The Poisson distribution

A discrete random variable N has an Poisson distribution with a parameter $\lambda > 0$, denoted by $N \sim \mathrm{Po}(\lambda)$, if its probability mass function is given by

$$p_n = \begin{cases} \frac{e^{-\lambda}\lambda^n}{n!}, & \text{for } n = 0, 1, \ldots \\ 0, & \text{for } n < 0. \end{cases}$$

The mean and variance of N are both equal to λ.

The exponential distribution

A random variable T has an exponential distribution with parameter $\lambda > 0$ (called the rate), denoted by $T \sim \exp(\lambda)$, if its distribution function is

$$F_T(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{for } t \geq 0, \\ 0, & \text{for } t < 0. \end{cases}$$

It follows that the probability density function of T is

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{for } t \geq 0, \\ 0, & \text{for } t < 0. \end{cases}$$

The mean of T is $1/\lambda$ and the variance of T is $1/\lambda^2$.

The Poisson Process

The Poisson distribution arises as the limit of the binomial distribution.

- If $X_n \stackrel{d}{=} \mathrm{Bin}(n, \lambda/n)$ and $N_\lambda \stackrel{d}{=} \mathrm{Po}(\lambda)$, then for $k = 0, 1, \ldots,$

$$\lim_{n\to\infty} P(X_n = k) = P(N_\lambda = k).$$

The exponential distribution arises as the limit of the geometric distribution.

- If $Y_n \stackrel{d}{=} \mathrm{Geo}(\lambda/n)$ and $T_\lambda \stackrel{d}{=} \mathrm{Exp}(\lambda)$, then for $t > 0$,

$$\lim_{n\to\infty} P(Y_n/n \leq t) = P(T_\lambda \leq t).$$

The Poisson Process

Definition:

A nonnegative integer-valued process $\{N_t : t \geq 0\}$ is a Poisson process with a rate λ if

- it has independent increments on disjoint intervals: for $k \geq 2$ and $0 \leq s_1 < t_1 \leq s_2 < \cdots < t_k$,

$$N_{t_1} - N_{s_1}, \ldots, N_{t_k} - N_{s_k}$$

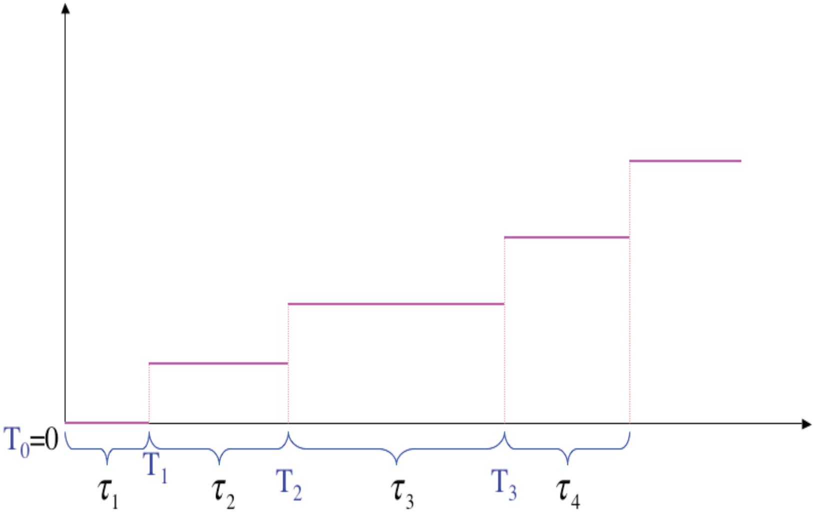
  are independent variables.
- For each $t \geq 0$, $N_t \sim \mathrm{Po}(\lambda t)$

Properties of the Poisson Process

- $N_{t+s} - N_t \sim \mathrm{Po}(\lambda s)$.
- Furthermore if, for fixed h, we define $N_t^* = N_{t+h} - N_h$, then N^* is a Poisson process with rate λ. [This is true even if h is a random variable independent of $(N_s, s > 0)$.]

The Poisson Process

A trajectory



Poisson Process Empirical Data

- Earthquakes
- Grazing animals head raises
- Goals in the world cup
- Horse kick deaths in past wars

The Poisson Process

Let $T_j = \min\{t : N_t = j\}$, the time of jth jump and define $\tau_j = T_j - T_{j-1}$ the time between the $(j-1)^{\mathrm{st}}$ jump and the j^{th} jump.

Theorem

$\{N_t : t \geq 0\}$ is a Poisson process with rate λ if and only if $\{\tau_j\}$ are independent $\exp(\lambda)$ random variables.

