MS-C2111 Stochastic Processes

Lecture 7

Random point patterns and counting processes

Jukka Kohonen
Aalto University

Contents

Contents

Study object: Unpredictable events

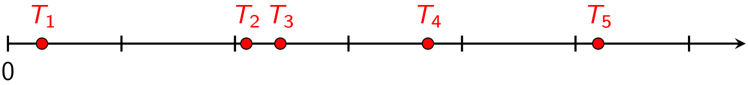
Time instants which cannot be accurately predicted:

- Occurrences of earthquakes
- Service requests in data centers
- Press releases of publicly traded companies

Study goals

- Learn to model unpredictable time instants as random point patterns
- Derive the distribution of point counts of an independently scattered pattern
- Learn to analyze independently scattered point patterns using a Poisson process

Random point patterns



Random point pattern
= Countable set of points $\{T_1, T_2, \dots\}$ on the time line $\mathbb{R}_+$ such that the location of each point is a random variable

- Countable $\implies$ The points can be listed using 1,2,...

Distribution of a random point pattern
= The joint probability distribution of *all* random variables $T_1, T_2, \dots$

- Usually implicitly defined via a generating mechanism
- What is a natural mechanism?

Independent scattering

Maximally random time instants: Information about $[0, t]$ is irrelevant for predicting what happens in (t, ∞)

A random point pattern X with counting measure $N(A) = |X \cap A|$ is

- Independently scattered if

$$A_1, \dots, A_n \text{ disjoint} \quad \Longrightarrow \quad N(A_1), \dots, N(A_n) \text{ independent}$$

- Homogeneous if

$$N(A + t) =_{\mathrm{st}} N(A) \quad \text{for all } t \geq 0,$$

where $A + t = \{a + t : a \in A\}$ is the t-shifted version of A

Thought experiment

How to construct an independently scattered point pattern using $U_1, U_2, \ldots =_{\mathrm{st}} \mathsf{Unif}(0,1)$ as building blocks?

Observations

In an independently scattered point pattern:

- The number of points on every nonempty interval is unlimited
- The distance between two points can be arbitrarily small and arbitrarily large

Distribution of point counts

Theorem

Let X be a homogeneous and independently scattered random point pattern on $(0, \infty)$ with intensity

$$\lambda = \mathbb{E}N((0,1]).$$

Then the number of points of X in the interval $(0, t]$ follows a Poisson distribution with parameter λt:

$$\mathbb{P}\Big(N\big((0,t]\big) = k\Big) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}, \qquad k = 0, 1, \ldots$$

Distribution of point counts: Proof (1/2)

Let $v(t) = \mathbb{P}(N(0,t] = 0)$ be the probability that $(0,t]$ contains no points of X.

$$
\begin{aligned}
v(s+t) &= \mathbb{P}(N(0,s+t] = 0) \\
&= \mathbb{P}(N(0,s] = 0,\ N(s,s+t] = 0) \\
&= \mathbb{P}(N(0,s] = 0)\,\mathbb{P}(N(s,s+t] = 0) \\
&= \mathbb{P}(N(0,s] = 0)\,\mathbb{P}(N(0,t] = 0) \\
&= v(s)v(t)
\end{aligned}
$$

$t \mapsto v(t)$ is decreasing $\implies$

$$v(t) = e^{-\alpha t} \quad \text{for some } \alpha > 0.$$

Distribution of point counts: Proof (2/2)

Divide $(0, t]$ into small subintervals $I_{n,j} = (\frac{j-1}{n}t, \frac{j}{n}t]$ and denote

$$Z_n \;=\; \sum_{j=1}^{n} \theta_j, \quad \theta_j \;=\; 1(N(I_{n,j}) \geq 1).$$

Independent scattering & homogeneity $\implies$ $Z_n =_{\mathrm{st}} \mathrm{Bin}(n, q_n)$ where

$$q_n \;=\; 1 - v(t/n) \;=\; 1 - e^{-\alpha t/n}.$$

$$\text{l'Hôpital} \implies \qquad nq_n = \frac{1 - e^{-\alpha t/n}}{1/n} \to \alpha t, \quad n \to \infty.$$

Law of small numbers (+ a density argument):
$\mathbb{P}(N(0, t] = k) \approx \mathbb{P}(Z_n = k) \to e^{-\alpha t} \frac{(\alpha t)^k}{k!}$.
Moreover, $\alpha = \mathbb{E}N(0, 1] = \lambda$.

