

Markov-modulált Poisson folyamatok statisztikai vizsgálata

Diplomamunka

Írta: Nagy Tamás

Matematikus szak

Témavezető:

Arató Miklós, egyetemi docens
Valószínűségelméleti és Statisztika Tanszék
Eötvös Loránd Tudományegyetem, Természettudományi Kar



Eötvös Loránd Tudományegyetem
Természettudományi Kar

2006

Tartalomjegyzék

| | |
|---|-----------|
| 1. Bevezető | 1 |
| 1.1. Bevezető | 1 |
| 1.2. A használt jelölések és rövidítések | 3 |
| 2. A Markov-modulált Poisson folyamat | 4 |
| 2.1. A pontfolyamatok elméletének rövid áttekintése | 4 |
| 2.2. Duplán sztochasztikus (Cox) folyamatok | 14 |
| 2.3. Néhány szó mátrix értékű függvények differenciálásáról és a mátrix exponenciális függvényről | 21 |
| 2.4. Véges állapotterű, folytonos idejű Markov folyamatok | 26 |
| 2.5. Kvázi-Markov folyamatok és Markov felújítási sorozatok | 28 |
| 2.6. A Markov-modulált Poisson folyamat, mint speciális Cox folyamat | 32 |
| 3. Az MMPP paramétereinek becslése az EM algoritmus segítségével | 37 |
| 3.1. Az EM algoritmus | 37 |
| 3.2. Egy EM algoritmus Markov-modulált Poisson folyamatokra | 39 |
| 3.3. Az algoritmus implementálása | 43 |
| 4. Markov-modulált Poisson folyamat a gyakorlatban | 50 |
| 4.1. Minta generálása adott paraméterű MMPP-ből | 50 |
| 4.2. Egy biztosításmatematikai feladat | 52 |
| 4.3. Numerikus eredmények | 58 |
| 5. Záró megjegyzések | 62 |
| 5.1. Összegzés | 62 |
| 5.2. További nyitott kérdések | 62 |
| 5.3. Köszönetnyilvánítás | 63 |
| Függelék | 63 |
| A. A futási eredmények összefoglaló táblázatai | 64 |
| A.1. A valós paraméterek és az EM algoritmus által kapott becslés | 64 |
| A.2. A várható kárszám becslése | 67 |
| A.3. Kvantilisbecslések | 69 |
| Irodalomjegyzék | 72 |

Ábrák jegyzéke

| | |
|--|----|
| 2.1. A $t \mapsto N_t$ trajektóriák egy-egy realizációja $\mu = 1$ és $\mu = 1.5$ esetén | 13 |
| 2.2. Egy SMP lehetséges realizációja | 30 |
| 2.3. Egy MMPP lehetséges realizációja | 33 |
| 4.1. Napi kárszámok egy MMPP egyik realizációja esetén | 53 |

Táblázatok jegyzéke

| | |
|--|----|
| 1.1. Főbb jelölések és rövidítések | 3 |
| 3.1. A kiinduló becslés pontossága | 46 |
| 3.2. A kiinduló becslés pontossága (folytatás) | 46 |
| 4.1. A Rydén által vizsgált 2A és 2B eset | 58 |
| 4.2. Szimulált kárszámok a 2A és 2B esetben | 58 |
| 4.3. Rydén becslése és a mostani becslés | 59 |
| 4.4. Az $\mathbb{E}(\zeta_{k+1})$ előrejelzések a három különböző modell esetén | 59 |
| 4.5. A 0.95-os megbízhatósági szinthez tartozó kvantilisbecslések | 59 |
| 4.6. A 0.99-os megbízhatósági szinthez tartozó kvantilisbecslések | 60 |
| 4.7. Az MMPP becslésből számított kvantiliseltéréseinek a gyakorisága | 60 |
| 4.8. A kevert Poisson becslésből számított kvantiliseltéréseinek a gyakorisága | 60 |
| 4.9. MMPP eltérések gyakorisága 0.99-es mb. szinten | 61 |
| 4.10. Kevert Poisson gyakoriságok 0.99-es mb. szinten | 61 |
| A.1. A becsülendő paraméterek és a kiinduló becslések | 64 |
| A.2. A becsülendő paraméterek és a kiinduló becslések (folytatás) | 65 |
| A.3. A becsülendő paraméterek és az EM algoritmus által kapott becslés | 66 |
| A.4. A becsülendő paraméterek és az EM algoritmus által kapott becslés (folytatás) | 67 |
| A.5. Az $\mathbb{E}(\zeta_{k+1})$ előrejelzések valós és becsült értékei | 67 |
| A.6. Az $\mathbb{E}(\zeta_{k+1})$ előrejelzések valós és becsült értékei (folytatás) | 68 |
| A.7. A 95%-os megbízhatósági szinthez tartozó kvantilisbecslések | 69 |
| A.8. A 95%-os megbízhatósági szinthez tartozó kvantilisbecslések (folytatás) | 70 |
| A.9. A 99%-os megbízhatósági szinthez tartozó kvantilisbecslések | 70 |
| A.10. A 99%-os megbízhatósági szinthez tartozó kvantilisbecslések (folytatás) | 71 |

1. fejezet

Bevezető

1.1. Bevezető

Az olyan pontfolyamatoknak, melyek intenzitása az időben véletlenszerűen változik, számos gyakorlati alkalmazása van, a sorbanállási és kiszolgálási rendszerektől kezdve az orvostudomány és pénzügyi modelleken át a számítógépes hálózatokig bezárólag. A Markov-modulált Poisson folyamat (MMPP) - szemléletesen kifejezve - egy olyan általánosított Poisson folyamat, melynek intenzitása egy adott időpillanatban egy folytonos időparaméterű és véges állapotterű Markov folyamat aktuális állapota által meghatározott, s mint ilyen, a véletlentől függő érték.

Az MMPP első használata a kiszolgálási rendszerek elméletében Naor és Yechiali [15] cikkéhez fűződik, melyet aztán a Neuts által írt [16] publikáció követett. Azóta több kiváló értekezés is született a témakörben. Az egyik legátfogóbb - és egyben talán a leggyakrabban hivatkozott - mű a W. Fisher és K. Meier-Hellstern által 1993-ban közölt cikk [6]. Szintén alapvetőnek számítanak a nem sokkal később napvilágot látó [22] és [23] publikációk.

Az MMPP gyakorlati modellezés során való alkalmazhatósága miatt természetes módon merül fel a paraméterek becslésének kérdése. Ez egyrészt a mögöttes Markov folyamat kiinduló valószínűségi vektorának és infinitezimális generátormátrixának, másrészt a különböző intenzitások becsléséből áll. A jelenleg ismert módszerek azonban nem koncentrálnak a kiinduló valószínűségi vektor becslésére, hanem azzal a feltételezéssel élnek, hogy az nem egy független paraméter, hanem jól definiált módon kapható a többi paraméterből, pl. ez az eset áll fenn akkor, ha az említett kiinduló valószínűségi vektor éppen a mögöttes Markov folyamat stacionárius eloszlása. Az MMPP rendjét szintén ismertnek feltételezzük. A becslés során alkalmazható technikákról jó áttekintés található [22] -ben, melyben T. Rydén egyebek mellett bemutatott egy általánosított EM (GEM) algoritmust a maximum-likelihood becslés kiszámítására a 2-állapotú MMPP estében. Ennek az algoritmusnak a hátránya az, hogy a maximalizációs - ún. M - lépés nem explicit. 2 évre rá azonban Rydén módosította korábbi algoritmusát (lásd [23]) úgy, hogy az már egy explicit M-lépést és egy valamivel bonyolultabb, ám még kezelhető E-lépést eredményezett. További fejlesztésként megszüntette a csupán 2-állapotú esetre való szorítkozást, az új algoritmus már tetszőleges rendű MMPP-re is alkalmazhatóvá vált.

A [23]-ben közölt algoritmusnak azonban továbbra is maradt két gyenge pontja. Egyrészt, az E-lépés kiszámításában szerepet játszó ún. előre tekintő, illetve hátráló rekurziók kalkulációja során olyan értékek merültek fel, melyek meghaladják egy „átlagos” személyi számítógép számítási kapacitását. Rydén az említett cikkben felhívta erre a figyelmet, megjegyezve, hogy a probléma

a rekurziók alkalmas átskálázásával elkerülhető, azonban nem közölte ennek konkrét kivitelezését. Másrészt - szintén az E-lépés kapcsán - mátrix exponenciális függvények bizonyos integráljainak kiértékelése merült fel, melyhez Rydén akkor mátrix-diagonalizációt használt, ami jelentősen lelassítja az algoritmust. Ezen problémás részek kezelésére csupán a közelmúltban született megoldás, W.J.J. Roberts, Y. Ephraim és E. Dieguez 2006-ban megjelent publikációjukban mindkét kérdés részletes tárgyalásra került, a szerzők bemutattak egy megfelelő skálázási eljárást az előre tekintő, illetve hátráló rekurziók normalizálására, illetve hatékony és szép megoldást adtak a másik problémára is, azaz a mátrix exponenciális függvény integráljainak kiszámítására C. Van Loan egy korábbi [29] eredménye alapján.

Jelen dolgozat egyik tárgya T. Rydén és W.J.J. Roberts et. al. eredményeinek áttekintése és ismertetése, illetve a Rydén féle algoritmus újbóli implementálásának részletes leírása a [21]-ben közölt eredmények ismeretében. Tekintünk továbbá egy olyan biztosításmatematikai problémát, ahol egy biztosítótársaság valamely adott, homogén kockázatú üzletágában felmerülő kárigények számának időbeli fejlődésének leírására MMPP modellt használ és a rendelkezésre álló minta alapján készített becslés ismeretében kíván előrejelzéseket adni egy soron következő időperiódusban bekövetkező károk várható számára, illetve az időperiódusra eső károk számának az eloszlásának a különböző elvárt megbízhatósági szintekhez tartozó kvantilis értékekre.

A dolgozat további felépítése a következők szerint alakul. A 2. fejezet a Markov-modulált Poisson folyamatok kellően megalapozott bevezetésére törekszik. A fő definíciók a 6. alfejezetben vannak, az ezt megelőző alfejezetek ezek előkészítését szolgálják. A 3. fejezetben kerül bemutatásra az EM algoritmus, mint általános iteratív paraméterbecslési eljárás nem megfigyelhető adattal rendelkező feladatok esetén, majd speciálisan a már korábban említett Rydén-féle EM algoritmus MMPP-k becslésére a W.J.J. Roberts et. al. által eszközölt javításokkal. A 4. fejezetben egy biztosításmatematikában felmerülő, MMPP modellt használó feladatot vizsgálunk a 3. fejezetben megismert algoritmus segítségével, s ennek kapcsán konkrét numerikus eredményeket is közlünk. Az 5. fejezetben összefoglaljuk a dolgozat eredményét, illetve megemlítjük az esetleges további kutatási lehetőségeket és nyitva maradt kérdéseket.

1.2. A használt jelölések és rövidítések

| Jelölés vagy rövidítés | Magyarázat |
|------------------------------------|---|
| $F(\infty)$ | $\lim_{x \rightarrow \infty} F(x)$ |
| f.a.e. | független, azonos eloszlású (valószínűségi változók) |
| SMP | kvázi-Markov folyamat |
| MRS | Markov felújítási sorozat |
| MMPP | Markov-modulált Poisson folyamat |
| m.b. | majdnem biztosan |
| M_{ij} | az M mátrix i -edik sorának j -edik eleme |
| (m_{ij}) | az a mátrix, melynek i . sorának j . oszlopában m_{ij} áll |
| M_i | Az M mátrix i -edik sora |
| $M^{(j)}$ | Az M mátrix j -edik oszlopa |
| $M_i.$ | Az M mátrix i -edik sorában lévő elemek összege |
| I | identitás mátrix |
| 0 | csupa nulla mátrix |
| $F(t+)$ | $\lim_{s \rightarrow 0+} F(t+s)$ |
| $\mathbf{1}$ | csupa egy oszlopvektor |
| $a \vee b$ | $\max(a, b)$ |
| $\mathcal{M}(S, \mathcal{B})$ | (S, \mathcal{B}) -pontmértékek tere |
| $\bar{\mathbb{N}}_0$ | $\mathbb{N}_0 \cup \{\infty\}$ |
| $\nu _{\mathcal{A}}$ | az \mathcal{A} σ -algebrán értelmezett mérték, vagy a ν mérték \mathcal{A} -ra vett megszorítása |
| $\frac{d\nu}{d\Lambda}$ | a ν mérték Radon-Nykodim deriváltja a Λ mértékre nézve |
| $\mathcal{M}_R(S, \mathcal{B})$ | véletlen mértékek tere |
| \underline{d} | eloszlásbeli egyenlőség |
| Q_X, F_X, f_X | az X valószínűségi változó eloszlása, eloszlásfüggvénye, illetve sűrűségfüggvénye |
| k.v.v. | kiindulási valószínűségi vektor |
| $\text{diag}(x)$, ahol x vektor | Az x vektorból képzett diagonális mátrix |
| $\text{diag}(M)$, ahol M mátrix | Az M mátrix diagonálisában álló elemekből képzett vektor |

1.1. táblázat. Főbb jelölések és rövidítések

2. fejezet

A Markov-modulált Poisson folyamat

Az 1. alfejezet egy általános bevezetést nyújt a pontfolyamatok elméletéről. Itt az alapvető definíciók R.-D. Reiss [20] munkájában leírt felépítésből származnak, az alfejezet túlnyomó része azonban saját munka, különös tekintettel a pontfolyamatok és számlálófolyamatok kapcsolatának megvilágítására. A 2. alfejezet az ún. duplán sztochasztikus folyamatok általános elméletét ismerteti, azon okból kifolyólag, hogy egy Markov-modulált Poisson folyamat egy speciális duplán sztochasztikus folyamat. Itt kétféle definíciót ismertetünk, a 2.2.2 definíció saját, míg a 2.2.4 definíció Reiss-től származik. Az alfejezet további része - mely megvizsgálja például a két definíció kapcsolatát - mind saját eredmény. Mátrixok exponenciálisával kapcsolatos összefüggések folyton előkerülnek majd az egész dolgozat során, ezért az ehhez kapcsolódó legfontosabb információkat a 3. alfejezetben gyűjtöttük össze. Itt a 2.3.8 lemma saját eredmény, mely C. Van Loan egy korábbi lemmájának (lásd [29], 2. fejezet, 1. tétel) analógiájára készült. Ez a lemma csak egy későbbi fejezetben kerül felhasználásra. A 4. alfejezet a folytonos idejű, véges állapotú Markov folyamatok elméletéből ismertet néhány szükséges fogalmat. Az 5. alfejezet a kvázi-Markov folyamatok és Markov felújítási sorozatok alapvető definícióit ismerteti R. Pyke [18] és [19] felépítése alapján. Az alfejezet szükségességét az indokolja, hogy - amint azt később látni fogjuk - az MPPP-k tárgyalása beilleszthető az így kapott általános elméletbe. A 6. alfejezet definálja a Markov-modulált Poisson folyamat fogalmát. Az alfejezetben saját eredményként levezetünk egy összefüggést a $\mathbb{P}(N_{T+t} - N_T = n)$ mennyiségekre (N_t definícióját lásd az alfejezetben) a W. Fisher és K. S. Meier-Hellstern által felírt Chapman-Kolmogorov egyenletek segítségével.

2.1. A pontfolyamatok elméletének rövid áttekintése

Jelen alfejezet a pontfolyamatok elméletének és a hozzájuk kapcsolódó fogalmak rövid áttekintését tűzi ki célul. Először ismertetünk egy általános, mértékelméleti felépítést, mely R.-D. Reiss [20] munkáját követi, majd ezeket ültetjük át a gyakorlati alkalmazásoknak megfelelő valós esetre. Nem törekszünk a pontfolyamatok szép, ám igencsak szerteágazó elméletéről teljes képet adni, csak azon elemi összetevők ismertetésére szorítkozunk, amelyek a további fejezetek megértéséhez szükségesek. Tegyük fel, hogy adott egy S alaphalmaz, melyet a továbbiakban *állapottérnek* nevezünk.

2.1.1. Definíció. Egy $C \subset S$ legfeljebb megszámlálható számosságú halmazt konfigurációnak hívunk.

2.1.2. Definíció. Legyen $C = \{x_i : i \in I\} \subset S$ egy adott konfiguráció az állapottéren. Ekkor a

$\mu_C(B) = |C \cap B|$ képlettel meghatározott diszkrét mértéket a B σ -algebrán értelmezett, és a C konfigurációhoz tartozó pontmértéknek nevezzük.

Könnyen láthatóan $\mu_C = \sum_{i \in I} \delta_{x_i}$, ahol δ_x a Dirac-delta.

2.1.3. Definíció. Legyen (S, \mathcal{B}) egy mérhető tér, és legyen μ egy \mathcal{B} -n értelmezett halmazfüggvény. Ha létezik olyan $C \subset S$ konfiguráció, hogy μ előáll $\mu(B) = |C \cap B|$ alakban, akkor azt mondjuk, hogy a μ halmazfüggvény (S, \mathcal{B}) -pontmérték.

2.1.1. Megjegyzés. Az elnevezés jogos, hiszen $\mu(\emptyset) = |C \cap \emptyset| = 0$, $\mu(B) \geq 0$ és páronként diszjunkt $B_1, B_2, \dots \in \mathcal{B}$ esetén

$$\mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \left|C \cap \left(\bigcup_{i=1}^{\infty} B_i\right)\right| = \left|\bigcup_{i=1}^{\infty} (C \cap B_i)\right| = \sum_{i=1}^{\infty} |C \cap B_i| = \sum_{i=1}^{\infty} \mu(B_i),$$

tehát μ valóban mérték.

Az S -en értelmezett konfigurációk terét jelölje $\mathcal{C}(S) = \{C \subset S : C \text{ konfiguráció}\}$, az (S, \mathcal{B}) -pontmértékek terét pedig $M \equiv M(S, \mathcal{B})$. Legyen \varkappa az a $\mathcal{C}(S) \mapsto M(S, \mathcal{B})$ leképezés, ami egy adott C konfigurációhoz hozzárendeli a $\mu_C = |C \cap \cdot|$ pontmértéket. Ekkor definíció szerint μ szürjektív, viszont nem feltétlen injektív. Ha azonban a \mathcal{B} σ -algebra kellően gazdag struktúrájú - például, ha \mathcal{B} tartalmazza az egy pontú halmazokat -, akkor az injektivitás igaz, tehát ezekben az esetekben \varkappa egy bijekció.

A következőkben az lesz a célunk, hogy alkalmas módon definiáljunk egy σ -algebrát az M téren. Ehhez tekintsük minden $B \in \mathcal{B}$ -re a

$$\begin{aligned} \pi_B : M &\mapsto \bar{\mathbb{N}}_0 \\ \mu &\mapsto \mu(B) \end{aligned}$$

leképezést - ahol $\bar{\mathbb{N}}_0 = \mathbb{N}_0 \cup \{\infty\}$ -, majd ezek után legyen \mathcal{M} a legszűkebb olyan σ -algebra, melyre nézve az összes π_B leképezés ($B \in \mathcal{B}$) - mint $(M, \mathcal{M}) \mapsto (\bar{\mathbb{N}}_0, P(\bar{\mathbb{N}}_0))$ leképezés mérhető, azaz

$$\mathcal{M} \equiv \mathcal{M}(S, \mathcal{B}) = \sigma\{\pi_B : B \in \mathcal{B}\} = \sigma\{\pi_B^{-1}(C) : B \in \mathcal{B}, C \subset \bar{\mathbb{N}}_0\}$$

A továbbiakban - ha külön nem jelezzük - $\bar{\mathbb{N}}_0$ -ba képező függvények mérhetőségét mindig $P(\bar{\mathbb{N}}_0)$ -ra nézve értjük.

Most már rátérhetünk a *pontfolyamat* fogalmának definíciójára. Ennek érdekében rögzítsünk egy $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mezőt.

2.1.4. Definíció. Egy $N : \Omega \mapsto M(S, \mathcal{B})$ leképezést (S, \mathcal{B}) -pontfolyamatnak nevezünk, ha $(\mathcal{A}, \mathcal{M})$ -mérhető.

N tehát egy véletlen mérték \mathcal{B} -n. Mielőtt megvizsgálánk egy konkrét példát, bebizonyítunk egy karakterizációs tételt, mely segítségével könnyebben ellenőrizhető lesz majd a definícióban szereplő mérhetőségi tulajdonság. Először is, minden $B \in \mathcal{B}$ -re definiálható az

$$\begin{aligned} N(B) : \Omega &\mapsto \bar{\mathbb{N}}_0 \\ \omega &\mapsto N(\omega)(B) \end{aligned}$$

leképezés. $N(B)$ tehát nem más, mint a B -be eső pontok véletlen száma. $N(S)$ pedig az N pontfolyamat *mintaelemszáma*. Könnyen láthatóan igaz a következő előállítás:

$$N(B) = \pi_B \circ N. \quad (2.1)$$

Mivel N minden $\omega \in \Omega$ -ra egy mérték, ezért fennállnak az alábbi tulajdonságok:

(1) $N \geq 0$

(2) Ha $A_1, A_2, \dots \in \mathcal{B}$ páronként diszjunktak, akkor $N\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} N(A_k)$

(3) Ha $B \subset C$, $N(B) < \infty$, akkor $N(C \setminus B) = N(C) - N(B)$

(4) Ha $B \subset C$, akkor $N(B) \leq N(C)$

A továbbiakban $N(B)$ -re, mint a N leképezés B -re vonatkozó *egydimenziós marginálisára* fogunk hivatkozni.

2.1.2. Megjegyzés. Természetesen a marginálisok definiálásához nem szükséges feltenni az $(\mathcal{A}, \mathcal{M})$ -mérhetőséget, illetve az sem szükséges, hogy N minden ω -ra véges értékű mértéket vegyen fel.

Ezek után az alábbi egyszerű, ám fundamentális jellegű kritérium fogalmazható meg:

2.1.1. Állítás. Legyen adott az $N : \Omega \mapsto \mathcal{M}(S, \mathcal{B})$ leképezés. Ekkor a következő két állítás ekvivalens:

(i) N pontfolyamat.

(ii) $N(B) : \Omega \mapsto \overline{\mathbb{N}}_0$ $(\mathcal{A}, P(\overline{\mathbb{N}}_0))$ -mérhető minden $B \in \mathcal{B}$ -re.

Bizonyítás. (i) \Rightarrow (ii) \mathcal{M} konstrukciója miatt triviális, hiszen az (2.1) előállítás értelmében $N(B)$ két mérhető leképezés kompozíciója.

(ii) \Rightarrow (i) A $\{\pi_B^{-1}(C) : B \in \mathcal{B}, C \subset \overline{\mathbb{N}}_0\}$ halmazrendszer \mathcal{M} egy generátorrendszere, ezért N mérhetőségéhez elég belátni, hogy minden $B \in \mathcal{B}$ és $C \subset \overline{\mathbb{N}}_0$ -ra $N^{-1}(\pi_B^{-1}(C)) = (\pi_B \circ N)^{-1}(C) = N(B)^{-1}(C) \in \mathcal{A}$, ami viszont $N(B)$ mérhetőségi tulajdonsága miatt nyilvánvaló. ■

2.1.1. Példa. Legyenek adottak az $X_1, X_2, \dots, X_k : \Omega \mapsto (S, \mathcal{B})$ valószínűségi változók, és N legyen egyenlő $\sum_{i=1}^k \delta_{X_i}$ -vel. Így pontfolyamatot kapunk, hiszen ha $B \in \mathcal{B}$, akkor tekintve, hogy N a δ_{X_i} , $i = 1, 2, \dots$ függvények összege, elég belátni a δ_{X_i} függvények mérhetőségét, ami viszont nyilvánvaló abból a tényből kifolyólag, hogy mérhető halmaz indikátora is mérhető:

$$\delta_{X_i}(B) = \chi(X_i \in B).$$

A következőkben azt tűzzük ki célul, hogy a 2.1.1 állítás második pontjának feltevését gyengítsük oly módon, hogy azt a \mathcal{B} halmazrendszer elemei helyett elegendő legyen egy kevésbé gazdag struktúrájú halmazrendszer elemeire leellenőrizni. Ehhez először igazolunk egy lemmát, amit a későbbiekben is többször felhasználunk.

2.1.1. Lemma. Legyen adott egy $\mathcal{F} \subset P(X)$ halmazrendszer, mely rendelkezik az alábbi három tulajdonsággal:

1. X előállítható \mathcal{F} -beli halmazok megszámlálható uniójaként,
2. \mathcal{F} különbségzárt és

3. zárt a megszámlálható diszjunkt unióra.

Ekkor \mathcal{F} σ -algebra.

Bizonyítás. Elég belátni, hogy \mathcal{F} σ -uniózárt és komplementerzárt. Az előbbihez tekintsük a tetszőleges $A_i \in \mathcal{F}, i = 1, 2, \dots$ halmazokat. Ezeket diszjunktizáljuk, azaz megadunk egy olyan páronként diszjunkt halmazokból álló $B_i \in \mathcal{F}, i = 1, 2, \dots$ halmazrendszert, amire $\cup_{i=1}^{\infty} B_i = \cup_{i=1}^{\infty} A_i$. Ezt a következő módon tehetjük meg: $B_1 := A_1 \in \mathcal{F}$, majd $i \geq 2$ -re legyen $B_i := A_i \setminus \cup_{k=1}^{i-1} B_k \in \mathcal{F}$, ez nyilván teljesíti a kívánt tulajdonságot. Tehát az A_i halmazok unióját előállítottuk a páronként diszjunkt B_i halmazok uniójaként, ami viszont \mathcal{F} -beli, hiszen \mathcal{F} zárt a megszámlálható diszjunkt unióra. Tehát \mathcal{F} valóban σ -uniózárt. A lemma 1. feltétele és a most bizonyított uniózártság miatt $X \in \mathcal{F}$. Ezután, ha $A \in \mathcal{F}$, akkor a különbségzártság miatt $A^C = X \setminus A \in \mathcal{F}$, tehát \mathcal{F} valóban különbségzárt. Ezt kellett bizonyítanunk. ■

Ezek után már belátható az alábbi állítás.

2.1.2. Állítás. Legyen adott az $N : \Omega \mapsto M(S, \mathcal{B})$ leképezés. Legyen a $\mathcal{G} \subset \mathcal{P}(S)$ halmazrendszer a \mathcal{B} σ -algebra egy generátorrendszere és tegyük fel, hogy S előállítható \mathcal{G} -beli halmazok megszámlálható uniójaként. Ekkor az alábbi három állítás ekvivalens.

(i) N pontfolyamat.

(ii) $N(B) : \Omega \mapsto \bar{\mathbb{N}}_0$ ($\mathcal{A}, P(\bar{\mathbb{N}}_0)$)-mérhető minden $B \in \mathcal{B}$ -re.

(iii) $N(B) : \Omega \mapsto \bar{\mathbb{N}}_0$ ($\mathcal{A}, P(\bar{\mathbb{N}}_0)$)-mérhető minden $B \in \mathcal{G}$ -re.

Bizonyítás. (i) \Rightarrow (ii) \Rightarrow (iii) nyilvánvaló, be kell látnunk még a (iii) \Rightarrow (i) irányt. Ehhez legyen

$$\mathcal{E} = \sigma\{\pi_B : B \in \mathcal{G}\}.$$

$\{\pi_B : B \in \mathcal{G}\} \subset \{\pi_B : B \in \mathcal{B}\}$ -ből következik, hogy $\mathcal{E} \subset \sigma\{\pi_B : B \in \mathcal{B}\}$. Legyen $\mathcal{F} = \{B \in \mathcal{B} : \pi_B \mathcal{E}\text{-mérhető}\}$. Ekkor nyilvánvalóan $\mathcal{G} \subset \mathcal{F} \subset \mathcal{B}$. Másrészt fenállnak az alábbiak.

1. S előállítható megszámlálható sok \mathcal{G} -beli - s így egyúttal \mathcal{F} -beli - halmaz uniójaként.

2. \mathcal{F} különbségzárt, hiszen $A, B \in \mathcal{F}, A \subset B$ esetén

$$\pi_{B \setminus A}(\nu) = \nu(B \setminus A) = (\nu(B) - \nu(A))\chi(\nu(A) < \infty) = (\pi_B(\nu) - \pi_A(\nu))\chi(\pi_A(\nu) < \infty),$$

ami viszont mérhető, hiszen $\nu \mapsto \pi_B(\nu), \nu \mapsto \pi_A(\nu)$, illetve $\nu \mapsto \chi(\pi_A(\nu) < \infty)$ egyaránt mérhető.

3. \mathcal{F} zárt a megszámlálható diszjunkt unióra, hiszen ha $B_i \in \mathcal{F}, i = 1, 2, \dots$ páronként diszjunkt halmazok, akkor

$$\pi_{\cup B_i}(\nu) = \nu(\cup B_i) = \sum_{i=1}^{\infty} \nu(B_i) = \sum_{i=1}^{\infty} \pi_{B_i}(\nu),$$

azaz $\pi_{\cup B_i} = \sum_{i=1}^{\infty} \pi_{B_i}$, megszámlálható sok mérhető függvény összege pedig megint csak mérhető.

Így a 2.1.2 lemma segítségével azt kapjuk, hogy \mathcal{F} σ -algebra, ekkor azonban csak az $\mathcal{F} \equiv \mathcal{B}$ eset lehetséges. Ez éppen azt jelenti, hogy $\pi_B \mathcal{E}$ -mérhető minden $B \in \mathcal{B}$ -re, vagyis $\sigma\{\pi_B : B \in \mathcal{B}\} \subset \mathcal{E}$. Ekkor tehát

$$\sigma\{\pi_B : B \in \mathcal{G}\} = \sigma\{\pi_B : B \in \mathcal{B}\},$$

vagy másképpen írva

$$\sigma\{\pi_B^{-1}(C) : B \in \mathcal{G}, C \subset \bar{\mathbb{N}}_0\} = \sigma\{\pi_B^{-1}(C) : B \in \mathcal{B}, C \subset \bar{\mathbb{N}}_0\},$$

azaz a $\{\pi_B^{-1}(C) : B \in \mathcal{G}, C \subset \bar{\mathbb{N}}_0\}$ halmazrendszer is \mathcal{M} egy generátorrendszere. Így elég belátnunk, hogy minden $B \in \mathcal{G}$ és $C \subset \bar{\mathbb{N}}_0$ -ra $N^{-1}(\pi_B^{-1}(C)) = (\pi_B \circ N)^{-1}(C) = N(B)^{-1}(C) \in \mathcal{A}$, ami viszont $N(B)$ mérhetőségi tulajdonsága miatt nyilvánvaló. ■

A következő definíciók szintén alapvetőek a pontfolyamatok elméletében:

2.1.5. Definíció. Az $N(S, \mathcal{B})$ -pontfolyamat intenzitásmértéke alatt a $\nu(B) = \mathbb{E}N(B)$, $(B \in \mathcal{B})$ összefüggéssel definiált mértéket értjük.

$\nu(B)$ tehát a B halmazba eső pontok várható számát adja meg. Könnyű ellenőrizni, hogy ν valóban mérték \mathcal{B} -n.

2.1.2. Példa. A 2.1.1 példában szereplő folyamat intenzitásmértékére

$$\nu(B) = \mathbb{E}N(B) = \sum_{i=1}^k \mathbb{E}(\delta_{X_i}(B)) = \sum_{i=1}^k \mathbb{P}(X_i \in B) = \sum_{i=1}^k \mathbb{P}(X_i \in B) = \sum_{i=1}^k Q_{X_i}(B)$$

adódik, azaz $\nu = \sum_{i=1}^k Q_{X_i}$.

2.1.3. Példa. Legyen $\nu|\mathcal{B}$ egy egészen σ -véges mérték. Azt mondjuk, hogy az $N : \Omega \mapsto M(S, \mathcal{B})$ pontfolyamat Poisson folyamat, ha az alábbi két tulajdonság teljesül:

1. $N(B)$ Poisson eloszlású $\nu(B)$ paraméterrel $\forall B \in \mathcal{B}, \nu(B) < \infty$ -re, és
2. $N(B_1), N(B_2), \dots, N(B_k)$ függetlenek minden $k \in \mathbb{N}_0$ és diszjunkt $B_1, B_2, \dots, B_k \in \mathcal{B}, \nu(B_i) < \infty$ esetén.

Ekkor az intenzitásmérték éppen $\mathbb{E}N(B) = \nu(B)$, amennyiben $\nu(B) < \infty$. ν egészen σ -véges tulajdonságát kihasználva egyszerűen megmutatható, hogy ekkor a Poisson folyamat előállítható megszámlálható sok páronként független, véges intenzitású Poisson folyamat összegeként.

Koncentráljunk most arra az esetre, ami a tipikusan felmerülő szituáció a gyakorlati alkalmazásokban, vagyis amikor az S állapottér a pozitív valós félegyenes - azaz $S = \mathbb{R}^+ = (0, \infty)$ - és \mathcal{B} az S Borel-halmazainak a σ -algebrája: $\mathcal{B} = \mathcal{B}(\mathbb{R}^+) \equiv \mathcal{B}_0$. Mivel ekkor \mathcal{B}_0 tartalmazza az egy pontú halmazokat, ezért korábbi észrevételünk alapján ekkor a $\varkappa : \mathcal{C}(\mathbb{R}^+) \mapsto M(\mathbb{R}^+, \mathcal{B}_0)$ leképezés bijektív, azaz egy pontmérték és az őt meghatározó konfiguráció kölcsönösen egyértelműen meghatározzák egymást. Tekintsük most a következő definíciókat.

2.1.6. Definíció. Azt mondjuk, hogy a $\mu \in M(\mathbb{R}^+, \mathcal{B}_0)$ pontmérték torlódó, ha a hozzá tartozó konfigurációnak létezik ∞ -tól különböző torlódási pontja. Egyéb esetben a μ pontmértéket nem torlódónak mondjuk.

Jelölje $M_0(\mathbb{R}^+, \mathcal{B}_0)$ a nem torlódó pontmértékek terét, azaz

$$M_0 := \{\mu \in M : \varkappa^{-1}(\mu)\text{-nek nem létezik } \infty\text{-tól különböző torlódási pontja}\}.$$

Tehát $M_0(\mathbb{R}^+, \mathcal{B}_0)$ azon pontmértékek terét jelöli, amelyekre igaz az, hogy a hozzájuk tartozó konfigurációnak nem létezik ∞ -tól különböző torlódási pontja. Nyilvánvaló, hogy $\mu \in M_0$ akkor és csak akkor, ha μ korlátos intervallumokhoz véges értéket rendel. Kevés utánagondolással belátható, hogy fennáll az alábbi karakterizáció is.

