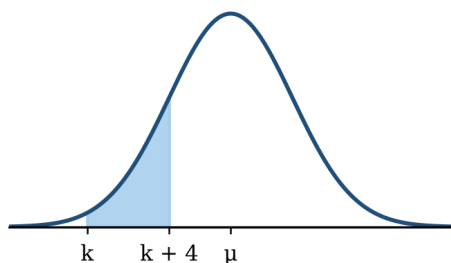
- (a) The machine is set so that 2% of bottles contain less than 500 ml, as shown in the diagram. Find  $\mu$ . [3]
- (b) With this value of  $\mu$ , find the proportion of bottles containing more than 503 ml. [2]
- (c) The manufacturer wants at most 2% of bottles below 500 ml and at most 1% above 503 ml. Find the largest standard deviation, and the corresponding mean, that achieves this. [4]
- 
- 23** Marks in an examination are normally distributed with mean 58 and standard deviation 14. The pass mark is set so that 85% of candidates pass, and a distinction is awarded to the top 12%.
- (a) Find the pass mark and the minimum mark for a distinction, to the nearest whole number. [5]
- (b) Find the probability that a candidate, chosen at random from those who passed, gained a distinction. [3]
- 
- 24** The independent random variables  $X \sim N(80, 5^2)$  and  $Y \sim N(50, 10^2)$  are such that  $P(X < k) = P(Y > k)$  for a particular value  $k$ .
- (a) Find the value of  $k$ . [4]
- (b) Find the common value of  $P(X < k)$  and  $P(Y > k)$ . [2]
- (c) A value is chosen at random from  $X$  and a value from  $Y$ . Find the probability that both are less than  $k$ . [2]
- 
- 25** The lengths of planks produced by a sawmill are normally distributed with mean 2.00 m and standard deviation 0.04 m. A plank is rejected if it is shorter than 1.95 m or longer than 2.08 m; all other planks are accepted.
- (a) Find the proportion of planks that are accepted. [3]
- (b) Find the median length of the accepted planks, to the nearest millimetre. [4]
- (c) Find the lower quartile of the lengths of the accepted planks, to the nearest millimetre. [3]
- 
- 26**  $X \sim N(\mu, \sigma^2)$ . The lower and upper quartiles are  $Q_1$  and  $Q_3$ , and a value is classed as an outlier if it lies more than  $1.5(Q_3 - Q_1)$  beyond a quartile.
- (a) Show that  $Q_3 - Q_1 = 1.349\sigma$ , to 4 significant figures. [3]
- (b) Find the outlier boundaries in terms of  $\mu$  and  $\sigma$ . [2]
- (c) Find the probability that an observation from a normal distribution is classed as an outlier. [2]

- 27** A commuter's journey time  $T$  minutes is normally distributed with mean 34 and standard deviation 6.



- Find  $P(T > 40 \mid T > 35)$ . [3]
- The commuter makes 5 journeys. Given that at least one takes more than 40 minutes, find the probability that exactly one does. [4]
- Find the value  $t$  such that  $P(T < t \mid T > 30) = 0.5$ . [4]

- 28** The random variable  $X \sim N(\mu, \sigma^2)$ . For a fixed width 4, the probability  $P(k < X < k + 4)$  depends on  $k$ , as the shaded region in the diagram suggests.



- Let  $\phi$  be the probability density function of  $X$ . By differentiating  $P(k < X < k + 4)$  with respect to  $k$ , show that the probability is greatest when the interval is symmetrical about  $\mu$ , that is when  $k = \mu - 2$ . [5]
- For  $X \sim N(50, 3^2)$ , find the greatest value of  $P(k < X < k + 4)$ . [2]
- Find the width of the shortest interval which contains 90% of the distribution when  $\sigma = 3$ . [4]

- 29** In a large group of students, 55% are girls and 45% are boys. The heights of girls are normally distributed with mean 165 cm and standard deviation 6 cm, and those of boys with mean 177 cm and standard deviation 7 cm.

- Find the probability that a randomly chosen student is taller than 175 cm. [5]
- Given that a student is taller than 175 cm, find the probability that the student is a boy. [3]
- Find the median height of the whole group, giving your answer to 1 decimal place, and find the height exceeded by 10% of the group, to 1 decimal place. [6]

- 30**  $X \sim B(40, 0.5)$ .

- State a normal approximation for  $X$ , with the values of the mean and variance. [2]
- Find  $P(X \leq 17)$  exactly, then estimate it using the normal approximation (i) without a continuity correction and (ii) with a continuity correction. [5]
- Explain why the continuity correction gives the better estimate. [2]
- Use the approximation with a continuity correction to estimate  $P(X = 20)$ , and find the percentage error compared with the exact value. [5]

## Continuous Random Variables

- 1** The continuous random variable  $X$  has probability density function  $f(x) = kx$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise. Find the value of  $k$ . [2]

- 2 The continuous random variable  $X$  has probability density function  $f(x) = \frac{3}{8}x^2$  for  $0 \leq x \leq 2$ , and 0 otherwise. Find  $P(X < 1)$ . [3]

- 3 Each function below is defined on the stated interval and is zero elsewhere. State, with a reason, whether it could be the probability density function of a continuous random variable.

