91586

915860

3

SUPERVISOR'S USE ONLY

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.

Achievement

TOTAL 11

ASSESSOR'S USE ONLY

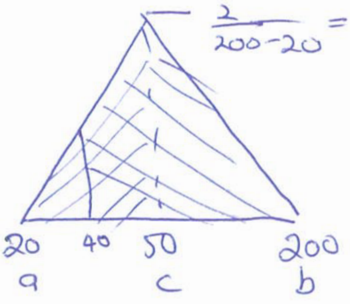
QUESTION ONE

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.

$\frac{2}{200-20} = \frac{1}{90}$.



$P = \frac{1}{2} \times b \times h$

$= \frac{1}{2} \times 30 \times \frac{1}{90}$

$= \frac{1}{6}$ so 16.67% of showers use less than 50 litres

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

Between 20 & 40 $f(x) = \frac{2(40-20)}{(200-20)(50-20)} = 0.0074$

Area $= \frac{1}{2} \times 180 \times 0.0074 = 0.666$

so 66.6% of showers

(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".

$\pi = 15\%$ $n = 10$ $x < 4$

$P(x < 4) = 0.99$

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

Binomial used because
- only 10 trips
- only two possible outcomes
- probability is constant
- events are independent

(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. Because 87% of trips had differences less than five minutes, so most values would be at one end of the distribution

2. Time cannot stretch on infinety in either direction

A3

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 $VAR(X)$.

$Var(X) = 1.62$

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

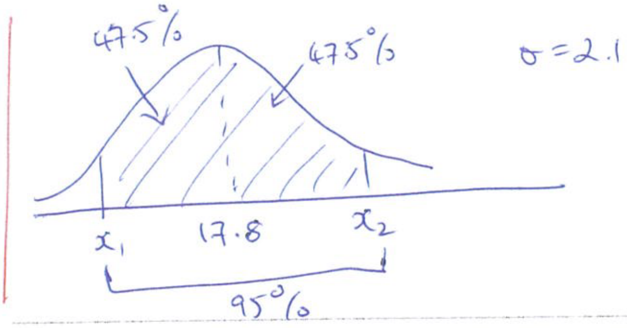
$VAR(X + Y) = 5.5696$.

Are X and Y independent?

Support your answer with appropriate statistical statements.

(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?



$x_1 = 13.68$

$x_2 = 21.92$

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

A factor is whether people have heaters

(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.

Binomial $x \geq 4$ $p = 0.8 = \frac{4}{5}$

$P(x \leq 3) = 0.2672$

$P(x \geq 4) = 1 - 0.2672$

$= 0.7372$

ASSESSOR'S USE ONLY

A4

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.

$P(x < 5) = 0.4946$

(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**.

As Poisson is a discrete variable (toilet flushes) inside continuous space (time) toilet flushes should be discrete numbers $4.7 \div 6 = 0.783$ so 0.783 flushes in 4 hours This average is less than one flush so you can't use Poisson

(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.

$E(x) = 4.74 / 10$
$= 0.474$

(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.

