STA347 Problem Set

Problem 1. Using the definition of probability measure, prove that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Solution. Theorem in lecture note.

Problem 2. Let A_n be a sequence of events. Prove Boole's inequality, that is,

$$P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n).$$

Solution. Theorem in lecture note.

Problem 3. A_n is a monotone decreasing event to $\emptyset$, that is, $A_1 \supset \cdots \supset A_{n-1} \supset A_n \supset A_{n+1} \supset \cdots$ and $\cap_{n=1}^{\infty} A_n = \emptyset$. Prove that $\lim_{n\to\infty} P(A_n) = 0$.

Solution. Theorem in lecture note.

Problem 4. For $P(A) > 0$, define $Q_A(B) = P(B \mid A) = P(B \cap A)/P(A)$. Prove that Q_A is a probability measure.

Solution. (a) $Q_A(B) = P(B \cap A)/P(A) \geq 0$ for $B \in \mathcal{F}$ and $Q_A(\emptyset) = P(B \cap \emptyset)/P(A) = 0$
(b) Let $B_1, B_2, \ldots$ be disjoint events. Then $A \cap B_1, A \cap B_2, \ldots$ are disjoint events too. Hence $Q_A(\cup_{n=1}^{\infty} B_n) = P(\cup_{n=1}^{\infty} B_n \cap A)/P(A) = P(\cup_{n=1}^{\infty}(B_n \cap A))/P(A) = \sum_{n=1}^{\infty} P(B_n \cap A)/P(A) = \sum_{n=1}^{\infty} Q_A(B_n)$.
(c) $Q_A(S) = P(S \cap A)/P(A) = P(A)/P(A) = 1$.
Hence Q_A is probability.

Problem 5. Show that the distribution function F of X satisfies the followings.
(a) F is non-decreasing.
(b) $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$.
(c) F is right-continuous, that is, $F(x+) = \lim_{h\searrow 0} F(x+h) = F(x)$.

Solution. Theorem in lecture note

Problem 6. Suppose $X \sim \text{Poisson}(\mu)$. Prove that

$$\sum_{n=0}^{\infty} P(X > n) = \mathbb{E}(X).$$

Solution.

$$\sum_{n=0}^{\infty} P(X > n) = \sum_{n=0}^{\infty} \sum_{k=n+1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} \sum_{n=0}^{k-1} P(X = k) = \sum_{k=1}^{\infty} kP(X = k) = \mathbb{E}(X).$$

In the second equality, Fubini's theorem was used.

Problem 7. Let X be a random variable. Define $X_+ = \max(0, X)$ and $X_- = \max(0, -X)$. Show that $X = X_+ - X_-$ and $|X| = X_+ + X_-$.

Solution. If $X \geq 0$, then $X_+ = X$ and $X_- = 0$ imply $X = X_+ - X_-$ as well as $|X| = X = X_+ + X_-$. If $X \leq 0$, then $X_+ = 0$ and $X_- = -X$ imply $X = X_+ - X_-$ and $|X| = -X = X_+ + X_-$. Hence $X = X_+ - X_-$ and $|X| = X_+ + X_-$.

Problem 8. If $X_n \geq 0$, then

$$\mathbb{E}(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} \mathbb{E}(X_n).$$

Solution. See Theorem in lecture note.

Problem 9. Write the probability density/mass function, probability generating function, moment generating function, cumulative generating function and characteristic function of the following distributions.
(a) $Bernoulli(p)$; (b) $Binomial(n, p)$; (c) Poisson(λ); (d) $Geometric(p)$; (e) $Negative - Binomial(r, p)$; (f) $N(\mu, \sigma^2)$; (g) $Uniform(\theta_1, \theta_2)$; (h) $Gamma(\alpha, \beta)$; (i) $Beta(\alpha, \beta)$ and (j) $Exponential(\mu) \sim Gamma(1, \mu)$.

Problem 10. Suppose $\mathbb{E}(|X|) < \infty$, F is the distribution function of X and $f(x) = F'(x)$ is the derivative of F. Prove that

$$\int_0^\infty (1 - F(x))dx - \int_{-\infty}^0 F(x)dx = \mathbb{E}(X).$$

Solution. Note that $F(x) = P(X \leq x) = \int_{-\infty}^x f(z)\, dz$. Suppose $X \geq 0$.

$$\begin{aligned}\mathbb{E}(X) &= \int_0^\infty x\, f(x)\, dx = \int_0^\infty \int_0^x 1(z < x)\, dz\, f(x)\, dx = \int_0^\infty \int_0^\infty 1(z < x) f(x)\, dx\, dz \\ &= \int_0^\infty \int_z^\infty 1(z < x) f(x)\, dx\, dz = \int_0^\infty P(X > z)\, dz = \int_0^\infty (1 - F(x))\, dx.\end{aligned}$$

For general X, $X = X_+ - X_-$. Then

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X_+) - \mathbb{E}(X_-) = \int_0^\infty P(X_+ > x)\, dx - \int_0^\infty P(X_- > x)\, dx \\ &= \int_0^\infty P(X > x)\, dx - \int_0^\infty P(-X > x)\, dx = \int_0^\infty (1 - F(X))\, dx - \int_{-\infty}^0 P(X < x)\, dx \\ &= \int_0^\infty (1 - F(X))\, dx - \int_{-\infty}^0 F(x)\, dx\end{aligned}$$

Note that $\int_0^\infty P(X = -x)\, dx = 0$ because $P(X = -x) = 0$ almost everywhere.

Problem 11. Suppose X, Y have the joint density $f(x, y)$. Prove that $W = \int x f(x, y)/f_Y(y) dx$ satisfies that for any set B, $\mathbb{E}(1_A W) = \mathbb{E}(1_A X)$ where $A = (Y \in B)$.

Solution. $\mathbb{E}(1_A W) = \int 1(y \in B) \int x f(x,y)/f_Y(y)\, dx\ f_Y(y)\, dy = \int\int 1(y \in B) x f(x,y)\, dx\, dy = \mathbb{E}(1_A X)$.

Problem 12. Prove Slutsky's Theorem.

Solution. See Theorem in lecture notes.

Problem 13. Suppose X_i's are i.i.d. random variables having $\mathbb{E}(X_i) = \mu$ and $\mathbb{E}(X_i^2) = \mu_2$. Let $\overline{X} = (X_1 + \cdots + X_n)/n$ and $S^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \overline{X})^2$. Prove that

$$\mathbb{E}(\overline{X}) = \mu \quad \text{and} \quad \mathbb{E}(S^2) = \mathbb{V}\text{ar}(X_i).$$

Solution.

$$\begin{aligned}
\mathbb{E}(\overline{X}) &= \mathbb{E}[(X_1 + \cdots + X_n)/n] = n^{-1}[\mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)] = n^{-1}[n\mathbb{E}(X_1)] = \mathbb{E}(X_1) = \mu, \\
(n-1)S^2 &= \sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n [X_i^2 - 2\overline{X}X_i + \overline{X}^2] = \sum_{i=1}^n X_i^2 - 2n\overline{X}^2 + n\overline{X}^2 = \sum_{i=1}^n X_i^2 - n\overline{X}^2, \\
\mathbb{E}[\sum_{i=1}^n X_i^2] &= \sum_{i=1}^n \mathbb{E}(X_i^2) = n\mu_2, \\
\mathbb{E}[\overline{X}^2] &= \mathbb{E}[\sum_{i=1}^n X_i/n \sum_{j=1}^n X_j/n] = n^{-2} \sum_{i,j=1}^n \mathbb{E}[X_i X_j] = n^{-2} \sum_{i=1}^n \mathbb{E}(X_i^2) + n^{-2} \sum_{i \neq j} \mathbb{E}(X_i X_j) \\
&= n^{-2} \cdot n\mu_2 + n^{-2} \cdot n(n-1)\mu^2 = \mu_2/n + (n-1)\mu^2/n, \\
\mathbb{E}(S^2) &= (n-1)^{-1}(n\mu_2 - n \cdot (\mu_2/n + (n-1)\mu^2/n)) = (n-1)^{-1}((n-1)\mu_2 - (n-1)\mu^2) \\
&= \mu_2 - \mu^2 = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = \mathbb{V}\text{ar}(X_i).
\end{aligned}$$

Problem 14. Let $X_1, \ldots, X_n$ be a random sample from Poisson(θ).
(a) Find the distribution of $T = X_1 + \cdots + X_n$.
(b) Compute $\mathbb{E}(T), \mathbb{E}(T^2), \mathbb{E}(T^3)$ and $\mathbb{E}(T^4)$.

Solution. (a) Suppose $X \sim$ Poisson(μ) and $Y \sim$ Poisson(λ) are independent. Then

$$\begin{aligned}
P(X + Y = n) &= \sum_{x=0}^n P(X = x, Y = n - x) = \sum_{x=0}^n e^{-\mu} \frac{\mu^x}{x!} \times e^{-\lambda} \frac{\lambda^{n-x}}{(n-x)!} \\
&= \frac{\exp(-(\mu+\lambda))}{n!} \sum_{x=0}^n \binom{n}{x} \mu^x \lambda^{n-x} = \frac{\exp(-(\mu+\lambda))}{n!} (\mu+\lambda)^n \sim \text{Poisson}(\mu+\lambda).
\end{aligned}$$

Hence $X_1 + X_2 \sim \text{Poisson}(\theta + \theta) \sim \text{Poisson}(2\theta)$, $X_1 + X_2 + X_3 = (X_1 + X_2) + X_3 \sim \text{Poisson}(2\theta + \theta) \sim \text{Poisson}(3\theta)$ and by mathematical induction, $X_1 + \cdots + X_n \sim \text{Poisson}(n\theta)$.
(b) Note that $\mathbb{E}(T(T-1)\cdots(T-k+1)) = (n\theta)^k$. Hence $\mathbb{E}(T) = (n\theta)$, $\mathbb{E}(T^2) = \mathbb{E}(T(T-1) + T) = (n\theta)^2 + (n\theta)$, $\mathbb{E}(T^3) = \mathbb{E}[T(T-1)(T-2) + 3T(T-1) + T] = (n\theta)^3 + 3(n\theta)^2 + n\theta$ and finally $\mathbb{E}(T^4) = \mathbb{E}[T(T-1)(T-2)(T-3) + 6T(T-1)(T-2) + 7T(T-1) + T] = (n\theta)^4 + 6(n\theta)^3 + 7(n\theta)^2 + n\theta$.

Problem 15. Let $X_1, \ldots, X_n$ be a random sample from a distribution having a density

$$f_\theta(x) = I(x \geq \theta)c(\theta)/x^4 \quad \text{for } \theta > 0.$$

(a) Find $c(\theta)$.
(b) Find the density of $X_{(1)} = \min(X_1, \ldots, X_n)$.
(c) Compute mean and variance of X_1 and $X_{(1)}$.

