Felsőbb Matematika Informatikusoknak A,D és Villamosmérnököknek A,B
házi feladatok a „Sztochasztika 2" részhez
2013 ősz

Minden héten összesen egy pontot érnek a kitűzött feladatok.

1.HF: (Beadási határidő: 2013.11.11.)

HF 1.1 Pistike és Móricka a következő játékot játsszák: van két, ránézésre egyforma hatoldalú dobókockájuk, melyek közül az egyik szabályos, azaz $\frac{1}{6}-\frac{1}{6}$ valószínűséggel lesz felül az 1, 2, 3, 4, 5, 6 számok bármelyike, a másik viszont cinkelt: a 6-osnak $\frac{1}{2}$ a valószínűsége, a többi számnak $\frac{1}{10}-\frac{1}{10}$. Találomra elveszi az egyik kockát Pistike, a másikat Móricka, majd elkezdenek dobálni.

(a) Mekkora valószínűséggel választotta Pistike a cinkelt kockát, feltéve, hogy mindkét dobása 6-os lett?

(b) Jelölje X Móricka első dobásának értékét és Y Pistike első dobásának értékét. Mennyi $\mathrm{cov}(X, Y)$? (Tipp: használjunk teljes várható érték tételt.)

Megoldás: Jelölje A_1 azt az eseményt, hogy Pistike a cinkelt kockát választotta, A_2 pedig azt, hogy a szabályosat. Ekkor $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \frac{1}{2}$.

(a) Jelölje B azt az eseményt, hogy Pistike mindkét dobása 6-os. $\mathbb{P}(B|A_1) = \frac{1}{4}$, $\mathbb{P}(B|A_2) = \frac{1}{36}$, mi pedig a $\mathbb{P}(A_1|B)$ feltételes valószínűséget keressük. A Bayes tétel szerint

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(A_1)\mathbb{P}(B|A_1)}{\mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2)} = \frac{\frac{1}{2}\cdot\frac{1}{4}}{\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2}\cdot\frac{1}{36}} = \frac{9}{9+1} = \frac{9}{10}.$$

Megjegyzés: $\mathbb{P}(B|A_1)$ *azért* $\frac{1}{4}$, *mert Pistike két dobása külön-külön* $\frac{1}{2}$ ***feltételes*** *valószínűséggel* 6*-os (feltéve, hogy a cinkelt kockával dob), és a két dobás eredménye* ***feltételesen független*** *(megint csak feltéve, hogy a cinkelt kockával dob).* $\mathbb{P}(B|A_2) = \frac{1}{36}$ *ugyanilyen okból. Az viszont* ***nem igaz****, hogy a két dobás eredménye független lenne.*

(b) Jelölje U a szabályos, V pedig a cinkelt kockán dobott első számot. Ekkor

$$\begin{aligned}\mathbb{E}U &= \frac{1}{6}\cdot 1+\frac{1}{6}\cdot 2+\frac{1}{6}\cdot 3+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=\frac{7}{2},\\ \mathbb{E}V &= \frac{1}{10}\cdot 1+\frac{1}{10}\cdot 2+\frac{1}{10}\cdot 3+\frac{1}{10}\cdot 4+\frac{1}{10}\cdot 5+\frac{1}{2}\cdot 6=\frac{9}{2}.\end{aligned}$$

Egyúttal $\mathbb{E}(X|A_1) = \mathbb{E}(Y|A_2) = \mathbb{E}U$ és $\mathbb{E}(X|A_2) = \mathbb{E}(Y|A_1) = \mathbb{E}V$. Továbbá $\mathbb{E}(XY|A_1) = \mathbb{E}(XY|A_2) = \mathbb{E}(UV) = \mathbb{E}U\cdot\mathbb{E}V$ (megint a dobások feltételes függetlensége miatt). Így a teljes várható érték tételből az jön ki, hogy

$$\begin{aligned}\mathbb{E}X &= \mathbb{P}(A_1)\mathbb{E}(X|A_1)+\mathbb{P}(A_2)\mathbb{E}(X|A_2) = \frac{1}{2}\mathbb{E}U+\frac{1}{2}\mathbb{E}V=\frac{8}{2},\\ \mathbb{E}Y &= \mathbb{P}(A_1)\mathbb{E}(Y|A_1)+\mathbb{P}(A_2)\mathbb{E}(Y|A_2) = \frac{1}{2}\mathbb{E}V+\frac{1}{2}\mathbb{E}U=\frac{8}{2},\\ \mathbb{E}(XY) &= \mathbb{P}(A_1)\mathbb{E}(XY|A_1)+\mathbb{P}(A_2)\mathbb{E}(XY|A_2) = \frac{1}{2}\mathbb{E}(UV)+\frac{1}{2}\mathbb{E}(UV)=\frac{63}{4},\end{aligned}$$

vagyis

$$\mathrm{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y = -\frac{1}{4}.$$

