



SEKOITETUSTA POISSON-PROSESSISTA

Aku Ojala

Pro gradu -tutkielma
Joulukuu 2023

Tarkastajat:
Apulaisprofessori J. Lempa
Tutkijatohtori H. Saarinen

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Turun yliopiston laatu järjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä

TURUN YLIOPISTO

Matematiikan ja tilastotieteen laitos

AKU OJALA: Sekoitetusta Poisson-prosessista

Pro gradu -tutkielma, 25 s., 1 liites.

Matematiikka

Joulukuu 2023

Tässä työssä käsittelen sekoitettuja Poisson-prosesseja ja niiden karakterisointia. Käyn läpi aiheen historiaa ja kehitystä ajan myötä. Tarkastelen aiempien tutkijoiden työtä ja niiden vaikutusta nykyiseen ymmärrykseemme sekoitetuista Poisson-prosesseista. Esittelen luvussa 2 keskeiset käsitteet ja teoreettisen taustan liittyen todennäköisyysteoriaan ja stokastisiin prosesseihin. Seuraavassa luvussa tarkastelen laskuriprosesseja ja niiden ominaisuuksia. Käyn läpi laskuriprosessien määritelmät, ominaisuudet ja erilaiset esimerkit. Palautan tutkielmassa mieleen myös Markov-ketjujen peruskäsitteitä. Samassa luvussa keskityn Poisson-prosessiin, joka on yksi keskeinen laskuriprosessien muoto. Esittelen sen ominaisuudet, sovellukset ja matemaattisen teorian taustalla. Poisson-prosesseista jatkan tarkastelua sekoitettuun Poisson-prosessiin. Tämäkin tapaus käydään läpi määritelmä kerrallaan, aloittaen havainnollistavasta esimerkkitilanteesta. Työn keskeisin osa on sekoitettujen Poisson-prosessien karakterisointi. Tarkastelen kahta eri tapaa, joilla näitä prosesseja voidaan luonnehtia tietyillä ominaisuuksilla ja jakaumilla. Tekstin kulku tulee olemaan erittäin pohjustavaa, ja jokainen uusi käsite selitetään omana määritelmänä. Tämä lähestymistapa mahdollistaa sen, että lukijat voivat silmäillä alkuosan läpi saaden yleiskuvan aiheesta, ja vasta työn loppuvaiheessa he voivat paneutua yksityiskohtiin ja lukea asiat ajatuksen kanssa.

Poisson-prosessi, laskuriprosessi, sekoitettu Poisson-jakauma, karakterisointilauseet

Sisällys

1	Johdanto	1
2	Yleisiä määritelmiä	1
3	Laskuri prosessit	8
3.1	Poisson-prosessi	12
4	Sekoitettu Poisson -prosessi	13
4.1	Sekoitettu todennäköisyysjakauma	13
4.2	Sekoitettu Poisson -jakauma	13
5	Sekoitetun Poisson-prosessin karakterisaatio	17
5.1	Karakterisaatio syntymäprosessina	18
5.2	Karakterisaatio laskuri prosessina	22
	Kirjallisuutta	25
6	Liitteet	26

1 Johdanto

Työn pohjustana käytetään Jan Grandellin teosta *Mixed Poisson Processes* (1997). Jan Grandell työskenteli teosta tehdessään Tukholman kuninkaallisessa teknillisessä korkeakoulussa tilastotieteen professorina. Teos käsittelee Ove Lundbergin tutkielmaa *On random processes and their application to sickness and accident statistics* (1964). Lundberg käsitteli tutkielmassaan Markov-prosesseja, ja vielä tarkemmin, sekoitettua Poisson-prosessia. Sekoitettun Poisson-prosessin idea esitettiin vuonna 1938 Dubourdieuun toimesta. Lundbergin perusajatuksena oli esittää matemaattinen malli sairaus- ja tapaturmavakuutuksiin liittyviin korvausvaatimuksiin. Ongelman lähestymistapa oli siis hyvin käytännönläheinen. 1920-1930-luvulla oli jo tutkittu, että negatiivibinomijakauma sopi korvausvaatimusten mallintamiseen paremmin kuin Poisson jakauma. Lundberg halusi ymmärtää edellä mainitut tutkimuskysymykset teoreettisella näkökannalla, mikä johti hänen tutkielmansa syntyyn.

Grandell pohjustaa omaa teostaan vain yksityiskohtaisena Lundbergin tutkielman tarkasteluna, eikä juurikaan tuo esille uusia näkökulmia. Hän esittelee yhdistettyä Poisson-prosessia sekä laskuriprosessien, että vakuutusmatematiikan näkökulmasta.

2 Yleisiä määritelmiä

Määritellään aluksi todennäköisyyslaskentaan liittyviä peruskäsitteitä ja yleisimpiä lauseita, joita tässä työssä käytetään. Käsitteet löytyvät kirjasta *Todennäköisyyslaskennan ja tilastomatematiikan perusteet*, (1998) [2].

Määritelmä 2.1. Yleistä joukkoa $\Omega = \omega_1, \dots, \omega_n$ kutsutaan otos- tai tapausavaruudeksi. Alkioita $\omega_1, \dots, \omega_n$ kutsutaan alkeistapauksiksi.

Käytetään tapahtuman A komplementtitapahtumasta merkintää \bar{A} .

Määritelmä 2.2. Otosavaruuden Ω osajoukkojen kokoelma \mathcal{F} on *sigma-algebra*, jos

- (i) $\Omega \in \mathcal{F}$

$$(ii) A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$$

$$(iii) A_i \in \mathcal{F}, i = 1, 2, 3, \dots \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}.$$

Sigma-algebraa $\mathcal{B}(R)$ kutsutaan Borelin sigma-algebraksi ja sen alkioita Borelin joukoiksi.

Määritelmä 2.3. Kuvausta $P : \mathcal{F} \rightarrow \mathbb{R}$ kutsutaan todennäköisyysmitaksi tai todennäköisyysdeksi, jos

$$(i) P(A) \geq 0, \text{ kaikille } A \in \mathcal{F}$$

$$(ii) P(\Omega) = 1$$

$$(iii) \text{ Jos } A_i \in \mathcal{F}, i = 1, 2, 3, \dots \text{ ja } A_i \cap A_j = \emptyset, \text{ kun } i \neq j \\ \Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Kolmikolle (Ω, \mathcal{F}, P) käytetään tällöin nimitystä **todennäköisyysavaruus**. Seuraavaan lauseeseen on koottu todennäköisyysmitan perusominaisuuksia.

Lause 2.4. Olkoon kolmikko (Ω, \mathcal{F}, P) todennäköisyysavaruus. Tällöin

$$(i) 0 \leq P(A) \leq 1$$

$$(ii) P(\Omega) = 1$$

$$(iii) P(\emptyset) = 0$$

$$(iv) P(\bar{A}) = 1 - P(A)$$

$$(v) P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$(vi) \text{ Jos } A \cap B = \emptyset, \text{ niin } P(A \cup B) = P(A) + P(B)$$

$$(vii) P(A \setminus B) = P(A) - P(A \cap B)$$

$$(viii) \text{ Jos } A \subset B, \text{ niin } P(A) \leq P(B)$$

Todistus. Todistetaan kohta (v). Jaetaan joukot A ja B aksioomien pohjalta eri osiin:

$$B = (B \cap A) \cup (B \cap \bar{A}), \quad A \cup B = (\bar{A} \cap B) \cup (A \cap \bar{B}) \cup (A \cap B).$$

Nyt saadaan yhtälön (vii) nojalla

$$P(A) = (A \cap B) + P(A \cap \bar{B}) \text{ ja } P(B) = P(B \cap A) + P(B \cap \bar{A}).$$

josta edelleen

$$\begin{aligned} P(A \cup B) &= P(\bar{A} \cap B) + P(A \cap \bar{B}) + P(A \cap B) \\ &= P(B) - P(B \cap A) + P(A) - P(A \cap B) + P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

□

Määritelmä 2.5. Olkoon (D, ζ) ja (F, η) mitta-avaruuksia. Funktiota $X : D \rightarrow F$ kutsutaan **mitalliseksi** sigma-algebroiden ζ ja η suhteen, jos $X^{-1}(\Gamma) \in \zeta$ jokaisella $\Gamma \in \eta$.