2.1.3. Állítás. Legyen adott a μ pontmérték és tekintsük \mathbb{R}^+ felbontását megszámlálható sok korlátos intervallumra: $\mathbb{R}^+ = \bigcup_{i=1}^{\infty} B_i$. Ekkor

$$\mu \in \mathbf{M}_0 \Leftrightarrow \mu(B_i) < \infty \quad \forall B_i, i = 1, 2, \dots \text{ esetén.}$$

Hogy \mathbf{M}_0 is mérhető tér legyen, definiáljunk rajta egy \mathcal{M}_0 σ -algebrát, mégpedig az $\mathcal{M}(\mathbb{R}^+, \mathcal{B}_0)$ σ -algebra $\mathbf{M}_0(\mathbb{R}^+, \mathcal{B}_0)$ -ra vett nyomaként, azaz:

$$\mathcal{M}_0 := \mathcal{M}|_{\mathbf{M}_0} = \{H \cap \mathbf{M}_0 : H \in \mathcal{M}\}.$$

Persze az is világos, hogy $\mathcal{M}_0 = \sigma\{\pi_B|_{\mathbf{M}_0} : B \in \mathcal{B}_0\}$. Ekkor tekintsük a következő definíciót:

2.1.7. Definíció. Azt mondjuk, hogy az $N : \Omega \mapsto \mathbf{M}(\mathbb{R}^+, \mathcal{B}_0)$ pontfolyamat nem torlódó, ha $N \in \mathbf{M}_0(\mathbb{R}^+, \mathcal{B}_0)$ majdnem biztosan.

A 2.1.3 állítás segítségével belátható a következő állítás.

2.1.4. Állítás. Legyen adott az \mathbb{R}^+ állapotter $\mathbb{R}^+ = \bigcup_{i=1}^{\infty} B_i$ felbontása megszámlálható sok korlátos intervallumra és tekintsük az $N : \Omega \mapsto \mathbf{M}(\mathbb{R}^+, \mathcal{B}_0)$ pontfolyamatot. Ekkor

$$N \text{ nem torlódó} \Leftrightarrow \mathbb{P}(N(B_i) < \infty) = 1 \quad \forall B_i, i = 1, 2, \dots \text{ esetén.}$$

Bizonyítás. N nem torlódó $\Leftrightarrow \mathbb{P}(N \in \mathbf{M}_0) = 1 \Leftrightarrow \mathbb{P}(\bigcap_{i=1}^{\infty} (N(B_i) < \infty)) = 1$
 $\Leftrightarrow \mathbb{P}(\bigcup_{i=1}^{\infty} (N(B_i) = \infty)) = 0 \Leftrightarrow \mathbb{P}(N(B_i) = \infty) = 0, i = 1, 2, \dots \Leftrightarrow \mathbb{P}(N(B_i) < \infty) = 1,$
 $i = 1, 2, \dots$ esetén. ■

Az alábbiakban következő valós esetre vonatkozó definíciókban, illetve állításokban ahol szükséges, mindig feltesszük, hogy a pontfolyamatunk nem torlódó. Ez nem jelent számunkra nagy megszorítást, hiszen a későbbi vizsgálódásaink tárgyát képező pontfolyamatok úgyis nem torlódók lesznek majd.

Tekintsük most az alábbi függvényt:

$$\begin{aligned} \mu_{(\cdot)} : \mathbb{R}^+ &\mapsto \bar{\mathbb{N}}_0 \\ t &\mapsto \mu((0, t]) \end{aligned}$$

Ekkor nyilvánvaló, hogy $\mu_{(\cdot)}$ jobbról folytonos módon kiterjeszthető \mathbb{R}_0^+ -ra a $\mu_0 = \lim_{t \rightarrow 0^+} \mu_t$ értelmezéssel. Egyben az is világos, hogy μ_0 csak ∞ vagy 0 lehet attól függően, hogy a μ pontmérték torlódó-e vagy sem. Vezessünk is be erre egy definíciót.

2.1.8. Definíció. A $\mu \in \mathbf{M}(\mathbb{R}_0^+, \mathcal{B}_0)$ pontmérték által meghatározott számlálófüggvény alatt azt a $\mu_{(\cdot)} : \mathbb{R}_0^+ \mapsto \bar{\mathbb{N}}_0$ függvényt értjük, amelyre

$$\mu_t = \begin{cases} \mu((0, t]) & , \text{ ha } t > 0 \\ \lim_{t \rightarrow 0^+} \mu_t & , \text{ ha } t = 0 \end{cases} .$$

2.1.5. Állítás. Tegyük fel, hogy a $\mu \in \mathbf{M}_0(\mathbb{R}_0^+, \mathcal{B}_0)$. Ekkor μ számlálófüggvénye rendelkezik az alábbi tulajdonságokkal.

1. $\mu_t \in \bar{\mathbb{N}}_0$.
2. $\mu_0 = 0$.
3. $\mu_{(\cdot)}$ monoton növény, jobbról folytonos, balról határértékkel rendelkező függvény 1 nagyságú ugrásokkal.

Bizonyítás.

1. Korlátos intervallum mértéke véges.
2. Világos, hiszen $\mu \in \mathbf{M}_0$, így μ nem torlódik a nullában, így létezik $s > 0$, hogy $\mu_t = \mu((0, t]) = 0$ minden $0 < t < s$ esetén.
3. $t \geq s$ esetén

$$\mu_t - \mu_s = \mu((0, t]) - \mu((0, s]) = \mu((s, t]) \geq 0,$$

hiszen μ korlátos intervallumokon véges értéket vesz fel, így igaz a szubtraktív tulajdonság. Szintén ez okból kifolyólag μ monoton folytonos is, tehát $\lim_{s \searrow 0} \mu_{t+s} = \lim_{s \searrow 0} \mu((0, t+s]) = \mu((0, t]) = \mu_t$, tehát $\mu(\cdot)$ jobbról folytonos. Az $x > 0$ -beli baloldali határérték vizsgálatához pedig $\lim_{t \searrow 0} \mu_x - \mu_{x-t} = \lim_{t \searrow 0} \mu_x - \mu_{x-t} = \lim_{t \searrow 0} \mu((x-t, x]) = \mu(\{x\}) = \begin{cases} 0 & ,\text{ha } x \notin C \\ 1 & ,\text{ha } x \in C \end{cases}$. Ezzel beláttuk a baloldali határérték létezését, illetve azt is, hogy az ugrások csak 1 nagyságúak lehetnek.

■

Nyilván a 2. és 3. tulajdonság együttesen implikálja az elsőt. Ezzel az iménti lemma úgy foglалható össze, hogy egy $\mu \in \mathbf{M}_0$ pontmérték által meghatározott számlálófüggvény egy olyan nullában eltűnő monoton növekvő függvény, amely jobbról folytonos, balról határértékkel rendelkezik és ugrásai 1 nagyságúak. Egy ilyen tulajdonságú $f : \mathbb{R}_0^+ \mapsto \mathbb{N}_0$ függvényt a továbbiakban *korlátos számlálófüggvénynek* nevezünk. A korlátos számlálófüggvények terére bevezetjük a **BCF** jelölést. Megmutatjuk, hogy minden korlátos számlálófüggvény nem torlódó pontmérték által meghatározott és azt, hogy különböző nem torlódó pontmértékekhez különböző számlálófüggvény tartozik. Jelölje α azt a leképezést, ami egy $\mu \in \mathbf{M}$ pontmértékhez hozzárendeli az általa meghatározott számlálófüggvényt, akkor az előzőek alapján világos, hogy az $\alpha|_{\mathbf{M}_0}$ függvény **BCF**-be képez. Az egyszerűség kedvéért a továbbiakban a megszorított leképezést is α -val jelöljük, amennyiben nem okoz félreértést. Ekkor fennáll az alábbi állítás.

2.1.6. Állítás. *Az $\alpha : \mathbf{M}_0(\mathbb{R}_0^+, \mathcal{B}_0) \mapsto \mathbf{BCF}$ leképezés bijekció.*

Bizonyítás. $\alpha(\mu_1) = \alpha(\mu_2) \Leftrightarrow \mu_1((0, t]) = \mu_2((0, t])$ minden $t > 0$ -ra $\Leftrightarrow \mu_1((s, t]) = \mu_2((s, t])$ minden $0 < s \leq t$ -re. Tekintsük a $\mathcal{G} = \{(s, t] : 0 \leq s \leq t\}$ félgűrűt. Ekkor $\sigma(\mathcal{G}) = \mathcal{B}_0$ és közsímert, hogy félgűrűn értelmezett egészen σ -véges mérték egyértelműen terjed ki a generált σ -algebrára (márpedig μ_1 és μ_2 \mathbf{M}_0 -beli elemek lévén egyaránt egészen σ -végesek), így az ekvivalens átalakítások fenti sora így folytatható:

$$\Leftrightarrow \mu_1|_{\mathcal{G}} = \mu_2|_{\mathcal{G}} \Leftrightarrow \overline{\mu_1|_{\mathcal{G}}} = \overline{\mu_2|_{\mathcal{G}}} \Leftrightarrow \mu_1 = \mu_2.$$

Ezzel beláttuk az injektivitást.

Legyen most f egy korlátos számlálófüggvény és legyen \mathcal{G} az iménti intervallum-félgűrű. Értelmezzük a μ pontmértéket az alábbiak szerint: $0 < s \leq t$ esetén $\mu((s, t]) := f(t) - f(s)$. Legyen $C = \text{dis}(f)$, ahol $\text{dis}(f)$ jelöli az f függvény szakadási helyeinek a halmazát. Nyilván C legfeljebb megszámlálható számosságú halmaz. Ekkor $\mu((s, t]) = \mu_t - \mu_s = \sum_{x \in C \cap (s, t]} 1 = \sum_{x \in C} \delta_x((s, t])$. Vagyis ezzel egy \mathcal{G} -n értelmezett mértéket definiáltunk. C -nek nincs torlódási pontja végesben, ezért μ egészen σ -véges. Mivel félgűrűről egyértelmű a kiterjesztés $\sigma(\mathcal{G}) = \mathcal{B}_0$ -ra - és ez az egyértelmű kiterjesztés éppen a $\mu = \sum_{x \in C} \delta_x$ képlettel meghatározott -, azért ezzel hozzá tudunk rendelni

egy $\mu \in \mathbf{M}_0(\mathbb{R}_0^+, \mathcal{B}_0)$ pontmértéket az f függvényhez. Most már csak az van hátra, hogy megmutassuk, hogy a μ által meghatározott számlálófüggvény éppen f . Ez azonban világos, hiszen $t > 0$ esetén $\mu_t = \mu((0, t]) = \mu(\bigcup_{n=1}^{\infty} (\frac{1}{n}, t]) = \lim_{n \rightarrow \infty} \mu((\frac{1}{n}, t]) = \lim_{n \rightarrow \infty} (f(t) - f(\frac{1}{n})) = f(t) - \lim_{n \rightarrow \infty} f(\frac{1}{n}) = f(t) - f(0) = f(t)$. Innen már $\mu_0 = f(0)$ is világos, amivel bebizonyítottuk az állítást. ■

Ha pedig N egy $\mathbf{M}(\mathbb{R}^+, \mathcal{B}_0)$ -beli véletlen elem, akkor a fenti definícióval összhangban N -hez hozzárendelhető az alábbi sztochasztikus folyamat.

2.1.9. Definíció. Az $N : \Omega \mapsto \mathbf{M}(\mathbb{R}^+, \mathcal{B}_0)$ pontfolyamat által meghatározott számlálófolyamat alatt az $N_t(\omega) := (N(\omega))_t, (t \geq 0)$ képlettel értelmezett sztochasztikus folyamatot értjük.

Világos, hogy így valóban egy sztochasztikus folyamatot értelmeztünk, hiszen 2.1.2 alapján $N_t = N((0, t]) : \Omega \mapsto \bar{\mathbb{N}}_0$ mérhető leképezés. Ha most t -re, mint időparaméterre gondolunk, akkor tehát N_t szemléletesen a t időpontig megfigyelt érkezések véletlen számát fejezi ki, ami jól magyarázza az elnevezést. Az alábbi állítás a számlálófolyamatok tulajdonságait vizsgálja.

2.1.7. Állítás. Tegyük fel, hogy az N pontfolyamat nem torlódó és legyen $N_t, t \geq 0$ az általa meghatározott számlálófolyamat. Ekkor $N_{(\cdot)}$ 1 valószínűséggel az alábbi tulajdonságokkal rendelkezik:

1. $N_0 \in 0$,
2. $N_{(\cdot)} : \Omega \mapsto \mathbb{N}_0$,
3. $N_{(\cdot)}$ trajektóriái monoton növekvő, jobbról folytonos, balról határértékkel rendelkező függvények 1 nagyságú ugrásokkal.
4. $N_t : \Omega \mapsto \mathbb{N}_0$ valószínűségi változó minden $t \geq 0$ -ra.

Bizonyítás. Az 1.-3. pontok bizonyítása világos a determinisztikus esetre vonatkozó korábbi állítás alapján. Csupán az szorul igazolásra, hogy $N_t = N((0, t]) : \Omega \mapsto \mathbb{N}_0$ mérhető leképezés minden $t > 0$ -ra, ez pedig a 2.1.2 fennállása miatt világos. Ebből N_0 mérhetősége is következik, mivel mérhető függvények limesze. ■

2.1.10. Definíció. Az $N_t, t \geq 0$ sztochasztikus folyamatot korlátos számlálófolyamatnak nevezük, ha majdnem biztosan teljesül rá, hogy $N_0 = 0$ és, hogy a trajektóriái monoton növekvő, jobbról folytonos, balról határértékkel rendelkező függvények 1 nagyságú ugrásokkal.

A korlátos számlálófolyamatok terére bevezetjük a **BCP** jelölést. Az előzőek alapján világos, hogy ha az N pontfolyamat nem torlódó, akkor az általa meghatározott számlálófolyamat eleme lesz **BCP**-nek. Jelölje β azt a leképezést, ami egy adott \mathbf{M} -beli N pontfolyamathoz hozzárendeli az általa meghatározott számlálófolyamatot. Ekkor $\beta(N) = \alpha \circ N$

$$\begin{array}{ccc} \mathbf{M}_0(\mathbb{R}^+, \mathcal{B}_0) & \xrightarrow{\alpha} & \mathbf{BCF} \\ \cdot \uparrow N & & \\ \Omega & & \end{array}$$

2.1.8. Állítás. β bijekció a nem torlódó pontfolyamatok és **BCP** között.

2.1.3. Megjegyzés. Itt abban az értelemben beszélünk bijekcióról, hogy két valószínűségi változót azonosnak tekintünk, ha azok 1 valószínűséggel megegyeznek.

Bizonyítás. Legyen N_1 és N_2 nem torlódó. Ekkor $\beta(N_1) = \beta(N_2)$ m.b. $\Leftrightarrow \alpha(N_1) = \alpha(N_2)$ m.b. $\Leftrightarrow \alpha(N_1) = \alpha(N_2)$ és $N_1, N_2 \in \mathbf{M}_0$ m.b. $\Rightarrow N_1 = N_2$ és $N_1, N_2 \in \mathbf{M}_0$ m.b. $\Rightarrow N_1 = N_2$ m.b.

Ha pedig $\xi \in \mathbf{BCP}$, akkor módosítsuk úgy ξ -t, hogy a fennmaradó nullmértékű halma-
zon legyen egyenlő az azonosan nulla függvénnyel. Jelölje a módosított folyamatot ξ' . Ekkor $N(\omega) = \alpha^{-1}(\xi'(\omega))$ olyan nem torlódó pontfolyamatot értelméz, amely által meghatározott számlálófüggvény $\beta(N) = \alpha \circ N = \alpha \circ \alpha^{-1} \circ \xi' = \xi'$, ez pedig m.b. egyenlő ξ -vel. ■

Újabb definíciókat ismertetünk az intenzitásmértékkel kapcsolatban, továbbra is a valós esetben maradvá.

2.1.11. Definíció. Egy pontfolyamat G középértékfüggvénye alatt a $G(t) = \nu((0, t])$ függvényt értjük.

Világos, hogy ekkor $\nu((a, b]) = (G(b) - G(a))\chi(G(a) < \infty)$ és $G(t) = \mathbb{E}N((0, t]) = \mathbb{E}N_t$.

2.1.12. Definíció. A ν mérték Radon-Nikodym deriváltját a Lebesgue-mértékre nézve - amennyiben létezik - intenzitásfüggvénynek, vagy röviden intenzitásnak hívjuk. Ha az intenzitásfüggvény konstans és egyenlő α -val, akkor az α értéket intenzitáshányadnak nevezzük.

Megjegyezzük, hogy az általános esetben is definiálhatunk intenzitásfüggvényt a Lebesgue-mérték helyett egyéb alkalmas domináló mértéket választva.

Ha $G(t)$ abszolút folytonos, akkor a $\mu(t)$ intenzitásfüggvényre $\mu(t) = G'(t)$ - azaz $\mu = \frac{d\nu}{d\Lambda}$, vagy másképpen $G(b) - G(a) = \int_a^b \mu(t)\Lambda(dt)$ - adódik, ahol $\Lambda|\mathcal{B}_0$ jelöli a Lebesgue-mértéket.

Tekintsünk most egy N Poisson folyamatot a pozitív valós számokon. Ha N egy olyan $\nu(B)$ mérték által meghatározott, ami a korlátos intervallumokon véges, akkor

$$P(N(B) < \infty) = \sum_{k=1}^{\infty} \frac{\nu(B)^k}{k!} e^{-\nu(B)} = 1$$

minden B korlátos intervallumra. Legyen $B_i = (i - 1, i]$, $i = 1, 2, \dots$. Ekkor a 2.1.4 állítás alapján azt kapjuk, hogy a ν által meghatározott N Poisson folyamat nem torlódó. Ekkor azonban, mint ahogy azt korábban láttuk, kölcsönösen egyértelműen meghatározzák egymást a számlálófolyamatával. Emiatt ekvivalens módon beszélhetünk a két matematikai objektumról. Ez alapján ebben az esetben a Poisson folyamat számlálófolyamata a következőképpen jellemezhető.

1. $\mathbb{P}(N_0 = 0) = 1$,
2. $N_t - N_s$ Poisson eloszlású $G(t) - G(s) = \int_s^t \mu(x)\Lambda(dx)$ paraméterrel $\forall 0 \leq s \leq t$ -re és
3. N_t független növekményű.

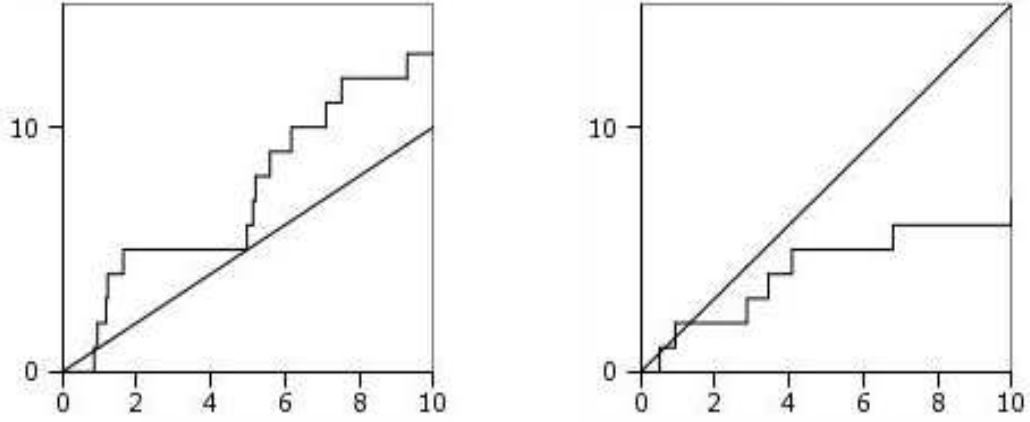
Szokás a Poisson folyamatot a számlálófolyamatával definiálni az imént látott módon. Ha most az intenzitásmérték $\nu(B) = \lambda\Lambda(B)$, ahol $\Lambda|\mathcal{B}$, akkor nyilván fenáll, hogy Λ véges a korlátos intervallumokon. Továbbá $G(t) = \lambda t$, és

$$\mathbb{P}(N_t - N_s = k) = \frac{[\lambda(t - s)]^k}{k!} e^{-\lambda(t-s)}.$$

Ebben az esetben *homogén* Poisson folyamatról beszélünk. A 2.1 ábra egy homogén Poisson számlálófolyamat realizációit illetve a hozzájuk tartozó középértékfüggvényeket mutatja két különböző intenzitáshányad esetén.

Tetszőleges N valós állapotterű pontfolyamat esetén definiálhatók a

$$\tau_i = \inf\{t : N_t \geq i\}, i = 0, 1, \dots$$



2.1. ábra. A $t \mapsto N_t$ trajektóriák egy-egy realizációja $\mu = 1$ és $\mu = 1.5$ esetén

valószínűségi változók, melyet a pontfolyamat i . érkezési idejének nevezünk. Könnyen látható, hogy nem torlódó pontfolyamat esetén τ_i nem más, mint a $t \mapsto N_t$ leképezés i -edik ugrásának "időpontja". Ilyenkor a számlálófolyamat trajektóriáinak jobbról folytonossága miatt inf helyett min is írható. Ekkor $(\tau_i < x) = (N_x \geq i)$, így τ_i eloszlásfüggvénye

$$F_{\tau_i}(x) = \mathbb{P}(\tau_i < x) = \mathbb{P}(N_x \geq i) = \sum_{j=i}^{\infty} \mathbb{P}(N_x = j).$$

2.1.4. Példa. Legyen az N_t Poisson számlálófolyamat a véges értékű G középértékfüggvény által meghatározott, ekkor

$$F_{\tau_i}(x) = \mathbb{P}(\tau_i < x) = \mathbb{P}(N_x \geq i) = \sum_{j=i}^{\infty} \mathbb{P}(N((0, x]) = j) = \sum_{j=i}^{\infty} \frac{G(x)^j}{j!} e^{-G(x)}.$$

Speciálisan $F_{\tau_1}(x) = 1 - \mathbb{P}(N((0, x]) = 0) = 1 - e^{-G(x)}$. Innen látható, hogy amennyiben N homogén Poisson pontfolyamat λ intenzitáshányaddal, úgy τ_1 eloszlása λ paraméterű exponenciális.

2.1.13. Definíció. A $\xi_i = \tau_i - \tau_{i-1}$ valószínűségi változót az i -edik köztes időnek nevezzük ($i = 1, 2, \dots$).

Megmutatható, hogy λ intenzitáshányadú homogén Poisson folyamat esetén a köztes idők páronként függetlenek és exponenciális eloszlásúak λ paraméterrel.

2.2. Duplán sztochasztikus (Cox) folyamatok

Az előző fejezetben megismerkedtünk a Poisson folyamat fogalmával, most pedig megvizsgáljuk ennek egyfajta általánosítási lehetőségét, amikor is a ν intenzitásmérték nem determinisztikus, hanem szintén a véletlentől függő mennyiség. Ehhez szükségünk lesz az úgynevezett *véletlen mérték* fogalmára. Ezen általánosítás eredményeként a pontfolyamatok egy bővebb osztályát kapjuk, az ún. duplán sztochasztikus, vagy más néven Cox folyamatokét. Ennek segítségével olyan modellek építhetők, melyek bizonyos esetekben pontosabban közelítik a valóságot a hagyományos Poisson modelleknél. A duplán sztochasztikus folyamat elnevezést Cox mutatta be először 1955-ös cikkében [3].

Legyen (S, \mathcal{B}) mérhető tér és jelölje $\mathbf{M}_R \equiv \mathbf{M}_R(S, \mathcal{B})$ a \mathcal{B} -n értelmezett egészen σ -véges mértékek terét. Ahogy azt az előző fejezetben tettük, itt is definiálhatóak minden $B \in \mathcal{B}$ -re a

$$\begin{aligned} \pi_B : \mathbf{M}_R &\mapsto [0, \infty) \\ \mu &\mapsto \mu(B) \end{aligned}$$

projekciók. Ezután szintén az előző fejezetben látottakkal analóg módon legyen \mathcal{M}_R a legszűkebb olyan σ -algebra, melyre nézve az összes π_B , $B \in \mathcal{B}$ leképezés $(\mathcal{M}_R, \mathcal{B}([0, \infty)))$ -mérhető, azaz

$$\mathcal{M}_R \equiv \mathcal{M}_R(S, \mathcal{B}) = \sigma\{\pi_B : B \in \mathcal{B}\}.$$

A továbbiakban használni fogjuk a $\mathcal{B}' = \mathcal{B}([0, \infty))$ jelölést is \mathbb{R}_0^+ Borel halmazainak σ -algebrájára.

2.2.1. Definíció. Egy $\rho : \Omega \mapsto \mathbf{M}_R(S, \mathcal{B})$ mérhető leképezést véletlen mértéknek nevezünk.

Most is ugyanúgy definiálhatóak a $\rho(B) = \pi_B \circ \rho$ egydimenziós marginálisok, mely segítségével a 2.1.1 állítással összhangban az alábbi karakterizáció fogalmazható meg:

2.2.1. Állítás. A következő két állítás ekvivalens:

(i) $\rho : \Omega \mapsto \mathbf{M}_R$ egy véletlen mérték

(ii) $\rho(B) : \Omega \mapsto [0, \infty)$ $(\mathcal{A}, \mathcal{B}([0, \infty)))$ -mérhető minden $B \in \mathcal{B}$ -re.

Tekintsük a $0 < \lambda < \infty$ paraméterhez tartozó Poisson eloszlást a nemnegatív egészen. Ez tehát i -hez a $\frac{\lambda^k}{k!}e^{-\lambda}$ valószínűségeket rendeli. Ez kiterjeszthető az $\bar{\mathbb{N}}_0$ halmazra úgy, hogy a ∞ -hez 0 valószínűséget rendelünk. Ezután már értelmezhetjük a $\lambda = 0$ illetve a $\lambda = \infty$ -hez tartozó elfajuló eseteket a $\lim_{\lambda \rightarrow 0} \frac{\lambda^k}{k!}e^{-\lambda}$ és $\lim_{\lambda \rightarrow \infty} \frac{\lambda^k}{k!}e^{-\lambda}$ határértékek segítségével. Ezáltal a 0-ra illetve a ∞ -re koncentrált eloszlásokat kapjuk. Legyen tehát $0 < \lambda < \infty$ esetén $p_k(\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$, ha $k = 1, 2, \dots$ és $p_\infty(\lambda) = 0$. Legyen továbbá

$$p_k(0) = \begin{cases} 1 & , \text{ ha } k = 0 \\ 0 & \text{ egyébként} \end{cases}$$

és

$$p_k(\infty) = \begin{cases} 1 & , \text{ ha } k = \infty \\ 0 & \text{ egyébként} \end{cases}.$$

Ha most $\nu \in \mathbf{M}_R(S, \mathcal{B})$ tetszőleges, akkor bevezethetjük a $P_k^B(\nu) = p_k(\nu(B))$ jelölést. P_k^B tehát egy $\mathbf{M}_R \mapsto \mathbb{R}$ függvény, ami a π_B leképezések mérhetőségéből kifolyólag szintén mérhető. Ezek után meghatározzuk a Cox folyamat fogalmát:

2.2.2. Definíció. Legyen $\rho : \Omega \mapsto M_R(\mathcal{S}, \mathcal{B})$ egy véletlen mérték. Ekkor a $C : \Omega \mapsto M(\mathcal{S}, \mathcal{B})$ pontfolyamatot duplán sztochasztikus vagy más néven Cox folyamatnak nevezzük, ha a diszjunkt B_1, B_2, \dots, B_k halmazokhoz tartozó egydimenziós marginálisai - $C(B_1), C(B_2), \dots, C(B_k)$ - ρ -ra nézve feltételesen függetlenek és $C(B)$ ρ melletti feltételes eloszlása $\nu(B)$ partaméterű Poisson eloszlású, azaz

$$\mathbb{P}(C(B) = k | \rho) = P_k^B(\rho).$$

Ekkor a teljes várható érték tételével felírhatjuk $C(B)$ eloszlását:

$$\mathbb{P}(C(B) = k) = \mathbb{E}(\mathbb{P}(C(B) = k | \rho)) = \mathbb{E}(P_k^B(\rho)) = \int_{M_R} P_k^B(\nu) Q_\rho(d\nu), \quad (2.2)$$

Vagyis $C(B)$ eloszlása keverék Poisson eloszlás a ρ keverő valószínűségi változóval. Innen meghatározhatjuk a C Cox folyamat γ intenzitásmértékét, hiszen amennyiben $\mathbb{E}(C(B)) < \infty$, úgy igaz az alábbi összefüggés:

$$\begin{aligned} \gamma(B) = \mathbb{E}(C(B)) &= \sum_{k=0}^{\infty} k \mathbb{P}(C(B) = k) = \sum_{k=0}^{\infty} k \int_{M_R} P_k^B(\nu) Q_\rho(d\nu) = \\ &= \int_{M_R} \sum_{k=0}^{\infty} k P_k^B(\nu) Q_\rho(d\nu) = \int_{M_R} \nu(B) Q_\rho(d\nu) = \mathbb{E}(\rho(B)) \end{aligned}$$

A Cox folyamat fogalmának létezik egy másik definíciója, amely sokszor célszerűbb tárgyalást tesz lehetővé, ám kevésbé szemléletes, ezért most csak vázlatosan ismertetjük, rámutatva a két definíció közötti kapcsolatra.

2.2.3. Definíció. Legyen adott egy $\zeta : \Omega \mapsto \Theta$ valószínűségi változó és egy N^ϑ pontfolyamat $(\mathcal{S}, \mathcal{B})$ -n minden $\vartheta \in \Theta$ -ra úgy, hogy minden $M \in \mathcal{M}$ esetén a $\vartheta \mapsto Q_{N^\vartheta}(M)$ leképezés mérhető legyen. Azt mondjuk, hogy a

$$Q_N = \int_{\Theta} Q_{N^\vartheta}(\cdot) Q_\zeta(d\vartheta)$$

eloszlással értelmezett $N : \Omega \mapsto M(\mathcal{S}, \mathcal{B})$ pontfolyamat keverék pontfolyamat Q_ζ keverő eloszlással (vagy ζ keverő valószínűségi változóval).

2.2.4. Definíció. Legyen adott egy $\rho : \Omega \mapsto M_R$ véletlen mérték és jelölje N^ν minden $\nu \in M_R$ esetén azt a Poisson folyamatot, aminek intenzitásmértéke éppen ν . Ekkor az N^ν folyamatokból a ρ véletlen mérték, mint keverő valószínűségi változó segítségével képzett $C : \Omega \mapsto M(\mathcal{S}, \mathcal{B})$ keverék Poisson-folyamatot sztenderd paraméterezésű Cox folyamatnak nevezzük.

Ekkor tehát C eloszlása a

$$Q_C = \int_{M_R} Q_{N^\nu}(\cdot) Q_\rho(d\nu)$$

összefüggéssel definiált.

2.2.1. Megjegyzés. Még általánosabb tárgyaláshoz jutunk, ha nem követeljük meg, hogy a keverő valószínűségi változó véletlen mérték legyen, hanem tetszőleges $\zeta : \Omega \mapsto \Theta$ valószínűségi változó szerint keverjük az N^ϑ , $\vartheta \in \Theta$ Poisson folyamatokat, feltéve persze, hogy $\vartheta \mapsto Q_{N^\vartheta}(M)$ mérhető minden $M \in \mathcal{M}$ esetén.

Most megvizsgáljuk, hogy milyen viszonyban állnak egymással a 2.2.4 és a 2.2.2 definíciók. Ehhez definiáljuk a $\rho^* : \Omega \mapsto M_R$ véletlen mértéket úgy, hogy a (C, ρ^*) együttes eloszlása az alábbi legyen:

$$Q_{(C, \rho^*)}(M \times B) = \int_B Q_{N^\nu}(M) Q_\rho(d\nu) \quad (2.3)$$

minden $M \in \mathcal{M}$, $B \in \mathcal{M}_R$ esetén. Ekkor világos, hogy $Q_{\rho^*}(B) = Q_{(C, \rho^*)}(M \times B) = Q_\rho(B)$, vagyis $\rho^* \stackrel{d}{=} \rho$. Bebizonyítjuk a következő lemmát:

2.2.1. Lemma. *A fenti jelölések mellett C feltételes eloszlása a ρ^* valószínűségi változóra nézve éppen $Q_{N^{\rho^*}}(\cdot)$, azaz:*

$$\mathbb{P}(C \in M | \rho^* = \nu) = Q_{N^\nu}(M), \quad M \in \mathcal{M}.$$

Bizonyítás. A feltételes eloszlás definíciója szerint azt kell belátnunk, hogy

$$\mathbb{E}(Q_{N^{\rho^*}}(M) \chi(\rho^* \in B)) = \mathbb{P}(C \in M, \rho^* \in B),$$

$\forall M \in \mathcal{M}$, $B \in \mathcal{M}_R$ esetén. Ez azonban világos, hiszen a baloldal nem más, mint

$$\int_B Q_{N^\nu}(M) Q_{\rho^*}(d\nu),$$

míg a jobboldal 2.3 alapján $\int_B Q_{N^\nu}(M) Q_\rho(d\nu)$, de ez a két integrál egyenlő, hiszen $Q_{\rho^*} = Q_\rho$. ■

Ha most az N^ν paraméterezést megtartva a ρ^* véletlen mérték segítségével definiálok egy C^* Cox folyamatot, akkor nyilván $C^* \stackrel{d}{=} C$, ezért

$$\mathbb{P}(C^* \in M | \rho^* = \nu) = Q_{N^\nu}(M) \quad (2.4)$$

is igaz lesz. Egy ilyen tulajdonságú (C^*, ρ^*) vektort szokás *Cox-pár*nak nevezni. Határozzuk most meg a $C^*(B)$ valószínűségi változó $\rho^* = \nu$ feltétel melletti eloszlását:

$$\begin{aligned} \mathbb{P}(C^*(B) = k | \rho^* = \nu) &= \mathbb{P}(\pi_B \circ C^* = k | \rho^* = \nu) = \mathbb{P}(C^* \in \pi_B^{-1}(\{k\}) | \rho^* = \nu) = \\ &= Q_{N^\nu}(\pi_B^{-1}(\{k\})) = \mathbb{P}(N^\nu \in \pi_B^{-1}(\{k\})) = \mathbb{P}(\pi_B \circ N^\nu = k) = \mathbb{P}(N^\nu(B) = k) = P_k^B(\nu), \end{aligned}$$

vagyis éppen a Cox folyamat 2.2.2 meghatározásában szerepelt definiáló tulajdonságot kaptuk vissza, mely rámutat a 2.2.4 definíció megközelítésének eredetére. A továbbiakban kimondunk majd egy tételt, mellyel a 2.2.2. definícióban szereplő $\mathcal{F}^\rho = \sigma\{\rho\}$ σ -algebrát jellemezhetjük. Ehhez először szükségünk lesz néhány előkészítő lemmára, amelyeket az alábbiakban bizonyítunk.

