

**ANGLO-CHINESE JUNIOR COLLEGE
JC2 PRELIMINARY EXAMINATION**

Higher 2

FURTHER MATHEMATICS

9649/02

Paper 2

17 September 2026

3 hours

Additional Materials: Printed Answer Booklet
 List of Formulae and Results (MF27)

READ THESE INSTRUCTIONS FIRST

Answer **all** the questions.

Write your answers on the Printed Answer Booklet. Follow the instructions on the front cover of the answer booklet.

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 an approved 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 **10** printed pages.



Anglo-Chinese Junior College

[Turn Over

Section A: Pure Mathematics [50 marks]

1 For real k , let

$$S_k = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + y + z = 0 \text{ and } x + 2y + 3z = 0 \text{ and } 2x + 3y + kz = 0 \right\}.$$

(a) Determine the value of k for which S_k contains non-zero vectors. [2]

Let k take the value found in part **(a)**.

(b) Solve the equations defining S_k and give S_k in vector form. [1]

(c) Hence solve the system of equations

$$\begin{aligned} x + y + z &= 1, \\ x + 2y + 3z &= 1, \\ 2x + 3y + kz &= 2. \end{aligned}$$

Determine whether the solution set T is a subspace of $\mathbb{R}^3$. [2]

2 (a) Solve the differential equation

$$x \frac{dy}{dx} + y(1-x) = 1,$$

expressing y in terms of x . [3]

(b) Sketch, on separate diagrams, a member of the family of solution curves for each of the following conditions and state the equation of each curve. (You are not required to label the coordinates of any stationary points.)

(i) A solution curve with a vertical asymptote and no stationary point;

(ii) A solution curve with a vertical asymptote and at least one stationary point.

[3]

(c) State the equation of the solution curve that does not have any vertical asymptote, and explain why this curve does not have any vertical asymptote. [2]

- 3** The temperature T at any point (x, y) is given by $T(x, y) = x^2 + xy + y^2 - 4x - 5y$, for all real values of x and y .

A particle is at the point $P(1, 1)$.

- (a)** Find a direction in which the temperature increases most rapidly at P .

Hence, find the maximum rate of increase of temperature at P . [3]

A unit vector $\mathbf{u}$ in the xy -plane makes an angle θ with the positive x -axis.

- (b)** Express $\mathbf{u}$ in terms of θ . Given that the directional derivative of T at P in the direction of $\mathbf{u}$ is zero, find the value of θ for $0 \leq \theta \leq \pi$, giving your answer exactly.

[3]

The surface S has equation $z = T(x, y)$.

- (c)** Find the Cartesian equation of the tangent plane to S at the point $(1, 1, -6)$. [2]

- (d)** Find the vector equation of the normal line to S at the point $(1, 1, -6)$. [2]

- 4 It is known that the function $f(x) = 3x^5 - 8x^2 + 4$ has exactly one negative root, γ , and exactly two positive roots, α and β , such that $\alpha < \beta$.

- (a) Show that $0.75 < \alpha < 1 < \beta < 1.25$. [2]
- (b) Use Newton-Raphson iteration with an initial value -1 to estimate γ to 3 decimal places. [2]

A student, Annie, wishes to explore different iterative approaches to estimating α and β using initial values close to 1.

- (c) Explain why Newton-Raphson iteration may not be appropriate for estimating either α or β with an initial value is close to 1. [1]

Her classmate, Bhaven, proposes using the iterative formula

$$x_{n+1} = \frac{3}{8}x_n^4 + \frac{1}{2x_n}$$

with a positive initial term x_1 , to estimate α and β .

- (d) Show that if the sequence $\{x_n\}$ converges to a limit using Bhaven's formula, then it must converge to either α or β . [2]
- (e) Show that Bhaven's formula is suitable for estimating α , but not β . Hence, estimate α to 3 decimal places. [4]
- (f) Suggest an iterative method to estimate β with an initial value of 1, and carry out the first step of the suggested method. [2]

- 5 Researchers wish to model the amount of atmospheric carbon dioxide generated due to human activities, known as anthropogenic carbon. In recent years, technological advancements have been made to store anthropogenic carbon and mitigate climate change. These researchers have proposed the general model

$$C_{n+1} = \left(1 + \frac{\alpha}{100}\right) C_n - S_n \text{ for } n \geq 0, \text{ where}$$

- α is a positive constant denoting the percentage increase in anthropogenic carbon generated per year;
- C_n denotes the amount of anthropogenic carbon present at the start of the n th year;
- S_n denotes the amount of anthropogenic carbon stored during the n th year, and
- C_0 denotes the amount of anthropogenic carbon present initially.

