Lecture 17 : Poisson Processes – Part I

STAT205 *Lecturer: Jim Pitman* Scribe: *Matias Damian Cattaneo* <cattaneo@econ.berkeley.edu>

17.1 The Poisson Distribution

Define $S_n = X_1 + X_2 + ... + X_n$. In a very simple setup, the X_i are independent indicators with

$$\begin{aligned}\mathbb{P}[X_i = 1] &= p \\ \mathbb{P}[X_i = 0] &= 1 - p\end{aligned}$$

We know that the distribution of S_n is Binomial (n, p). So we have

$$\begin{aligned}\mathbb{P}[S_n = k] &= \binom{n}{k} p^k (1-p)^{n-k}, \\ \mathbb{E}[S_n] &= np, \text{ and} \\ \mathbf{V}ar[S_n] &= np(1-p).\end{aligned}$$

For fixed p, we have that

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)$$

as $n \to \infty$.

Now we let $n \to \infty$ and choose p_n small so that $np_n = \lambda$. We see that

$$\begin{aligned}\mathbb{P}[S_n = 0] &= (1-p)^n = \left(1 - \frac{\lambda}{n}\right)^n \longrightarrow e^{-\lambda}, \\ \mathbb{P}[S_n = 1] &= np(1-p)^{n-1} = \lambda \left(1 - \frac{\lambda}{n}\right)^{n-1} \longrightarrow \lambda e^{-\lambda},\end{aligned}$$

and in general

$$\begin{aligned}\mathbb{P}[S_n = k] &= \binom{n}{k} p^k (1-p)^{n-k} \\ &\approx \frac{n^k}{k!} p^k (1-p)^n \longrightarrow \frac{\lambda^k}{k!} e^{-\lambda},\end{aligned}$$

which is the mass function of a Poisson distribution.

Definition 17.1 *The* Poisson distribution *with parameter λ is given by mass function*

$$P_\lambda(k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Observe that

$$\sum_{k=0}^{\infty} P_\lambda(k) = 1$$

Summing up, we see that as $n \to \infty$, if we let $p \to 0$ such that $np = \lambda \in [0, \infty)$,

$$S_n \xrightarrow{d} \text{Poisson}(\lambda)$$

Observe that we can relax the assumption that the indicators are identically distributed and extend this result to the triangular array setup. Let the $X_{n,i}$s be taken such that

$$\mathbb{P}[X_{n,i} = 1] = p_{n,i}$$

Note we have $\mathbb{E}[S_n] = \sum_{i=1}^{n} p_{ni}$ and assuming that as $n \to \infty$ we have $\sum_{i=1}^{n} p_{n,i} \longrightarrow \lambda$ and $\max_i p_{n,i} \longrightarrow 0$, we obtain the result:

$$S_n \xrightarrow{d} \text{Poisson}(\lambda)\,.$$

The proof of this result is formalized in [1].

Observe the following facts about the Poisson distribution:

1. As shown above, it is the limit of properly chosen binomial distributions.

2. The sum of independent Poisson random variables with parameters λ and υ is another Poisson random variable. This result can be written as follows:

$$\text{Poisson}(\lambda) * \text{Poisson}(\upsilon) = \text{Poisson}(\lambda + \upsilon)\,,$$

where we use the fact that for discrete distributions P and Q on $\{0, 1, 2, ...\}$,

$$(P + Q)(n) = \sum_{k=0}^{n} P(k)\,Q(n-k)$$

for $n = 0, 1, 2, ...$, the convolution formula for the distribution of sums of independent random variables.

17.2 Basics of Poisson Processes

We discuss the basics of Poisson processes here; for details, see [1].

Consider the positive numbers divided up into intervals of length 10^{-6}. We can think of this as time broken up into very short intervals. Consider a process $(X_1, X_2, \ldots)$ where X_i is 1 if a certain thing happens in the ith such time interval and 0 otherwise. Suppose further that these X_is are independent and are 1 with probability $\lambda \cdot 10^{-6}$.

The waiting time until the first occurrence of this thing is defined as

$$T_1 = \{first\ n: \ X_i = 1\}$$

here $X_i = 1$ with probability $\frac{\lambda}{n}$.

