

# Probability Theory Final Exam Cheat Sheet

December 16, 2025 — 7:10pm-8:40pm — 3 Questions, 1.5 Hours

Open Book Exam - Focus on Post-Midterm 2 Material

## 0. QUICK REFERENCE GUIDE

### TERMINOLOGY TRAPS (CRITICAL!)

- “Gaussian” = Normal!  $N(\mu, \sigma^2)$
- “Gaussian vector” = MVN (Multivariate Normal)
- “Independent components” =  $\rho = 0$  (for MVN: independence!)
- “Mean  $\theta = 3$ ” (Exp)  $\Rightarrow \lambda = 1/3$  NOT 3!
- For MVN ONLY:  $\rho = 0 \Leftrightarrow$  independent
- $\psi(t)$  = MGF (professor’s notation)

### VISUAL DECISION TREE (Start Here!)

START: What type of problem?

```

|
|--[Named Distribution?]
|   |--"Gaussian/Normal" -> Sec 3.3 (single) or 4.5 (joint)
|   |--"Exponential" -> Sec 3.4 [lambda=1/mean!]
|   |--"Poisson" -> Sec 2.3
|   |--"Binomial" -> Sec 2.2
|   |--"Lognormal" -> Sec 7.3 (stock prices!)
|   |--"Beta" -> Sec 3.6 (priors!)
|
|--[Multiple Variables?]
|   |--"Joint PDF/PMF" -> Sec 4.1
|   |--"Bivariate Normal/Gaussian vector" -> Sec 4.5
|   |--"X+Y, sum" -> MGF (5.2) or direct
|   |--"Max/Min" -> Order Stats (4.7)
|   |--"X|Y=y" -> Conditional (4.2, 4.5)
|
|--[Approximation/Limits?]
|   |--"Large n/approximate" -> CLT (6.1)
|   |--"Sample mean" -> CLT (6.1)
|   |--"n games/trials" -> CLT (6.1) Template B
|
|--[Bayesian?]
|   |--"Prior/Posterior" -> Sec 7.2
|   |--"Update belief" -> Bayes (1.3, 7.2)
|   |--"Defective rate" -> Discrete Bayes (8.8)
|   |--"Monty Hall" -> Template J
|
|--[Expectation?]
|   |--"E[X|Y]" -> Cond. Expectation (7.1)

```

+--"Total Expectation" ->  $E[X]=E[E[X|Y]]$

### Emergency Quick Reference: Problem Phrase → Section

- “Gaussian” → Normal! Sec 3.3
- “Gaussian vector” → MVN! Sec 4.5
- “Independent components” →  $\rho = 0$  for MVN, Sec 4.5
- “Large n” / “Approximate” → CLT, Sec 6.1
- “i.i.d.” → Independence, maybe CLT
- “Prior/Posterior” → Bayesian, Sec 7.2
- “Update belief” → Bayes’ Theorem, Sec 1.3
- “Conjugate” → Beta-Binomial, Sec 7.2
- “Stock price” / “ $S_0 e^Z$ ” → Lognormal, Sec 7.3
- “Mean  $\theta$ ” (Exp) →  $\lambda = 1/\theta$ ! Sec 3.4
- “Memoryless” → Exponential, Sec 3.4
- “Arrival/Counting process” → Poisson, Sec 2.3
- “Max/Min of n” → Order Statistics, Sec 4.7
- “ $\psi(t)$ ” → MGF! Sec 5.1
- “Conditional distribution” → Sec 4.2, 4.5
- “ $E[X|Y]$ ” → Conditional Expectation, Sec 7.1
- “Total winnings/games” → CLT, Template B
- “Monty Hall” → Bayesian, Sec 8.1
- “Defective rate” → Bayesian, Sec 8.8

### Top 20 Critical Formulas

1.  $P(A|B) = \frac{P(A \cap B)}{P(B)}$
2.  $P(H|E) = \frac{P(E|H)P(H)}{P(E)}$  (Bayes)
3.  $P(A) = \sum P(A|B_i)P(B_i)$  (Total Prob)
4.  $E[X] = \sum xP(X = x)$  (Discrete)

$$5. E[X] = \int xf(x)dx \quad (\text{Continuous})$$

$$6. \text{Var}(X) = E[X^2] - (E[X])^2$$

$$7. \text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

$$8. \rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

$$9. P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (\text{Binomial})$$

$$10. P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad (\text{Poisson})$$

$$11. f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\text{Normal})$$

$$12. Z = \frac{X - \mu}{\sigma} \quad (\text{Standardization})$$

$$13. M(t) = E[e^{tX}] \quad (\text{MGF})$$

$$14. E[X] = E[E[X|Y]] \quad (\text{Total Expectation})$$

$$15. \text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E[X|Y])$$

$$16. Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \quad (\text{CLT})$$

$$17. \text{CI} : \bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$18. \pi(\theta|x) \propto L(x|\theta)\pi(\theta) \quad (\text{Bayes})$$

$$19. E[e^X] = e^{\mu + \sigma^2/2} \quad (\text{Lognormal})$$

$$20. f_{UV}(u, v) = f_{XY}(x, y)|J| \quad (\text{Jacobian})$$

## Common Mistakes Checklist

- ☐ Forgot continuity correction for discrete→normal
- ☐ Confused  $P(A|B)$  with  $P(B|A)$
- ☐ Didn't check independence before using formulas
- ☐ Wrong integration limits for marginals
- ☐ Forgot to normalize Bayesian posterior
- ☐ Used Binomial instead of Hypergeometric
- ☐ Forgot absolute value of Jacobian
- ☐ Assumed correlation implies causation

## 1. FUNDAMENTAL CONCEPTS

### 1.1 Probability Axioms

- **Definition:** A probability measure satisfies:
  1. Normalization:  $P(S) = 1$
  2. Non-negativity:  $P(A) \geq 0$
  3. Additivity:  $P(A \cup B) = P(A) + P(B)$  if  $A \cap B = \emptyset$
- **Key Formula:**  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **When to Use:** Basic probability calculations
- **Solution Steps:**
  1. Identify sample space  $S$
  2. Count favorable outcomes
  3. Apply formula
- **Example:** Two dice:  $P(\text{sum} = 7) = 6/36 = 1/6$
- **Common Pitfalls:** Forgetting the intersection term
- **Note:** Equally likely:  $P(A) = |A|/|S|$

### 1.2 Conditional Probability

- **Definition:** Probability of  $A$  given  $B$  occurred
- **Key Formula:**  $P(A|B) = \frac{P(A \cap B)}{P(B)}$ , if  $P(B) > 0$
- **When to Use:** "Given that", "if we know", "conditional on"
- **Solution Steps:**
  1. Identify condition  $B$  and target  $A$
  2. Find  $P(A \cap B)$  and  $P(B)$
  3. Apply formula
- **Example:** Roll dice, sum odd.  $P(\text{sum} < 8 | \text{odd}) = 2/3$
- **Common Pitfalls:** Confusing  $P(A|B)$  with  $P(B|A)$
- **Note:** *Multiplication Rule:*  $P(A \cap B) = P(B)P(A|B)$

### 1.3 Bayes' Theorem

- **Definition:** Update probability given evidence
- **Key Formula:**  $P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_j P(E|H_j)P(H_j)}$
- **When to Use:** "Update", "posterior", "given evidence"
- **Solution Steps:**
  1. List hypotheses  $H_i$  with priors  $P(H_i)$
  2. Find likelihoods  $P(E|H_i)$
  3. Apply Bayes' formula
  4. Normalize if needed
- **Example:** Monty Hall: Switch wins 2/3 of time
- **Common Pitfalls:** Wrong likelihood, forgetting to normalize
- **Note:** *Total Probability:*  $P(A) = \sum P(A|B_i)P(B_i)$

