Engineering Mathematics 2

Dedicated to Carl Friedrich Gauss

November 18, 2020

1 VECTOR CALCULUS

1.1 Rules of vector differentiation

1. (Assume all variables from 2-5 are vectors)
2. (u + v)' = u' + v'
3. (cu)' = cu'
4. $(u \cdot v)' = u' \cdot v + u \cdot v'$
5. $(u \times v)' = u' \times v + u \times v'$
6. $grad(\alpha f + \beta g) = \alpha gradf + \beta gradg$
7. $div(\alpha \mathbf{f} + \beta \mathbf{g}) = \alpha div\mathbf{f} + \beta div\mathbf{f}$
8. $curl(\alpha \mathbf{f} + \beta \mathbf{g}) = \alpha curl\mathbf{f} + \beta curl\mathbf{f}$
9. $grad(fg) = fgradg + ggradf$
10. $div(f\mathbf{v}) = fdiv\mathbf{v} + gradf \cdot \mathbf{v}$
11. $curl(f\mathbf{v}) = fcurl\mathbf{v} + gradf \cdot \mathbf{v}$
12. $div(\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot curl\mathbf{u} - curl\mathbf{v} \cdot \mathbf{u}$
13. $divcurl\mathbf{v} = 0$
14. $curlgrad\phi = 0$
15. If $div\mathbf{v} = 0, \Rightarrow \mathbf{v} = curl\mathbf{u}$ (where **u** is a *vector* potential)
16. $curl\mathbf{v} = 0 \Leftrightarrow \mathbf{v} = \nabla\phi$ (ϕ is the scalar potential of a conservative vector field)

1.2 Scalar Fields

- f(**r**) $[\mathbb{R}^N \to \mathbb{R}]$
- For contours (isobars - const. pressure) set $f(\mathbf{r}) = const.$
- **Gradient**: **Vector differential operator** $\Rightarrow \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ **Gradient vector** $\Rightarrow grad\phi = \bigtriangledown\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial z}\right)$, where ϕ is a scalar field. **General formula** $d\phi = \nabla\phi \cdot d\mathbf{r}$ Can derive cylindrical and spherical polar form from this: **Spherical polar grad (r, θ, z)** $\Rightarrow \nabla\phi = \left(\frac{\partial\phi}{dr}, \frac{1}{r}\frac{\partial\phi}{\partial\theta}, \frac{\partial\phi}{\partial z}\right)$ and **Spherical Polar Coordinates (r,θ, ψ)**: $\Rightarrow \nabla\phi = \left(\frac{\partial\phi}{dr}, \frac{1}{r}\frac{\partial\phi}{\partial\theta}, \frac{1}{rsin\theta}\frac{\partial\phi}{\partial\psi}\right)$ Following shows it's properties:
  - Grads direction is along **normal vector** to the contour surface $\phi = const.$
  - Its magnitude $|\bigtriangledown\phi|$ gives the maximum rate of change of the the scalar field ϕ
    - Recall from Eng Maths 1 the **directional derivative** $D_{\hat{a}} = \hat{a} \cdot \bigtriangledown\phi$ which from the standard dot product rule is also $= |\hat{a}||\bigtriangledown\phi|cos(\theta)$. Now, since $D_{\hat{a}}$ by definition $= \frac{d\phi(\mathbf{r}(t))}{dt}\left(= \frac{\partial\phi}{\partial x}\frac{dx}{dt} + \frac{\partial\phi}{\partial y}\frac{dy}{dt} + \frac{\partial\phi}{\partial z}\frac{dz}{dt}\right) = \hat{a} \cdot \bigtriangledown\phi$ from the chain rule (good to remember!) then ϕ's maximum rate of change, which is $\frac{d\phi(\mathbf{r}(t))}{dt}$ will be $|\bigtriangledown\phi|$ if $\hat{a}$ is parallel to $\bigtriangledown\phi$ since $|\hat{a}| = 1$ and $cos(0) = 1$. [note, $\mathbf{r}(t) = p_0 + t\hat{a}$, ($p_0$- point of evaluation, $t\hat{a}$- direction) is the direction vector, so it natural the directional **derivative** will be defined as above]. Note the distinction between 'to' and 'in' the direction of **a**
      - Thus, the direction of max rate of change of a scalar field at a point is just $\nabla\phi$ at that point
      - And Maximum value of directional derivative at that point is just $|\nabla\phi|$ at that point.
  - Not constant in space
  - Applications:

* Equation of tangent plane to a surface: Recall that the equation of a plane is $(\mathbf{r}-\mathbf{r}_0)\cdot\hat{\mathbf{n}} = 0$. Now since grad is in the same direction as the normal vector to the contour surface $\hat{\mathbf{n}} = \frac{\bigtriangledown\phi}{|\bigtriangledown\phi|}$, the plane is simply $(\mathbf{r}-\mathbf{r}_0)\cdot\frac{\bigtriangledown\phi}{|\bigtriangledown\phi|} = 0$, which also turns out to be equivalent to doing $(\mathbf{r}-\mathbf{r}_0)\cdot\bigtriangledown\phi|_{r=r_0} = 0$.
* If ϕ is a **temperature** field then **heat** flows in the direction$-\bigtriangledown\phi$.
* If ϕ is a **pressure** field then**wind** flows in the direction$-\bigtriangledown\phi$.
* $\mathbf{F} = \bigtriangledown V$ - force is the gradient of a potential

1.3 Vector Fields

- $\underline{\mathrm{f}}(\mathbf{r})$ $[\mathbb{R}^N \to \mathbb{R}^N,]$. Note, magnitude is $|\mathbf{f}| = \sqrt{v_1^2+v_2^2+v_3^2}$ and direction of field arrows is $\theta_1 = tan^{-1}\left(\frac{v_2}{v_1}\right)$ and $\theta_2 = tan^{-1}\left(\frac{v_3}{v_1}\right)$ [Essentially taking all angles relative to the x-axis (i)]
- **Div** $[\mathbb{R}^3 \to \mathbb{R}]$ $div\mathbf{f} = \bigtriangledown\cdot\mathbf{f} = \left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\cdot(f_1,f_2,f_3) = \frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z}$
  - If $\bigtriangledown\cdot\mathbf{f} = 0$ then the vector field is **incompressible** or **solenoidal**.
- **Curl** $[\mathbb{R}^3 \to \mathbb{R}^3]$ $curl\mathbf{f} = \nabla\times\mathbf{f} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} = \left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_x}{\partial z}-\frac{\partial F_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_y}{\partial x}-\frac{\partial F_x}{\partial y}\right)\mathbf{k}$
  - If $curl\mathbf{f} = 0$ then a flow is called **irrotational** and a force **conservative**.

1.4 Stationary points of multi-variable surfaces

- Find stationary points from $\bigtriangledown\phi = 0$
- Then compute Hessian (matrix of curvature) $\mathbf{H_{i,j}} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. If the Hessian is not diagonal, i.e. the eigenvalues can be found directly, then you can also use the transform $H = V^{-1}\Lambda V$, where $\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$ or just calculate the eigenvalues using the polynomial method.
  - If all eigenvalues of the Hessian are positive then the stationary point is a minimum.
  - If all eigenvalues of the Hessian are negative then the stationary point is a maximum.
  - If the eigenvalues of the Hessian are of opposite sign then the stationary point is a saddle point.

1.5 Lagrangian Multipliers

- Used as a method of optimising a scalar field to a given constraint. The idea is that the constraint and the function that is being optimised will have the same unit vector **gradient** at the optimal point. $\nabla f(\mathbf{r}) = \lambda\nabla g(\mathbf{r})$, where λ is the Lagrangian multiplier.

1.6 Curvilinear coordinates

1. Cylindrical Coordinates (r,θ,z)
   (a) $x = rcos\theta$
   (b) $y = rsin\theta$
   (c) $z = z$
2. Spherical Coordinates (r,θ,ψ)
   (a) $x = rsin\theta cos\psi$
   (b) $y = rsin\theta sin\psi$
   (c) $z = rcos\theta$

1.7 Integration along curves

- Path integral of a vector field (work integral): $\int_C \mathbf{v}(\mathbf{r}) \cdot d\mathbf{r} = \int_C \left[\mathbf{v}(\mathbf{r}) \cdot \frac{d\mathbf{r}(t)}{dt}\right] dt$
  - If $\mathbf{v}$ is a <u>conservative</u> vector field (and finite between the limits of the integral), then the integral is <u>independent</u> of the path taken and can be expressed as $\Rightarrow \int_C \mathbf{v}(\mathbf{r}) \cdot d\mathbf{r} = \int_a^b \nabla\phi d\mathbf{r} = \phi(b) - \phi(a)$. Furthermore, $\oint_C \mathbf{v} \cdot d\mathbf{r} = 0$
- Path integral of a scalar field $\int_C f(\mathbf{r})ds = \int_C f(\mathbf{r})|d\mathbf{r}| = \int_C f(\mathbf{r}) \left|\frac{d\mathbf{r}}{dt}\right| dt$. Note for arc-length $f(\mathbf{r}) = 1$
- Both are decomposable (if $C = C_1 + C_2$) and have linearity properties.

