

# Applied Probability

January 29, 2018

<i>CONTENTS</i>	2
-----------------	---

## Contents

<b>0</b> Miscellaneous	<b>3</b>
<b>1</b> Poisson process	<b>4</b>
<b>2</b> Continuous-time Markov chains	<b>8</b>

## 0 Miscellaneous

Some speech

Google lecture's name to find his homepage and example sheets or probably  
some notice of a change of room

## 1 Poisson process

Suppose we have a Geiger counter. We model the "click process" as a family  $\{N(t) : t \geq 0\}$ , where  $N(t)$  denotes the total number of ticks up to time  $t$ . Now note that  $N(t) \in \{0, 1, \dots\}$ ,  $N(s) \leq N(t)$  if  $s \leq t$ ,  $N$  increases by unit jumps, and  $N(0) = 0$ . We also assert that  $N$  is right-continuous, i.e.  $\lim_{x \rightarrow t^+} N(x) = N(t)$ .

**Definition.** (infinitesimal definition)

A *Poisson process* with intensity  $\lambda$  is a process  $N = (N(t) : t \geq 0)$  which takes values in  $S = \{0, 1, 2, \dots\}$ , s.t.:

- (a)  $N(0) = 0$ ,  $N(s) \leq N(t)$  if  $s \leq t$ ;
- (b)

$$\mathbb{P}(N(t+h) = n+m | N(t) = n) = \begin{cases} \lambda h + o(h) & m = 1 \\ o(h) & m > 1 \\ 1 - \lambda h & m = 0 \end{cases}$$

Recall that  $g(h) = o(h)$  means that  $\frac{g(h)}{h} \rightarrow 0$  as  $h \rightarrow 0$ ;

- (c) if  $s < t$ , then  $N(t) - N(s)$  is independent of all arrivals prior to  $s$ .

**Theorem.**  $N(t)$  has the Poisson distribution with parameter  $\lambda t$ .

*Proof.* Study  $N(t+h)$  given  $N(t)$ . We have

$$\begin{aligned} \mathbb{P}(N(t+h) = j) &= \sum_{i \leq j} \mathbb{P}(N(t+h) = j | N(t) = i) \mathbb{P}(N(t) = i) \\ &= (1 - \lambda h) \mathbb{P}(N(t) = j) + \lambda h \mathbb{P}(N(t) = j-1) + o(h) \end{aligned}$$

So

$$\frac{\mathbb{P}(N(t+h) = j) - \mathbb{P}(N(t) = j)}{h} = -\lambda \mathbb{P}(N(t) = j) + \lambda \mathbb{P}(N(t) = j-1) + \frac{o(h)}{h}$$

write  $p_n(t) = \mathbb{P}(N(t) = n)$ , then let  $h \rightarrow 0^+$  we get

$$\begin{aligned} p'_j(t) &= -\lambda p_j(t) + \lambda p_{j-1}(t) \quad j \geq 1 \\ p'_0(t) &= -\lambda p_0(t) \end{aligned}$$

with boundary condition  $p_0(0) = 1$ .

We solve  $p_0$  to get  $p_0(t) = e^{-\lambda(t)}$ . Then we can use this to inductively solve  $p_1, p_2, \dots$  to get the desired result.  $\square$

An alternative derivation from the differential equations:

Let  $G(s, t) = \sum_j s^j p_j(t)$ . Now we take the set of differential equation, multiplying each one by  $s^j$ , then we get

$$\frac{\partial G}{\partial t} = \lambda(s-1)G$$

Then we have

$$G(s, t) = A(s)e^{\lambda(s-1)t}$$

We also have  $G(s, 0) = 1$  so we should be able to plug in a suitable value of  $s$  to get the desired result (I probably missed that).

**Definition.** (Holding/interarrival times) In a poisson process (pp) with parameter  $\lambda$ , let  $N(t)$  denote the total number of "clicks". Define the arrival times  $T_0 = 0$ ,  $T_n = \inf\{t \geq 0 : N(t) = n\}$ , i.e. the first time  $t$  that  $N$  reaches  $n$  (note right continuity of  $N$ ). We also define the interarrival times  $X_n = T_n - T_{n-1}$ .



