Poisson Processes

Stochastic Processes

UC3M

Feb. 2012

Exponential random variables

A random variable T has exponential distribution with rate $\lambda > 0$ if its probability density function can been written as

$$f(t) = \lambda e^{-\lambda t} \mathbf{1}_{(0,+\infty)}(t)$$

We summarize the above by $T \sim exp(\lambda)$.

The cumulative distribution function of a exponential random variable is

$$F(t) = P(T \leq t) = 1 - e^{-\lambda t} \mathbf{1}_{(0,+\infty)}(t)$$

And the tail, expectation and variance are

$$P(T > t) = e^{-\lambda t}, \quad E[T] = \lambda^{-1}, \quad \text{and} \quad Var(T) = E[T] = \lambda^{-2}$$

The exponential random variable has the lack of memory property

$$P(T > t + s | T > t) = P(T > s)$$

Exponencial races

In what follows, $T_1, \ldots, T_n$ are independent r.v., with $T_i \sim exp(\lambda_i)$.

P1:

$$\min(T_1, \ldots, T_n) \sim exp(\lambda_1 + \cdots + \lambda_n)$$

.

P2

$$P(T_1 < T_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

P3:

$$P(T_i = \min(T_1, \ldots, T_n)) = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}$$

P4: If $\lambda_i = \lambda$ and $S_n = T_1 + \cdots + T_n \sim \Gamma(n, \lambda)$. That is, S_n has probability density function

$$f_{S_n}(s) = \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} \mathbf{1}_{(0,+\infty)}(s)$$

The Poisson Process as a renewal process

Let $T_1, T_2, \ldots$ be a sequence of i.i.d. nonnegative r.v. (interarrival times). Define the **arrival times**

$$S_n = T_1 + \cdots + T_n \text{ if } n \geq 1 \quad \text{and} \quad S_0 = 0.$$

The process

$$N(t) = \max\{n : S_n \leq t\},$$

is called Renewal Process.

If the common distribution of the times is the exponential distribution with rate λ then process is called Poisson Process of with rate λ.

Lemma.

$$N(t) \sim \text{Poisson}(\lambda t)$$

and

$$N(t+s) - N(s),\ t \geq 0, \text{ is a Poisson process independent of } N(s),\ t \geq 0$$

The Poisson Process as a Lévy Process

A stochastic process $\{X(t), t \geq 0\}$ is a Lévy Process if it verifies the following properties:

1. $X(0) = 0$
2. For all $s, t \geq 0, \quad X(t+s) - X(s) \sim X(t)$, and
3. $X(t)$ has independent increments. This is, for any sequence $0 \leq t_0 < t_1 < \cdots < t_m$ we have $X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$ are independent random variables.

If $\{X(t)\}$ is a Lévy Process then one may construct a version of such that is almost surely right continuous with left limits.

Theorem. $N(t)$ is Poisson Process with rate λ if and only if

1. $N(0) = 0$
2. $N(t+s) - N(s) \sim Poisson(\lambda t)$, and
3. $N(t)$ has independent increments.

The Poisson Process as a Birth Process

Many process in nature may change their values at any instant of time rather than at certain epochs only. Depending of the underlying random mechanism, thsee processes may or may not satisfy a Markov property. Before attempting to study a general theory of continuous Markov Processes we explore a one simple in detail.

Theorem. $N(t)$ is Poisson Process with rate λ if and only if it is a process taking values in $\{0, 1, 2, 3, \ldots\}$ such that

1. $N(0) = 0$ and, If $s < t$ then $N(s) \leq N(t)$
2. $$P(N(t+h) = n+m | N(t) = n) = \begin{cases} \lambda h + o(h) & \text{if } m = 1 \\ o(h) & \text{if } m > 1 \\ 1 - \lambda h + o(h) & \text{if } m = 0 \end{cases}$$
3. If $s < t$ then the number of births in the interval $(s, t]$, namely $N(t) - N(s)$, is independent of the number of births during $[0, s]$, this is $N(s)$.

Law of Rare Events and Poisson Process

The Poisson Process is important for applications partially by the Poisson approximation to the sum of independent Bernoulli random variables.

