Hypothesis Testing

Questions

Q1.

The number of heaters, H, bought during one day from *Warmup* supermarket can be modelled by a Poisson distribution with mean 0.7

(a) Calculate $P(H \geq 2)$

(1)

The number of heaters, G, bought during one day from *Pumraw*supermarket can be modelled by a Poisson distribution with mean 3, where G and H are independent.

(b) Show that the probability that a total of fewer than 4 heaters are bought from these two supermarkets in a day is 0.494 to 3 decimal places.

(2)

(c) Calculate the probability that a total of fewer than 4 heaters are bought from these two supermarkets on at least 5 out of 6 randomly chosen days.

(3)

December was particularly cold. Two days in December were selected at random and the total number of heaters bought from these two supermarkets was found to be 14

(d) Test whether or not the mean of the total number of heaters bought from these two supermarkets had increased. Use a 5% level of significance and state your hypotheses clearly.

(5)

(Total for question = 11 marks)

Q2.

During the summer, mountain rescue team A receives calls for help randomly with a rate of 0.4 per day.

(a) Find the probability that during the summer, mountain rescue team A receives at least 19 calls for help in 28 randomly selected days.

(2)

The leader of mountain rescue team A randomly selects 250 summer days from the last few years.
She records the number of calls for help received on each of these days.

(b) Using a Poisson approximation, estimate the probability of the leader finding at least 20 of these days when more than 1 call for help was received by mountain rescue team A.

(4)

Mountain rescue team A believes that the number of calls for help per day is lower in the winter than in the summer. The number of calls for help received in 42 randomly selected winter days is 8

(c) Use a suitable test, at the 5% level of significance, to assess whether or not there is evidence that the number of calls for help per day is lower in the winter than in the summer. State your hypotheses clearly.

(4)

During the summer, mountain rescue team B receives calls for help randomly with a rate of 0.2 per day, independently of calls to mountain rescue team A.

The random variable C is the total number of calls for help received by mountain rescue teams A and B during a period of n days in the summer.
On a Monday in the summer, mountain rescue teams A and B each receive a call for help.

Given that over the next n days $P(C = 0) < 0.001$

(d) calculate the minimum value of n

(3)

(e) Write down an assumption that needs to be made for the model to be appropriate.

(1)

(Total for question = 14 marks)

Q3.

Rowan and Alex are both check-in assistants for the same airline.
The number of passengers, R, checked in by Rowan during a 30-minute period can be modelled by a Poisson distribution with mean 28

(a) Calculate $P(R \geq 23)$

(1)

The number of passengers, A, checked in by Alex during a 30-minute period can be modelled by a Poisson distribution with mean 16, where R and A are independent.
A randomly selected 30-minute period is chosen.

(b) Calculate the probability that exactly 42 passengers in total are checked in by Rowan and Alex.

(2)

The company manager is investigating the rate at which passengers are checked in.
He randomly selects 150 non-overlapping 60-minute periods and records the total number of passengers checked in by Rowan and Alex, in each of these 60-minute periods.

(c) Using a Poisson approximation, find the probability that for at least 25 of these 60-minute periods Rowan and Alex check in a total of fewer than 80 passengers.

(4)

On a particular day, Alex complains to the manager that the check-in system is working slower than normal. To see if the complaint is valid the manager takes a random 90-minute period and finds that the total number of people **Rowan** checks in is 67

(d) Test, at the 5% level of significance, whether or not there is evidence that the system is working slower than normal. You should state your hypotheses and conclusion clearly and show your working.

(4)

(Total for question = 11 marks)

Q4.

On a weekday, a garage receives telephone calls randomly, at a mean rate of 1.25 per 10 minutes.

(a) Show that the probability that on a weekday at least 2 calls are received by the garage in a 30-minute period is 0.888 to 3 decimal places.

(2)

(b) Calculate the probability that at least 2 calls are received by the garage in fewer than 4 out of 6 randomly selected, non-overlapping 30-minute periods on a weekday.

(2)

The manager of the garage randomly selects 150 non-overlapping 30-minute periods on weekdays.
She records the number of calls received in each of these 30-minute periods.

(c) Using a Poisson approximation show that the probability of the manager finding at least 3 of these 30-minute periods when exactly 8 calls are received by the garage is 0.664 to 3 significant figures.

(4)

(d) Explain why the Poisson approximation may be reasonable in this case.

(1)

The manager of the garage decides to test whether the number of calls received on a Saturday is different from the number of calls received on a weekday. She selects a Saturday at random and records the number of telephone calls received by the garage in the first 4 hours.

(e) Write down the hypotheses for this test.

(1)

The manager found that there had been 40 telephone calls received by the garage in the first 4 hours.

(f) Carry out the test using a 5% level of significance.

(4)

(Total for question = 14 marks)

Q5.

