STA447/STA2006 Stochastic Processes

Gun Ho Jang

Lecture on February 6, 2014

Note

This lecture note is prepared for the course STA447/STA2006 Stochastic Processes. This lecture note may contain flaws. Please consult text book or reference books for confidence.

- Gun Ho Jang

* indicates graduate level. So you may skip those parts.

3 Poisson Process

3.1 Exponential Distribution

The exponential distribution with rate parameter $\lambda > 0$ is $F(x) = 1 - e^{-\lambda x}$ for $x > 0$. Hence the density is $\lambda \exp(-\lambda x)$ for $x > 0$. Let $X \sim \mathrm{Exp}(\lambda)$. Then $\mathbb{E}X = \int_0^\infty x\lambda e^{-\lambda x}\,dx = 1/\lambda$ and $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \int_0^\infty x^2\lambda e^{-\lambda x}\,dx - \lambda^{-2} = 1/\lambda^2$. The moment generating function and characteristic function are

$$\mathrm{mgf}_X(t) = \mathbb{E}e^{tX} = \int_0^\infty e^{tx}\lambda e^{-\lambda x}\,dx = \frac{\lambda}{\lambda - t}, \qquad \mathrm{chf}_X(t) = \mathbb{E}e^{itX} = \mathrm{mgf}_X(it) = \frac{\lambda}{\lambda - it}.$$

Lack of memory property Let $X \sim \mathrm{Exp}(\lambda)$. $P(X > t + s \,|\, X > t) = P(X > t + s)/P(X > t) = e^{-\lambda(t+s)}/e^{-\lambda t} = e^{-\lambda s} = P(X > s)$.

Exercise 25. Show the lack of memory property for the geometric random variables.

Theorem 51. Let $X_i \sim i.i.d.\ \mathrm{Exp}(\lambda)$. Then $S_n = X_1 + \cdots + X_n \sim \mathrm{Gamma}(n, \lambda)$, that is, $\mathrm{pdf}_{S_n}(x) = (\lambda^n/\Gamma(n))x^{n-1}\exp(-\lambda x)$ for $x > 0$.

Proof. Note that $\mathrm{mgf}_{S_n}(t) = \mathbb{E}e^{tS_n} = \mathbb{E}e^{t(X_1+\cdots+X_n)} = \mathbb{E}e^{tX_1} \times \cdots \times \mathbb{E}e^{tX_n} = [\mathbb{E}e^{tX_1}]^n = [\mathrm{mgf}_{X_1}(t)]^n = (1 - t/\lambda)^{-n}$ which is the MGF of $\mathrm{Gamma}(n, \lambda)$. □

Note. Let $X \sim \mathrm{Gamma}(\alpha, \beta)$. Then, $\mathrm{pdf}_X(x) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}e^{-\beta x}$ for $x > 0$, $\mathrm{mgf}_X(t) = (1 - t/\beta)^{-\alpha}$, $\mathrm{chf}_X(t) = (1 - it/\beta)^{-\alpha}$.

3.2 Homogeneous Poisson Processes

Question: How can we model arrival times of customers in a coffee shop?

Let $N(t)$ be the number of customers arrived in a coffee shop regardless of the service status (served or in the queue). Obviously at the opening moment, there is no customer, that is, $N(0) = 0$. It is easy to assume that the number of customers arrived between 0 and s do not affect on the number of customers arrived after s, that is, $N(t) - N(s)$ and $N(s) - N(0)$ are independent. In general, for times $0 \le t_0 < t_1 < \cdots < t_k$, the difference process $N(t-1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_k) - N(t_{k-1})$ are independent. This property is called *independent increment*.

Also we may assume the arrival distribution between time s and t is only dependant on the time gap $t - s$. Which is called *stationary increment*.

Definition 33. A homogeneous Poisson process $N(t)$ with rate λ is a continuous time non-negative valued stochastic process satisfying
(a) $N(0) = 0$,
(b) [Independent increment] For $t_1 < t_2 \le t_3 < t_4$, $N(t_2) - N(t_1), N(t_4) - N(t_3)$ are independent.
(c) [Stationary increment] The distribution of $N(t) - N(s)$ depends only on the length $t - s$.
(d) [Poisson distribution] $N(t) - N(s) \sim \text{Poisson}(\lambda(t-s))$ for any $0 \le s < t$.

Note. The condition (d) implies the condition (c). So the condition (c) can be dropped in the definition.

