Felsőbb Matematika Villamosmérnököknek - Sztochasztika
házi feladat megoldások, 2015 ősz

Minden héten összesen 2 pontot érnek a kitűzött feladatok.

1.HF: (Beadási határidő: 2015.09.28.)

HF 1.1 Egy lépcsőházban lévő (egyetlen) villanykörte folyamatosan ég. Élettartama véletlen, exponenciális eloszlású 1000 óra várható értékkel. Ha a körte kiég, a gondnok azonnal kicseréli egy ugyanolyanra.

a.) Közelítőleg mennyi a valószínűsége, hogy holnap 08:00 és 08:10 között kiég (legalább) egy körte?

b.) Mennyi a valószínűsége, hogy jövő héten pontosan 2 körte ég ki?

c.) Mennyi a valószínűsége, hogy 2016 februárban pontosan 2 körte ég ki, és mindkettő a hónap első hetében?

Megoldás: Az időt mérjük órákban. Az egymást követő körtecserék között eltelt idő mindig $\lambda = \frac{1}{1000}$ paraméterű exponenciális eloszlású és független az előzményektől, így a körtecserék időpontjai Poisson folyamatot alkotnak $\lambda = \frac{1}{1000}$ rátával.

a.) A 08:00 és 08:10 között eltelt $t = \frac{1}{6}$ (óra) rövid idő, így

$$\mathbb{P}(\text{t idő alatt történik körtecsere}) \approx \lambda t = \frac{1}{6000} \approx 0.00017$$

b.) Legyen X a jövő héten, vagyis $t = 168$ óra alatt kiégő körték száma. Így $X \sim Poi(\lambda t) = Poi(0.168)$, vagyis $\mathbb{P}(X = k) = e^{-0.168}\frac{0.168^k}{k!}$ minden $k = 0, 1, 2, \ldots$-re. Konkrétan

$$\mathbb{P}(X = 2) = e^{-0.168}\frac{0.168^2}{2!} \approx 0.0119$$

c.) A februárban kiégő körték száma és a február első hónapjában kiégő körték száma is Poisson eloszlású, de VIGYÁZAT, nem függetlenek, hiszen a két intervallum nem diszjunkt. (Hát persze, ha az első héten kiég két körte, akkor a hónapban is legalább kettőnek ki kell égni). Vagyis a célszerű felírás:

* Y legyen az első héten (168 óra alatt) kiégett körték száma
* Z legyen a hónap hátralévő 22 napjában (528 óra alatt) kiégett körték száma (Affene, 2016 szökőév :))

Így $Y \sim Poi(\lambda \cdot 168) = Poi(0.168)$, $Z \sim Poi(\lambda \cdot 528) = Poi(0.528)$ *és függetlenek.* Ezért

$$\mathbb{P}(Y = 2, Z = 0) = \mathbb{P}(Y = 2)\mathbb{P}(Z = 0) = e^{-0.168}\frac{0.168^2}{2!} \cdot e^{-0.528}\frac{0.528^0}{0!} \approx 0.0070$$

HF 1.2 Eldobunk egy szabályos dobókockát. Ezután feldobunk annyi szabályos érmét, amennyit a kockával dobtunk. Legyen X az érmékkel dobott fejek száma. Mennyi X várható értéke? *(Tipp: használjuk a teljes várható érték tételt!)*

Megoldás: Legyen N a kockával dobott szám. Az $\{N = k\}$ feltétel mellett X *feltételes* eloszlása binomiális $Bin(n, p)$ mégpedig $n = k$ és $p = \frac{1}{2}$ paraméterekkel. Ennek várható értéke $\mathbf{E}(Bin(n, p)) = np = \frac{k}{2}$, vagyis éppen kiszámoltuk az X *feltételes* várható értékét:

$$\mathbf{E}(X \mid N = k) = \frac{k}{2} \quad k = 1, 2, 3, 4, 5, 6\text{-ra.}$$

Mivel $\{N = 1\}, \{N = 2\}, \ldots, \{N = 6\}$ teljes eseményrendszer, a teljes várható érték tétel szerint

$$\mathbf{E}X = \sum_{k=1}^{6} \mathbb{P}(N = k)\mathbf{E}(X \mid N = k) = \sum_{k=1}^{6} \frac{1}{6}\frac{k}{2} = \frac{1 + 2 + 3 + 4 + 5 + 6}{12} = \frac{21}{12} = 1.75$$

2.HF: (Beadási határidő: 2015.10.12.)

