P8104 — Probability Midterm Cheatsheet

Yongyan Liu

October 14, 2025

Notation & Basics

- Law of Total Probability: $P(A) = \sum_i P(A \mid B_i)P(B_i)$ for a partition $\{B_i\}$.
- Bayes' rule: $P(B \mid A) = \dfrac{P(A \mid B)P(B)}{P(A)}$ with $P(A) = \sum_i P(A \mid B_i)P(B_i)$.

Expectation, Variance, Covariance, MGF

- Discrete: $E[g(X)] = \sum_x g(x)\, p_X(x)$. Continuous: $E[g(X)] = \int g(x)\, f_X(x)\, dx$.
- $E(X) = \mu$; $\mathrm{Var}(X) = E[(X-\mu)^2] = E(X^2) - [E(X)]^2$.
- Linear transform: $E(aX+b) = aE(X)+b$; $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$.
- Covariance: $\mathrm{Cov}(X,Y) = E[(X-EX)(Y-EY)] = E(XY) - E(X)E(Y)$.
- If $X \perp Y$ (independent) then
  - $Cov(X,Y) = 0$. (Zero cov does **not** imply independence in general.)
  - $E(\sum_i X_i) = \sum_i E(X_i)$.
  - $Var(\sum_i X_i) = \sum_i Var(X_i)$.
  - $Var(\bar{X}) = \sigma^2/n$ for i.i.d. with variance σ^2.
  - $M_{X+Y}(t) = M_X(t)M_Y(t)$
- Definition: $M_X(t) = E(e^{tX})$ for t in a neighborhood of 0. $M_X^{(n)}(0) = E(X^n)$

Common Distributions (quick formulas)

Name	Support	pmf/pdf	Mean	Var
Bernoulli(p)	$\{0,1\}$	$P(X=1)=p$	p	$p(1-p)$
Binomial(n,p)	$\{0,\dots,n\}$	$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Geometric(p)	$\{0,1,\dots\}$ or $\{1,2,\dots\}$	$P(X=k)=(1-p)^k p$ (defn. 1)	$(1-p)/p$ or $1/p$	$(1-p)/p^2$
Poisson(λ)	$\{0,1,\dots\}$	$e^{-\lambda}\lambda^k/k!$	λ	λ
Uniform(a,b)	$[a,b]$	$1/(b-a)$	$(a+b)/2$	$(b-a)^2/12$
Exponential(λ)	$[0,\infty)$	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
Gamma(α,θ)	$[0,\infty)$	$\dfrac{1}{\Gamma(\alpha)\theta^\alpha}x^{\alpha-1}e^{-x/\theta}$	$\alpha\theta$	$\alpha\theta^2$
Normal(μ,σ^2)	$\mathbb{R}$	$\dfrac{1}{\sqrt{2\pi}\sigma}\exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$	μ	σ^2
Cauchy($0,1$)	$\mathbb{R}$	$\dfrac{1}{\pi(1+x^2)}$	— (DNE)	— (DNE)

Useful mgfs

- Binomial(n,p): $M(t) = [(1-p)+pe^t]^n$. Poisson(λ): $M(t) = \exp\{\lambda(e^t-1)\}$.
- Exponential(λ): $M(t) = \dfrac{\lambda}{\lambda - t},\ t<\lambda$. Gamma($\alpha,\theta$): $M(t) = (1-\theta t)^{-\alpha},\ t<1/\theta$.
- Normal(μ,σ^2): $M(t) = \exp\{\mu t + \frac{1}{2}\sigma^2 t^2\}$.

Transformations

Monotone differentiable transform $Y = g(X)$ (continuous):
If g is one-to-one with inverse $x = g^{-1}(y)$,

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

If not one-to-one, sum over all branches x_i with $g(x_i) = y$:

$$f_Y(y) = \sum_i f_X(x_i) \left| \frac{dx_i}{dy} \right|.$$

Linear: $Y = aX + b \Rightarrow E(Y) = aE(X) + b,\ \operatorname{Var}(Y) = a^2 \operatorname{Var}(X)$.

