

No part of the candidate evidence in this exemplar material may be presented in an external assessment for the purpose of gaining credits towards an NCEA qualification.

3

91584



NEW ZEALAND QUALIFICATIONS AUTHORITY  
MANA TOHU MĀTAURANGA O AOTEAROA

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

SUPERVISOR'S USE ONLY

## 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.**

**Merit**

**TOTAL**

**15**

ASSESSOR'S USE ONLY

## QUESTION ONE

ASSESSOR'S  
USE ONLY

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.

An explanatory variable is one where the parameter is changed in this report the ~~parameter~~ explanatory variable is colour of the car involved in car crash resulting in serious injury. A response variable is the outcome that comes from changing the parameter in this case the response variable is risk of serious injury.

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

This ~~is~~ is an observational study because the researchers aren't intervening to change the parameter of car colour. They were just examining what colour cars resulting in serious injuries. They didn't have a control and treatment group.

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

An implication of using this study is they can't ~~compare~~ control confounding variables that may change the results of the study, because it is observational they only examine results they don't do anything to limit the confounding variables.

No clear reference to causal claims being made/not made.

No link do comparison  
of table data or do  
Auckland as the population

- (c) For this study, the researchers collected sample cars on Auckland roads.

- (i) Explain why the researchers compared the percentage of colours for cars involved in crashes to the percentage of colours of cars on Auckland roads.

The researchers would've compared percentages to see if their results from the Auckland study were representative of the whole New Zealand car crash population. By doing this they can then say it would be an effective way to reduce crashes by reducing silver cars as it is similar all over NZ not just in Auckland. //

- (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 can ensure a representative sample as it is obtaining data from all parts of the population. For example they were getting data on car crashes from places all over NZ instead of limiting it to just the North Island. //



- (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.

One potential issue is that in general there maybe more silver cars on the road than any other colour so they have an increased risk of crashing due to there being more of them.

By extending the results to the whole of New Zealand it means they are saying silver cars generally crash more often in all of NZ when the reason ~~silver~~ cars crash so much in Auckland is due to things such as traffic and busy roads. //

Link made to Auckland roads being unrepresentative of NZ roads

3x c

A4

## 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.

$$M.O.E = \frac{1}{\sqrt{1000}} = 0.0316$$

$$= 3.2\%$$

$$C.I = 59\% \pm 3.2\%$$

$$= 55.8\% - 62.2\%$$

I estimate  
With 95% confidence, the true value of the population parameter for <sup>people who</sup> changing the radio/iPod/MP3 whilst driving see it as distracting lies <sup>somewhere</sup> between 55.8% - 62.2% //

- (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.

One potential issue is a non-sampling error named behavioural considerations. This is where people don't answer truthfully because they know their answer is ~~too~~ socially unacceptable. In this case people may have chosen to lie and say they hadn't sent texts whilst driving because sending texts when driving is illegal. They may not want to get caught or may not want people knowing they do something illegal. //

Link made to being illegal and respondents lying only.



~~Suppose 234 respondents were aged 18-24, and 288 respondents were aged 25-34.~~

6

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

40%

33%

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.

$$M.O.E (18-24) = \frac{1}{\sqrt{234}} = 0.065 = 6.5\%$$

$$M.O.E (25-34) = \frac{1}{\sqrt{288}} = 0.059 = 5.9\%$$

$$1.5 \times \frac{6.5 + 5.9}{2} = 9.3\%$$

$$95\% \text{ C.I.} = (40 - 33) \pm 9.3\% \\ = -2.3\% - 16.3\%$$

Claim cannot be made to say higher percentage of NZ drivers aged 18-24 will admit to talking on mobile compared to NZ drivers aged 25-34 because 0% is included in the 95% confidence interval //

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

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

There is evidence in the report to conclude that New Zealanders drivers are texting whilst driving with 50% <sup>of people</sup> aged 18-24 admitting and ~~2~~ admitting to texting when driving. Although they have no supporting evidence to claim that this was a distraction for the drivers, it is pretty easy to gather that texting will be a distraction as you are watching your mobile phone screen instead of concentrating on the road. //

First statement is backed up with hearsay and not using the report 50% of the 20% ...

