Assessment Schedule – 2019

Mathematics and Statistics (Statistics): Apply probability concepts in solving problems (91585)

Evidence Statement

Q	Expected Coverage	Achievement (u)	Merit (r)	Excellence (t)
ONE (a)	(see table below)			
(i)	P(male | middle) $= \frac{73}{138}$ = 0.529	Correct conditional probability calculated.		
(ii)	P(male | junior) $= \frac{49}{80} = 0.6125$ P(male | senior) $= \frac{72}{125} = 0.576$ It is **more likely** for a student to be male if they are in Junior School [compared to if they are in Senior School]. [or can use Ratio = 1.06]	**At least one** correct conditional probability calculated.	Both conditional probabilities calculated. AND reach correct conclusion.	
(iii)	P(male) $= \frac{194}{343} = 0.5656$	Correct proportion calculated.		

Gender / School	Junior	Middle	Senior	Totals
Female	31	65	53	149
Male	49	73	72	194
Totals	80	138	125	343

(iv)	**Two reasons** why care should be taken when using this data, for example: Data is correct as at 1 July 2018. Predictions made based on this data may not be correct for years after 2018 as proportions of students of each gender may change over time. Data given is only for the one school. Predictions made based on this data may not be correct for all schools as proportions of students of each gender may vary from school to school. (The concept of a only a small sample is insufficient.) Parents of children within the area may choose to attend different schools, and this choice may be gender dependent. Predictions based on this data for the future may be incorrect if the popularity of different schools for different genders changes over time. The proportion of males may change over time, as suggested by the proportion of males in the Junior school being 61%, which is higher than the 57% for the whole school. Accept other valid possible reasons. Eg Accept discussions of non-binary gender if written in context of the effect on the prediction.		One reason identified and clearly explained. OR two reasons identified without sufficient detail.	Two reasons identified with implication about impact on proportion of males.
(b)	A simulation would allow the principal to see the **variation** in (or distribution of) the number or proportion of males in samples of size 343, based on the assumption that the proportion of males is 0.5. The principal could then **compare** what was observed (194 males or a probability of 0.5656) to this simulated distribution to **consider** the likelihood of the observed result (or a higher proportion) happening, assuming the proportion of males in the region is 50%.		Discussion of how a simulation would allow the principal to see that there is variability associated with estimates of gender proportions **without** numerical support given.	A clear discussion of how a simulation would allow the principal to see that they need to take into account sampling variation to make a decision on this region's gender proportions **with** numerical support.

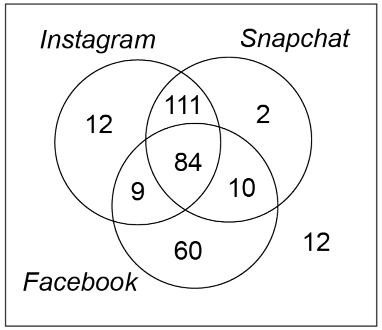
NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Reasonable start / attempt at one part of the question.	1 of u	2 of u	3 of u	1 of r	2 of r	1 of t	2 of t

Q	Expected Coverage	Achievement (u)	Merit (r)	Excellence (t)
TWO (a)(i)	$P(E \mid S) = \frac{91}{123} = 0.740$ $P(E) = \frac{91+68}{201} = \frac{159}{201} = 0.791$ **Different** answers suggest **non**-independence [of events *E* and *S*].	$P(E \mid S)$ or $P(E)$ correctly calculated.	Both probabilities correctly calculated AND explanation of non-independence of events.	
(ii)	Risk of receiving Not Achieved for snorers: $P(\text{not Ach} \mid \text{snorer}) = \frac{32}{123} = 0.260$ Risk of receiving Not Achieved for non-snorers: $P(\text{not Ach} \mid \text{non-snorer}) = \frac{10}{78} = 0.128$ Risk of receiving Not Achieved is 2.03 times as likely (or about double) for snorers compared to non-snorers.	Risk of Not Achieved for both groups is calculated.	Correct **ratio** calculated. AND correct conclusion as to effect of snoring on risk of Not Achieved in the exam.	
(iii)	P(don't snore and complete homework tasks) $= \frac{78}{78+123} \times 0.65 = \frac{78}{201} \times 0.65 = 0.388 \times 0.65 = 0.252$ P(all three students don't snore and complete homework tasks) = $0.252^3 = 0.016$ Assumptions: • Assumption made that the snoring and/or homework habits of the three students are independent. • Assumption made that the number of students is sufficiently large that sampling without replacement is not required so the probability remains relatively constant. • Assumption made that the probability of being a snorer is the same in this larger group as in the group of Year 13 students given at the beginning of the question. Accept other valid assumptions.	Probability correctly calculated for one student.	Probability correctly calculated for combined probability (three students).	Probability correctly calculated for combined probability (three students). AND One assumption stated clearly in context.