HF 2.1 Egy számítógépes programban egy véletlen, rekurzív rutin fut: minden egyes részfolyamat egységnyi időt vesz igénybe, ám ezen felül minden részfolyamat véletlen számú, önmagával megegyező al-folyamatot indít. Az így indított al-folyamatok száma 0, 1, 2 vagy 3 lehet, 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.*

Kezdetben egyetlen „gyökér" folyamat fut, ez alkotja egyedül a nulladik generációt. Az első generációt a „gyökér" által (közvetlenül) indított alfolyamatok alkotják, a második generációt az első generáció tagjai által indítottak, stb.

Jelölje Z_k a k-adik generáció tagjainak a számát ($k = 0, 1, 2, \dots$), N pedig a program futása során induló részfolyamatok teljes számát (vagyis $N = \sum_{k=0}^{\infty} Z_k$, ami egyben a program teljes futási ideje is).

Válaszoljuk meg az alábbi kérdéseket

I. $p = \frac{1}{2}$ esetén,

II. $p = \frac{1}{4}$ esetén:

a.) Mi Z_2 generátorfüggvénye?

b.) Mennyi Z_{10} várható értéke?

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

d.) Mennyi a valószínűsége annak, hogy a program előbb-utóbb lefut (vagyis hogy valamelyik generáció már üres)? *(Tipp: ha egy harmadfokú egyenlet egyik gyökét valahonnan ismerjük, akkor a többi is könnyű.)*

e.) Mennyi N várható értéke?

f.) Mi N generátorfüggvénye?

Megoldás: Z_k elágazó folyamat.

Egylépéses utódszám-eloszlása $\begin{array}{c|c|c|c|c} i & 0 & 1 & 2 & 3 \\ \hline \mathbb{P}(X=i) & p & \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-p \end{array}$. Ennek generátorfüggvénye

$$g(z) = pz^0 + \frac{1}{4}z^1 + \frac{1}{4}z^2 + \left(\frac{1}{2} - p\right) z^3,$$

várható értéke

$$m = p \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 2 + \left(\frac{1}{2} - p\right) \cdot 3 = \frac{9}{4} - 3p.$$

I. $p = \frac{1}{2}$-re $g(z) = \frac{1}{2} + \frac{1}{4}z + \frac{1}{4}z^2$ és $m = \frac{3}{4}$.

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

b.) $\mathbf{E}Z_{10} = m^{10} = \left(\frac{3}{4}\right)^{10} \approx 0.056$

c.) $\mathbb{P}(Z_3 = 0) = r_3$ ahol $r_0 = 0$ és $r_{k+1} = g(r_k)$. Esetünkben $r_1 = g(0) = \frac{1}{2}$, $r_2 = g(\frac{1}{2}) = \frac{11}{16}$, $r_3 = g\left(\frac{11}{16}\right) \approx 0.79$.

d.) $m < 1$, vagyis a folyamat szubkritikus, így a kihalás (=lefutás) valószínűsége 1.

e.) $m < 1$, vagyis a folyamat szubkritikus, így $\mathbf{E}N = \frac{1}{1-m} = 4$.

f.) N generátorfüggvénye, $g_N(z)$ a $g_N(z) = zg(g_N(z))$ egyenlet megoldása. Az átláthatóság kedvéért $g_N(z)$-t Y-nal jelölve $Y = zg(Y)$, vagyis

$$Y = z\left(\frac{1}{2} + \frac{1}{4}Y + \frac{1}{4}Y^2\right),$$

ami egy másodfokú egyenlet Y-ra:

$$\frac{z}{4}Y^2 + \left(\frac{z}{4} - 1\right) Y + \frac{z}{2} = 0.$$

Ezt megoldva

$$Y = \frac{1 - \frac{z}{4} \pm \sqrt{\left(\frac{z}{4} - 1\right)^2 - 4\frac{z}{4}\frac{z}{2}}}{2\frac{z}{4}} = \frac{4 - z \pm \sqrt{16 - 8z - 7z^2}}{2z}.$$

Hogy a két gyök közül a $+$-os vagy a $-$-os a jó, azt ki lehet találni pl. abból, hogy $z = 1$-ben minden generátorfüggvény 1 kell, hogy legyen, vagyis $1 = \frac{4-1\pm\sqrt{16-8-7}}{2} = \frac{3}{2} \pm \frac{1}{2}$, tehát a *minuszos* megoldás a helyes:

$$g_N(z) = \frac{4 - z - \sqrt{16 - 8z - 7z^2}}{2z}.$$

II. $p = \frac{1}{4}$-re $g(z) = \frac{1}{4} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{4}z^3$ és $m = \frac{3}{2}$.