2.2.5. Definíció. *Ha S nemüres halmaz és $\mathcal{H} \subset \mathcal{P}(S)$, akkor legyen*

$$\mathcal{H}^\delta = \{ \bigcap_{i=1}^\infty A_i : A_i \in \mathcal{H} (i = 1, 2, \dots) \} \text{ és } \mathcal{H}^\sigma = \{ \bigcup_{i=1}^\infty A_i : A_i \in \mathcal{H} (i = 1, 2, \dots) \}.$$

\mathcal{H}^δ és \mathcal{H}^σ tehát rendre a \mathcal{H} halmazrendszert tartalmazó legszűkebb σ -metszetzárt illetve σ -uniózárt halmazrendszer. Legyen $\mathcal{P}_0 = \mathcal{H}$, $\mathcal{Q}_0 = \mathcal{H}^C$. Ha pedig $\alpha > 0$ megszámlálható rendszám és \mathcal{P}_β , illetve \mathcal{Q}_β már definiált minden $\beta < \alpha$ esetén, akkor legyen

$$\mathcal{P}_\alpha = \left(\bigcup_{\beta < \alpha} \mathcal{Q}_\beta \right)^\delta \text{ és } \mathcal{Q}_\alpha = \left(\bigcup_{\beta < \alpha} \mathcal{P}_\beta \right)^\sigma.$$

Ekkor transzfinit indukcióval belátható, hogy $\mathcal{P}_\alpha \subset \mathcal{B}(\mathcal{H})$, $\mathcal{Q}_\alpha \subset \mathcal{B}(\mathcal{H})$ és $\mathcal{P}_\alpha^C = \mathcal{Q}_\alpha$ minden releváns α rendszámra. A definíció közvetlen következménye még, hogy $\beta < \alpha < \omega_1$ esetén $\mathcal{P}_\beta \subset \mathcal{Q}_\alpha$ és $\mathcal{Q}_\beta \subset \mathcal{P}_\alpha$, amiből az is nyilvánvalóan látható, hogy $\bigcup_{\alpha < \omega_1} \mathcal{P}_\alpha = \bigcup_{\alpha < \omega_1} \mathcal{Q}_\alpha$. Jelölje ezt a közös uniót \mathcal{F} . A fent említettek miatt $\mathcal{H} \subset \mathcal{F} \subset \mathcal{B}(\mathcal{H})$. Ekkor \mathcal{F} komplementerzárt, hiszen

$$H \in \mathcal{F} \Leftrightarrow \exists \alpha < \omega_1 : H \in \mathcal{P}_\alpha \Leftrightarrow \exists \alpha < \omega_1 : H^C \in \mathcal{P}_\alpha^C = \mathcal{Q}_\alpha \Leftrightarrow H^C \in \mathcal{F}.$$

Másrészt \mathcal{F} zárt a σ -unió-ra is, hiszen ha $H_i \in \mathcal{F}$ ($i = 1, 2, \dots$), akkor minden i -hez létezik α_i rendszám úgy, hogy $H_i \in \mathcal{P}_{\alpha_i}$. Ha most a γ megszámlálható rendszámot úgy választjuk meg, hogy nagyobb legyen minden α_i -nél (megtehetjük, hiszen például $\gamma = \alpha_1 + \alpha_2 + \dots + 1$ megfelelő), akkor $H_i \in \mathcal{Q}_\gamma \subset \mathcal{F}$. Tehát \mathcal{F} olyan σ -algebra, amely tartalmazza \mathcal{H} -t és része a legszűkebb \mathcal{H} -t tartalmazó σ -algebrának, ezért csak azonos lehet ez utóbbival, azaz $\mathcal{F} \equiv \mathcal{B}(\mathcal{H})$.

2.2.2. Lemma. *Ha $f : \Omega \mapsto X$ és $\mathcal{G} \subset P(X)$ egy nemüres halmazrendszer, akkor*

1. $\{f^{-1}(H) : H \in \mathcal{G}\}^\delta = \{f^{-1}(H) : H \in \mathcal{G}^\delta\}$ és $\{f^{-1}(H) : H \in \mathcal{G}\}^\sigma = \{f^{-1}(H) : H \in \mathcal{G}^\sigma\}$.
2. $\bigcup_{i \in I} \{f^{-1}(H) : H \in \mathcal{G}_i\} = \{f^{-1}(H) : H \in \bigcup_{i \in I} \mathcal{G}_i\}$ tetszőleges I indexhalmazra.

Bizonyítás.

1. Csak az állítás első felét mutatjuk meg, a második ezzel teljesen analóg módon kivitelezhető.
 $\{f^{-1}(H) : H \in \mathcal{G}\}^\delta = \{\bigcap_{i=1}^\infty f^{-1}(H_i) : H_i \in \mathcal{G} \ (i = 1, 2, \dots)\} = \{f^{-1}(\bigcap_{i=1}^\infty H_i) : H_i \in \mathcal{G} \ (i = 1, 2, \dots)\} = \{f^{-1}(H) : H \in \mathcal{G}^\delta\}$.
2. $\bigcup_{i \in I} \{f^{-1}(H) : H \in \mathcal{G}_i\} = \{A : \exists i \in I, \text{ amire } A = f^{-1}(H) : H \in \mathcal{G}_i\} = \{f^{-1}(H) : H \in \bigcup_{i \in I} \mathcal{G}_i\}$.

■

2.2.3. Lemma. *Legyen $f : \Omega \mapsto X$ és $\mathcal{G} \subset P(X)$ egy nemüres halmazrendszer. Ekkor $\sigma\{f^{-1}(H) : H \in \mathcal{G}\} = \{f^{-1}(H) : H \in \sigma(\mathcal{G})\}$*

Bizonyítás. Legyen $\mathcal{H} = \{f^{-1}(H) : H \in \mathcal{G}\}$. Az imént ismertett transzfinit indukciós eljárás segítségével előállítjuk a \mathcal{H} halmazrendszert tartalmazó legszűkebb σ -algebrát. Ehhez elkészítjük a $\mathcal{P}_0 = \mathcal{G}$, $\mathcal{Q}_0 = \mathcal{G}^C$, illetve ezzel párhuzamosan a $\mathcal{P}_0 = \mathcal{H} = \{f^{-1}(H) : H \in \mathcal{P}_0^*\}$ és $\mathcal{Q}_0 = \mathcal{H}^C = \{(f^{-1}(H))^C : H \in \mathcal{G}\} = \{f^{-1}(H^C) : H \in \mathcal{G}\} = \{f^{-1}(H) : H \in \mathcal{G}^C\} = \{f^{-1}(H) : H \in \mathcal{Q}_0^*\}$ halmazokat, majd a fent látottakhoz hasonlóan a \mathcal{P}_α , \mathcal{Q}_α illetve a \mathcal{P}_α és \mathcal{Q}_α halmazokat. Legyen most $\alpha > 0$ megszámlálható rendszám és tegyük fel, hogy minden $\beta < \alpha$ esetén teljesül a

$$\mathcal{P}_\beta = \{f^{-1}(H) : H \in \mathcal{P}_\beta^*\} \text{ és } \mathcal{Q}_\beta = \{f^{-1}(H) : H \in \mathcal{Q}_\beta^*\}$$

azonosság. Ekkor a 2.2.2. lemma alapján

$$\begin{aligned} \mathcal{P}_\alpha &= \left(\bigcup_{\beta < \alpha} \mathcal{Q}_\beta \right)^\delta = \left(\bigcup_{\beta < \alpha} \{f^{-1}(H) : H \in \mathcal{Q}_\beta^*\} \right)^\delta = \\ &= \left(\{f^{-1}(H) : H \in \bigcup_{\beta < \alpha} \mathcal{Q}_\beta^*\} \right)^\delta = \left\{ f^{-1}(H) : H \in \left(\bigcup_{\beta < \alpha} \mathcal{Q}_\beta^* \right)^\delta \right\} = \{f^{-1}(H) : H \in \mathcal{P}_\alpha^*\} \end{aligned}$$

és teljesen hasonló módon $\mathcal{Q}_\alpha = \{f^{-1}(H) : H \in \mathcal{Q}_\alpha^*\}$ is igaz. Innen

$$\begin{aligned} \sigma\{f^{-1}(H) : H \in \mathcal{G}\} &= \bigcup_{\alpha < \omega_1} \mathcal{P}_\alpha = \bigcup_{\alpha < \omega_1} \{f^{-1}(H) : H \in \mathcal{P}_\alpha^*\} = \\ &= \left\{ f^{-1}(H) : H \in \bigcup_{\alpha < \omega_1} \mathcal{P}_\alpha^* \right\} = \{f^{-1}(H) : H \in \sigma(\mathcal{G})\}. \end{aligned}$$

Ezzel bebizonyítottuk ezt a lemmát is. ■

Térjünk vissza most a 2.2.2 definícióra. Amennyiben a \mathcal{G} halmazrendszer a \mathcal{B} egy generátor-rendszere, és rendre \mathcal{F}^ρ , $\mathcal{F}^\mathcal{B}$, és $\mathcal{F}^\mathcal{G}$ jelöli a $\sigma\{\rho\}$, $\sigma\{\rho(B) : B \in \mathcal{B}\}$, illetve $\sigma\{\rho(B) : B \in \mathcal{G}\}$ σ -algebrákat, akkor az alábbi állítás fogalmazható meg.

2.2.1. Tétel. Tegyük fel, hogy a \mathcal{G} generátorrendszer olyan, hogy S előállítható megszámlálható sok \mathcal{G} -beli halmaz uniójaként. A fenti jelölések mellett:

$$\mathcal{F}^\rho \equiv \mathcal{F}^{\mathcal{B}} \equiv \mathcal{F}^{\mathcal{G}}.$$

Bizonyítás. $\mathcal{F}^{\mathcal{G}} = \sigma\{\rho(B) : B \in \mathcal{G}\} = \sigma\{\pi_B \circ \rho : B \in \mathcal{G}\} = \sigma\{\rho^{-1}(\pi_B^{-1}(C)) : B \in \mathcal{G}, C \in \mathcal{B}'\} = \sigma\{\rho^{-1}(H) : H \in \{\pi_B^{-1}(C) : B \in \mathcal{G}, C \in \mathcal{B}'\}\} = \{\rho^{-1}(H) : H \in \sigma\{\pi_B^{-1}(C) : B \in \mathcal{G}, C \in \mathcal{B}'\}\}$ a 2.2.3. lemma miatt. Megmutatjuk, hogy

$$\{\rho^{-1}(H) : H \in \sigma\{\pi_B^{-1}(C) : B \in \mathcal{G}, C \in \mathcal{B}'\}\} = \{\rho^{-1}(H) : H \in \sigma\{\pi_B^{-1}(C) : B \in \sigma(\mathcal{G}), C \in \mathcal{B}'\}\},$$

ehhez elég belátni, hogy $\sigma\{\pi_B^{-1}(C) : B \in \mathcal{G}, C \in \mathcal{B}'\} = \sigma\{\pi_B^{-1}(C) : B \in \sigma(\mathcal{G}), C \in \mathcal{B}'\}$, azaz, hogy $\sigma\{\pi_B : B \in \mathcal{G}\} = \sigma\{\pi_B : B \in \sigma(\mathcal{G})\}$. Az világos, hogy

$$\sigma\{\pi_B : B \in \mathcal{G}\} \subset \sigma\{\pi_B : B \in \sigma(\mathcal{G})\}, \quad (2.5)$$

látni kell még a fordított irányú tartalmazást. Ehhez legyen az $\mathcal{E} = \sigma\{\pi_B : B \in \mathcal{G}\}$ jelölés bevezetésével $\mathcal{F}^* = \{B \in \sigma(\mathcal{G}) : \pi_B \text{ } \mathcal{E}\text{-mérhető}\}$. Ekkor nyilván $\mathcal{G} \subset \mathcal{F}^* \subset \sigma(\mathcal{G})$. Megmutatjuk, hogy \mathcal{F}^* σ -algebra, amiből $\mathcal{F}^* = \sigma(\mathcal{G})$ adódik, azaz $\sigma\{\pi_B : B \in \sigma(\mathcal{G})\} \subset \mathcal{E}$, mely (2.5)-tel együtt adja a kívánt állítást.

1. S előállítható megszámlálható sok \mathcal{G} -beli halmaz uniójaként.
2. \mathcal{F}^* különbségzárt. Legyen $A, B \in \mathcal{F}^*$, $A \subset B$, ekkor $\pi_{B \setminus A}(\nu) = \nu(B \setminus A) = (\nu(B) - \nu(A))\chi(\nu(A) < \infty) = (\pi_B(\nu) - \pi_A(\nu))\chi(\pi_A(\nu) < \infty)$ és mivel mérhető függvények különbsége illetve szorzata is mérhető, ezért $B \setminus A \in \mathcal{F}^*$
3. \mathcal{F}^* zárt a megszámlálható diszjunkt unióra. Legyenek az $A_i \in \mathcal{F}^*$, ($i = 1, 2, \dots$) halmazok páronként diszjunktak, akkor $\pi_{\cup A_i}(\nu) = \nu(\cup A_i) = \sum_i \nu(A_i) = \sum_i \pi_{A_i}(\nu)$ és megszámlálható sok mérhető függvény szintén mérhető, így $\cup_i A_i \in \mathcal{F}^*$.

Tehát a 2.1.1 lemma miatt \mathcal{F}^* valóban σ -algebra. Folytatva a tétel bizonyítását, kimondhatjuk, hogy

$$\begin{aligned} \sigma\{\rho^{-1}(H) : H \in \{\pi_B^{-1}(C) : B \in \mathcal{G}, C \in \mathcal{B}'\}\} = \\ \sigma\{\rho^{-1}(H) : H \in \{\pi_B^{-1}(C) : B \in \sigma(\mathcal{G}), C \in \mathcal{B}'\}\} = \mathcal{F}^{\mathcal{B}}. \end{aligned}$$

Hátra van még az $\mathcal{F}^\rho \equiv \mathcal{F}^{\mathcal{B}}$ azonosság igazolása. $\sigma\{\rho\} = \sigma\{\rho^{-1}(H) : H \in \mathcal{M}_R\} = \sigma\{\rho^{-1}(H) : H \in \{\pi_B^{-1}(C) : B \in \sigma(\mathcal{G}), C \in \mathcal{B}'\}\}$ definíció szerint, erről viszont az imént láttuk, hogy egyenlő $\mathcal{F}^{\mathcal{B}}$ -vel. Ezzel a tételünket beláttuk. ■

Most térjünk rá arra az esetre, amikor $(S, \mathcal{B}) \equiv (\mathbb{R}^+, \mathcal{B}_0)$, a \mathcal{G} generátorrendszer pedig az alulról nyílt, felülről zárt intervallumok félgűrűje. Értelmezzük a $\rho_t, t \geq 0$ sztochasztikus folyamatot a következőképpen:

$$\rho_t = \rho((0, t]),$$

ha $t > 0$ és ρ_0 legyen ρ_{0+} .

2.2.6. Definíció. Az imént bemutatott sztochasztikus folyamatot a C Cox folyamat középérték-folyamatának nevezzük.

Ekkor

$$\sigma(\rho) = \sigma\{\rho(B) : B \in \mathcal{G}\} = \sigma\{\rho((s, t]) : 0 < s \leq t\} = \sigma\{\rho_t - \rho_s : 0 < s \leq t\} = : \mathcal{L}.$$

ρ_0 értelmezéséből adódóan könnyen látható, hogy $\mathcal{L} = \sigma\{\rho_t - \rho_s : 0 \leq s \leq t\}$. Másrészt, ha $\mathcal{L}^* = \sigma\{\rho_t : 0 \leq t\}$, akkor egyrészt $\mathcal{L}^* \subset \mathcal{L}$, másrészt mivel $\rho_t : 0 \leq t \leq \mathcal{L}^*$ -mérhető, azért $\rho_t - \rho_s : 0 \leq s \leq t$ is, így $\mathcal{L} \subset \mathcal{L}^*$ is igaz. Tehát

$$\sigma(\rho) = \sigma\{\rho_t : 0 \leq t\}. \quad (2.6)$$

Szintén természetes kérdésként vetődik fel, hogy mikor lesz egy Cox folyamat nem torlódó. Erre adunk most egy elégséges feltételt. A 2.1.4 állítás alapján elég lenne azt látni, hogy

$$\mathbb{P}(C(B) < \infty) = 1$$

minden B korlátos intervallumra. Fejezzük ki ezért ezt a valószínűséget (2.2) segítségével:

$$\begin{aligned} \mathbb{P}(C(B) < \infty) &= \sum_{k=0}^{\infty} \mathbb{P}(C(B) = k) = \sum_{k=0}^{\infty} \int_{M_R} P_k^B(\nu) Q_\rho(d\nu) = \\ &= \int_{M_R} \sum_{k=0}^{\infty} P_k^B(\nu) Q_\rho(d\nu) = \int_{M_R} \chi(\pi_B(\nu) < \infty) Q_\rho(d\nu) = \\ &= \int_{\{\pi_B < \infty\}} 1 Q_\rho = Q_\rho(\pi_B < \infty) = \mathbb{P}(\rho(B) < \infty). \end{aligned}$$

Ha tehát ρ olyan, hogy 1 valószínűséggel véges minden korlátos intervallumra, akkor a C Cox folyamat nem torlódó. Ebben az esetben tehát az előző alfejezetben látottak miatt ekvivalens módon beszélhetünk magáról a folyamatról és annak számlálófolyamatáról. Ekkor (2.6)-ot figyelembe véve a C Cox folyamatot definiáló tulajdonságok a $C_t, t \geq 0$ számlálófolyamatra az alábbi formában fogalmazhatóak meg.

1. C növekményei feltételesen függetlenek a $\sigma\{\rho_t : 0 \leq t\}$ σ -algebra mellett.
2. $\mathbb{P}(C_t - C_s = k | \rho_u, 0 \leq u) = \frac{(\rho_t - \rho_s)^k}{k!} e^{-(\rho_t - \rho_s)}$ minden $0 < s \leq t$ esetén.

Ha pedig ρ m.b. monoton folytonos, akkor még

$$\mathbb{P}(C_t = k | \rho_u, 0 \leq u) = \frac{\rho_t^k}{k!} e^{-\rho_t}$$

is igaz, természetesen az egyenlőséget majdnem biztosan értve.

Egy valós állapotterű Cox folyamat G középértékfüggvényére az intenzitásmértékre levezetett korábbi képlet alapján

$$G(t) = \gamma((0, t]) = \int_{M_R} \nu((0, t]) Q_\rho(d\nu) = \int_{M_R} G^\nu(t) Q_\rho(d\nu)$$

adódik valahányszor $\mathbb{E}(\rho_t) < \infty$, ahol $G^\nu(t)$ jelöli a ν intenzitásmértékű Poisson folyamat középértékfüggvényét, ρ_t pedig az imént bemutatott középértékfolyamatot.

2.2.7. Definíció. A C intenzitás-folyamata alatt az alábbi $\eta_t, t > 0$ sztochasztikus folyamatot értjük

$$\eta_t = \begin{cases} \frac{d\rho}{d\Lambda}(t) & , \text{ ha } \rho \ll \Lambda \\ 0 & \text{egyébként} \end{cases}.$$

2.2.2. Megjegyzés. A szokásos $\eta_0 := \eta_{0+}$ értelmezéssel az intenzitásfolyamat is kiterjeszhető \mathbb{R}_0^+ -ra.

Határozzuk most meg a C intenzitásfüggvényét. Tegyük fel, hogy a ρ véletlen mérték majdnem biztosan abszolút folytonos a Lebesgue-mértékre, azaz, hogy $\mathbb{P}(\rho \ll \Lambda) = 1$. Amennyiben a ν egészen σ -véges mérték abszolút folytonos a Λ Lebesgue mértékre, úgy jelöljük μ^ν -vel a $\frac{d\nu}{d\Lambda}$ Radon-Nykodim deriváltat, ellenkező esetben μ^ν legyen azonosan 0. Ha a $(\nu, t) \mapsto \mu^\nu(t)$ leképezés mérhető, akkor Fubini tételének értelmében fennáll a következő

$$\begin{aligned} \gamma(B) &= \int_{M_R} \nu(B) Q_\rho(d\nu) = \int_{\{\nu: \nu \ll \Lambda\}} \nu(B) Q_\rho(d\nu) = \\ & \int_{\{\nu: \nu \ll \Lambda\}} \int_B \mu^\nu(t) \Lambda(dt) Q_\rho(d\nu) = \int_{M_R} \int_B \mu^\nu(t) \Lambda(dt) Q_\rho(d\nu) = \\ & \int_B \int_{M_R} \mu^\nu(t) Q_\rho(d\nu) \Lambda(dt), \end{aligned}$$

vagyis a μ intenzitásfüggvény a

$$\mu(t) = \int_{M_R} \mu^\nu(t) Q_\rho(d\nu)$$

képlettel adható meg. Cox folyamatra példaként szolgál majd a második fejezet 6. alfejezetében bemutatásra kerülő Markov-modulált Poisson folyamat.

2.3. Néhány szó mátrix értékű függvények differenciálásáról és a mátrix exponenciális függvényről

Ebben az alfejezetben néhány alapvető fogalmat és azonosságot ismertetünk mátrix értékű függvények differenciálásával kapcsolatban, majd ezután mátrixokra vonatkozó lineáris differenciálegyenleteket vizsgálunk, melyek megoldásának explicit alakban történő felírásához értelmezzük az e^A kifejezést, ahol a kitevőben egy A négyzetes mátrix áll. Ennek kapcsán vizsgáljuk majd az ún. mátrix exponenciális függvényt és ismertetünk egy numerikus eljárást is, amely segítségével hatékonyan számolható $\exp(A)$. Ezek alapvető elemeit képezik majd a későbbi tárgyalásunknak.

2.3.1. Definíció. Az $A : \mathbb{R} \mapsto \mathbb{R}^{m \times n}$ mátrix értékű függvény deriváltja alatt az $A' : \mathbb{R} \mapsto \mathbb{R}^{m \times n}$ függvényt értjük, ahol $(A')_{ij} = A'_{ij}$.

2.3.1. Lemma. Legyen adott az $A : \mathbb{R} \mapsto \mathbb{R}^{m \times n}$ és $B : \mathbb{R} \mapsto \mathbb{R}^{n \times p}$ mátrix értékű függvény, ekkor $(AB)' = A'B + AB'$.

Bizonyítás.

$$\begin{aligned} [(AB)']_{ij} &= (AB)'_{ij} = \left(\sum_{p=1}^n A_{ip} B_{pj} \right)' = \sum_{p=1}^n (A_{ip} B_{pj})' \\ &= [(AB)']_{ij} = (AB)'_{ij} = \left(\sum_{p=1}^n A_{ip} B_{pj} \right)' = \sum_{p=1}^n (A_{ip} B_{pj})' \\ &= \sum_{p=1}^n A'_{ip} B_{pj} + \sum_{p=1}^n A_{ip} B'_{pj} = \sum_{p=1}^n (A')_{ip} B_{pj} + \sum_{p=1}^n A_{ip} (B')_{pj} \\ &= (A'B)_{ij} + (AB')_{ij} = (A'B + AB')_{ij}. \end{aligned}$$

■

2.3.1. Következmény. Ha az $A : \mathbb{R} \mapsto \mathbb{R}^{m \times n}$ és a $C : \mathbb{R} \mapsto \mathbb{R}^{p \times q}$ mátrix értékű függvény konstans, akkor - tekintve, hogy $A' = 0$ és $C' = 0$ - a $B : \mathbb{R} \mapsto \mathbb{R}^{n \times p}$ mátrix értékű függvény deriváltjára $(AB)' = AB'$ és $(BC)' = B'C$.

2.3.2. Lemma. Ha $A : \mathbb{R} \mapsto \mathbb{R}^{r \times r}$ mátrix értékű függvény, akkor $n \geq 2$ esetén

$$(A^n)' = \sum_{j=1}^{n-1} A^j A' A^{n-1-j}.$$

Bizonyítás. A 2.3.1 lemma alapján teljes indukcióval világos. ■

2.3.3. Lemma. Az $A : \mathbb{R} \mapsto \mathbb{R}^{r \times r}$ mátrix értékű függvény inverzének deriváltjára

$$(A^{-1})' = -A^{-1} A' A^{-1}.$$

Bizonyítás. $A^{-1}(t)A(t) = I$ minden $t \in \mathbb{R}$ esetén, ebből a 2.3.1 lemma alapján $(A^{-1})' A + A^{-1} A' = 0$, amiből átrendezés, majd A^{-1} -el való szorzás után adódik az állítás. ■

2.3.2. Definíció. Legyen A $r \times r$ -es valós mátrix, ekkor A exponenciálisa alatt az

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

hatványsorral értelmezett $r \times r$ -es mátrixot értjük.

Az így értelmezett hatványsor mindig konvergens – a konvergenciát tetszőleges normára nézve definiálhatjuk, mivel véges dimenziós vektortérben minden norma ekvivalens –, fennáll ugyanis az alábbi lemma.

2.3.4. Lemma. Ha A_0, A_1, \dots $m \times n$ -es mátrixok, $\|\cdot\|$ tetszőleges mátrix norma és a $\sum_{k=0}^{\infty} \|A_k\|$ numerikus sor konvergens, akkor $\sum_{k=0}^{\infty} A_k$ is az.

Bizonyítás. A Cauchy konvergenciakritérium és a háromszög-egyenlőtlenség miatt

$$\left\| \sum_{k=m+1}^n A_k \right\| \leq \sum_{k=m+1}^n \|A_k\| < \varepsilon,$$

ha $m, n \geq N(\varepsilon)$. ■

Innen már a $\sum_{k=0}^{\infty} \frac{A^k}{k!}$ hatványsor konvergenciája világos a $\sum_{k=0}^{\infty} \frac{\|A\|^k}{k!}$ sor konvergenciájából. Könnyen ellenőrizhetően igazak az alábbi lemmában felsorolt tulajdonságok.

2.3.5. Lemma. Az $A \mapsto \exp(A)$ leképezés rendelkezik az alábbi tulajdonságokkal.

1. A csupa nulla mátrixra $\exp(0) = I$.
2. $\exp(A)$ nonszinguláris, és inverze $\exp(-A)$.
3. $\exp(A) = \left[\exp\left(\frac{A}{m}\right) \right]^m$.
4. $\exp(A+B) = \exp(A)\exp(B)$ akkor és csak akkor, ha A és B felcserélhetők.

2.3.3. Definíció. Legyen adott az $A \in \mathbb{R}^{r \times r}$ mátrix, ekkor az A által meghatározott mátrix exponenciális függvény alatt a $t \mapsto \exp(At)$ összefüggéssel értelmezett mátrix értékű függvényt értjük.

2.3.6. Lemma. A mátrix exponenciális függvény rendelkezik az alábbi tulajdonságokkal.

1. $\exp[A(s+t)] = \exp(As)\exp(At)$.
2. $t \mapsto \exp(At)$ inverze $t \mapsto \exp(-At)$.
3. $\exp[(A+B)t] = \exp(At)\exp(Bt)$ akkor és csak akkor, ha A és B felcserélhetők.
4. $\frac{\partial}{\partial t} \exp(At) = A \exp(At) = \exp(At)A$.

Bizonyítás. Az 1-3. állítások közvetlen következményei a (2.3.5) lemmának, míg a 4. állítás szintén egyszerűen adódik a $\exp(At) = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$ hatványsor tagonkénti deriválásából. ■

A 2.3.6 lemma 4. tulajdonsága más szóval azt mondja, hogy ha A egy $r \times r$ -es mátrix és $X(t) = \exp(At)$, akkor az X mátrix értékű függvény kielégíti az $X(0) = I$ kezdeti feltétel melletti

$$X'(t) = AX(t)$$

lineáris differenciál-egyenletet. A közönséges differenciál-egyenletek elméletéből jól ismert, hogy ennek az egyenletnek az említett kezdeti feltétel mellett ez az egyetlen megoldása létezik.

A későbbiek során szükségünk lesz a következő lemmára, ami egy bizonyos speciális alakú mátrixhoz tartozó exponenciális függvény kiszámításában nyújt segítséget.

2.3.7. Lemma. Legyenek adottak az $r \times r$ -es A és B mátrixok, majd definiáljuk a $2r \times 2r$ -es C mátrixot a következő alakban:

$$C = \begin{pmatrix} A & B \\ 0 & A \end{pmatrix}.$$

Ekkor

$$\exp(Ct) = \begin{pmatrix} F(t) & G(t) \\ 0 & F(t) \end{pmatrix}, \quad (2.7)$$

ahol $F(t) = \exp(At)$ és $G(t) = \int_0^t \exp(A(t-s))B \exp(As)ds$.

Bizonyítás. A mátrix exponenciális függvény definíciójából világos, hogy $\exp(Ct)$ (2.7) alakú. Másrészt $\exp(Ct)$ az $X'(t) = CX(t)$ differenciál-egyenlet egyértelmű megoldása az $X(0) = I$ kezdeti feltétel mellett. Ezt felírva

$$\begin{pmatrix} F'(t) & G'(t) \\ 0 & F'(t) \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & A \end{pmatrix} \begin{pmatrix} F(t) & G(t) \\ 0 & F(t) \end{pmatrix} = \begin{pmatrix} AF(t) & AG(t) + BF(t) \\ 0 & AF(t) \end{pmatrix},$$

vagyis $F(t)$ és $G(t)$ az $F(0) = I$ és $G(0) = 0$ kezdeti feltételek mellett

$$F'(t) = AF(t)$$

$$G'(t) = AG(t) + BF(t)$$

differenciál-egyenletrendszer egyértelmű megoldásaként adódik. Ebből $F(t) = \exp(At)$ nyilvánvaló, $G(t) = \int_0^t \exp(A(t-s))B \exp(As)ds$ esetén pedig

$$\begin{aligned} G'(t) &= A \exp(At) \int_0^t \exp(-As)B \exp(As)ds + \exp(At) \exp(-At)B \exp(At) \\ &= AG(t) + BF(t). \end{aligned}$$

ezzel bebizonyítottuk az állítást. ■

A következő segédtevélyünk ugyanezt a lemmát általánosítja, bár egy későbbi alkalmazásra való tekintettel kissé más megközelítésben mondjuk ki. Előtte vezessük be a következő jelöléseket. Legyen adott egy $r \times r$ -es A mátrix, és az $\mathcal{I} = \{1, 2, \dots, r\}$ indexhalmaz I illetve J részhalmazai. Ekkor $A_{I,J}$ -vel fogjuk jelölni azt az $|I| \times |J|$ -es részmátrixot, amit az eredeti mátrixból kapunk a megadott indexhalmazok szerinti sorok és oszlopok kiválasztásával.

2.3.8. Lemma. *Legyen A és B két $r \times r$ -es mátrix és tegyük fel, hogy az $F_n(t)$ mátrix értékű függvényekből álló sorozat kielégíti a következő rekurziót:*

$$F_0(0) = I \text{ és } F_n(0) = 0, \text{ ha } n \geq 1,$$

$$F'_0(t) = F_0(t)A,$$

$$F'_n(t) = F_n(t)A + F_{n-1}(t)B, \text{ ha } n \geq 1.$$

Ekkor fenáll az alábbi két állítás.

1. $F_n(t)$ explicit módon felírható az alábbi formában:

$$F_0(t) = \exp(At)$$

$$F_n(t) = \int_0^t F_{n-1}(s)B \exp(A(t-s))ds, \text{ ha } n \geq 1.$$

2. Tekintsük az alábbi $(n+1)r \times (n+1)r$ -es mátrixot:

$$C = \begin{pmatrix} A & B & 0 & \cdots & 0 \\ 0 & A & B & \cdots & 0 \\ 0 & 0 & A & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & B \\ 0 & 0 & 0 & 0 & A \end{pmatrix}$$

Legyen $D(t) = \exp(Ct)$ és $0 \leq k \leq n$ esetén definiáljuk az $I_k = \{kr+1, kr+2, \dots, kr+r\}$ indexhalmazokat. Ekkor

$$F_k(t) = [D(t)]_{I_0, I_k}.$$

Bizonyítás.

1. $F_0(t) = \exp(At)$ nyilvánvaló, ha pedig $n \geq 1$, akkor tegyük fel, hogy F_{n-1} -et már meghatároztuk, ekkor az $F'_n(t) = F_n(t)A + F_{n-1}(t)B$ egyenletnek egyértelműen létezik megoldása az $F_n(0) = 0$ kezdeti feltétel mellett, mégpedig $F_n(t) = \int_0^t F_{n-1}(s)B \exp(A(t-s))ds$, hiszen erre a függvényre

$$\begin{aligned} F'_n(t) &= \int_0^t F_{n-1}(s)B \exp(-As)ds \exp(At)A + F_{n-1}(t)B \exp(-At) \exp(At) \\ &= F_n(t)A + F_{n-1}(t)B. \end{aligned}$$

2. Legyen n rögzített. Ekkor az $F_0(t), F_1(t), \dots, F_n(t)$ értékek egyértelműen meghatározhatók az első $n+1$ rekurziós egyenletből. Vezessük be az

$$X(t) = \begin{pmatrix} F_0(t) & F_1(t) & F_2(t) & \cdots & F_n(t) \\ 0 & F_0(t) & F_1(t) & \cdots & F_{n-1}(t) \\ 0 & 0 & F_0(t) & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & F_1(t) \\ 0 & 0 & 0 & 0 & F_0(t) \end{pmatrix}$$

új ismeretlent. Ekkor könnyen láthatóan a rekurziós egyenletek éppen azt mondják ki, hogy

$$X'(t) = X(t)C$$

és hogy $X(0) = I$. Ekkor tehát, $X(t) = \exp(Ct)$ és tekintve, hogy $F_k(t) = [X(t)]_{I_0, I_k}$ az állítás bebizonyított.

