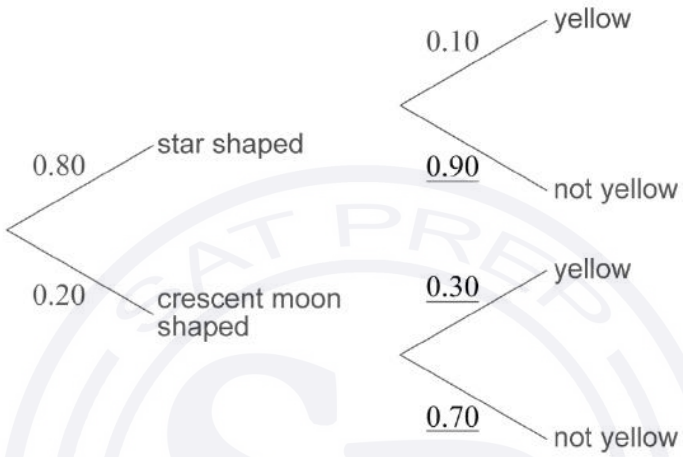
Subject – Math AI(Standard Level)
Topic - Statistics and Probability
Year - May 2021 – Nov 2022
Paper -2
Answers

Question 1

(a)



A1A1

[2 marks]

(b) (i) $P(Y) = 0.8 \times 0.1 + 0.2 \times 0.3$ ***M1***

$= 0.14$ ***A1***

(ii) $P(\text{Star} \mid Y) = \dfrac{0.8 \times 0.1}{0.14}$ ***M1***

$= 0.571 \left(\dfrac{4}{7}, 0.571428\ldots \right)$ ***A1***

[4 marks]

(c) the colours of the sweets are distributed according to manufacturer specifications ***A1***

[1 mark]

(d)

Colour	Brown	Red	Green	Orange	Yellow	Purple
Expected Frequency	12	20	16	16	8	8

A2

[2 marks]

(e) 5 **A1**

[1 mark]

(f) 0.469 (0.4688117…) **A2**

[2 marks]

(g) since $0.469 > 0.05$ **R1**

fail to reject the null hypothesis. There is insufficient evidence to reject the manufacturer's specifications **A1**

[2 marks]

Total [14 marks]

Question 2

(a) (i) evidence of correct probability *(M1)*
e.g sketch **OR** correct probability statement, $P(X < 6.5)$

0.0151 **A1**

(ii) 0.0228 **A1**

Note: Answers should be given to 4 decimal place.

[3 marks]

(b) (i) multiplying **their** probability by 1000 *(M1)*
451.7 **A1**

(ii) 510.5 **A1**

[3 marks]

Note: Answers should be given to 4 sf.

(c) H_0 : stopping distances can be modelled by $N(6.76, 0.12^2)$
H_1 : stopping distances cannot be modelled by $N(6.76, 0.12^2)$ **A1A1**

[2 marks]

(d) 15.1 or 22.8 seen *(M1)*

0.0727 (0.0726542…, 7.27%) **A2**

[3 marks]

(e) $0.05 < 0.0727$ **R1**
there is insufficient evidence to reject H_0 (or "accept H_0") **A1**

Note: Do not award ***R0A1***.

[2 marks]
Total [13 marks]

Question 3

(a) Quota sampling **A1**

[1 mark]

(b) 10 (hours) **A1**

[1 mark]

(c) $15-7$ **(M1)**

Note: Award ***M1*** for 15 and 7 seen.

8 **A1**

[2 marks]

(d) indication of a valid attempt to find the upper fence **(M1)**

$15+1.5\times 8$

27 **A1**

$25<27$ (accept equivalent answer in words) **R1**

Jason is correct **A1**

Note: Do not award ***R0A1***. Follow through **within** this part from *their* 27, but only if their value is supported by a valid attempt **or** clearly and correctly explains what their value represents.

[4 marks]

(e) "negative" seen **A1**

Note: Strength cannot be inferred visually; ignore "strong" or "weak".

[1 mark]

(f) correct substitution **(M1)**

$y=-1.54\times 1.5+98.8$

96.5 (%) (96.49) **A1**

[2 marks]

(g) not reliable **A1**

extrapolation **OR** outside the given range of the data **R1**

Note: Do not award ***A1R0***. Only accept reasoning that includes reference to the range of the data. Do not accept a contextual reason such as 1.5 hours is too short to read the book.

