

1. A company has a large number of regular users logging onto its website. On average 4 users every hour fail to connect to the company's website at their first attempt.

(a) Explain why the Poisson distribution may be a suitable model in this case.

(1)

Find the probability that, in a randomly chosen **2 hour** period,

(b) (i) all users connect at their first attempt,

(ii) at least 4 users fail to connect at their first attempt.

(5)

The company suffered from a virus infecting its computer system. During this infection it was found that the number of users failing to connect at their first attempt, over a 12 hour period, was 60.

(c) Using a suitable approximation, test whether or not the mean number of users per hour who failed to connect at their first attempt had increased. Use a 5% level of significance and state your hypotheses clearly.

(9)

(Total 15 marks)

2. A café serves breakfast every morning. Customers arrive for breakfast at random at a rate of 1 every 6 minutes.

Find the probability that

(a) fewer than 9 customers arrive for breakfast on a Monday morning between 10 am and 11 am.

(3)

The café serves breakfast every day between 8 am and 12 noon.

(b) Using a suitable approximation, estimate the probability that more than 50 customers arrive for breakfast next Tuesday.

(6)

(Total 9 marks)

3. An administrator makes errors in her typing randomly at a rate of 3 errors every 1000 words.
- (a) In a document of 2000 words find the probability that the administrator makes 4 or more errors.

(3)

The administrator is given an 8000 word report to type and she is told that the report will only be accepted if there are 20 or fewer errors.

- (b) Use a suitable approximation to calculate the probability that the report is accepted.

(7)

**(Total 10 marks)**

4. A web server is visited on weekdays, at a rate of 7 visits per minute. In a random one minute on a Saturday the web server is visited 10 times.

- (a) (i) Test, at the 10% level of significance, whether or not there is evidence that the rate of visits is greater on a Saturday than on weekdays. State your hypotheses clearly.
- (ii) State the minimum number of visits required to obtain a significant result.

(7)

- (b) State an assumption that has been made about the visits to the server.

(1)

In a random two minute period on a Saturday the web server is visited 20 times.

- (c) Using a suitable approximation, test at the 10% level of significance, whether or not the rate of visits is greater on a Saturday.

(6)

**(Total 14 marks)**

5. (a) State the condition under which the normal distribution may be used as an approximation to the Poisson distribution.

(1)

- (b) Explain why a continuity correction must be incorporated when using the normal distribution as an approximation to the Poisson distribution. (1)

A company has yachts that can only be hired for a week at a time. All hiring starts on a Saturday.

During the winter the mean number of yachts hired per week is 5.

- (c) Calculate the probability that fewer than 3 yachts are hired on a particular Saturday in winter. (2)

During the summer the mean number of yachts hired per week increases to 25.

The company has only 30 yachts for hire.

- (d) Using a suitable approximation find the probability that the demand for yachts cannot be met on a particular Saturday in the summer. (6)

In the summer there are 16 Saturdays on which a yacht can be hired.

- (e) Estimate the number of Saturdays in the summer that the company will not be able to meet the demand for yachts. (2)
- (Total 12 marks)**

6. An estate agent sells properties at a mean rate of 7 per week.

- (a) Suggest a suitable model to represent the number of properties sold in a randomly chosen week. Give two reasons to support your model. (3)

- (b) Find the probability that in any randomly chosen week the estate agent sells exactly 5 properties. (2)

- (c) Using a suitable approximation find the probability that during a 24 week period the estate agent sells more than 181 properties. (6)
- (Total 11 marks)**

7. A manufacturer produces large quantities of coloured mugs. It is known from previous records that 6% of the production will be green.

A random sample of 10 mugs was taken from the production line.

- (a) Define a suitable distribution to model the number of green mugs in this sample. (1)
- (b) Find the probability that there were exactly 3 green mugs in the sample. (3)

A random sample of 125 mugs was taken.

- (c) Find the probability that there were between 10 and 13 (inclusive) green mugs in this sample, using
- (i) a Poisson approximation, (3)
- (ii) a Normal approximation. (6)
- (Total 13 marks)**

8. A teacher thinks that 20% of the pupils in a school read the Deano comic regularly.

He chooses 20 pupils at random and finds 9 of them read Deano.

- (a) (i) Test, at the 5% level of significance, whether or not there is evidence that the percentage of pupils that read Deano is different from 20%. State your hypotheses clearly.
- (ii) State all the possible numbers of pupils that read Deano from a sample of size 20 that will make the test in part (a)(i) significant at the 5% level. (9)

The teacher takes another 4 random samples of size 20 and they contain 1, 3, 1 and 4 pupils that read Deano.

- (b) By combining all 5 samples and using a suitable approximation test, at the 5% level of significance, whether or not this provides evidence that the percentage of pupils in the school that read Deano is different from 20%. (8)

- (c) Comment on your results for the tests in part (a) and part (b).

(2)

(Total 19 marks)

9. The random variable  $X$  is the number of misprints per page in the first draft of a novel.

- (a) State two conditions under which a Poisson distribution is a suitable model for  $X$ .

(2)

The number of misprints per page has a Poisson distribution with mean 2.5. Find the probability that

- (b) a randomly chosen page has no misprints,

(2)

- (c) the total number of misprints on 2 randomly chosen pages is more than 7.

(3)

The first chapter contains 20 pages.

- (d) Using a suitable approximation find, to 2 decimal places, the probability that the chapter will contain less than 40 misprints.

