Lecture 13

§1 Compound Poisson Process

1. **Definition:** Compound Poisson distribution（复合泊松分布）

一个随机变量 X 被称为 compound Poisson random variable，若其满足

$$X = \sum_{i=1}^{N} Y_i$$

其中 $N \sim Poi(\lambda)$ 为一个 Poisson random variable，

$\{Y_i, i \geq 1\}$ 为一个独立于 N 的 a family of i.i.d. random variable

注：① 这是 random sum 的一个特例

② 由 conditioning method，有

- $E[X] = \lambda \cdot E(Y_1)$
- $Var[X] = \lambda E[Y_1^2]$

证明：（利用全期望公式）

- $E[X] = E[E[X|N]] = E[\sum_{i=1}^{N} E(Y_i|N)] = E[N E(Y_1)] = \lambda \cdot E(Y_1)$
- $Var[X] = E[X^2] - E[X]^2$

$$= E[E[X^2|N]] - \lambda^2 \cdot E(Y_1)^2$$

$$E[X^2|N] = E[\sum_{i=1}^{N} Y_i^2 + 2\sum_{i<j}^{N} Y_i Y_j \mid N]$$

$$= \sum_{i=1}^{N} E(Y_i^2) + 2\sum_{i<j}^{N} E(Y_i)E(Y_j)$$

$$= N \cdot E(Y_1^2) + (N^2 - N) E(Y_1)^2$$

$$Var[X] = E[N \cdot E(Y_1^2) + (N^2-N) E(Y_1)^2] - \lambda^2 \cdot E(Y_1)^2$$

$$= \lambda E(Y_1^2) + [\underbrace{Var(N)}_{\lambda} + \underbrace{E(N)^2}_{\lambda^2} - \underbrace{E(N)}_{\lambda}] E(Y_1^2) - \lambda^2 \cdot E(Y_1)^2$$

$$= \lambda E(Y_1^2)$$

2. **Definition:** Compound Poisson Process（复合泊松过程）

$\{X(t): t \geq 0\}$ 被称为 compound Poisson process，若其满足

$$X_t = \sum_{i=1}^{N_t} Y_i, \quad t \geq 0$$

其中 N_t 为一个 Poisson process，

$\{Y_i, i \geq 1\}$ 为一个独立于 $\{N_t: t \geq 0\}$ 的 a family of i.i.d. random variable

注：① 若 $Y_i = 1$，则 $X_t = N_t$，因此 Poisson process 可视作 compound Poisson process 的特例

② 对于任意 fixed time t，X_t 可视作一个 compound Poisson random variable，有

- $E[X_t] = \lambda t \cdot E(Y_1)$
- $Var[X_t] = \lambda t \cdot E[Y_1^2]$

证明：（利用全方差公式）

$$X_t = \sum_{i=1}^{N_t} Y_i, \quad t \geq 0$$

$$Var[X_t] = E[Var(X_t|N_t)] + Var(E[X_t|N_t])$$

$$= E\left[\sum_{i=1}^{N_t} Var(Y_i \mid N_t)\right] + Var\left(\sum_{i=1}^{N_t} E[Y_i]\right)$$
$$= E[N_t \cdot Var(Y_1)] + E(Y_1)^2 \, Var(N_t)$$
$$= Var(Y_1) \cdot \lambda t + E(Y_1)^2 \cdot \lambda t$$
$$= \lambda t \cdot E[Y_1^2]$$

e.g. ▶ Suppose that buses arrive at a sporting event in accordance with a Poisson process, and suppose that the numbers of fans in each bus are assumed to be i.i.d.
Then $\{X_t, t \geq 0\}$ is a compound Poisson process, where N_t represents the number of buses arriving at the sporting event by time t, Y_i represents the number of fans in the ith bus, and $X_t = \sum_{i=1}^{N_t} Y_i$ denotes the number of fans who have arrived by time t.

e.g. ▶ Suppose customers leave a supermarket in accordance with a Poisson process. If Y_i, the amount spent by the ith customer, $i = 1, 2, \ldots$, are i.i.d.
Then $\{X_t, t \geq 0\}$ is a compound Poisson process, where N_t denotes the number of customers leaving the supermarket by time t and $X_t = \sum_{i=1}^{N_t} Y_i$ denotes the total amount of money spent by time t.

例1: Suppose that

- ▶ families migrate to an area at a Poisson rate $\lambda = 2$ per week;
- ▶ the number of people in each family is independent and takes on the values of 1, 2, 3, 4 with respective probabilities $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{3}$, $\frac{1}{6}$.

Question.

1. What is the expected value and variance of the number of individuals migrating to this area during a fixed five-week period?
2. Find the approximate probability that at least 240 people migrate to the area within the next 50 weeks.

① 由题意可知, $X_t = \sum_{i=1}^{N_t} Y_i$

$$Y_i = \begin{cases} 1 & w.p. \ 1/6 \\ 2 & w.p. \ 1/3 \\ 3 & w.p. \ 1/3 \\ 4 & w.p. \ 1/6 \end{cases}$$

则 $E[X_5] = \lambda t \, E[Y_i] = 2 \cdot 5 \cdot 2.5 = 25$

$Var(X_5) = \lambda t \, E[Y_i^2] = 2 \cdot 5 \cdot \frac{43}{6} = \frac{215}{3}$

② 由 CLT, $X_{50} = \sum_{i=1}^{N_{50}} Y_i \overset{approx.}{\sim} N(E[X_{50}], Var(X_{50}))$

$E[X_{50}] = \lambda t \, E[Y_i] = 2 \cdot 50 \cdot 2.5 = 250$

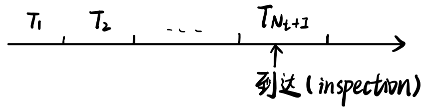
$Var(X_{50}) = \lambda t \, E[Y_i^2] = 2 \cdot 50 \cdot \frac{43}{6} = \frac{2150}{3}$

因此,

$$\frac{X_{50} - E[X_{50}]}{\sqrt{Var(X_{50})}} = \frac{X_{50} - 250}{\sqrt{2150/3}} \overset{approx.}{\sim} N(0, 1)$$

$$P(X_{50} \geq 240) \approx P(X_{50} \leq 260) \approx \Phi\left(\frac{260-250}{\sqrt{2150/3}} \approx 0.3735\right) \approx 0.6456$$

注*: inspection paradox: 平均而言, 坐上的公交车总是比较"磨蹭"



令 T_i 表示 i th interarrival time

$$
\begin{aligned}
E[T_{N_t+1}] &= E(\text{waiting time since last arrival}) + E(\text{waiting time until next arrival}) \\
&= E(\text{waiting time since last arrival}) + 1/\lambda \\
&> \frac{1}{\lambda}
\end{aligned}
$$