Problem 1

Random social networks. Let G = (V,E) be a random graph with m = |V| vertices and edge-set E. Write d_v for the degree of vertex v, that is, the number of edges meeting at v. Let Y be a uniformly chosen vertex, and Z a uniformly chosen neighbour of Y.
(a) Show that $Ed_Z \geq Ed_Y$.
(b) Interpret this inequality when the vertices represent people, and the edges represent friendship.

Proof. **(a) Show that $Ed_Z \geq Ed_Y$.**

First, we compute $E[d_Y]$. Since Y is uniformly chosen from V:

$$E[d_Y] = \sum_{v \in V} d_v \cdot P(Y = v) = \sum_{v \in V} d_v \cdot \frac{1}{m} = \frac{1}{m} \sum_{v \in V} d_v$$

Since $\sum_{v \in V} d_v = 2|E|$, we have $E[d_Y] = \frac{2|E|}{m}$.

For $E[d_Z]$, we need to analyze how Z is selected. For a vertex u:

$$\begin{aligned} P(Z = u) &= \sum_{v \in V} P(Z = u | Y = v) \cdot P(Y = v) \\ &= \frac{1}{m} \sum_{v:(u,v) \in E} \frac{1}{d_v} \end{aligned}$$

Therefore:

$$\begin{aligned} E[d_Z] &= \sum_{u \in V} d_u \cdot P(Z = u) \\ &= \frac{1}{m} \sum_{u \in V} d_u \sum_{v \in N(u)} \frac{1}{d_v} \\ &= \frac{1}{m} \sum_{(u,v) \in E} \frac{d_u}{d_v} \end{aligned}$$

By symmetry, this equals:

$$E[d_Z] = \frac{1}{2m} \sum_{(u,v) \in E} \left(\frac{d_u}{d_v} + \frac{d_v}{d_u} \right)$$

Since $\frac{d_u}{d_v} + \frac{d_v}{d_u} \geq 2$ by the AM-GM inequality:

$$E[d_Z] \geq \frac{1}{2m} \sum_{(u,v) \in E} 2 = \frac{2|E|}{2m} = \frac{2|E|}{m} = E[d_Y]$$

Thus, $E[d_Z] \geq E[d_Y]$.

(b) Interpret this inequality when the vertices represent people, and the edges represent friendship.

This inequality represents the ”Friendship Paradox” which states that, on average, your friends have more friends than you do. This occurs because individuals with many connections are overrepresented in friendship networks. When people compare their social circles to those of their friends, they often feel less connected, creating a perception bias. This mathematical property helps explain social phenomena like feelings of inadequacy on social media, and has practical applications in monitoring trends and information diffusion in populations. □

Problem 2

In your pocket is a random number N of coins, where N has the Poisson distribution with parameter λ. You toss each coin once, with heads showing with probability p each time. Show that the total number of heads has the Poisson distribution with parameter λp.

Proof. Let $N \sim \text{Poisson}(\lambda)$ represent the random number of coins in your pocket. Each coin is tossed once, with probability p of showing heads. Let H denote the total number of heads observed. We aim to show that $H \sim \text{Poisson}(\lambda p)$.

Using the law of total probability, we can write:

$$P(H = x) = \sum_{n=x}^{\infty} P(H = x|N = n) \cdot P(N = n)$$

Since $H|N = n \sim \text{Binomial}(n, p)$ and $N \sim \text{Poisson}(\lambda)$, we have:

$$\begin{aligned}
P(H = x) &= \sum_{n=x}^{\infty} \binom{n}{x} p^x (1-p)^{n-x} \cdot \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \sum_{n=x}^{\infty} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \cdot \frac{\lambda^n e^{-\lambda}}{n!} \\
&= \frac{p^x e^{-\lambda}}{x!} \sum_{n=x}^{\infty} \frac{\lambda^n (1-p)^{n-x}}{(n-x)!}
\end{aligned}$$

Let’s extract common factors:

$$\begin{aligned}
P(H = x) &= \frac{p^x \lambda^x e^{-\lambda}}{x!} \sum_{n=x}^{\infty} \frac{\lambda^{n-x} (1-p)^{n-x}}{(n-x)!} \\
&= \frac{(p\lambda)^x e^{-\lambda}}{x!} \sum_{n=x}^{\infty} \frac{[\lambda(1-p)]^{n-x}}{(n-x)!}
\end{aligned}$$

