Sztochasztika 2 vizsga megoldókulcs Felsőbb matematika tárgy.

2015. január 20. 13:00. Munkaidő: $\leq$60 perc.

1. *(7 pont)* Egy koncessziós pályázat 10 fejezetből áll, a pályázók minden fejezetre legfeljebb 5 pontot kaphatnak. A bírálók a megítélt pontszámot fejezetenként kockadobással döntik el, azonos esélyt adva a $0, 1, 2, 3, 4, 5$ pontszámoknak. A felhívásra 10000 pályázat érkezik. Adjunk nagy eltérés becslést annak valószínűségére, hogy a pályázatok átlagos pontszáma eléri a 26-ot. *(Vigyázat: hányszor is gurítják el a bírálók azt a dobókockát?)*

   (A p paraméterű Bernoulli eloszlás Cramér féle rátafüggvénye

$$I(x) = x \ln\left(\frac{x}{1-x}\frac{1-p}{p}\right) + \ln\left(\frac{1-x}{1-p}\right) \qquad \textit{(ha } 0 < x < 1\textit{).)}$$

Megoldás: $n = 100000$ kockadobás történik, ennyi pályázat-fejezet pontszámát kell összeadni, amik függetlenek és egyenletesek a $\{0, 1, 2, 3, 4, 5\}$ halmazon. Legyen X_i ezek közül az i-edik ($i = 1, 2, \ldots, n$). Így $S_n := X_1 + \cdots + X_n$ az össz-pontszám, és a kérdés

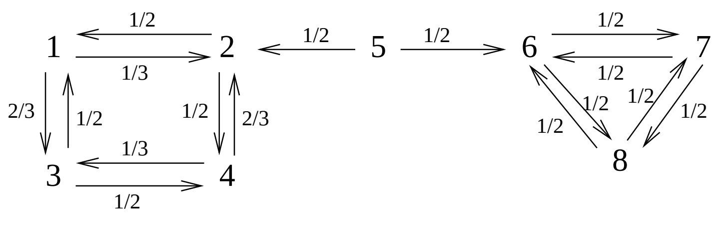
$$\mathbb{P}\left(\frac{S_n}{10000} \geq 26\right) = \mathbb{P}(S_n \geq 260000) =?$$

Az X_i-k függetlenek és korlátosak: $0 = a_i \leq X_i \leq b_i = 5$ minden i-re, vagyis a Hoeffding egyenlőtlenség használható a nagy eltérés becslésre. Ehhez $\sum_{i=1}^{n}(b_i - a_i)^2 = n(5-0)^2 = 25n = 2500000$, valamint $\mathbb{E}S_n = n\mathbb{E}X_i = n \cdot \frac{5}{2} = 250000$. Így $t := 10000$ választással

$$\begin{aligned} \mathbb{P}\left(\frac{S_n}{10000} \geq 26\right) &= \mathbb{P}(S_n \geq 260000) = \mathbb{P}(S_n \geq \mathbb{E}S_n + t) \leq \\ &\leq \exp\left\{\frac{-2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right\} = \exp\left\{\frac{-2 \cdot 10000^2}{2500000}\right\} = e^{-80} \approx 1.8 \cdot 10^{-35}. \end{aligned}$$

(Megjegyzés: a nagy eltérés becsléshez elvileg a Cramér tétel is használható lenne, de az itt szereplő, $\{0, \ldots, 5\}$-ön egyenletes eloszlás rátafüggvénye csúnya, és nem volt megadva. A feladatban megadott Bernoulli rátafüggvény itt nem jó semmire.)

