Preliminary Examination
Higher 1

MATHEMATICS 8864

Friday **8am – 11am**

16 September 2016 **3 hours**

READ THESE INSTRUCTIONS FIRST

Write your name and CT group on all the work you hand in.
Write in dark blue or black pen on both sides of the paper.
You may use a soft pencil for any diagrams or graphs.
Do not use staples, paper clips, highlighters, glue or correction fluid.

Answer **all** the questions.
Give non-exact numerical answers correct to 3 significant figures, or 1 decimal place in the case of angles in degrees, unless a different level of accuracy is specified in the question.
You are expected to use a graphic calculator.
Unsupported answers from a graphic calculator are allowed unless a question specifically states otherwise.
Where unsupported answers from a graphic calculator are not allowed in a question, you are required to present the mathematical steps using mathematical notations and not calculator commands.
You are reminded of the need for clear presentation in your answers.

At the end of the examination, fasten all your work securely together.
The number of marks is given in brackets [ ] at the end of each question or part question.

Section A: Pure Mathematics (35 marks)

1 The normal to the curve $y=\dfrac{ax+b}{\sqrt{x}}$ at $x=1$ is $2x-y=2$. Find the values of a and b. [4]

2 Given that $\int_1^n 2^{-\sqrt{x}}\,dx = 2.7551$ and $\int_2^n \left(2^{-\sqrt{x}}-\dfrac{1}{2x}\right) dx = 1.1727$, correct to 4 decimal places, find the value of n, giving your answer correct to the nearest integer. (Answer obtained by trial and improvement from a calculator will obtain no marks.) [5]

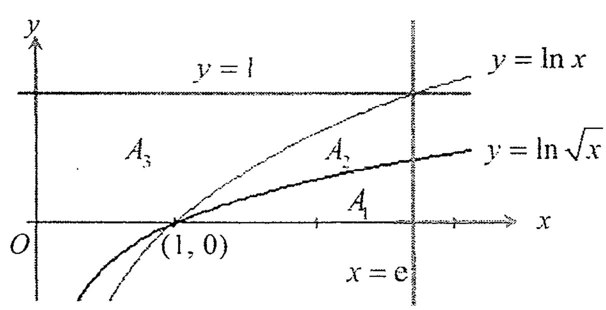
3 Sketch the curve $y=k+\dfrac{1}{x+1}$, where k is a positive constant, showing clearly the equations of any asymptotes and coordinates of any points of intersection with the axes. [2]

Show that the equation of the tangent to the curve at the point where $x=p$ is given by

$$y=-\frac{x-p}{(p+1)^2}+\frac{1}{p+1}+k.$$ [2]

Use this equation to explain why the curve will always have two tangents that pass through the point (2, 0). [4]

4 The region bounded by the curve $y=\ln x$, the $x-$axis, the lines $x=1$ and $x=\mathrm{e}$ is divided by the curve $y=\ln\sqrt{x}$ into two regions with areas A_1 and A_2 (see diagram). The region bounded by the curve $y=\ln x$, the axes, the lines $y=1$ and $x=\mathrm{e}$ has area A_3 (see diagram).

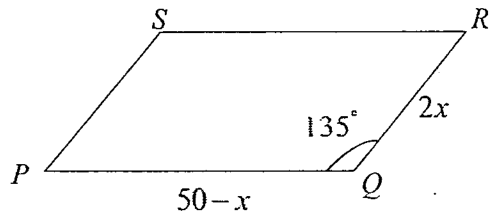


(i) Express each of A_1 and A_2 as integral. Without performing any integration, show that $A_1 = A_2$. [4]

(ii) Find the exact value of A_3. [2]

(iii) Using your results in parts (i) and (ii), find the value of A_1. [3]

5 The diagram shows the parallelogram $PQRS$ whose side $PQ = (50 - x)$ cm, $QR = 2x$ cm and $\angle PQR = 135°$.



Show that the area, A cm^2, of the parallelogram is given by

$$A = \sqrt{2}x(50 - x).$$ [3]

Using differentiation, find the exact maximum value of A. [3]

Find the largest integer value of x for which the value of A is greater than $100\sqrt{2}$. [3]

Section B: Statistics (60 marks)

6 In a particular shopping mall, the directory classifies all the shops into 4 main categories: fashion, health and wellness, food and beverage, and services. A manager of the shopping mall wishes to sample 30 shop owners to find their opinions of the mall facilities.

