Week 9

Poisson processes

25 Sum and decomposition of poisson process

Example 9.1. Suppose that the number of customers visiting a fast food restaurant in a given time interval I is $N \sim Poisson(\lambda)$. Assume that each customer purchases a drink with probability p, independently from other customers, and independently from the value of N. Let X be the number of customers who purchase drinks in that time interval. Also, let Y be the number of customers that do not purchase drinks; so $X + Y = N$.

- What are the distributions for X and Y?
- Find the joint distribution of X and Y.
- Are X and Y independent?

Solution

- First note that, given $N = n$, X is a sum of n independent $Bernoulli(p)$ random variables. Thus,

$$\begin{aligned} X|N = n &\sim Binomial(n, p), \\ Y|N = n &\sim Binomial(n, q = 1 - p). \end{aligned}$$

It follows from the law of total probability that

$$\begin{aligned} P(X = k) &= \sum_{n=0}^{\infty} P(X = k|N = n)P(N = n) \\ &= \sum_{n=k}^{\infty} \binom{n}{k} p^k q^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} \\ &= \frac{e^{-\lambda p}(\lambda p)^k}{k!}, \qquad \text{for } k = 0, 1, 2, ... \end{aligned}$$

Thus, we conclude:

$$X \quad \sim \quad Poisson(\lambda p).$$

Similarly, we obtain

$$Y \quad \sim \quad Poisson(\lambda q).$$

- To find the joint distribution of X and Y, we can also use the law of total probability:

$$P(X = i, Y = j) \;=\; \sum_{n=0}^{\infty} P(X = i, Y = j | N = n) P(N = n)$$

Note that $P(X = i, Y = j | N = n) = 0$ if $N \neq i + j$. It follows that

$$\begin{aligned}
P(X = i, Y = j) &= P(X = i, Y = j | N = i + j) P(N = i + j) \\
&= P(X = i | N = i + j) P(N = i + j) \\
&= \binom{i+j}{i} p^i q^j e^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!} \\
&= \frac{e^{-\lambda} (\lambda p)^i (\lambda q)^j}{i! j!} \\
&= \frac{e^{-\lambda p} (\lambda p)^i}{i!} . \frac{e^{-\lambda q} (\lambda q)^j}{j!} \\
&= P(X = i) P(Y = j).
\end{aligned}$$

- X and Y are independent since,

$$P(X = i, Y = j) = P(X = i) P(Y = j), \quad \text{for all } i, j \geq 0.$$

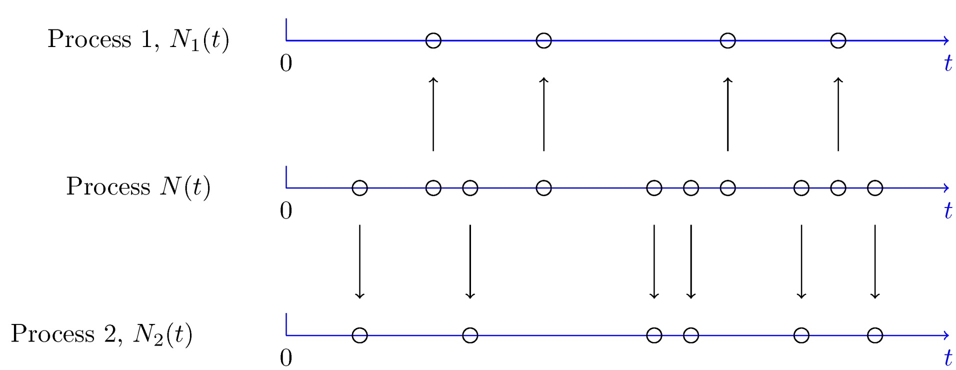
Example 9.1 is for Poisson variable. Result is still true for a Poisson process.

Splitting a Poisson process

Let $N_t, t \geq 0$, be a Poisson process with rate λ. Suppose that each arrival of N_t can be classified into "type I" or "type II" arrival with probability p and $1-p$, respectively. Denote by $N_1(t)$ or $N_2(t)$ the number of "type I" or "type II" arrivals, respectively.