### 1.4 Independence

- **Definition:**  $A$  and  $B$  independent if  $P(A \cap B) = P(A)P(B)$
- **Key Formula:**  $P(A|B) = P(A)$  iff independent
- **When to Use:** Testing if events affect each other
- **Solution Steps:**
  1. Calculate  $P(A)$ ,  $P(B)$ ,  $P(A \cap B)$
  2. Check if  $P(A \cap B) = P(A) \cdot P(B)$
  3. State conclusion
- **Example:** Card draws with replacement are independent
- **Common Pitfalls:** Assuming independence without checking
- **Note:** *Pairwise  $\neq$  Mutual independence*

### 1.5 Counting Methods

- **Permutations:** Order matters  $P(n, k) = \frac{n!}{(n-k)!}$
- **Combinations:** Order doesn't matter  $C(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
- **Multinomial:** Multiple categories  $\frac{n!}{n_1!n_2! \dots n_k!}$
- **When to Use:** "How many ways", "arrangements", "selections"
- **Example:** 6-card poker hands from 52 cards:  $\binom{52}{6}$
- **Note:** *With replacement:*  $n^k$ ; *Without:*  $P(n, k)$  or  $C(n, k)$

## 2. DISCRETE RANDOM VARIABLES

### 2.1 PMF and CDF

- **PMF:**  $p(x) = P(X = x)$ , where  $\sum p(x) = 1$
- **CDF:**  $F(x) = P(X \leq x) = \sum_{k \leq x} p(k)$
- **Expectation:**  $E[X] = \sum x \cdot P(X = x)$
- **Variance:**  $\text{Var}(X) = E[X^2] - (E[X])^2$
- **Properties:** CDF is right-continuous, non-decreasing
- *Note:*  $P(a < X \leq b) = F(b) - F(a)$

### 2.2 Binomial Distribution (a.k.a. Binomial( $n, p$ ), “n trials”)

- **SYNONYMS:** “n trials”, “success/failure”, “fixed number of trials”
- **Definition:** Number of successes in  $n$  independent trials
- **PMF:**  $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
- **Mean:**  $E[X] = np$
- **Variance:**  $\text{Var}(X) = np(1-p)$
- **MGF:**  $M(t) = (1-p + pe^t)^n$
- **When to Use:** Fixed  $n$ , constant  $p$ , independent trials
- **Example:** Flip coin 10 times,  $P(X = 6)$  heads with  $p = 0.5$
- **! Check conditions before using!**
- *Note:* Normal approximation when  $np(1-p) > 10$

### 2.3 Poisson Distribution (a.k.a. Poisson( $\lambda$ ), Counting Process)

- **SYNONYMS:** “arrival process”, “counting process”, “rare events”, “rate  $\lambda$ ”
- **Definition:** Count of rare events in fixed interval
- **PMF:**  $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$
- **Mean:**  $E[X] = \lambda$
- **Variance:**  $\text{Var}(X) = \lambda$
- **MGF:**  $M(t) = e^{\lambda(e^t - 1)}$
- **When to Use:** Rate  $\lambda$  per unit time/space
- **Example:** Arrivals per hour, defects per batch
- **Property:** Sum of independent Poissons is Poisson
- *Note:* Approximates Binomial when  $n$  large,  $p$  small,  $np = \lambda$

### 2.4 Geometric Distribution (a.k.a. “First success”, Memoryless Discrete)

- **SYNONYMS:** “first success”, “waiting for success”, “trials until success”
- **Definition:** Number of failures before first success
- **PMF:**  $P(X = k) = p(1-p)^k, \quad k = 0, 1, 2, \dots$
- **Mean:**  $E[X] = (1-p)/p$
- **Variance:**  $\text{Var}(X) = (1-p)/p^2$
- **Memoryless:**  $P(X = m + n | X \geq m) = P(X = n)$
- **When to Use:** “First success”, “waiting time”
- **Example:** Roll die until first 6 appears
- *Note:* Alternative parameterization: trials until success

### 2.5 Negative Binomial

- **Definition:** Failures before  $r$ -th success
- **PMF:**  $P(X = k) = \binom{k+r-1}{k} p^r (1-p)^k$
- **Mean:**  $E[X] = r(1-p)/p$
- **Variance:**  $\text{Var}(X) = r(1-p)/p^2$
- **When to Use:** “ $r$ -th success”, extended geometric
- *Note:* Geometric is special case with  $r = 1$

### 2.6 Hypergeometric Distribution

- **Definition:** Sampling without replacement
- **PMF:**  $P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$
- **Parameters:**  $N$  total,  $K$  success,  $n$  sample,  $k$  observed
- **Mean:**  $E[X] = n \cdot K/N$
- **When to Use:** Finite population, no replacement
- **Example:** Draw 5 cards, probability of 3 aces
- **! Different from Binomial (with replacement)**

## 3. CONTINUOUS RANDOM VARIABLES

### 3.1 PDF and CDF

- **PDF:**  $f(x) \geq 0, \int_{-\infty}^{\infty} f(x) dx = 1$
- **CDF:**  $F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$
- **Probability:**  $P(a < X < b) = \int_a^b f(x) dx$

- **Expectation:**  $E[X] = \int_{-\infty}^{\infty} xf(x)dx$
- **Variance:**  $\text{Var}(X) = \int (x - \mu)^2 f(x)dx$
- **Relation:**  $f(x) = F'(x)$  where derivative exists
- **Note:**  $P(X = a) = 0$  for continuous RV

### 3.2 Uniform Distribution

- **Definition:** Equally likely over interval  $[a, b]$
- **PDF:**  $f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$
- **CDF:**  $F(x) = \frac{x-a}{b-a}$  for  $a \leq x \leq b$
- **Mean:**  $E[X] = \frac{a+b}{2}$
- **Variance:**  $\text{Var}(X) = \frac{(b-a)^2}{12}$
- **When to Use:** "Equally likely", "random point"
- **Example:** Random number between 0 and 1
- **Note:** Probability proportional to interval length

### 3.3 Normal Distribution (a.k.a. Gaussian, $N(\mu, \sigma^2)$ )

#### \* High Priority!

- **SYNONYMS:** Gaussian = Normal =  $N(\mu, \sigma^2)$  = "bell curve"
- **Definition:** Bell curve, most important continuous distribution
- **PDF:**  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- **Notation:**  $X \sim N(\mu, \sigma^2)$
- **Standardization:**  $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$
- **Properties:**

- Linear combination:  $aX + b \sim N(a\mu + b, a^2\sigma^2)$
- Sum of normals: normal
- 68-95-99.7 rule for  $\pm 1, 2, 3$  std dev

- **MGF:**  $M(t) = e^{\mu t + \sigma^2 t^2 / 2}$
- **Example:** Heights, measurement errors, CLT limit
- **Note:** Use  $\Phi(z)$  table for standard normal CDF

