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.

2017

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 $\neq$ 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.

Assessment Schedule – 2018

Scholarship Statistics (93201)

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)
Evidence:

- Rural –
  - Trend: Linear, non-linear or piecewise
  - Strength: Very weak or weak or moderate
  - Direction: positive or changing direction.

  Eg for direction and piecewise trend: Initial increase in sales for advertising spend up to \$400 to \$600, then no further increase in sales for additional advertising spend.

  Urban – linear, strong and, positive correlation between advertising and sales for most data but with outlier at (12, 0.5).

Task Q1(b)
Evidence:

- For rural, predicted sales = 3.808 (000\$) \$3800 [Using $S = 0.0825A + 3.3955$ Due to small sample, considerable variation and no reason to expect a decrease in the context, the simpler linear model may be appropriate.]

  OR 4.141 (000\$) \$4140 [Using $S = -0.0356A^2 + 0.4783A + 2.6392$ Due to the fact that the trend is non-linear, a non-linear model might be more appropriate even though the parabola will not be a good fit for higher advertising spends].
- For urban, predicted sales = 4.51 (000\$) \$4500. [Using $S = 0.1263A + 3.8788$. Since the sample is small and we do not know the reason for the outlier, it is may not be valid to remove it in this context. When considering return on investment, information about losses shouldn't be ignored]

  OR 4.57 (000\$) \$4600 [Using $S = 0.29A + 3.12$ Since the outlier is so far from the other points, it will have a large effect on the trendline and it might be better to remove it].

Task Q1(c)
Evidence:

- For rural there is a change in the trend at about \$500, which means that it is difficult to model reliably for \$500.
- For rural, there is considerable variation in the scatter about the trend so the prediction is suspect. *Accept weak correlation for rural.*
- For urban (if outlier included) the effect of the outlier has been included in the model so the prediction may not be valid if the outlier turns out not to be valid data.
- For urban (if outlier excluded) the effect of the outlier has not been included in the model so the prediction may not be valid if the outlier turns out to be valid data.
- The small sample size is an issue for rural because of the variability in the data since variability may be due to sampling variation rather than representative of the underlying relationship. *(no mark for small sample size without explanation in terms of variability and sampling variation).*
- The small sample size is an issue for urban as the outlier will have a big effect on the trend for a small sample. We do not know whether the outlier represents an error or a real problem which means that sometimes the advertising spend does not result in sales. *(no credit for small sample size without explanation in terms of effect of outlier).*
- If discussing parabolic model, it does not make sense in terms of the context that an increase in advertising would result in a decrease in sales.

Task Q1(d)
Evidence:

- Number of competitors – expect a negative correlation.
- Score location from 1 = least favourable to 7 = most favourable – expect a positive correlation.
- Income of the population or economic changes, positive correlation as more sales when times or income are better, fewer sales when times/income are worse.
- Number of customers or size of store or number/total salary of staff positively correlated with sales.

  Note: *Accept other reasonable numeric variables with comment on direction of correlation.*

Task Q1(e)
Evidence:

- With the strong relationship between advertising spend and sales for most data from the urban outlets and $1,400 being not too far from the given data it, is possible further sales could be expected, though the outlier indicates that the prediction may not be reliable. Other factors like location and demand come into play. The claim of no increase is not supported.

  There is unlikely to be any further increase in sales in the rural sector since there is no increase in the sales for spends over $600 in the scatterplot. Claim of no increase would be supported in the rural sector.

  Note: *Sufficient to discuss claim for one with discussion of forecast for both. Discussion must be based on observation of the graph, not on the model.*

QUESTION THREE

Task Q3(a)(i)

Evidence:

- The sentiment score is bounded by 0 and 1, while the normal distribution is unbounded.
- Considering two standard deviations either side of the mean would result in an expected central 95% for sentiment scores of (–0.036,1.112), which include scores below 0 and above 1 *(or other probability calculation with interpretation*

  e.g. $P(X<0)=0.0304$) and $P(X>1)=0.0537$, so over 8% of the model is outside the observed distribution, which is a big proportion to exclude

  OR $P(X<0.8)=0.8194$, which is 7 percentage points higher than 0.749 observed, which is a large % difference.

Task Q3(a)(ii).