Making a change of variables with $m = n - x$, the sum becomes:

$$\begin{aligned} P(H = x) &= \frac{(p\lambda)^x e^{-\lambda}}{x!} \sum_{m=0}^{\infty} \frac{[\lambda(1-p)]^m}{m!} \\ &= \frac{(p\lambda)^x e^{-\lambda}}{x!} \cdot e^{\lambda(1-p)} \end{aligned}$$

Since $e^{\lambda(1-p)} = e^{\lambda - \lambda p}$, we get:

$$\begin{aligned} P(H = x) &= \frac{(p\lambda)^x e^{-\lambda}}{x!} \cdot e^{\lambda - \lambda p} \\ &= \frac{(p\lambda)^x e^{-\lambda p}}{x!} \end{aligned}$$

This is precisely the probability mass function of a Poisson distribution with parameter λp, completing our proof that $H \sim \text{Poisson}(\lambda p)$. $\square$

Problem 3

Compound Poisson distribution. Let Λ be a positive random variable with density function f and distribution function F, and let Y have the Poisson distribution with parameter Λ. Show for n = 0, 1, 2, ... that

$$\mathbb{P}(Y \le n) = \int_0^{\infty} p_n(\lambda) F(\lambda)\, d\lambda, \qquad \mathbb{P}(Y > n) = \int_0^{\infty} p_n(\lambda)[1 - F(\lambda)]\, d\lambda,$$

where $p_n(\lambda) = e^{-\lambda}\lambda^n/n!$.

Proof. First, let's calculate $\mathbb{P}(Y \le n)$ using conditional probability and the law of total probability. Given that $\Lambda = \lambda$, Y follows a Poisson distribution with parameter λ, so:

$$\begin{aligned} \mathbb{P}(Y \le n) &= \int_0^{\infty} \mathbb{P}(Y \le n | \Lambda = \lambda) f(\lambda)\, d\lambda \\ &= \int_0^{\infty} \sum_{r=0}^{n} \frac{e^{-\lambda}\lambda^r}{r!} f(\lambda)\, d\lambda \\ &= \sum_{r=0}^{n} \int_0^{\infty} \frac{e^{-\lambda}\lambda^r}{r!} f(\lambda)\, d\lambda \\ &= \sum_{r=0}^{n} \int_0^{\infty} p_r(\lambda) f(\lambda)\, d\lambda \end{aligned}$$

Now, we'll apply integration by parts to each term in the sum. For each r, let:

$$u = F(\lambda) \quad \text{and} \quad dv = p_r(\lambda)\, d\lambda$$

Then $du = f(\lambda)\,d\lambda$ and v needs to be determined by integrating $p_r(\lambda)$. We also know that:

$$\begin{aligned}\frac{d}{d\lambda}p_r(\lambda) &= \frac{d}{d\lambda}\left(\frac{e^{-\lambda}\lambda^r}{r!}\right)\\ &= \frac{1}{r!}\left(-e^{-\lambda}\lambda^r + e^{-\lambda}r\lambda^{r-1}\right)\\ &= -p_r(\lambda) + \frac{r}{\lambda}p_r(\lambda)\end{aligned}$$

For $r < n$, we can establish a recurrence relation:

$$p_{r+1}(\lambda) = \frac{\lambda}{r+1}p_r(\lambda)$$

Using integration by parts:

$$\begin{aligned}\int_0^\infty p_r(\lambda)f(\lambda)\,d\lambda &= [p_r(\lambda)F(\lambda)]_0^\infty - \int_0^\infty F(\lambda)\frac{d}{d\lambda}p_r(\lambda)\,d\lambda\\ &= [p_r(\lambda)F(\lambda)]_0^\infty - \int_0^\infty F(\lambda)\left(-p_r(\lambda) + \frac{r}{\lambda}p_r(\lambda)\right)\,d\lambda\\ &= [p_r(\lambda)F(\lambda)]_0^\infty + \int_0^\infty F(\lambda)p_r(\lambda)\,d\lambda - \int_0^\infty F(\lambda)\frac{r}{\lambda}p_r(\lambda)\,d\lambda\end{aligned}$$

The boundary term $[p_r(\lambda)F(\lambda)]_0^\infty$ evaluates to 0 because: - At $\lambda = 0$, $p_r(0) = 0$ for $r > 0$ and $F(0) = 0$ for a positive random variable - As $\lambda \to \infty$, $p_r(\lambda) \to 0$ exponentially while $F(\lambda) \to 1$

