National Certificate of Educational Achievement
TAUMATA MĀTAURANGA Ā-MOTU KUA TAEA

Exemplar for Internal Assessment Resource

Mathematics Level 3

Resource title: New Zealand crash statistics

This exemplar supports assessment against:

Achievement Standard 91582

Use statistical methods to make a formal inference

Date version published by   December 2012
Ministry of Education   To support internal assessment from 2013

	Grade Boundary: Low Excellence
1.	For Excellence the student is required to use statistical methods to make a formal inference, with statistical insight. This involves integrating statistical and contextual knowledge throughout the statistical enquiry cycle and reflecting on the process. There is evidence that statistical and contextual knowledge has been integrated. The student has used their research to develop a purpose and to pose an appropriate question (1). The Los Angeles research that is referenced is not directly related to drivers. Analysis comments are clearly linked to the context (highlighted) and there is some consideration of other explanations. For a more secure Excellence the student would need to investigate the data more fully. For example, a consideration of other variables and research findings needs to be more directly related to the question.

New Zealand Crash Statistics

Introduction

I wonder what the difference in median Blood Alcohol levels is between minor and severe crashes involving alcohol in New Zealand for the most (15-24 years old) and least (50-59 years old) likely people to crash groups. My analysis relates to recorded accidents in 2011.

I am doing this investigation as I have, as a teenager, recently started driving. By the end of this year, I will be old enough to drink. It is known that these two areas do not tend to mix very well, with it generally believed that alcohol is causing the crash, but to what extent does the amount of blood alcohol impact on the *severity* of the crash you have.

Contrary to what one would expect, a study from Cedars-Sinai Medical Center, in Los Angeles, suggests that if someone suffered a brain injury in a motor vehicle crash (MVC), they had a better chance of surviving if they had been drinking before the MVC. [1]

This is refuted by others, though, such as the Injury Research Center at the University of Michigan, who have found that alcohol can exacerbate the effects of an MVC, making things worse. [2]

While my analysis will not delve into such deep issues and as detailed conclusions, I do expect to find a higher blood alcohol level in MVCs that are more severe, and a lower blood alcohol level in MVCs that are less severe.

Analysis

	Min	Lower Q	Median	Mean	Upper Q	Max	Std Dev	Sample Size
Minor	32	111.0	135	133.7	158	214	36.142	189
Serious	47	119.2	160	154.5	185	336	50.703	128

Overall Visual Comparison
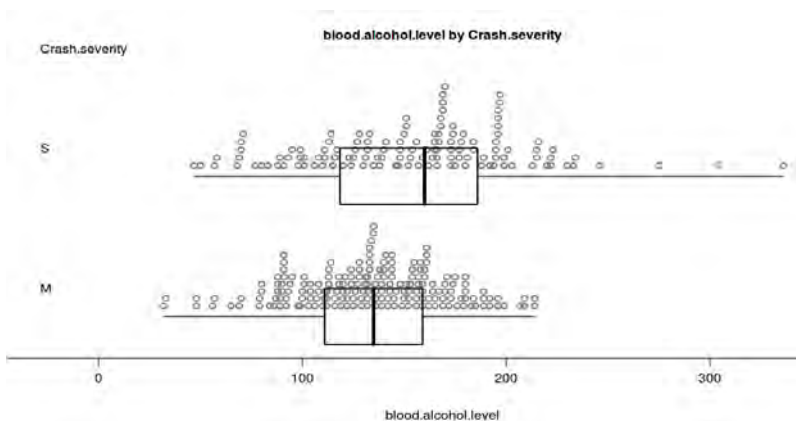
My first observation is the range of the data: the amount of alcohol present in the blood is much more varied in the Serious Crash data, than that in the Minor Crash data. The minor crash data is fairly symmetrical with the slightest hint of a skew to the left, but the severe crash data has a slightly clearer skew to the right. The severe crash data is shifted further up the graph, to a higher blood alcohol level, while the minor crashes sit lower on the graph, meaning there is a lower blood alcohol level in these crashes. The severe crash data has two modes, but this is most likely a peculiarity in the sample, rather than a feature we would see in the

Centres

The median value for blood alcohol level for these drivers is 25mg higher in serious crashes (160mg/100mL) than in minor crashes (135mg/100mL). This is a significant difference, as, for under 20 drivers, the legal limit for driving is 30mg/100mL (despite the official 'zero tolerance' stance — this allows for any medication influences, etc.). [3]

Shift, Overlap & Spread

The serious crashes for this group of drivers have a higher blood alcohol level, with a median of 160mg/100mL, so have shifted higher up the graph than the minor crashes, with a median of only 135mg/100mL. The median with this sample of severe crashes is higher than the upper quartile (i.e. 75% of the data) for the minor crash blood alcohol levels. Having said that, there is still quite a bit of overlap in the interquartile ranges, with almost 50% of the minor data sitting in the lower-middle 25% of the serious crash interquartile range. The serious crash blood alcohol levels interquartile range is 65.8mg, and the interquartile range for the minor crash blood alcohol levels is 47, so it is 18.8mg less spread out for the middle 50% of data.
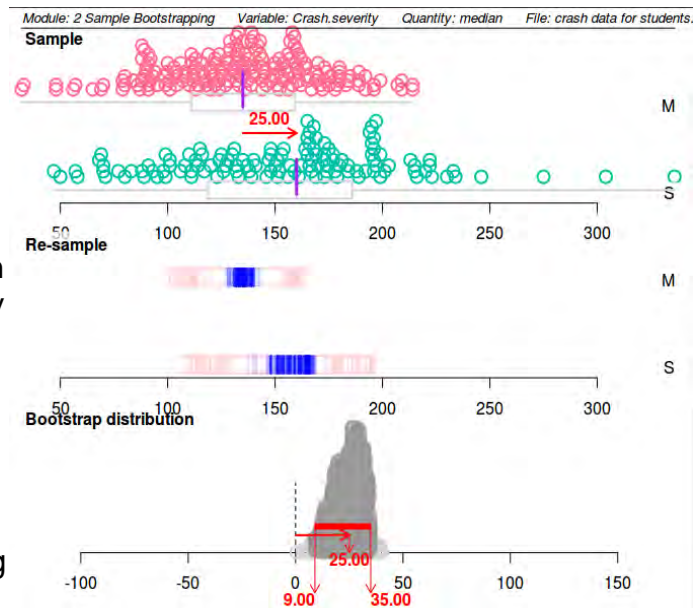
Unusual

The severe crashes have some very high data points for blood alcohol level. The highest point, 336 mg/100mL, is very high, and while not impossible, it is unlikely. At this point, death is a very real possibility, with passing out likely. To be able to attempt to drive at this level of inebriation is very unlikely. [4]Having said that, some heavy drinkers can, over time, develop a tolerance, so may be able to deal with that kind of drinking.[5]

Statistical Inference

It is a fairly safe bet that the median blood alcohol level in severe crashes for the most and least likely people to crashing New Zealand is between 9mg/100mL and 35mg/100mL higher than those the median blood alcohol level in minor crashes for the most and least likely people to crash in New Zealand. I would therefore be happy to make the call that there is a higher level of blood alcohol in serious crashes.