Theorem. If $np_n \to \lambda > 0$, when $n \to to + \infty$, then the *binomial*(n, p_n) distribution is approximately *Poisson*(λ). In particular, the *binomial*$(n, \lambda/n)$ distribution is approximately *Poisson*(λ). Moreover, Let $\varepsilon_1, \varepsilon_1, \ldots$ be independent Bernoulli random variables, with $P(\varepsilon_i = 1) = p_i$. Let $S_n = \varepsilon_1 + \cdots + \varepsilon_n$ and $\lambda = p_1 + \cdots + p_n$. Then

$$|P(S_n = k) - P(Poisson(\lambda = k)| \leq \sum_{i=1}^{n} p_i^2$$

The latter is consequence of the following important result: If X and Y are integer valued random variables then, for any set A

$$|P(X \in A) - P(Y \in A)| \leq \frac{1}{2} \sum_k |P(X = k) - P(Y = k)|$$

The right-hand side is called the **total variation distance** between X and Y.

Conditioning

Teorema. Let be the arrival times of a Poisson Process with rate λ. Let $V_1, \ldots, V_n$ be the order statistics of n i.i.d. random variables, uniformly distributed on $(0, t)$. Then, the conditional distribution of $(T_1, \ldots, T_n)$ given $N(t) = n$ is the distribution of $(V_1, \ldots, V_n)$. This is, for any $t_1 < \cdots < t_n < t$

$$P(S_1 \le t_1, \ldots, S_n \le t_n | N(t) = n) = P(V_1 \le t_1, \ldots, V_n \le t_n)$$

Note: The joint probability density of $(V_1, \ldots, V_n)$ is

$$n! t^{-n} \mathbf{1}_{\{0 < t_1 < \cdots < t_n < t\}}.$$

The above remarkable fact implies another important result

Teorema. If $0 < s < t$, and $0 \le m \le$,

$$P(N(s) = m | N(t) = n) = \binom{n}{k} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m} \quad m = 0, 1, \ldots, n$$

That is, the conditional distribution of $N(s)$ given $N(t) = n$ is $Binomial(n, s/t)$.

Nonhomogeneous Poisson Process

Consider $\lambda : \mathbb{R}^+ \to \mathbb{R}^+$. We say $\{N(t), t \geq 0\}$ is a Nonhomogeneous Poisson Process with rate $\lambda(r)$ if

1. $N(0) = 0$
2. $N(t) - N(s) \sim Poisson\left(\int_s^t \lambda(r)dr\right)$, para cualquier $0 \leq s < t$
3. $N(t)$ has independent increments.

Remarks:

1. For the Nonhomogeneous Poisson Process, the interarrival times are no longer exponentially distributed or independent.
2. Let $X(t)$ be a nonhomogeneous Poisson process of rate $\lambda(t)$ and define

$$\Lambda(t) = \int_0^t \lambda(u)du.$$

   Then, $S_1, S_2, \ldots$ are the arrival times corresponding to $X(t)$ if and only if $\Lambda(S_1), \Lambda(S_2), \ldots$ are the corresponding arrival times of a homogeneous Poisson process of rate one. This result provides a method of generation nonhomogeneous Poisson Processes.

Compound Poisson Process

Given a Poisson Process $X(t)$ of rate λ, suppose that each arrival time has associated with a random value or cost. Assume the successive values $Y_1, Y_2, \ldots$ are i.i.d. and independent of the Poisson Process. The process $S(t)$ defined by

$$S(t) = \begin{cases} 0 & \text{si } N(t) = 0 \\ Y_1 + \cdots + Y_{N(t)} & \text{si } N(t) \geq 1 \end{cases}$$

is called Compound Poisson Porcess.

The distribution function for a compound Poisson process can be represented explicitly. In particular, if $E[Y_i] = \mu$, $Var(Y_i) = \sigma^2$ and $Y_1 + \cdots + Y_n \sim S_n$ then

$$E[S(t)] = \lambda\mu t$$

$$Var(S(t)) = \lambda(\sigma^2 + \mu^2)t$$

and

$$P(S(t) \leq s) = \sum_{n=0}^{\infty} e^{\lambda t} \frac{(\lambda t)^n}{n!} P(S_n \leq s)$$

Collective Risk Model

The classical model describes an insurance company who experiences two opposing cash flows: incoming cash premiums and outgoing claims. Premiums arrive a constant rate $c > 0$ from customers and claims arrive according to a Poisson process with intensity λ and are independent and identically distributed non-negative random variables with distribution F and mean μ. So for an insurer who starts with initial surplus x

$$C(t) = x + ct - S(t)$$

$S(t)$ being a compound Poisson process t.