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

b.) $\mathbf{E}Z_{10} = m^{10} = \left(\frac{3}{2}\right)^{10} \approx 57.7$

c.) $\mathbb{P}(Z_3 = 0) = r_3$ ahol $r_0 = 0$ és $r_{k+1} = g(r_k)$. Esetünkben $r_1 = g(0) = \frac{1}{4}$, $r_2 = g(\frac{1}{4}) = \frac{85}{256}$, $r_3 = g\left(\frac{85}{256}\right) \approx 0.37$.

d.) $m > 1$, vagyis a folyamat szuperkritikus, így számolni kell: a kihalás (=lefutás) valószínűsége a $z = g(z)$ egyenlet legkisebb nemnegatív megoldása. Esetünkben

$$z = \frac{1}{4} + \frac{1}{4}z + \frac{1}{4}z^2 + \frac{1}{4}z^3,$$

ami egy harmadfokú egyenlet z-re. Átszorozva és nullára redukálva

$$z^3 + z^2 - 3z + 1 = 0.$$

Tudjuk, hogy $z = 1$ megoldás, mert $z = 1$ mindig fixpontja a generátorfüggvénynek, de mi nem ezt a megoldást keressük. Viszont ha egyszer tudjuk, hogy 1 gyöke a baloldali polinomnak, ki tudunk emelni $(z-1)$-et: $z^3 + z^2 - 3z + 1 = (z-1)(z^2 + 2z - 1)$, így az egyenlet

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

A másodfokú tényező gyökei $-1 - \sqrt{2}$ és $-1 + \sqrt{2}$, vagyis az egyenlet összes gyöke 1, $-1 - \sqrt{2}$ és $-1 + \sqrt{2}$.
Mivel a folyamat szuperkritikus, a kihalás valószínűsége 1-nél kisebb, vagyis ez egyetlen 1-nél kisebb nemnegatív gyököt keressük. Így hát

$$\mathbb{P}(\text{kihalás}) = \sqrt{2} - 1 \approx 0.41$$

e.) $m > 1$, vagyis a folyamat szuperkritikus, így $\mathbf{E}N = \infty$.

f.) $m > 1$, vagyis a folyamat szuperkritikus így N pozitív valószínűséggel végtelen. Az ilyen elfajult val.változók generátorfüggvényéről nem beszéltünk, úgyhogy inkább hagyjuk.

3.HF: (Beadási határidő: 2015.10.26.)

HF 3.1 Egy béka úgy ugrál a számegyenesen, hogy minden másodpercben $\frac{1}{3}$ valószínűséggel helyben ugrik, $\frac{1}{3}$ valószínűséggel egyet ugrik előre, $\frac{1}{3}$ valószínűséggel pedig kettőt, az előzményektől függetlenül. Kezdetben a 0-ban van. Móricka a centrális határeloszlás tétel segítségével szeretné megbecsülni annak valószínűségét, hogy a béka 1000 ugrással legalább 1300-ig eljut. Legfeljebb mennyi lesz Móricka becslésének hibája a Berry-Esseen tétel szerint? *(Megj.: A tételben szereplő konstans egy 2010-es eredmény szerint választható $C = 0.4748$-nak.)*

Megoldás: Legyen X_i a béka i-edik ugrásának hossza. A feladat szerint az X_i-k függetlenek egymástól és a közös eloszlásuk $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 2) = \frac{1}{3}$. $S_n := X_1 + X_2 + \cdots + X_n$ a béka helye n ugrás után. Kelleni fog nekünk, hogy $m := \mathbf{E}X_i = 1$, $\sigma^2 := \mathrm{Cov}X_i = \mathbf{E}(X_i^2) - m^2 = \frac{0^2+1^2+2^2}{3} - 1^2 = \frac{2}{3}$ és

$$\delta := \mathbf{E}(|X_i - m|^3) = \frac{|0-1|^3 + |1-1|^3 + |2-1|^3}{3} = \frac{2}{3}.$$

Móricka a $\mathbb{P}(S_n \geq K)$ valószínűséget becsülte az $n = 1000$, $K = 1300$ esetben. (És ezt állította, hogy $\mathbb{P}(S_n \geq K) \approx 1 - \Phi\left(\frac{K-nm}{\sigma\sqrt{n}}\right) \approx 1 - \Phi(11.619) \approx 1.65 \cdot 10^{-31}$, de ezt a feladat nem kérdezte.)

