AI2613 随机过程 2020–2021 春季学期

Lecture 7 – Poisson Process (I)

2021 年 4 月 12 日

Lecturer: 张驰豪 *Scribe*: 杨宽

1 Poisson Distribution

Example 1. Suppose that there exists a restaurant. How can we predict the nubmer of tomorrow's customers based on the number of customers in the past several days?
For instance, we assume that the number of customers in the past five days are: 100, 120, 80, 75 and 110. A natural idea is to use the average number (e.g., 97 in our instance) of the past. However, with probability 1/2 or even greater, the restaurant may not prepare sufficient food.

To analyze the distribution of the number of customers, we should make some assumptions first. Assume that there are n slots in a day. Every slot is sufficiently small s.t. at most one customer comes into the restaurant in a slot and the probability of coming in each slot is p independently of each other. Now let's compute the distribution of the number of customers X_n (where we denote $p \cdot n$ by λ):

$$\begin{aligned}\mathbf{Pr}[X_n = k] &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\ &= \binom{n}{k} \cdot \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \xrightarrow{n \to \infty} \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}.\end{aligned}$$

Here we use the facts that $\binom{n}{k}/n^k \to 1$, $(1-\lambda/n)^n \to \mathrm{e}^{-\lambda}$ and $(1-\lambda/n)^k \to 1$ as $n \to \infty$ and k is a *constant*. So X_n has a *Poisson distribution*.

Definition 2 (Poisson Distribution). A random variable X is said to have a *Poisson distribution* with mean λ, or $X \sim \mathrm{Pois}(\lambda)$, if

$$\mathbf{Pr}[X = k] = \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda}.$$

We can verify that the expectation of a Poisson with mean λ is indeed λ:

$$\mathbb{E}_{X \sim \mathrm{Pois}(\lambda)}[X] = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)} \cdot \mathrm{e}^{-\lambda} = \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \cdot \mathrm{e}^{-\lambda} = \lambda.$$

Remark. Let λ be a fixed *constant*. A Poisson distribution $\mathsf{Pois}(\lambda)$ is the limit of binomial distributions $\mathsf{Binom}(n, \lambda/n)$ as $n \to \infty$. If n is sufficiently large, $\mathsf{Binom}(n, \lambda/n)$ is approximately $\mathsf{Pois}(\lambda)$.

Question. What is the distribution of two days' customers?

Fact 1. *Suppose that $X_1 \sim \mathsf{Pois}(\lambda_1)$ and $X_2 \sim \mathsf{Pois}(\lambda_2)$ are independent. Then*

$$X_1 + X_2 \sim \mathsf{Pois}(\lambda_1 + \lambda_2).$$

Proof. We calculate the distribution of $X_1 + X_2$ directly:

$$\begin{aligned}
\mathbf{Pr}[X_1 + X_2 = n] &= \sum_{m=0}^{n} \mathbf{Pr}[X_1 = m \wedge X_2 = n - m] \\
&= \sum_{m=0}^{n} \mathbf{Pr}[X_1 = m] \cdot \mathbf{Pr}[X_2 = n - m] \\
&= \sum_{m=0}^{n} \frac{\lambda_1^m}{m!} \cdot \mathrm{e}^{-\lambda_1} \cdot \frac{\lambda_2^{n-m}}{(n-m)!} \cdot \mathrm{e}^{-\lambda_2} \\
&= \mathrm{e}^{-(\lambda_1+\lambda_2)} \sum_{m=0}^{n} \frac{\lambda_1^m \lambda_2^{n-m}}{n!} \cdot \binom{n}{m} \\
&= \frac{(\lambda_1 + \lambda_2)^n}{n!} \cdot \mathrm{e}^{-(\lambda_1+\lambda_2)}.
\end{aligned}$$

□

It is easy to extend the fact to a sequence of independent Poissons and yield the following proposition.

Proposition 2. *Suppose that $X_1, X_2, \ldots, X_n$ are n mutually independent random variables, where $X_i \sim \mathsf{Pois}(\lambda_i)$. Then*

$$\sum_{i=1}^{n} X_i \sim \mathsf{Pois}\Big(\sum_{i=1}^{n} \lambda_j\Big).$$

In particular, if $X_1, X_2, \ldots, X_n$ are i.i.d. Poissons with mean λ, then $\sum_{i=1}^{n} X_i \sim \mathsf{Pois}(n\lambda)$.