(a)  $f(x) = x - 1$  for  $0 \leq x \leq 3$ . [1]

(b)  $f(x) = \frac{x}{3}$  for  $0 \leq x \leq 2$ . [2]

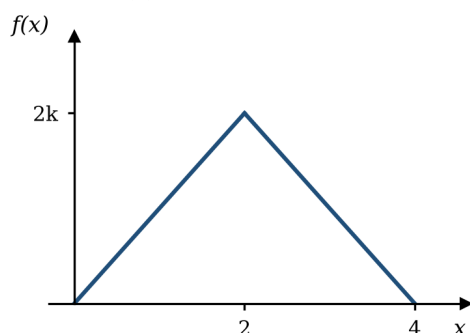
(c)  $f(x) = \frac{3x^2}{8}$  for  $0 \leq x \leq 2$ . [1]

- 4 The continuous random variable  $X$  has probability density function  $f(x) = \frac{x}{8}$  for  $0 \leq x \leq 4$ , and  $f(x) = 0$  otherwise. Find  $P(X < 2)$ . [2]

- 5 The cumulative distribution function of  $X$  is  $F(x) = \frac{x^3}{8}$  for  $0 \leq x \leq 2$  (with  $F(x) = 0$  for  $x < 0$  and  $F(x) = 1$  for  $x > 2$ ). Find the probability density function of  $X$  and  $P(X > 1)$ . [3]

- 6 The continuous random variable  $X$  has probability density function  $f(x) = \frac{1}{5}$  for  $2 \leq x \leq 7$ , and  $f(x) = 0$  otherwise. Find  $P(3 < X < 5)$  and  $E(X)$ . [3]

- 7 The diagram shows the probability density function of  $X$ :  $f(x) = kx$  for  $0 \leq x \leq 2$ ,  $f(x) = k(4 - x)$  for  $2 \leq x \leq 4$ , and  $f(x) = 0$  otherwise.



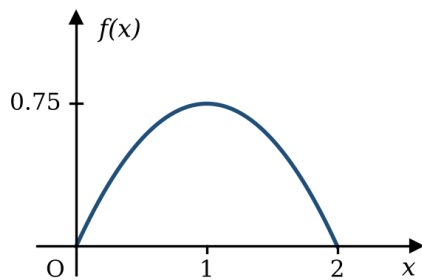
- (a) Find the value of  $k$ . [3]  
 (b) Find  $P(X < 1)$ . [1]  
 (c) Write down  $E(X)$ , giving a reason. [1]

- 8 The continuous random variable  $X$  has cumulative distribution function  $F(x) = \frac{x^2}{9}$  for  $0 \leq x \leq 3$ , with  $F(x) = 0$  for  $x < 0$  and  $F(x) = 1$  for  $x > 3$ .

- (a) Find the probability density function of  $X$ . [1]  
 (b) Find  $P(1 < X < 2)$ . [1]  
 (c) Find the median of  $X$ . [2]

- 9 The continuous random variable  $X$  has probability density function  $f(x) = k(x + 1)$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise. Find  $k$  and  $P(X > 1)$ . [4]

- 10** The diagram shows the probability density function  $f(x) = \frac{3}{4}(2x - x^2)$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise.

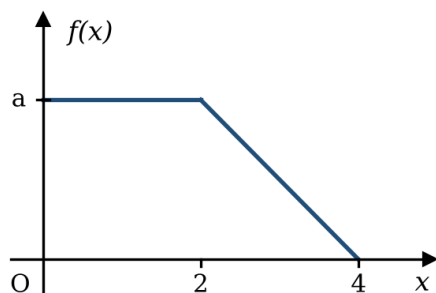


- (a) State  $E(X)$ , giving a reason. [2]  
 (b) Find  $\text{Var}(X)$ . [3]

- 11** The continuous random variable  $X$  has probability density function  $f(x) = \frac{x}{2}$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise. Find the exact value of the median of  $X$ . [4]

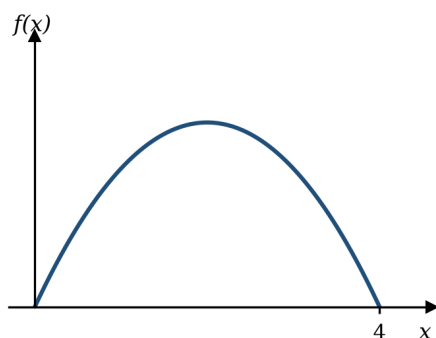
- 12** The continuous random variable  $X$  has probability density function  $f(x) = kx^2(3 - x)$  for  $0 \leq x \leq 3$ , and  $f(x) = 0$  otherwise. Show that  $k = \frac{4}{27}$ , find the mode of  $X$  and find  $P(X > 2)$ . [6]

- 13** The diagram shows the probability density function of  $X$ :  $f(x) = a$  for  $0 \leq x \leq 2$ ,  $f(x)$  decreases linearly from  $a$  at  $x = 2$  to  $0$  at  $x = 4$ , and  $f(x) = 0$  otherwise.



- (a) Find the value of  $a$ . [3]  
 (b) Find  $P(X > 2.5)$ . [3]

- 14** The continuous random variable  $X$  has probability density function  $f(x) = kx(4 - x)$  for  $0 \leq x \leq 4$ , and  $f(x) = 0$  otherwise, as sketched.



