SUPERVISOR'S USE ONLY

S

93201A

932011

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

+

# SCHOLARSHIP EXEMPLAR

Mana Tohu Mātauranga o Aotearoa
New Zealand Qualifications Authority

# Scholarship 2023
# Statistics

Time allowed: Three hours
Total score: 32

ANSWER BOOKLET

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

Write your answers in this booklet.

Make sure that you have Formulae Booklet S–STATF.

Show ALL working. Start your answer to each question on a new page. Carefully number each question.

Check that this booklet has pages 2–24 in the correct order and that none of these pages is blank.

Do not write in any cross-hatched area (). This area may be cut off when the booklet is marked.

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

1a) The use of electricity for heating & cooling across all 4 relative sectors differs – though both residential & commercial use similar proportions of approx ~60% of electricity usage for each residential & commercial respectively using electricity being for heating and cooling. For the industrial sector, only about ~20% of electricity use was for heating & cooling, while ~~the~~ ~80% of total electricity used in agriculture was used for heating and cooling. The use of electricity for electronics and lighting is similar between the residential & commercial sectors – about 20% of total electricity use for residential powers all 100% of electronic & lighting. 30% of total electricity use for the commercial sector is also used, ~~with~~ for lighting, and like commercial use 100% of electronics is powered by electricity. On the other hand industrial only uses about ~20% of their total electricity to electronics, and about 30% of total electricity use for agriculture was used for lighting. Agriculture has the greatest proportion of diesel use out of all sectors at about 45% of all total energy use being diesel, whereas industrial only uses about 10% diesel in total, commercial using about 10% as well, and residential using ~~5%~~ < 5% diesel energy in total, compared to their total energy use. Geothermal energy appears to be most prominently used in the commercial sector, with ~10% of total energy use being geothermal. Industrial uses about ~5% geothermal out of total energy use, and agricultural using about 2 or 3% geothermal out of total energy use. Residential energy use has the lowest proportion of geothermal energy use at about <1% ~~the~~ use relative to their total energy use types.

(b) difference $\pm$ 2 × mae

$= (\text{heat} - \text{electric}) \pm 2 \times \sqrt{\frac{1}{3039}}$

$= \cancel{21} \; \cancel{\times} \; \cancel{22}$

$= 21 \pm (2 \times 1.8139)$

$= [\ 17.37\%,\ 24.62\%\ ]$

With 95% confidence I can say that about ~~17.37%~~ ~~the~~ heat pump usage in NZ households will be at a range of 17.37% to 24.62% higher than electric plugin heaters. The entire confidence interval is positive, suggesting that heat pumps are verly highly likely to be greater in ~~percentage~~ percentage use in NZ homes by 17.37 to 24.62% than plug-in heaters.

(bii) In both before and after for the ~~winter~~ mean cost in \$ for ~~both~~ the winter graph, the total ~~number of hours that~~ mean cost in \$ has generally stayed the same. Notice in Figure 4, there is a generally increasing trend in winter of mean cost per hour ~~of about~~ starting at \$0.0156 mean cost at 0 hours to about \$0.3375 mean cost per hours at the peak at the 21st hour, giving a \$0.3219 increase, before heat pump was installed. After the heat pump is installed for winter, the start mean cost per hour is about \$0.0156 — same as pre-heatpump installation. The peak for the after heat pump graph at winter is at the 20th hour but is still similar to the peak at the 21st hour of winter's pre-installation graph at about \$0.3125 mean cost. This suggests a very little change in how much they spend in winter pre-and-post heatpump installation. In autumn, there appears to be a significant ~~decrease in the~~ decrease in the peak at the 16th hour from before and after heatpump installation — before it was at ~~\$0.22~~ approx \$0.218 mean cost per hour but comparing the peak at 16th hour post-heatpump installation, ~~this could be because~~ it is at \$0.15625, a near \$0.062 mean cost per hour decrease. ~~HOWEVER,~~ I notice that the ~~last month of autumn,~~ June, is ~~not recorded,~~ and This is shown in fig. 3 as the increase in total cost in \$ from april to may pre-heatpump used to be at about \$46 increase, but post-heatpump it is only a \$15 increase, suggesting the heatpump use in may also may have lead to decreased power use in the 18th hour in the autumn season. The total cost of electricity appeared to increase significantly from pre- ~~and~~ to post- heatpump use from a value of \$121 total in January 2021 to \$179 total in Jan 2022. The peak in mean cost in \$ for after heatpump installation is also significant, going from a \$0.104 peak in summer pre-heatpump to a \$0.212 peak in summer post-heatpump - a \$0.108 dollar difference in mean cost. This is likely because heatpumps also often double as air conditioners, so now the household uses more electricity in the summer now.

