Applied Mathematics Cheat Sheet

Prelims 2021

Jiaming (George) Yu
jiaming.yu@jesus.ox.ac.uk

February 7, 2022

Contents

Part I

Introductory Calculus

1 First-Order Differential Equations

(a) *First order differential equations* are of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y)$$

(b) For ODE in the form $\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)$, we integrate directly.

(c) For ODE in the form $\frac{\mathrm{d}y}{\mathrm{d}x} = a(x)b(y)$, we use separation of variables to obtain

$$\int \frac{\mathrm{d}y}{b(y)} = \int a(x)\,\mathrm{d}x \tag{1.1}$$

Sometimes we can reduce an ODE to a separable form by substitution.

(d) For ODE which are homogeneous, that is, $\frac{\mathrm{d}y}{\mathrm{d}x} = f\left(\frac{y}{x}\right)$, we can perform the substitution $y(x) = xv(x)$ to obtain

$$x\frac{\mathrm{d}v}{\mathrm{d}x} = f(v) - v \tag{1.2}$$

(e) For first-order inhomogeneous linear ODE in the form $\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = q(x)$, we multiply by the integrating factor

$$I(x) = e^{\int p(x)\,\mathrm{d}x} \tag{1.3}$$

to obtain

$$y = I(x)\left(\int I(x)q(x)\,\mathrm{d}x + c\right) \tag{1.4}$$

2 Second-Order Differential Equations

(a) Suppose $z(x) \neq 0$ is a solution to the second-order homogeneous linear ODE

$$p(x)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + q(x)\frac{\mathrm{d}y}{\mathrm{d}x} + r(x)y = 0$$

Perform the substitution $y(x) = v(x)z(x)$ and rearrange to obtain an ODE for v':

$$p(x)zv'' + \big(2p(x)z' + q(x)z\big)v' = 0 \tag{2.1}$$

(b) For second-order homogeneous ODE with constant coefficients

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + q\frac{\mathrm{d}y}{\mathrm{d}x} + ry = 0$$

where the auxiliary equation $\lambda^2 + q\lambda + r = 0$ has roots $\lambda_1, \lambda,$

if $\lambda_1 \neq \lambda_2$ are real, then

$$y(x) = C_1e^{\lambda_1 x} + C_2e^{\lambda_2 x} \tag{2.2}$$

if $\lambda_1 = \lambda_2 = \lambda$, then

$$y(x) = (C_1x + C_2)e^{\lambda x} \tag{2.3}$$

if $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta$, then

$$y(x) = e^{\alpha x}(C_1 \cos\beta x + C_2 \sin\beta x) \tag{2.4}$$

(c) To find y_p in an inhomogeneous case, we start by trying something in the form $f(x)$, and then try the next most complicated thing by multiplying by polynomials of x or trying a more general form.

3 Partial Differentiation

(a) For $F(x,y) = f(u(x,y), v(x,y))$, we have

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}, \quad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} \tag{3.1}$$

Part II

Probability

4 Events and Probability

(a) The number of arrangements of

$$\underbrace{\alpha_1, \dots, \alpha_1}_{m_1 \text{ times}}, \dots, \underbrace{\alpha_k, \dots, \alpha_k}_{m_k \text{ times}}$$

where $m_1 + \dots + m_k = n$, is

$$\frac{n!}{m_1! \cdots m_k!} \tag{4.1}$$

This is also the <u>*multinomial coefficient*</u> of $a_1^{m_1} \cdots a_k^{m_k}$ in the expansion of $(a_1 + \dots + a_k)^n$.

(b) **Vandermonde's Identity**

For $k, m, n \geqslant 0$,

$$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j} \binom{n}{k-j} \tag{4.2}$$

(c) **Law of Total Probability**

For a partition $\{B_1, B_2, \dots\}$ of Ω with each $\mathbb{P}(B_i) > 0$ and an event $A \in \mathcal{F}$,

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i) \tag{4.3}$$

(d) **Bayes' Theorem**

For a partition $\{B_1, B_2, \dots\}$ of Ω and an event $A \in \mathcal{F}$ with $\mathbb{P}(B_i), \mathbb{P}(A) > 0$,

$$\mathbb{P}(B_k \mid A) = \frac{\mathbb{P}(A \mid B_k)\, \mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid B_k)\, \mathbb{P}(B_k)}{\sum_i \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i)} \tag{4.4}$$

(e)
- Events A and B are <u>*independent*</u> if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\, \mathbb{P}(B)$.
- A family of events $\{A_i : i \in I\}$ is <u>*independent*</u> if for any finite subset J of I,

$$\mathbb{P}\left(\bigcap_{i \in J} A_i \right) = \prod_{i \in J} \mathbb{P}(A_i) \tag{4.5}$$

- A family of events $\{A_i : i \in I\}$ is <u>*pairwise independent*</u> if $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\, \mathbb{P}(A_j)$ whenever $i \neq j$.