- (a) Show that  $k = \frac{3}{32}$ . [3]  
 (b) Find  $E(X)$  and  $\text{Var}(X)$ . [4]

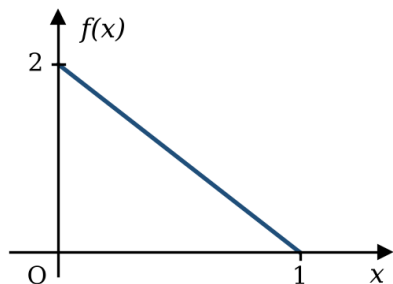
- 15** The continuous random variable  $X$  has probability density function  $f(x) = \frac{3}{x^4}$  for  $x \geq 1$ , and  $f(x) = 0$  otherwise.
- (a) Find the cumulative distribution function  $F(x)$ . [3]  
 (b) Find the median of  $X$ . [2]  
 (c) Find  $E(X)$ . [2]

- 16** The continuous random variable  $X$  has probability density function  $f(x) = \frac{k}{x^2}$  for  $1 \leq x \leq 4$ , and  $f(x) = 0$  otherwise. Find  $k$ , the median of  $X$  and  $E(X)$ . [6]

- 17** The continuous random variable  $X$  has cumulative distribution function  $F(x) = \frac{x^3 + x}{10}$  for  $0 \leq x \leq 2$ , with  $F(x) = 0$  for  $x < 0$  and  $F(x) = 1$  for  $x > 2$ .
- (a) Verify that  $F(2) = 1$  and find the probability density function  $f(x)$ . [3]  
 (b) Find  $P(\frac{1}{2} < X < \frac{3}{2})$ . [2]  
 (c) Show that the median  $m$  satisfies  $m^3 + m - 5 = 0$ , and find  $m$  to 3 significant figures. [3]

- 18** The continuous random variable  $X$  has probability density function  $f(x) = \frac{3x^2}{8}$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise. Find  $E(X)$ ,  $\text{Var}(X)$  and the standard deviation of  $X$ . [6]

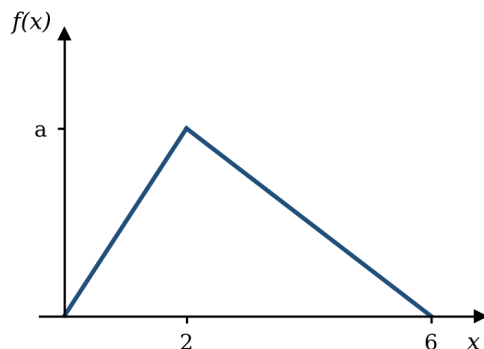
- 19** The continuous random variable  $X$  has probability density function  $f(x) = 2(1 - x)$  for  $0 \leq x \leq 1$ , and  $f(x) = 0$  otherwise, as shown.



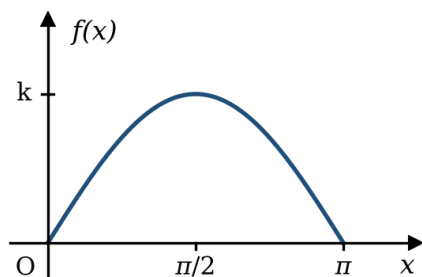
- (a) Write down the mode of  $X$ . [1]  
 (b) Find the median and the mean of  $X$ . [4]  
 (c) State the skewness of the distribution, with a reason. [1]

- 20** The lifetime  $X$ , in years, of a component has probability density function  $f(x) = \frac{3x^2}{64}$  for  $0 \leq x \leq 4$ , and  $f(x) = 0$  otherwise. A machine contains two such components, with independent lifetimes, and works only while both components are working.
- (a) Find  $P(X > 3)$ . [2]  
 (b) Find the probability that the machine stops working within 3 years. [3]  
 (c) Find the median lifetime of a single component. [2]

- 21** The continuous random variable  $X$  has probability density function  $f(x) = kx$  for  $0 \leq x \leq 2$ ,  $f(x) = \frac{k(6-x)}{2}$  for  $2 \leq x \leq 6$ , and  $f(x) = 0$  otherwise, where  $k$  is a constant. The graph of  $f$  is sketched, with maximum height  $a$ .



- (a) Find the values of  $k$  and  $a$ . [2]  
 (b) Find the median of  $X$ . [4]  
 (c) Show that  $E(X) = \frac{8}{3}$ . [3]
- 
- 22** The continuous random variable  $X$  has probability density function  $f(x) = a + bx$  for  $0 \leq x \leq 1$ , and  $f(x) = 0$  otherwise. It is given that  $E(X) = 0.6$ .
- (a) Find the values of  $a$  and  $b$ . [5]  
 (b) Find  $P(X > 0.5)$ . [2]  
 (c) State the mode of  $X$ , with a reason. [1]
- 
- 23** The continuous random variable  $X$  has probability density function  $f(x) = ax^2 + bx$  for  $0 \leq x \leq 1$ , and  $f(x) = 0$  otherwise. It is given that  $E(X) = 0.7$ .
- (a) Show that  $2a + 3b = 6$  and  $3a + 4b = 8.4$ , and hence find  $a$  and  $b$ . [6]  
 (b) Hence find  $P(X > 0.5)$ . [3]
- 
- 24** The continuous random variable  $X$  has probability density function  $f(x) = \frac{x}{8}$  for  $0 \leq x \leq 2$ ,  $f(x) = \frac{1}{4}$  for  $2 < x \leq 5$ , and  $f(x) = 0$  otherwise.
- (a) Find the cumulative distribution function  $F(x)$ , specifying it fully. [5]  
 (b) Find the median and the interquartile range of  $X$ , and show that  $E(X) = \frac{71}{24}$ . [6]
- 
- 25** The continuous random variable  $X$  has probability density function  $f(x) = k \sin x$  for  $0 \leq x \leq \pi$ , and  $f(x) = 0$  otherwise.

