Felsőbb Matematika Villamosmérnököknek – Sztochasztika
2. pótZH megoldások

2017 ősz, 2017.12.12 13:00

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

0. Írja rá a ZH-ra a *gyakorlatvezető nevét* és, ha ez nem egyértelmű, a *gyakorlat időpontját* (meg persze a saját nevét és Neptun-kódját is). Lehetséges helyes megoldások:

   a.) (0 pont) Horváth Illés, páratlan heteken (péntek, IB145)

   b.) (0 pont) Kói Tamás, páratlan heteken (péntek, QBF10)

   c.) (0 pont) Patkó Richárd, páratlan heteken (péntek, IB147)

   d.) (0 pont) Patkó Richárd, páros heteken (péntek, IB147)

1. Használható-e a Hoeffding-egyenlőtlenség, és használható-e a Cramér nagy eltérés tétel a $P(X_1+X_2+\cdots+X_n < K)$ valószínűség becslésére (trükközés nélkül) az alábbi esetekben? *A válaszokat indokoljuk!*

   a.) Az X_k-k független 1 paraméterű exponenciálisok.

   b.) Az X_k-k független és azonos, de ismeretlen eloszlásúak, viszont $P(2 \le X_k \le 5) = 1$, továbbá ismert a várható értékük és a szórásuk.

   c.) X_k egyenletes a $[0,1]$ intervallumon, és az X_k-k függetlenek.

   d.) X_k egyenletes a $[0,k]$ intervallumon, és az X_k-k függetlenek.

   e.) Jancsi egy szabályos érmét dobál. X_k legyen 1, ha a k-adik és a $k+1$-edik dobás is fej, egyébként pedig legyen 0.

   Megoldás: A Hoeffding-egyenlőtlenség akkor használható, ha az X_k-k függetlenek, korlátosak és az összeg várható értéke ismert. A Cramér tétel akkor használható, ha az X_k-k függetlenek, azonos eloszlásúak, és az eloszlásuk ismert (plusz egy enyhe technikai feltétel a momentum-generáló függvényről). Ennek alapján

	Hoeffding	Cramér
a.)	NEM, mert nem korlátosak	IGEN
b.)	IGEN	NEM, mert ismeretlen az eloszlás
c.)	IGEN	IGEN
d.)	IGEN	NEM, mert nem azonos eloszlásúak
e.)	NEM, mert nem függetlenek	NEM, mert nem függetlenek

2. 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 legfeljebb 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 $\mathbb{P}(S_n \le 300)$. Erre a Hoeffding-egyenlőtlenség *nem alkalmazható*, mert az X_k-k nem korlátosak. Marad a Cramér tétel. Ehhez a kérdéses valószínűséget $\mathbb{P}(S_n \le 300) = \mathbb{P}(\frac{S_n}{n} \in (0, \frac{3}{4}]) = \mathbb{P}(\frac{S_n}{n} \in (a, b])$ alakba írjuk. Mivel $\mathbb{E}X_k = m$-re $b < m$, a Cramér tétel szerint (az exponenciális eloszlás rátafüggvényét használva $\lambda = 1$-gyel)

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

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 $\mathbb{P}(S_n \ge 400)$. Mivel $\frac{4}{3} > m = \mathbb{E}X_k = 1$, a Cramér tétel szerint (a Poisson eloszlás rátafüggvényét használva $\lambda = 1$-gyel)

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

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 $\mathbb{P}(S_n \ge 400)$. Mivel $400 > m = \mathbb{E}X_1 = 300$, a Cramér tétel szerint (a Poisson eloszlás rátafüggvényét használva $\lambda = 300$-zal)

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

3. Jancsi és Juliska randit beszélt meg a Kököjszi utca és a Boborján utca kereszteződéséhez. Azt azonban nem beszélték meg, hogy a négy sarok közül melyiken találkozzanak. Jancsi pontban 11 órakor érkezik az északnyugati sarokhoz, majd keresni kezdi Juliskát. A négy gyalogos-lámpa percenként egyszer, egyszerre vált zöldre. Ilyenkor Jancsi $\frac{1}{4}$ valószínűséggel marad, ahol volt, $\frac{1}{4}$ valószínűséggel órajárás-irányba megy át a zebrán, $\frac{1}{2}$ valószínséggel pedig órajárással ellentétes irányban. Eközben Juliska órákat késik, így Jancsi hosszasan bolyong a négy sarok között. Jelölje X_n Jancsi helyét (vagyis hogy melyik sarkon áll) n perc elteltével.

   a.) (2 pont) Adjuk meg az X_n Markov lánc állapotterét és átmenet-valószínűség-mátrixát.

   b.) (2 pont) Mennyi a valószínűsége, hogy Jancsi két perc elteltével ugyanott van, mint a legelején?

   c.) (3 pont) Egy óra elteltével megközelítőleg mekkora valószínűséggel találjuk Jancsit a délkeleti sarkon?

   d.) (2 pont) A magas házak árnyékot vetnek a délkeleti és a délnyugati sarokra, az északkeleti és az északnyugati sarok viszont napos. Hosszú távon az idő hány százalékát tölti Jancsi napon?