Solution. (a) From the totality, $1 = \int f_\theta(x)\ dx = \int_\theta^\infty c(\theta)/x^4\ dx = [-c(\theta)x^{-3}/3]_\theta^\infty = c(\theta)/(3\theta^3)$. Hence $c(\theta) = 3\theta^3$.
(b) Note that

$$\begin{aligned} P(X_{(1)} \leq x) &= 1 - P(X_{(1)} > x) = 1 - P(X_1 > x, X_2 > x, \ldots, X_n > x) = 1 - P(X_1 > x) \cdots P(X_n > x) \\ &= 1 - [P(X_1 > x)]^n = 1 - [1 - P(X_1 \leq x)]^n = 1 - [1 - F(x)]^n. \end{aligned}$$

Hence $\text{pdf}_{X_{(1)}}(x) = \frac{d}{dx}P(X_{(1)} \leq x) = n(1 - F(x))^{n-1}\frac{d}{dx}F(x) = n(1 - F(x))^{n-1}f(x) = n(\theta^3/x^3)^{n-1} \cdot 3\theta^3 1(x \geq \theta)/x^4 = 3n1(x \geq \theta)\theta^{3n}/x^{3n+1}$.
(c) For $0 < k < 3$, $\mathbb{E}(X_1^k) = \int_\theta^\infty x^k \cdot 3\theta^3/x^4\ dx = [-3\theta^3/x^{3-k}/(3-k)]_\theta^\infty = 3\theta^3/\theta^{3-k}/(3-k) = \theta^k/(1 - k/3)$. Hence $\mathbb{E}(X_1) = \theta/(1 - 1/3) = 3\theta/2$ and $\mathbb{V}\text{ar}(X_1) = \mathbb{E}(X_1^2) - (\mathbb{E}(X_1))^2 = \theta^2/(1 - 2/3) - (\theta/(1 - 1/3))^2 = (3/4)\theta^2$. The density of $X_{(1)}$ is $n(1 - F(x))^{n-1}f(x) = n(\theta/x)^{3(n-1)}1(x \geq \theta)3\theta^3/x^4 = 3n\theta^{3n}1(x \geq \theta)/x^{3n+1}$. Hence, for $0 \leq k < 3n$, $\mathbb{E}(X_{(1)}^k) = \int_\theta^\infty x^k \cdot 3n\theta^{3n}/x^{3n+1}\ dx = \theta^k/(1 - k/(3n))$ and $\mathbb{E}(X_{(1)}) = \theta/(1 - 1/3n)$ and $\mathbb{V}\text{ar}(X_{(1)}) = \mathbb{E}(X_{(1)}^2) - (\mathbb{E}(X_{(1)}))^2 = \theta^2/(1 - 2/3n) - (\theta/(1 - 1/3n))^2 = \theta^2 3n/(3n-1)^2/(3n-2)$. As $n \to \infty$, $X_{(1)}$ concentrated around θ more and more.

Problem 16. Assume $X_1, \ldots, X_n$ are i.i.d. random variables sampled from Uniform$(\theta - 1, \theta + 1)$. Find the mean and variance of $\bar{X}_n = (X_1 + \cdots + X_n)$, $X_{(1)} = \min(X_1, \ldots, X_n)$ and $X_{(n)} = \max(X_1, \ldots, X_n)$.

Solution. Let F be the distribution function of X_n. The joint distribution function of $(X_{(1)}, X_{(n)})$ is

$$\begin{aligned} &\text{cdf}_{X_{(1)},X_{(n)}}(x, y) = P(\min(X_1, \ldots, X_n) \leq x, \max(X_1, \ldots, X_n) \leq y) \\ &= P(\max(X_1, \ldots, X_n) \leq y) - P(\min(X_1, \ldots, X_n) > x, \max(X_1, \ldots, X_n) \leq y) \\ &= P(X_1 \leq y, \ldots, X_n \leq y) - P(x < X_1 \leq y, \ldots, x < X_n \leq y) = P(X_1 \leq y)^n - P(x < X_1 \leq y)^n \\ &= (F(y))^n - (F(y) - F(x))^n. \end{aligned}$$

Hence the joint density function becomes

$$\begin{aligned} \text{pdf}_{X_{(1)},X_{(n)}}(x, y) &= \frac{d^2}{dydx}\text{cdf}_{X_{(1)},X_{(n)}}(x, y) = \frac{d}{dy}n(F(y) - F(x))^{n-1}f(x) \\ &= n(n-1)(F(y) - F(x))^{n-2}f(x)f(y). \end{aligned}$$

where $f(x) = F'(x)$. Hence the joint density function is $\text{pdf}_{X_{(1)},X_{(n)})}(x, y) = 1(\theta - 1 < x \leq y \leq \theta + 1)((y - x)/2)^{n-2}/4 = 2^{-n}1(\theta - 1 \leq x \leq y \leq \theta + 1)(y - x)^{n-2}$. Similarly, the marginal density functions are $\mathrm{pdf}_{X_{(1)}}(x) = 2^{-n}1(\theta-1 \le x \le \theta = 1)n(\theta+1-x)^{n-1}$ and $\mathrm{pdf}_{X_{(n)}}(y) = 2^{-n}(y-\theta+1)^{n-1}$. Note that $X_{(1)} - (\theta-1) \equiv^d (\theta+1) - X_{(n)}$ due to symmetry of the uniform distribution. Then,

$$\mathbb{E}((X_{(1)} - (\theta-1))^k) = \int_{\theta-1}^{\theta+1} (x-\theta+1)^k 2^{-n} \cdot n(\theta+1-x)^{n-1}\,dx = 2^k n \int_0^1 (1-z)^k z^{n-1}\,dz$$

Transformation $z = (\theta+1-x)/2$ or $x = \theta+1-2z$ is used in the previous equality. It is beta function.

$$= 2^k n\Gamma(k+1)\Gamma(n)/\Gamma(n+k+1) = 2^k k!n!/(n+k)! = 2^k k!/[(n+1)\cdots(n+k)].$$

Hence $\mathbb{E}(X_{(1)}) = (\theta-1)+2^1 1/(n+1) = \theta-1+2/(n+1)$ and $\mathbb{V}\mathrm{ar}(X_{(1)}) = \mathbb{V}\mathrm{ar}(X_{(1)}-\theta+1) = 2^2 2!/[(n+1)(n+2)] - (2/(n+1))^2 = 8/[(n+1)(n+2)] - 4/(n+1)^2 = 4[2(n+1)-(n+2)]/[(n+1)^2(n+2)] = 8n/[(n+1)^2(n+2)]$. Using symmetry, $\mathbb{E}(X_{(n)}) = \theta+1-2/(n+1)$ and $\mathbb{V}\mathrm{ar}(X_{(n)}) = 8n/[(n+1)^2(n+2)]$.

Problem 17. Suppose that $X_1 \sim N(\mu,\sigma^2), X_2 \sim N(3\mu, 4\sigma^2)$ are independent.

(a) Show $T_{a,b} = aX_1 + bX_2$ is a normal distribution.
(b) Compute mean and variance of $T_{a,b} = aX_1 + bX_2$.
(c) Find a condition for a, b to make $\mathbb{E}(T_{a,b}) = \mu$.
(d) Find a, b so that $\mathbb{V}\mathrm{ar}(T_{a,b})$ is the smallest satisfying $\mathbb{E}(T_{a,b}) = \mu$.

Solution. (a) $\mathrm{mgf}_{T_{a,b}}(t) = \mathbb{E}[e^{t(aX_1+bX_2)}] = \mathbb{E}[e^{(ta)X_1}]\mathbb{E}[e^{(tb)X_2}] = e^{(ta)\mu+(ta)^2\sigma^2/2}e^{(tb)(3\mu)+(tb)^2(4\sigma^2)/2} = e^{t(a+3b)\mu+t^2(a^2+4b^2)\sigma^2/2}$ implies $T_{a,b} \sim N((a+3b)\mu, (a^2+4b^2)\sigma^2)$.
(b) From part (a), $\mathbb{E}[T_{a,b}] = (a+3b)\mu$ and $\mathbb{V}\mathrm{ar}(T_{a,b}) = (a^2+4b^4)\sigma^2$.
(c) $a+3b=1$.
(d) From part (c) $a = 1-3b$. Hence $\mathbb{V}\mathrm{ar}(T_{a,b}) = \sigma^2(a^2+4b^4) = \sigma^2((1-3b)^2+4b^2) = \sigma^2(13b^2-6b+1)$ which is minimized at $b = 3/13$. Thus $a = 2/13$ and $b = 3/13$ minimizes the variance of $T_{a,b}$ among all $T_{a,b}$'s having mean μ.

Problem 18. Let $X_1,\dots,X_n$ be a random sample from the probability density function given by $f_\theta(x) = I(x > \mu)\sigma^{-1}\exp(-(x-\mu)/\sigma)$ where $\theta = (\mu,\sigma)$. Compute mean and variance of $X_{(1)}, \bar{X}$ where $X_{(1)} = \min(X_1,\dots,X_n)$.

Solution. Note that $X_i - \mu \sim i.i.d.$ Exponential$(1/\sigma)$. Hence $\mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = \mu + \mathbb{E}(X_1 - \mu) = \mu + 1/(1/\sigma) = \mu+\sigma$ and $\mathbb{V}\mathrm{ar}(\bar{X}) = \mathbb{V}\mathrm{ar}(X_1)/n = \mathbb{V}\mathrm{ar}(X_1-\mu)/n = 1/(1/\sigma)^2/n = \sigma^2/n$. Note that, for $x > \mu$, $P(X_{(1)} > x) = P(X_1 > x,\dots,X_n > x) = [P(X_1 > x)]^n = (e^{-(x-\mu)/\sigma})^n$. Hence $X_{(1)} - \mu \sim$ Exponential$(n/sigma)$. Then $\mathbb{E}(X_{(1)}) = \mu + 1/(n/\sigma) = \mu + \sigma/n$ and $\mathbb{V}\mathrm{ar}(X_{(1)} = \mathbb{V}\mathrm{ar}(X_{(1)} - \mu) = 1/(n/\sigma)^2 = \sigma^2/n^2$.

Problem 19. Let $X_1,\dots,X_n$ be a i.i.d. sample from a distribution having density $f_\theta(x) = I(x > 0)\theta^{-1}\exp(-x/\theta)$.
(a) Find the density function of $T = X_1 + \cdots + X_n$.
(b) Compute mean and variance of T.

Solution. Note that $X_i \sim i.i.d.$Exponential$(1/\theta)$. (a) $\text{mgf}_T(t) = \mathbb{E}[e^{t(X_1+\cdots+X_n)}] = [\mathbb{E}(e^{tX_1})]^n = [(1-t/(1/\theta))^{-1}]^n = (1-\theta t)^{-n}$ indicates $T \sim$ Gamma$(n, 1/\theta)$. Thus $\text{pdf}_T(x) = 1(x > 0)x^{n-1}e^{-x/\theta}$.
(b) $\mathbb{E}(T) = n/(1/\theta) = n\theta$ and $\mathbb{V}\text{ar}(T) = n/(1/\theta)^2 = n\theta^2$.

Problem 20. Let $X_1, \ldots, X_n$ be a random sample from Uniform(θ_1, θ_2) whose density is $I(\theta_1 \leq x \leq \theta_2)/(\theta_2 - \theta_1)$.
(a) Show that $n(X_{(1)} - \theta_1)$ converges in distribution.
(b) Show that $n(\theta_2 - X_{(n)})$ converges in distribution.
(c) Prove or disprove that $n(\theta_2 - X_{(n)} + X_{(1)} - \theta_1)$ converge in distribution.