The central object of the model is to investigate the probability that the insurer's surplus level eventually falls below zero (making the firm bankrupt). This quantity, called the probability of ultimate ruin, is defined as

$$\psi(x) = P(\tau < \infty | C(0) = x)$$

τ being the time of ruin defined by

$$\tau = \inf_t \{t > 0 : C(t) < 0\}$$

Thinning

As we did for the Compound Poisson Process, assume $Y_1, Y_2, \ldots$ are i.i.d. and independent of a Poisson Process $N(t)$ with rate λ. But now, Y_i is the type of value at the i^{th}-arrival time. We are interested in the number of $Y_i = y$ up to time t, that we will call $N_y(t)$.

This procedure, taking one Poisson process and splitting in two or more by using an i.i.d. sequence is called thinning.

Theorem. $N_y(t)$ are independent Poisson Processes with rate $\lambda P(Y_i = y)$

There are two surprises here:

1. The resulting processes are Poisson
2. They are independent

Superposition

Superposition goes in the inverse direction than thinning. Since a Poisson process can be split into independent Poisson processes, it should be intuitive that when independent Poisson processes are put together, the sum (superposition) is Poisson. Remember that if $X_1, \ldots, X_n$ are independent Poisson random variables, with $X_i \sim Poisson(\lambda_i)$, then

$$X_1 + \cdots + X_n \sim Poisson(\lambda_1 + \cdots + \lambda_n)$$

So,

Theorem. Suppose $N_1(t), \ldots, N_k(t)$ are independent Poisson processes with rates $\lambda_1, \ldots, \lambda_k$, then $N_1(t) + \cdots + N_k(t)$ is a Poisson process with rate $\lambda_1 + \cdots + \lambda_k$.

The ideas of compounding, thinning and superposition are very useful in simulation of continuous Markov Chains.

Spatial Poisson Process

Let $B \subset \mathbb{R}^d$ be a bounded set and let $U_1, \ldots, U_n$ be i.i.d. uniformly distributed on B. The process $\mathcal{U}_n = \{U_1, \ldots, U_n\}$ is called the binomial process on B and it describes a totally random set of points on B.

How to extend this notion of total randomness to points in the entire $\mathbb{R}^d$? The is no uniform probability measure on $\mathbb{R}^d$. The Poisson point process $\mathcal{P}$ gieves a way to solve this problem. Desiderata:

1. The number of points in a set B, namely $\mathcal{P}(B)$, is a Poisson random variable with expected value $= \lambda m_d(B)$, $m_d(B)$ being the Lebesgue measure of B (area for $d = 2$, volume for $d = 3$, etc.)
2. if A and B are disjoints then $\mathcal{P}(A)$ and $\mathcal{P}(B)$ should be independent.

Properties

1. Conditioning to have n points in a given set B we obtain the binomial process $\mathcal{U}_n$ (exhibit a complete spatial random pattern).
2. The translated process and the rotated process have the same distribution (invariant under translation and rotation).

Renewals

The major reason for studying renewal processes is that many complicated processes have randomly occurring instants at which the system returns to a state probabilistically equivalent to the starting state. These embedded renewal epochs allow us to separate the long term behavior of the process (which can be studied through renewal theory) from the behavior of the actual process within a renewal period.

The basic elements involved in a Renewal Process are:

- The interarrivals times $T_1, T_2, \ldots$. They are i.i.d. nonnegative r.v. with common distribution F.
- The arrival times $S_n = T_1 + \cdots + T_n$ if $n \geq 1$ and $S_0 = 0$.
- The counting process $N(t) = \max\{n : S_n \leq t\}$.

When the distribution of the interarrival times is not exponential (i.e. when the process is not a Poisson Process) then

1. the process has not independent increments and,
2. in general, it has not stationary increments.