(b)(i)	Expected number gaining Achievement (or better) $= \frac{91}{123} \times 127 + \frac{68}{78} \times 133$ = 94 + 116 = 210 students		Expected count calculated.	
(ii)	Assumption is not valid because: • Year 11 students are younger than Year 13 students; hence it is unlikely that they will have the same chance of gaining Achievement (or better) in their exams (e.g. they might not take their examinations as seriously as the Year 13 students). • Year 13 students sitting examinations may have developed better study skills; hence may be more likely to gain Achievement (or better) in their examinations than Year 11 students. • Year 11 students self-classified as snorers, while we do not know how the Year 13 students were classified. Less Year 13 students that got Not Achieved may have responded. • The proportion of snorers is different in the two groups, so it may be that the groups of snorers are not comparable (e.g. only very light snorers are included in only one group). • We do not know what the pass rates are for the different exams. Some NCEA standards may have high rates of Not Achieved. *Accept other valid possible assumptions.* A discussion about different sample sizes is insufficient.		One reason for assumption being invalid is described in context.	Two reasons for assumption being invalid is described in context.

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Reasonable start / 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)(i)	P(log in daily) $= 0.6 \times 0.2 + 0.475 \times 0.8$ $= 0.5$ Number $= 0.5 \times 1$ billion $= 0.5$ billion Hence the claim is **not** supported.	Correct proportion of users calculated. AND correct conclusion that claim is not supported.		
(ii)	P(non-US | log in daily) $= \frac{0.475 \times 0.8}{0.2 \times 0.6 + 0.475 \times 0.8} = \frac{0.38}{0.5} = 0.76$ It is more likely that users that log in daily are non-US users. Alternatively, also calculate P(US | log in daily) $= \frac{0.2 \times 0.6}{0.2 \times 0.6 + 0.475 \times 0.8} = \frac{0.12}{0.5} = 0.24$ It is less likely that users that log in daily are US users.	Correct probability calculated.	Correct proportion of users calculated. AND Correct conclusion in context.	Both of (a)(i) and (a)(ii) correct. AND correct reason for why model cannot be applied for all US users.
(iii)	This model can not be applied for all US Instagram users as: • Some US Instagram users will always log in daily (e.g. social media influencers). • Some US Instagram users may not have daily access to internet services (e.g. live in remote areas). *Accept other valid possible reasons.*			

(b)	 			
(i)	Number that use both Snapchat and Instagram but not Facebook $= 195 - 84 = 111$ Proportion $= \frac{111}{300} = \frac{32}{100} = 0.37$	Correct proportion calculated.		
(ii)	$P(SC \mid IG) = \frac{195}{216} = 0.9028$ $P(FB \mid IG) = \frac{93}{216} = 0.4306$ Ratio $\frac{195}{93} = 2.1$ or Ratio $= \frac{0.9028}{0.4306} = 2.1$ A randomly chosen young adult is more than twice as likely to be a Snapchat user as a FB user. Hence, the claim is **correct.**		The two correct relevant values or probabilities.	The two correct relevant values or probabilities compared. AND Claim confirmed.

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Reasonable start / 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 – 7	8 – 13	14– 18	19 – 24