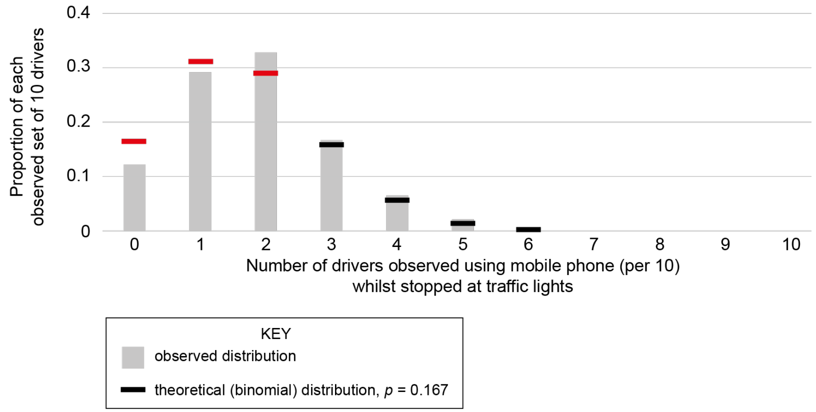
## Assessment Schedule – 2021

## Mathematics and Statistics (Statistics): Apply probability distributions in solving problems (91586)

### Evidence Statement

| Q | Expected coverage | Achievement (u) | Merit (r) | Excellence (t) |
|---|---|---|---|---|
| ONE (a)(i) | Binomial distribution model because:<br>• Fixed number of outcomes with 6 sets of traffic lights along Memorial Avenue.<br>• Two outcomes possible – a driver either is required to stop at a set of lights (amber / red light) or is not required to stop (green light).<br>• Fixed probability of success (chance of being required to stop) – (0.45).<br>• Whether a driver must stop at one set of lights is independent of whether they must stop at another set. | | • At least two conditions correctly identified with contextual justification (set of 6 traffic lights, 0.54 probability of stopping).<br>AND | |
| (ii) | Using $n = 6$ and $p = 0.45$<br>$P(X = 0)$ or $P(X = 6)$<br>$= 0.0277 + 0.0083$<br>$= 0.0360$ | • ONE correct probability. | Final probability correct with evidence. | |
| (b)(i) | Using $n = 6$ and $p = 0.60$<br>$P(X \geq 4) = 1 - P(X \leq 3)$<br>$= 1 - 0.456$<br>$= 0.544$ | • Correct probability. | | |

| | | | | |
|---|---|---|---|---|
| (ii) | Without proposed safety changes<br>Using $n = 6$ and $p = 0.45$<br>$P(X \geq 4) = 1 - P(X \leq 3)$<br>$= 1 - 0.7448$<br>$= 0.2552$<br>Through all working week probability of being unlucky each time without the proposed safety changes is<br>$= 0.2552^{10} = 0.000001172$ (4 sf).<br>With proposed safety changes through all working week probability of being unlucky each time $= 0.544^{10} = 0.002270$ (4 sf).<br>Both of these probabilities are very small so the changes aren't going to make any realistic difference and therefore the claim is correct.<br>OR<br>As a relative comparison, $\frac{0.002270}{0.000001172} = 1937$<br>So the claim is rejected as they are 1937 times as likely to have an unlucky 10 journeys with the safety changes compared to without the safety changes.<br>*Accept other reasonable rounding.* | • Correct probability seen for being unlucky with the new parameter for the safety changes. | • Both probabilities calculated for a full working week.<br>AND<br>A comparative comment made. | • Both probabilities calculated for the full working week.<br>AND<br>A comparitive comment made.<br>AND<br>An appropriate conclusion about the claim made. |

| | | | | |
|---|---|---|---|---|
| (c)(i) | Theoretical probabilities for binomial distribution model using parameters $n = 10$ and $p = 0.167$<br>*These theoretical probabilities are plotted on the graph*<br>$P(X = 0) = 0.161$<br>$P(X = 1) = 0.322$<br>$P(X = 2) = 0.291$<br> | • The three correct model probabilities are shown on the graph. | • The three correct model probabilities are shown on the graph.<br>AND<br>ONE feature of the experimental distribution is described and compared to the theoretical conclusion. | • The three distribution probabilities are shown on the graph.<br>AND<br>TWO features of the observed distribution are described and compared to the theoretical distribution.<br>AND<br>A valid conclusion stating the binomial model is suitable is made. |
| (ii) | **Discussion:** features of the observed distribution and the theoretical binomial distribution are compared.<br>Examples of possible comparisons<br>• The probability of the theoretical distribution for 0 and 1 drivers is more than the probabilities for the observed distribution, while the probability for the theoretical distribution for 2 drivers using mobile phones while stopped at traffic lights is less than the probability for the observed data. Both distributions have very similar probabilities for 3 or more drivers.<br>• The modal value for the theoretical distribution is 1 driver using a mobile phone while stopped at traffic lights. This is less than the observed distribution which is 2 drivers using a mobile phone.<br>• The mean of the theoretical distribution is 1.67 drivers using their mobile phone while stopped at traffic lights. This is slightly lower than the mean of the observed distribution which is approximately 1.83 drivers<br>• Both the theoretical and observed distributions of the number of drivers using their mobile phones while stopped at traffic lights are positively or right skewed<br>**Conclusion**: the binomial model appears to be a good fit for this situation on Memorial Avenue. | | | |