[2 marks]

(h)

	Book							
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Rank – Number of pages	**1**	3	5	2	6	8	4	7
Rank – Top 50 Rating	**1**	2	3	4	5	6	7	8

A1A1

Note: Award ***A1*** for correct ranks for 'number of pages'. Award ***A1*** for correct ranks for 'top 50 rating'.

[2 marks]

(i) (i) 0.714 (0.714285...) ***A2***

Note: ***FT*** from their table.

(ii) **EITHER**
there is a (strong/moderate) positive association between the number of pages and the top 50 rating. ***A1***

OR
there is a (strong/moderate) agreement between the rank order of number of pages and the rank order top 50 rating. ***A1***

OR
there is a (strong/moderate) positive (linear) correlation between the rank order of number of pages and the rank order top 50 rating. ***A1***

Note: Follow through from their value of r_s .

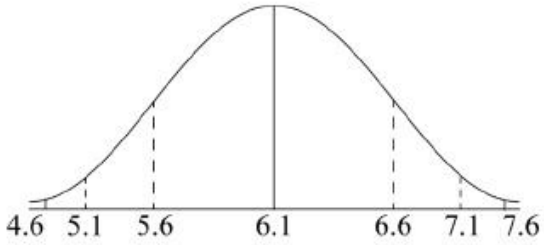
[3 marks]

Total [18 marks]

SATPREP www.satprep.co.in

Question 4

(a)



A1A1

Note: Award ***A1*** for a normal curve with mean labelled 6.1 or μ, ***A1*** for indication of SD (0.5): marks on horizontal axis at 5.6 and/or 6.6 **OR** $\mu-0.5$ and/or $\mu+0.5$ on the correct side and approximately correct position.

[2 marks]

(b) $X \sim \mathrm{N}(6.1,\ 0.5^2)$

$\mathrm{P}(5.5 < X < 6.5)$ **OR** labelled sketch of region ***(M1)***

$= 0.673\ \ (0.673074...)$ ***A1***

[2 marks]

(c) $(\mathrm{P}(X < 5.3) =)\ 0.0547992...$ ***(A1)***

$0.0547992... \times 80$ ***(M1)***

$= 4.38\ \ (4.38393...)$ ***A1***

[3 marks]

(d) 0.15 **OR** 0.85 ***(A1)***

$\mathrm{P}(X > x) = 0.15$ **OR** $\mathrm{P}(X < x) = 0.85$ **OR** labelled sketch of region ***(M1)***

$6.62\ \ (6.61821\ldots)$ ***A1***

[3 marks]

(e) $(\mathrm{P}(X > 6.25) =)\ 0.382088...$ ***(A1)***

recognition of binomial ***(M1)***

e.g. $\mathrm{B}(10,\ 0.382088...)$

$0.0502\ \ (0.0501768...)$ ***A2***

[4 marks]

Total [14 marks]

Question 5

(a) convenience sampling **A1**

[1 mark]

(b) (i) 95% **A1**

(ii) 1% **A1**

(iii) 2% **A1**

(iv) 98% **A1**

[4 marks]

(c) (i) 0.95×0.02 **(M1)**

0.019 **A1**

(ii) $0.05 \times 0.01 + 0.95 \times 0.98$ **(M1)(M1)**

Note: Award ***M1*** for summing two products and ***M1*** for correct products seen.

0.932 (0.9315) **A1**

(iii) recognition of conditional probability **(M1)**

$$\frac{0.05 \times 0.01}{0.05 \times 0.01 + 0.95 \times 0.98}$$ **A1**

0.000537 $(0.000536768\ldots)$ **A1**

Note: Accept 0.000536 if 0.932 used.

[8 marks]