Conclusion

I can make the call, with the sample that I am given, that the blood alcohol level is higher in serious crashes than minor crashes, for the most and least at-risk of crashing groups in New Zealand, as there is a significant gap in the median values of 135 and 160mg/100mL.

This is backed up by the bootstrapped confidence interval, as a difference of 0 is well

outside the confidence interval of 9mg/100ml to 35mg/100mL. The bootstrapped confidence interval was achieved by resampling the data set again, with replacement, for the same number of samples. This gives us many different samples that mimic the original population surprisingly well. This method will give us a range of blood alcohol levels that will cover the median of the population we sampled most of the time.

However, there is a causality issue to deal with. We cannot know in these crashes whether the driver who was drunk was indeed the cause of the accident, nor how the level of blood alcohol affected their injuries. As such, we cannot completely disprove the research saying that alcohol protects you in a crash: an identical crash with a person who was less intoxicated may have resulted in more injuries. This would require much more data, and a significantly more detailed analysis.

All of this is based on the assumption that my original sample was correct, and I have no reason to suspect it was not. The sample was of a fairly large number of accidents, and had a good spread of data from various crashes, so it should be a good representation of the population of crashes involving alcohol in New Zealand within the most- and least-likely to crash groups.

In this sample, the means are only slightly affected by outliers: so slightly, I would not immediately hesitate in using them as a second source of evidence. This may only apply to this sample of crashes, though, so median would still be a better average.

Bibliography

1. Alcohol Brain Injuries – http://articles.cnn.com/2009-09-25/health/alcohol.brain.injury_1_alcohol-brain-injuries-david-hovda?_s=PM:HEALTH, retrieved 23/6/2012

2. Alcohol's Double Threat – http://sci.rutgers.edu/forum/showthread.php?t=22530, retrieved 23/6/2012

3. Alcohol and Drugs Limit – http://www.nzta.govt.nz/resources/roadcode/about-limits/alcohol-and-drugs-limits.html, retrieved 24/6/2012

4. Blood Alcohol Content – http://en.wikipedia.org/wiki/Blood_alcohol_content, retrieved 24/6/2012

5. Grams per Litre of Alcohol to BAC – http://forum.onlineconversion.com/showthread.php?t=589#2, retrieved 24/6/2012

|	Grade Boundary: High Merit
2.	For Merit the student is required to use statistical methods to make a formal inference, with justification. This involves linking components of the statistical enquiry cycle to the context and making supporting statements which refer to evidence. There is evidence of some research related to the purpose and a comparison question has been posed (1). While the population is not clear in the question, it is clear in the conclusion. The discussion has been linked to the context throughout (highlighted) and the conclusion has been clearly justified (2). There is some evidence of insight in the comment "*Something I found unusual",* but in order to reach Excellence there needs to be more evidence of statistical insight.

Problem

As Year 13 Students and as teenagers, many of us are well into our 'learning to drive phase,' and therefore we are considered extremely prone to crashing. New Zealand has the highest road death rate in the OECD for 15-17 year olds, and the fourth highest road death rate for 18-20 year-olds, according to New Zealand's Transport Agency (NZTA). As drivers of all ages, citizens need to be aware of the consequences of poor decisions associated with driving, such as what they consume during or before driving. This is especially important for New Zealand, as we are known worldwide for our dangerous driving, with the NZTA saying 421 fatal road crashes occurred in 2007. The media are highlighting such stereotypical views with young drivers with dangerous driving such as alcohol consumption reflected in the crash severity, and how to avoid these crashes altogether, represented in the successful 'Legend' campaign aimed at young Maori drivers. I believe this stereotype is indeed correct, and that the higher the blood alcohol level is, the higher the crash severity. Therefore, I wonder what the difference in medians is between the blood alcohol level found in a severe crash ⓵ and a minor crash in New Zealand? This is where alcohol was deemed the cause of the crash.
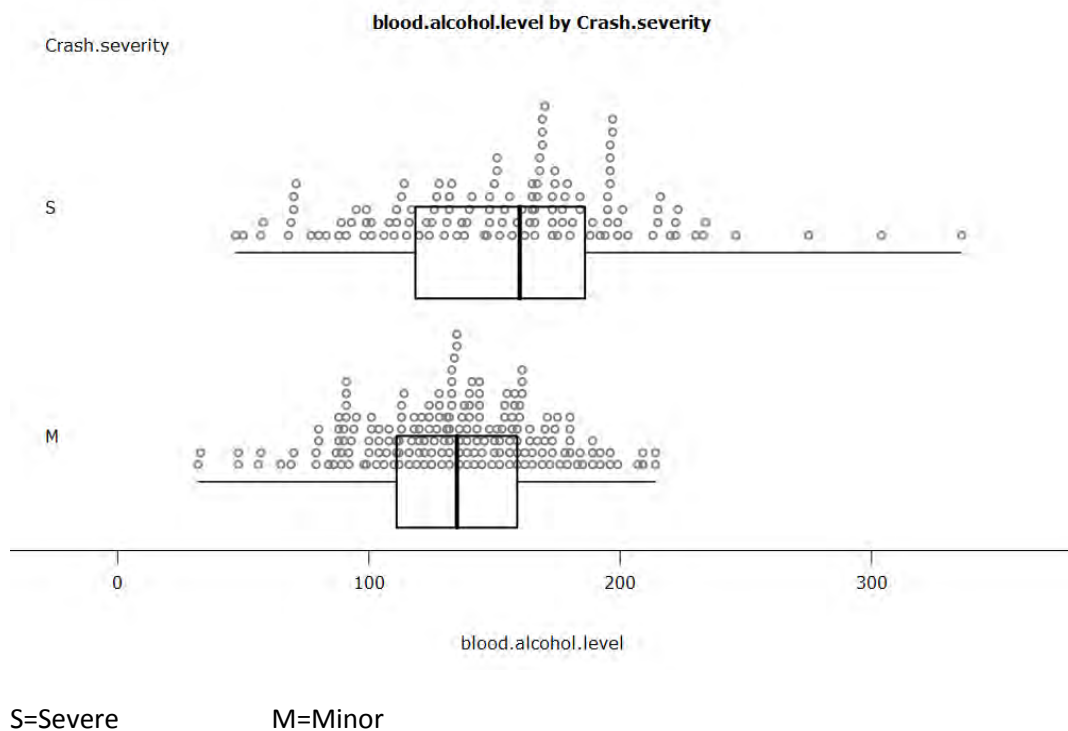
blood.alcohol.level by Crash.severity

