3

SUPERVISOR'S USE ONLY

91584

915840

NZQA

NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

QUALIFY FOR THE FUTURE WORLD
KIA NOHO TAKATŪ KI TŌ ĀMUA AO!

Level 3 Mathematics and Statistics (Statistics), 2015

91584 Evaluate statistically based reports

2.00 p.m. Thursday 19 November 2015
Credits: Four

Achievement	Achievement with Merit	Achievement with Excellence
Evaluate statistically based reports.	Evaluate statistically based reports, with justification.	Evaluate statistically based reports, with statistical insight.

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.

Pull out Resource Booklet 91584R from the centre of this booklet.

Show ALL working.

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

If you need more room for any answer, use the space provided at the back of this booklet and clearly number the question.

Check that this booklet has pages 2–11 in the correct order and that none of these pages is blank.

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

Excellence

TOTAL 20

ASSESSOR'S USE ONLY

QUESTION ONE

Refer to Report 1 in the resource booklet to answer the following questions.

(a) Identify and describe the explanatory and response variables for the study.

The explanatory variable is the colour of the car.

The response variable is the car that has been involved in a crash resulting in serious injury, and the proportion of cars that crashed in relation to the entire sample. //

(b) (i) Explain whether this study is an observational study or an experiment.

This is an observational study because the response variable of a car that has been involved in a crash resulting in serious injury would not be legally or ethically appropriate to base an experiment on, so researchers can only observe whether there is a relationship, but not run an experiment to find this out. The researchers cannot facilitate a car crash and resulting serious injury. //

(ii) Give an implication of using the type of study identified in part (i) for the specific relationship investigated.

Observational studies cannot result in causal claims to be made, as that would not be appropriate, since a number of confounding variables may have been acting to affect the results. In this specific relationship confounding variables such as whether the drivers were under the influence of drugs or alcohol would significantly impact on their driving ability negatively and perhaps be more likely to result in a crash because of this reason, not necessarily related to the colour of the car. //

(c) For this study, the researchers collected sam... all cars on Auckland roads.

No justify with example to illustrate link between the two data sets.

(i) Explain why the researchers compared ... n crashes to the percentage of colours of ...

It is important for the researches to investigate the percentage of colours of cars on Auckland roads to see which colours are most common, which may explain higher percentages of these coloured cars out of the total sample population have been involved in crashes, because there are more of these coloured cars distributed on Auckland roads, so both the proportion of each coloured car in the sample and the proportion of crashed cars need to be taken into consideration.

C

(ii) The researchers used cluster sampling to obtain their sample.

Discuss ONE example of how clusters may have been determined to ensure a representative sample.

Cluster sampling may have meant the researchers took into account groups of cars from central city areas, as well as groups of cars from more urban, suburb areas so that the sample was representative of all cars that travel on Auckland roads. It is vital for samples to be as representative of the population as possible so the demographics must be considered when it comes to the distribution of cars.

C

Cluster sample identified linking to population for representativeness

ASSESSOR'S USE ONLY

(d) The report states that researchers found "a significant reduction in the risk of serious injury in silver cars compared with white cars" and that "factors that could affect the results were taken into account in the analysis."

A potential issue with a statistical study is extending the results inappropriately.

Discuss ONE potential issue with extending the results of this study to all cars on New Zealand roads in 1998 – 1999.

A potential issue with extending the results of this study inappropriately to all cars on New Zealand roads in 1998-1999, as opposed to the actual sample of cars only on Auckland roads is that Auckland is the most populated city in the country but the colour of cars in this city will be independent of colours of cars around the country and in more rural areas of New Zealand, there will be different proportions of colours of cars due to farming etc. Car crashes occur independently and randomly throughout the country so inferences cannot be made appropriately from this study to apply to all of New Zealand. Although researchers said 'factors that could affect results were taken into account', these factors are still out of their control and no causal claims can be made from this observational study such as 'a significant reduction in the risk of serious injury in silver cars compared with white cars'. The proportion of silver cars and white cars will differ across the country, and so will the percentage of these involved in crashes.

No link to an example and transferability to NZ of roads.

5xC
1xU

E7

QUESTION TWO

ASSESSOR'S USE ONLY

Refer to Report 2 in the resource booklet to answer the following questions.

For parts (a) – (d), assume that the sample obtained is representative of all New Zealand drivers.

(a) The report states that "59% of the survey respondents rated changing the radio / iPod / MP3 player while driving as distracting."

Construct a confidence interval using this survey percentage and interpret this confidence interval.

Margin of Error = $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{1000}} \Rightarrow \pm 3.16\%$

95% Confidence Interval = $59\% \pm 3.16\% = [55.84\%, 62.16\%]$

We can infer with 95% confidence that the percentage of respondents who rated changing the radio/iPod/MP3 player while driving as distracting in this survey lies somewhere between the interval of 55.84% to 62.16%.

C

Not linked back to an inference on the population.

(b) The report states that 20% of survey respondents had sent texts while driving.

Discuss ONE potential issue with a survey question that asks respondents for this particular survey if they have sent texts while driving.

Many survey respondents may not wish to disclose this information because it is illegal (texting while driving ↑ an action), so they may provide false information in the survey which will skew the results and therefore the analysis will be less accurate.

This survey question effect must be approached carefully by researches as they require honest answers for validity of the survey but respondents may not be willing to admit they have illegally texted while driving.

C

No link made to underrepresentation of survey 20% result.