**Theorem.** Suppose  $X_1, X_2, \dots$  are known. Let  $T_n = \sum_{i=1}^n x_i$ , note  $N(t) = \max\{n : T_n \leq t\}$ . Then the random variables  $X_1, X_2, \dots$  are independent and they have the exponential distribution with parameter  $\lambda$  ( $Exp(\lambda)$ ).

*Proof.*

$$\mathbb{P}(X_1 > t) = \mathbb{P}(N(t) = 0) = e^{-\lambda t}$$

So  $X_1$  has  $Exp(\lambda)$  distribution. Now consider  $\mathbb{P}(X_2 > t | X_1 = t_1)$ . This doesn't look to make much sense as  $X_1$  has a continuous distribution so  $\mathbb{P}(X_1 = t_1) = 0$ ; however we could consider the conditional density as  $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ . Then  $\mathbb{P}(X_2 > t | X_1 = t_1) = \mathbb{P}(\text{no arrivals in } (t_1, t_1 + t) | X_1 = t_1) = \mathbb{P}(\text{no arrivals in } (t_1, t_1 + t))$  by independence. This is then equal to  $\mathbb{P}(\text{no arrivals in } (0, t)) = \mathbb{P}(N(t) = 0) = e^{-\lambda t}$ . Then continue by induction.  $\square$

**Proposition.** (properties of a poisson process  $N$ )

(a)  $N$  has stationary independent increments, i.e.:

(i) If  $0 < t_1 < \dots < t_n$ , then  $N(t_1), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$  are independent;

(ii)  $N(s+t) - N(s) \stackrel{d}{\rightarrow} N(t) - N(0)$ .

Amongst processes which are right continuous, non-decreasing, has only jump discontinuities of size 1, (i) and (ii) are characteristics of the Poisson process, meaning that Poisson process is the only process that has those two properties.

(b) Thinning:

Suppose insects arrive as a poisson process with parameter  $\lambda$ . Each insect is a mosquito with probability  $\alpha$ , or a skeet with probability  $1 - \alpha$ , and the occurrences of the two insects are independent. Then

(i) the mosquito-arrival process  $F$  is a  $PP(\alpha\lambda)$ , (ii) the skeet-arrival process is  $S$  a  $PP((1-\alpha)\lambda)$ , and (iii) these processes are independent.

*Proof.* (i) and (ii) are immediate by infinitesimal definition of a poisson process. For (iii), by independence we mean that  $\mathbb{P}(F(t_1) = f_1, S(t_1) = s_1, \dots, F(t_n) = f_n, S(t_n) = s_n) = \mathbb{P}(F(t_1) = f_1, \dots, F(t_n) = f_n) \mathbb{P}(S(t_1) = s_1, \dots, S(t_n) = s_n) \forall t_1, \dots, t_n, f_1, \dots, f_n, s_1, \dots, s_n$ .

The simple case is

$$\begin{aligned} \mathbb{P}(F(t) = f, S(t) = s) &= \frac{(\lambda t)^{f+s} e^{-\lambda t}}{(f+s)!} \binom{f+s}{f} \alpha^f (1-\alpha)^s \\ &= \frac{(\alpha \lambda t)^f}{f!} e^{-\alpha \lambda t} \frac{((1-\alpha)\lambda t)^s}{s!} e^{-(1-\alpha)\lambda t} \\ &= \mathbb{P}(F(t) = f) \mathbb{P}(S(t) = s) \end{aligned}$$

□

(c) Superposition:

$F$ : Flies arrive as  $PP(\lambda_1)$ ;

$S$ : Skeets arrive as  $PP(\lambda_2)$ , and these processes are independent. Then  $N = F + S$  is a  $PP(\lambda_1 + \lambda_2)$ . This follows by infinitesimal construction of  $PP$ .

(d) Given  $N(t) = n$ , write  $\mathbf{T} = (T_1, \dots, T_n)$ ,  $\mathbf{t} = (t_1, \dots, t_n)$ , we have  $f_{\mathbf{T}}(\mathbf{t}|N(t) = n) = \left(\frac{1}{t}\right)^n n! L(\mathbf{t})$ , where  $L(\mathbf{t}) = 1$  iff  $t_1 < t_2 < \dots < t_n$ .

*Proof.* Next time.

□

Let's complete the proof left last lecture.

