Assessment Schedule – 2022

Mathematics and Statistics (Statistics): Apply probability distributions in solving problems (91586)

Evidence Statement

Q	Expected coverage	Achievement (u)	Merit (r)	Excellence (t)
ONE (a)(i)	Binomial distribution, $n = 10$, p = 0.35 $P(X < 3)$ $= P(X \leq 2)$ = 0.2616	• Probability correctly calculated.	• Probability correctly calculated.	
(ii)	The binomial distribution was chosen because: *Possible Assumptions* • There is a constant probability of success. For each customer entering the store the probability of purchasing is the same p=0.35. • Each trial is independent – a customer's choice to purchase something or not is not affected by the purchasing behaviour of other customers.	OR TWO valid assumptions for choice of distribution.	AND Two valid assumptions for choice of distribution in context.	
(iii)	For one store: Binomial $n = 10$, $p = 0.35$ $P(X \geq 3) = 0.7384$ For four stores: $0.7384^4 = 0.2973$	• Probability $P(X \geq 3)$ correct for one store. OR Assumption in context.	• Probability correct for four stores. AND Correct assumption identified in context.	• Probability correct for four stores. AND Correct assumption identified in context.

(iv)	Possible assumptions **Combining Probabilities:** The assumption that the event "3 or more out of the 10 customers make a purchase" at one store is independent of the event "3 or more out of the 10 customers make a purchase" at each other store is likely to be valid. This is because any influences of buying behaviour in one store (eg a sale) is not likely to impact buying behaviour in another store as all stores are part of the same clothing retail company. If one store has a sale for example, it is unlikely that this will influence the probability that "3 or more out of the 10 customers make a purchase" in one of the other stores. **Binomial probability within one store:** • Independence When calculating the binomial probability, it is assumed that each trial is independent - a customer's choice to purchase or not is not affected by the purchasing behaviour of other customers. This may not be correct as if a group of friends are shopping, they often encourage each other to purchase items therefore influencing the likelihood of purchase. • Constant Probability There is a constant probability of success - For each customer entering the store the probability of purchasing is the same $p = 0.35$. This may not be correct as a customer might be spending birthday money so more likely to buy something, whereas another customer might just be browsing to fill in time and very unlikely to buy something. **Same binomial probability model for ALL FOUR stores** We are assuming that the same binomial probability distribution model will hold for all four stores so each store will have a constant probability of purchase of 0.35. This may not be correct as some stores in higher socio-economic areas will have more affluent customers who might be more likely to purchase an item, whereas stores in lower socio-economic areas will have less well-off customers who might be less likely to purchase an item. *Accept other contextual reasons for above assumption being valid or not.*			AND Correct discussion of the validity of the assumption in context.

(b)	For click and collect orders Uniform distribution $a = 24$, $b = 144$ $P(X > 120) = \frac{24}{120} = 0.2$ For home delivery orders Normal distribution $\mu = 96$, $\sigma = 38.4$ $P(X > 120) = 0.26599$ Estimate of probability that both friends have to wait for more than 120 hours $= 0.2 \times 0.266 = 0.0532$ *Possible Assumptions:* • In the uniform distribution, we are assuming that there is equal probability of a click and collect order being ready at any time between 24-144 hours. This may not be the case as there could be times when the business was more efficient and so the order is ready earlier; at other times it could be slower and take longer for the order to be ready. • In the uniform distribution, we are assuming that there is equal probability of a click and collect order being ready at any time between 24-144 hours. This may not be the case as it is unlikely that in New Zealand businesses are open 24 hours a day. • We are also assuming that the waiting time for receiving orders from click and collect is independent of the waiting time for the home delivery which may not be the case as the friends are ordering from the same store so the same workers could be doing both jobs and may finish one job first eg home delivery before completing the click and collect orders. *Accept other reasonable assumptions related to either model and context.*	• Correct probability $P(X > 120)$ for one method of ordering.	• Correct probability estimate for both friends waiting for more than 120 hours. OR One probability correct for one method and one assumption discussed in context.	• Correct probability estimate for both friends waiting for more than 120 hours. AND Discussion of at least one assumption and its validity.

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of u	2 of u	3 of u	1 of r	2 of r	1 of t	2 of t

Q	Expected coverage	Achievement (u)	Merit (r)	Excellence (t)
TWO (a)(i)	Poisson distribution, $\lambda = 2.8$ $P(X < 4) = 0.6919$	• Probability correct.	•	•
(ii)	Possible reasons the Poisson distribution may not be appropriate: • The Poisson distribution assumes **independence**: that the occurrence of items being purchased does not affect the probability that further items will be purchased in the 20-minute interval. This may not be the case if two friends shop together, one friend purchasing items might make it more likely that the other friend also purchases items. • The Poisson distribution assumes that the **average rate at which items are purchased is constant/proportional** in the given unit of time. This might not be the case as the rate may not apply at all times of the year. The number of items purchased per 20 minutes might be higher at certain times e.g., Christmas shopping, or at sale time than at other times. • The Poisson distribution assumes that when items are purchased is **random, and therefore unpredictable**. This may not be the case as some 20-minute time intervals will be when the shop is closed and so customers are unable to make an in-store purchase so the 20-minute time intervals when purchases are made will be predictable (limited to the store hours). • The Poisson distribution assumes that items being purchased cannot occur at the same time **(not simultaneous)**. This might not be the case because items could be purchased by two or more customers at essentially the same time (e.g., when the store has a number of check outs).	• One correct reason identified in context.	• TWO correct reasons identified in context. OR ONE correct reason linked strongly to the context with reasoning as to why the distribution is not appropriate.	• TWO correct reasons identified in context. AND Reasoning as to why the distribution is not appropriate for both.
(iii)	$P(X = 0) = 0.01$ and calculating λ: $e^{-\lambda} = 0.01$ $\lambda = -\ln 0.01$ $\lambda = 4.6$	• Setting up relevant equation $e^{-\lambda} = 0.01$	• Correct calculation for lambda (λ).	