Megjegyzés: megint jól látszik, hogy X és Y ***feltételesen függetlenek*** *(feltéve akár* A_1*-et, akár* A_2*-t), de nem függetlenek.*

2.HF: (Beadási határidő: 2013.11.18.)

HF 2.1 Egy 1000-oldalas könyvben 1500 sajtóhuba van, véletlenszerűen elszórva.

(a) Körülbelül mennyi annak a valószínűsége, hogy a 13-adik oldalon legalább 2 sajtóhuba van?

(b) Körülbelül mennyi annak a valószínűsége, hogy a 13-adik oldalon legalább 2, a 42-ediken pedig pontosan 2 sajtóhuba van?

(c) *Bónusz kérdés:* A sajtóhubáknak kb. $\frac{1}{3}$-a vesszőhiba (abban az értelemben, hogy minden sajtóhuba $\frac{1}{3}$ valószínűséggel vesszőhiba, a többitől függetlenül). Mennyi annak a valószínűsége, hogy a 13-adik oldalon legalább 2 vesszőhiba és pontosan 1 egyéb sajtóhuba van?

Megoldás: Az egyes oldalakra eső sajtóhubák száma Poisson eloszlással közelíthető, mivel sok sajtóhuba próbálkozik egymástól lényegében függetlenül, hogy pont oda essen, és ez mindegyiknek kicsi valószínűséggel sikerül. Sőt, az egyes oldalakon lévő sajtóhubák száma jó közelítéssel független, ugyanilyen megfontolásból. Ezek után csak a várható értékükre van szükség, ami persze az adott oldalszámra eső hubák átlagos száma.

(a) Jelölje X_{13} a 13-adik oldalra eső sajtóhubák számát. A fentiek alapján ez jó közelítéssel Poisson eloszlású $\lambda = \frac{1500}{1000} = 1.5$ paraméterrel, vagyis

$$\mathbb{P}(X_{13} = k) \approx e^{-\lambda}\frac{\lambda^k}{k!} = e^{-1.5}\frac{(1.5)^k}{k!} \quad (k = 0, 1, 2, \dots).$$

Emiatt

$$\mathbb{P}(X_{13} \geq 2) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \approx 1 - e^{-1.5} - e^{-1.5} \cdot 1.5 \approx 0.44.$$

(b) Legyen X_{42} a 42-edik oldalra eső sajtóhubák száma. A fentiek miatt X_{42} is $Poi(1.5)$ eloszlással közelíthető és X_{13}-tól jó közelítéssel független, vagyis

$$\mathbb{P}(X_{13} \geq 2 \text{ és } X_{42} = 2) \approx \mathbb{P}(X_{13} \geq 2)\mathbb{P}(X_{42} = 2) \approx 0.44 \cdot e^{-1.5}\frac{(1.5)^2}{2!} \approx 0.11.$$

(c) Sőtsőt, a vesszőhibák és az egyéb sajtóhubák száma jó közelítéssel külön-külön is Poisson eloszlású és egymástól független, ugyanilyen megfontolásból. Ezért ha Y_{13} a 13-adik oldalon lévő vesszőhibák száma, Z_{13} pedig a 13-adik oldalon lévő egyéb sajtóhubák száma, akkor jó közelítéssel $Y_{13} \sim Poi(\frac{1500 \cdot \frac{1}{3}}{1000})$, $Z_{13} \sim Poi(\frac{1500 \cdot \frac{2}{3}}{1000})$ és ezek függetlenek. Így

$$\mathbb{P}(Y_{13} \geq 2 \text{ és } Z_{13} = 1) \approx \left(1 - e^{-0.5}(1 + 0.5)\right)\left(e^{-1} \cdot 1\right) \approx 0.033.$$

HF 2.2 Egy vicc úgy terjed, hogy mindenki, aki meghallja, véletlen számú új embernek meséli el, éspedig 0, 1, 2 vagy 3 új embernek, rendre $p_0 = p$, $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{4}$ és $p_3 = \frac{1}{2} - p$ valószínűséggel, *az előzményektől függetlenül.*

A viccet Móricka találja ki, ő alkotja egyedül a nulladik generációt. Első generációnak nevezzük azokat, akinek Móricka maga meséli el a viccet, második generációnak azokat, akiknek az első generáció tagjai mesélik el, stb.

