H2 Mathematics

Ryan Joo

Last updated: October 29, 2024

Graphing Calculator

Basic input methods

- Change decimal to fraction: math 1:▶Frac
- Change fraction to decimal: math 2:▶Dec
- Cube root: math 4: $\sqrt[3]{}$
- Equality and inequality symbols: test conditions

Graphing

- Input equation: y=
- View graph: graph
- Restrict domain/range: window zoom ZoomFit
- Zoom in or out: zoom
- Find y value at specific x value: calc 1:value
- Find x-intercept: calc 2:zero
- Find point of intersection: calc 5:intersect
- Find min or max point: calc 3:minimum or 4:maximum
- Parametric functions: mode function parametric then proceed to graph
- Conic sections: apps 2:Conics
- Piecewise function: math B:piecewise
- Composite function: input Y1 and Y2, f4 to input Y2(Y1)

Equations and Inequalities

- Solve quadratic equation: apps 4:PlysmIt2
- Solve system of linear equations: apps 4:PlysmIt2
- Find solution(s) to a complicated equation: graph the functions, then find point(s) of intersection

Sequences and Series

- Generate sequence: list ops 5:seq
- Determine behaviour of sequence/series: input graph, then table
- Evaluate summation: math 0:summation

Calculus

- Evaluate derivative/gradient at a point (graph): calc 6:dy/dx
- Evaluate integral given the lower and upper limits (graph): calc 7:∫f(x)dx
- Evaluate derivative at an x value: math 8:nDeriv
- Evaluate definite integral: math 9:fnInt

Complex Numbers

- Changde mode settings: mode real a+bi
- Rectangular to polar: math cmplx 7:Polar
- Simplify complex expressions: just input
- Modulus: math cmplx 5:abs
- Argument: math cmplx 4:angle

Probability

- Permutation: math prob 2:nPr
- Combination: math prob 3:nCr
- Factorial: math prob 4:!

Distributions

- Binomial distribution (pdf): distr A:binompdf
- Binomial distribution (cdf): distr B:binomcdf
- Normal distribution (cdf): distr 2:normalcdf
- Inverse normal: distr 3:invNorm
- Get summary statistics: stat 1:edit to input data, stat 1-Var Stats
- Z-test: stat tests 1:Z-Test

Correlation and Regression

- Scatter diagram: input values into lists L1 and L2, stat plots 1:Plot1 Xlist Ylist 9:ZoomStat
- Value of r: input values, stat calc 8:LinReg(a+bx)
- Regression line on scatter diagram: F4 1:Y1 to store equation into Y1
- Linearise: bring cursor to L3, key in 1/L1 to generate values of $\frac{1}{x}$

Part I

Pure Mathematics

1 Functions

- **Domain** is set of possible inputs, **range** is set of possible outputs.

  Use **vertical line test** to check for functions: f is a function if vertical line $x = k$ cuts the graph at most once for all $k \in D_f$.

- **One-one function**: no two distinct elements in the given domain have the same image under f.

  $$\forall x_1, x_2 \in D_f, \quad x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

  Use **horizontal line test** to check for one-one functions: f is one-one if horizontal line $y = k$ cuts the graph at most once for all $k \in \mathbb{R}$ (or $y = k$ cuts the graph at exactly one point for all $k \in R_f$).

  Conversely, to show that a function is not one-one, provide a *specific counter-example* of a horizontal line that cuts the graph at more than one point.

- **Inverse function** of f is

  $$f^{-1}(y) = x \iff f(x) = y \text{ for all } x \in D_f.$$

  Domain and range: $D_{f^{-1}} = R_f$, $R_{f^{-1}} = D_f$

  For inverse function f^{-1} to exist, f is a one-one function (check using horizontal line test).

  To find the inverse of a given function, make x the *subject*.

  Remark. If a function is a quadratic, either complete the square, or apply quadratic formula, to make x the subject. When taking square root, consider both signs and then choose the sign by looking domain of f.

  Exercise

  Show that the inverse function of

  $$f(x) = (x-3)^2 + 1, \quad x \geq 3$$

  exists. Find f^{-1}.

  Solution. Since every horizontal line $y = k$, $k \geq 1$ cuts the graph $y = f(x)$ exactly once, f is one-one, so f^{-1} exists.

  Let $y = (x-3)^2 + 1$, then

  $$\begin{aligned} (x-3)^2 &= y - 1 \\ x - 3 &= \pm\sqrt{y-1} \\ x &= 3 \pm \sqrt{y-1} \end{aligned}$$

  Since $x \geq 3$, we choose the appropriate sign:

  $$x = 3 - \sqrt{y-1}.$$

  Hence $f^{-1}(x) = 3 - \sqrt{x-1}$, $x \geq 1$. □

  Graph sketching: If (a, b) is a point on $y = f(x)$, then (b, a) is a point on $y = f^{-1}(x)$. Hence f and f^{-1} are <u>reflections</u> of each other in the line $y = x$, and so $y = f(x)$, $y = f^{-1}(x)$ and $y = x$ intersect at one same point.

- Given functions f and g, the **composite function** is

  $$gf(x) := g\left(f(x)\right).$$

  For gf to exist, $R_f \subseteq D_g$.
  (The domain of function g must include values of the range of f, so that the function g is well-defined as every element in its domain, as well as in R_f, is mapped to something.)

  Domain and range: $D_{gf} = D_f$, $R_{gf} = R_g$

  To determine the range of a composite function gf,

  - **two-stage mapping**:

    $$D_{gf} = D_f \xrightarrow{f} R_f \xrightarrow{g} R_{gf}$$

    and check answers by graphing out f and g.
  - sketch the graph of gf, then find the range of gf based on the domain of gf (i.e. domain of f)

  Exercise

  Given

  $$\begin{aligned} f &: x \to \frac{2+x}{2-x}, \quad -2 \leq x \leq 1, \\ g &: x \to x^2 + 2x + 2, \quad x \in \mathbb{R} \end{aligned}$$

  show that gf exists and find the range of gf.

  Solution. Using GC, $R_f = [0, 3]$ and $R_g = [1, \infty)$. Since $[0, 3] = R_f \subseteq D_g = \mathbb{R}$, gf exists.

  By two-stage mapping,

  $$[-2, 1] \xrightarrow{f} [0, 3] \xrightarrow{g} [2, 17]$$

  Hence $R_{gf} = [2, 17]$. □

  Identity function: returns the same value, which was used as its input.

  $$f^{-1}f(x) = ff^{-1}(x) = x$$

  Remark. Even though the composite functions $f^{-1}f$ and ff^{-1} have the same rule, they may have different domains. $D_{f^{-1}f} = D_f$ but $D_{ff^{-1}} = D_{f^{-1}}$.

2 Graphs

- Features to include in graph sketch:
  1. **Stationary points**
     - Maximum point
     - Minimum point
     - Point of inflexion
  2. **Intercepts**
  3. **Asymptotes**
     - Horizontal asymptote: line $y = a$ where $x \to \pm\infty$, $y \to a$
     - Vertical asymptote: line $x = a$ where $x \to a$, $y \to \pm\infty$
     - Oblique asymptote: line $y = ax + b$ where $x \to \pm\infty$, $y - (ax + b) \to 0$

  To determine the restriction on possible values of x or y, use *discriminant*.

- Conic sections
  - **Circle** with centre (h, k) and radius r:
    $$(x-h)^2 + (y-k)^2 = r^2 \tag{1}$$
  - **Ellipse** with centre (h, k) and semi-major axis of a units, semi-minor axis of b units:
    $$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \tag{2}$$
  - **Horizontal hyperbola** with centre (h, k):
    $$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \tag{3}$$
    Distance from turning points to centre: a units in x-direction
    Equations of oblique asymptotes:
    $$y - k = \pm\frac{b}{a}(x-h)$$
  - **Vertical hyperbola** with centre (h, k):
    $$\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1 \tag{4}$$
    Distance from turning points to centre: b units in y-direction
    Equations of oblique asymptotes:
    $$y - k = \pm\frac{b}{a}(x-h)$$
  - **Regular parabola** with vertex (h, k):
    $$y = a(x-h)^2 + k \tag{5}$$
    Sideways parabola with vertex (h, k):
    $$x = a(y-k)^2 + h \tag{6}$$
    Rectangular hyperbola
    $$y = \frac{ax+b}{cx+d} \tag{7}$$
    has vertical asymptote $x = -\frac{d}{c}$, horizontal asymptote $y = \frac{a}{c}$.
    $$y = ax + b + \frac{c}{dx+e} \tag{8}$$
    has vertical asymptote $x = -\frac{e}{d}$, oblique asymptote $y = ax + b$.

- Parametric equations
  - To sketch graph, change mode from **function** to **parametric**, and take note of the domain of parameter t.
  - To find Cartesian equation (contain only x and y), eliminate the parameter t by solving the equations simultaneously.
  - To convert Cartesian equation to parametric, trigonometric expressions are involved, Pythagorean trigonometric identity $\sin^2 x + \cos^2 x = 1$ may be handy.

- **Transformations**

  Recommended order: **T**ranslation, **S**caling, **R**eflection

  To get $cf(bx+a)+d$ from $f(x)$,

$$f(x) \xrightarrow{1} f(x+a) \xrightarrow{2} f(bx+a) \xrightarrow{3} cf(bx+a) \xrightarrow{4} cf(bx+a)+d$$

  1. translate by a units in negative x direction,
  2. scale by factor of $\frac{1}{b}$ parallel to x-axis,
  3. scale by factor of c parallel to the y-axis,
  4. translate by d units in positive y direction.

  Remark. Transformations in x- and y-directions are independent of one another.

