JC2 Preliminary Examination

Higher 1

H1 Mathematics **8864/01**

Paper 1 **13 September 2016**

3 Hours

Additional Materials: Writing paper

List of Formulae (MF 15)

READ THESE INSTRUCTIONS FIRST

Write your name and civics 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 graphing calculator.
Unsupported answers from a graphing calculator are allowed unless a question specifically states otherwise.
Where unsupported answers from a graphing 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.

This document consists of **9** printed pages.

Section A: Pure Mathematics [35 marks]

1 **(a)** Find $\int \frac{e^{2x}(x-1)+2}{x-1}\,dx$. [3]

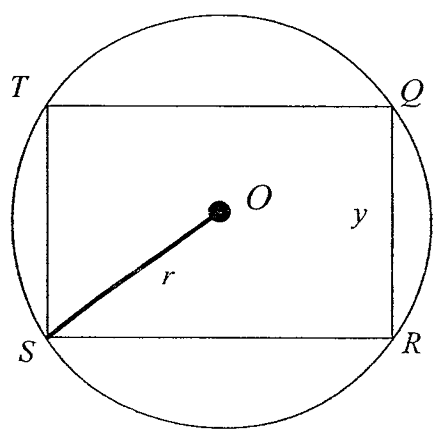
(b) Find the exact value of $\int_0^a (3x-a)^2\,dx$ in terms of a where $a>0$. [2]

2 A curve has equation $y=-2x^3+13x^2-13x-10$.

(i) Sketch the curve, stating clearly the coordinates of all points of intersection with the axes and the stationary points. [2]

(ii) Solve the inequality $-2x^3+13x^2-13x-10<0$. Hence find the exact solutions of the inequality $-2e^{3x}+13e^{2x}-13e^{x}-10<0$. [4]

3



The diagram shows a rectangle $QRST$ inscribed in a circle, centre at O, with fixed radius r cm. The four corners of the rectangle lie on the circumference of the circle. Given that $QR=y$ cm, show that the perimeter, P cm, of the rectangle $QRST$ is given by $P=2y+2\sqrt{(4r^2-y^2)}$. [2]

Without using a calculator, find the exact maximum value of P, in terms of r, as y varies, justifying that this value is a maximum. [5]

4 **(i)** On the same diagram, sketch the graphs of $y = 3\sqrt{2^x}$ and $y = \dfrac{3}{x+1}$, stating clearly the equations of the asymptotes and the coordinates of the point of intersection of the two graphs. [4]

(ii) Hence find the range of values of x such that $\sqrt{2^x} - \dfrac{1}{x+1} \geq 0$. [2]

(iii) Write down, as an integral, an expression for the area of the region bounded by $y = 3\sqrt{2^x}$, $y = \dfrac{3}{x+1}$ and the line $x = 3$. Evaluate this integral, giving your answer correct to 3 decimal places. [2]

5 A curve has equation $y = \dfrac{1}{x}, x \in \square, x \neq 0$.

The point A on the curve has coordinates $\left(a, \dfrac{1}{a}\right), a \in \square, a \neq 0$.

(i) Find the equation of the tangent to the curve at A, in terms of a. [3]

(ii) The tangent at A intersects the x-axis at P and the y-axis at Q. Show that the area of triangle OPQ is independent of a, where O is the origin. [2]

(iii) Write down the equation of the normal to the curve at A in terms of a. Hence find the value(s) of a if the normal passes through the origin. [4]

Section B: Statistics [60 marks]

6 A home furnishing company has a total of 1000 employees who are split into the following departments: Manufacturing Department, Assembly Department and Administration Department. Each employee works in only one department. The number of employees in each department of each level of seniority is summarized in the table below.

	Manufacturing Department	Assembly Department	Administration Department
Junior Level	300	240	100
Senior Level	150	60	150

A researcher wants to investigate the employees' number of working hours per week. He decides to do this by choosing a stratified random sample of size 100.