Megoldás:

a.) Az állapottér legyen $S = 1, 2, 3, 4$, és számozzuk a sarkokat az északnyugatitól kezdve, órajárásal ellentétes irányban. Így az átmenetmátrix

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

b.)

$$(P^2)_{11} = \begin{pmatrix} 1/4 & 1/2 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} 1/4 \\ 1/4 \\ 0 \\ 1/2 \end{pmatrix} = \frac{1}{4}\frac{1}{4} + \frac{1}{2}\frac{1}{4} + \frac{1}{4}\frac{1}{2} = \frac{5}{16},$$

amit persze úgy is el lehet mondani, hogy $\frac{1}{4} \cdot \frac{1}{4}$ valséggel marad végig ahol volt, $\frac{1}{2} \cdot \frac{1}{4}$ valséggel elmegy órairányba aztán visszajön, $\frac{1}{4} \cdot \frac{1}{2}$ valséggel pedig fordítva.

c.) A Markov lánc irreducibilis, aperiodikus és véges állapotterű, egy óra pedig hosszú idő. Így a Markov láncok alaptétele szerint közelíthetünk a stacionárius eloszlással, vagyis oldjuk meg a $(P^T - I)\pi^T = 0$ lineáris egyenletrendszert. **A transzponálás nagyon fontos.** Az egyenletrendszer mátrixos alakban, az áttekinthetőség kedvéért néggyel végigszorozva:

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

Innentől

i.) Szabad megsejteni, hogy szimmetria-okból a stacionárius eloszlás az egyenletes, aztán leellenőrizni, hogy a $\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4}$ tényleg kielégíti az egyenletrendszert, *vagy*

ii.) szabad megoldani az egyenletrendszert.

Mindenképpen arra jutunk, hogy $\pi = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$. A feladat kérdésére a válasz $\pi_3 = \frac{1}{4}$.

d.) Legyen $f : S \to \{0; 1\}$ a napon levés indikátorfüggvénye, vagyis

$$f(x) = \begin{cases} 1, & \text{ha } x = 1 \text{ vagy } x = 4, \\ 0, & \text{ha nem.} \end{cases}$$

A Markov lánc irreducibilis és véges állapotterű, így az ergodtétel szerint $f(X_n)$ időátlaga hosszú távon $\sum_{i \in S} \pi_i f(i) = \pi_1 + \pi_4 = \frac{1}{2}$.

4. Egy kisbolt parkolójában 3 autónak van hely. A parkolóhoz Poisson-folyamat szerint érkeznek az autós vevők, átlagosan 5 percenként. Ha a parkoló tele van, akkor továbbmennek, ha pedig van hely, akkor leparkolnak és bemennek a boltba, ahol exponenciális eloszlású véletlen időt töltenek el, 5 perc várható értékkel, egymástól függetlenül. Vásárlás után azonnal autóba ülnek és elhajtanak. Kezdetben a parkoló üres. Jelölje X_t $(t \geq 0)$ a parkolóban lévő autók számát t perc elteltével.

a.) (2 pont) Modellezzük X_t-t folytonos idejű Markov lánccal. Adjuk meg az állapotteret és az infinitezimális generátort. (Vigyázat: érdemes észnél lenni. Két bent lévő vevő *egyike* könnyebben elmegy, mint egy vevő önmaga.)

b.) (2 pont) Számoljuk ki X_t stacionárius eloszlását.

c.) (2 pont) Hosszú idő elteltével közelítőleg mennyi a valószínűsége, hogy a parkolót üresen találjuk?

d.) (2 pont) Hosszú idő átlagában hány autó áll a parkolóban?

e.) (1 pont) A potenciális autós vevők hány %-át veszíti el a bolt amiatt, hogy kicsi a parkolója?