S=Severe   M=Minor

```
Summary of blood.alcohol.level by Crash.severity
  Min. 1st Qu. Median  Mean 3rd Qu. Max. Std.dev Sample.Size
M  32   111.0    135  133.7   159  214  36.142     189
S  47   119.2    160  154.5   185  336  50.703     128
```

Bootstrap distribution

Mean: 24.51
SD: 6.65

25.00

12.00   35.50

-100   -50   0   50   100   150
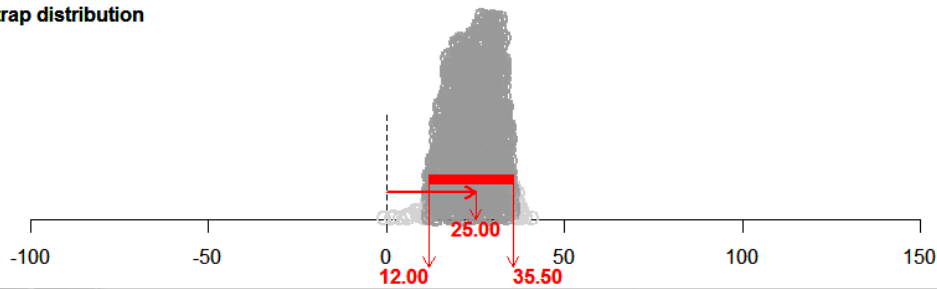
Analysis

My initial impressions of these samples of blood alcohol levels are that the severe crashes are more spread out than the minor crashes; while both severe and minor crashes have at least some portion of alcohol in the blood system during the crash, the bottom half of severe crashes is more spread than the bottom half of minor crashes. The severe crashes have many values that are higher than the minor crashes.

The median blood alcohol level for minor crashes is 25mg per 100ml of blood less than the median blood alcohol level, at 135mg, while the median blood alcohol level for severe crashes is 160mg. I was surprised at this as I thought the difference would be greater as the median for the serious crashes would be a lot higher.

The middle 50% of these blood alcohol levels of the severe crashes overlaps the middle 50% of the minor crashes by a reasonable amount. The median blood alcohol level for the severe crashes is 160mg/100mL, 1mg above the upper quartile of the minor crashes at 159mg/100mL. The lower quartile for serious crashes is 119.2 mg/100mL which is only 8.2 mg/100mL more than the lower quartile for the minor crashes. However the upper quartile for the serious crashes (185 mg/100mL) and minor crashes (159 mg/100mL) have a difference of 26 mg/100mL which results in the middle 50% of the serious crashes having a larger spread than the minor crashes. I was surprised at how closely they both overlapped but with the upper quartile for the serious crashes being a lot more than the minor crashes it fits with the view that the higher your blood alcohol level is the more serious the crash will be if it does take place.

Something I found unusual in this sample is that the highest blood alcohol level found in the severe crashes is 336mg/100mL, enough to kill that individual according to Wikipedia's blood alcohol content page that states 0.3-0.39 BAC is enough to cause an individual to become unconscious or kill them. The internet online conversion forum states that multiplying the BAC by 100 gives of alcohol for every 100ml of blood. I am surprised that he was only involved in a severe crash and not killed, either due to the alcohol itself or driving.

The distribution of severe crashes is right skewed, with most of the data at the lower end of the scale and heads towards the maximum blood alcohol level of 336/100mLmg, while the distribution of minor crashes is distributed evenly. I would have expected the shape for the serious crash group to be more left skewed with most of the data on the higher end of the scale for blood alcohol level.

Conclusion

From these samples, I can make the call that there is a difference in the blood alcohol level for "drunk" New Zealand drivers involved in a severe crash and New Zealand drivers involved in a minor crash. That is, I can make the call that the blood alcohol level for severe crashes is higher than the blood alcohol level for minor crashes. From these samples, I am pretty sure that the median blood alcohol level for severe crashes in New Zealand is somewhere between 12mg/100mL and 35.5mg/100mL more than the median blood alcohol level for minor crashes in New Zealand.

I can make the call from these samples that there is a difference in blood alcohol level for severe and minor crashes in New Zealand. This conclusion is also confirmed with my bootstrap confidence interval for the difference between the blood alcohol level for severe and minor crashes in New Zealand which does not contain zero, indicating that the severe crashes in New Zealand tend to have a higher blood alcohol level than minor crashes. Also the fact that the confidence interval is positive again proves the severe crashes in New Zealand tend to have a higher blood alcohol level than minor crashes.

I am basing this conclusion on the bootstrap confidence interval I calculated, which involves re-sampling from my original sample of 317 crash victims. The key assumption with this method is that my original sample was representative of the population of all crash victims in New Zealand where alcohol was deemed a factor of the crash. I believe that this sample is indeed representative of the population of New Zealand's "drunk" drivers. There are many other possible data categories that could be considered during this investigation, such as the age of the driver, the weight of the driver(especially when unusually high contents of blood alcohol levels are concerned) and their licence type, which could potentially provide a more accurate representation of the population as to the effects of blood alcohol level affecting crash severity. I also did not have many minor crash drivers that contained excessively high amounts of alcohol in their system. My sample did highlight that the top half of blood alcohol levels for severe crashes appears to be much higher than the top half of blood alcohol levels for minor crashes-this was the reason I chose to use the median as my summary statistic rather than mean. The mean would have been influenced by the large blood alcohol levels in severe crashes (sample mean for the severe crashes was 154.5mg/100mL whereas for the minor crashes it was only 133.7mg/100mL).

|	
	Grade Boundary: Low Merit
3.	For Merit the student is required to use statistical methods to make a formal inference, with justification. This involves linking components of the statistical enquiry cycle to the context and making supporting statements which refer to evidence. An explanation for the choice of variables has been given, and an appropriate question that identifies the population and the difference between the medians of the two groups has been posed (1). Some of the discussion on the distributions has been linked to the context or investigative question (highlighted). The comment on unusual features is unclear and the reference to the range is incorrect. The student has made a correct call ("*Based on my samples…*") that has been justified. For a more secure Merit the linking of the analysis discussion to the population needs to be developed, and more than a minimal statement on representativeness and sampling variability would be expected.

