3

SUPERVISOR'S USE ONLY

91586

915860

NZQA
NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

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

Level 3 Mathematics and Statistics (Statistics), 2017

91586 Apply probability distributions in solving problems

9.30 a.m. Monday 27 November 2017
Credits: Four

Achievement	Achievement with Merit	Achievement with Excellence
Apply probability distributions in solving problems.	Apply probability distributions, using relational thinking, in solving problems.	Apply probability distributions, using extended abstract thinking, in solving problems.

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.

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 2–8 in the correct order and that none of these pages is blank.

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

Excellence

TOTAL 23

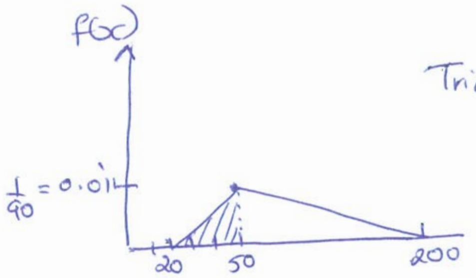
ASSESSOR'S USE ONLY

QUESTION ONE

ASSESSOR'S USE ONLY

(a) The amount of water used when taking a shower can be modelled by a random variable that takes on values between 20 litres and 200 litres. The most likely amount of water used when taking a shower is 50 litres.

(i) Using an appropriate probability distribution model, calculate an estimate for the percentage of showers that use less than 50 litres of water.



Triangular

$h = \frac{2}{180} = \frac{1}{90}$

$P(x<50) = \frac{1}{2} \times 30 \times \frac{1}{90}$

$= \frac{1}{6}$ Percentage $= 16.67\%$

(ii) Using an appropriate probability distribution model, calculate an estimate for the percentage of showers that use more than 40 litres of water.

$P(x>40) = 1 - P(x<40)$

$P(20<x<40) = \frac{1}{2} \times 20 \times \frac{1}{135}$

$= \frac{2}{27}$

$P(x>40) = 1 - \frac{2}{27}$

$= \frac{25}{27}$

Percentage $= 92.59\%$ use more than 40 litres of water

(b) Car drivers can use various mobile phone GPS navigation apps to get an estimate of the time it will take to travel to a destination. A study was carried out to investigate how accurate the travel time estimates were from one particular GPS navigation app. For each trip in the study, the estimated travel time was compared to the actual travel time, and the absolute difference calculated (see the table below).

Trip	Estimated travel time	Actual travel time	Absolute difference between estimated travel time and actual travel time
1	10.4 minutes	11.3 minutes	0.9 minutes
2	6.5 minutes	5.2 minutes	1.3 minutes
3	3.9 minutes	3.9 minutes	0 minutes
…	…	…	…

ASSESSOR'S USE ONLY

(i) Suppose 15% of trips made during the study were classified as "not accurate" using the absolute difference.

If ten trips from the study were chosen at random, using an appropriate model, calculate the probability that at most four of the trips were classified as "not accurate".

$\pi = 0.15$ $n = 10$ $x \leq 4$ Binomial

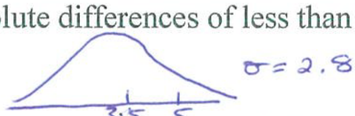
$P(x \leq 4) = 0.9901$

(ii) Justify the use of the probability distribution for your answer in part (i).

The probability distribution is Binomial because

- there are only two possible outcomes (either GPS was accurate or not accurate)
- There is a fixed number of trials (10 trips)
- The probability of a trip being 'not accurate' stays constant at 0.15
- ~~if~~ a trip is classified as 'accurate' or 'not accurate' is independent on the classification of another trip being 'accurate' or 'not accurate'

(iii) The mean absolute difference between the estimated travel time and the actual travel time for trips in this study was 3.5 minutes, with a standard deviation of 2.8 minutes. The study also found that 87% of trips had absolute differences of less than five minutes.



Discuss TWO reasons why it would be inappropriate to use a normal distribution to model the absolute differences between the estimated travel time and the actual travel time for trips.

1. It would be inappropriate to use normal distribution because there are no upper and lower bounds in normal distribution, however, in this context it would be impossible to have an absolute difference of below 0 minutes, and it would be unlikely to have extremely high absolute differences beyond a certain point (eg 24 hours) as people would simply give see extra paper // see p8

2. If you use a normal distribution to calculate the probability of trips having absolute differences of less than five minutes, then $P(x<5) = 0.7039 = 70\% \neq 87\%$ stated above. The real probability of trips having absolute differences of less than five minutes has a big difference with the result calculated by normal distribution so it would be inappropriate.

E7

ASSESSOR'S USE ONLY

QUESTION TWO

(a) The table below shows the probability distribution of the random variable X.

x	0	1	2	3	4
$P(X=x)$	0.11	0.21	0.24	0.25	0.19

(i) $E(X) = 2.2$.

Calculate $VAR(X)$.

