SUPERVISOR'S USE ONLY

3

91586

915860

Draw a cross through the box (☒) if you have NOT written in this booklet

+

NZQA | Mana Tohu Mātauranga o Aotearoa
New Zealand Qualifications Authority

Level 3 Mathematics and Statistics (Statistics) 2025

91586 Apply probability distributions in solving problems

Credits: Four

Achievement	Achievement with Merit	Achievement with Excellence
Apply probability distributions in solving problems.	Apply probability distributions, using relational thinking, in solving problems.	Apply probability distributions, using extended abstract thinking, in solving problems.

Check that the National Student Number (NSN) on your admission slip is the same as the number at the top of this page.

You should attempt ALL the questions in this booklet.

Make sure that you have the Formulae and Tables Booklet L3–STATF.

Show ALL working.

If you need more room for any answer, use the extra space provided at the back of this booklet.

Check that this booklet has pages 2–12 in the correct order and that none of these pages is blank.

Do not write in the margins (⁝⁝⁝). This area will be cut off when the booklet is marked.

YOU MUST HAND THIS BOOKLET TO THE SUPERVISOR AT THE END OF THE EXAMINATION.

QUESTION ONE

The length of posts on the social networking site *X* (previously known as *Twitter*) tends to follow a normal distribution, where most posts have average lengths, with few very short or very long posts.

The parameters for the normal distribution chosen to model post lengths on *X* will depend on the platform and the specific group of users; for one group of *X* users, the average post length is 150 characters, and the standard deviation is 30 characters.

(a) (i) A short post is a post that is fewer than 80 characters long.

Use this normal distribution model to estimate the probability that a post is NOT short.

(ii) Two posts are chosen at random.

Estimate the probability that one post is short and the other is not short.

(iii) State an assumption that you made when making this calculation and describe how violating this assumption could affect the accuracy of your estimate in part (ii).

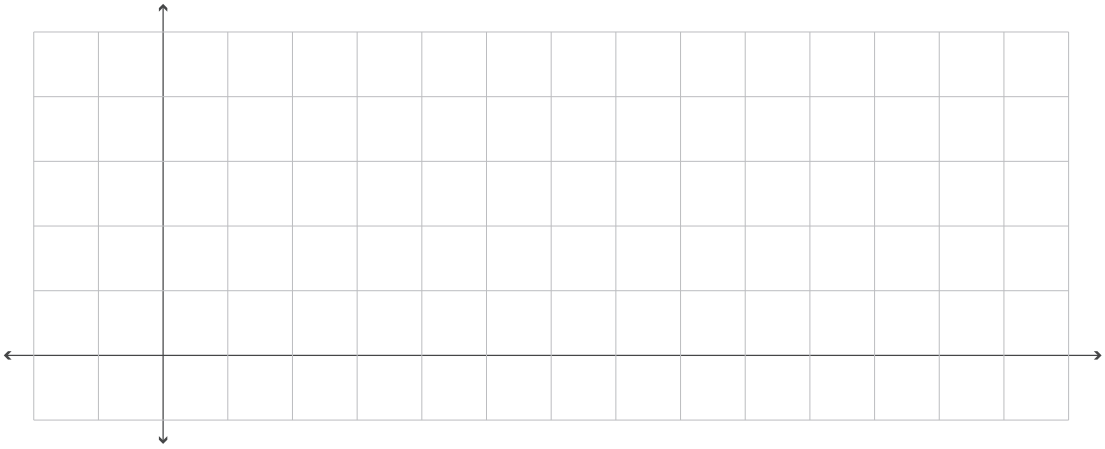
(iv) The maximum number of characters allowed per post for *X* users with a free account is 280.

Without using any further probability calculations, use statistical reasoning to explain why this restriction on the number of characters allowed per post is unlikely to affect this group of X users.

(b) The *Instagram* users from this group of *X* users create at least 10 posts per week, and no more than 130 posts per week.

(i) Sketch the distribution of the number of posts per week on the axes below.

If you need to redraw your response, use the grid on page 9.



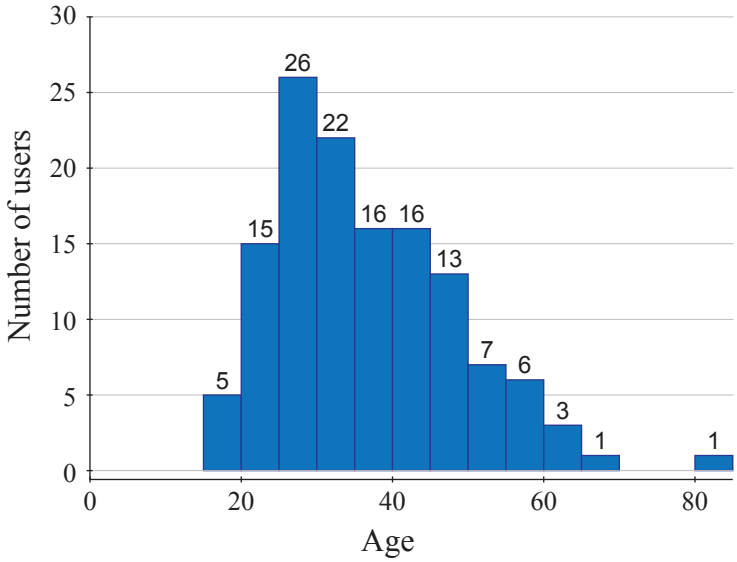
(ii) Given that a particular *Instagram* user creates fewer than 120 posts in one week, estimate the probability that they create more than 50 posts in one week.

