Felsőbb Matematika Informatikusoknak – Sztochasztika
házi feladatok a 2. részhez – megoldások
2015 tavasz

Minden héten összesen egy pontot érnek a kitűzött feladatok.

1.HF: (Beadási határidő: 2015.04.10.)

HF 1.1 Egy radioaktív sugárforrás nagyon sok bomlásra képes atommagból áll, melyek mindegyike valamilyen kis valószínűséggel bomlik el éppen az általunk megfigyelt időintervallumban (és bocsát ki észlelhető sugárzást), a többi atommagtól függetlenül. A minta aktivitása $0.1Bq$ (vagyis Becquerel), ami azt jelenti, hogy másodpercenként átlagosan 0.1 bomlás történik.

a.) Legyen X az egy perc alatt (mondjuk 08:00-tól 08:01-ig) történő bomlások száma. Milyen eloszlással jó ezt modellezni? Vagyis mennyi a $\mathbb{P}(X = k)$ valószínűség (és melyik k-kra)?

b.) Mennyi annak a valószínűsége, hogy 08:00-tól 08:01-ig pontosan 4 bomlás történik, 08:01-től 08:03-ig pedig pontosan 10?

c.) Mennyi annak a valószínűsége, hogy 09:00-tól a következő bomlásig legalább 10 másodpercet kell várni?

Megoldás:

a.) X a sikeres próbálkozók száma nagyon sok független próbálkozásból, nagyon kis sikerval.ség mellett, így $X \sim Poi(\lambda)$ ahol $\lambda = \mathbf{E}X$. Esetünkben $\lambda = \mathbf{E}X = 60s \cdot 0.1Bq = 6$, vagyis $X \sim Poi(6)$, ami azt jelenti, hogy

$$\mathbb{P}(X = k) = e^{-6}\frac{6^k}{k!}, \quad k = 0, 1, 2, \dots.$$

b.) Legyen $X_{[0,1]}$ a 08:00 és 08:01 közötti bomlások száma, $X_{[1,3]}$ a 08:01 és 08:03 közötti bomlások száma. Ezek diszjunkt intervallumok, ezért $X_{[0,1]}$ és $X_{[1,3]}$ függetlenek. Továbbá $X_{[0,1]} \sim Poi(6)$ az előző részből, és hasonlóan $X_{[1,3]} \sim Poi(12)$. Ezért

$$\mathbb{P}(X_{[0,1]} = 4 \text{ és } X_{[1,3]} = 10) = \mathbb{P}(X_{[0,1]} = 4)\mathbb{P}(X_{[1,3]} = 10) =$$
$$= e^{-6}\frac{6^4}{4!} \cdot e^{-12}\frac{12^{10}}{10!} \approx 0.134 \cdot 0.105 \approx 0.01403$$

c.) Legyen Y a 09:00:00 és 09:00:10 közötti bomlások száma. Az előzőekhez hasonlóan $Y \sim Poi(1)$. Vegyük észre, hogy pontosan akkor kell legalább 10 másdopercet várni, ha 10 másodperc alatt nem történik bomlás, vagyis a keresett valószínűség

$$\mathbb{P}(Y = 0) = e^{-1}\frac{1^0}{0!} = e^{-1} \approx 0.368.$$

Alternatív megoldás: Legyen τ a 09:00 után történő első bomlásig a várakozási idő másodpercben. A Poisson-folyamat tulajdonságai miatt $\tau \sim Exp(\lambda)$ ahol $\lambda = \frac{1}{10}$ a folyamat intenzitása. Így

$$\mathbb{P}(\tau \geq 10) = 1 - F_{Exp(\frac{1}{10})}(10) = 1 - \left(1 - e^{-\frac{1}{10}\cdot 10}\right) = e^{-1} \approx 0.368$$

2.HF: (Beadási határidő: 2015.04.17.)