■

2.3.1. Megjegyzés. A 2.3.7. lemmának más irányú általánosítása is létezik, lásd például [29] 1. Tételét.

A mátrix exponenciális kiszámítására rengeteg numerikus módszer létezik, a témakörrel alapos áttekintést nyújt a [30] alatti dolgozat. Ezek közül mutatunk most itt be egyet, amit használni fogunk egy későbbi algoritmusunk során.

2.3.4. Definíció. Legyen A adott $r \times r$ -es négyzetes mátrix, ekkor $\exp(A)$ (p, q) -Padé approximációja alatt az

$$R_{pq}(A) = [D_{pq}(A)]^{-1} N_{pq}(A)$$

mennyiséget értjük, ahol

$$N_{pq}(A) = \sum_{j=0}^p \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!} A^j$$

és

$$D_{pq}(A) = \sum_{j=0}^p \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!} (-A)^j.$$

Van Loan [30]-ban megemlíti, hogy $D_{pq}(A)$ regularitása biztosított, amennyiben p és q bizonyos értelemben elég nagyok vagy A sajátértékei negatívak. $\exp(A)$ Padé approximációval történő kiszámítása tehát lényegében egy megfelelő p és q érték meghatározásából, majd $R_{pq}(A)$ kiszámításából áll. [30] részletesen kifejti, hogy milyen módon érdemes választani p -t és q -t a különböző pontossági szintek eléréséhez. A MatlabTM-ban implementált "expm" eljárás a (6, 6) párhoz tartozó Padé approximációt használ. Először azonban átskálázza a mátrixot úgy, hogy ∞ normája $\frac{1}{2}$ -nél kisebb legyen, majd a 2.3.5 lemma 3. tulajdonságát használva visszaállítja az eredeti skálázást.

2.3.1. Algoritmus (Padé-approximáció a mátrix exponenciális kiszámítására).

1. *Bemenet:* A négyzetes mátrix.
2. Válaszd meg az $\frac{1}{2} \leq f < 1$ és e számokat úgy, hogy $\|A\|_\infty = f2^e$ legyen. Ekkor $s = 0 \vee (e+1)$ esetén $\frac{\|A\|_\infty}{2^s} < \frac{1}{2}$.
3. $E := R_{(6,6)}(A)$
4. $i = 1, 2, \dots, s$ esetén $E := E^2$.
5. *Kimenet:* E approximáció $\exp(A)$ -ra.

2.4. Véges állapotterű, folytonos idejű Markov folyamatok

2.4.1. Definíció. Azt mondjuk, hogy a ξ_t folyamat folytonos idejű és véges állapotterű Markov folyamat az $\mathcal{F}_t, t \geq 0$ filtrációra nézve, ha $t \geq 0$ esetén ξ_t az $E = \{1, 2, \dots, r\}$ értékek valamelyikét veszi fel és $i \in E$ illetve $t \geq s$ esetén

$$\mathbb{P}(\xi_t = i | \mathcal{F}_s) = \mathbb{P}(\xi_t = i | \xi_s).$$

A mi esetünkben az $\mathcal{F}_t, t \geq 0$ filtráció a ξ_t folyamat által generált filtráció lesz majd. Feltesszük, hogy az átmenet-valószínűségek stacionáriusak, azaz, hogy $P_{ij}(t) = \mathbb{P}(\xi_{t+s} = j | \xi_s = i) = P_{ij}(t) = \mathbb{P}(\xi_t = j | \xi_0 = i), \forall t, s \geq 0$ -ra.

2.4.1. Lemma. A fenti feltételek mellett

1. $P_{ij}(t) \geq 0$,
2. $\sum_{j=1}^r P_{ij}(t) = 1, i, j \in E$
3. $P_{ij}(s+t) = \sum_{u=1}^r P_{iu}(s)P_{uj}(t), t, s \geq 0$ -ra (Chapman-Kolmogorov összefüggés).

Bizonyítás. Az első két pont triviális, a harmadik a teljes valószínűség tételéből következik. ■

A továbbiakban feltesszük még, hogy $P_{ij}(t)$ a nulla pontban jobbról differenciálható és $P_{ij}(0+) = \delta_{ij}$. Legyen $P = (P_{ij})$. Ekkor a Chapman-Kolmogorov egyenletek eképpen írhatók fel mátrix alakban:

$$P(t+s) = P(t)P(s).$$

A deriválhatóság miatt létezik olyan Q mátrix, amire $P(t) = \exp(Qt)$.

2.4.2. Definíció. A Q mátrixot hívjuk a ξ_t Markov folyamat infinitezimális generátorának.

Mivel $P(t)$ sorösszege 1, ezért Q sorösszege 0. Emiatt a modell $r(r-1)$ független paraméterrel rendelkezik. A Q mátrix elemei a főátlóbelieket kivéve mind nemnegatívak, tehát a főátlóban lévő elemek nempozitívak. Ezekre szokás bevezetni a $q_i = -q_{ii}$ jelölést. Ezek segítségével a Markov folyamat fejlődése a következőképpen írható le. Ha a folyamat éppen az i -edik állapotban van, akkor ott töltött ideje q_i paraméterű exponenciális eloszlású, majd innen a $\frac{q_{ij}}{q_i}$ valószínűségek szerint vált át valamelyik másik $j \neq i$ állapotba.

2.4.3. Definíció. A ξ_t Markov folyamat irreducibilis, ha tetszőleges állapotból el lehet jutni bármely másikkba, azaz ha tetszőleges $i, j \in E$ esetén létezik $t \geq 0$, amire $P_{ij}(t) > 0$.

Definiáljuk egy adott i állapoton történő első áthaladás idejét a

$$T_i = \inf\{t > 0 : \xi_t = i, \xi_{t-} \neq i\}$$

összefüggéssel. Legyen továbbá $F_{ij} = \mathbb{P}(T_j < \infty | \xi_0 = i)$.

2.4.4. Definíció. Azt mondjuk, hogy az i állapot pozitív ismétlődő állapot, ha $F_{ii} < \infty$ és $\mathbb{E}(T_i | \xi_0 = i) < \infty$.

2.4.5. Definíció. A ξ_t Markov folyamat ergodikus, ha minden állapot pozitív ismétlődő.

2.4.1. Tétel. Ha a $\xi_t, t \geq 0$ Markov folyamat irreducibilis és legalább az egyik állapot pozitív ismétlődő, akkor minden állapot pozitív ismétlődő és $F_{ij} = 1$ minden $i, j \in E$ -re.

2.4.6. Definíció. Az r dimenziós $\pi = (\pi_1, \dots, \pi_r)$ sorvektort valószínűségi vektornak hívjuk, ha $\pi_i \geq 0$ minden i -re és $\sum_{i=1}^r \pi_i = 1$.

2.4.7. Definíció. A $p_i(t) = \mathbb{P}(\xi_t = i)$ függvényeket ($i = 1, 2, \dots, r$) állapot-valószínűség függvényeknek hívjuk.

Legyen $p(t) = (p_1(t), \dots, p_r(t))$. Ekkor nyilván $p(t)$ minden t -re valószínűségi vektor és a teljes valószínűség tétele szerint

$$p_i(t) = \mathbb{P}(\xi_t = i) = \sum_{u=1}^r P_{ui}(t)p_u = p(0) [P(t)]^{(i)},$$

vagyis $p(t) = p(0)P(t)$, amiből $p'(t) = p(0)P'(t) = p(0)P(t)Q = p(t)Q$, azaz $p(t)$ időbeli fejlődése a

$$p'(t) = p(t)Q$$

differenciál-egyenlettel írható le.

2.4.8. Definíció. A p r dimenziós valószínűségi vektor a $\xi_t, t \geq 0$ Markov folyamat stacionárius eloszlása, ha $\mathbb{P}(\xi_0 = i) = p_i$ maga után vonja, hogy $\mathbb{P}(\xi_t = i) = p_i$ minden $t \geq 0$ -re.

Ha létezik p stacionárius eloszlás, akkor arra nyilván $p = pP(t) \forall t$ -re, vagy ezzel ekvivalensen

$$pQ = 0. \tag{2.8}$$

Megfordítva, ha a p valószínűségi vektor kielégíti a 2.8 egyenlőséget, akkor p a $\xi_t, t \geq 0$ Markov folyamat stacionárius eloszlása. Fennáll a következő tétel.

2.4.2. Tétel. Legyen a $\xi_t, t \geq 0$ Markov folyamat irreducibilis és tegyük fel, hogy véges idő alatt 1 valószínűséggel csak véges sok állapotváltás történik. Ekkor ξ ergodikus akkor és csak akkor, ha létezik egy olyan p valószínűségi vektor, hogy $p_i > 0$ minden $i \in E$ -re és p megoldása a

$$\begin{aligned} pQ &= 0 \\ p\mathbf{1} &= 1 \end{aligned}$$

lineáris egyenletrendszernek. Ekkor p egyértelmű és

$$p_j = \lim_{t \rightarrow \infty} P_{ij}(t)$$

minden $i, j \in E$ esetén.

A későbbi vizsgálódásaink során felmerülő véges állapotterű Markov folyamatokról mindig feltezzük majd, hogy irreducibilisek, stacionárius átmenetvalószínűségeik és azt, hogy egyértelműen létezik stacionárius eloszlás. Végezetül megemlítjük, hogy könnyen ellenőrizhetően a 2×2 -es esetben a stacionárius eloszlás:

$$(p_1, p_2) = \left(\frac{q_2}{q_1 + q_2}, \frac{q_1}{q_1 + q_2} \right).$$

2.5. Kvázi-Markov folyamatok és Markov felújítási sorozatok

Szükségünk lesz a sztochasztikus folyamatok két újabb osztályának ismertetésére, melyekből később speciális esetként adódnak majd a számunkra érdekes Markov-modulált Poisson folyamatra vonatkozó következtetések. E két osztály pedig nevezetesen a kvázi-Markov folyamatok (SMP, az angol "semi-Markov process" elnevezésből) és a Markov felújítási sorozatok (MRS, az angol "Markov renewal sequence" elnevezésből) osztálya, habár tárgyalásunkból következni fog, hogy ez a két osztály tulajdonképpen egy és ugyanaz. Történeti kialakulását tekintve az SMP fogalma született meg előbb, Lévy[12], Smith[13], illetve Takács[27] egymástól függetlenül definiálták 1954-ben.

Egy SMP egy olyan sztochasztikus folyamat, ami egy adott időpillanatban az előre rögzített, legfeljebb megszámlálható sok állapot valamelyikében van oly módon, hogy az egymás után meglátogatott állapotok Markov-láncot alkotnak, és egy adott állapotban a folyamat által eltöltött idő egy olyan valószínűségi változó, melynek eloszlása függhet az adott és a legközelebb meglátogatásra kerülő állapottól egyaránt. Ez szemléletesen egy olyan Markov láncként fogható fel, aminek az időparamétere véletlenszerűen transzformált. Mi a továbbiakban csak olyan SMP-kel foglalkozunk majd, melyek csak véges sok állapottal rendelkeznek.

Egy MRS pedig egy olyan valószínűségi változókból álló (X_n, Y_n) sorozat, ahol X_n egy SMP i . állapotváltásának idejét rögzíti, Y_n pedig az $(i - 1)$. és az i . állapotváltás között eltelt időt. A szakirodalomban szokás még használni a Markov felújítási folyamat elnevezést, mi itt a *sorozat* szót használjuk, hangsúlyozva ezzel a folyamat diszkrét voltát. Az alfejezet további részeiben szabatosan is definiáljuk az előbbi szemléletes megfogalmazásokat, majd ismertetjük a terület néhány alapvető eredményét.

2.5.1. Definíció. Egy $F : \mathbb{R} \mapsto \mathbb{R}$ monoton növekvő, jobbról folytonos függvényt súlyfüggvénynek hívunk, ha $F(\infty) \leq 1$.

2.5.2. Definíció. Egy $Q : \mathbb{R} \mapsto \mathbb{R}^{r \times r}$ mátrix értékű függvényt átmeneteloszlás-mátrixnak hívunk, ha a Q_{ij} elemek olyan súlyfüggvények, melyek kielégítik az alábbi feltételeket:

- (i) $Q_{ij}(t) = 0$, ha $t \leq 0$ és
- (ii) $Q_{i \cdot}(t) = \sum_{j=1}^r Q_{ij}(\infty) = 1$, $i = 1, \dots, r$.

2.5.3. Definíció. Azt mondjuk, hogy a kétdimenziós (X_n, Y_n) , $n \geq 0$ sztochasztikus folyamat a π kiindulási valószínűségi vektor és a Q átmeneteloszlás-mátrix által meghatározott Markov felújítási sorozat, ha

1. $Y_0 = 0$,
2. $\mathbb{P}(X_0 = i) = \pi_i$, és
3. $\mathbb{P}(X_n = i, Y_n \leq y | X_0, X_1, Y_1, X_2, Y_2, \dots, X_{n-1}, Y_{n-1}) = \mathbb{P}(X_n = i, Y_n \leq y | X_{n-1}) = Q_{X_{n-1}, i}(y)$

minden $y \in \mathbb{R}$ és $i = 1, \dots, r$ esetén.

2.5.1. Megjegyzés. Az így definiált MRS-re röviden a (π, Q) pár segítségével hivatkozunk majd. Itt és a továbbiakban valószínűségi változókra vonatkozó egyenlőségeken mindig majdnem biztosan fennálló egyenlőséget értünk.

Erre a modellre úgy tekinthetünk, mint ami egy olyan rendszer időbeli fejlődését írja le, ami minden t időpillanatban az $E = \{1, \dots, r\}$ állapotok valamelyikében van és két állapot között véletlenszerűen (Y_n szerint) időzik. Ez a véletlenszerűen eltöltött idő pedig az aktuális és az éppen látogatásra kerülő állapottól függ. Hogy jobban megértsük a modellt, a következő példákban rámutatunk, hogy miként illeszthető be ebbe a tárgyalásba néhány igen jól ismert sztochasztikus folyamat.

2.5.1. Példa. Ha $r = 1$, akkor (X, Y) egy felújítási folyamatot határoz meg, ha az n . felújítás ideje $\sum_{i=0}^n Y_i$.

2.5.2. Példa. Ha $Y_n = 1$ minden $n \geq 1$ -re, akkor (X, Y) egy diszkrét idejű Markov láncot határoz meg.

2.5.3. Példa. Ha Y_n exponenciális eloszlású olyan paraméterrel, ami X_{n-1} -től függ, akkor egy folytonos idejű Markov láncot kapunk.

Vezessük be a $T_n = \sum_{i=0}^n Y_i$ jelölést, ekkor T_n az n -edik állapotváltás idejeként interpretálható.

2.5.2. Megjegyzés. Megjegyezzük, hogy itt az "állapotváltás" szót használtuk, ami azonban nem feltétlenül jelent a szó szoros értelmében vett állapotváltást, tekintve, hogy egyelőre nem zártuk ki a $Q_{ii} > 0$ lehetőséget.

2.5.1. Lemma. (X_n, T_n) , $n \geq 0$ kielégíti az alábbi egyenlőségeket:

$$\mathbb{P}(X_n = i, T_n \leq t | X_0, X_1, T_1, X_2, T_2, \dots, X_{n-1}, T_{n-1}) = Q_{X_{n-1}, i}(t - T_{n-1}) \quad (2.9)$$

és

$$\mathbb{P}(X_n = i | X_0, X_1, T_1, X_2, T_2, \dots, X_{n-1}, T_{n-1}) = Q_{X_{n-1}, i}(\infty). \quad (2.10)$$

Bizonyítás. (2.9) könnyen látható a

$$\mathbb{P}(X_n = i, Y_n \leq t - T_{n-1} | X_0, X_1, T_1, X_2, T_2, \dots, X_{n-1}, T_{n-1})$$

átrendezett alakból és abból a tényből, hogy a σ -algebra, amelyre a feltételt vesszük megegyezik a 2.5.3 definícióban szereplő σ -algebrával. (2.10) pedig a Lebesgue monoton konvergencia tétel egyszerű következménye. ■

Vezessünk még be néhány jelölést. Legyen $\tilde{Q} = Q(\infty)$. A 2.5.2 definíció alapján világos, hogy \tilde{Q} sztochasztikus mátrix. Jelölje $U_c(x)$ a c pontra koncentrált valószínűségi változó eloszlásfüggvényét. Defináljuk az F mátrixértékű függvényt az alábbi képlettel:

$$F_{ij} := \begin{cases} \frac{Q_{ij}}{\tilde{Q}_{ij}} & , \text{ ha } \tilde{Q}_{ij} > 0 \\ U_1 & , \text{ ha } \tilde{Q}_{ij} = 0 \end{cases}.$$

Legyen továbbá $F = (F_{ij})$. Ezek segítségével könnyen beláthatók az alábbi állítások.

2.5.2. Lemma. A fenti jelölések mellett fennállnak az alábbi összefüggések

1. $\mathbb{P}(Y_n \leq y | X_0, X_1, \dots, X_{n-1}) = Q_{X_{n-1}, \cdot}(y)$
2. $\mathbb{P}(Y_n \leq y | X_0, X_1, \dots, X_n) = F_{X_{n-1}, X_n}(y)$
3. $\mathbb{P}(X_n = i | X_0, X_1, \dots, X_{n-1}) = \tilde{Q}_{X_{n-1}, i}$

$$4. \mathbb{P}(Y_{n_1} \leq y_1, Y_{n_2} \leq y_2, \dots, Y_{n_k} \leq y_k | X_n, n \geq 0) \\ = \mathbb{P}(Y_{n_1} \leq y_1, Y_{n_2} \leq y_2, \dots, Y_{n_k} \leq y_k | X_{n_1}, X_{n_k}, \dots, X_{n_k}) = \prod_{i=1}^k F_{X_{n_{i-1}}, X_{n_i}}(y_i).$$

Vegyük észre, hogy a 3. állítás éppen azt jelenti, hogy X_0, X_1, \dots Markov lánc, melynek átmenet-valószínűség mátrixa \tilde{Q} .

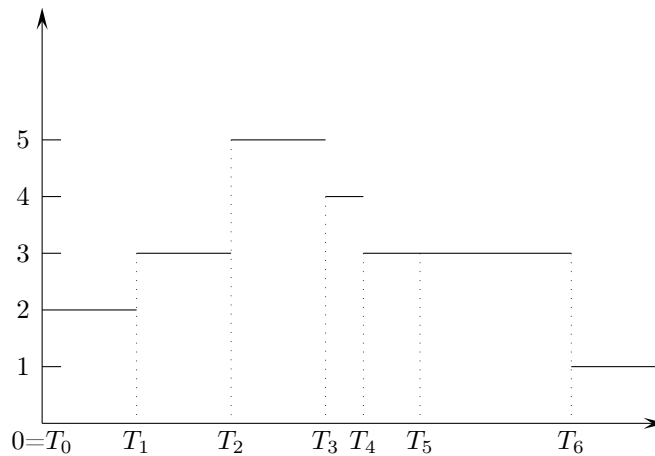
Értelmezzük az $N_t, t \geq 0$ egész értékű sztochasztikus folyamatot az $N_t = \sup\{n \geq 0 : T_n \leq t\}$ képlettel és. Ezenkívül legyen $i = 1, \dots, r$ esetén $N_t^i = \#\{k : X_k = i, k = 1, \dots, N_t\}$, vagyis N_t^i azt adja meg, hogy a $(0, t]$ intervallumban N_t hányszor került az i állapotba. Világos, hogy minden t -re $N_t = \sum_{i=1}^r N_t^i$.

2.5.4. Definíció. Az $N_t, t \geq 0$ sztochasztikus folyamatot az MRP számlálófolyamatának, míg az $N_t^i, t \geq 0$ sztochasztikus folyamatokat az MRP-hez tartozó i . típusú felújítási folyamatnak nevezzük.

Könnyen ellenőrizhető, hogy az elnevezés jogos, vagyis N_t^i valóban felújítási folyamat. Az iménti definíciók figyelembe vételével bevezethetünk egy másik sztochasztikus folyamatot, amely minden időpillanatban az MRP aktuális állapotának az értékét veszi fel.

2.5.5. Definíció. Azt mondjuk, hogy a $Z_t, t \geq 0$ folyamat a π k.v.v. és a Q átmeneteloszlás-mátrix által meghatározott kvázi-Markov folyamat, ha $Z_t = X_{N_t}$.

A 2.2 ábrán látható egy kvázi-Markov folyamat lehetséges realizációja.



2.2. ábra. Egy SMP lehetséges realizációja

Bizonyítás nélkül közöljük az alábbi tételeket, a részletek megtalálhatók R. Pyke [18] illetve [19] cikkében.

2.5.1. Tétel. $\mathbb{P}(N_t < \infty, \forall t \geq 0) = 1$.

Legyen most az MRS véges állapotterű, ekkor kimondható az alábbi tétel.

2.5.2. Tétel. Definiáljuk minden $t \geq 0$ és $n \geq 0$ esetén a

$$P_{ij}(n, t) = \mathbb{P}(N_t = n, Z_t = j | Z_0 = i)$$

mennyiségeket. Ekkor

$$P_{ij}(0, t) = \delta_{ij} U_0(t)(1 - Q_i(t))$$

és $n = 1, 2, \dots$ esetén

$$P_{ij}(n, t) = \sum_{k=1}^r P_{kj}(n-1, t) * Q_{ik}(t).$$

2.5.3. Megjegyzés. Itt $*$ konvolúciót jelöl, azaz

$$(f * g)(t) = \begin{cases} \int_{0-}^t f(t-y)dg(y) & , \text{ ha } t \geq 0 \\ 0 & , \text{ ha } t < 0 \end{cases}.$$

Ezután a konvolúció művelete mátrixértékű függvényekre is definiálható:

$$K * L = \sum_{k=1}^r K_{ik} * L_{kj},$$

mely formálisan éppen ugyanúgy van definiálva, mint a hagyományos mátrixszorzás azzal a különbséggel, hogy itt a szorzás operátor konvolúcióra van cserélve. Legyen továbbá $K^{(*0)} = I$ és $n = 1, 2, \dots$ esetén definiáljuk a $K^{(*n)}$ mátrixértékű függvényt a K mátrixértékű függvény önmagával vett n -szeres konvolúciójaként, azaz legyen $K^{(*n)} = K^{*(n-1)} * K$.

A 2.5.2 tétel így fogalmazható meg tömören mátrix jelölésekkel:

$$P(0, t) = \mathcal{I} - \mathcal{H}(t) \tag{2.11}$$

és

$$P(n, \cdot) = \sum_{n=0}^{\infty} Q^{(*n)} * (\mathcal{I} - \mathcal{H}), \tag{2.12}$$

ahol $\mathcal{I}_{ij} = \delta_{ij}U_0(t)$ és $\mathcal{H}_{ij}(t) = \delta_{ij}Q_i(t)$.

2.6. A Markov-modulált Poisson folyamat, mint speciális Cox folyamat

A Markov-modulált Poisson folyamat, vagy röviden MMPP (az angol Markov-modulated Poisson Process elnevezés alapján) egy olyan általános Poisson folyamatként képzelhető el, amelyben a $\mu(t)$ intenzitásfüggvény nem determinisztikus, hanem maga is egy sztochasztikus folyamat, mégpedig egy véges állapotterű folytonos idejű Markov-folyamat által meghatározott az alábbiak szerint: $\mu(t) = \lambda_i$, ha a Markov-folyamat éppen az i -edik állapotban van.

Hogy precízzé tegyük definíciónkat, legyen ξ_t egy véges állapotterű folytonos idejű Markov folyamat, melynek állapottere az egyszerűség kedvéért legyen $E = \{1, 2, \dots, r\}$. Tekintsük továbbá azt a $\lambda : E \mapsto \mathbb{R}^+$ függvényt, amelyre $\lambda(i) = \lambda_i$ minden $i \in E$ esetén az előre rögzített $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{R}^+$ számokkal. Az MMPP-t egy olyan Cox folyamatként fogjuk definiálni, melynek η intenzitás-folyamatára $\eta_t = \lambda(\xi_t)$, $t > 0$. Legyen tehát $\mathcal{G} = \{(s, t] : 0 < s \leq t\}$, majd definiáljuk a ρ véletlen mértéket a következőképpen: és $\rho((s, t]) := \int_{s+}^t \eta_t dt$. Ezzel egy egészen σ -véges mértéket értelmezünk, mely egyértelműen terjed ki a generált σ -algebrára, vagyis \mathcal{B}_0 -ra. Ha most a kiterjesztett mértéket továbbra is ρ jelöli, úgy ezzel megadtunk egy \mathcal{B}_0 -n értelmezett egészen σ -véges mértéket. Ezzel már definiálható a Markov-modulált Poisson folyamat fogalma.

2.6.1. Definíció. *Legyen a ρ véletlen mérték az előző bekezdésben leírtak szerint definiált, ekkor az általa meghatározott Cox folyamatot Markov-modulált Poisson folyamatnak (MMPP) hívjuk.*

Az előző alfejezetekre visszatekintve az adódik, hogy egy MMPP paramétereizhető a mögöttes Markov-folyamat Q infinitezimális generátormátrixával, illetve a $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_r]^T$ vektorral. Ez összesen $r(r-1) + r = r^2$ paramétert jelent. Később ezen paraméterek becslésére mutatunk majd be egy eljárást. Szokás még bevezetni a $\Lambda = \text{diag}(\lambda)$ jelölést, mellyel a paraméterter az alábbi formában írható le:

$$\Theta = \{(Q, \Lambda) : Q \text{ irreducibilis és } \Lambda = \text{diag}(\lambda), \text{ ahol } \lambda_i > 0 \forall i\}.$$

Igazából a mögöttes Markov folyamat kezdeti eloszlása is paraméter, ezt azonban nem fogjuk becsülni, hanem Q függvényeként választjuk majd meg. A Q és Λ paramétereket illetően fenáll az alábbi lemma.

2.6.1. Lemma. *Amennyiben Q irreducibilis, úgy a $\Lambda = 0$ triviális esetet leszámítva $Q - \Lambda$ nonszinguláris.*

Ahhoz, hogy egy MMPP-t teljesen leírjunk, szükséges, hogy megadjuk a mögöttes Markov folyamat $t = 0$ pillanatbeli állapotát. Jelöljük ezt π -vel. Az alábbiakban ismertetjük a (Q, Λ, π) paraméterekkel rendelkező MMPP fejlődését, ahol tehát π és Q rendre a mögöttes Markov folyamat kezdeti eloszlásának vektora illetve infinitezimális generátora. Ehhez előbb belátjuk az alábbi lemmát.

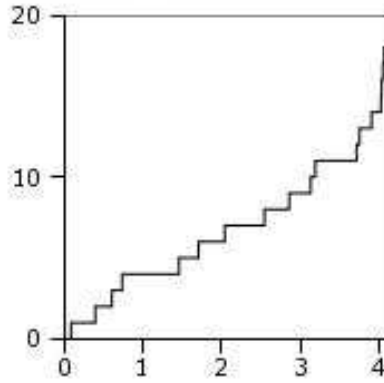
2.6.2. Lemma. *Legyenek ξ_1 és ξ_2 független exponenciális eloszlású valószínűségi változók λ_1 és λ_2 paraméterekkel. Legyen továbbá*

$$U = \min(\xi_1, \xi_2)$$

és

$$N = \begin{cases} 1 & , \text{ ha } \xi_1 < \xi_2 \\ 2 & , \text{ ha } \xi_1 \geq \xi_2 \end{cases}.$$

Ekkor fenáll a következő három állítás.



2.3. ábra. Egy MMPP lehetséges realizációja

1. $\mathbb{P}(N = 1) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$ és $\mathbb{P}(N = 2) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$.
2. U exponenciális eloszlású $\lambda_1 + \lambda_2$ paraméterrel.
3. N és U függetlenek.

Bizonyítás.

1. Mivel ξ_1 és ξ_2 függetlenek, azért $\xi_1 - \xi_2$ eloszlásfüggvénye felírható ξ_1 és $-\xi_2$ eloszlásfüggvényeinek konvolúciójaként:

$$\begin{aligned}
 \mathbb{P}(\xi_1 < \xi_2) &= F_{\xi_1 - \xi_2}(0) = F_{\xi_1} * F_{-\xi_2}(0) = \int_{-\infty}^{\infty} F_{\xi_1}(-y) dF_{-\xi_2}(y) \\
 &= \int_{-\infty}^0 F_{\xi_1}(-y) dF_{-\xi_2}(y) = \int_{-\infty}^0 (1 - e^{-\lambda_1 y}) d(e^{\lambda_2 y}) \\
 &= \int_{-\infty}^0 (1 - e^{-\lambda_1 y}) \lambda_2 e^{\lambda_2 y} dy = 1 - \lambda_2 \frac{1}{\lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_1 + \lambda_2}.
 \end{aligned}$$

2. Ismét ξ_1 és ξ_2 függetlenségéből adódóan $\mathbb{P}(U \geq x) = \mathbb{P}(\xi_1 \geq x, \xi_2 \geq x) = \mathbb{P}(\xi_1 \geq x)\mathbb{P}(\xi_2 \geq x) = e^{-\lambda_1 x} e^{-\lambda_2 x} = e^{-(\lambda_1 + \lambda_2)x}$, ha $x \geq 0$.
3. Az $f_Y(y) = f_{h(Y)}(h(y))|\mathcal{J}(h(y))|$ összefüggés alkalmazásával — ahol $|\mathcal{J}(h(\cdot))|$ jelöli a h függvény Jacobi-determinánsát — és ξ_1 illetve ξ_2 függetlenségének kihasználásával felírható, hogy

$$f_{(\xi_1 - \xi_2, \xi_2)}(s, t) = f_{(\xi_1, \xi_2)}(s + t, t) = f_{\xi_1}(s + t) f_{\xi_2}(t).$$

Emiatt

$$\begin{aligned}
\mathbb{P}(N = 2, U \geq x) &= \mathbb{P}(\xi_1 - \xi_2 \geq 0, \xi_2 \geq x) = \int_x^\infty \int_0^\infty f_{(\xi_1 - \xi_2, \xi_2)}(s, t) ds dt \\
&= \int_x^\infty \int_0^\infty f_{\xi_1}(s+t) f_{\xi_2}(t) ds dt = \int_x^\infty f_{\xi_2}(t) \int_0^\infty f_{\xi_1}(s+t) ds dt \\
&= \int_x^\infty f_{\xi_2}(t) \int_t^\infty f_{\xi_1}(u) du dt = \int_x^\infty f_{\xi_2}(t) e^{-\lambda_1 t} dt = \lambda_2 \int_x^\infty e^{-(\lambda_1 + \lambda_2)t} dt \\
&= \frac{\lambda_2}{\lambda_1 + \lambda_2} e^{-(\lambda_1 + \lambda_2)x} = \mathbb{P}(N = 2) \mathbb{P}(U \geq x),
\end{aligned}$$

ami bizonyítja a függetlenséget.

■

Ezek után egyaránt nevezzük eseménynek azt, ha ξ_t állapotot vált t -ben vagy ha t érkezési ideje az MMPP-nek. Ekkor az MMPP fejlődése az alábbiak szerint jellemezhető. Ha s -ben bekövetkezik egy esemény, akkor feltéve, hogy $\xi_{s+} = i$, a következő eseményig eltelt idő exponenciális eloszlású $q_i + \lambda_i$ paraméterrel és az eltelt időtől függetlenül az esemény $\frac{q_{ij}}{q_j + \lambda_i}$ valószínűséggel állapotváltás vagy $\frac{\lambda_i}{q_j + \lambda_i}$ valószínűséggel ugrás. Ez azt jelenti, hogy (N_t, ξ_t) egy megszámlálható állapotterű Markov folyamat, ezért a

$$P_{ij}(n, t) = \mathbb{P}(N_t = n, \xi_t = j | \xi_0 = i) \quad (2.13)$$

átmenet-valószínűségekre $n \geq 1$ esetén felírhatók az alábbi előre haladó Chapman-Kolmogorov egyenletek:

$$P'_{ij}(n, t) = \sum_{l \neq j} P_{il}(n, t) q_{lj} + P_{ij}(n, t) (q_j - \lambda_j) + P_{ij}(n-1, t) \lambda_j,$$

ami mátrix jelöléssel így írható tömören:

$$P'(n, t) = P(n, t)(Q - \Lambda) + P(n-1, t)\Lambda,$$

ahol nyilván $P(0, 0) = I$ és $P(n, 0) = I$, ha $n \geq 1$. Hasonlóan $n = 0$ -ra

$$P'(n, t) = P(n, t)(Q - \Lambda)$$

adódik. Ekkor a $P^*(z, t) = \sum_{n=0}^{\infty} P(n, t) z^n$ generátorfüggvény kielégíti a

$$\frac{\partial}{\partial t} P^*(z, t) = P^*(z, t)(Q - \Lambda) + z P^*(z, t) \Lambda$$

egyenlőséget, továbbá $P^*(z, 0) = I$, amiből már következik, hogy $P^*(z, t) = \exp[(Q - \Lambda + z\Lambda)t]$. Ennek segítségével meghatározzuk N_t várható értékét minden rögzített t -re. Ha az $M(t)$ mátrix az

$$M(t) = \left[\frac{\partial}{\partial z} P^*(z, t) \right]_{z=1}$$

képlettel definiált, akkor világos, hogy

$$M_{ij}(t) = \sum_{n=0}^{\infty} n P_{ij}(n, t),$$

ami azt jelenti, hogy

$$\begin{aligned}
E(N_t) &= \sum_{n=0}^{\infty} n P(N_t = n) = \sum_{n=0}^{\infty} \sum_{j=1}^r n P(N_t = n, \xi_t = j) \\
&= \sum_{n=0}^{\infty} \sum_{j=1}^r \sum_{i=1}^r n P_{ij}(n, t) \mathbb{P}(\xi_0 = i) \\
&= \sum_{j=1}^r \sum_{i=1}^r M_{ij}(t) \pi_i = \pi M(t) \mathbf{1}.
\end{aligned}$$

A mátrix differenciálás szabályait figyelembe véve

$$M(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n-1} Q^p \Lambda Q^{n-1-p}.$$

Kis számolással az is belátható, hogy ekkor

$$M(t)\mathbf{1} = \mathbf{1}\pi\lambda t + (\exp(Qt) - I)(Q + \mathbf{1}\pi)^{-1}\lambda,$$

amiből azt kapjuk, hogy

$$E(N_t) = \pi\mathbf{1}\pi\lambda t + \pi(\exp(Qt) - I)(Q + \mathbf{1}\pi)^{-1}\lambda.$$

Ha most π speciálisan a ξ_t Markov folyamat stacionárius eloszlása, akkor $\pi \exp(Qt) = \pi$ és így

$$E(N_t) = \pi\mathbf{1}\pi\lambda t = \pi\lambda t.$$

2.6.1. Megjegyzés. A $P(n, t)$ -re levezetett összefüggések kiszámolásához felhasználhattuk volna a 2.11 és 2.12 alatti összefüggéseket is.