Jos $(D, \zeta) = (\Omega, \mathcal{F})$, niin mitallista kuvausta X kutsutaan **satunnaismuuttujaksi**.

Määritelmä 2.6. Satunnaismuuttujan X kertymä- tai jakaumafunktio on muotoa:

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

Määritelmä 2.7. Olkoon $X : \Omega \rightarrow \mathbb{N}$ diskreetti satunnaismuuttuja. Kaikille mahdollisille arvoille $x \in \mathbb{N}$ tapaukset $\{X = x\}$ kuuluvat sigma-algebraan \mathcal{F} . **Pistetodennäköisyysfunktio** p_x määritellään seuraavasti:

$$p_X = P(X = x).$$

Pistetodennäköisyysfunktio antaa todennäköisyyden sille, että satunnaismuuttuja X saa tietyn arvon x . On selvää, että pistetodennäköisyyksien summa on yksi:

$$\sum_x P(X = x) = 1.$$

Lause 2.8. Pistetodennäköisyys- ja kertymäfunktioilla on seuraavat ominaisuudet:

$$1) p_X(x) \in [0, 1], x \in \Omega_X,$$

$$2) \sum_{x \in \Omega_X} p_X(x) = 1,$$

$$3) F_X(x) = \sum_{y \leq x} p_X(y),$$

$$4) \lim_{x \rightarrow -\infty} F_X(x) = 0$$

$$5) \lim_{x \rightarrow \infty} F_X(x) = 1$$

6) F_X on kasvava funktio.

Määritelmä 2.9. Olkoon $X : \Omega \rightarrow \mathbb{N}$ diskreetti satunnaismuuttuja. Tällöin X :n odotusarvo on muotoa

$$\mu = E[X] = \sum_x x p_X(x),$$

jossa ylläoleva sarja *suppenee itseisesti*. [ks. [7], s.302].

Varianssi on muotoa

$$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2] = \sum_x (x - \mu)^2 p_X(x),$$

jossa ylläoleva sarja *suppenee itseisesti*. [ks. [7], s.302].

Määritelmä 2.10. Tapahtumat $A, B \in \Omega$ ovat toisistaan **riippumattomia**, jos

$$P(A \cap B) = P(A) P(B).$$

Yleisesti tapahtumat $A_1, A_2, \dots, A_n \subset \Omega$ ovat toisistaan riippumattomia, jos jokaiselle $I \subset 1, \dots, n$

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

Lause 2.11. Satunnaismuuttujat X ja Y ovat toisistaan riippumattomia, jos

$$F_{X,Y}(x, y) = F_X(x) F_Y(y) \text{ kaikilla } x, y.$$

Määritelmä 2.12. Ehdollinen odotusarvo: Olkoon X ja Y satunnaismuuttujia. Tällöin Y :n odotusarvo ehdolla X on muotoa

$$E(Y|X = x) = \sum_y y P(Y = y|X = x).$$

Määritelmä 2.13. Stokastinen prosessi: Jos joukolla pisteitä on satunnainen rakenne ajan suhteen, niin tällöin voidaan puhua stokastisesta prosessista. Tällöin pisteet muodostavat polun ja koko joukosta puhutaan **polkureaalisaationa**. Olkoon $X(t)$ satunnaismuuttuja, missä $t \in T$, ($T \in \mathbb{Z}$). Tällöin joukko $\{X(t), t \in T\}$ on stokastinen prosessi ja muodostaa ajan mittaan polkureaalisaation.

Määritelmä 2.14. Stationaarinen prosessi: Jos stokastisen prosessin jakauma ei muutu ajan mittaan ennalta määrätyn ajanjakson aikana, tällöin tätä ominaisuutta kutsutaan stationaariseksi. Tällöin myös prosessin odotusarvo ja varianssi pysyvät samana yli ajan. Matemaattisin merkinnöin:

$$F_x(x_{t_1+\tau}, \dots, x_{t_n+\tau}) = F_x(x_{t_1}, \dots, x_{t_n}) \text{ kaikille } \tau, t_1, \dots, t_n \in \mathbb{R} \text{ ja kaikille } n \in \mathbb{N}_{>0}.$$

Poisson-jakauma on diskreetti todennäköisyysjakauma, joka kuvaa yksittäisten tapahtumien esiintymistiheyttä tai lukumäärää. Jakaumalla on monia käytännönläheisiä käyttötapoja esimerkiksi taloustieteessä, vakuutusmaailmassa, teleoperaattoreilla ja eri luonnontieteiden aloilla. Poisson-jakaumaa voidaan esimerkiksi käyttää tilanteessa, jossa ollaan kiinnostuneita myymälään saapuvista asiakkaista tietyn tunnin aikana.

Määritelmä 2.15. Satunnaismuuttuja X on **Poisson-jakautunut** parametrilla λ ($\lambda > 0$), jos X :n arvojoukko on $0, 1, 2, \dots$ ja

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \text{ missä } k = 0, 1, 2, 3, \dots$$

Merkitään $X \sim \text{Poisson}(\lambda)$.

Poisson-jakauman parametri λ kuvaa **intensiteettiä** eli tapahtumisnopeutta.

Määritelmä 2.16. Poisson-jakauman kertymäfunktio on muotoa:

$$F(x) = P(X \leq x) = e^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!}$$

Esimerkki 2.17. Erääseen myymälään saapuu keskimäärin 3 asiakasta minuutissa.

a) Millä todennäköisyydellä tasan 77 asiakasta saapuu seuraavan tunnin aikana myymälään? b) Mikä on todennäköisyys, että yhtään asiakasta ei saavu myymälään seuraavan minuutin aikana? c) Millä todennäköisyydellä ainakin 10 asiakasta saapuu myymälään seuraavan 5 minuutin aikana?

Oletetaan, että asiakkaat saapuvat myymälään toisistaan riippumattomasti.

a) Voimme mallintaa asiakkaiden saapumista myymälään Poisson-jakauman avulla, jossa parametri λ saa arvon $3 \times 60 = 180$ asiakasta tunnissa. Olkoon X satunnaismuuttuja, joka kuvastaa seuraavan tunnin aikana myymälään saapuvien asiakkaiden lukumäärää. Tällöin $X \sim Po(180)$, joten todennäköisyys 77 asiakkaan saapumiseen tunnin aikana saadaan Poisson-jakauman pistetodennäköisyysfunktioista:

$$P\{X = 77\} = \frac{e^{-\lambda} \lambda^{77}}{77!} = \frac{e^{-180} (180)^{77}}{77!} \approx 0.013$$

Siis seuraavan tunnin aikana myymälään saapuu tasan 77 asiakasta noin 1,3% todennäköisyydellä.

b) Olkoon satunnaismuuttuja X seuraavan minuutin aikana myymälään saapuvien asiakkaiden määrä. Tällöin $\lambda = 3$, jolloin

$$P\{X = 0\} = \frac{e^{-\lambda} \lambda^0}{0!} = e^{-3} \approx 0.05$$

Siis noin 5 prosentin mahdollisuudella myymälään ei saavu yhtään asiakasta seuraavan minuutin aikana.

c) Jälleen $X \sim Po(15)$, joten

$$\begin{aligned} P\{X \geq 10\} &= 1 - P\{X \leq 9\} \\ &= 1 - P\{X = 0\} + P\{X = 1\} + \dots + P\{X = 9\} \\ &= 1 - \frac{e^{-15} \cdot 15^0}{0!} + \frac{e^{-15} \cdot 15^1}{1!} + \dots + \frac{e^{-15} \cdot 15^9}{9!} \\ &\approx 0.97 \end{aligned}$$

Siis seuraavan 5 minuutin aikana myymälään saapuu 10 asiakasta 97 prosentin todennäköisyydellä.