HF 2.1 Legyenek $X_1, X_2, \ldots$ független valószínűségi változók, melyek eloszlása egyenletes a $[0,1]$ intervallumon. Legyen

$$M_n = \max\{X_1, \ldots, X_n\}$$

és legyen F_n az M_n eloszlásfüggvénye, vagyis $F_n(x) := \mathbb{P}(M_n < x)$.

a.) Számoljuk ki az F_n eloszlásfüggvényt. *Segítség: az $\{X_1, \ldots, X_n\}$ maximuma persze pontosan akkor kisebb x-nél, ha minden X_i külön-külön kisebb x-nél.*

b.) Legyen $Y_n = n(1 - M_n)$ és legyen G_n az Y_n eloszlásfüggvénye. Számoljuk ki a G_n eloszlásfüggvényt. *Segítség:* $\mathbb{P}(Y_n < y) = \mathbb{P}(n(1 - M_n) < y) = \mathbb{P}(M_n > 1 - \frac{y}{n})$.

c.) Számoljuk ki a $G(y) := \lim_{n\to\infty} G_n(y)$ határértéket.

d.) A valószínűségi változók gyenge konvergenciájának eloszlásfüggvényes definíciója alapján mutassuk meg, hogy az Y_n sorozat gyengén konvergens, és nevezzük meg a határeloszlást.

Megoldás:

a.) $F_n(x) = \mathbb{P}(X_1 < x, X_2 < x, \ldots, X_n < x)$ a függetlenség miatt faktorizálódik:

$$F_n(x) = \mathbb{P}(X_1 < x)\mathbb{P}(x_2 < x) \ldots \mathbb{P}(X_n < x) = \mathbb{P}(X_1 < x)^n = x^n,$$

legalábbis ha $0 < x < 1$.

b.) A segítség alapján

$$G_n(y) = \mathbb{P}\left(M_n > 1 - \frac{y}{n}\right) = 1 - F_n\left(1 - \frac{y}{n}\right) = 1 - \left(1 - \frac{y}{n}\right)^n,$$

legalábbis ha $0 < y < n$.

c.)

$$G(y) = \lim_{n\to\infty}\left[1 - \left(1 - \frac{y}{n}\right)^n\right] = 1 - e^{-y},$$

legalábbis ha $y > 0$.

d.) Ez a G éppen az $Exp(1)$ eloszlás eloszlásfüggvénye, vagyis a folytonossági tételből az következik, hogy

$$Y_n \Rightarrow Exp(1).$$

HF 2.2 Egy béka egy hosszú lépcsősoron ugrál felfelé, de egyre fárad: annak valószínűsége, hogy az n-edik ugrással sikerül egy lépcsővel feljebb kerülnie, p_n a korábbi ugrásoktól függetlenül, és sajnos p_n csökken. Mennyi a valószínűsége, hogy kitartóan próbálkozva mégis akármilyen magasra feljut, ha

a.) $p_n = \frac{1}{n^2}$

b.) $p_n = \frac{1}{n}$

Segítség: használjuk a Borel-Cantelli lemmákat.

Megoldás: Legyen A_n az az esemény, hogy a béka n-edik ugrása sikeres, vagyis egy lépcsőfokkal feljebb jut vele. Így $\limsup_n A_n$ az az esemény, hogy a béka akármilyen magasra feljut.

a.) Ha $p_n = \frac{1}{n^2}$, akkor $\sum_n p_n < \infty$, így az első Borel-Cantelli lemma miatt $\mathbb{P}(\limsup_n A_n) = 0$.

b.) $p_n = \frac{1}{n}$ akkor $\sum_n p_n = \infty$. Mivel az A_n-ek függetlenek, a második Borel-Cantelli lemma miatt $\mathbb{P}(\limsup_n A_n) = 1$.

3.HF: (Beadási határidő: 2015.04.24.)