Solution. (a) For $x > 0$, $P(n(X_{(1)} - \theta_1) > x) = P(X_{(1)} > \theta_1 + x/n) = [P(X_1 > \theta_1 + x/n)]^n = ((\theta_2 - \theta_1 - x/n)/(\theta_2 - \theta_1))^n = (1 - x/[n(\theta_2 - \theta_1)])^n \to e^{-x/(\theta_2-\theta_1)}$. Hence $n(X_{(1)} - \theta_1) \xrightarrow{d}$ Exponential$(1/(\theta_2 - \theta_1))$.
(b) Note that $X_{(1)} - \theta_1 \equiv^d \theta_2 - X_{(n)}$. Hence $n(\theta_2 - X_{(n)}) \xrightarrow{d}$ Exponential$(1/(\theta_2 - \theta_1))$.
(c) Let $Z_n = n(\theta_2 - X_{(n)} + X_{(1)} - \theta_1)$. Then

$$\mathbb{E}[Z_n^k] = \int_{\theta_1}^{\theta_2} \int_x^{\theta_2} [n(\theta_2 - y + x - \theta_1)]^k \cdot n(n-1)(\theta_2 - \theta_1)^{-n}(y-x)^{n-2}\, dydx$$

By taking $z = (y-x)/(\theta_2 - \theta_1)$

$$\begin{aligned}
&= n^k(\theta_2 - \theta_1)^{k-1} \int_{\theta_1}^{\theta_2} \int_0^1 1(z < (\theta_2 - x)/(\theta_2 - \theta_1))(1-z)^k n(n-1) z^{n-2}\, dz\, dx \\
&= n^{k+1}(n-1)(\theta_2 - \theta_1)^{k-1} \int_0^1 z^{n-2}(1-z)^k \int_0^1 1(x < \theta_2 - z(\theta_2 - \theta_1))\, dx\, dz \\
&= n^{k+1}(n-1)(\theta_2 - \theta_1)^{k-1} \int_0^1 z^{n-2}(1-z)^k \cdot (\theta_2 - \theta_1)(1-z)\, dz \\
&= n^{k+1}(n-1)(\theta_2 - \theta_1)^k \Gamma(k+2)\Gamma(n-1)/\Gamma(n+k+1) \\
&= (\theta_2 - \theta_1)^k \Gamma(k+2) \frac{n^{k+1}}{n(n+1)\cdots(n+k)} \longrightarrow (\theta_2 - \theta_1)^k \Gamma(k+2)/\Gamma(2)
\end{aligned}$$

which is the kth moment of Gamma$(2, 1/(\theta_2 - \theta_1))$. Hence $Z_n \xrightarrow{d}$ Gamma$(2, 1/(\theta_2 - \theta_1))$.

Problem 21. $X_i \sim i.i.d.$ Uniform$(0, \theta)$ for $i = 1, \ldots, n$.
(a) Find the distribution of Z such that $n(\theta - X_{(n)}) \xrightarrow{d} Z$.
(b) Find $c > 0$ such that $\mathbb{E}(T_c) = \theta$ where $T_c = cX_{(n)}$.

Solution. (a) For $x > 0$, $P(n(\theta - X_{(n)}) > x) = P(X_{(n)} < \theta - x/n) = [P(X_1 < \theta - x/n)]^n = [(\theta - x/n)/\theta]^n = [1 - x/(n\theta)]^n \to e^{-x/\theta}$. Hence $Z \sim$ Exponential$(1/\theta)$.
(b) From $\theta = \mathbb{E}[T_c] = c\mathbb{E}[X_{(n)}]$, the constant $c = \theta/\mathbb{E}[X_{(n)}]$. Note that $P(X_{(n)} > x) = 1 - [P(X_1 \leq x)]^n = 1(0 < x < \theta)(1 - (1 - x/\theta)^n)$. Hence $\mathbb{E}[X_{(n)}] = \int_0^\infty P(X_{(n)} > x)\, dx = \int_0^\theta 1 - (1 - x/\theta)^n\, dx = \theta - \theta/(n+1)$. Finally $c = \theta/[\theta(1 - 1/(n+1))] = 1 + 1/n$.

Problem 22. Suppose $\mathbb{E}(|X|) < \infty$, F is the distribution function of X and $f(x) = F'(x)$ is the derivative of F. Prove that

$$\int_0^\infty (1 - F(x))dx - \int_{-\infty}^0 F(x)dx = \mathbb{E}(X).$$

Solution. See Problem 10.

Problem 23. Suppose X_i's are i.i.d. random variables having $\mathbb{E}(X_i) = \mu$ and $\mathbb{E}(X_i^2) = \mu_2$ for $i = 1, \ldots, n$. Let $\overline{X} = (X_1 + \cdots + X_n)/n$ and $S^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \overline{X})^2$. Compute the mean of $\overline{X}$ and S^2.

Solution. See Problem 13.

Problem 24. Suppose φ_X and φ_Y are the characteristic functions of X and Y, respectively.
(a) Prove that X is symmetric if and only if φ_X is real-valued.
(b) Find the characteristic function of $aX + b$ using φ_X.
(c) Prove $|\varphi_X|^2$ is also a characteristic function.
(d) Prove $(\varphi_X + \varphi_Y)/2$ is also a characteristic function.

Solution. (a) If X is symmetric $\text{chf}_X(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX) + i\sin(tX)] = \mathbb{E}[\cos(t|X|)]$ is real because $\cos(tx)$ is even, $\sin(tx)$ is odd and X is symmetric. From inversion formula, $P(0 < X < a) + (P(X = 0) + P(X = a))/2 = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^T \frac{1-e^{-iat}}{it}\varphi(t)\ dt$ and $P(-a < X < 0) + (P(X = 0) + P(X = -a))/2 = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^T \frac{e^{iat}-1}{it}\varphi(t)\ dt = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^T \frac{1-e^{-iat}}{it}\varphi(-t)\ dt$ Hence X is symmetric if and only if $\varphi(-t) = \varphi(t)$.
(b) $\text{chf}_{aX+b}(t) = \mathbb{E}[e^{it(aX+b)}] = e^{itb}\mathbb{E}[e^{i(at)X}] = e^{itb}\text{chf}_X(at)$.
(c) Let X_1, X_2 be i.i.d. copy of X. Then $\text{chf}_{X_1-X_2}(t) = \mathbb{E}[e^{it(X_1-X_2)}] = \mathbb{E}[e^{itX_1}]\mathbb{E}[e^{-itX_2}] = \text{chf}_X(t)\overline{\text{chf}_X(t)} = |\varphi_X(t)|^2$.
(d) Let $W \sim \text{Bernoulli}(1/2)$. Let Z be a random variable the same to X if $W = 0$ and to Y if $W = 1$. Then $Z = X(1 - W) + YW$ and its characteristic function is $\text{chf}_Z(t) = \mathbb{E}[e^{itZ}] = \mathbb{E}[e^{it(X(1-W)+YW)}] = \mathbb{E}[e^{itX} \,|\, W = 0]P(W = 0) + \mathbb{E}[e^{itY} \,|\, W = 1]P(W = 1) = \varphi_X(t)/2 + \varphi_Y(t)/2$.

Problem 25. (a) Find the density function of X_1/X_2 when $X_i \sim i.i.d.\, N(0, \sigma^2)$ for $i = 1, 2$.
(b) Find the characteristic function of $X \sim \text{Cauchy}(0, \sigma^2)$ having density $(\sigma/\pi)/(x^2 + \sigma^2)$.
(c) Suppose $X_i \sim i.i.d.\, \text{Cauchy}(0, \sigma^2)$ for $i = 1, \ldots, n$. Find the distribution of $\overline{X}$.

Solution. (a) Consider a map $(x, y) \mapsto (u, v) = (x/y, y)$. Then $(x, y) = (uv, v)$ and the determinant is $\begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = |v|$. Hence $\text{pdf}_{U,V}(u, v) = (2\pi)^{-1}\exp(-(uv)^2/2 - v^2/2)|v|$ and $\text{pdf}_U(u) = \int_{-\infty}^\infty (2\pi)^{-1}|v|\exp(-v^2(u^2+1)/2)\ dv = \pi^{-1}\int_0^\infty \sqrt{\frac{2z}{u^2+1}}\exp(-z)(2(u^2+1)z)^{-1/2}\ dz$(by taking $z =$
$[\pi(u^2+1)]^{-1}\int_0^\infty \exp(-z)\ dz = 1/[\pi(u^2+1)]$ which is the density of $\text{Cauchy}(0, 1)$ distribution.
(b) Let $Z \sim \text{Cauchy}(0, \sigma)$. Then $\mathbb{E}[e^{itZ}] = \int_{-\infty}^\infty e^{itx} \cdot \frac{\sigma}{\pi(x^2+\sigma^2)}\ dx = e^{-|t|/\sigma}$ by complex analysis.
(c) Since $\text{chf}_{\bar{X}}(t) = \mathbb{E}[e^{it\bar{X}}] = [\text{chf}_{X_1}(t/n)]^n = [e^{-|t/n|/\sigma}]^n = e^{-|t|/\sigma}$, $\bar{X} \sim \text{Cauchy}(0, \sigma^2)$.

Problem 26. Two random variables X and Y are independent. If $X \sim \text{Poisson}(\lambda)$ and $X+Y \sim \text{Poisson}(\lambda+\mu)$, then find the distribution of Y.

Solution. Solution I.(Hard solution) Since both X and $X+Y$ take values among nonnegative integers, Y should take values only on integers. Suppose $P(Y=-n)>0$ for a positive integer $n>0$. Then $P(X+Y=-n) \geq P(X=0, Y=-n) = P(X=0)P(Y=-n) > 0$. But $P(X+Y=-n)=0$ because $X+Y \sim \text{Poisson}(\mu+\lambda)$. Hence Y takes values at most on nonnegative integers. Note that

$$e^{-(\mu+\lambda)}(\mu+\lambda)^n/n! = P(X+Y=n) = \sum_{k=0}^{\infty} P(X=n, Y=n-k) = \sum_{k=0}^{n} e^{-\mu}\mu^k/k! \times \text{pmf}_Y(n-k)$$

Hence

$$\sum_{k=0}^{n} P(Y=k)\mu^{n-k}/(n-k)! = e^{-\lambda}(\mu+\lambda)^n/n! = e^{-\lambda}\sum_{k=0}^{n}\binom{n}{k}\lambda^k\mu^{n-k}/n! = \sum_{k=0}^{n}[e^{-\lambda}\lambda^k/k!]\mu^{n-k}/(n-k)!.$$

If $n=0$, then $P(Y=0) = e^{-\lambda}\lambda^0/0! \cdot \mu^0/0!$ which implies $P(Y=0)=e^{-\lambda}$. If $n=1$, $P(Y=1)+P(Y=0)\mu/1! = e^{-\lambda}\lambda/1! + e^{-\lambda}\cdot\mu^1/1!$ implies $P(Y=1)=e^{-\lambda}\lambda^1/1!$. Using mathematical induction, $P(Y=k)=e^{-\lambda}\lambda^k/k! \sim \text{Poisson}(\lambda)$.