2 Defining the Poisson Process

Note that in Proposition 2, n is not necessary an integer. Now we introduce the Poisson process.

Definition 3 (Poisson Process). A *Poisson process* $\{N(s) \colon s \geq 0\}$ with rate λ satisfies that

1. $N(0) = 0$;

2. $\forall\, t, s \geq 0$, $N(t+s) - N(s) \sim \mathsf{Pois}(\lambda \cdot t)$;
3. $\forall\, t_0 \leq t_1 \leq \cdots \leq t_n$, $N(t_1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$ are mutually independent.

In fact, the Poisson process has another *constructive* definition. We first recall the exponential distribution mentioned in the first lecture.

Definition 4 (Exponential distribution). The probability density function of the exponential distribution with rate $\lambda > 0$ is given by

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0; \\ 0, & \text{otherwise}. \end{cases}$$

So the cumulative probability function of $X \sim \mathsf{Exponential}(\lambda)$ is

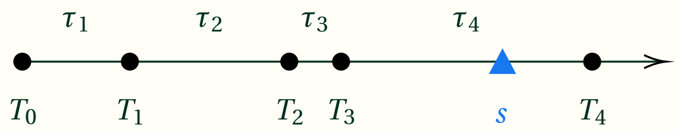
$$F_X(x) = \mathbf{Pr}[X \leq k] = \int_{-\infty}^{k} f(x)\, \mathrm{d}x = 1 - \mathrm{e}^{-\lambda k}.$$

Then the following proposition gives another definition of the Poisson process.

Proposition 3. *Suppose that $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ is a sequence of independent random variables that each r.v. has an exponential distribution with rate λ (i.e., $\tau_i \sim \mathsf{Exponential}(\lambda)$). Let $T_n = \sum_{i=1}^{n} \tau_i$ and*

$$N(s) \triangleq \max\{n \colon T_n \leq s\}.$$

Then $N(s)$ is a Poisson process with rate λ.



Before we prove this proposition (and the equivalence of the two definitions), we are going to discuss some properties of the exponential distribution.

Fact 4. *Let $X \sim \mathsf{Exponential}(\lambda)$. Then $\mathbb{E}[X] = 1/\lambda$.*

Proof. We calculate the expectation directly.

$$\begin{aligned} \mathbb{E}[X] &= \int_0^{\infty} t \cdot \lambda \mathrm{e}^{-\lambda t}\, \mathrm{d}t \; = -\int_0^{\infty} t\, \mathrm{d}\mathrm{e}^{-\lambda t} \\ &= \left(-t \cdot \mathrm{e}^{-\lambda t}\right)\Big|_0^{\infty} + \int_0^{\infty} \mathrm{e}^{-\lambda t}\, \mathrm{d}t \\ &= -\frac{1}{\lambda} \cdot \left(\mathrm{e}^{-\lambda t}\right)\Big|_0^{\infty} \; = -1/\lambda. \end{aligned}$$

□

Remark. We think of the τ_n as times between arrivals of customers at the restaurant, so $T_n = \tau_1 + \cdots + \tau_n$ is the arrival time of the n-th customer, and $N(s)$ is the number of arrivals by time s. So $1/\lambda$ measures the average of times between arrivals — the average of τ_i is $1/\lambda$.
Note that in Definition 3, the rate λ measures the average increments of arrivals in units of time. Intuitively, the quantity measured by λ in Proposition 3 is consistent with the quantity measured by λ in Definition 3.

Fact 5. *Let* $X \sim \text{Exponential}(\lambda)$*. Then* $\mathbb{E}[X^2] = 2/\lambda^2$.

Proof. Again, we calculate it directly.

$$\begin{aligned}\mathbb{E}[X^2] &= \int_0^\infty t^2 \cdot \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t = -\int_0^\infty t^2\,\mathrm{d}\mathrm{e}^{-\lambda t} \\ &= \left(-t^2 \cdot \mathrm{e}^{-\lambda t}\right)\Big|_0^\infty + \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t^2 \\ &= 2\int_0^\infty t \cdot \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \mathbb{E}[X] \cdot 2/\lambda = 2/\lambda^2.\end{aligned}$$

□

Corollary 6. *Let* $X \sim \text{Exponential}(\lambda)$*. Then* $\text{Var}[X] = 1/\lambda^2$.