Law of small numbers

Theorem

If $Z_n =_{\mathrm{st}} \mathrm{Bin}(n, q_n)$ *and* $nq_n \to \alpha \in (0, \infty)$ *as* $n \to \infty$, *then*

$$\mathbb{P}(Z_n = k) \;\to\; e^{-\alpha} \frac{\alpha^k}{k!}, \quad n \to \infty.$$

In other words: $\mathrm{Bin}(n, q_n) \to \mathrm{Poi}(\alpha)$ as $n \to \infty$.

Proof.

The probability that $Z_n = k$ can be written as

$$\binom{n}{k} q_n^k (1 - q_n)^{n-k} \;=\; \underbrace{\frac{n!}{n^k (n-k)!}}_{\to 1} \underbrace{\frac{1}{(1-q_n)^k}}_{\to 1} \underbrace{\frac{(nq_n)^k}{k!}}_{\to \frac{\alpha^k}{k!}} \underbrace{\left(1 - \frac{nq_n}{n}\right)^n}_{\to e^{-\alpha}}.$$

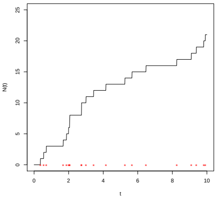
$\square$

Example: Teemu Selänne

- Scored 1430 points in 1387 games during NHL regular seasons 1993–2013
- $\lambda = 1.03$ points per game on average
- #Points/game $=_{\mathrm{st}}$ Poi(λ) if goal occurrence instants independently scattered

Points	Pr (predicted)	#Games (predicted)	#Games (realized)
0	0.35665	495	505
1	0.36771	510	506
2	0.18955	263	251
3	0.06514	90	87
4	0.01679	23	29
5	0.00346	5	9
6	0.00059	1	0
7	0.00009	0	0
> 7	0.00001	0	0

Poisson process



$N : \mathbb{R}_+ \to \mathbb{Z}_+$ is a Poisson process with intensity λ if

- $N(t) - N(s) =_{\mathrm{st}} \mathrm{Poi}(\lambda(t-s))$ for all $(s,t] \subset \mathbb{R}_+$
- N has independent increments:

$$N(t_1) - N(s_1), \dots, N(t_n) - N(s_n)$$

  are independent for disjoint $(s_1, t_1], \dots, (s_n, t_n] \subset \mathbb{R}_+$

The previous theorem rephrased:

Theorem

The counting process $N(t) = N(0,t]$ of a homogeneous independently scattered point pattern is a Poisson process with intensity $\lambda = \mathbb{E}N(0,1]$.

Contents

Exponential distribution

Random number $X \geq 0$ is $\text{Exp}(\lambda)$-distributed with rate parameter $\lambda > 0$ if

$$\mathbb{P}(X \leq t) \;=\; 1 - e^{-\lambda t}, \quad t \geq 0.$$

- X has a density $f(t) = \lambda e^{-\lambda t}$, $t \geq 0$
- $\mathbb{E}X = \frac{1}{\lambda}$

Memoryless property

Theorem

$X =_{\mathrm{st}} \mathrm{Exp}(\lambda)$ *satisfies*

$$\mathbb{P}(X > s + t \,|\, X > s) \;=\; \mathbb{P}(X > t) \quad \textit{for all } s, t \geq 0.$$

Proof:

$$\begin{aligned}
\mathbb{P}(X > s + t \,|\, X > s) \;&=\; \frac{\mathbb{P}(X > s + t,\, X > s)}{\mathbb{P}(X > s)} \\
&=\; \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} \\
&=\; \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \;=\; e^{-\lambda t} \;=\; \mathbb{P}(X > t).
\end{aligned}$$

Stochastic public traffic: Having waited for a bus for s minutes, the remaining waiting time is $\mathrm{Exp}(\lambda)$-distributed.

First jump instant of a Poisson process

Let $T_1 = \min\{t \geq 0 : N(t) = 1\}$ be the first jump instant of a Poisson process N with intensity λ. What is the distribution of T_1?

$$\mathbb{P}(T_1 > t) \;=\; \mathbb{P}(N(t) = 0) \;=\; e^{-\lambda t}\frac{\lambda^0}{0!} \;=\; e^{-\lambda t}$$

Because the cumulative distribution function $\mathbb{P}(T_1 \leq t) = 1 - e^{-\lambda t}, t \geq 0$ determines the distribution, we conclude that T_1 is $\mathrm{Exp}(\lambda)$-distributed.