In a *constant storage model*, it is taken that $S_n = k$, where k is a positive constant.

- (a) (i) Express C_n in terms of α , k and C_0 . [3]
- (ii) State a condition on α , in terms of k and C_0 , to prevent the long run amount of anthropogenic carbon from increasing without bound. [1]

Now, assume an *idealised technological model*, where it is taken that

- For $n = 0$, $S_0 = \frac{1}{200} \alpha C_0$,
- For $n \geq 1$, S_n is half of the sum of the excess anthropogenic carbon generated in the n th year and the year before.

- (b) (i) Find C_1 and show that $C_{n+1} - \left(1 + \frac{\alpha}{200}\right) C_n + \frac{\alpha}{200} C_{n-1} = 0$ for $n \geq 1$. [2]

- (ii) Find an expression for C_n , in terms of C_0 and α , and hence determine the exceptional value of α for which the model predicts a linear increase in anthropogenic carbon. Deduce exactly the range of values of α such that the long run amount of anthropogenic carbon does not exceed $1.5C_0$. [6]

- (iii) Give a criticism of the idealised technological model and suggest an improvement to the model to overcome the issue. [2]

Section B: Statistics [50 marks]

- 6** A study was conducted to investigate whether a dietary intervention programme or a standard exercise programme is more effective in reducing stress levels among university students in Singapore. Eighteen students with elevated stress scores were divided into nine pairs with similar baseline scores. For each pair, one member was randomly assigned to the dietary intervention programme and the other to the standard exercise programme. Eight weeks later, the following stress scores were recorded, all lower than their baseline scores. (Scores are measured on a standardised scale, where higher scores indicate greater stress.)

Pair	1	2	3	4	5	6	7	8	9
Dietary intervention	38	50	47	54	44	55	49	36	41
Standard exercise	41	46	52	61	53	55	47	42	51

Carry out the following tests at the 5% significance level to test whether the dietary intervention programme is more effective in reducing anxiety than the standard exercise programme.

(a) Sign test; [5]

(b) Wilcoxon matched-pair signed rank test. [3]

Comment on the use of the above tests. [1]

- 7 A streaming platform conducted a survey to study the genre preference of its subscribers, which are broadly divided into the student and working adult categories. A random sample of 580 subscribers was asked to identify their preferred genre. The results are summarised in the table below.

	Action	Romance	Documentary	Comedy	Horror	Anime
Student	84	52	22	63	43	56
Working adult	71	44	43	38	29	35

- (a) Perform a chi-squared test at the 1% significance level to examine whether the data provide any evidence that genre preference is associated with the type of subscriber [5]
- (b) State the contributions to the test statistic and hence comment on the significant differences in genre preferences between the two types of subscribers. [2]

- 8** The daily rainfall depth, X mm, recorded at a weather station is modelled as a continuous random variable with probability density function

$$f(x) = \begin{cases} kxe^{-\frac{x}{4}}, & \text{for } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

- (a)** Show that the value of k is $\frac{1}{16}$ and sketch the graph of f . [2]
- (b)** Find the mode of X and $E(X)$. Explain, in the context of the question, why the two values are quite different. [3]
- (c)** Find $\text{Var}(X)$. [3]

The flood warning index is related to the daily rainfall depth. A model by the weather station proposed the flood warning index, Y , to be given by

$$Y = \begin{cases} 0, & x < 10, \\ 2(X - 10), & x \geq 10. \end{cases}$$

- (d)** Use this model to find the expected flood warning index. [2]

- 9 Researchers at a wildlife centre monitors the number of sightings of a rare bird species at a nature reserve. The number of sightings in a 20-minute period is denoted by S .

- (a) State two assumptions required for S to be well modelled by a Poisson distribution. [2]

Now assume that S is modelled by a Poisson distribution with mean of 6.

- (b) Find the probability that in a 20-minute period, exactly 4 bird sightings are recorded. [1]
- (c) Find the probability that there are between 3 and 8 bird sightings inclusive recorded in an hour. [2]

The researchers want to determine the number of bird sightings in a 20-minute period that is most likely to occur.

- (d) By considering the values of

$$\frac{P(S = s + 1)}{P(S = s)},$$

determine the most likely number(s) of bird sightings. [3]

A researcher observes the reserve at a random time.