Proof

The key to the proof is to observe that the event $T_j \leq t$ is the same as $N_t \geq j$. That is the waiting time until the jth event is less than or equal to t if and only if there are j or more events in time t. Assume that $\{N_t : t \geq 0\}$ is a Poisson process. Then $P(T_1 \leq t) = P(N_t \geq 1) = 1 - P(N_t = 0) = 1 - e^{-\lambda t}$, from which we see that the waiting time until the first event is exponentially-distributed.

The Poisson Process

Furthermore, we have

$$\begin{aligned} P(T_j \leq t) &= P(N_t \geq j) \\ &= \sum_{k=j}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \\ &= 1 - \sum_{k=0}^{j-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}. \end{aligned}$$

This final expression is the distribution function for gamma distribution with parameter k and rate λ. You can check this by differentiating to get the density function
$f_{T_j}(t) = e^{-\lambda t} \lambda^j t^{j-1}/(j-1)!$.
So the waiting time until the jth event is the sum of j independent exponentially-distributed inter-event times with parameter λ.

The Poisson Process

This argument also holds in reverse.
Assuming that τ_1 is exponentially-distributed with parameter λ, we know that $P(T_1 \le t) = 1 - e^{-\lambda t}$, which tells us that $P(N_t = 0) = e^{-\lambda t}$.
Furthermore, for $j > 1$, if $\{\tau_1, \ldots, \tau_j\}$ are independent and exponentially-distributed, then T_j has an Gamma distribution with parameters λ and j. So

$$P(N_t \ge j) = P(T_j \le t) = 1 - \sum_{k=0}^{j-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!},$$

which tells us that N_t has a Poisson distribution with parameter λt.

The Poisson Process

We still need to show that $N_{t_i} - N_{s_i}$ are independent over sets $[s_i, t_i)$ of disjoint intervals.

This follows from the memoryless property of the exponential distribution (so the remaining time from s_i doesn't depend on $s_i - T_{N_{s_i}}$) and the independence of the τ_j.

Order statistics

For random variables $\xi_1, \xi_2, \ldots, \xi_k$, denote by $\xi_{(i)}$ the *i*th smallest of them. Then $\xi_{(1)}, \xi_{(2)}, \ldots, \xi_{(k)}$ are called the order statistics associated with $\xi_1, \xi_2, \ldots, \xi_k$.

For example, if we sample these random variables and find that $\xi_1 = 1.3, \xi_2 = 0.9, \xi_3 = 0.7, \xi_4 = 1.1$ and $\xi_5 = 1.5$, then $\xi_{(1)} = 0.7, \xi_{(2)} = 0.9, \ldots, \xi_{(5)} = 1.5$.

Order statistics play a very important role in applications. For example, the maximum likelihood estimator of θ for a sample $\xi_1, \xi_2, \ldots, \xi_k$ from the uniform $[0, \theta]$ distribution is $\xi_{(k)}$.

Examples

- If X_1, X_2 and X_3 are independent and identically-distributed random variables taking values 1, 2 and 3 each with probability 1/3, find the joint distribution of $(X_{(1)}, X_{(2)}, X_{(3)})$.
- If $Y_1, Y_2, \ldots, Y_k$ are independent random variables, uniformly distributed on $[0, t]$, show that, for $i \leq k$ and $x \leq t$, the distribution function of the order statistic $Y_{(i)}$ is given by

$$F_{Y_{(i)}}(x) = \sum_{\ell=i}^{k} \binom{k}{\ell} (x/t)^{\ell} (1 - x/t)^{k-\ell}.$$

Order Statistics

- In general if $Y_1, Y_2, \ldots, Y_k$ are independent random variables with distribution function F and density f, the distribution function of $Y_{(i)}$ is

$$F_{Y_{(i)}}(x) = \sum_{\ell=i}^{k} \binom{k}{\ell} F(x)^{\ell}(1 - F(x))^{k-\ell}.$$

  and the density is

$$\begin{aligned} f_{Y_{(i)}}(x) &= \binom{k}{i-1}(k-i+1)F(x)^{i-1}f(x)(1-F(x))^{k-i} \\ &= \binom{k}{i} iF(x)^{i-1}f(x)(1-F(x))^{k-i}. \end{aligned}$$

Order Statistics

Similarly the joint densities are for $1 \le r \le k$ and $x_1 < \cdots < x_r$,

$$
\begin{aligned}
& f_{Y_{(i_1)},\dots,Y_{(i_r)}}(x_1,\dots,x_r) \\
& \quad = \binom{k}{i_1-1,1,i_2-i_1-1,1\cdots,1,k-i_r} \\
& \qquad \times \prod_{j=1}^{r} f(x_j) \prod_{j=0}^{r} (F(x_{j+1}) - F(x_j))^{i_{j+1}-i_j-1},
\end{aligned}
$$

where $\binom{\ell}{a_1,\cdots,a_j}$ is the number of ways to choose subsets of sizes $a_1,\dots,a_j$ from a set of size ℓ and for the sake of brevity we set $x_0 = -\infty$ and $x_{r+1} = \infty$ so $F(x_0) = 0$ and $F(x_{r+1}) = 1$.