Then $N_t = N_1(t) + N_2(t)$, where

- $N_1(t)$ is a Poisson process with rate λp ;
- $N_2(t)$ is a Poisson process with rate $\lambda(1-p)$;
- $N_1(t)$ and $N_2(t)$ are independent.



Example 9.2. Suppose that people arrive at a service counter in accordance with a Poisson process with rate λ per hour. If customers are male with probability $p = 1/3$, given that 20 males arrived between $10:30$am and 11 am, how many females would we expect to have arrived in that time?

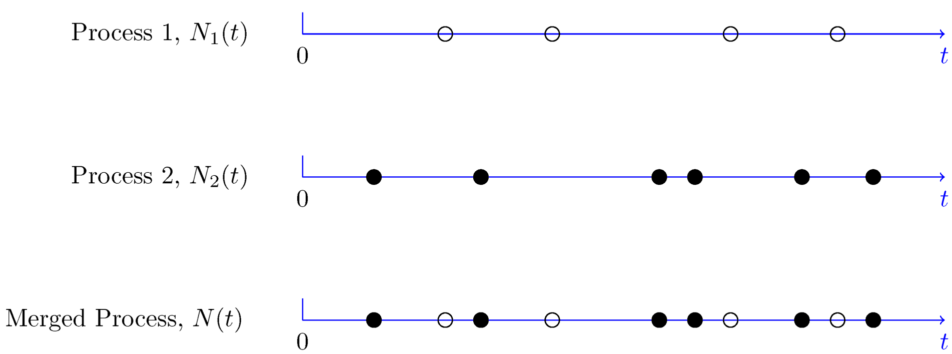
Solution. Let N be the number of customers arriving at a service counter between $10:30$am and 11 am. Let N_1 (N_2) be the number of male (female) customers arriving at a service counter between $10:30$am and 11 am.

It follows that $N \sim Poisson(\lambda/2), N_1 \sim Poisson(\lambda/6)$ and $N_2 \sim Poisson(2\lambda/6)$

Merging Independent Poisson Processes

Let $N_1(t)$ and $N_2(t)$ be two independent Poisson processes with rates λ_1 and λ_2 respectively. Define $N_t = N_1(t) + N_2(t)$.

N_t is obtained by combining the arrivals in $N_1(t)$ and $N_2(t)$ as below.



We claim that N_t is a Poisson process with rate $\lambda = \lambda_1 + \lambda_2$.
To see this, first note that

$$N_0 = N_1(0) + N_2(0) = 0.$$

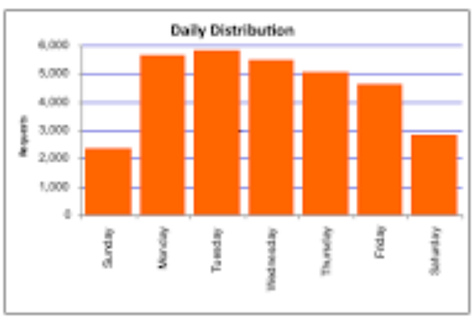
Next, since $N_1(t)$ and $N_2(t)$ are independent and both have independent increments, we conclude that N_t also has independent increments. Finally, consider an interval of length τ, i.e, $I = (t, t+\tau]$. Then the numbers of arrivals in I associated with $N_1(t)$ and $N_2(t)$ are $Poisson(\lambda_1\tau)$ and $Poisson(\lambda_2\tau)$ and they are independent. Therefore, the number of arrivals in I associated with N_t is $Poisson\big((\lambda_1+\lambda_2)\tau\big)$ (sum of two independent Poisson random variables).

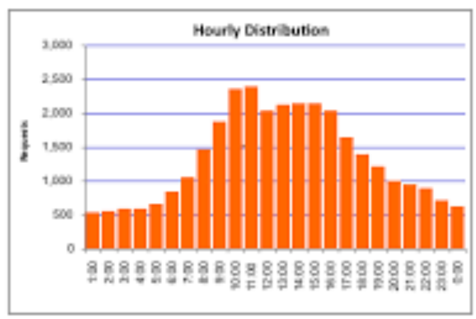
Example 9.3. Suppose the number of claims to an insurance company from smokers and non-smokers follow independent Poisson processes. Suppose the expected number of claims by non-smokers is 2 per unit time and the expected number of claims from smokers is 6 per unit time. Then the total number of claims is also Poisson with rate 8.