New Zealand Crash Statistics:

As year 13 students, many of us want to be more independent and plan to sit the first part of our driver's license as well as start our 'learning to drive campaign'. Despite being confident to begin driving amongst other vehicles, individuals between the ages of 15 – 24 are also known to often drink alcohol to excess. Unfortunately one of the side-effects of alcohol is an impairment of rational thought, which combined with a vehicle and / or peer pressure can have lethal consequences.

There are many studies on the effects of drinking and driving that conclude that "car crashes are the leading cause of death among people aged 15 to 20. " In New Zealand approximately "1,900 people under the age of 21 die every year from car crashes involving underage drinking".

Although a horrific statistic, my research has shown that as the age of the 'risk group' increases the severity of the crash decreases. This could be due to the amount of alcohol in blood levels, between the two age groups and is most likely a result of learned behaviour. i.e. the drinkers are much more experienced and have knowledge of the effects of alcohol on their bodies so drink relatively less than they may have done when in the higher risk age group. Alcohol is also a confidence booster and lowers peoples' perception of hazards and risk.

Therefore, I suspect that the level of blood alcohol has the greatest influence on crash severity. If the blood alcohol level is higher, then the severity of the crash is likely to be greater than if the blood alcohol level was lower.

I want to investigate if there is a difference in the medians for the crash severity under the influence of alcohol in blood level for New Zealand drivers.
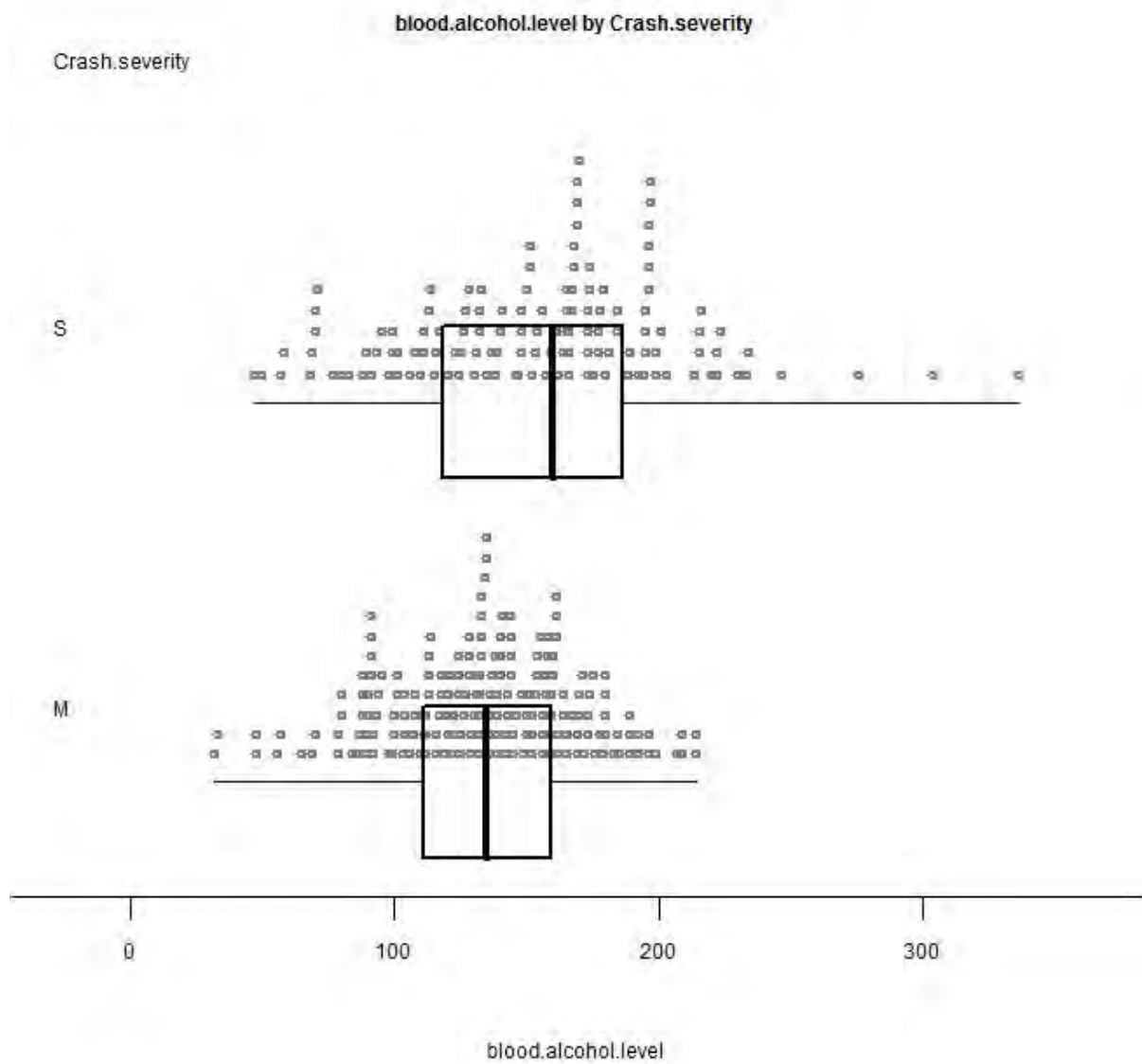
①

blood.alcohol.level by Crash.severity

blood.alcohol.level

Summary of blood alcohol level by Crash severity

	Min	1st Qu	Median	Mean	3rdQu	Max	Std dev	Sample.Size
M	32	111.0	135	133.7	159	214	36.142	189
S	47	119.2	160	154.5	185	336	50.703	128
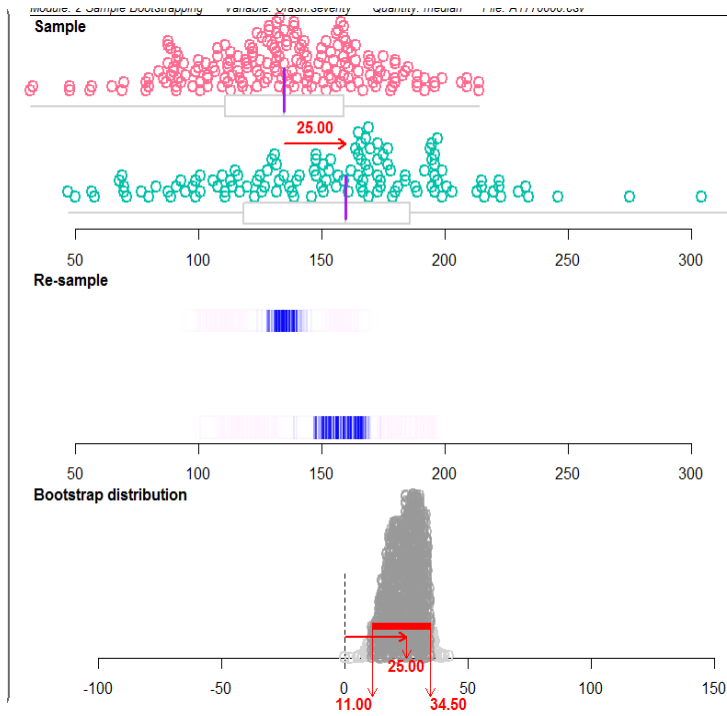
First overall impression:

