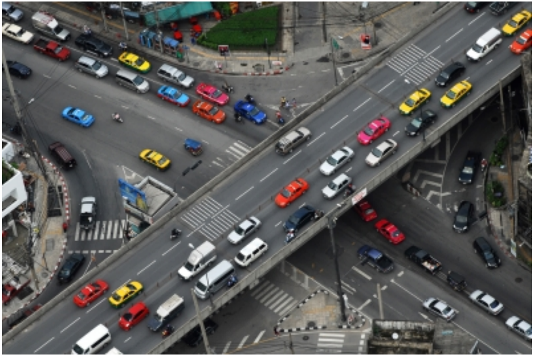
MS-C2111 Stochastic Processes



Lecture 8

Poisson processes

Jukka Kohonen
Aalto University

Contents

Contents

Poisson process

$N : \mathbb{R}_+ \to \mathbb{Z}_+$ is a Poisson process with intensity $\lambda > 0$ if

(i) $N(0) = 0$

(ii) $N(t) - N(s) =_{\mathrm{st}} \mathrm{Poi}(\lambda(t-s))$ for all $s < t$

(iii) N has independent increments:

$$(s_1, t_1], \dots, (s_k, t_k] \text{ disjoint}$$
$$\Longrightarrow$$
$$N(t_1) - N(s_1), \dots, N(t_k) - N(s_k) \text{ independent}$$

$N(s,t] = N(t) - N(s)$ equals the count of independently scattered time instants in $(s,t]$

$N(t+h) - N(t) =_{\mathrm{st}} \mathrm{Poi}(\lambda h) =_{\mathrm{st}} N(h) - N(0) =_{\mathrm{st}} N(h)$

The expected number of time instants on the unit interval $(t, t+1]$ is $\mathbb{E}(N(t,t+1]) = \mathbb{E}(N(1)) = \lambda$

Superposed Poisson processes

Theorem

If $N_1, N_2, \ldots$ are independent Poisson processes with intensities λ_j, then $N(t) = \sum_j N_j(t)$ is a Poisson process with intensity $\lambda = \sum_j \lambda_j$.

Proof.

(i) $N(0) = \sum_j N_j(0) = 0$. OK

(ii) $N(t) - N(s) = \sum_j (N_j(t) - N_j(s)) =_{\mathrm{st}} ?$ $\square$

Lemma

If $N_j =_{\mathrm{st}} \mathrm{Poi}(\lambda_j)$ are independent, then $\sum_j N_j =_{\mathrm{st}} \mathrm{Poi}(\sum_j \lambda_j)$.

Proof.

$$G_{N_j}(z) \;=\; \mathbb{E}(z^{N_j}) \;=\; \sum_{n=0}^{\infty} z^n \left(e^{-\lambda_j} \; \frac{\lambda_j^n}{n!} \right) \;=\; e^{-\lambda_j}\, e^{\lambda_j z} \;=\; e^{\lambda_j (z-1)}$$

$$G_{\sum_j N_j}(z) \;=\; \mathbb{E}(z^{\sum_j N_j}) \;=\; \prod_j \mathbb{E}(z^{N_j}) \;=\; \prod_j e^{\lambda_j (z-1)} \;=\; e^{\sum_j \lambda_j (z-1)}$$

Because pgf determines the distribution,
$\sum_j N_j =_{\mathrm{st}} \mathrm{Poi}(\sum_j \lambda_j)$. □

Let's continue the proof.

Superposed Poisson processes

Theorem
If $N_1, N_2, \dots$ are independent Poisson processes with intensities λ_j, then $N(t) = \sum_j N_j(t)$ is a Poisson process with intensity $\lambda = \sum_j \lambda_j$.

