



H1 Maths Preliminary Exam Solutions

Section A: Pure Mathematics [35 marks]

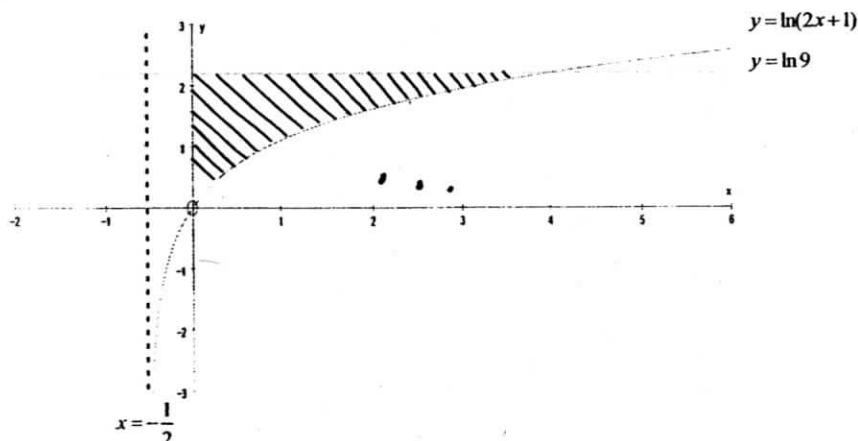
- 1 Find  $\frac{d}{dx}(\sqrt{(x-1)^2+3})$ , and simplify your answer.

Hence, evaluate  $\int_{-2}^3 \frac{1-x}{\sqrt{(x-1)^2+3}} dx$  giving your answer in the form  $\sqrt{a}-\sqrt{b}$  where  $a$  and  $b$  are positive constants to be determined. [4]

**Solution**

$$\begin{aligned}\frac{d}{dx}(\sqrt{(x-1)^2+3}) &= \frac{1}{2\sqrt{(x-1)^2+3}}(2)(x-1) \\ &= \frac{x-1}{\sqrt{(x-1)^2+3}} \\ \int_{-2}^3 \frac{1-x}{\sqrt{(x-1)^2+3}} dx &= -\int_{-2}^3 \frac{x-1}{\sqrt{(x-1)^2+3}} dx \\ &= -\left[\sqrt{(x-1)^2+3}\right]_{-2}^3 \\ &= -\sqrt{7} + \sqrt{12} \\ a &= 12, \quad b = 7\end{aligned}$$

- 2 Find the exact area bounded by the curve  $y = \ln(2x+1)$ , the  $y$ -axis and the line  $y = \ln 9$  (as shown in the diagram below) in the form  $p - \ln q$  where  $p$  and  $q$  are constants to be determined. [3]



**Solution**

$$y = \ln(2x+1) \Rightarrow x = \frac{e^y - 1}{2}$$

Area of shaded region by integrating with respect to the  $y$ -axis,

$$= \int_0^{\ln 9} \frac{e^y - 1}{2} dy = \left[ \frac{e^y}{2} - \frac{y}{2} \right]_0^{\ln 9}$$

$$= \frac{9}{2} - \frac{\ln 9}{2} - \frac{1}{2} = 4 - \frac{1}{2} \ln 9$$

$$= 4 - \ln 3$$

$$p=4, \quad q=3$$

- 3 Using a non-calculator method, find the coordinates of all the stationary points of the curve  $y = 3x^5 - 20x^3$ . Determine the nature of each of the stationary points, showing your working clearly. [6]

**Solution**

$$\frac{dy}{dx} = 15x^4 - 60x^2 = 0$$

$$15x^2(x^2 - 4) = 0$$

$$x = 0 \quad \text{or} \quad x = 2 \quad \text{or} \quad x = -2$$

$$y = 0 \quad \text{or} \quad y = -64 \quad \text{or} \quad y = 64$$

$x$	$0^-$	$0$	$0^+$
$\frac{dy}{dx}$	-ve	0	-ve

$x$	$2^-$	$2$	$2^+$
$\frac{dy}{dx}$	-ve	0	+ve

$x$	$-2^-$	$-2$	$-2^+$
$\frac{dy}{dx}$	+ve	0	-ve

$(0, 0)$  is a stationary point of inflexion,  $(2, -64)$  is a min turning point,  $(-2, 64)$  is a max turning point.