Order Statistics

In particular, for $r = k$, $x_1 < \cdots < x_r$,

$$f_{Y_{(1)},\ldots,Y_{(k)}}(x_1, \ldots, x_k) = k! \prod_{j=1}^{k} f(x_j).$$

Theorem

The conditional distribution of $(T_1, \cdots, T_k)$ given that $N_t = k$ is the same as the distribution of order statistics of a sample of k independent and identically-distributed random variables uniformly distributed on $[0, t]$. Thus,

$$(T_1, \cdots, T_k)|(N_t = k) \stackrel{d}{=} (U_{(1)}, \cdots, U_{(k)})$$

where $U_1, \cdots, U_k$ are independent Uniform $(0, t)$.
The same representation holds for the conditional distribution of $(T_1, \cdots, T_k)$ given that $T_{k+1} = t$.

Proof

According to our derivation for order statistics, $(U_{(1)}, \cdots, U_{(k)})$ has density $k!t^{-k}$ for $0 < x_1 < \cdots < x_k < t$. So we show the LHS has the same density:

$$
\begin{aligned}
&\mathbb{P}(T_1 \in dx_1, \cdots, T_k \in dx_k | N_t = k) \\
&\quad = \frac{\mathbb{P}(\tau_1 \in dx_1, \tau_2 \in d(x_2 - x_1), \ldots, \tau_k \in d(x_k - x_{k-1}), \tau_{k+1} > t - x_k)}{\mathbb{P}(N_t = k)} \\
&\quad = \frac{\lambda e^{-\lambda x_1} \lambda e^{-\lambda(x_2 - x_1)} \cdots \lambda e^{-\lambda(x_k - x_{k-1})} e^{-\lambda(t - x_k)}}{(\lambda t)^k e^{-\lambda t}/k!} \\
&\quad = k!t^{-k}.
\end{aligned}
$$

The proof of the last sentence is virtually the same.

The Poisson Process

The theorem implies that if $\tau_1, \ldots$ are iid exponential variables, then

$$\left(\frac{\tau_1}{\sum_{j=1}^{n+1} \tau_j}, \frac{\tau_1 + \tau_2}{\sum_{j=1}^{n+1} \tau_j}, \ldots, \frac{\sum_{j=1}^{n} \tau_j}{\sum_{j=1}^{n+1} \tau_j} \right)$$

have the same distribution as uniform order statistics.

Superposition of Poisson processes

Let $\{N_t : \ t \geq 0\}$ and $\{M_t : \ t \geq 0\}$ be two independent Poisson processes with rates λ and μ respectively and $L_t = N_t + M_t$. Then $\{L_t : \ t \geq 0\}$ is a Poisson process with rate $\lambda + \mu$.

Proof

- By independence, $L_t \sim \mathrm{Po}(\lambda t + \mu t)$.
- For disjoint $[s_1, t_1]$ and $[s_2, t_2]$,

$$\begin{aligned} L_{t_1} - L_{s_1} &= (N_{t_1} - N_{s_1}) + (M_{t_1} - M_{s_1}) \\ L_{t_2} - L_{s_2} &= (N_{t_2} - N_{s_2}) + (M_{t_2} - M_{s_2}) \end{aligned}$$

  which are independent because of the same property of $\{N_t : \ t \geq 0\}$ and $\{M_t : \ t \geq 0\}$.

Example

A shop has two entrances, one from East St, the other from West St. Flows of customers through the two entrances are independent Poisson processes with rates 0.5 and 1.5 per minute, respectively.

- What is the probability that no new customers enter the shop in a fixed three minute time interval?
- What is the mean time between arrivals of new customers?
- What is the probability that a given customer entered from West St?

Thinning of a Poisson process

Suppose in a Poisson process $\{N_t : \; t \geq 0\}$ each 'customer' is 'marked' independently with probability p. Let M_t count the number of 'marked customers' that arrive on $[0, t]$.

Theorem

The processes $\{M_t : \; t \geq 0\}$ and $\{N_t - M_t : \; t \geq 0\}$ are independent Poisson processes with rates λp and $\lambda(1 - p)$ respectively.