Evidence: *(Must include some comparison of theoretical vs observed probabilities or comparison of graph to Poisson distribution – could be overall evaluation of model).*

- Sample distribution is positively skewed (similar shape to Poisson distribution model for small λ)
- Calculate the mean number of letters per word as $(1\times0.05+2\times0.07 \ldots\ldots 12\times0.0025)=4.9975$ *(CAO)*
- Use the mean of 4 as an estimate of the Poisson parameter λ (since we are modelling $N-1$) to calculate a probability.
- Create a correct table of expected probabilities (at least 8 out of 12 correct).

E.g. calculating table or probabilities of events

Number of letters (N)	1	2	3	4	5	6	7	8	9	10	11	12
N – 1	0	1	2	3	4	5	6	7	8	9	10	11
Observed proportion	0.05	0.0875	0.145	0.1825	0.1325	0.1375	0.13	0.055	0.045	0.0225	0.01	0.0025
Expected proportion	0.0183	0.0733	0.1465	0.1954	0.1954	0.1563	0.1042	0.0595	0.0298	0.0132	0.0053	0.0019

- Specific comparison observed vs theoretical: e.g. Under a Poisson distribution with $\lambda=4$, $\Pr(X=3)=0.1954$ (using tables) and the observed proportion is 0.18 which is similar. e.g. Similarly, we would expect 0.156 of words to have 6 letters but the observed is 0.14 which is 12% less than expected so fairly different.
- General comparison of the observed proportions with the expected proportions: e.g. (under the proposed model), we get a close fit for most numbers of letters. This would suggest the Poisson distribution as being a reasonable model
- OR (under the proposed model), we get do not observe a close fit for most numbers of letters. This would suggest the Poisson distribution is not an appropriate model or justified by difference in probabilities for X = 5 or discussion of X<3 or X>7.
- Discussion of conditions for Poisson e.g. the rate is number of letters in 1 word, so probability of occurring is proportional to the number of words being measured, letters do not occupy the same space, so are not simultaneous, the occurrence of each letter after the first is unpredictable, so may be considered random. While letters are not independent, the word length is unpredictable, so letters may be considered to occur independently.

Task Q3(b)

Evidence:

1. The mean reading speed for the Website A sample was 11.62 words per minute higher than for the Website B sample.
2. The bootstrap distribution shows that the mean reading speed for articles from Website A is likely to be between 2.99 and 19.40 words per minute higher than that for Website B.
3. Zero is not included in interval; there is evidence to establish that Website A has a higher mean reading speed than the Website B.
4. Therefore, can conclude that the mean reading speed for website A is significantly different from the mean reading speed for website B.

Note: Mean reading speed must be mentioned in at least one place. Only accept points 3 or 4 if point 2 is correct. The full context may be considered holistically over the student's answer.

QUESTION FOUR

Tasks Q4(a)(i)

Evidence:

- The women who test positive for pregnancy will include those who are not pregnant (false positives).
- Calculation with interpretation (may be algebraic) For example, if overall 10% of women who take this test are pregnant,

then you would expect $0.1 \times 0.98 = 0.098$ true positives, and $0.9 \times 0.04 = 0.036$ false positives, and the proportion of women who are pregnant given a positive test result.

$= \frac{0.098}{0.098+0.036} = 0.731$, which is not 98%.

- The claims made by the developer would be based on the results from a study (or several studies). Therefore, the claims are based on estimates for the true specificity (true positive rate) and true sensitivity (true negative rate) of the pregnancy test. The estimates may differ from the true positive and negative rates because of sampling variation (or randomness or chance).

Task Q4(a)(ii)

Evidence:

- $\Pr(\text{pregnant} \mid \text{positive test}) = \frac{44}{55}$
- $\Pr(\text{positive test} \mid \text{pregnant}) = 0.94$
- $\Pr(\text{negative test} \mid \text{not pregnant}) = 0.81$
- Number of pregnant women in study $= \frac{44}{0.94} = 47$
- Number of not pregnant women in study $= \frac{11}{0.19} = 58$
- Therefore, $47/105 = 0.4476$ of the women in this study were pregnant.

Note:

Alternate methods are acceptable e.g. two-way table.