1.8 Central Force

- A force directed towards an origin. These are conservative forces and the following property holds: $\mathbf{F} = \frac{df(r)}{dr}\hat{\mathbf{r}}$, where $f(r)$ is a scalar function and also the potential $\phi(\mathbf{r})$

1.9 Integrating areas and volumes in scalar fields

- Double Integral: $\int_{y=c}^{d} \int_{x=p(y)}^{q(y)} f(x,y)dxdy$. Key thing are the limits! Properties: Linearity, Decomposability, Separable [ONLY if limits are constant (rectangle, or circle) and $f(x,y) = g(x)h(y)$]
- Applications: Area $A = \int\int_R dxdy$ (i.e. when $\rho = 1$, Mass $M = \int\int_R \rho(x,y)dxdy$, Centre of gravity $\underline{r}_c = \begin{pmatrix}\bar{x} \\ \bar{y}\end{pmatrix} = \frac{1}{M}\int\int_R \begin{pmatrix}x \\ y\end{pmatrix}\rho(x,y)dxdy$, Moment of Inertia $\begin{pmatrix}I_x \\ I_y\end{pmatrix} = \int\int_R \begin{pmatrix}y^2 \\ x^2\end{pmatrix}\rho(x,y)dxdy$, $I_z = \int\int_R [(x-a)^2 + (y-b)^2]\rho(x,y)dxdy$ (at point (a,b))
- Jacobian Matrix (Use for coordinate swapping) $dA = dxdy = \begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix} dudv$ (where $x = x(u,v), y = y(u,v)$)
- Exact same applies for triple integrals - except with one added variable. For moment of inertia you need to make sure to add z^2: $\begin{pmatrix}I_x \\ I_y \\ I_z\end{pmatrix} = \int\int\int_R \begin{pmatrix}y^2+z^2 \\ x^2+z^2 \\ x^2+y^2\end{pmatrix}\rho(x,y,z)dxdydz$

. Furthermore the Jacobian matrix is extended to $dV = dxdydz = \begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w}\end{vmatrix} dudvdw$

1.10 Surface Integrals

1.10.1 Parametrisation of surfaces

- Surfaces are 2-dimensional objects so we only need two parameters to describe them $\mathbf{r}(u,v) = x(u,v)\mathbf{i} + y(u,v)\mathbf{j} + z(u,v)\mathbf{k}$. (Recall for lines we only need one parameter $r(\mathbf{t})$)
- If $z = f(x,y)$ then $(x,y) = (u,v)$ and $r(u,v) = u\mathbf{i} + v\mathbf{j} + f(u,v)\mathbf{k}$. For cylinder of radius a: $r(u,v) = (acos(u), asin(u), v)$ $[(u,v) = (\theta, z)]$For a cone: $r(u,v) = (vcos(u), vsin(u), v)$. For a sphere of radius a: $r(u,v) = (asin(u)cos(v), asin(u)sin(v), acos(u))$ [standard polar coordinates $[(u,v) = (\theta,\phi)]$]. If a surface is off centre, calculate parametrisation as if it were at the centre and then shift it using by the required amount in each direction.

- Now you need to find the area element $\mathbf{dA} = (\mathbf{r}_u \times \mathbf{r}_v)dudv$ Use the formula sheet to find the area element for a sphere $r^2 sin\theta d\theta d\phi$ instead of working it out. You need to know the direction of the normal vector. For surfaces think about the whether the z-component is positive or negative.

- A surface is **orientable** if a label (e.g outwards) can be a given to a normal direction at any point continuously throughout the surface.

1.10.2 In Vector Fields

- $\int\int_S \mathbf{F}(\mathbf{r}) \cdot \mathbf{dA} = \int\int_S \mathbf{F}(\mathbf{r}) \cdot \hat{\mathbf{n}} dA$

1.10.3 In Scalar Fields

- $\int\int_S f(\mathbf{r})|\mathbf{dA}| = \int\int_S f(\mathbf{r}) dA$

- To answer surface integral do the following: 1) Sketch what is going on 2) Parametrise S as $\mathbf{r}(u, v)$ 3) Determine limits on u and v 4) Find $\mathbf{dA}$ (vector) or $|\mathbf{dA}|$ (scalar) 5) Solve double integral. Note if the surface consists of multiple surfaces, then parametrise each of them separately and evaluate the flux integrals separately and finally add their answers because $S = S_1 + S_2 ...$

1.11 The fundamental theorems

- The divergence theorem relates volumes to surfaces and Stokes theorem relates surfaces to curves...

