Poisson Processes Reading Group Notes

Gergely Flamich

November 15, 2023

1 Outline

- what defines a PP? independence, and mean measure
- simulation:
  1. count first, points second
  2. in time order
- thinning, mapping and restriction: maybe draw a triangle?
- superposition theorem
- equivalence with Gumbel processes in log-space
- emphasize dual view: all the points are there already, vs computational simulation
- the intensity transform for Poisson processes: processes of Poisson type
- Indexing processes
- I won't be dealing with:
  1. Markov Chains
  2. fitting the mean measure of Poisson processes

2 General Poisson processes

Let Π denote the set of points in space.
Define:

- $N(A) = \#(\Pi \cap A)$
- $\mu(A) = \mathbb{E}[N(A)]$

Two conditions:

1. if $A \cap B = \emptyset$, then $N(A) \perp N(B)$
2. $N(A)$ is Poisson distributed with mean $\mu(A)$.

Assume that the base measure is normalized, things also work when stuff is not normalized
PP is localized: whatever happens in set A is independent of whatever happens outside

2.1 Simulating General PPs

- no structure assumed on Ω
- pick a partition $B_1, \dots$ of Ω
- sample $N(B_i) \sim \text{Pois}(\mu(B_i))$
- sample $X_1, \dots, X_{N(B_i)} \sim \mu(\cdot)/\mu(B_i)$.

2.2 Modifying Poisson processes

1. **Superposition theorem:** Let Π_1, Π_2 be independent Poisson processes on the same space with mean measures μ_1 and μ_2. Then, $\Pi = \Pi_1 \cup \Pi_2$ is a Poisson process with mean measure $\mu(A) = \mu_1(A) + \mu_2(A)$. (Generalizes to countable superposition)

2. **Thinning theorem:** Let Π be a Poisson process over Ω with mean measure μ, and let $S(x) \sim \text{Bernoulli}(\rho(x))$ for $\rho : \Omega \to [0, 1]$. Let

$$S(\Pi) = \{X \in \Pi \mid S(X) = 1\}. \tag{1}$$

   Then, $S(\Pi)$ is a Poisson process with mean measure

$$\mu^*(A) = \int_A \rho \, d\mu \tag{2}$$

3. **Mapping Theorem:** Let Π be a Poisson process on Ω, and let $h : \Omega \to \Psi$ be one-to-one. Then,

$$h(\Pi) = \{h(X) \in \Psi \mid X \in \Pi\} \tag{3}$$

   is a Poisson process over Ψ with mean measure

$$h_*\mu(A) = \mu(h^{-1}(A)). \tag{4}$$

   This extends to any case where $h_*\mu$ is non-atomic, e.g. projections

4. **Restriction theorem:** Let Π be a process over Ω with mean measure μ. Let $U \subseteq \Omega$. Then $\Pi|_U = \Pi \cap U$ is a Poisson process with mean measure $\mu|_U(A) = \mu(A \cap U)$.

3 Using PPs for sampling: Exponential Races

Use spatio-temporal processes over $\Omega = \mathbb{R}^+ \times \mathcal{A}$, with mean measure

$$\mu(A) = \int_A p(x \mid t)\lambda(t) \, dx \, dt \tag{5}$$

Idea: we will want to sample from a distribution Q over $\mathcal{A}$, augment this space with time $\mathbb{R}^+$.

Time-projection:

$$\text{proj}(\Pi) = \{t \in \mathbb{R}^+ \mid \exists x \in \mathcal{A} : (t, x) \in \Pi\} \tag{6}$$

Then:

$$\text{proj}_* \mu(B) = \int_B \lambda(t)\, dt \tag{7}$$

Distribution of the first arrival:

$$\mathbb{P}[T \geq t] = \mathbb{P}[N(t) = 0] = e^{-\mu(t)} \tag{8}$$

In general:

$$\mathbb{P}[T_k \geq t \mid T_{k-1}] = \exp\left(-(\mu(t) - \mu(T_{k-1}))\right) \tag{9}$$

Define:

$$\Lambda(t) = \int_0^t \lambda(\tau)\, d\tau \tag{10}$$

Cumulative intensity transform:

$$(\Lambda \circ \text{proj})_* \mu(B) = \int_{\Lambda^{-1}(B)} \lambda(t)\, dt, \quad \text{set } u = \Lambda(t) \tag{11}$$

$$= \int_B du \tag{12}$$

Hence, we can always transform a spatiotemporal process to be time-homogeneous.

