Assessment Schedule – 2025
Scholarship Statistics (93201)
Evidence Statement
General Principles:

1 Ignore incorrect answers if alongside correct answers. The exception is contradictory statements.

2 Ignore minor copying errors.

3 When required in evidence, answers need to be contextual.

QUESTION ONE

Task Q1(a)(i)

Evidence:

- In Figure 1, colour has been used on a world map to show the share of global plastic waste emitted as a percentage range for individual countries.
- Lighter colours represent lower percentages, and darker colours represent higher percentages.
- In Figure 2, different colours have been used on a world map to show the amount of plastic waste (in kg) emitted into the ocean per capita.

Task Q1(a)(ii)

Evidence:

- Small island nations, with more coastline, such as those in Asia, tend to be darker in colour across both figures than larger nations like Australia or Canada.

Task Q1(b)(i)

Evidence:

- $\text{P(high level of microplastics)} = 0.6 \times \left(0.25 \times \frac{8}{12}\right) + \left(0.4 \times \frac{8}{12}\right) = \frac{1}{10} + \frac{4}{15} = \frac{11}{30}$

$$\text{P(midwater} \mid \text{high)} = \frac{\text{P(midwater} \cap \text{high)}}{\text{P(high)}} = \frac{\left(0.4 \times \frac{8}{12}\right)}{\frac{11}{30}} = \frac{8}{11}$$

Task Q1(b)(ii)

Evidence:

- It was assumed the midwater samples from the 12 locations in the original study were representative of all midwater locations.
- It was assumed the probability of a surface water sample having a high level of microplastics was ¼ the probability of a midwater sample having a high level of microplastics.

Task Q1(b)(iii)

Evidence:

- $\text{P}(X = 1) = 0.2707$
- The expected number of samples with exactly 1 piece = $500 \times 0.2707 = 135.35$
- The observed number of samples with exactly 1 piece is 98, which is much lower than the expected 135, and the suggested model may not be appropriate.

Task Q1(b)(iv)

Evidence:

- The two 25 mL samples are independent of each other, which introduces additional combinations.
- Finding a total of 2 microplastic pieces in the two 25 mL samples can be achieved through various combinations (e.g. 2 in sample A and 0 in sample B, 2 in sample B and 0 in sample A, or 1 in each sample), whereas finding exactly 2 microplastic pieces in 50 mL only allows for one successful combination.
- $\text{P(2 in 2} \times \text{25 mL)} = 0.3679^2 + 2(0.3679 \times 0.1839) = 0.2707$
  $\text{P(2 in 50 mL)} = 0.2707$

Scholarship level	Outstanding Scholarship level
• Describing graphical techniques such as colour. • Using data visualisations to evaluate a claim. • Calculation of a conditional probability. • Discussing assumptions of using study data to calculate a probability. • Evaluating a given probability distribution model using a suitable calculation. • Using calculations to support that the probability of success for two 25mL samples is equal to the probability of success for one 50mL sample. **Max 6 marks**	• Calculating a conditional probability and discussing the assumptions that have been made. • Discussing why the probability of success for two 25 mL samples is equal to the probability of success for one 50 mL sample. **Max 2 marks**

QUESTION TWO

Task Q2(a)

Evidence:

- Of the packaging assessed, an average of 57% of it was not practically recyclable per item.
- The claim that 57% of the packaging used in New Zealand is not recyclable in practice cannot be supported, as the items in the study were specifically chosen due to their popularity internationally, and cannot be assumed to be representative of New Zealand packaging.
- The claim also does not take into account the quantity of each product consumed within New Zealand.

Task Q2(b)

Evidence:

- Households without the app had a mean which was 4.16 percentage points higher than households with the app. Sending more recyclable waste to landfill on average than those households with the app.
- The tail proportion is 0%. As this is smaller than 5%, there is evidence to support a claim that the app reduced the amount of recyclable packaging sent to landfill.
- Participants were aware that they were in a study which was looking at their recycling, their behaviour may not have been representative of their usual recycling habits.
- Participants in the app group were instructed to separate their discarded packaging with the use of the app. In practice, having access to the app may not mean that people will use it.
- The experiment was only run over 3 months; recycling habits may change over time, for example participants my exhibit different recycling or purchasing habits in summer and winter that may affect results.