- (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.

a potential issue with using a random sample of AA insurance customers to make generalisations about all New Zealanders is that it could be unrepresentative of the whole New Zealand population. //

Issue identified of subgroup being used as representative only.

ASSESSOR'S  
USE ONLY

C

2x c  
1x j.  
1x u

E7



### QUESTION THREE

ASSESSOR  
USE ONLY

Refer to Report 3 in the resource booklet to answer the following questions.

- (a) The report states that "The data have been adjusted (weighted) according to 2013 Census data."

Explain why this was done.

They may have weighted the data according to 2013 Census data to see if the respondents were telling the truthfully. Or too see if their results were similar to the census.

No link to demographic specific to under/over representations

- (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 confidence intervals for never and ex smokers are shorter than those for current smokers because the sample size for never and ex smokers must have been greater than the sample size for the current smokers. A larger sample means a smaller margin of error and therefore smaller confidence interval.

- (ii) The 95% confidence interval for the proportion of people in New Zealand that engaged in risky behaviour is approximately 44% (95% CI: 44%, 62%).

Larger sample sizes identified for 'never' and 'ex-' groups only

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

$$62 - 53 = 9 \quad \therefore \text{M.o.E} = 9\% = 0.09$$

$$\text{M.o.E} = \frac{1}{\sqrt{n}}$$

$$0.09 = \frac{1}{\sqrt{n}}$$

$$n = \frac{1}{0.09^2}$$

MoE calculation only.



- (c) The report states that "After adjusting for confounding, ex-smokers were more likely than never smokers to have tobacco consumption in the last four weeks."

Identify ONE potential confounding variable that and discuss how this variable may have been confounded.

No identified confounding variable appropriate to the report.

one potential confounding variable may have been that people who said they were ex-smokers may have only just quit. so this may have been a confounding variable because

one potential confounding variable is that if somebody had only smoked once before and that's it they are classed as an 'ex-smoker' //

- (d) A potential non-sampling error for surveys is to call people who are not at home.

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

Only identifies behaviour of trying to interview.

when people are being surveyed they <sup>generally</sup> like to say things they think will be socially acceptable to the person they are talking to. This causes people to potentially change their answer to something they think the interviewer would like to hear so <sup>respondent</sup> ~~they don't~~ doesn't feel judged, this could be a non-sampling error for this survey because it <sup>means</sup> people <sup>may</sup> ~~are~~ changing their answers to 'never smoking' and 'no drinking' to feel more accepted as both are illegal under 18 (survey from 15 yrs). This will then create a <sup>larger</sup> difference between the estimates of the population parameter and the true value for the population parameter. It creates bias as people are more inclined to say they don't smoke and drink when they actually do. //