- (e) Find the probability that the researcher waits more than 4 minutes before the next bird sighting. [2]
- (f) The researcher has already waited 5 minutes without any bird sightings. Find the probability that the researcher waits a further 4 minutes before the next sighting. [2]

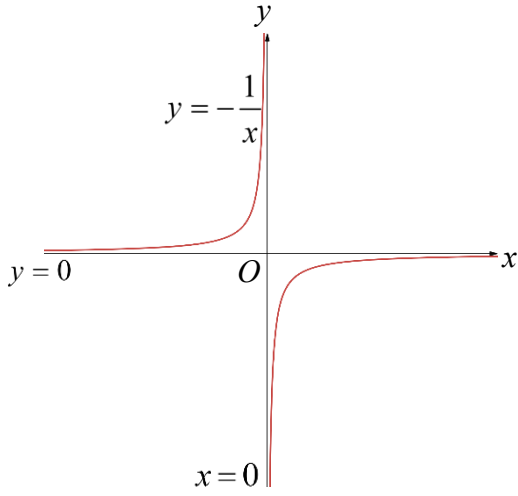
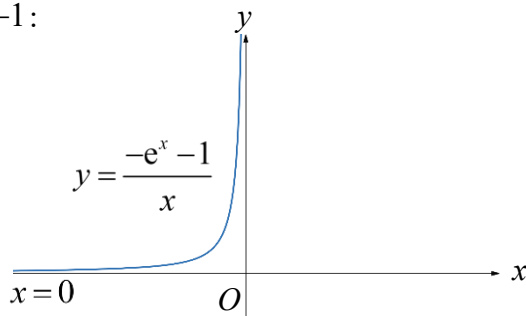
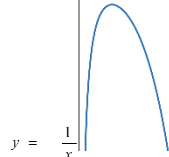
- 10** A sports scientist is investigating whether athletes complete a reaction-time task more quickly using their dominant hand or non-dominant hand. Eight athletes are selected at random. Each athlete performs the same task once using each hand.

The reaction times, in milliseconds (ms), are shown below.

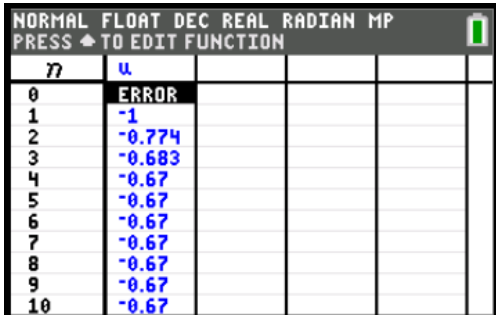
Athlete	1	2	3	4	5	6	7	8
Dominant hand	242	251	238	247	244	255	239	248
Non-dominant hand	249	257	242	253	249	261	244	255

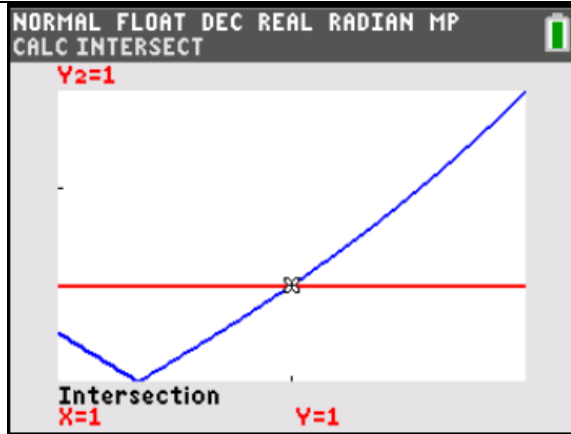
- (a) Construct a 96% confidence interval for the mean increase in reaction time when the athletes use their non-dominant hand instead of the dominant hand. Give your answer correct to 3 decimal places. [3]
- (b) Hence determine whether there is sufficient evidence, at the 4% significance level, to conclude that the mean reaction times for the two hands are different. [3]
- (c) State two assumptions about the increases in reaction time that are necessary to carry out a paired sample t -test. [2]
- (d) The scientist claims that the athletes take longer, on average, when using their non-dominant hand. State suitable hypotheses and determine the p -value for the test. Hence comment on the strength of the evidence to support the claim. [4]

