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

Question 1

[Maximum mark: 15]

The length, X mm, of a certain species of seashell is normally distributed with mean 25 and variance, σ^2.

The probability that X is less than 24.15 is 0.1446.

(a) Find $P(24.15 < X < 25)$. [2]

(b) (i) Find σ, the standard deviation of X.

(ii) Hence, find the probability that a seashell selected at random has a length greater than 26 mm. [5]

A random sample of 10 seashells is collected on a beach. Let Y represent the number of seashells with lengths greater than 26 mm.

(c) Find $E(Y)$. [3]

(d) Find the probability that exactly three of these seashells have a length greater than 26 mm. [2]

A seashell selected at random has a length less than 26 mm.

(e) Find the probability that its length is between 24.15 mm and 25 mm. [3]

Question 2

[Maximum mark: 5]

The following table below shows the marks scored by seven students on two different mathematics tests.

Test 1 (x)	15	23	25	30	34	34	40
Test 2 (y)	20	26	27	32	35	37	35

Let L_1 be the regression line of x on y. The equation of the line L_1 can be written in the form $x = ay + b$.

(a) Find the value of a and the value of b. [2]

Let L_2 be the regression line of y on x. The lines L_1 and L_2 pass through the same point with coordinates (p, q).

(b) Find the value of p and the value of q. [3]

Question 3

[Maximum mark: 6]

A six-sided biased die is weighted in such a way that the probability of obtaining a "six" is $\frac{7}{10}$.

The die is tossed five times. Find the probability of obtaining

(a) at most three "sixes". [3]

(b) the third "six" on the fifth toss. [3]

Question 4

[Maximum mark: 6]

A company produces bags of sugar whose masses, in grams, can be modelled by a normal distribution with mean 1000 and standard deviation 3.5. A bag of sugar is rejected for sale if its mass is less than 995 grams.

(a) Find the probability that a bag selected at random is rejected. [2]

(b) Estimate the number of bags which will be rejected from a random sample of 100 bags. [1]

(c) Given that a bag is not rejected, find the probability that it has a mass greater than 1005 grams. [3]

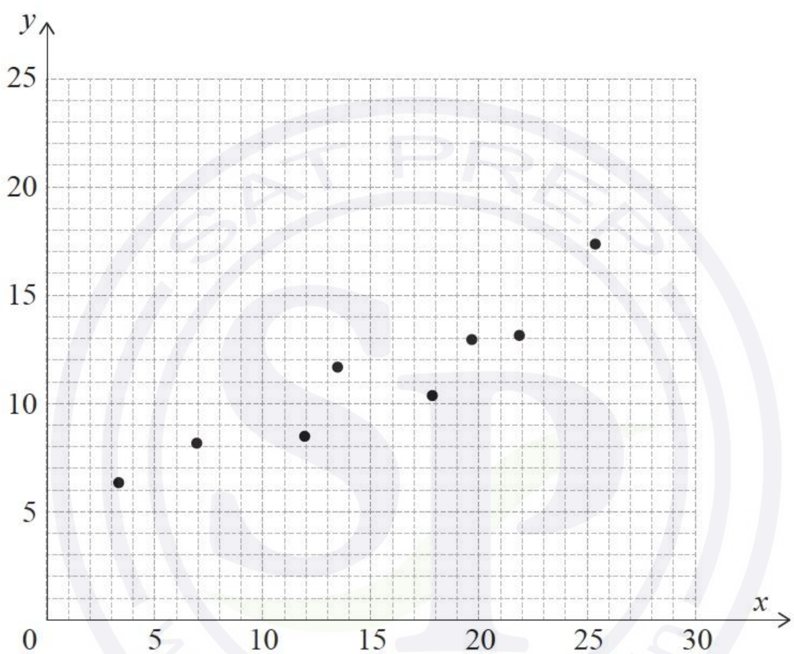
Question 5

[Maximum mark: 7]

The following table shows the data collected from an experiment.

x	3.3	6.9	11.9	13.4	17.8	19.6	21.8	25.3
y	6.3	8.1	8.4	11.6	10.3	12.9	13.1	17.3

The data is also represented on the following scatter diagram.



The relationship between x and y can be modelled by the regression line of y on x with equation $y = ax + b$, where $a, b \in \mathbb{R}$.

(a) Write down the value of a and the value of b. [2]

(b) Use this model to predict the value of y when $x = 18$. [2]

(c) Write down the value of $\bar{x}$ and the value of $\bar{y}$. [1]

(d) Draw the line of best fit on the scatter diagram. [2]

Question 6

[Maximum mark: 15]

All answers in this question should be given to four significant figures.