15

C

3 x C

A4

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

QUESTION  
NUMBER

ASSESSOR'S  
USE ONLY

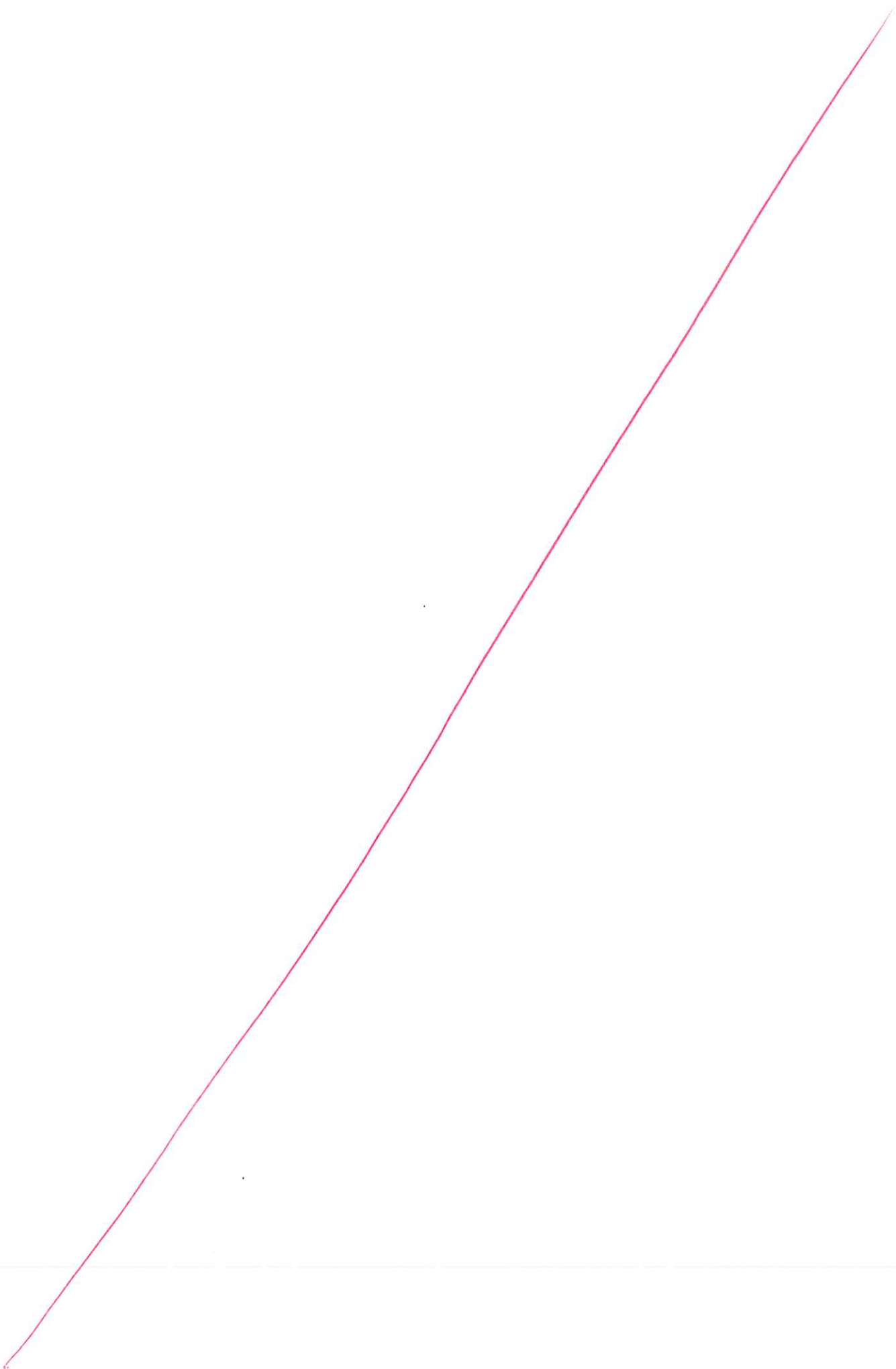
Extra paper if required.  
Write the question number(s) if applicable.

ASSESSOR'S  
USE ONLY

QUESTION  
NUMBER



91584



3

91584



915840



NEW ZEALAND QUALIFICATIONS AUTHORITY  
MANA TOHU MĀTAURANGA O AOTEAROA

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

SUPERVISOR'S USE ONLY

# 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.**

Merit

TOTAL

16

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 for this report is the colour of car that the participant drives, and the response variable is the number of crashes resulting in serious injury.

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

This study is an observational study because no variables were actively changed, therefore the participants didn't have to interrupt their daily lives in order to partake in the study (researchers only watched to collate their data).

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

The implication for using cluster sampling is that because researchers are only using small clusters of people, this may not be entirely representative of the entire population. The reason being because people were only chosen from one area (region) meaning this study and claims drawn may not accurately represent the whole population in New Zealand.

No statement linking an observational study and no causal claim can be made.