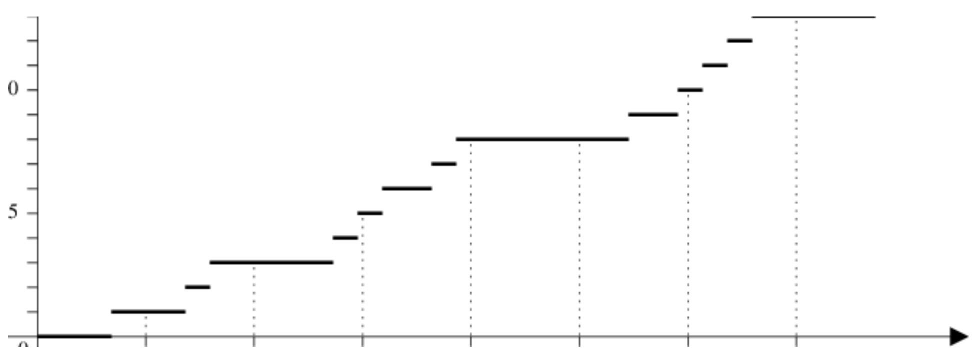
26 Nonhomogeneous poisson process

Motivation:

Let N_t be the number of customers arriving at a fast food restaurant by time t. As the customers arrive somewhat randomly, it is natural to model N_t as a Poisson process. This process, however, does not have stationary increments. In fact, the arrival rate of customers is larger during lunch time compared to, say, 4 p.m. In such scenarios, we might model N_t as a **nonhomogeneous Poisson process**. Such a process has all the properties of a Poisson process, except for the fact that its rate is a function of time, i.e., $\lambda = \lambda(t)$.







Definition:

Let $\lambda(t) : [0, \infty) \mapsto [0, \infty)$ be an integrable function. The counting process $\{N_t, t \geq 0)\}$ is called a **nonhomogeneous Poisson process** with rate function $\lambda(t)$ if all the following conditions hold.

- $N_0 = 0$;
- N_t has *independent* increments;
- for any $t \in [0, \infty)$, we have

$$\begin{aligned}
&P(N_{t+\delta} - N_t = 0) = 1 - \lambda(t)\delta + o(\delta), \\
&P(N_{t+\delta} - N_t = 1) = \lambda(t)\delta + o(\delta), \\
&P(N_{t+\delta} - N_t \geq 2) = o(\delta).
\end{aligned}$$

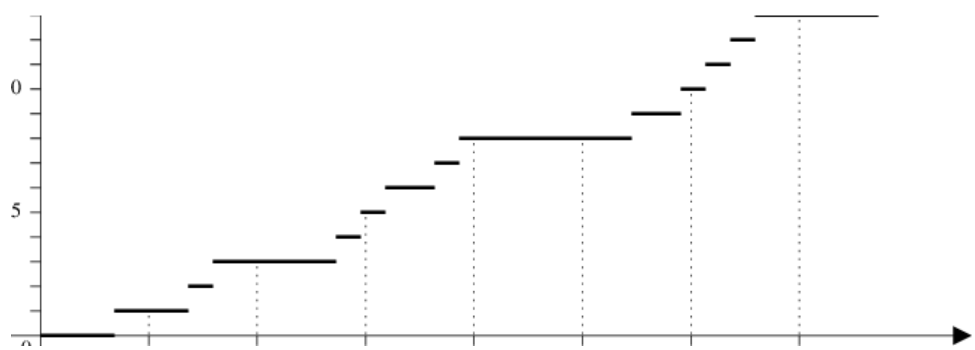
For a nonhomogeneous Poisson process with rate function $\lambda(t)$, the number of arrivals in any interval is a Poisson random variable. The parameter (or mean) of the Poisson variable, however, depends on the location of the interval. Indeed, we have

$$N_t - N_s \;\sim\; Poisson\left(\int_s^t \lambda(x)dx\right),$$

Denote by $m(u) = \int_0^u \lambda(x)dx$ ($m(u)$ **the mean value function**. It follows that

$$P\big[N_t - N_s = k\big] \;=\; \big[m(t) - m(s)\big]^k e^{-\big[m(t)-m(s)\big]}/k!$$