Lause 2.18. Olkoon $X \sim Po(\lambda)$. Tällöin odotusarvo ja varianssi ovat muotoa

$$\mu = \mathbb{E}(X) = \lambda \quad \text{ja} \quad Var(X) = \lambda$$

ja momentit generoiva funktio on muotoa

$$M(t) = \mathbb{E}(e^{tX}) = \exp(\lambda e^t - \lambda).$$

Todistus.

$$\begin{aligned} M(t) &= \mathbb{E}(e^{tX}) \\ &= \sum_{x=0}^{\infty} e^{tX} \frac{\lambda^x}{x!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\ &= e^{-\lambda} \cdot \exp(\lambda e^t) \\ &= \exp(\lambda e^t - \lambda). \end{aligned}$$

Ensimmäinen momentti saadaan derivoimalla momentit generoiva funktio ja tutkimalla sen arvoa nollassa. Tällöin

$$M'(t) = \exp(\lambda e^t - \lambda) \lambda e^t,$$

josta saadaan

$$\mathbb{E}(X) = M'(0) = \exp(\lambda \cdot 1 - \lambda) \cdot \lambda \cdot 1 = e^0 \cdot \lambda = \lambda.$$

Toinen momentti saadaan tällöin derivoimalla $M(t)$ kahdesti ja tutkimalla sen arvoa nollassa:

$$M''(t) = \exp(\lambda e^t - \lambda) \cdot \lambda e^t \cdot \lambda e^t + \exp(\lambda e^t - \lambda) \cdot \lambda e^t,$$

josta saadaan

$$\begin{aligned} M''(0) &= \exp(\lambda e^0 - \lambda) \cdot \lambda e^0 \cdot \lambda e^0 + \exp(\lambda e^0 - \lambda) \cdot \lambda e^0 \\ &= \exp(\lambda - \lambda) \cdot \lambda \cdot \lambda + \exp(\lambda - \lambda) \cdot \lambda \\ &= \lambda^2 + \lambda. \end{aligned}$$

Tällöin saadaan varianssiksi

$$\text{Var}(X) = M''(0) - [M'(0)]^2 = \lambda^2 + \lambda - (\lambda)^2 = \lambda.$$

□

3 Laskuriprosessit

Jos sama tapahtuma toistuu useasti, niin sitä voidaan kuvata **laskuriprosessin** avulla. Laskuriprosessi on stokastinen prosessi, joka seuraa kyseisen tapahtuman esiintymiskertojen määrää ajan kuluessa.

Määritelmä 3.1. Olkoon $N(t)$ hetkeen t mennessä tapahtuneiden tapausten lukumäärä. Tällöin muotoa $\{N(t)|t \geq 0\}$ olevaa prosessia kutsutaan laskuriprosessiksi.

Laskuriprosessilla on seuraavat ominaisuudet:

- (i) Tila-avaruus on diskreetti,
- (ii) Parametrijoukko R^+ on jatkuva,
- (iii) $N(0) = 0$ eli hetkeen 0 mennessä on tapahtunut 0 tapausta,
- (iv) Yhdellä ajanhetkellä ei voi sattua useampaa tapausta kuin yksi.

Määritelmä 3.2. Okoon N joukko funktioita $\nu = \{\nu(t); t \geq 0\}$, jolle on ominaista:

- (i) $\nu(0) = 0$;
- (ii) $\nu(t)$ on kokonaisluku kaikilla $t < \infty$;
- (iii) $\nu(\cdot)$ on kasvava ja oikealta jatkuva.

Funktiota $\nu \in N$ kutsutaan laskurifunktioksi.

Nyt saadaan laskurifunktion generoimaksi sigma-algebraksi:

$$\mathcal{B}(N) = \sigma\{\nu(t) \leq y; \nu \in N, t \geq 0, y < \infty\}.$$

Määritelmä 3.3. Olkoon X_i joukko pisteitä, jotka kuvaavat tapahtumien määrää tiettyinä diskreetteinä ajanhetkinä tila-avaruudessa R^+ . **Laskuriprosessi** N on kokoelma satunnaismuuttujia $\{N(B) : B \in \mathcal{B}(S)\}$, jossa $\mathcal{B}(S)$ on Borelin sigma-algebra tila-avaruudessa R^+ . Toisin sanoen $N(B)$ satunnaismuuttuja, joka antaa laskuriprosessin luvun Borelin joukossa B , ja laskuriprosessi voidaan määritellä mitallisena kuvauksena $(\Omega, \mathcal{F}, P) \rightarrow (N, \mathcal{B}(N))$.

Laskuriprosessi on siis stokastinen prosessi, joka mallintaa pisteiden esiintymisiä ajassa tai tilassa (esimerkiksi Eukleedisessa tilassa). Määritellään laskuriprosessin osat sanallisesti:

Määritelmä 3.4. Piste - Joukon osa, missä prosessi tapahtuu. Esimerkiksi jos laskuriprosessia mallinnetaan ajassa, niin 'pisteellä' tarkoitetaan sitä ajanhetkeä, kun tapahtuma ilmenee.

Alkeistapaus - Voidaan määritellä mielivaltaisesti, esimerkiksi 'minuutin aikana sattuu kolme tapahtumaa'.

Satunnaismuuttuja - Laskuriprosessien kontekstissa tarkoitetaan tällöin funktiota, joka kuvantaa alkeistapahtumaa reaalityyppiseksi. Esimerkiksi tapahtumien määrä voi olla satunnaismuuttuja.

Intensiteettifunktio - Funktio joka kertoo tapahtumien keskimääräisen nopeuden eli intensiteetin. Intensiteettifunktion avulla voidaan laskea todennäköisyyksiä liittyen laskuriprosessin tapahtumiin.

Laskuriprosessin sanotaan olevan **yksinkertainen**(simple), jos prosessin lisäykset ovat aina vakiosuuruisia. Laskuriprosessi on stationaarinen, jos termin $N(t) = N(s+t) - N(s)$ jakauma ei muutu kaikilla $s \geq 0$. Stationaariselle

laskuriprosessille on ominaista, että $P\{N(\infty) = 0 \text{ tai } \infty\} = 1$ eli prosessilla on pisteitä joko 0 tai ääretön määrä todennäköisyydellä 1.

Syntymä-kuolema prosessi on yksinkertainen laskuriprosessi, jossa on sekä lisäyksiä, että poistumisia. Kyseessä on jatkuva-aikainen prosessi. Tilojen muuttumiset tapahtuu vain viereisestä tilasta seuraavaan tai edelliseen. Syntymät tapahtuvat intensiteetillä λ ja kuolemat tapahtuvat intensiteetillä μ . Jos laskuriprosessi koostuu vain lisäyksistä, niin tällöin puhutaan syntymäprosessista. Poisson-prosessi on syntymäprosessi vakiointensiteetillä λ , ($\lambda > 0$).

Jos jatkuvassa laskuriprosessissa on vain lisäyksiä, ja lisäyksiä uuteen tilaan tapahtuu intensiteetillä $\lambda n = n \cdot \lambda$, ($n \in \mathbb{N}$) niin on kyse **Yule-prosessista**. Tällöin voidaan puhua myös yksinkertaisesta syntymäprosessista, koska mallissa jokainen yksilö synnyttää uusia jälkeläisiä tasaisella tahdilla.

Laskuriprosessin ominaisuuksia:

Siirtymä-operaattori $T_s : \mathcal{N} \rightarrow \mathcal{N}$, joka on määritelty laskuriprosessien tila-avaruudessa \mathcal{N} muodostetaan seuraavasti:

$$(T_s \nu)(t) = \nu(s + t) - \nu(s), \quad s \geq 0.$$

Siirtymäoperaattori T_s siis siirtää laskuriprosessia s ajanhetken verran eteenpäin. Olkoon $T_s^{-1}B = \{\nu \in \mathcal{N}; T_s \nu \in B\}$, mille tahansa $B \in \mathcal{B}(\mathcal{N})$. Joukkoa $B \in \mathcal{B}(\mathcal{N})$ kutsutaan **invariantiksi**(invariant), jos $B = T_s^{-1}B$ kaikille $s \geq 0$.