**Theorem.** Conditional on  $\{N(t) = n\}$ , the times  $T_1, \dots, T_n$  have joint pdf

$$f_{\mathbf{T}|N(t)=n}(\mathbf{t}) = \frac{n!}{t^n} L(\mathbf{t}) 1_{\{t_n \leq t\}}$$

where  $L(\mathbf{t}) = 1_{\{t_1 \leq t_2 \leq \dots \leq t_n\}}$ .

*Proof.* The interarrival times  $X_1, X_2, \dots, X_n$  have joint pdf

$$f_{\mathbf{X}}(\mathbf{x}) = \lambda^n \exp(-\lambda \sum_i x_i)$$

by change of variables, we now have (noting  $T_i = X_1 + \dots + X_i$ )

$$f_{\mathbf{T}}(\mathbf{t}) = \lambda^n e^{-\lambda t_n} L(\mathbf{t})$$

Now for  $C \subseteq \mathbb{R}^n$ , we have

$$\begin{aligned}\mathbb{P}(T \in C | N(t) = n) &= \frac{\mathbb{P}(T \subseteq C, N(t) = n)}{\mathbb{P}(N(t) = n)} \\ &= \frac{1}{\mathbb{P}(N(t) = n)} \int_C \mathbb{P}(N(t)) \\ &= n! \int_C f_{\mathbf{T}}(\mathbf{t}) d\mathbf{t} \\ &= \frac{1}{\mathbb{P}(N(t) = n)} \int_{t_n \leq t} e^{-\lambda(t-t_n)} \lambda^n e^{-\lambda t_n} L(\mathbf{t}) d\mathbf{t}\end{aligned}$$

the last equation is because we need there to be no arrival between  $t$  and  $t_n$ . Now the conditional pdf of  $\mathbf{T}$  given  $N(t) = n$  is

$$\frac{1}{(\lambda t)^n e^{-\lambda t} / n!} e^{-\lambda(t-t_n)} \lambda^n e^{-\lambda t_n} L(\mathbf{t}) = \frac{n! L(\mathbf{t})}{t^n} 1_{\{t_n \leq t\}}$$

I think somewhere in this proof we used  $\mathbb{P}(X \in C) = \int_C g(u) du \iff f_X(u) = g(u)$ , otherwise the lecture wouldn't have written this down on a separate board.  $\square$

## 2 Continuous-time Markov chains

This is actually quite a complicated topic, so we are going to make a lot of assumptions to simplify it.

Assume state space  $S$  is countable, and we often take  $S \subseteq \mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$  (sometimes useful to assume  $|S| < \infty$ ).

**Definition.** A process  $X = \{X(t) : t \geq 0\}$  taking values in  $S$  satisfies the *Markov property* if:

$$\begin{aligned}\mathbb{P}(X(t_n) = j | X(t_1) = i_1, \dots, X(t_{n-1}) = i_{n-1}) \\ = \mathbb{P}(X(t_n) = j | X(t_{n-1}) = i_{n-1})\end{aligned}$$

for all  $i_1, i_2, \dots, i_{n-1}, j \in S, t_1 < t_2 < \dots < t_n$ .

We have the transition probabilities  $p_{i,j}(s, t) = \mathbb{P}(X(t) = j | X(s) = i)$ . We, however, assume the process is homogeneous, i.e.

$$p_{i,j}(s, t) = p_{i,j}(0, t - s) := p_{i,j}(t - s) \forall s, t, i, j$$

so the transition probabilities only depend on the duration of time passed instead of the absolute time. We can then write this as a transition matrix  $(p_{i,j}(t))_{i,j \in S} = P_t$ .

**Proposition.** The family  $\{P_t : t \geq 0\}$  satisfies

- (a)  $P_0 = I$ ;
- (b)  $P_t$  is a *stochastic* matrix, i.e. a non-negative matrix with row sum 1;
- (c)  $P_{s+t} = P_s P_t$  for  $s, t \geq 0$ .

*Proof.* (of (c))

$p_{i,j}(s+t) = \sum_{k \in S} p_{i,k}(s) p_{k,j}(t)$  by Markov Property which is just the component form of  $P_{s+t} = P_s P_t$ .

$P_{s+t} = P_s P_t$  is sometimes called the *semigroup property* ( $s, t \geq 0$ ).

$(P_t : t \geq 0)$  is called a *stochastic semigroup*. □

General theory involves conditions of regularity.

We assume  $X$  is a right-continuous jump process.

Holding times for general chains:

Assume  $X(t_0) = i$ .