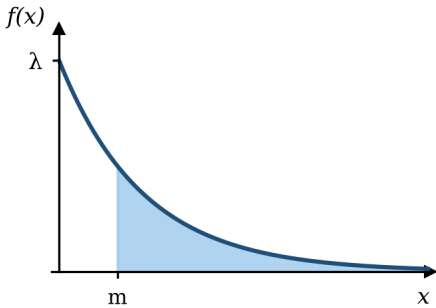


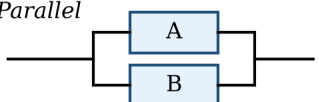
- (a) Show that  $k = \frac{1}{2}$ . [2]  
 (b) Find  $P(X < \frac{\pi}{3})$  and write down the median of  $X$ . [3]  
 (c) Show that  $\text{Var}(X) = \frac{\pi^2}{4} - 2$ . [5]

- 26** The time  $X$ , in hours, taken to repair a machine has probability density function  $f(x) = \frac{3x^2}{8}$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise. The cost of a repair is  $\pounds C$ , where  $C = 20 + 15X^2$ .
- Find the expected cost of a repair. [4]
  - Find the probability that a repair costs more than  $\pounds 50$ . [4]
- 
- 27** The continuous random variable  $X$  has probability density function  $f(x) = \frac{1}{x \ln 5}$  for  $1 \leq x \leq 5$ , and  $f(x) = 0$  otherwise.
- Show that  $f$  is a valid probability density function and find  $F(x)$  for  $1 \leq x \leq 5$ . [3]
  - Find the median, giving the exact value. [2]
  - Find  $E(X)$ , giving your answer to 3 significant figures. [3]
  - Let  $Y = \ln X$ . By finding  $P(Y \leq y)$  for  $0 \leq y \leq \ln 5$ , show that  $Y$  has a continuous uniform distribution, and state its range. [4]
- 
- 28** For a positive integer  $n$ , the continuous random variable  $X$  has probability density function  $f(x) = (n + 1)x^n$  for  $0 \leq x \leq 1$ , and  $f(x) = 0$  otherwise.
- Show that  $F(x) = x^{n+1}$  for  $0 \leq x \leq 1$ . [2]
  - Find the median of  $X$  in terms of  $n$ . [2]
  - Show that  $E(X) = \frac{n+1}{n+2}$ . [3]
  - Find the least integer  $n$  for which the median exceeds 0.95. [3]
- 
- 29** The continuous random variable  $X$  has probability density function  $f(x) = \frac{3x^2}{8}$  for  $0 \leq x \leq 2$ , and  $f(x) = 0$  otherwise. Let  $Y = X^3$ .
- Show that the cumulative distribution function of  $Y$  is  $G(y) = \frac{y}{8}$  for  $0 \leq y \leq 8$ , and name the distribution of  $Y$ . [5]
  - Write down  $E(Y)$  and  $\text{Var}(Y)$ , and hence find  $E(X^3)$  and  $E(X^6)$ . [5]
  - A computer produces  $U$ , uniform on  $[0, 1]$ . Find a function  $g$  such that  $g(U)$  has the same distribution as  $X$ . [3]
  - Check your answer to (c) by finding  $P(2U^{1/3} \leq 1)$ . [2]
- 
- 30**  $X$  is a continuous random variable with probability density function  $f(x)$  on  $0 \leq x \leq b$ , and cumulative distribution function  $F(x)$ .
- By integrating by parts, show that  $E(X) = \int_0^b [1 - F(x)] dx$ . [4]
  - For  $f(x) = \frac{x^2}{9}$  on  $0 \leq x \leq 3$ , find  $F(x)$  and use part (a) to find  $E(X)$ . [3]
  - Verify your answer to (b) by calculating  $E(X)$  directly from  $f(x)$ . [2]
  - Show similarly that  $E(X^2) = \int_0^b 2x[1 - F(x)] dx$ , and use it to find  $\text{Var}(X)$  for the distribution in (b). [6]

## Exponential Distribution

- 1** The random variable  $T$  has an exponential distribution with probability density function  $f(t) = 0.5e^{-0.5t}$  for  $t \geq 0$ . Find  $P(T > 2)$ . [2]
- 
- 2** The random variable  $T \sim \text{Exp}(4)$ , that is, it has probability density function  $f(t) = 4e^{-4t}$  for  $t \geq 0$ .
- Write down  $E(T)$ . [1]
  - Find  $P(T \leq 0.5)$ . [2]