During the morning, the number of cyclists passing a particular point on a cycle path in a 10-minute interval travelling eastbound can be modelled by a Poisson distribution with mean 8

The number of cyclists passing the same point in a 10-minute interval travelling westbound can be modelled by a Poisson distribution with mean 3

(a) Suggest a model for the total number of cyclists passing the point on the cycle path in a 10-minute interval, stating a necessary assumption.

(2)

Given that exactly 12 cyclists pass the point in a 10-minute interval,

(b) find the probability that at least 11 are travelling eastbound.

(3)

After some roadworks were completed, the total number of cyclists passing the point in a randomly selected 20-minute interval one morning is found to be 14

(c) Test, at the 5% level of significance, whether there is evidence of a decrease in the rate of cyclists passing the point.
State your hypotheses clearly.

(3)

(Total for question = 8 marks)

Q6.

Andreia's secretary makes random errors in his work at an average rate of 1.7 errors every 100 words.

(a) Find the probability that the secretary makes fewer than 2 errors in the next 100-word piece of work.

(2)

Andreia asks the secretary to produce a 250-word article for a magazine.

(b) Find the probability that there are exactly 5 errors in this article.

(2)

Andreia offers the secretary a choice of one of two bonus schemes, based on a random sample of 40 pieces of work each consisting of 100 words.

In scheme **A** the secretary will receive the bonus if more than 10 of the 40 pieces of work contain no errors.

In scheme **B** the bonus is awarded if the total number of errors in all 40 pieces of work is fewer than 56.

(c) Showing your calculations clearly, explain which bonus scheme you would advise the secretary to choose.

(5)

Following the bonus scheme, Andreia randomly selects a single 500-word piece of work from
the secretary to test if there is any evidence that the secretary's rate of errors has decreased.

(d) Stating your hypotheses clearly and using a 5% level of significance, find the critical region for this test.

(4)

(Total for question = 13 marks)

Q7.

A company receives telephone calls at random at a mean rate of 2.5 per hour.

(a) Find the probability that the company receives

(i) at least 4 telephone calls in the next hour,
(ii) exactly 3 telephone calls in the next 15 minutes.

(5)

(b) Find, to the nearest minute, the maximum length of time the telephone can be left unattended so that the probability of missing a telephone call is less than 0.2

(3)

The company puts an advert in the local newspaper. The number of telephone calls received in a randomly selected 2 hour period after the paper is published is 10

(c) Test at the 5% level of significance whether or not the mean rate of telephone calls has increased. State your hypotheses clearly.

(5)

(Total for question = 13 marks)

Q8.

The number of customers entering Jeff's supermarket each morning follows a Poisson distribution.

Past information shows that customers enter at an average rate of 2 every 5 minutes.

Using this information,

(a) (i) find the probability that exactly 26 customers enter Jeff's supermarket during a randomly selected 1-hour period one morning,

(2)

(ii) find the probability that at least 21 customers enter Jeff's supermarket during a randomly selected 1-hour period one morning.

(2)

A rival supermarket is opened nearby. Following its opening, the number of customers entering Jeff's supermarket over a randomly selected 40-minute period is found to be 10

(b) Test, at the 5% significance level, whether or not there is evidence of a decrease in the rate of customers entering Jeff's supermarket. State your hypotheses clearly.

(4)

A further randomly selected 20-minute period is observed and the hypothesis test is repeated.
Given that the true rate of customers entering Jeff's supermarket is now 1 every 5 minutes,

(c) calculate the probability of a Type II error.

(5)

(Total for question = 13 marks)

Q9.

Asha, Davinda and Jerry each have a bag containing a large number of counters, some of which are white and the rest are red.
Each person draws counters from their bag one at a time, notes the colour of the counter and returns it to their bag.

The probability of Asha getting a red counter on any one draw is 0.07

(a) Find the probability that Asha will draw at least 3 white counters before a red counter is drawn.

(2)

(b) Find the probability that Asha gets a red counter for the second time on her 9th draw.

(2)

The probability of Davinda getting a red counter on any one draw is p.
Davinda draws counters until she gets n red counters. The random variable D is the number of counters Davinda draws.

Given that the mean and the standard deviation of D are 4400 and 660 respectively,

(c) find the value of p.

(4)

Jerry believes that his bag contains a smaller proportion of red counters than Asha's bag. To test his belief, Jerry draws counters from his bag until he gets a red counter. Jerry defines the random variable J to be the number of counters drawn up to and including the first red counter.

(d) Stating your hypotheses clearly and using a 10% level of significance, find the critical region for this test.

(5)

Jerry gets a red counter for the first time on his 34th draw.

(e) Giving a reason for your answer, state whether or not there is evidence that Jerry's bag contains a smaller proportion of red counters than Asha's bag.

(2)

Given that the probability of Jerry getting a red counter on any one draw is 0.011

(f) show that the power of the test is 0.702 to 3 significant figures.

(3)

(Total for question = 18 marks)

Q10.

Bacteria are randomly distributed in a river at a rate of 5 per litre of water. A new factory opens and a scientist claims it is polluting the river with bacteria. He takes a sample of 0.5 litres of water from the river near the factory and finds that it contains 7 bacteria. Stating your hypotheses clearly test, at the 5% level of significance, whether there is evidence that the level of pollution has increased.