Proof.
(i) $N(0) = \sum_j N_j(0) = 0$. OK
(ii) $N(t) - N(s) = \sum_j (N_j(t) - N_j(s)) =_{\mathrm{st}} \mathrm{Poi}(\lambda(t-s))$. OK
(iii) Independent increments? If $(s_1, t_1]$ ja $(s_2, t_2]$ disjoint,

$$N_j(s_1, t_1] \perp\!\!\!\perp N_j(s_2, t_2] \quad \text{for all } j$$

$$\implies \sum_j N_j(s_1, t_1] \perp\!\!\!\perp \sum_j N_j(s_2, t_2] \quad \implies N(s_1, t_1] \perp\!\!\!\perp N(s_2, t_2]$$

Analogously, when $(s_1, t_1], \dots, (s_k, t_k]$ disjoint. OK $\square$

Contents

Example: Länsiväylä

The average flow of cars crossing the Helsinki–Espoo border on weekdays equals $\lambda = 40$ cars/min

- The average number of people per car is $m = 1.9$ with standard deviation $\sigma = 1.2$

What is the expectation and standard deviation of the number of people crossing the border per hour?

The number of cars crossing the border during $[0, t]$ is naturally modeled using a Poisson process $N(t)$ with intensity λ.

How to model the number of people crossing the border during $[0, t]$?

Compound Poisson process

We can add randomness to point pattern $X = \{ T_1, T_2, \dots \}$ by defining

$$\tilde{X} \ = \ \{(T_1, Z_1), (T_2, Z_2), \dots\},$$

where $Z_1, Z_2, \dots$ are independent of X and of each other.

We may interpret Z_i as the reward at time instant T_i
$\implies$ The cumulative reward from time interval $[0, t]$ is

$$S(t) \ = \ \sum_{i=1}^{\infty} Z_i 1(T_i \le t) \ = \ \sum_{i=1}^{N(t)} Z_i, \quad N(t) = \sum_{i=1}^{\infty} 1(T_i \le t)$$

$S(t)$ is a compound Poisson process when $N(t)$ is a Poisson process and $Z_1, Z_2, \dots$ are IID

Compound Poisson process

Theorem

The mean and variance of a compound Poisson process $S(t) = \sum_{i=1}^{N(t)} Z_i$ *at time instant* t *are given by*

$$\begin{aligned} \mathbb{E}(S(t)) &= \lambda m t, \\ \mathrm{Var}(S(t)) &= \lambda(m^2 + \sigma^2)t, \end{aligned}$$

where $\lambda = \mathbb{E}(N(1))$, $m = \mathbb{E}(Z_i)$ *and* $\sigma^2 = \mathrm{Var}(Z_i)$.

Proof.

By conditioning on the event $\{N(t) = n\}$ one can verify that

$$\begin{aligned} \mathbb{E}(S(t)) &= \mathbb{E}(N(t))\mathbb{E}(Z_i), \\ \mathrm{Var}(S(t)) &= \mathbb{E}(N(t))\,\mathrm{Var}(Z_i) + \mathrm{Var}(N(t))(\mathbb{E}(Z_i))^2. \end{aligned}$$

$\mathbb{E}(N(t)) = \lambda t$, $\mathrm{Var}(N(t)) = \lambda t$. $\square$

Compound Poisson process

Theorem

A compound Poisson process $S(t) = \sum_{i=1}^{N(t)} Z_i$ has independent increments.

Proof.

Choose disjoint $I_k = (s_k, t_k]$, $k = 1, 2$.
Under the occurrence of events
$A_k = \{N(s_k) = m_k, N(t_k) = m_k + r_k\}$ the random numbers

$$D_k \; = \; S(t_k) - S(s_k) \; = \sum_{i=m_k+1}^{m_k+r_k} Z_i \quad \text{are independent}$$

$$\mathbb{P}(D_1 \in B_1, D_2 \in B_2, A_1, A_2) = \cdots$$

The claim follows by summing over possible m_1, m_2, r_1, r_2. $\square$

Example: Länsiväylä

On average $\lambda = 40$ cars/min cross the Helsinki–Espoo border with $m = 1.9$ people per car on average (standard deviation $\sigma = 1.2$).

During $[0, t]$ the border is crossed by

- $N(t) \;=\; \sum_{i=1}^{\infty} 1(T_i \leq t)$ cars
- $S(t) \;=\; \sum_{i=1}^{\infty} Z_i \, 1(T_i \leq t)$ people

where

- $T_i =$ border crossing time of the i-th car
- $Z_i =$ number of people in the i-th car

Natural assumptions $\implies$ $S(t)$ is a compound Poisson process.