We see that

$$\mathbb{P}[T_1 > m] = (1-p)^m = [(1-p)^n]^{\frac{m}{n}} \longrightarrow e^{-\lambda \frac{m}{n}}$$

as $m \to \infty$, and thus

$$\mathbb{P}\left[\frac{T_1}{n} > t\right] = \mathbb{P}[T_1 > nt] \longrightarrow e^{-\lambda t}$$

as $m \to \infty$. Notice that this is an exponential distribution.

When $n = 10^6 \to \infty$ we obtain a *point process* (T_i), an increasing sequence of random variables where

$$\begin{aligned} T_1 &\sim \text{Exponential}(\lambda) \\ T_2 - T_1 &\sim \text{Exponential}(\lambda)\text{, independent of } T_1 \\ &\vdots \end{aligned}$$

so we get $T_1, (T_2 - T_1), (T_3 - T_2)$ are iid Exponential(λ).

Alternatively, the *counting process* is defined by

$$N_t = \#\{i : T_i \le t\} = \sum_{k=1}^{n} \mathbf{1}\{T_k \le t\}.$$

In general $\{N_t\}_{t\ge 0}$ is called a *Poisson process* with rate λ on $(0,\infty)$. It is not difficult to show that $N_t \sim \text{Poisson}(\lambda t)$.

It is important to note that almost by construction this process $(N_t, t \ge 0)$ has *stationary independent increments*: that is, for $0 < t_1 < t_2 < \ldots < t_n$, we have that $N(t_1), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$ are independent Poisson variables with means $\mu t_1, (t_2 - t_1)\mu, \ldots, (t_n - t_{n-1})\mu$.

17.3 Details of Poisson Processes

In this section we present details on Poisson processes.

17.3.1 The Poisson Process and its counting process

Definition 17.2 *A real valued process N_t, $t \geq 0$, is a* Poisson Process *with rate λ (or $PP(\lambda)$) if N_t has stationary independent increments, and for all $s, t \geq 0$, the increment $X_{t+s} - X_t$ is a Poisson random variable with parameter λs.*

Interpretation: Jumps of N_t are "arrivals" or "points".

Let $T_k = \inf\{t : \; N_t = k\}$. We say that $0 < T_1 < T_2 < \cdots$ are these arrival times. Note that $N_{T_k} = k$ and $N_{T_k^-} = k - 1$.

As a convention, we will always work with the right continuous version of $(N_t, t \geq 0)$ which is increasing and hence has a left limit a.s.

Theorem 17.3 *A counting process N_t, $t \geq 0$ (increasing, right continuous, left limits exists, jumps of 1 only) is a $PP(\lambda)$ if and only if the distribution of its jumps $(T_1, T_2, T_3, \ldots)$ satisfies that $(T_k - T_{k-1}, k = 1, 2, \ldots)$ is a sequence of i.i.d. Exponential(λ) random variables. ($T_0 = 0$.)*

Proof: For the "if" direction, see [1], Section 2.6 (c. Poisson process). The idea is given i.i.d. (W_k) such that $\mathbb{P}(W_k > t) = e^{-\lambda t}$, $t \geq 0$, we construct $(N_t, t \geq 0)$ by the formula:

$$\begin{aligned} T_k &= W_1 + \cdots + W_k \\ N_t &= \text{ number of } \{k : T_k \leq t\} \\ &= \sum_{k=1}^{\infty} \mathbf{1}\{T_k \leq t\} \end{aligned}$$

Informally, the $\{T_k\}$ are the points of the $PP(\lambda)$ denoted by N.[1]

The "only if" direction is easier. Suppose $N_t \sim$ Poisson (λ). By the Strong Markov Property (see [1, section 5.2]), we can deduce that $T_k - T_{k-1}$ are independent. We also can deduce: $\mathbb{P}(T_k > t) = \mathbb{P}(N_t < k) = \sum_{j=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$. Hence, by applying $\frac{d}{dt}$, we get $\mathbb{P}(T_k \in dt) = e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!} dt$.

[1]Friendly reference: "Probability", J.Pitman, Springer 1992.