Use calculations to support your answer.

QUESTION TWO

For adult social media users, some studies report that the most common age is 30 years old.
A sample of 131 adult social media users between the ages of 17 and 82 years old was taken.
The distribution of the ages of these social media users is shown in Figure 1.

Figure 1: Adult social media users by age



https://www.statcrunch.com/reports/view?reportid=45453&tab=preview

(a) It is claimed that 35% to 40% of social media users are between 25 and 35 years old.

Does this data support the claim?

(b) A triangular distribution model with parameters $a = 17$, $b = 82$, and $c = 30$ is used to model this data.

(i) Justify the choice of these parameters.

Include an observation from Figure 1 as part of your answer.

a = 17:

$b = 82$:

$c = 30$:

(ii) Use this triangular distribution model to estimate the probability that a randomly chosen social media user is more than 35 years old.

(iii) Use your answer from part (ii) and observations from Figure 1 to comment on the suitability of using this triangular distribution model for the data in this sample.

(c) The average age of the sample of social media users is 37 years. No individuals are aged less than 17 years, and only one individual is aged over 70 years.

If a normal distribution is to be used to model the ages of these social media users, calculate the minimum standard deviation that fulfils both of the following constraints: a probability of less than 0.001 of a randomly chosen social media user being less than 17 years of age, and a probability of less than 0.01 of a randomly chosen social media user being more than 70 years of age.

QUESTION THREE

A social media platform checks user-generated content using an automated system to identify potentially harmful or inappropriate posts. A population of posts where 1 in 10 posts contain harmful or inappropriate content is screened.

(a) The automated system is used to screen 20 posts created by one user for harmful or inappropriate content.

Use a probability distribution model to estimate the probability of observing fewer than 3 harmful or inappropriate posts.

State the name and parameters of this distribution as part of your answer.

(b) To apply the distribution used in part (a) to model the number of harmful or inappropriate posts, an assumption of independence was made.

Describe this assumption in context, and whether it is valid to make this assumption for this situation.

(c) The system reports false positives, meaning that a proportion of harmless content is mistakenly identified as inappropriate.

(i) The table below shows the probability distribution of the random variable, X, the false positives reported per batch of 100 posts.

x	0	1	2	3	4	5	6	7
$P(X = x)$	0.12	0.27	0.27	0.18	0.09	0.04	0.02	0.01

How many false positives, on average, are reported by the system for each batch of 100 posts?

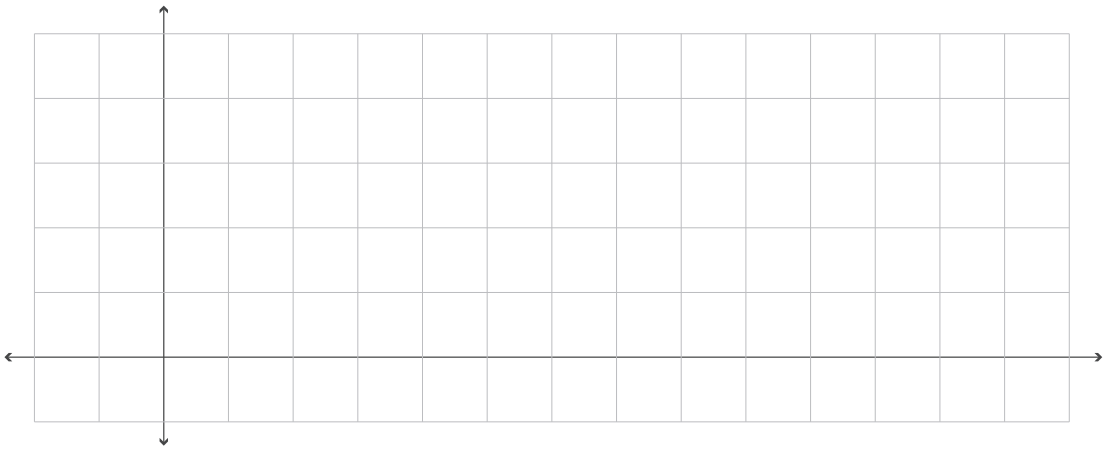
(ii) Discuss why using the Poisson distribution might be useful to model an event like false positives in very large populations.

(d) A competing social media screening system states that at least one false positive is found in 97% of their batches of 100 posts.

Calculate the false positive rate for this screening system, and use this calculation to comment on the accuracy of the competing system compared to the original system from part (c).

SPARE DIAGRAMS

If you need to redraw your response to Question One (b)(i), use the grid below. Make sure it is clear which answer you want marked.



Extra space if required.
Write the question number(s) if applicable.

QUESTION NUMBER

Extra space if required.
Write the question number(s) if applicable.

QUESTION NUMBER

Extra space if required.
Write the question number(s) if applicable.

QUESTION NUMBER