Importantly, pairwise independence $\not\Longrightarrow$ independence.

5 Discrete Random Variables

(a) A *discrete random variable* is a function $X : \Omega \to \mathbb{R}$ s.t.

(i) $\big\{\omega \in \Omega : X(\omega) = x\big\} \in \mathcal{F}$ for each $x \in \mathbb{R}$; (probability assignment)

(ii) $\operatorname{Im} X := \big\{X(\omega) : \omega \in \Omega\big\}$ is a countable subset of $\mathbb{R}$. (discreteness)

(b) The *probability mass function* (p.m.f.) of X is the function $p_X : \mathbb{R} \to [0, 1]$ defined by

$$p_X(x) = \mathbb{P}(X = x)$$

(c) The *expectation* of X is

$$\mathbb{E}[X] = \sum_{x \in \operatorname{Im} X} x\, p_X(x) \tag{5.1}$$

We have that

$$\mathbb{E}[h(X)] = \sum_{x \in \operatorname{Im} X} h(x)\, p_X(x) \tag{5.2}$$

(d) The *k^{th} moment* of X is $\mathbb{E}[X^k]$.

(e) The *variance* of X is

$$\operatorname{var}(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \tag{5.3}$$

We have that

$$\operatorname{var}(aX + b) = a^2 \operatorname{var}(X) \tag{5.4}$$

(f) The *conditional distribution* of X given B is given by

$$\mathbb{P}(X = x \mid B) = \frac{\mathbb{P}(\{X = x\} \cap B)}{\mathbb{P}(B)} \tag{5.5}$$

The *conditional expectation* of X given B is

$$\mathbb{E}[X \mid B] = \sum_x x\, \mathbb{P}(X = x \mid B) = \sum_x x\, p_{X|B}(x) \tag{5.6}$$

(g) **Partition Theorem for Expectations**

For a partition $\{B_1, B_2, \ldots\}$ of Ω with each $\mathbb{P}(B_i) > 0$,

$$\mathbb{E}[X] = \sum_i \mathbb{E}[X \mid B_i]\, \mathbb{P}(B_i) \tag{5.7}$$

(h) The *joint probability mass function* is given by

$$p_{X,Y}(x, y) = \mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big)$$

and the *marginal distributions* are obtained by

$$p_X(x) = \sum_y p_{X,Y}(x, y), \quad p_Y(y) = \sum_x p_{X,Y}(x, y) \tag{5.8}$$

(i) For two r.v.s, the *conditional distribution* of Y given that $X = x$ is

$$p_{Y|X=x}(y) = \mathbb{P}(Y = y \mid X = x) = \frac{p_{X,Y}(x, y)}{p_X(x)} \tag{5.9}$$

The *conditional expectation* of Y given that $X = x$ is then

$$\mathbb{E}[Y \mid X = x] = \sum_y y\, p_{Y|X=x}(y) = \frac{1}{p_X(x)} \sum_y y\, p_{X,Y}(x, y) \tag{5.10}$$

(j) • Discrete r.v.s X and Y are *independent* if

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\, \mathbb{P}(Y = y) \tag{5.11}$$

• A family of discrete r.v.s $\{X_i : i \in I\}$ is *independent* if for any finite subset J of I and collection $\big\{A_j : j \in J\big\}$ where each $A_j \subseteq \mathbb{R}$,

$$\mathbb{P}\left(\bigcap_{j \in J} \{X_j \in A_j\} \right) = \prod_{j \in J} \mathbb{P}(X_j \in A_j) \tag{5.12}$$