Task Q4(b)(i)

Evidence:

- People were surveyed over a period of four months to reduce bias since the views are not overly affected by a particular event in the short term.
- Stratified sampling was used to assist representativeness of sample for different regions in the population.
- Random sampling of telephone numbers to make it likely that the sample is representative of the population.
- CATI used to help with consistency of interviews *(do not accept if response implies that questions were asked by or answered on a computer)*
- Use of multi-choice on a scale to ensure consistency of answers and ability to be analysed.
- Sample size was sufficiently large to reduce variability in responses due to chance/randomness/sampling variation.
- Do not accept sample weighting as a strength since it is explained in the article.

Note: Any two strengths accepted.

Task Q4(b)(ii)

Evidence: *(must describe issue and explain how it can cause bias in this context)*

- Use of telephone directory (white pages) excludes those who do not have landlines or who have made their phone number private (potential selection bias). People with landlines may be richer or more conservative and have different attitudes to alcohol consumption than people without.
- The nature of the questions asked (drinking during pregnancy) could lead respondents to not be honest with their responses (potential behavioural considerations).
- With the nature of the questions asked (drinking during pregnancy), the gender of the interviewer could influence respondent's answers, for example a female interviewer asking the question could result in more or less honest answers than a male interviewer (interviewer effect).
- Telephone surveys tend to have high non-response, and those who are home to answer the phone may have a different lifestyle than those not home (may have different attitudes to alcohol consumption).
- Question effects such as not defining what "small amounts of alcohol" means, so non-drinkers might interpret it differently from heavy drinkers (do not accept criticism of standard questionnaire responses such as "agree" and "strongly agree").
- The time period of the survey included the Christmas season and the summer holidays when many people drink more alcohol, which might have affected their attitude to drinking compared to other times of year.

Note: *Any two points above or other reasonable non-sampling errors are acceptable. Must relate specifically to this survey, not be general points about possible leading questions etc*

Task Q4 (b)(iii)

Evidence:

The percentage of respondents who disagreed with the statement 'During pregnancy drinking small amounts of alcohol is OK' may have varied between the three survey years (i.e. they may not all have been 84%), but the differences between these percentages could be due just to chance / sampling variability.

Task Q4(b)(iv)

Evidence:

- The group size is smaller for "Other" than for "Degree", which means the margin of error will be larger for the "Other" group.
- The survey percentage (85%) is closer to 50% for the "Other" group, which means the margin of error will be larger for the "Other" group.

2019

QUESTION TWO

Task Q2(a)

Evidence:

- Both time series (total usage of Auckland Transport and estimated population of Auckland Region) show a positive trend over 2006 to 2018 which, due to the use of y-axis scales, looks like similar rates of increase for both series. However, in terms of comparing *rates of increase* for the two series, we need to use relative measures.
- Total usage has increased from around 52500(000) in 2006 to around 95100(000) in 2018, an increase of around 81%. Estimated population has increased from around 1370(000) in 2006 to around 1700(000) in 2018, an increase of around 24%. So total usage of Auckland Transport has increased at a faster rate than the estimated population of the Auckland Region, over the period 2006 to 2018.

  Note: Alternatively, in 2006, the rate of usage per person (based on estimated population) was around 52500(000) / 1370(000) = 38.3 but in 2018, the rate of usage per person was around 96000(000) / 1695(000) = 56.6. If the rates of increase were similar for the two series, then the rate of usage per person would be similar. Accept piecewise analysis considering separate trends before and after 2013 (both usage and population increase more steeply after 2013).

Task Q2(b)

Evidence:

General:

- More people use trains than ferries, with an average of around 5100(000) uses per quarter for trains by the end of 2018 and an average of around 1550(000) uses per quarter for ferries by the end of 2018.

Trends:

- Over the period 2006 to 2018, usage of both trains and ferries displays an overall positive trend.
- (Although more obvious for trains), after increasing over 2006 to 2011, the usage of trains and ferries plateaued during the period 2011 to 2013, before increasing again from 2014.

Seasonality:

- The usage of both ferries and trains displays a seasonal pattern.
- For trains, there are regular peaks in the 3rd quarter of each year and regular lows in the 4th quarter of each year, whereas for ferries there are regular peaks in the 1st quarter of each year and regular lows in the 3rd quarter of each year.
- The seasonal fluctuations for trains are similar in size between 2006 and 2018, however for ferries the seasonal fluctuations have increased in size from 2015 onwards.

Note: At least one of the points must contain numerical evidence.