1.11.1 Gauss's Divergence Theorem

- If S is a **closed** surface bounding a volume V in a vector field $\mathbf{F}$ then: $\boxed{\int\int\int_V \nabla \cdot \mathbf{F} dV = \int\int_S \mathbf{F} \cdot \mathbf{dA}}$

- If showing something holds for 'any' surface you can use the fact that smaller volumes can be concatenated to give a larger volume whose divergence is equal to the flux surface integral of the largest volume.

1.11.2 Stoke's Theorem

- If C is a **closed** curve and S is a surface with **open** boundary at C, then for any vector field $\mathbf{F}$: $\boxed{\int\int_S \nabla \times \mathbf{F} \cdot \mathbf{dA} = \oint_C \mathbf{F} \cdot \mathbf{dr}}$

- Convention is that the curve C is traversed anti-clockwise if the surface is looked at from above.

2 LINEAR SYSTEMS AND PARTIAL DIFFERENTIAL EQUATIONS

2.1 Fourier Series

- **Even extension** - reflect on y-axis, **Odd extension** reflect on y-axis then on x-axis.

- If reflection of f(t) positive is **odd** then $a_n = 0$ (since integral of $ODD \times EVEN(cos) = ODD = 0$) and $a_0 = 0$. If reflection of f(t) positive is **even** then $b_n = 0$ (since integral of $EVEN \times ODD(sin) = 0$)

  - If f(t) is odd then use $f(t) \approx \sum_{n=1}^{\infty} b_n sin(n\omega t)$, where $\omega = \frac{2\pi}{T}$ and then $b_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t) sin(n\omega t)dt$. If odd use a_n part.

- Formulae + (useful tricks):

1. Period $T = 2\pi/\omega$, $T = 2L$, $\omega = \pi/L$
2. $f(t) \approx \dfrac{a_0}{2} + \sum_i^{\infty}(a_n cos(n\omega t) + b_n sin(n\omega t))$
3. $a_n = \dfrac{2}{T}\int_{-T/2}^{T/2} f(t)cos(n\omega t)dt$
4. $b_n = \dfrac{2}{T}\int_{-T/2}^{T/2} f(t)sin(n\omega t)dt$
5. $cos(mt)cos(nt) = \frac{1}{2}(cos([m+n]t) - cos([m-n]t)$
6. $sin(mt)cos(nt) = \frac{1}{2}(sin([m+n]t) - sin([m-n]t)$
7. Check if $n \neq 1$!
8. If you see $(-1)^n - 1$ re-write in odd/even form.

- **Half-range**

1. Chose half-range based on which does not have **sharp edges** or **discontinuities** (leads to Gibb's phenomena (overshoot) (both lead to slow Fourier series convergence)
2. Cosine - for good even extension - $a_n = \dfrac{2}{L}\int_0^L f(t)cos(\dfrac{n\pi t}{L})dt$ for $t \in [0, L]$ and $f(t) \approx \dfrac{a_0}{2} + \sum_i^{\infty}(a_n cos(n\pi t/L)$
3. Sine - if odd extension good - $b_n = \dfrac{2}{L}\int_0^L f(t)sin(\dfrac{n\pi t}{L})dt$ and $f(t) \approx \sum_i^{\infty}(b_n sin(n\pi t/L)$

2.2 Fourier Transform

[Frequency analysis of non-periodic functions] (work done through worksheets, watch missed lectures if necessary)

- F.T: $F[f(t)] = F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt$
- I.F.T $F^{-1}[f(t)] = f(t) = \dfrac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}dt$
- Properties $F(t) \leftrightarrow 2\pi f(-\omega)$ (symmetry) and $f(t - t_0) =\leftrightarrow F(\omega)e^{-j\omega t}$ (time-delay)
- Convolution $F[x(t) * y(t)] = X(\omega)Y(\omega)$
- Differentiation: $F[f^n(t)] = (j\omega)^n F(\omega)$
- Amplitude/Magnitude spectrum $= |F(\omega)| = \sqrt{Re(F(\omega))^2 + Im(F(\omega))^2}$
- Phase spectrum $= arg(F(\omega)) = tan^{-1}\left(\frac{Im(F)}{Re(F)}\right)$ - recall that $arg(z^2) = 2arg(z)$, $arg(1/z^2) = -2arg(z)$ and $arg(z_1/z_2) = arg(z_1) - arg(z_2)$ (zeros/poles ;0)
- $Y(\omega) = G(\omega)U(\omega)$, where $Y(\omega)$ is the output, $U(\omega)$ is the input and $G(\omega)$ is the frequency transfer function.
- Tips: $e^{jwt} = cos(wt) + jsin(wt)$,