- 3** The random variable  $T$  has cumulative distribution function  $F(t) = 1 - e^{-0.8t}$  for  $t \geq 0$ .  
 (a) Find the probability density function of  $T$ . [2]  
 (b) Write down  $E(T)$ . [1]
- 
- 4**  $T \sim \text{Exp}(0.25)$ . Write down  $E(T)$ ,  $\text{Var}(T)$  and the standard deviation of  $T$ . [3]
- 
- 5** The time  $T$  minutes between arrivals of customers is modelled by an exponential distribution with mean 5. Write down the value of  $\lambda$  and find  $P(T > 5)$ . [2]
- 
- 6** The random variable  $T$  has probability density function  $f(t) = 3e^{-3t}$  for  $t \geq 0$ . Find the cumulative distribution function  $F(t)$ . [2]
- 
- 7** The waiting time,  $T$  minutes, at a bus stop is modelled by an exponential distribution with mean 8 minutes.  
 (a) Find the probability that a waiting time exceeds 10 minutes. [2]  
 (b) Find the probability that a waiting time is between 5 and 12 minutes. [2]
- 
- 8** The lifetime  $X$  of a component, in years, has cumulative distribution function  $F(x) = 1 - e^{-0.2x}$  for  $x \geq 0$ .  
 The diagram shows the probability density function with the median  $m$  marked; the shaded area is  $\frac{1}{2}$ .
- 
- (a) Find the median lifetime. [2]  
 (b) Find the lifetime which is exceeded by 10% of components. [2]
- 
- 9**  $T \sim \text{Exp}(\lambda)$  and  $P(T > 3) = 2P(T > 6)$ .  
 Find the exact value of  $\lambda$  and the mean of  $T$ . [4]
- 
- 10**  $T \sim \text{Exp}(0.1)$ . Find the median and the upper quartile of  $T$ . [4]
- 
- 11**  $T \sim \text{Exp}(\lambda)$  and  $P(T > 2) = 0.3$ . Find the value of  $\lambda$  and the mean of  $T$ . [3]
- 
- 12** The lifetime of a type of light bulb is exponentially distributed with mean 5 years. In a batch of 200 bulbs, estimate how many last longer than 8 years. [3]
- 
- 13**  $T \sim \text{Exp}(0.5)$ . Find the value  $t$  such that  $P(T > t) = 0.01$ . [3]
- 
- 14** The time  $T$ , in hours, for which a battery works has an exponential distribution with parameter  $\lambda$ . It is known that  $P(T > 3) = 0.6$ .  
 (a) Find  $\lambda$ . [3]  
 (b) Write down the mean lifetime. [1]  
 (c) Find  $P(T > 6 \mid T > 3)$ , and comment on your answer. [3]
- 
- 15** The random variable  $T$  has an exponential distribution with parameter  $\lambda$ .  
 (a) Show that  $P(T > s + t \mid T > s) = P(T > t)$  for  $s, t \geq 0$ . [4]  
 (b) A machine part has an exponential lifetime with mean 8000 hours. It has already worked for 5000 hours. Find the probability that it works for at least 2000 more hours. [3]

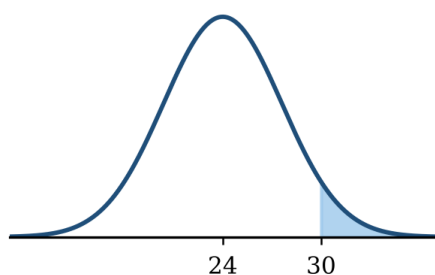
- 16** Customers at a service counter wait for a time  $T$  minutes, where  $T$  has an exponential distribution. A customer gives up and leaves if the wait would exceed 10 minutes, and 25% of customers leave in this way.
- Show that  $\lambda = \frac{\ln 4}{10}$  and find its value to 3 significant figures. [2]
  - Of the 40 customers who are served in one hour, estimate how many wait for less than 5 minutes. [3]
  - Find the probability that a customer who is served (does not leave) waited for less than 5 minutes. [3]
- 
- 17** Cars arrive at a toll gate as a Poisson process at an average rate of 12 per hour. Let  $T$  be the time in hours between consecutive arrivals.
- State the distribution of  $T$ . [1]
  - Find the probability that the gap between two cars exceeds 10 minutes. [2]
  - Find the probability that at least 2 cars arrive in a 30 minute period. [3]
- 
- 18** The lifetimes of 8 independent components are each exponentially distributed with mean 20 hours. Find the probability that exactly 3 of them last longer than 30 hours. [5]
- 
- 19** The durations of 100 telephone calls were recorded. Their mean was 3.2 minutes, and 17 of the calls lasted longer than 6 minutes. It is proposed to model call durations by an exponential distribution.
- Estimate the parameter  $\lambda$  using the sample mean. [2]
  - Use the model to find the expected number of calls, out of 100, that last longer than 6 minutes, and comment on how well the model fits. [3]
- 
- 20** The continuous random variable  $T$  has probability density function  $f(t) = ke^{-0.4t}$  for  $t \geq 0$ , and  $f(t) = 0$  otherwise.
- Show that  $k = 0.4$ . [3]
  - State the distribution of  $T$ , and find  $P(T > 3)$ . [3]
- 
- 21** The random variable  $T$  has probability density function  $f(t) = \lambda e^{-\lambda t}$  for  $t \geq 0$ , where  $\lambda > 0$ .
- Use integration by parts to show that  $E(T) = \frac{1}{\lambda}$ . [4]
  - Given that  $E(T^2) = \frac{2}{\lambda^2}$ , show that the standard deviation of  $T$  equals its mean. [2]
  - For a mean of 5, find the probability that  $T$  exceeds its mean plus one standard deviation. [2]
- 
- 22** A system has two independent components, A and B, with lifetimes in hours  $T_A \sim \text{Exp}(0.1)$  and  $T_B \sim \text{Exp}(0.2)$ . The components can be connected in series (the system works only while both work) or in parallel (the system works while at least one works), as shown.
- Series*
- 
- Parallel*
- 
- For the series system, show that the system lifetime  $S$  satisfies  $P(S > t) = e^{-0.3t}$ . [3]
  - State the mean lifetime of the series system. [2]
  - For the parallel system, find the probability that it lasts longer than 10 hours. [3]

