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

3

91586

915860

SUPERVISOR'S USE ONLY

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.

Merit

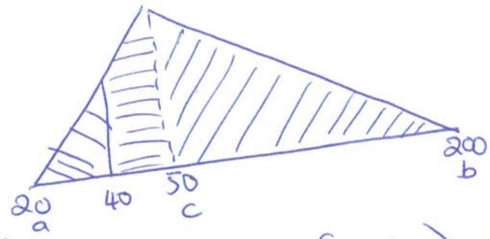
TOTAL 16

ASSESSOR'S USE ONLY

QUESTION ONE

(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.



$h = \frac{2}{b-a}$

$= \frac{20}{200-20}$

$= \frac{1}{90}$

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

$= \frac{1}{6}$

$= 0.167$ (3dp)

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

$p(x>50) = \frac{1}{2} \times 150 \times \frac{1}{90} = 0.833$ (3dp)

$P(40<x<50) = (\frac{2}{3} \times \frac{1}{90} \times \frac{1}{2} \times 20) = 0.074$ (3dp) $\quad 0.167 - 0.074 = 0.093$ (3dp)

$p(x>40) = 0.833 + 0.093 = 0.926$ (3dp)

(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
…	…	…	…

(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".

$P(x \leq 4) = 0.9901$ (4sf)

Using Binomial

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

I used binomial because at most four trips were classified as 'not accurate', as there is a set number of trials with having ten trips. There is a constant probability of success with 15% of trips classified as 'not accurate'. Also there is only two possible outcomes: the trips are either 'accurate' or 'not accurate' and it is assumed that the trips are independent. A

(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. A normal distribution model allows a small proportion of trials to go less than zero and up to infinity. However, this would not be accurate to model the absolute differences as the travel times are bounded above and below (cannot travel in negative times and absolute difference won't go below 0 so can't have negative difference and will not have an infinite difference)

2. Also for normal distribution it would not be appropriate as the absolute differences do not represent a bell shaped curve, with having 87% of trips having differences of less than five minutes and the mean being 3.5 with a standard deviation of 2.8 minutes

A

ASSESSOR'S USE ONLY

M6

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).

$E(X) = (0 \times 0.11) + (1 \times 0.21) + (2 \times 0.24) + (3 \times 0.25) + (4 \times 0.19)$

$Var(X) = E(X^2) - [E(X)]^2$

$Var(X) = 0.89$

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

VAR($X + Y$) = 5.5696.

Are X and Y independent?

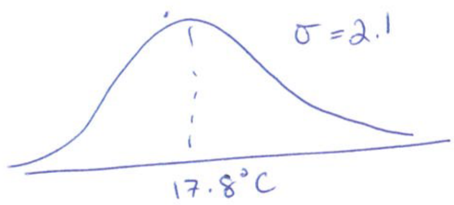
Support your answer with appropriate statistical statements.

$0.89 \times 1.5376 \neq 5.5696$

so X and 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?



$Z = \frac{X - \mu}{\sigma}$ $\quad 1.96 = \frac{X - 17.8}{2.1}$ $\quad 21.916 = x$.

$21.916 - 17.8 = 4.116$

$17.8 + 4.116 = 21.916°C$

$17.8 - 4.116 = 13.684°C$

expect between 13.684°C and 21.916°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

Type of insulation the house has may be different to others. This would make the temperatures in the living rooms of the houses different.

(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.

P (temp < 16°C) = 0.1957 (4dp)

P (four houses' temp < 16°C) = 0.0015 (4dp)

Yes, finding four or more households out of five with an average temperature of the living room below 16°C would be unlikely

M5

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.

$P(x<5)=0.496$ (4dp)

(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**.

The number of toilet flushes could be too small. Poisson is for rare events There might be 0 flushes in 4 hours It needs to be over a long period of time

(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.