Qn	Suggested Solution	
1(a)	Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & k \end{pmatrix}$ Nonzero solutions exist if $\mathbf{A}$ is not invertible, i.e. $\det(\mathbf{A}) = 0$. $\det(\mathbf{A}) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & k \end{vmatrix}$ $= 1 \cdot (2k - 9) - 1 \cdot (1 \cdot k - 6) + 1 \cdot (3 - 4)$ $= 2k - 9 - (k - 6) - 1 = k - 4$ $\det(\mathbf{A}) = 0$ when $k = 4$. Hence nonzero solutions exist when $k = 4$.	
(b)	When $k = 4$, RREF of $\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. $S_k = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, t \in \mathbb{R} \right\}$	
(c)	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in T$ but $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \notin T$ (can check that $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ is not a solution to the system of equations.) Since $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \notin T$, T is not closed under addition, hence T is not a vector subspace of $\mathbb{R}^3$. OR Since $2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \notin T$, T is not closed under scalar multiplication, hence T is not a vector subspace of $\mathbb{R}^3$. OR Check that $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is not a solution to the system of equations.)	

	Since $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \notin T$, hence T is not a vector subspace of $\mathbb{R}^3$.	
2(a)	$x \frac{dy}{dx} + y(1-x) = 1$ $\frac{dy}{dx} + y \left(\frac{1-x}{x} \right) = \frac{1}{x}$ Integrating factor: $I(x) = e^{\int \frac{1-x}{x} dx} = e^{\ln x - x} = x e^{-x}.$ Thus, $\frac{d}{dx} (y x e^{-x}) = \frac{1}{x} (x e^{-x})$ $y x e^{-x} = \int e^{-x} dx$ $= -e^{-x} + c$ $y = \frac{-e^{-x} + c}{x e^{-x}}$ $y = \frac{c e^x - 1}{x}.$	
(b)	E.g., $c = 0$:  E.g. $c = -1$:  	

(c)	The curve $y = \frac{e^x - 1}{x}$ does not have any vertical asymptote (despite having a denominator of x). This is because the Maclaurin series gives $y = \frac{e^x - 1}{x}$ $= \frac{1}{x} \left(-1 + 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \right)$ $= 1 + \frac{x}{2!} + \frac{x^2}{3!} + \dots$ As $x \rightarrow 0$, $\frac{e^x - 1}{x} \rightarrow 1 + \frac{0}{2!} + \frac{0^2}{3!} + \dots = 1$ which is a finite value.	
3(a)	Taking partial derivatives, we have $\frac{\partial T}{\partial x} = 2x + y - 4$ $\frac{\partial T}{\partial y} = x + 2y - 5$ At P, $\frac{\partial T}{\partial x} = 2(1) + (1) - 4 = -1$ and $\frac{\partial T}{\partial y} = (1) + 2(1) - 5 = -2, \text{ i.e., } \nabla T = \begin{pmatrix} -1 \\ -2 \end{pmatrix}.$ Maximum rate of increase is $\nabla T = \sqrt{5}$.	
3(b)	Since the directional derivative of T at P in the direction of $\mathbf{u}$ is zero, $\nabla T \cdot \mathbf{u} = 0$. Taking $\mathbf{u} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$, we have $\nabla T \cdot \mathbf{u} = -\cos \theta - 2 \sin \theta = 0$ Solving, $\tan \theta = -\frac{1}{2}$. Thus, $\theta = \pi - \tan^{-1} \left(\frac{1}{2} \right)$.	
3(c)	At $(1, 1, -6)$, $z - z_0 = T_x(x - x_0) + T_y(y - y_0)$, we have $z + 6 = -(x - 1) - 2(y - 1)$. Thus, $x + 2y + z = -3$.	
3(d)	Normal vector for S at $(1, 1, -6)$ is $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.	

