Mathematical Statistics Midterm Cheat Sheet

Xuange Liang
xl3493@cumc.columbia.edu

1. Probability Foundations

Sample Space and Events:

- Sample Space S: Set of all possible outcomes
- Event: Subset of S
- Simple Event: Single outcome; Compound Event: Multiple outcomes
- Discrete S: Finite/countably infinite; Continuous S: Uncountable

Probability Axioms:

- Axiom 1: $P(A) \geq 0$ for any event A
- Axiom 2: $P(S) = 1$
- Axiom 3: For disjoint $A_1, A_2, \ldots$: $P(\bigcup A_i) = \sum P(A_i)$

Basic Rules:

- Complement: $P(A^c) = 1 - P(A)$
- Addition: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Mutually Exclusive: $P(A \cup B) = P(A) + P(B)$

Sample-Point Method: $P(A) = \frac{\text{Number of outcomes in } A}{\text{Total outcomes in } S}$

Counting Techniques:

- Multiplication Principle: $n_1 \times n_2$ ways
- Permutations: $P_n = n!$, $P_{n,r} = \frac{n!}{(n-r)!}$
- Combinations: $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

Conditional Probability:

- $P(A|B) = \frac{P(A\cap B)}{P(B)}$ if $P(B) > 0$
- Multiplication Rule: $P(A \cap B) = P(A|B)P(B)$
- Chain Rule: $P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots$

Independence:

- $P(A \cap B) = P(A)P(B)$
- Equivalent: $P(A|B) = P(A)$, $P(B|A) = P(B)$
- Multiple: $P(A_1 \cap \cdots \cap A_k) = \prod P(A_i)$

Law of Total Probability: If $B_1, \ldots, B_n$ partition S:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

Bayes' Theorem:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}$$

Random Variables:

- Function $X : S \to \mathbb{R}$
- Discrete: Takes countable values
- Continuous: Takes uncountable values

Cumulative Distribution Function (CDF):

$$F_X(x) = P(X \leq x)$$

Properties: Non-decreasing, right-continuous, $F(-\infty) = 0$, $F(\infty) = 1$

Random Sampling:

- i.i.d.: Independent and identically distributed
- With replacement: Independent observations
- Without replacement: Dependent (unless $n \ll N$)

Expectation:

- Discrete: $\mathbb{E}[X] = \sum_x x \cdot p(x)$
- Continuous: $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

Properties of Expectation:

- Linearity: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$
- $\mathbb{E}[g(X)] = \sum g(x)p(x)$ (discrete) or $\int g(x)f(x)dx$ (continuous)
- $\mathbb{E}[c] = c$ (constant)
- If X, Y independent: $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. In general: $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] + \mathrm{Cov}(X, Y)$

Variance:

- Definition: $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$
- If independent: $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Covariance:

- $\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X, Y)$
- Correlation: $\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$, $-1 \leq \rho \leq 1$

Conditional Expectation:
$\mathbb{E}[Y|X = x] = \sum_y y \cdot p(y|x)$ or $\int y \cdot f(y|x)dy$
Law of Total Expectation: $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$

2. MGF

Definition: $M_X(t) = \mathbb{E}[e^{tX}]$

- Discrete: $M_X(t) = \sum_x e^{tx}p(x)$
- Continuous: $M_X(t) = \int_{-\infty}^{\infty} e^{tx}f(x)\,dx$

Properties of MGF:

- Uniqueness: MGF uniquely determines distribution
- Finding moments: $\mathbb{E}[X^n] = M_X^{(n)}(0)$ (nth derivative at 0)
- $\mathbb{E}[X] = M_X'(0)$, $\mathbb{E}[X^2] = M_X''(0)$
- $\mathrm{Var}(X) = M_X''(0) - [M_X'(0)]^2$
- Linear transformation: $M_{aX+b}(t) = e^{bt}M_X(at)$
- Independent sum: If X, Y independent, $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

3. Common Distributions & MGF

3.1 Discrete

Bernoulli Distribution: $X \sim \mathrm{Bern}(p)$

- PMF: $P(X = 1) = p$, $P(X = 0) = 1 - p$
- $\mathbb{E}[X] = p$, $\mathrm{Var}(X) = p(1 - p)$
- MGF: $M_X(t) = (1 - p) + pe^t$
- Derivation: $M_X(t) = \mathbb{E}[e^{tX}] = e^{t\cdot 0}(1 - p) + e^{t\cdot 1}p = 1 - p + pe^t$
- $\mathbb{E}[X] = M_X'(0) = pe^t|_{t=0} = p$
- $\mathbb{E}[X^2] = M_X''(0) = pe^t|_{t=0} = p$
- $\mathrm{Var}(X) = p - p^2 = p(1 - p)$