4 Given  $\log_3 2 = a$ ,  $\log_3 125 = b$ , express the following in terms of  $a$  and  $b$ .

(i)  $\log_3 \frac{4}{5}$ , [2]

(ii)  $\log_3 200$ , [2]

(iii)  $\log_9 10$ . [3]

**Solution**

$$(i) \log_3 \frac{4}{5} = \log_3 2^2 - \log_3 125^{\frac{1}{3}} = 2a - \frac{b}{3}$$

$$(ii) \log_3 200 = \log_3 (8 \times 25) = \log_3 2^3 + \log_3 125^{\frac{2}{3}} = 3a + \frac{2}{3}b$$

$$(iii) \log_9 10 = \frac{\log_3 10}{\log_3 9} = \frac{\log_3 (2 \times 5)}{\log_3 3^2} = \frac{\log_3 2 + \log_3 5}{2} = \frac{a + \frac{b}{3}}{2} = \frac{a}{2} + \frac{b}{6}$$

5 Find the range of values of  $m$  for which the graphs of  $y = x^2 - 2mx + 1$  and  $y = m - x$  do not intersect. [3]

For  $m = 1$ , find the coordinates of the two points at which the two graphs intersect. Hence find the equation of the normal to  $y = x^2 - 2mx + 1$  at each of the two points. [4]

**Solution**

$$x^2 - 2mx + 1 = m - x$$

$$x^2 - 2mx + 1 - m + x = 0$$

$$x^2 + (1 - 2m)x + 1 - m = 0$$

$$\text{Discriminant} = (1 - 2m)^2 - 4(1 - m) < 0$$

$$1 - 4m + 4m^2 - 4 + 4m < 0$$

$$4m^2 < 3$$

$$-\frac{\sqrt{3}}{2} < m < \frac{\sqrt{3}}{2}$$

$y = x^2 - 2x + 1$  and  $y = 1 - x$  intersect at  $(0, 1)$  and  $(1, 0)$

$$\text{For } y = x^2 - 2x + 1, \frac{dy}{dx} = 2x - 2$$

$$\text{At } (0, 1), \frac{dy}{dx} = -2, \text{ gradient of normal} = \frac{1}{2}$$

$$\text{Equation of normal: } y = \frac{1}{2}x + 1$$

$$\text{At } (1, 0), \frac{dy}{dx} = 0$$

$$\text{Equation of normal: } x = 1$$

6 The curve  $C$  has equation  $y = \frac{6-5x}{2-x}$ .

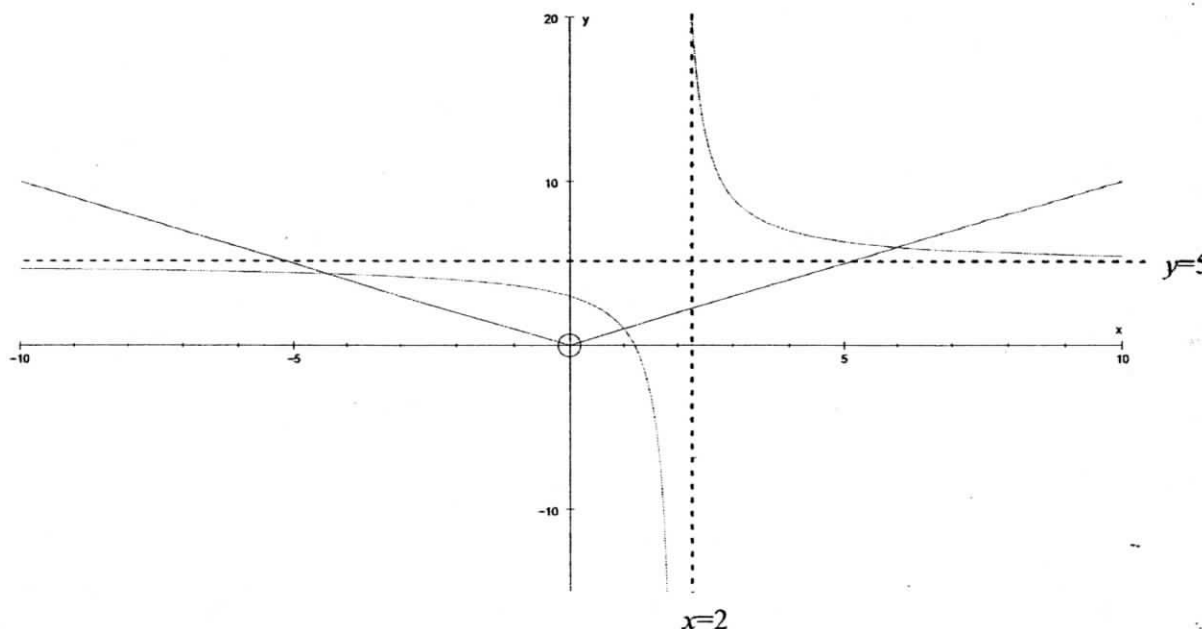
(i) Sketch the graph of  $C$ , showing clearly the intercepts with the axes and the equation(s) of any asymptote(s). [3]