(d) **EITHER**

sample may not be representative of population **A1**

OR

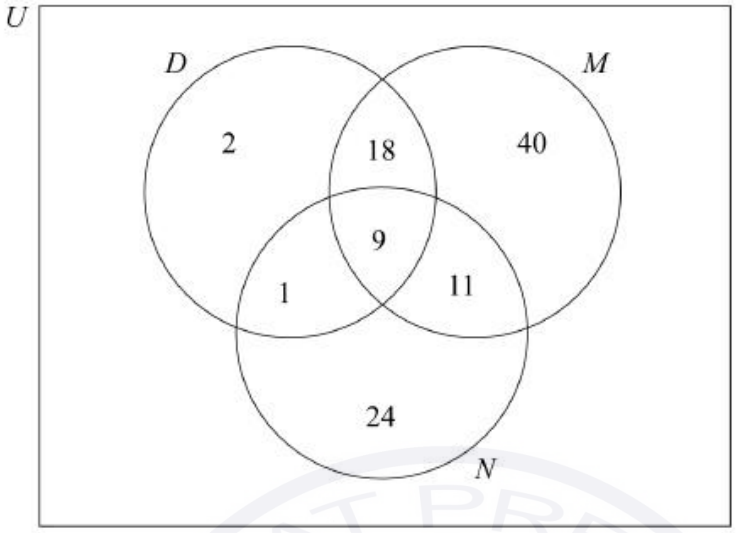
sample is not randomly selected **A1**

OR

unrealistic to think expected and observed values will be exactly equal **A1**

[1 mark]

(e)



A1A1A1

Note: Award ***A1*** for rectangle and 3 labelled circles and 9 in centre region; ***A1*** for 2, 40, 24; ***A1*** for 18, 1, and 11.

[3 marks]

(f) $18+9+1+11+2+40+24$ ***(M1)***

105 ***A1***

Note: Follow through from the entries on their Venn diagram in part (e). Working required for ***FT***.

[2 marks]

Total [19 marks]

SAT PREP www.satprep.co.in

Question 6

(a) (i) Let X be the random variable "distance from O".

$X \sim \mathrm{N}(10, 3^2)$

$\mathrm{P}(X < 13) = 0.841 \ (0.841344...)$ ***(M1)A1***

(ii) $(\mathrm{P}(X > 15) =) \ 0.0478 \ (0.0477903)$ ***A1***

[3 marks]

(b) $\mathrm{P}(X > 15) \times \mathrm{P}(X > 15)$ ***(M1)***

$= 0.00228 \ (0.00228391...)$ ***A1***

[2 marks]

(c) $1 - (0.8143)^3$ ***(M1)***

$0.460 \ (0.460050...)$ ***A1***

[2 marks]

(d) (i) **METHOD 1**

let Y be the random variable "number of points scored"

evidence of use of binomial distribution ***(M1)***

$Y \sim \mathrm{B}(10, 0.539949...)$ ***(A1)***

$(\mathrm{P}(Y \geq 5) =) \ 0.717 \ (0.716650...)$. ***A1***

METHOD 2

let Q be the random variable "number of times a point is not scored"

evidence of use of binomial distribution ***(M1)***

$Q \sim \mathrm{B}(10, 0.460050...)$ ***(A1)***

$(\mathrm{P}(Q \leq 5) =) \ 0.717 \ (0.716650...)$ ***A1***

(ii) $\mathrm{P}(5 \leq Y < 8)$ ***(M1)***

$0.628 \ (0.627788...)$ ***A1***

Note: Award ***M1*** for a correct probability statement or indication of correct lower and upper bounds, 5 and 7.

(iii) $\dfrac{\mathrm{P}(5 \leq Y < 8)}{\mathrm{P}(Y \geq 5)} \left(= \dfrac{0.627788...}{0.716650...} \right)$ ***(M1)***

$0.876 \ (0.876003...)$ ***A1***

[7 marks]

Total: [14 marks]

Question 7

(a) (i) $\frac{560}{1280}\left(\frac{7}{16}, 0.4375\right)$ **A1A1**

Note: Award ***A1*** for correct numerator, ***A1*** for correct denominator.

(ii) $\frac{72}{1280}\left(\frac{9}{160}, 0.05625\right)$ **A1A1**

Note: Award ***A1*** for correct numerator, ***A1*** for correct denominator.

(iii) $\frac{153}{348}\left(\frac{51}{116}, 0.439655...\right)$ **A1A1**

Note: Award ***A1*** for correct numerator, ***A1*** for correct denominator.

(iv) $160+224+128+205+131$ **OR** $560+512-224$ ***(M1)***

$\frac{848}{1280}\left(\frac{53}{80}, 0.6625\right)$ **A1A1**

Note: Award ***A1*** for correct denominator (1280) seen, ***(M1)*** for correct calculation of the numerator, ***A1*** for the correct answer.

[9 marks]

(b) H_0 : the variables are independent

H_1 : the variables are dependent **A1**

Note: Award ***A1*** for both hypotheses correct. Do not accept "not correlated" or "not related" in place of "independent".