Solution II.(Easy solution) The moment generation function of $X \sim \text{Poisson}(\mu)$ is $\text{mgf}_X(t) = \mathbb{E}[e^{tX}] = \sum_{n=0}^{\infty} e^{tn} \cdot e^{-\mu}\mu^n/n! = e^{-\mu}\exp(\mu e^t) = \exp(-\mu(1-e^t))$. Then the moment generation function of $X+Y$ is

$$\text{mgf}_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}e^{tY}] = \mathbb{E}(e^{tX}]\mathbb{E}(e^{tY}) = \text{mgf}_X(t)\text{mgf}_Y(t)$$

In the second equality, the independence of X and Y is used. Hence $\text{mgf}_Y(t) = \text{mgf}_{X+Y}(t)/\text{mgf}_X(t) = \exp(-(\mu+\lambda)(1-e^t))/\exp(-\mu(1-e^t)) = \exp(-\lambda(1-e^t))$ which is the moment generation function of $\text{Poisson}(\lambda)$. Hence $Y \sim \text{Poisson}(\lambda)$.

Problem 27. Suppose $X_n \sim \text{Binomial}(n, \lambda/n)$. Prove that $X_n \xrightarrow{d} \text{Poisson}(\lambda)$.

Solution. For a fixed nonnegative integer k, $P(X_n=k) = \binom{n}{k}(\lambda/n)^k(1-\lambda/n)^{n-k} = (\lambda^k/k!)1(1-1/n)\cdots(1-(k-1)/n)(1-\lambda/n)^{n-k} \approx (\lambda^k/k!)\exp(-\lambda/n\times(n-k)) \to e^{-\lambda}\lambda^k/k! = P(Z=k)$ where $Z \sim \text{Poisson}(\lambda)$.

Problem 28. Prove the following memoryless properties.
(a) If $X \sim \text{Exponential}(\lambda)$, then $P(X>a+b \mid X>a) = P(X>b)$ for all positive real numbers a and b.
(b) If $X \sim \text{Geometric}(\theta)$, then $P(X>a+b \mid X>a) = P(X>b)$ for all positive integers a and b.

Solution. (a) $P(X>x) = \int_x^{\infty} \lambda e^{-\lambda z}\, dz = [-e^{-\lambda}]_x^{\infty} = e^{-\lambda x}$. Hence $P(X>a+b \mid X>a) = P(X>a+b, X>a)/P(X>a) = P(X>a+b)/P(X>a) = e^{-\lambda(a+b)}/e^{-\lambda a} = e^{-\lambda b} = P(X>b)$ which is called memoryless property.
(b) $P(X>a) = \sum_{n=a+1}^{\infty} p(1-p)^{n-1} = p(1-p)^a(1+(1-p)+(1-p)^2+\cdots) = p(1-p)^a \times 1/(1-(1-p)) = (1-p)^a$. Thus $P(X>a+b \mid X>a) = P(X>a+b)/P(X>a) = (1-p)^{a+b}/(1-p)^a = (1-p)^b = P(X>b)$.

Problem 29. Suppose $X \,|\, Y = y \sim N(y, \sigma^2)$ and $Y \sim N(0, \tau^2)$.
(a) What is the marginal distribution of X?
(b) What is the conditional distribution of Y given $X = x$?

Solution. (a) $\mathrm{mgf}_X(t) = \mathbb{E}[\mathbb{E}[e^{tX} \,|\, Y]] = \mathbb{E}[e^{tY + t^2\sigma^2/2}] = e^{t^2\tau^2/2 + t^2\sigma^2/2} = e^{t^2(\sigma^2+\tau^2)/2}$ shows $X \sim N(0, \sigma^2 + \tau^2)$.
(b) $\mathrm{pdf}_{Y\,|\,X}(y \,|\, x) = \mathrm{pdf}_{X,Y}(x, y)/\mathrm{pdf}_X(x) = \mathrm{pdf}_{X\,|\,Y}(x \,|\, y)\mathrm{pdf}_Y(y)/\mathrm{pdf}_X(x) = (2\pi\sigma^2)^{-1/2}\exp(-(x-y)^2/(2\sigma^2))\cdot(2\pi\tau^2)^{-1/2}\exp(-y^2/(2\tau^2))/[(2\pi(\sigma^2+\tau^2))^{-1/2}\exp(-x^2/[2(\sigma^2+\tau^2)])] = (2\pi\sigma^2\tau^2/(\sigma^2+\tau^2))^{-1/2}\exp(-(y-x)^2/[2\sigma^2\tau^2/(\sigma^2+\tau^2)])$ indicates $Y \,|\, X = x \sim N(x, \sigma^2\tau^2/(\sigma^2+\tau^2))$.

Problem 30. Assume $X_i \sim i.i.d.\, N(0, \sigma^2)$ for $i = 1, 2, \ldots, n$.
(a) Show that $X_i^2/\sigma^2 \sim i.i.d.\, \chi^2(1) \sim \mathrm{Gamma}(1/2, 1/2)$.
(b) Find the kurtosis of X_i, i.e., $\mathbb{E}[(X - \mathbb{E}(X))^4]$.

Solution. Let $X \sim N(0, 1)$. Then $X_i/\sigma \sim i.i.d.\; N(0, 1)$.
(a) $\mathrm{mgf}_{X_i^2/\sigma^2}(t) = \mathbb{E}[e^{tX_i^2/\sigma^2}] = \mathbb{E}[e^{tX^2}] = \int_{-\infty}^{\infty} e^{tx^2}\cdot(2\pi)^{-1/2}\exp(-x^2/2)\, dx = (2/\pi)^{1/2}\int_0^\infty e^{-x^2(1/2-t)}\, dx = (2/\pi)^{1/2}\int_0^\infty e^{-z}\cdot(1/2)(1/2-t)^{-1/2}z^{-1/2}\; dz = (1-2t)^{-1/2}\pi^{-1/2}\int_0^\infty z^{1/2-1}e^{-z}\; dz = (1-2t)^{-1/2}\pi^{-1/2}\Gamma(1/2) = (1-2t)^{-1/2}$ which is the mgf of $\chi^2(10$ or $\mathrm{Gamma}(1/2, 1/2)$.
(b) Note that $\mathbb{E}[(X_i - \mathbb{E}(X_i))^4] = \mathbb{E}[X_i^4] = \mathbb{E}[(\sigma X)^4] = \sigma^4\mathbb{E}[X^4] = \sigma^4\frac{d^4}{dt^4}\mathrm{mgf}_X(0) = \sigma^4\frac{d^4}{dt^4}e^{-t^2/2}|_{t=0} = \sigma^4\frac{d^4}{dt^4}\sum_{k=0}^\infty(-1/2)^k t^{2k}/k!|_{t=0} = \sigma^4(-1/2)^2(4!)/2! = 3\sigma^4$.

Problem 31. Let $X_1, \ldots, X_n$ be a random sample from a distribution having a density

$$f_\theta(x) = c(\theta)x^2 I(0 \le x \le \theta).$$

(a) Compute $c(\theta)$.
(b) Show that $X_{(n)} = \max(X_1, \ldots, X_n)$ converges to θ in probability.
(c) Prove or disprove that $X_{(n)} = \max(X_1, \ldots, X_n)$ converges to θ almost surely.

Solution. (a) $1 = \int_0^\theta f_\theta(x)\, dx = c(\theta)\int_0^\theta x^2\, dx = c(\theta)\theta^3/3$ implies $c(\theta) = 3/\theta^3$.
(b) For $x \in (0, \theta)$, $P(X_{(n)} \le x) = P(X_1 \le x, \ldots, X_n \le x) = [P(X_1 \le x)]^n = [(3/\theta^3)\int_0^x z^2\, dz]^n = (x/\theta)^n$. Hence for any $\epsilon \in (0, \theta)$, $P(|X_{(n)} - \theta| > \epsilon) = P(X_{(n)} < \theta - \epsilon) = [(\theta - \epsilon)/\theta]^n = [1 - \epsilon/\theta]^n \to 0$. Thus $X_{(n)} \to \theta$ in probability.
(c) From part (b), $P(|X_{(n)} - \theta| > \epsilon) = (1 - \epsilon/\theta)^n$. Then $\sum_{n=1}^\infty P(|X_{(n)} - \theta| > \epsilon) = \sum_{n=1}^\infty(1 - \epsilon/\theta)^n = (1 - \epsilon/\theta)/[1 - (1 - \epsilon/\theta)] = \theta/\epsilon - 1 < \infty$ for any $0 < \epsilon < \theta$. Therefore $X_{(n)} \to \theta$ almost surely.

Problem 32. Assume $X_1, \ldots, X_n$ are i.i.d. random variables from $\mathrm{Poisson}(\theta)$.
(a) Find the moment generating function of X_i.
(b) Show that $T = X_1 + \cdots + X_n$ is also a Poisson distribution.
(c) Assume $X \sim \mathrm{Poisson}(\theta)$ and $Y \sim \mathrm{Poisson}(\mu)$ are independent. Show that the conditional distribution of X given $X + Y = t$ is $\mathrm{Binomial}(t, \theta/(\theta + \mu))$.

Solution. (a) $\text{mgf}_X(t) = \mathbb{E}[e^{tX_i}] = \sum_{n=0}^{\infty} e^{tn} e^{-\mu}\mu^n/n! = e^{-\mu+\mu e^t} = \exp(\mu(e^t - 1))$.
(b) If $X \sim \text{Poisson}(\mu)$ and $Y \sim \text{Poisson}(\lambda)$ are independent, probability can be obtained from

$$P(X+Y=n) = \sum_{x=0}^{n} P(X=x, Y=n-x) = \sum_{x=0}^{n} e^{-\mu}\frac{\mu^x}{x!} \times e^{-\lambda}\frac{\lambda^{n-x}}{(n-x)!}$$
$$= \frac{\exp(-(\mu+\lambda))}{n!}\sum_{x=0}^{n}\binom{n}{x}\mu^x\lambda^{n-x} = \frac{\exp(-(\mu+\lambda))}{n!}(\mu+\lambda)^n \sim \text{Poisson}(\mu+\lambda).$$

Or $\text{mgf}_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = \text{mgf}_X(t)\text{mgf}_Y(t) = \exp(\mu(e^t-1)) \cdot \exp(\lambda(e^t - 1)) = \exp((\mu+\lambda)(e^t-1)) \sim \text{Poisson}(\mu+\lambda)$. Hence $T = X_1 + \cdots + X_n \sim \text{Poisson}(n\theta)$.
(c) Note that $P(X=k \,|\, X+Y=n) = P(X=k, X+Y=n)/P(X+Y=n) = P(X=k, Y=n-k)/P(X+Y=n) = P(X=k)P(Y=n-k)/P(X+Y=n)$ which implies

$$P(X=k \,|\, X+Y=n) = \frac{e^{-\mu}\mu^k/k! \times e^{-\lambda}\lambda^{n-k}/(n-k)!}{e^{-\mu-\lambda}(\mu+\lambda)^n} = \binom{n}{k}\Big(\frac{\mu}{\mu+\lambda}\Big)^k\Big(\frac{\lambda}{\mu+\lambda}\Big)^{n-k} \sim \text{Binomial}(n, \frac{\mu}{\mu+\lambda$$

Problem 33. Determine whether the following statements are True or False.