(ii) By adding a suitable graph on the same diagram, solve the inequality,

$$\frac{6-5x}{2-x} < |x| \quad [3]$$

(iii) Hence, solve the inequality  $\frac{6+5x}{x+2} < |x|$ . [2]

**Solution**



$$x < -4.37 \text{ or } 1 < x < 2 \text{ or } x > 6$$

Replace  $x$  by  $-x$ ,

$$-x < -4.37 \text{ or } 1 < -x < 2 \text{ or } -x > 6$$

$$x > 4.37 \text{ or } -2 < x < -1 \text{ or } x < -6$$

**Section B: Statistics [60 marks]**

- 7 At an ASEAN Games Chess Championship, a reporter in the competition village is conducting a survey to find out what the participants think of the weather, accommodation and food provided. From a namelist of 300 participants from 5 countries, the reporter selects the 10<sup>th</sup>, 20<sup>th</sup>, 30<sup>th</sup> ... person to conduct the survey via the systematic sampling method. It is noted that each country has varying number of participants.

(i) How should he improve on the way he got his data for the systematic sampling method?

[1]

(ii) Give a reason why stratified sampling method might be preferable in this context and describe how it can be done.

[4]

**Solution**

(i) He should have randomly picked a number from 1 to 10 as the 1<sup>st</sup> person, and pick every 10<sup>th</sup> person on the namelist after that.

(ii) In stratified sampling, the participants are divided into non-overlapping groups of strata, for example nationality. As participants from different countries could have differing views on weather, food and accommodation based on their backgrounds, selecting the participants randomly from each stratum with sample size proportional to the relative size of the stratum will give a more representative sample and avoid any possible cyclical pattern from the systematic sampling.

For example if the 300 participants from the 5 countries were in the proportion a% from country A, b% from country B, c% from country C, d% from country D and e% from country E, then number of participants selected from each country for the survey should follow the same proportion. Furthermore, the participants within each country are randomly selected using simple random sampling or stratified sampling.

8. At a doctor's clinic, records show that 18% of patients who make an appointment fail to turn up for the afternoon consultation. During afternoon consultation hours, the doctor has time to see 20 patients within a 4 hour clinic.

There are 20 appointments to see the doctor on one afternoon.

- (i) Find the probability that all patients turn up. [2]

- (ii) Find the probability that more than 4 patients do not turn up. [2]

To improve efficiency, the doctor decides to make more than 20 appointments for the afternoon consultation, although there will still only be enough time to see 20 patients within a 4 hour clinic.

- (iii) Find the greatest number of appointments the doctor can make for afternoon consultation and have a probability of 0.9 of having enough time within the 4 hour clinic to see all the patients who turn up. [2]

#### Solution

Let  $X$  be the number of patients out of 20 who fail to turn up for their appointments.

$$Y \sim B(20, 0.18)$$

- (i)  $P(X=0) = 0.0189$  (3 s.f.)  
(ii)  $P(X > 4) = 1 - P(X \leq 4) = 0.285$  (3 s.f.)  
(iii) Let  $Y$  be the number of patients out of  $n$  who fail to turn up for their appointment.

$$\text{If } n = 21, P(Y \geq 1) = 1 - P(Y = 0) = 0.985$$

$$n = 22, P(Y \geq 2) = 1 - P(Y \leq 1) = 0.926$$

$$n = 23, P(Y \geq 3) = 1 - P(Y \leq 2) = 0.810$$

There the greatest number of appointments is 22.

- 9 Mark and June played four rounds of a one versus one online game, where there will always be a winner for each round. The probability that Mark wins an odd round is 0.7 and the probability that June wins an even round is 0.8.