[1 mark]

(c) 4 **A1**

[1 mark]

(d) (i) $(\chi^2 =)$ 23.3 (23.3258...) **A2**

(ii) 0.000109 (0.000108991...) **OR** 1.09×10^{-4} **A1**

(iii) **EITHER**

$23.3 > 13.277$ **R1**

OR

$0.000109 < 0.01$ **R1**

THEN

(there is sufficient evidence to accept H_1 that) preferred device and age group are not independent **A1**

Note: For the final ***A1*** the answer must be in context. Do not award ***A1R0***.

[5 marks]

Total: [16 marks]

Question 8

(a) (let T be the number of passengers who arrive)

$(\mathrm{P}(T > 72) =)\ \mathrm{P}(T \geq 73)$ **OR** $1 - \mathrm{P}(T \leq 72)$ ***(A1)***

$T \sim \mathrm{B}(74,\ 0.9)$ **OR** $n = 74$ ***(M1)***

$= 0.00379\ \ (0.00379124...)$ ***A1***

Note: Using the distribution $\mathrm{B}(74,\ 0.1)$, to work with the 10% that do not arrive for the flight, here and throughout this question, is a valid approach.

[3 marks]

(b) (i) 72×0.9 ***(M1)***

64.8 ***A1***

(ii) $n \times 0.9 = 72$ ***(M1)***

80 ***A1***

[4 marks]

(c) **METHOD 1**

EITHER

when selling 74 tickets

	$T \leq 72$	$T = 73$	$T = 74$
Income minus compensation (I)	11100	10800	10500
Probability	0.9962…	0.003380…	0.0004110…

top row ***A1A1***

bottom row ***A1A1***

Note: Award ***A1A1*** for each row correct. Award ***A1*** for one correct entry and ***A1*** for the remaining entries correct.

$\mathrm{E}(I) = 11100 \times 0.9962... + 10800 \times 0.00338... + 10500 \times 0.000411 \approx 11099$ ***(M1)A1***

OR

income is $74 \times 150 = 11100$ ***(A1)***

expected compensation is

$0.003380... \times 300 + 0.0004110... \times 600\ \ (= 1.26070...)$ ***(M1)A1A1***

expected income when selling 74 tickets is $11100 - 1.26070\ldots$ ***(M1)***

$= 11098.73..\ \ (= \$11099)$ ***A1***

THEN

income for 72 tickets $= 72 \times 150 = 10800$ ***(A1)***

so expected gain $\approx 11099 - 10800 = \$299$ ***A1***

Question 9

(a) (i) $\frac{370+472}{2}$ ***(M1)***

Note: This ***(M1)*** can also be awarded for either a correct Q_3 or a correct Q_1 in part (a)(ii).

$Q_3 = 421$ ***A1***

(ii) their part (a)(i) – their Q_1 (clearly stated) ***(M1)***

$IQR = (421-318=)\ 103$ ***A1***

[4 marks]

(b) $(Q_3 + 1.5(IQR) =)\ 421 + (1.5\times 103)$ ***(M1)***

$= 575.5$

since $498 < 575.5$ ***R1***

Netherlands is not an outlier ***A1***

Note: The ***R1*** is dependent on the ***(M1)***. Do not award ***R0A1***.

[3 marks]

(c) not appropriate ("no" is sufficient) ***A1***

as r is too close to zero / too weak a correlation ***R1***

[2 marks]

(d) (i) 6 ***A1***

(ii) 4.5 ***A1***

(iii) 4.5 ***A1***

[3 marks]

(e) (i) $r_s = 0.683\ (0.682646...)$ ***A2***

(ii) **EITHER**

there is a (positive) association between the population size and the score ***A1***

OR

there is a (positive) linear correlation between the ranks of the population size and the ranks of the scores (when compared with the PMCC of 0.249). ***A1***

[3 marks]

(f) lowering the top score by 20 does not change its rank so r_s is unchanged ***R1***

Note: Accept "this would not alter the rank" or "Netherlands still top rank" or similar. Condone any statement that clearly implies the ranks have not changed, for example: "The Netherlands still has the highest score."

[1 mark]

[Total 16 marks]

Question 10

(a) (i) 0.58 (s) ***A1***

(ii) $0.7 - 0.42$ ***(A1)(M1)***

Note: Award ***A1*** for correct quartiles seen, ***M1*** for subtraction of their quartiles.