2. *(8 pont)* Legyen az X_n diszkrét idejű Markov lánc gráf-reprezentációja a következő:



Adjuk meg közelítőleg az alábbi valószínűségeket. A válaszokat indokoljuk.

a.) *(2 pont)* $\mathbb{P}(X_{1000} = 7 \,|\, X_0 = 6) \approx ?$

b.) *(2 pont)* $\mathbb{P}(X_{1000} = 2 \,|\, X_0 = 1) \approx ?$

c.) *(2 pont)* $\mathbb{P}(X_{1000} = 2 \,|\, X_0 = 6) \approx ?$

d.) *(2 pont)* $\mathbb{P}(X_{1000} = 7 \,|\, X_0 = 5) \approx ?$

Megoldás: A Markov lánc NEM irreducibilis. Három osztálya közül kettő zárt: a $C_1 := \{1, 2, 3, 4\}$ osztály periódusa 2, a $C_2 := \{6, 7, 8\}$ osztály pedig aperiodikus, mert pl. 6-bó 6-ba vissza lehet jutni 2 és 3 lépésben is. Így a.) *(2 pont)* $\mathbb{P}(X_{1000} = 7 \,|\, X_0 = 6) \approx \frac{1}{3}$, mert 6-ból indulva örökre bent maradunk a C_2 osztályban. A Markov lánc ide megszorítva irreducibilis és aperiodikus, így a Markov láncok alaptétele szerint hosszú idő elteltével az eloszlás a stacionáriussal közelíthető. A C_2 irreducibilis komponensen a(z egyetlen) π stacionárius eloszlás szimmetria okból az egyenletes, így $\mathbb{P}(X_{1000} = 7 \,|\, X_0 = 6) \approx \pi_7 = \frac{1}{3}$.

b.) *(2 pont)* $\mathbb{P}(X_{1000} = 2 \,|\, X_0 = 1) = 0$, mert 1-ből indulva örökre bent maradunk a C_2 osztályban, ez viszont periodikus 2 periódussal, így páros sok lépésben csak 1-be és 3-ba juthatunk el.

c.) *(2 pont)* $\mathbb{P}(X_{1000} = 2 \,|\, X_0 = 6) = 0$, mert 6-ból indulva örökre bent maradunk a C_2 osztályban, vagyis 2-be nem lehet eljutni.

d.) *(2 pont)* $\mathbb{P}(X_{1000} = 7 \,|\, X_0 = 5) \approx \frac{1}{6}$, mert 5-ből indulva $\frac{1}{2}$ valószínűséggel az első lépésben a C_1 osztályba lépünk és ott is ragadunk, $\frac{1}{2}$ valószínűséggel viszont a C_2-be, és innen kezdve az a.) pont szerinti $\frac{1}{3}$ az esélyünk hosszú idő alatt 7-be érkezni.

3. *(10 pont)* Mérnök Mari újszülött gyermeke az édesanyja megfigyelése szerint háromféle állapotban lehet: 1 – „sír"; 2 – „alszik"; 3 – „eszik". A gyermek időnként véletlenszerűen ugrik át egyik állapotból a másikba, az előzményektől (a jelenre, mint feltételre nézve feltételesen) függetlenül, vagyis ő egy háromállapotú, folytonos idejű Markov lánc. Jelölje $X(t)$ a gyerek állapotát t időben. A beágyazott diszkrét idejű Markov-lánc Q átmenetvalószínűség mátrixa a következő:

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 0.8 & 0 & 0.2 \\ 0.8 & 0.2 & 0 \end{pmatrix}$$

Az állapotsorrend 1,2,3 balról-jobbra és felülről-lefelé. Feltesszük, hogy az 1-es állapotban marad $Exp(8)$ ideig, a 2-es állapotban $Exp(1)$ ideig és a 3-asban $Exp(5)$ ideig. (Mari az időt órában méri.)

a.) *(3 pont)* Írjuk fel a Markov lánc infinitezimális generátorát. Indokoljuk.

b.) *(3 pont)* Keressük meg a Markov lánc stacionárius eloszlásait.

c.) *(2 pont)* Az idő hány százalékában van az 1-es, 2-es, 3-as állapotokban? Miért?

d.) *(2 pont)* Ha a gyerek az 1-es állapotban van, Marinak óránként 100 hajszála hullik ki. Hasonlóan a 2-es állapotban 5, a 3-as állapotban 20 hajszálat veszít óránként. Körülbelül hány hajszála hullik ki Mérnök Marinak, mire a gyermek eléri a négyhetes kort? Miért?