- (c) For this study, the researchers collected sample data on cars on Auckland roads.

- (i) Explain why the researchers compared the percentage of crashes to the percentage of colours of cars

No example given to illustrate usefulness of the comparison to justify.

ASSESSOR'S  
USE ONLY

Because the researchers wanted to determine whether or not the colour of a participant's car had any correlation between the colour and the number of cars involved in crashes, eg to see if a certain percentage of cars (that were a particular colour) were more likely to crash. 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.

An example of how clusters may have been determined to ensure a representative sample could be taking measurements on multiple days (in different conditions eg rainy, sunny etc) and also on multiple roads (some being tight roads, motorways, large intersections etc) in order to test to see whether or not the sample is in fact representative of the entire population. i

Cluster — multiple roads  
Representativeness

- (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 of extending the results to all cars on New Zealand roads in 1988-1999 is that the roads in different areas and regions of New Zealand are very different. In the Auckland region there are not all of the possible roading layouts eg motorway, gravel roads in more rural areas, inner city etc. This means that there is a huge issue surrounding making this claim about all cars on New Zealand roads because Auckland only has a small percentage of roading variations / conditions meaning that the results of this study are not going to be representative of the entire population (as not all cars in New Zealand were included in the study).

Issue identified as 'roads', with example to illustrate Auckland as being unrepresentative of all NZ roads. No transferability

3xc  
1xj  
1xi

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.

$$\begin{aligned}\text{margin of error} &= \frac{1}{\sqrt{n}} \\ &= \frac{1}{\sqrt{1000}} \\ &= 3.16\%\end{aligned}$$

$$\begin{aligned}\text{confidence interval} &= 59 \pm 3.16 \\ &= [55.8\% - 62.2\%]\end{aligned}$$

Because the lowest value on the confidence interval (55.8%) is above 50%, it can be claimed that changing the radio/iPod/MP3 whilst driving is a distraction. //

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

Discuss ONE potential issue with a survey question if they have sent texts while driving.

No issue identified that links to lying.

One potential issue with this type of survey question is that the question is not very clear. The answer to the question is really open to individual interpretation eg someone may think that because they sent one text whilst driving a week ago that meant they didn't send texts whilst driving (so answered 'no' on the survey). A possible improvement to the survey would be to accurately pin point a number of texts that an individual would have to send before they put 'no' or 'yes' on the survey for this question. //



23.4

6

28.8

- (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 hands free) than drivers aged 25–34, based on the survey percentages?

Construct a confidence interval and interpret this.

$$\text{Difference} = 28.8 - 23.4$$

$$= 5.4$$

Incorrect Confidence Interval calculated. Only moe is correct 9.3%.

~~margin of error 12.3%~~

~~margin of error 7.7%~~

~~moe =  $\frac{1}{\sqrt{1000}}$~~

~~margin of error~~

margin of error 18–24

moe 25–34

$$\text{moe} = \frac{1}{\sqrt{234}}$$

$$\text{moe} = \frac{1}{\sqrt{288}}$$

$$= 6.54\%$$

$$= 5.89\%$$

$$\text{moe for difference} = 1.5 \times \frac{\text{moe}_1 + \text{moe}_2}{2}$$

$$= 1.5 \times \frac{6.54 + 5.89}{2}$$

$$= 9.3\%$$

confidence interval =  $5.4 \pm 9.3$

interval =  $[-3.9\% - 14.7\%]$

Because the lowest value of the confidence interval is below 0 the claim that more drivers 18–24 talk on their mobile phones whilst driving compared to drivers aged 25–34 is not valid, the claim cannot be made.

- (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.

$$\text{margin of error} = \frac{1}{\sqrt{1000}} = 3.16\%$$

$$\text{confidence interval} = 50 \pm 3.16$$

$$= [46.8 - 53.2]$$

Because the lowest value of the confidence interval is below 50% (46.8%) the claim made in the title is not appropriate "txting a top distraction for young drivers" and the claim can't even be made.