HF 3.1 a.) Amint azt órán kiszámoltuk, a standard normális eloszlás karakterisztikus függvénye $\Psi_{\mathcal{N}(0,1)}(t) = e^{-\frac{t^2}{2}}$. Ennek alapján számoljuk ki egy tetszőleges $\mathcal{N}(m, \sigma^2)$ normális eloszlás karakterisztikus függvényét. *(Segítség: Használjuk ki, hogy ha $X \sim \mathcal{N}(0,1)$ és $Y = m + \sigma X$, akkor $Y \sim \mathcal{N}(m, \sigma^2)$. Használjuk továbbá a karakterisztikus függvény definícióját: $\Psi_{\mathcal{N}(0,1)}(t) = \Psi_X(t) := \mathbf{E}(e^{itX})$ és $\Psi_{\mathcal{N}(m,\sigma^2)}(t) = \Psi_Y(t) := \mathbf{E}(e^{itY})$.)*

b.) Legyenek U és V független normális eloszlású valószínűségi változók: $U \sim \mathcal{N}(m_1, \sigma_1^2)$, $V \sim \mathcal{N}(m_2, \sigma_2^2)$. Legyen $Z = X + Y$. Számoljuk ki Z karakterisztikus függvényét, és ebből mondjuk meg Z eloszlását! *(Segítség: Használjuk ki, hogy független X-re és Y-ra $\Psi_{X+Y}(t) = \Psi_X(t) \cdot \Psi_Y(t)$, és hogy a karakterisztikus függvény egyértelműen meghatározza az eloszlást.)*

Megoldás:

a.) $\Psi_Y(t) = \mathbf{E}e^{it(m+\sigma X)} = e^{itm}\mathbf{E}e^{i(\sigma t)X} = e^{itm}\Psi_X(\sigma t) = e^{itm}e^{-\frac{\sigma^2 t^2}{2}}$.

b.)

$$\Psi_Z(t) = \Psi_U(t)\Psi_V(t) = e^{itm_1}e^{-\frac{\sigma_1^2 t^2}{2}} \cdot e^{itm_2}e^{-\frac{\sigma_2^2 t^2}{2}} = e^{it(m_1+m_2)}e^{-\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}},$$

vagyis $Z = U + V \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

HF 3.2 a.) Legyen az X valószínűségi változó pesszimista geometriai eloszlású $p \in (0,1)$ paraméterrel, vagyis $\mathbb{P}(X = k) = q^k p$ $(k = 0, 1, \dots)$, ahol $q := 1 - p$. Számoljuk ki X generátorfüggvényét – vagyis a $g_X(z) := \sum_{k=0}^{\infty} \mathbb{P}(X = k)z^k$ függvényt. *(Segítség: $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, ha $-1 < x < 1$.)*

b.) Egy Y nemnegatív egész értékű valószínűségi változó generátorfüggvénye $g_Y(z) = \frac{2+3z+4z^3+z^7}{10}$. Határozzuk meg ebből Y eloszlását, vagyis hogy melyik $k \in \mathbb{N}$ értéket mekkora valószínűséggel veszi fel. *(Segítség: ezt a generátorfüggvényt könnyen fel lehet írni $g_Y(z) = \sum_k p_k z^k$ alakban.)*

Megoldás:

a.) Csak a nemelfajult $0 < p < 1$, vagyis $0 < q < 1$ esettel foglalkozom. Ekkor

$$g_X(z) = \sum_{k=0}^{\infty} \mathbb{P}(X = k)z^k = \sum_{k=0}^{\infty} q^k p z^k = p\sum_{k=0}^{\infty}(qz)^k = \frac{p}{1 - qz},$$

legalábbis $0 \leq z \leq 1$-re, ahol minket a $g(z)$ értéke érdekel, mert ekkor biztosan $|qz| < 1$, vagyis a fenti mértani sor konvergens.

b.)

$$(1) \qquad g_Y(z) = \frac{2 + 3z + 4z^3 + z^7}{10} = \frac{2}{10}z^0 + \frac{3}{10}z^1 + \frac{4}{10}z^3 + \frac{1}{10}z^7 =$$

$$(2) \qquad = p_0 z^0 + p_1 z^1 + p_3 z^3 + p_7 z^7 = \sum_{k=0}^{\infty} p_k z^k,$$

ahol

k	0	1	3	7	egyébként
p_k	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{1}{10}$	0

, vagyis ez a táblázat éppen az Y eloszlástáblázata.

4.HF: (Beadási határidő: 2015.05.05.)