(5)

(Total for question = 5 marks)

Mark Scheme – Hypothesis Testing

Q1.

Question	Scheme		Marks	AOs
(a)	$\mathrm{P}(H \geq 2) = 0.1558$ awrt **0.156**		B1	1.1b
			(1)	
(b)	$H \sim \mathrm{Po}(0.7)$ $G \sim \mathrm{Po}(3)$			
	$Y = H + G \rightarrow Y \sim \mathrm{Po}(3.7)$		M1	3.4
	$\mathrm{P}(Y \leq 3) = 0.494*$		A1cso*	1.1b
			(2)	
(c)	$K \sim \mathrm{B}(6, 0.494)$		M1	3.3
	$\mathrm{P}(K \geq 5) = 1 - \mathrm{P}(K \leq 4)$		M1	1.1b
	$= 1 - 0.896...$			
	$= 0.1039...$ awrt **0.104**		A1	1.1b
			(3)	
(d)	$\mathrm{H_0}: \lambda = "3.7"$ $\mathrm{H_1}: \lambda > "3.7"$		B1ft	2.5
	$J \sim \mathrm{Po}(7.4)$		B1ft	1.1b
	Method 1	Method 2		
	$\mathrm{P}(J \geq 14) = 1 - \mathrm{P}(J \leq 13)$ $= 1 - 0.9804...$	$\mathrm{P}(J \geq 12) = 0.0735...$ $\mathrm{P}(J \geq 13) = 0.0391...$	M1	1.1b
	$= 0.0195...$	CR $J \geq 13$	A1	1.1b
	$0.0195 < 0.05$ or $14 \geq 13$ or 14 is in the critical region or 14 is significant or Reject $\mathrm{H_0}$. There is evidence at the 5% level of significance that the **number of heaters** brought in total from the two supermarkets has **increased**.		A1	2.2b
			(5)	
			(11 marks)	

Notes

(a) **B1:** awrt 0.156

(b) **M1:** For combining distributions and use of Po(3.7)
A1*cso: $\mathrm{P}(Y \leq 3) = 0.494$ we need to see $\mathrm{P}(Y \leq 3)$ or $\mathrm{P}(Y < 4)$ allow different letters.

(c) **M1:** Setting up a new model B(6, 0.494) may be implied by a correct answer or ${}^6C_n (0.494)^n (0.506)^{6-n}$
M1: Using $1 - \mathrm{P}(K \leq 4)$
A1: awrt 0.104

(d) **B1:** Both hypotheses correct using λ or μ. ft "3.7" from their 3.7 in part (b) and allow $2 \times$"their 3.7" Ignore any words
B1: Realising that Po($2 \times$"their 3.7") is to be used. This may be stated or used.
M1: writing or using $1 - \mathrm{P}(J \leq 13)$ or $1 - \mathrm{P}(J < 14)$
or if finding a CR for writing $\mathrm{P}(J \geq 12) = 0.0735...$ **and** $\mathrm{P}(J \geq 13) = 0.0391...$
A1: awrt 0.0195 or CR $J \geq 13$ or $J > 12$
A1: A fully correct solution and drawing a correct inference in context.

Q2.

Qu	Scheme	Marks	AOs
(a)	$W \sim \text{Po}(11.2)$ and $\text{P}(W \geqslant 19) = 1 - \text{P}(W \leqslant 18)$ <u>or</u> suitable 3sf probs $\text{P}(W \geqslant 19) = 0.020776...$ awrt <u>**0.021**</u>	M1 A1 (2)	3.4 1.1b
(b)	[S = # calls per day, $S \sim \text{Po}(0.4)$] $\text{P}(S > 1) = 0.061551...$ awrt 0.0616 $X \sim \text{B}(250,$ "0.061551..") $Y \sim \text{Po}($"15.3879..."$)$ [Accept Po(15.4) or better] <u>or</u> suitable 3sf probs $= 0.14751...$ awrt <u>**0.148**</u>	B1 M1 M1 A1 (4)	1.1b 3.3 3.4 1.1b
(c)	H_0: $\lambda = 16.8$ H_1: $\lambda < 16.8$ $U \sim \text{Po}(16.8)$ $\text{P}(U \leqslant 8) = 0.014$ [$0.014 < 0.05$ or there is sufficient evidence to reject H_0] There is sufficient evidence at the 5% level of significance that the <u>number of calls</u> received <u>per day</u> is <u>lower</u> in winter **or** <u>rate of calls</u> is <u>lower</u> in winter **or** <u>less calls per day</u> in <u>winter</u> (o.e.)	B1 B1 M1 A1 (4)	2.5 3.3 1.1b 2.2b
(d)	$C \sim \text{Po}(0.4 \times n + 0.2 \times n)$ [$= \text{Po}(0.6n)$] <u>or</u> $D \sim \text{B}(n,\ e^{-0.6}$ or awrt 0.549) $e^{-0.6n} < 0.001$ <u>or</u> $-0.6n < \ln(0.001)$ <u>or</u> $n > 11.5...$ $n =$ <u>**12**</u>	M1 M1 A1 (3)	3.1b 1.1b 1.1b
(e)	The <u>rate of calls</u> per day is <u>constant</u> or the <u>number of calls</u> occurring in non-overlapping time <u>intervals</u> is <u>independent</u>. **or** <u>number of calls per day is independent</u> (o.e.)	B1 (1)	2.4
		Total 14	