- 23** The independent random variables  $T_A \sim \text{Exp}(\lambda)$  and  $T_B \sim \text{Exp}(\mu)$  are the lifetimes of two components, A and B.
- (a) Show that  $P(T_A < T_B) = \int_0^\infty \lambda e^{-\lambda t} e^{-\mu t} dt$ . [3]
- (b) Hence show that  $P(T_A < T_B) = \frac{\lambda}{\lambda + \mu}$ . [3]
- (c) For  $\lambda = 0.3$  and  $\mu = 0.2$ , find the probability that A fails first, and the probability that the first failure occurs after time 4. [3]
- 
- 24**  $T \sim \text{Exp}(\lambda)$  and  $P(1 < T < 3) = 0.3$ .
- (a) Show that  $u^3 - u + 0.3 = 0$ , where  $u = e^{-\lambda}$ . [3]
- (b) The equation in (a) has two roots between 0 and 1. Find the corresponding values of  $\lambda$ . [4]
- (c) Given that the mean of  $T$  exceeds 2, find the value of  $\lambda$  and state the mean. [3]
- 
- 25** A machine fails at random, the time between failures being exponentially distributed with mean 100 hours. Each failure costs £250 to repair.
- (a) State the distribution of the number of failures  $N$  in 1000 hours of running, and find the expected repair cost. [3]
- (b) Find the probability that there are more than 12 failures in 1000 hours. [2]
- (c) Find the probability that no failure occurs in the first 150 hours, given that none occurs in the first 50 hours. [2]
- 
- 26** A part has lifetime  $T$  hours, where  $T \sim \text{Exp}(\lambda)$ . To avoid failures the part is replaced after  $c$  hours if it is still working. The time in service is  $M = T$  if  $T < c$ , and  $M = c$  otherwise.
- (a) Show that  $E(M) = \int_0^c t\lambda e^{-\lambda t} dt + ce^{-\lambda c}$ . [3]
- (b) Use integration by parts to show that  $E(M) = \frac{1 - e^{-\lambda c}}{\lambda}$ . [4]
- (c) For  $\lambda = 0.02$  and  $c = 50$ , find  $E(M)$  and the probability that the part is replaced rather than failing. [3]
- (d) Explain why  $E(M)$  is less than  $E(T) = 50$  hours, and say what happens to  $E(M)$  as  $c \rightarrow \infty$ . [2]
- 
- 27** Calls arrive at a helpline as a Poisson process at rate 6 per hour. Let  $T$  be the time, in hours, until the first call and let  $U$  be the time until the second call.
- (a) Show that  $T$  has cumulative distribution function  $F(t) = 1 - e^{-6t}$ , and state the distribution of  $T$ . [2]
- (b) Explain why  $P(U > t) = e^{-6t}(1 + 6t)$ . [2]
- (c) Find the probability that the second call arrives within 30 minutes. [2]
- (d) Hence find the probability density function of  $U$ , and verify that it integrates to 1. [4]
- 
- 28** The lifetime  $T$ , in hours, of a certain type of bulb is exponentially distributed with mean 40.
- (a) Find the probability that a bulb lasts more than 60 hours. [2]
- (b) Five bulbs are tested independently. Find the probability that at least 2 last more than 60 hours. [3]
- (c) Find the least number  $n$  of bulbs which must be tested so that the probability that at least one lasts more than 60 hours exceeds 0.99. [3]
- (d) A bulb has already lasted 30 hours. Find the probability that it fails before it has lasted 60 hours. [2]
- (e) Show that the median lifetime is less than the mean lifetime. [2]

- 29**  $T \sim \text{Exp}(\lambda)$ , and  $N$  is the integer part of  $T$ , so that  $N = n$  when  $n \leq T < n + 1$ , for  $n = 0, 1, 2, \dots$
- (a) Show that  $P(N = n) = (1 - e^{-\lambda})(e^{-\lambda})^n$ . [3]
- (b) Hence show that  $N$  has a geometric distribution (starting at 0) and write down its common ratio. [2]
- (c) Show that  $E(N) = \frac{e^{-\lambda}}{1 - e^{-\lambda}}$ . [3]
- (d) For  $\lambda = 0.5$  compare  $E(N)$  with  $E(T)$ , and explain why  $E(N) < E(T)$ . [3]
- 
- 30**  $U$  is uniformly distributed on the interval  $(0, 1)$ , and  $T = -\frac{1}{\lambda} \ln U$ , where  $\lambda > 0$ .
- (a) Show that  $T \sim \text{Exp}(\lambda)$ . [4]
- (b) For  $\lambda = 0.25$ , find the values of  $T$  that correspond to  $U = 0.3$  and  $U = 0.8$ . [2]
- (c)  $T_1, \dots, T_n$  are independent with  $T_i \sim \text{Exp}(\lambda)$ , and  $M$  is the smallest of them. Show that  $M \sim \text{Exp}(n\lambda)$ . [3]
- (d) Five machines each have a lifetime in hours with an exponential distribution of mean 100. Find the expected time until the first failure, and the probability that no machine fails during the first 30 hours. [4]