$Var(X) = (1\times0.21 + 4\times0.24 + 9\times0.25 + 16\times0.19) - 2.2^2$

$= 6.46 - 4.84$

$= 1.62$.

(ii) The random variable Y has $VAR(Y) = 1.5376$.

$VAR(X+Y) = 5.5696$.

Are X and Y independent?

Support your answer with appropriate statistical statements.

$Var(X) + Var(Y) = 1.62 + 1.5376$

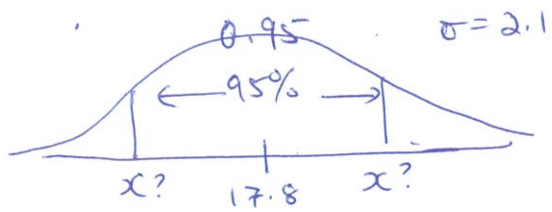
$= 3.1576$

$\neq 5.5696$

Because $Var(X) + Var(Y) \neq Var(X+Y)$, X & Y are NOT independent.

(b) The average temperature in a New Zealand living room on a winter evening can be modelled by a normal distribution, with mean 17.8°C and standard deviation 2.1°C.

(i) Using this model, between what two values would you expect the middle 95% of average temperatures for New Zealand living rooms on a winter evening to be?



Left $x = 13.684$

Right $x = 21.9159$

values of 13.7° and 21.9° C

(ii) Discuss ONE factor that should be considered when modelling the average temperature in a New Zealand living room on a winter evening.

ASSESSOR'S USE ONLY

The location of the livingroom should be considered as the average temperature may vary by region (for example South Island tends to be colder than North Island in winter)

(iii) Suppose that five New Zealand houses were selected at random, and it was found that the average temperature of the living room on a winter evening was below 16°C for four of these houses.

Would finding four or more houses out of five with an average temperature of the living room below 16°C be unlikely under the probability distribution model described above?

Support your answer with a calculation.

Using normal distribution: $P(x<16) = 0.1957$

Then using binomial: $n=5$, $\pi = 0.1957$ $x \geq 4$

$P(X \geq 4) = 1 - P(X \leq 3)$

$= 1 - 0.9938$

$= 0.006186$

So it can be seen that it is highly unlikely that four or more houses out of five with average temperature of living room below 16°C can be selected by chance alone (the probability of this occurring is very little at only 0.006186, which is well below 5%) So it is very unlikely to happen using the model

E8

ASSESSOR'S USE ONLY

QUESTION THREE

A study collected data on water use within New Zealand homes for the purpose of assisting councils, government agencies, and water suppliers to introduce water efficiency measures.

Prior to the study, it was estimated that each person in New Zealand flushes the toilet on average 4.7 times per 24-hour period.

(a) (i) Using a Poisson distribution model, calculate an estimate for the probability that a person flushes the toilet less than five times in any 24-hour period.

$\lambda = 4.7$

$P(X \leq 4) = 0.4946$

(ii) Give ONE reason why it may not be appropriate to use a Poisson distribution to model the number of toilet flushes for **any 4-hour period**.

Because the probability of a person flushing the toilet is not necessarily proportional to the size of the time interval, people tend to go to the toilet more or less times depending on the time of day (for example people may not go frequently to the toilet in the middle of the night)

(b) Data was collected on the number of times each person flushed the toilet during a 24-hour period. The data from 200 people from across 84 homes in the study is summarised in the table below.

Number of toilet flushes during 24-hour period	0	1	2	3	4	5	6	7	8	9	10
Proportion	0	0.01	0.1	0.17	0.26	0.14	0.12	0.11	0.05	0.01	0.03

(i) Calculate the mean number of toilet flushes made per 24-hour period for people in this study.

$E(X) = 1\times0.01 + 2\times0.1 + 3\times0.17 + 4\times0.26 + 5\times0.14 + 6\times0.12$

$+ 7\times0.11 + 8\times0.05 + 9\times0.01 + 10\times0.03$

$= 4.74$

(ii) Did most people in this study flush the toilet at least four times during a 24-hour period?

