Assessment Schedule – 2022

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

Q1(a)

Evidence:

- NZ farm area is decreasing, while the amount of fertiliser seems to be increasing (nitrogen). This suggests an intensification of agriculture as suggested in the first paragraph. This can also be seen in the increase in the total number in cattle over time.
- While the number of sheep has reduced from about 60 million from 1975 to about 28 million in 2015 (around 47% decrease), the number of dairy cattle has increased from about 3 million to about 6 million in the same time interval (around 100% increase). This is linked to the second paragraph, which describes how dairy cattle have a higher environmental impact than beef cattle and sheep.
- Both the total farm area and the number of farms in NZ have decreased from 2002 to 2019. Total farm area: 2 mil ha / 15.6 mil ha = 12.8% decrease, Total farm count: 21 000 / 70 000 = 30% decrease, this suggests that the average farm has increased in size. This and the generally increased used of fertiliser can be linked to the intensification of agriculture as suggested in the first paragraph.

Q1(b)

Evidence:

- The high and low values in both graphs have been coded with the same colour.
- For example, areas with a high nitrogen concentration are indicated with a yellow dot on the left graph and high proportions of human modified land cover in the upstream catchment are shown in yellow in the right graph.
- This allows people to identify patterns in these two metrics and identify areas in New Zealand where there seems to be a correlation between the two.

Q1(c)(i)

Evidence:

- The trend is that the value of imported fertiliser is increasing from about $10 million per quarter in 1988 to about $160 million in 2021 at about $5 million per year.
- There is evidence for seasonal variation throughout the year and the seasonal variation has been increasing over time, with the largest seasonal fluctuations between 2012 and 2021 (not including the “spike” in 2009).
- There are several spikes in the fertiliser data (in 2009, 2012 and 2021) where the import in one quarter was much higher (about $200 million) than elsewhere.

Q1(c)(ii)

Evidence:

- It seems that Q4 of 2021 had a much higher value (about $150 million higher) than the earlier quarters.
- The model seems to assume that this value is the new baseline, because the forecasted red line is much higher than the values of the years leading up to this point.
- These spikes in the value of imported fertiliser have happened before, around 2009 and 2012, but the following quarters came back to the earlier levels.
- The model is likely overestimating the value of imported fertiliser in Q4 2022 and should be used with caution.
- Some of the increase could be due to inflation or exchange rates, not actual increases in amount of fertiliser.

QUESTION TWO

Q2(a)(i)

Evidence:

- The cows received all three treatments to control the variation of the response variable between cows.
- For example, one cow may produce more methane than others, but by giving all treatments to all cows, the difference in the response variable compared to the baseline (control) can be measured for each cow and for each treatment.

Q2(a)(ii)

Evidence:

- Since each cow received all three treatments, including the control (0% seaweed), the methane emissions for each cow under each treatment would need to be used to calculate two paired differences:
  - low additive compared to control
  - high additive compare to control
- To determine that the decrease in methane emissions was significant, researchers would need to use these paired differences to test that they were larger than what could be expected due to chance alone.
- A randomisation test based on comparing two independent groups using a difference of two means / medians would **not** be appropriate for this study design.

Q2(b)

Evidence:

- This is a causal claim, so a randomised experiment is needed.
- The treatment variable could have two levels: whether the cow feeds on a paddock that is treated with seaweed fertiliser or whether the cow feeds on a paddock that is an untreated (or treated with existing fertiliser).
- The response variable is the amount of nitrogen in the cows' urine.
- The experimental units are the cows.
- It would be a good idea to control other variables, such as:
  - The differences in the different paddocks (perhaps different plants grow there).
  - The differences in the cows' natural urine content. This could be done by comparing each cow's urine content for both treatment and control i.e., using a paired comparison design.

Q2(c)

Evidence:

- The headlines were distributed reasonably uniformly across the months in the year, indicating articles about the climate are published somewhat regularly throughout the year.
- The distribution of sentiment scores has a median of about 0.55, and seems slightly negatively skewed, suggesting that the headlines tended to be more positive in sentiment.
- The distribution of sentiment score for headlines containing the word 'change' is shifted slightly higher than that for headlines not containing the word 'change', with respective medians of about 0.57 compared to 0.54.
- Most of the headlines (over 55%) containing the word "climate" also contained the word "change".

Note: The expected analysis is not inferential. All headlines from these years (2018 – 2020) with the word "climate" in their headline were used for analysis.

QUESTION THREE

Q3(a)(i)
Evidence:

- Horizon Research could have pre-screened participants by asking them questions related to gender, location, age, ethnicity, income, etc., and then only included participants who would ensure the weighting of these features within the sample was similar to the census population.
- Horizon Research could have included some questions related to gender, location, age, ethnicity, income, etc., and then "post-weighted" the survey data based on these features to reflect the census population.