More generally, the distance

$$\Delta \;=\; \min\{t \geq s : N(t) = N(s) + 1\}$$

from any point s to the next jump instant of the Poisson process is $\mathrm{Exp}(\lambda)$-distributed.

Winning time in an exponential race

Winning time of a race is $U = \min\{X_1, ..., X_n\}$, where the competitors' times $X_i =_{\mathrm{st}} \mathrm{Exp}(\lambda_i)$ are independent.

What is the distribution of U?

$$
\begin{aligned}
\mathbb{P}(U > t) &= \mathbb{P}(X_1 > t, ..., X_n > t) \\
&= \mathbb{P}(X_1 > t) \cdots \mathbb{P}(X_n > t) \\
&= e^{-\lambda_1 t} \cdots e^{-\lambda_n t} \\
&= e^{-(\sum_{i=1}^n \lambda_i) t}
\end{aligned}
$$

$$\Longrightarrow \quad U =_{\mathrm{st}} \mathrm{Exp}\left(\textstyle\sum_{i=1}^n \lambda_i\right)$$

The minimum of independent exponentially distributed random numbers is exponentially distributed.

Winning probability in an exponential race

What is the probability that 1 wins?

Competitor 1 wins if $X_1 < \min\{X_2, \ldots, X_n\} =: \tilde{U}$.

Because X_1 and $\tilde{U}$ are independent,

$$\mathbb{P}(X_1 < \tilde{U}) \;=\; \int_0^\infty \mathbb{P}(t < \tilde{U})\, \lambda_1 e^{-\lambda_1 t} dt.$$

Because $\tilde{U} =_{\mathrm{st}} \mathrm{Exp}(\sum_{i=2}^n \lambda_i)$, competitor 1 wins with probability

$$\begin{aligned}\mathbb{P}(X_1 < \tilde{U}) &= \int_0^\infty e^{-(\sum_{i=2}^n \lambda_i)t}\, \lambda_1 e^{-\lambda_1 t} dt \\ &= \lambda_1 \int_0^\infty e^{-(\sum_{i=1}^n \lambda_i)t} dt \;=\; \frac{\lambda_1}{\sum_{i=1}^n \lambda_i}\end{aligned}$$

Exponential race — Summary

Theorem

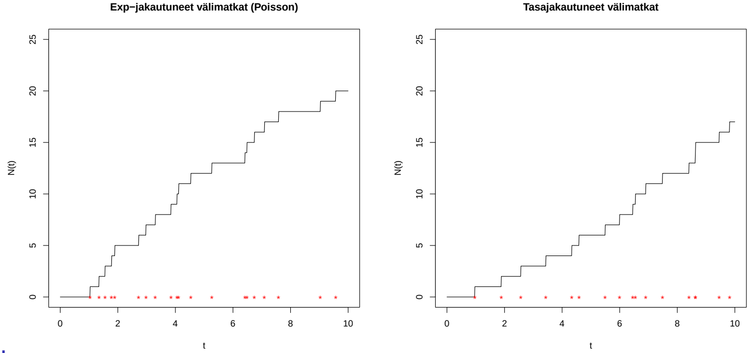
Let $U = \min\{X_1, \ldots, X_n\}$ *where* $X_i =_{\mathrm{st}} \mathsf{Exp}(\lambda_i)$ *are independent. Then:*

- $U =_{\mathrm{st}} \mathsf{Exp}(\sum_{i=1}^{n} \lambda_i)$
- $\mathbb{P}(X_i = U) = \frac{\lambda_i}{\lambda_1 + \cdots \lambda_n}$
- U *and* $I = \arg\min_i \{X_1, \ldots, X_n\}$ *are independent.*

Contents

Simulating independently scattered point patterns

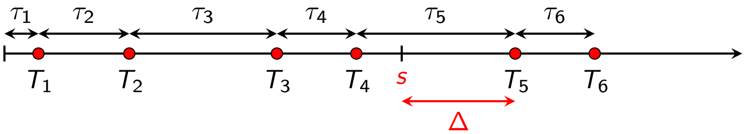
$$T_n = \tau_1 + \cdots \tau_n$$



Theorem

If the interpoint distances $\tau_1, \tau_2, \ldots$ are independent and Exp(λ)*-distributed, then the point pattern $X = \{T_1, T_2, \ldots\}$ is homogeneous and independently scattered, and the corresponding counting process $N(t)$ is a Poisson process with intensity λ.*