Distribution of $N(t)$

Note that

$$N(t) \geq n \quad \text{if and only if} \quad S_n \leq t.$$

It follows that

$$P(N(t) \geq n) = P(S_n \leq t) = F_n(t),$$

where $F_1(t) = F(x)$, and

$$F_{n+1}(t) = \int_0^t F_n(t-s)F(s)ds.$$

So, the **renewal function** $m(t) = E[N(t)]$ can be written as

$$m(t) = \sum_n F_n(t)$$

and it satisfies the **renewal equation**

$$m(t) = F(t) + \int_0^t m(t-x)dF(x)$$

Limit behavior

SLLN:

$$\frac{N(t)}{t} \to \frac{1}{\mu} \quad a.s., \quad \text{with} \quad \mu = E[T_i].$$

The asymptotic result also holds for the expectation and is known as **elementary renewal theorem:**

$$\frac{m(t)}{t} \to \frac{1}{\mu}$$

Second order properties of $m(t)$ are harder to find and require of an additional definition: Call a random variable X and its distribution F **arithmetic** with span $\lambda > 0$ if X take values in the set $\{m\lambda : m = 0, \pm 1, \pm 2, \dots\}$ with probability 1, and λ is maximal with this property.
Renewal Theorem: If the interarrival times are not arithmetic then

$$m(t+h) - m(t) \to \frac{h}{\mu}, \quad \text{for all } h.$$

If T_i is arithmetic with span λ then the above asymptotic results holds whenever h is a multiple of λ.

Renewal-Reward process

There are situations in which, along with a renewal process, *one can make some money* between arrival times. In this context, the interarrival times are called cycles.

The model: Suppose that a reward R_n is earned during the n^{th} cycle. Assume:

1. R_n may depend on the length of the n^{th} cycle.
2. $R_1, R_2, \ldots$ are i.i.d.

And consider the total reward earned up to time t, denoted $R(t)$.

Renewal-Reward Theorem:

$$\frac{R(t)}{t} \to \frac{E[R_i]}{\mu} \quad a.s.$$

Under bounding conditions, the limit holds also for the expectation.

$$\frac{E[R](t)}{t} \to \frac{E[R_i]}{\mu}$$

Regenerative process

The asymptotic result discussed above holds in very general setups. A very general framework can be described in terms of Regenerative Processes.

A stochastic process $\{X(t), t \geq 0\}$ is regenerative if there are random times $S_1 < S_2 < S_3 < \dots$ (called regeneration epochs) such that

1. $\{X(S_k + t), t \geq 0\}$ is independent of $\{X(t), 0 \leq t < S_k\}$
2. $\{X(S_k + t), t \geq 0\}$ has the same distribution as $\{X(t), t \geq 0\}$

It is very common to be interested in the long-run behavior of

$$\frac{1}{t}\int_0^t X(t)dt$$

In such situations, it is useful the modeling

- $C_n = S_n - S_{n-1}$ are cycles,
- $R_n = \int_{S_{n-1}}^{S_n} X(t)dt$ are rewards, and
- $N(t)$ is the number of regeneration epochs up to time t

Application: asymptotics for the excess variable

Consider the waiting time (or excess or residual life) until the next event, namely

$$\gamma_t = S(N(t)+1) - t$$

A well known result in Renewal theory is

$$P(\gamma_t \leq x) \to \frac{1}{\mu}\int_0^x P(T_i > y)dy \quad \text{and} \quad E[\gamma_t] \to \frac{E[T^2]}{2\mu}$$

The proof is advanced, outside scope of this course. But we can derivate more easily the next result

$$\frac{1}{t}\int_0^t P(\gamma_s \leq x)ds \to \frac{1}{\mu}\int_0^x P(T_i > y)dy \quad \text{and} \quad \frac{1}{t}\int_0^t E[\gamma_s]ds \to \frac{E[T^2]}{2\mu}$$

Hint: take, according each case,

$$R_n = \int_{S_{n-1}}^{S_n} \mathbf{1}_{[0,x]}(\gamma_t)dt \quad \text{and} \quad R_n = \int_{S_{n-1}}^{S_n} \gamma_t dt$$

and apply the asymptotics discussed for regenerative process

Applications to queueing systems