Tegyük fel most, hogy T rögzített és hogy a T időpontig megfigyelt legutolsó állapotváltás a $s < T$ időpontban történik, akkor az exponenciális eloszlás örökifjú tulajdonsága miatt a T utáni legelső esemény $q_i + \lambda_i$ paraméterű exponenciális eloszlás szerint következik be, feltéve, hogy ξ_t T -ben az i . állapotban volt. Ez azt jelenti, hogy az $N_t^* = N_{T-t} - N_T$ összefüggéssel egy MMPP-t értelmezünk, mely esetén $\xi_t^* = \xi_{T+t}$ a mögöttes Markov folyamat a $\pi_i^* = \mathbb{P}(\xi_0^* = i) = \mathbb{P}(\xi_T = i)$ kezdeti valószínűségekkel. Tehát míg N_t^* csupán a π^* kezdeti eloszlásvektorban különbözik N_t -től. Ha π a Q stacionárius eloszlása, akkor még $\pi^* = \pi$ is igaz. Ez alapján az 2.13 alatti mennyiségek segítségével meghatározhatjuk a $\mathbb{P}(N_{T+t} - N_T = n)$, $n \geq 0$ valószínűségeket.

$$\begin{aligned} \mathbb{P}(N_{T+t} - N_T = n) &= \mathbb{P}(N_t^* = n) = \sum_{j=1}^r \mathbb{P}(N_t^* = n, \xi_t^* = j) \\ &= \sum_{i=1}^r \sum_{j=1}^r \mathbb{P}(N_t^* = n, \xi_t^* = j | \xi_0^* = i) \mathbb{P}(\xi_0^* = i) \\ &= \sum_{i=1}^r \sum_{j=1}^r P_{ij}(n, t) \mathbb{P}(\xi_T = i) = \pi^* P(n, t) \mathbf{1}. \end{aligned}$$

A Markov-folyamatokról szóló fejezetben láttuk, hogy $\pi^* = \pi \exp(Qt)$, tehát

$$\mathbb{P}(N_{T+t} - N_T = n) = \pi \exp(Qt) P(n, t) \mathbf{1},$$

ha pedig π a mögöttes Markov folyamat stacionárius eloszlása, akkor

$$\mathbb{P}(N_{T+t} - N_T = n) = \pi P(n, t) \mathbf{1}.$$

A következőkben bemutatjuk, hogy miként gondolhatunk az MMPP-re úgy, mint egy MRS-re. Ennek érdekében legyen $X_n, n \geq 0$ a mögöttes Markov folyamat állapota az n . megfigyeléskor (X_0 legyen a $t = 0$ időpontbeli állapot). Legyen továbbá $Y_0 = 0$ és $n \geq 1$ esetén jelölje Y_n az n . és az $(n - 1)$. érkezés között eltelt időt. Ekkor $(X_n, Y_n), n \geq 0$ sorozat, ennek megokolása megtalálható például [6]-ben. Fenáll továbbá az alábbi tétel, melynek bizonyítása megtalálható [17]-ben ezért annak közlését most mellőzzük.

2.6.1. Tétel. Az imént definiált $(X_n, Y_n), n \geq 0$ Markov felújítási sorozat átmeneteloszlás-mátrixa:

$$\begin{aligned} Q(x) &= \int_0^x \exp[(Q - \Lambda)u] du \Lambda \\ &= \{I - \exp[(Q - \Lambda)x]\} (\Lambda - Q)^{-1} \Lambda \\ &= \{I - \exp[(Q - \Lambda)x]\} Q(\infty). \end{aligned}$$

2.6.2. Megjegyzés. Ha az időhorizont kezdete nem felel meg egy érkezési időpontnak, akkor a $\mathbb{P}(\xi_1 = j, Y_1 \leq y | \xi_0 = i)$ valószínűségeket külön definiálni kell.

A $\tilde{Q} = Q(\infty) = (\Lambda - Q)^{-1}\Lambda$ mátrix sztochasztikus, ahogy azt láttuk az 5. alfejezetben. Ráadásul \tilde{Q} azon Markov lánc átmenet-valószínűség mátrixa, ami rögzíti a folyamat állapotát az i . érkezéskor (beágyazott Markov lánc). Ha pedig minden λ_i pozitív - ahogy azt mi végig fel is tesszük az általunk vizsgált MMPP-k esetén - akkor \tilde{Q} egyben irreducibilis is. Ha π jelöli a Q mátrixhoz tartozó stacionárius eloszlást, akkor könnyű látni, hogy a

$$\tilde{\pi} = \frac{\pi \Lambda}{\pi \lambda}$$

valószínűségi vektor kielégíti a $\tilde{\pi} = \tilde{\pi} \tilde{Q}$ lineáris egyenletrendszer.

Térjünk rá a mögöttes Markov folyamat kiindulási vektorának megválasztásának kérdésére. Az MMPP *intervallum-stacionárius* verzióját kapjuk, ha a kiindulási valószínűségi vektort $\tilde{\pi}$ -nak választjuk. Az ún. *környezet-stacionárius* verzióhoz jutunk, ha a kiindulási valószínűségi vektort a Q mátrix π stacionárius vektorának választjuk. Ez a Markov felújítási sorozatként történő leírásban azt jelenti, hogy a $\mathbb{P}(X_0 = i)$ valószínűségeket π_i -re állítjuk és aztán a $\mathbb{P}(\xi_1 = j, Y_1 \leq y | \xi_0 = i)$ valószínűségeket úgy definiáljuk, hogy a belőlük alkotott mátrix értékű függvény az y helyen éppen $Q(y)$ -nal legyen egyenlő. Ekkor az időhorizont kezdete ugyan nem egy érkezési időpont, a mögöttes ξ Markov folyamat azonban stacionárius lesz. Mi ez utóbbi verziót fogjuk majd a későbbiekben használni.

Végezetül kimondunk egy tételt, mely a későbbiekben fontos szerepet játszik majd. A bizonyítást illetően lásd [31] vagy [17].

2.6.2. Tétel. Tegyük fel, hogy adott egy (Y_1, \dots, Y_n) minta egy $\vartheta = (Q, \Lambda)$ paraméterű MMPP-ből származó köztes időkből, ekkor (Y_1, \dots, Y_n) együttes sűrűségfüggvénye

$$f_{(Y_1, \dots, Y_n)}^{\vartheta}(y_1, \dots, y_n) = \pi^{\vartheta} \left(\prod_{k=1}^n f^{\vartheta}(y_k) \right) \mathbf{1},$$

ahol

$$f^{\vartheta}(t) = \exp[(Q - \Lambda)t] \Lambda. \quad (2.14)$$

3. fejezet

Az MMPP paramétereinek becslése az EM algoritmus segítségével

A fejezet első alfejezete egy rövid összefoglaló az EM algoritmusról, mint paraméterbecslési eljárásról. A 2. alfejezet a Rydén [23] által prezentált EM algoritmust hivatott bemutatni, míg a 3. alfejezet az algoritmus implementálásával foglalkozik, figyelembe véve a W.J.J. Roberts et. al. által közölt új eredményeket. Az alfejezet saját eredménye a kiinduló becslés meghatározására szolgáló k-means klaszterezést használó eljárás.

3.1. Az EM algoritmus

Az úgynevezett EM (expectation maximization) algoritmus egy széles körben használható iteratív paraméterbecslési eljárás olyan feladatok megoldására, ahol a minta egy része nem megfigyelhető. Ez a szituáció lép fel például akkor, ha (X, Y) egy véletlen vektor az f_ϑ együttes sűrűségfüggvénnyel és a $\vartheta \in \theta$ paramétert szeretnénk becsülni, viszont csak az Y valószínűségi változót tudjuk megfigyelni. Az alábbiakban megmutatjuk, hogy hogyan kezelhetők az ilyen esetek az EM algoritmus segítségével.

3.1.1. Definíció. Legyen $(\Omega, \mathcal{A}, \mathbb{P}_\vartheta)$, $(\vartheta \in \theta)$ statisztikai mező, $(\mathbf{X}, \mathcal{B})$ mérhető tér és $X : \Omega \mapsto \mathbf{X}$ valószínűségi változó. Tegyük fel, hogy a statisztikai mező dominált, azaz létezik $\Lambda | \mathcal{B}$ mérték, hogy $\mathbb{P}_\vartheta \circ X^{-1} \ll \Lambda$ minden $\vartheta \in \theta$ -ra. Legyen $x \in \mathbf{X}$ egy megfigyelés az X valószínűségi változóból. Ekkor az x megfigyelés loglikelihood-függvénye alatt a $\vartheta \mapsto \log f_\vartheta(x)$ függvényt értjük, ahol $f_\vartheta = \frac{d\Lambda}{d(\mathbb{P}_\vartheta \circ X^{-1})}$.

3.1.2. Definíció. Legyen $x \in \mathbf{X}$ egy megfigyelés az X valószínűségi változóból. Az előző definíció jelölései mellett azt mondjuk, hogy $\hat{\vartheta}$ a $\vartheta \in \theta$ paraméter maximum-likelihood (ML) becslése, ha maximalizálja a $\vartheta \mapsto \log f_\vartheta(x)$ loglikelihood-függvényt.

Ha most (X, Y) egy véletlen vektor és $\vartheta \in \theta$ a becsülendő paraméter, de csak Y -t tudjuk megfigyelni, akkor nem tudjuk kiszámolni az ML becslést, mivel a nem megfigyelhető adat nem áll rendelkezésünkre. Ehelyett keresünk egy olyan függvényt, ami bizonyos értelemben jól közelíti $\log f_\vartheta(X, Y)$ -t, de csak Y -től függ. Ehhez tekintsük az alábbi lemmát.

3.1.1. Lemma. Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ egy valószínűségi mező és tekintsük az $L^2(\Omega)$ -beli valós értékű X illetve Y valószínűségi változókat. Definiáljuk az \mathcal{F} függvényhalmazt az alábbi módon.

$$\mathcal{F} = \{f : \mathbb{R} \mapsto \mathbb{R} \text{ mérhető függvény, } \mathbb{E}f(Y)^2 < \infty\}$$

Ekkor

$$\min_{f \in \mathcal{F}} \|X - f(Y)\|_2 = \|X - \mathbb{E}(X|Y)\|_2.$$

- ahol $\|\cdot\|_2$ \mathcal{L}^2 -normát jelöl, - és a minimum a $f(y) = \mathbb{E}(X|Y = y)$ függvény esetén vétetik fel.

Ha tehát abból indulunk ki, hogy már van egy θ_0 feltételezésem arról, hogy mi lehet a valódi paraméter, akkor a $\log f_{\vartheta}(X, Y)$ függvény legjobb közelítése a fenti lemma szerinti értelemben $l(\vartheta_0, \vartheta, y) = \mathbb{E}_{\vartheta_0}(\log f_{\vartheta}(X, Y)|Y = y)$, ami már csak a y megfigyeléstől függ, így ennek már meg tudjuk határozni a ϑ -ban vett maximumát. Ha ez a maximum a ϑ_1 -ben vétetik fel, akkor azt reméljük, ϑ_1 már közelebb áll az ML becsléshez, mint ϑ_0 és ϑ_1 -re megismételhetjük az iménti eljárást. Tehát az algoritmus az alábbi módon implementálható, az egyszerűség kedvéért megállási kritérium nélkül.

3.1.1. Algoritmus (EM algoritmus). 1. bemenő adatok: ϑ_0 kiindulási becslés, y megfigyelés a Y valószínűségi változóból.

2. n . iteráció:

- a) Határozd meg az $l(\vartheta_{n-1}, \vartheta, y) = \mathbb{E}_{\vartheta_{n-1}}(\log f_{\vartheta}(X, Y)|Y = y)$ függvényt! (ún. E-lépés)
- b) $\vartheta_n := \arg \max_{\vartheta} l(\vartheta_{n-1}, \vartheta, y)$. (ún. M-lépés).

3. kimenő adat: $\{\vartheta_n\}$ végtelen sorozat.

A gyakorlatban általában az E és M lépések közül az egyik nehezen kivitelezhető, míg a másik könnyen, hogy mikor melyik, az az adott probléma sajátosságaitól függ. Megmutatható, hogy $l(\vartheta_0, y) \leq l(\vartheta_1, y) \leq \dots$, azaz, hogy minden egyes iterációs lépés növeli az Y -ből vett y minta loglikelihood-függvényét. Az is igazolható továbbá, hogy bizonyos meghatározott regularitási feltételek mellett (lásd Wu [32]) ϑ_n a $\vartheta \mapsto l(\vartheta, y)$ -ből számított ML-becsléshez konvergál.

3.1.1. Megjegyzés. Bár jelen dolgozatban nem kerül tárgyalásra, csupán nagy vonalakban megemlítjük, hogy létezik az EM algoritmusnak egy általánosabb változata az ún. GEM (generalized EM). A GEM esetében nem követeljük meg az EM iterációban leírt M lépés teljesülését, csupán az a törekvésünk, hogy

$$l(\vartheta_n, y) \leq l(\vartheta_{n+1}, y)$$

teljesüljön. Világos, hogy minden EM algoritmus egyben GEM is.

3.2. Egy EM algoritmus Markov-modulált Poisson folyamatokra

A következő fejezetben ismertetünk egy EM algoritmust, melyet először T. Rydén [23] mutatott be 1996-os cikkében. Tételezzük fel, hogy van egy n elemű megfigyelésünk az y_1, y_2, \dots, y_n köztes időkből. Ez nyilván nem teljes adat, hiszen a mögöttes Markov folyamatról nem kapunk kielégítő információt. Ebben a helyzetben tehát természetesen adódó lehetőségként merül fel az EM algoritmus alkalmazása, a kérdés csupán az, hogy mit válasszunk nem megfigyelhető adatnak. Itt elsősorban két lehetőség kínálja magát:

1. Választható hiányzó adatnak az $\{X_k\}$ beágyazott Markov-lánc, vagy
2. a mögöttes $\{X_t\}$ Markov-folyamat egész trajektóriája.

Rydén korábbi munkájában vizsgálta az első esetet és megmutatta, hogy bár ekkor az E-lépés egyszerű, az M-lépés ezzel ellentétben nem explicit és csak numerikus módszerek alkalmazása segítségével kivitelezhető. Majd 1996-os cikkében vizsgálta a második esetet, ahol éppen az E-lépés a bonyolultabb, az M-lépés viszont - ahogy azt nemsokára látni fogjuk - explicit módon kifejezhető. Jelen fejezetben ezt az algoritmust fogjuk ismertetni. Ekkor tehát a megfigyelésünk egy $Y = (Y_1, Y_2, \dots, Y_n)$ véletlen elem az $((\mathbb{R}^+)^n, \mathcal{B})$ mérhető térből (ahol \mathcal{B} a Borel-halmazok σ -algebrája). A hiányzó (vagy nem megfigyelhető) adat pedig az X Markov folyamat, ami szintén egy véletlen elem egy alkalmasan választott mérhető térben. Először vezessünk be néhány statisztikát és jelölést az Y megfigyelés függvényében.

1. $t_0 := 0$ és $k = 1, 2, \dots, n$ esetén $t_k := \sum_{i=1}^k y_i$, azaz t_k nem más, mint az i -edik érkezés időpontja, a $t_0 = 0$ konvencióval. Legyen továbbá $T = t_n$.
2. Tegyük fel, hogy a mögöttes Markov folyamat ugrásai a $0 < u_1 < u_2 < \dots < u_m < T$ időpontokban történnek, majd itt is állapodjunk meg az $u_0 = 0$ és $u_{m+1} = T$ jelölésekben.
3. $1 \leq k \leq m+1$ -re $I_k := [u_{k-1}, u_k)$, azaz I_k a mögöttes Markov folyamat $(k-1)$ -edik és k -adik ugrása közti időintervallumot jelöli.
4. Legyen továbbá $s_k = X_u$ (ahol $u \in I_k$ tetszőleges), és $z_k = \#\{k : t_k \in I_k, (k \neq 0)\}$, azaz s_k a Markov-folyamat állapota a $(k-1)$ -edik és k -adik ugrása között, míg z_k az érkezések száma a Markov-folyamat $(k-1)$ -edik és k -adik ugrása között.
5. A $t \leq T$ ideig bekövetkezett érkezések számát pedig jelöljük N_t -vel:

$$N_t := \#\{k : 0 < k \leq n, t_k \leq t\}.$$

3.2.1. Megjegyzés. *Világos, hogy (y_1, \dots, y_n) -ből megkonstruálható N_t és ugyanez fordítva is igaz.*

Ezek után - ahogy az [23]-ban bemutatásra kerül - az (Y, X) minta szerinti teljes L_C likelihood-függvény a következőképpen írható fel:

$$\begin{aligned}
L_C(Q, \Lambda) &= \pi_{s_1} \prod_{k=1}^m \left(q_{s_k} e^{-q_{s_k} \Delta u_k} \frac{q_{s_k, s_{k+1}}}{q_{s_k}} \right) e^{-q_{s_{m+1}} \Delta u_{m+1}} \\
&\times \prod_{k=1}^{m+1} \left(\frac{(\lambda_{s_k} \Delta u_k)^{z_k}}{z_k!} e^{-\lambda_{s_k} \Delta u_k} \right) \frac{z_k!}{(\Delta u_k)^{z_k}} \\
&= \pi_{s_1} \prod_{k=1}^m (e^{-q_{s_k} \Delta u_k} q_{s_k, s_{k+1}}) e^{-q_{s_{m+1}} \Delta u_{m+1}} \prod_{k=1}^{m+1} (\lambda_{s_k}^{z_k} e^{-\lambda_{s_k} \Delta u_k}).
\end{aligned}$$

Ebból a loglikelihood-függvény:

$$\begin{aligned}
l_C(Q, \Lambda) &= \log \pi_{s_1} - \sum_{k=1}^m (q_{s_k} \Delta u_k + \log q_{s_k, s_{k+1}}) - q_{s_{m+1}} \Delta u_{m+1} \\
&+ \sum_{k=1}^{m+1} (z_k \log \lambda_{s_k} - \lambda_{s_k} \Delta u_k).
\end{aligned} \tag{3.1}$$

Világos, hogy $\log \pi_{s_1} = \sum_{i=1}^r \chi(X_0 = i) \log \pi_i$. Határozzuk meg most q_i együtthatóját! Könnyen láthatóan ez éppen $\sum_{\{k: 1 \leq k \leq m+1, s_k = i\}} \Delta u_k$. Erre bevezethetjük a D_i jelölést. D_i tehát nem más, mint az az idő, amit az $\{X_t\}$ Markov folyamat az i állapotban tölt a $[0, T]$ intervallum alatt. Észrevetjük, hogy D_i felírható az alábbi, kissé kompaktabb formában is:

$$D_i = \int_0^T \chi(X_t = i).$$

Nyilván λ_i együtthatója is ugyanez lesz. A $\log q_{ij}$, ($i \neq j$) tag együtthatója pedig így fejezhető ki: $\#\{k : 1 \leq k \leq m, s_k = i, s_{k+1} = j\}$. Ez tehát nem más, mint az $\{X_t\}$ Markov folyamat (i, j) -állapotváltásainak a száma a $[0, T]$ intervallum alatt. Erre bevezetjük az m_{ij} jelölést. Másképpen kifejezve:

$$m_{ij} = \#\{t : 0 < t \leq T, X_{t-} = i, X_t = j\}.$$

Hátravan még $\log \lambda_i$ együtthatójának a meghatározása. Ha ezt n_i jelöli, akkor

$$n_i = \sum_{\{k: 1 \leq k \leq m+1, s_k = i\}} z_k = \int_0^T \chi(X_{t_k} = i).$$

Tehát n_i azon események száma, melyek az alatt az idő alatt következtek be, míg $\{X_t\}$ az i állapotban volt. Ezekkel tehát (3.1) az alábbi alakban írható fel:

$$\begin{aligned}
l_C(Q, \Lambda) &= \sum_{i=1}^r \chi(X_0 = i) \log \pi_i - \sum_{i=1}^r D_i q_i \\
&+ \sum_{i=1}^r \sum_{j=1, j \neq i}^r m_{ij} \log q_{ij} + \sum_{i=1}^r (n_i \log \lambda_i - \lambda_i q_i).
\end{aligned}$$

Tegyük fel most, hogy ϑ_0 rögzített és legyen $\mathbb{P}_0 = \mathbb{P}_{\vartheta_0}$. Ekkor

$$\begin{aligned}
\mathbb{E}_0(l_C(Q, \Lambda) | Y = y) &= \sum_{i=1}^r \chi(X_0 = i) \log \hat{\pi}_i - \sum_{i=1}^r \hat{D}_i q_i \\
&+ \sum_{i=1}^r \sum_{j=1, j \neq i}^r \hat{m}_{ij} \log q_{ij} + \sum_{i=1}^r (\hat{n}_i \log \lambda_i - \lambda_i q_i),
\end{aligned} \tag{3.2}$$

ahol

$$1. \hat{\pi}_i = \mathbb{P}_0(X_0 = i | Y = y),$$

2. $\widehat{m}_{ij} = \mathbb{E}_0(m_{ij}|Y = y)$,
3. $\widehat{D}_i = \mathbb{E}_0(D_i|Y = y) = \int_0^T \mathbb{P}_0(X_t = i|Y = y)$ és végül
4. $\widehat{n}_i = \mathbb{E}_0(n_i|Y = y) = \sum_{k=1}^n \mathbb{P}_0(X_{t_k} = i|Y = y)$.

A (3.2) összegben lévő első tagot elhagyjuk, mert a minta méretének növekedésével ennek mérete elhanyagolható a többihez képest. Ezek után egyszerűen ellenőrizhető, hogy a $q_i = \sum_{j \neq i} q_{ij}$ feltételek mellett a fennmaradó kifejezés egyértelmű maximuma a

$$\widehat{q}_{ij} = \frac{\widehat{m}_{ij}}{\widehat{D}_i}, i \neq j \text{ és a } \widehat{\lambda}_i = \frac{\widehat{n}_i}{\widehat{D}_i}, i = 1, \dots, r$$

helyeken vétetik fel. Ez lesz az algoritmus M-lépése. Az E-lépés meghatározása pedig \widehat{m}_{ij} , \widehat{D}_i és \widehat{n}_i feltételes várható értékek meghatározásából áll, ami a fenténél jóval bonyolultabb feladat. Vezessük be a következő jelöléseket. Legyen

$$\alpha_t(i, \vartheta) = \pi^\vartheta \left(\prod_{k=1}^{N_t} f^\vartheta(y_k) \right) \overline{F}_\vartheta(t - N_t) \mathbf{1}_i \quad (3.3)$$

és

$$\beta_t(i, \vartheta) = \mathbf{1}_j^\top f^\vartheta(t_{N_t+1} - t) \left(\prod_{k=N_t+2}^n f^\vartheta(y_k) \right) \mathbf{1}, \quad (3.4)$$

ahol f^ϑ a (2.14) alatt definiált és

$$\overline{F}_\vartheta(t) = \exp[(Q - \Lambda)t]. \quad (3.5)$$

Itt azzal a konvencióval élünk, hogy $\prod_{p=k}^n (\cdot) = I$, amennyiben $k > n$. Ezzel egyidejűleg legyen

$$L_\vartheta^*(k) = \begin{cases} \pi^\vartheta \prod_{p=1}^k f^\vartheta(y_p) & , k = 1, 2, \dots, n \\ \pi^\vartheta & , k = 0 \end{cases}$$

és

$$R_\vartheta^*(k) = \prod_{p=k}^n f^\vartheta(y_p) \mathbf{1}, k = 1, 2, \dots$$

Könnyen láthatóan $L'_{\vartheta^0}(k)$ kielégíti az

$$L_\vartheta^*(0) = \pi^\vartheta, L_\vartheta^*(k) = L_\vartheta^*(k-1) f^\vartheta(y_k), k = 1, 2, \dots, n, \quad (3.6)$$

$R_\vartheta^*(k)$ pedig az

$$R_\vartheta^*(n+1) = \mathbf{1}, R_\vartheta^*(k) = f^\vartheta(y_k) R_\vartheta^*(k+1), k = n, \dots, 1 \quad (3.7)$$

rekurziót. Ryden [23]-ben megmutatta, hogy ekkor

$$\widehat{m}_{ij} = \frac{q_{ij}^0}{L(\vartheta^0, y)} \int_0^T \alpha_t(i, \vartheta^0) \beta_t(j, \vartheta^0) dt, \quad (3.8)$$

továbbá, hogy

$$\widehat{D}_i = \frac{1}{L(\vartheta^0, y)} \int_0^T \alpha_t(i, \vartheta^0) \beta_t(i, \vartheta^0) dt \quad (3.9)$$

és, hogy

$$\widehat{n}_i = \frac{\pi^0}{L(\vartheta^0, y)} L_{\vartheta^0}^*(k) R_{\vartheta^0}^*(k+1), \quad (3.10)$$

ahol $L(\vartheta, y)$ az y mintához tartozó likelihood-függvény értéke a ϑ helyen, ami a 2.6.2 tétel alapján $\pi^\vartheta \left(\prod_{k=1}^n f^\vartheta(y_k) \right) \mathbf{1}$. Szokás a (3.3), illetve (3.4) alatti valószínűséget *előre tekintő*, illetve *hátráló*

sűrűségnek nevezni. Hasonlóan a (3.6) és (3.7) alatti rekurziókat előre tekintő, illetve hátráló rekurzióknak. A további átalakítások során a képletek egyszerűsítése érdekében elhagyjuk a ϑ -tól való függést, hiszen a továbbiakban úgymint minden kiértékelés ϑ^0 mellett folyik. Világos, hogy

$$\alpha_t(i) = L^*(N_t)\bar{F}(t - t_{N_t})\mathbf{1}_i \text{ és } \beta_t(i) = \mathbf{1}_i^\top f(t_{N_t+1} - t)R'(N_t + 2),$$

tehát \hat{m}_{ij} , \hat{D}_i és \hat{n}_i kifejezései az alábbiak szerint alakíthatók tovább, az $L = L(\vartheta^0, y)$ egyszerűsítő jelölés bevezetése mellett.

$$\begin{aligned} \frac{L}{q_{ij}}\hat{m}_{ij} &= \int_0^T L^*(N_t)\bar{F}(t - t_{N_t})\mathbf{1}_i\mathbf{1}_j^\top f(t_{N_t+1} - t)R^*(N_t + 2)dt \\ &= \sum_{k=1}^n \int_{t_{k-1}}^{t_k^-} L^*(N_t)\bar{F}(t - t_{N_t})\mathbf{1}_i\mathbf{1}_j^\top f(t_{N_t+1} - t)R^*(N_t + 2)dt, \end{aligned}$$

ami pedig - tekintve, hogy a $[t_{k-1}, t_k)$ intervallumon $N_t = k - 1$ -, így írható tovább:

$$\begin{aligned} &= \sum_{k=1}^n \int_{t_{k-1}}^{t_k^-} L^*(k - 1)\bar{F}(t - t_{k-1})\mathbf{1}_i\mathbf{1}_j^\top f(t_k - t)R^*(k + 1)dt \\ &= \sum_{k=1}^n L^*(k - 1) \int_{t_{k-1}}^{t_k^-} \bar{F}(t - t_{k-1})\mathbf{1}_i\mathbf{1}_j^\top f(t_k - t)dt R^*(k + 1) \end{aligned}$$

Ugyanezt elvégezve \hat{D}_i -re és \hat{n}_i -ra, azt kapjuk, hogy

$$L\hat{D}_i = \sum_{k=1}^n \int_{t_{k-1}}^{t_k^-} L^*(k - 1)\bar{F}(t - t_{k-1})\mathbf{1}_i\mathbf{1}_i^\top f(t_k - t)R^*(k + 1)dt,$$

illetve

$$L\hat{n}_i = \sum_{k=1}^n L^*(k)\mathbf{1}_i\mathbf{1}_i^\top R^*(k + 1) = \sum_{k=1}^n (L^*)^i(k)R_i^*(k + 1).$$

Ezzel megadtuk az algoritmus M, illetve E lépéseit, az implementációval kapcsolatos további gyakorlati problémákat a következő alfejezetben tárgyaljuk.

3.3. Az algoritmus implementálása

Az előző fejezetben képletszerűen felírtuk az algoritmus E, illetve M lépését, most rátérünk az algoritmus implementálásának tárgyalására. Ehhez kicsit át kell alakítanunk a képleteinket, ugyanis a jelenlegivel a következő probléma merül fel. Megmutatható, hogy az előretékintő illetve hátráló sűrűségek exponenciális sebességgel tartanak 0-hoz, amennyiben $n \mapsto \infty$. Emiatt az előretékintő illetve hátráló rekurziók nem feltétlenül adnak pontos eredményt, ha a valós kalkulációk során olyan számítógépet használunk, amelynek a számítási korlátait meghaladják az esetlegesen felmerülő kiugróan kicsi vagy nagy értékek. W.J.J. Roberts, Y. Ephraim és E. Dieguez 2006-os eredményükben - lásd [21] - bemutattak egy skálázási eljárást, melynek alkalmazása megóv minket a túlsordulástól. Először ezt ismertetjük. Ennek érdekében tetszőleges y vektor esetén $k = 1, \dots, \dim y$ -ra y^k jelölje az y első k koordinátájából álló vektort, míg y^0 legyen konstans 0. Mint ismeretes, két valószínűségi változó - monjuk X és Y - együttes és az egyiknek a másikra vonatkozó feltételes sűrűségfüggvénye között - amennyiben azok léteznek - az alábbi kapcsolat áll fenn:

$$f_{(X,Y)}(x, y) = f_{X|Y}(x|y)f_Y(y).$$

Ezt n -szer alkalmazva - kissé pongyola jelöléssel, elhagyva a ϑ -tól való függést - fennáll az $L(y) = \prod_{l=1}^n f_{Y^l|Y^{l-1}}(y_l|y^{l-1})$ azonosság, amit a $c_l = f_{Y^l|Y^{l-1}}(y_l|y^{l-1})$ jelöléssel együtt egyszerűen $L = \prod_{l=1}^n c_l$ alakba írhatunk. Ezzel az \hat{m}_{ij} -re vonatkozó korábbi alfejezetben szereplő előállítás az alábbiak szerint alakul át.

$$\hat{m}_{ij} = \sum_{k=1}^n \frac{q_{ij}^0}{c_k} \left(\pi \prod_{l=1}^{k-1} \frac{f(y_l)}{c_l} \right) \int_{t_{k-1}}^{t_k^-} \bar{F}(t - t_{k-1}) \mathbf{1}_i \mathbf{1}_j^\top f(t_k - t) dt \left(\prod_{l=k+1}^n \frac{f(y_l)}{c_l} \mathbf{1} \right).$$

Itt újabb jelöléseket vezetünk be a két zárójeles kifejezésre, legyen

$$L_\vartheta(k) = \pi^\vartheta \prod_{l=1}^{k-1} \frac{f_\vartheta(y_l)}{c_l(\vartheta)} \quad \text{és} \quad R_\vartheta(k) = \prod_{l=k}^n \frac{f_\vartheta(y_l)}{c_l(\vartheta)} \mathbf{1}. \quad (3.11)$$

Ezenkívül állapotdjunk még meg abban, hogy $L_\vartheta(0) = \pi^\vartheta$, illetve $R_\vartheta(n+1) = \mathbf{1}$. A továbbiakban itt sem írjuk majd ki mindig a ϑ -tól való függést, hogy a kiértékelés milyen ϑ mellett történik, az a kontextusból úgyszólván kiderül. Mivel egyrészt $f_{Y^k}(y^k) = \prod_{l=1}^k c_l$, másrészt $f_{Y^k}(y^k) = \pi \prod_{l=1}^k f(y_l) \mathbf{1}$, azért

$$c_k = \pi \prod_{l=1}^{k-1} \frac{f(y_l)}{c_l} f(y_k) \mathbf{1} = L(k-1) f(y_k) \mathbf{1}. \quad (3.12)$$

Ezzel (3.11) és (3.12) alapján a következő rekurziók írhatók fel:

$$L(k) = \frac{L(k-1) f(y_k)}{L(k-1) f(y_k) \mathbf{1}}, \quad k = 1, \dots, n$$

és

$$R(k) = \frac{f(y_k) R(k+1)}{L(k-1) f(y_k) \mathbf{1}}, \quad k = n, \dots, 1.$$

Megmutatható (lásd [21]), hogy az ily módon átskálázott rekurziók esetén egyrészt

$$[L(k)]_j = \mathbb{P}_0(X_{t_k} = j | Y^k = y^k),$$

tehát $[L(k)]_j$ legfeljebb 1, másrészt

$$[R(k)]_j < \frac{1}{[L(k-1)]_j},$$

azaz $[R(k)]_j$ felülről korlátos és $[R(k)]_j [L(k-1)]_j < 1$.