Joint, Marginal, Conditional

- Joint density/pmf $f_{X,Y}(x, y)$.
- Marginals: $f_X(x) = \int f_{X,Y}(x, y)\, dy$ (continuous) or $\sum_y f_{X,Y}(x, y)$ (discrete).
- Independence: $f_{X,Y}(x, y) = f_X(x) f_Y(y)$.
- Conditional density: $f_{X|Y}(x \mid y) = \dfrac{f_{X,Y}(x, y)}{f_Y(y)}$ when $f_Y(y) > 0$.
- Conditional expectation: $E[X \mid Y] = \int x\, f_{X|Y}(x \mid y)\, dx$.

Convergence & Limit Theorems (quick)

- **Markov**: For $X \geq 0$, $P(X \geq a) \leq E(X)/a$.
- **Chebyshev**: $P(|X - \mu| \geq k\sigma) \leq 1/k^2$.
- **Weak Law (WLLN)**: $\bar{X}_n \xrightarrow{P} \mu$ for i.i.d. with finite μ, σ^2.
- **CLT**: $\dfrac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$.

Common Confusions & Exam Traps

- **Independence vs uncorrelated**: Cov $= 0$ does not imply independence (except for Normal).
- **mgf existence**: heavy-tailed (e.g., Cauchy) has no mgf; don't use mgf arguments.
- **Support matters**: always state the support before integrating/jacobian transforms.
- **Poisson limit**: Binomial(n, p) with $np = \lambda$ and $p \to 0$ approximates Poisson(λ).
- **Exponential is memoryless**: $P(X > s + t \mid X > t) = P(X > s)$.

Mini Worked Reminders

1) Sum of independent Gammas
If $X_i \overset{iid}{\sim} \text{Gamma}(\alpha_i, \theta)$ with common scale θ, then $S = \sum_i X_i \sim \text{Gamma}(\sum_i \alpha_i, \theta)$ since mgfs multiply.
2) Linear transform
If $Y = 3X - 2$ and $\operatorname{Var}(X) = \sigma^2$, then $\operatorname{Var}(Y) = 9\sigma^2$.
3) Change of variables (one-to-one)
If $X \sim \text{Uniform}(0, 1)$, $Y = -\ln(1 - X) \Rightarrow Y \sim \text{Exp}(1)$ by jacobian rule.

P8104 Calculus Review

1. Definition: A function f is a rule that assigns to each element x in a set A exactly one element, called $f(x)$, in a set B, i.e. $f(x) = y$.

2. Definition: Inverse function: if f is a one-to-one function, $f(x) = y$, then its inverse function, f^{-1} is defined by $f^{-1}(y) = x$. Note: $f\left(f^{-1}(y)\right) = y$.

3. Suppose that the limits $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist, then

   (a) $\lim_{x\to a}\{f(x) \pm g(x)\} = \lim_{x\to a} f(x) \pm \lim_{x\to a} g(x)$

   (b) $\lim_{x\to a}\{f(x)\cdot g(x)\} = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$

   (c) $\lim_{x\to a}\frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$, if $\lim_{x\to a} g(x) \neq 0$

4. Theorem: $\lim_{x\to a} f(x) = L$ if and only if $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L$.

5. Definition:

   (a) A function is continuous at a number a if $\lim_{x\to a} f(x) = f(a)$.

   (b) A function is continuous from the right at a number a if $\lim_{x\to a^+} f(x) = f(a)$.

   (c) A function is continuous from the left at a number a if $\lim_{x\to a^-} f(x) = f(a)$.

6. Quadratic formula: if $ax^2 + bx + c = 0$, then $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$.

7. Definition: The derivative of $f(x)$ with respect to x is the function $f'(x)$ and is defined as $f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$.