(i) Describe how the researcher might choose his sample with three strata from the combined level-groups. [2]

(ii) State one advantage that stratified sampling would have compared to random sampling in this context, and state how a better stratified sample of size 100 could have been achieved, using the data in the above table. [2]

7 According to ABC driving school, the probability of a randomly chosen student passing the basic driving theory test at the first attempt is p, where $0 < p < 1$. Students who fail the test at the first attempt are allowed to take a second attempt. The probability of passing the basic driving theory test during the second attempt is 1.2 times of the probability of passing at the first attempt.

(i) Show that the probability that a randomly chosen student from ABC driving school passes the basic driving theory test during the first or second attempt is $2.2p - 1.2p^2$. Hence find the smallest value of p such that this probability is at least 0.99. Give your answer correct to one decimal place. [3]

(ii) Find the value of p such that the probability that a randomly chosen student passes the examination on the first attempt given that the student passes is 0.7. [3]

(iii) It is given now that p is 0.7. Two students are randomly chosen. Find the probability that one passes the examination on the first attempt and the other passes the examination on the second attempt. [2]

8 Melons and durians are sold by weight. The masses, in kg, of melons and durians are modelled as having independent normal distributions with means and standard deviations as shown in the table.

	Mean mass	Standard deviation
Melons	2.7	0.3
Durians	1.5	0.1

Melons are sold at \$5 per kg and durians at \$10 per kg.

(i) Find the probability of the event that 3 randomly chosen melons have a total selling price exceeding \$41 and a randomly chosen durian has a selling price exceeding \$17. [4]

(ii) Find the probability that total selling price of 3 randomly chosen melons and a randomly chosen durian exceeds \$58. [2]

(iii) Explain why the answer to part **(ii)** is greater than the answer to part **(i)**. [1]

9 **(a)** A random sample of eight pairs of values of variables x and y is obtained. Draw a sketch of a possible scatter diagram of the data for each of the following cases:

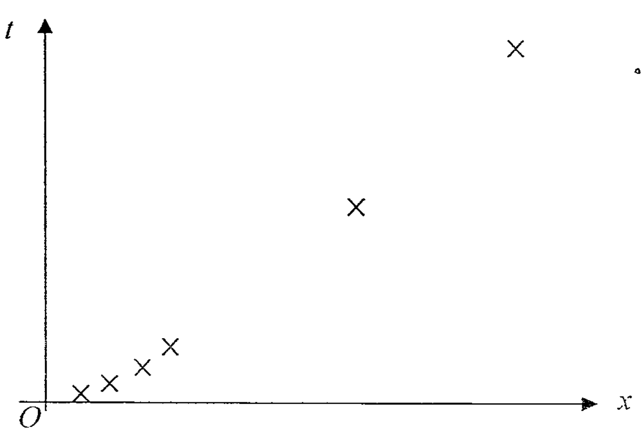
(i) the product moment correlation coefficient is approximately zero, [1]

(ii) the product moment correlation coefficient is approximately 0.99. [1]

(b) A research is conducted to investigate how the period of a planet's orbit around the Sun is related by its furthest distance from the Sun. The following table gives the furthest distance, x million kilometers, of the respective planet from the Sun and the period, t years, of the planet's orbit around the Sun.

Planet	Mercury	Venus	Earth	Mars	Jupiter	Saturn
x	58.3	108	150	227	778	1429
t	0.24	0.62	1.00	1.88	11.9	29.5

It is given that the value of the product moment correlation coefficient for this data is 0.993, correct to 3 decimal places. The scatter diagram for the data is shown below.