This shows $T_k \sim gamma(k, \lambda)$. Since $\phi_{T_k - T_{k-1}} = (\phi_{T_k})^{1/n}$, and the characteristic function uniquely determines the distribution, this determines the distributions of $T_k - T_{k-1}$. Since $T_k - T_{k-1} \sim exponential(\lambda)$ will ensure $T_k \sim gamma(k, \lambda)$ and vice versa, $T_k - T_{k-1} \sim exponential(\lambda)$. ■

17.3.2 The Poisson Point Process

Definition 17.4 *A* Poisson Point Process *(P.P.P.) with intensity measure μ on $(S, \mathcal{S})$ is a collection of random variables $N(B, \omega)$, $B \in \mathcal{S}$, $\omega \in \Omega$ defined on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ such that:*

1. *$N(B) = N(B, \omega)$, $B \in \mathcal{S}$, $\omega \in \Omega$;*
2. *$N(\cdot, \omega)$ is a non-negative integer or ∞ -valued measure on $(S, \mathcal{S})$ for each $\omega \in \Omega$;*
3. *$N(B, \cdot)$ is a r.v. with Poisson($\mu(B)$) distribution: $\mathbb{P}(N(B) = k) = \frac{e^{-\mu(B)}(\mu(B))^k}{k!}$* for all $B \in \mathcal{S}$; *and*
4. *If B_1, B_2, ... are disjoint sets then $N(B_1, \cdot)$, $N(B_2, \cdot)$, ... are independent random variables.*

Example 17.5 *Let $S = \mathbb{R}_+$, $\mathcal{S} = Borel(\mathbb{R}_+)$, $\mu = \lambda \cdot Lebesgue$. Let $N(B, \omega)$ be the measure of B, whose cumulative distribution function is defined as the counting process $(N_t, t \geq 0)$ for a $PP(\lambda)$ as described before. That is,*

$$N([0,t]) := N_t = \sum_{k=1}^{\infty} 1\{T_k \leq t\}$$

So,

$$\begin{aligned} N(B) &= \sum_{k=1}^{\infty} \mathbf{1}_{(T_k \in B)} \\ &= \text{the number of } T_k \text{ which fall in } B \end{aligned}$$

Example 17.6 *$S = \mathbb{R}$, $\mathcal{S} = Borel(\mathbb{R})$, $\mu = \lambda \cdot Lebesgue$. Stick together independent $PP(\lambda)$ on $\mathbb{R}_+$ and $\mathbb{R}_-$.*

Theorem 17.7 *Such a P.P.P. exists for any σ-finite measure space.*

Proof: The only convincing argument is to give an explicit construction from sequences of independent random variables. We begin by considering the case $\mu(S) < \infty$.

1. Take X_1, X_2, ... to be i.i.d. random variables with form $\mu(\cdot|S)$ so that $P(X_i \in B) = \frac{\mu(B)}{\mu(S)}$.
2. Take $N(S)$ to be a Poisson random variable with mean $\mu(S)$, independent of the X_i's. Assume all random variables are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
3. Define $N(B) = \sum_{i=1}^{N(S)} \mathbf{1}_{(X_i \in B)}$, for all $B \in \mathcal{S}$.

■

Exercise 17.8 *Verify that this $N(B, \omega)$ is a P.P.P. with intensity μ. (Use thinning property of Poisson distributions.)*

Example 17.9 (Poissonization of multinomial) *If B_1, B_2 are disjoint events and $B_1 \cup B_2 = S$*

$$\begin{aligned} P(N(B_1) = n_1,\ N(B_2) = n_2) &= P(N = n_1 + n_2)P(N(B_1) = n_1,\ N(B_2) = n_2 | N = n_1 + n_2) \\ &= e^{-\lambda}\frac{\lambda^{n_1+n_2}}{(n_1+n_2)!}(\frac{n_1+n_2}{n_1})p^{n_1}q^{n_2}, \end{aligned}$$

where $N = N(S)$, $p = \frac{\mu(B_1)}{\mu(S)}$, and $q = \frac{\mu(B_2)}{\mu(S)}$.