Theorem 52. A homogeneous Poisson process $N(t)$ with rate λ has the Markov property.

Proof. Let $0 \le t_1 < \cdots < t_k < t$. We will show that, for given $N(t_1), \ldots, N(t_k)$, the distribution of $N(t)$ only depend on $N(t_k)$. Using the independent increment, $N(t) - N(t_k)$ is independent from $N(t_1), N(t_2) - N(t_1), \ldots, N(t_k) - N(t_{k-1})$. Hence for $0 \le m_1 \le \cdots \le m_k \le m$, $P(N(t) = m \,|\, N(t_1) = m_1, \ldots, N(t_k) = m_k) = P(N(t) - N(t_k) = m - m_k \,|\, N(t_k) = m_k) = e^{-\lambda(t-t_k)}[\lambda(t-t_k)]^{m-m_k}/(m-m_k)!$ only depends on $N(t_k)$. Therefore the theorem follows. □

Note. A homogeneous Poisson process is a continuous time Markov chain having countably many states.

Exercise 26. Show that any process having independent increment also satisfies the Markov property.

Definition 34. The k-th arrival time T_k is defined by $T_k = \inf\{t \ge 0 : N(t) = k\}$ for any $k \ge 1$. The k-th *interarrival time* is the time gap between $(k-1)$-th and k-th arrival, that is, $\tau_k = T_k - T_{k-1}$.

Exercise 27. Show that T_k are stopping times.

Since $N(t)$ is a homogeneous Markov chain, the process $N(T_k + t) - N(T_k)$ does not depend on $N(T_k)$ and it behaves the save homogeneous Markov chain started from $N(0) = 0$. Hence $T_1, T_2 - T_1, \ldots, T_k - T_{k-1}$ are i.i.d.

Proposition 53. $T_1 \sim \text{Exp}(\lambda)$.

Proof. For any $x > 0$, $P(T_1 > x) = P(N(x) = 0) = e^{-\lambda x}(\lambda x)^0/0! = e^{-\lambda x}$. Thus the density function becomes $\text{pdf}_{T_1}(x) = \frac{d}{dx}(1 - e^{-\lambda x}) = \lambda e^{-\lambda x}$ which is the density of $\text{Exp}(\lambda)$. □

Exercise 28. Show that $T_k \sim \text{Gamma}(k, \lambda)$.

Exercise 29. Let X_k be independent Poisson(μ_k). Show that $X_1 + \cdots + X_n \sim \text{Poisson}(\mu_1 + \cdots + \mu_k)$.

Example 40 (Exercise 2.22)**.** Let $N(t)$ be a Poisson process with rate λ. Let T_k be the k-th arrival time. Note that the interarrival time $\tau_1, \tau_2, \ldots$ are i.i.d. and $T_k = \tau_1 + \cdots + \tau_k$. Hence $\mathbb{E}T_k = k\mathbb{E}T_1 = k/\lambda$ because $T_1 \sim \text{Exp}(\lambda)$.

For any $0 < m < n$ and $k > 0$, given $N(k) = m$, the distributions of $N(k+t) - N(k)$ and $N(t)$ are the same. Hence $\mathbb{E}(T_n \,|\, N(k) = m) = \mathbb{E}(k + T_{n-m}) = k + (n-m)/\lambda$.

For any $0 < m < n$ and $k > 0$, $N(n) - N(m) \sim \text{Poisson}(\lambda(n-m))$. Hence $\mathbb{E}(N(n) \,|\, N(m) = k) = \mathbb{E}(N(m) + (N(n) - N(m)) \,|\, N(m) = k) = k + \lambda(n-m)$.

Suppose $N(t)$ is a Poisson process with rate 3. Let T_n denote the time of the n-th arrival. Find (a) $\mathbb{E}(T_{12})$, (b) $\mathbb{E}(T_{12} \,|\, N(2) = 5)$, (c) $\mathbb{E}(N(5) \,|\, N(2) = 5)$.

(a) $\mathbb{E}(T_{12}) = 12/\lambda = 13/3 = 4$. (b) $\mathbb{E}(T_{12} \,|\, N(2) = 5) = 5 + \mathbb{E}(T_{10}) = 5 + 10/3 = 25/3$, (c) $\mathbb{E}(N(5) \,|\, N(2) = 5) = 5 + \mathbb{E}(N(5) - N(2)) = 5 + \lambda(5-2) = 5 + 3 \times 3 = 14$.