Jelölje Z_k a k-adik generáció tagjainak a számát ($k = 0, 1, 2, \dots$), N pedig a viccet megismerő emberek teljes számát (Mórickát is beleértve, vagyis $N = \sum_{k=0}^{\infty} Z_k$).

Válaszoljuk meg az alábbi kérdéseket

I. $p = \frac{1}{3}$ esetén,

II. $p = \frac{1}{6}$ esetén:

a.) Mi Z_2 generátorfüggvénye?

b.) Mennyi Z_{12} várható értéke?

c.) Mennyi a $\mathbb{P}(Z_3 = 0)$ valószínűség?

d.) Mennyi a valószínűsége annak, hogy a vicc terjedése előbb-utóbb megáll (vagyis hogy valamelyik generáció már üres)? *(Segítség: ha egy harmadfokú egyenletnek ismerjük egy gyökét, akkor a gyöktényezőt kiemelve a többi gyökre másodfokú egyenletet kapunk.)*

e.) Mennyi N várható értéke?

Megoldás: Z_n Galton-Watson elágazó folyamat, ahol az egylépéses utódszám-eloszlás a fenti $\mathbb{P}(k \text{ utód}) = p_k$ $(k = 0, 1, 2, 3)$. Jelöljük ennek várható értékét m-mel, generátorfüggvényét g-vel. Legyen továbbá Z_n várható érétke m_n, generátorfüggvénye g_n, a kihalás-valószínűségek pedig $r_n = \mathbb{P}(Z_n = 0) = g_n(0)$, $r = \mathbb{P}(\exists n : Z_n = 0) = \lim_{n\to\infty} r_n$.

I. $p = \frac{1}{3}$ esetén $m = \frac{1}{3} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{6} \cdot 3 = \frac{5}{4}$, $g(z) = \frac{1}{3} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{6}z^3$.

a.) $g_2(z) = g(g(z)) = \frac{1}{3} + \frac{1}{4}(\frac{1}{3} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{6}z^3) + \frac{1}{4}(\frac{1}{3} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{6}z^3)^2 + \frac{1}{6}(\frac{1}{3} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{6}z^3)^3$.

b.) $m_{12} = m^{12} = \left(\frac{5}{4}\right)^{12} \approx 14.55$.

c.) $r_0 = 0$; $r_1 = g(r_0) = g(0) = \frac{1}{3}$; $r_2 = g(r_1) = g(1/3) \approx 0.4506$; $r_3 = g(r_2) \approx g(0.4506) \approx 0.5120$.

d.) Mivel $m > 1$, a folyamat szuperkritikus, így a kihalás valószínűsége az $r = g(r)$ egyenlet egyetlen $[0, 1)$-beli gyöke. Vagyis

$$r = \frac{1}{3} + \frac{1}{4}r + \frac{1}{4}r^2 + \frac{1}{6}r^3.$$

Ezt nullára rendezve és 12-vel végigszorozva azt kapjuk, hogy

$$0 = 2r^3 + 3r^2 - 9r + 4.$$

Mivel a $g(1) = 1$ mindig teljesül, tudjuk, hogy ennek az egyenletnek $r = 1$ biztosan gyöke. (És mivel a folyamat szuperkritikus, azt is tudjuk, hogy mi nem ezt a gyököt keressük.) Vagyis a jobboldali polinomból ki lehet emelni $(r - 1)$-et, és az jön ki, hogy

$$0 = (r - 1)(2r^2 + 5r - 4),$$

tehát az 1-től különböző gyökökre

$$0 = 2r^2 + 5r - 4,$$

amiből $r = \frac{-5\pm\sqrt{57}}{4}$. Mivel $r \in [0, 1)$,

$$\mathbb{P}(\text{kihalás}) = r = \frac{-5 + \sqrt{57}}{4} \approx 0.637.$$

e.) Mivel a folyamat szuperkritikus, $\mathbb{E}N = \infty$. (Sőt, $\mathbb{P}(N = \infty) > 0$.)