(a)	M1 A1	For using the model Po(11.2) implied by sight of: 0.02077... or 0.9889.. or 0.9792.. awrt 0.021
(b)	B1	awrt 0.0616
	1st M1	Setting up a new model B(250, "0.0616") [condone B("0.0616", 250)]
	2nd M1	Seeing the model Po(their np) implied by sight of: 0.1475.. or 0.89975 or 0.8524...
	A1	awrt 0.148
SC		if **no approximation** used(and 1st M1 not seen) an answer of awrt 0.140 could get B1M1M0A0
(c)	1st B1	Both hypotheses correct using λ or μ and 16.8 or 0.4 [Accept their ans to 0.4×42]
	2nd B1	Realising Po(16.8) needs to be used. Sight or use of, implied by correct prob or CR
	M1	For 0.014 or better (0.0141..) or CR $X \leqslant 9$ oe must be CR and not probability. [Allow CR $X \leqslant 10$ with probability $\text{P}(X \leqslant 10) = 0.054$ or better]
	A1	**Indep of 1st B1** (must see 2nd B1 and M1 scored) for a correct inference in context
(d)	1stM1	Selecting a suitable model. Sight of Po($0.6n$) <u>or</u> B(n, $e^{-0.6}$) or implied by 2nd M1
	2ndM1	For a correct inequality or equality involving n [Condone slips in solving] Allow **MR** i.e. misread of 0.01 for 0.001 (or similar) to score M1M1A0
	A1	$n = 12$ cao [Correct answer with no incorrect working seen scores 3/3]
(e)	B1	Allow equivalent statements. Underlined words required.

Q3.

Question	Scheme		Marks	AOs
(a)	$P(R \geq 23) = 0.8517...$ awrt **0.852**		B1	1.1b
			(1)	
(b)	$R \sim Po(28)$ $A \sim Po(16)$			
	$Y = R + A \rightarrow Y \sim Po(44)$		M1	3.4
	$P(Y = 42) = 0.05866...$ awrt **0.0587**		A1	1.1b
			(2)	
(c)	P(less than 80 passengers checked in) = 0.183...		B1	1.1b
	$X \sim B(150, "0.183...")$ mean $= 150 \times "0.183..."$ [= 27.48…]		M1	3.3
	$T \sim Po("27.4...")$ and $1 - P(T \leq 24)$		M1	3.4
	$= 1 - 0.2922...$ awrt **0.708**		A1	2.1
			(4)	
(d)	$H_0: \lambda = 84$ $H_1: \lambda < 84$ (allow 28 for both)		B1	2.5
	$J \sim Po(84)$		M1	1.1b
	Method 1	Method 2		
	$P(J \leq 67) = 0.03[246...]$	CR $J \leq 68$	A1	1.1b
	$0.03... < 0.05$ or $67 \leq 68$ or 67 is in the critical region or 67 is significant or Reject H_0. There is evidence at the 5% level of significance that the system is working **slower** than normal.		A1cao	2.2b
			(4)	
			(11 marks)	

Notes:		
(a)	B1:	awrt 0.852
(b)	M1:	For combining distributions and sight or use of Po(28 + 16[=44]) Condone 28 + 16 = 42 followed by awrt 0.061
	A1:	awrt 0.0587
(c)	B1:	awrt 0.18 may be implied by awrt 27.5 for the mean
	M1:	Setting up a new model B(150, "0.183") and using np to calculate the mean.
	M1:	Using the model Po(their np) and using or writing $1 - P(T \leq 24)$
	A1:	awrt 0.708
(d)	B1:	Both hypotheses correct using λ or μ. Allow 28 instead of 84
	M1:	Writing or using Po(84)
	A1:	awrt 0.03 or $J \leq 68$
	A1cao	dep on previous M mark awarded and a probability found. Drawing a correct inference in context – need the word slower or support for Alex's complaint

Q4.