for $k = 0, 1, 2, ...$



Example 9.4. Siegbert runs a hot dog stand that opens at 8 A.M. From 8 until 11 A.M. customers seem to arrive, on the average, at a steadily increasing rate that starts with an initial rate of 5 customers per hour at 8 A.M. and reaches a maximum of 20 customers per hour at 11 A.M. From 11 A.M. until 1 P.M. the (average) rate seems to remain constant at 20 customers per hour. However, the (average) arrival rate then drops steadily from 1 P.M. until closing time at 5 P.M. at which time it has the value of 12 customers per hour.

- If we assume that the numbers of customers arriving at Siegbert's stand during disjoint time periods are independent, then what is a good probability model for the preceding?
- What is the probability that no customers arrive between 8:30 A.M. and 9:30 A.M. on Monday morning?
- What is the expected number of arrivals in this period?

Solution:

- Let N_t represent the number of arrivals during the first t hours that the store is open. A good model for the preceding would be to assume that arrivals constitute a nonhomogeneous Poisson process with intensity function $\lambda(t)$ given by

$$\lambda(t) \;=\; \begin{cases} 5+5t, & 0 \le t \le 3 \\ 20, & 3 < t \le 5 \\ 20-2(t-5), & 5 < t \le 9 \end{cases}$$

and

$$\lambda(t) = 0 \quad t > 9.$$

- Let $m(t) = \int_0^t \lambda(x)dx$. The number of arrivals between 8:30 A.M. and 9:30 A.M. is Poisson with mean $m(3/2) - m(1/2)$. Hence the required probability is

$$P(N_{3/2} - N_{1/2} = 0) \;=\; e^{-\int_{1/2}^{3/2} \lambda(u)du} = e^{-10}.$$

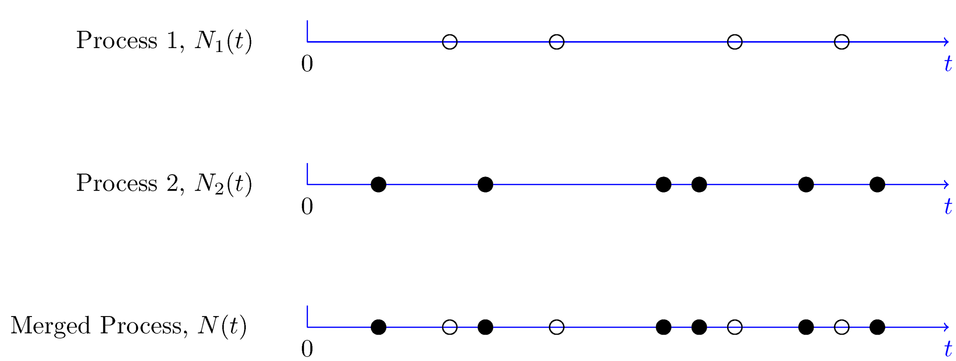
- The mean number of arrivals is $\int_{1/2}^{3/2} \lambda(u)du = 10$.

Sum of nonhomogeneous Poisson processes

Let $N_1(t)$ and $N_2(t)$ be two independent Poisson processes with rates $\lambda_1(t)$ and $\lambda_2(t)$ respectively. Define $N_t = N_1(t) + N_2(t)$. We claim that N_t is a Poisson process with rate $\lambda(t) = \lambda_1(t) + \lambda_2(t)$.