A becslés hibája a Berry-Esseen tétel szerint legfeljebb

$$\frac{C\delta}{\sigma^3\sqrt{n}} = \frac{0.4748 \cdot \frac{2}{3}}{\sqrt{\frac{2}{3}}^3\sqrt{1000}} \approx 0.018 = 1.8\%.$$

HF 3.2 Egy béka úgy ugrál a számegyenesen, hogy minden másodpercben $\frac{1}{3}$ valószínűséggel helyben ugrik, $\frac{1}{3}$ valószínűséggel egyet ugrik előre, $\frac{1}{3}$ valószínűséggel pedig kettőt, az előzményektől függetlenül. Kezdetben a 0-ban van. Adjunk nagy eltérés becslést annak valószínűségére, hogy a béka 1000 ugrással legalább 1300-ig eljut.

Megoldás: Legyen X_i a béka i-edik ugrásának hossza. A feladat szerint az X_i-k függetlenek egymástól és a közös eloszlásuk $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 2) = \frac{1}{3}$. Legyen $n = 1000$ és $S_n := X_1 + X_2 + \cdots + X_n$ a béka helye n ugrás után. A feladat így a $\mathbb{P}(S_n \geq 1300)$ valószínűség becslése.

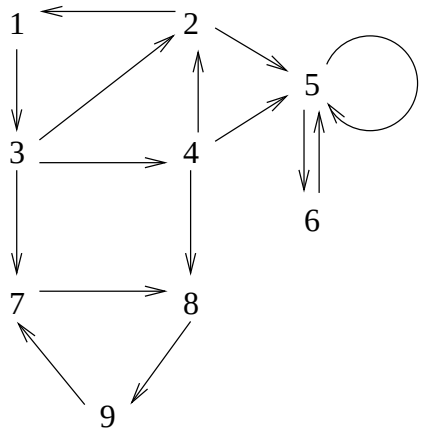
Mivel az X_i-k függetlenek és korlátosak, legkönnyebb a Hoeffding egyenlőtlenséggel becsülni. Ehhez $\mathbf{E}S_n = n\mathbf{E}X_i = 1000 \cdot 1 = 1000$, a korlátok $a_i := 0 \leq X_i \leq b_i := 2$, amiből $\sum_{i=1}^n (b_i - a_i)^2 = n(2-1)^2 = 1000 \cdot 4 = 4000$. Vagyis $t = 300$ választással a Hoeffding egyenlőtlenség azt mondja, hogy

$$\begin{aligned}
\mathbb{P}(S_n \geq 1300) &= \mathbb{P}(S_n \geq \mathbf{E}S_n + t) \leq \\
&\leq \exp\left\{-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right\} = e^{-\frac{2\cdot 300^2}{4000}} = \\
&= e^{-45} \approx 2.86 \cdot 10^{-20}.
\end{aligned}$$

4.HF: (Beadási határidő: 2015.11.09.)

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 meg, 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:

* Az $\{1,2,3,4\}$ osztály nyílt, ezért lényegtelen és átmeneti (azaz tranziens). Periódusa 1, mert pl. 1-ből 1-be 3 éa 4 lépésben is vissza lehet térni, ezeknek legnagyobb közös osztója pedig 1. Más szóval ez az osztály aperiodikus.
* Az $\{5,6\}$ osztály zárt (és véges), ezért lényeges és visszatérő, avagy rekurrens. Periódusa 1, vagyis aperiodikus.
* A $\{7,8,9\}$ osztály zárt (és véges), ezért lényeges és visszatérő, avagy rekurrens. Periódusa 3.

HF 4.2 **Második változat, egy sajtóhiba javítása után.** Pistike olyan számítógépes kalandjátékot játszik, ahol öt pálya van 1-től 5-ig számozva. Minden menetben az előzményektől függetlenül $\frac{1}{4}$ valószínűséggel sikerül neki 1 pályával feljebb jutni, $\frac{1}{2}$ valószínűséggel azonban egy pályával lejjebb csúszik - a maradék $\frac{1}{4}$ val.séggel marad a régi pályán. Kivétel ez alól, ha a legalsó (1-es) pályán van, mert akkor $\frac{3}{4}$ valószínűséggel marad (és továbbra is $\frac{1}{4}$ valószínűséggel jut eggyel feljebb), illetve ha a legfelső pályán van, mert akkor $\frac{1}{2}$ valószínűséggel marad (és továbbra is $\frac{1}{2}$ valószínűséggel csúszik eggyel lejjebb).