HF 4.1 Egy sort kiszolgáló számítógép a hozzá érkező, sorra kerülő feladatokat pontosan egységnyi idő alatt végzi el. Ezen időegység alatt azonban újabb feladat(ok) érkez(het)nek, melyek száma véletlen, és az előzményektől független. Ezek a feladatok beállnak a sorba, és ott várakoznak, amíg az előttük érkezett feladatokat a gép el nem végzi. Az egy időegység alatt érkező új feladatok számának eloszlása:

k	0	1	2	3
$P(k$ igény$)$	$4/10$	p	$5/10-p$	$1/10$

A sor kezdetben üres, az első feladat érkezésének pillanatában kezdjük az időt mérni. Nevezzük a feladatok „nulladik generációjának” a legelőször érkezett feladatot. Nevezzük „első generációnak” azokat a feladatokat, amik az ő elvégzése alatt érkeznek. Nevezzük „második generációnak” azokat a feladatokat, amik az első generációba tartozó feladatok elvégzése alatt érkeznek, stb ... Nevezzük „(n+1)-edik generációnak” azokat a feladatokat, amik az n-edik generációba tartozó feladatok elvégzése alatt érkeznek. Jelölje Z_n az n-edik generáció elemszámát, $n = 0, 1, 2, \dots$. Válaszoljuk meg az alábbi kérdéseket

I. ha $p = 4/10$,

II. ha $p = 2/10$.

a.) Mennyi Z_1 várható értéke?

b.) Mi Z_1 generátorfüggvénye?

c.) Mi Z_2 generátorfüggvénye?

d.) Mennyi Z_{72} várható értéke?

e.) Mennyi az $r_3 := \mathbb{P}(Z_3 = 0)$ valószínűség?

f.) Mennyi annak a valószínűsége, hogy a sor előbb-utóbb újra üres lesz? *(Segítség: egy harmadfokú egyenletet könnyű megoldani, ha ismerjük az egyik gyökét.)*

g.) Mennyi a foglaltsági periódus hosszának várható értéke? (A foglaltsági periódus az a véletlen időszak, aminek a végén először lesz a sor újra üres.)

Megoldás: Z_n Galton-Watson elágazó folyamat, amire $Z_0 = 1$ és az egylépéses utódszám-eloszlás éppen a feladatban szereplő táblázatban megadott eloszlás. A kérdések megválaszolásához ennek m várható értékére és $g(z)$ generátorfüggvényére lesz szükség:

I. ha $p = 4/10$, akkor $g(z) = \frac{4}{10} \cdot z^0 + \frac{4}{10} \cdot z^1 + \frac{1}{10} \cdot z^2 + \frac{1}{10} \cdot z^3 = \frac{4+4z+z^2+z^3}{10}$, $m = \frac{4}{10} \cdot 0 + \frac{4}{10} \cdot 1 + \frac{1}{10} \cdot 2 + \frac{1}{10} \cdot 3 = \frac{9}{10} < 1$, a folyamat szubkritikus.

II. ha $p = 2/10$, akkor $g(z) = \frac{4}{10} \cdot z^0 + \frac{2}{10} \cdot z^1 + \frac{3}{10} \cdot z^2 + \frac{1}{10} \cdot z^3 = \frac{4+2z+3z^2+z^3}{10}$, $m = \frac{4}{10} \cdot 0 + \frac{2}{10} \cdot 1 + \frac{3}{10} \cdot 2 + \frac{1}{10} \cdot 3 = \frac{11}{10} > 1$, a folyamat szuperkritikus.

Ezek alapján, a szokásos jelölésekkel

I. ha $p = 4/10$,

a.) $\mathbf{E}Z_1 = m = \frac{9}{10}$

b.) Z_1 generátorfüggvénye $g_1(z) = g(z) = \frac{4+4z+z^2+z^3}{10}$

c.) Z_2 generátorfüggvénye

$$\begin{aligned} g_2(z) &= g(g(z)) = g\left(\frac{4+4z+z^2+z^3}{10}\right) = \\ &= \frac{4 + 4\left(\frac{4+4z+z^2+z^3}{10}\right) + \left(\frac{4+4z+z^2+z^3}{10}\right)^2 + \left(\frac{4+4z+z^2+z^3}{10}\right)^3}{10} \end{aligned}$$