Type	Equation	Replacement	Graph
Translation	$y = f(x) + a$	$y \mapsto y - a$	Translate a units in positive y-direction
	$y = f(x) - a$	$y \mapsto y + a$	Translate a units in negative y-direction
	$y = f(x - a)$	$x \mapsto x - a$	Translate a units in positive x-direction
	$y = f(x + a)$	$x \mapsto x + a$	Translate a units in negative x-direction
Reflection	$y = -f(x)$	$y \mapsto -y$	Reflect in x-axis
	$y = f(-x)$	$x \mapsto -x$	Reflect in y-axis
Scaling	$y = af(x)$	$y \mapsto \dfrac{y}{a}$	Scale by factor of a parallel to y-axis
	$y = f\left(\dfrac{x}{a}\right)$	$x \mapsto \dfrac{x}{a}$	Scale by factor of a parallel to x-axis
Modulus	$y = \lvert f(x) \rvert$	$y \mapsto \lvert y \rvert$	Reflect $y < 0$ in the x-axis
	$y = f(\lvert x \rvert)$	$x \mapsto \lvert x \rvert$	Ignore $x < 0$, keep and reflect $x \geq 0$ in the y-axis
Reciprocal	$y = \dfrac{1}{f(x)}$	$y \mapsto \dfrac{1}{y}$	x-intercept becomes vertical asymptote (and vice versa), maximum point becomes minimum point (and vice versa), horizontal asymptote $y = a$ becomes horizontal asymptote $y = \frac{1}{a}$
Derivative	$y = f'(x)$	–	Vertical asymptote remains the same, horizontal aymptote $y = a$ becomes horizontal asymptote $y = 0$ (x-axis), oblique asymptote $y = mx + c$ becomes horizontal asymptote $y = m$, stationary point (a, b) becomes x-intercept $x = a$, point of inflexion (increasing/decreasing function) becomes turning point (max/min)

Problem. The plane p_2 is obtained by first translating $p_1 : \mathbf{r} \cdot \begin{pmatrix} 8 \\ -1 \\ 4 \end{pmatrix} = -5$ by 2 units in the positive z-direction, and then reflecting in the xz-plane. Obtain an equation of p_2 in scalar product form.

Solution. Since transformations are involved, consider replacement of variables.
Cartesian form of p_1 is $8x - y + 4z = -5$. Then

$$8x - y + 4z = -5 \xrightarrow{y \mapsto -y} 8x - (-y) + 4z = -5 \xrightarrow{z \mapsto z-2} 8x - (-y) + 4(z-2) = -5$$

which simplifies to $8x + y + 4z = 3$, or, in scalar product form,

$$\mathbf{r} \cdot \begin{pmatrix} 8 \\ 1 \\ 4 \end{pmatrix} = 3.$$

□

3 Equations and Inequalities

Equations

- **Systems of linear equations** can be solved using GC PlySmlt2.
- Types of solutions:
  1. Unique solution
  2. Infinitely many solutions
  3. No solutions
- Questions often involve practical problems, from which systems of linear equations are set up.

Inequalities

- Steps to solve inequalities:
  1. Move all terms to one side
  2. Either numerator or denominator is always positive (need to show working)
  3. If (2) does not apply, use **test-value method**
     - Indicate the critical value(s) on a number line.
     - Choose an x-value within each interval as the *test-value*.
     - Plug in the test-value to evaluate whether the polynomial is positive or negative within that interval.

     Use **graphical method** only if question permits use of GC
     - For inequalities in the form of $P(x) > 0$, sketch graph of $y = P(x)$ and its x-intercepts, identify region of graph where $P(x) > 0$.
     - More generally, for inequalities in the form of $P(x) > Q(x)$, sketch graphs of $y = P(x)$ and $y = Q(x)$, then identify region of graph where the inequality holds.
  4. An alternative method is to multiply the square of a term in the denominator.
  5. **Solutions of related inequalities**
     Replace x with some expression of x using the solutions of inequalities solved in earlier parts of the question.

  Remark. Important points to take note when solving inequalities:

  - Do not cross multiply without knowing whether terms are positive or not.
  - Know the difference between "and" and "or", i.e. intersection and union of sets.
  - Solutions should not be equal to roots of denominator.
- **Modulus function**:
$$|x| = \begin{cases} x & (x \geq 0) \\ -x & (x < 0) \end{cases} \tag{9}$$
  From the definition of modulus, we have
$$\begin{aligned} |x| < a &\iff -a < x < a \\ |x| > a &\iff x > a \text{ or } x < -a \end{aligned}$$

Problem. Without using a graphic calculator, solve the inequality

$$\frac{2x^2+1}{2-x^2} \le \frac{2+x^2}{x^2}.$$

[4]

Solution. Let $u = x^2$. Then the given inequality reduces to

$$\frac{2u+1}{2-u} \le \frac{2+u}{u}.$$

Moving terms to one side and factorising,

$$\frac{(3u+4)(u-1)}{u(2-u)} \le 0.$$

Idenfity critical values

$$u \le -\frac{4}{3} \text{ (rej.)} \quad \text{or} \quad 0 < u \le 1 \quad \text{or} \quad u > 2$$

thus

$$x^2 > 2 \quad \text{or} \quad x^2 \le 1$$

and so

$$\boxed{x > \sqrt{2} \quad \text{or} \quad x < -\sqrt{2} \quad \text{or} \quad -1 \le x \le 1, x \ne 0}$$

□

Problem. The inequality $\dfrac{ax^2+bx+c}{x^2-3x+9} < 0$ is satisfied for all real values of x such that a, b and c are constants. State a relation involving a, b and c and also the range of values of a. [4]

Solution. Since denominator

$$x^2 - 3x + 9 = \left(x - \frac{3}{2}\right)^2 + \frac{27}{4} \ge \frac{27}{4} > 0 \quad (\forall x \in \mathbb{R})$$

thus numerator $ax^2 + bx + c < 0$ for all $x \in \mathbb{R}$. By discriminant, $\boxed{b^2 - 4ac < 0}$. Obvious that $\boxed{a < 0}$. □

4 Sequences and Series

- A **sequence** is a set of numbers, denoted by (u_n). A sequence can be generated by giving a formula for the n-th term, e.g. $u_n = f(n)$.

  To describe the behaviour of a sequence,

  1. **Trend**:
     (strictly) increasing: $u_{n+1} > u_n$
     (strictly) decreasing: $u_{n+1} < u_n$
     alternating
  2. **Convergence**:
     convergent: as $n \to +\infty$, $u_n \to L$
     divergent: as $n \to +\infty$, $u_n \to +\infty$ or $u_n \to -\infty$

  To find $\lim\limits_{n\to\infty} \dfrac{P(n)}{Q(n)}$ where $P(n)$ and $Q(n)$ are polynomials, divide numerator and denominator by $\max\{\deg P, \deg Q\}$.

- A **series** is the sum of terms of a sequence. Sum of n terms is denoted by S_n.

  To find the term for sequence when given the series,

  $$u_n = S_n - S_{n-1}. \tag{10}$$

  For sum to infinity S_∞ to exist, the *series* converges; conversely, S_∞ does not exist if the *series* diverges.

- **Arithmetic progression**

  $$\begin{aligned} u_n &= a + (n-1)d \\ S_n &= \frac{n}{2}[2a + (n-1)d] \end{aligned} \tag{11}$$

  To show that u_n is an AP, show that $u_n - u_{n-1}$ is a constant.

- **Geometric progression**

  $$\begin{aligned} u_n &= ar^{n-1} \\ S_n &= \frac{a(1-r^n)}{1-r}, \quad |r| < 1 \\ &= \frac{a(r^n-1)}{r-1}, \quad |r| > 1 \end{aligned} \tag{12}$$

  To show that u_n is a GP, show that $\dfrac{u_n}{u_{n-1}}$ is a constant.

  For $|r| < 1$, taking limits, sum to infinity is

  $$S_\infty = \frac{a}{1-r}. \tag{13}$$

  For sum to infinity of a GP to exist, GP converges so common ratio $|r| < 1$.

- Summation series

  For standard algebraic series,

  $$\begin{aligned} \sum_{r=1}^{n} r &= \frac{n(n+1)}{2} \\ \sum_{r=1}^{n} r^2 &= \frac{n(n+1)(2n+1)}{6} \\ \sum_{r=1}^{n} r^3 &= \frac{n^2(n+1)^2}{4} \end{aligned}$$

  Remark. If lower limit is not 1, change lower limit using $\sum\limits_{r=m}^{n} u_r = \sum\limits_{r=1}^{n} u_r - \sum\limits_{r=1}^{m-1} u_r$.

  Method of differences: given general term $u_r = f(r) - f(r-1)$,

  $$\begin{aligned} \sum_{r=1}^{n} u_r =& \sum_{r=1}^{n} (f(r) - f(r-1)) \\ =& \cancel{f(1)} - f(0) + \\ & \cancel{f(2)} - \cancel{f(1)} + \\ & \vdots \\ & \cancel{f(n-1)} - \cancel{f(n-2)} + \\ & f(n) - \cancel{f(n-1)} \\ =& f(n) - f(0) \end{aligned}$$

  Remark. Must show **diagonal** cancellation of intermediate terms in the working.

- When solving questions with practical scenarios, use a table to tabulate values e.g. n and u_n.

5 Differentiation

- Differentiation rules
  - Scalar multiplication: $(kf)' = kf'$
  - Sum/Difference rule: $(f \pm g)' = f' \pm g'$
  - Product rule: $(fg)' = f'g + fg'$
  - Quotient rule: $\left(\dfrac{f}{g}\right)' = \dfrac{gf' - fg'}{g^2}$
  - Chain rule: $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\mathrm{d}y}{\mathrm{d}u}\dfrac{\mathrm{d}u}{\mathrm{d}x}$
- New functions
  - Exponential functions

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x \ln a$$

  - Logarithmic functions

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \frac{1}{x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\log_a x = \frac{1}{x \ln a}$$

  - Trigonometric functions

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos x = -\sin x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan x = \sec^2 x$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{cosec} x = -\operatorname{cosec} x \cot x \qquad \text{(MF26)}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\sec x = \sec x \tan x \qquad \text{(MF26)}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cot x = -\operatorname{cosec}^2 x$$

  - Inverse trigonometric functions

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}} \qquad \text{(MF26)}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos^{-1} x = -\frac{1}{\sqrt{1-x^2}} \qquad \text{(MF26)}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan^{-1} x = \frac{1}{1+x^2} \qquad \text{(MF26)}$$

Remark. Remember to apply chain rule whenever a function $f(x)$, instead of x, is involved.

- Applications
  - Increasing and decreasing function

$f' > 0$	f increasing
$f' < 0$	f decreasing
$f' = 0$	f stationary

  - Concavity

$f'' > 0$	f concave
$f'' < 0$	f convex

  - Nature of stationary points

First derivative test:

$f'(a^-)$	$f'(a)$	$f'(a^+)$	nature
< 0	0	> 0	minimum point
> 0	0	< 0	maximum point
> 0	0	> 0	inflexion point
< 0	0	< 0	inflexion point

Second derivative test:

$f''(a)$	nature
> 0	minimum point
0	unable to determine
< 0	maximum point

  - Tangents and normals

Equation of **tangent** to $y = f(x)$ at (a, b) is

$$y - b = f'(a)(x - a).$$

Equation of **normal** to $y = f(x)$ at (a, b) is

$$y - b = -\frac{1}{f'(a)}(x - a).$$

  - Optimisation problems

To maximise/minimise $A(x)$ as x varies,

1. Solve $\frac{\mathrm{d}A}{\mathrm{d}x}$ to find stationary values of A. (If there are two variables involved, reduce the equation to a function of 1 variable, through the use of another restriction equation obtained.)
2. Use 1st or 2nd derivative test to check for the nature of stationary points.