Task Q2(c)(i)

Evidence:

- $0.05 \pm \frac{2}{\sqrt{1000}} = [-0.0132, 0.1132]$
- We can be pretty sure that the proportion of New Zealanders who consider recycling to be more effective at reducing the effects of a climate crisis than reducing packaging, is somewhere between 1.32% less and 11.32% more than the proportion of New Zealanders who do not consider recycling to be more effective at reducing the effects of a climate crisis than reducing packaging.
- Because this confidence interval contains 0, it does not support the claim that a higher proportion of New Zealanders believe recycling is more effective at reducing the effects of a climate crisis than reducing packaging.
- Because the proportion for less packaging is below 30%, using the rule-of-thumb to calculate the margin of error will result in an overestimate.
- *Alternatively, an exact confidence interval could be found using* $Z \times \sqrt{\frac{p(1-p)}{n}}$ *as the margin of error.*

Task Q2(c)(ii)

Evidence:

- Study A will enable the Ministry for the Environment to make a sample-to-population inference about the difference in energy usage in recently renovated New Zealand homes, which were originally (pre-renovation) ‘energy efficient’, compared with those that were originally ‘not energy efficient’.
- Study B is an experiment which will enable the Ministry for the Environment to make a causal claim about whether or not energy efficient renovations cause a reduction in energy usage.

Scholarship level	Outstanding Scholarship level
• Explaining how the given sample is non-representative. • Comparing relevant features of the sample distributions in context. • Interpreting the tail proportion in context. • Calculating an appropriate estimate for the margin of error for the difference of two independent proportions. • Interpreting a confidence interval for the difference of two proportions in context. • Identifying and explaining the use of a given sample-to-population inference study design. • Identifying and explaining the use of a given experimental study design. Max 6 marks	• Discussing the effect of the study design on the extension of results. • Identifying that the margin of error will be over-estimated due to one poll proportion being less than 30%. Max 2 marks

QUESTION THREE

Task Q3(a)

Evidence:

- There is a positive relationship between GDP per capita and per capita plastic waste. The strength of the relationship decreases as GDP per capita increases.
- Countries in Africa tend to have lower per capita plastic waste, with much less variation than countries in North and Central America or Europe.
- Oceania and Asia have similar variation in their per capita plastic waste output, with the exception of Kuwait (in Asia) producing around 0.3 kg more plastic waste per capita than any other Asian or Oceanic country.
- Countries with larger populations tend to have lower per capita plastic waste, with the exception of the USA.

Task Q3(b)(i)

Evidence:

- E(X) = 551μm

Task Q3(b)(ii)

Evidence:

- Normal distribution with mean of 551 and standard deviation of 150–225.
- Triangular distribution with a = 150, b = 1050, c = 551.

Task Q3(b)(iii)

Evidence:

- Using sample data $\mathrm{P}(X < 550) = \frac{31}{74} = 0.4189$.
- Comparison of tail probability using sample data, and both distributions, for example:
  - The normal distribution model overestimates P(X < 550), giving 0.4911-0.4972.
  - The triangular distribution slightly overestimates P(X < 550), giving 0.4433.
- Probability estimate using the triangular distribution is closer to the sample probability, the triangular distribution model is recommended.

Task Q3(c)

Evidence:

- The sample mean number of macroplastics in a 20 m^2 area on exposed beaches on the Coromandel Peninsula is 0.37 higher than for sheltered beaches on the Coromandel Peninsula.
- The variation in the sample number of macroplastics in a 20 m^2 area on exposed beaches on the Coromandel Peninsula is greater than on sheltered beaches.
- Based on the bootstrap confidence interval, we can be fairly confident that the mean number of macroplastics in a 20 m^2 area on exposed beaches on the Coromandel Peninsula is somewhere between 0.79 pieces less and 1.49 pieces more than on sheltered beaches on the Coromandel Peninsula.
- Because this confidence interval is both negative and positive, there is not enough evidence to conclude that there is any difference in the number of macroplastics found on exposed beaches on the Coromandel Peninsula in comparison to sheltered beaches on the Coromandel Peninsula.