d.) $\mathbf{E}Z_{72} = m_{72} = m^{72} = \left(\frac{9}{10}\right)^{72} \approx 0.000507$

e.) Az $r_n := \mathbb{P}(Z_n = 0)$ jelöléssel

* $r_0 = 0$
* $r_1 = g(r_0) = g(0) = \frac{4}{10}$
* $r_2 = g(r_1) = g\left(\frac{4}{10}\right) = 0.5824$
* $r_3 = g(r_2) = g(0.5824) \approx 0.68663$

f.) $\mathbb{P}$(a sor előbb-utóbb újra üres lesz) $= \mathbb{P}$(kihalás) $= r_\infty = 1$, mert a folyamat szubkritikus.

g.) A foglaltsági periódus hossza éppen $N := \sum_{n=0}^{\infty} Z_n$, amire $\mathbf{E}N = \frac{1}{1-m} = 10$ (mert $m < 1$).

II. ha $p = 2/10$,

a.) $\mathbf{E}Z_1 = m = \frac{11}{10}$

b.) Z_1 generátorfüggvénye $g_1(z) = g(z) = \frac{4+2z+3z^2+z^3}{10}$

c.) Z_2 generátorfüggvénye

$$
\begin{aligned}
g_2(z) &= g(g(z)) = g\left(\frac{4+2z+3z^2+z^3}{10}\right) = \\
&= \frac{4 + 2\left(\frac{4+2z+3z^2+z^3}{10}\right) + 3\left(\frac{4+2z+3z^2+z^3}{10}\right)^2 + \left(\frac{4+2z+3z^2+z^3}{10}\right)^3}{10}
\end{aligned}
$$

d.) $\mathbf{E}Z_{72} = m_{72} = m^{72} = \left(\frac{11}{10}\right)^{72} \approx 995.59$

e.) Az $r_n := \mathbb{P}(Z_n = 0)$ jelöléssel

* $r_0 = 0$
* $r_1 = g(r_0) = g(0) = \frac{4}{10}$
* $r_2 = g(r_1) = g\left(\frac{4}{10}\right) = 0.5344$
* $r_3 = g(r_2) = g\left(0.5344\right) \approx 0.60782$

f.) A kérdés $\mathbb{P}$(a sor előbb-utóbb újra üres lesz) $= \mathbb{P}$(kihalás) $= r_\infty$. Mivel a folyamat szuperkritikus, tudjuk, hogy $r_\infty < 1$ és számolni kell: meg kell oldani a $g(z) = z$ egyenletet. Esetünkben az egyenlet

$$\frac{4+2z+3z^2+z^3}{10} = z,$$

átszorozva és nullára redukálva

$$z^3 + 3z^2 - 8z + 4 = 0.$$

Ez egy harmadfokú egyenlet, de szerencsére tudjuk, hogy $z = 1$ mindig gyöke, vagyis a baloldali polinomból ki lehet emelni a $(z-1)$ gyöktényezőt. Valóban,

$$z^3 + 3z^2 - 8z + 4 = (z-1)(z^2 + 4z - 4),$$

amit az egyenletbe visszaírva

$$(z-1)(z^2 + 4z - 4) = 0.$$

Mivel mi NEM a $z = 1$ megoldást keressük, nyogodtan vévgig lehet osztani $(z-1)$-gyel, és marad, hogy

$$z^2 + 4z - 4 = 0,$$

ami már másodfokú. A megoldásai

$$z = \frac{-4 \pm \sqrt{(-4)^2 - 4 \cdot 1 \cdot (-4)}}{2} = \frac{-4 \pm \sqrt{32}}{2} = -2(1 \pm \sqrt{2}),$$

vagyis a gyökök $z = -2(1 - \sqrt{2}) \approx -4.4$ és $z = -2(1 - \sqrt{2}) \approx 0.8284$. Mi az egyetlen $[0, 1)$-beli megoldást keressük, vagyis

$$\mathbb{P}(\text{kihalás}) = r_\infty = 2(\sqrt{2} - 1) \approx 0.8284$$

g.) A foglaltsági periódus hossza éppen $N := \sum_{n=0}^{\infty} Z_n$, amire $\mathbf{E}N = \infty$, mert $m > 1$.