2.3 Laplace Transform W13/14

- Tells us about 'growth' and 'decay' rates and stability.
- The Laplace Transform of a function $f(t)$ into the Laplace space "s" is defined as: $L[f(t)] = F(s) = \int_0^{\infty} e^{-st}f(t)dt$, where s is a (usually positive) complex variable.
- Inverse $L^{-1}[L[f(t)]] = L^{-1}[F(s)] = f(t)$
- Linearity $L[af(t) + bg(t)] = aL[f(t)] + bL[g(t)]$
- Common transforms: $L[1] = \frac{1}{s}$, $L[t^n] = \frac{n!}{s^{n+1}}$ $(L^{-1}[\frac{1}{s^{n+1}}] = \frac{t^n}{n!})$, $L[e^{-at}] = \frac{1}{s+a}$ if $Re(s+a) > 0$, $L[cos(\omega t)] = \frac{s}{s^2+\omega^2}$, $L[sin(\omega t)] = \frac{\omega}{s^2+\omega^2}$
- **Derivative Theorem:** $L[tf(t)] = -\frac{dL[f(t)]}{ds} \Rightarrow L[t^n f(t)] = (-1)^n\frac{d^n F(s)}{ds^n}$
- $L[\frac{df(t)}{dt}] = sF(s) - f(0) \Rightarrow L[\frac{d^n f}{dt^n}] = s^n F(s) - s^{n-1}f(0) - s^{n-2}f'(0) - ... - f^{n-1}(0)$

- Solving ODE's - 1. $x(t) \to X(s)$ 2. Solve for $X(s)$ 3. $X(s) \to x(t)$ (using primarily partial fractions!)
- (Heaviside function $H(t-a) = 0, t < a$ and $1, t \geq a$) - useful tool to simplify things.
- The Shifting Theorems:
  - **First shifting theorem:** $L[e^{-at}f(t)] = F(s+a)$ or as the inverse $L^{-1}[F(s+a)] = e^{-at}L^{-1}[F(s)]$
  - **Second shifting theorem:** $L[H(t-a)f(t-a)] = e^{-as}F(s)$
- Stability
  - Can express transfer function as $G(s) = \frac{Q(s)}{P(s)}$, where Q and P are polynomials in s. A system is **asymptotically stable** if the zeros of $P(s)$ are in the left-half plane (the left side of the complex Argand graph, i.e. $Re(z) < 0$). If $Re(z) > 0$, then the system is unstable.
- CONVOLUTION Theorem: $(f * g)(t) = \int_0^t f(t-\tau)g(\tau)d\tau = \int_0^t f(\tau)g(t-\tau)d\tau = F(s)G(s)$

2.4 Partial Differential Equations W15/16

- The key equations:
  - The Heat Equation. $\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2}$. Generally: $\frac{\partial u}{\partial t} = \alpha^2 \nabla^2 u$. **PARABOLIC** $\to B^2 - 4AC = 0$
    - Conditions: 2 boundary conditions + 1 initial condition
  - The Wave Equation. $\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$. $(c = \frac{T}{l})$ Generally: $\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u$. **HYPERBOLIC.**$\to B^2 - 4AC > 0$
    - Conditions: 2 boundary conditions + 2 initial condition
  - Laplace Equation. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$. Generally $\nabla^2 u = 0$. **ELLIPTIC** $\to B^2 - 4AC < 0$
    - Conditions: no initial condition, just boundary conditions for each boundary [Derichlet] (or derivative of u at boundary [Neumann] (because spacial not time based).
  - (Where A,B,C come from using $Au_{xx} + Bu_{xy} + Cu_{yy} = f(x, y, u, u_x, u_y)$ to sort 2nd order semi-linear PDE's)
- Classifying PDE's: Linear - dependent variables and their derivatives only appear in linear combinations. Linear Homogeneous - all terms contain the dependent variable. Semi-linear - Linear in highest derivative terms. Non-linear - Not linear nor semi-linear.
- **SOLVING PDE's**
  - Separation of variables method
    - Separate $u(x,t)$ into $u(x,t) = X(x)T(t)$
    - Solve ODES by setting them equal to $-k^2$ (heat and wave) or $\pm k^2$ (Laplace - depending on whether there are more boundary conditions in x or in y (need those to sums of sin and cos))find out which!
    - Solve ODE's See Eng maths 1 and practise!!
    - Find BDCs and applicable ICs for $X(x)T(t)$
    - Set $u(x,t) = \sum X(x)T(t)$ and then use final boundary condition. This usually leads to having to use a half range sine or cosine series ;)
    - Look up 'principle of linear superposition'
  - D'Alambert method
    - First differentiate between whether this is a 'semi-infinite' or 'infinite' problem.
    - Start by stating the solution $u(x,t) = f(x-ct) + g(x+ct)$
    - For an infinite domain use the two initial conditions and then solve the simultaneous equations for $f(x)$ and $g(x)$. Then substitute that back into $u(x,t)$ and changing $f(x)$ to $f(x-ct)$ and $g(x)$ to $g(x+ct)$