(7)

(Total 14 marks)

10. The random variables  $R$ ,  $S$  and  $T$  are distributed as follows

$$R \sim B(15, 0.3), \quad S \sim \text{Po}(7.5), \quad T \sim N(8, 2^2).$$

Find

- (a)  $P(R = 5)$ ,

(2)

- (b)  $P(S = 5)$ ,

(1)

(c)  $P(T = 5)$ .

(1)

(Total 4 marks)

11. In an experiment, there are 250 trials and each trial results in a success or a failure.

(a) Write down two other conditions needed to make this into a binomial experiment.

(2)

It is claimed that 10% of students can tell the difference between two brands of baked beans. In a random sample of 250 students, 40 of them were able to distinguish the difference between the two brands.

(b) Using a normal approximation, test at the 1% level of significance whether or not the claim is justified. Use a one-tailed test.

(6)

(c) Comment on the acceptability of the assumptions you needed to carry out the test.

(2)

(Total 10 marks)

12. Minor defects occur in a particular make of carpet at a mean rate of 0.05 per  $\text{m}^2$ .

(a) Suggest a suitable model for the distribution of the number of defects in this make of carpet. Give a reason for your answer.

(3)

A carpet fitter has a contract to fit this carpet in a small hotel. The hotel foyer requires  $30 \text{ m}^2$  of this carpet. Find the probability that the foyer carpet contains

(b) exactly 2 defects,

(3)

(c) more than 5 defects.

(3)

The carpet fitter orders a total of  $355 \text{ m}^2$  of the carpet for the whole hotel.

- (d) Using a suitable approximation, find the probability that this total area of carpet contains 22 or more defects.

(6)

(Total 14 marks)

13. The random variable  $R$  has the binomial distribution  $B(12, 0.35)$ .

- (a) Find  $P(R \geq 4)$ .

(2)

The random variable  $S$  has the Poisson distribution with mean 2.71.

- (b) Find  $P(S \leq 1)$ .

(3)

The random variable  $T$  has the normal distribution  $N(25, 5^2)$ .

- (c) Find  $P(T \leq 18)$ .

(2)

(Total 7 marks)

14. (a) Write down the condition needed to approximate a Poisson distribution by a Normal distribution.

(1)

The random variable  $Y \sim \text{Po}(30)$ .

- (b) Estimate  $P(Y > 28)$ .

(6)

(Total 7 marks)

The random variable  $T$  has the normal distribution  $N(25, 5^2)$ .

- (c) Find  $P(T \leq 18)$ .

(2)

(Total 7 marks)

15. A farmer noticed that some of the eggs laid by his hens had double yolks. He estimated the probability of this happening to be 0.05. Eggs are packed in boxes of 12.

Find the probability that in a box, the number of eggs with double yolks will be

- (a) exactly one,

(3)

- (b) more than three.

(2)

A customer bought three boxes.

- (c) Find the probability that only 2 of the boxes contained exactly 1 egg with a double yolk.

(3)

The farmer delivered 10 boxes to a local shop.

- (d) Using a suitable approximation, find the probability that the delivery contained at least 9 eggs with double yolks.

(4)

The weight of an individual egg can be modelled by a normal distribution with mean 65 g and standard deviation 2.4 g.

- (e) Find the probability that a randomly chosen egg weighs more than 68 g.

(3)

(Total 15 marks)

16. An athletics teacher has kept careful records over the past 20 years of results from school sports days. There are always 10 competitors in the javelin competition. Each competitor is allowed 3 attempts and the teacher has a record of the distances thrown by each competitor at each attempt. The random variable  $D$  represents the greatest distance thrown by each competitor and the random variable  $A$  represents the number of the attempt in which the competitor achieved their greatest distance.

- (a) State which of the two random variables  $D$  or  $A$  is continuous.

(1)

A new athletics coach wishes to take a random sample of the records of 36 javelin competitors.



- (b) Specify a suitable sampling frame and explain how such a sample could be taken. (2)

The coach assumes that  $P(A = 2) = \frac{1}{3}$ , and is therefore surprised to find that 20 of the 36 competitors in the sample achieved their greatest distance on their second attempt.

Using a suitable approximation, and assuming that  $P(A = 2) = \frac{1}{3}$ ,

- (c) find the probability that at least 20 of the competitors achieved their greatest distance on their second attempt. (6)

- (d) Comment on the assumption that  $P(A = 2) = \frac{1}{3}$ . (2)
- (Total 11 marks)**

17. On a typical weekday morning customers arrive at a village post office independently and at a rate of 3 per 10 minute period.

Find the probability that

- (a) at least 4 customers arrive in the next 10 minutes, (2)

- (b) no more than 7 customers arrive between 11.00 a.m. and 11.30 a.m. (3)

The period from 11.00 a.m. to 11.30 a.m. next Tuesday morning will be divided into 6 periods of 5 minutes each.

- (c) Find the probability that no customers arrive in at most one of these periods. (6)

The post office is open for  $3\frac{1}{2}$  hours on Wednesday mornings.