Question	Scheme	Marks	AOs
(a)	$C \sim$ Poisson (3.75)	M1	3.3
	$P(C \geqslant 2) = 0.88829...*$. awrt 0.8883*	A1*cso	1.1b
		(2)	
(b)	$D \sim B(6, \text{"0.888"})$	M1	3.3
	$P(D \leqslant 3) = 0.02163...$ awrt 0.0216 / 0.0215	A1	1.1b
		(2)	
(c)	$P(C = 8) = 0.02281...$	B1	1.1b
	$E \sim B(150, \text{"0.02281.."}) \Rightarrow \text{mean} = 150 \times \text{"0.02281..."}[= 3.4215...]$	M1	3.3
	$E \sim Po(\text{"3.4215..."}) \Rightarrow P(E \geq 3) = [1 - P(E \leqslant 2)]$	M1	3.4
	$= 0.664$ *	A1*cso	2.1
		(4)	
(d)	The number of periods is large and the probability of receiving 8 calls in 30-minutes is small.	B1 (1)	2.4
(e)	H_0: $\lambda = 30$ H_1: $\lambda \neq 30$	B1	2.5
		(1)	
(f)	$X \sim Po(30)$	B1	3.3
	$P(X \geqslant 40) = 1 - P(X \leqslant 39)$	M1	1.1b
	$= 0.04625...$	A1	1.1b
	$0.046... > 0.025$ or no evidence to reject H_0 There is insufficient evidence at the 5% level of significance that the number of calls received is different on a Saturday	A1 (4)	2.2b
			(14 marks)

Notes:		
(a)	M1:	For calculating the mean and setting up the correct model. Poisson may be implied by 0.8883 or better or 1 – awrt 0.1117 but must see 3.75 or 1.25×3
	A1*cso:	$P(C \geqslant 2) =$ awrt 0.8883 or 1 – awrt 0.1117 = 0.888 Must see $P(C \geqslant 2)$ oe
(b)	M1:	Setting up a new model using their answer to (a) Implied by correct answer
	A1:	awrt 0.0216 or awrt 0.0215
(c)	B1:	awrt 0.0228
	M1:	Setting up a new model B(150, "0.0228") and using np (working seen if incorrect)
	M1:	Using the model Po(their np) Must be clearly stated and $P(E \geqslant 3)$ oe seen
	A1*cso:	Only award if the previous 3 marks have been awarded and 0.664 is stated. NB Use of B(150 0.02281) gives 0.668
(d)	B1:	Idea that $n = 150$ (number of periods selected) is large and p is 0.022... (exactly 8 calls in the time period) is small.
(e)	B1:	Both hypotheses correct using λ or μ allow 1.25 or 3.75
(f)	B1:	Realising Po(30) needs to be used. NB Implied by correct answer or $P(X = 40) = 0.0139...$
	M1:	Writing or using $1 - P(X \leqslant 39)$ or if CR method for $P(X \geqslant 42) = 0.0221...$
	A1:	0.04... or awrt 0.05 or CR $X \geqslant 42$ oe must be CR and not probability
	A1:	A fully correct solution and correct inference in context. Calls required If put this prob but then give Cr X >= 40 M1A1A0

Q5.

Question	Scheme	Marks	AOs
(a)	$[X \sim \text{Po}(8) \quad Y \sim \text{Po}(3)]$ $[X + Y \sim]$ Po (11)	B1	3.3
	The number of cyclists travelling eastbound is independent of the number of cyclists travelling westbound.	B1	3.5b
		(2)	
(b)	$\frac{\text{P}(X=11)\times\text{P}(Y=1)+\text{P}(X=12)\times\text{P}(Y=0)}{\text{P}(X+Y=12)}$	M1 M1	2.1 1.1b
	$=0.1204\ldots$ awrt **0.120**	A1	1.1b
		(3)	
(c)	H_0: $\lambda = 11$ or $\mu = 22$ H_1: $\lambda < 11$ or $\mu < 22$	B1	2.5
	$(E + W) \sim \text{Po}(22)$ $\text{P}(E + W \leq 14)$ [= awrt 0.048]	M1	3.3
	(Reject H_0.) There is evidence that the **rate**(oe) of **cyclists**(oe) has decreased.	A1	2.2b
		(3)	
			(8 marks)

Notes	
(a)	**B1:** Correct model **B1:** Correct modelling assumption in context (must mention cyclists oe)
(b)	**M1:** Attempt at ratio expression with denominator $\text{P}(X+Y=12)$ (may see 0.10942…) **M1:** Probability expression for numerator (may be implied by 0.01317…) **A1:** awrt 0.120 accept 0.12 with correct working seen Alternative use of binomial: **M1:** Use of $C \sim \text{B}(12, \frac{8}{11})$ **M1:** $\text{P}(C \geq 11) = 1 - \text{P}(C \leq 10)$ **A1:** awrt 0.120 accept 0.12 with correct working seen
(c)	**B1:** Both hypotheses with λ or μ **M1:** Using Po(22) to calculate $\text{P}(E + W \leq 14)$ **A1:** A fully correct conclusion with awrt 0.048 or CR: $E + W \leq 14$ drawing an inference in context.

Q6.