(i) Describe, in the context of the question, how a stratified sample of 30 shop owners might be obtained. [2]

The manager also wishes to get the opinions of mall shoppers.

(ii) Give a reason why it would be difficult to use a stratified sample. [1]

(iii) Name a sampling method that would be appropriate and explain whether the manager would be able to obtain a random sample. [2]

7 In a probability experiment, two boxes have the following contents.

Box A contains 4 red balls and 2 blue balls.
Box B contains 4 red balls, 3 blue balls and 1 white ball.

A fair die is tossed. If the score on the die is a multiple of 3, Box A is selected; otherwise, Box B is selected.

(a) One ball is taken from the selected box at random and the colour of the ball is noted.

(i) Draw a tree diagram to represent this situation. [2]

(ii) Given that a blue ball is chosen, find the probability that it was from Box A. [3]

(b) Instead, three balls are taken from the selected box, at random and without replacement. Find the probability that all three balls are of the same colour. [3]

[Turn over

8 A chicken farm produces a large number of eggs. The eggs are randomly packed into cartons with ten eggs in each carton. The probability that none of the eggs in a carton is spoiled is 0.48. Show that the probability that exactly one egg is spoiled is 0.3656, correct to 4 decimal places. [4]

Three cartons of eggs are selected at random. Find the probability that all of them have at least twice the expected number of spoiled eggs in each carton. [3]

Thirty cartons are chosen at random. Using a suitable approximation, find the probability that there are more than 20 cartons but at most 25 cartons with no spoiled egg in each carton. [3]

9 A machine produces nails whose lengths are normally distributed with mean μ cm and standard deviation σ cm. A sample of 60 nails is selected at random and each length, x cm, is measured. Given that $\sum(x-3)=45$ and $\sum(x-3)^2=425$, calculate unbiased estimates of μ and σ^2. [3]

When the machine is working correctly, $\mu=3$, but occasionally the machine goes wrong, in which case $\mu>3$. Based on the above sample, determine whether the machine is working correctly at the 5% significance level. State, with a reason, whether the test is valid if the lengths of nails are not normally distributed. [5]

The machine is subsequently adjusted and $\sigma=0.1$. In order to determine whether the machine is working correctly, the lengths of nails in a random sample of size n are measured and the sample mean $\bar{x}$ is found. If $\bar{x}>a$, then it is concluded that the machine has gone wrong. Find, in terms of n, the value of a, if the probability of concluding that the machine has gone wrong when in fact it is working correctly is 0.01. [3]

10 The masses of Butternut pumpkins have a normal distribution for which the mean mass is eight times the standard deviation. Find the percentage of Butternut pumpkins such that its mass is at least 0.9 of the mean mass. [3]

It is now given that the mean mass of the Butternut pumpkins is 0.8 kg.

(i) Five Butternut pumpkins are chosen at random. Find the probability that the mean mass of these pumpkins is at most 0.9 kg. [3]

(ii) The farmer also grows Japanese pumpkins. The masses of these pumpkins have a normal distribution with mean 1 kg and standard deviation 0.15 kg. Butternut pumpkins cost \$2.50 per kg and Japanese pumpkins cost \$1.67 per kg. Joel bought two Butternut pumpkins and three Japanese pumpkins. Find the probability that he spends between \$8.50 and \$9.50. [5]

11 A city council attempted to reduce traffic congestion by introducing a congestion charge when a car drives into the city. The charge was set at \$4.00 per entry for the first year and was then subsequently increased by \$2.00 each year. For each of the first eight years, the council recorded the average number of vehicles entering the city centre per day. The results are shown in the table.