### 3.4 Exponential Distribution (a.k.a. $\text{Exp}(\lambda)$ , Memoryless)

- **SYNONYMS:**  $\text{Exp}(\lambda)$ , "waiting time", "memoryless", "inter-arrival time"
- **! "Mean  $\theta = 3$ " means  $\lambda = 1/3$  NOT  $\lambda = 3$ !**
- **Definition:** Waiting time until event
- **PDF:**  $f(x) = \lambda e^{-\lambda x}, \quad x > 0$
- **CDF:**  $F(x) = 1 - e^{-\lambda x}$
- **Mean:**  $E[X] = 1/\lambda$
- **Variance:**  $\text{Var}(X) = 1/\lambda^2$
- **Memoryless:**  $P(X > s + t | X > s) = P(X > t)$
- **MGF:**  $M(t) = \frac{\lambda}{\lambda - t}$  for  $t < \lambda$
- **When to Use:** Time between Poisson events
- **Example:** Service times, component lifetime
- **Note:** Min of exponentials is exponential

### 3.5 Gamma Distribution (a.k.a. $\text{Gamma}(r, \lambda)$ , Erlang)

- **SYNONYMS:** "sum of exponentials", "time until  $r$ -th event", Erlang (integer  $r$ )
- **Definition:** Sum of exponentials, generalization
- **PDF:**  $f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, \quad x > 0$

- **Mean:**  $E[X] = r/\lambda$
- **Variance:**  $\text{Var}(X) = r/\lambda^2$
- **Special Cases:**
  - $r = 1$ : Exponential( $\lambda$ )
  - $r = n/2, \lambda = 1/2$ : Chi-square with  $n$  df

- **When to Use:** Time until  $r$ -th event

- **Note:**  $\Gamma(n) = (n-1)!$  for integer  $n$

### 3.6 Beta Distribution (a.k.a. $\text{Beta}(\alpha, \beta)$ , Conjugate Prior)

- **SYNONYMS:** "prior for probability", "proportion model", "conjugate to Binomial"
- **Definition:** Models probabilities/proportions
- **PDF:**  $f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$
- **Support:**  $0 < x < 1$
- **Mean:**  $E[X] = \frac{\alpha}{\alpha + \beta}$
- **Variance:**  $\text{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$
- **Special Cases:**
  - $\alpha = \beta = 1$ : Uniform(0,1)
  - Conjugate prior for Binomial

- **[\$] Used in Bayesian statistics**

## 4. MULTIVARIATE DISTRIBUTIONS

#### \* Post-M2

## 4.1 Joint Distributions

- **Joint PMF (Discrete):**  $p(x, y) = P(X = x, Y = y)$
- **Joint PDF (Continuous):**  $f(x, y) \geq 0$
- **Normalization:**  $\int \int f(x, y) dx dy = 1$
- **Joint CDF:**  $F(x, y) = P(X \leq x, Y \leq y)$
- **When to Use:** Two or more random variables together
- **Solution Steps:**
  1. Verify normalization (integral = 1)
  2. Find constant  $c$  if needed
  3. Calculate probabilities over regions
- **Example:**  $f(x, y) = c(x^2 + xy)$  on  $[0, 1]^2$ , find  $c = 12/7$

• **! Check bounds carefully for integration!**

## 4.2 Marginal and Conditional Distributions

- **Marginal PDF:**  $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$
- **Marginal PMF:**  $p_X(x) = \sum_y p(x, y)$
- **Conditional PDF:**  $f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$
- **Properties:** Conditional is a valid PDF/PMF
- **Solution Steps:**
  1. Find marginal by integrating/summing out other variable
  2. For conditional, divide joint by marginal
  3. Verify it integrates to 1
- **Example:** Uniform on triangle, find conditional
- **Note:** Bounds change for conditional distributions

## 4.3 Independence of Random Variables

- **Definition:**  $X, Y$  independent iff  $f(x, y) = f_X(x) \cdot f_Y(y)$
- **Test for Independence:**
  1. Find joint distribution
  2. Find both marginals
  3. Check if product equals joint for ALL  $(x, y)$
- **Consequences of Independence:**
  - $E[XY] = E[X]E[Y]$
  - $\text{Cov}(X, Y) = 0$  (but not vice versa!)
  - $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
- **! Zero covariance  $\neq$  independence (except for normal)**
- **Note:** For normal: independent  $\Leftrightarrow \rho = 0$

## 4.4 Covariance and Correlation

- **Covariance:**  $\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - E[X]E[Y]$
- **Correlation:**  $\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \quad -1 \leq \rho \leq 1$
- **Properties:**
  - $\text{Cov}(X, X) = \text{Var}(X)$
  - $\text{Cov}(aX + b, cY + d) = ac \cdot \text{Cov}(X, Y)$
  - $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$
- **Solution Steps:**
  1. Find  $E[X], E[Y], E[XY]$
  2. Apply covariance formula
  3. For correlation, also find  $\sigma_X, \sigma_Y$

- **Example:** HW4: Joint PDF, find  $\rho$
- **[\$] Portfolio variance uses covariance matrix**

## 4.5 Bivariate Normal (a.k.a. Gaussian Vector, MVN, Jointly Normal)

\* **Critical!**

- **SYNONYMS:** Gaussian vector = MVN = Multivariate Normal = “Jointly Normal”
- **! “Independent components” =  $\rho = 0$  = independence (for MVN ONLY!)**
- **Definition:**  $(X, Y)$  jointly normal with parameters  $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho$
- **Properties:**
  - Linear combinations are normal
  - Marginals are normal:  $X \sim N(\mu_X, \sigma_X^2)$
  - Independence  $\Leftrightarrow \rho = 0$  (unique to normal!)
  - Conditional is normal:  $X|Y = y \sim N(\mu_{X|Y}, \sigma_{X|Y}^2)$

• **Conditional Mean:**  $\mu_{X|Y} = \mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y)$

• **Conditional Variance:**  $\sigma_{X|Y}^2 = \sigma_X^2 (1 - \rho^2)$

- **Example:** HW5 Problem 1: Find  $P(X + Y > 0)$

• **Note:**  $aX + bY$  is normal with specific mean/variance

## 4.6 Transformations (a.k.a. Jacobian Method, CDF Method)

### \* Complex!

- **SYNONYMS:** “change of variables”, “find distribution of  $Y = g(X)$ ”, “Jacobian”

- **Single Variable:**  $Y = g(X)$

- CDF Method: Find  $F_Y(y) = P(g(X) \leq y)$
- PDF Method:  $f_Y(y) = f_X(g^{-1}(y))|dg^{-1}/dy|$

- **Jacobian Method:**  $(U, V) = g(X, Y)$

$$f_{UV}(u, v) = f_{XY}(x(u, v), y(u, v)) \cdot |J| \quad \text{where}$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

- **Solution Steps:**

1. Define transformation
2. Find inverse transformation
3. Compute Jacobian determinant
4. Apply formula with absolute value
5. Check new bounds

- **! Don't forget absolute value of Jacobian!**

- **Example:** Polar coordinates:  $X = R \cos \Theta$ ,  $Y = R \sin \Theta$

## 4.7 Order Statistics (a.k.a. Max/Min of i.i.d., $X_{(k)}$ )

- **SYNONYMS:** “maximum”, “minimum”, “ $k$ -th smallest”, “range”

- **Definition:**  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$

- **Maximum:**  $F_{X_{(n)}}(x) = [F(x)]^n$

- **Minimum:**  $F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n$

- **PDF of  $k$ -th order statistic:**  $f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1-F(x)]^{n-k} f(x)$