Stationaarista laskuriprosessia N jollakin mielekkäällä jakaumalla Π kutsutaan **ergodiseksi**(ergodic), jos $\Pi\{B\} = 0$ tai 1 kaikilla invariantteilla joukoilla $B \in \mathcal{B}(\mathcal{N})$.

Olkoon laskuriprosessi N stationaarinen. Tällöin $E[N(t)] = \alpha \cdot t$, missä parametria α kutsutaan laskuriprosessin **intensiteetiksi**.

On aina olemassa satunnaismuuttuja \bar{N} , jolle $E[\bar{N}] = \alpha$ ja jota kutsutaan **yksilölliseksi intensiteetiksi**(individual intensity), jolle $\frac{N(t)}{t} \rightarrow \bar{N}$ P-m. v., kun $t \rightarrow \infty$. Jos $\alpha < \infty$, niin siitä seuraa että \bar{N} on äärellinen P-m. v.. Koska $\{\nu \in \mathcal{N}; \bar{\nu} \leq x\}$ on invariantti jokaiselle x , niin \bar{N} on **deterministinen**(eli sen arvo ei riipu satunnaisista tekijöistä), jos N on ergodinen.

Laskuriprosessin **harventamisella** tarkoitetaan alkuperäisen prosessin

muuttamista siten, että siitä poistetaan osa pisteistä jollakin todennäköisyydellä p . Sekä jäljelle jäävät pisteet, että poistetut pisteet muodostavat uuden laskuri-prosessin. Samalla periaatteella voidaan myös yhdistää laskuri-prosesseja yhdeksi. Tästä käytetään nimitystä **superpositio**. Olkoon N_1, N_2, N_3 laskuri-prosesseja. Niiden superpositio voidaan ilmaista seuraavasti:

$$\mathbf{N} = N_1 + N_2 + N_3,$$

missä \mathbf{N} edustaa superpositiota ja se koostuu kaikkien kolmen laskuri-prosessin pisteistä yhdistettynä yhdeksi laskuri-prosessiksi.

Markov-ketju ja Markov-prosessi ovat käsitteitä, joita käytetään kuvaamaan stokastisia prosesseja, joissa tilan kehittyminen riippuu vain edellisestä tilasta, eivätkä kaikista menneistä tiloista. Tästä käytetään nimitystä **Markov-ominaisuus**. Toinen kuvaava termi, joka tarkoittaa samaa asiaa on Markov-prosessin **muistittomuus**. Markov-ketju viittaa diskreettiin aikaan ja diskreetteihin tiloihin, kun taas Markov-prosessi voi olla joko diskreetti tai jatkuva-aikaista ja voi sisältää myös jatkuvia tiloja. Näitä käsitteitä käytetään laajalti monilla tieteenaloilla, kuten tilastotieteessä, kemiassa, taloustieteessä ja fysiikassa analysoimaan ja mallintamaan tapahtumien ja tilojen dynamiikkaa.

Määritelmä 3.5. Olkoon (Ω, \mathcal{F}, P) todennäköisyysavaruus ja $S, \mathcal{B}(S)$ mittallinen tila, jossa S on tila-avaruus ja $\mathcal{B}(S)$ on borelin sigma-algebra S :ssä. Stokastista prosessia $X = \{X_t : t \in T\}$ kutsutaan Markov-prosessiksi, jos kaikilla $t \in T$ ja kaikilla $A \in \mathcal{B}(S)$ pätee:

$$P(X_{t+s} \in A | \mathcal{F}_t) = P(X_{t+s} \in A | X_t),$$

missä \mathcal{F}_t on aikaisempia tiloja kuvaavien satunnaismuuttujien generoima sigma-algebra $\sigma(X_s), s \leq t$.

Tietyissä Markov-ketjun tilassa i kulunutta aikaa t merkitään viipymisajalla: diskreeteissä tapauksissa \mathbf{T}_i (holding time), ja jatkuva-aikaisessa tapauksessa \mathbf{R}_i (residence time). Diskreetissä tapauksessa jokaisella ajanhetkellä tila muuttuu (myös nykyisestä tilasta siirtyminen samaan tilaan tulkitaan muutokseksi). Jatkuva-aikaisessa tapauksessa tila ei muutu ennen kuin siirrytään nykyisestä tilasta johonkin muuhun tilaan.

Määritelmä 3.6. Siirtymätodennäköisyys Diskreettiaikaisessa Markovin ketjussa tilan siirtymät kuvataan todennäköisyyksillä, joita kutsutaan siirtymätodennäköisyyksiksi. Siirtymätodennäköisyys $P_{ij}^{(n)}$ ilmaisee todennäköisyyden siirtyä tilasta i tilaan j käyttäen n askelta. Se voidaan määritellä seuraavasti:

$$P_{ij}^{(n)} = P\{X_n = j | X_0 = i\}.$$

Bernoulli-prosessi on diskreettiaikainen stokastinen prosessi, jossa seurataan satunnaismuttujia X_1, X_2, \dots , joille $X_i \in \{0, 1\}$ eli joiden arvot voivat olla joko 0 tai 1. Näissä prosesseissa todennäköisyydet pysyvät vakioina eli $P(X_i = 1) = p$ ja $P(X_i = 0) = 1 - p$. Tästä saamme laskuriprosessin $N(t)$, joka laskee kuinka monta tapahtumaa on tapahtunut hetkeen t mennessä. Tällöin saadaan laskuriksi $\sum_{i=1}^t X_i$. Bernoulli-prosessi on yksi esimerkki diskreettiaikaisesta prosessista. Poisson-prosessi perustuu samoihin periaatteisiin kuin Bernoulli-prosessi, mutta se on sovellettu jatkuva-aikaiseen ympäristöön.

3.1 Poisson-prosessi

Jos laskuriprosessin tilat ovat Poisson-jakautuneita, niin tällöin puhutaan Poisson-prosessista. Poisson-prosessi on yksinkertainen laskuriprosessi, joka laskee vakiointensiteetillä tapahtuvien riippumattomien tapausten sattumusten lukumäärää jatkuvassa ajassa. Tällöin siis tapahtumien lukumäärä on Poisson-jakautunut, tapahtumien välinen aika on eksponentiaalisesti jakautunut ja n :nnen tapahtuman sattumishetki on gamma-jakautunut.

Määritelmä 3.7. Olkoon $\lambda > 0$. Laskuriprosessia $\{N(t), t \in [0, \infty)\}$ kutsutaan Poisson-prosessiksi intensiteetillä λ , jos seuraavat ehdot täyttyvät:

- (i) $N(0) = 0$;
- (ii) Prosessin $N(t)$ lisäykset ovat riippumattomia;
- (iii) Lisäysten määrä aikavälillä, jonka pituus on τ , ($\tau > 0$) on Poisson-jakautunut parametrilla $\lambda\tau$.

4 Sekoitettu Poisson -prosessi

4.1 Sekoitettu todennäköisyysjakauma

Sekoitetussa todennäköisyysjakaumassa satunnaismuuttujan X todennäköisyysjakaumassa jokin parametri on satunnaismuuttuja. Satunnaismuuttuja X on siis jakautunut jakauman F mukaan, jossa F :n parametri θ on jakautunut jakauman G mukaan. Saadaan tällöin jakauma H , joka koostuu jakaumista F ja G . Jakaumaa G kutsutaan tällöin sekoitusjakaumaksi ja parametria θ kutsutaan sekoitusmuuttujaksi. Tällöin sekoitetun jakauman H :n tiheysfunktio on muotoa:

$$p_H(x) = \int p_F(x|\theta) \cdot p_G(\theta) d\theta$$

4.2 Sekoitettu Poisson -jakauma

Esitellään sekoitettu Poisson-jakauma avaavan esimerkkitilanteen kautta. Grandell käytti kirjassaan kyseistä esimerkkiä kuvaamaan ihmisten yksilöllisten erojen satunnaisuuden mallintamista.