  - Connected rates of change – use chain rule
  - Implicit differentiation – use chain rule when differentiating variables involving y
  - Parametric differentiation: given $x(t)$ and $y(t)$, use chain rule to obtain

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}}.$$

Problem (N2022/I/10(a)). A curve C has equation $y = ax + b + \dfrac{a+2b}{x-1}$, where a and b are real constants such that $a > 0$, $b \neq \dfrac{1}{2}a$.

Given that C has no stationary points, use differentiation to find the relationship between a and b. [3]

Solution.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a - \frac{a+2b}{(x-1)^2}$$

Since C has no stationary point, $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$ has no solution.

$$(x-1)^2 = \frac{a+2b}{a}$$

has no solution, so

$$\frac{a+2b}{a} < 0$$

which simplifies to $\boxed{a < -2b}$. ☐

6 Maclaurin Series

- **Maclaurin series**:
$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdots \tag{14}$$
To obtain Maclaurin series of a function $f(x)$,
    1. find $f'(x), f''(x), \ldots$ until the desired order,
    2. substitute $x = 0$ into $f'(x), f''(x), \ldots,$
    3. substitute $f'(0), f''(0), \ldots$ into the Maclaurin expression above.

  Remark. It is often useful to apply implicit differentiation in obtaining Maclaurin series.

- Expansion of **standard series** and their validity range:
$$\begin{aligned} (1+x)^n &= 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!} + \cdots \\ e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \\ \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \\ \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \end{aligned} \tag{MF26}$$

Expanding in terms of ascending powers of x,
$$(a+x)^n = a^n\left(1+\frac{x}{a}\right)^n = a^n\left[1 + n\left(\frac{x}{a}\right) + \frac{n(n-1)}{2!}\left(\frac{x}{a}\right)^2 + \frac{n(n-1)(n-2)}{3!}\left(\frac{x}{a}\right)^3 + \cdots\right],$$
which has validity range $\left|\frac{x}{a}\right| < 1$ i.e. $|x| < |a|$.
Expanding in terms of descending powers of x,
$$(a+x)^n = x^n\left(1+\frac{a}{x}\right)^n = x^n\left[1 + n\left(\frac{a}{x}\right) + \frac{n(n-1)}{2!}\left(\frac{a}{x}\right)^2 + \frac{n(n-1)(n-2)}{3!}\left(\frac{a}{x}\right)^3 + \cdots\right],$$
which has validity range $\left|\frac{a}{x}\right| < 1$ i.e. $|x| > |a|$.

- **Small angle approximation** for sufficiently small x such that *higher powers of x can be neglected.*
$$\begin{aligned} \sin x &\approx x \\ \cos x &\approx 1 - \frac{x^2}{2} \\ \tan x &\approx x \end{aligned} \tag{15}$$

Remark. When x is small, $\sin\left(x + \frac{\pi}{3}\right) \neq x + \frac{\pi}{3}$ as $x + \frac{\pi}{3}$ is not small; so need to use compound angle formula.

Remark. When there is reciprocal, binomial series is involved. For example,
$$\frac{\sin x}{\cos x + 1} \approx \frac{x}{2 - \frac{x^2}{2}} = x\left(2 - \frac{x^2}{2}\right)^{-1}$$
and then apply binomial expansion.

Remark. Be familiar with sine and cosine rule.

Problem (HCI 2009 P1 Q3)**.** Expand $\dfrac{x^2+2x}{2x^2+1}$ in ascending powers of x up to and including the term in x^5. State the range of values of x for which this expansion is valid. [3]

Find, in the simplest form, the coefficient of x^{2017} in this expansion. [2]

Solution.

$$\begin{aligned}
&(x^2+2x)(2x^2+1)^{-1}\\
&=(x^2+2x)(1+2x^2)^{-1}\\
&=(x^2+2x)\left(1-2x^2+\frac{(-1)(-2)}{2!}(2x^2)^2+\cdots\right)\\
&=(x^2+2x)(1-(2x^2)^1+(2x^2)^2-(2x^2)^3+\cdots)\\
&=(x^2+2x)(1-2x^2+4x^4+\cdots)\\
&=x^2-2x^4+2x-4x^3+8x^5+\cdots\\
&=2x+x^2-4x^3-2x^4+8x^5+\cdots\\
&\approx \boxed{2x+x^2-4x^3-2x^4+8x^5}
\end{aligned}$$

Validity range: $\left|2x^2\right|<1$, or $\boxed{-\dfrac{1}{\sqrt{2}}<x<\dfrac{1}{\sqrt{2}}}$.

Finding coefficient of x^{2017}, an odd power:

$$(x^2+2x)(\cdots+(2x^2)^{1008}+\cdots)$$

The only way to obtain x^{2017} is through the product of $2x$ and $(2x^2)^{1008}$. Thus coefficient of x^{2017} is $\boxed{2^{1009}}$. □

Problem. Using the standard series in MF26, show

$$e^{i\theta}=\cos\theta+i\sin\theta.$$

Solution.

$$\begin{aligned}
e^{i\theta}&=1+(i\theta)+\frac{(i\theta)^2}{2!}+\frac{(i\theta)^3}{3!}+\frac{(i\theta)^4}{4!}+\frac{(i\theta)^5}{5!}+\cdots\\
&=\left(1+\frac{(i\theta)^2}{2!}+\frac{(i\theta)^4}{4!}+\frac{(i\theta)^6}{6!}\right)+\left(i\theta+\frac{(i\theta)^3}{3!}+\frac{(i\theta)^5}{5!}+\frac{(i\theta)^7}{7!}\right)\\
&=\left(1-\frac{\theta^2}{2!}+\frac{\theta^4}{4!}-\frac{\theta^6}{6!}+\cdots\right)+i\left(\theta-\frac{\theta^3}{3!}+\frac{\theta^5}{5!}-\frac{\theta^7}{7!}\right)\\
&=\cos\theta+i\sin\theta.
\end{aligned}$$

□

7 Integration Techniques

Note that the arbitrary constant C will be omitted throughout this text.

Systematic integration

$$\int f'(x)[f(x)]^n \, \mathrm{d}x = \frac{[f(x)]^{n+1}}{n+1}$$

$$\int \frac{f'(x)}{f(x)} \, \mathrm{d}x = \ln |f(x)|$$

$$\int f'(x) e^{f(x)} \, \mathrm{d}x = e^{f(x)}$$

- Trigonometric functions

$$\int \sin x \, \mathrm{d}x = -\cos x$$

$$\int \cos x \, \mathrm{d}x = \sin x$$

$$\int \sec^2 x \, \mathrm{d}x = \tan x$$

$$\int \operatorname{cosec}^2 x \, \mathrm{d}x = -\cot x$$

$$\int \sec x \tan x \, \mathrm{d}x = \sec x$$

$$\int \operatorname{cosec} x \cot x \, \mathrm{d}x = -\operatorname{cosec} x$$

$$\int \tan x \, \mathrm{d}x = \ln |\sec x| \qquad \text{(MF26)}$$

$$\int \sec x \, \mathrm{d}x = \ln |\tan x + \sec x| \qquad \text{(MF26)}$$

$$\int \operatorname{cosec} x \, \mathrm{d}x = \ln |\operatorname{cosec} x - \cot x| \qquad \text{(MF26)}$$

$$\int \cot x \, \mathrm{d}x = \ln |\sin x| \qquad \text{(MF26)}$$

Some transformations (double-angle formula and Pythagorean formula):

$$\sin 2x = 2 \sin x \cos x \qquad \text{(MF26)}$$

$$\cos 2x = 2\cos^2 x - 1 = 1 - 2\sin^2 x \qquad \text{(MF26)}$$

$$\sin^2 x + \cos^2 x = 1$$

$$\tan^2 x + 1 = \sec^2 x$$

More importantly,

$$\cos^2 x = \frac{1 + \cos 2x}{2}$$

$$\sin^2 x = \frac{1 - \cos 2x}{2}$$

Sum to product:

$$\sin P + \sin Q = 2 \sin \frac{P+Q}{2} \cos \frac{P-Q}{2} \qquad \text{(MF26)}$$

$$\sin P - \sin Q = 2 \cos \frac{P+Q}{2} \sin \frac{P-Q}{2} \qquad \text{(MF26)}$$

$$\cos P + \cos Q = 2 \cos \frac{P+Q}{2} \cos \frac{P-Q}{2} \qquad \text{(MF26)}$$

$$\cos P - \cos Q = -2 \sin \frac{P+Q}{2} \sin \frac{P-Q}{2} \qquad \text{(MF26)}$$

- Algebraic fractions

$$\int \frac{1}{\sqrt{a^2 - x^2}} \, \mathrm{d}x = \sin^{-1} \frac{x}{a} \qquad \text{(MF26)}$$

$$\int \frac{1}{a^2 + x^2} \, \mathrm{d}x = \frac{1}{a} \tan^{-1} \frac{x}{a} \qquad \text{(MF26)}$$

- Partial fractions

$$\int \frac{1}{x^2 - a^2} \, \mathrm{d}x = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| \qquad \text{(MF26)}$$

$$\int \frac{1}{a^2 - x^2} \, \mathrm{d}x = \frac{1}{2a} \ln \left| \frac{a+x}{a-x} \right| \qquad \text{(MF26)}$$

- **Integration by substitution**

Given $x = \phi(t)$, $f(x)$ is a function of t, and

$$\int f(x) \, \mathrm{d}x = \int g(t) \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t \qquad (16)$$

where $g = f \circ \phi$.

Remark. For definite integrals, remember to change the limits after making the substitution.

- **Integration by parts**

$$\int uv' = uv - \int u'v \qquad (17)$$

Guideline on choosing "u":

1. **L**ogarithmic
2. **I**nverse trigonometric
3. **A**lgebraic
4. **T**rigonometric
5. **E**xponential

DI method:

Exercise

Evaluate

$$\int x^2 e^x \, \mathrm{d}x \, .$$

Solution. Choose x^2 as "D", and e^x as "I".

sign	D	I
$+$	x^2	e^x
$-$	$2x$	e^x
$+$	2	e^x
$-$	0	e^x

Multiplying terms diagonally and summing gives

$$\int x^2 e^x \, \mathrm{d}x = x^2 e^x - 2xe^x + 2e^x + C.$$

□

8 Applications of Integration

- Integral is a **limit of sum**.

  To approximate area under $y = f(x)$ over $[0,1]$, taking right hand heights,

  $$\int_0^1 f(x)\,\mathrm{d}x = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right). \tag{18}$$

- Area under curve

  Area of required region should always be positive, so add a negative sign to an integral that is negative.