Sketch of proof



$\Delta = T_{N(s)+1} - s$

Condition on $A_s = \{N((0,s]) = 4, T_1 = t_1, \ldots, T_4 = t_4\}$

$$
\begin{aligned}
\mathbb{P}(\Delta > t \,|\, A_s) &= \mathbb{P}(T_5 - s > t \,|\, A_s) \\
&= \mathbb{P}(t_4 + \tau_5 - s > t \,|\, A_s) \\
&= \mathbb{P}(\tau_5 > s - t_4 + t \,|\, \tau_5 > s - t_4, A_s) \\
&= \mathbb{P}(\tau_5 > s - t_4 + t \,|\, \tau_5 > s - t_4) \\
&= \mathbb{P}(\tau_5 > t) \\
&= e^{-\lambda t}
\end{aligned}
$$

$\Delta =_{\mathrm{st}} \mathrm{Exp}(\lambda)$ and independent of $N(s)$

$\implies$ The points in $[s, \infty)$ are independent of the points in $[0, s]$

(We conditioned on a zero-probability event.)

What is the distribution of the n-th time instant T_n?

Gamma distribution

A random number $X \geq 0$ is $\mathrm{Gam}(n, \lambda)$-distributed with shape parameter n and rate parameter λ if it has a density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Lemma

If $\tau_1, \ldots, \tau_n$ are independent and $\mathrm{Exp}(\lambda)$*-distributed, then* $T_n = \sum_{i=1}^{n} \tau_i$ *is* $\mathrm{Gam}(n, \lambda)$*-distributed.*

Proof.

Classroom problem 8.1. $\square$

Distribution of $N(t)$

Let $T_n = \tau_1 + \cdots + \tau_n$ and let $N(t)$ be the counting process of $\{T_1, T_2, \ldots\}$.

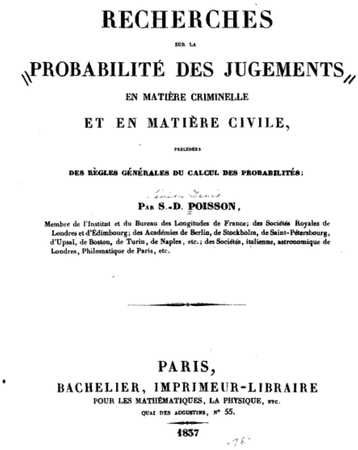
$$
\begin{aligned}
\mathbb{P}(N(t) = n) &= \mathbb{P}(N(t) \geq n) - \mathbb{P}(N(t) \geq n+1) \\
&= \mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t) \\
&= \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} ds - \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^n}{(n)!} ds \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^n}{(n)!} ds - \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^n}{(n)!} ds \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!}.
\end{aligned}
$$

Hence $N(t) =_{\mathrm{st}} \text{Poisson}(\lambda t)$.

Summary

- Random point pattern on $\mathbb{R}$ = model for unpredictable time instants
- Independently scattered point patterns have Poisson distributed point counts and exponentially distributed interpoint distances.
- Independently scattered point patterns can be constructed using Exp-distributed interpoint distances

Siméon Denis Poisson (1781–1840)



RECHERCHES

SUR LA

PROBABILITÉ DES JUGEMENTS

EN MATIÈRE CRIMINELLE

ET EN MATIÈRE CIVILE,

PRÉCÉDÉES

DES RÈGLES GÉNÉRALES DU CALCUL DES PROBABILITÉS;

Siméon Denis

PAR S.-D. POISSON,

Membre de l'Institut et du Bureau des Longitudes de France; des Sociétés Royales de Londres et d'Édimbourg; des Académies de Berlin, de Stockholm, de Saint-Pétersbourg, d'Upsal, de Boston, de Turin, de Naples, etc.; des Sociétés, italienne, astronomique de Londres, Philomatique de Paris, etc.

PARIS,

BACHELIER, IMPRIMEUR-LIBRAIRE

POUR LES MATHÉMATIQUES, LA PHYSIQUE, ETC.

QUAI DES AUGUSTINS, N° 55.

1837

References

Photos

1. Image courtesy of think4photop at FreeDigitalPhotos.net
2. Image courtesy of Hockeybroad/Cheryl Adams at Wikipedia.