Sztochasztika 2 félvizsga megoldások

Felsőbb matematika informatikusoknak – Sztochasztika

2015. június 2. 8:00. Munkaidő: 70 perc.

1. Legyen Z_k Galton-Watson elágazó folyamat, ahol az egylépéses utódszám-eloszlás generátorfüggvénye $g(z) = e^{z-1}$. Mennyi a valószínűsége, hogy a folyamat előbb-utóbb kihal?

   Megoldás: Az egylépéses utódszámeloszlás várható értéke $m = g'(1) = 1$, vagyis a folyamat kritikus. Ezért a kihalás valószínűsége 1.

2. Juliska a körmét minden nap más színűre festi. Vörös, narancs és barna között váltogat. Narancs után mindig barna következik, barna után viszont érmedobással dönt arról, hogy vörös vagy narancs következzen-e. Vörös után kockát dob: ha az eredmény 6-os, akkor barna következik, egyébként narancs.

   a.) Írjuk fel Juliska körme színének, mint Markov láncnak az átmenetmátrixát!

   b.) Ha Juliska körme május 1-én vörös, mennyi a valószínűsége, hogy május 5-én is vörös?

   c.) A napok hanyad részében lesz vörös, narancs illetve barna Juliska körme hosszú távon?

   Megoldás:

   a.) Jelöljük az állapotokat számokkal, mondjuk 1: vörös; 2: narancs; 3: barna. Így az átmenetmátrix

   $$P = \begin{pmatrix} 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

   b.) Jelöljük a Markov láncot X_n-nel. Ha mondjuk május 1-e a nulladik nap, akkor május 5-e a negyedik, vagyis a kérdés $\mathbb{P}(X_4 = 1 | X_0 = 1) = ?$ Ez a P^4 mátrix $(1,1)$ eleme: $\mathbb{P}(X_4 = 1 | X_0 = 1) = P^4_{11}$. Ezt kiszámolhatjuk mondjuk úgy, hogy $P^4 = (P^2)^2$, Ebből persze nem kell minden elemet kiszámolni - ami nem kell, $*$-gal jelölöm:

   $$P^2 = \begin{pmatrix} \frac{1}{12} & \frac{1}{12} & \frac{5}{6} \\ \frac{1}{2} & * & * \\ 0 & * & * \end{pmatrix}, \quad P^4 = \begin{pmatrix} \frac{7}{144} & * & * \\ * & * & * \\ * & * & * \end{pmatrix}.$$

   Ebből $\mathbb{P}(X_4 = 1 | X_0 = 1) = P^4_{11} = \frac{7}{144} \approx 0.048611$.

   c.) Az ergodtételt fogjuk használni, ehhez szükség van a Markov lánc stacionárius eloszlására. (Mivel a Markov lánc irreducibilis és végea állapotterű, pontosan egy stacionárius eloszlása van.) Meg kell oldanunk a $(P^T - I)\pi^T$ homogén lineáris egyenletrendszert. A szokásos mátrix-jelöléssel

   $$\left(\begin{array}{ccc|c} -1 & 0 & \frac{1}{2} & 0 \\ \frac{5}{6} & -1 & \frac{1}{2} & 0 \\ \frac{1}{6} & 1 & -1 & 0 \end{array}\right).$$

   Ennek egy lehetséges megoldása pl. $\tilde{\pi} = (6\,11\,12)$, egyetlen normált megoldása pedig

   $$\pi = \begin{pmatrix} \frac{6}{29} & \frac{11}{29} & \frac{12}{29} \end{pmatrix} \approx \begin{pmatrix} 0.20690 & 0.37931 & 0.41379 \end{pmatrix}.$$

   Mivel a Markov lánc irreducibilis és véges állapotterű, az ergodtételt az egyes állapotok indikátoraira alkalmazva azt kapjuk, hogy az i állapotban hosszú távon az idő π_i hányadát tölti. Vagyis Juliska körme az idő $\pi_1 \approx 20.7\%$-ában vörös, $\pi_2 \approx 37.9\%$-ában naracs és $\pi_3 \approx 41.4\%$-ában barna.