  For area between curve and x-axis,

  1. Sketch the curve
  2. Observe where the curve cuts the x-axis.
  3. Split the area into one above x-axis (positive), one below x-axis (negative).

  For area between curve and y-axis,

  1. Sketch the curve
  2. Observe where the curve cuts the y-axis.
  3. Split the area into one right of y-axis (positive), one left of y-axis (negative).

  When 2 or more curves are involved,

  1. Evaluate point(s) of intersection.
  2. Split the area into different parts.

  Area between curves $f(x) \geq g(x)$ is

  $$\int_a^b [f(x) - g(x)]\,\mathrm{d}x \tag{19}$$

  Area between curves $f(y) \geq g(y)$ is

  $$\int_a^b [f(y) - g(y)]\,\mathrm{d}y \tag{20}$$

  Remember this as: “upper” minus “lower” curve

  Parametrically, given $x = h(t)$ and $y = g(t)$,

  $$\int_a^b y\,\mathrm{d}x = \int_{t_1}^{t_2} g(t)\frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{d}t \tag{21}$$

  where t_1 and t_2 are values of t when $x = a$ and $x = b$ respectively. Similarly,

  $$\int_c^d x\,\mathrm{d}y = \int_{t_3}^{t_4} h(t)\frac{\mathrm{d}y}{\mathrm{d}t}\,\mathrm{d}t \tag{22}$$

  where t_3 and t_4 are values of t when $y = c$ and $y = d$ respectively.

- **Solid of revolution**

  Rotate region bounded by $y = f(x)$, x-axis, $x = a$, $x = b$ by angle of 2π around x-axis:

  $$\pi \int_a^b y^2\,\mathrm{d}x = \pi \int_a^b \left[f(x)\right]^2 \mathrm{d}x \tag{23}$$

  Rotate region bounded by $x = f(y)$, y-axis, $y = a$, $y = b$ by angle of 2π around y-axis:

  $$\pi \int_a^b x^2\,\mathrm{d}y = \pi \int_a^b \left[f(y)\right]^2 \mathrm{d}y \tag{24}$$

  Remark. Standard shapes: volume of cone is $\frac{1}{3}\pi r^2 h$, volume of cylinder is $\pi r^2 h$.

  Rotate region bounded by two curves $y = f(x)$ and $y = g(x)$ by angle of 2π around x-axis:

  $$\pi \int_a^b \left[f(x)\right]^2 - \left[g(x)\right]^2 \mathrm{d}x \tag{25}$$

  Rotate region bounded by two curves $x = f(y)$ and $x = g(y)$ by angle of 2π around y-axis:

  $$\pi \int_a^b \left[f(y)\right]^2 - \left[g(y)\right]^2 \mathrm{d}y \tag{26}$$

  Remark. Take note if solid of resolution is hollow; in that case, subtract away some volume.

9 Differential Equations

- 1st order differential equation
  - **Direct integration**, if in the form $\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)$
    One arbitrary constant expected.
  - **Separation of variables**, if in the form $\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y)$
    Manipulate to give $\int \frac{1}{g(y)}\,\mathrm{d}y = \int f(x)\,\mathrm{d}x$.
  - **Substitution** (will be provided in question)
- 2nd order differential equation
  - **Direct integration**, if in the form $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = f(x)$
    Two arbitrary constants expected.

Exercise
Find the general solution of

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y}{2x(x+1)}.$$

Solution.

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{y}{2x(x+1)} \\
\int \frac{2}{y}\,\mathrm{d}y &= \int \frac{1}{x(x+1)}\,\mathrm{d}x = \int \left(\frac{1}{x} - \frac{1}{x+1}\right)\mathrm{d}x \\
2\ln y &= \ln x - \ln(x+1) + C \\
&\boxed{y^2 = \frac{Cx}{x+1}}
\end{aligned}$$

for some arbitrary constant C. □

Exercise
Use the substitution $z = x + y$ to solve the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \cos(x+y).$$

Solution. From $z = x + y$,

$$\frac{\mathrm{d}z}{\mathrm{d}x} = 1 + \frac{\mathrm{d}y}{\mathrm{d}x} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}x} - 1$$

Substituting this and solving by separable variables,

$$\begin{aligned}
\frac{\mathrm{d}z}{\mathrm{d}x} &= 1 + \cos z \\
\int \frac{1}{1+\cos z}\,\mathrm{d}z &= \int \mathrm{d}x \\
\frac{1}{2}\sec^2\frac{z}{2}\,\mathrm{d}z &= x + C \quad \text{[double angle formula]} \\
\tan\frac{z}{2} &= x + C
\end{aligned}$$

Substituting back gives us

$$\boxed{\tan\frac{x+y}{2} = x + C}$$

where C is an arbitrary constant. □

10 Vectors

Basic properties

- **Magnitude** of vector $\mathbf{a}$ is $|\mathbf{a}|$. **Unit vector** of $\mathbf{a}$ is
$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|}.$$
Convention: $\mathbf{i}$ is unit vector along x-axis, $\mathbf{j}$ is unit vector along y-axis, $\mathbf{k}$ is unit vector along z-axis.
Ratio theorem: point P divides AB in ratio of $\lambda : \mu$
$$\overrightarrow{OP} = \frac{\mu \cdot \overrightarrow{OA} + \lambda \cdot \overrightarrow{OB}}{\mu + \lambda} \tag{27}$$
Midpoint theorem: special case where P is midpoint
$$\overrightarrow{OP} = \frac{\overrightarrow{OA} + \overrightarrow{OB}}{2}$$
Equal vectors: $\mathbf{a} = \mathbf{b}$ iff they have the same magnitude and direction.
Parallel vectors: $\mathbf{a} \parallel \mathbf{b}$ iff $\mathbf{a} = \lambda\mathbf{b}$ for some non-zero $\lambda \in \mathbb{R}$.
Collinear points: A, B, C are collinear iff $\overrightarrow{AB} \parallel \overrightarrow{AC}$.
Coplanar vectors: $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are coplanar iff $\mathbf{a} = \lambda\mathbf{b} + \mu\mathbf{c}$ for some $\lambda, mu \in \mathbb{R}$ (one of the vectors can be expressed as a *unique linear combination* of the other non-parallel, non-zero two vectors).
Parallelogram: $OACB$ is parallogram iff $\overrightarrow{OA} = \overrightarrow{BC}$ (two opposite sides are equal).

- **Dot product** (or scalar product) of $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ is defined as
$$\mathbf{a} \cdot \mathbf{b} := \sum_{i=1}^{n} a_i b_i. \tag{28}$$
For 3D vectors, it can be shown that
$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta. \tag{29}$$
Remark. The angle between two vectors which either *both* point outwards or inwards.
Remark. Geometrically, dot product measures the *alignment* between two vectors.
Properties:
  - Commutative: $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
  - Distributive:
  $(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{c} + \mathbf{d}) = \mathbf{a} \cdot \mathbf{c} + \mathbf{a} \cdot \mathbf{d} + \mathbf{b} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{d}$
  $(\lambda\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\lambda\mathbf{b}) = \lambda(\mathbf{a} \cdot \mathbf{b})$
  - Associative: $(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c})$

- For 3D vectors, **cross product** is defined as
$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta\, \hat{\mathbf{n}}. \tag{30}$$
In terms of computation,
$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}. \tag{MF26}$$
Remark. After computing $\mathbf{c} = \mathbf{a} \times \mathbf{b} = \mathbf{c}$, check that $\mathbf{a} \cdot \mathbf{c} = 0$ and $\mathbf{b} \cdot \mathbf{c} = 0$.
Remark. Geometrically, cross product produces a *new vector* perpendicular to the two vectors.
Properties:
  - Not commutative; in fact $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. Also $|\mathbf{a} \times \mathbf{b}| = |\mathbf{b} \times \mathbf{a}|$.
  - Distributive: $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c})$
  $(\lambda\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda\mathbf{b}) = \lambda(\mathbf{a} \times \mathbf{b})$
  - Not associative: $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$

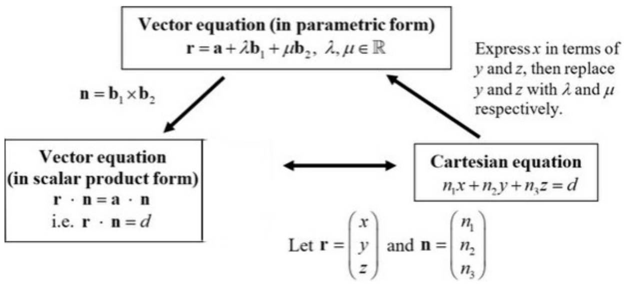
- Applications
  - Length of vector:
  $$\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$$
  - Perpendicular vectors:
  $$\mathbf{a} \cdot \mathbf{b} = 0 \quad \text{or} \quad |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}|$$
  - Parallel vectors:
  $$\mathbf{a} \times \mathbf{b} = \mathbf{0} \quad \text{or} \quad \mathbf{a} \cdot \mathbf{b} = \pm|\mathbf{a}||\mathbf{b}|$$
  where positive sign implies same direction, negative sign implies opposite directions.
  - Angle between two vectors:
  $$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$
  If $\mathbf{a} \cdot \mathbf{b} > 0$, θ is acute; if $\mathbf{a} \cdot \mathbf{b} < 0$, θ is obtuse.
  - Length of projection of $\mathbf{a}$ onto $\mathbf{b}$:
  $$|\mathbf{a} \cdot \hat{\mathbf{b}}|$$
  Vector projection of $\mathbf{a}$ onto $\mathbf{b}$
  $$(\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}$$
  Shortest distance from point to vector or line:
  $$|\mathbf{a} \times \hat{\mathbf{b}}|$$
  - Area of parallelogram:
  $$|\mathbf{a} \times \mathbf{b}|$$
  Area of triangle:
  $$\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$$
  In both cases, $\mathbf{a}$ and $\mathbf{b}$ are two adjacent sides of parallelogram or triangle.