The number of people during one hour $(t = 60)$ satisfies

$$\mathbb{E}(S(60)) \;=\; \lambda m t \;=\; 40 \times 1.9 \times 60 \;=\; 4560$$

$$(\mathrm{Var}(S(60)))^{1/2} \;=\; (\lambda(m^2+\sigma^2)t)^{1/2} \;=\; (40\times(1.9^2+1.2^2)\times 60)^{1/2} \;=\; 110.09$$

Contents

Example: Länsiväylä

On average $\lambda = 40$ cars/min cross the Helsinki–Espoo border with $p_1 = 30\%$ of the cars only carrying the driver.

During $[0, t]$ the border is crossed by

- $N(t) \ = \ \sum_{i=1}^{\infty} 1(T_i \le t)$ cars
- $N_1(t) = \ \sum_{i=1}^{\infty} \theta_i \, 1(T_i \le t)$ solo drivers
- $N_2(t) = \ \sum_{i=1}^{\infty} (1 - \theta_i) \, 1(T_i \le t)$ other drivers

where $\theta_i \ = \ 1(Z_i = 1) =_{\mathrm{st}} \mathrm{Ber}(p_1)$

What is the probability of $\{N_2(1) \le 20\}$ given $\{N_1(1) \ge 30\}$?

$N_1(t)$ is a thinned (harvennettu) Poisson process which is obtained by removing 70% of the events of $N(t)$.

Thinned Poisson process

Theorem

If $\theta_1, \theta_2, \ldots$ are IID and independent of a Poisson process $N(t)$, then $N_1(t) = \sum_{i=1}^{\infty} \theta_i 1(T_i \le t)$ and $N_2(t) = \sum_{i=1}^{\infty} (1-\theta_i) 1(T_i \le t)$ are mutually independent Poisson processes.

Proof.

$N_1(t)$ is a compound Poisson process $\implies \perp\!\!\!\perp$ increments

$$G_{\theta_i}(z) \;=\; \mathbb{E}(z^{\theta_i}) \;=\; (1-p_1)z^0 + p_1 z^1$$

$$G_{N_1(t)}(z) \;=\; G_{N(t)}(G_{\theta_i}(z)) \;=\; e^{\lambda t (G_{\theta_i}(z)-1)} \;=\; e^{\lambda t p_1 (z-1)}$$

$N_1(t) =_{\mathrm{st}} \mathrm{Poi}(\lambda p_1 t)$.
N_1 is a Poisson process with intensity λp_1.
N_2 is a Poisson process with intensity $\lambda(1-p_1)$.
Are N_1 and N_2 independent? (They appear not.) $\square$

Thinned Poisson process

Proof.

Are N_1 and N_2 independent? The event $N_1(s,t] = j$ ja $N_2(s,t] = k$ occurs precisely when the interval $(s,t]$ contains $N(s,t] = j+k$ events, out which j are selected into N_1.

$$\begin{aligned}
\mathbb{P}(N_1(t) = j, N_2(t) = k) &= \mathbb{P}\Big(N(t) = j+k\Big) \cdot \binom{j+k}{j} p_1^j (1-p_1)^k \\
&= e^{-\lambda t} \frac{(\lambda t)^{j+k}}{(j+k)!} \binom{j+k}{j} p_1^j (1-p_1)^k \\
&= \cdots \\
&= \mathbb{P}(N_1(t) = j)\, \mathbb{P}(N_2(t) = k)
\end{aligned}$$

□

Example: Länsiväylä

On average $\lambda = 40$ cars/min cross the Helsinki–Espoo border with $p_1 = 30\%$ of the cars only carrying the driver.

