Assessment Schedule – 2016

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

Tasks Q1 (a)
Evidence:
Figure 1

Tasks Q1 (a)
Evidence:

- For State X, the ozone concentration displays a weak to moderate positive (linear) relationship with the city population size.
- For State Y, the ozone concentration has a strong, positive, linear relationship with the city population size. There is an outlier at (2,120).
- For State Z, there is a moderate positive non-linear relationship between ozone concentrations within a small range of city population size (0.3 to 1.1 million).

Note:

1. Credit is given for only one dot point per graph.
2. For State Y, both statements are required. Must have linear.
3. For State Z no marks for just stating "moderate positive". Need to have something more like non-linear.
4. If no context, penalise one mark overall.

Task Q1(b)
Evidence:

- Ozone (p.p.b.) = 4.1495*1.4 + 123.74 = 130

  Validity is suspect because of the variation shown in the scatterplot near 1.4 million population. Correlation is moderate suggesting other factors influence ozone concentration as well as population size.

Note:

1. One mark for making a prediction, then either two marks for reservation about its validity or one mark supporting prediction and one mark for reservation.

Task Q1(c)
Evidence:

- Sample size of 10 cities is small from each state.
- Coverage of each sample focusses only on large population centres within each state and will probably not be representative of all locations within the state.
- No data available about state population nor how many cities in each state.

Task Q1(d)
Evidence:

- A large representative sample of ozone concentration levels would be taken from each state. A statistical summary of the ozone concentrations would be constructed for each state and a comparison made regarding the differences in ozone concentrations. A bootstrap distribution would be constructed in order to estimate the size of the difference.

Note:

1. One mark for mentioning overlap in confidence intervals.
2. One mark for how sample was taken and one mark for an analysis comment.

Overall Judgement for Q1
Max 3 for (a), max 3 for (b), **max 5 for (a) + (b)** then max 2 per reservation for (c) and max 2 for (d).
Overall **max 8 marks**.

QUESTION TWO

Task Q2(a)(i)
Evidence:

- At the end of the study, the mean score was 61.68% for M1 and 62.52% for M2 giving a difference of 0.84% (refer Fig 5).
- The re-randomisation distribution for the mean of the difference between the test scores for M1 and the test scores for M2 has a tail proportion of 0.432. A difference of 0.84% or more could occur by chance alone with a probability of 0.432 **so** this study gives insufficient evidence (43.2% is much greater than a 10% threshold) to claim that there is a difference in the effectiveness of the two methods of teaching.

Note:

1. Not necessary to mention actual mean scores for M1 and M2 but if not mentioned, difference must be stated.

Task Q2(a)(ii)
Evidence:

- Depending on which teaching method the student had been assigned to, either the mean or the median score of the other 24 students in the group could be used to estimate the student's mark.
- If another set of scores was available for the group, then a fitted regression line could be used to estimate the student's mark.

Note:

1. Choice of mean or median is determined by the distribution, symmetrical or skewed respectively, of the other scores.

Task Q2(a)(iii)
Evidence:

- Over the teaching period of the course the students could have had additional influence from various exposures to issues to do with global warming which could have influenced the outcome of the test.
- Students taught by M1 share notes with those taught by M2.

Note:

1. Quality of teachers not accepted as an influencing factor.

Task Q2(b)
Evidence:
To carry out the experiment, the following steps could be performed:

- Students would be tested before the course to ascertain background knowledge.
- The course would then be taught.
- Students would be tested at the end of the course with a similar test to the pre-test.
- For each student the difference between the pre-test and post-test score would be calculated.
- A statistical summary of these differences in scores would be produced for analysis.
- A randomisation test would be carried out on these differences.
- We would conclude from the randomisation test how likely the mean of the differences in scores could occur by chance alone, in order to reach a conclusion about the effect of the course.

Note:

1. Allocate three marks for: Pre/Post-test, the analysis and how you interpret?
2. A range of other methods are accepted for the analysis.

Task Q2(c)

Evidence:

- The mean test score for Group G1 was 5.04% more than the mean score for Group G2.
- The test scores for G2 are more consistent than for G1. (IQR 19% versus 22% or standard deviation 11.3% versus 16.0% or range was 40% to 73% versus 32% to 90%).
- The mean test score % for Group G1 is likely to be between 13.03% and –2.43% more than the mean test score for Group G2, OR zero is included in interval so no evidence to establish that Group G1 had the greater background knowledge.

Note:

1. Any two distinct comments of a central measure comparison, a spread comparison **plus** a bootstrap comment is acceptable for full marks.

Overall Judgement for Q2

Max 2 for (a)(i), 1 mark for each of (a)(ii) and (a)(iii), max 3 for (b). **Overall (a) + (b) max 6.** For (c) 3 marks max.