3.1 Simulating exponential races

Time homogeneous process:

$$\mathbb{P}[T \geq t] = e^{-t} \quad \Rightarrow \quad T \sim \text{Exp}(1) \tag{13}$$

Hence, simulate $(T_k - T_{k-1}) \sim \text{Exp}$, then compute T_k, then simulate $X_k \sim p(x \mid T_k)$.

3.1.1 Numerical stability: Gumbel processes

Simulate stuff in the log-domain! Let $T_1 \sim \mathbb{E}(\lambda)$

$$\text{Exp}(\lambda) \sim \frac{1}{\lambda} \cdot \text{Exp}(1) \tag{14}$$

$$\Rightarrow -\log T_1 = \log \lambda - \log \mathbb{E}(1) = \log \lambda + G_1 \sim \text{Gumbel}(\log \lambda) \tag{15}$$

Where $\mathbb{P}[G_1 \leq g] = e^{e^{-g}}$. Furthermore:

$$G_k \mid G_{k-1} \sim \text{Gumbel}(0)|_{(-\infty, G_{k-1})} \tag{16}$$

4 Basic Applications

We now have our hammer, let's hit some nails!

4.1 Superposition theorem: Gumbel-max Trick

Let's use a discrete alphabet $|\mathcal{A}| = K$. Pick $\lambda_1, \ldots, \lambda_K \in \mathbb{R}^+$ as our rates, and let $\Pi_1, \ldots, \Pi_K$ be Poisson processes with intensities $\lambda_1 \cdot \delta(x = 1), \ldots, \lambda_K \cdot \delta(x = K)$. Now, define

$$\Pi = \bigcup_k \Pi_k. \tag{17}$$

Then

$$\lambda(t, x) = \sum_k \lambda_k \cdot \delta(x = k). \tag{18}$$

Then, the projected intensity is

$$\begin{aligned} \lambda(t) &= \sum_j \sum_k \lambda_k \cdot \delta(j = k) & (19) \\ &= \sum_k \lambda_k & (20) \end{aligned}$$

From which

$$\begin{aligned} p(x \mid t) &= \frac{\lambda(x, t)}{\lambda(t)} & (21) \\ &= \sum_k \frac{\lambda_k}{\sum_j \lambda_j} \cdot \delta(x = k) & (22) \end{aligned}$$

Thus,

$$p(x = k \mid t) = \frac{\lambda_k}{\sum_j \lambda_j} \tag{23}$$

How can we simulate the first arrival of Π?

1. first arrival of Π only depends on the first arrivals of Π_k
2. simulate first arrival of each process separately: $\frac{1}{\lambda_k} \cdot E_k$, where $E_k \sim \text{Exp}(1)$
3. first arrival time of Π is earliest arrival across all of Π_k:

$$T_1 = \min_k \left\{ \frac{E_k}{\lambda_k} \right\} \tag{24}$$

4. first arrival coordinate of Π is

$$\begin{aligned} X_1 &= \arg\min_k \left\{ \frac{E_k}{\lambda_k} \right\} & (25) \\ &= \arg\max \{G_k + \log \lambda_k\} & (26) \end{aligned}$$

4.2 Thinning theorem: Rejection Sampling

REMINDER: from here onwards, divide the board into picture of PP and calculation
From now: assume we have a target Q and a proposal P. We can sample from P and can evaluate $r = q/p$.

Idea: Set Π as the base process over $\mathbb{R}^+ \times \mathcal{A}$ with mean measure $\lambda \times P$. Then, use one of the theorems to modify Π, find first arrival of modified process.