Olkoon $N \sim Po(\cdot)$ tietylle yksilölle sattuvat liikenneonnettomuudet vuoden aikana. Tällöin

$$P\{N = k\} = \frac{\alpha^k}{k!} e^{-\alpha}, \quad k = 0, 1, \dots$$

Tässä parametrin α arvo riippuu:

- Ajetusta matkasta
- Liikenneympäristöstä
- Yksilön ajotaidoista

Voidaan siis kutsua parametria α yksilön **tapaturma-alttiudeksi** (accident proneness).

On luontevaa olettaa, että tarkasteltaessa yksilöä eri vuosina myös tapaturma-alttius α muuttuu. Toisaalta jos oletetaan α :n pysyvän samana,

eli riippumattomana sekä $\alpha \sim Po(\cdot)$, niin tällöin yksilön tapaturmia voitaisiin mallintaa **Poisson-prosessilla**. Laajennetaan tarkastelua useammalle henkilölle vuoden aikana, ja oletetaan näiden henkilöiden ajavan yhtä paljon samassa ympäristössä. Heidän tapaturmiensa määrät eroavat silti toisistaan ajotaitojen erojen takia. On helppo olettaa siis, että eri yksilöillä on siis eri α :n arvo. Kuitenkaan tätä arvoa ei voida numeerisesti määrittää. Oletetaan siis jokaisella yksilöllä olevan vastaava **Poisson-prosessi**, joka mallintaa heidän onnettomuuksiaan tulevaisuudessa. Parametrin α arvosta voidaan tietää edes vähän, varsinkin jos yksilöt muodostavat homogeenisen joukon.

Nyt tarkastellaan tapaturma-alttiutta α tietyn satunnaismuuttujan Λ tuloksena, jonka jakauma U tunnetaan. Jakaumaa U kutsutaan **rakennejakaumaksi** (structure distribution). Nyt saadaan tulkittua ylläolevaa yhtälöä N :n ehdollisena jakaumana ehdolla $\Lambda = \alpha$. Tällöin saadaan N :n jakaumaksi:

$$P\{N = k\} = \int_{-\infty}^{\infty} \frac{l^k}{k!} e^{-l} dU(l), \quad k = 0, 1, \dots$$

Tästä jakaumasta käytetään nimeä **sekoitettu Poisson jakauma**. Sekoitettulle Poisson jakaumalle on ominaista:

$$E(N) = E(\Lambda) \quad \text{ja} \quad Var(N) = E(\Lambda) + Var(\Lambda) \geq E(\Lambda)$$

Vertaamalla näitä Poisson jakauman tunnuslukuihin huomataan, että $Var(\Lambda) = 0$, kun $E(N) = Var(N)$. Sekoitettu Poisson jakauma on siis **ylihajottunut** (over-dispersed) verrattuna Poisson jakaumaan. Intuitiivisesti tämä on helppo ymmärtää, sillä $Var(\Lambda)$ voidaan tulkita intensiteetin variaation mittana ja $E(\Lambda)$ tulkitaan samoin kuin Poisson jakauman tapauksessa. Siis $Var[N - \Lambda] = E[\Lambda]$.

Greenwood ja Yule (1920) ja Newbold (1926) olivat tutkineet edellä mainittua aihetta läpi alan pioneereina. He tutkivat α parametria Γ -jakauman tuloksena tiheydellä u , jossa u saadaan kaavasta:

$$u(l) = \frac{\beta^\gamma}{\Gamma(\gamma)} l^{\gamma-1} e^{-\beta l}, \quad l \geq 0, \beta > 0, \gamma > 0. \quad (1)$$

N on tällöin negBin-jakaunut.

$$P\{N = k\} = \binom{\gamma + k - 1}{k} \left(\frac{\beta}{\beta + 1} \right)^\gamma \left(\frac{1}{\beta + 1} \right)^k, \quad k = 0, 1, \dots$$

Sekoitettu Poisson-jakauma on siis selvästi monimutkaisempi Poisson-jakaumaan verrattuna, mutta huomataan Λ olevan eksponentiaalisesti jakautunut tilanteessa $\gamma = 1$:

$$P\{N = k\} = \left(\frac{\beta}{\beta + 1}\right) \left(\frac{1}{\beta + 1}\right)^k, \quad k = 0, 1, \dots$$

Tässä tapauksessa huomataan N olevan geometrisesti jakautunut, joka on jopa yksinkertaisempi jakauma kuin Poisson jakauma.

Jatkamalla aikaisemmin läpikäytyä tarkastelua yksilöllisillä Poisson-prosesseilla, huomataan hieman erilainen määritelmä, josta voidaan käyttää nimitystä **sekoitettu Poisson-prosessi**. Jos sekoitettu Poisson jakauma saa rakennejakaumakseen ylläolevan yhtälön (1), niin tällöin puhutaan **Pólya-prosessista**.

Tutkitaan nyt yksilöä, jolla on liikennevakuutus ajasta $t = 0$ eteenpäin. Aikaan t asti henkilö on tehnyt $N(t)$ korvausvaatimusta. Vakuutusyhtiö haluaa ennustaa kuinka monta korvausvaatimusta seuraavalla laskutuskaudella $(t, t + h]$ tulee olemaan. Siis vakuutusyhtiö haluaa tietää $N(t + h) - N(t)$ käyttäen yksilöön liittyvää historiaa tai dataa. Tästä saadaan:

$$\begin{aligned} & E[N(t + h) - N(t) | N(t) = n] \\ &= E[E[N(t + h) - N(t) | \Lambda, N(t) = n] | N(t) = n] \\ &= hE[\Lambda | N(t) = n]. \end{aligned}$$

Sekoitetun Poisson-prosessin tutkimisen motivaationa voidaan tiivistää Newboldin kolmeen periaatteeseen:

- N1** Yksilöiden tapaturma-alttiudet α eroavat toisistaan.
- N2** Menneisyyden tapaturmat eivät vaikuta tulevaisuuden tapaturmiin.
- N3** Yksilön tapaturma-alttius pysyy vakiona yli ajan.

On helppo nähdä, että **N1** ja **N2** viittaavat riippumattomuuteen ja **N3** stationaarisuuteen, mitkä ovat tärkeitä käsitteitä todennäköisysteoriassa stokastisilla prosesseilla.

Monia jakaumia voidaan käyttää sekoitusmuuttujan θ jakaumana, mutta todellisuudessa mielekästä on käyttää vain kolmea jakaumaa: Gamma-jakaumaa, logNorm-jakaumaa ja käänteistä gaussista jakaumaa. Näiden sekoitus antaa täten (samassa järjestyksessä): negBin-jakauman, Poisson-log-normal -jakauman ja Sichel-jakauman.

Määritelmä 4.1. Olkoon X **Gamma-jakautunut**. Tällöin merkitään $X \sim \Gamma(\alpha, \beta)$, jossa Γ on **Eulerin gammafunktio**. Gammajakautuneen satunnaismuuttujan tiheysfunktio on muotoa:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}},$$

kun $x > 0$ $\alpha, \beta > 0$.

Määritelmä 4.2. Olkoon **logaritmisien normaalijakauman** parametri $X = \ln Y$ logNorm-jakautunut. Parametri Y on tällöin normaalijakautunut. Log-normaalijakautuneen satunnaismuuttujan tiheysfunktio on muotoa:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_Y} \frac{1}{x} \exp\left(-\frac{(\ln x - \mu_Y)^2}{2\sigma_Y^2}\right)$$

Esimerkki 4.3. Tutkitaan sekoitettua Poisson-jakaumaa, missä sekoitusjakaumana on logNormaalijakauma.