In fact, we have

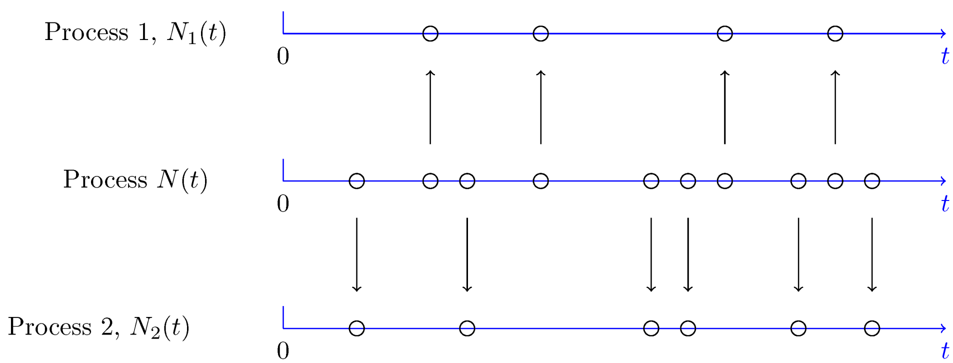
$$
\begin{aligned}
P(N_t = k) &= \sum_{j=0}^{k} P(N_1(t) = j, N_2(t) = k - j) \\
&= \sum_{j=0}^{k} \frac{m_1^j(t)}{j!} e^{-m_1(t)} \frac{m_2^{k-j}(t)}{(k-j)!} e^{-m_2(t)} \\
&= \frac{1}{k!} e^{-m_1(t) - m_2(t)} \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} m_1^j(t) m_2^{k-j}(t) \\
&= \frac{\left[m_1(t) + m_2(t)\right]^k}{k!} e^{-m_1(t) - m_2(t)}.
\end{aligned}
$$



Splitting nonhomogeneous Poisson process

Let $N_t, t \geq 0$, be a Poisson process with rate λ.

Suppose that events occur according to a Poisson process with rate λ, and suppose that, independent of what has previously occurred, an event at time s is a type 1 event with probability $p_1(s)$ or a type 2 event with probability $p_2(s) = 1 - p_1(s)$. If $N_i(t)$ denotes the number of type i events by time t, then $N_1(t), t \geq 0$ and $N_2(t), t \geq 0$ are independent nonhomogeneous Poisson processes with respective intensity functions $\lambda_i(t) = \lambda\, p_i(t), i = 1, 2$.



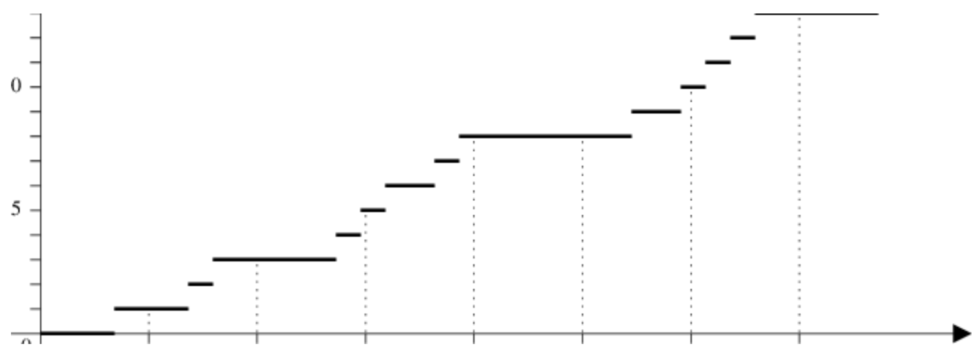
Similarly, let $N_t, t \geq 0$, be a nonhomogeneous Poisson process with rate $\lambda(t)$. Suppose that, independent of what has previously occurred, an event at time s is a type 1 event with probability p or a type 2 event with probability $q = 1 - p$. If $N_i(t)$ denotes the number of type i events by time t, then $N_1(t), t \geq 0$ and $N_2(t), t \geq 0$ are independent nonhomogeneous Poisson processes with respective intensity functions $\lambda_1(t) = p\,\lambda(t)$ and $\lambda_2(t) = (1 - p)\,\lambda(t)$.

Remark

For a **nonhomogeneous Poisson process** $N_t, t \geq 0$, with rate $\lambda(u)$ (or mean value function $m(u) = \int_0^u \lambda(x)dx$), the arrival times T_n and the interarrival times E_n are still well-defined, where $T_0 = 0$,

$$T_n = \inf\{s > 0 : N_s = n\}, \quad n = 1, 2, ...,$$

and $E_n = T_n - T_{n-1}$, $n = 1, 2, ...$ (hence $T_n = \sum_{j=1}^n E_j$).