Proof

$$
\begin{aligned}
P(M_t = j, N_t - M_t = k) &= P(M_t = j, N_t = k + j) \\
&= P(M_t = j | N_t = k + j) P(N_t = k + j) \\
&= \binom{k+j}{j} p^j (1-p)^k \frac{e^{-\lambda t} (\lambda t)^{k+j}}{(k+j)!} \\
&= \frac{e^{-p\lambda t} (p\lambda t)^j}{j!} \frac{e^{-(1-p)\lambda t} ((1-p)\lambda t)^k}{k!}.
\end{aligned}
$$

Example

The flow of customers to a shop is a Poisson process with rate 25 customers per hour. Each of the customers has a probability $p = 0.8$ of making a purchase.

- What is the probability that all customers who enter the shop during the time interval from 11.00 am to 11.15 am make a purchase?
- What is the probability that, conditional on there being two customers that made a purchase during that period, all customers who enter the shop during the time interval make a purchase?

The Poisson Process

Consider an integer-valued stochastic system $\{X_t\}$ fed by 'customers' arriving according to a counting process $\{N_t\}$. A queue would be a good example.

Assume that $\{X_t\}$ is ergodic, so that there is a stationary regime with $P_{\text{st}}(X_t = j) = \pi_j$ for $j \geq 0$. We saw that π_j is also the proportion of the time the system spends in state j in the long run. That is, with probability one, the proportion of time spent in state j during $[0, t]$ approaches π_j as $t \to \infty$.

The Poisson Process

Now assume that an arriving 'customer' sees the system in state j with prob π_j^* for $j \geq 0$.

An interesting question is 'When is $\pi_j = \pi_j^*$?

That is, when does an arriving customer observe the system in its stationary state?

The Poisson Process

To see the issues around this question, consider the following example.

- You own a PC and you are the only user.
- The PC has two states: $0 =$ free and $1 =$ occupied.
- The counting process $\{N_t\}$ records the number of moments that you come to use the PC in $[0, t]$.
- Since it is your own PC, when you arrive, it is always ready for you, so $\pi_0^* = 1$ and $\pi_1^* = 0$.
- However, in general, π_0 is the proportion of time that the PC is free is which is less than one (unless you never use the PC).
- Hence, in this case, $\pi_j \neq \pi_j^*$.

PASTA: (Poisson Arrivals See Time Averages)

Now consider a stationary version of such a system $\{X_t\}$ where arrivals constitute a Poisson process with intensity λ. In this case, $\pi_j = \pi_j^*$, that is Poisson arrivals see time averages.

Why?

- Suppose an arriving customer A arrives at time $s > 0$.
- The state of the system at time s, that is X_s, depends on A's view of the customer flow before herself.
- In A's eyes (that is, conditional on $N_s - N_{s^-} = 1$), the customer flow of others is the same as the original $\{N_t\}$, with herself being an outsider.
- Thus she observes the system as if she is an outsider.

The Poisson Process

To see this, observe that, for all $k \geq 1$ and
$a_1 < b_1 \leq a_2 < b_2 \leq ... \leq a_k < b_k < s$,

$$
\begin{aligned}
& \lim_{t \downarrow 0} P\left(\bigcap_{i=1}^{k} \{N_{b_i} - N_{a_i} = x_i\} \middle| N_{s+t} - N_{s-t} = 1\right) \\
= \; & P\left(\bigcap_{i=1}^{k} \{N_{b_i} - N_{a_i} = x_i\}\right) \\
= \; & \prod_{i=1}^{k} \frac{e^{-\lambda(b_i - a_i)}(\lambda(b_i - a_i))^{x_i}}{x_i!}
\end{aligned}
$$

for all non-negative integers $x_1, ..., x_k$.
This is the same as the distribution of a process in which A had not arrived at time s.

What about the PC example?

The arrival process is not Poisson: when an arrival has occurred (that is, you have started to work with you PC) no other arrival will occur until you switch off and come back again.

So the process of arrivals is not a Poisson process, and does not have independent increments.

The Compound Poisson Process

Suppose that $\{N_t : t \geq 0\}$ is a Poisson process and $\{X_i : i \geq 1\}$ are independent and identically-distributed random variables, which are also independent of $\{N_t : t \geq 0\}$.

For $t \geq 0$, define $Y_t = \sum_{j \leq N_t} X_j$. Then $\{Y_t : t \geq 0\}$ is called a compound Poisson process.

It can be shown that $\{Y_t : t \geq 0\}$ has independent increments and it is possible to compute the distribution of Y_t by conditioning on N_t.

Example

Suppose claims made on an insurance company occur according to a Poisson process with rate λ, and each policy holder carries a policy for an amount X_k.

Assume $X_1, X_2, \ldots$ are independent and identically-distributed, and the number of claims and the size of claims are independent.

Calculate the mean and variance of the total amount of claims on the company up to time t.