Using the recurrence relation and simplifying:

$$\int_0^\infty p_r(\lambda)f(\lambda)\,d\lambda = \int_0^\infty F(\lambda)p_r(\lambda)\,d\lambda - \int_0^\infty F(\lambda)\frac{r+1}{\lambda}p_{r+1}(\lambda)\,d\lambda$$

Summing over all r from 0 to n:

$$\begin{aligned}\sum_{r=0}^{n}\int_0^\infty p_r(\lambda)f(\lambda)\,d\lambda &= \sum_{r=0}^{n}\int_0^\infty F(\lambda)p_r(\lambda)\,d\lambda - \sum_{r=0}^{n-1}\int_0^\infty F(\lambda)p_{r+1}(\lambda)\,d\lambda\\ &= \int_0^\infty F(\lambda)p_n(\lambda)\,d\lambda\end{aligned}$$

The telescoping sum cancels out all terms except the last one with $p_n(\lambda)$. Therefore:

$$\mathbb{P}(Y \leq n) = \int_0^\infty p_n(\lambda)F(\lambda)\,d\lambda$$

For the second part, we have:

$$\mathbb{P}(Y > n) = 1 - \mathbb{P}(Y \leq n)$$

$$= 1 - \int_0^\infty p_n(\lambda)F(\lambda)\,d\lambda$$

Since $\int_0^\infty p_n(\lambda)\,d\lambda = 1$ (as $p_n(\lambda)$ can be viewed as a gamma distribution density with appropriate normalization), we have:

$$\begin{aligned}\mathbb{P}(Y > n) &= \int_0^\infty p_n(\lambda)\,d\lambda - \int_0^\infty p_n(\lambda)F(\lambda)\,d\lambda \\ &= \int_0^\infty p_n(\lambda)(1 - F(\lambda))\,d\lambda\end{aligned}$$

□

Problem 4

Mutual information. Let X and Y be discrete random variables with joint mass function f. (a) Show that E(log $f_X(X)$)   E(log $f_Y(X)$). (b) Show that the mutual information

$$I = \mathbb{E}\left(\log\left\{\frac{f(X,Y)}{f_X(X)f_Y(Y)}\right\}\right)$$

satisfies I   0, with equality if and only if X and Y are independent.

Proof. **(a)** We'll use the inequality $\log y \leq y - 1$ with equality if and only if $y = 1$.

Setting $y = \frac{f_Y(X)}{f_X(X)}$ and taking expectations:

$$\begin{aligned}\mathbb{E}\left[\log\left(\frac{f_Y(X)}{f_X(X)}\right)\right] &\leq \mathbb{E}\left[\frac{f_Y(X)}{f_X(X)} - 1\right] \\ &= \sum_x f_X(x)\left(\frac{f_Y(x)}{f_X(x)} - 1\right) \\ &= \sum_x f_Y(x) - \sum_x f_X(x) \\ &= 1 - 1 = 0\end{aligned}$$

Therefore:

$$\mathbb{E}[\log f_X(X)] \geq \mathbb{E}[\log f_Y(X)]$$

Equality holds if and only if $f_Y = f_X$.

(b) For the mutual information:

$$I = \mathbb{E}\left[\log\left(\frac{f(X,Y)}{f_X(X)f_Y(Y)}\right)\right]$$

Applying the same inequality with $y = \frac{f_X(X)f_Y(Y)}{f(X,Y)}$:

$$-\log\left(\frac{f(X,Y)}{f_X(X)f_Y(Y)}\right) \leq \frac{f_X(X)f_Y(Y)}{f(X,Y)} - 1$$

Taking the expectation and multiplying by -1:

$$\begin{aligned} I &\geq \mathbb{E}\left[1 - \frac{f_X(X)f_Y(Y)}{f(X,Y)}\right] \\ &= 1 - \sum_{x,y} f_X(x)f_Y(y) = 0 \end{aligned}$$

Therefore, $I \geq 0$, with equality if and only if $f(x,y) = f_X(x)f_Y(y)$ for all (x,y), which is equivalent to X and Y being independent. □

Problem 5

Let X be a non-negative random variable with density function f. Show that

$$\mathbb{E}(X^r) = \int_0^\infty rx^{r-1}\mathbb{P}(X > x)dx$$

for any $r \geq 1$ for which the expectation is finite.