Lines

- Equations of lines
  1. **Vector form**
  $$\ell : \mathbf{r} = \mathbf{a} + \lambda\mathbf{m}, \lambda \in \mathbb{R} \tag{31}$$
  2. **Cartesian form**
  $$\frac{x-a_1}{m_1} = \frac{y-a_2}{m_2} = \frac{z-a_3}{m_3} (= \lambda) \tag{32}$$
  3. **Parametric form**
  $$\begin{cases} x = a_1 + \lambda m_1 \\ y = a_2 + \lambda m_2 \\ z = a_3 + \lambda m_3 \end{cases} \tag{33}$$
- Relationships between two lines $\ell_1 : \mathbf{r} = \mathbf{a}_1 + \lambda\mathbf{m}_1$ and $\ell_2 : \mathbf{r} = \mathbf{a}_2 + \lambda\mathbf{m}_2$
  1. **Parallel** (coplanar)
     Direction vectors are parallel: $\mathbf{m}_1 \parallel \mathbf{m}_2$
  2. **Intersecting** (coplanar)
     Direction vectors are not parallel: $\mathbf{m}_1 \nparallel \mathbf{m}_2$
     Solving simultaneously gives one unique solution (λ, μ)
  3. **Skew** (non-coplanar)
     Direction vectors are not parallel: $\mathbf{m}_1 \nparallel \mathbf{m}_2$
     Solving simultaneously does not give unique solution (λ, μ)
- Applications
  - (Acute) angle between two lines:
    Using dot product,
    $$\cos\theta = \frac{|\mathbf{m}_1 \cdot \mathbf{m}_2|}{|\mathbf{m}_1||\mathbf{m}_2|}.$$
  - Foot of perpendicular F from point P to line ℓ:
    Since F lies on ℓ, $\overrightarrow{OF} = \mathbf{a} + \lambda\mathbf{m}$ for some λ.
    Find $\overrightarrow{PF}$ in terms of λ. Since $\overrightarrow{PF} \perp \mathbf{m}$, using dot product,
    $$\overrightarrow{PF} \cdot \mathbf{m} = 0,$$
    solve for λ. Then substitute the value of λ into expression for $\overrightarrow{OF}$.
  - Perpendicular distance from point P to line ℓ:
    Using cross product,
    $$h = \left|\overrightarrow{AP} \times \hat{\mathbf{m}}\right|.$$
  - Length of projection of vector $\mathbf{v}$ onto line ℓ:
    Using dot product,
    $$|\mathbf{v} \cdot \hat{\mathbf{m}}|.$$
  - Point of reflection of point in line:
    Use ratio theorem

Planes

- Equations of planes
  1. **Vector equation (parametric form)**
  $$\pi : \mathbf{r} = \mathbf{a} + \lambda\mathbf{m}_1 + \mu\mathbf{m}_2, \quad \lambda, \mu \in \mathbb{R} \tag{34}$$
  2. **Vector equation (scalar product form)**
  $$\pi : \mathbf{r} \cdot \mathbf{n} = D, \quad D = \mathbf{a} \cdot \mathbf{n} \tag{35}$$
     where $\mathbf{n} = \mathbf{m}_1 \times \mathbf{m}_2$.
  3. **Cartesian equation**
  $$\pi : n_1x + n_2y + n_3z = D \tag{36}$$



- Relationships between line $\ell : \mathbf{r} = \mathbf{a} + \lambda\mathbf{m}$ and plane $\pi : \mathbf{r} \cdot \mathbf{n} = D$
  1. **Parallel and do not intersect**
     $\mathbf{m} \cdot \mathbf{n} = 0$ (ℓ parallel to π)
     $\mathbf{a} \cdot \mathbf{n} \neq D$ (no common point)
     When solving line and plane simultaneously, no solution.
  2. **Parallel and lies on plane**
     $\mathbf{m} \cdot \mathbf{n} = 0$ (ℓ parallel to π)
     $\mathbf{a} \cdot \mathbf{n} = D$ (infinitely many common points)
     When solving line and plane simultaneously, infinitely many solutions.
  3. **Intersect**
     $\mathbf{m} \cdot \mathbf{n} \neq 0$ (ℓ not parallel to π)
     When solving line and plane simultaneously, one solution.
- Relationship between two planes
  1. **Parallel**: $\mathbf{n}_1 \parallel \mathbf{n}_2$
  2. **Intersecting**: $\mathbf{n}_1 \nparallel \mathbf{n}_2$

  To find line of intersection, express both equations in cartesian form, then use GC to solve.
- Applications
  - Perpendicular distance from point P to plane
    Given point A lies on π. Using length of projection (dot product),
    $$\left|\overrightarrow{AP} \cdot \hat{\mathbf{n}}\right| = \left|\frac{D - \mathbf{a} \cdot \mathbf{n}}{|\mathbf{n}|}\right|$$

Remark. If plane passes through origin, $D = 0$ (since dot product of $\mathbf{0}$ with any vector is 0). Then perp distance from origin to plane is $\frac{|D|}{|\mathbf{n}|}$.

Alternative: Find foot of perpendicular, find vector, take magnitude

– Foot of perpendicular F from point Q to plane

Consider line ℓ_{QF} passing through Q and F, with direction vector $\mathbf{n}$.

$$\begin{cases} \pi : \mathbf{r} \cdot \mathbf{n} = D \\ \ell_{QF} : \mathbf{r} = \overrightarrow{OQ} + \lambda \mathbf{n} \end{cases}$$

Since line and plane intersect at P, solve simultaneously to find λ, substitute back to find $\overrightarrow{OF}$.

– (Acute) angle between line and plane

$$\sin\theta = \frac{|\mathbf{n} \cdot \mathbf{m}|}{|\mathbf{n}||\mathbf{m}|}.$$

– (Acute) angle between two intersecting planes

$$\cos\theta = \frac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{|\mathbf{n}_1||\mathbf{n}_2|}$$

– Relationship among three planes

Reduce to relationship between a line and a plane, **OR**

Solve simultaneously using GC:

1. unique solution: intersect at one unique point
2. infinitely many solutions: intersect at a line
3. no solution: no common point

Problem (N2023/I/3)**.** Vectors $\mathbf{a}$ and $\mathbf{b}$ are such that $\mathbf{a} \cdot \mathbf{b} = -1$. It is also given that $\mathbf{a} \times \mathbf{b} + \mathbf{a}$ is perpendicular $\mathbf{a} \times \mathbf{b} + \mathbf{b}$.

(a) Show that $|\mathbf{a} \times \mathbf{b}| = 1$. [3]

(b) Hence find the angle between the direction of $\mathbf{a}$ and the direction of $\mathbf{b}$. [3]

Solution.

(a) Since $\mathbf{a} \times \mathbf{b}$ is perpendicular to $\mathbf{a}$ and $\mathbf{b}$,

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0, \quad (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0.$$

Hence

$$\begin{aligned} (\mathbf{a} \times \mathbf{b} + \mathbf{a}) \cdot (\mathbf{a} \times \mathbf{b} + \mathbf{b}) &= 0 \\ (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} + \mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) + \mathbf{a} \cdot \mathbf{b} &= 0 \\ |\mathbf{a} \times \mathbf{b}|^2 + 0 + 0 + (-1) &= 0 \\ |\mathbf{a} \times \mathbf{b}|^2 &= 1 \end{aligned}$$

$|\mathbf{a} \times \mathbf{b}| = 1$ since $|\mathbf{a} \times \mathbf{b}| \geq 0$.

(b) Since $|\mathbf{a} \times \mathbf{b}| = 1$,

$$|\mathbf{a}||\mathbf{b}| \sin\theta = 1. \tag{1}$$

Since $\mathbf{a} \cdot \mathbf{b} = -1$ (given),

$$|\mathbf{a}||\mathbf{b}| \cos\theta = -1. \tag{2}$$

Dividing (1) by (2) gives $\tan\theta = -1$, so $\theta = 135°$.

□

Problem. Given that $|\mathbf{a}| = 1$ and $|\mathbf{b}| = 2$, and the angle between $\mathbf{a}$ and $\mathbf{b}$ is $60°$, find $|2\mathbf{a} + 3\mathbf{b}|$.

Solution.

$$\begin{aligned} |2\mathbf{a} + 3\mathbf{b}|^2 &= (2\mathbf{a} + 3\mathbf{b}) \cdot (2\mathbf{a} + 3\mathbf{b}) \\ &= 4\mathbf{a} \cdot \mathbf{a} + 6\mathbf{a} \cdot \mathbf{b} + 6\mathbf{b} \cdot \mathbf{a} + 9\mathbf{b} \cdot \mathbf{b} \\ &= 4|\mathbf{a}|^2 + 12\mathbf{a} \cdot \mathbf{b} + 9|\mathbf{b}|^2 \\ &= 4|\mathbf{a}|^2 + 12|\mathbf{a}||\mathbf{b}| \cos 60° + 9|\mathbf{b}|^2 \end{aligned}$$

and the rest follows. Hence $|2\mathbf{a} + 3\mathbf{b}| = 2\sqrt{13}$. □

11 Complex Numbers

- **Cartesian form** $a + bi$
  - Equality: equal if and only if real and imaginary parts are equal
  $$a + bi = c + di \iff a = c \text{ and } b = d$$
  - Addition and subtraction: add or subtract real and imaginary parts
  - Multiplication: expand brackets in the usual fashion, apply $i^2 = -1$
  - **Conjugate** is $\bar{z} = a - bi$
  $$z\bar{z} = |z|^2 = a^2 + b^2$$
  - Division: multiply numerator and denominator by the conjugate of denominator
  - **Argand diagram**: consists of real axis and imaginary axis. The point (a, b) represents complex number $a + bi$.
- Complex roots of polynomial equations

  Fundamental Theorem of Algebra: a degree n polynomial has n roots in $\mathbb{C}$ (including repetitions).

  Conjugate Root Theorem: complex roots of a polynomial with *real* coefficients occur in conjugate pairs.

- **Polar form**: r is **modulus** $|z|$, θ is **argument** $\arg(z)$
  - Trigonometric form
  $$z = r(\cos\theta + i\sin\theta) \tag{37}$$
  where
  $$|z| = r = \sqrt{a^2 + b^2}$$
  $$\arg(z) = \theta = \tan^{-1}\frac{b}{a} \quad (-\pi < \theta \le \pi).$$
  - Exponential form

    Euler's Formula:
  $$z = re^{i\theta} \tag{38}$$
  Euler's Identity:
  $$e^{i\pi} + 1 = 0 \tag{39}$$
  - Some useful properties:
    * $|z_1 z_2| = |z_1||z_2|$, $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}$
    * $\arg z_1 z_2 = \arg z_1 + \arg z_2$,
      $\arg \dfrac{z_1}{z_2} = \arg z_1 - \arg z_2$,
      $\arg z^n = n \arg z$
    * $\mathrm{Re}(z) = r\cos\theta = \dfrac{z + \bar{z}}{2}$,
      $\mathrm{Im}(z) = r\sin\theta = \dfrac{z - \bar{z}}{2i}$

Summation of trigonometric series

To find $\cos\theta + \cos 2\theta + \cdots + \cos n\theta = \displaystyle\sum_{k=1}^{n} \cos k\theta$, consider

$$\begin{aligned}
\sum_{k=1}^{n} (\cos k\theta + i\sin k\theta) &= \sum_{k=1}^{n} e^{ik\theta} \\
&= \frac{e^{i\theta}\left(e^{in\theta} - 1\right)}{e^{i\theta} - 1} \quad \text{[sum of geometric progression]} \\
&= e^{i\theta}\frac{e^{\frac{1}{2}n\theta}\left(e^{\frac{1}{2}in\theta} - e^{-\frac{1}{2}in\theta}\right)}{e^{\frac{1}{2}i\theta}\left(e^{\frac{1}{2}i\theta} - e^{-\frac{1}{2}i\theta}\right)} \quad \text{[technique of dividing by half of the power]} \\
&= e^{\frac{1}{2}i(n+1)\theta}\frac{\cancel{2i}\sin\frac{n}{2}\theta}{\cancel{2i}\sin\frac{1}{2}\theta}
\end{aligned}$$

Then equating real parts gives us

$$\sum_{k=1}^{n} \cos k\theta = \cos\frac{n+1}{2}\theta \times \frac{\sin\frac{n}{2}\theta}{\sin\frac{1}{2}\theta}.$$

Problem (N2022/II/3(c)). It is given that $|z| = \sqrt{3}$, $\arg z = \dfrac{7\pi}{30}$.