Ezek után tehát felírható az

$$\begin{aligned}
\widehat{m}_{ij} &= \sum_{k=1}^n \frac{q_{ij}^0}{c_k} \int_{t_{k-1}}^{t_k^-} L(k-1) \overline{F}(t-t_{k-1}) \mathbf{1}_i \mathbf{1}_j^\top f(t_k-t) R(k+1) dt \\
&= \sum_{k=1}^n \frac{q_{ij}^0}{c_k} \int_{t_{k-1}}^{t_k^-} [\mathbf{1}_i^\top \overline{F}(t-t_{k-1})^\top L(k-1)^\top] [R(k+1)^\top f(t_k-t)^\top \mathbf{1}_j] dt \\
&= \sum_{k=1}^n \frac{q_{ij}^0}{c_k} \int_{t_{k-1}}^{t_k^-} [\overline{F}(t-t_{k-1})^\top L(k-1)^\top R(k+1)^\top f(t_k-t)^\top]_{ij} dt \\
&= \sum_{k=1}^n \frac{q_{ij}^0}{c_k} \int_{t_{k-1}}^{t_k^-} \left\{ [f(t_k-t) R(k+1) L(k-1) \overline{F}(t-t_{k-1})]^\top \right\}_{ij} dt \tag{3.13}
\end{aligned}$$

összefüggés, valamint amennyiben \odot jelöli az elemenkénti mátrix szorzást és \widehat{n} jelöli azt az $r \times 1$ -es vektort, aminek i . eleme \widehat{n}_i , akkor

$$\widehat{n} = \sum_{k=1}^n L(k)^\top \odot R(k+1),$$

továbbá világos, hogy

$$\widehat{D}_i = \frac{\widehat{m}_{ii}}{q_{ii}^0}.$$

Most már csak az maradt hátra, hogy megmutassuk, miként lehet hatékonyan kiértékelni a (3.13) előállításban szereplő integrálokat. Ennek érdekében jelölje \widehat{m} azt az $r \times r$ -es mátrixot, amelyben az i . sor j . eleme \widehat{m}_{ij} . Ekkor, könnyen láthatóan

$$\widehat{m} = Q^0 \odot \sum_{k=1}^n \frac{1}{c_k} \int_{t_{k-1}}^{t_k^-} [f(t_k-t) R(k+1) L(k-1) \overline{F}(t-t_{k-1})]^\top dt.$$

Ebből \widehat{m} transzponáltjára a következő adódik:

$$\widehat{m}^\top = (Q^0)^\top \odot \sum_{k=1}^n \frac{1}{c_k} \int_{t_{k-1}}^{t_k^-} f(t_k-t) R(k+1) L(k-1) \overline{F}(t-t_{k-1}) dt,$$

ami az $\mathcal{I}_k = \int_{t_{k-1}}^{t_k^-} f(t_k-t) R(k+1) L(k-1) \overline{F}(t-t_{k-1}) dt$ jelölés bevezetésével tovább egyszerűsödik:

$$\widehat{m} = Q^0 \odot \sum_{k=1}^n \frac{\mathcal{I}_k^\top}{c_k} \tag{3.14}$$

Itt figyelembe véve a (2.14) és (3.5) előállításokat és alkalmazva az $s = t-t_{k-1}$ integrál-transzformációt, azt kapjuk, hogy

$$\mathcal{I}_k = \int_0^{y_k} \exp((Q^0 - \Lambda^0)(y_k - y)) \Lambda^0 R(k+1) L(k-1) \exp((Q^0 - \Lambda^0)y) dy.$$

Ennek kiszámításához alkalmazzuk a 2.3.7 lemmát az $A = Q^0 - \Lambda^0$ és $B = \Lambda^0 R(k+1) L(k-1)$ szereposztással, legyen ennek érdekében

$$S_k = \begin{pmatrix} Q^0 - \Lambda^0 & \Lambda^0 R(k+1) L(k-1) \\ 0 & Q^0 - \Lambda^0 \end{pmatrix}$$

Ekkor tehát a 2.3.7 lemmából az következik, hogy \mathcal{I}_k nem más, mint az $\exp(S_k y_k)$ mátrix jobb felső $r \times r$ -es blokkja. Így tehát \mathcal{I}_k kiszámítását visszavezettük egy magasabb dimenziós mátrix exponenciálisának meghatározására. Ezt pedig az 2.3 fejezetben ismertetett Padé-féle diagonális approximáció módszerének segítségével tesszük meg, megjegyezve, hogy a mátrix exponenciális kiszámítására alkalmas "expm" MatlabTM függvény éppen ezt az algoritmust használja.

Most már minden adva van ahhoz, hogy megfogalmazzhassuk az algoritmusunkat.

3.3.1. Algoritmus (Rydén EM algoritmusa). 1. *Bemenő adat: $y = (y_1, y_2, \dots, y_n)$ megfigyelés egymás utáni köztes időkből.*

2. *Kezdeti $\vartheta^0 = (Q^0, \Lambda^0)$ becslés meghatározása a minta alapján.*

3. *Tegyük fel, hogy ϑ^{l-1} -et már meghatároztuk, ekkor az l . iterációs lépés:*

(a) $\pi = \pi(Q^{l-1}, \Lambda^{l-1})$ kezdeti eloszlás meghatározása.

(b) $L(0) := \pi$ és $k = 1, \dots, n$ -ig $L(k) := \frac{L(k-1)f(y_k)}{L(k-1)f(y_k)\mathbf{1}}$. (Itt, és a további iterációs lépésekben is f -et ϑ^{l-1} mellett kell kiértékelni.)

(c) $R(n+1) := \mathbf{1}$ (ahol $\mathbf{1}$ jelöli az $r \times 1$ -es csupa 1 vektort) és $k = 1, \dots, n$ -ig $R(k) := \frac{f(y_k)R(k+1)}{L(k-1)f(y_k)\mathbf{1}}$.

(d) $k = 1, \dots, n$ -ig $c_k := L(k-1)f(y_k)\mathbf{1}$.

(e) $k = 1, \dots, n$ -ig

$$S_k := \begin{pmatrix} Q^{l-1} - \Lambda^{l-1} & \Lambda^{l-1}R(k+1)L(k-1) \\ 0 & Q^{l-1} - \Lambda^{l-1} \end{pmatrix}$$

$R_k := \exp(S_k y_k)$ (Padé mátrix approximáció) és

$$(\mathcal{I}_k)_{ij} := (R_k)_{i,r+j}.$$

(f)

$$\hat{m} := Q^{l-1} \odot \sum_{k=1}^n \frac{\mathcal{I}_k^\top}{c_k}$$

(g)

$$\hat{n} := \sum_{k=1}^n L(k)^\top \odot R(k+1)$$

(h) Ekkor

$$l(\vartheta^{(l-1)}) = l(Q^{(l-1)}, \Lambda^{(l-1)}) = \sum_{k=1}^n \log c_k$$

és az új ϑ^l becslés pedig:

$$q_{ij}^l = q_{ii}^{l-1} \frac{\hat{m}_{ij}}{\hat{m}_{ii}}, i, j = 1, \dots, r, i \neq j \text{ és } \lambda_i^l = q_{ii}^{l-1} \frac{\hat{n}_i}{\hat{m}_{ii}}, i = 1, \dots, r.$$

4. *Kimenő adat: $\{\vartheta^l\}$ becsléssorozat.*

Nyitva maradt még az algoritmus 2. pontjának, azaz a kiindulási becslés kezdeti értékének megválasztásának a kérdése. Erre Rydén [23]-ban javasol egy eljárást, aminek segítségével kellően közel lehet kerülni a kezdeti paraméterekhez. Ennek azonban az a hátránya, hogy több fázisból áll és használja például a rejtett Markov modell paramétereinek a becslésére vonatkozó EM algoritmust, ami önmagában egy, az imént bemutatott algoritmushoz hasonló kaliberű procedúra. Bár a rejtett Markov modell paramétereire vonatkozó becslés kiszámítására létezik egy eljárás a MatlabTM "Statistics Toolbox" csomagjában, most azonban mégsem ezt az utat választjuk, hanem bemutatunk egy egyszerűbb intuitív eljárást, mely jóval pontatlanabb, ám egyszerűségéből adódóan rendkívül gyorsan számolható és legalább a paraméterek egymáshoz való nagyságrendi viszonya eltalálható vele.

Egy MMPP esetén a λ_i , ($i = 1, \dots, r$) intenzitáshányadok véletlenül váltakoznak. Ha a mögöttes Markov folyamat éppen az i . állapotban van, akkor az ez alatt bekövetkező érkezések $\frac{1}{\lambda_i}$

| | | | | | | |
|----------------------------|--------|--------|--------|--------|--------|--------|
| köztes idő sorszáma | 1 | 2 | 3 | 4 | 5 | 6 |
| köztes idő | 0.0042 | 0.0118 | 0.0106 | 0.0380 | 0.0254 | 0.0200 |
| valódi állapot | 1 | 1 | 1 | 2 | 2 | 1 |
| becsült állapot | 1 | 1 | 1 | 2 | 2 | 1 |

3.1. táblázat. A kiinduló becslés pontossága

| | | | | | | |
|----------------------------|--------|--------|--------|--------|--------|--------|
| köztes idő sorszáma | 7 | 8 | 9 | 10 | 11 | 12 |
| köztes idő | 0.0050 | 0.0064 | 0.0035 | 0.0284 | 0.0105 | 0.0413 |
| valódi állapot | 1 | 1 | 1 | 1 | 1 | 1 |
| becsült állapot | 1 | 1 | 1 | 2 | 1 | 2 |

3.2. táblázat. A kiinduló becslés pontossága (folytatás)

várható értékű és $\frac{1}{\lambda_i}$ szórású exponenciális eloszlású valószínűségi változók, ezért úgy képzeljük, hogy az ez alatt az idő alatt bekövetkező érkezések közti időkből vett megfigyeléseink $\frac{1}{\lambda_i}$ környéken kell, hogy szóródjanak. Ez alapján megpróbáljuk r csoportra osztani az Y megfigyelésünket, mégpedig úgy, hogy az egyes csoportokon belüli elemek az euklideszi távolságra nézve közel legyenek egymáshoz. Ezt pedig az euklideszi távolságra vonatkozó k -means klaszterező eljárással - egy magyar nyelvű bevezető található például az [1] alatti jegyzetben - tesszük meg, megjegyezve, hogy ez az eljárás már szintén előre programozottan rendelkezésünkre áll a MatlabTM-ban. Ha elkészültek a klaszterek, akkor azt feltételezzük, hogy az egy klaszteren belüli megfigyelések f.a.e. exponenciális valószínűségi változók. Ekkor a paraméter maximum-likelihood becslése a mintaátlag reciproka, amit könnyen megkaphatunk, hiszen az egy klaszteren belüli elemek mintaátlaga éppen a klaszterközéppont. Összefoglalva, legyen adott az $Y = (Y_1, Y_2, \dots, Y_n)$ köztes időkből vett minta, és tegyük fel, hogy a klaszterezés az $\{1, 2, \dots, n\}$ halmaz alábbi partícióját eredményezi:

$$\{1, 2, \dots, n\} = \bigcup_{i=1}^r K_i.$$

Ekkor a λ_i -k becslésére a

$$\hat{\lambda}_i = \frac{n}{\sum_{j \in K_i} Y_j}, \quad i = 1, 2, \dots, r$$

összefüggés adódik. Ennek az eljárásnak a kritikus pontja az, hogy a köztes idők közül valójában nem sorolható mindegyik köztes idő tisztán valamelyik állapothoz, ugyanis van olyan, ami egy adott állapotban indul, de már egy másik állapotban ér véget. Ettől a nehézségtől most az egyszerűség kedvéért eltekintünk. Illusztráció gyanánt álljon itt a következő példa.

3.3.1. Példa. *9000 elemű mintát szimuláltunk a $q_{12} = 1$ és $q_{21} = 10$, illetve $\lambda_1 = 100$ és $\lambda_2 = 10$ paraméterekkel rendelkező két állapotú MMPP-ből. A minta első néhány értékét a (3.1) és (3.2) táblázatokban foglaltuk össze. A "valódi állapot" mező mutatja a mögöttes Markov-folyamat állapotát a megfelelő érkezés bekövetkeztekor, a "becsült állapot" mező pedig a klaszterezés során kapott javaslatot.*

Látjuk tehát, hogy a klaszterező eljárás több ízben is hibázik a köztes idők klasszifikálása során. A klaszterközéppontokra $c_1 = 0.0075$ illetve $c_2 = 0.041$ adódik, melyből a $\hat{\lambda}_1 = \frac{1}{c_1} = 133.6267$ illetve a $\hat{\lambda}_2 = \frac{1}{c_2} = 24.3681$ kiinduló becsléseket kapjuk.

A Q generátormátrix becslésénél visszautalunk a korábbi fejezetekben ismertetett összefüggésekre, nevezetesen arra, hogy amennyiben a mögöttes Markov folyamat éppen az i . állapotban van,

úgy az ott töltött ideje q_i paraméterű exponenciális eloszlású valószínűségi változó, majd ezek után a $\frac{q_{ij}}{q_j}$ valószínűségek szerint ugrik át egy másik $j \neq i$ állapotba. Ebből kiindulva a klaszterezett minta alapján megpróbáljuk reprodukálni az egyes állapotokban eltöltött időket.

3.3.2. Példa. Például a (3.1) és (3.2) táblázatokban összefoglalt minta esetén a klaszterezés szerint a Markov folyamat kezdetben az 1. állapotban volt, majd a 4. érkezési idő már a 2. állapot szerinti intenzitással történt. Ez azt jelenti, hogy az állapotváltás valahol T_3 és T_4 között volt. Mivel a 2. állapot szerinti érkezésre az állapotváltástól számítva várhatóan $\frac{1}{\lambda_2}$ időt kell várni, ezért a Markov folyamat 1. állapotban eltöltött idejét - úgy értve, hogy az első állapotváltásig - a $\xi_{11} = T_4 - \frac{1}{\lambda_2}$ kifejezéssel becsüljük. Ezután megkeressük a következő csupa 1-esből álló szakaszt. Látható, hogy ez a 6. megfigyeléstől kezdődik és a 9. megfigyelésnél ér véget. Tehát itt azt feltételezzük, hogy az állapotváltás T_9 és T_{10} között volt. Így az ebben az állapotban töltött időt most a $\xi_{12} = T_{10} - \frac{1}{\lambda_2} - T_5 - (y_6 - \frac{1}{\lambda_1})$ kifejezéssel becsülhetjük. Ezzel egy $\xi_{11}, \xi_{12}, \dots$ mintát kapunk q_1 paraméterű exponenciális eloszlásból, majd ezután ismét ezek mintaátlagának reciprokával becsüljük a q_1 paramétert.

Mіндеzt precízebben megfogalmazva, jelölje $K_i \in \{1, 2, \dots, r\}$ az Y_i megfigyelés klaszterének sorszámát. Legyen H_{ij} a j -edik maximális méretű csupa i -ből álló blokk elemeinek az indexe K -ban ($i = 1, \dots, r$ és $H_{ij} = \emptyset$, ha már nincs több ilyen blokk). Legyen $n_i = \#\{j : H_{ij} \neq \emptyset\}$. Ezzel tehát ismét partíciónáltuk az $\{1, 2, \dots, n\}$ halmazt:

$$\{1, 2, \dots, n\} = \bigcup_{i=1}^r \bigcup_{j=1}^{\infty} H_{ij}.$$

H_{ij} legkisebb eleme legyen m_{ij} , legnagyobb eleme pedig M_{ij} . Ekkor legyen

$$\begin{aligned} \hat{\xi}_{ij} &= \left(T_{M_{ij}+1} - T_{m_{ij}-1} - \frac{1}{\hat{\lambda}_{K_{M_{ij}+1}}} \right) - \left(Y_{m_{ij}} - \frac{1}{\hat{\lambda}_{K_{M_{ij}}}} \right) \\ &= T_{M_{ij}+1} - T_{m_{ij}} + \frac{1}{\hat{\lambda}_{K_{M_{ij}}}} - \frac{1}{\hat{\lambda}_{K_{M_{ij}+1}}}, \quad j = 1, \dots, n_i \end{aligned}$$

Ebből már megkapható q_i becslése:

$$\hat{q}_i = \frac{n_i}{\sum_{j=1}^{n_i} \hat{\xi}_{ij}}.$$

Ezzel megbecsültük a Q mátrix főátlóját, hátravan még a többi elem becslése. Ehhez rögzített $i \in \{1, \dots, r\}$ esetén minden $j \neq i$ -hez készítsük el az $A_{ij} = \#\{0 \leq k \leq n-1 : K_k = i, K_{k+1} = j\}$. Itt tehát A_{ij} adja meg a klaszterezés alapján várt (i, j) állapotváltások számát. Legyen

$$A_{i\cdot} = \sum_{j=1, j \neq i}^r A_{ij},$$

vagyis $A_{i\cdot}$ azt adja meg, hogy összesen hány állapotváltás volt az i állapotból. Ekkor a korábban említettek miatt $j \neq i$ esetén q_{ij} természetes becsléseként

$$\hat{q}_{ij} = \hat{q}_i \frac{A_{ij}}{A_{i\cdot}}$$

adódik. $i = 1, \dots, r$ esetén legyen $A_{ii} = -A_{i\cdot}$ és jelöljük B -vel azt az $r \times r$ -es mátrixot, ahol az i -edik sor j -edik eleme $\frac{A_{ij}}{A_{i\cdot}}$. Ekkor tehát a Q mátrix becslése

$$\hat{Q} = CB,$$

ahol $C = \text{diag}(q)$, vagyis a q elemeiből képzett diagonális mátrix.

3.3.3. Példa. A 3.3.1 példa esetén ekkor $\hat{q}_{12} = 5.2394$ illetve $\hat{q}_{21} = 14.2649$ adódik kiinduló becslésként.

Összefoglalásképpen, bár a módszer nem szolgáltat túl nagy pontosságot és érzékeny arra, hogy például a λ_i paraméterek mennyire jól szeparálhatók egymástól - tehát kellő pontosságú eredmény inkább akkor várható, ha előre tudjuk, hogy a λ_i intenzitáshányadok között nincsenek nagyon közeli - , arra azonban mindenképpen alkalmas, hogy gyorsan jussunk intuitív kezdeti becslésekhez. A következőkben ismertetjük e becslés kiszámításának konkrét implementációját.

3.3.2. Algoritmus (Kiinduló becslés keresése). 1. *Bemenő adatok: y n dimenziós vektor, r (az illesztendő modell rendje)*

2. *Klaszterezés k -means eljárás segítségével, kimenetek:*

(a) *k n dimenziós vektor, k_i jelenti y_i klaszterének sorszámát (1-től r -ig terjedhet)*

(b) *$c = (c_1, \dots, c_r)$ vektor, amely tartalmazza a klaszterközpontokat*

3. **1. kimenet:** *a λ_i paraméterek becslése: $\hat{\lambda}_i := \frac{1}{c_i}, i = 1, \dots, r$.*

4. *Levágjuk az utolsó blokkot a mintából (az ugyanis egy csonka megfigyelést adna), $n := \dim y$*

5. *inicializálás:*

(a) *S legyen $r \times 1$ -es vektor $S(i) := \begin{cases} 0 & , \text{ha } i \neq k_1 \\ y_1 & , \text{ha } i = k_1 \end{cases}$*

(b) *N legyen $r \times 1$ -es vektor, csupa nulla elemekkel.*

6. *Első iteráció: $i = 2, \dots, n$ -ig:*

(a) *Ha $k_i = k_{i-1}$, akkor $S(k_{i-1}) := S(k_{i-1}) + y_i$, különben*

(b) *$S(k_{i-1}) := S(k_{i-1}) + y_i - c_{k_i}$;*

$S(k_i) := S(k_i) + c_{k_i}$;

$N(k_{i-1}) := N(k_{i-1}) + 1$.

7. *Utólagos módosítás: $S(k_n) := S(k_n) - c_{k_n}$.*

8. *A q_i paraméterek becslése $\hat{q}_i := \frac{S(i)}{N(i)}, i = 1, \dots, r$.*

9. *Az A mátrix inicializálása $r \times r$ -es csupa nulla mátrixként, majd A feltöltése, 2. iteráció: $i = 1, \dots, n - 1$ -ig:*

(a) *Ha $k_i \neq k_{i+1}$, akkor $A_{k_i, k_{i+1}} := A_{k_i, k_{i+1}} + 1$.*

10. *$i = 1, \dots, r$ -ig $A_i := \sum_{j=1, j \neq i}^r A_{ij}$ és $A_{ii} := -A_i$.*

11. *$B := \text{diag}(\frac{1}{A_1}, \dots, \frac{1}{A_r})A$ és $C := \text{diag}(q_1, \dots, q_r)$.*

12. **2. kimenet:** *Q becslése: $\hat{Q} = CB$*

Ezzel tehát befejeztük az implementáció kérdésének tárgyalását, mielőtt azonban rátérnénk az EM algoritmus gyakorlati alkalmazására, megemlítjük még annak egy fontos tulajdonságát. Nevezetesen azt, hogy megőrzi a nullákat, vagyis, ha a ϑ^0 kezdeti becslés esetén $q_{ij}^0 = 0$ valamilyen i -re és j -re, akkor a következő iteráció során kiszámolt becslés is rendelkezni fog ezzel a tulajdonsággal, azaz $q_{ij}^1 = 0$ szintén fenáll. Ez könnyen látható abból a tényből, hogy $q_{ij}^1 = \frac{\hat{m}_{ij}}{\hat{D}_i}$ és a (3.14) előállítás alapján $\hat{m}_{ij} = q_{ij}^0 \left(\sum_{k=1}^n \frac{x_k^\top}{c_k} \right)_{ij}$. Ez a tulajdonság igen hasznos lehet bizonyos speciális struktúrájú MMPP-k becslésekor, nevezetesen amikor tudjuk, hogy a Q mátrix bizonyos elemei nullával egyenlők.

4. fejezet

Markov-modulált Poisson folyamat a gyakorlatban

Jelen fejezet az előző fejezetben bemutatott algoritmus alkalmazása egy gyakorlati példára. Az 1. alfejezetben ismertetjük, hogy milyen eljárást használtunk MMPP-ből vett minta szimulálására. A 2. alfejezet fontos önálló eredménye a 4.2.1 lemma, - ami a 2.3.8 lemmán alapul -, enélkül ugyanis igencsak nehézkes lenne az alfejezet során bemutatásra kerülő $\mathbb{P}(\zeta_{k+1} < K)$ farokvalószínűségek numerikus számolása. A 3. alfejezetben közöljük vizsgálódásaink numerikus eredményeit.

4.1. Minta generálása adott paraméterű MMPP-ből

Ebben az alfejezetben bemutatjuk, hogy miként tudunk mintát generálni egy megadott paraméterű Markov modulált Poisson folyamatból vett köztes időkből. E célból legyen tehát a $\vartheta = (Q, \Lambda)$ paraméter rögzített. Kihhasználva a mögöttes Markov folyamat fejlődését leíró tulajdonságokat és azt, hogy egy λ intenzitású Poisson folyamat köztes idői független, λ paraméterű exponenciális eloszlásúak, az alábbi módon járhatunk el. Véletlenszerűen kiválasztjuk a Markov folyamat kezdeti állapotát a π stacionárius eloszlás szerint. Ekkor a Markov folyamat a megadott állapottól függő - nevezetesen q_i paraméterű, amennyiben i ez az állapot - exponenciális eloszlás szerinti időt tölt itt el, majd szintén a megadott állapottól függően átvált egy másik állapotba. Rögzítsük ezt az időt és jelöljük τ_1 -gyel. Ez alatt az időintervallum alatt a folyamatunk szintén az adott állapottól függő - nevezetesen λ_i intenzitáshányadú, amennyiben i ez az állapot - homogén Poisson folyamatként viselkedik. Ez azt jelenti, hogy a megfigyelések az adott paraméter szerinti, f.a.a. exponenciális eloszlású valószínűségi változók szerint érkeznek. Generáljunk tehát ezekből egy mintát egészen addig, amíg a mintaelemek összege még kisebb, mint τ_1 . Ezután határozzuk meg a $\frac{q_i}{q_i}$ valószínűségek szerint a Markov folyamat új állapotát, majd ismét szimuláljuk az itt eltöltött τ_2 időt, mely alatt ismét konstans a Poisson folyamat intenzitáshányada. Ezt folytassuk egészen addig, amíg elő nem állt a kellő elemszámú minta. Ugyanez a következő módon fogalmazható meg egzakt módon.

4.1.1. Algoritmus (MMPP szimuláció). 1. *Bemenet:* $\vartheta = (Q, \Lambda)$ paraméter, n mintaelemszám.

2. π stacionárius eloszlás meghatározása a $\pi Q = 0, \pi \mathbf{1} = 1$ egyenletrendszer megoldásával.

3. Válasszunk egy s_0 számot véletlenszerűen az $E = \{1, 2, \dots, r\}$ halmazból π szerint. (i valószínűsége π_i).

4. A Markov folyamat aktuális állapotát az s változóban tároljuk, s inicializálása:

$$s := s_0.$$

.

5. Inicializálás: $y := 0$, $t := \text{vélexp}(q_s)$ és $a := \text{vélexp}(\lambda_s)$.

6. Ciklus: amíg $\dim y < n + 1$:

(a) Ciklus: amíg $(a < t)$ és $(\dim y < n + 1)$

i. $y := (y, a - \sum_{i=1}^{\dim y} y_i)$.

ii. $a := a + \text{vélexp}(\lambda_s)$.

(b) $a := t$.

(c) $e \in E \setminus s$ véletlen állapot választása a $\frac{q_{sj}}{q_s}$ valószínűségek szerint

(d) $s := e$.

(e) $t := t + \text{vélexp}(q_s)$ és $a := a + \text{vélexp}(\lambda_s)$.

7. Ciklus: $i = 1, \dots, n$ -re $y_i := y_{i+1}$

8. kimenet: n hosszú y sorvektor MMPP köztes időkből.

Itt $\text{vélexp}(\lambda)$ alatt egy λ paraméterű exponenciális eloszlás szerint generált véletlen számot értünk. Ha van egy olyan generátorunk, ami egyenletes eloszlású U véletlen számot állít elő a $(0, 1)$ intervallumból, akkor tetszőleges abszolút folytonos F eloszlásból vett véletlen minta generálható vele, ugyanis $F^{-1}(U)$ eloszlása éppen F lesz. Mi a Matlab beépített "randexp" függvényét használtuk exponenciális eloszlású minta generálására.

4.2. Egy biztosításmatematikai feladat

A Markov-modulált Poisson folyamatnak rengeteg gyakorlati alkalmazása van, használata igen közkedvelt az olyan sorbanállási illetve kiszolgálási modellek esetén, ahol az egyes kiszolgálási idők intenzitása az időben véletlenszerűen váltakozik. Jelen dolgozat a továbbiakban a következő problémát vizsgálja. Tételezzük fel, hogy adott egy biztosító társaság, akinek a biztosítottjaival meghatározott káresemények történhetnek. Most nem koncentrálnak a kárnagyságra, csupán a károk száma érdekel bennünket. Ha egy adott pillanattól kezdve (ezt választjuk origónak) elkezdjük megfigyelni a károk bekövetkezésének idejét, akkor ezek meghatároznak egy konfigurációt az \mathbb{R}^+ halmazon. Ha $n(B)$ jelöli a konfigurációból az \mathbb{R}^+ egy B Borel-halmazába eső pontok számát, akkor ezáltal egy N pontfolyamat n realizációját kapjuk. A gyakorlatban persze a B Borel-halmaz nyilván többnyire egy intervallum, azaz a biztosítót az érdekli, hogy egy meghatározott összefüggő időperiódusban hány kár következik be. Persze ennek nem kell feltétlenül így lennie, hiszen érdekelheti például a biztosítót a csak tavasszal bekövetkező károk statisztikája, ami máris egy nem összefüggő halmazt eredményez, bár ekkor is vissza tudjuk vezetni a problémát diszjunkt intervallumok uniójára.

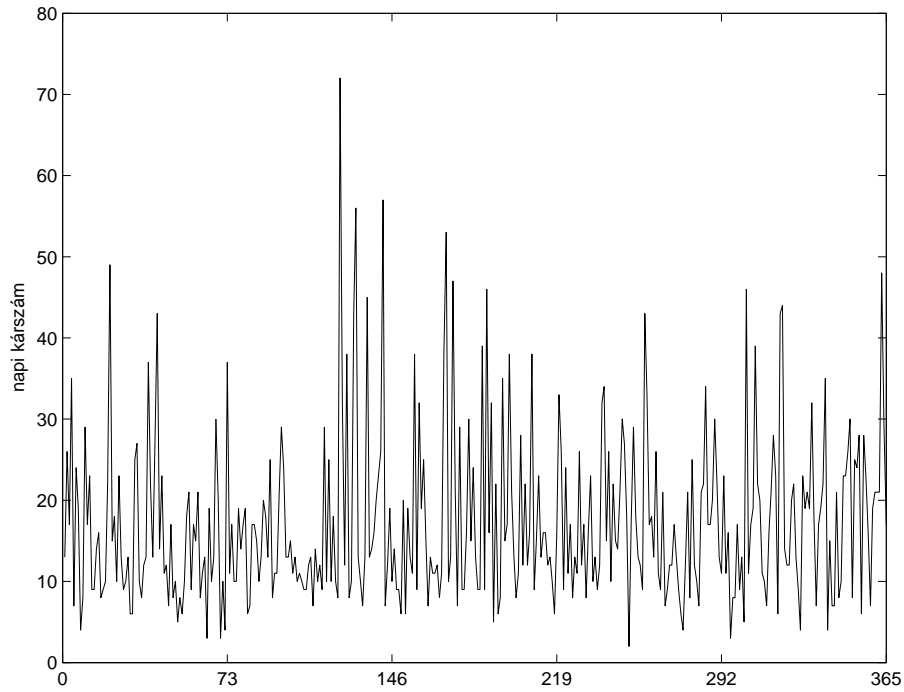
Az alapsituáció a következő: a társaság a kárfolyamatot MMPP-vel próbálja modellezni. Képzeljük el például, hogy a biztosító portfóliója gépjármű felelősség-biztosításokból áll, ekkor gépkocsibalesetek bekövetkezési idejét kísérjük figyelemmel. Köztudott, hogy esős időszak esetén több káresemény történik, azon kívül kevesebb. Az, hogy mikor van esős időszak, nem determinisztikus, feltehető például, hogy az esős illetve napos időszakok egy kétállapotú Markov folyamat szerint váltakoznak. Világos, hogy ez esetben természetes feltevésként adódik az MMPP, mint kárfolyamat. Nyilván ennél több állapot is elképzelhető, például az előbbi esetnél maradván, több időjárástól függő kategóriát is bevezethetnénk.

Egyéb modellre is gondolhatunk, például egy egészségbiztosításokból álló portfólió esetén lehet a nagyobb intenzitással járó véletlentől függő időszak egy járvány által okozott. A továbbiakban nem kívánunk különösebb hangsúlyt fektetni a konkrét interpretációra, csupán magára a modellre fókuszálunk, azaz a kárfolyamatról csak mint matematikai objektumról beszélünk és ezt modellezzük egy MMPP-vel. A feladatunk a következő lesz. Feltesszük, hogy a biztosító több egymást követő egyenlő időintervallumon keresztül (pl. nap, hónap, év) rögzíti a kárstatisztikát.

A cél előrejelzést adni az éppen soron következő időintervallumban bekövetkező károk ζ számára, illetve az előre rögzített α megbízhatósági szintekhez tartozó $K(\zeta) = \inf\{K \in \mathbb{N}_0 : \mathbb{P}(\zeta < K) \geq 1 - \alpha\}$ kvantilis értékekre. Előbbi egyfajta legjobb tudásunk szerint elvárható becslést (biztosítási szaknyelven Best Estimate) szolgáltat, míg az utóbbi érték arról informál, hogy mi az a szám, aminél nagyobb kárszám már egyhez közeli valószínűséggel nem következik be, tehát ez tulajdonképpen egy bizonyos értelemben a legnagyobb lehetséges veszteséget jelenti.

Tegyük fel, hogy az említett egyenlő nagyságú időperiódusok hossza h , az egyszerűség kedvéért hívjuk ezeket *napoknak*. Tegyük fel, hogy k darab napból álló megfigyeléssel rendelkezünk. Az l . nap alatt egészen pontosan az $((l-1)h, lh]$ intervallumot értjük, erre a továbbiakban, mint I_l is hivatkozunk majd. Jelölje az l . nap alatt bekövetkező károk számát ζ_l , ekkor $\zeta_l = N(((l-1)h, lh])$. Mi az $\mathbb{E}(\zeta_{k+1})$ illetve az $\inf\{K \in \mathbb{N}_0 : \mathbb{P}(\zeta_{k+1} < K) \geq 1 - \alpha\}$ értékeket akarjuk meghatározni az MMPP paramétereire kapott becslés alapján. Itt két különböző esetet fogunk vizsgálni.

Az egyik esetben feltesszük, hogy a biztosítónak nemcsak a napi kárszámok, hanem a pontos káridőpontok is rendelkezésére állnak az egész $(0, kh]$ intervallumon. Ekkor az előző fejezetben ismerttetett EM algoritmus segítségével megbecsüljük az MMPP paramétereit, majd az így kapott paraméterekkel kiszámoljuk a minket érdeklő előrejelzéseket.



4.1. ábra. Napi kárszámok egy MMPP egyik realizációja esetén

A másik esetben azt tesszük fel, hogy csak a napi $(\zeta_1, \zeta_2, \dots, \zeta_k)$ kárszámokból álló megfigyelés áll a rendelkezésünkre. Ekkor itt is az lenne a cél, hogy az ezen mintából számított egzakt ML-becslést felírva megbecsüljük az MMPP paramétereit, majd kiszámoljuk a megfelelő előrejelzéseket. A likelihood-függvény explicit alakban való megadásának bonyolultsága miatt azonban ehelyett egyszerűsített közelítő modelleket alkalmazunk, név szerint az alábbiakat tekintjük majd:

1. A napi kárszámok f.a.e. Poisson valószínűségi változók.
2. A napi kárszámok f.a.e. keverék Poisson valószínűségi változók.

A további célunk az lesz, hogy numerikus példákat megvizsgálva megállapítsuk, hogy egyrészt mennyire tér el a valóságtól a teljes trajektória ismeretében számított előrejelzés, másrészt, hogy mennyire tér el a valóságtól illetve a teljes trajektória ismeretében számított becsléstől az egyszerűsített és csupán napi kárszámokat felhasználó előrejelzés.

A továbbiakban feltesszük, hogy a kárfolyamat által meghatározott MMPP valós paramétereit Q és Λ , és, hogy a π kezdeti eloszlás a Q stacionárius eloszlása. Az előző fejezetben látottak alapján

$$\begin{aligned} \mathbb{E}(\zeta_{k+1}) &= \mathbb{E}(N_{kh+h} - N_{kh}) = \mathbb{E}(N_{kh+h}) - \mathbb{E}(N_{kh}) \\ &= \pi\lambda(kh + h) - \pi\lambda kh = \pi\lambda h, \end{aligned}$$

így az előrejelzésünk a kárszámra $\pi\lambda h$, ha ismerjük a valós paramétereiket.

Ha ζ tetszőleges diszkrét valószínűségi változó, akkor

$$\mathbb{P}(\zeta < K) = \sum_{n=0}^{K-1} \mathbb{P}(\zeta = n),$$

ezért az $1 - \alpha$ megbízhatósági szinthez tartozó kvantilis kiszámítása az alábbi módon algoritmizálható.

4.2.1. Algoritmus. 1. *Bemenet:* ζ v.v., $1 - \alpha$.

2. *Inicializálás:* $K := 1$.

3. *Ciklus:* Amíg $\mathbb{P}(\zeta < K) < 1 - \alpha$

(a) $K := K + 1$

4. *Kimenet:* K kvantilis.