In a local weekly lottery, tickets cost $\$2$ each.

In the first week of the lottery, a player will receive $\$D$ for each ticket, with the probability distribution shown in the following table. For example, the probability of a player receiving $\$10$ is 0.03. The grand prize in the first week of the lottery is $\$1000$.

d	0	2	10	50	Grand Prize
$\mathrm{P}(D=d)$	0.85	c	0.03	0.002	0.0001

(a) Find the value of c. [2]

(b) Determine whether this lottery is a fair game in the first week. Justify your answer. [4]

If nobody wins the grand prize in the first week, the probabilities will remain the same, but the value of the grand prize will be $\$2000$ in the second week, and the value of the grand prize will continue to double each week until it is won. All other prize amounts will remain the same.

(c) Given that the grand prize is not won and the grand prize continues to double, write an expression in terms of n for the value of the grand prize in the nth week of the lottery. [2]

The wth week is the first week in which the player is expected to make a profit. Ryan knows that if he buys a lottery ticket in the wth week, his expected profit is $\$p$.

(d) Find the value of p. [7]

Question 7

[Maximum mark: 15]

The flight times, T minutes, between two cities can be modelled by a normal distribution with a mean of 75 minutes and a standard deviation of σ minutes.

(a) Given that 2% of the flight times are longer than 82 minutes, find the value of σ. [3]

(b) Find the probability that a randomly selected flight will have a flight time of more than 80 minutes. [2]

(c) Given that a flight between the two cities takes longer than 80 minutes, find the probability that it takes less than 82 minutes. [4]

On a particular day, there are 64 flights scheduled between these two cities.

(d) Find the expected number of flights that will have a flight time of more than 80 minutes. [3]

(e) Find the probability that more than 6 of the flights on this particular day will have a flight time of more than 80 minutes. [3]

Question 8

[Maximum mark: 8]

At a school, 70% of the students play a sport and 20% of the students are involved in theatre. 18% of the students do neither activity.

A student is selected at random.

(a) Find the probability that the student plays a sport and is involved in theatre. [2]

(b) Find the probability that the student is involved in theatre, but does not play a sport. [2]

At the school 48% of the students are girls, and 25% of the girls are involved in theatre.

A student is selected at random. Let G be the event "the student is a girl" and let T be the event "the student is involved in theatre".

(c) Find $\mathrm{P}(G \cap T)$. [2]

(d) Determine if the events G and T are independent. Justify your answer. [2]

Question 9

[Maximum mark: 6]

At a café, the waiting time between ordering and receiving a cup of coffee is dependent upon the number of customers who have already ordered their coffee and are waiting to receive it.

Sarah, a regular customer, visited the café on five consecutive days. The following table shows the number of customers, x, ahead of Sarah who have already ordered and are waiting to receive their coffee and Sarah's waiting time, y minutes.

Number of customers (x)	3	9	11	10	5
Sarah's waiting time (y)	6	10	12	11	6

The relationship between x and y can be modelled by the regression line of y on x with equation $y = ax + b$.

(a) (i) Find the value of a and the value of b.

(ii) Write down the value of Pearson's product-moment correlation coefficient, r. [3]

(b) Interpret, in context, the value of a found in part (a)(i). [1]

On another day, Sarah visits the café to order a coffee. Seven customers have already ordered their coffee and are waiting to receive it.

(c) Use the result from part (a)(i) to estimate Sarah's waiting time to receive her coffee. [2]

Question 10

[Maximum mark: 16]

The random variable X follows a normal distribution with mean μ and standard deviation σ.

(a) Find $\mathrm{P}(\mu - 1.5\sigma < X < \mu + 1.5\sigma)$. [3]

The avocados grown on a farm have weights, in grams, that are normally distributed with mean μ and standard deviation σ. Avocados are categorized as small, medium, large or premium, according to their weight. The following table shows the probability an avocado grown on the farm is classified as small, medium, large or premium.

Category	Small	Medium	Large	Premium
Probability	0.04	0.576	0.288	0.096

The maximum weight of a small avocado is 106.2 grams.

The minimum weight of a premium avocado is 182.6 grams.

(b) Find the value of μ and of σ. [5]

A supermarket purchases all the avocados from the farm that weigh more than 106.2 grams.