But the interarrival times E_i are no longer iid with $E_i \sim Exp(\lambda)$. To see this, let T_1 be the first arrival, then

$$\mathbb{P}(T_1 > t) = \mathbb{P}(0 \text{ arrivals in } [0, t]) = P(N_t = 0) = e^{-m(t)}.$$

Since the distribution of T_1 depends on $\lambda(x)$, $E_1 = T_1$ is no longer an exponential random variable. On the other hand, the interarrival times are no longer independent in general, since if E_2 denotes the next interarrival time after T_1 then

$$\mathbb{P}(E_2 > t \mid T_1 = k) = e^{-\left[m(k+t)-m(k)\right]}.$$

Whenever $\lambda(x)$ is not a constant function, this conditional probability depends on $T_1 = E_1$, and thus (E_1, E_2) are not independent.

27 Compound poisson process

Example 9.5. Suppose that customers arrive at a store since 8 am in accordance with a Poisson process with rate $\lambda = 10$ per hour. Suppose that the amounts of money spent by the k-th customer is $X_k, k = 1, 2, \ldots$. What is the best model to describe the total amount spent in the store by all customers?

Solution: Let N_t be the number of customers arriving at a store in t hours since 8 am. Then $N_t, t \geq 0$, is a Poisson process with rate $\lambda = 10$.

$S_t = \sum_{j=1}^{N_t} X_j$ is the total amount spent in the store by all customers in t hours since 8 am.

Compound poisson process

Let $\{N_t, t \geq 0)\}$ be a homogeneous Poisson process with rate λ. Let $X_1, X_2, ...$ be i.i.d random variables independent of $N_t, t \geq 0$. The process

$$Z_t \ := \ \sum_{k=1}^{N_t} X_k, \quad t \geq 0,$$

where $\sum_{k=1}^{0} \equiv 0$, is called a **Compound poisson process**.

A compound poisson process Z_t has stationary and independent increments, i.e.,

$$Z_{t+s} - Z_t = \sum_{k=N_t+1}^{N_{t+s}} X_k,$$

has the same distribution as that of $Z_s = \sum_{k=1}^{N_s} X_k$; and $Z_{t+s} - Z_t, s, t > 0$ is independent of $Z_t, t > 0$.

Draft proof.

$$\begin{aligned}
P(Z_{t+s} - Z_t \leq x) &= E\big[P\big(Z_{t+s} - Z_t \leq x \mid N_t, N_{t+s}\big)\big] \\
&= \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} P\Big(\sum_{k=j+1}^{i} X_k \leq x\Big)\, P(N_t = j, N_{t+s} = i) \\
&= \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} P\Big(\sum_{k=1}^{i-j} X_k \leq x\Big)\, P(N_t = j, N_{t+s} - N_t = i - j) \\
&\qquad (X_k \ iid) \\
&= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} P\Big(\sum_{k=1}^{i} X_k \leq x\Big)\, P(N_t = j, N_{t+s} - N_t = i) \\
&= \sum_{i=0}^{\infty} P\Big(\sum_{k=1}^{i} X_k \leq x\Big)\, P(N_s = i) \\
&\qquad (N_u \text{ has indep and stationary increments}) \\
&= P(Z_s \leq x).
\end{aligned}$$

For the independence, we have to show: for any $x, y \in R$,

$$P(Z_{t+s} - Z_t \le x, Z_t \le y) \;=\; P(Z_{t+s} - Z_t \le x)\, P(Z_t \le y).$$

The idea is similar:

$$\begin{aligned}
& P(Z_{t+s} - Z_t \le x, Z_t \le y) \;=\; E\Big[P\Big(Z_{t+s} - Z_t \le x, Z_t \le y \mid N_t, N_{t+s}\Big)\Big] \\
=\; & \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} P\Big(\sum_{k=j+1}^{i} X_k \le x, \sum_{k=1}^{j} X_k \le y\Big)\, P(N_t = j, N_{t+s} = i) \\
=\; & \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} P\Big(\sum_{k=1}^{i-j} X_k \le x\Big)\, P\Big(\sum_{k=1}^{j} X_k \le y\Big) P(N_t = j, N_{t+s} - N_t = i - j) \\
& \quad (X_k \; iid) \\
=\; & \sum_{j=0}^{\infty} P\Big(\sum_{k=1}^{j} X_k \le y\Big) P(N_t = j) \sum_{i=0}^{\infty} P\Big(\sum_{k=1}^{i} X_k \le x\Big)\, P(N_{t+s} - N_t = i) \\
=\; & P(Z_t \le y) \sum_{i=0}^{\infty} P\Big(\sum_{k=1}^{i} X_k \le x\Big)\, P(N_s = i) \\
& \quad (N_u \text{ has indep and stationary increments}) \\
=\; & P(Z_t \le y) P(Z_s \le x) = P(Z_t \le y) P(Z_{t+s} - Z_t \le x).
\end{aligned}$$

Note that Z_t is not a Poisson process.