Explain why it is advisable to plot a scatter diagram before interpreting the correlation coefficient calculated. [1]

(c) A researcher conducted an experiment to investigate the fish population for different plankton concentration in the fish farms. The following table gives the concentration of plankton, x, and fish population, y. Both x and y are measured in appropriate units. A statistician is given the following data to analyse.

x	49	51	54	58	62	64	66	70
y	90	88	84	83	82	79	77	78

(i) Draw the scatter diagram for these values, labelling the axes clearly. [2]

(ii) Calculate the product moment correlation coefficient. [1]

(iii) Calculate the least squares regression line of y on x and use it to estimate the value of y when $x = 80$. Comment on the reliability of the estimate. [3]

10 **(a)** In a poultry farm, there is a large number of chickens. The mean mass for all the chickens is recorded as 1.5 kg and the variance is 0.68 kg^2. Give a reason why a normal distribution, with this mean and variance, would not be a suitable model.[1]

(b) In a dairy farm, there is a large number of cows. The daily milk production of each cow, in litres, is normally distributed with mean μ and standard deviation σ. The probability that the daily milk production of a randomly chosen cow is more than 35 litres is 0.25. The probability that the average daily milk production of 16 randomly chosen cows is at most 25 litres is 0.089.

(i) Show that $\mu = 28.3$ and $\sigma = 9.89$ correct to 3 significant figures. [6]

(ii) A cow is labelled grade A if it produces more than 35 litres of milk daily. Find the probability that in a new sample of 20 cows, the 20th cow selected is the fourth grade A cow. [3]

11 The mass of a tin of a particular type of milk powder is claimed to have a mass of 2kg. A food regulator takes a random sample of 100 tins and the mass, x kg, of each tin of milk powder is measured. The results are summarised by

$$\sum x = 212.8 \qquad \sum x^2 = 490.65$$

A second random sample of 50 tins is taken and the mass, y kg, of each tin of milk powder is measured, with results summarised by

$$\sum y = 102.6 \qquad \sum y^2 = 230.125$$

Both samples are combined into a single sample.

(i) Calculate unbiased estimates of the population mean and variance. [3]

(ii) The population mean mass is denoted by μ kg. The null hypothesis of $\mu = 2$ is to be tested against the alternative hypothesis of $\mu > 2$. Carry out a test at the 5% level of significance and state the meaning of the p-value in context of the question. [4]

(iii) With brief explanation, deduce the conclusion of the test if the null hypothesis of $\mu = 2$ is to be tested against the alternative hypothesis of $\mu \neq 2$ at the 5% level of significance. [1]

(iv) Suppose the null hypothesis is now given as $\mu = \mu_0$ against the alternative hypothesis of $\mu > \mu_0$. Use an algebraic method to find the range of values of μ_0 for which the null hypothesis would be rejected at the 5% level of significance. [3]

12 A bakery produces two kinds of cookies, plain and chocolate. The cookies are sold in packs of n, where n is a positive integer. The probability that a randomly chosen cookie is a chocolate cookie is constant. Each pack has a random selection of the two kinds of cookies.

(a) For the case where $n = 12$, the mean number of chocolate cookies in a randomly chosen pack is 5.4,

(i) find the probability that a pack chosen at random has exactly 3 chocolate cookies, [2]

(ii) find the most likely number of chocolate cookies in a randomly chosen pack. [2]

(b) It is found that the probability that a randomly chosen cookie is a chocolate cookie is 0.3.