(a) Assume $X \sim \text{Poisson}(\lambda), Y \sim \text{Poisson}(\mu)$. If X and Y are independent, then $X+Y \sim \text{Poisson}(\lambda+\mu)$.

(b) $T = \theta + (X_1 - X_2)/2$ is a statistic.

(c) The maximum likelihood estimate and the method of moments estimate are always the same.

(d) $\{N(\mu, \sigma^2)$ or $\text{Poisson}(\mu)\}$ is not a model.

(e) If $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} 1$, then $X_n/Y_n \xrightarrow{p} X$.

Solution. (a) T, (b) F, (c) F, (d) F, (e) T

Problem 34. Assume $X \sim \text{Gamma}(\alpha, \beta)$ having density $(\Gamma(\alpha)\beta^\alpha)^{-1}x^{\alpha-1}\exp(-x/\beta)I(x \geq 0)$ where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}\exp(-x)dx$.
(a) Prove $\mathbb{E}(X^r) = \beta^r\Gamma(\alpha+r)/\Gamma(\alpha)$ for $r > -\alpha$.
(b) Compute mean and variance of X.

Solution. See Example 44 in lecture note.

Problem 35. The conditional expectation $\mathbb{E}(X \,|\, Y)$ of X given Y is the random variable $Z = Z(Y)$ such that $\mathbb{E}[(X - Z(Y))g(Y)] = 0$ for all bounded function g.
(a) Prove that $\mathbb{E}(\mathbb{E}(X \,|\, Y)) = \mathbb{E}(X)$.
(b) Show that $\mathbb{V}\text{ar}(X) = \mathbb{V}\text{ar}(\mathbb{E}(X \,|\, Y)) + \mathbb{E}(\mathbb{V}\text{ar}(X \,|\, Y))$ where $\mathbb{V}\text{ar}(X \,|\, Y) = \mathbb{E}(X^2 \,|\, Y) - [\mathbb{E}(X \,|\, Y)]^2$.

Solution. (a) Let $Z = Z(Y) = \mathbb{E}(X \mid Y)$. Take $g(Y) = 1$ the constant function 1. Then $0 = \mathbb{E}[(X - Z(Y))g(Y)] = \mathbb{E}[X - Z(Y)] = \mathbb{E}(X) - \mathbb{E}[Z(Y)] = \mathbb{E}(X) - \mathbb{E}[\mathbb{E}(X \mid Y)]$ implies $\mathbb{E}[\mathbb{E}(X \mid Y)] = \mathbb{E}(X)$.
(b) Let $Z = \mathbb{E}(X \mid Y)$. Then $\mathbb{V}\mathrm{ar}(\mathbb{E}(X \mid Y)) = \mathbb{E}(Z^2) - (\mathbb{E}(Z))^2 = \mathbb{E}(Z^2) - (\mathbb{E}(X))^2$ and $\mathbb{E}[\mathbb{V}\mathrm{ar}(X \mid Y)] = \mathbb{E}[\mathbb{E}(X^2 \mid Y) - (\mathbb{E}(X \mid Y))^2] = \mathbb{E}[\mathbb{E}(X^2 \mid Y)] - \mathbb{E}(Z^2) = \mathbb{E}(X^2) - \mathbb{E}(Z^2)$. Hence $\mathbb{V}\mathrm{ar}(\mathbb{E}(X \mid, Y)) + \mathbb{E}(\mathbb{V}\mathrm{ar}(X \mid Y)) = \mathbb{E}(Z^2) - (\mathbb{E}(X))^2 + \mathbb{E}(X^2) - \mathbb{E}(Z^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \mathbb{V}\mathrm{ar}(X)$.

Problem 36. Consider a probability density function

$$f(x, y \mid \theta) = \begin{cases} c(\theta)x^2 y & \text{if } 0 \le x \le y \le \theta \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $c(\theta)$.
(b) Prove or disprove that X and Y are independent.

Solution. (a) $1 = \int\int f(x, y \mid \theta)\ dx\ dy = \int_0^\theta \int_0^y c(\theta)x^2 y\ dx\ dy = \int_0^\theta c(\theta)(y^3/3)y\ dy = c(\theta)y^5/15\_0^\theta = c(\theta)\theta^5/15$. Gives $c(\theta) = 15/\theta^5$.

(b) The marginal densities of Y and X are $\mathrm{pdf}_Y(y) = \int_0^y c(\theta)x^2 y\ dx = c(\theta)y^4/3 = 5y^4/\theta^5$ for $0 \le y \le \theta$ and $\mathrm{pdf}_X(x) = \int_x^\theta c(\theta)x^2 y\ dy = c(\theta)x^2 y^2/2|_x^\theta = (15/2\theta^5)x^2(\theta^2 - x^2)$ for $0 \le x \le \theta$. Hence $\mathrm{pdf}_{X,Y}(x, y) = c(\theta)x^2 y 1(0 \le x \le y) \ne (15/2\theta^5)x^2(\theta^2 - x^2)c(\theta)y^4/3 = \mathrm{pdf}_X(x)\mathrm{pdf}_Y(y)$ implies X and Y are not independent.

Problem 37. Assume that $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$. Using Slutsky's theorem, prove the followings.
(a) $X_n + Y_n \xrightarrow{p} X + Y$.
(b) $X_n Y_n \xrightarrow{p} XY$.

Solution. Note that $X_n - X \xrightarrow{d} 0$ and $Y_n - Y \xrightarrow{d} 0$. Hence $X_n + Y_n - (X + Y) = (X_n - X) + (Y_n - Y) \xrightarrow{d} 0$ implies $X_n + Y_n - (X + Y) \xrightarrow{p} 0$ as well as $X_n + Y_n \xrightarrow{p} X + Y$. Similarly, $X_n Y_n - XY = X_n(Y_n - Y) + (X_n - X)Y \xrightarrow{d} X \cdot 0 + 0 \cdot Y = 0$. Hence $X_n Y_n - XY \xrightarrow{p} 0$ and $X_n Y_n \xrightarrow{p} XY$.

Problem 38. Find a continuous function f and a sequence $X_n \to X$ in L^p but $f(X_n) \not\to f(X)$.

Solution. Make $X_n \to X$ in L^p but not in L^q for all $q > p$. Take $f(x) = |x|^{1+\delta}$ for any $\delta > 0$.

Problem 39. Monte Carlo integration Let f be a measurable function on $[0, 1]$ with $\int_0^1 |f(x)|^2\ dx < \infty$. Let $U_1, U_2, \ldots$ be i.i.d Uniform$[0, 1]$, and $I_n = (f(U_1) + \cdots + f(U_n))/n$. Show that $I_n \to I = \int_0^1 f(x)\ dx$ in probability and compute a convergence rate $P(|I_n - I| > \epsilon/n^{1/2})$ using Chebyshev's inequality.

Solution. Using the law of large numbers $I_n \to I$ almost surely. Besides the central limit theorem implies $\sqrt{n}(I_n - I) \to N(0, \tau^2)$ where $\tau^2 = \mathbb{V}\text{ar}(f(U_1))$. In other words, $I_n = I + O_p(n^{-1/2})$.

Problem 40. Let X_n is an AR (autoregressive) process satisfying $X_0 = \mu$ and $X_n = (1-\rho)\mu + \rho X_{n-1} + \epsilon_n$ where $|\rho| < 1$, $\epsilon_n \sim i.i.d.N(0, \sigma^2)$. Prove that $\overline{X}_n = (X_1 + \cdots + X_n)/n \to \mu$ in probability.

Solution. Note that $\mathbb{E}(X_n) = (1-\rho)\mu + \rho\mathbb{E}(X_{n-1}) = (1-\rho)\mu(1+\rho) + \rho^2\mathbb{E}(X_{n-2}) = \mu(1-\rho^2) + \rho^2\mathbb{E}(X_{n-2}) = \mu(1-\rho^n) + \rho^n\mathbb{E}(X_0) = \mu$, $\mathbb{V}\text{ar}(X_n) = \rho^2\mathbb{V}\text{ar}(X_{n-1}) + \sigma_\epsilon^2 = \sigma_\epsilon^2(1+\rho^2+\cdots+\rho^{2(n-1)}) + \rho^{2n}\mathbb{V}\text{ar}(X_0) = \sigma_\epsilon^2(1-\rho^{2n})/(1-\rho^2) \le \sigma_\epsilon^2/(1-\rho^2)$. Let $\tilde{X}_n = X_n - \mu$ so that $\tilde{X}_n = \rho\tilde{X}_{n-1} + \epsilon_n$. Then, for $i \le j$, $\text{Cov}(X_i, X_j) = \text{Cov}(\tilde{X}_i, \tilde{X}_j) = \mathbb{E}[\tilde{X}_i(\epsilon_j + \rho\epsilon_{j-1} + \cdots + \rho^{j-i+1}\epsilon_{i+1} + \rho^{j-i}\tilde{X}_i)] = \rho^{j-i}\mathbb{V}\text{ar}(\tilde{X}_i)$ and $\mathbb{V}\text{ar}(\bar{X}_n) = n^{-2}\sum_{i,j}\text{Cov}(X_i, X_j) = n^{-2}\sum_{i=1}^n \mathbb{V}\text{ar}(X_i) + 2n^{-2}\sum_{i<j}\rho^{2(j-i)}\mathbb{V}\text{ar}(X_i) \le n^{-2}\sum_{i=1}^n \sigma_\epsilon^2/(1-\rho^2) + 2n^{-2}\sum_{i<j}\rho^{2(j-i)}\sigma_\epsilon^2/(1-\rho^2) \le n^{-1}\sigma_\epsilon^2/(1-\rho^2)[1+\sum_{j=1}^n \rho^{2j}] \le n^{-1}\sigma_\epsilon^2/(1-\rho^2)[1/(1-\rho^2)]$. Using Chebychev's inequality

$$\mathbb{E}[|\bar{X}_n - \mu|^2] = \mathbb{V}\text{ar}(\bar{X}_n) \le n^{-1}\sigma_\epsilon^2/(1-\rho^2)^2 \to 0.$$

Hence $\bar{X}_n \to \mu$ in L^2 as well as in probability.

Problem 41. Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}|X_1| < \infty$. Show that $\max(X_1, \ldots, X_n)/n \to 0$ in probability.

Solution. Let $M_n = \max(X_1, \ldots, X_n)$. Fix $\epsilon > 0$, $P(M_n/n > \epsilon) = P(M_n > \epsilon n) = P(X_1 > \epsilon n \text{ or } \cdots \text{ or } X_n > \epsilon n) \le nP(X_1 > \epsilon n) = n\mathbb{E}[1(X_1 > \epsilon n)] = \epsilon^{-1}\mathbb{E}[\epsilon n 1(X_1 > \epsilon n)] \le \epsilon^{-1}\mathbb{E}[|X_1| 1(X_1 > \epsilon n)] \to 0$ by dominated convergence theorem. Note that $|X_1| 1(X_1 > \epsilon n) \to 0$ almost surely and bounded by $|X_1|$.

Problem 42. Prove that $X_n \to X$ in probability if and only if there exist $\epsilon_n \searrow 0$ such that $P(|X_n - X| > \epsilon_n) \le \epsilon_n$. Compare $X_n \to X$ a.s. if $\sum_{n=0}^\infty P(|X_n - X| > \epsilon_n) < \infty$ for a sequence $\epsilon_n \searrow 0$.