Overall max 8 marks.

QUESTION THREE

Task Q3(a)(i)
Evidence:

- We have Normal Distribution with mean 0.64% and standard deviation 0.12%.
  So pr(> 0.7) = 0.3085
- We have a Binomial Distribution with n = 10 and p = 0.3085; so if x = no of years where there has been at least a 0.7% increase, then pr($x \geq 5$) = 0.1653

Note:

1. One mark for each answer.

Task Q3(a)(ii)
Evidence:

- We create a triangular distribution with range 0.28% to 1.00% with the height calculated to be 2.78 at the mode of 0.82%. The probability of the yearly increase being at least 0.7% is 0.5466. Hence the pr ($x \geq 5$) is 0.7312.

Task Q3(a)(iii)
Evidence:
The triangular distribution for the following reasons:

- Estimated number of years where there has been at least a 0.7% increase in major weather events is 46.5. So probability is 46.5 / 70 = 0.66, which is closer to 0.5466 than 0.3085 obtained using the normal distribution.
- Modal interval is 0.76% to 0.88%, which contains the mode 0.82% for the triangular but not the mode of 0.64% for the normal.
- Distribution has a negative skew (mean = 0.75 < median = 0.79) whereas the normal distribution is symmetrical with mean = median = 0.64%.

  Note:

  1. No data required in skew reason.
  2. Two distinct reasons are required.

Task Q3(b)(i)
Evidence:
Take the following steps:

- Calculate the mean number of floods per annum as 154 / 70 = 2.2
- This mean of 2.2 is used as an estimate of the Poisson parameter λ.
- Calculate firstly the Poisson probabilities, then the expected number of floods for each number of floods per annum, to get the following:

Number of Floods	0	1	2	3	4	5	6	7
Frequency	6	18	22	12	7	3	1	1
Poisson Frequency	7.76	17.07	18.77	13.76	7.57	3.33	1.22	0.39

- Comparing the observed frequency with the Poisson frequency we get a close fit. This would suggest the Poisson distribution as being an appropriate model fit.

Note

1. Only one mark for the following: Expected number of floods per annum is 2.20. Variance of the number of floods per annum is 2.10. Close agreement between mean and variance, which is a property of the Poisson distribution.
2. Four matches of frequency or probability are sufficient for two marks otherwise one mark for fewer matches.
3. Can match probabilities rather than frequencies as below:

Number of Floods	0	1	2	3	4	5	6	7
Probability	0.0857	0.2571	0.3143	0.1714	0.1	0.0429	0.0143	0.0143
Poisson Probability	0.1108	0.2438	0.2681	0.1966	0.1081	0.0476	0.0174	0.0055

4. No marks for either cumulative probability or cumulative frequency matching.

Task Q3(b)(ii)
Evidence:
We assume the floods occur randomly and independently from one year to the next.
Note:

1. One mark for any two of: randomly, independently and not simultaneously.

Overall Judgement for Q3
Max 3 for (a)(i), max 2 for (a)(ii), max 2 for (a)(iii). **Overall so far max 6.** For (b)(i) 2 marks max, and (b)(ii) 1 mark.
Overall max 8 marks.

QUESTION FOUR

Task Q4 (a)
Evidence:

- **Transport**:
  Transport has consistently produced the highest amount of CO_2 emissions. There was an increasing trend in CO_2 emissions due to transport in NZ overall from about 8.5 million tonnes in 1990 to 14 million tonnes by 2012. The greatest increase occurred from 1990 – 2004 and after that the trend has plateaued.
- **Manufacturing and Construction industries:**
  Over the period 1990 to 2012, the overall trend in CO_2 emissions has remained about constant, and apart from a few years, manufacturing and construction industries have recorded the lowest emissions. The emissions did rise between1990 to 2002 (5 million tonnes – 7 million tonnes). Between 2002 and 2004, there was a sharp decline back to 5 million tonnes, and since then the amount of CO_2 emitted due to manufacturing industries has fluctuated around the 5 million tonnes level.
- **Energy industries:**
  The CO_2 emissions from energy industries has shown a slight increase overall but has shown the most fluctuations during this time period, possibly due to changes in the economy. In 1990 energy accounted for about 6 million tonnes of CO_2 emission: by 1996 this had dipped to the lowest at about 4.5 million tonnes. In 2005, the highest emission was recorded (just over 10 million tonnes), after which the emission from energy has decreased again to about 7 million tonnes by 2012.
- **Proportional Comparison:**
  In 1990, manufacturing industries produced about half the CO_2 emissions that transport did, energy industries about 2 / 3 the amount that transport produced. By 2012, manufacturing produced 35% the CO_2 that transport produced and energy industries produced about half the emissions of transport.

  Note:
  1. No further mark for repeated comments.