Binomial Distribution: $X \sim B(n, p)$

- PMF: $P(X = k) = \binom{n}{k}p^k(1 - p)^{n-k}$
- $\mathbb{E}[X] = np$, $\mathrm{Var}(X) = np(1 - p)$
- MGF: $M_X(t) = (1 - p + pe^t)^n = [pe^t + (1 - p)]^n$
- Derivation: $X = \sum_{i=1}^n X_i$, $X_i \sim \mathrm{Bern}(p)$ independent
- $M_X(t) = [M_{X_1}(t)]^n = (1 - p + pe^t)^n$
- $\mathbb{E}[X] = M_X'(0) = n(pe^t + 1 - p)^{n-1}pe^t|_{t=0} = np$
- $\mathbb{E}[X^2] = M_X''(0)$, calculating gives $\mathrm{Var}(X) = np(1 - p)$

Poisson Distribution: $X \sim \mathrm{Pois}(\lambda)$

- PMF: $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$
- $\mathbb{E}[X] = \lambda$, $\mathrm{Var}(X) = \lambda$
- MGF: $M_X(t) = e^{\lambda(e^t-1)}$
- Derivation: $M_X(t) = \sum_{k=0}^{\infty} e^{tk}\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)}$
- $\mathbb{E}[X] = M_X'(0) = \lambda e^t \cdot e^{\lambda(e^t-1)}|_{t=0} = \lambda$
- $\mathbb{E}[X^2] = M_X''(0) = [\lambda e^t + \lambda^2 e^{2t}]e^{\lambda(e^t-1)}|_{t=0} = \lambda + \lambda^2$
- $\mathrm{Var}(X) = \lambda + \lambda^2 - \lambda^2 = \lambda$

3.2 Continuous Distributions

Uniform Distribution: $X \sim U(a, b)$

- PDF: $f(x) = \frac{1}{b-a}$, $a \leq x \leq b$
- $\mathbb{E}[X] = \frac{a+b}{2}$, $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$
- MGF: $M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}$, $t \neq 0$
- Derivation: $M_X(t) = \int_a^b e^{tx}\frac{1}{b-a}dx = \frac{1}{b-a}\frac{e^{tb}-e^{ta}}{t}$
- $\mathbb{E}[X] = \lim_{t\to 0} M_X'(t) = \frac{a+b}{2}$ (using L'Hôpital's rule)

Exponential Distribution: $X \sim \mathrm{Exp}(\lambda)$

- PDF: $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$
- $\mathbb{E}[X] = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$
- MGF: $M_X(t) = \frac{\lambda}{\lambda-t}$, $t < \lambda$
- Derivation: $M_X(t) = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}dx = \lambda\int_0^{\infty} e^{-(\lambda-t)x}dx = \frac{\lambda}{\lambda-t}$
- $\mathbb{E}[X] = M_X'(0) = \frac{\lambda}{(\lambda-t)^2}|_{t=0} = \frac{1}{\lambda}$
- $\mathbb{E}[X^2] = M_X''(0) = \frac{2\lambda}{(\lambda-t)^3}|_{t=0} = \frac{2}{\lambda^2}$
- $\mathrm{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$
- Memoryless property: $P(X > s + t|X > s) = P(X > t)$

Normal Distribution: $X \sim N(\mu, \sigma^2)$

- PDF: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- $\mathbb{E}[X] = \mu$, $\mathrm{Var}(X) = \sigma^2$
- MGF: $M_X(t) = e^{\mu t+\frac{\sigma^2t^2}{2}}$
- Derivation: Let $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$
- $M_Z(t) = \int_{-\infty}^{\infty} e^{tz}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz = e^{t^2/2}$ (completing square)
- $M_X(t) = \mathbb{E}[e^{t(\sigma Z+\mu)}] = e^{\mu t}M_Z(\sigma t) = e^{\mu t+\sigma^2t^2/2}$
- $\mathbb{E}[X] = M_X'(0) = (\mu + \sigma^2 t)e^{\mu t+\sigma^2t^2/2}|_{t=0} = \mu$
- $\mathbb{E}[X^2] = M_X''(0) = \mu^2 + \sigma^2$
- Linear combination: $aX + b \sim N(a\mu + b, a^2\sigma^2)$

Gamma Distribution: $X \sim \mathrm{Gamma}(\alpha, \beta)$

- PDF: $f(x) = \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}$, $x > 0$
- Here α is shape parameter, β is scale parameter
- $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}dt$, $\Gamma(n) = (n - 1)!$
- $\mathbb{E}[X] = \alpha\beta$, $\mathrm{Var}(X) = \alpha\beta^2$
- MGF: $M_X(t) = (1 - \beta t)^{-\alpha}$, $t < 1/\beta$
- Derivation: $M_X(t) = \int_0^{\infty} e^{tx}\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}dx = \frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-x(1/\beta-t)}dx$
- Let $u = x(1/\beta - t)$: $= \frac{1}{\beta^\alpha\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha)}{(1/\beta-t)^\alpha} = \frac{1}{\beta^\alpha(1/\beta-t)^\alpha} = \left(\frac{1}{\beta(1/\beta-t)}\right)^\alpha = (1 - \beta t)^{-\alpha}$
- $\mathbb{E}[X] = M_X'(0) = -\alpha(1 - \beta t)^{-\alpha-1}(-\beta)|_{t=0} = \alpha\beta(1)^{-\alpha-1} = \alpha\beta$
- $\mathrm{Var}(X) = \alpha\beta^2$