- (d) Using a suitable approximation, estimate the probability that more than 49 customers arrive at the post office next Wednesday morning. (7)
- (Total 18 marks)**

1. (a) Connecting occurs at random/independently, singly or at a constant rate B1 1

**Note**

**B1** Any one of randomly/independently/singly/constant rate. Must have context of connection/logging on/fail

- (b) Po (8) B1

**Note**

**B1** Writing or using Po(8) in (i) or (ii)

- (i)  $P(X = 0) = 0.0003$  M1 A1

**Note**

**M1** for writing or finding  $P(X = 0)$

**A1** awrt 0.0003

- (ii)  $P(X \geq 4) = 1 - P(X \leq 3)$  M1  
 $= 1 - 0.0424$   
 $= 0.9576$  A1 5

**Note**

**M1** for writing or finding  $1 - P(X \leq 3)$

**A1** awrt 0.958

- (c)  $H_0 : \lambda = 4$  (48)  $H_1 : \lambda > 4$  (48) B1  
 $N(48, 48)$  M1 A1

Method 1

Method 2

$$P(X \geq 59.5) = P\left(Z \geq \frac{59.5 - 48}{\sqrt{48}}\right) \quad \frac{x - 0.5 - 48}{\sqrt{48}} = 1.6449 \quad \text{M1 M1 A1}$$

$$= P(Z \geq 1.66)$$

$$= 1 - 0.9515$$

$$= 0.0485$$

$$x = 59.9 \quad \text{A1}$$

$$0.0485 < 0.05$$

Reject  $H_0$ . Significant. 60 lies in the Critical region M1

The number of failed connections at the first attempt has increased. A1 ft 9

**Note**

**B1** both hypotheses correct. Must use  $\lambda$  or  $\mu$

**M1** identifying normal

**A1** using or seeing mean and variance of 48

These first two marks may be given if the following are seen in the standardisation formula : 48 and  $\sqrt{48}$  or awrt 6.93

**M1** for attempting a continuity correction (Method 1:  $60 \pm 0.5$  / Method 2:  $x \pm 0.5$ )

**M1** for standardising using their mean and their standard deviation and using either Method 1 [59.5, 60 or 60.5. accept  $\pm z$ .] Method 2 [( $x \pm 0.5$ ) and equal to a  $\pm z$  value)

**A1** correct  $z$  value awrt  $\pm 1.66$  or  $\pm \frac{59.5-48}{\sqrt{48}}$ , or  $\frac{x-0.5-48}{\sqrt{48}}=1.6449$

**A1** awrt 3 sig fig in range 0.0484 – 0.0485, awrt 59.9

**M1** for “reject  $H_0$ ” or “significant” maybe implied by “correct contextual comment”

If one tail hypotheses given follow through “their prob” and  $0.05, p < 0.5$

If two tail hypotheses given follow through “their prob” with  $0.025, p < 0.5$

If one tail hypotheses given follow through “their prob” and  $0.95, p > 0.5$

If two tail hypotheses given follow through “their prob” with  $0.975, p > 0.5$

If no  $H_1$  given they get M0

**A1** fit correct contextual statement followed through from their prob and  $H_1$ . need the words number of failed connections/log ons has increased o.e.

Allow “there are more failed connections”

NB A correct contextual statement alone followed through from their prob and  $H_1$  gets M1 A1

[15]

- |    |     |                          |         |
|----|-----|--------------------------|---------|
| 2. | (a) | $X \sim \text{Po}(10)$   | B1      |
|    |     | $P(X < 9) = P(X \leq 8)$ | M1      |
|    |     | $= 0.3328$               | A1    3 |

**Note**

**B1** for using  $\text{Po}(10)$

**M1** for attempting to find  $P(X \leq 8)$ : useful values

$P(X \leq 9)$  is 0.4579(M0), using  $\text{Po}(6)$  gives 0.8472, (M1).

**A1** awrt 0.333 but do not accept  $\frac{1}{3}$

(b) $Y \sim \text{Po}(40)$		M1 A1	
$Y$ is approximately $N(40,40)$			
$P(Y > 50)$	$= 1 - P(Y \leq 50)$	M1	
	$= 1 - P\left(Z < \frac{50.5 - 40}{\sqrt{40}}\right)$	M1	
	$= 1 - P(Z < 1.660..)$	A1	
	$= 1 - 0.9515$		
	$= 0.0485$	A1	6

N.B. Calculator gives 0.048437.

Poisson gives 0.0526 (but scores nothing)

### Note

**1<sup>st</sup> M1** for identifying the normal approximation

**1<sup>st</sup> A1** for [mean = 40] and [sd =  $\sqrt{40}$  or var = 40]

**NB These two marks are B1 M1 on ePEN**

These first two marks may be given if the following are seen in the standardisation formula : 40 and  $\sqrt{40}$  or awrt 6.32

**2<sup>nd</sup> M1** for attempting a continuity correction (50 or 30  $\pm$  0.5 is acceptable)

**3<sup>rd</sup> M1** for standardising using their mean and their standard deviation and using either 49.5, 50 or 50.5. (29.5, 30, 30.5) accept  $\pm$

**2<sup>nd</sup> A1** correct z value awrt  $\pm 1.66$  or this may be awarded

if see  $\pm \frac{50.5 - 40}{\sqrt{40}}$  or  $\pm \frac{29.5 - 40}{\sqrt{40}}$

**3<sup>rd</sup> A1** awrt 3 sig fig in range 0.0484 – 0.0485

[9]

3. (a)  $X$  = the number of errors in 2000 words so  $X \sim \text{Po}(6)$  B1  
 $P(X \geq 4) = 1 - P(X \leq 3)$  M1  
 $= 1 - 0.1512 = 0.8488$  awrt 0.849 A1 3

**Note**

B1 for seeing or using Po(6)

