Assessment Schedule – 2017

Scholarship Statistics (93201)

Evidence Statement

Evidence Statement

General Principles:

1. Ignore incorrect answers if alongside correct answers. The exception is contradictory statements.
2. Ignore minor copying errors.
3. When required in evidence, answers need to be contextual.

QUESTION ONE

Tasks Q1 (a)(i)
Evidence:

- Overall the blood alcohol level displays no linear relationship with age OR there is an outlier at (52,300).Two distinct subgroups are evident. For the younger age grouping less than 26 there is a strong positive linear correlation. However for the older age grouping, 29 years and over no relationship is evident.

Note:

1. Accept three groups.
2. One mark is allocated for mention of subgroups.

Task Q1(a)(ii)
Evidence:

- Prediction is 123 mg/100mL. Reservations are that a line isn't an appropriate fit to the overall data. Also when the age is 27 there is a wide fluctuation in the scatterplot and that the relationship isn't linear at that age point.

Note:

1. Any level of accuracy was accepted for the prediction**.**

Task Q1(a)(iii)
Evidence:

- Two further variables could be a person's weight and how much food had been consumed prior to testing. In both cases the expected relationship would be negative.

Note:

1. Must have variable with relationship to get mark. E.g. negative: time between drinking and driving.
2. Categorical variables with be accepted with appropriate descriptions. E.g. type of alcoholic drink. Gender wasn't accepted.

Task Q1(b)
Evidence:

- The blood alcohol level for the older group was 45.75 mg/100mL more than for the younger group.
- The bootstrap distribution shows that the mean blood alcohol level for the older group is likely to be between 71.55 and 23.10 more than for the younger group.
- Zero is not included in interval there is evidence to establish that the older group had a greater mean blood alcohol level than the younger group.

Overall Judgement for Q1
Max 2 for (a)(i), max 2 for (a)(ii) and max 2 for (a)(iii). Max 5 for (a). Max 3 for (b).
Overall max 8 marks.

QUESTION TWO

Task Q2(a)
Evidence:
From Figure 3:

- Fatal crashes involving alcohol or drugs are most common on between 8pm Fri to 4 am Sat, where the average number of crashes peaked at 6 in each 4-hour time period. There were no fatal crashes involving alcohol or drugs recorded for both Tues and Wed between 8am and midday.
- The seasonal pattern shows an increasing trend from Tues through to Sat then declining through Sun and Mon.
- On Sat 0000- 0359 the average of 6 crashes involving alcohol or drugs is approximately 70% of the fatal crashes of all causes, whereas on Sun 0000-0359 the 1.2 average is about 65% of all fatal crashes, meaning that the average number of fatal crashes (all causes) at this time on Sat is approximately 8.5 whereas on Sun at this time the average is approximately 2.

From Figure 4:

- Of all drivers involved in fatal crashes, the 20-24 age group are the most likely to be affected by alcohol/drugs. This is shown in both the highest annual average number of crashes (~17) and highest percentage of the age group (~37%)
- The age group 35-39 has a similarly high percentage of the age group (~35%) involved in a fatal crash where alcohol or drugs were involved but the annual average number is much lower (~12).
- For drivers aged 40 or more, crashes involving alcohol/drugs remains between 5-10% of the age group and around an annual average of 5 fatal crashes.

From Figure 5:

- The overall trend in total number of fatalities in crashes involving alcohol or drugs each year has declined from about 320 in 1990 to 95 in 2015 – a decrease of roughly 8.6 per year on average.
- Between 1990 and 2000, there was a steady decline of approximately 19 fatalities per year, between 2000 and 2010 the number of fatalities hovered between 100 and 150 and after that the number of fatalities has lowered to just under 100.

Note:

1. One mark per comment with at least one from each figure.
2. Peaks/lows was scored as one point.

Task Q2(b)
Evidence:

- Sat 40% fatal crashes involve alcohol/drugs => 60% do not

  So 60% of N = 4.5 fatal road crashes on average where N = 7.5. Note: 40% corresponds to 3.

  Sun 20% fatal crashes involve alcohol/drugs=> 80% do not

  So 80% of M = 8.0 fatal road crashes on average where M = 10.0. Note: 20% corresponds to 2.

  Hence $\frac{8.0}{4.5}$ = 1.8. It is 1.8 times as likely for a fatal crash on Sun 4pm to 8pm to not have alcohol or drugs as a factor as it is for a fatal crash on Sat in the same time interval.

Note:

1. One mark for indication of correct method.
2. Twice as likely was accepted.

Task Q2(c)
Evidence:

- New cars have better safety features, decreasing the likelihood of a fatality.
- Road design has improved (e.g. more safety barriers), decreasing the likelihood of crashes.
- Driving drunk campaigns has made it socially unacceptable to drive after consuming alcohol- decreasing the likelihood of crashes and fatalities with alcohol as a cause.