Q3(a)(ii)
Evidence:

- The smallest sample percentage point increase was from 2012 to 2014: 2%.
- Estimated maximum (reported) margin of error (MOE) = $\frac{1}{\sqrt{1097}} = 3.0\%$
- Two independent samples, so MOE for the difference = $1.5 \times 3.0\% = 4.5\%$
- CI for the difference of two proportions: (–2.5%, 6.5%)
- As the CI contains both positive and negative differences of proportions, there is not enough evidence to suggest that there has been an increase every time in the proportion of ALL NZers that consider climate change a problem now.
- Because of this, the claim that this proportion has increased each year it was surveyed from 2012 to 2021 is not supported.

Q3(b)(i)
Evidence:

- P(problem) = P(urgent) + P(future)
- P(Y1–6) = 0.217, P(Y7–8) = 0.326, P(Y9–13) = 0.457
- P(problem | Y1–6) = 0.492, P(problem | Y7–8) = 0.737, P(problem | Y9–13) = 0.670
- P(problem) = $0.492 \times 0.217 + 0.737 \times 0.326 + 0.670 \times 0.457 = 0.653$
- P(Y1–6 | problem) = $0.492 \times \frac{0.217}{0.653} = 0.163$
- P(Y7–8 | problem) = $0.737 \times \frac{0.326}{0.653} = 0.368$
- P(Y9–13 | problem) = $0.670 \times \frac{0.457}{0.653} = 0.469$
- For those candidates who think that climate change is a problem, the highest proportion in is the Y9 – Y13 range and the smallest proportion is in the Y1 – Y6 range.
- The conditional proportions increase as the year levels of the groups increase, so the statement is supported.

Note: Other explanations are possible, and candidates may use a probability tree diagram to support their reasoning.

Q3(b)(ii)
Evidence:

- If the sample was representative of all candidates in NZ schools, the proportions of candidates in the different year groups in the sample would be very similar to the proportions of candidates in the different year groups in the population.
- In the sample, the proportion of candidates in the Y1 to Y6 group is only about 2 / 3 of the proportion of candidates in the Y7 to Y8 group, but the proportion should be much larger as it contains candidates from three times as many year levels. Therefore, it appears that the sample is not representative in terms of year levels.

QUESTION FOUR

Q4(a)

Evidence:

- Every person in each photo would be recorded as either being a school-aged child or not.
  - The data from these photos would then be used to calculate a sample proportion i.e., how many people in the photos would be the denominator, and how many people who were school-aged would be the numerator.
  - Bootstrapping could then be used to construct a confidence interval for the underlying / true proportion of all school-aged children at the event.
- There are several potential limitations with this method, for example
  - It is difficult to classify candidates correctly as school-aged or not. This is particularly difficult for people aged between 17 and 20.
  - It is likely that people in photos are either leading the march or hold more sophisticated signs, because this makes for more interesting photos. These people, however, are likely to be older than those in the march in general.
  - Some people might appear in more than one photo, and so might be counted several times in the data / analysis.
  - Younger children are shorter and so might be harder to see in photos.

Q4(b)(i)

Evidence:

- One assumption is that people in the crowd is distributed reasonably uniformly. This is probably not true, because people tend to stand closer together in some areas, for example close to the stage.
- Another assumption is that the area the crowd occupies is entirely "free space" and isn't limited by trees, buildings, or other obstacles.
- Another assumption is that the perimeter of the crowd is clear.

Q4(b)(ii)

Evidence:

- Dividing the total area by the 5 by 5 metre squares results in 720 "squares", $\frac{18000}{25} = 720$
- Multiplying the limits of the confidence interval by 720 to obtain estimates for total crowd size, $(20.4 \times 720, 23.2 \times 720) = (14\,688, 16\,704)$
- As 17000 is outside the confidence interval for the total number of people attending (so is not a plausible value), the claim is not supported.

Q4(c)

Evidence:

- The Poisson model makes sense in this context as it is about an average number of discrete occurrences (natural disasters) in a continuous interval (1 year).
- The shape of the distribution of frequency of weather-related natural disasters is positively skewed which is an expected shape for a Poisson model with $\lambda = 4.2$.
- The higher-than-expected number of 8 disasters per year can be explained by random variation, given the relatively small sample size ($n = 41$).
- One issue with using a Poisson model with $\lambda = 4.2$ is that the number of weather-related natural disasters seems to be increasing over time, which could mean that the assumption that λ has a constant value is not appropriate.
- Independence of events is also a potential issue, as weather events can be related for different natural disasters.

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
20 – 25	27 – 32