II. $p = \frac{1}{6}$ esetén $m = \frac{1}{6} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{3} \cdot 3 = \frac{7}{4}$, $g(z) = \frac{1}{6} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{3}z^3$.

a.) $g_2(z) = g(g(z)) = \frac{1}{6} + \frac{1}{4}(\frac{1}{6} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{3}z^3) + \frac{1}{4}(\frac{1}{6} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{3}z^3)^2 + \frac{1}{3}(\frac{1}{6} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{3}z^3)^3$.

b.) $m_{12} = m^{12} = \left(\frac{7}{4}\right)^{12} \approx 825$.

c.) $r_0 = 0$; $r_1 = g(r_0) = g(0) = \frac{1}{6}$; $r_2 = g(r_1) = g(1/6) \approx 0.2168$; $r_3 = g(r_2) \approx g(0.2168) \approx 0.2360$.

d.) Mivel $m > 1$, a folyamat szuperkritikus, így a kihalás valószínűsége az $r = g(r)$ egyenlet egyetlen $[0, 1)$-beli gyöke. Vagyis

$$r = \frac{1}{6} + \frac{1}{4}r + \frac{1}{4}r^2 + \frac{1}{3}r^3.$$

Az előzőhöz hasonlóan ezt nullára rendezve és 12-vel végigszorozva azt kapjuk, hogy

$$0 = 4r^3 + 3r^2 - 9r + 2 = (r - 1)(4r^2 + 7r - 2),$$

amiből $r = 1$ vagy $r = \frac{-7\pm\sqrt{81}}{8}$, és mivel $r \in [0, 1)$,

$$\mathbb{P}(\text{kihalás}) = r = \frac{-7 + \sqrt{81}}{8} = \frac{1}{4}.$$

e.) Mivel a folyamat szuperkritikus, $\mathbb{E}N = \infty$. (Sőt, $\mathbb{P}(N = \infty) > 0$.)

3.HF: (Beadási határidő: 2013.12.02.)

HF 3.1 Egy épülő szennyvíztisztító üzem kapacitása 320000 liter/nap. A szennyvíztisztító 1500 háztartást szolgál ki, melyeket a következő kategóriákba sorolnak:

* kicsi, amelynek átlagos napi szennyvíztermelése 100 liter, de semmiképpen nem több, mint 300 liter;
* közepes, amelynek átlagos napi szennyvíztermelése 200 liter, de semmiképpen nem több, mint 500 liter;
* nagy, amelynek átlagos napi szennyvíztermelése 300 liter, de semmiképpen nem több, mint 800 liter.

Az egyes kategóriákba tartozó háztartások száma rendre 400, 800 illetve 300. Az egyes háztartások napi szennyvíztermelése függetlennek tekinthető.

a.) Adjunk felső becslést annak a valószínűségére, hogy egy adott napon az üzem nem képes a termelt szennyvizet megtisztítani.

b.) Az üzemeltető elégedetlen az előző részben kijött eredménnyel. Adjunk felső becslést arra, hogy mekkorára növeljék az üzem kapacitását ahhoz, hogy a túllépés kockázata 10^{-8} alá csökkenjen.

Megoldás: Jelöljük X_i-vel az i-edik háztartás egynapi szennyvíztermelését literben, legyen $n = 1500$ és $S_n = X_1 + X_2 + \cdots + X_n$. Az X_i-k függetlenek és korlátosak: egy valószínűséggel $a_{\le} X_i \le b_i$, ahol minden $a_i = 0$, továbbá 400 db i-re $b_i = 300$, 800 db i-re $b_i = 500$ és 300 db i-re $b_i = 800$. Az átlagokat figyelembe véve $\mathbb{E}S_n = 400 \cdot 100 + 800 \cdot 200 + 300 \cdot 300 = 290000$.