	Scheme	Marks	AO
(a)	[X = number of errors in 100-word piece] $X \sim \text{Po}(1.7)$	M1	3.3
	$P(X < 2) = P(X \leqslant 1) = 0.49324...$ awrt **0.493**	A1	1.1b
		(2)	
(b)	[R = number of errors in the article] $R \sim \text{Po}(4.25)$	M1	3.3
	$P(R = 5) = 0.16482...$ awrt **0.165**	A1	1.1b
		(2)	
(c)	Scheme **A**: Let $A \sim B(40,\ e^{-1.7})$ or $B(40, 0.18268...)$	M1	3.3
	$P(A > 10) = 1 - P(A \leqslant 10)$	M1	1.1b
	$= 0.0995591...$ awrt **0.0996**	A1	1.1b
	Scheme **B** : Let $B \sim \text{Po}(40 \times 1.7)$ or $\text{Po}(68)$	M1	3.3
	$P(B < 56) = P(B \leqslant 55) = 0.061133...$		
	So choose scheme **A** (since the probability of a bonus is greater)	A1	2.4
		(5)	
(d)	$H_0 : \lambda = 1.7$ (or $\mu = 8.5$) $H_1 : \lambda < 1.7$ (or $\mu < 8.5$)	B1	2.5
	[E = no. of errors in the piece of work] $E \sim \text{Po}(8.5)$	M1	3.3
	$P(E \leqslant 3) = 0.0301$ or $P(E \leqslant 4) = 0.0744$	A1	1.1b
	So critical region is $E \leqslant 3$	A1	2.2a
		(4)	
		(13 marks)	

	Notes
(a)	M1 for selecting the correct Poisson distribution A1 for awrt 0.493
(b)	M1 for selecting the correct Poisson distribution A1 for awrt 0.165
(c)	1st M1 for choosing a correct model for scheme **A** i.e. $B(40, P(X = 0))$, where $X \sim \text{Po}(1.7)$ Allow use of awrt 0.183 for $P(X = 0)$ … 0.183 gives answer awrt 0.101 Condone B(0.183, 40) (o.e.) if it leads to a prob rounding to range (0.09~0.1) otherwise M0 2nd M1 for $1 - P(A \leqslant 10)$ 1st A1 for awrt 0.0996 [NB use of 0.183 will give awrt 0.101 and scores M1M1A0] 3rd M1 for selecting a correct Poisson model for scheme **B** i.e. $\text{Po}(40 \times 1.7)$ or better 2nd A1 for a correct conclusion based on comparing two probs: awrt 0.1 vs 0.061 or better So can allow $0.1 > 0.061$ leading to choosing **A** [Probably scores M1M1A0M1A1]
NB	[Normal approx.(not on spec) leading to 0.06477… might score 3rd M1 if Po(68) seen but 2nd A0]
(d)	B1 for both hypotheses in terms of λ or μ (can be interchanged) M1 for selecting Po(8.5) (sight of or use of e.g. may be implied by 1st A1) 1st A1 for some evidence of correct use of Po(8.5) i.e. either of these probs (2dp or better) May be implied by a correct critical region 2nd A1 for a correct critical region. Allow $E < 4$ and allow any letter for E. Two different regions (e.g. from 2 tail test) is 2nd A0
SC	**Use of binomial throughout:** (with hypotheses $H_0 : p = 0.017$ and $H_1 : p < 0.017$ in (d)) Scores 0 in (a) 0 in (b) possibly just 2nd M1 in (c) But allow all 4 marks in (d): B1 hypotheses, M1 for $Y \sim B(500, 0.017)$, 1st A1 for $P(Y \leqslant 3) = 0.02913...$ or $P(Y \leqslant 4) = 0.07266...$ 2nd A1 $Y \leqslant 3$ Allow probs to be to 2dp or better so 0.03 and 0.07 as in main scheme.

Q7.