Scholarship level	**Outstanding Scholarship level**
• Describing the relationship in context. • Comparing and interpreting features of groups in context. • Describing trend(s) and key features for time series data in context. • Calculating E(X). • Providing suitable parameters for a probability distribution models. • Comparing relevant features of the sample distributions in context. • Interpreting the confidence interval in context. **Max 6 marks**	• Evaluating the goodness of fit of a given model with appropriate probability calculations. • Discussing conclusions that can be made including the appropriate extension of results. **Max 2 marks**

QUESTION FOUR

Task Q4(a)(i)

Evidence:

- The median age of people who always or almost always actively recycle (53 years-old) is higher than the median age of people who sometimes or never actively recycle at (39 years-old).
- Of those who always or almost always recycle, the median age of respondents who use the recycle bin when uncertain (56 years-old) is higher than those who use the rubbish bin when uncertain (52 years-old).
- When uncertain which bin to use, a higher proportion of those who sometimes or never recycle use the rubbish bin (95%) than the corresponding proportion of those who always or almost always recycle (78%). However, both groups are more likely to use the rubbish bin when uncertain.
- People who do not have kerbside recycling are less likely to always or almost always actively recycle (29%) than people who do have kerbside recycling (71%).
- Far more people have kerbside recycling (89%) than do not (11%).

Task Q4(a)(ii)

Evidence:

- Figure 12 shows only the behaviours of a representative sample of New Zealanders with kerbside recycling.
- Because this is not an experiment, it would be inappropriate to use these results to make a causal claim or to extend these results to other forms of waste beyond recycling.
- The group who does not have kerbside recycling, as shown in Fig. 12, appears to be very small, which may make the comparison (inference) less valid.

Task Q4(b)

Evidence:

- The total mean weekly weight of rubbish produced by Christchurch City Council households:
- $\frac{1522/2}{0.83} + \frac{936.1/2}{0.73} + \frac{92.7}{0.6} = 1712.53$ tonnes
- The number of households:

  $1712.5 \div \left(\frac{11.46}{1000}\right) = 149\,389$ households

Task Q4(c)(i)

Evidence:

- Normal Distribution: $\mu = 10.5$, $\sigma = 2.75$
  P(bin distance < 12 m) = 0.7073
- Binomial Distribution: $n = 5$, $\pi = 0.7073$
- P(all 5 group members recycle) = 0.1770

Task Q4(c)(ii)

Evidence:

- Assume that the five people are independent in their distance from a recycling bin. One person's location will not have an effect on another person's location.
- Assume that the probability of any person being < 12 m from a recycling bin remains constant at 0.7073.

Scholarship level	Outstanding Scholarship level
• Linking features of data to contextual information. • Discussing the inappropriate extension of results. • Identifying the need for an experiment to support a causal claim. • Calculating the number of households. • Selecting and using a suitable probability distribution model to calculate one relevant probability. • Discussing one relevant assumption in context. **Max 6 marks**	• Calculating an appropriate probability using two probability distribution models. • Discussing two or more relevant assumptions in context. **Max 2 marks**

Sufficiency Statement

For each question:

Score 1–4 No award	5–6 Scholarship level	7–8 Outstanding Scholarship level
Shows understanding of relevant statistical and probability concepts and methods, and some progress towards applying this in context.	Application of high-level statistical analysis and critical thinking, knowledge and skills, to complex situations. Shows logical development, precision, and clarity of ideas.	In addition to the requirements of Scholarship, demonstration of perception and insight, sophisticated integration and abstraction of ideas, independent reflection and extrapolation, and convincing communication.

Cut Scores

Scholarship	Outstanding Scholarship
18 – 27	28 – 32