| NØ | N1 | N2 | A3 | A4 | M5 | M6 | E7 | E8 |
|---|---|---|---|---|---|---|---|---|
| No response; no relevant evidence. | Attempt at one part of the question. | 1 of u | 2 of u | 3 of u | 1 of r | 2 of r | 1 of t | 2 of t |

| Q | Expected coverage | Achievement (u) | Merit (r) | Excellence (t) |
|---|---|---|---|---|
| TWO (a)(i) | Using Poisson model with $\lambda = 1.8$<br>$P(X > 4) = 1 - P(X \leq 4)$<br>$= 1 - 0.963359$<br>$= 0.03641$ | • Correct probability calculated. | | |
| (ii) | Need $P(X = 0) = \frac{2}{3}$<br>So $e^{-\lambda} = \frac{2}{3}$<br>$\lambda = 0.4055$ | • Correct initial set up to solve. | • Correct value of $\lambda$ with clear and logical steps as to how the value was found. | |
| (iii) | Examples of possible responses are<br>• The Poisson distribution assumes independence of events, that is the occurrence of one wilding pine tree growing does not affect the probability that a second wilding pine will grow. This may not be the case as a tree will drop seeds therefore increasing the likelihood that more trees will grow around it.<br>• The Poisson distribution assumes that the rate that wilding pines grow is proportional to the area. This may not be the case as the average number of wilding pines per square kilometre may not be proportional to smaller areas since trees require adequate space to grow. This means there would be a limit on how small this scaling down can go.<br>• The Poisson distribution assumes that the location of a wilding pine growing is random and unpredictable. This may not be the case because some conditions e.g. better soil in some areas of the island and volcanic rocks in others, may be more conducive to growing pines than others. Therefore, growth in some areas will be more predictable.<br>*Do not accept arguments about simultaneous.* | • TWO correct reasons in context. | • At least ONE correct reason linked to the context.<br><br>AND<br><br>Explanation of why it is not appropriate to use a Poisson model. | • At least TWO different, correct reasons linked to the context.<br><br>AND<br><br>Explanation of why each reason is not appropriate to use a Poisson model. |

| | | | | |
|---|---|---|---|---|
| (b)(i) |  | • One graph correctly drawn including correct height and clearly identified. | • Both graphs correctly drawn including correct heights and clearly identified. | |
| (ii) | P(NI re-emerge within 8 months) = $\frac{1}{15} \times 5 = \frac{1}{3}$<br>P(SI re-emerge within 8 months) = $0.5 \times 5 \times \frac{5}{108} = \frac{25}{216} (0.1157\ldots)$<br>Probability that randomly selected North Island and South Island areas have re-emergence within 8 months = $\frac{1}{3} \times \frac{25}{216} = \frac{25}{648} (0.0386\ (3\text{ s. f.}))$ | • EITHER<br>North island or South Island probability for re-emergence within 8 months correctly calculated. | • Each of the North Island and the South Island probabilities for re-emergence within 8 months correctly calculated.<br>AND<br>EITHER<br>The correct probability that the two areas see a re-emergence within 8 months. | • Each of the North Island and the South Island probabilities for re-emergence within 8 months correctly calculated.<br>AND<br>The correct probability that the two areas see a re-emergence within 8 months. |
| (iii) | The assumption is independence of wilding pine re-emergence in an area in the North Island and an area in the South Island. This may not be valid because an area of wilding pines from the bottom of the North Island and one from the top of the South Island may be close enough to share characteristics (e.g. broad weather patterns).<br>*Accept other valid reasons for independence or that they are not independent.* | OR<br>A correct discussion over the validity of the assumption of independence. | OR<br>A correct discussion over the validity of the assumption of independence. | AND<br>A correct discussion over the validity of the assumption of independence |

| NØ | N1 | N2 | A3 | A4 | M5 | M6 | E7 | E8 |
|---|---|---|---|---|---|---|---|---|
| No response; no relevant evidence. | Attempt at one part of the question. | 1 of u | 2 of u | 3 of u | 1 of r | 2 of r | 1 of t | 2 of t |

