POISSONPROCESSEN
Extra material till TMS137

J. Björnberg
september 2018

Poissonprocessen används ofta vid modellering av vissa tidsförlopp, såsom exempelvis:

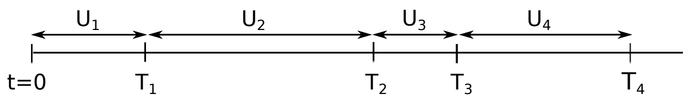
- besök till en hemsida,
- ankomster av kunder till en butik,
- trafik som passerar en viss korsning.

Gemensamt för dessa exempel är att tidsförloppet består av ett slumpmässigt antal ”händelser” eller ”ankomster”, med en slumpmässig mängd tid mellan. Poissonprocessen skrivs som $N(t)$ där t betecknar tiden och $N(t)$ är antalet ”händelser” eller ”ankomster” inom tid t. Så $N(3.4)$ skulle exempelvis kunna beteckna antal bilar som passerat inom tid 3.4 timmar.

1 Notation och definition

För att beskriva Poissonprocessen används följande beteckningar:

- Tiderna för ankomsterna skrivs $T_1, T_2, \ldots$ Av konvention startar vår tidräkning vid $t = 0$, och detta innebär att $0 < T_1 < T_2 < \cdots$ Dessa ankomsttider är slumpvariabler och T_i betecknar tidpunkten för ankomst nummer i.
- Tiderna *mellan* ankomster skrivs $U_1, U_2, \ldots$ och kallas för mellantider. Dessa är också slumpmässiga. Per konvention är U_1 tiden mellan 0 och första ankomsten, dvs $U_1 = T_1$. Sedan är U_2 tiden mellan första och andra ankomsten, dvs $U_2 = T_2 - T_1$. På samma vis gäller $U_3 = T_3 - T_2$ osv.



Själva Poissonprocessen $N(t)$ räknar *antal ankomster inom tid* t, dvs i tidsintervallet $[0, t]$. Genom att fundera en stund kan man se att man kan skriva det i formler som följer:

$$N(t) = \max\{i \geq 0 : T_i \leq t\}.$$

Detta säger att $N(t)$ är det sista värde på i så att ankomsttiden T_i ligger före t, vilket är samma som antal ankomster inom tid t. Det är dock inte så hjälpsamt att memorera detta uttryck, utan det är bättre att komma ihåg vad $N(t)$, T_i samt U_i innebär.

Om man tänker sig att man successivt ökar t från 0 och uppåt så kommer $N(t)$ börja vid 0, och sedan kommer $N(t)$ öka med 1 i taget varje gång t passerar en av ankomsttiderna T_i.

Vi har ännu inte definierat Poissonprocessen. Detta görs genom att specificera fördelningen hos de slumpvariabler (T_i, U_i) som ingår. Smidigast är att specificera fördelningen hos mellantiderna:

Definition: En Poissonprocess $N(t)$ med intensitet λ, förkortat $N(t) \sim \mathrm{PP}(\lambda)$, ges av att mellantiderna $U_1, U_2, \ldots$ är oberoende och alla har fördelning $\mathrm{Exp}(\lambda)$.

(Kom ihåg att U har fördelning $\mathrm{Exp}(\lambda)$ om $P(U \leq t) = 1 - e^{-\lambda t}$ för $t \geq 0$.) I princip skulle man kunna ha valt någon annan fördelning för mellantiderna U_i än just exponentialfördelning. Man väljer exponentialfördelningen för att den visar sig mest användbar i praktiken.

Förutom $N(t)$ själv så tittar man ofta på skillnader $N(t) - N(s)$ är $s < t$. Detta räknar precis antalet ankomster i tidsintervallet $(s, t]$.

2 Egenskaper

För att kunna göra uträkningar med Poissonprocessen så behöver vi följande resultat.

Sats: Låt $N(t) \sim \mathrm{PP}(\lambda)$. Då gäller:

1. För $0 \leq s < t$ så har $N(t) - N(s)$ fördelningen $\mathrm{Po}(\lambda(t-s))$.
2. För $0 \leq t_1 < t_2 < \cdots < t_n$ så är slumpvariablerna

   $$N(t_2) - N(t_1), \quad N(t_3) - N(t_2), \quad \ldots \quad N(t_n) - N(t_{n-1})$$

   alla oberoende.

Satsen säger dels att antal ankomster i ett intervall är Poissonfördelat, dels att antal ankomster i flera *icke-överlappande intervall* är oberoende.

Ett specialfall av första delen av satsen är när $s = 0$. Då säger den att $N(t) = N(t) - N(0)$ har fördelning $\mathrm{Po}(\lambda t)$. I synnerhet gäller därför att $E(N(t)) = \lambda t$. Detta ger en tolkning av λ: det är förväntat antal ankomster per tidsenhet.

2.1 Exempel

Om $N(t) \sim \mathrm{PP}(2)$, vad är

1. Sannolikheten att det är en ankomst mellan tid 0 och 1, samt 3 ankomster mellan tid 1 och 5?
2. Sannolikheten att det är en ankomst mellan tid 0 och 1, *givet* att det är 4 ankomster mellan tid 0 och 5?

Lösningar:

1. $$P\big(N(1) = 1, N(5) - N(1) = 3\big) = P\big(N(1) = 1)P(N(5) - N(1) = 3\big) = \frac{e^{-2}2^1}{1!}\frac{e^{-8}8^3}{3!} \approx 0.0077.$$

2. $$\begin{aligned} P\big(N(1) = 1 \mid N(5) = 4\big) &= \frac{P\big(N(1) = 1, N(5) - N(1) = 3\big)}{P\big(N(5) = 4\big)} \\ &= \frac{P\big(N(1) = 1)P(N(5) - N(1) = 3\big)}{P\big(N(5) = 4\big)} \\ &= \frac{\frac{e^{-2}2^1}{1!}\frac{e^{-8}8^3}{3!}}{\frac{e^{-10}10^4}{4!}} \\ &\approx 0.41. \end{aligned}$$