- **Example:** Max of 3 uniform(0,1) variables

- **Note:** Range =  $X_{(n)} - X_{(1)}$

## 5. MOMENT GENERATING FUNCTIONS

### \* Post-M2

### 5.1 Definition and Properties (Prof. uses $\psi(t)$ for MGF)

- **SYNONYMS:** MGF =  $M_X(t) = \psi(t)$  (professor's notation)

- **Definition:**

$$M_X(t) = \psi(t) = E[e^{tX}] = \begin{cases} \sum e^{tx} p(x) & \text{discrete} \\ \int e^{tx} f(x) dx & \text{continuous} \end{cases}$$

- **Moments:**  $E[X^k] = M^{(k)}(0)$  (k-th derivative at 0)

- **Properties:**

- Uniqueness: Same MGF  $\Rightarrow$  same distribution
- Linear:  $M_{aX+b}(t) = e^{bt} M_X(at)$
- Sum of independent:  $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

- **Common MGFs:**

- Binomial:  $(1 - p + pe^t)^n$
- Poisson:  $e^{\lambda(e^t - 1)}$
- Normal:  $e^{\mu t + \sigma^2 t^2 / 2}$
- Exponential:  $\lambda / (\lambda - t)$ ,  $t < \lambda$

- **Example:** Find distribution of sum using MGFs

- **Note:** Not all distributions have MGF (e.g., Cauchy)

### 5.2 Using MGFs for Sums

- **Method:** For independent  $X_1, \dots, X_n$ :

1. Find individual MGFs:  $M_{X_i}(t)$
2. Multiply:  $M_S(t) = \prod M_{X_i}(t)$
3. Match with known MGF to identify distribution

- **Example Applications:**

- Sum of normals is normal
- Sum of Poissons is Poisson
- Sum of gammas (same  $\lambda$ ) is gamma

- **Example:**  $X_i \sim \text{Exp}(\lambda)$  independent, then  $\sum X_i \sim \text{Gamma}(n, \lambda)$

- **Note:** Powerful for proving CLT

## 6. LIMIT THEOREMS

### \* Critical for Final!

### 6.1 Central Limit Theorem (a.k.a. CLT, Normal Approximation)

#### \* Most Important!

- **SYNONYMS:** CLT, “approximate”, “large  $n$ ”, “as  $n \rightarrow \infty$ ”, “normal approximation”

- **TRIGGER WORDS:** “i.i.d.”, “sample mean”, “total/sum of  $n$  games”, “average of  $n$ ”

- **Statement:** If  $X_1, \dots, X_n$  are iid with mean  $\mu$ , variance  $\sigma^2$ :

$$\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} N(0, 1) \text{ as } n \rightarrow \infty$$

- **Equivalent:**  $\bar{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right)$  for large  $n$

- **When to Use:**

- Large sample size (typically  $n \geq 30$ )
- Sum or average of many random variables
- Approximating discrete by continuous

- **Solution Steps:**

1. Verify conditions (iid, finite variance)
2. Identify  $\mu = E[X_i]$ ,  $\sigma^2 = \text{Var}(X_i)$
3. Standardize:  $Z = (\bar{X} - \mu) / (\sigma / \sqrt{n})$
4. Use normal table



- **Example:** 400 games, win \$3 with  $p = 0.25$ , find  $P(\text{total} > 240)$

- **! Apply continuity correction for discrete!**

## 6.2 Normal Approximations

- **Binomial Approximation:** If  $X \sim \text{Binomial}(n, p)$  with  $np(1-p) > 10$ :  

$$X \approx N(np, np(1-p))$$

- **Poisson Approximation:** If  $X \sim \text{Poisson}(\lambda)$  with  $\lambda > 30$ :  

$$X \approx N(\lambda, \lambda)$$

- **Continuity Correction:** For discrete  $X$ :

- $P(X = k) \approx P(k - 0.5 < Y < k + 0.5)$
- $P(X \leq k) \approx P(Y < k + 0.5)$
- $P(X < k) \approx P(Y < k - 0.5)$

- **Example:** Binomial(100, 0.3), find  $P(X > 35)$

- *Note: Correction improves accuracy significantly*

## 6.3 Law of Large Numbers (LLN)

- **Weak LLN:**  $\bar{X}_n \xrightarrow{P} \mu$  (convergence in probability)
- **Strong LLN:**  $\bar{X}_n \xrightarrow{a.s.} \mu$  (almost sure convergence)
- **Interpretation:** Sample mean converges to true mean
- **Conditions:**
  - Weak: Finite mean, pairwise uncorrelated
  - Strong: Finite mean (iid case)
- **Example:** Casino games, long-run frequency
- *Note: Foundation for frequentist probability*

## 6.4 Confidence Intervals

- **Definition:** Interval estimate with specified confidence level

- **For Mean (known  $\sigma$ ):** 
$$\bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

- **Common Values:**

- 90% CI:  $z_{0.05} = 1.645$
- 95% CI:  $z_{0.025} = 1.96$
- 99% CI:  $z_{0.005} = 2.576$

- **Interpretation:** 95% of such intervals contain true parameter

- **Width:**  $2 \cdot z_{\alpha/2} \cdot \sigma / \sqrt{n}$

- **! NOT "95% chance parameter is in interval"**

- *Note: Larger  $n \Rightarrow$  narrower interval*

## 7. SPECIAL TOPICS & APPLICATIONS

### \* Post-M2

### 7.1 Conditional Expectation (a.k.a. $E[X|Y]$ , Total Expectation)

#### \* Conceptual!

- **SYNONYMS:** “ $E[X|Y]$ ”, “average given”, “expected value given”

- **TRIGGER WORDS:** “ $E[X|Y = y]$ ”, “break down by cases”, “tower property”

- **Definition:**

- Discrete:  $E[X|Y = y] = \sum x \cdot P(X = x|Y = y)$
- Continuous:  $E[X|Y = y] = \int x \cdot f_{X|Y}(x|y)dx$

- **Law of Total Expectation:** 
$$E[X] = E[E[X|Y]]$$

- **Properties:**

- Linearity:  $E[aX + bZ|Y] = aE[X|Y] + bE[Z|Y]$
- Taking out known:  $E[h(Y)X|Y] = h(Y)E[X|Y]$
- Independence:  $E[X|Y] = E[X]$  if independent

- **Law of Total Variance:** 
$$\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E[X|Y])$$

- **Example:** Breaking sticks problem

- *Note:  $E[X|Y]$  is a function of  $Y$ , not a number!*

### 7.2 Bayesian Statistics (a.k.a. Prior/Posterior, Conjugate Priors)

#### \* Professor's Favorite!

- **SYNONYMS:** “prior”, “posterior”, “update belief”, “given evidence”, “conjugate”

- **TRIGGER WORDS:** “defective rate”, “unknown parameter”, “given data”, “Monty Hall”

- **Bayesian Framework:** 
$$\pi(\theta|x) = \frac{L(x|\theta)\pi(\theta)}{\int L(x|\theta)\pi(\theta)d\theta}$$
  
 where Prior  $\times$  Likelihood  $\rightarrow$  Posterior

- **Conjugate Priors:**

- Beta-Binomial:  $\text{Beta}(\alpha, \beta) \rightarrow \text{Beta}(\alpha + x, \beta + n - x)$
- Gamma-Poisson:  $\text{Gamma}(\alpha, \beta) \rightarrow \text{Gamma}(\alpha + \sum x_i, \beta + n)$
- Normal-Normal: With known variance

- **Solution Steps:**

1. Specify prior  $\pi(\theta)$
2. Write likelihood  $L(x|\theta)$
3. Compute posterior (use conjugacy if possible)
4. Normalize if needed