3. Egy egyszerű jelfeldolgozó eszköz az egyes beérkező jeleket független, exponenciális eloszlású véletlen idők alatt dolgozza fel. A feldolgozási idő várható értéke 1 másodperc (vagyis $\frac{1}{60}$ perc). Amíg egy bejövő jel feldolgozása zajlik, addig az esetlegesen beérkező újabb jeleket az eszköz figyelmen kívül hagyja (vagyis nincs feldolgozási sor). A beérkező jelek Poisson folyamat szerint érkeznek, percenként átlagosan 2. Az eszköz így kétféle állapotban lehet: „szabad, passzív, jelre vár”, illetve „foglalt, feldolgozás folyamatban, nem figyel”.

   Modellezzük az eszköz állapotát folytonos idejű Markov lánccal. Az időt mérjük percben.

a.) Írjuk fel a Markov lánc infinitezimális generátorát. Indokoljuk.

b.) Az eszköz a működése első pillanatában szabad. Közelítőleg mennyi a valószínűsége, hogy tíz óra elteltével éppen foglalt lesz? Miért?

c.) Az eszköz teljesítményfelvétele passzív állapotban $1W$, feldogozás során viszont $10W$. Mennyi az átlagos teljesítményfelvétel hosszú távon? Miért?

Megoldás: Az állapotokat jelöljük számokkal: legyen $S = \{0, 1\}$ ahol 0 a passzív, 1 pedig az aktív állapot. Jelöljük a rendszer állapotát t idő elteltével X_t-vel. Mivel csak két állapot van, ugrani persze 0-ból csak 1-be, 1-ből pedig csak 0-ba lehet.

a.) Az 1-ből 0-ba ugrás rátája $\lambda_{10} = 60$, mert a feldolgozással eltöltött idő várható értéke $\frac{1}{\lambda_1} = \frac{1}{\lambda_{10}} = \frac{1}{60}$. A 0-ból 1-be ugrás rátája $\lambda_{01} = 2$, mert ilyen rátájú Poisson folyamat szerint érkeznek a jelek. Ezek a λ_{ij}-k lesznek a G infinitezimális generátor főátlón kívüli elemei. A főátlót pedig úgy töltjük ki, hogy minden sorösszeg 0 legyen. Így

$$G = \begin{pmatrix} -2 & 2 \\ 60 & -60 \end{pmatrix}.$$

b.) Mivel a Markov lánc véges állapotterű, irreducibilis és folytonos idejű, a Markov láncok alaptétele szerint hosszú idő elteltével az egyes állapotok valószínűségei tartanak a (z egyetlen) stacionárius eloszlás szerinti súlyokhoz. 10 óra azaz 600 perc pedig (ilyen ráták mellett) hosszú idő. Ezért keressük a $\pi = (\pi_0\ \pi_1)$ stacionárius eloszlást a $G^T \pi^T = 0$ lineáris egyenletrendszer megoldásával:

$$\begin{pmatrix} -2 & 60 \\ 2 & -60 \end{pmatrix} \begin{pmatrix} \pi_0 \\ \pi_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

avagy a lineáris algebrában szokásos tömör jelöléssel

$$\left(\begin{array}{cc|c} -2 & 60 & 0 \\ 2 & -60 & 0 \end{array}\right).$$

A két egyenlet egymásnak -1-szerese, így elég mondjuk az elsőt nézni: $-2\pi_0 + 60\pi_1 = 0$, amiből $\pi_0 = 30\pi_1$. Például $\pi_1 = 1$ választással is megkapjuk az egyenletrendszer egy lehetséges megoldását: $\tilde{\pi} = (30\ 1)$. A keresett stacionárius eloszlás ennek olyan konstansszorosa, amiben az elemek összege 1:

$$\pi = \begin{pmatrix} \frac{30}{31} & \frac{1}{31} \end{pmatrix}.$$

A Markov láncok alaptétele szerint tehát $\mathbb{P}(X_{600} = 1) \approx \pi_1 = \frac{1}{31} \approx 0.0323$.

c.) A pillanatnyi teljesítményfelvétel a t időpontban $f(X_t)$, ahol az $f : S \to \mathbb{R}$ megfigyelhető mennyiség olyan, hogy $f(0) = 1$ és $f(1) = 10$. Ezt célszerű oszlopvektorként írni:

$$f = \begin{pmatrix} 1 \\ 10 \end{pmatrix}.$$

Mivel a Markov lánc véges állapotterű és irreducibilis, az ergodtétel értelmében $f(X_t)$ időátlaga hosszú távon (1 valószínűséggel) tart az egyetlen π stacionárius eloszlás szerinti várható értékhez:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X_t)\,\mathrm{d}\,t = \sum_{i\in S} \pi_i f(i) = \pi f = \begin{pmatrix} \frac{30}{31} & \frac{1}{31} \end{pmatrix} \begin{pmatrix} 1 \\ 10 \end{pmatrix} = \frac{30}{31} \cdot 1 + \frac{1}{31} \cdot 10 = \frac{40}{31} \approx 1.29.$$

