Felsőbb Matematika Villamosmérnököknek - Sztochasztika
2. ZH megoldások

2017 ősz, 2017.12.07 18:00

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.)** Horváth Illés (páratlan heteken, IB145) **b.)** Kói Tamás (páratlan heteken, QBF10) **c.)** Patkó Richárd páratlan heteken (IB147) **d.)** Patkó Richárd páros heteken (IB147).

1. Jancsika tapasztalatai szerint egy kilenc pontos ZH-feladatra a hallgatói átlagosan 5 pontot szoktak szerezni, az egyes hallgatók által elért pontszámok pedig függetlenek. Adjunk nagy eltérés becslést annak valószínűségére, hogy a következő ilyen feladatra a 100 hallgatója összesen több, mint 680 pontot szerez.

   Megoldás: Legyen $n = 100$ és legyen $i = 1, 2, \dots, n$-re X_i az i-edik hallgató pontszáma. Legyen $S_n = X_1 + X_2 + \cdots + X_n$ a hallgatók összpontszáma. Mivel átlagosan 5 pontot szoktak elérni, tudjuk, hogy $\mathbb{E}S_n = 5n = 500$. A feladat szövege szerint az X_i-k függetlenek és persze korlátosak: $a_i := 0 \le X_i \le b_i := 9$, így a Hoeffding egyenlőtlenség alkalmazható. Ehhez először kiszámoljuk, hogy

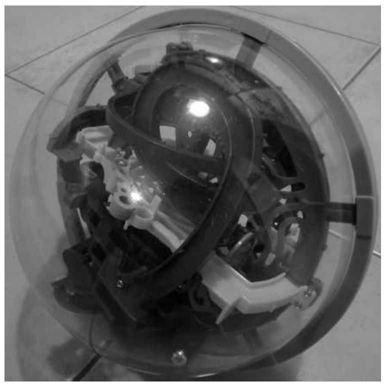
$$\sum_{i=1}^{n}(b_i - a_i)^2 = \sum_{i=1}^{n}(9-0)^2 = n \cdot 81 = 8100.$$

   Így a Hoeffding egyenlőtlenség $t = 180$ választással azt adja, hogy

$$\mathbb{P}(S_n \ge 680) = \mathbb{P}(S_n \ge \mathbb{E}S_n + t) \le \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right) = \exp\left(-\frac{2 \cdot 180^2}{8100}\right) = e^{-8} \approx 0.00034.$$

   (Megjegyzés: Erre a feladatra a Cramér nagy eltérés tétel nem alkalmazható, két okból sem. Egyrészt nincs okunk feltételezni, hogy az X_i-k azonos eloszlásúak lennének. Másrészt, még ha azok is lennének, a feladat semmit nem mond az eloszlásukról – túl azon, hogy mennyi a minimum, maximum és átlagos pontszám. Így a rátafüggvényt sem tudjuk kiszámolni.)

2. Móricka egy golyós ügyességi játékot játszik, ahol egy csapágygolyót kell végigvezetni egy akadálypályán. Az első pályát gyakorolja, ahol 3 nehéz akadályon kell átjutni. Móricka az első akadályon $\frac{1}{4}$, a másodikon $\frac{1}{3}$, a harmadikon $\frac{1}{2}$ valószínűséggel *bukik el*, az előzményektől függetlenül. Ilyenkor a golyó „leesik”, és Móricka kezdheti az egészet elölről. Ellenkező esetben továbbjut a következő akadályhoz. Ha véletlenül mindhárom akadályon sikerül túljutnia, akkor szintén újrakezdi a legelejéről.
   Jelölje X_n azt, hogy n lépés után Móricka éppen hány akadályon van túl – így X_n lehetséges értékei $0, 1, 2, 3$.



Magical Intellect Ball

   a.) Írjuk fel az X_n Markov lánc átmenetmátrixát. *(2 pont)*

   b.) Hosszú távon melyik állapotban lesz a Markov lánc legtöbbször, és a lépések mekkora hányadát tölti Móricka ezzel a leggyakoribb akadállyal? *(5 pont)*

   c.) Hosszú távon hanyadik akadályon *bukik el* legtöbbször Móricka, és a bukások mekkora hányada történik ezen az akadályon? *(2 pont)*

   Megoldás:

   a.) Az $S = \{0, 1, 2, 3\}$ állapottérrel

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

b.) Keressük a π stacionárius eloszlást, amihez megoldjuk a $(P-I)^T\pi^T = 0$ lineáris egyenletrendszert:

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

Ennek megoldása az $\sum_{i\in S}\pi_i$ normálási feltételt is figyelembe véve

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

vagyis a Markov lánc a 0 állapotban van legtöbbször (hát persze), éspedig az ergodtétel értelmében hosszú távon a lépések $\frac{4}{10}$-ében. (A lánc irreducibilis és aperiodikus, az ergodtételt az egyes állapotok indikátorfüggvényeire alkalmazhatjuk.)

c.) A lépések $\frac{4}{10}$-ében próbálkozik Móricka az 1-es akadállyal, ezen belül mindig $\frac{1}{4}$ valószínűséggel bukik el, vagyis a lépések $\frac{4}{10}\cdot\frac{1}{4} = \frac{1}{10}$-ében éppen az 1-es akadályt bukja. Hasonlóan a 2-es és 3-as akadályt is a lépések $\frac{1}{10}$-ében bukja, vagyis **hosszú távon mindhárom akadályon ugyanannyiszor, az összes bukás $\frac{1}{3}$-ában bukik.**

