Felsőbb Matematika Villamosmérnököknek - Sztochasztika
2. ZH – megoldások

2019 ősz, 2019.12.04 18:00

Munkaidő: 90 perc. A nulladik feladat 0 pontos, a többi mind 9 pontot ér.

0. Írja rá a ZH-ra a *gyakorlat időpontját* (meg persze a saját nevét és Neptun-kódját is). Lehetséges helyes megoldások: csütörtök 8-10 (R504); csütörtök 10-12 (R517); péntek 10-12 (R516).

1. Egy véletlen rekurzív algoritmus futása során kezdetben elindul egyetlen folyamat, melynek sorsa kétféle lehet:

   i.) Eredményesen lefut és kilép anélkül, hogy újabb folyamatot indítana. Ennek valószínűsége $\frac{6}{10}$.

   ii.) Mielőtt lefutna, véletlen számú újabb folyamatot indít, amik ugyanolyanok, mint ő maga. Ezek száma lehet 1, 2, 3 vagy 4, azonos (vagyis $\frac{1}{10}$) valószínűséggel. Az eredeti folyamat csak akkor lép ki, amikor ezek mind lefutottak.

   Az algoritmus futása során elinduló egyes folyamatok mind ugyanolyan eloszlással hoznak létre újabb folyamatokat, függetlenül egymástól és az előzményektől. Mennyi a valószínűsége annak, hogy az eredeti folyamat előbb-utóbb kilép?

   Megoldás:

   Kezdetben egyetlen folyamat indul el – legyen ez egymaga a „nulladik generáció”. Az „első generáció” álljon a legelső folyamat által indított folyamatokból (ha van ilyen). A „második generáció” álljon az első generáció tagjai által indított folyamatokból. És így tovább, az $n+1$-edik generáció álljon az n-edik generáció tagjai által indított folyamatokból, $n = 0, 1, 2, \ldots$-re. Jelölje Z_n az n-edik generáció elemeinek számát. Így Z_n Galton-Watson elágazó folyamat, és a kérdés ennek kihalási valószínűsége.

   Az egylépéses utódszám-eloszlás várható értéke

   $$m = \frac{6}{10} \cdot 0 + \frac{1}{10} \cdot 1 + \frac{1}{10} \cdot 2 + \frac{1}{10} \cdot 3 + \frac{1}{10} \cdot 4 = 1.$$

   Vagyis a Galton-Watson folyamat kritikus (és nem elfajult), ezért a kihalás valószínűsége 1.

2. Móricka azt tervezi, hogy a zsebszámológépével generál $n = 3888$ darab független, $[0, 1]$-en egyenletes eloszlású véletlen számot, és ezekből átlagot számol. A centrális határeloszlás tétel (CHT) segítségével megbecsülte annak a valószínűségét, hogy az átlag legalább 0.54 lesz. Legfeljebb mennyi lehet ennek a becslésnek a hibája a ~~Hoeffding-egyenlőtlenség~~ Berry-Essen tétel szerint? (A tételbeli konstans választható $C = 0.48$-nak.)

   Megoldás: Ha X egyenletes eloszlású $[0, 1]$-en, akkor várható értéke $m := \mathbb{E}X = \frac{1}{2}$, szórásnégyzete

   $$\sigma^2 = \mathbb{E}\left((X-m)^2\right) = \int_0^1 \left(x - \frac{1}{2}\right)^2 \mathrm{d}x = \frac{1}{12},$$

   amiből a szórás

   $$\sigma = \sqrt{\frac{1}{12}},$$

harmadik centrált abszolút momentuma pedig

$$\delta := \mathbb{E}\left(|X-m|^3\right) = \int_0^1 \left|x-\frac{1}{2}\right|^3 \mathrm{d}x = \int_{-\frac{1}{2}}^{\frac{1}{2}} |y|^3 \mathrm{d}y = 2\int_0^{\frac{1}{2}} y^3 \mathrm{d}y = \frac{1}{32}.$$

Így a CHT becslés hibájára a Berry-Esseen tétel szerint

$$\text{hiba} \leq \frac{C\delta}{\sqrt{n}\sigma^3} = \frac{0.48 \cdot \frac{1}{32}}{\sqrt{3888}\sqrt{\frac{1}{12}}^3} = 0.01$$

3. Móricka azt tervezi, hogy a zsebszámológépével generál $n = 3888$ darab független, $[0,1]$-en egyenletes eloszlású véletlen számot, és ezekből átlagot számol. Adjunk nagy eltérés becslést annak valószínűségére, hogy az átlag legalább 0.54 lesz!