Task Q2(c)(i)

Evidence:

- Model 1 forecast 2020 Q2: 19549.90 (17850.78, 21249.01).
- Model 2 forecast 2020 Q2: 20543.03 (19680.01, 21406.06).
- Forecast from Model 1 is lower than forecast from Model 2.
- Forecast interval from Model 1 is wider than forecast interval from Model 2.

Task Q2(c)(ii)

Evidence:

- Trend: The short-term trend for the last three years of bus usage was steeper than previous years, hence Model 2 gave a higher forecast than Model 1.
- Variation in data: Model 1 was based on more data which varied more than the data on which Model 2 was based, which has contributed to the wider forecast interval for Model 1.

QUESTION THREE
Task Q3(a)
Evidence:

Data distribution-based comments:

- These UK students have provided more "continuous" time values for their travel times compared with these NZ students which have been recorded as either 5, 15, 25 or 45 minutes.
- A higher percentage of these 400 UK students sampled travel to school by bus or walking (287 / 400 vs 198 / 400).
- The sample distributions of walking and bus times are positively skewed for these UK students.
- The variability for walking and bus times is greater for these UK students compared to these NZ students, particularly for bus times.

Confidence interval-based comments:

- For UK school students, it's a fairly safe bet that the mean time to travel to school by bus is between 7.9 and 15.2 minutes longer than the mean time to travel to school by walking.
- For NZ school students, it's a fairly safe bet that the mean time to travel to school by bus is between 10.4 and 17.3 minutes longer than the mean time to travel to school by walking.
- Based on these samples we can claim that it takes longer, on average, to travel to school by bus than by walking for both UK and NZ students, as both confidence intervals contain positive values for the estimated differences.

Task Q3(b)(i)
Evidence:

- Overall design: Comparison of two independent groups.
- Units / participants: 136 university students from an environment club.
- Treatment groups: Received daily messages, did not receive daily messages (control group).
- Response variable: Number of days each student walked or cycled to university.
- Random allocation used to allocate treatments to units / participants.

Task Q3(b)(ii)
Evidence:

- The tail proportion is large (0.266), which does not give evidence to support a claim that the daily text messages encouraged the students to cycle or walk to university more often.
- Looking at the data collected, the results look similar both in terms of centre and spread (the difference between the group means is 1.2 days).
- As all participants were tracking each day whether they walked or cycled to university, we could expect to see similar behaviour irrespective of the applied treatment.

Task Q3(b)(iii)
Evidence:

Study:

- Did participants give informed consent?
  It is unethical to carry out an experiment without informed consent of the participants.
- Did participants know the treatments for the experiment?
  Participants may change their behaviour because they know they are in the treatment group; blinding should be used.
- How were participants recruited for the experiment? Had they already expressed interest in cycling or walking more?
  If participants were already keen to increase their walking or cycling to university, this limits the scope of generalisation of a "treatment effect".
- What was the wording of the messages sent to participants?
  The treatment applied was the wording of the message sent to participants, so it is important to know what the exact wording used was.

Data:

- Did all participants track data for every day during the study period? How were missing days managed by the researchers?
  It is important that the data used for each participant is consistently measured, e.g. perhaps the percentage of days tracked where a participant walked or cycled to university might be a better measure.

Participants:

- How many days did each participant typically walk or cycle to university before the study? For each participant, how did this change during the study period?
  If the study wants to conclude that the treatment increased the number of days participants walked or cycled to university, then they should compare each participant's walking and cycling rate before the experiment to those observed during the experiment.
- What were the reasons people didn't walk or cycle before the study?
  The reasons people didn't walk or cycle before the study might be related to practical reasons like needing to drop children off at school or the distance they live from the university.

Note: Accept other questions about study, data or participants, with valid statistical or contextual justification.

2019

QUESTION FOUR

Task Q4(a)

Evidence:

Strengths of the study design:

- Selecting a variety of bus and train routes, based on length (long / short), location (downtown / suburban, wealthier / poorer), times (morning / midday / evening / night).
- Consistency and range with data collection e.g. four-minute periods using established 12 pre-set codes for behaviour.

Challenges with how the data was collected:

- Activity measured in terms of "ever-observed" rather than how long the passenger spent.
- No data collecting on late night trips, due to safety reasons for the two researchers.

Note: Accept other strengths and other challenges based on information provided in the report.