M1 for  $1 - P(X \leq 3)$  or  $1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3)]$ 

A1 awrt 0.849

**SC**

If B(2000, 0.003) is used and leads to awrt 0.849 allow B0 M1 A1

If no distribution indicated awrt 0.8488 scores B1M1A1 but any other awrt 0.849 scores B0M1A1

- (b)  $Y$  = the number of errors in 8000 words. M1  
 $Y \sim \text{Po}(24)$  so use a Normal approx
- $$Y \approx N(24, \sqrt{24}^2)$$
- A1
- Require  $P(Y \leq 20) = P\left(Z < \frac{20.5 - 24}{\sqrt{24}}\right)$  M1
- $$= P(Z < -0.714\dots)$$
- M1
- 
- $$= 1 - 0.7611$$
- A1
- 
- $$= 0.2389$$
- M1
- 
- awrt (0.237~0.239) A1 7

[N.B. Exact Po gives 0.242 and no  $\pm 0.5$  gives 0.207]



- (c) [ $H_0 : \lambda = 7$      $H_1 : \lambda > 7$     (or  $H_0 : \lambda = 14$      $H_1 : \lambda > 14$ )]  
 $X \sim N(14, 14)$  B1; B1
- $P(X \geq 20) = P\left(z \geq \frac{19.5 - 14}{\sqrt{14}}\right)$  +/- 0.5, stand    M1 M1
- $= P(z \geq 1.47)$   
 $= 0.0708$     or  $z = 1.2816$  A1 dep both  
M
- 0.0708 < 0.10 therefore significant. The rate of visits is greater on a Saturday A1 dep 2<sup>nd</sup> M    6

[14]

5. (a)  $\lambda > 10$  or large  $\mu$  ok    B1    1
- (b) The Poisson is discrete and the normal is continuous. B1    1
- (c) Let  $Y$  represent the number of yachts hired in winter  
 $P(Y < 3) = P(Y \leq 2)$  P( $Y \leq 2$ ) & Po(5)    M1  
 $= 0.1247$  awrt 0.125    A1    2
- (d) Let  $X$  represent the number of yachts hired in summer  $X \sim \text{Po}(25)$ .  
 $N(25, 25)$  all correct, can be implied by standardisation below B1  
 $P(X > 30) \approx P\left(Z > \frac{30.5 - 25}{5}\right) \pm$  standardise with 25 & 5;  $\pm 0.5$  c.c. M1; M1  
 $\approx P(Z > 1.1)$  1.1    A1  
 $\approx 1 - 0.8643$  'one minus'    M1  
 $\approx 0.1357$  awrt 0.136    A1    6
- (e) no. of weeks =  $0.1357 \times 16$  ANS (d)  $\times 16$     M1  
 $= 2.17$  or 2 or 3 ans > 16 M0A0    A1ft    2

[12]

6. Let  $X$  represent the number of properties sold in a week
- (a)  $\therefore X \sim P_0(7)$  B1  
*must be in part a*
- Sales occur independently / randomly, singly, at a constant rate B1 B1    3  
*context needed once*

(b)  $P(X = 5) = P(X \leq 5) - P(X \leq 4)$  or  $\frac{7^5 e^{-7}}{5!}$  M1  
 $= 0.3007 - 0.1730$   
 $= 0.1277$  A1 2  
*awrt 0.128*

(c)  $P(X > 181) \approx P(Y \geq 181.5)$  where  $Y \sim N(168, 168)$  N(168, 168) B1  
 $= P\left(z \geq \frac{181.5 - 168}{\sqrt{168}}\right)$   
 $\pm 0.5$  M1  
*stand with  $\mu$  and  $\sigma$*  M1  
*Give A1 for 1.04 or correct expression* A1  
 $= P(z \geq 1.04)$   
 $= 1 - 0.8508$  M1  
*attempt correct area*  
*1 - p where  $p > 0.5$*   
 $= 0.1492$  A1 6  
*awrt 0.149*

[11]

7. (a) Binomial B1 1  
 Let  $X$  represent the number of green mugs in a sample

(b)  $X \sim B(10, 0.06)$  B1  
*may be implied or seen in part a*  
 $P(X = 3) = {}^{10}C_3(0.06)^3(0.94)^7$  M1  
 ${}^{10}C_3(p)^3(1-p)^7$   
 $0.016808\dots$  A1 3  
*awrt 0.0168*



- (c) Let  $X$  represent number of green mugs in a sample of size 125
- (i)  $X \sim P_0(125 \times 0.06 = 7.5)$  may be implied B1  
 $(10 \leq X \leq 13) = P(X \leq 13) - P(X \leq 9)$  M1  
 $0.9784 - 0.7764$   
 $= 0.2020$  A1 3  
*awrt 0.202*
- (ii)  $P(10 \leq X \leq 13) \approx P(9.5 \leq Y \leq 13.5)$  where  $Y \sim N(7.5, 7.05)$  B1  
 $7.05$  B1  
 $9.5, 13.5$   
 $= P\left(\frac{9.5-7.5}{\sqrt{7.05}} \leq z \leq \frac{13.5-7.5}{\sqrt{7.05}}\right)$   
 $\pm 0.5$  M1  
*stand.* M1  
*both values or both correct expressions.*  
 $= P(0.75 \leq z \leq 2.26)$  A1  
*awrt 0.75 and 2.26*  
 $= 0.2147$  A1 6  
*awrt 0.214 or 0.215*

[13]