5.HF: (Beadási határidő: 2015.05.15.)

HF 5.1 Egy diszkrét idejű, időben homogén X_n Markov lánc állapottere $S = \{1, 2, 3\}$. A Markov lánc az 1-es állapotból $50 - 50\%$ valószínűséggel ugrik a 2-es és 3-as állapotba. Ha a 2-es állapotban van, akkor 50% valószínűséggel ott is marad, 50% valószínűséggel pedig a 3-as állapotba ugrik. A 3-as állapotból mindig az 1-esbe ugrik. A Markov lánc X_0 kezdeti állapotát kockadobással sorsoljuk, egyenlő esélyt adva mindhárom állapotnak.

a.) Rajzoljuk fel a Markov lánc gráf-reprezentációját.

b.) Írjuk fel a Markov lánc átmenetmátrixát.

c.) Írjuk fel a Markov lánc kezdeti eloszlás vektorát.

d.) Mennyi a valószínűsége, hogy a folyamat kezdetén a 131223 állapot-sorozatot figyeljük meg (a 0-dik (kezdő) állapotot is beleértve)?

e.) Mennyi a $\mathbb{P}(X_4 = 1 \mid X_0 = 1)$ átmenetvalószínűség?

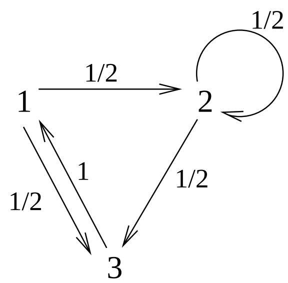
f.) Keressük meg a Markov lánc stacionárius eloszlásait.

g.) Közelítőleg mennyi a valószínűsége, hogy 100 lépés után a Markov lánc a 2-es állapotban lesz?

h.) Legyen az $f : S \to \mathbb{R}$ függvény olyan, hogy $f(1) = 0$, $f(2) = 1$ és $f(3) = 5$. Mennyi lesz az $f(X_n)$ sorozat (idő)átlaga hosszú távon?

Megoldás:

a.) A gráf-reprezentáció



b.) Az átmenetmártix $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}$

c.) A kezdeti eloszlás vektor $\pi(0) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, mert mindhárom állapotnak azonos esélyt adtunk.

d.)

$$\begin{aligned} \mathbb{P}(131223) &= \mathbb{P}(X_0 = 1, X_1 = 3, X_2 = 1, X_3 = 2, X_4 = 2, X_5 = 3) = \\ &= \pi_1(0) P_{13} P_{31} P_{12} P_{22} P_{23} = \frac{1}{3} \cdot \frac{1}{2} \cdot 1 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{48}. \end{aligned}$$

e.) Az 1-es állapotból az 1-esbe négy lépésben visszajutni csak kétféleképpen lehet: $1 \to 2 \to 2 \to 3 \to 1$, vagy $1 \to 3 \to 1 \to 3 \to 1$. Ezek feltételes val.sége (feltéve, hogy $X_0 = 1$) $P_{12} P_{22} P_{23} P_{31} = \frac{1}{8}$, illetve $P_{13} P_{31} P_{13} P_{31} = \frac{1}{4}$, így

$$\mathbb{P}(X_4 = 1 \mid X_0 = 1) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}.$$

f.) Mivel a Markov lánc irreducibilis és véges állapotterű, pontosan egy π stacionárius eloszlás (sorvektor) van, éspedig a

$$(P^T - I)\pi^T = 0$$

lineáris egyenletrendszer egyetlen olyan megoldása, ahol a sorösszeg 1. (Itt I az egységmátrix.) Jelen esetben

$$P^T - I = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 \end{pmatrix},$$

vagyis a lineáris egyenletrendszer a szokásos mátrix-reprezentációval

$$\left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 & 0 \end{array}\right).$$