   (A p paraméterű Bernoulli eloszlás Cramér féle rátafüggvénye

   $$I_{B(p)}(x) = x\ln\frac{x}{p} + (1-x)\ln\frac{1-x}{q} \quad \textit{(ahol } q = 1-p \textit{ és } 0 < x < 1\textit{)}.$$

   A μ paraméterű exponenciális eloszlás Cramér féle rátafüggvénye

   $$I_{Exp(\mu)}(x) = \mu x - 1 - \ln(\mu x) \quad \textit{(ahol } x > 0\textit{)}.$$

)

   Megoldás: Legyen $i = 1, 2, \ldots, n$-re X_i az i-edik generált szám. Legyen $S_n = X_1 + \cdots + X_n$ a számok összege. Így a feladat a

   $$\mathbb{P}\left(\frac{S_n}{n} \geq 0.54\right) = \mathbb{P}(S_n \geq 0.54n)$$

   valószínűség becslése. Mivel az X_i-k függetlenek, azonos eloszlásúak, korlátosak és az eloszlásuk is ismert, elvileg a Hoeffding egyenlőtlenség és a Cramér tétel egyaránt használható lenne nagy eltérés becslésre. Sajnos azonban az egyenletes eloszlás Cramér féle rátafüggvénye nincs megadva, így ezt nem tudjuk használni. (Meg lehet próbálni kiszámolni, de az derül ki, hogy nincs rá szép képlet.)

   Így marad a Hoeffding egyenlőtlenség. Az X_i-kre korlátnak jó lesz $a_i := 0 \leq X_i \leq b_i := 1$, így

   $$\sum_{i=1}^{n}(b_i - a_i)^2 = n(1-0)^2 = n = 3888.$$

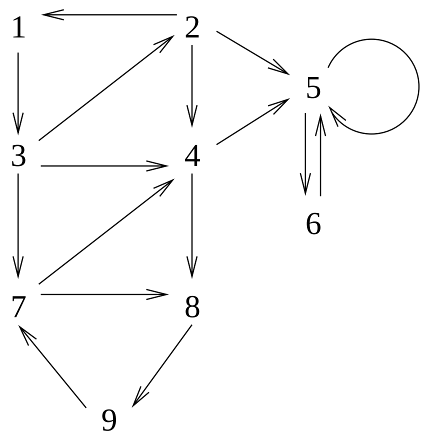
   Mivel $\mathbb{E}X_i = 0.5n = 1944$, $0.54n = \mathbb{E}S_n + 0.04n$, a keresett valószínűség pedig

   $$\mathbb{P}\left(\frac{S_n}{n} \geq 0.54\right) = \mathbb{P}(S_n \geq 0.54n) = \mathbb{P}(S_n \geq \mathbb{E}S_n + t)$$

   ahol $t = 0.04n = 116.64$. A Hoeffding egyenlőtlenség szerint pedig

   $$\mathbb{P}(S_n \geq \mathbb{E}S_n + t) \leq \exp\left\{-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right\} = \exp\left\{-\frac{2\cdot 116.64^2}{3888}\right\} \approx e^{-12.44} \approx 0.000004$$

4. 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



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

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

Megoldás:

osztály	zárt/nyílt	lényeges/lényegtelen	visszatérő/átmeneti	periódus
$\{1,2,3\}$	nyílt	lényegtelen	átmeneti	3
$\{4,7,8,9\}$	nyílt	lényegtelen	átmeneti	1
$\{5,6\}$	zárt	lényeges	visszatérő	1

Indoklás: Két állapot akkor tartozik ugyanazon osztályba, ha mindkettőből el lehet jutni a másikba. Egy osztály akkor zárt, ha nem vezet ki belőle él, egyébként nyílt. Így $\{1,2,3\}$ nyílt, mert kivezet belőle pl. a $2 \to 5$ él, $\{4,7,8,9\}$ nyílt, mert kivezet belőle a $4 \to 5$ él, viszont $\{5,6\}$ zárt. Véges osztályokra zárt=lényeges=visszatérő, nyílt=lényegtelen=átmeneti.

Az $\{1,2,3\}$ osztály periódusa 3, mert pl. 1-ből 1-be visszajutni csak $3,6,9,12,15,\ldots$ lépésben lehet, és ezek legnagyobb közös osztója 3. A $\{4,7,8,9\}$ osztály periódusa 1, mert pl. 4-ből 4-be vissza lehet térni 3 és 4 lépésben is, így a legnagyobb közös osztó 1. Az $\{5,6\}$ osztály periódusa 1, mert pl. 5-ből 5-be 1 lépésben is vissza lehet térni (így a visszatérési idők legnagyobb közös osztója 1).

5. Egy lépcsőházban 3 villanykörte van, és folyamatosan égnek – ha csak nincsenek éppen kiégve. Az egyes villanykörték élettartama független és exponenciális eloszlású, 1 év várható értékkel. A lépcsőházban évente átlag kétszer megjelenik a gondnok (Poisson folyamat szerint), és az összes kiégett körtét újra cseréli. Jelöljük $X(t)$-vel a t idő elteltével működő körték számát. Az időt mérjük években.

   a.) Adjuk meg az $X(t)$ Markov lánc állapotterét és az átmenetrátákat (ráta-mátrixot). *(Vigyázat: ha éppen 2 körte működik, milyen rátával ég ki közülük* ***valamelyik****?)*

   b.) Írjuk fel az infinitezimális generátort!

   c.) Keressük meg a Markov lánc stacionárius eloszlásait!

   d.) Tegnap délben minden körte működött. Körülbelül mennyi a valószínűsége, hogy pont 20 évvel később egy se fog működni?

   e.) A lépcsőházban akkor van zavaróan sötét, ha legfeljebb 1 körte világít. Hosszú távon az idő hány százalékában van zavaróan sötét?