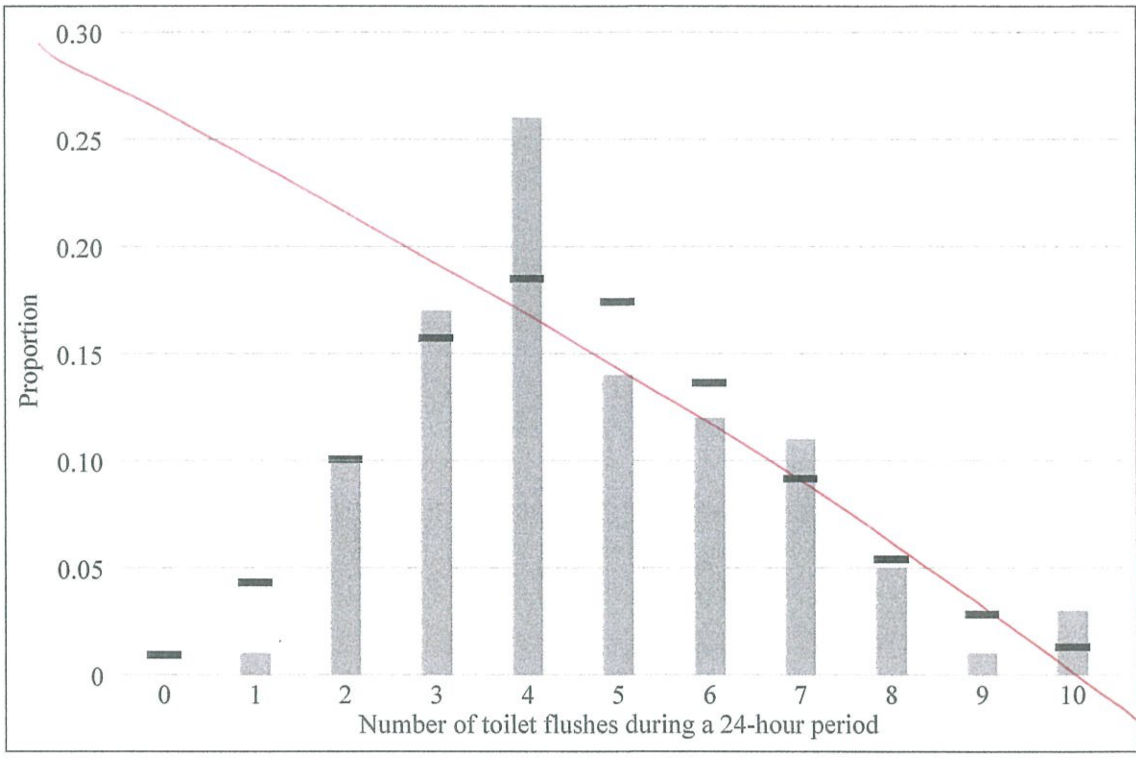
mean $=(0\times0)+(1\times0.01)+(2\times0.1)+(3\times0.17)+(4\times0.26)+(5\times0.14)$
$+(6\times0.12)+(7\times0.11)+(8\times0.05)+(9\times0.01)+(10\times0.03)$
$=4.74$ flushes per 24 hours

(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. $P(X \geq 4) = 0.72$.

Yes, 72% of the people in this study (144 people out of the 200) flushed the toilet at least 4 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. The mean and variance are different. In a Poisson model they should be equal but in this case they aren't.

2. A condition of Poisson is that events are independent, and the probability of one event doesn't effect the other. However, one person's flush could be effected by another person's flush so they aren't independent.

M5

Subject:	Mathematics & Statistics	Standard:	AS91586	Total score:	16

Q	Grade score	Annotation
1	M6	(b) (ii) the context for the condition of independence is not sufficient. However, only 2 conditions in context are required for r. (b) (iii) Reason 2 has not provided evidence so it is n. If the candidate had provide a calculation using the Normal distribution for $P(x<5)$ and compared that probability to the stated 87% of the model they may have scored an r . Reason 1 is sufficient for a Merit grade and with Reason 2 as detailed above would score E7. To gain E8 the candidate would need to expand on why absolute differences can't go below zero.
2	M5	(a) (ii) The candidate has not used the correct formula from their formula sheet. (b) (ii) The candidate has identified the factor and explained the effect it would have on the temperature of the living room. (b) (iii) The candidate has correctly calculated the probability the temperature is below 16 degrees using the Normal Distribution but has not used this probability correctly in a Binomial calculation for $P(x\leq3)$. Only $P(x<16)$ is correct so u.
3	M5	(a) (i) and (ii) are both necessary for r. (i) is correct so the grade is u. To reach r, the candidate needed to discuss in (ii) the idea that the rate a person would flush the toilet at will change depending on which 4 hour block they are in i.e. the lambda would not remain constant but change depending on what 4 hour block of time they were in. (b) (i) and (ii) are both necessary for r. Both parts are correct. In (ii) the candidate has used the table on page 6 to calculate correct probability and used that to answer the question. The use of incorrect inequality sign > instead of ≥ has been ignored in marking. (b) (iii) no evidence has been provided for either reason so n. Reason 1: to reach r the mean and variance for the experimental, and Poisson would have had to be stated and compared. Reason 2: The candidate has identified a criterion for the Poisson but has misunderstood the context, which is the number of flushes per person. They would need argue something like this: *"someone*

		may flush twice or may clean the toilet flushing it multiple times therefore the Poisson may not be an appropriate model."