(k) For $h : \mathbb{R}^2 \to \mathbb{R}$ and discrete r.v.s X and Y,

$$\mathbb{E}[h(X, Y)] = \sum_x \sum_y h(x, y)\, p_{X,Y}(x, y) \tag{5.13}$$

(l) For two discrete r.v.s X and Y,

$$\mathbb{E}[aX + bY] = a\, \mathbb{E}[X] + b\, \mathbb{E}[Y] \tag{5.14}$$

and when they are independent,

$$\mathbb{E}[XY] = \mathbb{E}[X]\, \mathbb{E}[Y] \tag{5.15}$$

Notation	**Parameters**	**Image**	**P.m.f.**	$\mathbb{E}[\boldsymbol{X}]$	**var**$(\boldsymbol{X})$	$\boldsymbol{G_X(s)}$
$\mathrm{Ber}(p)$	$p \in [0, 1]$	$\{0, 1\}$	$\begin{cases} p, & k = 1 \\ 1 - p, & k = 0 \end{cases}$	p	pq	$1 - p + ps$
$\mathrm{Bin}(n, p)$	$n \in \mathbb{Z}^+, p \in [0, 1]$	$\{0, \dots, n\}$	$\binom{n}{k} p^k (1 - p)^{n-k}$	np	npq	$(1 - p + ps)^n$
$\mathrm{Geom}(p)$	$p \in [0, 1]$	$\mathbb{Z}^+$	$p(1 - p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1 - p}{p^2}$	$e^{\lambda(s-1)}$
$\mathrm{Po}(\lambda)$	$\lambda \geqslant 0$	$\mathbb{Z}_0^+$	$\dfrac{\lambda^k e^{-\lambda}}{k!}$	λ	λ	$\dfrac{ps}{1 - (1 - p)s}$

6 Difference Equations and Random Walks

(a) A k^{th}-order linear *recurrence relation* or *difference equation* has the form

$$\sum_{j=0}^{k} a_j u_{n+j} = f(n), \quad a_0, a_k \neq 0$$

(b) Given $u_{n+1} = au_n + b$.

If $a = 1$,

$$u_n = A + bn \tag{6.1}$$

Otherwise,

$$u_n = Aa^n + \frac{b}{1-a} \tag{6.2}$$

(c) Given $u_{n+1} = au_n + bn$.

If $a = 1$,

$$u_n = A + \frac{bn(n-1)}{2} \tag{6.3}$$

Otherwise,

$$u_n = Aa^n + \frac{bn}{1-a} - \frac{b}{(1-a)^2} \tag{6.4}$$

(d) Given $u_{n+1} + au_n + bu_{n-1} = f(n)$.

If its *auxiliary equation* $\lambda^2 + a\lambda + b = 0$ has distinct roots λ_1, λ_2, then the solution to its homogeneous equation is

$$w_n = A_1\lambda_1^n + A_2\lambda_2^n \tag{6.5}$$

If the auxiliary equation has repeated roots λ,

$$w_n = (A + Bn)\lambda^n \tag{6.6}$$

To find a particular solution, we start by trying something in the same form as f, but omitting any terms which is included in the solution to the homogeneous equation. If this doesn't work, we then try the next most complicated thing.

As an example, if $w_n = A \cdot 2^n + B$ and $f = 1$, we could start with a constant, but that is a special case of w_n, so we should start with $v_n = Cn$.

As another example, if $w_n = An + B$ and $f = 1$, we should start with Cn^2 as the linear and constant terms are all included in w_n.

7 Probability Generating Functions

(a) For a non-negative integer-valued r.v. X, its probability generating function is

$$G_X(s) = \mathbb{E}\left[s^X\right] = \sum_{k=0}^{\infty} s^k \, \mathbb{P}(X = k) \tag{7.1}$$

with domain of absolute convergence:

$$\mathcal{S} = \left\{ s \in \mathbb{R} : \sum_{k=0}^{\infty} |s|^k \, \mathbb{P}(X = k) < \infty \right\}$$

(b) Abbreviating $\mathbb{P}(X = k)$ as p_k, we have

$$G_X(0) = p_0, \; G'_X(0) = p_1, \text{ more generally, } G_X^{(k)}(0) = k! p_k \tag{7.2}$$

and

$$\mathbb{E}[X] = G'_X(1) \tag{7.3}$$

$$\operatorname{var}(X) = G''_X(1) + G'_X(1) - \left(G'_X(1)\right)^2 \tag{7.4}$$

(c) If X and Y are independent, then

$$G_{X+Y}(s) = G_X(s) G_Y(s) \tag{7.5}$$

(d) For independent r.v.s with $X_i \sim \text{Po}(\lambda_i)$,

$$\sum_{i=1}^{n} X_i \sim \text{Po}\left(\sum_{i=1}^{n} \lambda_i\right) \tag{7.6}$$

(e) For i.i.d. r.v.s $X_1, X_2, \ldots$ with $X_i \sim \text{Ber}(p)$ and $N \sim \text{Po}(\lambda)$ independently of X_i,

$$\sum_{i=1}^{N} X_i \sim \text{Po}(\lambda p) \tag{7.7}$$

(f) Let $X_1, X_2, \ldots$ be i.i.d. non-negative integer-valued r.v.s with p.g.f. G_X and N be a non-negative integer-valued r.v. independent of X_i and with p.g.f. G_N. Then the p.g.f. of $\sum_{i=1}^{N} X_i$ is $G_N \circ G_X$.