- (i) Find the probability that Mark wins the first three rounds. [2]

- (ii) Find the probability that June wins two rounds out of the first three rounds. [2]

- (iii) Find the probability that after four rounds, they each have won two rounds. [3]

- (iv) Given that Mark wins two rounds in the first three rounds, find the probability that Mark wins three consecutive rounds in the four rounds. [3]

**Solution**

$$(i) P(MMM) = (0.7)(0.2)(0.7) = 0.098$$

$$(ii) P(JJM) + P(JMJ) + P(MJJ) = (0.3)(0.8)(0.7) + (0.3)(0.2)(0.3) + (0.7)(0.8)(0.3) = 0.354$$

$$(iii) P(MMJJ) + P(JJMM) + P(MJMJ) + P(JMJM) + P(JMMJ) + P(MJMJ) \\ = (0.7)(0.2)(0.3)(0.8) + (0.3)(0.8)(0.7)(0.2) + (0.7)(0.8)(0.7)(0.8) + (0.3)(0.2)(0.3)(0.2) \\ + (0.3)(0.2)(0.7)(0.8) + (0.7)(0.8)(0.3)(0.2) = 0.4516$$

$$(iv) \frac{P(JMMM)}{P(JJM) + P(MMJ) + P(MJM)} = \frac{(0.3)(0.2)(0.7)(0.2)}{(0.3)(0.2)(0.7) + (0.7)(0.2)(0.3) + (0.7)(0.8)(0.7)} \\ = 0.176 \text{ (3s.f.)}$$

- 10 A coffee company claims that a new coffee product can lower cholesterol level by as much as 20 mg/dL. Over a period of time, a random sample of 60 volunteers drinks this new coffee product. The decrease in cholesterol level,  $x$  in mg/dL are summarised by

$$\sum x = 1020, \quad \sum x^2 = 21164$$

- (i) Find the unbiased estimates of the population mean and variance. [3]
- (ii) Test, at the 5% significance level, whether the mean decrease in cholesterol level is less than 20 mg/dL. [4]
- (iii) State, giving a reason, whether any assumptions about the population are needed in order for the test to be valid. [1]
- (iv) In another test, using the same data and at the 1% significance level, the hypotheses are as follows:

Null hypothesis: the population mean decrease in cholesterol level is equal to  $\mu_0$ ,

Alternative hypothesis: the population mean decrease in cholesterol level is not equal to  $\mu_0$ .

Given that the null hypothesis is rejected in favour of the alternative hypothesis, find the range of possible values of  $\mu_0$ . [4]

**Solution**

$$\sum x = 1020, \quad \sum x^2 = 21164$$

$$\bar{x} = \frac{\sum x}{n} = \frac{1020}{60} = 17$$

$$(i) \quad s^2 = \frac{1}{n-1} \left[ \sum x^2 - \frac{(\sum x)^2}{n} \right] = \frac{1}{59} \left[ 21164 - \frac{(1020)^2}{60} \right] = 64.814$$

$$(ii) \quad H_0: \mu = 20 \quad \text{vs} \quad H_1: \mu < 20$$

Perform a 1-tail test at 5% level of significance.

Under  $H_0$ ,  $\bar{X} \sim N\left(\mu_0, \frac{s^2}{n}\right)$  approximately, by Central Limit Theorem,

where  $\mu_0 = 20$ ,  $s = \sqrt{64.814}$ ,  $\bar{x} = 17$ ,  $n = 60$

Using a  $z$ -test,  $p$ -value = 0.00195 (3 s.f.)

**Conclusion:**

Since the  $p$ -value = 0.00195 < 0.05, we reject  $H_0$  and conclude that there is sufficient evidence, at 5% significance level, that the mean decrease in cholesterol level is less than 20.

(iii) No assumption is needed as  $n$  is large and the sample mean follows a normal distribution by Central Limit Theorem.

(iv) For  $H_0$  to be rejected,

$$P(\bar{X} \leq \bar{x}) \leq 0.005 \quad \text{OR} \quad P(\bar{X} \geq \bar{x}) \leq 0.005$$

$$\text{From GC, } P(Z \leq -2.57583) = 0.005 \quad P(Z \leq 2.57583) = 0.98$$

$$\text{Hence, } \frac{17 - \mu_0}{\sqrt{64.814/60}} \leq -2.57583 \quad \text{OR} \quad \frac{17 - \mu_0}{\sqrt{64.814/60}} \geq 2.57583$$

$$\text{Solving, } \mu_0 \geq 19.7 \text{ (3sf)} \quad \text{OR} \quad \mu_0 \leq 14.3 \text{ (3sf)}$$

- 11 For two years, a factory records the number of components,  $x$  thousands, it produces each quarter and the total cost,  $y$  thousands of dollars, of producing the components. The table below summarises the results.

$x$	15	17	13	21	16	22	14	18
$y$	29	35	27	43	34	41	30	36

- (i) Give a sketch of the scatter diagram for the data as shown on your calculator. [2]
- (ii) Calculate the product moment correlation coefficient, and comment on its value in relation to the scatter diagram. [2]
- (iii) Calculate the equation of the regression line of  $y$  on  $x$ , and draw this line on your scatter diagram.