4. Egy kis telefonközpontba érkező, egymást követő hívások között eltelt idő mindig exponenciális eloszlású 1 perc várható értékkel, és független az előzményektől. Adjunk nagy eltérés becslést annak valószínűségére, hogy reggel 8 órától számítva a 400-adik hívásra kevesebb, mint 5 órát kell várni.

(Segítség: a λ paraméterű exponenciális eloszlás Cramér féle rátafüggvénye

$$I(x) = \lambda x - \ln(\lambda x) - 1 \quad \textit{(ha } x > 0\textit{)}.$$

A λ paraméterű Poisson eloszlás Cramér féle rátafüggvénye

$$I(x) = x\ln(x/\lambda) - x + \lambda \quad \textit{(ha } x > 0\textit{).)}$$

1. Megoldás: Legyen $n = 400$ és $X_1, X_2, \ldots, X_n$ független azonos 1 paraméterű exponenciális eloszlású valószínűségi változók: azt jelentik, hogy az egyes hívások között mennyi idő telik el (percben). Így $S_n := X_1 + \cdots + X_n$ a 400-adik hívás ideje, és a kérdés $\mathbb{P}(S_n \leq 300)$. Erre a Hoeffding-egyenlőtlenség *nem alkalmazható*, mert az X_k-k nem korlátosak. Marad a Cramér tétel. Ehhez a kérdéses valószínűséget $\mathbb{P}(S_n \leq 300) = \mathbb{P}(\frac{S_n}{n} \in (0, \frac{3}{4}]) = \mathbb{P}(\frac{S_n}{n} \in (a, b])$ alakba írjuk. Mivel $\mathbb{E}X_k = m$-re $b < m$, a Cramér tétel szerint (az exponenciális eloszlás rátafüggvényét használva $\lambda = 1$-gyel)

$$\mathbb{P}(\frac{S_n}{n} \in (a, b]) \lessapprox e^{-nI(b)} = e^{-400I(\frac{3}{4})} \approx e^{-15.07} \approx 2.84 \cdot 10^{-7}.$$

2. Megoldás: Vegyük észre, hogy a hívások Poisson folyamat szerint érkeznek, ezért az 1 perc alatt érkező hívások száma Poisson eloszlású $\lambda = 1$ várható értékkel, és az egyes percek függetlenek. Így ha $n = 300$ és $S_n = X_1 + \cdots + X_n$ az 5 óra alatt befutott hívások száma, ahol $X_k \sim Poi(1)$, akkor a kérdés $\mathbb{P}(S_n \geq 400)$. Mivel $\frac{4}{3} > m = \mathbb{E}X_k = 1$, a Cramér tétel szerint (a Poisson eloszlás rátafüggvényét használva $\lambda = 1$-gyel)

$$\mathbb{P}(S_n \geq 400) = \mathbb{P}(\frac{S_n}{n} \in [\frac{4}{3}, \infty)) \lessapprox e^{-300 \cdot I(\frac{4}{3})} \approx e^{-15.07} \approx 2.84 \cdot 10^{-7}.$$

3. Megoldás: Pontosan ugyanezt kapjuk akkor is, ha egybevesszük az 5 óra alatt érkező összes hívást: a 300 perc alatt érkező hívások száma Poisson eloszlású $\lambda = 300$ várható értékkel. Így alkalmazhatjuk a Cramér tételt az $S_n = X_1$ egytagú összegre ($n = 1$), ahol $X_1 \sim Poi(300)$, és a kérdés $\mathbb{P}(S_n \geq 400)$. Mivel $400 > m = \mathbb{E}X_1 = 300$, a Cramér tétel szerint (a Poisson eloszlás rátafüggvényét használva $\lambda = 300$-zal)

$$\mathbb{P}(S_n \geq 400) = \mathbb{P}(\frac{S_n}{n} \in [400, \infty)) \leq e^{-1 \cdot I(400)} \approx e^{-15.07} \approx 2.84 \cdot 10^{-7}.$$