Question Number	Scheme		Marks
(a)(i)	$X \sim \text{Po}(2.5)$		
	$\text{P}(X \geq 4) = 1 - \text{P}(X \leq 3)$ $= 1 - 0.7576$	**M1** writing or using $1 - \text{P}(X \leq 3)$ implied by awrt 0.242	**M1**
	$= 0.2424$	**A1** awrt 0.242	**A1**
(ii)	$X \sim \text{Po}(0.625)$	**B1** Using Po(0.625)	**B1**
	$\text{P}(X = 3) = \dfrac{e^{-0.625}0.625^3}{3!}$	**M1** finding $\text{P}(X = 3)$ with any λ e.g $\dfrac{e^{-\lambda}\lambda^3}{3!}$ or $\text{P}(X \leq 3) - \text{P}(X \leq 2)$ – may be implied by awrt 0.0218	**M1**
	$= 0.02177...$	**A1** awrt 0.0218	**A1** **(5)**
(b)	$1 - \text{P}(X = 0) < 0.2$ $\text{P}(X = 0) > 0.8$	**1st M1** for writing or using $1 - \text{P}(X = 0) < 0.2$ or $\text{P}(X = 0) > 0.8$ **oe** allow use of = instead of > or <. May be implied by $e^{-\lambda} = 0.8$ or $e^{-\lambda} > 0.8$ or by awrt 5.36 or 0.089	**M1**
	$e^{-2.5t} > 0.8$ $t < 0.089...$ hours = 5.36 mins	**2nd M1** writing an inequality of the form $e^{-\lambda} > 0.8$ using any λ. May be implied by or by awrt 5.36 or 0.089 Do not allow $e^{-\lambda} = 0.8$	**M1**
	[$t <$] 5 mins	**A1cso** both the method marks must be awarded. Accept 5 or $t = 5$ or $t < 5$	**A1cso** **(3)**
(c)	H_0: $\lambda = 2.5$ $(\lambda = 5)$ H_1: $\lambda > 2.5$ $(\lambda > 5)$	**B1** both hypotheses using λ or μ - allow 5 or 2.5 and it must be clear which is H_0 and which is H_1	**B1**
	$\text{P}(X \geq 10) = 1 - \text{P}(X \leq 9)$ $= 1 - 0.9682$	**1st M1** writing or using Po(5) and $1 - \text{P}(X \leq 9)$ May be implied by a correct CR. Do not allow for writing $\text{P}(X \geq 10)$	**M1**
	$= 0.0318$	**1st A1** awrt 0.0318. Allow CR $X \geq 10$ or $X > 9$	**A1**
		NB allow M1A1 if not using CR route for $\text{P}(X \leq 9) =$ awrt 0.968	
	Sufficient evidence to reject H_0, Accept H_1, significant. 10 does lie in the Critical region.	**2nd M1** dependent on previous M being awarded. A correct statement (do not allow if there are contradicting non-contextual statements). ft their Prob/CR compared with 0.05/10 (0.95 if using 0.968)	**M1d**
	There is sufficient evidence that the mean rate of telephone **calls** has increased **(oe)**	**2nd A1** A correct contextual statement must include the word **calls** and the idea the rate has increased. (do not allow "it has changed" on its own **oe**). All previous marks must be awarded for this mark to be awarded. M1A1 is awarded for a correct contextual statement on its own provided previous marks have been awarded	**A1cso** **(5)**
			(Total 13)

Q8.

Question	Scheme	Marks	AOs
(a)(i)	$X \sim \text{Po}(24)$	B1	3.4
	$P(X=26) = 0.071912...$ awrt **0.0719**	B1	1.1b
		(2)	
(ii)	$P(X \geq 21) = 1 - P(X \leq 20)$ [$= 1 - 0.24263...$]	M1	3.4
	$= 0.75736...$ awrt **0.757**	A1	1.1b
		(2)	
(b)	$H_0: \lambda = 2$ $[\mu = 16]$ $H_1: \lambda < 2$ $[\mu < 16]$	B1	2.5
	$P(Y \leq 10 \mid Y \sim \text{Po}(16)) = 0.077396...$ awrt **0.0774**	B1	1.1b
	Not significant / Do not reject H_0 / 10 is not in the CR	M1	1.1b
	There is not sufficient evidence to suggest a decrease/change in the rate of customers entering Jeff's supermarket.	A1	2.2b
		(4)	
(c)	Use of Po(8) to attempt critical region	M1	2.1
	Critical region is $Y \leq 3$ / H_0 is not rejected when $Y \geq 4$	A1	1.1b
	True distribution is $W \sim \text{Po}(4)$	B1	2.1
	$P(W \geq 4 \mid W \sim \text{Po}(4)) = 1 - P(W \leq 3)$ [$= 1 - 0.43347...$]	M1	1.1b
	$=0.56652...$ awrt **0.567**	A1	1.1b
		(5)	
			(13 marks)

Notes	
(a)(i) (ii)	**B1:** For realising the distribution is Po(24) (May be seen or implied in part (ii)) **B1:** awrt 0.0719 **M1:** Writing or using $1 - P(X \leq 20)$ **A1:** awrt 0.757
(b)	**B1:** Both hypotheses correct (must use μ or λ) **B1:** awrt 0.0774 Allow awrt 0.08 from a correct probability statement. allow CR: $X \leq 9$ **M1:** Correct non-contextual conclusion (may be implied by correct contextual conclusion). Allow a f.t. comparison of 'their p' with 0.05 (Ignore any contradictory contextual comments for this mark) **A1:** A fully correct solution drawing a correct inference in context **with all previous marks in (b) scored.**
(c)	**M1:** Use of Po(8) to attempt critical region [$P(Y \leq 3)$=0.0423.. $P(Y \leq 4)$=0.0996..] **A1:** Finding critical region for the test $Y \leq 3$ which must come from Po(8). **B1:** Identifying the need to use Po(4) as the true distribution. Allow Po(4) seen or used for this mark. **M1:** Writing or using $P(W \geq$ '4') or $1 - P(W \leq$ '3') from Po(4). Allow f.t. on their **identified** CR but must be using Po(4) **A1:** awrt 0.567

Q9.