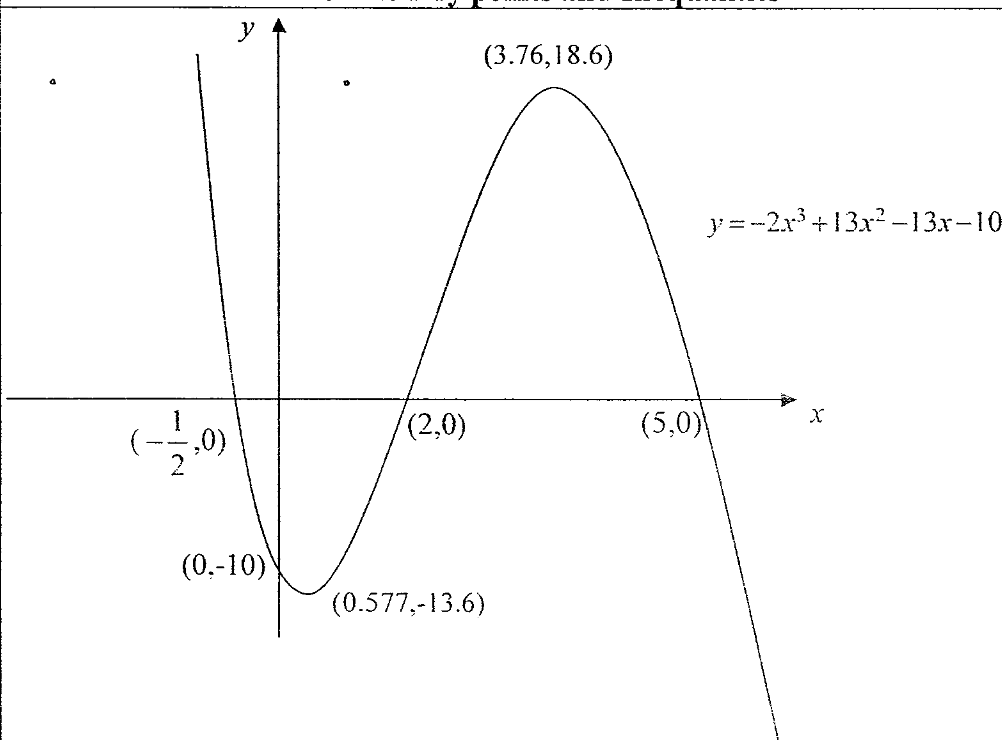
(i) For a pack of n cookies, the bakery wants to ensure that the probability that a pack contains at most 8 chocolate cookies does not exceed 0.9. Find the least value of n. [3]

(ii) A customer buys 10 packs of 12 cookies each. Use a suitable approximation to estimate the probability that there are more than 30 chocolate cookies in these 10 packs. You should state the mean and variance of any distribution that you use. [4]

END OF PAPER

2016 H1 MATH (8864/01) JC 2 PRELIM EXAMINATION – MARKING SCHEME

Qn	Solution
1	**Techniques of Integration**
(a)	$\int \frac{e^{2x}(x-1)+2}{x-1}\ dx$ $=\int\left(e^{2x}+\frac{2}{x-1}\right)\ dx$ $=\frac{1}{2}e^{2x}+2\ln\lvert x-1\rvert+c$
(b)	$\int_0^a (3x-a)^2 dx$ $=\left[\frac{(3x-a)^3}{3(3)}\right]_0^a$ $=\frac{1}{9}(8a^3+a^3)$ $=a^3$

Qn	Solution
2	**Stationary points and Inequalities**
(i)	
(ii)	For $-2x^3+13x^2-13x-10<0$. $-\frac{1}{2}<x<2$ or $x>5$ For $-2e^{3x}+13e^{2x}-13e^{x}-10<0$ Replace x with e^x in earlier part,

	$-\frac{1}{2} < e^x < 2$ or $e^x > 5$ Since $e^x > 0$, $0 < e^x < 2$ or $e^x > 5$ $\therefore\ x < \ln 2$ or $x > \ln 5$

Qn	Solution
3	**Application of Differentiation**

(i)

Using Pythagoras Theorem,

$RS = \sqrt{(2r)^2 - y^2}$

$\therefore P = 2QR + 2RS$

$\quad = 2y + 2\sqrt{4r^2 - y^2}$

When P attains maximum value,

$$\frac{dP}{dy} = 2 + 2\left(\frac{1}{2}\right)\left(4r^2 - y^2\right)^{-\frac{1}{2}}(-2y) = 0$$