1biii) Figure 2 uses 2 seperate graphs to show pre- and post electricity usage for each season, making it visually hard to compare differences in peaks and troughs for mean cost in \$ per season. So a better way to visualise cost data would be to create 4 seperate time series graph, with same x and y axis but seperated for each season. Two different colours can be used to differentiate 2021 and 2022 data and this way we can clearly see the differences in mean cost for electricity between certain hours of the day for each season, before and after the heatpump is installed.

2ai) The way the deviation of 0.57% may have been calculated was by taking the percentage difference between actual - predicted solar consumption, for each month dividing by the actual solar consumption, adding all together and dividing by 12 (amt of months).

Example: December's predicted: 5250 kWH, actual is ~~5~~ 5500 kWH.

$\frac{5500-5250}{5500}$ = 4.5% difference.

If actual < predicted, then we will get a -ve percentage, make this positive instead as we dont care if it is an under or over prediction - just how much the prediction is innaccurate. Add all values like the example above and divide by 12 months to ~~get the~~ get the average deviation.

aii) use binomial as:
Two outcomes - either yes predicted is higher than actual, or no,
Fixed probability - of 7/12 chance predicted is greater than production.
$P(X > 6) = 1 - P(X \leq 6)$ : Ncd
probability = 7/12
x = 6
ntrial = 12
A assuming 12 months

$P(X > 76) = 0.6203$

This suggests that 62.03% of the time, predicted production will be greater than actual production in the timespan of about a year.

more than half the time

2bi) The data shows the original difference between mean alertness score for incandescent group is 0.45 greater than the LED group. The rerandomization tests randomly assigns either treatment or control results ~~result groups~~ to either treatment or result groups and records the difference in their means 1000 times. This test determines if the results of our experiment were due to our treatment and not simply by chance. If the proportion of our original difference between the means is less than 10% (tail end of the graph) then our experiment's treatment has an effect on the response. Our randomization test shows a proportion of

17.8% for our results, hence we cannot make the call that our treatment of incandescent lights help visually than LED lights.

2bii) Using a paired comparison experiment, each individual student will do the test (3 news articles) for both treatment (incandescent) and control (LED). We record their response variable for each trial (treatment and control) and plot their results into 2 seperate ~~bar~~ box and whisker plots. An arrow graph can be used to show each individual student's change in alertness score ~~no. of news articles proofread~~ to see how much they improved in reading ~~pre and~~ with each condition. This can also show ~~how much~~ if the treatment was more effective ~~than~~ for those of high process ability compared to low ones. Our experimental units are students, response variable is alertness score, treatment and control is LED and incandescent lights respectively. If a ~~m~~ majority of students did see an increase in alertness score then we can make the causal claim. If only students of a certain threshold of allertness did, then further investig-ation is needed. ~~Random allocation is still needed for single treatment and control~~

No rerandomization test is needed.



01183

3ai) In both Sydney and Vancouver graphs there is a slight right skew suggesting a greater clustering of the lower 50% of values than the upper 50%. Using the mean would result in the measure of center being less than the median and wouldn't accurately capture the higher 50% of ~~value~~ values above the median.

3aii) Standard deviation is a measure of spread from the mean. about $3\sigma \pm$ mean cover 99.7% of the whole data.
Using estimates:
Sydney: 60
0 to ~~67.50~~% $\approx$ 99.7% of all data.