17.4 General Theory Behind Poisson Processes

It is important to note that we have presented two different examples of *convolution semigroups* (convolution 1/2groups) of probability distributions on $\mathbb{R}$. In general, we have a family of probability distributions $(F_t, t \geq 0)$ such that:

$$F_s * F_t = F_{s+t}$$

where $*$ is convolution. Examples are:

1. $F_t = \mathcal{N}(0, \sigma^2 t)$
2. $F_t = \text{Poisson}(\lambda t)$
3. $F_t = \delta_{ct}$ for some real $c \in \mathbb{R}$

For any such family we can create a *stochastic process* $(X_t, t \geq 0)$ so that X has stationary independent increments, that is:

$$\begin{aligned} X_t &\sim F_t \\ X_t - X_s &\sim F_{t-s}, \text{ for } 0 < s < t \end{aligned}$$

and so on. Notice that we did this by explicit construction for $F_t = \text{Poisson}(\lambda t)$. For the case of $F_t = \mathcal{N}(0, \sigma^2 t)$ we need to be clever, and in general we will need to appeal to *Kolmogorov's Extension Theorem* (see [1] for details). In particular,

1. the Poisson case gives a Poisson Process; and
2. the Normal case gives a *Brownian motion* (also known as a *Wiener Process*).

In particular, Brownian motion has the special feature that it is possible to define the process to have *continuous paths*; that is, for almost every ω, $t \to X_t(\omega)$ is continuous.

Definition 17.10 *A probability distribution F on $\mathbb{R}$ is called* infinitely divisible *if for every n, there exist i.i.d. random variables $X_{n,1}, X_{n,2}, ..., X_{n,n}$ such that*

$$S_n = X_{n,1} + X_{n.2} + ... + X_{n,n} \sim F;$$

that is, there exists a distribution $F_{\frac{1}{n}}$ so that

$$F_{\frac{1}{n}} * F_{\frac{1}{n}} * ... * F_{\frac{1}{n}} = F,$$

so F has a convolution "nth root" for every n.

Definition 17.11 *Let $(X_t, t \geq 0)$ be a process with stationary independent increments and distribution function F_t at time t. F_t is said to be* weakly continuous in t *as $t \downarrow 0$ if*

$$\lim_{t \downarrow 0} \mathbb{P}\left[|X_t| > \varepsilon\right] = 0$$

for all $\varepsilon > 0$

Theorem 17.12 (Lévy-Khinchine) *There is a 1-1 correspondence between infinitely divisible distributions F and weakly continuous convolution semigroups $(F_t, t \geq 0)$ with $F_1 = F$.*

Corollary 17.13 *Every infinitely divisible law is associated with a process with stationary independent increments which is continuous in P as t varies.*

For a proof see [2] or [3]

Now we present another example. Consider accidents that occur at times of Poisson Process with rate λ. At the time of the kth accident let there be some variable X_k like the damage costs covered by insurance companies. We have

$$Y_t = \sum_{k=1}^{N_t} X_k$$

where Y_t is the total cost/magnitude up to time t. It is not difficult to show that $(Y_t, t \geq 0)$ has stationary independent increments.

We look at the distribution of Y_t. This is called a *compound Poisson process*. It is hard to describe explicitly, but we can compute its characteristic function.

Exercise 17.14 *Find a formula for the characteristic function of Y_t in terms of the rate λ of N_t, and the distribution F of X_t, that is,*

$$F(B) = \mathbb{P}[X_k \in B]\,.$$

Observe we have

$$\begin{aligned}
\mathbb{E}[\exp\{i\theta Y_t\}] &= \sum_{n=1}^{N_t} \mathbb{E}[\exp\{i\theta Y_t\} \cdot \mathbf{1}\{N_t = n\}] \\
&= \sum_{n=1}^{N_t} \mathbb{E}[\exp\{i\theta\,(X_1 + X_2 + \ldots + X_n)\} \cdot \mathbf{1}\{N_t = n\}] \\
&= \sum_{n=1}^{N_t} \mathbb{E}[\exp\{i\theta\,(X_1 + X_2 + \ldots + X_n)\}] \cdot \mathbb{E}[\mathbf{1}\{N_t = n\}] \\
&= \sum_{n=1}^{N_t} (\mathbb{E}[\exp\{i\theta X\}])^n \cdot \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= e^{-\lambda t} \exp\left\{\lambda t \mathbb{E}\left[e^{i\theta X}\right]\right\} \\
&= \exp\left\{\lambda t \int_{\mathbb{R}} \left(e^{i\theta x} - 1\right) \mathbb{P}[X \in dx]\right\} \\
&= \exp\left\{t \int_{\mathbb{R}} \left(e^{i\theta x} - 1\right) L(dx)\right\}
\end{aligned}$$