$$\left(4r^2 - y^2\right)^{-\frac{1}{2}} y = 1$$

$$y = \left(4r^2 - y^2\right)^{\frac{1}{2}}$$

$$y^2 = 4r^2 - y^2$$

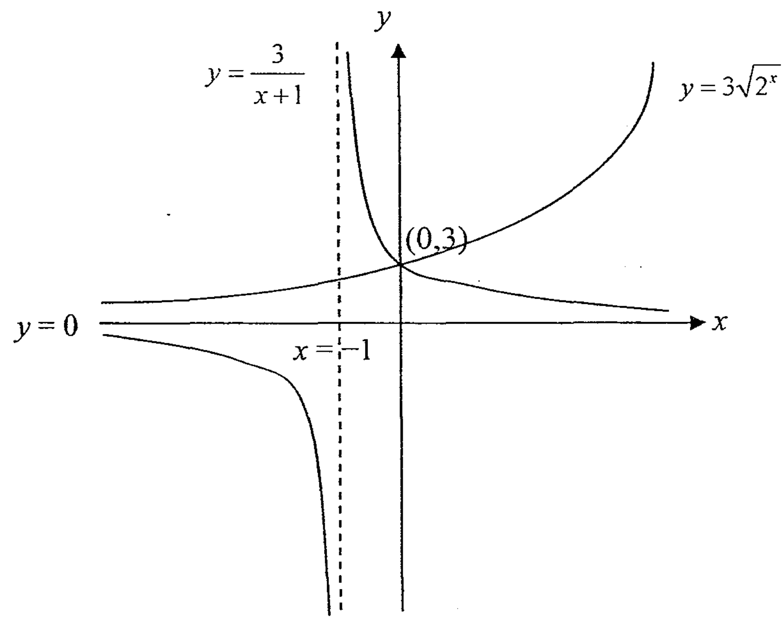
$$2y^2 = 4r^2$$

$y = \sqrt{2}r$ (reject $y = -\sqrt{2}r$ since $y > 0$)

$$P = 2\left(\sqrt{2}r\right) + 2\sqrt{4r^2 - \left(\sqrt{2}r\right)^2} = 2\sqrt{2}r + 2\sqrt{2}r = 4\sqrt{2}r$$

y	$\left(\sqrt{2}r\right)^-$	$\sqrt{2}r$	$\left(\sqrt{2}r\right)^+$
$\frac{dP}{dy}$	/	—	\

Therefore P is a maximum when $y = \sqrt{2}r$.

Qn	Solution
4	**Curve Sketching + Application of Integration**
(i)	
(ii)	$\sqrt{2^x}-\frac{1}{x+1}\geq 0$ $x<-1$ or $x\geq 0$
(iii)	$\int_0^3 3\sqrt{2^x}-\frac{3}{x+1}\,dx=11.668$ (3 d.p.)

Qn	Solution
5	**Applications of Differentiation (Tangent and Normal)**
(i)	$y=\frac{1}{x}$ $\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{1}{x^2}$ At $x=a, y=\frac{1}{a}, \frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{1}{a^2}$, Equation of tangent at A: $y-\frac{1}{a}=-\frac{1}{a^2}(x-a)$ $y=-\frac{1}{a^2}(x-a)+\frac{1}{a}$ $y=-\frac{1}{a^2}x+\frac{2}{a}$
(ii)	When $x=0,\ y=\frac{2}{a}$ When $y=0,\ \frac{1}{a^2}x=\frac{2}{a}\Rightarrow x=2a$ Area of triangle OPQ = $\frac{1}{2}(2a)\left(\frac{2}{a}\right)=2$ (independent of A)
(iii)	Equation of normal at A: $y-\frac{1}{a}=a^2(x-a)$ $y=a^2(x-a)+\frac{1}{a}$ $y=a^2x-a^3+\frac{1}{a}$ Since the normal passes through the origin, $0=a^2(0)-a^3+\frac{1}{a}$ $\Rightarrow a^3=\frac{1}{a}$ $\Rightarrow a=\pm 1$