	Thus, $\mathbf{r} = \begin{pmatrix} 1 \\ 1 \\ -6 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \lambda \in \mathbb{R}.$	
4(a)	$f(0.75) = 0.21191 > 0$ $f(1) = -1 < 0$ $f(1.25) = 0.65527 > 0$ As f is a continuous function, the sign change between 0.75 and 1 implies that there is at least 1 root between 0.75 and 1. Likewise, at least 1 root between 1 and 1.25.	
4(b)	$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{3x_n^5 - 8x_n^2 + 4}{15x_n^4 - 16x_n}$  $\therefore \gamma = -0.670$ (to 3 dp)	
4(c)	The value of $x = 1$ is near to a turning point of the curve where $f'(x) \approx 0$, which may cause iterates to diverge or converge to another root.	
4(d)	As $n \rightarrow \infty$, if $x_n \rightarrow L$, then $x_{n+1} \rightarrow L$. Thus, $L = \frac{3}{8}L^4 + \frac{1}{2L} \Rightarrow 8L^2 = 3L^5 + 4 \Rightarrow 3L^5 - 8L^2 + 4 = 0$ From the given info, $L = \alpha$ or β or γ. However, if $x_n > 0$, then $x_{n+1} = \frac{3}{8}x_n^4 + \frac{1}{2x_n} > 0$. So, L cannot be negative, and $L = \alpha$ or β.	
4(e)	Let $g(x_n) = \frac{3}{8}x_n^4 + \frac{1}{2x_n}$, so $g'(x_n) = \frac{3}{2}x_n^3 - \frac{1}{2x_n^2}$. Use the GC to plot the graph of $y = g'(x_n)$ for $x_n \in (0, 0.75, 1.25)$	



Thus, $|g'(x_n)| < 1$ for $0.75 < x_n < 1$ around the neighbourhood of α . Hence, the iteration would converge to α .

However, $|g'(x_n)| > 1$ for $1 < x_n < 1.25$ around the neighbourhood of β . Hence, the iteration would not converge to β .

Using an initial value of $x_0 = 0.9$ (for example),

NORMAL FLOAT DEC REAL Radian MP PRESS + FOR Δ Tb1					
n	u				
0	ERROR				
1	0.9				
2	0.8016				
3	0.7786				
4	0.78				
5	0.7798				
6	0.7799				
7	0.7798				
8	0.7798				
9	0.7798				
10	0.7798				
$n=0$					

$\alpha = 0.780$ (to 3 dp)

4(f)

Method 1: Bisection

$$\beta \approx \frac{1+1.25}{2} = 1.125$$

Method 2: Linear interpolation

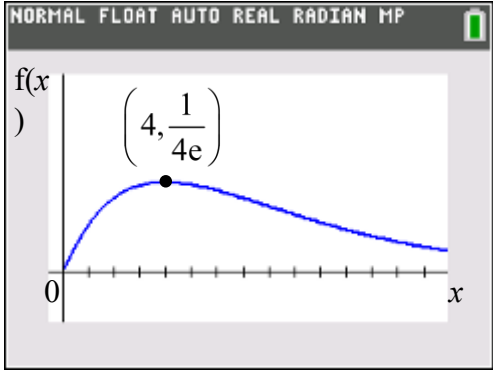
	$\beta \approx \frac{1 f(1.25) + 1.25 f(1) }{ f(1.25) + f(1) }$ $= \frac{0.65527 + 1.25}{0.65527 + 1}$ $= 1.15 \text{ (to 3 sf)}$	
5(ai)	$C_n = \left(1 + \frac{\alpha}{100}\right) C_{n-1} - k$ $= \left(1 + \frac{\alpha}{100}\right)^2 C_{n-2} - \left(1 + \frac{\alpha}{100}\right) k - k$ $= \dots$ $= \left(1 + \frac{\alpha}{100}\right)^n C_0 - k \left[1 + \left(1 + \frac{\alpha}{100}\right) + \dots + \left(1 + \frac{\alpha}{100}\right)^{n-1}\right]$ $= \left(1 + \frac{\alpha}{100}\right)^n C_0 - k \times \frac{\left(1 + \frac{\alpha}{100}\right)^n - 1}{\frac{\alpha}{100}}$ $= \left(C_0 - \frac{100k}{\alpha}\right) \left(1 + \frac{\alpha}{100}\right)^n + \frac{100k}{\alpha}$	
5(aii)	Since $1 + \frac{\alpha}{100} > 1$, $C_n \rightarrow \infty$ as $n \rightarrow \infty$, unless $C_0 - \frac{100k}{\alpha} \leq 0.$ $\therefore 0 < \alpha \leq \frac{100k}{C_0}$	
5(bi)	$C_1 = \left(1 + \frac{\alpha}{100}\right) C_0 - S_0$ $= \left(1 + \frac{\alpha}{100}\right) C_0 - \frac{\alpha}{200} C_0$ $= \left(1 + \frac{\alpha}{200}\right) C_0$ For $n \geq 1$, $C_{n+1} = \left(1 + \frac{\alpha}{100}\right) C_n - \left[\frac{\alpha}{200} C_n + \frac{\alpha}{200} C_{n-1}\right]$ $C_{n+1} = \left(1 + \frac{\alpha}{200}\right) C_n - \frac{\alpha}{200} C_{n-1}$ $C_{n+1} - \left(1 + \frac{\alpha}{200}\right) C_n + \frac{\alpha}{200} C_{n-1} = 0$	