a.) A feladat szerint a $\mathbb{P}(S_n \ge 320000)$ valószínűségre keresünk becslést. A Hoeffding egyenlőtlenség szerint $t = 30000$ választással

$$\mathbb{P}(S_n \ge 320000) = \mathbb{P}(S_n \ge \mathbb{E}S_n + t) \le \exp\left\{-\frac{2t^2}{\sum_{i=1}^n (b_i - a_1)^2}\right\} =$$

$$= \exp\left\{-\frac{2 \cdot 30000^2}{400 \cdot (300 - 0)^2 + 800 \cdot (500 - 0)^2 + 300 \cdot (800 - 0)^2}\right\} =$$

$$= \exp\left\{-\frac{2 \cdot 30000^2}{428000000}\right\} \approx$$

$$\approx e^{-4.21} \approx 0.015.$$

b.) Ha azt akarjuk, hogy a túllépés kockázata egészen biztosan 10^{-8} alatt legyen, akkor válasszuk a kapacitást $\mathbb{E}S_n + t$-nek, ahol t olyan, hogy

$$\exp\left\{-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_1)^2}\right\} = 10^{-8}.$$

Ezt az egyenletet megoldva

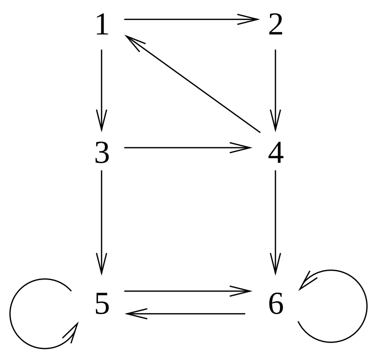
$$t = \sqrt{-214000000 \ln(10^{-8})} \approx 62786,$$

vagyis $\mathbb{E}S_n + t = 290000 + 62786 = 352786$ literes napi kapacitás biztosan elegendő. *Megjegyzés: Ahogy a feladat szövege is fogalmaz, ez egy* ***felső becslés*** *arra a kritikus kapacitásra, amivel a 10^{-8}-os hibavalószínűség biztosítható.*

4.HF: (Beadási határidő: 2013.12.09.)

HF 4.1 Az 1. ábrán látható gráf egy diszkrét idejű, időben homogén Markov lánc pozitív valószínűségű egylépéses átmeneteit mutatja. Osztályozzuk az állapotokat aszerint, hogy melyik melyikkel érintkezik! Minden osztályról állapítsuk megy, hogy

* zárt-e vagy nyílt,
* lényeges-e vagy lényegtelen,
* visszatérő-e vagy átmeneti,
* mennyi a periódusa.



1. ábra. Markov lánc gráf-reprezentációja (valószínűségek nélkül)

Megoldás:

* $\{1, 2, 3, 4\}$ nyílt, lényegtelen, átmeneti, periódusa 3.
* $\{5, 6\}$ zárt, lényeges, visszatérő, periódusa 1 (vagyis aperiodikus).

HF 4.2 Mari néni szeret beszélgetni, és befolyásolható. Minden este elmegy egy szomszédjához beszélgetni, és átveszi annak pártállását. Hat szomszédja van, ebből 2 fűpárti, 1 fapárti, 3 pedig virágpárti. Mari néni minden este vaktában választ beszélgetőpartnert azon 5 közül, akinél előző este *nem járt*. Jelöljük Mari néni lehetséges pártállásait $\{1, 2, 3\}$-mal, ahol „1" jelentése „fűpárti", „2" jelentése „fapárti", „3" jelentése „virágpárti". X_n pedig jelölje Mari néni pártállását n nap elteltével.

Modellezzük Mari néni állapotait időben homogén Markov láccal.

a.) Adjuk meg a Markov lánc átmenetmátrixát.

b.) Ha tudjuk, hogy a 0-dik napon Mari néni fűpárti volt, mi a valószínűsége az 123123 állapot-sorozatnak (a nulladik napot is beleértve)?

c.) Ha tudjuk, hogy a 0-dik napon Mari néni fűpárti volt, mi a valószínűsége, hogy a 2-dik napon is az?

d.) Hosszú idő elteltével közelítőleg mekkora valószínűséggel lesz Mari néni éppen fapárti?

e.) Mari néni minden nap elmegy a gazdaboltba, és ha éppen fűpárti, akkor fűnyíródamilt vesz 500 Ft-ért, ha éppen fapárti, akkor permetszert 3000 Ft-ért, ha pedig virágpárti, akkor tápoldatot 1000 Ft-ért. Napi átlagban mennyit költ a gazdaboltban hosszú távon?