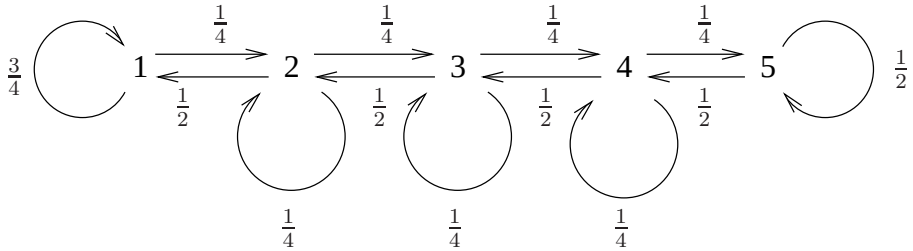
a.) Pistike az 1-es pályáról indul. Mennyi a valószínűsége, hogy 4 menet után már az 5-ös pályán játszik?

b.) Körülbelül mekkora a valószínűsége, hogy 100 menet után Pistike pont a 3-as pályán van? *(Szabad észrevenni, hogy az aktuális pálya száma születési-halálozási folyamat.)*

c.) Pistike a k-adik pályán mindig 2^k pontot szerez. Átlagosan hány pontot szerez menetenként hosszú távon?

d.) **Bónusz kérdés:** Mi a válasz a fenti kérdésekre, ha az elérhető pályák számára nincs felső korlát?

Megoldás: Jelöljük X_n-nel azt, hogy hanyadik pályán van Pistike n menet után. Ez



2. ábra. Pistike pályaszám-folyamatának gráf-reprezentációja

diszkrét idejű Markov lánc, állapottere $S = \{1,2,3,4,5\}$, gráf-reprezentációja a 2. ábrán látható, átmenetmátrixa

$$P = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

a.) $\mathbb{P}(X_4 = 5 \,|\, X_0 = 1) = P_{12}P_{23}P_{34}P_{45} = \frac{1}{4}\frac{1}{4}\frac{1}{4}\frac{1}{4} = \frac{1}{256} \approx 0.0039$, mert 4 lépésben 1-ből 5-be az egyetlen lehetséges út az $1 \to 2 \to 3 \to 4 \to 5$, vagyis minden lépésben fel kell jutnia.

b.) A Markov lánc véges állapotterű, irreducibilis és aperiodikus. 100 lépés hosszú idő, ezért a Markov láncok alaptétele szerint $\mathbb{P}(X_{100} = 3) \approx \pi_3$, függetlenül a kiinduló állapottól, ahol π az egyetlen stacionárius eloszlás. *(Megjegyzés: az indoklás fontos!).* Mivel ez a Markov lánc ugrani csak helyben vagy szomszédos állapotokba tud, ez egy születési halálozási folyamat, aminek a stacionárius eloszlása könnyen leolvasható a gráfról: minden $i = 1,2,3,4$-re

$$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i}$$

vagyis

$$\frac{\pi_1}{\pi_2} = \frac{P_{21}}{P_{12}} = \frac{1/2}{1/4} = 2 \quad , \quad \frac{\pi_2}{\pi_3} = \frac{P_{32}}{P_{23}} = \frac{1/2}{1/4} = 2$$
$$\frac{\pi_3}{\pi_4} = \frac{P_{43}}{P_{34}} = \frac{1/2}{1/4} = 2 \quad , \quad \frac{\pi_4}{\pi_5} = \frac{P_{54}}{P_{45}} = \frac{1/2}{1/4} = 2.$$

Ebből $\pi = c\begin{pmatrix}16 & 8 & 4 & 2 & 1\end{pmatrix}$, ahol a c normáló konstans olyan, hogy a sorösszeg 1 legyen, azaz $c = \frac{1}{31}$. Vagyis

$$\pi = \begin{pmatrix}\frac{16}{31} & \frac{8}{31} & \frac{4}{31} & \frac{2}{31} & \frac{1}{31}\end{pmatrix}.$$

A kérdésre a válasz tehát

$$\mathbb{P}(X_{100} = 3) \approx \pi_3 = \frac{4}{31} \approx 0.129$$

(Megjegyzés: A stacionárius eloszlás most is, mint mindig, megkereshető a $(P^T - I)\pi^T = 0$ lineáris egyenletrendszer megoldásával is, ahol I az egységmátrix. Esetünkben az egyenletrendszer

$$\begin{pmatrix} -\frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

ami a szokásos bővített mátrix jelöléssel

$$\left(\begin{array}{ccccc|c} -\frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{3}{4} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{4} & -\frac{1}{2} & 0 \end{array}\right).$$

Ennek a megoldása is arra vezet, hogy $\pi = c\begin{pmatrix}16 & 8 & 4 & 2 & 1\end{pmatrix}$, és nem sokkal több munka, mint a fenti (bár sokkal több helyet foglal a papíron).)