8. (a) (i) Two tail B1 B1  
 $H_0: p = 0.2, H_1: p \neq 0.2$   $p =$   
 $P(X \geq 9) = 1 - P(X \leq 8)$  or attempt critical value/region  
 $= 1 - 0.9900 = 0.01$  CR  $X \geq 9$  A1  
 $0.01 < 0.025$  or  $9 \geq 9$  or  $0.99 > 0.975$  or  $0.02 < 0.05$  or lies in interval with correct interval stated.  
 Evidence that the percentage of pupils that read Deano is not 20% A1
- (ii)  $X \sim \text{Bin}(20, 0.2)$  may be implied or seen in (i) or (ii) B1  
 So 0 or [9,20] make test significant.  
 0,9, between “their 9” and 20 B1 B1 B1 9

- (b)  $H_0 : p = 0.2, H_1 : p \neq 0.2$  B1  
 $W \sim \text{Bin}(100, 0.2)$   
 $W \sim N(20, 16)$  normal; 20 and 16 B1; B1
- $P(X \leq 18) = P(Z \leq \frac{18.5 - 20}{4})$  or  $\frac{x(+\frac{1}{2}) - 20}{4} = \pm 1.96$   
 $\pm$  cc, standardise or use z value, standardise M1 M1 A1  
 $= P(Z \leq -0.375)$   
 $= 0.352 - 0.354$  CR  $X < 12.16$  or  $11.66$  for  $\frac{1}{2}$  A1  
 [0.352 > 0.025 or 18 > 12.16 therefore insufficient evidence to reject  $H_0$   
 Combined numbers of Deano readers suggests 20% of pupils  
 read Deano A1 8
- (c) Conclusion that they are different. B1  
 Either large sample size gives better result  
 Or  
 Looks as though they are not all drawn from the same population. B1 2
- [19]**
- (a) (i) One tail  
 $H_0: p = 0.2, H_1 : p \neq 0.2$  B1B1  
 $P(X \geq 9) = 1 - P(X \leq 8)$  or attempt critical value/region M1  
 $= 1 - 0.9900 = 0.01$  CR  $X \geq 8$  A1  
 0.01 < 0.025 or  $9 \geq 9$  or  $0.99 > 0.975$  or  $0.02 < 0.05$  or lies in  
 interval with correct interval stated.  
 Evidence that the percentage of pupils that read  
 Deano is not 20% A1
- (ii)  $X \sim \text{Bin}(20, 0.2)$  may be implied or seen in (i) or (ii) B1  
 So 0 or [9,20] make test significant.  
0,9, between “their 9” and 20 B1 B1 B1 9

(b)	$H_0 : p = 0.2, H_1 : p \neq 0.2$		B1	
	$W \sim \text{Bin}(100, 0.2)$			
	$W \sim N(20, 16)$	normal; 20 and 16	B1; B1	
	$P(X \leq 18) = P(Z \leq \frac{18.5 - 20}{4})$ or $\frac{x - 20}{4} = -1.6449$			
	$\pm$ cc, standardise or standardise, use z value		M1 M1 A1	
	$= P(Z \leq -0.375)$			
	$= 0.3520$	CR $X < 13.4$ or $12.9$	awrt 0.352	A1
	[0.352 > 0.025 or 18 > 12.16 therefore insufficient evidence to reject $H_0$			
	Combined numbers of Deano readers suggests 20% of pupils read Deano			
			A1	8
(c)	Conclusion that they are different.		B1	
	Either large sample size gives better result			
	Or			
	Looks as though they are not all drawn from the same population.		B1	2
				<b>[19]</b>
9.	(a) Misprints are random / independent, occur singly in space and at a constant rate		B1, B1	2
	<i>Context, any 2</i>			
	(b) $P(X = 0) = e^{-2.5}$		M1	
	<i>Po (2.5)</i>			
	$= 0.08208\dots = 0.0821$		A1	2
	(c) $Y \sim \text{Po}(5)$ for 2 pages		B1	
	<i>Implied</i>			
	$P(Y > 7) = 1 - P(X \leq 7)$		M1	
	<i>Use of 1 - and correct inequality</i>			
	$= 1 - 0.8666 = 0.1334$		A1	3

- (d) For 20 pages,  $Y \sim P_0(50)$  B1  
 $Y \sim N(50, 50)$  approx B1  
 $P(Y < 40) = P(Y \leq 39.5)$  M1  
*cc  $\neq 0.5$*   
 $= P\left(Z \leq \frac{39.5 - 50}{\sqrt{50}}\right)$   
*standardise above* M1  
*all correct* A1  
 $= P(Z \leq -1.4849)$  A1  
*awrt - 1.48*  
 $= 1 - 0.93 = 0.07$  A1 7  
*0.07*

**[14]**

10. (a)  $P(R = 5) = P(R \leq 5) - P(R \leq 4) = 0.7216 - 0.5155$  M1  
*Can be implied*  
 $= 0.2061$  A1 2  
*Answer 0.2061*  
 (OR:  ${}^{15}C_5 (0.3)^5 (0.7)^{10} = 0.206130\dots$ )  
 (b)  $P(S = 5) = 0.2414 - 0.1321 = 0.1093$  B1 1  
*Accept 0.1093 (AWRT) or 0.1094 (AWRT)*  
 (OR:  $\frac{7.5^5 e^{-7.5}}{5!} = 0.10937459\dots$ )  
 (c)  $P(T = 5) = 0$  B1 1  
*cao*