(c) Find the probability that an avocado chosen at random from this purchase is categorized as

(i) medium;

(ii) large;

(iii) premium. [4]

The selling prices of the different categories of avocado at this supermarket are shown in the following table:

Category	Medium	Large	Premium
Selling price ($) per avocado	1.10	1.29	1.96

The supermarket pays the farm $\$200$ for the avocados and assumes it will then sell them in exactly the same proportion as purchased from the farm.

(d) According to this model, find the minimum number of avocados that must be sold so that the net profit for the supermarket is at least $\$438$. [4]

Question 11

[Maximum mark: 7]

A factory manufactures lamps. It is known that the probability that a lamp is found to be defective is 0.05. A random sample of 30 lamps is tested.

(a) Find the probability that there is at least one defective lamp in the sample. [3]

(b) Given that there is at least one defective lamp in the sample, find the probability that there are at most two defective lamps. [4]

Question 12

[Maximum mark: 5]

In Lucy's music academy, eight students took their piano diploma examination and achieved scores out of 150. For her records, Lucy decided to record the average number of hours per week each student reported practising in the weeks prior to their examination. These results are summarized in the table below.

Average weekly practice time (h)	28	13	45	33	17	29	39	36
Diploma score (D)	115	82	120	116	79	101	110	121

(a) Find Pearson's product-moment correlation coefficient, r, for these data. [2]

(b) The relationship between the variables can be modelled by the regression equation $D = ah + b$. Write down the value of a and the value of b. [1]

(c) One of these eight students was disappointed with her result and wished she had practised more. Based on the given data, determine how her score could have been expected to alter had she practised an extra five hours per week. [2]

Question 13

[Maximum mark: 16]

A bakery makes two types of muffins: chocolate muffins and banana muffins.

The weights, C grams, of the chocolate muffins are normally distributed with a mean of 62 g and standard deviation of 2.9 g.

(a) Find the probability that a randomly selected chocolate muffin weighs less than 61 g. [2]

(b) In a random selection of 12 chocolate muffins, find the probability that exactly 5 weigh less than 61 g. [2]

The weights, B grams, of the banana muffins are normally distributed with a mean of 68 g and standard deviation of 3.4 g.

Each day 60% of the muffins made are chocolate.

On a particular day, a muffin is randomly selected from all those made at the bakery.

(c) (i) Find the probability that the randomly selected muffin weighs less than 61 g.

(ii) Given that a randomly selected muffin weighs less than 61 g, find the probability that it is chocolate. [7]

The machine that makes the chocolate muffins is adjusted so that the mean weight of the chocolate muffins remains the same but their standard deviation changes to $\sigma\text{ g}$. The machine that makes the banana muffins is not adjusted. The probability that the weight of a randomly selected muffin from these machines is less than 61 g is now 0.157.

(d) Find the value of σ. [5]

Question 14

[Maximum mark: 6]

Let A and B be two independent events such that $\mathrm{P}(A \cap B') = 0.16$ and $\mathrm{P}(A' \cap B) = 0.36$.

(a) Given that $\mathrm{P}(A \cap B) = x$, find the value of x. [4]

(b) Find $\mathrm{P}(A' \mid B')$. [2]

Question 15

[Maximum mark: 6]

A discrete random variable, X, has the following probability distribution:

x	0	1	2	3
$P(X=x)$	0.41	$k-0.28$	0.46	$0.29-2k^2$

(a) Show that $2k^2 - k + 0.12 = 0$. [1]

(b) Find the value of k, giving a reason for your answer. [3]

(c) Hence, find $E(X)$. [2]

Question 16

[Maximum mark: 4]

The number of hours spent exercising each week by a group of students is shown in the following table.

Exercising time (in hours)	**Number of students**
2	5
3	1
4	4
5	3
6	x

The median is 4.5 hours.

(a) Find the value of x. [2]

(b) Find the standard deviation. [2]

Question 17

[Maximum mark: 18]

The time it takes Suzi to drive from home to work each morning is normally distributed with a mean of 35 minutes and a standard deviation of σ minutes.

On 25% of days, it takes Suzi longer than 40 minutes to drive to work.

(a) Find the value of σ. [4]

(b) On a randomly selected day, find the probability that Suzi's drive to work will take longer than 45 minutes. [2]

Suzi will be late to work if it takes her longer than 45 minutes to drive to work. The time it takes to drive to work each day is independent of any other day.

Suzi will work five days next week.

(c) Find the probability that she will be late to work at least one day next week. [3]

(d) Given that Suzi will be late to work at least one day next week, find the probability that she will be late less than three times. [5]

Suzi will work 22 days this month. She will receive a bonus if she is on time at least 20 of those days.

