Poisson process

Problem 1. The count of students dropping the course "Probability and Stochastic Processes" is known to be a Poisson process of rate 0.1 drops per day. Starting with day 0, the first day of the semester, let D_t denote the number of students that have dropped after t days. What is $P\{D_t = k\}$?

Problem 2. The arrivals of new telephone calls at a telephone switching office is a Poisson process X_t with an arrival rate of $\lambda = 4$ calls per second. An experiment consists of monitoring the switching office and recording X_t over a 10-second interval.

(a) What is $P\{X_1 = 0\}$, the probability of no phone calls in the first second of observation?

(b) What is $P\{X_1 = 4\}$, the probability of exactly four calls arriving in the first second of observation?

(c) What is $P\{X_2 = 2\}$, the probability of exactly two calls arriving in the first two seconds?

Problem 3. The stochastic process $\{X_t, t \geq 0\}$ is defined by

$$X_t = N_{t+1} - N_1 \text{ for } t \geq 0$$

where $\{N_t, t \geq 0\}$ is a Poisson process with rate $\lambda > 0$. Find a $K_X(s,t)$, for $0 \leq s \leq t$.

Problem 4. Let $X(t)$ be a Poisson process with parameter λ. Find

(a) $E[X^2(t)]$

(b) $E[(X(t) - X(s))^2]$, for $t > s$.

Problem 5. The number of failures N_t, which occur in a computer network over the time interval $[0,t)$, can be described by a homogeneous Poisson process $\{N_t, t \geq 0\}$. On an average, there is a failure after every 4 hours, i.e. the intensity of the process is equal to $\lambda = 0.25$.

(a) What is the probability of at most 1 failure in $[0,8)$, at least 2 failures in $[8,16)$ and at most 1 failure in $[16,24)$ (time unit: hour)?

(b) What is the probability that the third occurs after 8 hours?

Problem 6. A sequence of queries are made to a database system. The response time of the system, T seconds, is an exponential random variable with mean 8. As soon as the system responds to a query, the next query is made. Assuming the first query is made at time zero, let N_t denote the number of queries made by time t.

(a) What is $P\{T \geq 4\}$, the probability that a single query will last at least four seconds?

(b) If the database user has been waiting five seconds for a response, what is $P\{T \geq 13 | T \geq 5\}$, the probability that it will last at least eight more seconds?

(c) What is the PMF of N_t?

Problem 7. Suppose that people immigrate into a territory at a Poisson rate $\lambda = 1$ per day.

(a) What is the expected time until the tenth immigrant arrives?

(b) What is the probability that the elapsed time between the tenth and the eleventh arrival exceeds two days?

Problem 8. Customers arrive at a casino as a Poisson process of rate 100 customers per hour. Upon arriving, each customer must flip a coin, and only those customers who flip heads actually enter the casino. Let X_t denote the process of customers entering the casino. Find the PMF of N, the number of customers who arrive between 5 p.m. and 7 p.m..

Problem 9. If immigrants to area A arrive at a Poisson rate of ten per week, and if each immigrant is of English descent with probability $1/12$, then what is the probability that no people of English descent will emigrate to area A during the month of February?

Problem 10. Consider an elevator that starts in the basement and travels upward. Let N_i denote the number of people that get in the elevator at floor i. Assume the N_i are independent and that N_i is Poisson random variable with mean λ_i. Each person entering at i will, independent of everything else, get off at j with probability p_{ij}, $\sum_{j>i} p_{ij} = 1$. Let O_j denote the number of people getting off the elevator at floor j.

(a) Compute $E(O_j)$.

(b) What is the distribution of O_j?

(c) What is the joint distribution of O_j and O_k?

We know that if each event of a Poisson process is independently classified as a type I event with probability p and as a type II event with probability $1-p$ then the counting processes of type I and type II events are independent Poisson processes with respective rates λp and $\lambda(1-p)$. Suppose now, however, that there are k possible types of events and that the probability that an event is classified as a type i event, $i = 1, \ldots, k$, depends on the time the event occurs. Specifically, suppose that if an event occurs at time y then it will be classified as a type i event, independently of anything that has previously occurred, with probability $P_i(y)$, $i = 1, \ldots, k$ where $\sum_{i=1}^{k} P_i(y) = 1$. Now we have

Proposition 0.1. *If $X_i(t)$, $i = 1, \ldots, k$, represents the number of type i events occurring by time t then $X_i(t)$, $i = 1, \ldots, k$, are independent Poisson random variables having means*

$$E[X_i(t)] = \lambda \int_0^t P_i(s) ds.$$

Problem 11. [Tracking the Number of HIV Infections] There is a relatively long incubation period from the time when an individual becomes infected with the HIV virus, which causes AIDS, until the symptoms of the disease appear. As a result, it is difficult for public health officials to be certain of the number of members of the population that are infected at any given time. Find an approximation model for this phenomenon, which can be used to obtain a rough estimate of the number of infected individuals. Suppose that

- individuals contract the HIV virus in accordance with a Poisson process whose rate λ is unknown,
- the time from when an individual becomes infected until symptoms of the disease appear is a random variable having a known distribution G,
- the incubation times of different infected individuals are independent.

Let $N_1(t)$ denote the number of individuals who have shown symptoms of the disease by time t . Also, let $N_2(t)$ denote the number who are HIV positive but have not yet shown any symptoms by time t.