By the Wald's equation, it is easy to obtain that

$$E(Z_t) \;=\; \lambda t\, EX_1, \quad Var(Z_t) = \lambda t\, EX_1^2.$$

$$\begin{aligned}
EZ_t^2 &= \sum_{j=1}^{\infty} E\big(\sum_{k=1}^{j} X_k\big)^2 P(N_t = j) \\
&= \sum_{j=1}^{\infty} \big[j^2 (EX_1)^2 + j\, Vax(X_1)\big] P(N_t = j) \\
&= (EX_1)^2 EN_t^2 + Var(X_1)\, EN_t \\
&= (EX_1)^2 \big[\lambda t + (\lambda t)^2\big] + \big[EX_1^2 - (EX_1)^2\big]\, \lambda t \\
&= (EX_1)^2\, (\lambda t)^2 + EX_1^2\, \lambda t
\end{aligned}$$

A central limit theorem

Suppose that $EX_1^2 < \infty$. When $t \to \infty$, we have

$$\frac{Z_t - EZ_t}{\sqrt{var(Z_t)}} = \frac{Z_t - \lambda t EX_1}{\sqrt{\lambda t\, EX_1^2}} \to_D N(0,1).$$

Central limit theorem usually provides an approximate estimate for certain probability.

Example 9.6 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 $1/6, 1/3, 1/3, 1/6$.

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

Solution: Let X_i denote the number of people in i-th family and $\{N_t, t \geq 0)\}$ be a Poisson process with rate $\lambda = 2$. Then $Z_t = \sum_{i=1}^{N_t} X_i$ is the total number of people migrating to an area by the time week t.

It follows that

$$\begin{aligned} E(Z_{50} &= 50 \times 2 \times EX_1 = 250; \\ var(Z_{50}) &= 50 \times 2 \times EX_1^2 = 50 \times 2 \times 43/6. \end{aligned}$$

We want to find the probability:

$$\begin{aligned} P(Z_{50} > 240) &= P\Big(\frac{Z_{50} - 250}{\sqrt{4300/6}} \geq \frac{240 - 250}{\sqrt{4300/6}}\Big) \\ &\sim 1 - \Phi(-0.39) \sim 0.65 \end{aligned}$$

The following two compound poisson processes are often used in modelling repairable inventory system by US Air Force where the interarrival time of successive demand occurrences follows a Poisson distribution, and the number of units (for instance, aircraft engines) demanded at each demand occurrence is the random variable X_j.

Example 9.6. Let $\{N_t, t \geq 0)\}$ be a homogeneous Poisson process with rate λ. Let $X, X_1, X_2, ...$ be i.i.d random variables independent of $N_t, t \geq 0$. Consider the compound Poisson process

$$Z_t = \sum_{k=1}^{N_t} X_k, \quad t \geq 0$$

1. X is a geometric distribution: i.e.,

$$P(X = k) = (1-p)p^{k-1}, \quad k = 1, 2, ..,$$

   In this situation, Z_t is called the geometric Poisson process.

   We have $EX = 1/p$ and $var(X) = (1-p)/p^2$.

2. X has the distribution:

$$P(X = k) = \frac{1}{k \log q} \left(p/q\right)^k, \quad k = 1, 2, ..,$$

   where $q = p + 1 > 1$. In this situation, Z_t is called the Logarithmic Poisson process, which is convenient for modelling demand arrival processes of repairable items.

   We have $EX = \frac{1}{\log q}(p/q)\,(1-p/q)^{-1}$ and $var(X) = \frac{1}{\log^2 q}\,(1-p/q)^{-1}$.

For more details, we refer to Kao (1997), pages 74-76.