Proof. We'll use integration by parts. First, by definition:

$$\mathbb{E}(X^r) = \int_0^\infty x^r f(x)dx$$

Now, let's work with the right side of the equation to be proven:

$$\int_0^\infty rx^{r-1}\mathbb{P}(X > x)dx$$

Note that $\mathbb{P}(X > x) = \int_x^\infty f(t)dt$ and set:

$$\begin{aligned} u &= \mathbb{P}(X > x) = \int_x^\infty f(t)dt \\ dv &= rx^{r-1}dx \end{aligned}$$

Then:

$$\begin{aligned} du &= -f(x)dx \\ v &= x^r \end{aligned}$$

Using the integration by parts formula $\int udv = uv - \int vdu$:

$$\int_0^\infty rx^{r-1}\mathbb{P}(X > x)dx = [x^r\mathbb{P}(X > x)]_0^\infty - \int_0^\infty x^r(-f(x))dx$$

$$= \lim_{a\to\infty} a^r \mathbb{P}(X > a) - 0 \cdot \mathbb{P}(X > 0) + \int_0^\infty x^r f(x)dx$$

To complete the proof, we need to show that $\lim_{a\to\infty} a^r \mathbb{P}(X > a) = 0$.

Since $\mathbb{E}(X^r)$ is finite, we can apply Markov's inequality:

$$\mathbb{P}(X > a) \leq \frac{\mathbb{E}(X^r)}{a^r}$$

Therefore:

$$a^r \mathbb{P}(X > a) \leq \mathbb{E}(X^r)$$

As $a \to \infty$, we must have $a^r \mathbb{P}(X > a) \to 0$, otherwise $\mathbb{E}(X^r)$ would not be finite.

Thus:

$$\begin{aligned} \int_0^\infty r x^{r-1} \mathbb{P}(X > x)dx &= 0 + \int_0^\infty x^r f(x)dx \\ &= \mathbb{E}(X^r) \end{aligned}$$

This completes the proof that $\mathbb{E}(X^r) = \int_0^\infty r x^{r-1} \mathbb{P}(X > x)dx$ for any $r \geq 1$ where $\mathbb{E}(X^r) < \infty$. □

Problem 6

Log-normal distribution. Let $Y = e^X$ where X has the $N(0, 1)$ distribution. Find the density function of Y.

Proof. Let $Y = e^X$ where $X \sim N(0, 1)$. To find the density function of Y, we use the change of variable technique.

The CDF of Y is:

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \ln y) = \Phi(\ln y)$$

Differentiating to get the PDF:

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy} \Phi(\ln y) \\ &= \phi(\ln y) \cdot \frac{1}{y} \end{aligned}$$

Substituting $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$:

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(\ln y)^2}{2}} \cdot \frac{1}{y} \\ &= \frac{1}{\sqrt{2\pi} \cdot y} e^{-\frac{(\ln y)^2}{2}}, \quad y > 0 \end{aligned}$$

This is the standard log-normal distribution. □

Problem 7

Show that $E|X| < \infty$ if and only if the following holds: for all $\epsilon > 0$, there exists $\delta > 0$, such that $E(|X|I_A) < \epsilon$ for all A such that $P(A) < \delta$.

Proof. We need to show that $E|X| < \infty$ if and only if for all $\epsilon > 0$, there exists $\delta > 0$ such that $E(|X|I_A) < \epsilon$ for all A with $P(A) < \delta$.

(Forward Direction) Assume $E|X| < \infty$.

Since $E|X| < \infty$, by monotone convergence, $E(|X|I_{\{|X|>z\}}) \to 0$ as $z \to \infty$.

For any $\epsilon > 0$, choose z large enough so that $E(|X|I_{\{|X|>z\}}) < \frac{\epsilon}{2}$.

For any event A, we can write:

$$\begin{aligned} E(|X|I_A) &= E(|X|I_A I_{\{|X|\le z\}}) + E(|X|I_A I_{\{|X|>z\}}) \\ &\le z \cdot P(A) + E(|X|I_{\{|X|>z\}}) \\ &< z \cdot P(A) + \frac{\epsilon}{2} \end{aligned}$$

Now set $\delta = \frac{\epsilon}{2z}$. Then for any A with $P(A) < \delta$:

$$E(|X|I_A) < z \cdot \frac{\epsilon}{2z} + \frac{\epsilon}{2} = \epsilon$$

(Reverse Direction) Assume the condition holds.

For a fixed $\epsilon > 0$, there exists $\delta > 0$ such that $E(|X|I_A) < \epsilon$ whenever $P(A) < \delta$.

By Markov's inequality, $P(|X| > t) \le \frac{E|X|}{t}$, so $P(|X| > t) \to 0$ as $t \to \infty$ (regardless of whether $E|X|$ is finite).