- **Example:** HW6 Monty Hall Bayesian analysis

- **[\$] Used in risk assessment, portfolio optimization**

### 7.3 Lognormal Distribution (a.k.a. $\ln X \sim N(\mu, \sigma^2)$ )

#### \* Finance Applications!

- **SYNONYMS:** Lognormal, “log  $X$  is normal”, “ $e^X$  where  $X \sim N$ ”, “stock price model”
- **TRIGGER WORDS:** “stock price”, “ $S = S_0 e^Z$ ”, “log returns”, “always positive”
- **Definition:**  $Y = e^X$  where  $X \sim N(\mu, \sigma^2)$
- **Properties:**
  - Always positive (good for prices)
  - Right-skewed
  - Mean:  $E[Y] = e^{\mu + \sigma^2/2}$
  - Variance:  $\text{Var}(Y) = e^{2\mu + 2\sigma^2}(e^{\sigma^2} - 1)$
  - Median:  $e^\mu$
- **Stock Price Model:**  $S_t = S_0 \exp(X_t)$  where  $X_t \sim N(\mu t, \sigma^2 t)$
- **Example:** HW5 Problem 2, Practice Final stock problems
- **[\$] Black-Scholes model foundation**
- *Note: Log returns are normal, prices are lognormal*

### 7.4 Additional Important Concepts

- **Indicator Random Variables:**
  - $I_A = 1$  if  $A$  occurs, 0 otherwise
  - $E[I_A] = P(A)$
  - Useful for counting:  $\sum I_{A_i}$
- **Jensen’s Inequality:** For convex  $g$ :  $E[g(X)] \geq g(E[X])$
- **Chebyshev’s Inequality:**  $P(|X - \mu| \geq k\sigma) \leq 1/k^2$
- **Probability Integral Transform:**  $F(X) \sim \text{Uniform}(0, 1)$

## 8. PRACTICE PROBLEM COMPENDIUM

### 8.1 Bayesian Problems

\* **High Frequency!** Problem [HW6-1]: **Monty Hall (Sober vs Dizzy)**  
*Contestant picks door A. Monty opens door B showing goat.*

**Solution:**

1. **Problem Type:** Bayesian update with different likelihoods
2. **Required Concepts:** Bayes’ theorem, conditional probability
3. **Sober Monty:**
  - Prior:  $P(H_A) = P(H_B) = P(H_C) = 1/3$
  - Likelihood:  $P(\text{open B}|H_A) = 1/2$ ,  $P(\text{open B}|H_B) = 0$ ,  $P(\text{open B}|H_C) = 1$
  - Posterior:  $P(H_A|\text{data}) = 1/3$ ,  $P(H_C|\text{data}) = 2/3$
  - **Strategy: Switch!** (doubles probability)
4. **Dizzy Monty:**
  - Likelihood:  $P(\text{open B}|H_A) = 1/2$ ,  $P(\text{open B}|H_B) = 1/2$ ,  $P(\text{open B}|H_C) = 1/2$
  - Posterior: All equal at  $1/3$
  - **Strategy: Doesn’t matter!**

5. **Key Insight:** Knowledge affects likelihood function

### 8.2 CLT Applications

\* **Guaranteed on Final!** Problem [Practice Final-1]: **Coin Game with 400 Plays**

*Win \$3 if HH, lose \$1 if TT, else \$0. Play 400 times.*

**Solution:**

1. **Problem Type:** CLT with discrete outcomes
2. **Step 1:** Find distribution of single game

- $P(X = 3) = 1/4$  (HH)
- $P(X = -1) = 1/4$  (TT)
- $P(X = 0) = 1/2$  (HT or TH)

3. **Step 2:** Calculate  $\mu$  and  $\sigma^2$

- $E[X] = 3(1/4) + (-1)(1/4) + 0(1/2) = 0.5$
- $E[X^2] = 9(1/4) + 1(1/4) + 0 = 2.5$
- $\text{Var}(X) = 2.5 - 0.25 = 2.25$ , so  $\sigma = 1.5$

4. **Step 3:** Apply CLT for  $n = 400$

- Total:  $S_{400} \approx N(400 \cdot 0.5, 400 \cdot 2.25) = N(200, 900)$
- $P(S_{400} \geq 240) = P(Z \geq \frac{240-200}{30}) = P(Z \geq 1.33) \approx 0.092$

5. **Key Insight:** Use continuity correction:  $P(S \geq 240) \approx P(S > 239.5)$

### 8.3 Bivariate Normal

\* **Complex but Common!** Problem [HW5-1]: **Joint Normal with Correlation**

$X \sim N(1, 2)$ ,  $Y \sim N(-2, 3)$ ,  $\rho = -2/3$ . Find  $P(X + Y > 0)$

**Solution:**

1. **Problem Type:** Linear combination of bivariate normal
2. **Key Property:**  $X + Y$  is normal
3. **Parameters of  $Z = X + Y$ :**

- $\mu_Z = \mu_X + \mu_Y = 1 + (-2) = -1$
- $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y = 2 + 3 + 2(-2/3)\sqrt{6} = 5 - \frac{4\sqrt{6}}{3}$

4. **Standardize and compute:**

- $P(Z > 0) = P\left(\frac{Z+1}{\sigma_Z} > \frac{1}{\sigma_Z}\right) = 1 - \Phi(0.759) \approx 0.224$

5. **Key Insight:** Always check if linear combination, use properties of bivariate normal



## 8.4 Joint Distributions

### Problem [HW4-1]: Joint PDF Analysis

$f(x, y) = c(x^2 + xy)$  on  $[0, 1] \times [0, 1]$

**Solution:**

1. **Find constant  $c$ :**

- $\int_0^1 \int_0^1 (x^2 + xy) dx dy = \int_0^1 [\frac{x^3}{3} + \frac{x^2 y}{2}]_0^1 dy = \int_0^1 (\frac{1}{3} + \frac{y}{2}) dy = \frac{7}{12}$
- Therefore  $c = \frac{12}{7}$

2. **Marginal of  $X$ :**

- $f_X(x) = \int_0^1 \frac{12}{7}(x^2 + xy) dy = \frac{12}{7}x^2 + \frac{6x}{7}$

3. **Check independence:**

- Need  $f(x, y) = f_X(x) \cdot f_Y(y)$  for all  $(x, y)$
- Since  $f(x, y)$  has  $xy$  term, NOT independent

4. **Key Insight:** Cross-product terms indicate dependence

## 8.5 Lognormal Distribution

### [\$ Finance Focus! Problem [Practice Final-4]: Stock Price Model

$S = S_0 e^Z$  where  $Z \sim N((r - \sigma^2/2), \sigma^2)$ ,  $S_0 = 100$ ,  $r = 0.05$ ,  $\sigma = 0.2$

**Solution:**

1. **Problem Type:** Lognormal application

2. **Part (a):** Find  $E[e^{-r}S] = E[S_0 e^{Z-r}]$

- $Z - r \sim N(-\sigma^2/2, \sigma^2)$
- $E[e^{Z-r}] = \exp(-\sigma^2/2 + \sigma^2/2) = 1$
- Therefore  $E[e^{-r}S] = S_0 = 100$

3. **Part (b):** Find  $P(S > 100)$

- $P(S > 100) = P(e^Z > 1) = P(Z > 0)$
- $Z \sim N(-0.02, 0.04)$
- $P(Z > 0) = P(\frac{Z+0.02}{0.2} > 0.1) = 1 - \Phi(0.1) \approx 0.46$