Task Q2(d)
Evidence:

- An estimated weekly average is just under 2 (1.8). Tues average is (0.2, 1.2, 0, 1, 1.2, 1.2) / 6= 0.8

  Seasonal effect for Tues 0.8 – 2 = -1.2

  Seasonal effect for Tues is approx. 1.2 below the weekly average

Note:

1. One mark for justification of working and one mark for answer.

Overall Judgement for Q2
Max 4 for (a), max 2 for (b), max 2 for (c)and max 2 for (d). Overall max 8 marks.

QUESTION THREE

Task Q3(a)(i)
Evidence:
Construct the following frequency/contingency table:

	Number of Occupants in Car			
	One	**Two - Three**	**Four or more**	**TOTAL**
Fatal	33	15	47	95
Serious	182	78	195	455
Minor	810	645	45	1500
TOTAL	1025	738	287	2050

Thirty-three crashes were classified as fatal when there was only one occupant in the car.

Task Q3(a)(ii)
Evidence:
(i) 645/2050 = 0.3146
(ii) 195/455 = 0.4286

Task Q3(a)(iii)
Evidence:
Pr(Fatal Crash and One Occupant in car) = 33/2050 = 0.0161
Pr(Fatal Crash) x Pr (One Occupant in Car) = 95/2050 x 1025/2050 =0.0232
As 0.0161 ≠ 0.0232 events are not statistically independent. (Events need to be defined in answer).

Task Q3(a)(iv)
Evidence:
Ratio is 810:33 which is equivalent to 24.5:1. So 24.5 times more likely.

Task Q3(b)(i)
Evidence:

Number of Serious Crashes per Quarter	90 – 109	110 – 129	130 – 149	150 – 169	170 – 189	190 – 209	210 – 229	230 – 249
Probability	0.04	0.08	0.16	0.20	0.24	0.16	0.10	0.02
Normal Probability	0.0262	0.0765	0.1574	0.2267	0.2288	0.1616	0.0800	0.0277

Probabilities match closely with the raw data, which suggests that a normal distribution with mean = 170 and standard deviation 33 would be a suitable fit to these data.

1. No marks for either cumulative probability or cumulative frequency matching. Minimum of 4 matches.
2. Can use no continuity correction in calculating normal probabilities to get 0.0246, 0.0725, 0.1495, 0.2157, 0.2176, 0.1536, 0.0759, and 0.0262.
3. Full reasoning backed up with calculations using the bell-shape distribution of the data is acceptable. Just matching mean and standard deviation was deemed insufficient. Must have two and three sigma limit matching.

Task Q3(b)(ii)
Evidence:
Probability of fewer than 160 crashes in one quarter = $\frac{2+4+8+5}{50} = 0.38$ (or can use normal model to get 0.3753).
Over 4 quarters using binomial distribution, probability = $(0.38)^4 = 0.0209$.

- The assumption is that the event of fewer than 160 crashes occurring in each quarter is independent of any other quarter.
- The probability of fewer than 160 crashes is the same in each quarter.

Overall Judgement for Q3
Max 2 for (a)(i), max 2 for (a)(ii), max 2 for (a)(iii), and 1 mark for (a)(iv). Overall so far max 5. For (b)(i), 2 marks max, and (b)(ii) 2 marks (one for answer and one for assumption).
Overall max 8 marks.

QUESTION FOUR

Task Q4 (a)
Evidence:

The group given no alcohol is a control group. Having a group of people having consumed no alcohol allows the researchers to compare the driving of people who have consumed alcohol with those who have not. The researchers would see the effect of consumption of alcohol on driving performance.

Note:

1. An explanation is required. Just stating the first sentence was deemed insufficient.

Task Q4 (b)
Evidence:
Different blood alcohol levels allow the researches to investigate if a greater blood alcohol level has a greater impact on driving performance.

Task Q4 (c)
Evidence:
Drivers were tested over a period of time to see how long the effects of alcohol on driving performance lasted and if driving performance changed over time after consuming alcohol.

Task Q4 (d)
Evidence:
Drivers with no alcohol still had to react to hazards on the road. Also, it is possible that drivers were unfamiliar with a driving simulator and so might not perform as well as if in a real car.

Task Q4 (e)
Evidence:
Left edge line crossings
Drivers with no alcohol averaged around 1 edge line crossing in all of the 5 test blocks. Drivers who had consumed alcohol had an average of 2 or fewer edge line crossings on the first test, this rose to about 5 by test 3 and then decline to about 4 by test 5. In all cases the drivers with the high dose had slightly fewer edge line crossings than drivers with the medium alcohol dose.