3.3 Non-Homogeneous Poisson Processes

Question: Is it possible to have two customers arriving at the same time?

Proposition 54. $P(\sup_{0\le s\le t} N(s) - N(s-) \ge 2) = 0.$

Proof. Note that

$$P(\sup_{0\le s\le t} N(s) - N(s-) \ge 2) = \lim_{n\to\infty} P(\max_{1\le j\le n}\{N(jt/n) - N((j-1)t/n)\} \ge 2)$$

Note that $N(jt/n) - N((j-1)t/n)$ are i.i.d. Poisson$(t\lambda/n)$.

$$= 1 - \lim_{n\to\infty} P(N(jt/n) - N((j-1)t/n) \le 1, j = 1,\dots,n) = 1 - \lim_{n\to\infty} P(N(t/n) - N(0) \le 1)^n$$
$$= 1 - \lim_{n\to\infty}[e^{-\lambda t/n}(1 + \lambda t/n)]^n = 1 - \lim_{n\to\infty} e^{-\lambda t}e^{\lambda t/n\times n} = 1 - 1 = 0.$$

□

Note. In Poisson process, there are no time points having arrival bigger than 1.

Note (Some limit calculus)**.** $\log(1 + z_n) \approx z_n - z_n^2/2 + O(|z_n|^3)$ when $|z_n| < 1$. $\log(1 + \lambda t/n)^n = n\log(1 + \lambda t/n) = n(\lambda t/n - \lambda^2t^2/2n^2 + O(n^{-3})) = \lambda t + O(n^{-1})$.

Question: In reality, customers arrive frequently around noon and very rarely in early in the morning and late at night. Can we replace the stationary increment condition?

Definition 35. A nonhomogeneous Poisson process $N(t)$ with rate $\lambda(t)$ is a continuous time non-negative valued stochastic process satisfying
(a) $N(0) = 0$,
(b) [Independent increment] For $t_1 < t_2 \le t_3 < t_4$, $N(t_2) - N(t_1), N(t_4) - N(t_3)$ are independent.
(c) [Poisson distribution] $N(t) - N(s) \sim$ Poisson$(\int_s^t \lambda(r)\,dr)$ for any $0 \le s < t$.

Note. The interarrival times are not exponential unless $\lambda(t)$ is constant because $P(\tau_1 > t) = P(T_1 > t) = P(N(t) = 0) = e^{-\int_0^t \lambda(r)\,dr}$.

Question: In many cases, customers arrive as groups. The number of subjects in each group follows a distribution. How can we model the total number of subjects arrive?

Theorem 55. Let $Y_1, Y_2, \dots$ be a sequence of i.i.d. finite first moment and T be a stopping time with $P(T < \infty) = 1$. Define $S_0 = 0$ and $S_n = Y_1 + \cdots + Y_n$.
(a) [Wald equation] If $\mathbb{E}|Y_n| < \infty, \mathbb{E}T < \infty$, then $\mathbb{E}S_T = \mathbb{E}T\mathbb{E}Y_n$.
(b) If $\mathbb{E}Y_n^2 < \infty, \mathbb{E}T^2 < \infty$, then $\mathbb{V}\mathrm{ar}(S_T) = \mathbb{E}T\mathbb{V}\mathrm{ar}(Y_n) + \mathbb{V}\mathrm{ar}(T)\mathbb{E}Y_n^2$.
(c) If T is Poisson(λ), then $\mathbb{V}\mathrm{ar}(S_T) = \lambda\mathbb{E}Y_n^2$.