(c) Suppose 234 respondents were aged 18–24, and 288 respondents were aged 25–34.

Can a claim be made that a higher percentage of New Zealand drivers aged 18–24 will admit to talking on a mobile phone while driving (without a hands-free kit) than New Zealand drivers aged 25–34, based on the survey percentages presented in the report?

Construct a confidence interval and interpret this interval as part of your answer.

Aged 18-24: MoE = $\frac{1}{\sqrt{234}}$ = ± 6.54%

Aged 25-34: MoE = $\frac{1}{\sqrt{288}}$ = ± 5.89%

Average MoE = $\frac{6.54 + 5.89}{2}$ = ± 6.22%

Difference in poll % = 40% - 33% = 7%

1.5 x Average MoE = ± 9.3%

95% CI difference = 7% ± 9.3% = [-2.3%, 16.3%]

We can infer that the percentage of New Zealand drivers aged 18-24 who admit to talking on a mobile phone while driving (without a hands-free kit) will be somewhere between 2.3% lower and 16.3% higher than the percentage of New Zealand drivers aged 25-34 so the claim can NOT be supported because the lower end of the CI is below 0 (-2.3%).

(d) The headline for this report is "Txting a top distraction for young drivers".

Evaluate what statistical evidence, if any, has been presented in the report to support this claim.

Statistical evidence has only been provided for New Zealand drivers aged 18 and over. 'Young drivers' may only mean New Zealand drivers under 18, ~~but of~~ in which case, no statistical evidence has been presented otherwise 'Young drivers' may mean the 18-24 age group in which the study reports that 'nearly 50% sent texts while driving'. The headline is misleading as it only mentions texting - and not other factors discussed such as mobile phone use - and uses the words 'young drivers' when the overall report looks at drivers aged 18 and over.

No evidence clearly linked to not supporting the claim given.

(e) It is not clear in the article whether or not the survey of over 1000 New Zealand drivers involved only AA Insurance customers.

Describe ONE potential issue with using a random sample of AA Insurance customers to make generalisations about all New Zealand drivers.

This random sample of AA Insurance customers cannot be used to make generalisations about all New Zealand drivers appropriately because this group does not represent all New Zealand drivers.

Some drivers may not have insurance in the first place, and miss out on submitting information towards the survey for that reason.

This potential issue could be seen as selection bias of the sample because people, only AA Insurance customers, have been selected for the survey, when making inferences about the entire driving population of New Zealand //

ASSESSOR USE ONLY

j

E7

QUESTION THREE

Refer to Report 3 in the resource booklet to answ

Clearly links representative with sample matching census proportions for age.

(a) The report states that "The data have been ... data."

Explain why this was done.

Weighting samples ensures that groups of a population are not over or under-represented in the sample so that it can be representative of the population as a whole. This can include the demographics proportions within a population, such as age, gender and ethnicity. Since this survey is meant to represent New Zealanders aged 15 years and over, this large population will have specific demographics that weighting according to the 2013 Census will help to account for.

(b) Figure 1 uses vertical lines (error bars) to represent 95% confidence intervals.

(i) Give ONE reason why the vertical lines for the "never smokers" and "ex-smokers" are shorter than those for "current smokers".

The population of 'current smokes' is larger than the populations of 'never smokers' and 'ex-smokers' so there is more room for error with a large margin of error for larger populations like the 'current smokers', so the vertical lines are longer for this group and shorter for 'never smokers' and 'ex smokers'.

Incorrect statement: Increase n implies larger error bars.

(ii) The 95% confidence inter... New Zealand that engaged in risky alcol... roximately (44%, 62%).

Use the margin of error associated with this confidence interval to estimate the number of people in the sample who were current smokers.

95% CI = [44%, 62%]

Percentage = $\frac{44+62}{2}$ = 53%

Margin of Error = ± 9% = 0.09

$\frac{i}{\sqrt{n}} = 0.09$

$n = 123.46 \Rightarrow$ 123 people in the sample who were current smokers.

(c) The report states that "After adjusting for confounding variables, current smokers and ex-smokers were more likely than never smokers to report engaging in risky alcohol consumption in the last four weeks."

Identify ONE potential confounding variable that may have needed to be taken into account, and discuss how this variable may have been confounding.

A potential confounding variable that may have been needed to be taken into account is whether the respondents drink alcohol at all because this would significantly impact results, especially for 'never smokers' to report engaging in risky alcohol consumption if they never drink in the first place.

This is confounding in the sense that it can impact on all respondents in the survey, especially since it questions people as young as 15, for whom it is not legal to smoke or to drink (without parent's permission).

No clear confounding variable identified.

ns -

(d) A potential non-sampling error for surveys is to consider how people behave when surveyed.

Fully describe how the behaviour of people when surveyed could be a potential non-sampling error for this survey, and discuss how it could cause bias.

People in this survey had to remember events of their alcohol consumption for the past four weeks which is quite a long time to recall this information from ~~so~~ their memory may be a potential non-sampling error that causes bias because they may then choose to answer to the survey a response that they think the interviewer is wanting, especially in a face-to-face survey, so the interviewer effect can cause concern. The interviewer may also affect the respondent's answer by the way they behave or ask the question in a different way for different respondents so the results can provide bias and be less accurate.

Issue of Interviewer bias
F2F identified
No link back to population made.

M6

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

QUESTION NUMBER

ASSESSOR'S USE ONLY

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

QUESTION NUMBER

ASSESSOR'S USE ONLY