So far this month, she has worked 16 days and been on time 15 of those days.

(e) Find the probability that Suzi will receive a bonus. [4]

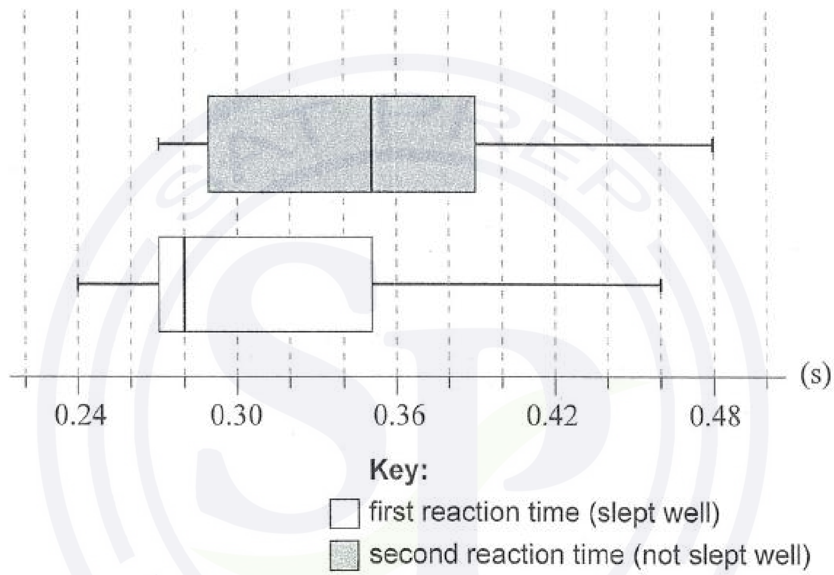
Question 18

[Maximum mark: 6]

A random sample of nine adults were selected to see whether sleeping well affected their reaction times to a visual stimulus. Each adult's reaction time was measured twice.

The first measurement for reaction time was taken on a morning after the adult had slept well. The second measurement was taken on a morning after the same adult had not slept well.

The box and whisker diagrams for the reaction times, measured in seconds, are shown below.



Consider the box and whisker diagram representing the reaction times after sleeping well.

(a) State the median reaction time after sleeping well. [1]

(b) Verify that the measurement of 0.46 seconds is not an outlier. [3]

(c) State why it appears that the mean reaction time is greater than the median reaction time. [1]

Now consider the two box and whisker diagrams.

(d) Comment on whether these box and whisker diagrams provide any evidence that might suggest that not sleeping well causes an increase in reaction time. [1]

Question 19

[Maximum mark: 6]

Events A and B are independent and $\mathrm{P}(A) = 3\mathrm{P}(B)$.

Given that $\mathrm{P}(A \cup B) = 0.68$, find $\mathrm{P}(B)$.

Question 20

[Maximum mark: 16]

The time worked, T, in hours per week by employees of a large company is normally distributed with a mean of 42 and standard deviation 10.7.

(a) Find the probability that an employee selected at random works more than 40 hours per week. [2]

(b) A group of four employees is selected at random. Each employee is asked in turn whether they work more than 40 hours per week. Find the probability that the fourth employee is the only one in the group who works more than 40 hours per week. [3]

(c) A large group of employees work more than 40 hours per week.

(i) An employee is selected at random from this large group.

Find the probability that this employee works less than 55 hours per week.

(ii) Ten employees are selected at random from this large group.

Find the probability that exactly five of them work less than 55 hours per week. [7]

It is known that $\mathrm{P}(a \le T \le b) = 0.904$ and that $\mathrm{P}(T > b) = 2\mathrm{P}(T < a)$, where a and b are numbers of hours worked per week. An employee who works fewer than a hours per week is considered to be a part-time employee.

(d) Find the maximum time, in hours per week, that an employee can work and still be considered part-time. [4]

Question 20

[Maximum mark: 5]

The following table shows the Mathematics test scores (x) and the Science test scores (y) for a group of eight students.

Mathematics scores (x)	64	68	72	75	80	82	85	86
Science scores (y)	67	72	77	76	84	83	89	91

The regression line of y on x for this data can be written in the form $y = ax + b$.

(a) Find the value of a and the value of b. [2]

(b) Write down the value of the Pearson's product-moment correlation coefficient, r. [1]

(c) Use the equation of your regression line to predict the Science test score for a student who has a score of 78 on the Mathematics test. Express your answer to the nearest integer. [2]