Solution. Sufficiency ($\Longrightarrow$). Let $n_0 = 0$ and take $n_k > n_{k-1}$ so that $P(|X_n - X| > 1/k) < 1/k$ for all $n \ge n_k$. Then define $\epsilon_n = 1$ if $n < n_2$ and $\epsilon_n = 1/k$ if $n_k \le n < n_{k+1}$ for some $k > 2$. Then ϵ_n is nonincreasing and $P(|X_n - X| > \epsilon_n) \le \epsilon_n$ for any n.
Necessity ($\Longleftarrow$). For any $\epsilon > 0$, there exists $N > 0$ such that $\epsilon_n \le \epsilon$ for all $n \ge N$. Then $P(|X_n - X| > \epsilon) < \epsilon$ for all $n \ge N$. Hence $X_n \xrightarrow{p} X$.
In general $\sum_n \epsilon_n < \infty$ is not guaranteed, that is, the convergence in probability may not imply almost sure convergence.

Problem 43. Suppose $X \ge 0$ and $\mathbb{E}(X^2) < \infty$. Prove that $P(X > 0) \ge (\mathbb{E}(X))^2/\mathbb{E}(X^2)$.

Solution. Let $Y = 1(X > 0$ and $Z = X$. Then $(\mathbb{E}(X))^2 = |\mathbb{E}(YZ)|^2 \le \mathbb{E}(Y^2)\mathbb{E}(Z^2) = \mathbb{E}(1(X > 0))\mathbb{E}(X^2)$ (by Cauchy-Schwartz' inequality) $= P(X > 0)\mathbb{E}(X^2)$. Hence $P(X > 0) \ge (\mathbb{E}(X))^2/\mathbb{E}(X^2)$.

Problem 44. Let $X_1, X_2, \ldots$ be independent random variables with $\mathbb{E}(X_n) = 0$ and $\mathbb{V}\text{ar}(X_n)/n \to 0$ as $n \to \infty$. Show that $\overline{X}_n = (X_1 + \cdots + X_n)/n \to 0$ in L^2.

Solution. We use the following claim.
Claim: If $x_n \to x$, then $\bar{x}_n = (x_1 + \cdots + x_n)/n \to x$.
For any $\epsilon > 0$, there exists $N > 0$ such that $|x_n - x| < \epsilon/3$ for all $n \geq N$. Take $M > N(1 + |x|)3/\epsilon$ so that $|x_1 + \cdots + x_N| < M\epsilon/3$ and $N|x|/M < \epsilon/3$. Then for any $n \geq M$, $|x_1 + \cdots + x_n - nx| \leq |x_1 + \cdots + x_N| + |x_{N+1} - x| + \cdots + |x_n - x| + N|x| \leq M\epsilon/3 + (n - N)\epsilon/3 + M\epsilon/3 \leq n\epsilon$. Hence $|\bar{x}_n - x| < \epsilon$ for all $n \geq M$. By taking $\epsilon > 0$ arbitrarily small, $\lim_{n\to\infty} \bar{x}_n = x$.

Note that $\mathbb{E}(\bar{X}_n) = 0$ and $\mathbb{V}\text{ar}(\bar{X}_n) = n^{-2}(\mathbb{V}\text{ar}(X_1) + \cdots + \mathbb{V}\text{ar}(X_n)) \leq n^{-1}(\mathbb{V}\text{ar}(X_1)/1 + \cdots + \mathbb{V}\text{ar}(X_n)/n)$. Then the second moment of $\bar{X}_n$ is

$$\mathbb{E}[(\bar{X}_n - 0)^2] \leq \mathbb{V}\text{ar}(\bar{X}_n) \leq n^{-1}(\mathbb{V}\text{ar}(X_1)/1 + \cdots + \mathbb{V}\text{ar}(X_n)/n) \to \lim_{n\to\infty} \mathbb{V}\text{ar}(X_n)/n = 0.$$

Therefore $\bar{X}_n \to 0$ in probability.

Problem 45. A sequence of random variables X_n is *uniformly integrable* if $\lim_{\alpha\to\infty} \sup_n \mathbb{E}(|X_n|1(|X_n| \geq \alpha)) = 0$. Suppose $X_n \to X$ almost surely. Show the following conditions are equivalent:
(a) X_n are uniformly integrable,
(b) $\mathbb{E}(|X_n - X|) \to 0$,
(c) $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|)$.

Solution. (a) $\Longrightarrow$ (b). There exists $\alpha > 0$ such that $\mathbb{E}[|X_n|1(|X_n| \geq \alpha)] \leq 1$ for all n. Note that $\mathbb{E}[|X_n|] = \mathbb{E}[|X_n|1(|X_n| < \alpha)] + \mathbb{E}[|X_n|1(|X_n| \geq \alpha)] \leq \alpha + 1$. Then using Fatou's lemma, $\mathbb{E}[|X|] = \mathbb{E}[\liminf_{n\to\infty} |X_n|] \leq \liminf_{n\to\infty} \mathbb{E}[|X_n|] \leq \liminf_{n\to\infty}(\alpha + 1) = \alpha + 1 < \infty$. Let $h_m(x) = \min(m, |x|)$ is bounded and continuous. Since $X_n - X \xrightarrow{p} 0$ implies $X_n - X \xrightarrow{d} 0$, the expectations $\mathbb{E}[h_m(|X_n - X|)] \to \mathbb{E}[h_m(|X - X|)] = 0$ as $n \to \infty$. Then

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}[|X_n - X|] = \mathbb{E}[\lim_{m\to\infty} h_m(|X_n - X|)] = \lim_{n\to\infty} \lim_{m\to\infty} \mathbb{E}[h_m(|X_n - X|)] \quad \text{(by MCT)} \\
&= \lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}[h_m(|X_n - X|)] \quad \text{(because all expectations are nonnegative)} \\
&= \lim_{m\to\infty} \mathbb{E}[h_m(|X - X|)] \quad \text{(by convergence in distribution)} = 0.
\end{aligned}$$

(b) $\Longrightarrow$ (c). If $\mathbb{E}(|X|) = \infty$, then the theorem is trivial. Suppose $\mathbb{E}(|X|) < \infty$. Note that $|X_n| \leq |X_n - X| + |X|$ with $\mathbb{E}[|X_n - X| + |X|] \to \mathbb{E}[|X|]$. By the generalized dominated convergence theorem, $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|)$.
(c) $\Longrightarrow$ (a). Define for $\alpha > 0$, $g_\alpha(x) = |x|1(|x| < \alpha - 1) + (\alpha - |x|)1(\alpha - 1 \leq |x| < \alpha)$ is a continuous function satisfying $g_\alpha(x) \leq |x|1(|x| < \alpha)$. Then $|x|1(|x| \geq \alpha) = |x| - |x|1(|x| < \alpha) \leq |x| - g_\alpha(|x|)$ and $\mathbb{E}[|X_n|1(|X_n| \geq \alpha)] \leq \mathbb{E}(|X_n|) - \mathbb{E}(g_\alpha(|X_n|)) \to \mathbb{E}(|X|) - \mathbb{E}(g_\alpha(|X|)) \leq \mathbb{E}(|X|1(|X| \geq \alpha - 1)$.

For any $\epsilon > 0$. Take $\alpha_0 > 1$ so that $\mathbb{E}(|X|1(|X| > \alpha_0 - 1)) < \epsilon/3$. Then there exists $N > 0$ such that $|\mathbb{E}[g_{\alpha_0}(|X_n|)] - \mathbb{E}[g_{\alpha_0}(|X|)]| < \epsilon/3$ and $|\mathbb{E}(|X_n|) - \mathbb{E}(|X|)| < \epsilon/3$ for all $n \geq N$ so that $\mathbb{E}(|X_n|1(|X_n| > \alpha_0)) \leq \mathbb{E}(|X_n|) - \mathbb{E}(g_{\alpha_0}(|X_n|)) \leq \mathbb{E}(|X|) + \epsilon/3 - \mathbb{E}(g_{\alpha_0}(|X|)) + \epsilon/3 \leq \epsilon$.

Also there exist $\alpha_i > 0$ such that $\mathbb{E}(|X_i| 1(|X_i| > \alpha_i)) < \epsilon$ for $i = 1, \ldots, N-1$. Let $\alpha = \max(1 + \alpha_0, \ldots, \alpha_{N-1})$. Then $\mathbb{E}(|X_n| 1(|X_n| > \alpha)) < \epsilon$ for all $n = 1, \ldots, N$ and $n = N+1, N+2, \ldots$. Hence $\lim_{\alpha\to\infty} \sup_n \mathbb{E}(|X_n| 1(|X_n| > \alpha) < \epsilon$ for any $\epsilon > 0$. Therefore the result holds.

Problem 46. Let X_n be a martingale with $\mathbb{E}(X_1) = 0$ and $\sum_{n=1}^{\infty} \mathbb{E}[(X_n - X_{n-1})^2] < \infty$. Show that X_n converges almost surely. [Hint: X_n^2 is a submatingale and $\mathbb{E}(X_n^2) = \mathbb{E}(X_0^2) + \sum_{k=1}^{n} \mathbb{E}[(X_k - X_{k-1})^2]$.]

Solution. Note that $\mathbb{E}(X_n^2) = \mathbb{E}[(X_{n-1} + X_n - X_{n-1})^2] = \mathbb{E}[X_{n-1}^2 + 2X_{n-1}(X_n - X_{n-1}) + (X_n - X_{n-1})^2] = \mathbb{E}[X_{n-1}^2] + 2\mathbb{E}[X_{n-1}\mathbb{E}(X_n - X_{n-1} \,|\, X_0, \ldots, X_{n-1})] + \mathbb{E}[(X_n - X_{n-1})^2] = \mathbb{E}[X_{n-1}^2] + \mathbb{E}[(X_n - X_{n-1})^2] = \cdots = \mathbb{E}[X_0^2] + \sum_{k=1}^{n} \mathbb{E}[(X_k - X_{k-1})^2] \le \mathbb{E}[X_0^2] + \sum_{n=1}^{\infty} \mathbb{E}[(X_n - X_{n-1})^2] < \infty$. Using the martingale convergence theorem, X_n converges almost surely.

Problem 47. Let $X_{n,i}$ be i.i.d. nonnegative integer valued random variables with mean $\mu \ge 0$. Define $Z_0 = 1$ and $Z_{n+1} = (X_{n+1,1} + \cdots + X_{n+1,Z_n})$ if $Z_n > 0$ and $Z_{n+1} = 0$ if $Z_n = 0$.
(a) Show that Z_n/μ^n is a martingale.
(b) Show that $Z_n \to 0$ if $\mu < 1$.
(c) Show that $Z_n \to 0$ if $\mu = 1$ and $P(X_{n,i} = 1) < 1$.