c.) Legyen az $f : S \to \mathbb{R}$ függvény olyan, hogy $f(k) = 2^k$, amit oszlopvektornak képzelünk:

$$f = \begin{pmatrix} 2 \\ 4 \\ 8 \\ 16 \\ 32 \end{pmatrix}.$$

A kérdés az

$$\frac{S_n}{n} = \frac{f(X_0) + f(X_1) + \cdots + f(X_{n-1})}{n}$$

időátlag határértéke, amint $n \to \infty$. Mivel a Markov lánc véges állapotterű és irreducibilis, az ergodtétel értelmében ez a határérték megegyezik az egyetlen stacionárius eloszlás szerint várható értékkel *(Az indoklás fontos!)*, vagyis a válasz

$$\begin{aligned} \mathbf{E}_\pi f &= \sum_{i \in S} \pi_i f(i) = \pi f = \left(\tfrac{16}{31} \quad \tfrac{8}{31} \quad \tfrac{4}{31} \quad \tfrac{2}{31} \quad \tfrac{1}{31}\right) \begin{pmatrix} 2 \\ 4 \\ 8 \\ 16 \\ 32 \end{pmatrix} = \\ &= \frac{16 \cdot 2 + 8 \cdot 4 + 4 \cdot 8 + 2 \cdot 16 + 1 \cdot 32}{31} = \frac{160}{31} \approx 5.16 \end{aligned}$$

d.) **Bónusz válasz:**

a.) Nem változik.

b.) A Markov lánc most is egy irreducibilis születési halálozási folyamat, az állapottér $S = \{1, 2, 3, \dots\}$ végtelen. Az egyetlen stacionárius eloszlás, *ha létezik*, most is olyan, hogy

$$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i},$$

ezúttal minden $i = 1, 2, 3, \dots$ -re. Ezt most úgy célszerű felírni, hogy $\pi_{i+1} = \frac{1}{2}\pi_i$ minden i-re, amiből

$$\pi = c\left(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\right),$$

már ha létezik. Előadásról pedig azt tudjuk, hogy a stacionárius eloszlás pontosan akkor létezik, ha a c normáló konstans megválasztható úgy, hogy a sorösszeg 1 legyen - vagyis pontosan akkor, ha a normálás nélkül a sorösszeg véges. Esetünkben $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1}{1-\frac{1}{2}} = 2 < \infty$, vagyis a stacionárius eloszlás létezik, $c = \frac{1}{2}$ és

$$\pi = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\right).$$

A Markov láncok alaptételének állítása most is érvényes, így

$$\mathbb{P}(X_{100} = 3) \approx \pi_3 = \frac{1}{8}.$$

c.) Most is $f(k) := 2^k$, ezúttal minden $k = 1, 2, 3, \dots$-re. Mivel az irreducibilis születési-halálozási folyamatunk stacionárius eloszlása létezik, az ergodtétel állítása most is érvényes (bár ezt nem mondtam), vagyis 1 valószínűséggel

$$\lim_{n \to \infty} \frac{f(X_0) + f(X_1) + \cdots + f(X_{n-1})}{n} = \mathbf{E}_\pi f = \sum_{i \in S} \pi_i f(i),$$

már amennyiben a jobboldali végtelen összeg értelmes. Esetünkben $\pi_i = \left(\frac{1}{2}\right)^i$, amiből

$$\sum_{i\in S} \pi_i f(i) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i 2^i = \sum_{i=1}^{\infty} 1 = \infty$$

értelmes, vagyis Pistike egy menetre eső átlagpontszáma hosszú távon végtelenhez tart.

5.HF: (Beadási határidő: 2015.11.30.)

HF 5.1 Legyen $X(t)$ folytonos idejű Markov lánc az $S = \{1, 2, 3\}$ állapottéren, ami egy rendszer állapotát modellezi. *Az időt mérjük percben.* A rendszer az 1-es állapotban átlagosan 10 másodpercet tartózkodik, majd $\frac{1}{2}$ valószínűséggel a 2-es, $\frac{1}{2}$ valószínűséggel pedig a 3-as állapotba ugrik. A 2-es állapotban átlagosan 20 másodpercet tartózkodik, majd $\frac{1}{3}$ valószínűséggel az 1-es, $\frac{2}{3}$ valószínűséggel pedig a 3-as állapotba ugrik. Végül a 3-as állapotból átlagosan 15 másodperc után ugrik el, $\frac{1}{4}$ valószínűséggel az 1-es, $\frac{3}{4}$ valószínűséggel pedig a 2-as állapotba.

a.) Írjuk fel a Markov lánc λ tartózkodási idő paraméter vektorát és a beépített diszkrét idejű Markov lánc Q átmenetmátrixát.

b.) Írjuk fel a Markov lánc G infinitezimális generátorát.

c.) Keressük meg a stacionárius eloszlás(ok)at.

d.) Ha a rendszer $t = 0$ -kor az 1-es állapotból indul, körülbelül mekkora valószínűséggel lesz 60 perc elteltével a 2-es állapotban? Miért?

e.) Hosszú távon az idő hány százalékát tölti a rendszer a 3-as állapotban? Miért?