Megoldás:

a.) Egy példa: ha ma éppen fűpárti, akkor a lehetséges 5 beszélgetőpartnere közül 1 fűpárti, 1 fapárti és 3 virágpárti. Ezért $P_{11} = \frac{1}{5}$, $P_{12} = \frac{1}{5}$, $P_{13} = \frac{3}{5}$. Hasonlóan végiggondolva

$$P = \begin{pmatrix} 1/5 & 1/5 & 3/5 \\ 2/5 & 0/5 & 3/5 \\ 2/5 & 1/5 & 2/5 \end{pmatrix}.$$

b.) $\mathbb{P}((X_1, X_2, X_3, X_4, X_5) = (2, 3, 1, 2, 3) \mid X_0 = 1) = P_{12}P_{23}P_{31}P_{12}P_{23} = \frac{1}{5}\frac{3}{5}\frac{2}{5}\frac{1}{5}\frac{3}{5} = \frac{18}{3125} = 0.00576.$

c.) $\mathbb{P}(X_2 = 1 \mid X_0 = 1) = (P^2)_{11} = \frac{9}{25} = 0.36.$

d.) Megkeressük a stacionárius eloszlászt, vagyis megoldjuk a $(P^T - I)\pi^T = 0$ lineáris egyenletrendszert:

$$\left(\begin{array}{ccc|c} -4/5 & 2/5 & 2/5 & 0 \\ 1/5 & -5/5 & 1/5 & 0 \\ 3/5 & 3/5 & -3/5 & 0 \end{array}\right)$$

Ennek megoldása (pontosabban: végtelen sok megoldása közül az az egy, ami elege tesz a $\pi_1 + \pi_2 + \pi_3 = 1$ normálási feltételnek is): $\pi = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \end{pmatrix} = \begin{pmatrix} 2/6 & 1/6 & 3/6 \end{pmatrix}$. Ez nem meglepő módon éppen a megfelelő pártállású szomszédok aránya.

Mivel a Markov lánc véges állapotterű, irreducibilis és aperiodikus, a Markov láncok alaptétele szerint $\lim_{n\to\infty} \mathbb{P}(X_n = i) \to \pi_i$, vagyis ha n nagy, akkor $\mathbb{P}(X_n = 2) \approx \pi_2 = \frac{1}{6}$.

e.) Mari néni költségfüggvénye

$$f = \begin{pmatrix} 500 \\ 3000 \\ 1000 \end{pmatrix}.$$

Az ergodtétel szerint ennek időátlaga 1 valószínűséggel

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{i\in S} \pi_i f(i) = \pi f = \begin{pmatrix} \frac{2}{6} & \frac{1}{6} & \frac{3}{6} \end{pmatrix} \cdot \begin{pmatrix} 500 \\ 3000 \\ 1000 \end{pmatrix} = \frac{7000}{6} \approx 1167,$$

vagyis Mari néni hosszú távon napi átlagban 1167 Ft-ot hagy a gazdaboltban.

5.HF: (Beadási határidő: 2013.12.16.)

HF 5.1 Egy hálózati kiszolgálóhoz Poisson-folyamat szerint érkeznek a feladatok, másodpercenként átlagosan kettő, és beállnak a sorba. Az egyes igények kiszolgálása egymástól és a beérkezésektől is független, exponenciális eloszlású véletlen ideig tart, aminek várható értéke $\frac{1}{4}$ másodperc. Az egyszerűség kedvéért tegyük fel, hogy a sorban legfeljebb 5 feladat lehet (azzal együtt, amelyik éppen kiszolgálás alatt áll), ami ezen felül esetleg érkezik, az elvész. Jelölje X_t a t időben a sorban álló feladatok számát.

a.) Modellezzük a folyamatot folytonos idejű Markov lánccal! Adjuk meg az állapotteret, és rajzoljuk fel a gráf-reprezentációt (az egyes átmenetek rátáival).

b.) Adjuk meg a ráta-mátrixot, a tartózkodási idő paraméter vektort és a beágyazott diszkrét idejű Markov lánc átmenetmátrixát.

c.) Írjuk fel a folyamat infinitezimális generátorát.

d.) Kezdetben a sor üres. Közelítőleg mekkora valószínűséggel lesz a sorban 120 másodperc elteltével pontosan 2 feladat?

e.) Az idő mekkora hányadát tölti a kiszolgáló üresjáratban hosszú távon?

f.) A beérkező feladatok mekkora hányada vész el hosszú távon?

g.) *Bónusz kérdés:* Mi lehet a válasz ugyanezekre a kérdésekre, ha a sor hossza nincsen korlátozva? A stacionárius eloszlás megsejtéséhez szabad kihasználni (mint a véges esetben is), hogy X_t születési-halálozási folyamat.