Moreover, the exponential distribution has the following property (that may be a bit surprising). This property somewhat explains the *mutual independence* in Definition 3.

Proposition 7 (Lack of Memory Property). *Let* $X \sim \text{Exponential}(\lambda)$*. Then for all* $t, s > 0$,

$$\mathbf{Pr}[X > t + s \mid X > s] = \mathbf{Pr}[X > t]\,.$$

Proof. It is easy to verify that

$$\begin{aligned}\mathbf{Pr}[X > t + s \mid X > s] &= \frac{\mathbf{Pr}[X > t + s \wedge X > s]}{\mathbf{Pr}[X > s]} = \frac{\mathbf{Pr}[X > t + s]}{\mathbf{Pr}[X > s]} \\ &= \frac{\mathrm{e}^{-\lambda(t+s)}}{\mathrm{e}^{-\lambda s}} = \mathrm{e}^{-\lambda t} = \mathbf{Pr}[X > t]\,.\end{aligned}$$

□

Now we introduce another property of the exponential distribution. It will be useful in the next lecture.

Proposition 8 (Exponential Races). *Let* $X_1 \sim \text{Exponential}(\lambda_1)$ *and* $X_2 \sim \text{Exponential}(\lambda_2)$ *be two independent random variables. Then* $Y \triangleq \min\{X_1, X_2\} \sim \text{Exponential}(\lambda_1 + \lambda_2)$.

Proof. Using the independence we can have

$$\begin{aligned}\mathbf{Pr}[Y > t] &= \mathbf{Pr}[\min\{X_1, X_2\} > t] \\ &= \mathbf{Pr}[X_1 > t \wedge X_2 > t] \\ &= \mathbf{Pr}[X_1 > t] \cdot \mathbf{Pr}[X_2 > t] \\ &= \mathrm{e}^{-\lambda_1 t} \cdot \mathrm{e}^{-\lambda_2 t} = \mathrm{e}^{-(\lambda_1 + \lambda_2)t}.\end{aligned}$$

□

Corollary 9. *Let $X_1, X_2, \ldots, X_n$ be n mutually independent random variables where X_i has an exponential distribution with rate λ_i. Then $Y \triangleq \min\{X_1, X_2, \ldots, X_n\}$ has an exponential distribution with rate $\lambda_1 + \lambda_2 + \cdots + \lambda_n$.*

We now consider the problem "who wins the race?". We first assume that there are only two random variables. Using the law of total probability, we can compute the probability that X_1 wins the race as follows:

$$\begin{aligned}
\mathbf{Pr}[X_1 \le X_2] &= \int_0^\infty \mathbf{Pr}[X_1 = t \wedge X_2 \ge t]\,\mathrm{d}t \\
&= \int_0^\infty \mathbf{Pr}[X_1 = t] \cdot \mathbf{Pr}[X_2 \ge t]\,\mathrm{d}t \\
&= \int_0^\infty \lambda_1 \mathrm{e}^{-\lambda_1 t} \cdot \mathrm{e}^{-\lambda_2 t}\,\mathrm{d}t \\
&= \lambda_1 \int_0^\infty \mathrm{e}^{-(\lambda_1 + \lambda_2) t}\,\mathrm{d}t \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}.
\end{aligned}$$

Thus, clearly, the probability that X_i wins the race among n random variables is $\frac{\lambda_i}{\lambda_1+\lambda_2+\cdots+\lambda_n}$.

3 Proof of Proposition 3

Now we are ready to prove Proposition 3. We first consider the distribution of T_n.

Theorem 10. *Let $\tau_1, \tau_2, \ldots, \tau_n$ be n mutually independent random variables where each τ_i has an distribution with rate λ. Then $T_n \triangleq \tau_1 + \cdots + \tau_n$ has a gamma distribution* Gamma(n, λ)*, where the probability density function of* Gamma(n, λ) *is given by*

$$f_{n,\lambda}(t) = \begin{cases} \lambda \mathrm{e}^{-\lambda t} \cdot \frac{(\lambda t)^{n-1}}{(n-1)!}, & t \ge 0; \\ 0, & t < 0. \end{cases}$$

Proof. We prove it by induction.