Chi-squared Distribution: $X \sim \chi^2(n)$ (special Gamma)

- Definition: $\chi^2(n) = \text{Gamma}(n/2, 2)$
- If $Z_i \sim N(0,1)$ independent, then $\sum_{i=1}^n Z_i^2 \sim \chi^2(n)$
- PDF: $f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}$, $x > 0$
- $\mathbb{E}[X] = n$, $\text{Var}(X) = 2n$
- MGF: $M_X(t) = (1-2t)^{-n/2}$, $t < 1/2$
- Derivation: From Gamma MGF, $\alpha = n/2$, $\beta = 2$
- $M_X(t) = (1-2t)^{-n/2}$
- Additivity: $\chi^2(n_1) + \chi^2(n_2) = \chi^2(n_1 + n_2)$ (independent)

4. Multivariate RVs

Joint Distribution:

- Discrete: $p(x,y) = P(X = x, Y = y)$, $\sum_x \sum_y p(x,y) = 1$
- Continuous: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$

Marginal Distribution:

- Discrete: $p_X(x) = \sum_y p(x,y)$
- Continuous: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$

Conditional Distribution:

- Discrete: $p(y|x) = \frac{p(x,y)}{p_X(x)}$
- Continuous: $f(y|x) = \frac{f(x,y)}{f_X(x)}$

Independence:

- X, Y independent $\Leftrightarrow f(x,y) = f_X(x)f_Y(y)$
- $\Leftrightarrow M_{X,Y}(s,t) = M_X(s)M_Y(t)$
- If independent: $\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$
- $\text{Cov}(X,Y) = 0$, but converse not necessarily true

Covariance Matrix: For random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$:

$$\Sigma = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1,X_2) & \cdots & \text{Cov}(X_1,X_n) \\ \text{Cov}(X_2,X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n,X_1) & \text{Cov}(X_n,X_2) & \cdots & \text{Var}(X_n) \end{pmatrix}$$

Property: Σ is symmetric and positive semidefinite

Markov's Inequality: If $X \geq 0$ and $\mathbb{E}[X]$ exists:

$$P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}, \quad a > 0$$

Derivation: $\mathbb{E}[X] = \int_0^{\infty} xf(x)dx \geq \int_a^{\infty} xf(x)dx \geq a\int_a^{\infty} f(x)dx = aP(X \geq a)$

Chebyshev's Inequality: For any random variable X with $\mathbb{E}[X] = \mu$, $\text{Var}(X) = \sigma^2$:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Or: $P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$

Derivation: Let $Y = (X-\mu)^2 \geq 0$, by Markov's inequality:

$P((X-\mu)^2 \geq k^2\sigma^2) \leq \frac{\mathbb{E}[(X-\mu)^2]}{k^2\sigma^2} = \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$

Application: Shows most data is near the mean

5. Transformations

Method 1: CDF Method

1. Find CDF of $Y = g(X)$: $F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$
2. Differentiate to get PDF: $f_Y(y) = F_Y'(y)$

Method 2: Transformation Formula (Monotonic) If $Y = g(X)$ is monotonic, g differentiable:

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dx}{dy}\right| = f_X(x)\left|\frac{dx}{dy}\right|_{x=g^{-1}(y)}$$

General form (Jacobian): If $Y_1 = g_1(X_1, X_2)$, $Y_2 = g_2(X_1, X_2)$:

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(x_1,x_2)|J|$$

where $J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$

Method 3: MGF Method If $M_Y(t) = M_Z(t)$, then $Y \stackrel{d}{=} Z$ (same distribution)

Method 4: Convolution (Independent Sum) If X, Y independent, $Z = X + Y$:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx = (f_X * f_Y)(z)$$

Discrete: $p_Z(z) = \sum_x p_X(x)p_Y(z-x)$

6. Limit Theorems & Sampling

Law of Large Numbers (LLN):

- Weak LLN: If $X_1, X_2, \ldots$ i.i.d., $\mathbb{E}[X_i] = \mu$, $\text{Var}(X_i) = \sigma^2 < \infty$:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{P} \mu$$

  i.e.: $\forall \epsilon > 0$, $\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \epsilon) = 0$

- Proof using Chebyshev: $P(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2/n}{\epsilon^2} \to 0$