Megoldás:

a.) Az állapottér $S = \{0; 1; 2; 3; 4; 5\}$. Átmenet csak a közvetlen szomszédok között lehet, A felfelé ugrás rátája mindig éppen a beérkező feladatok folyamatának rátája, vagyis 2. A lefelé urgás rátája az exponenciális kiszolgálási idő paramétere, vagyis várható érétkének reciproka, esetünkben 4. Kivétel a 0 állapo, ahonnan nincs lefelé ugrás, és az 5 állapot, ahonnan nincs felfelé ugrás. A gráf-reprezentáció:

$$0 \underset{2}{\overset{4}{\leftrightarrows}} 1 \underset{2}{\overset{4}{\leftrightarrows}} 2 \underset{2}{\overset{4}{\leftrightarrows}} 3 \underset{2}{\overset{4}{\leftrightarrows}} 4 \underset{2}{\overset{4}{\leftrightarrows}} 5$$

b.) A ráta-mátrix

$$\underline{\underline{\lambda}} = \begin{pmatrix} * & 2 & 0 & 0 & 0 & 0 \\ 4 & * & 2 & 0 & 0 & 0 \\ 0 & 4 & * & 2 & 0 & 0 \\ 0 & 0 & 4 & * & 2 & 0 \\ 0 & 0 & 0 & 4 & * & 2 \\ 0 & 0 & 0 & 0 & 4 & * \end{pmatrix},$$

a tartózkodási idő paraméter vektor

$$\underline{\lambda} = \begin{pmatrix} 2 & 6 & 6 & 6 & 6 & 4 \end{pmatrix},$$

a beágyazott diszkrét idejű Markov lánc átmenetmátrixa

$$Q = \begin{pmatrix} 0 & 2/2 & 0 & 0 & 0 & 0 \\ 4/6 & 0 & 2/6 & 0 & 0 & 0 \\ 0 & 4/6 & 0 & 2/6 & 0 & 0 \\ 0 & 0 & 4/6 & 0 & 2/6 & 0 \\ 0 & 0 & 0 & 4/6 & 0 & 2/6 \\ 0 & 0 & 0 & 0 & 4/4 & 0 \end{pmatrix}.$$

c.) Az infinitezimális generátor

$$G = \begin{pmatrix} -2 & 2 & 0 & 0 & 0 & 0 \\ 4 & -6 & 2 & 0 & 0 & 0 \\ 0 & 4 & -6 & 2 & 0 & 0 \\ 0 & 0 & 4 & -6 & 2 & 0 \\ 0 & 0 & 0 & 4 & -6 & 2 \\ 0 & 0 & 0 & 0 & 4 & -4 \end{pmatrix}.$$

d.) Mivel a Markov lánc véges állapotterű és irreducibilis, a Markov láncok alaptétele szerint hosszú idő elteltével a tartózkodási valószínűségek az egyetlen stacionárius eloszlással közelíthetők. Ezért kiszámoljuk a stacionárius eloszlást, vagyis megoldjuk a $G^T\pi^T = 0$ homogén lineáris egyenletrendszert, ahol a π sorvektor a stacionárius eloszlás (a transzponáltja pedig oszlopvektor). Pontosabban: ennek a lineáris egyenletrendszernek a végtelen sok megoldása közül azt az egyet keressük, amelyik valószínűségeloszlás (vagyis az elemek összege 1). Az egyenletrendszer:

$$\left(\begin{array}{cccccc|c} -2 & 4 & 0 & 0 & 0 & 0 & 0 \\ 2 & -6 & 4 & 0 & 0 & 0 & 0 \\ 0 & 2 & -6 & 4 & 0 & 0 & 0 \\ 0 & 0 & 2 & -6 & 4 & 0 & 0 \\ 0 & 0 & 0 & 2 & -6 & 4 & 0 \\ 0 & 0 & 0 & 0 & 2 & -4 & 0 \end{array}\right),$$

ennek normált megoldása pedig

$$\pi = \begin{pmatrix} \frac{32}{63} & \frac{16}{63} & \frac{8}{63} & \frac{4}{63} & \frac{2}{63} & \frac{1}{63} \end{pmatrix}.$$