8. Basic properties and formulas: if $f(x)$ and $g(x)$ are differentiable functions, c and n are any real numbers, $\frac{d}{dx}(c) = 0$, $(cf(x))' = cf'(x)$, $\frac{d}{dx}(x^n) = nx^{n-1}$ (Power Rule), $(f(x)\pm g(x))' = f'(x) \pm g'(x)$, $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$ (Product Rule), $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x)-f(x)g'(x)}{\{g(x)\}^2}$ (Quotient Rule), $\frac{d}{dx}\left[f\left(g(x)\right)\right] = f'\left(g(x)\right)g'(x)$ (Chain Rule)

9. Common derivatives: $\frac{d}{dx}(x) = 1$, $\frac{d}{dx}\sin x = \cos x$, $\frac{d}{dx}\cos x = -\sin x$, $\frac{d}{dx}\tan x = \sec^2 x$, $\frac{d}{dx}\sec x = \sec x\tan x$, $\frac{d}{dx}\csc x = -\csc x\cot x$, $\frac{d}{dx}\cot x = -\csc^2 x$, $\frac{d}{dx}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$, $\frac{d}{dx}\cos^{-1} x = -\frac{1}{\sqrt{1-x^2}}$, $\frac{d}{dx}\tan^{-1} x = \frac{1}{1+x^2}$, $\frac{d}{dx}\cot^{-1} x = -\frac{1}{1+x^2}$, $\frac{d}{dx}a^x = a^x\ln a$, $\frac{d}{dx}e^x = e^x$, $\frac{d}{dx}(\ln x) = \frac{1}{x}$ $(x > 0)$, $\frac{d}{dx}\log_a(x) = \frac{1}{x\ln a}$ $(x > 0)$.

10. Important facts: $\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e$, $\lim_{x\to\infty}(1+\frac{1}{x})^x = e$.

11. L'Hopital's Rule: Suppose $f'(x)$ and $g'(x)$ exist, and $g'(x) \neq 0$, then

    (a) if $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, then $\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}$.

    (b) if $\lim_{x\to a} f(x) = +\infty$ and $\lim_{x\to a} g(x) = +\infty$, then $\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}$.

(c) Example: $\lim_{x\to+\infty} \frac{\ln x}{x} = \lim_{x\to+\infty} \frac{1/x}{1} = 0$.

12. $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$, $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

13. Indefinite integrals: $\int x^n dx = \frac{1}{n+1}x^{n+1} + C$ $(n \neq -1)$, $\int \frac{1}{x} dx = \ln|x| + C$, $\int e^x dx = e^x + C$, $\int a^x dx = \frac{1}{\ln a}a^x + C$, $\int \sin x dx = -\cos x + C$, $\int \cos x dx = \sin x + C$, $\int \sec^2 x dx = \tan x + C$, $\int \csc^2 x dx = -\cot x + C$, $\int \frac{1}{1+x^2} dx = \tan^{-1} x + C$, $\int \frac{1}{\sqrt{1-x^2}} dx = \sin^{-1} x + C$.

14. The substitution rule: if $u = g(x)$ is a differentiable function, $\int f(g(x))\, g'(x) dx = \int f(u) du$.
Example: $\int_1^e \frac{\ln x}{x} dx = \int_0^1 u du = \frac{1}{2}$, where $u = \ln x$.

15. Integral by part: $\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$.
Example: $\int xe^x dx = xe^x - \int e^x dx = xe^x - e^x + C$.

16. Geometric series: $\sum_{n=1}^{\infty} ar^{n-1} = a + ar + ar^2 + \dots$
Let $S_n = a + ar + ar^2 + \cdots + ar^{n-1}$ and $rS_n = ar + ar^2 + ar^3 + \cdots + ar^n$. Taking a difference, we obtain $S_n = \frac{a(1-r^n)}{1-r}$.
If $-1 < r < 1$, $\sum_{n=1}^{\infty} ar^{n-1} = \lim_{n\to\infty} \frac{a(1-r^n)}{1-r} = \frac{a}{1-r}$.
Notes: $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$, if $|r| < 1$. $\sum_{n=1}^{\infty} ar^n = \frac{a}{1-r} - a$, if $|r| < 1$.

17. Taylor series of the function f at a: $\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$, where $f^{(n)}(a)$ is the nth order derivative at a. Example, Taylor series of function $f(x) = e^x$ at $a = 0$ is $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.