Central Limit Theorem (CLT): If $X_1, X_2, \ldots$ i.i.d., $\mathbb{E}[X_i] = \mu$, $\text{Var}(X_i) = \sigma^2 < \infty$:

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1)$$

Application: When n large, $\bar{X}_n \approx N(\mu, \sigma^2/n)$ or $\sum X_i \approx N(n\mu, n\sigma^2)$

Normal Distribution Properties:

- If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$
- If $X_i \sim N(\mu_i, \sigma_i^2)$ independent, then $\sum a_i X_i \sim N(\sum a_i\mu_i, \sum a_i^2\sigma_i^2)$
- If $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$:
  - $\bar{X} \sim N(\mu, \sigma^2/n)$
  - $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, where $S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$ (S^2 is the sample **variance**, S is the sample **standard deviation**)
  - $\bar{X}$ and S^2 are independent

t-Distribution: $T \sim t(n)$

- Definition: If $Z \sim N(0,1)$, $V \sim \chi^2(n)$ independent, then:

$$T = \frac{Z}{\sqrt{V/n}} \sim t(n)$$

- PDF: $f(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$
- $\mathbb{E}[T] = 0$ ($n > 1$), $\text{Var}(T) = \frac{n}{n-2}$ ($n > 2$)
- As $n \to \infty$, $t(n) \to N(0,1)$
- Application: If $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$:

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

  Derivation: $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, independent

$$\frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

F-Distribution: $F \sim F(n_1, n_2)$

- Definition: If $V_1 \sim \chi^2(n_1)$, $V_2 \sim \chi^2(n_2)$ independent:

$$F = \frac{V_1/n_1}{V_2/n_2} \sim F(n_1, n_2)$$

- $\mathbb{E}[F] = \frac{n_2}{n_2-2}$ ($n_2 > 2$)
- If $F \sim F(n_1, n_2)$, then $\frac{1}{F} \sim F(n_2, n_1)$
- If $T \sim t(n)$, then $T^2 \sim F(1, n)$
- Application: Comparing variances of two normal populations

Chi-squared Derivation: If $Z \sim N(0,1)$, find distribution of $Y = Z^2$:

- $F_Y(y) = P(Z^2 \leq y) = P(-\sqrt{y} \leq Z \leq \sqrt{y}) = 2\Phi(\sqrt{y}) - 1$
- $f_Y(y) = 2\phi(\sqrt{y})\frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}}e^{-y/2} = \frac{1}{2^{1/2}\Gamma(1/2)}y^{1/2-1}e^{-y/2}$
- Thus $Z^2 \sim \chi^2(1) = \text{Gamma}(1/2, 2)$
- If $Z_1, \ldots, Z_n$ i.i.d. $N(0,1)$, by MGF additivity: $\sum Z_i^2 \sim \chi^2(n)$

Important Formulas Summary:

- $\bar{X} \sim N(\mu, \sigma^2/n)$ (normal population)
- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ (σ known)
- $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ (σ unknown)
- $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$ (two populations)

7. Order Statistics

Definition: Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with CDF $F(x)$. The order statistics are the sorted values:

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$$

CDF of Order Statistics:

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

Mean and Variance:

- $\mathbb{E}[X_{(k)}] = \int_{-\infty}^{\infty} x \cdot f_{X_{(k)}}(x)\,dx$
- $\text{Var}(X_{(k)}) = \mathbb{E}[X_{(k)}^2] - (\mathbb{E}[X_{(k)}])^2$

Special Cases:

- **Minimum:** $X_{(1)}$: CDF: $F_{X_{(1)}}(x) = 1 - [1-F(x)]^n$, PDF: $f_{X_{(1)}}(x) = n[1-F(x)]^{n-1}f(x)$
- **Maximum:** $X_{(n)}$: CDF: $F_{X_{(n)}}(x) = [F(x)]^n$, PDF: $f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x)$
- **Median:** For odd $n = 2m+1$, sample median is $X_{(m+1)}$
  - For even $n = 2m$, sample median is usually $(X_{(m)} + X_{(m+1)})/2$

Sample Quantiles:

- p-th sample quantile: $X_{([np]+1)}$ where $[np]$ is the greatest integer less than or equal to np
- Sample median: 50th percentile
- Sample quartiles: 25th, 50th, 75th percentiles

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

- CDF: $F_R(r) = n\int_{-\infty}^{\infty}[F(x+r) - F(x)]^{n-1}f(x)\,dx$
- For uniform distribution $U(0,1)$: $X_{(k)} \sim \text{Beta}(k, n-k+1)$