Solution. (a) $\mathbb{E}[Z_{n+1}/\mu^{n+1} \,|\, Z_0, \ldots, Z_n] = \mathbb{E}[X_{n+1,1} + \cdots + X_{n+1,Z_n} \,|\, Z_n]/\mu^{n+1} = \mu Z_n/\mu^{n+1} = Z_n/\mu^n$ indicates Z_n/μ^n is a martingale.
(b) Note that $\mathbb{E}(Z_n) = \mu^n \mathbb{E}(W_n) = \mu^n \mathbb{E}(W_0) = \mu^n$. Hence for any $\epsilon \in (0,1)$, $\sum_{n=1}^{\infty} P(Z_n > \epsilon) = \sum_{n=1}^{\infty} P(Z_n > 0) \le \sum_{n=1}^{\infty} \mathbb{E}(Z_n) = \sum_{n=1}^{\infty} \mu^n = \mu/(1-\mu) < \infty$. Hence $Z_n \to 0$ almost surely.
(c) The martingale convergence theorem implies $Z_n = W_n \to W$ almost surely. If $P(X_{1,1} = 0) = 0$, then $X_{1,1} \ge 1$ implies $X_{1,1} - 1 \ge 0$ and $\mathbb{E}(|X_{1,1} - 1|) = \mathbb{E}(X_{1,1} - 1) = \mathbb{E}(X_{1,1}) - 1 = \mu - 1 = 0$. Hence $X_{1,1} = 1$ almost surely. It contradicts to the assumption $P(X_{1,1}) < 1$. Hence $p_0 = P(X_{1,1} = 0) > 0$. The state 0 is absolving and all other states are transient because $p(i,0) = P(Z_2 = 0 \,|\, Z_1 = i) = P(X_{2,1} = 0, \ldots, X_{2,i} = 0 \,|\, Z_1 = i) = \{P(X_{2,1} = 0)\}^i = p_0^i > 0$. While $p(0,i) = 0$ for all $i > 0$. Let $q_n = P(T_0 < \infty \,|\, Z_0 = n) = [P(T_0 < \infty \,|\, Z_0 = 1)]^n = q_1^n$. Then $q_1 = P(T_0 < \infty \,|\, Z_0 = 1) = p(1,0) + \sum_{k=1}^{\infty} p(1,k) q_k = \sum_{k=0}^{\infty} p(1,k) q_1^k = \mathbb{E}[q_1^{X_{1,1}}] = g(q_1)$ where $g(s) = \mathrm{pgf}_X(s)$. Note that $g''(s) = \mathbb{E}[X(X-1)s^{X-2}] = \sum_{k=2}^{\infty} k(k-1)s^{k-2} P(X = k) \ge 0$ implies g is convex. Since $g(1) = 1, g'(1) = \mathbb{E}(X) = 1$, $g(s) \ge g(1) + g'(1)(s-1) = 1 + (s-1) = s$ for all $s \ge 0$. Hence $g(s) = s$ has unique solution $s = 1$. That means, $q_1 = 1$ and $q_n = q_1^n = 1$ for all n. Then $P(Z_n > 0) = P(T_0 > n) \to 0$. Hence $Z_n \to 0$ in probability, that means, $W = 0$ and the martingale convergence theorem implies $Z_n \to 0$ almost surely as well as in L^1.

Problem 48. Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}X_n = 0$ and $\mathbb{E}|X_n|^p < \infty$ for some $1 < p < 2$. Show that $(X_1 + \cdots + X_n)/n^{p/2}$ converges to 0 a.s.

Please ignore it is a bit beyond our scope.

Problem 49. Show that $X_n \to X$ in probability if and only if $\mathbb{E}[|X_n X|/(1 + |X_n X|)] \to 0$.

Solution. If $X_n \xrightarrow{p} X$, then for $\epsilon \in (0, 1/2]$,

$$\mathbb{E}\Big[\frac{|X_n - X|}{1+|X_n - X|}\Big] = \mathbb{E}\Big[\frac{|X_n - X|}{1+|X_n - X|}1(|X_n - X| < \epsilon)\Big] + \mathbb{E}\Big[\frac{|X_n - X|}{1+|X_n - X|}1(|X_n - X| \geq \epsilon)\Big]$$
$$\leq \epsilon P(|X_n - X| < \epsilon) + P(|X_n - X| \geq \epsilon) = \epsilon + (1-\epsilon)P(|X_n - X| \geq \epsilon) \leq \epsilon + P(|X_n - X| \geq \epsilon)$$

send $n \to \infty$ and then $\epsilon \to 0$ to obtain

$$\mathbb{E}\Big[\frac{|X_n - X|}{1+|X_n - X|}\Big] \leq \lim_{\epsilon \searrow 0} \lim_{n\to\infty} [\epsilon + P(|X_n - X| \geq \epsilon)] = 0.$$

If $\mathbb{E}[|X_n - X|/(1+|X_n - X|)] \to 0$, then by noting for $x \geq 0$, $x \mapsto x/(1+x)$ is increasing and bounded by 1, we get

$$P(|X_n - X| \geq \epsilon) = P(|X_n - X|/(1+|X_n - X|) \geq \epsilon/(1+\epsilon))$$
$$= \mathbb{E}[1(|X_n - X|/(1+|X_n - X|) \geq \epsilon/(1+\epsilon))] \leq \frac{1+\epsilon}{\epsilon}\mathbb{E}\Big[\frac{|X_n - X|}{1+|X_n - X|}\Big] \to 0.$$

Problem 50. Let X and Y be i.i.d. from a distribution having finite second moment. Also $(X+Y)/\sqrt{2}$ and X have the same distribution. Find the distribution of X.

Solution. Let F be the distribution of X. Let $X_1, X_2, \dots$ be i.i.d. sequence from F.

From $\mathbb{E}((X+Y)/\sqrt{2}) = \sqrt{2}\mathbb{E}(X) = \mathbb{E}(X)$, $\mathbb{E}(X) = 0$ is obtained. Using the central limit theorem, X and $(X_1 + \cdots + X_{2^k})/2^{k/2}$ have the same distribution. Send $k \to \infty$ to obtain $X \equiv^d (X_1 + \cdots + X_{2^k})/2^{k/2} \to N(0, \sigma^2)$ by the central limit theorem where $\sigma^2 = \mathbb{V}\mathrm{ar}(X_1^2)$. Therefore $X \sim N(0, \sigma^2)$.

Problem 51. Show that $X_n + Y_n \to X + Y$ in L^p if $X_n \to X, Y_n \to Y$ in L^p.

Solution. Note that $|x+y|^p \leq c_p(|x|^p + |y|^p)$ where $c_p = 2^{\max(1,p-1)}$. Hence $\mathbb{E}(|X+Y|^p) \leq c_p\mathbb{E}(|X|^p + |Y|^p) < \infty$ and $\mathbb{E}(|X_n + Y_n - (X+Y)|^p) = \mathbb{E}(|(X_n - X) + (Y_n - Y)|^p) \leq c_p\mathbb{E}(|X_x - X|^p) + |Y_n - Y|^p) \to 0$. Hence $X_n + Y_n \to X + Y$ in L^p.

Problem 52. Let $X_1, X_2, \dots$ be an i.i.d. random variables satisfying $\mathbb{E}(|X_n|) < \infty$. Show that $\bar{X}_n \to \mathbb{E}(X_1)$ in L^1.

Solution. We prove a generalized problem, that is, if $X_n \to X$ in probability and $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|)$, then $X_n \to X$ in L^1.
Let $Y_n = |X_n|$ and $Y = |X|$ so that $Y_n \to Y$ in probability and L^1. Then $|X_n - X| \leq Y_n + Y$ for all n and $Y_n + Y \to 2Y$ in probability, $\mathbb{E}(Y_n + Y) \to \mathbb{E}(2Y)$. Using the generalized dominated convergence theorem, $\mathbb{E}(|X_n - X|) \to \mathbb{E}(0) = 0$.

Problem 53. Let $X_1, X_2, \dots$ be i.i.d. with finite second moment.
(a) Show that $\bar{X}_n = (X_1 + \cdots + X_n)/n$ converges to $\mathbb{E}(X_1)$ in probability, in L^2 and almost surely.
(b) Show that $S_n^2 = [(X_1 - \bar{X}_n)^2 + \cdots + (X_n - \bar{X}_n)^2]/n$ converges to $\mathbb{V}\mathrm{ar}(X_1)$ in probability, in L^1 and almost surely.

Solution. (a) Theorems in lecture note and the above problem.
(b) Note that $S_n^2 = (X_1^2 + \cdots + X_n^2)/n - \bar{X}_n^2$. It is known that $(X_1^2 + \cdots + X_n^2)/n \to \mathbb{E}(X_1^2)$ in L^1 (by the above problem) and almost surely (by strong law of large numbers).

Since $\bar{X}_n \to \mu$ in L^2, $\mathbb{E}[|\bar{X}_n - \mu|^2] \to 0$ and $\mathbb{E}[\bar{X}_n^2] \to \mu^2$. Then $\mathbb{E}[|S_n^2|] = \mathbb{E}[S_n^2] = \mathbb{E}[(X_1^2 + \cdots + X_n^2)/n] - \mathbb{E}[\bar{X}_n^2] \to \mathbb{E}(X_1^2) - \mu^2 = \mathbb{V}\mathrm{ar}(X_1) = \sigma^2$. Hence $S_n^2 \to \sigma^2$ in L^1.

Since $\bar{X}_n \to \mu$ almost surely, the continuous mapping theorem implies $(\bar{X}_n)^2 \to \mu^2$ almost surely. Then $S_n^2 = (X_1^2 + \cdots + X_n^2)/n - \bar{X}_n^2 \to \mathbb{E}(X_1^2) - (\mathbb{E}(X_1))^2 = \mathbb{V}\mathrm{ar}(X_1) = \sigma^2$ almost surely.

Problem 54. Let X_n be a homogeneous Markov chaing of which transition matrix is

$$p = \begin{array}{cccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} \\ \mathbf{a} & 0 & 1 & 0 \\ \mathbf{b} & 1 & 0 & 0 \\ \mathbf{c} & 0.2 & 0.3 & 0.5 \end{array}$$

(a) Find an irreducible set.
(b) Determine whether each state is recurrent or not.
(c) Find the period of each state.
(d) Prove or disprove the uniqueness of a stationary distribution.

Solution. (a) Note that $a \to b \to a$ implies $\{a, b\}$ is irreducible.
(b) Note that $a \to b \to a$ implies $\{a, b\}$ is irreducible and finite. Hence it is recurrent. Besides $\rho_{ca} \geq p(c, a) = 0.2 > 0$ while $\rho_{ac} = 0$. Thus c is transient.
(c) $p(c, c) = 0.5$ implies that the period of c is 1. States a, b are in an irreducible set. Hence periods are the same. Consider $a \xrightarrow{w.p.1} b \xrightarrow{w.p.1} a \xrightarrow{w.p.1} b \xrightarrow{w.p.1} a \cdots$. Hence periods of a and b are 2.
(d) The size of state space is finite. Hence there exists at least a stationary distribution π. Note that $\pi(x) = 0$ for any transient state x. Hence $\pi(c) = 0$ and π can be considered as a stationary distribution restricted on recurrent states. The set of recurrent states is $\{a, b\}$ which is irreducible. There exists the unique stationary distribution on $\{a, b\}$, that is, $\pi(a) = 1/(1+1), \pi(b) = 1/(1+1)$ is the unique stationary distribution on $\{a, b\}$. In sum, $\pi(a) = 1/2, \pi(b) = 1/2, \pi(c) = 0$ is the unique stationary distribution satisfying $\pi p = \pi$.