Tehát

$$\mathbb{P}(X_{120} = 2 \,|\, X_0 = 0) \approx \pi_2 = \frac{8}{63} \approx 0.127$$

e.) Legyen az $f : S \to \mathbb{R}$ függvény annak az indikátora, hogy a rendszer a 0 állapotban, vagyis üresjáratban van. Oszlopvektor formájában írva

$$f = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Mivel a Markov lánc véges állapotterű és irreducibilis, az ergodtétel értelmében f időátlaga majdnem biztosan konvergál az egyetlen stacionárius eloszlás szerinti várható értékhez, vagyis $\sum_{i\in S} \pi_i f_i = \pi f = \frac{32}{63}$-hoz. Tehát a rendszer hosszú távon az idő $\frac{32}{63}$-át tölti üresjáratban.

f.) Azt kell kitalálnunk, hogy a beérkező feladatok hanyad része érkezik pont olyankor, amikor a rendszer az 5 állapotban van. Mivel a feladatok érkezési rátája független a rendszer állapotától, ez pontosan annyi lesz, amekkora hányadát az időnek a rendszer az 5 állapotban tölti. Ez pedig az előző ponthoz hasonlóan $\pi_5 = \frac{1}{63}$, vagyis hosszú távon a beérkező feladatok $\frac{1}{63}$-ada vész el.

g.) Az állapottér ezúttal $S = \{0, 1, 2, \dots\} = \mathbb{N}$. A gráf-reprezentáció végtelen hosszú:

$$0 \underset{2}{\overset{4}{\leftrightarrows}} 1 \underset{2}{\overset{4}{\leftrightarrows}} 2 \underset{2}{\overset{4}{\leftrightarrows}} 3 \underset{2}{\overset{4}{\leftrightarrows}} 4 \underset{2}{\overset{4}{\leftrightarrows}} \cdots$$

a tartózkodási idő paraméter vektor

$$\underline{\lambda} = \begin{pmatrix} 2 & 6 & 6 & 6 & 6 & \dots \end{pmatrix},$$

a $\underline{\underline{\lambda}}$, a Q és a G is felírhatók végtelen mátrix formában.

A π stacionárius eloszlás is végtelen sorvektor, és a $G^T\pi^T = 0$ egyenlet miatt (vagy még inkább a grafikus reprezentációból kiolvashatóan) most is eleget kell tennie a $2\pi_k = 4\pi_{k+1}$ egyenletnek *minden* $k = 0, 1, 2\dots$*-re*. Ebből kiolvasható, hogy *ha van stacionárius eloszlás*, akkor az csak olyan lehet, hogy minden $k \in \mathbb{N}$-re

(1) $$\pi_k = \frac{\pi_0}{2^k}.$$

A kulcskérdés, ami megkülönbözteti a végtelen állapottér esetét a végestől, az, hogy vajon van-e olyan 1-*re felösszegződő* nemnegatív számsorozat, ami ennek eleget tesz, vagyis π_0 megválasztható-e úgy, hogy

$$\sum_{k\in S}\pi_k = \sum_{k=0}^{\infty}\frac{\pi_0}{2^k} = \pi_0\sum_{k=0}^{\infty}\frac{1}{2^k} = 1$$

legyen. Szerencsére esetünkben a $\sum_{k=0}^{\infty}\frac{1}{2^k}$ sor *véges*, konkrétan $= 2$, vagyis $\pi_0 = \frac{1}{2}$ választással az (1) egyenlet tényleg stacionárius eloszlást ad. Azt kaptuk tehát, hogy

$$\pi = \begin{pmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \frac{1}{32} & \cdots\end{pmatrix}.$$

Egy általunk nem tanult, végtelen állapotterű születési-halálozási folyamatokról szóló tétel szerint ez tényleg az egyetlen stacionárius eloszlás, a t időbeli eloszlások ehhez konvergálnak, az időátlagok pedig az ezen eloszlás szerinti várható értékhez (ha az átlagolandó függvény olyan, hogy ez a várható érték létezik).
Így hosszú idő elteltével kb. $\pi_2 = \frac{1}{8}$ valószínűséggel lesz pont 2 feladat a sorban, és hosszú távon az idő $\pi_0 = \frac{1}{2}$ részében lesz a gép üresjáratban. Elveszett feladat természetesen nem lesz.