**[4]**

11. (a) Probability of success/failure is constant B1  
Trials are independent B1 2

- (b) Let  $p$  represent proportion of students who can distinguish between brands  
 $H_0: p = 0.1; H_1: p > 0.1$  B1  
*both*
- $\alpha = 0.01; CR: \delta > 2.3263$  B1  
 $2.3263$
- $np = 25; npq = 22.5$  B1  
*both*  
*Can be implied*
- $\delta = \frac{39.5 - 25}{\sqrt{22.5}} = 3.0568\dots$  M1 A1  
*Standardisation with  $\pm 0.5$  their  $\sqrt{npq}$*   
*AWRT 3.06*
- Reject  $H_0$ : claim cannot be accepted A1ft 6  
*Based on clear evidence from  $\delta$  or  $p$*
- (c) eg:–  $np, nq$  both 75 – true or acceptable B1  
 $p$  close to 0.5 – not true, assumption not met B1 2  
 success/failure not clear cut necessarily  
 independence – one student influences another

[10]

- (b) Aliter  $\delta = 3.06 \Rightarrow p = 0.9989 > 0.99$   
 or  $p 0.0011 < 0.01$   
*B1 eqn to 2.3263*

12. (a) No of defects in carpet area  $a$  sq m is distributed  $Po(0.05a)$  B1B1  
*Poisson,  $0.05a$*
- Defects occur at a constant rate, independent, singly, randomly B1 3  
*Any 1*
- (b)  $X \sim P(30 \times 0.05) = P(1.5)$  B1  
 $P(X = 2) = \frac{e^{-1.5} \times 1.5^2}{2} = 0.2510$  M1A1 3  
*Tables or calc 0.251(0)*
- (c)  $P(X > 5) = 1 - P(X \leq 5) = 1 - 0.9955 = 0.0045$  M1M1A1 3  
*Strict inequality,  $1 - 0.9955, 0.0045$*

(d)	$X \sim P(17.75)$		B1	
	<i>Implied</i>			
	$X \sim N(17.75, 17.75)$		B1	
	<i>Normal, 17.75</i>			
	$P(X \geq 22) = P\left(Z > \frac{21.5 - 17.75}{\sqrt{17.75}}\right)$		M1M1	
	<i>Standardise, accept 22 or <math>\pm 0.5</math></i>			
	$= -P(Z > 0.89)$		A1	
	<i>awrt 0.89</i>			
	$= 0.1867$		A1	6
				<b>[15]</b>
<b>13.</b>	(a) $P(R \geq 4) = 1 - P(R \leq 3) = 0.6533$		M1A1	2
	<i>Require 1 minus and correct inequality</i>			
	(b) $P(S \leq 1) = P(S = 0) + P(S = 1), e^{-2.71} + 2.71e^{-2.71}, = 0.2469$		M1,A1,A1	3
	<i>awrt 0.247</i>			
	(c) $P(T \leq 18) = P(Z \leq -1.4) = 0.0808$		M1,A1	2
	<i>4 dp, cc no marks</i>			
				<b>[7]</b>
<b>14.</b>	(a) $\lambda$ is large or $\lambda > 10$		B1	1
	(b) $Y \sim N(30, 30)$ may be implied		B1	
	$P(Y > 28) = 1 - P(Y \leq 28.5)$			
	$= 1 - P\left(Z \leq \frac{28.5 - 30}{\sqrt{30}}\right)$		M1 A1	
	$= 1 - P(Z \leq -0.273)$		A1	
	<i>completely correct</i>			
	$= \underline{0.606 - 0.608}$		A1	6
	<i>must be 3 or 4 dp</i>			
				<b>[7]</b>
<b>15.</b>	(a) Let $X$ represent the number of double yolks in a box of eggs		B1	
	$\therefore X \sim B(12, 0.05)$		B1	
	$P(X = 1) = P(X \leq 1) - P(X \leq 0) = 0.8816 - 0.5404 = 0.3412$		M1 A1	3

- (b)  $P(X > 3) = 1 - P(X \leq 3) = 1 - 0.9978 = 0.0022$  M1 A1 2
- (c)  $P(\text{only 2}) = C_2^3 (0.3412)^2 (0.6588)^2$  M1 A1  
 $= 0.230087$  A1 3
- (d) Let  $X$  represent the number of double yolks in 10 dozen eggs  
 $\therefore X \sim B(120, 0.05) \Rightarrow X = \text{Po}(6)$  B1  
 $P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.8472$  M1 A1  
 $= 0.1528$  A1 4
- (e) Let  $X$  represent the weight of an egg  $\therefore W \sim N(65, 2.4^2)$  M1  
 $P(X > 68) = P\left(Z > \frac{68 - 65}{2.4}\right)$  A1  
 $= P(Z > 1.25)$  A1  
 $= 0.1056$  A1 3

[15]

16. (a)  $D$  is continuous B1 1
- (b) Sampling Frame is the list of competitors or their results, B1  
 e.g. label the results 1—200 and randomly select 36 of them B1 2
- (c)  $X = \text{no. of competitors with } A = 2$   $X \sim B(36, \frac{1}{3})$   
 $X \approx \sim N(12, 8)$  M1 A1  
 $P(X \geq 20) \approx P\left(Z \geq \frac{19.5 - 12}{\sqrt{8}}\right)$   $\pm \frac{1}{2}, 'z'$  M1 M1  
 $= P(Z \geq 2.65\dots)$  A1  
 $= 1 - 0.9960 = 0.004$  A1 6
- (d) Probability is very low, so assumption of  $P(A = 2) = \frac{1}{3}$  is unlikely. B1 B1 2  
 (Suggests  $P(A = 2)$  might be higher.)