5(bii)	Auxiliary equation: $\lambda^2 - \left(1 + \frac{\alpha}{200}\right)\lambda + \frac{\alpha}{200} = 0$ $(\lambda - 1)\left(\lambda - \frac{\alpha}{200}\right) = 0$ Thus, $\lambda = 1$ or $\lambda = \frac{\alpha}{200}$. We have repeated roots when $\frac{\alpha}{200} = 1 \Rightarrow \alpha = 200$. Case 1: When $\alpha = 200$, then $C_n = A + Bn$. Let $n = 0: C_0 = A$ Let $n = 1: C_1 = C_0 + B \Rightarrow B = (1 + 1)C_0 - C_0 = C_0$ $\therefore C_n = C_0 + C_0n = C_0(1 + n)$ Case 2: $\alpha \neq 200$: $C_n = A + B\left(\frac{\alpha}{200}\right)^n$ Let $n = 0: C_0 = A + B$ --- (1) Let $n = 1: C_1 = A + B\left(\frac{\alpha}{200}\right)$ --- (2) $(2) - (1): C_1 - C_0 = \left(\frac{\alpha}{200} - 1\right)B$ $\Rightarrow \frac{\alpha}{200}C_0 = \frac{\alpha - 200}{200}B$ $\Rightarrow B = \frac{\alpha}{\alpha - 200}C_0$ From (1): $A = C_0 - \frac{\alpha}{\alpha - 200}C_0 = \frac{200C_0}{200 - \alpha}$ $\therefore C_n = \frac{200C_0}{200 - \alpha} + \frac{\alpha C_0}{\alpha - 200}\left(\frac{\alpha}{200}\right)^n$ If $\alpha > 200$, $B > 0$ and $\frac{\alpha}{200} > 1$, so $\{C_n\}$ diverges to infinity.	
--------	---	--

	For $0 < \alpha < 200$, $0 < \frac{\alpha}{200} < 1$, so $\{C_n\}$ converges to $\frac{200C_0}{200-\alpha}$. Therefore, $\frac{200C_0}{200-\alpha} \leq \frac{3}{2} C_0 \Rightarrow 400 \leq 600 - 3\alpha \Rightarrow 0 < \alpha \leq \frac{200}{3}$																															
5(biii)	Criticism 1: The idealised technological model seems to assert that ever more anthropogenic carbon can be stored, especially if $\{C_n\}$ is increasing. This is <u>unrealistic as one would expect a limit to the capacity of the technology.</u> Improvement 1: Set a cap on S_n, e.g. $S_n = \min \left\{ \frac{\alpha}{200} C_n + \frac{\alpha}{200} C_{n-1}, k \right\}$ Criticism 2: The idealised technological model operates on the basis that a constant $\alpha\%$ of anthropogenic carbon is generated per year <u>when it may decrease due to improvements in carbon capture technology OR increase due to continued development.</u> Improvement 2: Allow α to depend on n or prior terms in the recurrence.																															
6(a)	Let X and Y be the anxiety scores of students on the dietary intervention programme and the standard exercise programme respectively. Let $D = X - Y$ be the difference in anxiety scores, and m_D be the population median difference. Sign Test: To test $H_0: m_D = 0$ against $H_1: m_D < 0$ <table><tr><td>Pair</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>d</td><td>-3</td><td>+4</td><td>-5</td><td>-7</td><td>-9</td><td>0</td><td>+2</td><td>-6</td><td>-10</td></tr><tr><td>sign</td><td>-</td><td>+</td><td>-</td><td>-</td><td>-</td><td>---</td><td>+</td><td>-</td><td>-</td></tr></table> (Pair 6 is excluded as $d = 0$, leaving $n = 8$ valid pairs.) Let S_+ be the number of + signs out of 8.	Pair	1	2	3	4	5	6	7	8	9	d	-3	+4	-5	-7	-9	0	+2	-6	-10	sign	-	+	-	-	-	---	+	-	-	
Pair	1	2	3	4	5	6	7	8	9																							
d	-3	+4	-5	-7	-9	0	+2	-6	-10																							
sign	-	+	-	-	-	---	+	-	-																							