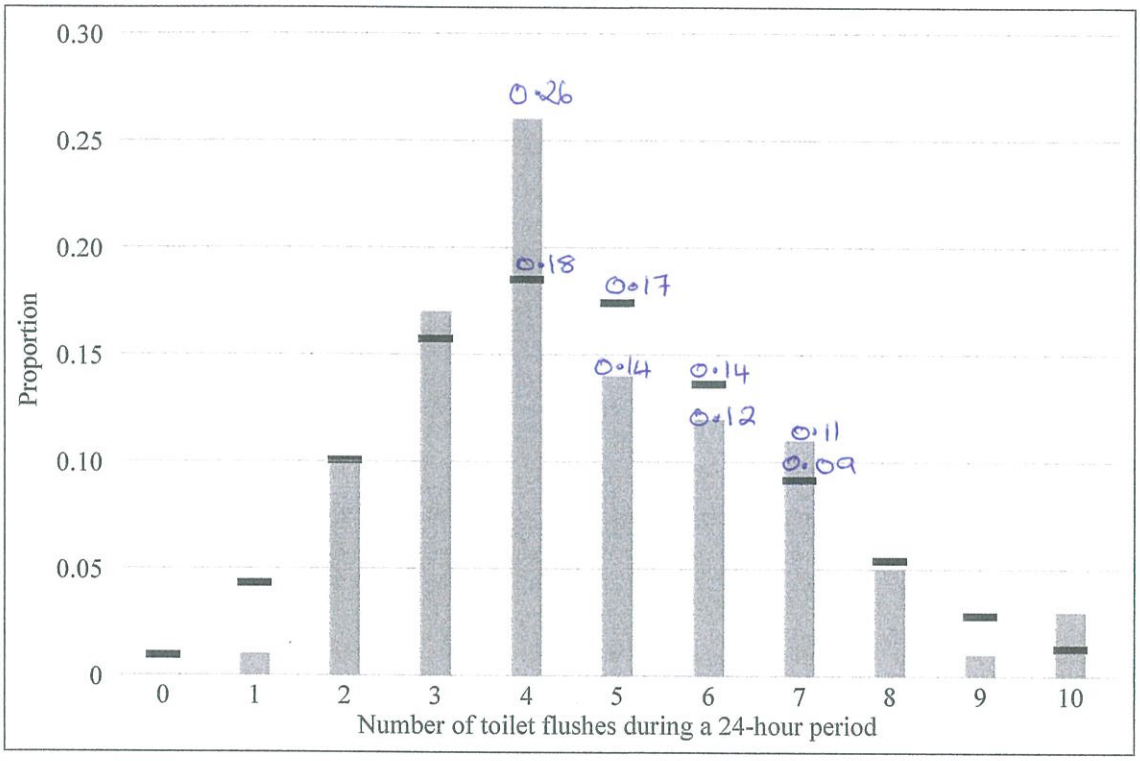
Support your answer with a calculation.

$P(X \leq 3) = 0 + 0.01 + 0.1 + 0.17 = 0.28$

$P(X \geq 4) = 1 - 0.28 = 0.72$

Yes, most people (around 72% which is over 50%) did flush the toilet at least four times in a 24 hour period.

(iii) The graph below shows the experimental distribution (shaded bars) and a Poisson distribution with λ = 4.7 (the model distribution shown in black).



Discuss TWO reasons why a Poisson distribution with λ = 4.7 may not be a good model for the number of toilet flushes for any 24-hour period.

1. There are multiple probabilities where the Poisson model does not match the probability from the study - for example from the study the probability (proportion) of someone flushing 4 times in a 24 hour period is 0.26 but the theoretical is 0.18 and for 5 flushes the experimental

See extra paper Seen

2. The Poisson distribution also does not really fit the shape of the distribution as it underestimates the proportion for $P(X \leq 4)$. The experimental probability is 0.54 while the probability modelled by Poisson is only 0.49, so its roughly 0.05 off. Therefore

see extra paper Seen

ASSESSOR'S USE ONLY

E8

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

QUESTION NUMBER

ASSESSOR'S USE ONLY

1 b(iii) continued

up driving to their destination and it would not be physically feasible ✓ seen

3 b iii.

1. continued is 0.14 while Poisson estimates it at 0.17. This suggests Poisson is not a good model. ✓

2. it also over estimates the probability of flushes greater than or equal to five. The Poisson Distribution estimates that $P(X \geq 5) = 0.5054$ but the study proportion is $P(X \geq 5) = 0.46$ which is much lower than the Poisson. Suggesting the Poisson is not a good model.

Also in a Poisson model the mean = variance. In this study of toilet flushes the mean is 4.74 flushes in 24hr period and the variance is 3.85 flushes. So they are not equal and the Poisson distribution may not be a good model ✓✓ seen

91586

Subject:	Mathematics & Statistics	Standard:	As91586	Total score:	23

Q	Grade score	Annotation
1	E7	(b) (iii) E7 is awarded rather than E8 as Reason 1's explanation of why the absolute differences could not be below zero needs to be explained more in terms of absolute differences rather than trips. Reason 2 is a good explanation of how the probabilities do not match.
2	E8	(b) (iii) candidate has correctly found $P(X<16)$ using normal distribution and then used that in a Binomial calculation for $P(X\geq4)$. They have identified they are using Binomial and answered the question.
3	E8	(b) (iii) The candidate has provided three reasons even though 2 only have been asked for. All three reasons provide evidence for the E8 grade. Reason 1 A comparison of the experimental and Possion models as shown on the graph with numerical evidence to support statements Reason 2: A good comparison of the models in terms of probabilities. Reason 3 : A good comparison of the variance and mean of the experimental data.