Solution 5. Since an individual who contracts the virus at time s will have symptoms by time t with probability $G(t-s)$ and will not with probability $1 - G(t-s) = \bar{G}(t-s)$, it follows that $N_1(t)$ and $N_2(t)$ are independent Poisson random variables with respective means

$$E(N_1(t)) = \lambda \int_0^t G(t-s)\, ds = \lambda \int_0^t G(y)\, dy,$$

$$E(N_2(t)) = \lambda \int_0^t \bar{G}(t-s)\, ds = \lambda \int_0^t \bar{G}(y)\, dy.$$

Now, if we knew λ, then we could use it to estimate $N_2(t)$, the number of individuals infected but without any outward symptoms at time t, by its mean value $E[N_2(t)]$. However, since λ is unknown, we must first estimate it. Now, we will presumably know the value of $N_1(t)$, and so we can use its known value as an estimate of its mean $E[N_1(t)]$. That is, if the number of individuals who have exhibited symptoms by time t is n_1, then we can estimate that

$$n_1 \approx E(N_1(t)) = \lambda \int_0^t \bar{G}(y)\, dy.$$

Therefore, we can estimate λ by the quantity $\hat{\lambda}$ given by

$$\hat{\lambda} = \frac{n_1}{\int_0^t G(y)\, dy}.$$

Using this estimate of λ, we can estimate the number of infected but symptomless individuals at time t by

$$\text{estimate of } N_2(t) = \hat{\lambda} \int_0^t \bar{G}(y)\, dy = n_1 \frac{\int_0^t \bar{G}(y)\, dy}{\int_0^t G(y)\, dy}.$$

For example, suppose that G is exponential with mean μ. Then $\bar{G}(y) = e^{-\frac{y}{\mu}}$, and a simple integration gives that

$$\text{estimate of } N_2(t) = \frac{n_1 \mu (1 - e^{-\frac{t}{\mu}})}{t - \mu(1 - e^{-\frac{t}{\mu}})}.$$

If we suppose that $t = 16$ years, $\mu = 10$ years, and $n_1 = 220$ thousand, then the estimate of the number of infected but symptomless individuals at time 16 is

$$\text{estimate of } N_2(16) = \frac{220 \cdot 10(1 - e^{-1.6})}{16 - 10(1 - e^{-1.6})} = 218.96.$$

That is, if we suppose that the foregoing model is approximately correct (and we should be aware that the assumption of a constant infection rate λ that is unchanging over time is almost certainly a weak point of the model), then if the incubation period is exponential with mean 10 years and if the total number of individuals who have exhibited AIDS symptoms during the first 16 years of the epidemic is 220 thousand, then we can expect that approximately 219 thousand individuals are HIV positive though symptomless at time 16.

Definition 0.1. *The counting process* $\{X_t, t \geq 0\}$ *is said to be a* **nonstationary or nonhomogeneous Poisson process** *with intensity function* $\lambda(t)$, $t \geq 0$ *if*

(i) $X_0 = 0$

(ii) $\{X_t, t \geq 0\}$ *has independent increments*

(iii) $P\{X_{t+h} - X_t = 1\} = \lambda(t)h + o(h), h \to 0$

(iv) $P\{X_{t+h} - X_t \geq 2\} = o(h), h \to 0$

If we let

$$m(t) = \int_0^t \lambda(s) ds,$$

then it can be shown

$$P\{X_{s+t} - X_t = n\} = e^{-(m(s+t) - m(s))} \frac{(m(s+t) - m(s))^n}{n!}.$$

That is, $X_{s+t} - X_t$ is Poisson distributed with mean $m(s+t) - m(t)$.

Problem 12. For a nonhomogeneous Poisson process N_t the intensity function is given by

$$\lambda(t) = \begin{cases} 5, & \text{if t is in } (1,2], (3,4], \dots \\ 3, & \text{if t is in } (0,1], (2,3], \dots \end{cases}.$$

Find the probability that the number of observed occurrences in the time period $(1.25, 3]$ is more than two.

Problem 13. A store opens at 8 a.m. From 8 until 10 customers arrive at a Poisson rate of four per hour. Between 10 and 12 they arrive at a Poisson rate of eight per hour. From 12 to 2 p.m. the arrival rate increases steadily from eight per hour at 12 to ten per hour at 2 p.m. and from 2 to 5 p.m. the arrival rate drops steadily from ten per hour at 2 to four per hour at 5 p.m.. Determine the probability distribution of the number of customers that enter the store on a given day.

Definition 0.2. *A Stochastic process* $\{X_t,\, t \geq 0\}$ *is said to be a* **compound Poisson process** *if it can be represented, for* $t \geq 0$, *by*

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

where $\{N_t, t \geq 0\}$ *is a Poisson process, and* $\{Y_i, i = 1, 2, \dots\}$ *is a family of independent and identically distributed random variables that is independent of the process* $\{N_t, t \geq 0\}$. *Thus, if* $\{X_t,\, t \geq 0\}$ *is a compound Poisson process then* X_t *is a compound Poisson random variable.*

We have

$$E(X_t) = E(N_t)E(Y_1), \quad D(X_t) = E(N_t)E(Y_1^2).$$

Problem 14. Suppose that health claims are filed with a health insurer at the Poisson rate $\lambda = 20$ per day, and that the independent severities of each claim are Exponential random variables with mean $\theta = 500$. Find the expected value and variance of an aggregate of claims during the first 10 days. Estimate the probability that the aggregate claims during the first 10 days exceed 120 000.

Problem 15. Suppose that families migrate to an area at a Poisson rate $\lambda = 2$ per week. If the number of people in each family is independent and takes on the values 1, 2, 3, 4 with respective probabilities $\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}$, then what is the expected value and variance of the number of individuals migrating to this area during a fixed five-week period?