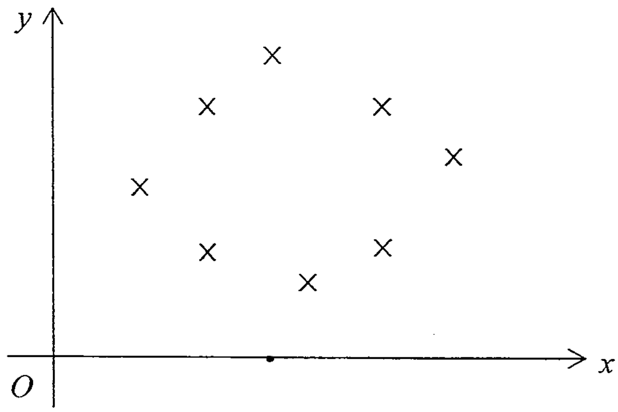
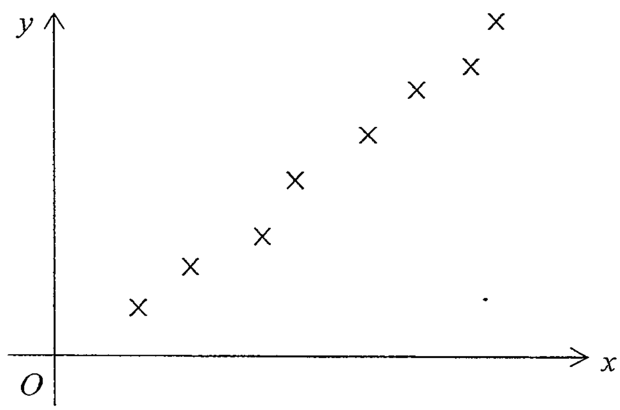
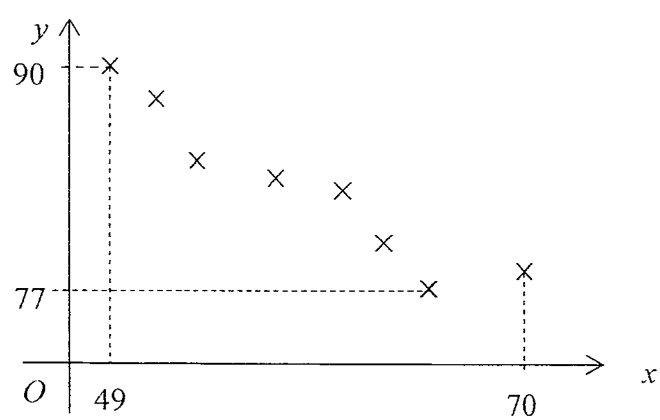
Qn	Solution
6	**Sampling Methods**
(i)	1. Divide the employees in the company into mutually exclusive strata according to the department they work in. 2. (see table below) 3. Using a Random Number Generator, **randomly** select 45 employees from the Manufacturing Department, 30 employees from the Assembly Department, and 25 employees from the Administrative Department. 4. The sample from each stratum is then combined to form a sample size of 100.
(ii)	Using stratified sampling with the strata above will give a **good representative** sample as compared to random sampling because the **working hours of employees from different departments have been considered** and the sample size of each stratum is **proportional** to the size of the stratum. A better stratified sample of size 100 could have been achieved by dividing the employees into **6 strata according to their departments and level of seniority**, as follows: Junior level employees in Manufacturing Department, Senior level employees in Manufacturing Department Junior level employees in Assembly Department, Senior level employees in Assembly Department Junior level employees in Delivery Department, Senior level employees in Delivery Department

Table for step 2 of (i):

	Manufacturing Department	**Assembly Department**	**Administrative Department**	**Total**
Sample size	$100\times\frac{450}{1000}=45$	$100\times\frac{300}{1000}=30$	$100\times\frac{250}{1000}=25$	100