Megoldás:
X_t véges állapotterű születési-halálozási folyamat. Az időt percben mérjük, így a felfelé ugrás rátája (autó jön) mindig $\frac{1}{5}$, hacsak nem tele van a parkoló, a lefele ugrás rátája (autó megy) pedig az i állapotból $i \cdot \frac{1}{5}$, $i = 0, 1, 2, 3$.

a.) Az állapottér $S = \{0; 1; 2; 3\}$, a generátor

$$G = \begin{pmatrix} -1/5 & 1/5 & 0 & 0 \\ 1/5 & -2/5 & 1/5 & 0 \\ 0 & 2/5 & -3/5 & 1/5 \\ 0 & 0 & 3/5 & -3/5 \end{pmatrix}.$$

b.) A születési-halálozási folyamat stacionárius eloszlása a szomszédos állapotoknak olyan relatív súlyt ad, ami reciproka az egymásba való átugrások rátái arányának. Vagyis $\pi_0 : \pi_1 = 1 : 1$, $\pi_1 : \pi_2 = 2 : 1$, $\pi_2 : \pi_3 = 3 : 1$. Összesítve $\pi_0 : \pi_1 : \pi_2 : \pi_3 = 6 : 6 : 3 : 1$. Az aránysort lenormálva

$$\pi = \begin{pmatrix} \frac{6}{16} & \frac{6}{16} & \frac{3}{16} & \frac{1}{16} \end{pmatrix}.$$

Persze ugyanez jön ki, ha megoldjuk az $G^T \pi^T = 0$ egyenletrendszert (**a transzponálás nagyon fontos**), vagyis azt, hogy (az átláthatóság kedvéért 5-tel végigszorozva)

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

c.) A Markov lánc folytonos idejű, irreducibilis és véges állapotterű, egy óra pedig hosszú idő. Így a Markov láncok alaptétele szerint hosszú idő elteltével a kiindulási állapottól függetlenül a stacionárius eloszlással közelítünk: $\mathbb{P}(X_t = 0 \,|\, X_0 = 0) \approx \pi_0 = \frac{6}{16} = 37.5\%$.

d.) A Markov lánc irreducibilis és véges állapotterű. Így az ergodtétel értelmében az időátlag a stacionárius eloszlás szerinti várható érték, vagyis

$$\sum_{i \in S} i \cdot \pi_i = 0 \cdot \frac{6}{16} + 1 \cdot \frac{6}{16} + 2 \cdot \frac{3}{16} + 3 \cdot \frac{1}{16} = \frac{15}{16} \approx 0.94.$$

(Ha valaki mindenáron az állapottéren értelmezett valós értékű függvényre akarja az ergodtételt alkalmazni, tekintse az $f : S \to \mathbb{R}$, $f(i) = i$ függvényt.)

e.) A parkoló az idő $\pi_3 = \frac{1}{16}$-ában van tele, tehát az autósoknak pontosan azt az $\frac{1}{16}$-át azaz 6.25%-át veszítjük el, aki ezalatt jön. Másképpen számolva: percenként átlagosan $\frac{1}{5}$ autós jön arra, de a Markov lánc felfelé ugrásainak száma (vagyis a ténylegesen leparkoló autók száma) időátlagban csak $\pi_0 A_{01} + \pi_1 A_{12} + \pi_2 A_{23} = (\pi_0 + \pi_1 + \pi_2) \cdot \frac{1}{5} = \frac{15}{16} \cdot \frac{1}{5}$, vagyis az arra járó autók $\frac{1}{16}$-oda nem parkol le.

5. Mintát vettünk egy X normális eloszlású valószínűségi változóból, melynek várható értéke *ismert*: $m = 1000$, de szórása ismeretlen. Azt kaptuk, hogy 997, 1002, 998, 1003, 996, 1001, 998, 1004, 1005. Adjunk maximum likelihood becslést az eloszlás szórására.

Megoldás: $n = 9$, $m = 1000$ ismert és a likelihood-függvény

$$L(\sigma) = \prod_{i=1}^{n} f_{m,\sigma}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-m)^2}{2\sigma^2}},$$

amiből a log-likelihood függvény

$$l(\sigma) = \ln L(\sigma) = \sum_{i=1}^{n} \left[-\ln\sqrt{2\pi} - \ln\sigma - \frac{(x_i-m)^2}{2\sigma^2} \right] = c - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2.$$

Ennek maximumát keressük, ehhez megoldjuk a $l'(\sigma) = 0$ egyenletet:

$$0 = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-m)^2,$$

amiből

$$\sigma_{ML} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-m)^2},$$

éppen a (korrigálatlan) tapasztalati szórás. Esetünkben

$$\sum_{i=1}^{n}(x_i-m)^2 = (-3)^2 + 2^2 + (-2)^2 + 3^3 + (-4)^2 + 1^2 + (-2)^2 + 4^2 + 5^2 = 88,$$

amiből

$$\sigma_{ML} = \sqrt{\frac{88}{9}} \approx 3.13\,.$$