Fall 2023 Midterm Solutions

Problem 1 (15 pts): A random variable Y has pdf $f(y) = \frac{1}{10}$ for $10 < y < 20$ (and zero otherwise). Find the moment generating function for Y.
Solution:

$$\begin{aligned} M_Y(t) = \mathbb{E}[e^{tY}] &= \int_{10}^{20} e^{ty} \cdot \frac{1}{10} dy \\ &= \frac{1}{10} \int_{10}^{20} e^{ty} dy \\ &= \frac{1}{10} \left[\frac{e^{ty}}{t} \right]_{10}^{20} \\ &= \frac{1}{10t}(e^{20t} - e^{10t}) \end{aligned}$$

$$\boxed{M_Y(t) = \frac{e^{20t} - e^{10t}}{10t}, \quad t \neq 0}$$

Problem 2 (15 pts): A random variable Y has mgf $m(t) = (1-3t)^{-5/2}$ for $t < 1/3$. **(a)** Calculate $E[Y]$. **(b)** Calculate $Var(Y)$.
Solution: Recognize this as Gamma distribution: $Y \sim \text{Gamma}(\alpha, \beta)$ with MGF $(1-\beta t)^{-\alpha}$.
Comparing $(1-3t)^{-5/2}$ with $(1-\beta t)^{-\alpha}$: $\beta = 3$, $\alpha = 5/2$.
(a) $\mathbb{E}[Y] = \alpha\beta = \frac{5}{2} \cdot 3 = \frac{15}{2}$
Alternatively using MGF: $M_Y'(t) = \frac{5}{2}(1-3t)^{-7/2} \cdot 3 = \frac{15}{2}(1-3t)^{-7/2}$
$\mathbb{E}[Y] = M_Y'(0) = \frac{15}{2}$

$$\boxed{\mathbb{E}[Y] = \frac{15}{2}}$$

(b) $\text{Var}(Y) = \alpha\beta^2 = \frac{5}{2} \cdot 9 = \frac{45}{2}$
Alternatively: $M_Y''(t) = \frac{15}{2} \cdot \frac{7}{2}(1-3t)^{-9/2} \cdot 3 = \frac{315}{4}(1-3t)^{-9/2}$
$\mathbb{E}[Y^2] = M_Y''(0) = \frac{315}{4}$
$\text{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = \frac{315}{4} - \frac{225}{4} = \frac{90}{4} = \frac{45}{2}$

$$\boxed{\text{Var}(Y) = \frac{45}{2}}$$

Problem 3 (10 pts): Seven patients were recruited for a clinical trial, 4 women and 3 men. They are ordered randomly for screening. What is the probability that all four women will be screened before any of the men?
Solution: Total number of orderings: 7!
Favorable orderings: Women in first 4 positions (in any order), men in last 3 positions (in any order).
Number of favorable orderings: $4! \cdot 3! = 24 \cdot 6 = 144$
$P(\text{all women first}) = \frac{4! \cdot 3!}{7!} = \frac{144}{5040} = \frac{1}{35}$
Alternatively: $P = \frac{1}{\binom{7}{4}} = \frac{1}{35}$

$$\boxed{P = \frac{1}{35}}$$

Problem 4 (15 pts): In a clinical trial, 30% of participants are randomly assigned to receive a placebo, 20% are assigned to Treatment A and the remaining 50% are assigned to Treatment B. Patients receiving placebo get better with probability 3/8 and patients receiving Treatment A get better with probability 5/8. If the probability that ANY patient entering the clinical trial will get better is 23/40, what is the probability that a patient assigned to Treatment B will get better?
Solution: Law of Total Probability:

$$\begin{aligned} P(\text{better}) = &P(\text{better}|\text{Placebo})P(\text{Placebo}) \\ &+ P(\text{better}|A)P(A) + P(\text{better}|B)P(B) \end{aligned}$$

$$\begin{aligned} \frac{23}{40} &= \frac{3}{8} \cdot 0.3 + \frac{5}{8} \cdot 0.2 + p \cdot 0.5 \\ \frac{23}{40} &= \frac{9}{80} + \frac{10}{80} + \frac{p}{2} \\ \frac{46}{80} &= \frac{19}{80} + \frac{p}{2} \\ \frac{27}{80} &= \frac{p}{2} \end{aligned}$$

$$\boxed{p = \frac{27}{40}}$$

Problem 5 (25 pts): Random variables Y_1 and Y_2 have joint pdf $f(y_1, y_2) = \frac{1}{2}e^{-(y_1+y_2)}$ for $0 < y_1 < y_2 < \infty$. **(a)** Find the marginal distribution of Y_2. **(b)** Find the distribution of Y_2 conditional on $Y_1 = 1$. **(c)** Are Y_1 and Y_2 independent? Explain.
Solution:

(a) Marginal distribution of Y_2:

$$\begin{aligned} f_{Y_2}(y_2) &= \int_0^{y_2} f(y_1, y_2) dy_1 \\ &= \int_0^{y_2} \frac{1}{2}e^{-(y_1+y_2)} dy_1 \\ &= \frac{1}{2}e^{-y_2} \int_0^{y_2} e^{-y_1} dy_1 \\ &= \frac{1}{2}e^{-y_2}[-e^{-y_1}]_0^{y_2} \\ &= \frac{1}{2}e^{-y_2}(1 - e^{-y_2}) \end{aligned}$$

$$\boxed{f_{Y_2}(y_2) = \frac{1}{2}(e^{-y_2} - e^{-2y_2}), \quad y_2 > 0}$$

(b) First find marginal of Y_1:

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{y_1}^{\infty} \frac{1}{2}e^{-(y_1+y_2)} dy_2 \\ &= \frac{1}{2}e^{-y_1} \int_{y_1}^{\infty} e^{-y_2} dy_2 \\ &= \frac{1}{2}e^{-y_1} \cdot e^{-y_1} = \frac{1}{2}e^{-2y_1} \end{aligned}$$

At $y_1 = 1$: $f_{Y_1}(1) = \frac{1}{2}e^{-2}$
Conditional pdf:

$$f_{Y_2|Y_1}(y_2|1) = \frac{f(1, y_2)}{f_{Y_1}(1)} = \frac{\frac{1}{2}e^{-(1+y_2)}}{\frac{1}{2}e^{-2}} = e^{1-y_2}$$

$$\boxed{f_{Y_2|Y_1}(y_2|1) = e^{1-y_2}, \quad y_2 > 1}$$

(c) Y_1 and Y_2 are $\boxed{\text{NOT independent}}$
Reason: The support constraint $y_1 < y_2$ shows dependence.

Problem 6 (20 pts): Random variables Y_1 and Y_2 have joint pdf $f(y_1, y_2) = \frac{1}{12}(y_1 + 2y_2)$ for $0 \le y_1 \le 2$ and $0 \le y_2 \le 2$. **(a)** Calculate $P(Y_1 > 1)$. **(b)** Calculate $P(Y_2 > 1|Y_1 > 1)$.
Solution:
(a) $P(Y_1 > 1)$:

$$\begin{aligned} P(Y_1 > 1) &= \int_1^2 \int_0^2 \frac{1}{12}(y_1 + 2y_2) dy_2 dy_1 \\ &= \int_1^2 \frac{1}{12}[y_1 y_2 + y_2^2]_0^2 dy_1 \\ &= \int_1^2 \frac{1}{12}(2y_1 + 4) dy_1 \\ &= \frac{1}{12}[y_1^2 + 4y_1]_1^2 \\ &= \frac{1}{12}[(4+8) - (1+4)] = \frac{7}{12} \end{aligned}$$

$$\boxed{P(Y_1 > 1) = \frac{7}{12}}$$

(b) $P(Y_2 > 1|Y_1 > 1)$:

$$\begin{aligned} P(Y_2 > 1, Y_1 > 1) &= \int_1^2 \int_1^2 \frac{1}{12}(y_1 + 2y_2) dy_2 dy_1 \\ &= \int_1^2 \frac{1}{12}[y_1 y_2 + y_2^2]_1^2 dy_1 \\ &= \int_1^2 \frac{1}{12}[(2y_1 + 4) - (y_1 + 1)] dy_1 \\ &= \int_1^2 \frac{y_1 + 3}{12} dy_1 \\ &= \frac{1}{12}[\frac{y_1^2}{2} + 3y_1]_1^2 \\ &= \frac{1}{12}[(2+6) - (\frac{1}{2} + 3)] = \frac{3}{8} \end{aligned}$$

$P(Y_2 > 1|Y_1 > 1) = \frac{P(Y_2>1, Y_1>1)}{P(Y_1>1)} = \frac{3/8}{7/12} = \frac{9}{14}$

$$\boxed{P(Y_2 > 1|Y_1 > 1) = \frac{9}{14}}$$

Fall 2024 Midterm Solutions

Problem 1 (15 pts): Events A and B are independent. Given that $P(A^c \cap B^c) = 0.6$ and $P(B) = 0.3$, find $P(A)$.
Solution: Since A and B are independent, A^c and B^c are also independent. Therefore:

$$\begin{aligned} P(A^c \cap B^c) &= P(A^c) \cdot P(B^c) \\ 0.6 &= P(A^c) \cdot (1 - P(B)) \\ 0.6 &= P(A^c) \cdot 0.7 \\ P(A^c) &= \frac{0.6}{0.7} = \frac{6}{7} \end{aligned}$$

Therefore: $\boxed{P(A) = 1 - P(A^c) = 1 - \frac{6}{7} = \frac{1}{7}}$

Problem 2 (15 pts): Let $X \sim N(\mu_1, \sigma_1^2)$ be independent of $Y \sim N(\mu_2, \sigma_2^2)$. Find the distribution of $3Y - 2X$ using moment generating functions.
Solution: Let $W = 3Y - 2X$. The MGF of W is:

$$\begin{aligned} M_W(t) &= \mathbb{E}[e^{t(3Y-2X)}] = \mathbb{E}[e^{3tY - 2tX}] \\ &= \mathbb{E}[e^{3tY}] \cdot \mathbb{E}[e^{-2tX}] \quad \text{(independence)} \\ &= M_Y(3t) \cdot M_X(-2t) \end{aligned}$$