Question	Scheme	Marks	AOs
(a)	P(at least 3 whites) = $(1-0.07)^3$ or $1-0.07-0.93\times0.07-0.93^2\times0.07$	M1	1.1b
	$=0.8043\ldots$ awrt 0.804	A1	1.1b
		(2)	
(b)	P(2nd red on 9th draw) $=\binom{8}{1}0.93^7\times0.07^2$	M1	3.3
	$=0.02358\ldots$ awrt 0.0236	A1	1.1b
		(2)	
(c)	$\frac{n}{p}=4400$ and $\frac{n(1-p)}{p^2}=660^2$	M1 A1	3.1b 1.1b
	$1-p=99p$ oe	M1	1.1b
	$p=0.01$	A1	1.1b
		(4)	
(d)	$H_0: p=0.07$ $H_1: p<0.07$	B1	2.5
	$J\sim \text{Geo}(0.07)$	M1	3.3
	$P(J\geq c)<0.1 \Rightarrow (1-0.07)^{c-1}<0.1$	M1	3.4
	$c-1>\frac{\log 0.1}{\log 0.93}$	M1	1.1b
	$c>32.72\ldots$ $\therefore$ CR $J\geq 33$	A1	1.1b
		(5)	
(e)	34 is in the Critical region	M1	1.1b
	There is evidence to suggest that Jerry's bag contains a smaller proportion of red counters than Asha's bag.	A1	2.2b
		(2)	
(f)	Power of test = $P(J\geq 33 \mid p=0.011)$	M1	2.1
	$=(1-0.011)^{32}$ oe	M1	1.1b
	$=0.7019\ldots$*	A1*	1.1b
		(3)	
			(18 marks)

Notes:		
(a)	M1:	A correct method to find $P(X \geqslant 3)$
	A1:	awrt 0.804
(b)	M1:	For selecting the appropriate model negative binomial or binomial with an extra trial
	A1:	awrt 0.0236
(c)	M1:	Forming an equation for the mean and variance. At least one correct.
	A1:	Both equations correct
		Allow M1 A1 if both equations correct with the same number subst for n
	M1:	Solving the 2 equations leading to $1-p=99p$ oe Allow $p-p^2=99p^2$ ft their 4400 and 660 Allow $1-p=0.15p$
	A1:	0.01
(d)	M1:	Both hypotheses correct using correct notation allow eg $p > 0.93$
	M1:	Realising the need to use Geo(0.07) ft their Hypotheses
	M1:	Using the model to find $P(J \geqslant c)$ Condone $(1-0.07)^c < 0.1$ ft their $0.07 \neq 0.93$ ALT $P(J \geqslant 32) = 0.1[054...\}]$or $P(J \geqslant 33) = 0.09[8...]$ Implied by correct CR
	M1:	For a valid method to solve the inequality or $P(J \geqslant 32) = 0.1[054]$ and $P(J \geqslant 33) = 0.09[81]$ Implied by correct CR
	A1:	Correct CR(any letter) A0 if given as a probability statement. Must be integer
(e)	M1:	Comparing 34 with their CR eg $34 > 33$ $34 \geqslant 33$ or $P(J \geqslant 34) = 0.09[12]$
	A1:	Fully correct conclusion in context. Allow Jerry's belief is true. Allow probability for proportion
(f)	M1:	Realising they need to find P(their CR in (d)) Allow $1 - P(J \leqslant 32)$
	M1:	For a Correct method. Allow $1 - 0.2981...$ May be implied by $0.7019...$ If the CR is incorrect $(1-0.011)^{"CR"-1}$ or $1 - \{1-(1-0.011)^{"CR"-1}\}$ must be seen
	A1*:	Only award if both method marks awarded.

Q10.

Question	Scheme		Marks	AOs
	$H_0 : \lambda = 5$ $(\lambda = 2.5)$ $H_1 : \lambda > 5$ $(\lambda > 2.5)$		B1	2.5
	$X \sim$ Po (2.5)		B1	3.3
	Method 1	Method 2		
	$P(X \geqslant 7) = 1 - P(X \leqslant 6)$ $= 1 - 0.9858$	$P(X \geqslant 5) = 0.1088$ $P(X \geqslant 6) = 0.042$	M1	1.1b
	$= 0.0142$	CR $X \geqslant 6$	A1	1.1b
	$0.0142 < 0.05$ $7 \geqslant 6$ or 7 is in critical region or 7 is significant Reject H_0. There is evidence at the 5% significance level that the level of pollution has increased. or There is evidence to support the scientists claim is justified		A1cso	2.2b
				(5 marks)

Notes	
	B1: Both hypotheses correct using λ **or** μ **and** 5 or 2.5 **B1:** Realising that the model Po(2.5) is to be used. This may be stated or used. **M1:** Using or writing $1 - P(X \leqslant 6)$ or $1 - P(X < 7)$ a correct CR **or** $P(X \geqslant 5) =$ awrt 0.109 **and** $P(X \geqslant 6) =$ awrt 0.042 **A1:** awrt 0.0142 or CR $X \geqslant 6$ or $X > 5$. **A1:** A fully correct solution and drawing a correct inference in context