Year, w	1	2	3	4	5	6	7	8
Charge, \$$x$	4	6	8	10	12	14	16	18
Average number of vehicles per day, y million	2.4	2.5	2.2	2.3	2.0	1.8	1.7	1.5

(i) Draw a scatter diagram for x and y as shown in your calculator. [2]

(ii) Calculate the product moment correlation coefficient between x and y, and comment on its value in the context of the data. [3]

(iii) State, giving a reason, which of the least squares regression lines, y on x or x on y, should be used to express a possible linear relation between x and y. [2]

(iv) Calculate the equation of the regression line chosen in part **(iii)**. [1]

(v) From the equation of the regression line, state the average number of vehicles which will enter the city centre per day when there is no congestion charge. Explain why this value does not necessarily give the expected average number of vehicles entering the city centre per day when there is no congestion charge. [2]

(vi) Write down the relationship between x and w in the form $x = k + kw$, where k is a constant. Without further calculations, state, giving reasons, the value of the product moment correlation coefficient between y and w. [4]

Question 1

$$y=\frac{ax+b}{\sqrt{x}}=ax^{1/2}+bx^{-1/2}$$

$$\frac{dy}{dx}=\frac{a}{2\sqrt{x}}-\frac{b}{2x^{3/2}}$$

$\Rightarrow$ at $x=1, \frac{dy}{dx}=\frac{a}{2}-\frac{b}{2}$

Given equation of normal at $x=1$ is $y=2x-2$

$\Rightarrow \frac{dy}{dx}=-\frac{1}{2}$ and $y=0$ when $x=1$.

$\therefore a-b=-1 \quad \cdots(1)$

$(1,0)$ is on the curve $\Rightarrow a+b=0 \quad \cdots(2)$

Solving (1) and (2), we have $a=-\frac{1}{2}, b=\frac{1}{2}$.

Question 2

$$\int_2^n \left(2^{-\sqrt{x}}-\frac{1}{2x}\right) dx = 1.1727$$

$$\Rightarrow \int_1^n 2^{-\sqrt{x}}\,dx-\int_1^2 2^{-\sqrt{x}}\,dx-\int_2^n \frac{1}{2x}dx=1.1727$$

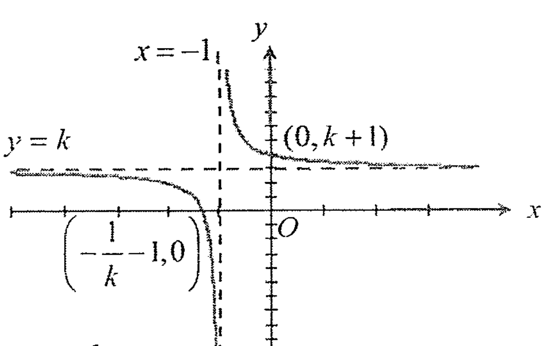
$\int_1^n 2^{-\sqrt{x}}\ dx = 2.7551$ and $\Rightarrow 2.7551-0.431062-\frac{1}{2}\ln x\Big|_2^n=1.1727$

$\Rightarrow \ln n-\ln 2=2.302676$

$\Rightarrow \ln n \approx 2.9958$

$\Rightarrow n \approx e^{2.9958} \approx 20.001 = 20$ (nearest integer)

Question 3



$$y=k+\frac{1}{x+1}$$

$$\frac{dy}{dx}=-\frac{1}{(x+1)^2}$$

$\therefore$ equation of tangent at $x=p$ is $y-\left(k+\frac{1}{p+1}\right)=-\frac{1}{(p+1)^2}(x-p)$

$$y=-\frac{x-p}{(p+1)^2}+k+\frac{1}{p+1}$$

Tangent passes through (2, 0) $\Rightarrow$ $0=-\frac{2-p}{(p+1)^2}+k+\frac{1}{p+1}$

$$-\frac{1}{(p+1)^2}(2-p)+k+\frac{1}{p+1}=0$$
$$-(2-p)+k(p+1)^2+(p+1)=0$$
$$-2+p+kp^2+2kp+k+p+1=0$$
$$kp^2+(2+2k)p+k-1=0$$

discriminant $=(2k+2)^2-4k(k-1)$
$=12k+4>0$ for all $k>0$

Hence, there are 2 distinct values of p.

This implies there will always be 2 tangents to the curve that passes through (2, 0).