0.28 s ***A1***

[4 marks]

(b) 9 (people have reaction time ≤ 0.4) ***(A1)***

31 (people have reaction time > 0.4) ***A1***

[2 marks]

(c) $(90\% \times 40 =)$ 36 **OR** 4 ***(A1)***

0.8 s ***A1***

[2 marks]

(d) (i) $(a =)$ 6 ***A1***

(ii) $(b =)$ 4 ***A1***

[2 marks]

(e) $0.6 < t \leq 0.8$ ***A1***

[1 mark]

(f) 0.55 s ***A2***

[2 marks]

(g) the mean will increase ***A1***

because the incorrect reaction times are moving from a lower interval to a higher interval which will increase the numerator of the mean calculation ***R1***

the median will stay the same ***A1***

because the median or middle of the data is greater than both intervals being changed ***R1***

Note: Do not award ***A1R0.***

[4 marks]

Total [17 marks]

Question 10

(a) (i) $(m =)\ 54(\%)$ **A1**

(ii) $(n =)\ 14(\%)$ **A1**

(iii) $(p =)\ 22(\%)$ **A1**

(iv) $(q =)\ 10(\%)$ **A1**

Note: Based on their n, follow through for parts (i) and (iii), but only if it does not contradict the given information. Follow through for part (iv) but only if the total is 100%.

[4 marks]

(b) 90 (%) **A1**

Note: Award ***A0*** for a decimal answer.

[1 mark]

(c) (i) $0.54\ \left(\frac{54}{100}, \frac{27}{50}, 54\%\right)$ **A1**

(ii) $\frac{54}{64}\ \left(0.844, \frac{27}{32}, 84.4\%, 0.84375\right)$ **A1A1**

Note: Award ***A1*** for a correct denominator (0.64 or 64 seen), ***A1*** for the correct final answer.

[3 marks]

(d) (i) recognizing Binomial distribution with correct parameters ***(M1)***

$X \sim \mathrm{B}(10, 0.68)$

$(\mathrm{P}(X = 5) =)\ 0.123\ (0.122940..., 12.3\%)$ **A1**

(ii) $1 - \mathrm{P}(X \leq 3)$ **OR** $\mathrm{P}(X \geq 4)$ **OR** $\mathrm{P}(4 \leq X \leq 10)$ ***(M1)***

0.984 (0.984497..., 98.4%) **A1**

(iii) $(0.68)^9 \times 0.32$ ***(M1)***

recognition of two possible cases ***(M1)***

$2 \times \left((0.68)^9 \times 0.32\right)$

0.0199 (0.0198957..., 1.99%) **A1**

[7 marks]

(e) **EITHER**

the probability is not constant **A1**

OR

the events are not independent **A1**

OR

the events should be modelled by the hypergeometric distribution instead **A1**

[1 mark]

Total [16 marks]

Question 11

(a) continuous **A1**

[1 mark]

(b) $160-50-62-14-8$ **(M1)**

$(k=)\ 26$ **A1**

[2 marks]

(c) (i) $20 \le T < 40$ **A1**

(ii) 30 **A1**

[2 marks]

(d) 33.5 minutes **A2**

Note: ***FT*** from their value of k and their mid-interval value. Follow through from part (c)(ii) but only if mid-interval value lies in their interval.

[2 marks]

(e) 112 **A1**

[1 mark]

(f) $\frac{22}{160}\ \left[0.138,\ 0.1375,\ 13.75\%,\ \frac{11}{80}\right]$ **A1A1**

Note: Award ***A1*** for correct numerator, ***A1*** for correct denominator.

[2 marks]

(g) 26 minutes **A1**

[1 mark]

(h) $50-16$ **(M1)**

Note: Award ***M1*** for both correct quartiles seen.

34 minutes **A1**

[2 marks]

SATPREP www.satprep.co.in

(i) correct substitution into outlier formula ***(M1)***

$50+1.5\times34$

$=101$ ***A1***

$92<101$ **OR** highest value on diagram <101 ***R1***

not an outlier ***AG***

Note: Award ***R1*** for their correct comparison. Follow through from their part (h).
Award ***R0*** if their conclusion is "it is an outlier", this contradicts Elsie's belief.

[3 marks]

(j) **EITHER**

the diagram is not symmetric or equivalent

e.g the median is not in the center of the box or

the lengths of the whiskers are (very) different or (positive or right) skew

OR

the mean and median are (very) different; ***A1***

[1 mark]

Total [17 marks]

SAT PREP SP www.satprep.co.in