Qn	Solution
7	**Probability**
(i)	Required probability $= p+(1-p)1.2p$ $= p+1.2p-1.2p^2$ $= 2.2p-1.2p^2$ (shown) $2.2p-1.2p^2 \geq 0.99$ Using GC, the smallest value of $p = 0.8$ (1 d.p.)
(ii)	$\frac{p}{2.2p-1.2p^2}=0.7$ $p=1.54p-0.84p^2$ $0.84p^2-0.54p=0$ $p(0.84p-0.54)=0$ $p=0$ (rej) or $p=0.643$ (3 s.f.)
(iii)	Required probability $=0.7\times(1-0.7)1.2(0.7)\times 2$ $=0.3528$

Qn	Solution
8	**Sampling**
(i)	Let X be the price of 3 randomly chosen melons. Let Y be the price of a randomly chosen durian. $X = 5(M_1 + M_2 + M_3)$ $X \sim \mathrm{N}(5\times3\times2.7, 5^2\times3\times0.3^2)$ $X \sim \mathrm{N}(40.5, 6.75)$ $Y = 10D$ $Y \sim \mathrm{N}(10\times1.5, 10^2\times0.1^2)$ $Y \sim \mathrm{N}(15,1)$ $\mathrm{P}(X > 41 \text{ and } Y > 17) = \mathrm{P}(X > 41)P(Y > 17)$ $= (0.42369)(0.02275)$ $= 0.0096389 \approx 0.00964$ (3 s.f.)
(ii)	$X + Y \sim \mathrm{N}(55.5, 7.75)$ $\mathrm{P}(X + Y > 58) = 0.18459 \approx 0.185$ (3 s.f.)
(iii)	Part **(ii)** includes cases that are not considered in part **(i)**. For example, $X = 30$ and $Y = 28$ and $X = 48$ and $Y = 10$. In fact, cases in **(i)** form a proper subset of cases in **(ii)**, which explains why the answer to part **(ii)** is greater than the answer to part **(i)**.

Qn	Solution
9	**Correlation and Regression**
(a)(i)	
(a)(ii)	
(b)	Scatter diagram allows us to assess whether a linear relationship between the two variables exist first or The scatter diagram can be used to identify outliers, if they exist. or Even though $r = 0.993$ indicates a high positive linear correlation between the variables, x and t, the scatter diagram shows that as x increases, t increases at an increasing rate, hence a linear model may not be suitable.
(c)(i)	
b(ii)	$r = -0.958$

b(iii)	$y=-0.59530x+117.90$ (5 s.f) $y=-0.595x+118$ (3 s.f) When $x=80$, $y=-0.59530(80)+117.90$ $y=70.276$ $y=70.3$(3 s.f) The estimate is unreliable as $x=80$ is **outside the data range of x** and thus **linear relation between x and y may no longer hold**.

Qn	Solution
10	**Normal Distribution**
(a)	If the mass of chickens has a normal distribution, there would be a possibility that a chicken has mass less than 0 kg, since $\text{P}(\text{mass of chicken} < 0) = 0.034454$. This means that 3.45% of the chickens would have mass less than 0 kg which is impossible.
(b) (i)	Let X be the daily milk production of a randomly chosen cow in litres $X \sim \text{N}(\mu, \sigma^2)$ and $\overline{X} = \frac{X_1 + X_2 + \ldots + X_{16}}{16} \sim \text{N}\left(\mu, \frac{\sigma^2}{16}\right)$ $\text{P}(X > 35) = 0.25$ $1 - \text{P}(X \le 35) = 0.25$ $\text{P}(X \le 35) = 0.75$ $\text{P}\left(Z \le \frac{35-\mu}{\sigma}\right) = 0.75$ $\Rightarrow \frac{35-\mu}{\sigma} = 0.67449$ $\Rightarrow \mu + 0.67449\sigma = 35 \quad ----(1)$ $\text{P}(\overline{X} \le 25) = 0.089$ $\text{P}\left(Z \le \frac{25-\mu}{\frac{\sigma}{4}}\right) = 0.089$ $\Rightarrow \frac{4(25-\mu)}{\sigma} = -1.3469$ $\Rightarrow 4\mu - 1.3469\sigma = 100 \quad ----(2)$ Using GC to solve (1) and (2). $\mu = 28.3 \quad (3 \text{ s.f.})$ $\sigma = 9.89 \quad (3 \text{ s.f.})$
(b) (ii)	Let Y be the number of cows, out of 19, that is grade A. $Y \sim \text{B}(19, 0.25)$ Required probability $= \text{P}(Y = 3)\text{P}(20\text{th cow is grade A})$ $= 0.15175 \times 0.25$ $= 0.0379 \quad (3 \text{ s.f.})$