Find the smallest positive integer value of n for which z^n is purely imaginary. State the modulus and argument of z^n in this case, giving the modulus in the form $k\sqrt{3}$, where k is an integer. [2]

Solution. $\arg(z^n) = n \arg z = \dfrac{7n\pi}{30}$.

For z^n to be purely imaginary,

$$\begin{aligned} \arg(z^n) &= (2k+1)\frac{\pi}{2}, \quad k \in \mathbb{Z} \\ \frac{7n\pi}{30} &= (2k+1)\frac{\pi}{2} \\ n &= \frac{15(2k+1)}{7} \end{aligned}$$

For n to be a positive integer, $2k+1$ must be positive and divisible by 7, so smallest value of $k = 3$. Thus smallest $n = \dfrac{15(7)}{7} = 15$.

Modulus of z^n is given by

$$|z^{15}| = |z|^{15} = \left(\sqrt{3}\right)^{15} = 2187\sqrt{3}.$$

Argument of z^n is given by

$$\arg(z^{15}) = \frac{7(15)\pi}{30} = \frac{7\pi}{2} \equiv \frac{7\pi}{2} - 4\pi = -\frac{\pi}{2}.$$

□

Problem (N2021/II/1). One of the roots of the equation $x^3 + 2x^2 + ax + b = 0$, where a and b are real, is $1 + \frac{1}{2}i$. Find the other roots of the equation and the values of a and b. [5]

Solution. Since coefficients are all real, by conjugate root theorem, $1 - \frac{1}{2}i$ is also a root. Hence $\left[x - (1 + \frac{1}{2}i)\right]\left[x - (1 - \frac{1}{2}i)\right] = x^2 - 2x + \frac{5}{4}$ is a factor of $x^3 + 2x^2 + ax + b$. By long division, $x + 4$ is another factor, thus -4 is a root. □

Problem (N2020/I/6). The complex number z satisfies the equation

$$z^2(2+i) - 8iz + t = 0,$$

where t is a real number. It is given that one root is of the form $k + ki$, where k is real and non-zero.

Find t and k, and the other root of the equation. [8]

Solution. Since $k + ki$ is a root,

$$(k+ki)^2(2+i) - 8i(k+ki) + t = 0$$

which simplifies to

$$(-2k^2 + 8k + t) + i(4k^2 - 8k) = 0.$$

Comparing imaginary parts gives us $k = 2$ (reject $k = 0$). Comparing real parts gives us $t = -8$.

Hence

$$z^2(2+i) - 8iz + t = (2+i)\left(z - (2+2i)\right)(z - c)$$

where c is the other root. Comparing constants,

$$-8 = (2+i)\left(z - (2+2i)\right)(-c)$$

so $c = \frac{1}{5}(-2+6i)$. □

Problem (N2019/I/9).

(i) The complex number w can be expressed as $\cos\theta + i\sin\theta$.

(a) Show that $w + \dfrac{1}{w}$ is a real number. [2]

(b) Show that $\dfrac{w-1}{w+1}$ can be expressed as $k\tan\frac{1}{2}\theta$, where k is a complex number to be found. [4]

(ii) The complex number z has modulus 1. Find the modulus of the complex number $\dfrac{z-3i}{1+3iz}$. [5]

Solution.

(i) (a) We see that in exponential form, $w = e^{i\theta}$. Thus

$$\begin{aligned} w + \frac{1}{w} &= e^{i\theta} + \frac{1}{e^{i\theta}} \\ &= e^{i\theta} + e^{-i\theta} \\ &= (\cos\theta + i\sin\theta) + (\cos\theta - i\sin\theta) \\ &= 2\cos\theta \end{aligned}$$

which is real.

(b) Note the technique of dividing by half of the power.

$$\begin{aligned} \frac{w-1}{w+1} &= \frac{e^{i\theta}-1}{e^{i\theta}+1} \\ &= \frac{e^{i\frac{\theta}{2}}\left(e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}}\right)}{e^{i\frac{\theta}{2}}\left(e^{i\frac{\theta}{2}} + e^{-i\frac{\theta}{2}}\right)} \\ &= \frac{\left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right) - \left(\cos\frac{\theta}{2} - i\sin\frac{\theta}{2}\right)}{\left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right) + \left(\cos\frac{\theta}{2} - i\sin\frac{\theta}{2}\right)} \\ &= \frac{2i\sin\frac{\theta}{2}}{2\cos\frac{\theta}{2}} \\ &= i\tan\frac{\theta}{2} \end{aligned}$$

thus $k = i$.

(ii) Since $|z| = 1$, let $z = e^{i\theta}$. Then

$$\frac{1}{z} = \frac{1}{e^{i\theta}} = e^{-i\theta} = \overline{e^{i\theta}} = \bar{z}.$$

Hence

$$\begin{aligned} \left|\frac{z-3i}{1+3iz}\right| &= \frac{|z-3i|}{\left|z\left(\frac{1}{z}+3i\right)\right|} \\ &= \frac{|z-3i|}{|z|\left|\frac{1}{z}+3i\right|} \\ &= \frac{|z-3i|}{(1)|\bar{z}+3i|} \\ &= \frac{|z-3i|}{\left|\overline{z-3i}\right|} \\ &= \frac{|z-3i|}{|z-3i|} \\ &= 1. \end{aligned}$$

□

Part II

Statistics

12 Probability

Permutation and Combination

- Counting principles
  - **Addition principle**
  - **Multiplication principle**
- **Permutation**: ordered arrangement
  - Arrange n distinct objects in a row: $n!$
  - Arrange n objects in a row, where n_1 of them are identical, n_2 of them are identical, and so on:

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

  - Arrange r of n distinct objects in a row:

$$\binom{n}{r} \times r! =^n \mathrm{P}_r$$

  - Arrange n distinct objects in a circle, if positions are distinguishable:

$$\frac{n!}{n} = (n-1)!$$

    If positions are distinguishable,

$$n \times (n-1)! = n!$$

    where there are n ways to number the positions.

    Remark. Use $(n-1)!$ only for the *first* group of n objects you place at the circular table.
- **Combination**: unordered selection
  - Choose r from n distinct objects without replacement:

$$\binom{n}{r} =^n \mathrm{C}_r = \frac{n!}{(n-r)!r!}$$

  - Choose at least 1 from n distinct objects:

$$2^n - 1$$

  Remark. Any restrictions must be satisfied *first.*
- Useful techniques:
  1. **Complement**
  2. **Grouping**: if certain objects must be adjacent, remember to consider order within the group
  3. **Slotting**: if certain objects must not be adjacent, remember to consider order of slotting
  4. **Cases**: if there is no straightforward method

  Remark. When subdividing into groups of equal number, remember to divide by the number of groups.

Probability

- Operations on events: **union** $A \cup B$, **intersection** $A \cap B$, **complement** A'

  A and B are **mutually exclusive** (or disjoint) if they cannot occur simultaneously: $A \cap B = \emptyset$.
- **Probability** of an event: $P(A) = \dfrac{n(A)}{n(S)}$

  For combined events, PIE:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

  For complement, $P(A') = 1 - P(A)$
- **Conditional probability** of A given B: probability of A occurring given that B has occurred.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{40}$$

  A and B are **independent** if probability of A occuring is not affected by occurrence of B:

$$P(A \mid B) = P(A)$$

  or,

$$P(A \cap B) = P(A) \times P(B).$$

- Useful techniques:
  1. **Tree diagram**: few stages with few outcomes, events are dependent
  2. **Venn diagram**: operations on events involved
  3. **Table of outcomes**: small sample space
  4. **P&C method**: only WITHOUT replacement
  5. **Sequences and series**: for turn-by-turn situations (GP could be involved)

Problem (RI 2023 Q6)**.** A group of 5 boys and 3 girls sit at random at a round table. Find the number of arrangements so that

(a) no 2 girls are adjacent to each other, [3]

(b) all 3 girls are seated together, [3]

(c) exactly 2 of the 3 girls are adjacent to each other. [2]

Solution.

(a) Arrange the 5 boys in $(5-1)! = 24$ ways.

Then slot in each of the 3 girls into the 5 spaces between the boys in ${}^5\mathrm{P}_3 = 60$ ways.

Total number of arrangements with no 2 girls being adjacent to each other is $24 \times 60 = \boxed{1440}$.

(b) Arrange the 3 girls within a unit in $3! = 6$ ways.

Then arrange the unit of 3 girls with the 5 boys in $(6-1)! = 120$ ways.

Total number of arrangements with all 3 girls seated together is $6 \times 120 = \boxed{720}$.

(c) Total number of arrangements with exactly 2 of the 3 girls adjacent to each other is $(8-1)! - 1440 - 720 = \boxed{2880}$.

□

Problem (HCI 2022 Q8)**.** A school canteen committee consists of 4 parents, 2 student leaders and 4 teachers, chosen from 10 parents, 5 student leaders and 8 teachers.

(a) There is a married couple amongst the 10 parents. How many different canteen committees can be formed if the couple cannot serve on the committee together? [3]

The school canteen committee of 10 members has been formed.

(b) All members are to stand in a row to take a group photo with the Vice-Principal. Find the number of arrangements such that the Vice-Principal stands at the centre, both ends of the row are occupied by the students' leaders, and no two parents stand next to each other. [3]

(c) The committee members, together with the Vice-Principal, are seated at a round table with 11 chairs during lunch time. Find the probability that the parents are seated together and the teachers are separated. [3]

Solution.