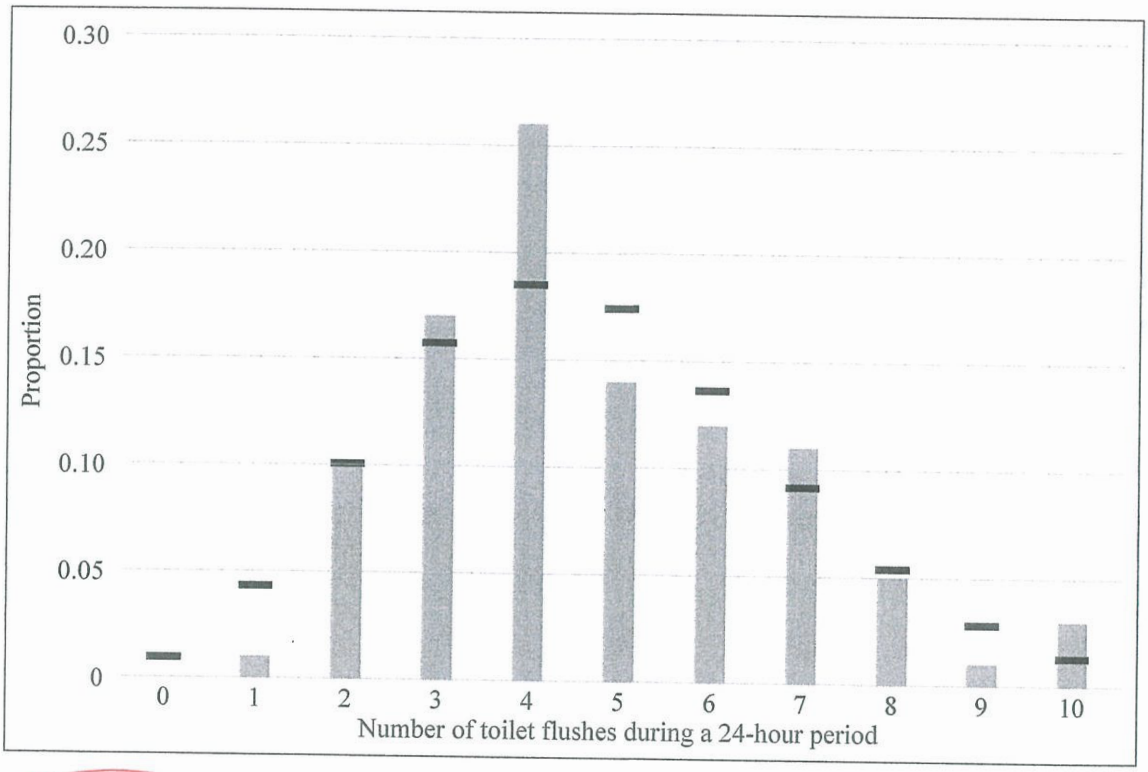
$P(X \geq 4) = 0.26 + 0.14 + 0.12 + 0.11 + 0.05 + 0.01 + 0.03$

$= 0.72$

So yes most people (72%) did.

ASSESSOR'S USE ONLY

(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 shaded bars and the black lines don't match well. The only place the Poisson and Experimental match is at 2 flushes. The Poisson model gives larger proportions at 1, 5, 6, 8 and 9 flushes.

2. Poisson Distribution is not a good model as you can flush two toilets at the same time (simultaneously)

A4

Subject:	Mathematics & Statistics	Standard:	91586	Total score:	11

Q	Grade score	Annotation
1	A3	(a) (ii) The shape the candidate is calculating is not a triangle so the triangular distribution is not appropriate. To reach r, the candidate needed to find P(x<40) and subtract it from 1, to do this the base of the triangle should be 20 not 180. (b) Both part (i) and (ii) are necessary for r. In (ii) the conditions of the binomial are not given in context. To reach r at least 2 need to be in context e.g. ‘there are only two possible outcomes: the estimate is accurate or not accurate.’ (b) (iii) evidence for each reason has not been provided. The candidate has also not linked their statements to the normal distribution. To reach u a link to the normal distribution is required e.g. ‘*Normal distribution has no upper or lower limits so they would not be able to travel for endless hours or travel for a negative amount of time.*‘ To reach r one limitation of the normal clearly linked to the context of absolute differences is needed e.g. ‘*Normal distribution has no lower and upper bound. “* For t, two limitations clearly linked and in context are required.
2	A4	(b) Both (i) and (ii) are needed for r. (ii) the factor has been identified but there is no explanation of how a heater will affect or change the living room temperature. (b) (iii) the student has not used the Normal Distribution to calculate the probability of less than 16 degrees and then used this as their probability in the Binomial calculation. This is necessary for r and t
3	A4	(a) Both part (i) and (ii) are necessary for r. (i) to reach r the candidate needed to discuss the idea that the rate people flush at will change depending on which 4 hour block they are in (b) both part (i) and (ii) are necessary for r. Part (ii) has been correctly answered. Part (i) is incorrect, as the candidate has divided the correct answer by 10. b) (iii) Reason 1. For u, the candidate has identified that the Poisson and Experimental distributions don’t match and places where the Poisson is overestimating the proportion of flushes.

		For r, the candidate needed to identify the specific numerical value(s) of the Poisson and the experimental distributions for at least one amount of flushes e.g. ‘ *the Poisson estimates that the proportion of flushes for 4 flushes per 24hour period is 0.18 while the experimental study had 0.26. This is a large difference.’* Reason 2: This is not a valid reason. The context is about the number of times one person flushes in 24 hours, not the number of toilets that can be flushed in the context of this question it is not possible for the same person to flush two toilets simultaneously.