$\frac{\cancel{67.25}\ 60}{6} = \cancel{11.25}\ 10 \quad \sigma = \cancel{4.25}\ 10\%$

Vancouver

$\frac{67.5}{6} = 11.25 \quad \sigma = 11.25\%$

iii) Use ~~normal distrib~~ ~~poisson, as lower limit = 0%,~~ normal dist.
mean is: $3 \times SD$ =
Vancouver: 33.75 %
Sydney: 30 %

Vancouver:

$P(x < 10) = \cancel{1 - P(x \geq 9)}$

Vancouver : P C.D

$\mu\ \cancel{x} = 33.75\%$ $\sigma = 11.25$

upper = ~~9~~ 10 lower = −9999

$P(x < 10) = 0.0173$

sydney: P C.D

$\mu\ \cancel{x} = 30\%$ $\sigma = 10$

upper ~~x~~ = ~~9~~ 10 lower = −9999

$P(x < 10) = 0.0227$

3b) In an observational study, we would have to randomly select a ~~representative~~ sample from both groups, light + dark coloured roof houses, accounting also for variability such as tree shade, location in city (some areas could be hotter than others), etc. this allows us to create a representative sample. However in an experiment, we would have to randomly allocate houses from Wellington either treatment or control groups - either colour rooves light or colour them dark, this prevents bias, like if all houses in a hotter area had dark rooves, then the affect of colour of roof on indoor temperature may be significantly higher than it actually is, which could just be because of the ~~lurking~~ confounding variable of the houses being in a hot area, hence random allocation is needed to prevent these biases. Both studies will measure the same variable of avg. indoor temperatures, but only the experiment can make a causal claim. Before a causal or suggested claim (for sample-to-population) can be made, a bootstrap confidence interval must be made for the observational study to determine whether or not lighter roofed houses do/dont tend to have lower indoor temperatures than dark ones. If the confidence interval is entirely positive then we can make a call that house roof colour tends to influence indoor temperatures. Similarly a rerandomization test must also be used for the experiment to ensure the results were not from chance but from the treatment. If the proportion of our mean/median difference in indoor temperatures is ~~above~~ <10% ~~then we~~ of the rerandomization test then we can make the call.

4a) One sampling strategy was to ~~randomly choose~~ use ~~either Auckland ad~~
use Google Maps to generate an aerial view of the suburbs in both Auckland and Christchurch. One issue would be first, when was the image taken? If the image is old then it may not be representative of

the present-day greeness score of the two cities. Another sampling method was randomly taking 1000 pixels from whatever coordinates were provided for the given suburb. the issue is that there is no indication of how these coordinates are generated - if the coordinates focus on the "town" part of a suburb then naturally the greeness score would decrease. ~~compared to a more~~ This is hardly a reliable measure of greeness. And it would increase if the Google Maps suburb coordinates landed on the park of the ~~town,~~ suburb though this is likely not representative of the greeness of the suburb. Finally a random sample of 100 suburbs were selected accross both cities, but Auckland City may be significantly larger in size with more suburbs compared to Christchurch ~~hence~~ hence there is a grea-ter chance that Auckland ~~cities~~ suburbs have more variation in their greeness scores than Christchurch.

3b) I can conclude that in our original sample, Auckland has a greater ~~greenes~~ greeness score than Christchurch with about 263.31 greeness to 220 greeness in Christchurch. Auckland appears to lead by 43.31 points in our sample. However to determine if the results are applicable to the population of all Auckland suburbs, the bootstrap confidence interval is done to determine the likelihood of Auckland being greater in greeness than Christchurch via sampling with replacement 1000 times. The confidence interval clearly shows that ~~there are~~ 0 is included in the confidence interval, suggesting that ~~there are~~ it is likely that Christchurch may have a greater greeness score than Auckland, even over 51.71 points greater. This suggests there is no difference between greeness scores in both Auckland ~~&~~ & Christchurch. ~~due~~ Another thing to note is that Christchurch only has about 20 data points ~~t~~ auckland appears to have a significantly larger number of data points. So perhaps christchurch's dataset is NOT representative ~~due~~ due to its small sample size, and another sample must be taken.