(a) Total number of committees formed $= \binom{5}{2} \times \binom{10}{4} \times \binom{8}{4} = 147000$

Number of committees with the couple serving together $= \binom{5}{2} \times \binom{8}{2} \times \binom{8}{4} = 19600$

Required number of committees formed $= 147000 - 19600 = \boxed{127400}$

(b) Number of arrangements if no two parents are to stand next to each other $= 2! \times 4! \times 4! \times 3 \times 3 = \boxed{10368}$

(c) No. of circular arrangements if all parents are together and teachers are separated $= 3! \times 4! \times 4! = 3456$

Required probability $= \dfrac{3456}{10!} = \boxed{\dfrac{1}{1050}}$

□

Problem (NYJC 2022 Q6)**.** Jean has forgotten the six-character login password for her laptop. She remembers that the password consists of four distinct letters from the twenty-six letters of the alphabet A-Z and two distinct digits from the ten digits 0-9.

(i) Assuming that Jean keys in a six-character password for all her login attempts and she never repeats the same incorrect password, find the largest number of unsuccessful login attempts. [2]

(ii) Find the number of possible six-character passwords if the first four characters are distinct letters in alphabetical order. [2]

(iii) Given that the first four characters are distinct letters, and the last two characters are distinct digits, find the probability that exactly one of the four letters is a vowel. [3]

Solution.

(i) Largest number of unsuccessful login attempts $= {}^{26}\mathrm{C}_4 \times^{10} \mathrm{C}_2 \times 6! - 1 = \boxed{484379999}$

(ii) Number of passwords $= {}^{26}\mathrm{C}_4 \times^{10} \mathrm{C}_2 \times 2! = \boxed{1345500}$

(iii) Number of passwords with exactly one vowel in the first four distinct letters and last two digits are distinct $= {}^{21}\mathrm{C}_3 \times^{5} \mathrm{C}_1 \times 4! \times^{10} \mathrm{C}_2 \times 2!$

Number of passwords such that the first four letters are distinct and the last two digits are distinct $= {}^{26}\mathrm{C}_4 \times 4! \times^{10} \mathrm{C}_2 \times 2!$

Required probability $= \dfrac{{}^{21}\mathrm{C}_3 \times^{5} \mathrm{C}_1 \times 4! \times^{10} \mathrm{C}_2 \times 2!}{{}^{26}\mathrm{C}_4 \times 4! \times^{10} \mathrm{C}_2 \times 2!} = \boxed{\dfrac{133}{299}}$

□

Problem (N2023/II/6)**.** A bag contains r red counters and b blue counters, where $r > 12$ and $b > 12$. Mei randomly removes 12 counters from the bag. The probability that there are 4 red counters among Mei's 12 counters is the same as the probability that there are 3 red counters.

(a) Show that $9r + 5 = 4b$. [2]

The probabiity that there are 3 red counters among Mei's 12 counters is $\frac{5}{3}$ times the probability that there are 2 red counters.

(b) Derive an equation similar to the equation in (a) and hence find the probability that just one of the 12 counters removed is red. [6]

Solution.

(a) Since $P(4 \text{ red } 8 \text{ blue}) = P(3 \text{ red } 9 \text{ blue})$,

$$\frac{{}^{r}\mathrm{C}_4 \times^{b} \mathrm{C}_8}{{}^{r+b}\mathrm{C}_{12}} = \frac{{}^{r}\mathrm{C}_3 \times^{b} \mathrm{C}_9}{{}^{r+b}\mathrm{C}_{12}},$$

which eventually simplifies to the desired expression.

(b) Since $P(3 \text{ red } 9 \text{ blue}) = \frac{5}{3}P(2 \text{ red } 10 \text{ blue})$,

$$\frac{{}^{r}\mathrm{C}_3 \times^{b} \mathrm{C}_9}{{}^{r+b}\mathrm{C}_{12}} = \frac{5}{3} \times \frac{{}^{r}\mathrm{C}_2 \times^{b} \mathrm{C}_{10}}{{}^{r+b}\mathrm{C}_{12}},$$

which simplifies to $10r - 5b = -25$. Together with (a) we get the system of equations

$$\begin{cases} 10r - 5b = -25 \\ 95 - 4b = -5 \end{cases} \implies \begin{cases} r = 15 \\ b = 35 \end{cases}$$

Hence $P(\text{just 1 red removed}) = \dfrac{{}^{15}\mathrm{C}_1 \times^{35} \mathrm{C}_{11}}{{}^{15+35}\mathrm{C}_{12}} = \boxed{0.0516}$.

□

13 Discrete Random Variables

- Let X be a discrete random variable taking values $x_1, x_2, \dots, x_n$. The **probability distribution function** of X is the function f that maps each value x_k to the probability that $X = x_k$:

  $$f(x) = P(X = x) \quad (x = x_1, x_2, \dots, x_n)$$

  Probability distribution of X can be presented using a table:

x	x_1	x_2	$\cdots$	x_n
$P(X = x)$				

  The **cumulative distribution function** of X is

  $$P(X \le x) = \sum_{r \le x} P(X = r).$$

- **Expectation** μ:

  $$\mathrm{E}(X) = \sum_{\forall x} x P(X = x) \tag{41}$$

  or more generally,

  $$\mathrm{E}\big(g(X)\big) = \sum_{\forall x} g(x) P(X = x).$$

  For constants a and b,

  - $\mathrm{E}(a) = a$ (average of a constant is itself)
  - $\mathrm{E}(aX) = a\,\mathrm{E}(X)$
  - $\mathrm{E}(aX \pm b) = a\,\mathrm{E}(X) \pm b$
  - $\mathrm{E}(aX \pm bY) = a\,\mathrm{E}(X) \pm b\,\mathrm{E}(Y)$

- **Variance** σ^2:

  $$\mathrm{Var}(X) = \sum (x_i - \mu)^2 p_i \tag{42}$$

  A more useful equation is

  $$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \tag{43}$$

  Standard deviation: $\sigma = \sqrt{\mathrm{Var}(X)}$

  For constants a and b,

  - $\mathrm{Var}(a) = 0$ (variance of a constant is zero)
  - $\mathrm{Var}(aX) = a^2\,\mathrm{Var}(X)$
  - $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$
  - $\mathrm{Var}(aX \pm bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y)$ provided X and Y are independent of each other

Binomial distribution

- $X \sim B(n, p)$, where n is number of trials, p is probability of success.

  $$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \tag{MF26}$$

  To find **mode**, use **TABLE** function in GC to obtain value of x that gives the highest value of $P(X = x)$.

- Conditions:

  1. Finite number of trials
  2. Outcome of each trial is termed a "success" or "failure"
  3. Probability of success is same for each trial
  4. Trials are independent of each other

- Expectation: $\mathrm{E}(X) = np$

  Variance: $\mathrm{Var}(X) = np(1-p)$

Problem. $X \sim B\left(30, \frac{41}{62}\right)$. Find $P(X^3 - 22X^2 + 129X - 108 > 0)$. [3]

Solution. The cubic polynomial inequality needs to be solved before finding the required probability.

$$\begin{aligned}
&P(X^3 - 22X^2 + 129X - 108 > 0) \\
&= P(1 < X < 9) + P(X > 12) \\
&= P(X \leq 8) - P(X \leq 1) + 1 - P(X \leq 12) \\
&= \boxed{0.997}
\end{aligned}$$

Remark. Note that $P(1 < X < 9) + P(X > 12)$ is a union, and care is needed to calculate the cumulative probabilities.

□

Problem (N1982/I/13)**.** The random variable X is the number of successes in n independent trials of an experiment in which the probability of success in any one trial is p.

Show that

$$\frac{P(X = k+1)}{P(X = k)} = \frac{(n-k)p}{(k+1)(1-p)}, \quad k = 0, 1, 2, \ldots, n-1.$$

Find the most probable number of success when $n = 10$ and $p = \frac{1}{4}$.

Solution. Given that $X \sim B(n, p)$, we have

$$\frac{P(X = k+1)}{P(X = k)} = \frac{\binom{n}{k+1} p^{k+1}(1-p)^{n-(k+1)}}{\binom{n}{k} p^k (1-p)^{n-k}} = \frac{(n-k)p}{(k+1)(1-p)}$$

Given $n = 10$, $p = \frac{1}{4}$. If k is the mode,

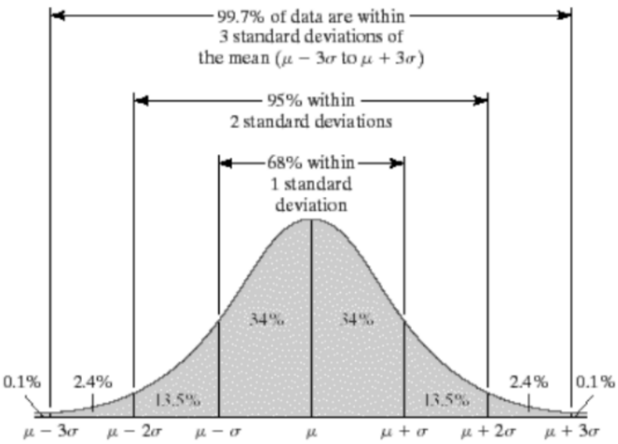
$$\begin{aligned}
&P(X = k+1) < P(X = k) \\
&\frac{P(X = k+1)}{P(X = k)} < 1 \\
&\frac{(10-k)\left(\frac{1}{4}\right)}{(k+1)\left(\frac{3}{4}\right)} < 1 \\
&10 - k < 3(k+1) \\
&7 < 4k \\
&\frac{7}{4} < k
\end{aligned}$$

$$\begin{aligned}
&P(X = k-1) < P(X = k) \\
&1 < \frac{P(X = k)}{P(X = k-1)} \\
&1 < \frac{(11-k)\left(\frac{1}{4}\right)}{(k)\left(\frac{3}{4}\right)} \\
&3k < 11 - k \\
&4k < 11 \\
&k < \frac{11}{4}
\end{aligned}$$

Hence $\frac{7}{4} < k < \frac{11}{4}$ and thus mode is $\boxed{k = 2}$. □

14 Normal Distribution

- **Normal distribution**: $X \sim N(\mu, \sigma^2)$, where μ is mean, σ^2 is variance.
- Normal curve is symmetrical:



By symmetry,

  - $P(X > a) = 1 - P(X < a)$
  - $P(X < \mu - a) = P(X > \mu + a)$
  - $P(X < \mu + a) = P(X > \mu - a)$

- Use normalcdf to evaluate $P(X < a)$.

  Use invNorm to find value of a, given $P(X < a) = p$.

- **Standard normal distribution**: $Z \sim N(0, 1)$, where $\mu = 0$, $\sigma = 1$. To standardise,

$$Z = \frac{X - \mu}{\sigma}$$

Standardise to find unknown μ or σ^2:

Exercise

$X \sim N(\mu, 5^2)$. Given that $P(X < 18) = 0.9032$, find μ.