Megoldás:

a.) Az állapottér $S = \{0, 1, 2, 3\}$. Ha egy körte ég, annak kiégési rátája a várható élettartam reciproka, esetünkben $1/1 = 1$, vagyis az 1 állapotból a 0 állapotba $\lambda_{10} = 1$ rátával ugrik a rendszer. Ha 2 körte ég, akkor már $1 + 1 = 2$ rátával fog kiégni valamelyik, így $\lambda_{21} = 2$. Ugyanígy $\lambda_{32} = 3$. A gondnok látogatásainak rátája 2, így a 0, 1 és 2 állapotból is $\lambda_{03} = \lambda_{13} = \lambda_{23} = 2$ rátával ugrik a rendszer 3-ba. (Ha a rendszer éppen a 3 állapotban van és jön a gondnok, akkor nem történik semmi.) Más átmenet nem lehetséges, így a ráta-mátrix

$$\underline{\underline{\lambda}} = \begin{pmatrix} * & 0 & 0 & 2 \\ 1 & * & 0 & 2 \\ 0 & 2 & * & 2 \\ 0 & 0 & 3 & * \end{pmatrix}.$$

(Ennek a főátlójában nincs semmi, mert helyben ugrás nincs.)

b.) Az infinitezimális generátort úgy kapjuk, hogy a $\underline{\underline{\lambda}}$ főátlóját kitöltjük úgy, hogy minden sorösszeg nulla legyen:

$$G = \begin{pmatrix} -2 & 0 & 0 & 2 \\ 1 & -3 & 0 & 2 \\ 0 & 2 & -4 & 2 \\ 0 & 0 & 3 & -3 \end{pmatrix}.$$

c.) a stacionárius eloszlás(ok) kereséséhez a $G^T \pi^T = 0$ lineáris egyenletrendszert kell megoldani:

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

Ezt kiolvasva $\pi_0 = \frac{1}{2}\pi_1$, $\pi_1 = \frac{2}{3}\pi_2$, $\pi_2 = \frac{3}{4}\pi_4$, vagyis pl. $\pi_3 = 4$ választással kapjuk, hogy egy megoldás a $\tilde{\pi} = (1, 2, 3, 4)$ vektor, és persze megoldás ennek minden számszorosa is. Minket az egyetlen normált megoldás érdekel (amire $\sum_{k \in S} \pi_k = 1$), vagyis

$$\pi = \frac{1}{10}\tilde{\pi} = \begin{pmatrix} \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} \end{pmatrix}.$$

d.) A kérdés

$$\mathbb{P}(X(t) = 0 \,|\, X(0) = 3) = P_{30}(t) = ?$$

ahol $t = 20$ hosszú idő. A Markov lánc véges állapotterű és irreducibilis (valamint folytonos idejű, így periodikus nem lehet), ezért a Markov láncok alaptétele szerint a $\mathbb{P}(X(t) = 0)$ valószínűséget a stacionárius eloszlás szerinti π_0 valószínűséggel közelíthetjük (a kiinduló állapottól függetlenül), ahol π az egyetlen stacionárius eloszlás. Így a kérdésre a válasz

$$\mathbb{P}(X(20) = 0 \,|\, X(0) = 3) \approx \pi_0 = \frac{1}{10}.$$

e.) Legyen $f : S \to \mathbb{R}$ a „zavaróan sötét van" esemény indikátora, vagyis $f(0) = f(1) = 1$, $f(2) = f(3) = 0$. Vektor-jelöléssel

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

Így a $t \in [0, T]$ időtartam alatt a 0 és 1 állapotok valamelyikében eltöltött idő $\int\limits_0^T f(X(t))\mathrm{d}t$, a kérdés pedig, hogy az idő hány százalékát tölti a rendszer a 0 és 1 állapotok valamelyikében, az $f(X(t))$ időátlaga:

$$\overline{f(X)} = \lim_{T\to\infty} \frac{1}{T} \int\limits_0^T f(X(t))\mathrm{d}t = ?$$

A Markov lánc véges állapotterű és irreducibilis, ezért az ergodtétel szerint egy valószínűséggel

$$\overline{f(X)} = \mathbb{E}_\pi f = \sum_{i\in S} \pi_i f(i) = \pi f,$$

ahol π az egyetlen stacionárius eloszlás. Esetünkben

$$\overline{f(X)} = \pi f = \begin{pmatrix} \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{3}{10}.$$