[11]

17. (a)  $X = \text{no. of customers arriving in 10 minute period}$   
 $X \sim \text{Po}(3)$   $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.6472 = 0.3528$  M1 A1 2
- (b)  $Y = \text{no. of customers in 30 minute period } Y \sim \text{Po}(9)$  B1  
 $P(Y \leq 7) = 0.3239$  M1 A1 3

- (c)  $p = \text{probability of no customers in 5 minute period} = e^{-1.5}$  B1  
 $C = \text{number of 5 minute periods with no customers}$   
 $C \sim B(6, p)$  M1  
 $P(C \leq 1) = (1-p)^6 + 6(1-p)^5 p$  M1, M1 A1  
 $= 0.59866\dots$   
*(accept awrt 0.599)* A16
- (d)  $W = \text{no. of customers on Wednesday morning}$   
 $3\frac{1}{2}$  hours = 210 minutes  $\therefore W \sim \text{Po}(63)$  '63' B1  
Normal approximation  $W \approx \sim N(63, (\sqrt{63})^2)$  M1 A1  
 $P(W > 49) \approx P(W \geq 49.5)$   $\pm \frac{1}{2}$  M1  

$$= P\left(Z \geq \frac{49.5 - 63}{\sqrt{63}}\right)$$
*standardising* M1  
 $= P(Z \geq -1.7008)$  A1  
 $= 0.9554$  (tables)  
*(accept awrt 0.955 or 0.956)* A17

[18]



1. The majority of candidates were familiar with the technical terms in part (a), but failed to establish any context.

Part (b) was a useful source of marks for a large proportion of the candidates. The only problems were occasional errors in detail. In part (i) a few did not spot the change in time scale and used  $Po(4)$  rather than  $Po(8)$ . Some were confused by the wording and calculated  $P(X = 8)$  rather than  $P(X = 0)$ . The main source of error for (ii) was to find  $1 - P(X \leq 4)$  instead of  $1 - P(X \leq 3)$ .

In part (c) the Normal distribution was a well-rehearsed routine for many candidates with many candidates concluding the question with a clear statement in context.

The main errors were

- Some other letter (or none) in place of  $\lambda$  or  $\mu$
- Incorrect Normal distribution: e.g.  $N(60, 60)$
- Omission of (or an incorrect) continuity correction
- Using 48 instead of 60
- Calculation errors

A minority of candidates who used the wrong distribution (usually Poisson) were still able to earn the final two marks in the many cases when clear working was shown. This question was generally well done with many candidates scoring full marks.

2. This question was accessible to the majority of candidates, with many gaining full marks. Most recognised the need to use a Poisson distribution in part (a) and translated the time of one hour successfully to a mean of 10. Common errors included using a mean of 6 or misinterpreting  $P(X < 9)$  as  $P(X \leq 9)$  or using  $1 - P(X \leq 8)$ . In part (b), a high percentage of candidates gained full marks for using a Normal approximation with correct working. Marks lost in this part were mainly due to using a 49.5 instead of 50.5 or no continuity correction at all. A small number of candidates wrote the distribution as  $B(240, 1/6)$  and translated this to  $N(40, 100/3)$ .

3. Part (a) was answered well with the majority of candidates gaining full marks.

Part (b) was also a good source of marks for a large majority of the candidates. Common errors included using 23.9...for variance and 19.5 instead of 20.5. A sizeable minority of candidates used 21.5 after applying the continuity correction. A few candidates had correct working up to the very end when they failed to find the correct probability by not subtracting the tables' probability from 1.

4. In Part (a) there are a sizeable number of candidates who are not using the correct symbols in defining their hypotheses although the majority of candidates recognised  $Po(7)$ .

For candidates who attempted a critical region there were still a number who struggled to define it correctly for a number of reasons:

- Looking at the wrong tail and concluding  $X \leq 3$ .
- Incorrect use of  $>$  sign when concluding 11 - not appreciating that this means  $\geq 12$  for a discrete variable.

- Not knowing how to use probabilities to define the region correctly and concluding 10 or 12 instead of 11.

The candidates who opted to calculate the probability were generally more successful.

A minority still try to calculate a probability to compare with 0.9. This proved to be the most difficult route with the majority of students unable to calculate the probability or critical region correctly. We must once again advise that this is not the recommended way to do this question. There are still a significant number who failed to give an answer in context although fewer than in previous sessions.

Giving the minimum number of visits needed to obtain a significant result proved challenging to some and it was noticeable that many did not use their working from part (a) or see the connection between the answer for (i) and (ii) and there were also number of candidates who did not recognise inconsistencies in their answers.

A number of candidates simply missed answering part (b) but those who did usually scored well.

There were many excellent responses in part (c) with a high proportion of candidates showing competence in using a Normal approximation, finding the mean and variance and realising that a continuity correction was needed. Marks were lost, however, for not including 20, and for not writing the conclusion in context in terms of the **rate** of visits being **greater**. Some candidates attempted to find a critical value for  $X$  using methods from S3 but failing to use 1.2816. There were a number of candidates who calculated  $P(X = 20)$  in error.