C I calculated that is not appropriate to statistics 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.

One potential issue with using a sample of AA Insurance customers is that people may actually be more likely to be careless eg texting whilst driving because they know that they're insured and therefore are prepared to take more risks, because if they crash or cause damage then they can just claim it on their insurance. Other New Zealand drivers may be less careless (possibly because they're not insured) and won't text and drive because they can't afford the damage costs if they do crash. This means that due to this reason the sample collected may not be entirely representative of the population. //

ASSESSOR'S  
USE ONLY

j

1x c  
2x j

M5



## QUESTION THREE

ASSESSOR'S  
USE ONLY

Refer to Report 3 in the resource booklet to answer the following questions.

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

Link made to lack of responses from particular groups. No example given for under/over representation.

Explain why this was done.

The reason that the data has been weighted is so that the results more clearly show the <sup>true</sup> population's results. Weighting takes into consideration the lack of responses from a particular group or region so that the overall results are more representative of the true population eg multiplying responses/percentages by that of the true population to get more accurate results eg how many times the

- (b) Figure 1 uses vertical lines (error bars) to represent 95% confidence intervals. result eg how many times the individual results need to be multiplied
- (i) Give ONE reason why the vertical lines for the "never smokers" and "ex-smokers" are shorter than those for "current smokers".

A potential reason why the vertical lines for the "never smokers" and "ex-smokers" are shorter than "current smokers" is because the sample size for never smokers and ex smokers are larger than current smokers, meaning that the confidence interval is smaller

- (ii) The 95% confidence interval for the percentage of current smokers in New Zealand because that engaged in risky alcohol consumption in the last four weeks is approximately the sample sizes are larger. (44%, 62%).

Use the margin of error of people in the sample who were current smokers.

MoE is incorrect.

estimate the number

53% = 1359 (4sf) people

margin of error =  $\frac{1}{\sqrt{n}} \times 100 = \frac{1}{\sqrt{1359}} \times 100 = 2.7\%$

number of people

who are current

smokers = 1372 (4sf)

$$\frac{1}{\sqrt{n}} = 1.96$$

$$2.7 = \frac{1}{\sqrt{n}}$$

$$\frac{1}{\sqrt{n}} \times 100 = 2.7$$

$$\frac{1}{\sqrt{n}} \times 100 = 1.96$$

$$\frac{1}{\sqrt{n}} = \frac{2.7}{100}$$

$$\sqrt{n} = \frac{100}{2.7}$$

$$n = \left(\frac{100}{2.7}\right)^2 = n = 1371.7$$



- (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 and discuss how this variable may have been con

No clear confounding variable identified.

One potential ~~confounding variable~~ (variable that was not taken into consideration when the data was being collected) was looking at medical history of the people who are non-smokers. Potentially those non-smokers may actually have medical reasons why they cannot smoke or engage in risky alcohol consumption in the last 4 weeks. This will effect the results of the survey and therefore the results may not be representative of the entire population (as maybe not this many people

- (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.

Due to the survey results being gathered in a face to face in-home survey and only 73.2% of people responded this means there is quite a lot of potential bias. An example of this could be the fact that people may be embarrassed about answering questions face-to-face, and therefore are more inclined to say one answer over another. Doing an at home survey may also mean that the person reading the questions to the participant may accidentally record data incorrectly, or read out a question, giving emphasis to

ASSESSOR'S  
USE ONLY

ns

have medical illnesses prevent them from engaging in the sort of behaviour in the true population

duppy 1e  
x → c  
ns

3xc

A4

Extra paper if required.

Write the question number(s) if applicable.

ASSESSOR'S  
USE ONLY

QUESTION  
NUMBER

a particular point, therefore swaying the participants answer and ultimately causing bias. //

Non sampling error identified  
linking to face to face  
interviews.

No link to population  
estimate made.