where we let $L(\cdot) = \lambda \mathbb{P}[X \in \cdot]$, which is called the *Lévy Measure* associated with the process.

It is important to note that this is an instance of the *Lévy-Khinchine* formula which gives the characteristic function of the most general ∞-divisible law.

In the next subsection we include more details on this.

17.4.1 Computation of an instance of LK Formula

Given a P.P.P. N, say $N(B) = \sum_i \mathbf{1}\{X_i \in B\}$. For some sequence of random variables X_i. Consider positive and measurable f:

$$\int f dN = \sum_i f(X_i) \tag{17.1}$$

has clear intuitive meanings and many applications.

Example 17.15 *X_i could be the arrival time, location, and magnitude of an earthquake:*

$$X_i = (T_i, M_i, Y_i).$$

$f(t, m, y)$ represents the cost to the insurance company incurred by an earthquake at time t with magnitude m in place y.

How do we describe its distribution? Consider the case where f is a simple function. Say $f = \sum_{i=1}^m x_i \mathbf{1}\{B_i\}$ where the B_i's are disjoint events and cover the space. Then

$$\int f dN = \sum_i x_i N(B_i) \tag{17.2}$$

where $N(B_i)$ are independent r.v.s with Poisson($\mu(B_i)$) distribution.
This is some new infinitely divisible distribution.

Now we need a transformation. Because $f \geq 0$, it is natural to look first at the Laplace transformation. Take $\theta > 0$:

$$\begin{aligned}
\mathbb{E}\left[e^{-\theta \int f dN}\right] &= \mathbb{E}\left[e^{-\theta \sum_i x_i N(B_i)}\right] && (17.3)\\
&= \prod_i E\left[e^{-\theta x_i N(B_i)}\right] && (17.4)\\
&= \prod_i \exp\left[-\mu(B_i)(1 - e^{-\theta x_i})\right] && (17.5)\\
&= \exp\left[-\sum_i \mu(B_i)(1 - e^{-\theta x_i})\right] && (17.6)\\
&= \exp\left[-\int (1 - e^{-\theta f(s)})\mu(ds)\right] && (17.7)
\end{aligned}$$

17.4$implies$17.5 because $N(B_i) \sim$ Poisson($\mu(B_i)$), and if $N \sim$ Poisson(λ), then

$$\begin{aligned}
\mathbb{E}(e^{-\theta N}) &= \sum_{n=0}^{\infty} (e^{-\theta})^n \frac{\lambda^n}{n!} e^{-\lambda}\\
&= e^{-\lambda} e^{(e^{-\theta}\lambda)}\\
&= \exp\left[-\lambda(1 - e^{-\theta})\right]
\end{aligned}$$

Theorem 17.16 *For every non-negative measurable function f, we have, writing $e^{-\infty} := 0$:*

$$\mathbb{E}\left[e^{-\theta \int f dN}\right] = \exp\left[\int (e^{-\theta f(s)} - 1)\mu(ds)\right] \tag{17.8}$$

This formula is an instance of the Lévy-Khinchine equation.

Proof: We have shown that the result holds when f is a simple function. In general, there exist sequence of simple functions f_n such that $f_n \uparrow f$. Then apply the Monotone Convergence Theorem and Dominated Convergence Theorem. ∎

References

[1] Richard Durrett. *Probability: theory and examples, 3rd edition.* Thomson Brooks/Cole, 2005.

[2] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[3] Olav Kallenberg. *Foundations of Modern Probability.* Second edition. Springer, Berlin, 2001.