4. **Key Insight:** Stock prices lognormal  $\Rightarrow$  log returns normal

## 8.6 Exponential Memoryless

### Problem [Practice Final-3]: Average of Exponentials

$X_1, \dots, X_{100}$  iid  $\text{Exp}(1/3)$ . Find  $P(\bar{X}/(\bar{X} + 3) < 0.5)$

**Solution:**

1. **Problem Type:** CLT for exponentials

2. **Setup:**  $E[X_i] = 3$ ,  $\text{Var}(X_i) = 9$

3. **Apply CLT:**  $\bar{X} \approx N(3, 9/100) = N(3, 0.09)$

4. **Transform inequality:**

- $\frac{\bar{X}}{\bar{X}+3} < 0.5 \Rightarrow \bar{X} < 0.5(\bar{X} + 3) \Rightarrow \bar{X} < 3$

5. **Calculate:**  $P(\bar{X} < 3) = 0.5$  (by symmetry of normal)

6. **Key Insight:** Transform inequality first, then apply CLT

## 8.7 Order Statistics

### Problem: Max and Min of Uniform(0,1)

$X_1, \dots, X_n$  iid  $\text{Uniform}(0,1)$ . Find distribution of max and min.

**Solution:**

1. **Maximum  $X_{(n)}$ :**

- $F_{\max}(x) = P(\text{all} \leq x) = x^n$
- $f_{\max}(x) = nx^{n-1}$  for  $0 < x < 1$
- $E[X_{(n)}] = \frac{n}{n+1}$

2. **Minimum  $X_{(1)}$ :**

- $F_{\min}(x) = 1 - P(\text{all} > x) = 1 - (1-x)^n$
- $f_{\min}(x) = n(1-x)^{n-1}$  for  $0 < x < 1$
- $E[X_{(1)}] = \frac{1}{n+1}$

3. **Key Insight:** Use complement for min, direct for max

## 8.8 Conjugate Priors

### \* Bayesian Favorite! Problem [Practice Final-5]: Beta-Binomial Update

Prior:  $\theta \in \{1/2, 3/4\}$  equally likely. Data: 0 defects in 10 items.

**Solution:**

1. **Problem Type:** Discrete prior Bayesian update

2. **Likelihoods:**

- $P(0 \text{ defects} | \theta = 1/2) = (1/2)^{10} = 1/1024$
- $P(0 \text{ defects} | \theta = 3/4) = (1/4)^{10} = 1/1048576$

3. **Posterior:**

- $P(\theta = 1/2 | \text{data}) \propto (1/2) \cdot 1/1024 = 1/2048$
- $P(\theta = 3/4 | \text{data}) \propto (1/2) \cdot 1/1048576 \approx 0$
- After normalization:  $P(\theta = 1/2 | \text{data}) \approx 0.999$

4. **Key Insight:** Extreme data strongly favors lower defect rate

## 8.9 Conditional Expectation

### Problem: Breaking Sticks

Break at  $X \sim U(0, \ell)$ , then break smaller piece at  $Y | X \sim U(0, X)$

**Solution:**

1. **Joint density:**  $f(x, y) = \frac{1}{\ell} \cdot \frac{1}{x} = \frac{1}{\ell x}$  for  $0 < y < x < \ell$

2. **Marginal of  $Y$ :**  $f_Y(y) = \int_y^\ell \frac{1}{\ell x} dx = \frac{1}{\ell} \ln(\ell/y)$

3. **Conditional expectation:**  $E[Y | X] = X/2$

4. **Total expectation:**  $E[Y] = E[E[Y | X]] = E[X/2] = \ell/4$

5. **Total variance:** Use  $\text{Var}(Y) = E[\text{Var}(Y | X)] + \text{Var}(E[Y | X])$

6. **Key Insight:** Hierarchical structure leads to law of total expectation

## 8.10 Hypothesis Testing & Confidence Intervals

**Problem:** Test Average with CLT

Sample mean  $\bar{X} = 52$  from  $n = 100$ , known  $\sigma = 10$ .

Test  $H_0 : \mu = 50$

**Solution:**

1. **Test statistic:**  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{52 - 50}{10/10} = 2$
2. **P-value:**  $P(|Z| > 2) = 2(1 - \Phi(2)) = 0.0455$
3. **95% CI:**  $\bar{X} \pm 1.96 \cdot \sigma/\sqrt{n} = 52 \pm 1.96 = [50.04, 53.96]$
4. **Decision:** Reject  $H_0$  at 5% level (barely)
5. **Key Insight:** CI excludes 50, consistent with rejection

## 9. MULTI-STEP PROBLEM TEMPLATES

\* Critical!

### Template A: Gaussian Vector Problems

**When you see:** “Gaussian vector”, “independent components”, “MVN”

**Steps:**

1. **Recognize:** “Gaussian” = Normal!
2. **“Independent components”** means  $\rho = 0$  and for MVN: **independent!**
3. **Find Cov:** Set  $\text{Cov}(Y_1, Y_2) = 0$  to find parameters
4. **Joint density:** Product of marginals (since independent)

**Key Formula:** For  $Y_1 = aX_1 + X_2, Y_2 = X_1 + bX_2$  (iid  $N(0, 1)$ ):

$$\text{Cov}(Y_1, Y_2) = a\text{Var}(X_1) + b\text{Var}(X_2) = a + b$$

Independence requires:  $b = -a$

### Template B: CLT Game/Coin Problems

**When you see:** “400 games”, “total winnings”, “approximate”

**Steps:**

1. **Define:**  $X_i$  = single trial outcome
2. **PMF:** List outcomes and probabilities
3. **Compute:**  $E[X_i] = \sum x \cdot P(X = x)$
4. **Compute:**  $\text{Var}(X_i) = E[X^2] - (E[X])^2$
5. **CLT:**  $S_n = \sum X_i \approx N(n\mu, n\sigma^2)$
6. **Standardize:**  $Z = \frac{S_n - n\mu}{\sigma\sqrt{n}}$

**Example:** Win \$3 if HH ( $p = 1/4$ ), lose \$1 if TT ( $p = 1/4$ ), else \$0 ( $p = 1/2$ )  
 $E[X] = 3(1/4) - 1(1/4) + 0(1/2) = 1/2$   
 $E[X^2] = 9(1/4) + 1(1/4) = 10/4, \text{Var}(X) = 10/4 - 1/4 = 9/4$

### Template C: Exponential + CLT

**When you see:** “i.i.d. Exp”, “mean  $\theta$ ”, “average”

**! If “mean  $\theta = 3$ ” then  $\lambda = 1/3$  NOT 3!**

**Steps:**

1. **Parameters:**  $E[X_i] = 1/\lambda, \text{Var}(X_i) = 1/\lambda^2$
2. **For  $\bar{X}$ :**  $E[\bar{X}] = 1/\lambda, \text{Var}(\bar{X}) = 1/(n\lambda^2)$
3. **CLT:**  $\bar{X} \approx N(1/\lambda, 1/(n\lambda^2))$
4. **Transform inequality first:** e.g.,  $\bar{X}/(\bar{X}+3) < 0.5 \Rightarrow \bar{X} < 3$