Choose t_0 large enough that $P(|X| > t_0) < \delta$. Let $A = \{|X| > t_0\}$.

Then $E(|X|I_A) = E(|X|I_{\{|X|>t_0\}}) < \epsilon$.

Now we can show that $\int_{-y}^{y} |u|dF_X(u)$ forms a Cauchy sequence as $y \to \infty$:

For $y > t_0$:

$$\begin{aligned} \int_{-y}^{y} |u|dF_X(u) &\le \int_{-t_0}^{t_0} |u|dF_X(u) + \int_{\{|u|>t_0\}\cap\{|u|\le y\}} |u|dF_X(u) \\ &\le \int_{-t_0}^{t_0} |u|dF_X(u) + E(|X|I_{\{|X|>t_0\}}) \\ &< \int_{-t_0}^{t_0} |u|dF_X(u) + \epsilon \end{aligned}$$

This shows that $\int_{-y}^{y} |u|dF_X(u)$ converges as $y \to \infty$, implying that $E|X| < \infty$.

Therefore, $E|X| < \infty$ if and only if the given condition holds. $\square$

Problem 8

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, f be a non-negative simple function. Show that $\nu(A) := \int_A f d\mu$ defines a new measure.

Proof. Since f can be written as $f = \sum_{i=1}^n a_i I_{E_i}$ where $a_i \geq 0$ are constants and $E_i \in \mathcal{F}$ are measurable sets.

For any $A \in \mathcal{F}$, we have:

$$\begin{aligned}
\nu(A) &= \int_A f d\mu \\
&= \int_A \sum_{i=1}^n a_i I_{E_i} d\mu \\
&= \sum_{i=1}^n a_i \int_A I_{E_i} d\mu \\
&= \sum_{i=1}^n a_i \mu(A \cap E_i)
\end{aligned}$$

To prove ν is a measure, we need to verify three properties:

1. Non-negativity: For any $A \in \mathcal{F}$, $\nu(A) \geq 0$.

This is clear since $f \geq 0$, so $\int_A f d\mu \geq 0$ for any measurable set A.

2. Empty set: $\nu(\emptyset) = 0$.

We have:

$$\nu(\emptyset) = \int_\emptyset f d\mu = 0$$

since the integral over a set of measure zero is zero.

3. Countable additivity: If $\{A_j\}_{j=1}^\infty$ is a sequence of disjoint measurable sets, then $\nu\left(\bigcup_{j=1}^\infty A_j\right) = \sum_{j=1}^\infty \nu(A_j)$.

Let $\{A_j\}_{j=1}^\infty$ be a sequence of disjoint sets in $\mathcal{F}$. We have:

$$\begin{aligned}
\nu\left(\bigcup_{j=1}^\infty A_j\right) &= \int_{\bigcup_{j=1}^\infty A_j} f d\mu \\
&= \int_{\bigcup_{j=1}^\infty A_j} \sum_{i=1}^n a_i I_{E_i} d\mu \\
&= \sum_{i=1}^n a_i \int_{\bigcup_{j=1}^\infty A_j} I_{E_i} d\mu
\end{aligned}$$

$$= \sum_{i=1}^{n} a_i \mu \left(E_i \cap \bigcup_{j=1}^{\infty} A_j \right)$$

Since μ is a measure and $E_i \cap A_j$ are disjoint for different j, we have:

$$\begin{aligned} \mu \left(E_i \cap \bigcup_{j=1}^{\infty} A_j \right) &= \mu \left(\bigcup_{j=1}^{\infty} (E_i \cap A_j) \right) \\ &= \sum_{j=1}^{\infty} \mu(E_i \cap A_j) \end{aligned}$$

Therefore:

$$\begin{aligned} \nu \left(\bigcup_{j=1}^{\infty} A_j \right) &= \sum_{i=1}^{n} a_i \sum_{j=1}^{\infty} \mu(E_i \cap A_j) \\ &= \sum_{j=1}^{\infty} \sum_{i=1}^{n} a_i \mu(E_i \cap A_j) \\ &= \sum_{j=1}^{\infty} \int_{A_j} \sum_{i=1}^{n} a_i I_{E_i} d\mu \\ &= \sum_{j=1}^{\infty} \int_{A_j} f d\mu \\ &= \sum_{j=1}^{\infty} \nu(A_j) \end{aligned}$$

We have verified all three properties, so $\nu(A) = \int_A f d\mu$ defines a measure on $(\Omega, \mathcal{F})$.

□