Problem 55. Let X_n be a HMC with state space $\mathcal{S} = \{A, B\}$ and the transition probability

$$p = \begin{array}{ccc} & \mathbf{A} & \mathbf{B} \\ \mathbf{A} & 1-\alpha & \alpha \\ \mathbf{B} & \beta & 1-\beta \end{array}$$

(a) Compute $P_A(T_A = n)$ where T_A is the first returning time to A, that is, $T_A = \inf\{n \geq 1 : X_n = A\}$.
(b) Compute $\mathbb{E}_A T_A$.

Solution. (a) For $n = 1$, $P_A(T_A = n) = P_A(X_1 = A) = p(A, A) = 1 - \alpha$. For $n > 1$,

$$\begin{aligned}
P_A(T_A = n) &= P_A(X_1 \neq A, \ldots, X_{n-1} \neq A, X_n = A) = P_A(X_1 = B, \ldots, X_{n-1} = B, X_n = A)\\
&= p(A, B)p(B, B) \cdots p(B, B)p(B, A) = p(A, B)p(B, B)^{n-2}p(B, A) = \alpha(1 - \beta)^{n-2}\beta.
\end{aligned}$$

(b) (Solution 1) A bit hard solution.

$$\begin{aligned}
\mathbb{E}_A(T_A) &= P_A(T_A = 1) + \sum_{n=2}^{\infty} nP_A(T_A = n) = 1 - \alpha + \sum_{n=2}^{\infty} n\alpha\beta(1 - \beta)^{n-2}\\
&= 1 - \alpha + 2\alpha\beta \sum_{n=2}^{\infty}(1 - \beta)^{n-2} + \alpha\beta \sum_{n=2}^{\infty}(n - 2)(1 - \beta)^{n-2}\\
&= 1 - \alpha + 2\alpha\beta/(1 - (1 - \beta)) + \alpha\beta(1 - \beta)/(1 - (1 - \beta))^2 = 1 - \alpha + 2\alpha + \alpha(1/\beta - 1)\\
&= 1 + \alpha/\beta.
\end{aligned}$$

(Solution 2) Very easy solution. Note that $\pi(A) = 1/\mathbb{E}_A(T_A) = \beta/(\alpha + \beta)$. Hence $\mathbb{E}_A(T_A) = (\alpha + \beta)/\beta = 1 + \alpha/\beta$.

Problem 56. There is a plant species blooming three different colors (red, white and pink). If pollinated within the same flower color group, the flower color of the offspring follows a homogeneous Markov chain having the transition probability

$$p = \begin{array}{c|ccc}
 & \textbf{red} & \textbf{white} & \textbf{pink}\\
\textbf{red} & 1 & 0 & 0\\
\textbf{white} & 0 & 1 & 0\\
\textbf{pink} & 1/4 & 1/4 & 1/2
\end{array}$$

(a) Compute the probability that pink color flower eventually absorbed into the red color flower group.
(b) Compute the expected time (generation) that pink color flower eventually absorbed into either red or white color flower group.

Solution. (a) $P_{\text{pink}}(T_{\text{red}} < \infty) = p(\text{pink}, \text{red}) + p(\text{plink}, \text{pink})P_{\text{pink}}(T_{\text{red}} < \infty)$. Hence $P_{\text{pink}}(T_{\text{red}} < \infty) = p(\text{pink}, \text{red})/(1 - p(\text{pink}, \text{pink})) = 1/4/(1 - 1/2) = 1/2$.
(b) Let $T = \min(T_{\text{red}}, T_{\text{white}})$. Since pink is a transient state, $P_{\text{pink}}(T < \infty) = 1$. $\mathbb{E}_{\text{pink}}T1(T < \infty) = \mathbb{E}_{\text{pink}}T = \mathbb{E}_{\text{pink}} \sum_y T1(X_1 = y) = p(\text{pink}, \text{red}) + p(\text{pink}, \text{white}) + p(\text{pink}, \text{pink})\mathbb{E}_{\text{pink}}(1 + T) = 1/4 + 1/4 + (1/2)(1 + \mathbb{E}_{\text{pink}}T)$. Hence $\mathbb{E}_{\text{pink}}T1(T < \infty) = (1/4 + 1/4 + 1/2)/(1/2) = 2$.

Problem 57. Let X_n be a homogeneous Markov chain. Let A be a closed subset of recurrent states and B be the set of recurrent states not in A. Assume both A and B are nonempty. Define $h(x) = 1$ for all $x \in A$, $h(x) = 0$ for all $x \in B$ and $h(x) = \sum_{y \in \mathcal{S}} p(x, y)h(y)$ for all $x \notin A \cup B$ where p is the transition probability. Show that $h(X_n)$ is a martingale.

Solution. Note that $\mathbb{E}[h(X_{n+1}) \,|\, X_0, \ldots, X_n] = \mathbb{E}[h(X_{n+1}) \,|\, X_n] = \sum_{y \in \mathcal{S}} p(X_n, y)h(y) = h(X_n)$. Hence $h(X_n)$ is a martingale.

Problem 58. John is playing a gamble. He gains a dollar when he tosses a fair coin and it lands head. Otherwise he loses a dollar. He starts the gamble with \$3 and will stop the gamble when either he loses all money or his wealth becomes \$5. Let X_n be the wealth of John at time n which is known to be a homogeneous Markov chain.
(a) Specify the state space and transition probability.
(b) Compute the probability John's wealth reaches \$5 before it reaches \$0.
(c) Compute the expected time for John to stop the gambling.

Solution. (a) The possible wealth states are $\{\$0, \$1, \$2, \$3, \$4, \$5\}$. The transition probability is $p(0, x) = I(x = 0)$, $p(5, x) = I(x = 5)$ and $p(x, x-1) = p(x, x+1) = 1/2$ for $x = 1, 2, 3, 4$.
(b) There are two absorbing states $0, 5$. All other states are transient since $\rho_{x,0} \geq 1/2^x > 0$ while $\rho_{0,x} = 0$ for $x = 1, 2, 3, 4$.

Let $H_x = \inf\{t \geq 0 : X_t = x\}$ and $h(x) = P_x(H_5 < H_0)$. Then $h(5) = 1$, $h(0) = 0$ and $h(x) = \sum_{y=0}^{5} P_x(X_1 = y, H_5 < H_0) = p(x, x-1)P_{x-1}(H_5 < H_0) + p(x, x+1)P_{x+1}(H_5 < H_0) = (h(x-1) + h(x+1))/2$ for $x = 1, 2, 3, 4$. It solves $h(x+1) - h(x) = h(x) - h(x-1) = h(1) - h(0) = h(1)$ and $h(x) = xh(1)$ with $h(5) = 1$. Thus $h(x) = x/5$. Finally $P_3(H_5 < H_0) = h(3) = 3/5 = 0.6$.
(c) Let $g(x)$ be the expected time to stop the gambling when started at state x, that is, $g(x) = \mathbb{E}_x \min(H_0, H_5)$. Then it satisfies $g(0) = g(5) = 0$ and $g(x) = p(x, x-1)(1 + g(x-1)) + p(x, x+1)(1 + g(x+1)) = 1 + (g(x-1) + g(x+1))/2$. It solves $g(x+1) - g(x) = g(x) - g(x-1) - 2 = g(1) - g(0) - 2x = g(1) - 2x$. It gives $g(x) = g(x-1) + g(1) - 2(x-1) = g(x-2) + 2g(1) - 2((x-1) + (x-2)) = g(1) + (x-1)g(10 - 2((x-1) + (x-2) + \cdots + 1) = xg(1) - (x-1)x$ with $0 = g(5) = 5g(1) - 20$. Hence $g(1) = 4$ and $g(x) = 4x - (x-1)x = x(5-x)$. Therefore $g(3) = 3 \times 2 = 6$.

Problem 59. Let X_n and Y_n be two positive stochastic processes satisfying $\mathbb{E}(X_{n+1} \mid X_0, \ldots, X_n) \leq X_n Y_n$. Assume that Y_n's are functions of $X_0, \ldots, X_n$, that is, $Y_n = g_n(X_0, \ldots, X_n)$ for some functions g_n. Show that Z_n defined by $Z_1 = X_1$ and $Z_n = X_n / \prod_{k=1}^{n-1} Y_k$ for $n \geq 2$ is supermartingale.

Solution. Note that

$$\mathbb{E}[Z_{n+1} \mid X_0, \ldots, X_n] = \mathbb{E}[X_{n+1} / \prod_{k=1}^{n} Y_k \mid X_0, \ldots, X_n] = \mathbb{E}[X_{n+1} \mid X_0, \ldots, X_n] / \prod_{k=1}^{n} Y_k \leq (X_n Y_n) / \prod_{k=1}^{n} Y_k = Z_n.$$

Hence Z_n is supermartingale.

Problem 60. Let X_n and Y_n be two positive stochastic processes satisfying $\mathbb{E}(X_{n+1} \mid X_0, \ldots, X_n) \leq X_n + Y_n$. Assume that Y_n's are functions of $X_0, \ldots, X_n$, that is, $Y_n = g_n(X_0, \ldots, X_n)$ for some functions g_n. Show that Z_n defined by $Z_1 = X_1$ and $Z_n = X_n - \sum k = 1^{n-1} Y_k$ for $n \geq 2$ is supermartingale.

Solution. A simple computation gives $\mathbb{E}[Z_{n+1} \mid X_0, \ldots, X_n] = \mathbb{E}[X_{n+1} - \sum_{k=1}^{n} Y_k \mid X_0, \ldots, X_n] = \mathbb{E}[X_{n+1} \mid X_0, \ldots, X_n] - \sum_{k=1}^{n} Y_k \leq (X_n + Y_n) - \sum_{k=1}^{n} Y_k = Z_n$. Hence Z_n is supermartingale.

Problem 61. Let $X_1, X_2, \ldots$ be an i.i.d. sequence of random variables with $\mathbb{E}(|X_n|^k) < \infty$ for a positive integer k. Let $\mu_k = \mathbb{E}(X_n^k)$. For a sequence of posivie number a_n with $a_n \to \infty$, show that $Y_n = (X_1^k - \mu_k)/a_1 + \cdots + (X_n^k - \mu_k)/a_n$ is a martingale.

Solution. Note that $Y_{n+1} = Y_n + (X_{n+1}^k - \mu_k)/a_{n+1}$. Then $\mathbb{E}(Y_{n+1} \mid X_0, \ldots, X_n) = Y_n + \mathbb{E}((X_{n+1}^k - \mu_k)/a_{n+1})) = Y_n$. Hence Y_n is a martingale.

The followings might be useful

1. $e^x = \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!}$ for any real number x.
2. $\log(1-z) = -\displaystyle\sum_{n=1}^{\infty} \frac{z^n}{n}$ for $|z| < 1$.
3. $\displaystyle\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}, \sum_{n=0}^{\infty} n r^n = \frac{r}{(1-r)^2}$ for $|r| < 1$.
4. $\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}^n = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix} + (1-a-b)^n \begin{pmatrix} \pi_2 & -\pi_2 \\ -\pi_1 & \pi_1 \end{pmatrix}$ where $\pi_1 = \frac{b}{a+b}, \pi_2 = \frac{a}{a+b}$.
5. $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \dfrac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ when $ad - bc \neq 0$.