

Assessment Schedule – 2025

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

Evidence Statement

Q	Expected coverage	Achievement (u)	Achievement with Merit (r)	Achievement with Excellence (t)
<p>ONE (a)(i)</p> <p>(ii)</p> <p>(iii)</p> <p>(iv)</p>	<p>Post is short: $P(X < 80) = 0.0098$</p> <p>Post is not short: $P(X > 80) = 1 - 0.0098 = 0.9902$</p> <p>Probability that one post is short and the other is not short: $P(X < 80) = 0.0099$ $P(X > 80) = 0.9902$ P(one short, one not) $= 2 \times P(X < 80) \times P(X > 80)$ $= 2 \times 0.0098 \times 0.9902$ $= 2 \times 0.009704$ $= 0.0194 \approx 0.02$</p> <p>Assumption: Independence: one post being short doesn't affect the length of another post. Impact of violation: One post's length could affect the length of another. (e.g. posts from the same user could all be around the same length or posts on a specific topic could be responses to an initial post and so all be cluster around similar lengths), The actual probability of a short post could be higher or lower than the estimate.</p> <p>For this group of X users, the probability of them creating a post longer than 280 characters is very small. The normal distribution assumes that 99% of all posts by this group of X users (150) will be within 3 standard deviations of the mean, In this case between 60 to 180 characters long. The 280-character limit is considerably more than this. The likelihood of observing an event more than three standard deviations away from the mean is so small that it is effectively negligible.</p>	<ul style="list-style-type: none"> • Probability correct. • Probability of one combination correct. • States the independence assumption. 	<ul style="list-style-type: none"> • Combined probability correct. • States the independence assumption in context. AND Gives an example of dependence in context. OR Explains the impact of violating the assumption on the probability estimate. • Identifies that the probability of a post longer than 280 characters is very small and unlikely, linked to the normal distribution. 	<ul style="list-style-type: none"> • States the independence assumption.in content. AND Explains why this might not be valid in context. AND Explains the impact of violating the assumption on the probability estimate. • Identifies that the probability of a post longer than 280 characters is very small and unlikely justified by referring to a feature of the normal distribution.

(c)	$P(X > 70) < 0.01$ $\frac{(70 - 37)}{s} < 2.33$ $\frac{33}{s} < 2.33$ $s < 14.16$ <p>Using $s = 6.47$</p> $P(X < 17) = 0.000998 \text{ which is less than } 0.001$ $P(X > 70) = 0.000000169 \text{ which is less than } 0.01$ <p>So, the minimum standard deviation to fulfil both of these criteria is less than or equal to 6.47 years.</p>	<ul style="list-style-type: none"> • One correct calculation of s. 	<ul style="list-style-type: none"> • Two correct calculations s. 	<ul style="list-style-type: none"> • Two correct calculations of s. AND Correct minimum value justified.
-----	--	--	--	--

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempts at least 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)	Achievement with Merit (r)	Achievement with Excellence (t)
THREE (a)	Binomial: $n = 20, p = 0.1$ $P(X < 3) = P(X \leq 2) = 0.6769268$ Alternatively: $P(X < 3) = P(X = 0,1,2) = 0.1216 + 0.2702 + 0.2852 = 0.6769$	<ul style="list-style-type: none"> Identifies correct distribution and parameters. AND Correct probability.		
(b)	Independence assumption: the identification of inappropriate or harmful content in one post does not affect the probability of inappropriate or harmful content being identified in another post. This assumption is unlikely to be valid because all 20 posts have been created by the same user and are likely to use similar language or grammar – if the model detects inappropriate content in one post it may be more likely to detect inappropriate content in another.	<ul style="list-style-type: none"> Independence assumption described in context. 	<ul style="list-style-type: none"> Independence assumption described in context with reasoning as to why it is unlikely to be valid. 	

<p>(c)(i)</p> <p>(ii)</p>	<p>$E(X) = 2.1$ The system reports 2.1 false positives, on average, in a batch of 100 posts.</p> <p>The Poisson model assumes that false positives are typically independent events whether one post is reported as a false positive in a batch does not affect another post being reported as being a false positive in another batch. The automated system being used has no memory of previous reports so its reporting of one false positive will be independent of the next report.</p> <p>The Poisson model assumes that false positives in a batch would be expected to occur randomly and unpredictably. We assume the automated system selects posts for checking at random and assume that false positives are reported randomly within those selected.</p> <p>The Poisson model assumes that the rate is proportional to the interval. It is assumed that the rate of reporting false positives in a batch of posts is 2.1 false positives per batch of 100 posts which is scalable to give a rate in a batch of 10, 50 100, 200, 500, 1000 posts and so on because of the automated nature of the system.</p> <p>The Poisson model assumes that events cannot occur simultaneously. In this context, we are assuming the system is automated and processes one item at a time very quickly. Therefore, two false positives cannot be identified at exactly the same moment.</p> <p>The Poisson model also assumes that the rate of events occurring remains constant over the given time or space interval. Since the automated system follows a fixed set of rules to identify harmful posts, we must assume that the rate at which it detects false positives within a batch must remain constant.</p>	<ul style="list-style-type: none"> • Correct calculation. • ONE correct reason identified in context 	<ul style="list-style-type: none"> • Justifies applicability of Poisson distribution to modelling false positives. Either TWO correct reasons identified in context Or • ONE correct reason justified in context. 	<ul style="list-style-type: none"> • TWO correct reasons justified in context
<p>(d)</p>	<p>$P(X = 0) = 0.03$ $\lambda = -\ln(0.03) = 3.507 = 3.5$ false positives per 100 posts. With the original system, the false positive rate is 2.1 false positives per 100 posts. The competing system appears to be less accurate less reliable as it has a higher false positive rate i.e. it is identifying a post as having harmful content when it doesn't have.</p>	<ul style="list-style-type: none"> • Calculates λ for the competing system. 	<ul style="list-style-type: none"> • Compares systems using λ. 	<ul style="list-style-type: none"> • Compares systems using λ. AND Conclusion regarding accuracy in context.

N0	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempts at least 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
0–7	8–13	14–18	19–24