### Template D: Lognormal Stock Price

**When you see:** “ $S = S_0 e^Z$ ”, “ $Z \sim N(\mu, \sigma^2)$ ”

**Key Formulas:**

- $E[e^Z] = e^{\mu + \sigma^2/2}$  when  $Z \sim N(\mu, \sigma^2)$
- $E[S] = S_0 e^{\mu + \sigma^2/2}$
- $P(S > K) = P(Z > \ln(K/S_0)) = 1 - \Phi\left(\frac{\ln(K/S_0) - \mu}{\sigma}\right)$

**For  $E[e^{-r}S]$  with  $Z \sim N(r - \sigma^2/2, \sigma^2)$ :**  
 $E[e^{-r}S] = e^{-r}S_0 E[e^Z] = e^{-r}S_0 e^{(r - \sigma^2/2) + \sigma^2/2} = S_0$

### Template E: Bayesian Discrete Prior

**When you see:** “prior”, “posterior”, “defective rate”

**Steps:**

1. **List hypotheses:**  $\theta_1, \theta_2, \dots$
2. **Priors:**  $P(\theta_i)$
3. **Likelihoods:**  $P(\text{data}|\theta_i)$
4. **Bayes:**  $P(\theta_i|\text{data}) = \frac{P(\text{data}|\theta_i)P(\theta_i)}{\sum_j P(\text{data}|\theta_j)P(\theta_j)}$
5. **Normalize:** Make sure posteriors sum to 1

### Template F: Bivariate Normal Conditional

**When you see:** “bivariate normal”, “ $Y|X = x$ ”, “conditional distribution”

**Formula:**  $Y|X = x \sim N\left(\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X), (1 - \rho^2)\sigma_Y^2\right)$

**Special Case:** If  $\rho = 0$ , then  $Y|X = x \sim N(\mu_Y, \sigma_Y^2)$  (unchanged!)

### Template G: Linear Combination Independence

**When you see:** “ $Y_1 = aX_1 + X_2$ ”, “ $Y_2 = X_1 + bX_2$ ”, “independent components”

**Steps:**

1. **Setup:**  $X_1, X_2$  i.i.d.  $N(0, 1)$
2. **Key insight:** For linear combinations of Gaussians, independence  $\Leftrightarrow$  zero covariance
3. **Calculate:**  $\text{Cov}(Y_1, Y_2) = a \cdot 1 + b \cdot 1 = a + b$
4. **Solve:**  $a + b = 0 \Rightarrow b = -a$
5. **Joint density:** Product of marginals (since independent)

### Template H: Predictive Distributions (Bayesian)

**When you see:** “predict next outcome”, “predictive probability”, “posterior predictive”

**Steps:**

1. **Prior predictive:**  $P(X_{n+1} = x) = \sum_{\theta} P(X = x|\theta)P(\theta)$
2. **Posterior predictive:**  $P(X_{n+1} = x|\text{data}) = \sum_{\theta} P(X = x|\theta)P(\theta|\text{data})$
3. **Use:** Posterior from Bayesian update in Step 2

**Example:** Dice problem: Prior  $P(\theta)$  for die type, observe data, predict next roll.

## Template I: Product of Lognormals

**When you see:** “ $XY$  where  $X, Y$  are lognormal”, “product of independent”

**Key Insight:**  $\ln(XY) = \ln X + \ln Y$

**Steps:**

1. If  $\ln X \sim N(\mu_1, \sigma_1^2)$  and  $\ln Y \sim N(\mu_2, \sigma_2^2)$  independent
2. Then  $\ln(XY) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
3. So  $XY$  is lognormal with parameters  $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
4.  $E[XY] = E[X]E[Y]$  (independence)  $= e^{\mu_1 + \sigma_1^2/2} \cdot e^{\mu_2 + \sigma_2^2/2}$

## Template J: Monty Hall Variants

**When you see:** “Monty Hall”, “contestant picks”, “host opens”

**Steps:**

1. **Define hypotheses:**  $H_A, H_B, H_C$  = car behind door A, B, C
2. **Priors:** Usually uniform  $P(H_i) = 1/3$
3. **Key:** Likelihoods depend on host behavior!
  - **Sober:** Knows car location, opens non-car door
  - **Dizzy:** Opens random door (50-50)
4. **Sober Monty:** Switch doubles probability (2/3 vs 1/3)
5. **Dizzy Monty:** No advantage to switching

## Template K: Finding n for CLT Probability

**When you see:** “smallest n such that”, “how many samples needed”

**Steps:**

1. **Setup:** Want  $P(\bar{X} > c) > p$  or  $P(\bar{X} < c) > p$
2. **CLT:**  $\bar{X} \approx N(\mu, \sigma^2/n)$
3. **Standardize:**  $P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > z\right) = \text{target}$
4. **Solve for n:** Using z-table, find  $z^*$  then solve  $\frac{c - \mu}{\sigma/\sqrt{n}} = z^*$
5. **Result:**  $n \geq \left(\frac{z^* \sigma}{c - \mu}\right)^2$

## Template L: Max/Min of i.i.d. Variables

**When you see:** “maximum of n”, “minimum of n”, “largest/smallest”

**Formulas:**

- $P(\max < a) = P(\text{all} < a) = [F(a)]^n$  (if i.i.d.)
- $P(\max > a) = 1 - [F(a)]^n$
- $P(\min < a) = 1 - [1 - F(a)]^n$
- $P(\min > a) = [1 - F(a)]^n$

**For Uniform(0,1):**  $E[X_{(n)}] = \frac{n}{n+1}$ ,  $E[X_{(1)}] = \frac{1}{n+1}$

## Template M: Conditioning on Event

**When you see:** “given that  $X > a$ ”, “conditional on event”

**Steps:**

1. **Conditional CDF:**  $P(X \leq x | X > a) = \frac{P(a < X \leq x)}{P(X > a)}$  for  $x > a$
2. **For Exponential:** Memoryless!  $P(X > s + t | X > s) = P(X > t)$
3. **General:** Use  $f_{X|A}(x) = f_X(x)/P(A)$  for  $x \in A$

## Template N: Bivariate Normal from Conditions

**When you see:** “ $E[Y|X = x] = \dots$ ”, “ $\text{Var}(Y|X) = \dots$ ”, “find parameters”

**Key Formulas:**

- $E[Y|X = x] = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$
- $\text{Var}(Y|X) = \sigma_Y^2(1 - \rho^2)$  (constant!)

**Method:** Match coefficients to extract  $\mu_Y, \rho \frac{\sigma_Y}{\sigma_X}$ , and  $\sigma_Y^2(1 - \rho^2)$

## Template O: Probability Involving Sample Average

**When you see:** “ $P(\bar{X}/(\bar{X} + c) < p)$ ”, “ratio with sample mean”

**Steps:**

1. **Transform:** Simplify inequality algebraically first!
2. **Example:**  $\frac{\bar{X}}{\bar{X}+3} < 0.5 \Leftrightarrow \bar{X} < 3$
3. **Apply CLT:**  $\bar{X} \approx N(\mu, \sigma^2/n)$
4. **Calculate:** Standard normal probability

## Template P: Sum of Independent Poissons

**When you see:** “sum of Poisson”, “combined arrivals”

**Key Property:** If  $X \sim \text{Poisson}(\lambda_1)$ ,  $Y \sim \text{Poisson}(\lambda_2)$  independent:  $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$

**Method:** Use MGF:  $M_{X+Y}(t) = e^{\lambda_1(e^t - 1)} \cdot e^{\lambda_2(e^t - 1)} = e^{(\lambda_1 + \lambda_2)(e^t - 1)}$