Qn	Solution
11	**Hypothesis Testing**
(i)	Let $m = x + y$ $\sum m = 212.8 + 102.6 = 315.4 \qquad \sum m^2 = 490.65 + 230.125 = 720.775$ Unbiased estimate of population mean using combined sample, $\overline{m} = \frac{315.4}{150} = \frac{1577}{750} = 2.1027 = 2.10$ (3.s.f) Unbiased estimate of population variance using combined sample, $s_m^{\ 2} = \frac{1}{149}\left(720.775 - \frac{315.4^2}{150}\right) = 0.38654 = 0.387$ (3 s.f.)
	Let M be the mass of a randomly chosen tin of milk powder. Let μ denote the population mean mass of a tin of milk powder. H_0: $\mu = 2$ H_1: $\mu > 2$ Since $n = 150$ is large, by Central Limit Theorem, $\therefore \overline{M} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ approximately. Test statistic: $Z = \frac{\overline{M} - \mu}{S/\sqrt{n}}$ Level of Significance: 5% Reject H_0 if p-value < 0.05 Under H_0, using G.C., p-value $= 0.021565 = 0.0216$ (3.s.f) Since p-value $= 0.0216 < 0.05$, we reject H_0 and conclude that at 5 % level of significance, there is sufficient evidence that the mean mass of a tin of milk powder is more than 2 kg. 0.0216 is the probability of obtaining a sample mean greater than or equal to 2.1027kg, assuming that the population mean mass of a tin of milk powder is 2kg.
	The null hypothesis is still rejected in favour of the alternative hypothesis, as p-value $= 2(0.021565) < 0.05$. Therefore there is sufficient evidence that the mean mass of a tin of milk powder is not 2 kg.
	Level of Significance: 5% Reject H_0 if z-value > 1.6449 $\frac{2.1027 - \mu_0}{\sqrt{0.38654}/\sqrt{150}} > 1.6449$ $\mu_0 < 2.1027 - (1.6449)\left(\sqrt{0.38654}/\sqrt{150}\right)$

Qn	Solution
12	**Binomial Distribution**
(a)(i)	Let X be the number of cookies, out of 12, that are chocolate cookies. $X \sim B(12, p)$ $E(X) = np = 5.4$ $\therefore p = \frac{5.4}{12} = 0.45$ $P(X = 3) = 0.092326$ ≈ 0.0923 (to 3 s.f.)
(ii)	$X \sim B(12, 0.45)$ Using GC, $P(X = 4) = 0.16996$ $P(X = 5) = 0.2225$ (highest probability) $P(X = 6) = 0.21238$ $\therefore$ most likely number is 5 cookies.
(b)(i)	Let Y be the number of cookies, out of n, that are chocolate cookies. $Y \sim B(n, 0.3)$ $P(Y \leq 8) \leq 0.9$ Using GC, When $n = 19$, $P(Y \leq 8) = 0.91608 > 0.9$ When $n = 20$, $P(Y \leq 8) = 0.88667 < 0.9$ $\therefore$ least n is 20.
(ii)	Total number of cookies $= 12 \times 10 = 120$ Let W be the number of cookies, out of 120, that are chocolate cookies. $W \sim B(120, 0.3)$ Since n is large such that $np = 36 > 5$ and $n(1-p) = 84 > 5$, $W \sim N(36, 25.2)$ approximately. $P(W > 30) = P(W > 30.5)$ after continuity correction $= 0.863378$ $= 0.863$ (to 3 s.f.)