Task Q4 (b)(i)
Evidence:
A scatterplot would give a quick visual comparison of methane and total emissions for each country. It would be obvious if a country was high in one but not in the other or high / low in both.
Note:
1. Multiple or Component-Part Bar Chart with description was acceptable.

Task Q4 (b)(ii)
Evidence:
Australia: Methane 5 × 22.7 million = 113.5 million kg
GHG 24 × 22.7 million = 544.8 million kg

NZ: Methane 6.5 × 4.4 million = 28.6 million kg
GHG 17 × 4.4 million = 74.8 million kg

Note:
1. Accept variations in answers due to estimates from graph.
2. Three correct calculations out of the four are required for the mark.
3. Not acceptable to conclude that Australia's Methane was 84.9 million kg more than NZ's.
4. Comparing methane and GHG emissions per person was acceptable.

Conclusion:
In absolute terms, Australia produced 7 times the amount of GHG and about 4 times the amount of methane that NZ did.

Task Q4 (c)(i)
Evidence:
Each year the CO_2 in the atmosphere is reasonably constant through spring and summer, then increases sharply over autumn and winter.
Note:
1. In 2001, the increase was not so sharp (up about 1 ppm) but in 2002 the increase was greater than any other year (up by 3 ppm).

Task Q4 (c)(ii)
Evidence:
Double = 560 ppm. At 2005, at 374.8 ppm.
Pessimistic: Grad = steep grad based on 2002–2005:
Grad = (374.8 – 368.1) / 3 = 2.23
Increase would take (560-374.8) / 2.23 = 79 years so double by 2084

Optimistic: Grad = average over all graph
Grad – (374.8-364.2) / 6 = 1.66
Increase would take (560-374.8) / 1.66 = 111 years so double by 2116.

Roughly between 80 – 110 years to reach double the mid-1800s level.

Note:

1. Accept a mid-point analysis based on 2ppm per year for one mark.
2. Must use two different gradients to get two marks.

- Answer is highly suspect, as the Baring Head data spans only six years, covers only a small region of the globe, and doesn't go back to pre-industrial times.

  Note:

  1. Must have two justification points.

Overall Judgement for Q4
Max 3 excluding proportional comparison for (a), max 1 for (b)(i), max 1 for calculation excluding conclusion in (b)(ii), max 1 for (c)(i) and max 2 excluding justification for (c)(ii). **Max total so far 6.**
Then 1 for proportional comparison in (a), 1 for conclusion in (b)(ii) and 1 for justification in (c)(ii).
Overall max 8 marks.

QUESTION FIVE

Tasks Q5 (a)

Evidence:

- Distance from the coast – closer to the coast the temperatures do not fluctuate as much as inland. The sea has a moderating effect on temperature.
- Altitude – higher altitudes would tend to be cooler
- Latitude – the further North, the warmer it tends to be
- Urban / rural – the more building / asphalt / human activity, the higher the temperature
- How large is the area that the station represents in terms of temperature – smaller microclimates should have a smaller weighting in the average.

Note:

1. Any two with comment are required for full marks.

Tasks Q5 (b)

Evidence:

NZ graph:

- There are more red bars than blue after 1970, meaning more years where the temperature was above the base average.
- Over the past 12 years, 9 years have had above average temperatures.
- **Interpretation of gradient:**
- The trend suggests an average increase in temperature of approx. 1°C per century (most likely between 0.67°C and 1.25°C).

Global and Antarctic graph:

- Global average temperatures have been above the base temperature since 1980.
- **Comparison of Gradient:**
  Antarctic average temperatures show a greater increase in trend than global temperatures (index changes by1.93°C, 0.57°C which means global temp risen by 0.0057°C / year or 0. 57°C / century, Antarctic temperature risen by 0.0193°C / year or 1.93°C / century).

Note:

1. One mark per point excluding comparison of gradient.

Tasks Q5 (c)

Evidence:

Prior to 1960, the mean annual temperature anomalies for Antarctica fluctuate greatly, with a difference in the index of close to 300 points some years. After 1960 the fluctuations are smaller, with the largest fluctuation from 1 year to the next of just over 100 points.

Note:

1. One mark for general statement about changes in fluctuations and one mark for actual values.

Explanation:

A possible reason for this could be improved collection of data – more stations used to record data, stations scattered more widely over the Antarctic region, technology allowing remote sensing.

Note:

1. One reason is sufficient.
2. No marks given for any duplication of answer to (b).
3. A non-statistical explanation is acceptable with context.

Overall Judgement for Q5

Max 2 for (a), max 4 for (b), max 2 for describing difference in (c). **Max total so far 6.**
Then 1 for interpretation of gradient in (b), plus 1 for explanation in (c).
Overall max 8 marks.