Let  $H = \inf\{t > t_0 : X(t) \neq i\}$ . We have

$$\mathbb{P}(H > u + v | H > u) = \mathbb{P}(H > v) \quad (*)$$

by Markov Property ( $u, v \geq 0$ ).

Let  $G(u) = \mathbb{P}(H > u)$ . By (\*), we get  $\frac{G(u+v)}{G(u)} = G(v)$ , so  $G(u+v) = G(u)G(v)$ .

We know  $G(0) = 1$ , and  $G$  is non-increasing.

Solution:  $G(n) = G(1)G(n-1) = G(1)^n \quad \forall n \in \mathbb{N}$ . Also  $G(p/q) \dots G(p/q) =$



$G(p) = G(1)^p$  so  $G(p/q) = G(1)^{p/q}$ , hence  $G(u) = G(1)^u$  for  $u \geq 0$ . We deduce that  $G(u) = e^{-\alpha u}$  for some  $\alpha > 0$ .

**Lemma.** A random variable  $X > 0$  has an exponential distribution iff it has the *lack of memory property*:  $\mathbb{P}(X > u + v | X > u) = \mathbb{P}(X > v) \forall u, v > 0$ .

A MC is a combination of exponential-distribution holding times, and a transition matrix for the *jump chain*  $Y = (Y_n)$  given by  $Y_0 = X(0)$ ,  $Y_1 = X(T_1)$ , where  $T_1 = \inf\{t : X(t) \neq X(0)\}$ , and  $Y_n = X(T_n)$  where  $T_{n+1} = \inf\{t > T_n : X(t) \neq X(T_n)\}$ .  $Y$  is a *discrete-time Markov chain*.

If in state  $i$ , want  $H$ , we jump to state  $j \neq i$  with probability  $\frac{g_{ij}}{\alpha_i}$ . Intensity of a jump is  $\alpha_i$ , and intensity of a jump to state  $j$  is  $g_{ij}$ .

Note a transition from  $i$  to itself is not deemed to be a transition.

We have  $p_{ij}(h) = g_{ij}h + o(h)$  ( $j \neq i$ ),  $p_{ij}(h) = 1 - \sum_{j \neq i} p_{ij}(h) = 1 - h \sum_{j \neq i} g_{ij} + o(h) = 1 - \alpha_i h + o(h) = 1 + g_{ii}h + o(h)$ , where we let  $g_{ii} = -\alpha_i$ . Now we let  $G$  be the matrix  $(g_{ij})$ , with the off-diagonal terms the previous  $g_{ij}$ 's, but the diagonal terms  $g_{ii}$  as defined just now (so as to make row sums 0). Now the off-diagonal terms are non-negative, and diagonal terms are non-positive. We call  $G$  the *generator* of the chain (otherwise known as the  $Q$ -matrix).

Conclusion:  $\frac{P_t - I}{t} \xrightarrow{t \rightarrow 0^+} G$ .

Questions of regularity: OK if  $|S| < \infty$ .

(?)

$$\begin{aligned} p_{ij}(t+h) &= \sum_k p_{ik}(t) p_{kj}(h) \\ &= \sum_{k \neq j} p_{ik}(t) [g_{kj}h + o(h)] + p_{ij}(t)(1 + g_{jj}h + o(h)) \\ &= \sum_k p_{ik}(t) g_{kj} \\ &= P_t \cdot G \end{aligned}$$

this is the (Kolmogorov) Forward Equation.

$p_{ij}(t+h) = \sum_k p_{ik}(h) p_{kj}(t)$ , so  $P' t = G P_t$ , called the K-Backward equation.

Interchange of limits requires justification – it's OK if  $|S| < \infty$ .

Now  $P'_t = P_t G$ , we can rewrite this as  $f' = f g$ , which gives  $f(t) = A e^{gt}$ . So the solution should be  $P_t = P_0 (= I) e^{tG}$  (i.e.  $= \sum_{k=0}^{\infty} \frac{t^k}{k!} G^k$ ).

In many cases, the solution to the forward and/or backward equation is the function  $P = e^{tG}$ .

A mistake has been made! The definition for holding time is wrong. It should be  $H = \inf\{t - t_0 : X(t) \neq X(t_0), t > t_0\}$  (the length rather than the absolute time).

Let's look at an example now.