Task Q4(b)(i)

Evidence:

For confidence interval 1: The relevant percentages are 76.5 and 56.6 (a point estimate for the difference of 19.9 percentage points). The respective sample sizes are 353 and 459 (e.g. $\frac{270}{0.765} = 353$).

For confidence interval 2: The relevant percentages are 76.5 and 12.5 (a point estimate for the difference of 64.0 percentage points). The sample size is 353.

Task Q4(b)(ii)

Evidence:

Confidence interval 1 involves estimating the difference between two percentages from independent groups.

Confidence interval 2 involves estimating the difference between two percentages from within the same group.

Task Q4(b)(iii)

Evidence:

- The study was conducted sometime before 2011, which is over eight years ago, and so the behaviour of passengers may have changed and findings would not be applicable to current bus and train passengers.
- Due to the way the data was collected (non-random samples), we would not be able to assume that the behaviour of the passengers observed in this study is representative of the behaviour of all NZ bus and train passengers.
- Wellington may have a higher percentage of commuters travelling by train, so the passengers observed in this study may not be representative of all NZ.

Note: Accept other reservations based on valid statistical or contextual points.

Task Q4(c)

Evidence:

Pr(observed talking | female) / Pr(observed talking | male) = 2.1
125 passengers observed talking
402 women, so 410 men (total passengers observed 812)

Let a be the number of women talking and b be the number of men talking

$$\frac{a}{402} \bigg/ \frac{b}{410} = 2.1$$

$a + b = 125$

$a = 84, b = 41$

Pr(male | observed talking) $= \frac{41}{125} = 0.328$

Note, due to similar group sizes (402 / 410), an incorrect method can generate an almost correct partial answer e.g. 125 / 3.1 = 40 males. The answer needs to be supported by a correct method. [credit may be given for this technique if evidence of student's understanding of the equal proportion of men and women]

QUESTION FIVE

Task Q5(a)(i)
Evidence:

- the number of stops the bus stops at – the more stops, the longer the time
- the number of passengers at each stop – the more passengers who get on, the longer the time to load passengers
- the number of red lights at each of the five intersections – the more red lights, the longer the time
- the amount of traffic on the road – the more traffic, the longer the time
- the speed the bus travels at – the faster the average speed, the shorter the time
- the time of day she leaves, during rush hour – the more traffic, the longer the time
- how full the bus is – the more full when she gets on, the fewer stops made, the shorter the time.

Note: Accept other realistic reasons based on information provided in question or related to prior contextual knowledge.

Task Q5(a)(ii)
Evidence:

- Normal distribution appears to be an appropriate model for Amelia's future bus commute times.
- The distribution of the sample data is unimodal, symmetric, bell shaped, the variable (time) is continuous.
- There are many possible factors for why the times vary and the interactions between these factors are complex; times can be viewed as unpredictable and suitable for modelling by a probability model.

Task Q5(b)(i)
Evidence:

Parameters of triangular distribution appear to be $a = 10$, $b = 30$, $c = 20$

Jacob appears to have added:

- the minimum values of both distributions to get a
- the maximum values of both distributions to get b
- the mean / median of the uniform distribution (3) with the modal value of the triangular distribution (17) to get c.

These parameters are consistent with the simulated data.

Task Q5(b)(ii)
Evidence:

Using a triangular distribution with $a = 10$, $b = 30$, $c = 20$, (height = 0.1)
$P(X > 25) = 0.125$
$P(X > 28) = 0.02$
$P(X > 28 \mid X > 25) = 0.16$

Task Q5(b)(iii)
Evidence:

- Jacob's model is a triangular distribution. However, the data collected over three years does not have a clear triangular shape.
- Due to the amount of data Jacob has collected, the non-triangular shape cannot be explained only by sampling variation.
- Jacob's model assumes independence of the two random variables used, which is invalid as the length of time he waits for the bus will not be independent from the length of the time for the bus commute for many reasons, including traffic conditions, number of buses on the same route, number of passengers, etc. e.g. if traffic is heavy then this could delay buses arriving at Jacob's bus stop and also cause buses to take longer to get to his destination from his stop.

Note: Jacob's model doesn't take into account whether the resulting distribution when combining a uniform and a triangular distribution, assuming independence, is triangular. However, candidates are not expected to combine continuous random variables, nor have knowledge of combinations of non-uniform discrete random variables in terms of resultant distributions, so this is not expected evidence (but could be accepted if given).