Megoldás:

a.) Az egyes állapotokban a tartózkodási idő várható értéke percben mérve rendre $\frac{1}{6}$, $\frac{1}{3}$ illetve $\frac{1}{4}$. A tartózkodási idő paraméter, vagyis az elugrás rátája ennek reciproka, vagyis

$$\underline{\lambda} = (6; 3; 4).$$

A beépített diszkrét idejű Markov lánc átmenetmátrixa a szöveg alapján

$$Q = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{4} & \frac{3}{4} & 0 \end{pmatrix}.$$

b.) A G generátor főátlón kívüli elemei ($i \neq j$-re) $G_{ij} = \lambda_{ij} = \lambda_i Q_{ij}$, a főátló elemei pedig $G_{ii} = -\lambda_i$ (így a sorösszegek nullák). Jelen esetben

$$G = \begin{pmatrix} -6 & 3 & 3 \\ 1 & -3 & 2 \\ 1 & 3 & -4 \end{pmatrix}.$$

c.) A $G^T \pi^T = 0$ homogén lineáris egyenletrendszert kell megoldani, ahol az ismeretlen a $\pi = (\pi_1; \pi_2; \pi_3)$ sorvektor (így π^T oszlopvektor). A szokásos bővített mátrix jelöléssel az egyenletrendszer

$$\left(\begin{array}{ccc|c} -6 & 1 & 1 & 0 \\ 3 & -3 & 3 & 0 \\ 3 & 2 & -4 & 0 \end{array}\right).$$

Ezt Gauss eliminációval (vagy bárhogy máshogy) megoldva azt kapjuk, hogy megoldás a $\tilde{\pi} = (2; 7; 5)$ vektor, illetve ennek minden számszorosa. A végtelen sok

megoldás közül az lesz (a jelen esetben egyetlen) stacionárius megoldás, amelyikre a sorösszeg 1, vagyis

$$\pi = \left(\frac{2}{14}; \frac{7}{14}; \frac{5}{14}\right).$$

(Megjegyzés: az egyenletrendszert elvileg sokféleképpen meg lehet oldani, de mindig Gauss eliminációval érdemes – ha másért nem, hát azért, mert máshogy áttekinthetetlen, és senki nem fogja benne a hibát megkeresni.)

d.) Az állapottér véges és irreducibilis, a Markov lánc folytonos idejű (így periodicitási probléma nem lehet), $t = 60$ perc pedig hosszú idő (nagy valószínűséggel sok ugrás történik). Így a Markov láncok alaptétele szerint

$$\mathbb{P}(X(60) = 2 \,|\, X(0) = 1) \approx \pi_2 = \frac{7}{14} = \frac{1}{2}.$$

e.) Legyen az $f : S \to \mathbb{R}$ megfigyelhető mennyiség a 3-as állapot indikátorfüggvénye, vagyis $f(1) = f(2) = 0$, $f(3) = 1$. Ezt oszlovektorként célszerű írni:

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

Így az $f(X(t))$ véletlen függvény $\frac{1}{T}\int_0^T f(X(t))\,\mathrm{d}t$ időátlagára vagyunk kiváncsiak amikor T nagy. Mivel az állapottér véges és irreducibilis, az ergodtétel szerint ennek határértéke $T \to \infty$-re (majdnem biztosan) megegyezik az f függvény várható értékével az egyetlen stacionárius eloszlás szerint, ami sokféleképpen felírható:

$$\mathbf{E}_\pi f = \sum_{i \in S} \pi_i f(i) = \pi f = \begin{pmatrix} \frac{2}{14} & \frac{7}{14} & \frac{5}{14} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

és ha kiszámoljuk, persze az jön ki, hogy $\mathbf{E}_\pi f = \pi_3 = \frac{5}{14} \approx 0.357$. Vagyis hosszú távon a 3-as állapotban töltött idő aránya az összidőhöz 1 valószínűséggel

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X(t))\,\mathrm{d}t = \mathbf{E}_\pi f = \pi_3 = \frac{5}{14}.$$

HF 5.2 Két áramköri elem ellenállását szeretnénk összehasonlítani, de sajnos az ellenállást csak hibával terhelten tudjuk mérni. Tegyük fel, hogy egy mérés eredménye normális eloszlású valószínűségi változó, aminek várható értéke a tényleges ellenállás, szórása pedig a mérőműszer jellemzője és ismeretlen, de minden mérés esetén azonos. Az egyes mérések hibái függetlenek. Mindkét ellenállást megmértük néhányszor, és a következő adatokat kaptuk ($M\Omega$-ban):

„A” ellenállás: 1000, 1023, 983, 1009, 990, 987, 988, 1009, 996

„B” ellenállás: 996, 995, 1004, 994, 1017, 1003, 1014

a.) Döntsünk 95%-os szinten arról a nullhipotézisről, hogy az „A” ellenállás legalább akkora, mint a „B”.

b.) Döntsünk 95%-os szinten arról a nullhipotézisről, hogy a „B” ellenállás legalább akkora, mint az „A”.