**Example.** Let  $S = \{1, 2\}$ ,  $G = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$ , where  $\alpha\beta > 0$ . What are the  $p_{ij}(t)$ ? There are only two states, so let's find

$$\begin{aligned} p'_{11}(t) &= -\alpha p_{11} + \beta p_{12}, \\ p'_{12}(t) &= \alpha p_{11}(t) - \beta p_{12}(t) \end{aligned}$$

...

So we get a bunch of differential equations, with boundary conditions  $p_{11}(0) = p_{22}(0) = 1$ . This has a unique solution (check). We get  $P_t = e^{tG} = \sum_n \frac{t^n}{n!} A \Lambda^n A^{-1}$  where we diagonalize  $G = A \Lambda A^{-1}$ . But  $e^{t\Lambda}$  is equal to  $\text{Diag}(e^{\lambda_1 t}, e^{\lambda_2 t} \dots)$ . Each  $p_{ij}(t)$  has the form  $\sum_k e^{\lambda_k t} c_k(i, j)$ , where we need to find the constants  $c_k(i, j)$ .

**Lemma.** Let  $i, s \in S$ . Then either  $p_{ij}(t) = 0 \forall t > 0$ , or  $p_{ij}(t) > 0 \forall t > 0$  (this is because our time here is continuous; we can fit in any number of jumps in any time length with positive probability).

*Proof.* Assume  $p_{ij}(T) > 0$  for some  $T > 0$ . Then for  $t > T$ ,  $p_{ij}(t) \geq p_{ij}(T) \mathbb{P}_j(X_j > t - T) > 0$  (start in  $j$  and holding time  $> t - T$ , which is positive).

For  $t < T$ : there exists finite sequence of jumps from  $i$  to  $j$  in time  $T$ , so there exists  $i_1, i_2, \dots, i_n \in S$  with  $g_{i, i_1} g_{i_1, i_2} \dots g_{i_n, j} > 0$ , where  $g_{i, j}$  is as defined previously. Then we can just divide the time  $(0, t)$  into  $n + 1$  intervals. Then  $p_{i, j}(t) \geq p_{i, i_1}(t/(n+1)) \dots p_{i_n, j}(t/(n+1)) > 0$  (this is an applied course, so we don't care that much).  $\square$

**Definition.** The chain  $X$  on state space is *irreducible* if  $\forall i, j \in S, \forall t > 0, p_{i, j}(t) > 0$  ( $\iff \exists t > 0, p_{i, j}(t) > 0$ ).

A distribution  $\pi$  on  $S$  is invariant (or stationary, or equilibrium distribution) if  $\pi = \pi P_t$  for all  $t \geq 0$ .

Note: if  $X(0)$  has distribution  $\mu_0$ , then  $X(t)$  has distribution  $\mu_t = \mu_0 P_t$  ( $\mu_t(j) = \sum_i \mu_0(i) p_{i, j}(t)$ ).

Note: (a) Differentiate  $\pi = \pi P_t$  to get  $0 = \pi G$ ;

(b) If  $P_t = e^{tG}$  then  $\pi G = 0$  iff  $\pi G^n = 0$  for  $n \geq 1$  iff  $\pi \sum_n \frac{t^n}{n!} G^n = \pi$  iff  $\pi P_t = \pi$ .

**Theorem.** Let  $X$  be irreducible. If there exist an invariant distribution  $\pi$ , then it is unique (what a surprise), and  $p_{ij}(t) \rightarrow \pi_j$  as  $t \rightarrow \infty$ . Can we prove this in the remaining 11 minutes? Let's try:

*Proof.* Let  $h > 0$ . Then  $Y_n = X(nh)$ . So  $Y$  is a 'skeleton' of  $X$ .  $Y$  is a markov chain. Since  $X$  is irreducible, so is  $Y$ .  $Y$  has invariant distribution  $\pi$ , hence  $\pi$  is unique. (?) Since  $X$  is irreducible,  $Y$  is aperiodic(?). Hence  $p_{ij}(nh) \rightarrow \pi_j$  as  $n \rightarrow \infty$ . Since this holds for all  $h \in \mathbb{Q}^+$ , we deduce that  $p_{ij}(t) \rightarrow \pi_j$  as  $t \rightarrow \infty$  through the rationals. The conclusion follows by continuity of  $p_{ij}(\cdot)$ .  $\square$

**Lemma.** The functions  $p_{ij}(\cdot)$  are continuous.