8 Continuous Random Variables

(a) A *continuous random variable* is a function $X : \Omega \to \mathbb{R}$ s.t. for each $x \in \mathbb{R}$,

$$\{\omega \in \Omega : X(\omega) \leqslant x\} \in \mathcal{F}$$

(b) The *cumulative distribution function* of X is the function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) = \mathbb{P}(X \leqslant x)$$

The *probability density function* of X is the function $f_X : \mathbb{R} \to \mathbb{R}$ s.t.

(i) $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(u)\,\mathrm{d}u$

(ii) $\displaystyle\int_{-\infty}^{\infty} f_X(u)\,\mathrm{d}u = 1$

(iii) $f_X(u) \geqslant 0$ for all $u \in \mathbb{R}$

Importantly, f_X is not a probability.

(c) The *expectation* of X is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x)\,\mathrm{d}x \tag{8.1}$$

We still have that

$$\mathbb{E}[h(X)] = \int_{-\infty}^{\infty} h(x)\, f_X(x)\,\mathrm{d}x \tag{8.2}$$

The variance is still defined as $\mathrm{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ and expectation and variance both behave identically to the discrete case.

(d) The *joint cumulative distribution function* is given by

$$F_{X,Y}(x, y) = \mathbb{P}\big(\{X \leqslant x\} \cap \{Y \leqslant y\}\big)$$

The *joint density function* is the function $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ s.t.

(i) $f_{X,Y}(x, y) \geqslant 0$ for all $x, y \in \mathbb{R}$

(ii) $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1$

Given sufficient smoothness of $f_{X,Y}$, we have

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y) \tag{8.3}$$

(e) The *marginal distributions* are obtained by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}y, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}x \tag{8.4}$$

(f) • Jointly continuous r.v.s X and Y are *independent* if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \tag{8.5}$$

- Jointly continuous r.v.s $X_1, \ldots, X_n$ are *independent* if

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{8.6}$$

(g) In both discrete and continuous cases,

$$\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y) + 2\operatorname{cov}(X,Y) \tag{8.7}$$

where

$$\operatorname{cov}(X,Y) = \mathbb{E}\big[\big(X - \mathbb{E}[X]\big)\big(Y - \mathbb{E}[Y]\big)\big] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] \tag{8.8}$$

Note independence implies covariance is 0, but not vice versa.

Notation	**Parameters**	**Image**	**P.d.f.**	$\mathbb{E}[\boldsymbol{X}]$	**var**$(\boldsymbol{X})$
$\mathrm{U}[a,b]$	$a, b \in \mathbb{R}$	$[a,b]$	$\begin{cases} \frac{1}{b-a}, & a \leqslant x \leqslant b \\ 0, & \text{otherwise} \end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$\operatorname{Exp}(\lambda)$	$\lambda \geqslant 0$	$\mathbb{R}_0^+$	$\lambda e^{-\lambda x}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$\operatorname{Gamma}(\alpha,\lambda)$	$\alpha > 0, \lambda \geqslant 0$	$\mathbb{R}_0^+$	$\dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$	$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$
$\operatorname{Beta}(\alpha,\beta)$	$\alpha, \beta > 0$	$\mathbb{R}^+$	$\dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}$	$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
$\mathrm{N}(\mu,\sigma^2)$	$\mu \in \mathbb{R}, \sigma^2 > 0$	$\mathbb{R}$	$\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

9 Weak Law of Large Numbers

(a) The <u>*sample mean*</u> of a random sample of size n from a distribution with mean μ and variance σ^2 is

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{9.1}$$

where

$$\mathbb{E}\big[\overline{X}_n\big] = \mu, \quad \operatorname{var}\big(\overline{X}_n\big) = \frac{1}{n}\sigma^2 \tag{9.2}$$

(b) **Weak Law of Large Numbers**

For i.i.d. r.v.s $X_1, X_2, \ldots$ with mean μ, and some $\varepsilon > 0$, as $n \to \infty$,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| > \varepsilon\right) \to 0 \tag{9.3}$$

(c) **Markov's Inequality**

For a non-negative r.v. Y whose expectation exists, and any $t > 0$,

$$\mathbb{P}(Y \geqslant t) \leqslant \frac{\mathbb{E}[Y]}{t} \tag{9.4}$$

(d) **Chebyshev's Inequality**

For a r.v. Z with finite variance, and any $t > 0$,

$$\mathbb{P}\Big(\big|Z - \mathbb{E}[Z]\big| \geqslant t\Big) \leqslant \frac{\operatorname{var}(Z)}{t^2} \tag{9.5}$$