Olkoon Λ logNorm-jakautunut. Tällöin

$$u(l) = \frac{1}{\sqrt{2\pi}\sigma_G l} e^{-\frac{(\log l - \mu_G)^2}{2\sigma_G^2}}$$

ja

$$\mu_\Lambda = e^{\mu_G + \frac{\sigma_G^2}{2}}, \quad \sigma_\Lambda^2 = e^{2\mu_G + \sigma_G^2} (e^{\sigma_G^2} - 1).$$

Käänteinen normaalijakauma, tai toiselta nimeltään **Waldin** jakauma, on kahden parametrin jakauma, jonka tiheysfunktio on muotoa:

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right),$$

kaikille $x > 0$, jossa $\mu > 0$ on keskiarvo- ja $\lambda > 0$ muotoparametri.

Martingaalit ovat jono satunnaismuuttujia, joille missä tahansa ajanjaksossa seuraavan jonon jäsenen ehdollinen odotusarvo on sama kuin nykyinen arvo.

Määritelmä 4.4. **F-martingaali** $M = \{M(t); t \geq 0\}$ on reaaliarvoinen prosessi jolle,

1. $M(t)$ on \mathcal{F}_t -mitallinen kun $t \geq 0$,
2. $E[|M(t)|] < \infty$ kun $t \geq 0$,
3. $[M(t)|\mathcal{F}_s] = M(s)$ P-m.v. kun $t \geq s$.

Määritelmä 4.5. **F-martingaali** M on **oikealta jatkuva** jos:

1. Kaikki polut $M(t)$ ovat oikealta jatkuvia
2. Filtraatio \mathbf{F} on oikealta jatkuva, eli

$$\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s \quad \text{jos } t \geq 0.$$

5 Sekoitettun Poisson-prosessin karakterisaatio

Karakterisointilauseet ovat teoreemoja, jotka osoittavat yhteyden satunnaismuuttujien jakauman ja tietyissä yleisissä funktioissa ilmenevien ominaisuuksien välillä. Karakterisointia voidaan ajatella siten, että tarkasteltaessa eri ryhmiä A, B, C, jotka sisältävät yleisiä funktioita ja määritelmiä, voimme yrittää karakterisoida sekoitetun Poisson-prosessin tietyn ryhmän sisään. Jos karakterisaatio esimerkiksi ryhmään C on mahdollista, voimme suoraan käyttää näitä ryhmän C ominaisuuksia sekoitettuun Poisson-prosessiin sen sijaan, että täytyisi todistaa kaikki sen ominaisuuksiin liittyvät teoreemat erikseen.

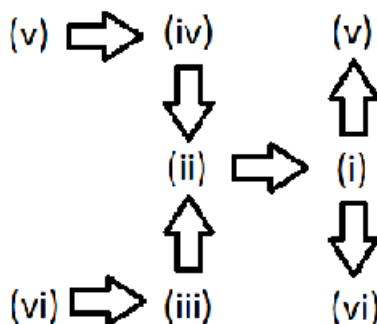
5.1 Karakterisaatio syntymäprosessina

Koska syntymäprosessi on hyvin yksinkertainen laskuriproessi, niin tällä karakterisaatiolla saadaan paljon määritelmiä avattua sekoitetulle Poisson-prosessille. Seuraavan lauseen todisti Lundberg tutkielmassaan *On random processes and their application to sickness and accident statistics* (1964).

Lause 5.1. Olkoon N syntymäprosessi intensiteetillä $\kappa_n(t)$ ja reunajakamalla $p_n(t)$. Seuraavat toteamukset ovat ekvivalentteja:

- (i) N on sekoitettu Poisson-prosessi
- (ii) $\kappa_n(t)$ täyttää ehdon $\kappa_{n+1}(t) = \kappa_n(t) - \frac{\kappa_n'(t)}{\kappa_n(t)}$, $n = 0, 1, \dots$
- (iii) $\kappa_n(t)$ ja $p_n(t)$ täyttävät ehdon $p_n(t) = \frac{t}{n} \kappa_{n-1}(t) p_{n-1}(t)$, $n = 1, 2, \dots$
- (iv) $E[\kappa_{N(t)}(t) | N(s) = m] = \kappa_m(s)$, kun $0 < s \leq t$ ja $m = 0, 1, \dots$
- (v) $E[N(t) - N(s) | N(s) = m] = \kappa_m(s)(t - s)$, kun $s \leq t$ ja $m = 0, 1, \dots$
- (vi) $P\{N(s) = m | N(t) = n\} = \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}$, kun $s \leq t$ ja $m \leq n$.

Todistus. Seurataan Lundbergin alkuperäistä todistusta vuodelta 1964. Todistuksen kulku näkyy kuvasta.



Kuva 1: Todistuksen kulun havainnollistaminen. Kuvan lähde: Grandell, (1997) [1].

Lähdetään liikkeelle kohdasta (ii).

(ii) \rightarrow (i)

Olkoon $\hat{u}(t) = p_0(t)$, jossa $\hat{u} \geq 0$ kun $t \geq 0$. yhtälöstä (6.1) seuraa:

$$\kappa_0(t) = -\frac{\hat{u}^{(1)}(t)}{\hat{u}(t)},$$

ja tällöin $-\hat{u}^{(1)}(t) \geq 0$, kun $t \geq 0$, sillä $\kappa_0(t) \geq 0$. Oletetaan nyt, että

$$\kappa_k(t) = -\frac{\hat{u}^{(k+1)}(t)}{\hat{u}^{(k)}(t)} \quad (2)$$

ja, että

$$(-1)^{k+1}\hat{u}^{(k+1)}(t) \geq 0, \text{ kun } t \geq 0, \quad k = 0, \dots, n. \quad (3)$$

Ottamalla logaritminen derivaatta kohdasta (ii), saadaan:

$$\kappa_{n+1}(t) = -\frac{\hat{u}^{(n+1)}(t)}{\hat{u}^{(n)}(t)} - \frac{\hat{u}^{(n+2)}(t)}{\hat{u}^{(n+1)}(t)} + \frac{\hat{u}^{(n+1)}(t)}{\hat{u}^{(n)}(t)} = -\frac{\hat{u}^{(n+2)}(t)}{\hat{u}^{(n+1)}(t)}.$$

Käyttämällä yhtälöä (3) ja $\kappa_{n+1}(t) \geq 0$, saadaan seurauksena, että yhtälöt (2) ja (3) pitävät paikkansa kaikille $k \geq 0$. Siis \hat{u} on täysin monotoninen. Koska $\hat{u}(0) = p_0(0) = 1$ niin tästä seuraa, että $\hat{u}(t)$ on Laplace muunnos ja tällöin intensiteetit määrittelevät sekoitetun Poisson-prosessin.

(iii) \Rightarrow (ii).

Ottamalla logaritminen derivaatta kohdasta (iii), saadaan:

$$\frac{p'_n(t)}{p_n(t)} = \frac{1}{t} + \frac{\kappa'_{n-1}(t)}{\kappa_{n-1}(t)} + \frac{p'_{n-1}(t)}{p_{n-1}(t)}. \quad (4)$$

yhtälöstä (6.1) ja (iii) saadaan:

$$\frac{p'_k(t)}{p_k(t)} = -\kappa_k(t) + \kappa_{k-1}(t) \frac{p_{k-1}(t)}{p_k(t)} = -\kappa_k(t) + \frac{k}{t}.$$

Edellinen yhtälö on myös tosi tilanteessa, kun $k = 0$. Asettamalla yhtälöön $k = n$ ja asettamalla $n - 1$ yhtälöön (4) saadaan (ii).

(iv) \Rightarrow (ii)

(iv) kohdasta saadaan:

$$\kappa_m(s) = \sum_{n=m}^{\infty} \kappa_n(t) p_{m,n}(s, t).$$

Josta yhtälöiden (6.3) ja (iv) kanssa saadaan:

$$\begin{aligned} \kappa'_m(s) &= \sum_{n=m}^{\infty} \kappa_n(t) \frac{\partial p_{m,n}(s, t)}{\partial s} \\ &= \kappa_m(s) \cdot \left(\kappa_m(t) p_{m,n}(s, t) + \sum_{n=m+1}^{\infty} \kappa_n(t) [p_{m,n}(s, t) - p_{m+1,n}(s, t)] \right) \\ &= \kappa_m(s) [\kappa_m(s) - \kappa_{m+1}(s)], \end{aligned}$$

joka on sama kuin (ii).