Question 4

(i)

$A_1=\int_1^e \ln\sqrt{x}\ \mathrm{d}x;\quad A_2=\int_1^e\left(\ln x-\ln\sqrt{x}\right)\ \mathrm{d}x$

Since $\ln\sqrt{x}=\frac{1}{2}\ln x$, $A_1=\frac{1}{2}\int_1^e \ln x\ \mathrm{d}x$ and

$A_2=\int_1^e\left(\ln x-\frac{1}{2}\ln x\right)\ \mathrm{d}x=\frac{1}{2}\int_1^e \ln x\ \mathrm{d}x=A_1$ (shown)

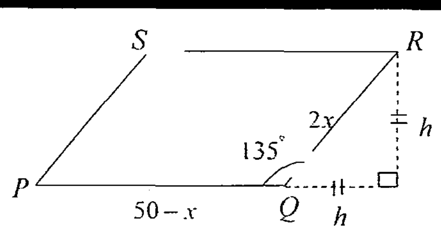
(ii)

$$A_3=\int_0^1 x\ \mathrm{d}y$$
$$=\int_0^1 \mathrm{e}^y\ \mathrm{d}y$$
$$=\mathrm{e}^y\Big|_0^1=\mathrm{e}-1$$

(iii)

$$A_1+A_2+A_3=\mathrm{e}\times 1$$
$$2A_1+\mathrm{e}-1=\mathrm{e}$$
$$\therefore A_1=\frac{1}{2}$$

Question 5



$2h^2=(2x)^2$

$\therefore h=\sqrt{2}x$

$A=\text{base}\times\text{height}=(50-x)\sqrt{2}x=\sqrt{2}x(50-x)$ (shown)

$A=\sqrt{2}x(50-x)=\sqrt{2}(50x-x^2)$ $\frac{\mathrm{d}A}{\mathrm{d}x}=\sqrt{2}(50-2x)=0\Rightarrow x=25$

Sign test on $\frac{dA}{dx}$:

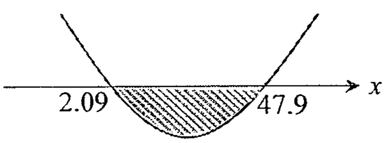
x	25^-	25	25^+
Sign of $\frac{dA}{dx}$	+	0	-
Slope	/	—	\

When $x = 25, A = 625\sqrt{2}$ is a maximum value.

$A > 100\sqrt{2} \Rightarrow 50x - x^2 > 100$

$x^2 - 50x + 100 < 0$

From GC, $x^2 - 50x + 100 = 0$ when $x = 2.09$ or 47.9



$2.09 < x < 47.9$

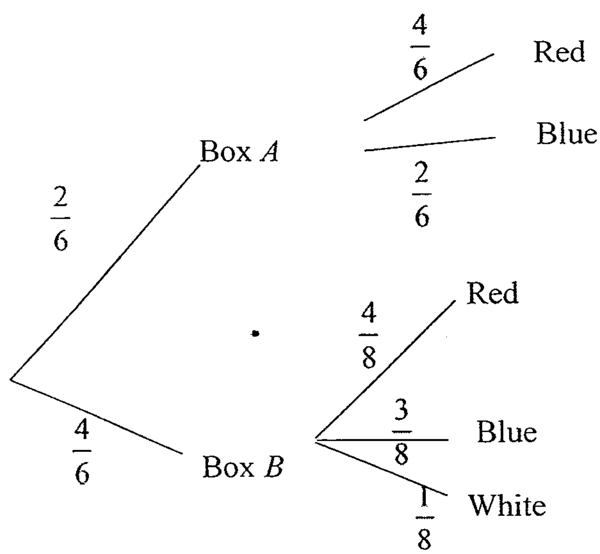
The largest value of x is 47.

Section B

Question 6
(i) Obtain a list of all the shops in each category in the shopping mall from the directory. Use a random sampling to select from each category a number which is proportional to the number of shops in the category. For example, if there are 60 shops in the fashion category, select 10 fashion shops. A stratified sample of 10 fashion shops can be obtained by using a random number generator to obtain 10 distinct numbers and then select the 10 shops which correspond to the numbers generated. This procedure is repeated for the remaining 3 categories.
(ii) It is difficult to obtain the sampling frame i.e. the number of shoppers in the shopping mall, thus, it would be difficult to use a stratified sampling.
(iii) Quota sampling. The manager would not be able to obtain a random sample as the manager might select shoppers based on his preference. Hence not everyone has an equal chance of being selected.