Ezt megoldjuk pl. Gauss eliminációval:

$$\left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \end{array}\right) \sim$$
$$\sim \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right),$$

amiből $\pi_1 = \pi_3$ és $\pi_1 = \pi_2$. Így pl. a $\pi_3 := 1$ önkényes választással megkapjuk a lineáris egyenletrendszer egy megoldását: $\tilde{\pi} = (1\,1\,1)$. Ez még nem az, amit keresünk, mert az elemek összege nem 1, hanem $1+1+1 = 3$, ezért a keresett megoldást úgy kapjuk, hogy ezt a $\tilde{\pi}$-t lenormáljuk, vagyis leosztjuk 3-mal:

$$\pi = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

g.) A Markov lánc irreducibilis, aperiodikus és véges állapotterű, $n = 100$ lépés pedig hosszú idő, ezért a Markov láncok alaptétele értelmében a kezdeti eloszlástól függetlenül

$$\mathbb{P}(X_{100} = 2) \approx \pi_2 = \frac{1}{3}.$$

h.) Az f függvényt célszerű oszlopvektor formájába írni:

$$f = \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}.$$

A Markov lánc irreducibilis és véges állapotterű, ezért az ergodtétel értelmében az $f(X_n)$ sorozat (idő)átlaga hosszú távon 1 valószínűséggel a stacionárius eloszlás szerinti várható értékhez tart:

$$\lim_{n\to\infty} \frac{f(X_0) + f(X_1) + \cdots + f(X_{n-1})}{n} = \mathbf{E}_\pi f = \sum_{i\in S} \pi_i f(i) = \pi f =$$
$$= \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix} = \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 5 = 2.$$

HF 5.2 Egy fagyis bácsi előtt gyerekek állank sorba. Ő minden sorra kerülő gyereket exponenciális eloszlású véletlen idő alatt szolgál ki, fél perc várható értékkel, az előzményektől és a sorban állók számától függetlenül. A gyerekek Poisson folyamat szerint érkeznek, percenként átlagosan 1, a múlttól és a sor hosszától függetlenül. Kivétel, ha a sorban már 3 gyerek áll, mert akkor több nem állhat be (az apukája elvonszolja). Jelölje $X(t)$ a sorban állók számát t idő elteltével. Az időt mérjük percben. Modellezzük a sor hosszát folytonos idejű Markov lánccal.

a.) Mi a Markov lánc állapottere?

b.) Rajzoljuk fel a Markov lánc gráf-reprezentációját – vagyis a lehetséges átmenetek irányított gráfját az egyes átmenetek rátáival.

c.) Írjuk fel a Markov lánc ráta-mátrixát.

d.) Írjuk fel a Markov lánc tartózkodási idő paraméter vektorát.

e.) Ha a sor hossza éppen 2, várhatóan mennyi idő múlva fog megváltozni?

f.) Ha a sor hossza éppen 2, mennyi a valószínűsége, hogy a következő állapot a 3 lesz?

g.) Írjuk fel a Markov lánc infinitezimális generátorát.

h.) Keressük meg a Markov lánc stacionárius eloszlásait.

i.) Ha a sor $t = 0$-kor üres, közelítőleg mennyi a valószínűsége, hogy 2 óra elteltével 3 lesz a hossza?

j.) Az idő hány százalékát tölti a fagyis bácsi tétlenül (mert üres a sor) hosszú távon?

Megoldás:

a.) A sorban állók száma 0,1,2 vagy 3 lehet, így az állapottér $S = \{0, 1, 2, 3\}$.

b.) Egyszerre csak egy gyerek mehet el, és csak egy érkezhet, így ugrani minden állapotból csak szomszédos állapotokba lehet. A felfelé ugrás rátája mindig a gyerekek érkezési folyamatának rátája, vagyis 1 (mert percenként átlag 1 gyerek érkezik). A lefelé ugrás rátája pedig a kiszolgálás rátája, vagyis 2, mert a bácsi percenként átlag 2 gyereket szolgál ki. *FONTOS, hogy a ráta, ami 2, nem egyenlő a kiszolgálási idő várható értékével, ami $\frac{1}{2}$, hanem annak a reciproka.* Így a gráf-reprezentáció