* For a 'semi-infinite' domain you need to be aware of the boundary conditions (namely $x \geq 0$). When substituting back into $u(x,t) = f(x-ct) + g(x+ct)$ you have to check the following conditions $x - ct > 0$ and $x - ct < 0$ (note, $x + ct > 0$ since $t > 0, x > 0$). For $x < ct$, i.e. $x - ct < 0$ use $z = -ct$ as a substitution and the final boundary condition. Finally bring it all back together and you should get a piecewise answer.

3 APPLIED STATISTICS

3.1 Rules

1. $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ then $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ and $X_1 - X_2 = N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$
2. $aX \sim N(a\mu, a^2\sigma^2)$
3. If $X_1 \sim \chi^2_{k1}$ and $X_2 \sim \chi^2_{k2}$ then $X_1 + X_2 \sim X^2_{k1+k2}$
4. From Binomial to Normal $\Rightarrow \mu = np, \sigma^2 = np(1-p) \frac{\bar{x}-\mu}{\sigma} \sim N(np, np(1-p))$ Make sure to add $(P(Z \geq))$ or subtract 0.5 $(P(Z \leq))$.

3.2 Little Eng Maths 1 things

- You can estimate a binomial distribution with a Poisson distribution if p is small and with a normal distribution if n is large.

3.3 Statistical Inference

3.3.1 Hypothesis Testing

- Key things:
  1. If the observations are from a 'simple random sample' and consist of fewer than 10% of the population, then the observations are independent.
  2. The distribution of the sample mean is well approximated by a normal model if the sample consists of more than 30 independent and identically distributed observations. (Central Limit Theorem, informal)
  3. The larger the sample size the more lenient one can be with **skewed** data.
- Let $\bar{x}$ = **sample mean** (also referred to as a point estimation of the population mean); $z*$ = standard deviation of a confidence level, e.g. 95% $\to z* = 1.96, -z* = -1.96$, 99% $\to z* = 2.58, -z* = -2.58$; $\boxed{SE = \frac{\sigma}{\sqrt{n}}}$ (Given n number of independent observations from a population with population standard deviation of σ.) $\boxed{SE \approx \frac{s}{\sqrt{n}}}$ (when the population standard deviation is unknown we can use the **sample standard deviation** s. SE is the **Standard Error** of a point estimate ($\bar{x}$) (it is the standard deviation associated with an estimate)
- Confidence Interval $= \bar{x} \pm z * SE$. If the null hypothesis is within the confidence level, we fail to reject it.
- $H_0 : \mu = \#$ (innocent unless guilty), $H_A : \mu \neq, \leq, \geq \#$ (guilty)
- Using p-values: 1) Find Z-score using $z = \frac{\bar{x} - H_0:(\mu)}{SE}$. 2) Decided on whether its a right-tailed ($p = P(z \geq \frac{\bar{x}-H_0:\mu}{SE})$), left-tailed ($p = P(z \leq \frac{\bar{x}-H_0:\mu}{SE})$), or two-tailed problem ($p = 2min(P(z \geq \frac{\bar{x}-H_0:\mu}{SE}), P(z \leq \frac{\bar{x}-H_0:\mu}{SE}))$). 3) Evaluate p and see if it is less than the significance level (usually 5%). If it is reject H_0 in favour of H_A
- Type 1 error - H_0 is true, but it has been rejected. Type 2 error - H_A is true but has been rejected.
  - Note, a 5% significance level implies that a type 1 error will occur 5% of the time.

3.3.2 Test statistics Go over understanding the normal and t-tables again!

- A test statistic is particular data summary that is useful for computing p-values and doing hypothesis tests.
- Z-score test $\Rightarrow$ Test statistic for the mean is of the general form $\frac{\bar{x}-\mu}{SE}$, where the standard error for the mean is given by $\frac{\sigma}{\sqrt{n}}$. You can use this as a way of testing if the **population standard deviation** is known.
- T-test $\Rightarrow$ Test statistic for mean when population std.dev. is unknown but sample std.dev is known $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim T_{n-1}$ where T_{n-1} is the Student's t distribution for $(n-1)$ d.o.f (n is the number of observations in the sample). Use for when sample size is less than $n = 30$, e.g. when testing if a sample is reflective of a population and H_0 is that it is representative, if there are few samples!