Question 7

(i) (a)



(ii) $P(\text{box A} \mid \text{blue}) = \dfrac{P(\text{box A} \cap \text{blue})}{P(\text{blue})}$

$$= \frac{\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)}{\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)+\left(\frac{4}{6}\right)\left(\frac{3}{8}\right)} = \frac{\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)}{\frac{13}{36}} = \frac{4}{13}$$

Required probability $=\left(\frac{2}{6}\right)\left(\frac{4}{6}\right)\left(\frac{3}{5}\right)\left(\frac{2}{4}\right)+\left(\frac{4}{6}\right)\left(\frac{4}{8}\right)\left(\frac{3}{7}\right)\left(\frac{2}{6}\right)+\left(\frac{4}{6}\right)\left(\frac{3}{8}\right)\left(\frac{2}{7}\right)\left(\frac{1}{6}\right)=\frac{53}{420}$

Question 8

Let X be the number of spoiled eggs in a carton. $X \sim B(10, p)$.

$P(X=0)=0.48$

$\binom{10}{0} p^0 (1-p)^{10} = 0.48$

$(1-p)^{10} = 0.48$

$p = 0.07076807$

$P(X=1)$

$= \binom{10}{1}(0.07076807)^1 (1-0.07076807)^9$

≈ 0.365556

$= 0.3656$ (4 d.p)

$X \sim B(10,\ 0.07076807)$

$2E(X) = 2(10 \times 0.07076807) = 1.41536$

$P(X \geq 1.41536)$
$= P(X \geq 2)$
$= 1 - P(X \leq 1)$
$= 0.1544435$
Required prob = $(0.1544435)^3 = 0.0036839 = 0.00368$(3 s.f.)

Let Y be the number of cartons, out of 30, with no spoiled eggs.
$Y \sim B(30, 0.48)$
$np = 14.4 > 5$; $n(1-p) = 15.6$; $np(1-p) = 7.488$
Since $np > 5$ and $n(1-p) > 5$, $Y \sim N(14.4, 7.488)$ approximately.
$P(20 < Y \leq 25)$
$= P(20.5 < Y < 25.5)$
$= 0.0128757$
$= 0.0129$ (3 s.f)

Question 9

Let $w = x - 3$, then, $\sum w = 45$, $\sum w^2 = 425$ and

Unbiased estimate of μ is $\bar{x} = 3 + \bar{w} = 3 + \dfrac{45}{60} = 3.75$

Unbiased estimate of σ^2 is $s_x^2 = s_w^2 = \dfrac{1}{n-1}\left[\sum x^2 - \dfrac{\left(\sum x\right)^2}{n}\right]$

$$= \frac{1}{59}\left[425 - \frac{(45)^2}{60}\right] = 6.631356 \approx 6.63$$

$H_0 : \mu = 3$
$H_1 : \mu > 3$
Level of significance: 5%
Test Statistic: Since $n = 60$ is sufficiently large, so s_x^2 is a good estimate of σ^2 and by Central Limit Theorem, $\bar{X}$ is approximately normal.

$\therefore \bar{X} \sim N\left(3, \dfrac{s_x^2}{n}\right)$ approximately when H_0 is true.

$\therefore Z = \dfrac{\bar{X} - 3}{s_x / \sqrt{n}} \sim N(0;1)$.

Rejection region: $z \geq 1.6449$
Computation: $\bar{x} = 3.75$, $n = 60$, $s_x = \sqrt{6.631356}$
$\therefore z = 2.25598 \approx 2.26$
$p -$ value $= 0.0120358 \approx 0.0120$

Conclusion: Since $p -$ value $= 0.0120 < 0.05$, $\therefore H_0$ is rejected at 5% significance level. Hence there is sufficient evidence to conclude that the machine is not working correctly at the 5% significance level.

Yes. The test is valid since $n = 60$ is sufficiently large, by Central Limit Theorem, the sample mean length of a nail ($\bar{X}$) is approximately normally distributed.

