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)

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

$P(X \geq 59.5) = P\left(Z \geq \frac{59.5 - 48}{\sqrt{48}}\right)$

$= P(Z \geq 1.66)$
$= 1 - 0.9515$
$= 0.0485$

Method 2

$\frac{x - 0.5 - 48}{\sqrt{48}} = 1.6449$ M1 M1 A1

$x = 59.9$ 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 λ or μ
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 ± 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 ± z.] Method 2 [($x \pm 0.5$) and equal to a ± z value)
A1 correct z value awrt ±1.66 or ± $\frac{59.5-48}{\sqrt{48}}$, 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 ft 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]

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 λ or μ
- 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.