Proof. (a) $\mathbb{E}S_T = \mathbb{E}\sum_{n=1}^\infty S_T 1(T = n) = \mathbb{E}\sum_{n=1}^\infty\sum_{k=1}^n Y_k 1(T = n) = \mathbb{E}\sum_{k=1}^\infty\sum_{n=k}^\infty Y_k 1(T = n) = \mathbb{E}\sum_{k=1}^\infty Y_k 1(T \ge k) = \mathbb{E}\sum_{k=1}^\infty Y_k(1 - 1(T \le k-1)) = \sum_{k=1}^\infty \mathbb{E}1(T \ge k)\mathbb{E}Y_1 = \sum_{k=1}^\infty P(T \ge k)\mathbb{E}Y_1 = \mathbb{E}T\mathbb{E}Y_1$.
(b) We only prove when T and $Y_1, Y_2, \dots$ are independent. Note that $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}(\mathbb{V}\mathrm{ar}(X \mid Z)) + \mathbb{V}\mathrm{ar}(\mathbb{E}(X \mid Z))$. Then $\mathbb{V}\mathrm{ar}(S_T) = \mathbb{E}\mathbb{V}\mathrm{ar}(S_T \mid T) + \mathbb{V}\mathrm{ar}(\mathbb{E}(S_T \mid T)) = \mathbb{E}(T\mathbb{V}\mathrm{ar}(Y_1)) + \mathbb{V}\mathrm{ar}(T\mathbb{E}(Y_1)) = \mathbb{E}(T)\mathbb{V}\mathrm{ar}(Y_1) + \mathbb{V}\mathrm{ar}(T)(\mathbb{E}Y_1)^2$.
(c) If $T \sim$ Poisson(λ), then $\mathbb{E}T = \lambda$ and $\mathbb{V}\mathrm{ar}(T) = \lambda$. Hence $\mathbb{V}\mathrm{ar}(S_T) = \lambda(\mathbb{V}\mathrm{ar}(Y_1) + (\mathbb{E}Y_1)^2) = \lambda\mathbb{E}Y_1^2$. □

Exercise 30. Prove part (b) for general stopping time T with $P(T < \infty) = 1$.

Example 41 (Liquor store)**.** The number of customer visiting the store follows a Poisson distribution with mean 81. Each customer spends in the store on average \$8 with the standard deviation \$6.

The income of the store is mean $\mathbb{E}S_T = \mathbb{E}T\mathbb{E}Y_1 = 81 \times 8 = \$648$ and the variance is $\mathbb{V}\mathrm{ar}(S_T) = \lambda\mathbb{E}Y_1^2 = 81(6^2 + 8^2) = \$8100$. Considering the standard deviation is $(8100)^{1/2} = 90$.

3.4 Thinning, Superposition, Conditioning

Let $N(t)$ be a homogeneous Poisson process and $Y_1, Y_2, \ldots$ be the associated random variables with the arrivals. Using Y_i's the Poisson process can be splitted. Suppose Y_i's are positive integer valued, i.i.d. random variables. Define $N_j(t)$ be the number of $i \le N(t)$ with $Y_i = j$, that is, $N_j(t) = \sum_{i=1}^{N(t)} 1(Y_i = j)$.

Theorem 56. $N_j(t)$ are independent Poisson processes with rate $\lambda P(Y_i = j)$.

Proof. Note that $N_j(0) = 0$ for all j. For $t_0 < t_1 < \ldots < t_n$, $N_j(t_k) - N_j(t_{k-1}) = \sum_{N(t_{k-1})<l\le N(t_k)} 1(Y_l = j)$ are independent. Hence $N_j(t)$ has independent increment. For any $m \ge 0$, $P(N_j(t) - N_j(s) = m) = \sum_{n=m}^{\infty} P(N_j(t) - N_j(s) = m \,|\, N(t) - N(s) = n) P(N(t) - N(s) = n) = \sum_{n=m}^{\infty} \binom{n}{m} P(Y_1 = j)^m P(Y_1 \ne j)^{n-m} e^{-\lambda(t-s)}\{\lambda(t-s)\}^n/n! = e^{-\lambda(t-s)}[\{\lambda P(Y_1 = j)(t-s)\}^m/m!] \sum_{n=m}^{\infty}\{\lambda P(Y_1 \ne j)(t-s)\}^{n-m}/(n-m)! = e^{-\lambda P(Y_1=j)(t-s)}\{\lambda P(Y_1 = j)(t-s)\}^m/m!$, that is, $N_j(t) - N_j(s) \sim \text{Poisson}(\lambda P(Y_1 = j))$. Hence $N_j(t)$ are Poisson processes.

Let $n_j \ge 0$ be a sequence sum up $n = n_1 + n_2 + \cdots < \infty$. Then $N_j(t) - N_j(s) = n_j$ for all j implies $N(t) - N(s) = n$. Given $N(t) - N(s) = n$, the conditional distributions of $N_j(t) - N_j(s)$ is a multinomial distribution with n trial and success probability $P(Y_1 = j)$ respectively. Hence,