(v) \rightarrow (iv).

Olkoon m ja s vakioita ja $e(t)$ muotoa:

$$e(t) = \sum_{n=m}^{\infty} (n - m) p_{m,n}(s, t).$$

Käyttämällä etenevää yhtälöä (6.2) saadaan:

$$\begin{aligned} e'(t) &= \sum_{n=m}^{\infty} (n - m) (-\kappa_n(t) p_{m,n}(s, t) + \kappa_{n-1}(t) p_{m,n-1}(s, t)) \\ &= \sum_{n=m}^{\infty} \kappa_n(t) p_{m,n}(s, t) \end{aligned}$$

Tällöin $e'(t) = \kappa_m(s)$, jolloin (iv) pitää paikkansa.

(vi) \Rightarrow (iii).

Koska

$$P\{N(s) = m, N(t) = n\} = P\{N(s) = m | N(t) = n\} p_n(t) = p_{m,n}(s, t) p_m(s),$$

niin kohdasta (vi) seuraa, että yhtälö (6.4) pätee. Otetaan logaritminen derivaatta yhtälöstä (6.4) t :n suhteen ja päästään muotoon:

$$\frac{\partial p_{m,n}(s, t)}{\partial t} = \left(-\frac{n}{t} + \frac{n-m}{t-s} + \frac{p'_n(t)}{p_n(t)} \right) p_{m,n}(s, t). \quad (5)$$

Yhtälöistä (6.1) ja (6.2) kun $m < n$, saadaan yhtälö (5) sievennettyä muotoon:

$$\kappa_{n-1}(t) \frac{p_{m,n-1}(s, t)}{p_{m,n}(s, t)} = -\frac{n}{t} + \frac{n-m}{t-s} + \kappa_{n-1}(t) \frac{p_{n-1}(t)}{p_n(t)}. \quad (6)$$

Käyttämällä edelleen (6.4) saadaan:

$$\frac{p_{m,n-1}(s, t)}{p_{m,n}(s, t)} = \frac{n-m}{n} \frac{t}{t-s} \frac{p_{n-1}(t)}{p_n(t)}$$

Tällöin yhtälö (6) sievenee muotoon:

$$\kappa_{n-1}(t) \left(\frac{n-m}{n} \frac{t}{t-s} - 1 \right) \frac{p_{n-1}(t)}{p_n(t)} = -\frac{n}{t} + \frac{n-m}{t-s}$$

tai eri tavalla merkittävästi

$$\kappa_{n-1}(t) \frac{p_{n-1}(t)}{p_n(t)} = \frac{n}{t},$$

joka on kohta (iii).

(i) \Rightarrow (v).

Olkoon $U^*(x) = \frac{\int_{0_-}^x l^m e^{-sl} dU(l)}{\int_{0_-}^{\infty} l^m e^{-sl} dU(l)}$. Nyt

$$E[N(t) - N(s) | N(s) = m] = \int_0^{\infty} (t - s) l dU^*(l) = (t - s) \kappa_m(s).$$

(i) \Rightarrow (vi).

Tämä seuraa suoraan siitä, että (vi) pitää paikkansa Poisson-prosessille riippumatta intensiteetistä α . Tällöin se pitää myös paikkansa, kun intensiteetti on satunnaisesti generoitu. Täten olemme todistaneet lauseen 5.1. \square

Edellisen jokseenkin pitkän todistuksen huomattavin kohta on (iv). Oletetaan, että $\kappa_0(0) < \infty$. Koska N on markovilainen, saadaan:

$$E[\kappa_{N(t)}(t) | N(s)] = E[\kappa_{N(t)}(t) | \mathcal{F}_s^N].$$

Ylläoleva yhtälö voitaisiin siis kirjoittaa seuraavaan tapaan.

Lause 5.2. Syntymäprosessi N , jolle $\kappa_0(0) < \infty$, on sekoitettu Poisson-prosessi, jos ja vain jos $\kappa_{N(t)}(t)$ on \mathbf{F}^N -martingaali.

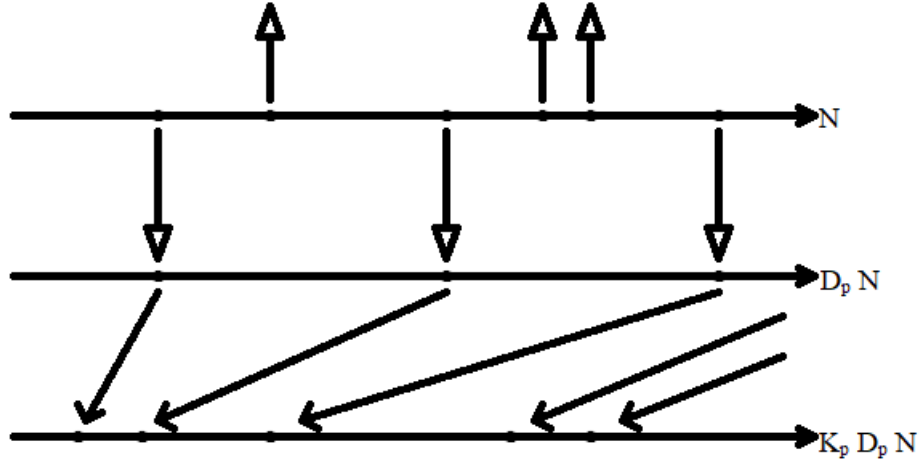
Tämä määrittely on ensimmäinen martingaali-karakterisaatio laskuri-prosessiteoriassa.

5.2 Karakterisaatio laskuri-prosessina

Palautetaan mieleen laskuri-prosessin stationaarisuuden määritelmä. laskuri-prosessia N kutsutaan stationaariseksi, jos $N(t) = N(s+t) - N(s)$ jakauma ei muutu kaikilla $s \geq 0$.

Olkoon N sekoitettu Poisson-prosessi jakaumalla U , missä U on ei-negatiivisen satunnaismuuttujan Λ jakauma ja olkoon U_c satunnaismuuttujan $c\Lambda$ jakauma. Nyt **ohennusoperaattori** (thinning operator) $D_p N$ on sekoitettu Poisson-prosessi jakaumalla U_p . Voimme nyt kompensoida ohentamista tiivistämällä aikaikkunaa.

Määritellään **kompresio-operaattori** K_p , $K_p N(t) = N(t/p)$. Tästä seuraa aikakompressoitu prosessi $K_p N$, joka on sekoitettu Poisson-prosessi jakaumalla $U_{1/p}$. Täten $K_p D_p \Pi_U = \Pi_U$ mille tahansa sekoitetulle Poisson-prosessille. Operaattori $K_p D_p$ on kuvattu seuraavassa kuvassa.



Kuva 2: Alkuperäisestä prosessista N ohennetaan operaattorilla $D_p N$ ja kompressoidaan operaattorilla $K_p D_p N$. Kuvan lähde: Grandell, (1997) [1].

Lause 5.3. Olkoon N stationaarinen laskuri prosessi jollain sopivalla jakaumalla Π . Seuraavat väitteet ovat ekvivalentit:

- (i) N on sekoitettu Poisson-prosessi.
- (ii) $K_p D_p \Pi = \Pi$ jollekin $p \in (0, 1)$.
- (iii) Hyppyjen määrä välillä $[0, t]$ annettuna $N(t) = n$ on tasajakautunut, kun $n \geq 1$ ja $t > 0$.

Todistus. (ii) \rightarrow (i).