Megoldás: $n_1 = 9$ elemű mintát vettünk az $X \sim \mathcal{N}(m_1, \sigma_1^2)$ valószínűségi változóból, $n_2 = 7$ eleműt pedig az $Y \sim \mathcal{N}(m_2, \sigma_2^2)$-ből. m_1, m_2, σ_1 és σ_2 is ismeretlen, de tudjuk, hogy $\sigma_1 = \sigma_2$. A nullhipotézis az a.) esetben $m_1 \geq m_2$, a b.) esetben pedig $m_2 \geq m_1$. Gyorsan kiszámoljuk az átlagokat:

$$\begin{aligned} \bar{x} &= \frac{1}{n_1}\sum_{i=1}^{n_1} x_i = \frac{8985}{9} \approx 998.33 \\ \bar{y} &= \frac{1}{n_2}\sum_{i=1}^{n_2} y_i = \frac{7023}{7} \approx 1003.29 \end{aligned}$$

Mivel $\bar{y} > \bar{x}$, a b.) feladat esetén az adatok megerősítik a nullhiptézist, vagyis ekkor a hipotézist további számolás és gondolkodás nélkül *elfogadjuk*.

Az a.) esetben viszont sajnos az $\bar{x}$ és $\bar{y}$ között tapasztalt eltérés ellentmondani látszik a hipotézisnek, így számolni kell, hogy eldőljön: szignifikáns-e ez az eltérés. Ehhez kétmintás egyoldali t-próbát kell végeznünk:

* Kétmintásat, mert a nullhipotézis *két* val.változó várható értékének összehasonlításáról szól.
* Egyoldalit, mert a nullhipotézis egyenlő*tlen*ség.
* t-próbát, mert a szórás ismeretlen. (De szerencsére a két szórás azonos: az általunk tanult kétmintás t-próbák csak ebben az esetben működnek.)

A képletgyűjtemény szerint a tesz-statisztika

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)s_1^{*2} + (n_2-1)s_2^{*2}}{n_1+n_2-2}}} \sqrt{\frac{n_1 n_2}{n_1+n_2}}.$$

Ennek kiszámolásához szükség lesz a korrigált tapasztalati szórás*négyzet*ekre: az adatokból

$$\begin{aligned} \sum_i x_i^2 &= 8971409 \\ s_1^2 &= \frac{1}{n_1}\sum_i x_i^2 - \bar{x}^2 \approx \frac{8971409}{9} - 998.33^2 \approx 157.78 \\ s_1^{*2} &= \frac{n_1}{n_1-1} s_1^2 \approx \frac{9}{8} 157.78 \approx 173 \\ \sum_i y_i^2 &= 7046587 \\ s_2^2 &= \frac{1}{n_2}\sum_i y_i^2 - \bar{y}^2 \approx \frac{7046587}{7} - 1003.29^2 \approx 73.06 \\ s_2^{*2} &= \frac{n_2}{n_2-1} s_2^2 \approx \frac{7}{6} 73.06 \approx 85.24 \end{aligned}$$

Ezeket felhasználva a teszt-statisztika

$$t \approx \frac{998.33 - 1003.29}{\sqrt{\frac{(9-1)173 + (7-1)85.24}{9+7-2}}} \sqrt{\frac{9 \cdot 7}{9+7}} \approx -1.242,$$

az elfogadási küszöb pedig a $df = n_1 + n_2 - 2 = 9 + 7 - 2 = 14$ szabadsági fokú t eloszlás $\varepsilon = 0.05$ höz tartozó $1 - \varepsilon$ kvantilise. (Mivel a hipotézist 95%-os szinten kell

vizsgálnunk, $\varepsilon = 0.05$. Mivel egyoldali próbáról van szó, az eloszlás *egyoldali* farkának kell ε-nak lenni: a táblázatban „one-tail”.) A táblázatból kiolvasva

$$K = t_\varepsilon = 1.761.$$

Döntés: $t \geq -K$, ezért a nullhipotézist az a.) esetben is *elfogadjuk*.