3.4 Inference for Numerical and Categorical data

3.4.1 Difference of two (sample) means

- $\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim T_{n-1}$ and $H_0 : \bar{x}_\Delta = \bar{x}_1 = \bar{x}_2 = 0$. Also $n = min(n_1, n_2)$ where $n_{1,2}, s_{1,2}$ are the number of observations and variance of the sets 1,2 of data. Note, if the populations standard deviations are known then this follows a normal distribution $N(0, 1)$

3.4.2 Paired data

- Used for **dependant** data.
- $T = \frac{\bar{x}_\Delta}{s/\sqrt{n}} \sim T_{n-1}$ where $\bar{x}_\Delta = \frac{1}{n}\sum X_\Delta$ and $s^2 = \frac{1}{n-1}\sum(X_\Delta - \bar{x}_\Delta)^2$, where $X_\Delta = X_i - X_j$, where X_i, X_j are paired random variables.
- $H_0 : \bar{x}_\Delta = 0$

3.4.3 Pearson's χ^2 test

- $\boxed{\sum_{i=1}^{m} \frac{(O_i - E_i)^2}{E_i} \sim \chi_d^2}$, where O_i are the observed outcomes and E_i are the expected outcomes.
  - For independence testing then d.o.f for χ^2 becomes d =(rows(number of items) - 1)(col(number of categories) - 1)
  - For categorical data $d = m - 1$, where m is the number of categories
  - For goodness-of-fit testing d = number of categories - number of calculated quantities (e.g mean when using poisson as fit) - 1.
    - H_0 - assume that the data is a good fit.

3.5 Joint distributions

- Definition $P_{X,Y}(x, y) = P((a_X \leq X \leq b_X) \cap (a_Y \leq X \leq b_Y)) = \int_{a_X}^{b_X}\int_{a_Y}^{b_Y} P_{X,Y}(x, y)dydx$ If X,Y are independent then $P_{X,Y} = P_X(x)P_Y(y)$
- Marginal distribution: $P_X(x) = \int_y P_{X,Y}(x, y)dy$ (basically means you sum the probabilities of the random variable that you are ignoring)
- Conditional distribution: $P_X(x|Y = y) = \frac{P_{X,Y}}{P_Y(y)}$ (basically the joint distribution at $Y = y$ divided by the marginal distribution for $Y = y$

- Multivariate Normal distribution $\mathbf{X} \sim N(\mu, \mathbf{\Sigma})$, where $\mathbf{\Sigma} = \frac{1}{N}\sum_i (\mathbf{Z}_{-}\mu)(\mathbf{Z}-\mu)^T$ and $\mu = [\mu_X, \mu_Y, ...]^T$ and $Z = [X, Y, ...]^T$. Unbiased sample covariance matrix is $Q = \frac{1}{n-1}\sum_i (\mathbf{u}_i - \bar{\mathbf{u}})(\mathbf{u_i} - \bar{\mathbf{u}})^T$, where $\mathbf{u}_i = [X_i, Y_i]^T$ and $\bar{u}$ is the sample mean.
- Base rate fallacy - When the mind focuses on specific information and not general information. Stop this from happening using Bayes's theorem or Joint probability

3.5.1 Correlation

- Pearson's correlation coefficient $\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$ such that $-1 \leq p \leq 1$ which tells us about the linear correlation between two random variables. Sample correlation is given by $r = \frac{q_{X,Y}}{s_X S_Y}$
- Fisher transformation $\left[\frac{1}{2}ln\left(\frac{1+r}{1-r}\right) - \frac{1}{2}ln\left(\frac{1+p}{1-p}\right)\right]\sqrt{n-3} \sim N(O, 1)$. Use this to compare the sample correlation (r) with hypothesised population correlation (p). If no correlation is claimed then $H_0 : p = 0$
- For a confidence interval (p_-, p_+), let $c = e^{2z_{\alpha/2}/\sqrt{n-3}}$ the $p_- = \frac{(1+r)-c(1-r)}{(1+r)+c(1-r)}$ and $p_+ = \frac{(1+r)-(1/c)(1-r)}{(1+r)+(1/c)(1-r)}$ - because c just switches when z changes sign.