(b)(i) (ii)	For Supplier A Mean = 0.4 Standard deviation = 0.7348 For Supplier B Mean number of defective items = 0.49 Standard deviation = 0.7937 Supplier B has a greater on average number of defective items per box (0.49) compared to Supplier A (0.4). The variation in the number of defective items per box is also greater for Supplier B (standard deviation for Supplier B 0.7937 is greater than that for Supplier A, 0.7348) which means that they are less consistent. Supplier A better fulfils the requirements as it has a lower average number of defects per box and is more consistent in the number of defective items per box.	• Mean and standard deviation correct for either Supplier A or Supplier B. OR Both means correct and compared. OR Both SD correct and compared.	• Mean and standard deviation correct for both Supplier A and Supplier B. OR Calculation and comparison in context of means or standard deviations and Conclusion as to which supplier best fulfills the requirements.	• Mean and standard deviation correct for both Supplier A and Supplier B. AND Calculation and comparison in context of means and standard deviations. AND Conclusion as to which supplier best fulfills the requirements.

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of u	2 of u	3 of u	1 of r	2 of r	1 of t	2 of t

Q	Expected coverage	Achievement (u)	Merit (r)	Excellence (t)
THREE (a)	33 out of the 100 transactions (0.33 or 33%) involve a spend of more than \$150. This is more than 25%; therefore the claim is supported.	• Proportion correctly calculated. AND A statement that claim is correct.		
(b)	Normal distribution: $\mu = 127.75$, $\sigma = 46.22$ $P(X > 150) = 0.3151$	• Correct probability.		
(c)	Triangular distribution: $a = 0$ (minimum spend) $b = 250$ (maximum spend) $c = 130$ (most common spend)	• Correct choice of distribution with parameters stated.		
(d)	$P(X > 130)$ Height at $X = 130$: $h = \frac{2}{250-0} = \frac{1}{125} = 0.008$ $P(X > 130) = \frac{1}{2}(250-130) \times \frac{1}{125}$ $= \frac{12}{25} = 0.48$ Height at $X = 150$: $h = \frac{2(250-130)}{(250-0)(250-130)} = \frac{1}{150}$ $(= 0.0067)$ $P(X > 150) = \frac{1}{2}(250-150) \times \frac{1}{150}$ $= \frac{1}{3} = 0.3333$ $P(X > 150 \mid X > 130 = \frac{\frac{1}{3}}{\frac{12}{25}}$ $= \frac{25}{36} = 0.6944$	• Probability correctly calculated P(X > 130). OR P(X > 150) correctly calculated.	• Probability correctly calculated P(X > 130). AND P(X > 150) correctly calculated.	• Conditional probability correctly calculated.

(e)	Observed probability: $P(X > 150) = \frac{33}{100} = 0.33$ Using a normal model: $P(X > 150) = 0.3151$ Using a triangular model: $P(X > 150) = \frac{1}{3}$ Comparing Models: $\frac{1}{3}$ is closer to 0.33 than 0.3151 The estimate of the probability of a customer spending more than $150 is closer to the observed probability when using a triangular model compared to using a normal model.	• Comparison of observed probability and normal model. OR Comparison of observed probability and triangular model.	• Comparison of observed probability with both models probability estimates. OR Comparison of - Observed probability and normal model and a conclusion referring to the features of the normal distribution and why it is appropriate to the context. OR Observed probability and triangular model and a conclusion referring to the features of the triangular distribution and why it is appropriate to the context.	• Comparison of observed probability with both models probability estimates. AND At least one reason given, with reference to features of the distributions and why their chosen distribution is more appropriate to the context.
(f)	*Recommends a model based on discussions of features of the model(s) and why they are the more appropriate model given the context* **Triangular distribution** The triangular distribution has fixed endpoints –it gives no probability of spending less than $0 or more than $250. This is a better fit to the observed data. It also allows for a non-symmetrical distribution (mean is not equal to mode) of amount spent unlike the normal model that should be symmetrical. The observed data is slightly skewed to the left meaning there are more customers who spend more than $125 than there are who spend less than $125 so the triangular is a better fit to this data. **Normal distribution** • The normal distribution has no fixed endpoints and gives some probability of spending less than $0 and some probability of spending more than $250 which is not possible in this context, so this model is not as appropriate to this context. • The normal distribution assumes a symmetrical distribution of amount spent by customers (50% of the customers should spend less than the mean of $127.75 and 50% more) but the observed data shows that more than 50% of the customers spend slightly more than this so this model is not as appropriate. • In the normal distribution the mean = median = mode. In the observed distribution the mean is $127.75, and the mode is $130 which aren't the same so the triangular distribution is the more appropriate model.			

Field Code Changed

Field Code Changed

Field Code Changed

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of u	2 of u	3 of u	1 of r	2 of r	1 of t	2 of t

Cut Scores

Not Achieved	Achievement	Achievement with Merit	Achievement with Excellence
0 – 6	7 – 12	13 – 18	19 – 24