# APPENDIX A: COMPLETE FORMULA SHEET

# APPENDIX B: DISTRIBUTION CHEAT SHEET

# APPENDIX D: LAST-MINUTE REVIEW CHECKLIST

## Probability Formulas

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A^c) = 1 - P(A)$
- $P(A|B) = P(A \cap B)/P(B)$
- $P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$
- $P(A) = \sum P(A|B_i)P(B_i)$  (Total Probability)
- $P(H|E) = P(E|H)P(H)/P(E)$  (Bayes)
- $C(n, k) = n!/(k!(n - k)!)$
- $P(n, k) = n!/(n - k)!$

## Expectation & Variance

- $E[X] = \sum xp(x)$  or  $\int xf(x)dx$
- $E[g(X)] = \sum g(x)p(x)$  or  $\int g(x)f(x)dx$
- $E[aX + b] = aE[X] + b$
- $E[X + Y] = E[X] + E[Y]$
- $\text{Var}(X) = E[X^2] - (E[X])^2$

- $\text{Var}(aX + b) = a^2\text{Var}(X)$
- $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$
- $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$
- $\rho = \text{Cov}(X, Y)/(\sigma_X\sigma_Y)$

## Conditional Expectation

- $E[X|Y = y] = \sum xP(X = x|Y = y)$
- $E[X] = E[E[X|Y]]$  (Total Expectation)
- $E[h(Y)X|Y] = h(Y)E[X|Y]$
- $\text{Var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E[X|Y])$

## MGF & Limit Theorems

- $M_X(t) = E[e^{tX}]$
- $E[X^k] = M^{(k)}(0)$
- $M_{X+Y}(t) = M_X(t)M_Y(t)$  (if independent)
- CLT:  $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1)$
- CI:  $\bar{X} \pm z_{\alpha/2}\sigma/\sqrt{n}$

## APPENDIX C: PROFESSOR'S NOTATION GUIDE

- Uses  $\psi(t)$  for MGF (not  $M(t)$ )
- Writes  $\text{Var}(X)$  not  $\sigma_X^2$
- Uses  $f(x)$  for both PMF and PDF
- $g_1(x|y)$  for conditional PDF of  $X|Y$
- $\pi(\theta)$  for prior,  $\pi(\theta|x)$  for posterior
- $L(x|\theta)$  for likelihood
- $H_i$  for hypotheses in Bayes problems
- $\bar{X}$  or  $\bar{X}_n$  for sample mean
- $X_{(k)}$  for  $k$ -th order statistic
- $I_A$  for indicator of event  $A$
- $\xrightarrow{d}$  for convergence in distribution
- $\xrightarrow{P}$  for convergence in probability
- $\Phi(z)$  for standard normal CDF
- $z_\alpha$  for quantile where  $P(Z > z_\alpha) = \alpha$
- Finance:  $S_t$  for stock price at time  $t$

## Time Management (90 minutes, 3 questions)

Distribution	PMF/PDF	Mean	Variance	MGF	Notes
Discrete Distributions					
Bernoulli( $p$ )	$p^x(1-p)^{1-x}$	$p$	$p(1-p)$	$1-p+pe^t$	Read all problems, identify types using decision tree
Binomial( $n, p$ )	$\binom{n}{k}p^k(1-p)^{n-k}$	$np$	$np(1-p)$	$(1-p+pe^t)^n$	$k = 0, \dots, n$
Poisson( $\lambda$ )	$e^{-\lambda}\lambda^k/k!$	$\lambda$	$\lambda$	$e^{\lambda(e^t-1)}$	$k = 0, 1, 2, \dots$
Geometric( $p$ )	$p(1-p)^k$	$(1-p)/p$	$(1-p)/p^2$	$p/(1-(1-p)e^t)$	Question 1 (aim for 30 min max)
Neg. Binomial( $r, p$ )	$\binom{k+r-1}{k}p^r(1-p)^k$	$r(1-p)/p$	$r(1-p)/p^2$	$(pe^t/(1-(1-p)e^t))^r$	Question 2 (aim for 30 min max)
Hypergeometric	Complex	$nK/N$	Complex	Complex	Question 3 (aim for 20 min)
Continuous Distributions					
Uniform( $a, b$ )	$1/(b-a)$	$(a+b)/2$	$(b-a)^2/12$	$(e^{bt}-e^{at})/(t(b-a))$	Review, check arithmetic
Normal( $\mu, \sigma^2$ )	$(2\pi\sigma^2)^{-1/2}e^{-(x-\mu)^2/(2\sigma^2)}$	$\mu$	$\sigma^2$	$e^{\mu t + \sigma^2 t^2/2}$	High-Yield Topics to Review (Post-M2 Focus)
Exponential( $\lambda$ )	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$	$\lambda/(1-t)$	
Gamma( $r, \lambda$ )	$\lambda^r x^{r-1} e^{-\lambda x} / \Gamma(r)$	$r/\lambda$	$r/\lambda^2$	$(\lambda/(\lambda-t))^r$	
Beta( $\alpha, \beta$ )	$x^{\alpha-1}(1-x)^{\beta-1}/B(\alpha, \beta)$	$\alpha/(\alpha+\beta)$	Complex	-	$x \in (0, 1)$

- \* Central Limit Theorem applications
- \* Bivariate Normal problems
- \* Bayesian updates (especially Monty Hall variants)
- \* Conditional Expectation and Total Expectation
- \* Lognormal/Finance applications
- Joint distributions (finding marginals, checking independence)
- Covariance and correlation calculations
- MGF for finding distributions of sums
- Normal approximations with continuity correction
- Confidence intervals using CLT

## What to Memorize vs Look Up

### MEMORIZE:

- “Gaussian” = Normal, “Gaussian vector” = MVN
- Normal standardization:  $Z = (X - \mu)/\sigma$
- CLT:  $(\bar{X} - \mu)/(\sigma/\sqrt{n}) \rightarrow N(0, 1)$
- Lognormal:  $E[e^X] = e^{\mu + \sigma^2/2}$  for  $X \sim N(\mu, \sigma^2)$
- BVN Conditional:  $\mu_{Y|X} = \mu_Y + \rho(\sigma_Y/\sigma_X)(x - \mu_X)$

- For MVN ONLY:  $\rho = 0 \Leftrightarrow$  independent

### LOOK UP:

- Distribution tables (PMF/PDF formulas)
- MGF formulas:  $\psi(t)$  values
- Jacobian details
- Normal table ( $\Phi$  values)

### PARAMETER TRAP CHECKLIST

- ☐ “Mean  $\theta = 3$ ” (Exp)  $\Rightarrow \lambda = 1/3$
- ☐ “Rate  $\lambda = 2$ ”  $\Rightarrow$  Mean =  $1/2$
- ☐ Check: Is it Geom(failures) or Geom(trials)?

- ☐ BVN: Is variance  $\sigma^2$  or std dev  $\sigma$ ?

### Common Professor Patterns

- Part (a): Basic setup/calculation
- Part (b): Extension requiring part (a)
- Part (c): Conceptual twist or limiting behavior
- Finance context in at least one problem
- One Bayesian problem guaranteed
- One CLT/approximation problem guaranteed

### Final Tips

- **! Always check if variables are independent before using simplified formulas**
- **! Apply continuity correction for discrete  $\rightarrow$  continuous**
- **! Verify bounds of integration match the region**
- Start with problems you recognize immediately
- Show all work - partial credit is generous
- If stuck, write down relevant formulas and what you know
- Check units/reasonableness of final answers