## Combinations of Random Variables

- 1**  $X$  and  $Y$  are independent random variables with  $E(X) = 5$ ,  $E(Y) = 8$ ,  $\text{Var}(X) = 2$  and  $\text{Var}(Y) = 3$ .  
Find  $E(X + Y)$  and  $\text{Var}(X + Y)$ . [2]
- 
- 2**  $X$  and  $Y$  are independent random variables with  $E(X) = 4$ ,  $E(Y) = 6$ ,  $\text{Var}(X) = 1$  and  $\text{Var}(Y) = 2$ .  
Find the mean and variance of  $3X - 2Y$ . [3]
- 
- 3**  $X \sim \text{Po}(4)$  and  $Y \sim B(10, 0.3)$  are independent.  
Find  $E(X + Y)$ ,  $\text{Var}(X + Y)$  and  $\text{Var}(X - Y)$ . [5]
- 
- 4**  $X \sim N(5, 2^2)$  and  $Y \sim N(3, 1^2)$  are independent. State the distributions of  $X + Y$  and  $X - Y$ . [4]
- 
- 5**  $X$  and  $Y$  are independent with  $\text{Var}(X) = 9$  and  $\text{Var}(Y) = 16$ . Find  $\text{Var}(X - Y)$  and the standard deviation of  $X - Y$ . [3]
- 
- 6**  $X_1, X_2, X_3, X_4$  are independent observations of  $X \sim N(20, 3^2)$ . State the distribution of  $S = X_1 + X_2 + X_3 + X_4$ . [2]
- 
- 7**  $X \sim N(10, 4)$  and  $Y \sim N(14, 9)$  are independent. The diagram shows the distribution of  $S = X + Y$ .  
Find  $P(X + Y > 30)$ .



[4]

- 8** The times, in minutes, taken by two runners over a fixed route are independent and normally distributed: Anna's time  $X \sim N(32, 4)$  and Ben's time  $Y \sim N(36, 9)$ . Find the probability that Anna runs the route faster than Ben. [4]
- 9**  $X \sim N(6, 5)$  and  $Y \sim N(10, 2)$  are independent, and  $W = 3X - 2Y$ . Find the distribution of  $W$  and hence find  $P(W > 5)$ . [5]
- 10** Three bags of flour, each of mass  $N(2.5, 0.1^2)$  kg, are placed independently in a box of mass  $N(0.4, 0.05^2)$  kg. Find the probability that the total mass exceeds 8 kg. [5]
- 11** The time  $X$  minutes taken to complete a task is  $N(10, 2^2)$ .  
 (a) Find the probability that one task takes more than 12.5 minutes. [2]  
 (b) Two tasks are completed, with times  $X_1$  and  $X_2$ , independently. Find the probability that their total time exceeds 25 minutes. [3]  
 (c) Find the probability that  $2X$  exceeds 25 minutes, and explain why this differs from (b). [3]
- 12**  $X \sim \text{Po}(3)$  and  $Y \sim \text{Po}(5)$  are independent. Find  $P(X + Y = 7)$ ,  $E(X - Y)$  and  $\text{Var}(X - Y)$ . [4]
- 13**  $X \sim N(40, 6^2)$ , and  $X_1$  and  $X_2$  are independent observations of  $X$ . Find the probability that  $X_2$  exceeds  $X_1$  by at least 8. [4]
- 14** Each chocolate bar has mass, in grams, normally distributed with mean 45 and standard deviation 2. A gift box has mass  $N(60, 3^2)$  when empty. A box is filled with 4 bars chosen at random. All masses are independent.  
 (a) Find the distribution of the total mass of the filled box. [3]  
 (b) Find the probability that the filled box has mass more than 250 g. [2]  
 (c) A student suggests using  $\text{Var} = 4^2 \times 2^2$  for the four bars. Explain why this is wrong. [2]
- 15** The random variable  $X$  has the distribution shown.  $X_1$  and  $X_2$  are independent observations of  $X$ , and  $S = X_1 + X_2$ .
- |            |          |          |          |
|------------|----------|----------|----------|
| $x$        | <b>1</b> | <b>2</b> | <b>3</b> |
| $P(X = x)$ | 0.2      | 0.5      | 0.3      |
- (a) Find  $E(X)$  and  $\text{Var}(X)$ . [3]  
 (b) Find the probability distribution of  $S$ . [3]  
 (c) Find  $\text{Var}(X_1 - X_2)$  and compare it with  $\text{Var}(2X)$ . [1]
- 16**  $X \sim N(20, \sigma^2)$  and  $Y \sim \text{Po}(4)$  are independent random variables, and  $W = 2X - 5Y + 9$ . It is given that  $\text{Var}(W) = 164$ .  
 (a) Find the value of  $\sigma$ . [3]  
 (b) Find  $E(W)$  and hence find  $P(X < E(W))$ . [4]
- 17** Apples have independent masses  $N(120, 10^2)$  g and oranges have independent masses  $N(180, 15^2)$  g. A bag contains 6 apples, another contains 4 oranges. Show that the two bags have the same expected mass, and find the probability that their masses differ by less than 30 g. [6]
- 18** A bolt has diameter  $D \sim N(9.95, 0.02^2)$  mm and the hole it must fit through has diameter  $H \sim N(10.00, 0.03^2)$  mm.  $D$  and  $H$  are independent. The bolt fits if  $H > D$ .  
 (a) Find the probability that a randomly chosen bolt fits a randomly chosen hole. [4]  
 (b) In 10 independent assemblies, find the probability that at least 9 fit. [3]