3.6 Linear Regression

- Means of estimating parameters of a model.
- Generalised form $\mathbf{y} = \mathbf{U}\beta$, where u can be any function in x (e.g. x^2) and β are the parameters that need to be found. The residual or error is $e = \mathbf{y} - \mathbf{U}\beta$, where $||e|| = \left(\sum_i e_i^2\right)^{1/2}$. To find beta we need to minimise the error term, thus we solve $\frac{d||e||}{d\beta_i} = 0$ for $i = 1, ..., n$. E.g for a straight line fitting $y = \beta_1 + \beta_2 x$ then we minimise $||e||^2 = \sum_{i=1}^n (y_i - (\beta_1 + \beta_2 x_i))^2$ (the squared error doesn't change anything, but makes the calculations easier so just do $\frac{d||e||^2}{d\beta_i} = 0$ instead). This will then produce *Normal Equations* which you can plug into matrix form and solve for the βs.
- Note, there is a difference when in the result if you consider the errors in the x-direction or in the y-direction. Only if r (sample correlation) $= \pm 1$ will they be the same. Indeed one can show that if $y = \beta_1 + \beta_2 x$ and $x = \hat{\beta}_1 + \hat{\beta}_2 y$ then $\beta_2 \cdot \hat{\beta}_2 = r^2$
- **General Least Squares:**
  - $\mathbf{y} = \mathbf{X}\mathbf{b}$, where $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} 1 & x_1 & sin(x_2) & WE \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & sin(x_n) & WE \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$,where WE is WHATEVER function you like.
  - Error vector is thus $||\mathbf{e}||^2 = (\mathbf{y} - \mathbf{X}\mathbf{b})^T(\mathbf{y} - \mathbf{X}\mathbf{b})$ (=square matrix) and from this general solution is $\mathbf{b} = (\mathbf{X^T X})^{-1}\mathbf{X^T y}$

3.7 Inference from linear regression

- In order to use linear regression to make statistical inferences there are three conditions that need to be checked:
  - 1. Normal residues:
    - Residual def: $e_i = y_i - \sum_j \beta_j x_{i,j}$ (i.e. the difference between the **line of best fit** and the actual observation). Plot a histogram of residuals.
    - The residuals needs to follow a normal distribution well - if there are any outliers this indicates that the line of best fit using least-squares will be biased too.
  - 2. Constant variability (homoscedasticity)
    - Plot the residuals against the **independent variable**

* There should be no obvious trends in the residuals (should be normal) and variability (e.g standard deviation) should not change as the independent variable does.

- 3. Independence of observations
  - Plot i-th residual against i + 1-th residual.
  - The value of each residual should not depend on the other. This means there should be no covariance in the plot, i.e. no elliptical appearance.

3.8 Monte-Carlo Methods

- Generate random numbers according to the distribution you want → put them into your model → Calculate desired statistics → Repeat millions of times.
- Law of Large Numbers (LLN) - "the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed". In mathematical terms, the (weak) law of large numbers states that given $X_1, X_2, ...$ is an infinite sequence of random variables with the same mean $E[X_1] = E[X_2] = ... = \mu$ then for any, real number $\epsilon > 0$ we have $P(|\bar{X}_n - \mu| > \epsilon) \to 0 as n \to \infty$

Things to be aware of!

- Fallacies
  - Gamblers fallacy - 'thinking the universe will even things out' - thinking the likely distribution drives the events, when the events are by nature purely random
  - Correlation is not causation!
  - Regression to the mean - things will generally always be closer to the mean.
  - Conjunction fallacy - Assuming specific conditions are more likely than general ones.
  - Base rate fallacy! (Prosecutor's fallacy) Forgetting two or more random variables are involved
  - Ludic fallacy - thinking mathematical models are a true representation of reality.
- Also beware of Biases (sample bias, lack of data bias, publication, bias etc...) and Misleading representations!

Calculator tricks

- Sample Mean ($\bar{x}$, (unbiased) Sample standard deviation sx: $mode - stat - 1 \to inputdata \to AC \to SHIFT - stat(1) - 4 \to sx, \bar{x}$
- Sample correlation r, linear regression $\sum xetc...$: $mode - stat - 2 \to inputdata \to AC \to SHIFT - stat(1) - 5, 3 \to r, \sum xetc..$

A few extra gems

- Sum of random variables, e.g. $Z = X + Y$ has a probability density function equal to the convolution of the pdf's being summed. For example if X, Y are uniformly distributed between $[0, 1]$then the pdf of Z is $f_Z(z) = \begin{cases} z & 0 \le z \le 1 \\ 2 - z & 1 \le z \le 1 \\ 0 & otherwise \end{cases}$
- For an exponential distribution of the form $f_X(x) = ae^{-ax}$, then the mean is $\mu = 1/\alpha$ and the variance $\sigma^2 = 1/\alpha^2$
- Integration: Don't give up - try change of coordinate system, change of order of integration, integration by parts and SUBSTITUTION.