My preliminary thought of these samples is that the serious crashes are more spread out than minor crashes. The serious crashes starts around the same mark as the minor crashes do for the minimum alcohol blood level, however it is reasonably higher for the maximum blood alcohol level, which confirms my suspicion above that an increase in blood alcohol level does have an effect on crash severity.

Shift and overlap:

From viewing the graph, we are able to see that the median of 160mg/100ml of blood alcohol for serious crashes is slightly higher than the upper quartile for the minor crashes which is around 155mg/100ml of blood alcohol. The upper quartile of the serious crashes is substantially greater than the middle 50% of the minor crashes, with the upper quartile at 190mg/100ml, 30 more than the middle 50% for the minor crashes.

The serious crash data is more spread out in comparison to the minor crashes with the IQR of the serious crashes at 80mg/100ml blood alcohol level, whilst the IQR for the minor crashes is around 55mls of blood alcohol level.

Centres:

The median for blood alcohol levels in serious crashes is 160mg/100ml, 25mg/100ml more than that of blood alcohol levels in minor crashes (135mls).

<u>Unusual:</u>

I believe that an unusual feature of the box plot graph is the fact that the alcohol blood level in the serious crashes is quite a lot smaller than many would expect it to be. The range for the serious crash blood alcohol level is quite narrow, only spread between 100mls and about 190mls. I expected the mean of the population sample to have a blood alcohol level of at least 250mls. While it is plausible to say that the more alcohol you have in your blood the slower your reaction time will be and therefore, the 'bigger the mess', of more significance is the fact that the two graphs are much closer than expected, demonstrating how people can easily become a higher risk by deciding to have "just one more drink".

<u>Shape:</u>

The distribution for the serious crash is right skewed; however the distribution for the minor crash is left skewed. The majority of this data from these data sets lies between the amounts of 100mg/100ml and 200mg/100ml and tails off towards the maximum amounts of blood alcohol of 336mg/100ml for the serious crashes and 214mg/100ml for minor crashes. I would have expected the distribution of the severe crashes to be more of a left skew, as the more alcohol that is consumed the slower the reaction time and consequently the likelihood of an accident increases.

<u>Make a formal statistical inference:</u>

From these samples, I suggest that there is a difference in crash severity (serious and minor crashes in New Zealand) when the blood alcohol level is either increased or decreased. Furthermore, I am pretty sure that the difference in median for blood alcohol is somewhere between 11.00mg/100mls and 34.50mg/100mls. Therefore, my investigation shows that serious crashes have relatively more alcohol per 100mls of blood, than minor crashes.

<u>Conclusion:</u>

From the sample data of crash severity I have concluded that there is a difference in the seriousness of the crash, which is dependent on the amount of alcohol absorbed into the blood stream (measured in milligrams per 100mls)

As the median of the serious crash data is greater than the upper quartile of the minor crash data, 160mg/100mls compared with 159mg/100mls of blood, this supports the argument that the seriousness of the crash is directly related to the amount of alcohol consumed. This conclusion is also confirmed with my bootstrap confidence interval for the difference between the population medians of serious and minor crashes in New Zealand, which does not contain zero, indicating that serious crashes have more alcohol per 100mls of blood than minor crashes.

I am basing this conclusion on the bootstrap confidence interval that I calculated. This involves re-sampling from my original sample of 317 drivers. The main assumption with this method is that my original sample was representative of the population of drivers between the ages of 15-24 and 50-59 in New Zealand. If I were to choose another identical age group sample, the results may have differed as that sample could have contained individuals that were relatively more responsible than individuals in this particular sample.

	Grade Boundary: High Achieved
4.	For Achieved the student is required to use statistical methods to make a formal inference, showing evidence of using each component of the statistical enquiry cycle. The student has provided evidence for each component of the statistical enquiry cycle to make a formal inference. While there is some linking of the analysis comments back to the context (highlighted), the evidence for this aspect of the cycle needs to be clearer in order to reach Merit.

Background

On October 23rd 2011 TVNZ first aired what is being hailed as the NZTA best drink driving add in decades, commonly known 'The Ghost Chips Add'. Although on a world scale NZ's road toll stats are pretty poor we are making a dramatic improvement, for example over the past 4 year our road toll has almost halved and a road toll of 0 over the 2012 Easter holidays. So it's fair to say that one way or another NZTA's messages of driving safely and making the right decisions is getting out there.

Never the less drink driving still remains a major problem in New Zealand. Between 2006 and 2008 26% of drivers killed in a fatal accident were under the influence of Alcohol and a further 2% under the influence of both alcohol and drugs. Although it's true that this means 72% of drivers killed in such accidents were sober, consider the fact that for every 100 people under the influence that died, 54 passengers in the cars of the drunk drivers died and a shocking 27 sober road users were also killed in the accidents. This means for every 100 drunk drivers killed, approximately 81 other people (or 44%) killed in drink driving accidents were killed because somebody else made the wrong decision.

Between 2008 and 2010 40% of drunk drivers involved in crashes were under the age of 24 and 82% of all drunk drivers involved in serious crashes were male. Being 17, having my restricted licence, a car to drive, and also going to parties, this means I fit the profile of a high risk (potential) drink driver perfect and it is therefore a matter that concerns me greatly. As well as the responsibility to make sensible decisions and be a responsible road user myself, I'm now also at risk of all those who don't make the right decisions and choose to drive drunk anyway. This is why I have chosen to do an investigation on how blood alcohol levels are reflected in the seriousness of a crash involving alcohol.

Based on prior knowledge of the nature of crashes involving alcohol I hypothesise that the blood alcohol levels recorded for major crashes will me much greater than those recorded for minor crashes.

Investigation

I will compare recorded blood alcohol levels (mg/100ml blood) of the drivers for crashes deemed as being serious (fatal or otherwise) and crashes deemed as minor (non fatal, no serious injury) from data recorded for crashes in New Zealand involving alcohol.