During $[0, t]$ the border is crossed by

- $N(t) \ = \ \sum_{i=1}^{\infty} 1(T_i \le t)$ cars
- $N_1(t) = \ \sum_{i=1}^{\infty} \theta_i \, 1(T_i \le t)$ solo drivers
- $N_2(t) = \ \sum_{i=1}^{\infty} (1-\theta_i) \, 1(T_i \le t)$ other drivers

where $\theta_i \ = \ 1(Z_i = 1) =_{\mathrm{st}} \mathrm{Ber}(p_1)$

What is the probability of $\{N_2(1) \le 20\}$ given $\{N_1(1) \ge 30\}$?

N_1 and N_2 are independent Poisson processes, so that

$$\mathbb{P}(N_2(1) \le 20 \,|\, N_1(1) \ge 30) \ = \ \mathbb{P}(N_2(1) \le 20)$$

Information about other types of cars does not help in predicting type-2 cars.

General thinning

Theorem

If N is a Poisson process with intensity λ, and $Z_1, Z_2, \ldots$ are IID random variables, independent of N, then the thinned processes

$$N_x(t) \; = \; \sum_{i=1}^{\infty} 1(Z_i = x) 1(T_i \le t)$$

are mutually independent Poisson processes with intensities $\lambda_x = \lambda \mathbb{P}(Z_i = x)$.

Contents

Example: Bus stop

The interarrival times of buses $\tau_1, \tau_2, \ldots$ are assumed independent and distributed according to a probability density f. What is the expected waiting time for a passenger who arrives to the bus stop independently and uniformly at random?

Renewal process

A renewal process is the counting process of a random point pattern $\{T_1, T_2, \dots\}$ defined by $T_n = \sum_{k=1}^{n} \tau_k$ where the interpoint distances $\tau_1, \tau_2, \dots \geq 0$ are IID.

Example: $\tau_k =_{\mathrm{st}} \mathrm{Exp}(\lambda) \implies$ Poisson process

Forward recurrence time

Let $B_t = T_{N(t)+1} - t$ be distance from t to the next point of random point pattern.

Under sufficient regularity: $t \mapsto B_t$ is a continuous-time Markov process on state space $\mathbb{R}_+$

What is the invariant distribution?

Forward recurrence time process

Assume that (B_t) has an invariant distribution on $\mathbb{R}_+$ with probability density $f_+(x)$. (Draw a picture.)
In statistical equilibrium (assuming such exists for large t):

$$\mathbb{E}(\#\text{UCR of } x \text{ during } (t, t+h)) \approx \mathbb{P}(B_t \in (0,h))\,\mathbb{P}(\tau > x) \approx f_+(0)h\,\mathbb{P}(\tau > x)$$

$$\mathbb{E}(\#\text{DCR of } x \text{ during } (t, t+h)) \approx \mathbb{P}(B_t \in (x, x+h)) \approx f_+(x)h$$

$$\implies f_+(0)\,\mathbb{P}(\tau > x) = f_+(x)$$
$$\implies \dots \implies f_+(x) = \frac{\mathbb{P}(\tau > x)}{\mathbb{E}(\tau)}.$$

Rigorous analysis can be done by applying Lotka–Volterra type differential equations, see [Asm03].

Theorem

In a statistical equilibrium, the remaining waiting for the next time instant is a random variable τ_+ which has a distribution characterized by the density function

$$f_+(t) = \frac{\mathbb{P}(\tau_k > t)}{\mathbb{E}(\tau_k)}.$$

Example

If $\tau_k = c$ is nonrandom, then $f_+(t) = \frac{1}{c} 1(0 < t \le c)$ is the uniform distribution on $[0, c]$.

Example

If $\tau_k =_{\mathrm{st}} \mathrm{Exp}(\lambda)$, then $f_+(t) = \frac{\mathbb{P}(\tau_k > t)}{\mathbb{E}(\tau_k)} = \frac{e^{-\lambda t}}{1/\lambda} = \lambda e^{-\lambda t}$.
Remaining waiting time $\tau_+ =_{\mathrm{st}} \mathrm{Exp}(\lambda)$, mean λ^{-1} time units.
(Waiting time paradox!)

References

Søren Asmussen.
Applied Probability and Queues.
Springer, second edition, 2003.

Sources

Photos

1. Image courtesy of think4photop at FreeDigitalPhotos.net