Látjuk, hogy az algoritmus egyetlen kérdéses pontja, hogy milyen könnyen tudjuk számolni a $\mathbb{P}(\zeta < K)$ mennyiségeket. Ha ezt ζ_{k+1} -re akarjuk alkalmazni, akkor tehát a $\mathbb{P}(\zeta_{k+1} < K)$ mennyiségek számítására kell hatékony módszert adnunk.

$$\begin{aligned} \mathbb{P}(\zeta_{k+1} < K) &= \sum_{n=0}^{K-1} \mathbb{P}(\zeta_{k+1} = n) = \sum_{n=0}^{K-1} \mathbb{P}(N_{kh+h} - N_{kh} = n) \\ &= \sum_{n=0}^{K-1} \pi P(n, h) \mathbf{1} = \pi \left(\sum_{n=0}^{K-1} P(n, h) \right) \mathbf{1}, \end{aligned}$$

tehát a célunk elérése érdekében az $S(K, h) = \sum_{n=0}^{K-1} P(n, h)$ kiszámítása szükségeltetik. Erre a következő állítás szolgáltat megoldást.

4.2.1. Lemma. *Tekintsük az alábbi $nr \times nr$ -es mátrixot:*

$$C = \begin{pmatrix} Q - \Lambda & \Lambda & 0 & \cdots & 0 \\ 0 & Q - \Lambda & \Lambda & \cdots & 0 \\ 0 & 0 & Q - \Lambda & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \Lambda \\ 0 & 0 & 0 & Q - \Lambda & Q - \Lambda \end{pmatrix}$$

Legyen $D(h) = \exp(Ch)$, ekkor $S(K, h)$ nem más, mint a $D(h)$ mátrixból a legelső r sorának megtartásával és a többi sorának elhagyásával képzett $E(h)$ mátrix, így $S(K, h) \mathbf{1}$ éppen ezen mátrix sorösszegeiből képzett oszlopvektor.

Bizonyítás. A második fejezet MMPP-kről szóló alfejezete alapján tudjuk, hogy $P(n, h)$, $n = 0, 1, \dots, K - 1$ kielégíti az alábbi differenciál-egyenletrendszer:

$$\begin{aligned} P(0, 0) &= I \text{ és } P(n, 0) = 0, \text{ ha } n \geq 1, \\ P'(0, t) &= P(0, t)(Q - \Lambda), \\ P'(n, t) &= P(n, t)(Q - \Lambda) + P(n - 1, t)\Lambda, \text{ ha } n \geq 1. \end{aligned}$$

Emiatt a 2.3.8 lemma szerint $P(n, h) = [D(h)]_{I_0, I_n}$, amiből már egyszerűen következik az állítás.

■

4.2.1. Megjegyzés. *A mátrix exponenciális kiszámítása itt is a már korábban megismert Padé approximáció segítségével történik.*

Ezzel tehát megmutattuk, hogy hogyan számolhatjuk ki a kérdéses előrejelzéseket abban az esetben, ha az összes érkezési időpont ismert. Most meghatározzuk ugyanezeket az egyszerűsített modellek esetén. Először tekintsük azt az esetet, amikor azt feltételezzük, hogy a $\zeta_1, \zeta_2, \dots, \zeta_k$ kárszámok f.a.e. exponenciálisak. Jelölje a paramétert μ . Az világos, hogy ekkor $\mathbb{E}(\zeta_{k+1}) = \mu$. A kvantilis meghatározása a 4.2.1 algoritmus alapján semmilyen nehézséget nem okoz, tekintve, hogy

a $\mathbb{P}(\zeta_{k+1} = n) = \frac{\mu^n}{n!} e^{-\mu}$ valószínűségek explicit alakban adóttak. Már csak a paraméter becslése maradt hátra, ez azonban szintén könnyen elvégezhető, hiszen ekkor a likelihood-függvény

$$L(\mu) = \prod_{i=1}^k \frac{\mu^{\zeta_i}}{\zeta_i!} e^{-\mu},$$

vagyis a loglikelihood-függvény

$$l(\mu) = \log \mu \sum_{i=1}^k \zeta_i - k\mu - c = k(\bar{\zeta} \log \mu - \mu) - c,$$

ahol $\bar{\zeta} = \frac{1}{k} \sum_{i=1}^k \zeta_i$ és c alkalmas $\zeta_1, \zeta_2, \dots, \zeta_k$ -től függő konstans. Ennek a kifejezésnek pedig $\mu = \bar{\zeta}$ esetén van maximuma, tehát a paraméter ML becslése nem más, mint $\hat{\mu} = \bar{\zeta}$.

Valamivel nehezebb dolgunk van, ha a $\zeta_1, \zeta_2, \dots, \zeta_k$ kárszámokat független, azonos eloszlásúaknak feltételezzük ugyan, de nem tisztán Poisson, hanem kevert Poisson eloszlást tételezünk fel. Itt kevert Poisson eloszlás alatt a következőt értjük.

4.2.1. Definíció. Legyen $\tau : \Omega \mapsto \{\mu_1, \mu_2, \dots, \mu_r\}$ valószínűségi változó és vezessük be a $p_i = \mathbb{P}(\tau = \mu_i)$ jelöléseket. Azt mondjuk, hogy a ζ valószínűségi változó a τ keverő valószínűségi változó (vagy a p keverő eloszlás) szerinti kevert Poisson eloszlás, ha

$$\mathbb{P}(\zeta = i | \tau) = \frac{\tau^i}{i!} e^{-\tau}.$$

Durván fogalmazva ez azt jelenti, hogy a ζ valószínűségi változó p_i valószínűséggel μ_i paraméterű Poisson eloszlású.

4.2.2. Megjegyzés. Természetesen nem szükséges, hogy τ csak véges sok értéket vegyen fel, lehet akár folytonos eloszlás is, mi azonban most csak ezt az esetet fogjuk használni.

Ebben az esetben tehát a paraméterek a $\mu = (\mu_1, \mu_2, \dots, \mu_r)$ és a $p = (p_1, p_2, \dots, p_r)$ vektorok. Először ezek segítségével megadjuk a $\mathbb{P}(\zeta_{k+1} = i)$ illetve az $\mathbb{E}(\zeta_{k+1})$ értékeket. A teljes valószínűség tételének értelmében

$$\mathbb{P}(\zeta_{k+1} = i) = \sum_{j=1}^r \mathbb{P}(\zeta = i | \tau = \mu_j) \mathbb{P}(\tau = \mu_j) = \sum_{j=1}^r \frac{\mu_j^i}{i!} e^{-\mu_j} p_j,$$

a várható érték pedig a teljes várható érték tételének figyelembe vételével

$$\mathbb{E}(\zeta_{k+1}) = \sum_{j=1}^r \mathbb{E}(\zeta_{k+1} | \tau = \mu_j) \mathbb{P}(\tau = \mu_j) = \sum_{j=1}^r \mu_j p_j.$$

Tehát a kvantilisre és a kárszámra vonatkozó előrejelzéseinket most is könnyen el tudjuk végezni, ha a μ és p paraméterek becslése már ismert. A likelihood-függvény most

$$L(\mu, p) = \prod_{i=1}^k \left(\sum_{j=1}^r \frac{\mu_j^{\zeta_i}}{\zeta_i!} e^{-\mu_j} p_j \right),$$

vagyis a loglikelihood

$$l(\mu, p) = \sum_{i=1}^k \log \left(\sum_{j=1}^r \frac{\mu_j^{\zeta_i}}{\zeta_i!} e^{-\mu_j} p_j \right)$$

alakot ölt. Ez esetben a likelihood egyenletek

$$0 = \frac{\partial}{\partial \mu_s} l(\mu, p) = \sum_{i=1}^k \left(\sum_{j=1}^r \frac{\mu_j^{\zeta_i}}{\zeta_i!} e^{-\mu_j} p_j \right)^{-1} \frac{\mu_s^{\zeta_i}}{\zeta_i!} e^{-\mu_s} p_s \left(\frac{\zeta_i}{\mu_s} - 1 \right),$$

$$0 = \frac{\partial}{\partial p_s} l(\mu, p) = \sum_{i=1}^k \left(\sum_{j=1}^r \frac{\mu_j^{\zeta_i}}{\zeta_i!} e^{-\mu_j} p_j \right)^{-1} \frac{\mu_s^{\zeta_i}}{\zeta_i!} e^{-\mu_s},$$

$$0 = \sum_{i=1}^r p_i - 1, s = 1, \dots, r$$

alakúak, amivel nyilván ekvivalens a

$$\begin{aligned}
0 &= \frac{\partial}{\partial \mu_s} l(\mu, p) = \sum_{i=1}^k \left(\sum_{j=1}^r \frac{\mu_j^{\zeta_i}}{\zeta_i!} e^{-\mu_j} p_j \right)^{-1} \frac{\mu_s^{\zeta_i-1}}{(\zeta_i-1)!}, \\
0 &= \frac{\partial}{\partial p_s} l(\mu, p) = \sum_{i=1}^k \left(\sum_{j=1}^r \frac{\mu_j^{\zeta_i}}{\zeta_i!} e^{-\mu_j} p_j \right)^{-1} \frac{\mu_s^{\zeta_i}}{\zeta_i!}, \\
0 &= \sum_{i=1}^r p_i - 1, s = 1, \dots, r
\end{aligned}$$

rendszer. Ebben az esetben az ML becslés egzakt felírása tehát nem lehetséges, a megoldás olyan numerikus módszerek alkalmazása árán érhető csak el, amely túlmutat a jelen dolgozat keretein. Ehelyett momentum módszert alkalmazunk a paraméterek becslésére, vagyis felírunk annyi összefüggést a momentumok és a paraméterek között, amelyekből már egyértelműen kifejezhetőek a paraméterek, majd az így adódó kifejezésekben a momentumokat a tapasztalati megfelelőjünkkel helyettesítjük. Mivel egy η paraméterű ζ Poisson-eloszlás generátorfüggvénye

$$G(z) = e^{\eta(z-1)},$$

azért az eloszlás $f_k = \mathbb{E}(\zeta(\zeta-1) \cdot \dots \cdot (\zeta-k+1))$ — az $f_0 = 1$ megállapodással — faktoriális momentumai könnyen meghatározhatók deriválás segítségével

$$f_k = \frac{\partial}{\partial z^k} [G(z)]_{z=1} = \eta^k.$$

Kevés utánagondolással látható, hogy a mi esetünkben a momentum módszerrel felírt összefüggések ekvivalensek a faktoriális momentumokra felírt egyenletekkel. Az i -edik faktoriális momentum:

$$\begin{aligned}
f_i &= \mathbb{E}(\zeta_{k+1}(\zeta_{k+1}-1) \cdot \dots \cdot (\zeta_{k+1}-i+1)) \\
&= \sum_{j=1}^r \mathbb{E}(\zeta_{k+1}(\zeta_{k+1}-1) \cdot \dots \cdot (\zeta_{k+1}-i+1) | \tau = \mu_j) \mathbb{P}(\tau = \mu_j) \\
&= \sum_{j=1}^r p_j \mu_j^i,
\end{aligned}$$

tehát a

$$\sum_{j=1}^r p_j \mu_j^i = f_i, i = 0, \dots, 2r-1$$

egyenletrendszert kell megoldanunk. Legyen $\vartheta = (\mu_1, \dots, \mu_r, p_1, \dots, p_r)$ és tegyük fel, hogy a megoldás $\vartheta = F(f_0, \dots, f_{2r-1})$, ekkor a ϑ becslése

$$\widehat{\vartheta} = F(\widehat{f}_0, \dots, \widehat{f}_{2r-1}),$$

ahol $\widehat{f}_0, \dots, \widehat{f}_{2r-1}$ az f_0, \dots, f_{2r-1} faktoriális momentumok ζ_1, \dots, ζ_k függvényében számított tapasztalati becslése. Az alábbiakban megmutatjuk, hogy miként számolható a $\widehat{\vartheta}$ becslés abban az esetben, amikor $r = 2$. Ekkor a következő egyenletrendszert kell megoldanunk:

$$\begin{aligned}
p_1 + p_2 &= 1 \\
p_1 \mu_1 + p_2 \mu_2 &= \widehat{f}_1 \\
p_1 \mu_1^2 + p_2 \mu_2^2 &= \widehat{f}_2 \\
p_1 \mu_1^3 + p_2 \mu_2^3 &= \widehat{f}_3,
\end{aligned}$$

ahol

$$\begin{aligned}\widehat{f}_1 &= \frac{1}{k} \sum_{i=1}^k \zeta_i = \bar{\zeta}, \\ \widehat{f}_2 &= \frac{1}{k} \sum_{i=1}^k \zeta_i(\zeta_i - 1) = \frac{1}{k} \sum_{i=1}^k \zeta_i^2 - \frac{1}{k} \sum_{i=1}^k \zeta_i,\end{aligned}$$

és

$$\widehat{f}_3 = \frac{1}{k} \sum_{i=1}^k \zeta_i(\zeta_i - 1)(\zeta_i - 2) = \frac{1}{k} \sum_{i=1}^k \zeta_i^3 - \frac{3}{k} \sum_{i=1}^k \zeta_i^2 - \frac{1}{k} \sum_{i=1}^k \zeta_i.$$

Az egyenletrendszer első két egyenletéből kapjuk, hogy

$$p_1 = \frac{\widehat{f}_1 - \mu_2}{\mu_1 - \mu_2} \text{ és } p_2 = \frac{\mu_1 - \widehat{f}_1}{\mu_1 - \mu_2}. \quad (4.1)$$

Írjunk, mindenütt $1 - p_1$ -et p_2 helyére, ekkor a rendszer így módosul:

$$\begin{aligned}p_1(\mu_1 - \mu_2) + \mu_2 &= \widehat{f}_1 \\ p_1(\mu_1^2 - \mu_2^2) + \mu_2^2 &= \widehat{f}_2 \\ p_1(\mu_1^3 - \mu_2^3) + \mu_2^3 &= \widehat{f}_3,\end{aligned} \quad (4.2)$$

A második egyenletbe p_1 kifejezését behelyettesítve $(\widehat{f}_1 - \mu_2)(\mu_1 + \mu_2) + \mu_2^2 = \widehat{f}_2$, azaz

$$(\widehat{f}_1 - \mu_2)\mu_1 = \widehat{f}_2 - \widehat{f}_1\mu_2 \quad (4.3)$$

adódik. A harmadik egyenlet bal oldala pedig

$$\begin{aligned}p_1(\mu_1^3 - \mu_2^3) + \mu_2^3 &= p_1(\mu_1 - \mu_2)(\mu_1^2 + \mu_1\mu_2 + \mu_2^2) + \mu_2^3 \\ &= (\widehat{f}_1 - \mu_2)(\mu_1^2 + \mu_1\mu_2 + \mu_2^2) + \mu_2^3 \\ &= (\widehat{f}_1 - \mu_2)\mu_1^2 + (\widehat{f}_2 - \widehat{f}_1\mu_2)\mu_2 + \widehat{f}_1\mu_2^2 \\ &= (\widehat{f}_1 - \mu_2)\mu_1^2 + \widehat{f}_2\mu_2,\end{aligned}$$

vagyis

$$(\widehat{f}_1 - \mu_2)\mu_1^2 + \widehat{f}_2\mu_2 = \widehat{f}_3.$$

Szorozva az egyenlet mindkét oldalát $(\widehat{f}_1 - \mu_2)$ -vel, azt kapjuk, hogy

$$(\widehat{f}_2 - \widehat{f}_1\mu_2)^2 + (\widehat{f}_1 - \mu_2)\widehat{f}_2\mu_2 = \widehat{f}_3(\widehat{f}_1 - \mu_2),$$

majd ezt rendezve végül az

$$(\widehat{f}_1^2 - \widehat{f}_2)\mu_2^2 + (\widehat{f}_3 - \widehat{f}_1\widehat{f}_2)\mu_2 + \widehat{f}_2^2 - \widehat{f}_1\widehat{f}_3 = 0 \quad (4.4)$$

másodfokú egyenlethez jutunk. Ha μ_2 kielégíti ezt az egyenletet, akkor μ_1 is, hiszen az (4.3) összefüggés miatt

$$\mu_1 + \mu_2 = \frac{\widehat{f}_2 - \widehat{f}_1\mu_2}{(\widehat{f}_1 - \mu_2)} + \mu_2 = \frac{\widehat{f}_2 - \widehat{f}_1\mu_2 + (\widehat{f}_1 - \mu_2)\mu_2}{(\widehat{f}_1 - \mu_2)} = \frac{\widehat{f}_2 - \mu_2^2}{(\widehat{f}_1 - \mu_2)} = -\frac{\widehat{f}_3 - \widehat{f}_1\widehat{f}_2}{(\widehat{f}_1^2 - \widehat{f}_2)},$$

és

$$\mu_1\mu_2 = \frac{(\widehat{f}_2 - \widehat{f}_1\mu_2)\mu_2}{(\widehat{f}_1 - \mu_2)} = \frac{\widehat{f}_2^2 - \widehat{f}_1\widehat{f}_3}{(\widehat{f}_1^2 - \widehat{f}_2)},$$

ahol az utolsó egyenlőség mindkét átalakítás-sorozat esetén csupán (4.4) átrendezése. Ez azt jelenti, hogy μ_1 és μ_2 kielégítik az

$$(\widehat{f}_1^2 - \widehat{f}_2)x^2 + (\widehat{f}_3 - \widehat{f}_1\widehat{f}_2)x + \widehat{f}_2^2 - \widehat{f}_1\widehat{f}_3 = 0 \quad (4.5)$$

másodfokú egyenlethez tartozó Viète-formulákat, vagyis a másik gyök csakis μ_1 lehet. Összefoglalva tehát, μ_1 és μ_2 becslései a (4.5) másodfokú egyenlet két gyökeként adódnak, ami után már p_1 és p_2 a (4.1) kifejezések alapján egyértelműen meghatározott.

4.3. Numerikus eredmények

Numerikus példákon teszteltük az EM algoritmust, illetve a korábbi alfejezetben leírt becslési modellek alkalmazásának hatékonyságát. A most következő szakaszban ezen vizsgálódásaink eredményéről számolunk be.

Ryden a már korábban említett 1996-os cikkében két különböző 2 rendű MMPP esetén közölt numerikus eredményeket az általa prezentált algoritmus segítségével. Ezek rendre az általa 2A, illetve 2B-vel jelölt esetek. Az 4.1 táblázat mutatja az ezeknek megfelelő paramétereket. Ezek jól láthatóan úgy lettek megválasztva, hogy az egyik esetben λ_1 és λ_2 jól elkülöníthető legyen egymástól, míg a másikban kevésbé.

| Eset | q_{12} | q_{21} | λ_1 | λ_2 |
|------|----------|----------|-------------|-------------|
| 2A | 10 | 1 | 100 | 10 |
| 2B | 5 | 2 | 100 | 50 |

4.1. táblázat. A Rydén által vizsgált 2A és 2B eset

Mi újrainplementáltuk az algoritmust a [21]-ben ajánlott javításoknak - és a 3.3 alfejezetben leírtaknak - megfelelően, tehát alkalmunk nyílik összevetni az eredményeket a Rydén által kapottakkal. Az összehasonlítás persze nem teljesen korrekt, mivel más trajektórából készítettük a becslést és más kiindulási becslést használtunk. Mi különösen az előző alfejezetben felvetett problémára fektettünk hangsúlyt, ezért a mintagenerálásnál a következő módon jártunk el. A korábban tárgyalt azonos hosszúsági időperiódusok (amiket az egyszerűség kedvéért napoknak kereszteltünk el) hosszát egységnyinek választottuk, majd feltettük, hogy $k = 365$. Persze k egész más érték is lehetne - ügyelve arra, hogy a köztes időkből vett minta kellően nagy maradjon -, mi most önkényesen választottuk ezt az értéket, azt az elképzelt esetet szemlélítve, amikor a biztosítónak egy évnyi megfigyelése van napi kárszámokból (vagy a jobbik esetben a kárfolyamat teljes trajektóriájából). Ez az eljárás a 2A esetben 6828 dimenziós, míg a 2B esetben egy jóval nagyobb 23483 dimenziós köztes időkből vett mintát eredményezett. Rydén a becsléseit 5000 elemű mintából számította. A 4.2 táblázat az első néhány nap szimulált kárszámait mutatja.

| Nap | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----|----|----|----|----|----|----|----|----|----|----|
| 2A | 18 | 14 | 16 | 9 | 10 | 9 | 41 | 17 | 24 | 30 |
| 2B | 68 | 72 | 56 | 71 | 77 | 72 | 65 | 56 | 47 | 74 |

4.2. táblázat. Szimulált kárszámok a 2A és 2B esetben

Bár mindkét esetben rosszabb kezdeti értékről indítottuk az EM algoritmust, mint Rydén, az általunk kapott becslés nagyjából ugyanolyan jól közelítette a valós paramétereket, sőt, némelyik paraméter esetén jobb is volt annál. A kapott becsléseket a 4.3 táblázatban foglaltuk össze.

A továbbiakban e két példán keresztül bemutatjuk, hogy milyen értékeket kaptunk az előző alfejezetben ismertett előrejelzésekre a szintén ott tárgyalt teljes trajektóriából számoló (tehát a Rydén-féle EM algoritmust használó), illetve az egyszerűsített, f.a.e. Poisson illetve kevert Poisson kárszámot feltételező modellek esetén. Először térjünk rá az $\mathbb{E}(\zeta_{k+1})$ előrejelzésre (emlékezve, hogy esetünkben $k = 365$). Az ide vonatkozó eredmények az 4.4 táblázatban láthatók. A "Valós paraméter" fejlécű oszlopban az igazi paraméterek szerint számított érték van, míg az "MMPP", "Tiszta Poisson", illetve "Kevert Poisson" fejlécű oszlopokban rendre a már korábban említett módon becsült paraméterekből számított értékek állnak. Itt persze a gyakorlati életben nyilván csak a

| Eset | Rydén becslése | | | | Új becslés | | | |
|------|----------------|----------|-------------|-------------|------------|----------|-------------|-------------|
| | q_{12} | q_{21} | λ_1 | λ_2 | q_{12} | q_{21} | λ_1 | λ_2 |
| 2A | 9.61 | 0.921 | 96.0 | 9.63 | 10.733 | 1.095 | 99.278 | 10.102 |
| 2B | 5.94 | 2.26 | 105 | 50.8 | 4.930 | 2.148 | 99.457 | 49.041 |

4.3. táblázat. Rydén becslése és a mostani becslés

valamilyen módon egészekre kerekített értékek az érdekesek — tekintve, hogy a kárszám csak egész értékű lehet —, most azonban kiírtuk a tizedesjegyeket is, hogy jobban összehasonlíthatóak legyenek a különböző modellek.

| Előrejelzés | Valós paraméter | MMPP | Tiszta Poisson | Kevert Poisson |
|-------------|-----------------|--------|----------------|----------------|
| 2A | 18.182 | 18.360 | 18.348 | 18.348 |
| 2B | 64.286 | 64.341 | 64.337 | 64.337 |

4.4. táblázat. Az $\mathbb{E}(\zeta_{k+1})$ előrejelzések a három különböző modell esetén

Azt látjuk tehát, hogy a 3 különböző modell esetén a becült előrejelzés csak minimálisan tér el a valós paraméterek ismeretében számítottól, a tisztán Poisson és kevert Poisson modellek esetén pedig teljesen megegyezik összhangban az előző alfejezet végén tett megjegyzésünkkel. Ebből az a sejtés adódik, hogy amennyiben az $\mathbb{E}(\zeta_{k+1})$ előrejelzésre vagyunk kíváncsiak, úgy elegendő csupán a napi kárszámok ismerete és az egyszerűsített Poisson modell használata.

Vizsgáljuk meg, hogy milyen becslések adódtak a kvantilisekre az egyes modellekből. Kétféle megbízhatósági szint - 0.95 és 0.99 - esetére határoztuk meg a becsléseket. Az eredmények az 4.5 és 4.6 táblákban láthatók.

| Szint=0.95 | Valós érték | MMPP | Tiszta Poisson | Kevert Poisson |
|------------|-------------|------|----------------|----------------|
| 2A | 42 | 42 | 27 | 43 |
| 2B | 90 | 91 | 79 | 90 |

4.5. táblázat. A 0.95-os megbízhatósági szinthez tartozó kvantilisbecslések

Látható, hogy a teljes trajektória ismeretében számított becslés nagyon jól közelít. A tisztán Poisson modell mindkét megbízhatósági szint esetén teljesen rossz eredményt ad, ráadásul számkra kedvezőtlen irányban, ugyanis alábecsüli a kvantilis értéket. Ezzel szemben a kevert Poisson modell 0.95-os megbízhatósági szinten meglepően jó eredményt ad. A magasabb megbízhatósági szinten azonban már ez a modell is jóval nagyobbat téved. Ez alapján az a sejtés fogalmazható meg, hogy a kvantilis becslése esetén az egyszerűsített Poisson modell alkalmazása teljesen elvethető, ugyanis az közel sem szolgáltatja az általunk elvárt eredményt. A kevert Poisson modell azonban alkalmas tűnik arra, hogy egyszerű eljárással viszonylag pontos értéket kapjunk vele.

Az imént említett 2 fő példán kívül még további 52 kísérletet végeztünk. A valós paramétereket véletlenszerűen választottuk, ügyelve arra, hogy az így keletkező minta „élethű” legyen, azaz a napi kárszámok alakulása reális értéket tükrözzön. Az egyes futási eredmények számszerű értékeit a függelékben található táblázatokban foglaltuk össze. Az alábbiakban ismertetjük a tesztelés során összegyűlt tapasztalatainkat.

A Ryden-féle becslés pontosságáról azt tapasztaltuk, hogy az euklideszi norma szerint vett átlagos eltérés a valós illetve a becült paraméterek között 3.04, ami 3.26%-a a valós paraméterek átlagos euklideszi normájának. A futások részletes jegyzőkönyve megtalálható az A.1, A.2, A.3

| Szint=0.99 | Valós érték | MMPP | Tiszta Poisson | Kevert Poisson |
|------------|-------------|------|----------------|----------------|
| 2A | 58 | 56 | 30 | 49 |
| 2B | 102 | 102 | 85 | 98 |

4.6. táblázat. A 0.99-os megbízhatósági szinthez tartozó kvantilisbecslések

és A.4 táblázatokban. Az $\mathbb{E}(\zeta_{k+1})$ értékekre kapott becslések esetén a közelítő Poisson modellből származó becslés egyetlen eset kivételével mindig jobb eredményt adott, mint a jóval kifinomultabb MMPP modellből a teljes trajektória ismeretében számított becslés. Ez utóbbi esetén az eltérések 0.02% és 7.07% százalék között mozogtak, az átlagos eltérés pedig 1,84% volt. Ugyanezen százalékok a Poisson modell esetén 0.000023%, 0.2% és 0.05%. A 7.07% százalékos eltérést produkáló MMPP futás esetén a Poisson futás eredményének eltérése csupán 0.02% volt, vagyis még mindig legfeljebb akkora, mint az MMPP legjobb esetében. A részletes futási eredményeket az A.5 és az A.6 táblázat tartalmazza. Összességében elmondható tehát, hogy a Poisson közelítés igen jól teljesít és a megvizsgált esetekben a teljes trajektóriából származó plusz információ sem tudott javítani az így kapott becslésen.

A 95%-os megbízhatósági szinthez tartozó kvantilis vizsgálatok az MMPP szerinti becslések legnagyobb tévedése 5 volt, de a legtöbb esetben pontosan eltalálták a valós értéket. A hibák a 4.7 táblázatban összefoglalt eloszlás szerint alakultak.

| Eltérés | 0 | 1 | 2 | 3 | 4 | 5 |
|---------------------------|-------|-------|-------|------|----|------|
| A futások számának %-ában | 46.3% | 31.5% | 11.1% | 7.4% | 0% | 3.7% |

4.7. táblázat. Az MMPP becslésből számított kvantilis eltéréseinek a gyakorisága

Elmondható tehát, hogy az esetek 96%-ában az így becsült kvantilis legfeljebb 3-mal tértek el a tényleges kvantilistól. A kevert Poisson modellből számított kvantilisbecslések csupán 3 futás esetén volt az eltérés nagyobb, mint 5, ezek értéke azonban kiugróan magas volt: két ízben 7, illetve egyszer 14. Ebben az esetben szintén teljesült az, hogy 0 eltérésnek a legnagyobb a gyakorisága, összehasonlításképpen közöljük az 4.7 táblázatnak megfelelő összegzést a kevert Poisson esetben.

| Eltérés | 0 | 1 | 2 | 3 | 4 | 5 | 5< |
|---------|-------|-------|-------|------|------|------|------|
| % | 38.9% | 25.9% | 16.7% | 5.6% | 5.6% | 1.9% | 5.9% |

4.8. táblázat. A kevert Poisson becslésből számított kvantilis eltéréseinek a gyakorisága

Tehát a kevert Poisson modell pontossága elmarad a szofisztikáltabb MMPP modelltől, bár az esetek 87%-ában nem volt nagyobb a hiba, mint 3, ami azért még mindig kellően magas arány. Persze hátrányként jegyezhető meg, hogy azokban az esetekben, amikor ez a hiba meghaladta az 5-öt, a tévedés mértéke nagyon nagy volt. Ezt valamelyest enyhíti az a tény, hogy a kiugróan nagy hibák pozitív irányúak voltak, azaz nem becsülték alul a lehetséges legnagyobb kárt.

Ezzel ellentétben - ahogy az a korábbi 2 fő példa esetén már körvonalazódott - a tisztán Poisson közelítés a megvizsgált esetek közül egyszer sem találta el pontosan a valós kvantilis értéket, és a hibák szórása is igen nagy volt, a legkisebb érték 2, a legnagyobb 39 volt. Ezen túlmenően az eltérés mindig negatív irányú volt, azaz alábecsültük a valós kvantilis értékeket, aminek a hatása végzetes lehet a biztosítótársaság számára, ha eszerint képzeli a minimális szavatoló tőke szükségletét. Összehasonlításképpen megemlíjtjük, hogy itt csak az esetek 24%-ában volt a hiba mértéke legfeljebb 3, ami lényegesen elmarad az előbbi két modell eredményétől. Itt az átlagos eltérés 10.79

volt. A futások részletes jegyzőkönyvét az A.7 és az A.8 táblázat tartalmazza.

Végezetül közöljük ugyanezeket az adatokat a 99%-os megbízhatósági szint esetére is. Az MMPP modell esetén az eltérési gyakoriságok ebben az esetben a 4.9 táblázatban összefoglaltak szerint alakultak.

| Eltérés | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---------|-----|-----|-------|------|------|------|----|------|
| % | 37% | 37% | 14.8% | 3.7% | 3.7% | 1.9% | 0% | 1.9% |

4.9. táblázat. MMPP eltérések gyakorisága 0.99-es mb. szinten

Ez már nem olyan jó eredmény, mint az alacsonyabb megbízhatósági szint esetén, de azért az esetek túlnyomó részében itt is csak 0 vagy 1 hiba volt. Legfeljebb 3 hiba volt az esetek 93%-ában. Sokkal szembetűnőbb a romlás a kevert Poisson modell eredményében, itt ugyanis meglehetősen nagy tévedések is előfordultak, a legrosszabb esetben 27-tel kisebb volt a kevert Poisson becslésből számított kvantilis, mint a valós érték. Itt csak egyetlen egy olyan eset volt, amikor a becült kvantilis pontos volt és csak valamivel több, mint az esetek felében (52%) volt a tévedés mértéke legfeljebb 3. Szemléltetésképpen közöljük a gyakoriságok táblázatát.

| Eltérés | 0 | 1 | 2 | 3 | 4 | 5 | 5< |
|---------|------|-------|------|------|------|-------|-------|
| % | 1.9% | 24.1% | 4.8% | 9.3% | 7.4% | 11.1% | 31.5% |

4.10. táblázat. Kevert Poisson gyakoriságok 0.99-es mb. szinten

A tisztán Poisson modell ismét nem került közel a valós értékekhez, az átlagos eltérés a valós és becült értékek között 18.07 volt, a legkisebb eltérés 4, a legnagyobb 65. Ennél a legnagyobb hibát eredményező futásnál az MMPP becslés hibája 0 volt, míg a kevert Poisson-é a meglehetősen nagy 27 érték. Továbbra is fenáll, hogy a tisztán Poisson modell végig alábecsülte az értékeket. A további részletek megtalálhatók az A.9 és az A.10 táblázatokban.

Összességében elmondható, hogy a tapasztalatok alapján az $\mathbb{E}(\zeta_{k+1})$ becslései közül a legjobb eredményt az egyszerűsített Poisson modell alkalmazása adja, vagyis a napi kárszámok mintaát-laga. Ezzel szemben a Poisson modell alkalmazása téves eredményeket szolgáltat mindkét vizsgált megbízhatósági szint melletti kvantilisbecslésekre. A kevert Poisson becslés az esetek túlnyomó részében megfelelő hibahatáron belül mozgott a 95%-os megbízhatósági szinten, az egyenletesen jó közelítésben azonban elvétve akadt egy-két kiugróan rossz érték, ami miatt ezen modell alkalmazása magában rejt némi bizonytalansági tényezőt. Ráadásul a megbízhatósági szint növelésével jelentősen romlottak a hibastatisztikák. Tehát a kísérletek alapján azt sejtjük, hogy a kevert Poisson közelítés alkalmazásának van létjogosultsága az alacsonyabb megbízhatósági szint esetén, 0.99%-os megbízhatósági szinten azonban már csak korlátozottan javasolt a használata. Úgy tűnik, hogy a kvantilisbecslések már jóval érzékenyebbek arra, hogy az egész trajektóriából számítsuk a becslést, vagy csupán napi kárszámokra szorítkozva, tehát amennyiben rendelkezésünkre áll az egész tra-jektóriából - vagyis a köztes időkből - vett minta, úgy akkor kapjuk a legmegfelelőbb eredményt, ha az - egyszerűsített modellekhez képest jóval bonyolultabb - EM algoritmus által megbecsült paraméterekből számítsuk becsléseinket.

5. fejezet

Záró megjegyzések

5.1. Összegzés

A dolgozat során ismertettük a Markov-modulált Poisson folyamat fő tulajdonságait. Áttekintettük a paramétereinek becslésére szolgáló Rydén által bevezetett EM algoritmust, majd módosítottuk és implementáltuk azt a Matlab programozási környezetben a W.J.J. Roberts et. al. által eszközölt új skálázási technikának és a mátrix exponenciális függvényekhez kapcsolódó integrálok kiértékelésében elért javításoknak megfelelően. Így már egy egyszerűen programozható, hatékony algoritmushoz jutottunk. Kiegészítettük az algoritmust egy k-means klaszterezésen alapuló, kiindulási becslés előállítására alkalmas eljárással. Ezután numerikus példákat vizsgáltunk egy biztossítmatematikai feladaton keresztül, melynek során C. Van Loan egyik lemmájának analógiájára konstruáltunk egy algoritmust az MMPP modell esetén egy rögzített hosszúságú intervallumba eső megfigyelések számának az eloszlása által meghatározott, különböző megbízhatósági szintekhez tartozó kvantilis értékek numerikus számítására.