Solution. Standardising,

$$P(X < 18) = P\left(Z < \frac{18 - \mu}{5}\right) = 0.9032.$$

Using invNorm,

$$\frac{18 - \mu}{5} = 1.30 \implies \mu = 11.5$$

$\square$

Exercise

$X \sim N(\mu, \sigma^2)$. Given that $P(X < 17) = 0.8159$ and $P(X < 25) = 0.9970$, find μ and σ.

Solution. Standardising,

$$P(X < 17) = P\left(Z < \frac{17 - \mu}{\sigma}\right) = 0.8159$$

$$P(X < 25) = P\left(Z < \frac{25 - \mu}{\sigma}\right) = 0.9970$$

Using invNorm,

$$\begin{cases} \dfrac{17 - \mu}{\sigma} = -0.8998 & \implies 17 - \mu = -0.8998\sigma \\ \dfrac{25 - \mu}{\sigma} = 2.748 & \implies 25 - \mu = 2.748\sigma \end{cases}$$

Solving simultaneously gives $\mu = 19.0$, $\sigma = 2.20$. $\square$

- Properties of expectation and variance

  Given $X \sim N(\mu_X, {\sigma_X}^2)$ and $Y \sim N(\mu_Y, {\sigma_Y}^2)$,

  - $X + Y \sim N(\mu_X + \mu_Y, {\sigma_X}^2 + {\sigma_Y}^2)$
  - $\mathrm{E}(X_1 + \cdots + X_n) = \mathrm{E}(X_1) + \cdots + \mathrm{E}(X_n) = n\,\mathrm{E}(X)$

    $\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = n\,\mathrm{Var}(X)$
  - $\mathrm{E}(aX \pm bY) = a\,\mathrm{E}(X) \pm b\,\mathrm{E}(Y)$

    $\mathrm{Var}(aX \pm bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y)$

15 Sampling

- **Random sampling**: every member of population has equal chance of being selected AND selections are independent of each other.

  Simple random sampling: each possible sample of size n has the same chance of being chosen from population of size N.

  1. Create a list of the population (sampling frame), number members of population from 1 to N.
  2. Select n distinct members using *random number generator*.

- If $X \sim N(\mu, \sigma^2)$, then

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

  where sample mean $\overline{X} = \dfrac{X_1 + \cdots + X_n}{n}$.

- If $X_1, \ldots, X_n$ is random sample of size n taken from non-normal or unknown distribution with mean μ, variance σ^2, then for sufficiently large $n \geq 50$, by **Central Limit Theorem** (CLT),

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{approximately.}$$

  Remark. Central Limit Theorem states that $\overline{X}$ is approximately normally distributed, NOT X!

  If sample size is small,

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \text{assume that } X \text{ is normally distributed.}$$

  Remark. Common mistake: remember to divide variance by n to get σ^2/n.

It is not necessary to assume that X follows a normal distribution.

Since sample size is large, Central Limit Theorem can be applied such that the distribution of the *sample mean* $\overline{X}$ is approximately normal.

- Estimation

  $\overline{x}$ is unbiased estimate of μ:

$$\overline{x} = \frac{\sum x}{n} \tag{44}$$

  s^2 is unbiased estimate of σ^2:

$$\begin{aligned} s^2 &= \frac{n}{n-1} \times \text{sample variance} \\ &= \frac{n}{n-1}\left(\frac{\sum(x-\overline{x})^2}{n}\right) \\ &= \frac{1}{n-1}\left(\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right) \end{aligned} \tag{MF26}$$

  Unbiased estimate of common population variance from two samples:

$$s^2 = \frac{\sum(x_1 - \overline{x}_1)^2 + \sum(x_2 - \overline{x}_2)^2}{n_1 + n_2 - 2} \tag{MF26}$$

  If data is represented in the form $x - c$,

$$\overline{x} = \frac{\sum(x-c)}{n} + c$$

  and

$$s^2 = \frac{1}{n-1}\left(\sum(x-c)^2 - \frac{\left(\sum(x-c)\right)^2}{n}\right)$$

  Remark. Include units for $\overline{x}$ and s^2.

16 Hypothesis Testing

- **Null hypothesis** H_0: particular claim for a *value* for the population mean (status quo claim).

  Alternative hypothesis H_1: range of values that excludes the value specified by null hypothesis (suspicion).

  p**-value**: p-value is also the lowest significance level at which H_0 is rejected (try to graph this out!)

$$\begin{aligned} p\text{-value} \leq \alpha &\implies \text{reject } H_0 \\ p\text{-value} > \alpha &\implies \text{do not reject } H_0 \end{aligned}$$

  Critical region: range of values of test statistic that leads to the rejection of H_0. The value of c which determines the critical region is known as the **critical value**.

$$\begin{aligned} \overline{x} \text{ lies in critical region} &\implies \text{reject } H_0 \\ \overline{x} \text{ does not lie in critical region} &\implies \text{do not reject } H_0 \end{aligned}$$

  Significance level α: probability of rejecting H_0 when it is actually true.

  Remark. Remember to contextualise.

- **1-tail test**: H_1 looks for increase/decrease in μ.
  - For an increase, $H_1 : \mu > \mu_0$, critical region and p-value are in **right tail**.

$$p\text{-value} = P(\overline{X} \geq \overline{x}), \quad \alpha = P(\overline{X} \geq c)$$

  - For a decrease, $H_1 : \mu < \mu_0$, critical region and p-value are in **left tail**.

$$p\text{-value} = P(\overline{X} \leq \overline{x}), \quad \alpha = P(\overline{X} \leq c)$$

  2-tail test: $H_1 : \mu \neq \mu_0$ looks for a *change* in μ, without specifying whether it is an increase or decrease.

$$\alpha = P(\overline{X} \leq c_1) + P(\overline{X} \geq c_2) = 2P(\overline{X} \leq c_1) = 2P(\overline{X} \geq c_2),$$

$$p\text{-value} = \begin{cases} 2P(\overline{X} \leq \overline{x}) & \text{if } \overline{x} < \mu_0, \\ 2P(\overline{X} \geq \overline{x}) & \text{if } \overline{x} > \mu_0. \end{cases}$$

- **Answering format for hypothesis test**:
  1. State hypotheses, define test statistics
  2. Under H_0, consider distribution of test statistic $\overline{X}$

  Critical value approach:

  3. Calculate critical value based on significance level, and test value based on sample data.
  4. If test value falls in critical region, reject H_0; otherwise, do not reject H_0.

  p-value approach:

  5. Calculate p-value based on sample data.
  6. If $p \leq \alpha$, reject H_0; otherwise, do not reject H_0.

  7. Write down conclusion in the context of question.

  Remark. When concluding a hypothesis test, it is either "reject H_0" or "do not reject H_0".

  Note that "do not reject H_0" is not equivalent to "accept H_0". This is because under the framework of hypothesis testing, there is no way to prove H_0 is true; we can only assess whether there is sufficient evidence *against* it.

17 Correlation and Regression

- **Scatter diagram**: a sketch where each axis represents a variable, each point represents an observation.
  - Need not start from $(0, 0)$.
  - Label axes according to context.
  - Indicate range of data values (minimum and maximum values).
  - Relative position of points should be accurate.

  Interpreting scatter diagram

  1. Direction: positive / negative direction
  2. Form: points lie on straight line (linear) / curve

  Example. Positive linear relationship, negative linear relationship, curvilinear relationship, no clear relationship.

- **Product moment correlation coefficient**: measures strength and direction of linear correlation between two variables

  $|r| \approx 1$ means strong linear correlation.

$$
\begin{aligned}
r &= \frac{\sum(x-\overline{x})(y-\overline{y})}{\sqrt{\left(\sum(x-\overline{x})\right)\left(\sum(y-\overline{y})\right)}} \\
&= \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right)\left(\sum y^2 - \frac{\left(\sum y\right)^2}{n}\right)}}
\end{aligned}
$$

  Remark. r is independent of units of measurement.

  Identify **outlier**(s) from scatter diagram, remove them to calculate more accurate value of r.

  Correlation does not imply causation.

- **Least squares method**
  - Regression line of y on x: line which minimises sum of squares of vertical distances from points to line

$$y - \overline{y} = b(x - \overline{x}), \quad b = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(x-\overline{x})^2}$$

(MF26)

  - Regression line of x on y: line which minimises sum of squares of horizontal distances from points to line

$$x - \overline{x} = d(y - \overline{y}), \quad d = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(y-\overline{y})^2}$$

    replace x with y and vice versa.

  Remark. Regression lines pass through mean point $(\overline{x}, \overline{y})$.

  Remark. The stronger the linear correlation, the closer the two regression lines are to each other.

  Choice of regression line is to minimise error

Case	**Estimate y given x**	**Estimate x given y**
x independent, y dependent	y on x	
y independent, x dependent	x on y	
not specified	y on x	x on y

  Reliability of estimate

  1. Appropriateness of regression line used
  2. Strength of linear correlation: $|r|$ should be close to 1 for the estimate to be reliable
  3. Interpolation or extrapolation: interpolation is likely to give a more reliable estimate than extrapolation

     Example. Since $x = 10$ lies outside the given data range ($31 \le x \le 98$), it is an extrapolation, so the linear model might not hold out of this range. Hence estimated value is unreliable.

     Since $y = 75$ lies within the given data range ($54 \le y \le 96$), it is an interpolation. Hence estimated value is reliable.

- **Transformation to linearity**
  - square transformation: $y = a + bx^2$
  - reciprocal transformation: $y = ab^x \implies \ln y = \ln a + x \ln b$
  - logarithmic transformation: $y = ax^b \implies \ln y = \ln a + b \ln x$

Problem. The table below gives the observed values of bivariate x and y.

x	20	30	34	35	36	40	42
y	32	25	a	22	26	18	19

It is given that the equation of the regression line y on x is $y = 43.5 - 0.602x$.

(a) Find the value of a correct to the nearest integer. [2]

(b) Using the result in part (a), write down the equation of regression line x on y and the value of product moment correlation coefficient between x and y. [2]

Solution.

(a) Regression line passes through mean point $(\overline{x}, \overline{y})$. $\overline{x} = \dfrac{237}{7}$, $\overline{y} = \dfrac{142+a}{7}$.

$$\begin{aligned}\overline{y} &= 43.5 - 0.602\overline{x} \\ \frac{142+a}{7} &= 43.5 - 0.602\left(\frac{237}{7}\right) \\ a &= 20 \quad \text{(nearest integer)}\end{aligned}$$

(b) From GC, $x = -1.32y + 64.3$, $r = -0.891$.

□