	Under H_0, $S_+ \sim B(8, 0.5)$ From the table, $s_- = 6$ and $s_+ = 2$. $p\text{-value} = P(S_+ \leq 2) = 0.1445 > 0.05$ We do not reject H_0. There is insufficient evidence at 5% significance level that the dietary intervention programme is more effective than the standard exercise programme in reducing anxiety.																															
6(b)	Wilcoxon Matched-Pair Signed Rank Test: <table><tr><td>Pair</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>d</td><td>-3</td><td>+4</td><td>-5</td><td>-7</td><td>-9</td><td>0</td><td>+2</td><td>-6</td><td>-10</td></tr><tr><td>rank</td><td>2</td><td>3</td><td>4</td><td>6</td><td>7</td><td>---</td><td>1</td><td>5</td><td>8</td></tr></table> To test $H_0: m_D = 0$ against $H_1: m_D < 0$ Sum of positive ranks, $P = 1 + 3 = 4$ Sum of negative ranks, $Q = 2 + 4 + 5 + 6 + 7 + 8 = 32$ $T = \min(P, Q) = 4$ From table, since $T \leq 5$, we reject H_0. There is sufficient evidence at 5% significance level that the dietary intervention programme is more effective than the standard exercise programme in reducing anxiety.	Pair	1	2	3	4	5	6	7	8	9	d	-3	+4	-5	-7	-9	0	+2	-6	-10	rank	2	3	4	6	7	---	1	5	8	
Pair	1	2	3	4	5	6	7	8	9																							
d	-3	+4	-5	-7	-9	0	+2	-6	-10																							
rank	2	3	4	6	7	---	1	5	8																							
	The conclusions of the two non-parametric tests are different. The Wilcoxon test makes use of the magnitude of the differences between matched pairs of the anxiety scores and is hence a better test compared to the sign test as it incorporates more information about the data. Hence the conclusion of the Wilcoxon matched-pair signed rank test is more reliable.																															
7(a)	Test H_0: Preferred content genre is independent of subscriber type. Against H_1: Preferred content genre is dependent on subscriber type. Using a χ^2 test for independence. Expected frequencies: <table><tr><td></td><td>Action</td><td>Romance</td><td>Documentary</td><td>Comedy</td><td>Horror</td><td>Anime</td></tr><tr><td>Student</td><td>85.52</td><td>52.96</td><td>35.86</td><td>55.72</td><td>39.72</td><td>50.21</td></tr></table>		Action	Romance	Documentary	Comedy	Horror	Anime	Student	85.52	52.96	35.86	55.72	39.72	50.21																	
	Action	Romance	Documentary	Comedy	Horror	Anime																										
Student	85.52	52.96	35.86	55.72	39.72	50.21																										

	<table><tr><td>Working adult</td><td>69.48</td><td>43.03</td><td>29.14</td><td>45.28</td><td>32.28</td><td>40.79</td></tr></table> Degree of freedom = (2 – 1)(6 – 1) = 5 $\chi^2 = 16.265$ $p\text{-value} = 0.00613 < 0.01$, we reject H_0. There is sufficient evidence at 1% significance level that genre preference is dependent on the type of subscribers.	Working adult	69.48	43.03	29.14	45.28	32.28	40.79															
Working adult	69.48	43.03	29.14	45.28	32.28	40.79																	
7(b)	Contributions of each cell to the test statistic: <table><tr><td></td><td>Action</td><td>Romance</td><td>Documentary</td><td>Comedy</td><td>Horror</td><td>Anime</td></tr><tr><td>Student</td><td>0.0270</td><td>0.0174</td><td>5.36</td><td>0.951</td><td>0.271</td><td>0.668</td></tr><tr><td>Working adult</td><td>0.0333</td><td>0.0219</td><td>6.59</td><td>1.17</td><td>0.333</td><td>0.822</td></tr></table> The largest contributions come from the "Documentary" column. Notably, far fewer students watch documentaries than expected while far more working adults do, suggesting that documentary viewing is more prevalent among working adults.		Action	Romance	Documentary	Comedy	Horror	Anime	Student	0.0270	0.0174	5.36	0.951	0.271	0.668	Working adult	0.0333	0.0219	6.59	1.17	0.333	0.822	
	Action	Romance	Documentary	Comedy	Horror	Anime																	
Student	0.0270	0.0174	5.36	0.951	0.271	0.668																	
Working adult	0.0333	0.0219	6.59	1.17	0.333	0.822																	
8(a)	$\int_0^{\infty} kxe^{-\frac{x}{4}}dx = k \left\{ \left[-4xe^{-\frac{x}{4}} \right]_0^{\infty} + \int_0^{\infty} 4e^{-\frac{x}{4}}dx \right\}$ $= k \left\{ 0 - \left[16e^{-\frac{x}{4}} \right]_0^{\infty} \right\}$ $= 16k$ Total probability = 1 $\Rightarrow 16k = 1$, i.e. $k = \frac{1}{16}$ (shown) $\therefore f(x) = \frac{x}{16}e^{-\frac{x}{4}}, x > 0$																						