Problem

I wonder what the difference in median blood alcohol levels is between recorded blood alcohol levels in serious crashes and recorded blood alcohol levels in minor accidents, for drivers in New Zealand who were involved in a crash involving alcohol?

Analysis

Sample
Module: 2 Sample Bootstrapping
Variable: Crash.severity

Spread
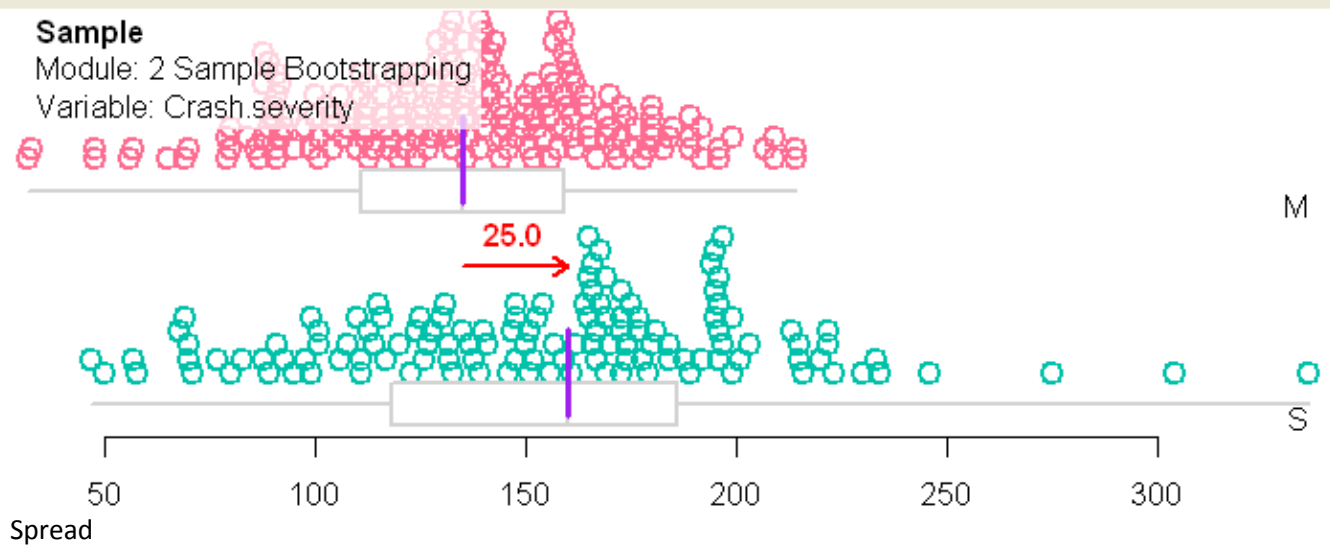
As you can see in the box plots the range in blood alcohol levels recorded in major crashes is much greater than that recorded in minor crashes ranging from between 45mg and 350mg, a range of 305mg as opposed minor crashes which range between only 25mg and 220mg a range of 195mg. However both graphs shows that both minor and serious alcohol related crashes begin occurring as early as 45ml alcohol per 100ml of blood and alarmingly this data also shows that beyond 220ml alcohol per 100ml of blood only serious crashes occur.

Centres

The median blood alcohol level recorded at 165ml for serious crashes is a mere 25ml more than the median blood alcohol levels recorded for minor crashes which is recorded at 140ml.

Centre 50%: Overlap

As the upper quartile of blood alcohol levels for minor crashes occurs directly in line with the median of the blood alcohol levels for serious crashes, 75% of blood alcohol levels recorded in minor crashed occur below 50% of blood alcohol levels recorded for serious crashes. Also the lower quartile for blood alcohol levels recorded in minor crashes falls just below that of the lower quartile for serious crashes meaning that not only does almost 50% of blood alcohol levels for minor crashes fall within 50% of that of major crashes but it almost fit completely within the lower quartile for that of serious crashes.

Centre 50%: Shape

The shape of the data's distribution for serious crashes is much different compared to that of minor crashes is significantly different. For minor crashes there seems to be a large peak of samples with a blood alcohol level close to the median level of 135mg/100ml (as would be expected) and then a second peak right above the upper quartile mark of about 160mg/100ml. For serious crashes the

data is distributed rather evenly giving a flat shape until it gets to the median of 160mg/100ml where there appears to be a low, followed by and abrupt peak just on the right side of the median. Over all the data for minor crashes appears to be slightly left skewed whilst the data for serious crashes appears to be right skewed

Inference
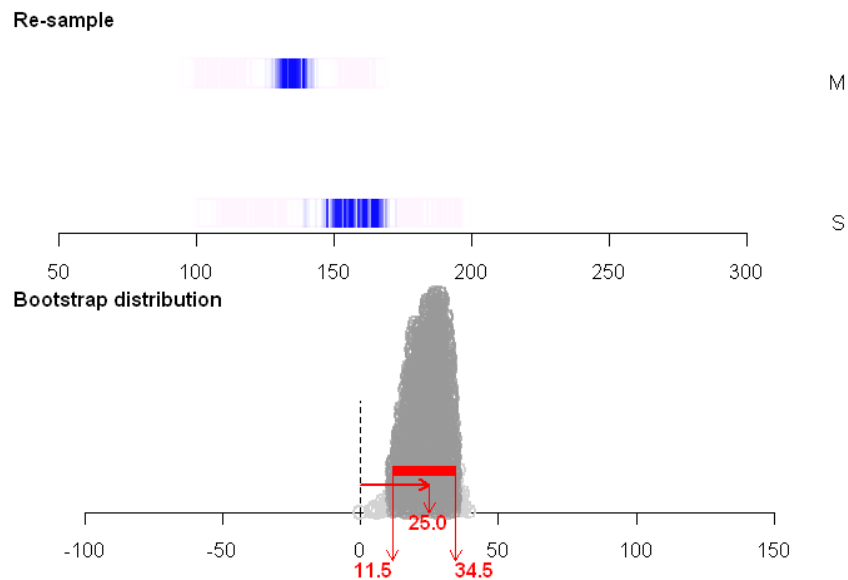
I'm fairly sure that the population median blood alcohol level recorded for serious crashes greater than the blood alcohol levels recorded for minor crashes, for crashes in New Zealand involving alcohol. From this sample of crashes involving alcohol I'm pretty sure that the population median of blood alcohol levels recorded for major crashes is between 11.5mg/100ml and 34.5mg/100ml greater than the population median blood alcohol level recorded for minor crashes (assuming this sample is representative of the population).