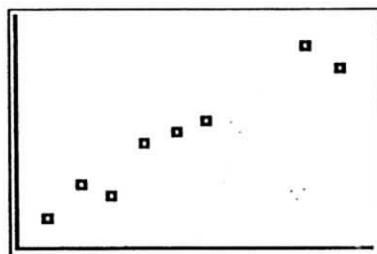
[2]



- (iv) Find the total cost of production (to the nearest thousand) if the number of components produced is 40000 for next three-month period. Comment on the validity of this prediction. [2]
- (v) For the next three-month period, the budget for the total cost of producing the components is \$38000. Using an appropriate equation of the regression line, estimate the number of components (to the nearest thousand) that will be produced. [2]
- (vi) The selling price of each component is \$2.05. Indicate on the same scatter diagram the additional line you have to include to estimate the number of components it has to produce so that the total cost and total income balances. Write down the number of components (correct to 2 s.f.) and give a brief interpretation what this value means. [3]

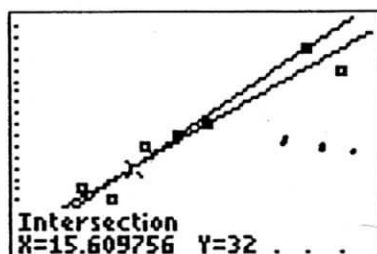
### Solution

(i)



- (ii)  $r = 0.969$ . This value is consistent with the scatter diagram which shows that  $x$  and  $y$  have a strong positive correlation.
- (iii)  $y = 5.3333 + 1.7083x$   
 $y = 5.33 + 1.708x$  (3 s.f.)
- (iv) When  $x = 40$ ,  
 $y = 5.3333 + 1.7083(40) = 73.6653 = 73.7$  (3 s.f.)  
 This value may not be accurate because we are extrapolating. The linear relation may not exist outside the range of values given.
- (v)  $x = 0.549y - 1.89$   
 When  $y = 38$ ,  
 $x = 0.549(38) - 1.89 = 19$   
 The number of components is 19000.

(vi)



The number of components (correct to 2 significant figures) it has to produce so that the total cost and total income balances = 16000.  
 The factory makes a profit if the output is above this level. If the output is below this level, it suffers a loss.

- 12 (a) The random variable  $X$  has the distribution  $N(\mu, \sigma^2)$ . Given that,  $P(X < 65) = P(X > 95) = 0.345$  find the values of  $\mu$  and  $\sigma$ . [4]
- (b) The total playing time of a football match may be modelled using a normal distribution with mean 94 minutes and standard deviation 1.3 minutes.
- (i) Two football matches are chosen at random. Find the probability that one has a total playing time of more than 95 minutes and one has a total playing time of less than 95 minutes. [3]
- (ii) Using a suitable approximation, find the probability that, in fifty football matches, there is at least 10 matches in which the total playing time is more than 95 minutes. [4]
- (iii) Find the probability that, in a random sample of five football matches, the average total playing time is more than 95 minutes. [3]

**Solution**

(a)  $P(X < 65) = P(X > 95) = 0.345$

$$\mu = \frac{65 + 95}{2} = 80$$

$$P(X < 65) = 0.345$$

$$P\left(Z < \frac{65 - 80}{\sigma}\right) = 0.345$$

$$\frac{-15}{\sigma} = -0.39886$$

$$\sigma = 37.6 \text{ (3 s.f.)}$$

(b) Let  $X$  be the total playing time of a football match.

$$X \sim N(94, 1.3^2)$$

(i)

$$\begin{aligned} 2P(X < 95)P(X > 95) &= 2(0.77912)(1 - 0.77912) \\ &= 0.344 \text{ (3s.f.)} \end{aligned}$$

(ii) Let  $Y$  be the number of matches out of 50 where the total playing time is more than 95 minutes.

$$Y \sim B(50, 0.22088)$$

Since  $np > 5$  and  $n(1-p) > 5$  and  $n$  is large,

$$Y \sim N(11.044, 8.6046)$$

$$\begin{aligned} P(Y \geq 10) &= P(Y \geq 10.5) \text{ by continuity correction} \\ &= 0.701 \text{ (3s.f.)} \end{aligned}$$

(iii) Let  $W$  be the average total playing time of a random sample of 5 football matches.

$$W = \frac{X_1 + X_2 + X_3 + X_4}{5} \sim N\left(94, \frac{1.3^2}{5}\right)$$

$$P(W > 95) = 0.0427 \text{ (3s.f.)}$$