$$P(N_j(t) - N_j(s) = n_j, j = 1, 2, \ldots) = P(N_j(t) - N_j(s) = n_j, j = 1, 2, \ldots \,|\, N(t) - N(s) = n) P(N(t) - N(s) = n)$$
$$= n! \prod_{j=1}^{\infty} \frac{P(Y_1 = j)^{n_j}}{n_j!} \times e^{-\lambda(t-s)} \frac{(\lambda(t-s))^n}{n!} = \prod_{j=1}^{\infty} \Big[\frac{P(Y_1 = j)^{n_j}}{n_j!} \times e^{-\lambda P(Y_1=j)(t-s)} (\lambda(t-s))^{n_j}\Big]$$
$$= \prod_{j=1}^{\infty} e^{-\lambda P(Y_1=j)(t-s)} (\lambda P(Y_1 = j)(t-s))^{n_j}/n_j! = \prod_{j=1}^{\infty} P(N_j(t) - N_j(s) = n_j).$$

Thus $N_j(t) - N_j(s)$ are independent. □

Theorem 57. Let $N_j(t)$ are independent homogeneous Poisson process with rates λ_j. Then, $N(t) = N_1(t) + \cdots + N_k(t)$ is a homogeneous Poisson process with rate $\lambda_1 + \cdots + \lambda_k$.

Proof. Note that $N(0) = 0$. For any $0 \le t_0 < \cdots < t_l$, $N(t_n) - N(t_{n-1}) = \sum_{j=1}^{k} N_j(t_n) - N_j(t_{n-1})$ are independent. Hence $N(t)$ has independent increment. Note $N_j(t) - N_j(s)$ are independently $\text{Poisson}(\lambda_j(t-s))$. Thus $N(t) - N(s) = \sum_{j=1}^{k} N_j(t) - N_j(s) \sim \text{Poisson}(\sum_{j=1}^{k} \lambda_j(t-s))$. Therefore $N(t)$ is a homogeneous Poisson process with rate $\lambda_1 + \cdots + \lambda_k$. □

Independent separation of Poisson processes is called *thinning* and the merge of independent Poisson processes is called *superposition*. The following theorem is called *conditioning*.

Theorem 58. Let $U_1, \ldots, U_n$ be i.i.d Uniform$(0, t)$ and $V_1, \ldots, V_n$ be the order statistic of $U_1, \ldots, U_n$. Given $N(t) = n$, $(T_1, \ldots, T_n)$ and $(V_1, \ldots, V_n)$ have the same distribution.

Proof. Assume $N(t) = n$. Let $0 = t_0 < t_1 < \cdots < t_n = t$. The conditional density is

$$\text{pdf}_{T_1,\ldots,T_n \,|\, N(t)=n}(t_1, \ldots, t_n) = \prod_{k=1}^{n} \lambda e^{-\lambda(t_k - t_{k-1})} \Big/ [e^{-\lambda t}(\lambda t)^n/n!] = n!/t^n.$$

Also the density of $(V_1, \ldots, V_n)$ is $\text{pdf}_{V_1,\ldots,V_n}(v_1, \ldots, v_n) = n! \prod_{k=1}^{n} \text{pdf}_{U_k}(v_k) = n!(1/t)^n = n!/t^n$. □

Example 42. For $0 \le m \le n$ and $0 \le s < t$, the conditional distribution of $N(s)$ given $N(t) = n$ is Binomial$(n, s/t)$, that is, $P(N(s) = m \,|\, N(t) = n) = \binom{n}{m}(s/t)^m(1 - s/t)^{n-m}$.

Example 43. The number of clients per hour in a coffee shop follows a homogeneous Poisson process with rate 20. There were 16 clients entered the coffee shop between 6pm to 7pm. What is the probability that 10 of them entered between 6pm and 6:20pm?

The probability is $P(N(1/3) = 10 \,|\, N(1) = 16) = \binom{16}{10}(1/3)^1 0(2/3)^6 = 8008 * 64/e^1 6 = 0.0119$.

Example 44 (Exercise 2.59)**.** Customers arrive at a bank according to a Poisson process with rate 10 per hour. Given that two customers arrived in the first 5 min, what is the probability that (a) both arrived in the first 2 min. (b) at least one arrived in the first 2 min.

(a) $P(N(2) = 2 \mid N(5) = 2) = \binom{2}{2}(2/5)^2(3/5)^0 = 4/25 = 0.16$. (b) $P(N(2) \geq 1 \mid N(5) = 2) = 1 - P(N(2) < 1 \mid N(5) = 2) = 1 - P(N(2) = 0 \mid N(5) = 2) = 1 - \binom{2}{0}(2/5)^0(3/5)^2 = 1 - 9/25 = 16/25 = 0.64$.