For $X \sim N(\mu_1, \sigma_1^2)$: $M_X(t) = e^{\mu_1 t + \frac{\sigma_1^2 t^2}{2}}$
For $Y \sim N(\mu_2, \sigma_2^2)$: $M_Y(t) = e^{\mu_2 t + \frac{\sigma_2^2 t^2}{2}}$

$$\begin{aligned} M_W(t) &= e^{3\mu_2 t + \frac{9\sigma_2^2 t^2}{2}} \cdot e^{-2\mu_1 t + \frac{4\sigma_1^2 t^2}{2}} \\ &= e^{(3\mu_2 - 2\mu_1)t + \frac{1}{2}(9\sigma_2^2 + 4\sigma_1^2)t^2} \end{aligned}$$

This is the MGF of $N(3\mu_2 - 2\mu_1, 4\sigma_1^2 + 9\sigma_2^2)$.

$\boxed{3Y - 2X \sim N(3\mu_2 - 2\mu_1, 4\sigma_1^2 + 9\sigma_2^2)}$

Problem 3 (10 pts): A population has mean $\mu = 95$ and standard deviation $\sigma = 15$. A random sample of size $n = 100$ is taken. Use the Central Limit Theorem to find $P(93 < \bar{X} < 99)$.
Solution: By the Central Limit Theorem:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(95, \frac{225}{100}\right) = N(95, 2.25)$$

Standard deviation: $\sigma_{\bar{X}} = \frac{15}{\sqrt{100}} = 1.5$
Standardize:

$$\begin{aligned} P(93 < \bar{X} < 99) &= P\left(\frac{93 - 95}{1.5} < Z < \frac{99 - 95}{1.5}\right) \\ &= P\left(-\frac{4}{3} < Z < \frac{8}{3}\right) \\ &= \Phi\left(\frac{8}{3}\right) - \Phi\left(-\frac{4}{3}\right) \end{aligned}$$

$\boxed{P(93 < \bar{X} < 99) = \Phi(8/3) - \Phi(-4/3)}$

Problem 4 (20 pts): Let $X_1, \ldots, X_{10}$ be iid $N(\mu_1, \sigma^2)$ independent of $Y_1, \ldots, Y_{20}$ iid $N(\mu_2, \sigma^2)$. Let s_1^2 and s_2^2 be the sample variances of the X_i's and Y_i's, respectively. **(a)** Calculate the expectation and variance of $\frac{1}{2}(s_1^2 + s_2^2)$. **(b)** Calculate the expectation and variance of $\frac{1}{28}(9s_1^2 + 19s_2^2)$.

Solution: Recall: $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$
For X sample: $\frac{9s_1^2}{\sigma^2} \sim \chi^2(9)$
For Y sample: $\frac{19s_2^2}{\sigma^2} \sim \chi^2(19)$
(a) Expectation:

$$\begin{aligned} \mathbb{E}\left[\frac{1}{2}(s_1^2 + s_2^2)\right] &= \frac{1}{2}(\mathbb{E}[s_1^2] + \mathbb{E}[s_2^2]) \\ &= \frac{1}{2}(\sigma^2 + \sigma^2) \\ &= \boxed{\sigma^2} \end{aligned}$$

Variance: For $\chi^2(k)$, variance is $2k$.
$\text{Var}\left(\frac{9s_1^2}{\sigma^2}\right) = 2(9) = 18 \Rightarrow \text{Var}(s_1^2) = \frac{18\sigma^4}{81} = \frac{2\sigma^4}{9}$
$\text{Var}\left(\frac{19s_2^2}{\sigma^2}\right) = 2(19) = 38 \Rightarrow \text{Var}(s_2^2) = \frac{38\sigma^4}{361} = \frac{2\sigma^4}{19}$

$$\begin{aligned} \text{Var}\left[\frac{1}{2}(s_1^2 + s_2^2)\right] &= \frac{1}{4}[\text{Var}(s_1^2) + \text{Var}(s_2^2)] \\ &= \frac{1}{4}\left(\frac{2\sigma^4}{9} + \frac{2\sigma^4}{19}\right) \\ &= \boxed{\frac{\sigma^4}{2}\left(\frac{1}{9} + \frac{1}{19}\right) = \frac{\sigma^4}{2} \cdot \frac{28}{171} = \frac{14\sigma^4}{171}} \end{aligned}$$

(b) Expectation:

$$\mathbb{E}\left[\frac{1}{28}(9s_1^2 + 19s_2^2)\right] = \frac{1}{28}(9\sigma^2 + 19\sigma^2) = \sigma^2$$