Conclusion

Based on my samples I can conclude that there is in fact a difference between median blood alcohol levels recorded for serious crashes and median blood alcohol level recorded for minor crashes. This is backed up by the bootstrapped confidence interval which shows that the population blood alcohol levels recorded for serious accidents do tend to be higher than those recorded at minor crashes by between 11.5 and 34.5mg/100ml blood.

With a bootstrapped confidence interval as large as 11.5mg/100ml to 34.5mg/100ml, I am fairly certain that should another sample be taken and processed the resulting confidence interval would still remain above zero and indicate that blood alcohol levels recorded for serious crashes are greater than blood alcohol levels recorded for minor crashes.

|	
	Grade Boundary: Low Achieved
5.	For Achieved the student is required to use statistical methods to make a formal inference, showing evidence of using each component of the statistical enquiry cycle.
	A comparison question ("*Do high risk…*") has been posed, but it does not include looking at the difference in the medians between the two population groups. The population in the question is unclear, but it is given in the background information and there is a reference to the difference in the medians in the conclusion (1).
	Summary statistics have not been given.
	In the analysis comments there is an attempt to link back to the context (highlighted).
	The confidence interval has been correctly interpreted and used to make a correct call (1).
	For a more secure Achieved the question needs to identify the population and make reference to the difference in the medians.

Assessment
New Zealand Crash Data

Problem:

In 2011 there were 1409 serious or minor crashes where alcohol or drugs were recorded as a factor. A random sample was taken from these drivers and they were interviewed in person by researchers. 317 samples were taken to analyse. The variables that I will be using are different risk groups and the blood alcohol level. The Two risk groups are a high risk group which is an age group between 15-24 and a low risk group which is between the ages of 50-59.

Do the high risk age group tend to have a higher blood alcohol level to that of people in the low sick age group?

Data:
A random sample was taken from these drivers and they were interviewed in person by researchers. 317 samples were taken to analyse.
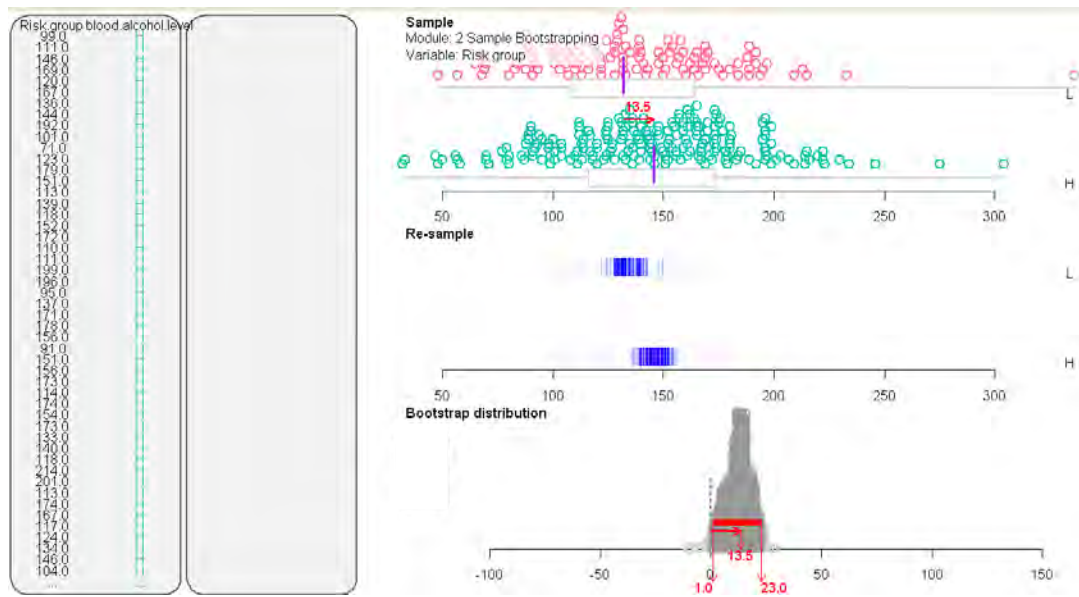
Analysis:

Overall Visual Comparisons:
My initial impression of this sample is that people aged 15-24 (high risk group) have their blood alcohol levels more spread out than people aged 50-59 (low risk group). Also there are more people in the high risk group than there are in the low risk group. You can see this clearly in the sample.

Centres:
The sample median blood alcohol level for people aged 50-59 (low risk group) is 13.5 milligrams of alcohol per 100 millilitres of blood lower than people aged 15-24 (high risk group).

I was surprised how many high risk drivers there were involved in serious and minor crashes in 2011. Out of the 317 people involved in this sample, 206 were high risk drivers and only 111 were low risk drivers.

Shape:

The distribution of both high and low risk groups are slightly right skewed, with most of the data between 0 milligrams of alcohol per 100 millilitres of blood and 150 milligrams of alcohol per 100 millilitres of blood.

Unusual:

In the sample the person who has the highest blood alcohol level is 336 milligrams of alcohol per 100 millilitres of blood yet they are in the low risk group. There must have been some extreme circumstances which is why they are an outlier.

Shift/Overlap:

There is large overlap between the middle 50% of the blood alcohol level for both risk groups with the median not too far away from each other. Still, this overlap is not extreme and with the medians being different I can make a clear call on what risk group has the higher blood alcohol level.

I will calculate a bootstrapping confidence interval anyway to confirm this.

Statistical Inference:

From the bootstrapping confidence interval you can see that the median blood alcohol level for high risk drivers is somewhere between 1 milligram of alcohol per 100 millilitres of blood and 23 milligrams of alcohol per 100 millilitres of blood higher than the blood alcohol level for those in the low risk group.

Conclusion:

From these samples, I can make the call that there is a difference in the blood alcohol levels for people in the low risk group (people aged 50-59) and the people in the high risk group (people aged 15-24). That is, I can make the call that people in the high risk group in New Zealand have a higher blood alcohol level.

I am pretty sure that the median blood alcohol level for high risk drivers is somewhere between 1 milligram of alcohol per 100 millilitres of blood and 23 milligrams of alcohol per 100 millilitres of blood more than the median blood alcohol level for low risk drivers.

I can make the call from these samples that there is a difference in the blood alcohol level for low and high risk drivers involved in crashes in New Zealand.

My sample did highlight that the middle section of the high risk group had more drivers involved to that of driver in the low risk group. The median would have been influenced by having more high risk driver than low risk drivers.