5.2. További nyitott kérdések

Bár a numerikus példák alapján a Rydén-féle EM algoritmus efficiens eszköznek bizonyult az MMPP-k becslésére, továbbra is nyitva maradt kérdés az algoritmus konvergencia-tulajdonságainak egzakt vizsgálata, mellyel se [23], se [21] nem foglalkozik. Annak ellenére, hogy az általunk implementált EM algoritmus működik tetszőleges rendű MMPP esetén, mi mégis csak 2-állapotú eseteket vizsgáltunk. További vizsgálódásokra volna szükség a magasabb dimenziós esetben is. A numerikus eseteken túlmenően szükség volna a 4. fejezet 2. alfejezetében leírt kvantilisbecslések korrekt matematikai érzékenységvizsgálatára, meg kellene vizsgálni, hogy mekkorát változik a kvantilis, a paraméterek kis változtatásának hatására. Amikor a napi kárszámokból álló $(\zeta_1, \dots, \zeta_k)$ elemű mintából becslünk, akkor közelítő módszereket alkalmaztunk. Ehelyett korrektebb lenne a $(\zeta_1, \dots, \zeta_k)$ együttes sűrűségfüggvényének felírása, s abból a loglikelihood-becslés meghatározása. Ez azonban a következő problémát veti fel. Az együttes sűrűségfüggvény:

$$\begin{aligned} f_{(\zeta_1, \dots, \zeta_k)}(z_1, \dots, z_k) &= \mathbb{P}(\zeta_1 = z_1, \dots, \zeta_k = z_k) \\ &= \mathbb{E}(\mathbb{P}(\zeta_1 = z_1, \dots, \zeta_k = z_k | \rho_u, u \geq 0)), \end{aligned}$$

ahol ρ_t az MMPP intenzitásfolyamata. Egységnyi hosszúságú napokkal számolva ekkor $\zeta_i = N((i-1, i])$, ahol N jelöli a szóban forgó MMPP-t. Az előző átalakításban kihasználva a növekmények feltételes függetlenségét, azt kapjuk, hogy

$$\begin{aligned} f_{(\zeta_1, \dots, \zeta_k)}(z_1, \dots, z_k) &= \mathbb{E} \left(\prod_{j=1}^k \mathbb{P}(N((j-1, j]) = z_j | \rho_u, u \geq 0) \right) \\ &= \mathbb{E} \left(\prod_{j=1}^k \frac{(\rho_j - \rho_{j-1})^{z_j}}{z_j!} e^{-(\rho_j - \rho_{j-1})} \right) \\ &= \mathbb{E} \left(\prod_{j=1}^k \frac{(\rho_j - \rho_{j-1})^{z_j}}{z_j!} e^{-(\rho_j - \rho_{j-1})} \right) \\ &= \mathbb{E} \left(e^{-\rho_k} \prod_{j=1}^k \frac{(\rho_j - \rho_{j-1})^{z_j}}{z_j!} \right), \end{aligned}$$

ahol, mint tudjuk $\rho_j - \rho_{j-1} = \int_{j-1+}^j \lambda(\xi_s) ds$, ahol ξ_s a mögöttes Markov folyamat. Ahhoz, hogy meghatározzuk a fenti várható értéket, a $(\rho_1, \rho_2, \dots, \rho_k)$ változók együttes eloszlását kellene felírni, ami nem tűnik túl könnyű feladatnak.

5.3. Köszönetnyilvánítás

Szeretném hálámat kifejezni témavezetőmnek, Arató Miklósnak, aki számomra e kutatási területet kijelölte és útmutatásával, illetve hasznos tanácsaival segített munkám sikeres elvégzésében. További köszönet illeti Fodor Erzsébetet, aki időt és fáradságot nem kímélve gondosan átnézte a dolgozatomat - nagy szerepet vállalva ezzel a felmerülő nyelvtani hibák kiküszöbölésében -, valamint felbecsülhetetlen értékű számítógép-használati lehetőség biztosításával is hozzájárult a bemutatott futási eredmények elkészültéhez.

A. Függelék

A futási eredmények összefoglaló táblázatai

A.1. A valós paraméterek és az EM algoritmus által kapott becslés

| Futás | Mintaelemszám | Valós paraméterek | | | | Kiinduló becslés | | | |
|-------|---------------|-------------------|----------|-------------|-------------|------------------|----------|-------------|-------------|
| | | q_{12} | q_{21} | λ_1 | λ_1 | q_{12} | q_{21} | λ_1 | λ_1 |
| 2A | 6828 | 10 | 1 | 100 | 10 | 5.641 | 3.407 | 38.287 | 4.880 |
| 2B | 23483 | 5 | 2 | 100 | 50 | 17.699 | 17.088 | 107.055 | 21.921 |
| 1 | 8100 | 2.2 | 9.7 | 26.3 | 3.8 | 4.422 | 5.408 | 34.548 | 6.666 |
| 2 | 4766 | 10 | 1 | 50 | 10 | 4.158 | 3.335 | 23.160 | 4.294 |
| 3 | 5953 | 5 | 1 | 50 | 10 | 4.360 | 3.405 | 29.812 | 4.702 |
| 4 | 5488 | 5 | 1 | 40 | 10 | 3.813 | 3.365 | 26.159 | 4.504 |
| 5 | 5524 | 3 | 1 | 30 | 10 | 4.067 | 3.594 | 26.246 | 4.793 |
| 7 | 5934 | 2 | 1 | 30 | 10 | 3.140 | 3.370 | 26.691 | 4.511 |
| 8 | 6537 | 50 | 5 | 100 | 10 | 6.179 | 4.179 | 34.821 | 5.494 |
| 9 | 6371 | 55 | 5 | 100 | 10 | 6.290 | 4.207 | 34.036 | 5.627 |
| 10 | 6343 | 60 | 5 | 100 | 10 | 6.021 | 4.154 | 33.853 | 5.449 |
| 11 | 5946 | 65 | 5 | 100 | 10 | 5.085 | 4.020 | 29.886 | 5.191 |
| 12 | 5232 | 70 | 4 | 90 | 10 | 4.760 | 3.828 | 25.965 | 4.879 |
| 13 | 5476 | 60 | 4 | 90 | 10 | 5.839 | 3.868 | 29.340 | 5.092 |
| 14 | 5020 | 65 | 4 | 80 | 10 | 5.227 | 3.717 | 25.586 | 4.835 |
| 15 | 7096 | 60 | 5 | 80 | 15 | 6.724 | 5.285 | 34.796 | 6.927 |
| 16 | 7940 | 10 | 1 | 100 | 15 | 6.789 | 4.769 | 41.527 | 6.478 |
| 17 | 10431 | 10 | 1 | 100 | 20 | 7.888 | 6.571 | 50.927 | 8.860 |
| 18 | 7478 | 10 | 1 | 125 | 10 | 5.494 | 3.253 | 43.500 | 4.771 |
| 19 | 8053 | 10 | 1 | 150 | 10 | 7.281 | 3.456 | 50.903 | 5.163 |
| 20 | 8479 | 10 | 1 | 175 | 10 | 5.529 | 3.233 | 50.802 | 4.721 |
| 21 | 10232 | 10 | 1 | 200 | 10 | 6.314 | 3.254 | 64.513 | 4.895 |

A.1. táblázat. A becslendő paraméterek és a kiinduló becslések

| Futás | Mintaelemszám | Valós paraméterek | | | | Kiinduló becslés | | | |
|-------|---------------|-------------------|----------|-------------|-------------|------------------|----------|-------------|-------------|
| | | q_{12} | q_{21} | λ_1 | λ_1 | q_{12} | q_{21} | λ_1 | λ_1 |
| 22 | 9318 | 10 | 2 | 100 | 10 | 5.769 | 3.754 | 52.542 | 5.428 |
| 23 | 5499 | 15 | 1 | 100 | 10 | 5.461 | 3.451 | 30.048 | 4.795 |
| 24 | 5072 | 20 | 1 | 100 | 10 | 4.285 | 3.144 | 26.379 | 4.263 |
| 25 | 9452 | 5 | 1 | 100 | 10 | 4.445 | 3.332 | 49.220 | 4.797 |
| 26 | 10439 | 5 | 1 | 100 | 15 | 6.503 | 4.909 | 53.498 | 7.051 |
| 27 | 7688 | 5 | 1 | 75 | 10 | 4.502 | 3.354 | 40.639 | 4.678 |
| 28 | 10495 | 5 | 1 | 125 | 10 | 5.620 | 3.330 | 61.084 | 5.091 |
| 29 | 6039 | 10 | 1 | 90 | 10 | 4.965 | 3.342 | 32.660 | 4.597 |
| 30 | 6071 | 10 | 1 | 80 | 10 | 5.366 | 3.424 | 33.209 | 4.809 |
| 31 | 5635 | 10 | 1 | 70 | 10 | 4.698 | 3.403 | 29.230 | 4.612 |
| 32 | 5390 | 10 | 1 | 60 | 10 | 5.465 | 3.540 | 29.061 | 4.907 |
| 33 | 5103 | 30 | 2 | 150 | 10 | 4.935 | 3.302 | 27.013 | 4.494 |
| 34 | 5927 | 20 | 1 | 150 | 10 | 5.800 | 3.207 | 35.229 | 4.677 |
| 35 | 9080 | 25 | 2 | 145 | 15 | 7.689 | 5.241 | 48.364 | 7.129 |
| 36 | 9643 | 15 | 1 | 130 | 20 | 8.311 | 6.451 | 47.957 | 8.527 |
| 37 | 5853 | 15 | 3 | 70 | 5 | 4.638 | 2.689 | 35.391 | 3.629 |
| 38 | 9252 | 10 | 3 | 80 | 10 | 5.163 | 3.949 | 48.224 | 5.419 |
| 39 | 10609 | 15 | 3 | 100 | 15 | 7.974 | 5.659 | 56.416 | 7.710 |
| 40 | 10819 | 10 | 3 | 80 | 15 | 6.287 | 5.279 | 53.672 | 7.267 |
| 41 | 12366 | 10 | 5 | 70 | 15 | 6.865 | 6.454 | 58.884 | 8.677 |
| 42 | 13107 | 10 | 7 | 60 | 20 | 8.105 | 8.088 | 59.801 | 10.791 |
| 43 | 10754 | 20 | 6 | 60 | 20 | 9.341 | 7.561 | 52.550 | 9.916 |
| 44 | 9763 | 25 | 6 | 70 | 15 | 7.795 | 6.306 | 48.484 | 8.194 |
| 45 | 8971 | 30 | 6 | 70 | 15 | 7.831 | 6.125 | 44.815 | 7.994 |
| 46 | 8512 | 30 | 5 | 70 | 15 | 7.140 | 5.702 | 42.502 | 7.504 |
| 47 | 6125 | 35 | 5 | 65 | 10 | 5.763 | 4.241 | 31.444 | 5.450 |
| 48 | 7478 | 30 | 4 | 65 | 15 | 7.466 | 5.464 | 37.856 | 7.234 |
| 49 | 10550 | 20 | 4 | 75 | 20 | 8.152 | 7.046 | 50.771 | 9.254 |
| 50 | 10225 | 20 | 5 | 60 | 20 | 7.940 | 7.241 | 48.296 | 9.340 |
| 51 | 9541 | 20 | 3 | 65 | 20 | 7.991 | 6.963 | 45.252 | 8.880 |
| 52 | 10804 | 20 | 2 | 70 | 25 | 9.080 | 8.278 | 49.781 | 10.558 |

A.2. táblázat. A becsülendő paraméterek és a kiinduló becslések (folytatás)

| Futás | Valós paraméterek | | | | EM becslés | | | |
|-------|-------------------|----------|-------------|-------------|------------|----------|-------------|-------------|
| | q_{12} | q_{21} | λ_1 | λ_1 | q_{12} | q_{21} | λ_1 | λ_1 |
| 2A | 10 | 1 | 100 | 10 | 10.733 | 1.095 | 99.278 | 10.102 |
| 2B | 5 | 2 | 100 | 50 | 4.930 | 2.148 | 99.457 | 49.041 |
| 1 | 2.2 | 9.7 | 26.3 | 3.8 | 2.053 | 8.348 | 26.310 | 5.461 |
| 2 | 10 | 1 | 50 | 10 | 10.570 | 1.041 | 45.980 | 9.826 |
| 3 | 5 | 1 | 50 | 10 | 5.644 | 1.074 | 49.898 | 9.944 |
| 4 | 5 | 1 | 40 | 10 | 5.739 | 1.061 | 41.569 | 10.121 |
| 5 | 3 | 1 | 30 | 10 | 2.874 | 1.053 | 29.123 | 10.022 |
| 6 | 1 | 1 | 20 | 10 | 1.335 | 1.045 | 20.449 | 10.513 |
| 7 | 2 | 1 | 30 | 10 | 1.966 | 0.997 | 29.579 | 9.502 |
| 8 | 50 | 5 | 100 | 10 | 55.794 | 5.315 | 102.427 | 9.862 |
| 9 | 55 | 5 | 100 | 10 | 62.465 | 6.101 | 95.444 | 9.844 |
| 10 | 60 | 5 | 100 | 10 | 62.678 | 5.652 | 103.602 | 9.608 |
| 11 | 65 | 5 | 100 | 10 | 74.598 | 5.004 | 106.530 | 10.246 |
| 12 | 70 | 4 | 90 | 10 | 70.758 | 5.352 | 78.119 | 9.512 |
| 13 | 60 | 4 | 90 | 10 | 60.706 | 4.184 | 91.077 | 9.761 |
| 14 | 65 | 4 | 80 | 10 | 63.104 | 3.092 | 87.467 | 10.145 |
| 15 | 60 | 5 | 80 | 15 | 58.291 | 5.076 | 71.170 | 14.951 |
| 16 | 10 | 1 | 100 | 15 | 12.554 | 1.083 | 103.653 | 14.704 |
| 17 | 10 | 1 | 100 | 20 | 9.074 | 1.012 | 99.493 | 20.694 |
| 18 | 10 | 1 | 125 | 10 | 9.725 | 0.987 | 124.496 | 9.947 |
| 19 | 10 | 1 | 150 | 10 | 9.946 | 0.958 | 147.670 | 10.009 |
| 20 | 10 | 1 | 175 | 10 | 11.092 | 0.962 | 177.055 | 9.888 |
| 21 | 10 | 1 | 200 | 10 | 10.218 | 1.102 | 197.928 | 9.711 |
| 22 | 10 | 2 | 100 | 10 | 9.646 | 1.943 | 102.143 | 10.077 |
| 23 | 15 | 1 | 100 | 10 | 17.092 | 0.971 | 101.196 | 10.176 |
| 24 | 20 | 1 | 100 | 10 | 20.551 | 0.996 | 100.011 | 9.750 |
| 25 | 5 | 1 | 100 | 10 | 4.970 | 1.053 | 100.234 | 10.202 |
| 26 | 5 | 1 | 100 | 15 | 5.142 | 0.992 | 98.231 | 15.178 |
| 27 | 5 | 1 | 75 | 10 | 5.107 | 1.076 | 74.611 | 9.788 |
| 28 | 5 | 1 | 125 | 10 | 4.945 | 0.957 | 126.361 | 9.930 |
| 29 | 10 | 1 | 90 | 10 | 10.085 | 0.927 | 88.062 | 9.978 |
| 30 | 10 | 1 | 80 | 10 | 10.406 | 1.042 | 82.846 | 10.006 |
| 31 | 10 | 1 | 70 | 10 | 11.414 | 1.028 | 72.281 | 10.341 |
| 32 | 10 | 1 | 60 | 10 | 11.212 | 1.041 | 63.726 | 10.231 |
| 33 | 30 | 2 | 150 | 10 | 19.278 | 0.532 | 151.836 | 10.177 |
| 34 | 20 | 1 | 150 | 10 | 20.643 | 1.058 | 144.530 | 9.650 |
| 35 | 25 | 2 | 145 | 15 | 25.910 | 2.091 | 146.129 | 15.091 |
| 36 | 15 | 1 | 130 | 20 | 17.087 | 0.992 | 129.806 | 20.418 |
| 37 | 15 | 3 | 70 | 5 | 15.710 | 3.063 | 71.107 | 5.304 |
| 38 | 10 | 3 | 80 | 10 | 10.183 | 2.884 | 80.160 | 9.829 |
| 39 | 15 | 3 | 100 | 15 | 15.094 | 3.122 | 97.627 | 14.890 |
| 40 | 10 | 3 | 80 | 15 | 9.896 | 2.999 | 80.088 | 14.363 |

A.3. táblázat. A becsülendő paraméterek és az EM algoritmus által kapott becslés

| Futás | Valós paraméterek | | | | EM becslés | | | |
|-------|-------------------|----------|-------------|-------------|------------|----------|-------------|-------------|
| | q_{12} | q_{21} | λ_1 | λ_1 | q_{12} | q_{21} | λ_1 | λ_1 |
| 41 | 10 | 5 | 70 | 15 | 10.470 | 5.269 | 70.948 | 15.232 |
| 42 | 10 | 7 | 60 | 20 | 9.749 | 7.175 | 59.275 | 18.720 |
| 43 | 20 | 6 | 60 | 20 | 22.396 | 8.116 | 59.556 | 18.557 |
| 44 | 25 | 6 | 70 | 15 | 24.131 | 6.585 | 69.926 | 14.973 |
| 45 | 30 | 6 | 70 | 15 | 30.715 | 5.565 | 74.287 | 15.577 |
| 46 | 30 | 5 | 70 | 15 | 29.833 | 5.355 | 70.803 | 14.802 |
| 47 | 35 | 5 | 65 | 10 | 31.168 | 4.005 | 64.055 | 10.718 |
| 48 | 30 | 4 | 65 | 15 | 31.987 | 3.900 | 66.252 | 14.914 |
| 49 | 20 | 4 | 75 | 20 | 20.601 | 4.224 | 74.448 | 19.563 |
| 50 | 20 | 5 | 60 | 20 | 18.984 | 5.438 | 57.875 | 19.463 |
| 51 | 20 | 3 | 65 | 20 | 19.159 | 2.857 | 64.783 | 20.379 |
| 52 | 20 | 2 | 70 | 25 | 17.048 | 2.370 | 62.820 | 24.989 |

A.4. táblázat. A becslendő paraméterek és az EM algoritmus által kapott becslés (folytatás)

A.2. A várható kárszám becslése

| Futás | Valós érték | Becsült értékek | | |
|-------|-------------|-----------------|----------------|-----------------------|
| | | MMPP modell | Poisson modell | Kevert Poisson modell |
| 2A | 18.182 | 18.360 | 18.348 | 18.348 |
| 2B | 64.286 | 64.341 | 64.337 | 64.337 |
| 1 | 22.140 | 22.195 | 22.192 | 22.192 |
| 2 | 13.636 | 13.066 | 13.058 | 13.058 |
| 3 | 16.667 | 16.330 | 16.310 | 16.310 |
| 4 | 15.000 | 15.030 | 15.036 | 15.036 |
| 5 | 15.000 | 15.144 | 15.134 | 15.134 |
| 6 | 15.000 | 14.874 | 14.863 | 14.863 |
| 7 | 16.667 | 16.259 | 16.258 | 16.258 |
| 8 | 18.182 | 17.913 | 17.910 | 17.910 |
| 9 | 17.500 | 17.461 | 17.455 | 17.455 |
| 10 | 16.923 | 17.383 | 17.378 | 17.378 |
| 11 | 16.429 | 16.298 | 16.290 | 16.290 |
| 12 | 14.324 | 14.337 | 14.334 | 14.334 |
| 13 | 15.000 | 15.003 | 15.003 | 15.003 |
| 14 | 14.058 | 13.757 | 13.753 | 13.753 |
| 15 | 20.000 | 19.454 | 19.441 | 19.441 |
| 16 | 22.727 | 21.770 | 21.753 | 21.753 |
| 17 | 27.273 | 28.599 | 28.578 | 28.578 |
| 18 | 20.455 | 20.498 | 20.488 | 20.488 |
| 19 | 22.727 | 22.099 | 22.063 | 22.063 |

A.5. táblázat. Az $\mathbb{E}(\zeta_{k+1})$ előrejelzések valós és becsült értékei

| Futás | Valós érték | Becsült értékek | | |
|-------|-------------|-----------------|----------------|-----------------------|
| | | MMPP modell | Poisson modell | Kevert Poisson modell |
| 20 | 25.000 | 23.234 | 23.230 | 23.230 |
| 21 | 27.273 | 28.041 | 28.033 | 28.033 |
| 22 | 25.000 | 25.512 | 25.529 | 25.529 |
| 23 | 15.625 | 15.069 | 15.066 | 15.066 |
| 24 | 14.286 | 13.924 | 13.896 | 13.896 |
| 25 | 25.000 | 25.939 | 25.893 | 25.893 |
| 26 | 29.167 | 28.607 | 28.600 | 28.600 |
| 27 | 20.833 | 21.071 | 21.063 | 21.063 |
| 28 | 29.167 | 28.806 | 28.753 | 28.753 |
| 29 | 17.273 | 16.551 | 16.545 | 16.545 |
| 30 | 16.364 | 16.638 | 16.633 | 16.633 |
| 31 | 15.455 | 15.457 | 15.438 | 15.438 |
| 32 | 14.545 | 14.775 | 14.767 | 14.767 |
| 33 | 13.415 | 13.984 | 13.981 | 13.981 |
| 34 | 16.667 | 16.225 | 16.238 | 16.238 |
| 35 | 24.630 | 24.879 | 24.877 | 24.877 |
| 36 | 26.875 | 26.419 | 26.419 | 26.419 |
| 37 | 15.833 | 16.040 | 16.036 | 16.036 |
| 38 | 26.154 | 25.350 | 25.348 | 25.348 |
| 39 | 29.167 | 29.070 | 29.066 | 29.066 |
| 40 | 30.000 | 29.648 | 29.641 | 29.641 |
| 41 | 33.333 | 33.884 | 33.879 | 33.879 |
| 42 | 36.471 | 35.913 | 35.910 | 35.910 |
| 43 | 29.231 | 29.463 | 29.463 | 29.463 |
| 44 | 25.645 | 26.754 | 26.748 | 26.748 |
| 45 | 24.167 | 24.583 | 24.578 | 24.578 |
| 46 | 22.857 | 23.325 | 23.321 | 23.321 |
| 47 | 16.875 | 16.791 | 16.781 | 16.781 |
| 48 | 20.882 | 20.494 | 20.488 | 20.488 |
| 49 | 29.167 | 28.902 | 28.904 | 28.904 |
| 50 | 28.000 | 28.016 | 28.014 | 28.014 |
| 51 | 25.870 | 26.142 | 26.140 | 26.140 |
| 52 | 29.091 | 29.605 | 29.600 | 29.600 |

A.6. táblázat. Az $\mathbb{E}(\zeta_{k+1})$ előrejelzések valós és becült értékei (folytatás)

A.3. Kvantilisbecslések

| Megbízhatósági szint: 95% | | Becsült értékek | | |
|---------------------------|-------------|-----------------|----------------|-----------------------|
| Futás | Valós érték | MMPP modell | Poisson modell | Kevert Poisson modell |
| 2A | 42 | 42 | 27 | 43 |
| 2B | 90 | 91 | 79 | 90 |
| 1 | 33 | 33 | 31 | 32 |
| 2 | 26 | 24 | 20 | 25 |
| 3 | 35 | 34 | 24 | 34 |
| 4 | 30 | 30 | 23 | 30 |
| 5 | 29 | 29 | 23 | 29 |
| 6 | 25 | 25 | 22 | 25 |
| 7 | 32 | 32 | 24 | 32 |
| 8 | 31 | 30 | 26 | 31 |
| 9 | 30 | 29 | 26 | 30 |
| 10 | 29 | 29 | 25 | 29 |
| 11 | 28 | 27 | 24 | 27 |
| 12 | 24 | 24 | 22 | 24 |
| 13 | 25 | 25 | 23 | 25 |
| 14 | 28 | 23 | 21 | 24 |
| 15 | 30 | 29 | 28 | 30 |
| 16 | 46 | 43 | 31 | 50 |
| 17 | 49 | 52 | 39 | 51 |
| 18 | 50 | 51 | 29 | 54 |
| 19 | 59 | 57 | 31 | 59 |
| 20 | 67 | 62 | 32 | 72 |
| 21 | 75 | 76 | 38 | 89 |
| 22 | 53 | 55 | 35 | 53 |
| 23 | 34 | 31 | 23 | 32 |
| 24 | 29 | 28 | 21 | 30 |
| 25 | 64 | 65 | 36 | 71 |
| 26 | 66 | 64 | 39 | 66 |
| 27 | 50 | 50 | 30 | 49 |
| 28 | 78 | 78 | 39 | 77 |
| 29 | 39 | 37 | 24 | 40 |
| 30 | 36 | 36 | 25 | 39 |
| 31 | 32 | 32 | 23 | 32 |
| 32 | 29 | 29 | 22 | 31 |
| 33 | 30 | 31 | 21 | 37 |
| 34 | 38 | 36 | 24 | 40 |
| 35 | 46 | 46 | 34 | 49 |
| 36 | 49 | 46 | 36 | 49 |
| 37 | 33 | 33 | 24 | 33 |

A.7. táblázat. A 95%-os megbízhatósági szinthez tartozó kvantilisbecslések

| Megbízhatósági szint: 95% | | Becsült értékek | | |
|---------------------------|-------------|-----------------|----------------|-----------------------|
| Futás | Valós érték | MMPP modell | Poisson modell | Kevert Poisson modell |
| 38 | 50 | 49 | 35 | 48 |
| 39 | 52 | 51 | 39 | 51 |
| 40 | 53 | 53 | 40 | 50 |
| 41 | 54 | 54 | 45 | 52 |
| 42 | 53 | 52 | 47 | 51 |
| 43 | 43 | 43 | 40 | 44 |
| 44 | 40 | 41 | 37 | 40 |
| 45 | 37 | 38 | 34 | 38 |
| 46 | 36 | 36 | 33 | 38 |
| 47 | 28 | 28 | 25 | 26 |
| 48 | 33 | 32 | 29 | 33 |
| 49 | 44 | 44 | 39 | 43 |
| 50 | 41 | 41 | 38 | 41 |
| 51 | 39 | 39 | 36 | 40 |
| 52 | 42 | 42 | 40 | 42 |

A.8. táblázat. A 95%-os megbízhatósági szinthez tartozó kvantilisbecslések (folytatás)

| Megbízhatósági szint: 99% | | Becsült értékek | | |
|---------------------------|-------------|-----------------|----------------|-----------------------|
| Futás | Valós érték | MMPP modell | Poisson modell | Kevert Poisson modell |
| 2A | 58 | 56 | 30 | 49 |
| 2B | 102 | 102 | 85 | 98 |
| 1 | 37 | 37 | 35 | 36 |
| 2 | 34 | 31 | 23 | 29 |
| 3 | 46 | 44 | 27 | 40 |
| 4 | 38 | 38 | 26 | 35 |
| 5 | 35 | 35 | 26 | 33 |
| 6 | 29 | 29 | 26 | 29 |
| 7 | 38 | 37 | 27 | 37 |
| 8 | 37 | 37 | 29 | 35 |
| 9 | 36 | 35 | 29 | 35 |
| 10 | 34 | 35 | 29 | 33 |
| 11 | 33 | 32 | 27 | 31 |
| 12 | 29 | 28 | 25 | 27 |
| 13 | 30 | 31 | 26 | 29 |
| 14 | 28 | 28 | 24 | 29 |
| 15 | 35 | 34 | 31 | 31 |
| 16 | 61 | 56 | 34 | 58 |
| 17 | 63 | 67 | 43 | 58 |
| 18 | 70 | 71 | 33 | 62 |
| 19 | 82 | 80 | 35 | 66 |
| 20 | 95 | 88 | 36 | 84 |

A.9. táblázat. A 99%-os megbízhatósági szinthez tartozó kvantilisbecslések

| Megbízhatósági szint: 99% | | Becsült értékek | | |
|---------------------------|-------------|-----------------|----------------|-----------------------|
| Futás | Valós érték | MMPP modell | Poisson modell | Kevert Poisson modell |
| 21 | 107 | 107 | 42 | 80 |
| 22 | 68 | 70 | 39 | 59 |
| 23 | 45 | 42 | 26 | 38 |
| 24 | 38 | 38 | 24 | 35 |
| 25 | 85 | 86 | 39 | 80 |
| 26 | 86 | 84 | 43 | 73 |
| 27 | 65 | 65 | 33 | 50 |
| 28 | 105 | 106 | 43 | 85 |
| 29 | 53 | 51 | 28 | 47 |
| 30 | 48 | 49 | 28 | 45 |
| 31 | 43 | 42 | 26 | 38 |
| 32 | 38 | 38 | 25 | 36 |
| 33 | 43 | 45 | 24 | 45 |
| 34 | 52 | 50 | 27 | 47 |
| 35 | 58 | 58 | 38 | 56 |
| 36 | 63 | 59 | 40 | 56 |
| 37 | 43 | 43 | 27 | 37 |
| 38 | 62 | 61 | 39 | 54 |
| 39 | 64 | 63 | 43 | 57 |
| 40 | 64 | 64 | 44 | 56 |
| 41 | 63 | 63 | 49 | 58 |
| 42 | 60 | 59 | 52 | 57 |
| 43 | 49 | 49 | 44 | 50 |
| 44 | 47 | 48 | 40 | 45 |
| 45 | 44 | 44 | 38 | 43 |
| 46 | 42 | 43 | 36 | 43 |
| 47 | 34 | 34 | 28 | 30 |
| 48 | 38 | 37 | 33 | 39 |
| 49 | 52 | 51 | 43 | 48 |
| 50 | 47 | 47 | 42 | 46 |
| 51 | 45 | 46 | 40 | 46 |
| 52 | 48 | 48 | 44 | 47 |

A.10. táblázat. A 99%-os megbízhatósági szinthez tartozó kvantilisbecslések (folytatás)

Irodalomjegyzék

- [1] F. Bodon. *Adatbányászati algoritmusok*. <http://www.cs.bme.hu/bodon/magyar/adatbanyaszat/tanulmany/adatbanyaszat.pdf>, 2006.
- [2] E. Çinlar. Markov renewal theory. *Adv. Appl. Prob.*, 1:123–187, 1969.
- [3] D.R. Cox. Some statistical methods connected with series of events (with discussion). *J. R. Statist. Soc.*, B 17(6):129., 1955.
- [4] L. Czách. Mértékelmélet. *kézirat*, 2001.
- [5] D. Daley, D.J. és Vere-Jones. *An introduction to the theory of point processes*. Springer Series in Statistics. New York: Springer.
- [6] Wolfgang Fischer and Kathleen S. Meier-Hellstern. The markov-modulated poisson process (mmp) cookbook. *Perform. Eval.*, 18(2):149–171, 1993.
- [7] Rice-J.A. Fredkin, D.R. Maximum likelihood estimation and identification directly from single-channel recordings. *Proc. R. Soc. Lond.*, 249:125–132, 1992.
- [8] L.A. Freed, D. S.-Shepp. A poisson process whose rate is a hidden markov process. *Adv. Appl. Prob.*, 14:21–36, 1982.
- [9] M. Laczkovich. *Valós függvénytan*. ELTE egyetemi jegyzet, Budapest, 1995.
- [10] M.L. Leroux, B.G. Puterman. Maximum-penalized-likelihood estimation for independent and markov-dependent mixture models. *Biometrics*, 48:545–558, 1992.
- [11] Rabiner L.R. és Sondhi M.M. Levinson, S.E. An introduction to the application of the theory of probabilistic functions of a markov process in automatic speech recognition. *Bell Syst. Tech. J.*, 62:1035–1074, 1983.
- [12] P. Lévy. Systèmes semi-markoviens à au plus une infinité dénombrable d'états possibles. *Proc. Int. Congr. Math.*, 2:294., 1954.
- [13] P. Lévy. Regenerative stochastic processes. *Proc. Roy. Soc. (London)*, 232:6–31, 1955.
- [14] Gy. Mihaletzky. *Kockázati folyamatok*. ELTE Eötvös kiadó, 2001.
- [15] U. Naor, P. és Yechiali. Queueing problems with heterogenous arrivals and service. *Oper. Res.*, 19.
- [16] M.F. Neuts. A queue subject to extraneous phase changes. *Adv. Appl. Probab.*, 3.
- [17] M.F. Neuts. *Structured Stochastic Matrices of M/G/1-type and their Applications*. Marcel Dekker, New York, NY, 1989.
- [18] R. Pyke. Markov renewal processes: definitions and preliminary properties. *Annals of Mathematical Statistics*, 32:1231–1242, 1961.
- [19] R. Pyke. Markov renewal processes with finitely many states. *Annals of Mathematical Statistics*, 32:1242–1259, 1961.
- [20] R.-D. Reiss. *A course on point processes*. Springer-Verlag, 1993.

- [21] Ephraim Y. és Dieguez E. Roberts, W.J.J. On rydén's em algorithm for estimating mmpps. *IEEE SIGNAL PROCESSING LETTERS*, 13(6):373–376, 2006.
- [22] T. Rydén. Parameter estimation for markov-modulated poisson processes. *Stochastic Models*, 10(4), 1994.
- [23] T. Rydén. An em algorithm for estimation in markov-modulated poisson processes. *Comput. Stat. Data Anal.*, 21(4):431–447, 1996.
- [24] G.A.F. Seber. *Multivariate Observations*. Wiley, New York, 1984.
- [25] H. Spath. *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples*. Halsted Press, New York, 1985.
- [26] Matlab szoftver. <http://www.mathworks.com>.
- [27] L. Takács. Some investigations concerning recurrent stochastic processes of a certain type. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 3:115–128, 1954.
- [28] S. Karlin H. M. Taylor. *Sztochasztikus folyamatok*. Gondolat, Budapest, 1985.
- [29] C. Van Loan. Computing integrals involving the matrix exponential. *IEEE transactions on automatic control*, 23(3):395–404, 1978.
- [30] C. Van Loan, C.-Moller. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45:3–49, 2003.
- [31] D. Vere-Jones. The e-m algorithm for continuous-time point process examples. *kézirat*, 2005.
- [32] C.F.J. Wu. On the convergence properties of the em algorithm. *Ann. Statist.*, 11:95–103, 1992.