Megoldás:

a.) Az egyes álapotokban a tartózkodási idők a feladat szövege szerint exponenciálisak, ahogy annak egy folytonos idejű Markov láncban lenni kell. Ezek paraméterei (rátái) éppen a tartózkodási idő paraméter vektort adják: $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3) = (8, 1, 5)$. Ezek a ráták kerülnek negatív előjellel a G infinitezimális generátor főátlójába. A főátlón kívüli elemekre $G_{ij} = \lambda_{ij} = \lambda_i Q_{ij}$. Ezeket mind beírva

$$G = \begin{pmatrix} -8 & 8 & 0 \\ 0.8 & -1 & 0.2 \\ 4 & 1 & -5 \end{pmatrix}.$$

b.) Meg kell oldani a $G^T \pi^T = 0$ lineáris egyenletrendszert. A lineáris algebrában szokásos mátrix-jelöléssel

$$\left(\begin{array}{ccc|c} -8 & 0.8 & 4 & 0 \\ 8 & -1 & 1 & 0 \\ 0 & 0.2 & -5 & 0 \end{array}\right) \overset{\text{2. sor+=1. sor}}{\sim} \left(\begin{array}{ccc|c} -8 & 0.8 & 4 & 0 \\ 0 & -0.2 & 5 & 0 \\ 0 & 0.2 & -5 & 0 \end{array}\right) \sim$$

Az utolsó egyenlet elhagyható, mert azonos a másodikkal. Az elsőhöz hozzáadjuk a másodikat 4-szer, majd mindkét sort leosztjuk a főátlóbeli elem abszolút értékével:

$$\sim \left(\begin{array}{ccc|c} -8 & 0 & 24 & 0 \\ 0 & -0.2 & 5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & 0 & 3 & 0 \\ 0 & -1 & 25 & 0 \end{array}\right).$$

Vagyis $\pi_1 = 3\pi_3$ és $\pi_2 = 25\pi_3$, amiből az egyenletrendszer egy lehetséges megoldása $\bar{\pi} = (3, 25, 1)$. Ezt lenormálva (hogy az elemek összege 1 legyen)

$$\pi = \left(\frac{3}{29}, \frac{25}{29}, \frac{1}{29}\right).$$

c.) Legyen $S = \{1, 2, 3\}$ az állapottér. Az, hogy az 1-es állapotban eltöltött idő az összes időnek hanyad része, az $f : S \to \mathbb{R}$,

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

megfigyelhető mennyiség időátlaga. Mivel a Markov lánc irreducibilis és véges állapotterű, az ergodtétel szerint az időátlag hosszú távon (1 valószínűséggel) megegyezik a stacionárius eloszlás szerinti átlaggal, vagyis $\pi \cdot f = \pi_1 = \frac{3}{29}$. Hasonlóan hosszú távon a 2-es állapotban az idő $\pi_2 = \frac{25}{29}$, a 3-asban pedig $\pi_2 = \frac{1}{29}$ hányadát tölti.

d.) Ezúttal először a $g : S \to \mathbb{R}$,

$$g = \begin{pmatrix} 100 \\ 5 \\ 20 \end{pmatrix}$$

függvény időátlagát keressük. Megint csak az ergodtétel értelmében az időátlag hosszú távon egy valószínűséggel $\pi \cdot g = 100\pi_1 + 5\pi_2 + 20\pi_3 = \frac{445}{29}$ (hajszál/óra). Négy hét az $4 \cdot 7 \cdot 24 = 672$ óra, vagyis Mari ezalatt körülbelül $672 \cdot \frac{445}{29} \approx 10312$ hajszálat veszít.