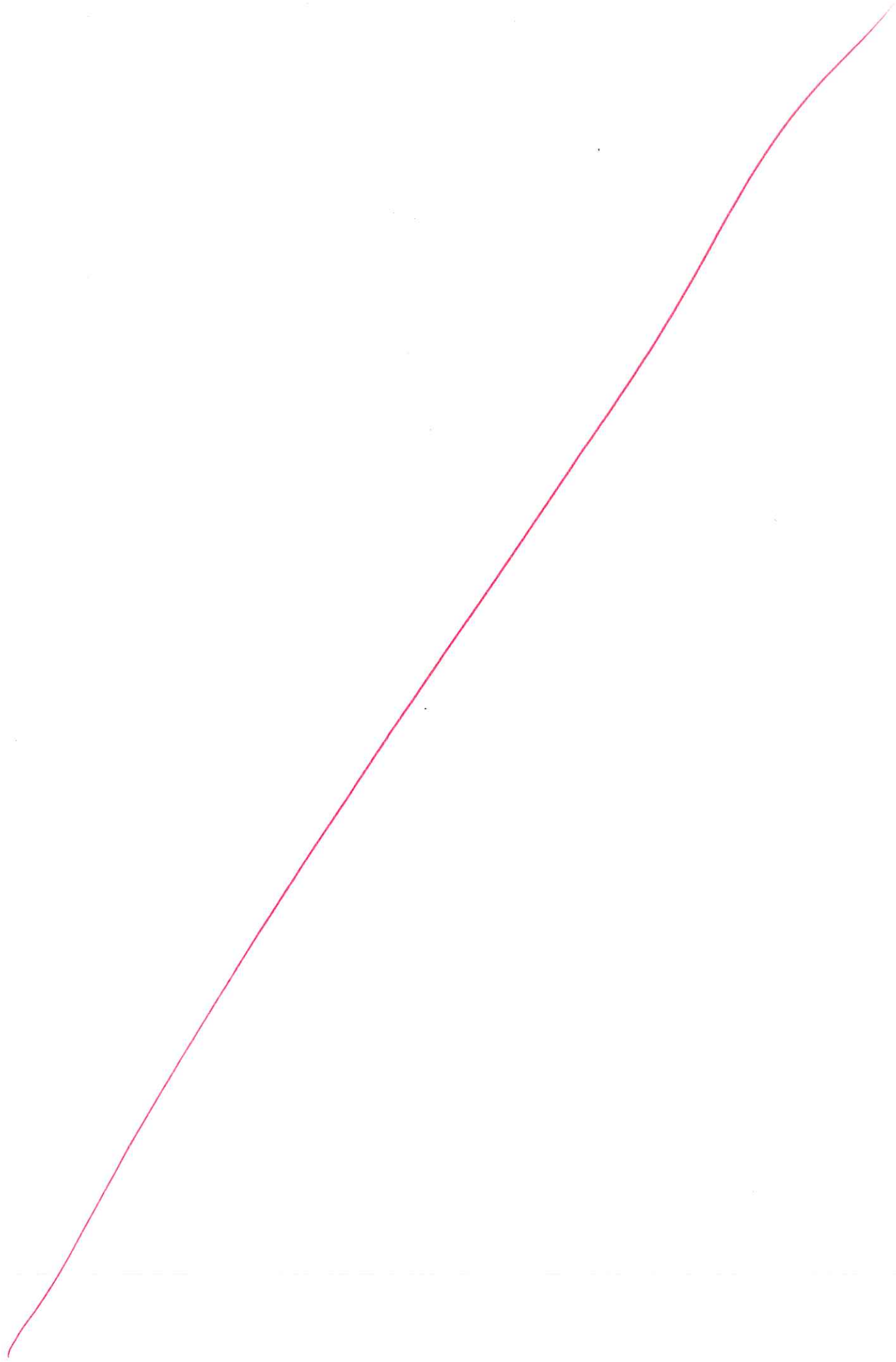


Extra paper if required.  
Write the question number(s) if applicable.

ASSESSOR'S  
USE ONLY

QUESTION  
NUMBER





91584