Suppose that $n = 1$. Then $T_1 = \tau_1 \sim$ Exponential(λ). On the other hand, the probability density function of Gamma$(1, \lambda)$ is

$$f_{1,\lambda}(t) = \lambda \mathrm{e}^{-\lambda t} \text{ for } t \ge 0,$$

which is the same as the probability function of Exponential(λ).

Suppose that $T_n \sim \text{Gamma}(n, \lambda)$ for some $n \geq 1$. We now consider the distribution of T_{n+1}. Using the independence of T_n and τ_{n+1}, we have

$$\begin{aligned}
\mathbf{Pr}[T_{n+1} = t] &= \int_0^t \mathbf{Pr}[T_n = s] \cdot \mathbf{Pr}[\tau_{n+1} = t - s] \, \mathrm{d}s \\
&= \int_0^t f_{n,\lambda}(t) \cdot \lambda \mathrm{e}^{-\lambda(t-s)} \, \mathrm{d}s \\
&= \int_0^t \lambda \mathrm{e}^{-\lambda s} \cdot \frac{(\lambda s)^{n-1}}{(n-1)!} \cdot \lambda \mathrm{e}^{-\lambda(t-s)} \, \mathrm{d}s \\
&= \lambda \mathrm{e}^{-\lambda t} \cdot \frac{\lambda^n}{(n-1)!} \cdot \int_0^t s^{n-1} \, \mathrm{d}s \\
&= \lambda \mathrm{e}^{-\lambda t} \cdot \frac{\lambda^n}{(n-1)!} \cdot \frac{t^n}{n} = f_{n+1,\lambda}(t) \, .
\end{aligned}$$

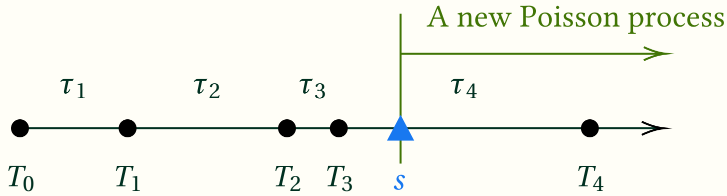
□

Next, we are going to give a proof of Proposition 3.

Proof of Proposition 3. Let $N(s)$ be the random variable constructed in Proposition 3. It is clear that $N(0) = 0$ with probability 1. So we begin our proof by verifying that $N(t) \sim \text{Pois}(\lambda t)$:

$$\begin{aligned}
\mathbf{Pr}[N(t) = n] &= \mathbf{Pr}[T_n \leq t \wedge T_{n+1} > t] \\
&= \int_0^t \mathbf{Pr}[T_n = s] \cdot \mathbf{Pr}[\tau_{n+1} > t - s] \, \mathrm{d}s \\
&= \int_0^t \lambda \mathrm{e}^{-\lambda s} \cdot \frac{(\lambda s)^{n-1}}{(n-1)!} \cdot \mathrm{e}^{-\lambda(t-s)} \, \mathrm{d}s \\
&= \lambda \mathrm{e}^{-\lambda t} \cdot \frac{\lambda^{n-1}}{(n-1)!} \int_0^t s^{n-1} \, \mathrm{d}s \\
&= \lambda^n \mathrm{e}^{-\lambda t} \cdot \frac{t^n}{n!} \, .
\end{aligned}$$

Thus $N(t) \sim \text{Pois}(\lambda t)$. For $s > 0$, we would like to show that $N(t+s) - N(s) \sim \text{Pois}(\lambda t)$.



A Poisson process

Let $n = N(s)$. Using the lack of memory property, it is clear that for any $t > 0$,

$$\mathbf{Pr}[\tau_{n+1} > s + t - T_n \mid \tau_{n+1} > s - T_n] = \mathbf{Pr}[\tau_{n+1} > t] = \mathbf{Pr}[\tau_1 > t] \, .$$

So the distribution of the first arrival after s is the same as Exponential(λ). Applying the independence of $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ we conclude that the process starting from time s has the same distribution as the original process. That is, $N(t+s) - N(s) \sim \text{Pois}(\lambda t)$.

Furthermore, it is easy to see that $N(t+s) - N(s)$ is independent of $N(r)$ for all $r \leq s$ since for any s, $N(t+s) - N(s) \sim \text{Pois}(\lambda t)$. It implies that $N(s)$ has independent increments, and hence completes our proof of Proposition 3. □