If $\sigma = 0.1$, then when H₀ is true, $\bar{X} \sim N\left(3, \frac{\sigma^2}{n}\right)$

P (presuming machine has gone wrong when in fact it is working correctly) = 0.01

$\therefore P\left(\bar{X} > a \text{ when } H_0 \text{ is true}\right) = 0.01$

$\Rightarrow P\left(Z > \frac{a-3}{0.1/\sqrt{n}}\right) = 0.01$

From GC: $P\left(Z > 2.32635\right) = 0.01$

$\therefore \frac{a-3}{0.1/\sqrt{n}} = 2.32635$

$\Rightarrow a = 3 + 2.32635\left(\frac{0.1}{\sqrt{n}}\right)$

$\approx 3 + \frac{0.233}{\sqrt{n}}$

Question 10

Let B kg be the mass of a randomly chosen Butternut pumpkin.

$B \sim N(\mu, \left(\frac{\mu}{8}\right)^2)$

$P(B \geq 0.9\mu)$

$= P\left(Z \geq \frac{0.9\mu - \mu}{\mu / 8}\right)$

$= P(Z \geq -0.8)$

$= 0.7881447$

$= 0.788$ (3 s.f.)

78.8% of the Butternut pumpkins have mass at least 0.9 of the mean mass.

$B \sim N(0.8, 0.1^2)$

$\bar{B} = \frac{B_1 + B_2 + \ldots + B_5}{5}$

$\bar{B} \sim N\left(0.8, \frac{0.1^2}{5}\right)$

$P\left(\bar{B} \leq 0.9\right) = 0.987326 = 0.987$ (3 s.f.)

Let J kg be the mass of a randomly chosen Japanese pumpkin.

$J \sim N(1, 0.15^2)$

Cost of one Butternut pumpkin, $X = 2.5B$

$E(X) = 2.5(0.8) = 2$ $\quad$ $Var(X) = 2.5^2(0.1^2) = 0.0625$

$X \sim N(2, 0.0625)$

Cost of one Japanese pumpkin, $Y = 1.67J$

$E(Y) = 1.67(1) = 1.67$

$Var(Y) = 1.67^2(0.15^2) = 0.06275025$

$Y \sim N(1.67, 0.06275025)$

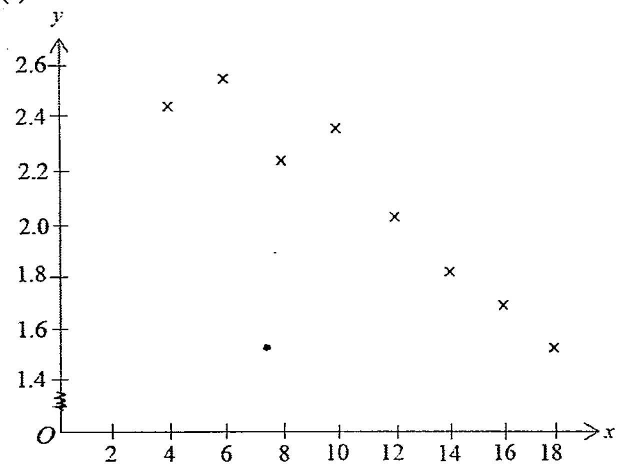
$C = X_1 + X_2 + Y_1 + Y_2 + Y_3 \sim N\left(9.01, 0.31325075\right)$

$P\left(8.50 < C < 9.50\right) = 0.628$ (3 s.f.)

Assume that the masses of all the pumpkins are independent of one other.

Question 11

(i)



(ii) $r = -0.960$ (to 3 s.f.)

Since $r = -0.960$ is close to -1 and the points seem to lie close to a straight line with negative gradient are indications of a strong negative linear relationship between the charge (x) and the average number of vehicles entering the city centre per day (y). This means that as x increases, y tends to decrease at a constant rate.

(iii) Since the values of x are fixed (or controlled), hence x is an independent variable. So the least squares regression lines, y on x should be used.

(iv) The equation of the regression line of y on x is $y = 2.8226 - 0.070238x$ i.e $y = 2.82 - 0.0702x$ (to 3 s.f.)

(v) When there is no congestion charge i.e. $x = 0$, so the average number of vehicles which will enter the city centre per day is 2820000 (or 2.82 million).
Since $x = 0$ is out of the range of the data $4 \le x \le 18$, $\therefore$ 2.82 million does not necessarily give the expected average number of vehicles entering the city centre per day.

(vi) $x = 2 + 2w$
$\therefore r = -0.960$ (same as the product moment correlation coefficient between x and y found in **(i)**) as the product moment correlation coefficient is unaffected by change of scale and location.