Q4

3c) There is a linear positive moderate-to-weak trend between distance to city and greeness score for ~~both the combined~~ Auckland city. However for Christchurch it appears to be more of a non-linear strong graph with a positive trend. I notice Auckland has S a greater dist. to city center at about 65km being the largest distance compared to Christchurch's 35km - likely because Auckland is a bigger city. Using a linear trend for the Auckland graph, I noticed ~~some~~ moderate-to-lots of non constant scatter ~~from~~ all throughout. The vertical variation at the 10km range was at about ± 112.5 greeness score from my trend line ~~,~~ and about ± 235 greeness score at 57km. Christchurch has little constant scatter at about a maximum of ± 100 greeness score at 6km, excluding the extreme x and y points, at ~(4, 620) and ~ (35, 610). Due to this, I can say Auckland tends to have greater variation in their ~~mean~~ greeness scores than Christchurch. Christchurch also generally only has nearby suburbs to the city center, but where there are datapoints they are generally similar to the scatter ~~of~~ of Auckland at those points at 2-12km. As distances increase from city I can say greeness is likely to increase, at a less steep rate for Auckland than Christchurch. ~~However~~ Though I would not use my model for extrapolating data for potential greeness scores at >20km for Christchurch as at these points greeness score would be >1000 which is not even possible for the furthest of Auckland's suburbs, as the max greeness score was about 920 greeness score at 59km for auckland. ~~It is unlikely~~

Q4

3a) One way the sampling process may be changed is to use stratified sampling in order to represent the ~~greater~~ suburbs with greater distance from the city center in Christchurch. As observed, Christchurch has no datapoints >40km, whilst auckland has over 10. Also, at 12 to 35 km Christchurch again has no data. this makes it difficult to compare greeness at these points. Stratification would identify subgroups to represent in the main population, randomly select from each strata, and then ensure each represented strata in the sample is proportional to the population. This way, if ~~a~~ certain percent of all Auckland suburbs were from 10-15 km, then only that percentage of our total sample would be in our sample. this makes it more fair for us to fully represent each suburb per city.

Another isse is that christchurch has about 20 datapoints only to Auckland's ~~[illegible]~~ ~40+. This larger sample size means Christchurch is more likely to be underrepresented and also its greeness score may not be as accurate compared to Auckland.



01183

If we increased Christchurch's sample size, this could lead to better predictions in greeness score, especially at ranges 40+km, allowing for a more accurate trend line and better comparison in greeness score with auckland, especially in ranges 12-35km and ~~40+km.~~ 40+km.

## Scholarship

**Subject:** Statistics

**Standard:** 93201

**Total score:** 23

| Q | Score | Marker commentary |
|---|---|---|
| 1 | 7o | This question provides evidence for 7o because the candidate has gained at least 1 o grade and at least 7 marks overall.<br>The candidate has two sufficient comparisons across different diagrams of fuel and fuel use including estimated proportions. They have identified and applied the correct margin of error calculation with an indication of some uncertainty. The candidate has identified two appropriate differences or similarities in trends before and after heatpump installation across figures 3 and 4 with numerical evidence. They have not identified a valid difficulty when interpreting figure 4, however, they have identified an appropriate way to visualise the data for easier comparison. |
| 2 | 6s | This question provides evidence for 6s because the candidate has gained at least 6 marks overall.<br>The candidate has provided a description and example of an appropriate deviation calculation. The candidate has correctly identified the use of binomial distribution and has given two key features in context. They have used an incorrect probability so have not calculated an appropriate probability. The candidate has correctly interpreted the difference between the two means and has correctly used the tail proportion to reject the claim. The candidate has discussed a paired comparison experiment but has not discussed random allocation of treatment order or discussed factors that may need to be controlled. |
| 3 | 5s | This question provides evidence for 5s because the candidate has gained at least 5 marks overall.<br>The candidate has not demonstrated an understanding of measures of central tendency and variation in context. They have identified parameters and have calculated an appropriate probability for Sydney. They have not identified an appropriate distribution model for Vancouver, however, they have gained a mark for continuity as they used appropriate parameters with the model to calculate a probability for Vancouver. The candidate has demonstrated understanding of the need for random allocation in an experiment and that an observational study cannot make a causal claim. They have not discussed the limitations of causal claims made from an experiment or random selection and the extension of results from an observational study. |

| | | |
|---|---|---|
| 4 | 5s | This question provides evidence for 5s because the candidate has gained at least 5 marks overall.<br>The candidate has identified two of the sampling scenarios and has discussed potential issues with one of these scenarios. They have not adequately compared sample means and have not interpreted the bootstrap confidence interval. There is some confusion around what it means when a bootstrap confidence interval contains zero so they have not gained credit for this question. The candidate has described the relationships of two cities from a bivariate data display separately and has made a valid comparison between the two cities. They have discussed how to change the sampling process to make comparisons fairer between the two cities. |