Problem 1. **Linearity of Expectation, Independence**

Let $X_1, X_2, X_3$ be discrete independent random variables with mean 0. Find $E[(X_1 + X_2)(X_2 + X_3)(X_3 + X_1)]$.

**Solution 1.** $E[(X_1 + X_2)(X_2 + X_3)(X_3 + X_1)]$ contains terms of the form $E(X_i^2X_j), i \neq j$ and $E(X_iX_jX_k), i \neq j \neq k$. Since $X_1, X_2, X_3$ are independent and zero mean, $E(X_i^2X_j) = E(X_i^2)E(X_j) = 0$, and from the same logic, $E(X_iX_jX_k) = 0$. Hence, $E[(X_1 + X_2)(X_2 + X_3)(X_3 + X_1)] = 0$.

Problem 2. **Poisson Packet Routing**

Packets arriving at a switch are routed to either destination $A$ (with probability $p$) or destination $B$ (with probability $1-p$). The destination of each packet is chosen independently of each other. In the time interval $[0,1]$, the number of arriving packets is Pois($\lambda$).

1. Show that the number of packets routed to $A$ is Poisson distributed. With what parameter?

2. Are the number of packets routed to $A$ and to $B$ independent?

**Solution 2.** 1. Let $X, Y$ be random variables which are equal to the number of packets routed to the destinations $A, B$ respectively. Let $Z = X + Y$. We are given that $Z \sim$ Pois($\lambda$). We prove that $X$ has the Poisson distribution with mean $p\lambda$.

\[
\Pr(X = x) = \sum_{z=x}^{\infty} \Pr(X = x, Z = z) \\
= \sum_{z=x}^{\infty} \Pr(Z = z) \Pr(X = x \mid Z = z) \\
= \sum_{z=x}^{\infty} \frac{e^{-\lambda} \lambda^z}{z!} \frac{\left(\frac{z}{x}\right) P^x (1-p)^{z-x}}{x!} \\
= e^{-\lambda} \sum_{z=x}^{\infty} \frac{\lambda^z}{z!} \frac{1}{x!} \frac{1}{(z-x)!} \frac{1}{(z-x)!} \frac{P^x (1-p)^{z-x}}{x!} \\
= e^{-\lambda} \frac{(\lambda P)^x}{x!} \sum_{z=x}^{\infty} \frac{(\lambda (1-p))^{z-x}}{(z-x)!} \\
= e^{-\lambda} \frac{(\lambda p)^x}{x!} e^{\lambda (1-p)} \\
= \frac{e^{-\lambda p} (\lambda p)^x}{x!}.
\]
2. We prove that $X$ and $Y$ are independent.

$$\Pr(X = x, Y = y) = \sum_{z=0}^{\infty} \Pr(X = x, Y = y, Z = z)$$

$$= \sum_{z=0}^{\infty} \Pr(X = x, Y = y \mid Z = z) \Pr(Z = z)$$

$$= \Pr(X = x, Y = y \mid Z = x + y) \Pr(Z = x + y)$$

$$= \frac{(x+y)!}{x!y!} p^x (1 - p)^y \frac{e^{-\lambda} \lambda^{x+y}}{(x+y)!}$$

$$= e^{-\lambda} (\lambda p)^x \frac{\lambda^{x+y}}{y!}$$

$$= \Pr(X = x) \Pr(Y = y).$$

**Problem 3. Poisson Merging**

The Poisson distribution is used to model rare events, such as the number of customers who enter a store in the next hour. The theoretical justification for this modeling assumption is that the limit of the binomial distribution, as the number of trials $n$ goes to $\infty$ and the probability of success per trial $p$ goes to 0, such that $np \to \lambda > 0$, is the Poisson distribution with mean $\lambda$.

Now, suppose we have two independent streams of rare events (for instance, the number of female customers and male customers entering a store), and we do not care to distinguish between the two types of rare events. Can the combined stream of events be modeled as a Poisson distribution?

Mathematically, let $X$ and $Y$ be independent Poisson random variables with means $\lambda$ and $\mu$ respectively. Prove that $X + Y \sim \text{Pois}(\lambda + \mu)$. (This is known as Poisson merging.) Note that it is not sufficient to use linearity of expectation to say that $X + Y$ has mean $\lambda + \mu$. You are asked to prove that the distribution of $X + Y$ is Poisson.