5. It was disappointing to find that a large number of candidates failed to attain both of the first 2 marks available. These were often the only marks lost by some, since the majority of candidates achieved most or all marks. In part (d) most candidates did attempt an approximation, although a minority calculated an exact binomial. Again, the common errors were to fail to use a continuity correction and the standard deviation when using the approximation and then not using the  $1 - \Phi(z)$ . The simple calculation of 16 x the answer to part (d) was performed correctly by the majority of candidates attempting this part of the question. A common error was to attempt a binomial probability.
6. The majority of candidates knew the conditions for the Poisson distribution but many did not get the marks because they failed to put them into context. As in many previous series, it was very common for candidates to repeat at least some of these conditions parrot-fashion preceded by “events occur” or “it occurs”. Other common errors were listing randomness and independence as separate reasons and citing the fixed time period and lack of an upper limit as reasons. Quite a number failed to mention the parameter for the distribution. The majority of candidates answered part (b) correctly. Most candidates answered part (c) correctly. Where marks were lost it was usually through failing to use a continuity correction rather than applying it wrongly.
7. This was well answered by almost all candidates and many correct solutions were seen. A few candidates tried to use Poisson rather than Binomial for parts (a) and (b). In part (b) a few candidates used  $B(10, 0.6)$  instead of  $B(10, 0.06)$ . In part (c)(i) most errors occurred because candidates did not understand what was meant by “between 10 and 13 inclusive” The most common wrong answer was in using  $P(10 \leq X \leq 13) = P(X \leq 13) - P(X \leq 10)$  instead of  $P(X \leq 13) - P(X \leq 9)$  Another fairly common error was using  $P(X \leq 13) - (1 - P(X \leq 10))$  Some candidates tried to use a continuity correction in this Poisson approximation. Part (c)(ii) was

often correct the most common errors being to use 7.5 instead of 7.05 for the variance and to use an incorrect continuity correction.

8. In part (a)(i) the null and alternative hypotheses were stated correctly by most candidates but then many had difficulties in either calculating the probability or obtaining the correct critical region and then comparing it to the significance level or given value. Most of those obtaining a result were able to place this in context but not always accurately or fully. Candidates still do not seem to realise that just saying accept or reject the hypothesis is inadequate. In (a)(ii) although some candidates obtained the critical regions the list of values was not always given. Many candidates got the 9 but forgot the 0 and a minority gave a value of  $\geq 9$  but did not give the upper limit.

In part (b) there was a wide variety of errors in the solutions provided including using the incorrect approximation, failing to include the original sample in the calculations, not using a continuity correction and errors in using the normal tables. Again in this part many candidates lost the interpretation mark.

Most candidates attempting part (c) of the question noted that the results for the two hypothesis tests were different but few suggested that either the populations were possibly not the same for the samples or that larger samples are likely to yield better results.

9. Many candidates did not achieve any marks in part (a) as they failed to give conditions in context, events being the most common error seen. Part (b) was done well and if candidates lost a mark then it was usually the final mark due to accuracy. A common answer was 0.082. Part (c) was generally completed satisfactorily, but there were a number of candidates who struggled with the inequalities. A common error was  $P(Y > 7) = 1 - P(T \leq 6)$ . Diagrams were again in evidence; these candidates were perhaps the least likely to make mistakes with the inequalities. Some candidates calculated a probability using  $P(2.5)$  and then squared their answer. The overall response to part (d) was good. However, only a minority scored full marks. Most candidates failed to implement the instruction to write their answer “to 2 decimal places”. There were other errors; omission of the continuity correction or the wrong version (40.5), confusion between variance and standard deviation, and problems dealing with a negative  $z$ -value.
10. Candidates knew how to answer parts (a) and (b) but many did not work to sufficient accuracy. If they used their calculator instead of the tables they were expected to give their answer to the same accuracy as the tables. Too many of them did not read part(c) carefully enough. The random variable  $T$  was defined to be normally distributed and thus  $P(T = 5) = 0$ .
11. Most candidates wrote down two other conditions associated with the binomial experiment but too many did not use ‘trials’ when referring to independence. The alternative hypothesis was often wrongly defined and far too many of those using the normal approximation ignored the need to use the continuity correction. The conclusion needed to be in context but many did not do this. Few candidates made any sensible attempt to answer part (c).

12. This was a well answered question and high marks were frequently being scored. In part (a), many candidates chose the correct Poisson model and gave a correct reason in the light of the problem posed. Many correct solutions were seen in parts (b) and (c). In part (d), many candidates found the correct Normal approximation to the Poisson distribution. The most common error was the incorrect application of the continuity correction with weaker candidates generally losing the final two accuracy marks. In part (a), a small minority of candidates incorrectly chose a Binomial model and applied this model throughout the question, thereby losing a considerable number of marks.
13. The whole question was generally very well answered, by far the most common error was the use of a continuity correction in part (c). A small number of candidates didn't realise that part (b) required the sum of two probabilities.
14. This question was generally well answered with many candidates gaining full marks. In part (b) a minority of candidates standardized correctly but then found the incorrect area.
15. This was the question which most candidates found to their liking and which gave them ample opportunity to score high marks. Most were successful in parts (a) and (b), although some chose to exercise their calculator in part (b) rather than use tables. Part (c) was usually done well, but here again a mark was easily lost by not following the rubric. Part (d) held no terrors for most, but in part (e) too many candidates, including some of the better ones, fell into the continuity correction trap.
16. No Report available for this question.
17. No Report available for this question.