Koska $K_{p^{n+1}} D_{p^{n+1}} \Pi = K_{p^n} D_{p^n} K_p D_p \Pi = K_p D_p \Pi$, niin tästä seuraa, että (ii) pitää paikkansa kun p, p^2, p^3, \dots . Asetetaan

$$p_n = p^n \quad \text{ja} \quad N_n = K_{p_n} N.$$

Koska $D_{p_n} N_n = N$, saadaan triviaalisti $D_{p_n} N_n \xrightarrow{d} N$ (\xrightarrow{d} tarkoittaa konvergenssia jakautumisessa eli niiden jakautumat lähenevät toisiaan, kun otoskoko kasvaa). Nyt yhtälöstä (6.5) seuraa, että on olemassa satunnaismitta Λ , jolle $p_n N_n \xrightarrow{d} \Lambda$. Tällöin N on Cox prosessi satunnaismittalla Λ . Tästä seuraa

$$p_n N_n(1) = p_n N(1/p_n) \xrightarrow{d} \Lambda(1), \quad \text{kun } n \rightarrow \infty. \quad (7)$$

Koska N on stationaarinen, niin on olemassa satunnaismuuttuja \bar{N} , jota kutsutaan **yksilölliseksi intensiteetiksi** jolle

$$N(s)/s \rightarrow \bar{N} \quad \text{P-m. v. kun } s \rightarrow \infty. \quad (8)$$

Yhtälöstä (7) seuraa, että $\bar{N} \stackrel{d}{=} \Lambda(1)$ ($\stackrel{d}{=}$ tarkoittaa yhtäläisesti jakautunutta) ja tällöin $P\{\bar{N} < \infty\} = 1$. Jatkaen yhtälöstä (8)

$$p_n N_n(t) = p_n N(t/p_n) = t \cdot \frac{p_n}{t} N(t/p_n) \rightarrow t \cdot \bar{N} \quad \text{P-m. v., kun } n \rightarrow \infty.$$

Täten $(\Lambda(t_1), \dots, \Lambda(t_n)) \stackrel{d}{=} (t_1 \cdot \bar{N}, \dots, t_n \cdot \bar{N})$ ja kohta (i) seuraa tästä.

(iii) \rightarrow (i)

Koska tasajakauma on jatkuva jakauma, niin ainoastaan yksi hyppy voi tapahtua tietynä ajanhetkenä. Toisin sanoen N on yksinkertainen laskuri-prosessi. Koska (6.6), niin on riittävä osoittaa, että:

$$P\{N\{A\} = 0\} = \int_{0-}^{\infty} e^{-|A|} dU(l), \quad A \in \mathcal{I}_U, \quad (9)$$

missä $|A|$ on A :n välien pituuksien summa. Tässä \mathcal{I}_U kuvastaa välin $[0, \infty)$ äärellisten unionien joukkoa.

Valitaan $A \in \mathcal{I}_U$ ja $t > \sup\{x | x \in A\}$ ja asetetaan $s = |A|$. Nyt

$$\begin{aligned} P\{N\{A\} = 0\} &= \sum_{n=0}^{\infty} P\{N\{A\} = 0 | N(t) = n\} P\{N(t) = n\} \\ &= \sum_{n=0}^{\infty} \left(\frac{t-s}{t}\right)^n P\{N(t) = n\}, \end{aligned}$$

ja nähdään, että $P\{N\{A\} = 0\} = P\{N(s) = 0\}$.

Satunnaismuuttuja $N(s)$ syntyy satunnaismuuttujasta $N(t)$ ohentamalla.

Koska diskreettiä satunnaismuuttujaa voidaan pitää laskuri-prosessina yksidimensioisella tila-avaruudella, niin käytetään edelleen lausetta (6.5). Valitaan t_1, t_2, \dots siten, että $\lim_{k \rightarrow \infty} t_k = \infty$.

Käyttämällä ohentamisoperaattoria D_p satunnaismuuttujaan $N(t_k)$, saadaan $N(s) \stackrel{d}{=} D_{s/t_k} N(t_k)$ ja lauseesta (6.5) seuraa

$$\frac{s}{t_k} N(t_k) \xrightarrow{d} \text{jokin satunnaismuuttuja } \Lambda_s \quad (10)$$

ja $N(s)$ on $SP(U_s)$, missä U_s on satunnaismuuttujan Λ_s jakauma ja $SP(\cdot)$ tarkoittaa sekoitettua Poisson-prosessia.

Mikä tahansa A valitaan niin $t_k s$ termi tulee olemaan tarpeeksi suuri, eikä t_1, t_2, \dots valintaa tarvitse verrata A :n valintaan. Siis yllä olevat perustelut pätevät kaikille A .

Asettamalla yhtälöön (10) $s = 1$, seuraa että $\Lambda_s \stackrel{d}{=} s \cdot \Lambda_1$. Tällöin $N(|A|)$ kaikilla $A \in \mathcal{I}_U$ on $SP(|A|, U_1)$ ja yhtälö 9 seuraa. Täten olemme todistaneet lauseen (5.3).

□

Kirjallisuutta

- [1] Jan Grandell (1997). *Mixed Poisson Processes*, Chapman and Hall.
- [2] Antti Niemi (1998). *Todennäköisyyslaskennan ja tilastomatematiikan perusteet*, Kolmas painos, Hakapaino Oy.
- [3] Jukka Lempa (kevät 2019). *Todennäköisyyslaskentaa ja -teoriaa, luentomoniste*. Turun Yliopisto, 2019.
- [4] Christian Webb (syksy 2020). *Markov Chains - lecture notes*. Åbo Akademi University, 2020.
- [5] Kalle Parvinen (syksy 2003). *Riskiteoria, luentomoniste*. Turun Yliopisto, 2003.
- [6] Ove Lundberg (1964). *On random processes and their application to sickness and accident statistics*, Almqvist & Wiksells Boktryckeri AB.
- [7] Petteri Harjulehto, Riku Klén, Mika Koskenoja (2016). *Analyysiiä reaalityyppillä*, Kolmas painos.

6 Liitteet

Lemma 6.1.

$$\begin{aligned} p_0(t) &= -\kappa_0(t)p_0(t), \\ p_n(t) &= -\kappa_n(t)p_n(t) + \kappa_{n-1}(t)p_{n-1}(t), \quad n > 0. \end{aligned}$$

Lemma 6.2.

$$\begin{aligned} \frac{\partial p_{n,n}(s,t)}{\partial t} &= -\kappa_n(t)p_{n,n}(s,t), \\ \frac{\partial p_{m,n}(s,t)}{\partial t} &= -\kappa_n(t)p_{m,n}(s,t) + \kappa_{n-1}(t)p_{m,n-1}(s,t), \quad \text{kaikilla } m < n. \end{aligned}$$

Lemma 6.3.

$$\begin{aligned} \frac{\partial p_{m,m}(s,t)}{\partial s} &= \kappa_m(s)p_{m,m}(s,t), \\ \frac{\partial p_{m,n}(s,t)}{\partial s} &= \kappa_m(s)[p_{m,n}(s,t) - p_{m+1,n}(s,t)], \quad \text{kaikilla } m < n. \end{aligned}$$

Lemma 6.4.

$$\begin{aligned} p_{m,n}(s,t) &= p_{n-m}^*(t-s) \\ &= \int_{0-}^{\infty} \frac{(l \cdot (t-s))^{n-m}}{(n-m)!} e^{-l \cdot (t-s)} dU^*(l) \\ &= \frac{\int_{0-}^{\infty} \frac{l^n \cdot (t-s)^{n-m}}{(n-m)!} e^{-lt} dU(l)}{\int_{0-}^{\infty} l^m e^{-ls} dU(l)} \\ &= \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m} \frac{p_n(t)}{p_m(s)}. \end{aligned}$$

Lemma 6.5.

$$p_k(t) = \int_{0-}^{\infty} \frac{(lt)^k}{k!} e^{-lt} dU_t(l), \quad \text{kaikilla } k = 0, 1, \dots$$

Lemma 6.6. Yksinkertaisen laskuriprosessin jakauma Π määräytyy kaikkien B :n muotojen $\Pi\{B\}$ perusteella, missä

$$B = \{\nu \in \mathcal{N}; \nu\{A\} = 0\}, \quad A \in \mathcal{I}_U.$$