$$0 \underset{2}{\overset{1}{\rightleftarrows}} 1 \underset{2}{\overset{1}{\rightleftarrows}} 2 \underset{2}{\overset{1}{\rightleftarrows}} 3$$

c.) Ez alapján a ráta-mátrix

$$\underline{\underline{\lambda}} = \begin{pmatrix} * & 1 & 0 & 0 \\ 2 & * & 1 & 0 \\ 0 & 2 & * & 1 \\ 0 & 0 & 2 & * \end{pmatrix}.$$

d.) A tartózkodási idő paraméter vektor a ráta-mátrix sorösszegeiből áll, vagy más szóval az egyes állapotokból való elugrás rátáiból:

$$\underline{\lambda} = \begin{pmatrix} 1 & 3 & 3 & 2 \end{pmatrix}.$$

e.) A 2 állapotból való elugrás rátája $\lambda_2 = 3$ vagyis a tartózkodási idő várható értéke $\frac{1}{\lambda_2} = \frac{1}{3}$.

f.) Az egyes irányokba való ugrások valószínűségeit a beépített diszkrét idejű Markov lánc átmenetmátrixa adja meg: $Q_{ij} = \frac{\lambda_{ij}}{\lambda_1}$ ha $i \neq j$, és $Q_{ij} = 0$, ha $i = j$. Most a kérdés csak a 2-ből 3-ba ugrás val.ségére vonatkozott, vagyis a válasz $Q_{23} = \frac{\lambda_{23}}{\lambda_2} = \frac{1}{3}$. De ha már itt tartunk, felírom az egész mátrixot:

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

g.) Az infinitezimális generátort úgy kapjuk a ráta-mátrixból, hogy a főátlóba beírjuk a tartózkodási idő paraméter vektor eleminek ellentettjét (vagyis annyit, hogy a sorösszegek nullák legyenek):

$$G = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 2 & -3 & 1 & 0 \\ 0 & 2 & -3 & 1 \\ 0 & 0 & 2 & -2 \end{pmatrix}.$$

h.) Mivel a Markov lánc irreducibilis és véges állapotterű, pontosan egy π stacionárius eloszlás (sorvektor) van, éspedig a $G^T \pi^T = 0$ lineáris egyenletrendszer egyetlen olyan megoldása, ahol a sorösszeg 1. Jelen esetben az egyenletrendszer a szokásos mátrix-reprezentációval

$$\left(\begin{array}{cccc|c} -1 & 2 & 0 & 0 & 0 \\ 1 & -3 & 2 & 0 & 0 \\ 0 & 1 & -3 & 2 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right),$$

aminek a normált megoldása

$$\pi = \begin{pmatrix} \frac{8}{15} & \frac{4}{15} & \frac{2}{15} & \frac{1}{15} \end{pmatrix}.$$

(Az egyenletrendszer megoldásának és a megoldás lenormálásának menetét illetően lásd az előző feladat megoldását, vagy bármelyik lineáris algebra könyvet.)

i.) A Markov lánc folytonos idejű, irreducibilis és véges állapotterű, $t = 120$ perc pedig hosszú idő, ezért a Markov láncok alaptétele értelmében a kezdeti eloszlástól függetlenül

$$\mathbb{P}(X(120) = 3) \approx \pi_3 = \frac{1}{15}.$$

j.) Legyen

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

függvény az állapotéren. Ez éppen a 0 állapot indikátora, így $f(X(t))$ időátlaga éppen azt mondja meg, hogy a sor az idő mekkora hányadában üres. A Markov lánc irreducibilis és véges állapotterű, ezért az ergodtétel értelmében az $f(X(t))$ függvény (idő)átlaga hosszú távon 1 valószínűséggel a stacionárius eloszlás szerinti várható értékhez tart:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X(t))\,\mathrm{d}\,t = \mathbf{E}_\pi f = \sum_{i\in S} \pi_i f(i) = \pi f =$$

$$= \begin{pmatrix} \frac{8}{15} & \frac{4}{15} & \frac{2}{15} & \frac{1}{15} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{8}{15}.$$