**Note:** This property will be extensively used when we discuss Poisson processes.

**Solution 3.** For $z \in \mathbb{N},$

$$\Pr(X + Y = z) = \sum_{j=0}^{z} \Pr(X = j, Y = z-j) = \sum_{j=0}^{z} \frac{e^{-\lambda} \lambda^j e^{-\mu} \mu^{z-j}}{j! (z-j)!}$$

$$= \frac{e^{-\lambda} \lambda^z}{z!} \sum_{j=0}^{z} \frac{z!}{j! (z-j)!} \frac{1}{\lambda^j \mu^{z-j}}$$

$$= \frac{e^{-\lambda} \lambda^z}{z!} \sum_{j=0}^{z} \binom{z}{j} \frac{(\lambda)^j (\mu)^{z-j}}{z!}$$

Here is some intuition for why Poisson merging holds. If we are interested in the number of customers entering a store in the next hour, we can discretize the hour into $n$ time intervals, where $n$ is a positive integer. In each time interval, independently of other time intervals, the probability that a female customer enters the store is $\lambda/n$ and the probability that a male customer enters the store is $\mu/n$. Since the two types of customers are assumed to be independent, the probability that a
customer, disregarding gender, enters the store is \( \frac{\lambda}{n} + \frac{\mu}{n} - \frac{\lambda\mu}{n^2} \). As \( n \to \infty \), the number of customers who enter the store in the hour is Poisson with mean \( \lim_{n \to \infty} n[\frac{\lambda}{n} + \frac{\mu}{n} - \frac{\lambda\mu}{n^2}] = \lambda + \mu \).

We will be able to give a much easier proof of this result after we introduce transforms of random variables.

**Problem 4. [Extra] Clustering Coefficient**

This problem will explore an important probabilistic concept of clustering that is widely used in machine learning applications today. Consider \( n \) students, where \( n \) is a positive integer. For each pair of students \( i, j \in \{1, \ldots, n\}, i \neq j \), they are friends with probability \( p \), independently of other pairs. We assume that friendship is mutual. We can see that the friendship among the \( n \) students can be represented by an undirected graph \( G \). Let \( N(i) \) be the number of friends of student \( i \) and \( T(i) \) be the number of triangles attached to student \( i \). We define the clustering coefficient \( C(i) \) for student \( i \) as follows:

\[
C(i) = \frac{T(i)}{\left(\begin{array}{c} N(i) \\ 2 \end{array}\right)}.
\]

The clustering coefficient is not defined for the students who have no friends. An example is shown in Figure [1]. Student 3 has 4 friends (1, 2, 4, 5) and there are two triangles attached to student 3, i.e., triangle 1-2-3 and triangle 2-3-4. Therefore \( C(3) = \frac{2}{\left(\begin{array}{c} 4 \\ 2 \end{array}\right)} = \frac{1}{3} \).

Find \( \mathbb{E}[C(i) \mid N(i) \geq 2] \).

**Solution 4.** First, we compute \( \mathbb{E}[C(i) \mid N(i) = k] \), for \( k \in \{2, \ldots, n - 1\} \). Suppose that student \( i \) has friends \( f_1, \ldots, f_k \). We can see that \( T(i) \) equals the number of friend pairs among \( \{f_1, \ldots, f_k\} \). Since there are \( \left(\begin{array}{c} k \\ 2 \end{array}\right) \) possible pairs and each pair of students are friends with probability \( p \), independently of other pairs, we know that the expected number of friend pairs among \( \{f_1, \ldots, f_k\} \) is \( \left(\begin{array}{c} k \\ 2 \end{array}\right) p \). Then we have

\[
\mathbb{E}[C(i) \mid N(i) = k] = \frac{\left(\begin{array}{c} k \\ 2 \end{array}\right) p}{\left(\begin{array}{c} k \\ 2 \end{array}\right)} = p.
\]

Since this is true for all \( k \geq 2 \), we have \( \mathbb{E}[C(i) \mid N(i) \geq 2] = p \).

![Figure 1: Friendship and clustering coefficient.](image)