1. Thinning theorem:

$$\mu^*(A) = \int_A \rho(x)p(x)\,dx\,dt. \tag{27}$$

   Want:

$$\mu^*(A) = \int_A q(x)\,dx\,dt. \tag{28}$$

   Therefore, set

$$\rho(x) = q(x)/p(x). \tag{29}$$

2. However, we need to ensure that $0 \leq \rho(x) \leq 1$. So instead, set

$$\rho(x) = \frac{q(x)}{M \cdot p(x)}, \tag{30}$$

   where $M = \sup\{q(x)/p(x)\}$

3. Hence, to sample, simulate Π, and for an arrival (T, X), delete it with probability $r(x)/M$. Return the first point that wasn't deleted.

4.3 Mapping theorem: A* Sampling

Similar idea as before.

1. Mapping theorem: for h one-to-one, we have

$$h_*\mu([0,s] \times B) = \int_{h^{-1}([0,s]\times B)} p(x)\,dx\,dt \tag{31}$$

   Want:

$$h_*\mu([0,s] \times B) = \int_{[0,s]\times B} q(x)\,dx\,dt \tag{32}$$

$$= s\int_B q(x)dx \tag{33}$$

Idea: Restrict to only temporal shifts. Especially good idea, since $\mathbb{R}^+$ is guaranteed to have "more" structure. Thus, let $h(t, x) = (f(x, t), x)$. Then:

$$h_*\mu([0,s] \times B) = \int_{h^{-1}([0,s]\times B)} p(x)\,dx\,dt \tag{34}$$
$$= \int_B \int_{f^{-1}([0,s],x)} p(x)\,dt\,dx \tag{35}$$
$$= \int_B \int_{f^{-1}(0,x)}^{f^{-1}(s,x)} p(x)\,dt\,dx \tag{36}$$
$$= \int_B (f^{-1}(s,x) - f^{-1}(0,x))p(x)\,dt\,dx \tag{37}$$

Therefore, we want $f^{-1}(0, x) = 0$, and $f^{-1}(s, x)p(x) = s \cdot q(x)$. Thus,

$$f^{-1}(s,x) = s \cdot \frac{q(x)}{p(x)} \tag{38}$$

From which

$$f(t,x) = t \cdot \frac{p(x)}{q(x)} \tag{39}$$
$$h(t,x) = \left(t \cdot \frac{p(x)}{q(x)}, x\right) \tag{40}$$

2. runtime is geometric
3. **Simulation:**
4. **depth-limitation possible**

4.4 Restriction theorem: Greedy Poisson Rejection Sampling

Again, similar idea as before. We pick a function $\varphi : \mathcal{A} \to \mathbb{R}^+$ and let

$$U = \{(t,x) \in \mathbb{R}^+ \times \mathcal{A} \mid t \leq \varphi(x)\} \tag{41}$$

Then, by the restriction theorem, $\Pi|_U$ is a Poisson process with mean measure

$$\mu|_U(A) = \mu(A \cap U) \tag{42}$$

Let $(\tilde{T}, \tilde{X})$ be the first arrival off $\Pi|_U$, and let $\tilde{X} \sim q_\varphi$. Then,

$$\frac{q_\varphi(x)}{p(x)} = \int_0^{\varphi(x)} \mathbb{P}[\tilde{T} \geq t]\,dt. \tag{43}$$

WLOG, we can decompose $\varphi = \sigma \circ r$. Then, we want to pick σ such that $q_\varphi = q$. It turns out, that to achieve this, σ^{-1} must solve

$$\left(\sigma^{-1}\right)' = w_Q\left(\sigma^{-1}\right) - \sigma^{-1} \cdot w_P\left(\sigma^{-1}\right), \tag{44}$$

where

$$w_P(h) = \mathbb{P}_{Z\sim P}\left[\frac{q(Z)}{p(Z)} \geq h\right]. \tag{45}$$

For the triangular-uniform case, where the triangle has base ℓ:

$$\sigma(h) = \frac{2h}{2 - \ell \cdot h} \tag{46}$$