$\boxed{\mathbb{E}\left[\frac{1}{28}(9s_1^2 + 19s_2^2)\right] = \sigma^2}$

Variance:

$$\begin{aligned} \text{Var}\left[\frac{1}{28}(9s_1^2 + 19s_2^2)\right] &= \frac{1}{784}[81\text{Var}(s_1^2) + 361\text{Var}(s_2^2)] \\ &= \frac{1}{784}\left(81 \cdot \frac{2\sigma^4}{9} + 361 \cdot \frac{2\sigma^4}{19}\right) \\ &= \frac{1}{784}(18\sigma^4 + 38\sigma^4) = \frac{56\sigma^4}{784} = \frac{\sigma^4}{14} \end{aligned}$$

$\boxed{\text{Var}\left[\frac{1}{28}(9s_1^2 + 19s_2^2)\right] = \frac{\sigma^4}{14}}$

Problem 5 (20 pts): Random variables X and Y have $\mathbb{E}[X] = 120$, $\sigma_X = 10$, $\mathbb{E}[Y] = 180$, $\sigma_Y = 20$, and correlation $\rho = 1/4$. Let $Z = \frac{2}{5}X + \frac{1}{5}Y$. **(a)** Calculate $\mathbb{E}[Z]$. **(b)** Calculate $\text{Var}(Z)$. **(c)** Calculate $\text{Var}(Z)$ if X and Y are independent.
Solution:
(a) Expectation:

$$\begin{aligned} \mathbb{E}[Z] &= \mathbb{E}\left[\frac{2}{5}X + \frac{1}{5}Y\right] \\ &= \frac{2}{5}\mathbb{E}[X] + \frac{1}{5}\mathbb{E}[Y] \\ &= \frac{2}{5}(120) + \frac{1}{5}(180) \\ &= 48 + 36 = 84 \end{aligned}$$

$\boxed{\mathbb{E}[Z] = 84}$

(b) Variance: First find covariance: $\text{Cov}(X, Y) = \rho\sigma_X\sigma_Y = \frac{1}{4}(10)(20) = 50$

$$\begin{aligned} \text{Var}(Z) &= \text{Var}\left(\frac{2}{5}X + \frac{1}{5}Y\right) \\ &= \frac{4}{25}\text{Var}(X) + \frac{1}{25}\text{Var}(Y) + 2 \cdot \frac{2}{5} \cdot \frac{1}{5}\text{Cov}(X, Y) \\ &= \frac{4}{25}(100) + \frac{1}{25}(400) + \frac{4}{25}(50) \\ &= \frac{400 + 400 + 200}{25} = \frac{1000}{25} = 40 \end{aligned}$$

$\boxed{\text{Var}(Z) = 40}$

(c) If independent, $\text{Cov}(X, Y) = 0$:

$$\begin{aligned} \text{Var}(Z) &= \frac{4}{25}(100) + \frac{1}{25}(400) + 0 \\ &= \frac{800}{25} = 32 \end{aligned}$$

$\boxed{\text{Var}(Z) = 32}$

Problem 6 (20 pts): Let Y_1, Y_2, Y_3 be iid random variables with CDF $F(y) = y^2$ for $0 < y < 1$. Find $\mathbb{E}[Y_{(3)} - Y_3]$, where $Y_{(3)} = \max\{Y_1, Y_2, Y_3\}$.
Solution: PDF: $f(y) = F'(y) = 2y$ for $0 < y < 1$
First find $\mathbb{E}[Y_3]$ (any single Y_i):

$$\begin{aligned} \mathbb{E}[Y_3] &= \int_0^1 y \cdot 2y\, dy = \int_0^1 2y^2 dy \\ &= 2 \cdot \frac{y^3}{3}\Big|_0^1 = \frac{2}{3} \end{aligned}$$

For maximum $Y_{(3)}$, the PDF is:

$$\begin{aligned} f_{Y_{(3)}}(y) &= n[F(y)]^{n-1} f(y) \\ &= 3[y^2]^2 \cdot 2y = 3y^4 \cdot 2y = 6y^5 \end{aligned}$$

Then:

$$\begin{aligned} \mathbb{E}[Y_{(3)}] &= \int_0^1 y \cdot 6y^5 dy = \int_0^1 6y^6 dy \\ &= 6 \cdot \frac{y^7}{7}\Big|_0^1 = \frac{6}{7} \end{aligned}$$

Therefore:

$$\begin{aligned} \mathbb{E}[Y_{(3)} - Y_3] &= \mathbb{E}[Y_{(3)}] - \mathbb{E}[Y_3] \\ &= \frac{6}{7} - \frac{2}{3} = \frac{18 - 14}{21} = \frac{4}{21} \end{aligned}$$

$\boxed{\mathbb{E}[Y_{(3)} - Y_3] = \frac{4}{21}}$