	Grade Boundary: High Not Achieved
6.	For Achieved the student is required to use statistical methods to make a formal inference, showing evidence of using each component of the statistical enquiry cycle. The comparison question does not identify looking at the difference in the population medians and the population is not completely clear (1). While there are descriptive comments about the sample distributions, the units ("*milligrams*") are incorrect, and there is no attempt to discuss what the observations mean. The interpretation of the confidence interval ("*After 1000 bootstrapping…*") in the first paragraph of the analysis is incorrect. The student has made an appropriate formal statistical inference and used the inference to make a correct call (2). To reach Achieved the analysis discussion would need to be more clearly linked back to the context and the question would be improved.

Problem: The purpose of my investigation is to see if high risk groups tend to have a higher blood alcohol level than low risk groups involved in their car crashes. ①

Plan: The data is of 145 females and 172 males involved in a crash that has alcohol involved. Blood alcohol level is per 100 millilitres of blood. High risk age group is 15-24 years low risk age group is 50-59 years. The blood alcohol level in New Zealand is under is 20 years zero blood alcohol level over 20 years is 80 milligrams blood alcohol level per 100 millilitres refer to http://www.alac.org.nz/alcohol-you/drinking-driving.

Analysis: Level of alcohol in blood of high risk group and low risk group in their car crashes. The sample median of the data shows high risk group median is 13.5 milligrams bigger than low risk groups sample median. After 1000 bootstrapping high risk groups sample median is between 0.0 and 23 milligrams bigger than low risk groups sample median this shows high risk groups tend to have higher blood alcohol level than low risk groups.
The middle 50% of alcohol in blood of high risk group overlaps with the middle 50% of and low risk groups, low risk groups upper quartile is 164 milligrams and lower quartile of high risk group is 116.2 milligrams, an overlap of 47.8 milligrams. The shape of the data is normal distributed (bell curve). Low risk group has one outlier of 336 milligrams it does not fit with the other data.
The middle 50% of high risk groups slightly more spread out than the middle 50% of low risk groups.

Summary of blood.alcohol.level by Risk.group

	Min.	1st Qu.	Median	Mean	3rd Qu.
H	32	116.2	145.5	144.8	173
L	48	108.0	132.0	137.0	164

	Max.	Std.dev	Sample.Size
H	304	44.12	206
L	336	42.80	111

Sample
Module: 2 Sample Bootstrapping
Variable: blood.alcohol.level
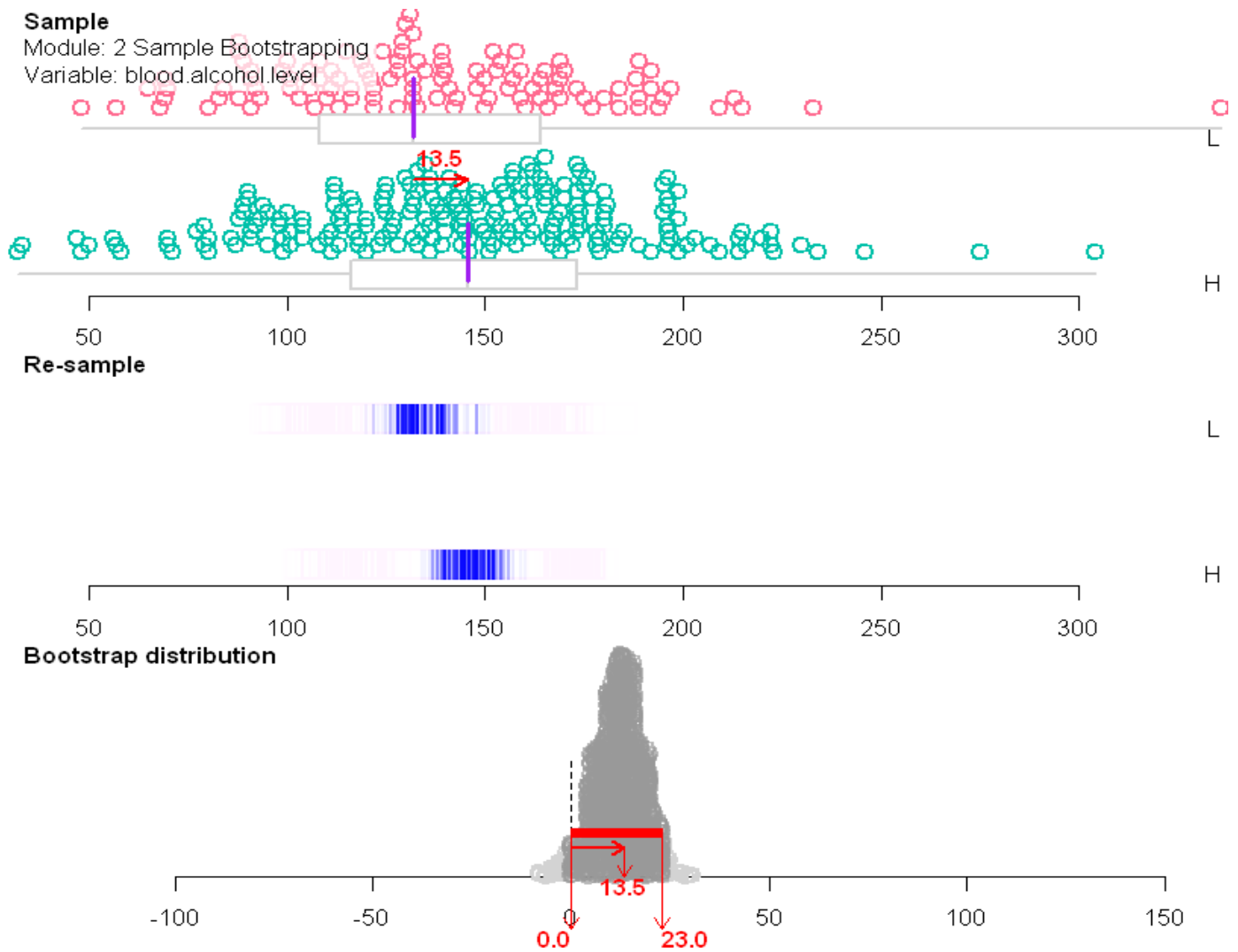
Conclusion I am pretty sure that the population median of blood alcohol level for high risk groups involved in car crashes in New Zealand is somewhere between 0 milligrams and 23 milligrams higher than the population median low risk groups

2

Based on these bootstrapping 1000 samples I would make the call that there is a difference in blood alcohol level between high risk groups in New Zealand and low risk groups. High risk groups tend to have a higher blood alcohol level than low risk groups in New Zealand in car crashes.

The bootstrap confidence interval is positive, indicating that blood alcohol level between high risk groups in New Zealand and low risk groups tend to be higher for high risk groups.