Centre line crossings:
Drivers who had no alcohol averaged between 7 and 8 centre line crossings over the test blocks. Drivers who had alcohol started at a similar average at the first test, and increased to about 10 by test 2. The medium dose group continued to increase to 11 by test 3 and then decreased back to 10 by test 5. The high dose group remained at an average of 10 at test 3, followed by a sharp rise to 11 for tests 4 and 5.

Task Q4 (f)
Evidence:
Randomisation test.

- The mean / median number of centre line crossings drivers who had no alcohol and the mean / median number of centre line crossings of those who had consumed alcohol is calculated.
- The difference of the means / medians is calculated.
- The number of centre lines for each person is re-randomised and assigned at random to one of the two groups – alcohol or none.
- The difference of the re-randomised means / medians is calculated.
- The re-randomisation process is repeated many times (say 1000), producing a distribution of differences of means / medians.
- The difference in means / medians of the actual experiment is compared with this distribution to see how likely such a difference would occur by chance.
- If there is a small chance (less than 5% or 10%) that the experimental difference would occur, this is evidence that the observed difference has come about because of the different treatment of the groups (alcohol or not).

Note

1. Hypothesis test or bootstrap confidence interval test descriptions in context are acceptable alternatives.
2. One mark is allocated for name and two marks for description.

Overall Judgement for Q4
Max 1 mark for (a), (b) max 1 mark, (c) max 1 mark, (d) max 1mark, (e) max 3 marks: Overall so far max 6. (f) max 3 marks.
Overall max 8 marks.

QUESTION FIVE

Tasks Q5 (a)(i)

Evidence:

Percentage holding driver's licences = **91.7%** (1527/1666)

Tasks Q5 (a)(ii)

Evidence:

	Year		
	2014	2015	2016
Women	90%	94%	89%
Men	73%	76%	62%
Difference	17%	18%	27%

- For every year, the percentage of women (90%) that felt comfortable driving when they limited themselves to one or no drinks always exceeded that of the men (70%). The difference stayed virtually unchanged between 2014 and 2015, however it rose to 27% in 2016.

Task Q5 (a)(iii)

Evidence:

Use a multiple bar chart with percentages on the vertical scale. Three categories: one or no drinks, two drinks, and three or more drinks, would feature on the horizontal axis. The legend would distinguish between "admitted to driving while slightly intoxicated" and "not admitting to driving while slightly intoxicated".

Note:

1. Stacked bar charts or two pie charts with appropriate description was acceptable.

Task Q5 (a)(iv)

Evidence:

- In figure 7, the % of those who felt that the drink-driving laws aimed at reducing the road toll were effective, stayed steady at about 60% over the 20 years, 1995 to 2015.
- The percentages of drivers who were stopped at checkpoints in the last year, and the percentage of drivers who said that the legal blood alcohol limit should be lower, both stayed at about 40% between 1995 and 2006.
- After 2006, both percentages rose to peak in 2010, with the percentage wanting a lower legal blood alcohol limit being the highest peak at 63%. This peak was 13% higher than the peak at 50% for the percentage of those stopped at checkpoints last year. After that both percentages levelled off.

Task Q5 (b)(i)

Evidence:

- Between 2010 and 2013 the percentage who felt the legal blood limit should be lower were within the margin of error for the population percentage estimate of those who felt the legal blood limit should be lower than was calculated in 2010.

Note:

1. Discussion of sampling variability and / or sampling error is acceptable.
2. A confidence interval based answer is acceptable.
3. "Due to chance alone" not acceptable.

Task Q5 (b)(ii)

Evidence:

- Both variables are increasing at a similar rate so relationship would be positive, moderate to high, and linear. It's likely that both variables would be correlated as a result of increasing public awareness of the benefits of a lower blood alcohol limit, and with an increasing number of checkpoints.

Note:

1. Must have linear otherwise only one mark.

Task Q5 (b)(iii)

Evidence:

Let A = event of being slightly intoxicated.

Let B = event of being comfortable driving when only one or no drinks had been consumed.

We have:
Pr(A) = 0.18, pr(B given A) = 0.49 and pr(B given not A) = 0.83.

So pr(B) = $0.18 \times 0.49 + 0.82 \times 0.83 = 0.7688$

Hence pr(A given B) = $\dfrac{0.18 \times 0.49}{0.7688}$ = 0.115 i.e. **12%**.

Note:
1. One mark for some idea of conditional.

Overall Judgement for Q5
Max 1 mark for (a)(i), max 2 marks for (a)(ii), max 1 mark for description (a)(iii), max 2 marks for (a)(iv). Overall so far max 4 marks. Max 1 mark for (b) (i), max 2 marks for (b)(ii), max 2 marks for (b)(iii).
Overall max 8 marks.