- 19** Method A gives results  $X \sim N(50, 8^2)$  and method B gives results  $Y \sim N(48, 6^2)$ . A mean  $\bar{X}$  is found from 10 independent results of A and a mean  $\bar{Y}$  from 15 independent results of B. Find the probability that  $\bar{X}$  is less than  $\bar{Y}$ . [6]
- 
- 20** X and Y are independent random variables with  $E(X) = 3$ ,  $\text{Var}(X) = 4$ ,  $E(Y) = 2$  and  $\text{Var}(Y) = 1$ . Find  $E(XY)$ ,  $E[(X + Y)^2]$  and  $\text{Var}(XY)$ . [7]
- 
- 21** The mass of sugar in a bag, in grams, is normally distributed with mean 1000 and standard deviation 8. Five bags are chosen at random. Let  $\bar{X}$  be the mean mass of sugar in the five bags and T the total mass.
- (a) Find the distribution of  $\bar{X}$ . [3]
- (b) Find  $P(\bar{X} < 995)$ . [2]
- (c) Find  $P(T > 5030)$ . [3]
- 
- 22** A lift is designed to carry a maximum load of 600 kg. The masses of passengers are independent and normally distributed with mean 72 kg and standard deviation 10 kg.
- (a) Eight passengers are in the lift. Find the probability that the load exceeds 600 kg. [4]
- (b) A colleague calculates the variance of the total as  $8^2 \times 10^2 = 6400$ . Explain the mistake. [2]
- (c) Find the greatest number of passengers for which the probability of exceeding 600 kg is below 0.05. [3]
- 
- 23** A quantity is measured with results normally distributed with unknown mean  $\mu$  and standard deviation 6. The mean  $\bar{X}$  of  $n$  independent measurements is used to estimate  $\mu$ .
- (a) State the distribution of  $\bar{X}$ . [2]
- (b) Find the smallest  $n$  such that  $P(|\bar{X} - \mu| < 2) \geq 0.95$ . [4]
- 
- 24** A taxi fare is £F, where  $F = 3 + 1.5D$  and  $D$  is the distance in km, with  $D \sim N(8, 2^2)$ .
- (a) State the distribution of  $F$ . [2]
- (b) Find the probability that the total of 5 independent fares exceeds £80. [4]
- (c) Two passengers take independent journeys. Find the probability that the first pays more than £6 more than the second. [4]
- 
- 25**  $X \sim B(50, 0.6)$  and  $Y \sim \text{Po}(16)$  are independent random variables.
- (a) Explain why  $X - Y$  can be approximated by a normal distribution, and state its mean and variance. [4]
- (b) Use your approximation, with a continuity correction, to estimate  $P(X - Y \geq 18)$ . [4]
- 
- 26**  $X \sim N(30, 4^2)$  and  $Y \sim N(95, 5^2)$ .  $X_1, X_2, X_3$  are independent observations of  $X$ , and  $Y$  is independent of them.
- (a) Find  $P(X_1 + X_2 + X_3 < Y)$ . [4]
- (b) The mean of  $Y$  is changed to  $\mu_Y$  (variance unchanged). Find the value of  $\mu_Y$  for which  $P(X_1 + X_2 + X_3 < Y) = 0.1$ . [4]
- 
- 27**  $X \sim N(20, 5^2)$  and  $Y \sim N(\mu, 3^2)$  are independent random variables.
- (a) State the distribution of  $D = X - Y$ . [2]
- (b) Given that  $P(X > Y) = 0.25$ , find  $\mu$ . [4]
- (c) Using this value of  $\mu$ , find  $k$  such that  $P(X + Y > k) = 0.1$ . [3]
- (d)  $X_1, X_2, X_3$  are independent observations of  $X$ . Find  $\text{Var}(3X) - \text{Var}(X_1 + X_2 + X_3)$  and interpret the result. [2]
-

- 28**  $X_1, X_2, \dots, X_n$  are independent observations of  $X \sim N(\mu, \sigma^2)$ , and  $\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$  is the sample mean.
- (a) Show that  $\text{Var}(\bar{X}) = \frac{\sigma^2}{n}$ . [2]
- (b) Show that  $\text{Var}(X_1 - \bar{X}) = \frac{(n-1)\sigma^2}{n}$ . [4]
- (c) For  $n = 5$  and  $\sigma = 4$ , find  $P(|X_1 - \bar{X}| < 3)$ . [3]
- (d) A student says  $\text{Var}(X_1 - \bar{X}) = \text{Var}(X_1) + \text{Var}(\bar{X}) = \sigma^2(1 + \frac{1}{n})$ . Explain why this is wrong. [2]

- 29** The joint probability distribution of  $X$  and  $Y$ , each taking the values 0 and 1, is shown in the table.

|         | $Y = 0$ | $Y = 1$ |
|---------|---------|---------|
| $X = 0$ | 0.1     | 0.2     |
| $X = 1$ | 0.3     | 0.4     |

- (a) Find the distributions of  $X$  and  $Y$  separately and show that  $X$  and  $Y$  are not independent. [4]
- (b) Find  $E(X)$ ,  $E(Y)$ ,  $\text{Var}(X)$ ,  $\text{Var}(Y)$  and  $E(XY)$ . [4]
- (c) Find  $E(XY) - E(X)E(Y)$  and hence  $\text{Var}(X + Y)$ , using  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2[E(XY) - E(X)E(Y)]$ . [4]
- (d) Verify  $\text{Var}(X + Y)$  from the distribution of  $X + Y$ , and say why it differs from  $\text{Var}(X) + \text{Var}(Y)$ . [3]

- 30** A machine makes rods whose lengths are normally distributed with mean 100 mm and standard deviation 4 mm when working correctly. Every hour a sample of 5 rods is measured and the sample mean  $\bar{X}$  is calculated. An alarm is raised if  $\bar{X}$  lies outside the limits  $100 \pm 3$  standard errors.
- (a) State the distribution of  $\bar{X}$  when the machine is correct and find the limits. [4]
- (b) Find the probability of a false alarm in one hour. [2]
- (c) The machine drifts so that the mean becomes 102 mm (standard deviation unchanged). Find the probability that the alarm is raised in a given hour. [5]
- (d) After the drift, the number of hours until the first alarm has a geometric distribution. Find the expected number of hours and the probability that the alarm is raised within the first 3 hours. [4]