| Q | Expected coverage | Achievement (u) | Merit (r) | Excellence (t) |
|---|---|---|---|---|
| THREE (a)(i) | Reason(s) for similar expected values:<br>• The probabilities across both systems have symmetry in the middle around 3 to 4 extras purchased. System A "achieves this" with a peak concentration of proportions in the middle and System W via bi-modality at either extremes, so the expected value will still fall close to the middle.<br>Reason(s) for different standard deviations:<br>• System A with the proportions concentrated in the middle has less variation. With the concentration of proportions for System W at either end there will be more variation.<br>Note: No credit awarded for calculations. Explanations must come from description of distributions. | • Correct explanation for standard deviation | • Correct explanation for expected value. | Correct explanation for BOTH expected value and standard deviation. |
| (ii) | For System A, P(5 or more) = 0.08 + 0.04 + 0.02<br>= 0.14<br>For System W, P(5 or more) = 0.18 + 0.14 + 0.15<br>= 0.47<br>It is much more likely (3.36 times as likely) to purchase 5 or more extras for system W than for system A.<br>OR<br>0.47 > 0.14 so it is more likely that owners of System W will buy more extras than owners of system A. | • BOTH probabilities calculated. | • BOTH probabilities calculated.<br>AND<br>A comparative comment made. | |
| (b)(i) | Using a normal distribution, N = 5.5, $\sigma = 0.9$<br>$P(X > 6) = 0.289$<br>$P(X > 7.5) = 0.0131$<br>So $P(X > 7.5 \mid X > 6) = \dfrac{0.0131}{0.289} = 0.0453$ | • One correct normal distribution probability correct. | • Conditional probability correctly calculated. | |

| | | | | |
|---|---|---|---|---|
| (ii) | Using parameters given of N = 5.5, $\sigma = 0.9$<br>$P(X < 2.5) = 0.00042906$<br>$P(X > 8.5) = 0.00042906$<br>0.00042906 + 0.00042906 = 0.00085817 or 0.08581%<br>0.08591% is much less than 4%<br>Therefore, the parameters of the normal distribution cannot be used to model the time to serious error for system L<br>Alternative solution show that $1 - P(2.5 < X < 8.5) < 4\%$, therefore the parameters used for System A and W are not appropriate for System L. | • EITHER<br>Both tail end probabilities calculated.<br>OR<br>$1 - P(2.5 < X < 8.5)$ correctly calculated. | • Relevant probability correctly calculated.<br>AND<br>compared to 4% with correct conclusion regarding use of given parameters. | |
| (c) | The Normal distribution model would be an appropriate model for the time to serious error because:<br>• *Centre:* The mean for the sample of 800 computers (6.236) is very close to the mean (6.25) of the normal model<br>• *Shape:* The normal distribution is bell-shaped, unimodal and symmetrical about the mean, this graph is reasonably bell-shaped and symmetrical about the centre, so it appears to fit the normal model.<br>• *Spread:* In the normal distribution, 99% of the data should be within ±3 standard deviations of the mean. In this data the number of computers which developed a serious error outside ±3 standard deviations of the mean is very small, i.e. there is very little data before 3.9 years or after 8.56 years.<br>• *Probability:* In the normal model the middle 50% of the data should fall between 5.7 years and 6.8 years. In the observed data it is between 5.7years and 6.7 years so these are very similar.<br>• *Accept other valid responses that clearly compare the features of observed data with the proposed theoretical model.* | • EITHER<br>ONE relevant feature of the distribution is described and compared to the features of the normal distribution.<br>OR<br>A relevant probability is calculated using the normal distribution and compared to the given distribution for the sample. | • EITHER<br>ONE relevant feature of the distribution described and compared to the features of the normal distribution.<br>AND<br>a relevant probability calculated using the normal distribution and compared to the given distribution.<br>OR<br>TWO different relevant features of the distribution are described and compared to the features of the normal distribution. | • TWO different relevant features of the distribution are described in context and compared to the features of the normal distribution.<br>AND<br>A relevant probability is calculated using the normal distribution and compared to the given distribution. |

| NØ | N1 | N2 | A3 | A4 | M5 | M6 | E7 | E8 |
|---|---|---|---|---|---|---|---|---|
| No response; no relevant evidence. | Attempt at one part of the question. | 1 of u | 2 of u | 3 of u | 1 of r | 2 of r | 1 of t | 2 of t |

## Cut Scores

| Not Achieved | Achievement | Achievement with Merit | Achievement with Excellence |
|---|---|---|---|
| 0 – 6 | 7 – 13 | 14 – 18 | 19 – 24 |