		
8(b)	Mode of $X = 4$ $E(X) = \int_0^{\infty} \frac{x^2}{16} e^{-\frac{x}{4}} dx = 8$ The mode represents the most commonly occurring daily rainfall depth and light-to-moderate rain events are the most probable. However, the mean is pulled considerably higher because of rare but intense rainfall events. These events contribute disproportionately large depths. (The mean is sensitive to these high values in a way that the mode is not.)	
8(c)	$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ &= \int_0^{\infty} \frac{x^3}{16} e^{-\frac{x}{4}} dx - (8)^2 \\ &= 32 \end{aligned}$	
8(d)	$\begin{aligned} E(Y) &= \int_{10}^{\infty} (2x - 20) \frac{x}{16} e^{-\frac{x}{4}} dx \\ &= 2.96 \text{ (to 3sf)} \end{aligned}$	
9(a)	Sightings occur independently of one another. OR Mean rate of sighting is a constant. OR Probability of two or more sightings occurring in a very small time interval is negligible.	
9(b)	$S \sim \text{Po}(6)$ $P(S = 4) = 0.134$	
9(c)	Let R be the number of sightings in 1 hour i.e. 60 min $R \sim \text{Po}(18)$ $P(3 \leq R \leq 8)$ $= P(R \leq 8) - P(R \leq 2)$ $= 0.00705$	

9(d)	$\frac{P(S = s + 1)}{P(S = s)} = \frac{e^{-6} \frac{6^{s+1}}{(s+1)!}}{e^{-6} \frac{6^s}{s!}} = \frac{6}{s+1}$ Probabilities increase when $\frac{6}{s+1} > 1 \Rightarrow s < 5$ Probabilities decrease when $s > 5$ Thus maximum prob occurs when $s = 5$ However when $s = 5$, $\frac{P(S = 6)}{P(S = 5)} = \frac{6}{6} = 1 \Rightarrow P(S = 5) = P(S = 6)$ Thus modes are 5 and 6.	
9(e)	Let T be the waiting time between 2 consecutive sightings $T \sim \text{Exp}\left(\frac{6}{20}\right) = \text{Exp}(0.3)$ $P(T > 4) = e^{-0.3(4)} = 0.301$	
9(f)	$P(T > 9 \mid T > 5)$ $= P(T > 4)$ $= 0.301$	
10(a)	Let D = reaction time of non-dominant hand – reaction time of dominant hand $d = 7, 6, 4, 6, 5, 6, 5, 7$ $\bar{d} = 5.75, s^2 = 1.0351^2$ A 96% confidence interval for the mean increase in reaction time when athletes use their non-dominant hand instead of the dominant hand is $\left(\bar{d} \pm t_{(7, 0.02)} \frac{s}{\sqrt{8}} \right) = (4.829, 6.671)$	
10(b)	Let μ_D be the mean increase in reaction time. Test $H_0 : \mu_D = 0$, Against $H_1 : \mu_D \neq 0$, at 4% significance level. Since the 96% confidence interval for the increase (4.829, 6.671) does not contain $\mu_D = 0$, we conclude that there is sufficient evidence at 4% significance level that the population mean reaction times for the two hands are different.	
10(c)	(1) The increases are approximately normally distributed.	

	(2) The increases are independent of one another across athletes.	
10(d)	Test $H_0 : \mu_D = 0$ Against $H_1 : \mu_D > 0$ Using a paired sample t-test, Under H_0, Test statistic $T = \frac{\bar{D}}{\frac{s}{\sqrt{8}}} \sim t(7)$ $p\text{-value} = 5.12 \times 10^{-7}$ Since the p-value is very close to 0, the test will reject H_0 if significance level is at least 0.0000512% . There is very strong evidence to support the claim that the athletes take longer, on average, when using their non-dominant hand.	