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. *(3 pont)*

   b.) Az eszköz a működése első pillanatában szabad. Közelítőleg mennyi a valószínűsége, hogy tíz óra elteltével éppen foglalt lesz? Miért? *(4 pont)*

   c.) Az eszköz teljesítmény-felvétele passzív állapotban $1W$, feldolgozás során viszont $10W$. Mennyi az átlagos teljesítmény-felvétel hosszú távon? Miért? *(2 pont)*

Megoldás: Legyen az állapottér $S = \{0, 1\}$, ahol 0 jelentése „szabad, passzív, jelre vár”, és 1 jelentése „foglalt, feldolgozás folyamatban, nem figyel”. *(Megjegyzés: a kétállapotú Markov lánc neve ON-OFF folyamat.)* Legyen $X(t) \in S$ az eszköz állapota t perc elteltével. Így $X(t)$ folytonos idejű, időben homogén Markov lánc. 0-ból 1-be akkor ugrik, ha érkezik egy jel, vagyis $\lambda_{01} = 2$ rátával (mert percenként átlag 2 jel jön). 1-ből 0-ba akkor ugrik, ha az eszköz végez a jel feldolgozásával, vagyis $\lambda_{10} = 60$ rátával (mert a feldolgozási idő várható értéke $\frac{1}{60}$).

a.) A fenti szöveg alapján az ugrási ráták mátrixa

$$\underline{\underline{\lambda}} = \begin{pmatrix} * & 2 \\ 60 & * \end{pmatrix}.$$

(Ennek a főátlójában nincs semmi.) Ebből a G infinitezimális generátort úgy kapjuk , hogy a főátlót úgy töltjük ki, hogy minden sorösszeg nulla legyen:

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

b.) Az X_t Markov lánc véges állapotterű és irreducibilis. (Mivel folytonos idejű, periodikus nem lehet.) A $t = 600$ (perc) hosszú idő, így a Markov láncok alaptétele szerint

$$\mathbb{P}(X(t) = 1 | X(0) = 0) \approx \pi_1,$$

ahol π az egyetlen stacionárius eloszlás. Ennek kereséséhez megoldjuk a $G^T\pi^T = 0$ lineáris egyenletrendszert:

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

Ennek egyetlen (normált) megoldása

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

Így a keresett valószínűség

$$\mathbb{P}(X(t) = 1 | X(0) = 0) \approx \pi_1 = \frac{1}{31} \approx 0.032$$

(Megjegyzés: a stacionárius eloszlás úgy is megtalálható, ha kihasználjuk, hogy $X(t)$ születési-halálozási folyamat.))

c.) Az $f(X(t))$ teljesítmény időátlagát keressük, ahol $f(0) = 1$ és $f(1) = 10$ (Watt-ban). Vektor-jelöléssel $f = \begin{pmatrix} 1 \\ 10 \end{pmatrix}$. Mivel a Markov lánc véges állapotterű és irreducibilis, az ergodtétel szerint $f(X(t))$ időátlaga

$$\overline{f(X(t))} = \mathbb{E}_\pi f = \sum_{i \in S} \pi_i f(i) = \pi f = \begin{pmatrix} \frac{30}{31} & \frac{1}{31} \end{pmatrix} \begin{pmatrix} 1 \\ 10 \end{pmatrix} = \frac{40}{31} \approx 1.29 \text{ (Watt).}$$

4. Egy (esetleg) hamis dobókockán a 6-os valószínűsége valami ismeretlen $p \in (0;1)$, az összes többi szám valószínűsége pedig azonos, $\frac{1-p}{5}$. A kockával 10-szer dobva mintát vettünk az eloszlásból, és azt kaptuk, hogy $5; 6; 4; 3; 4; 6; 3; 1; 6; 3$. Adjunk maximum likelihood becslést p értékére.

Megoldás: A diszkrét valószínűségeloszlás $p(x) = p$, ha $x = 6$ és $p(x) = \frac{1-p}{5}$, ha $x = 1, 2, 3, 4, 5$. Így ha $n = 10$ és a minta $x_1, \ldots, x_n$, akkor a likelihood-függvény

$$L(p) = \prod_{i=1}^{n} p(x_i),$$

esetünkben

$$L(p) = \left(\frac{1-p}{5}\right)^7 \cdot p^3,$$

ennek logaritmusa

$$l(p) = 7\ln(1-p) - 7\ln 5 + 3\ln p.$$

A maximum likelihood becsléshez az $l'(p) = 0$ egyenletet megoldva

$$p_{ML} = \frac{3}{10}.$$

További deriválással (pl.) ellenőrizhető, hogy ez tényleg globális maximumhelye l-nek.

5. Ha egy ember kitölt egy IQ-tesztet, az eredmény normális eloszlású valószínűségi változó. Ennek várható értékét definíció szerint az illető ember *intelligencia-hányadosának* nevezzük, szórása pedig 3. Pistike kitöltött néhány független IQ-tesztet. Pontszámai: 97, 101, 96, 98. Döntsünk 90%-os szinten arról a hipotézisről, hogy Pistike intelligencia-hányadosa 100.

Megoldás: Legyen x_i az i-edik teszt eredménye ($i = 1, \ldots, n$ és $n = 4$). Így $x_i, \ldots, x_n$ minta az (m, σ^2) paraméterű normális eloszlásból, ahol $\sigma = 3$, m pedig Pistike (számunkra ismeretlen) intelligencia-hányadosa. A $m = \mu$ nullhipotézisről kell döntenünk, ahol $\mu = 100$ a hipotetikus várható érték. Ehhez egymintás kétoldali u-próbát végzünk: a teszt-statisztika

$$u = \frac{\bar{x} - \mu}{\sigma}\sqrt{n} = \frac{\frac{97+101+96+98}{4} - 100}{3}\sqrt{4} = \frac{-2}{3} \cdot 2 = -\frac{4}{3} \approx -1.333$$

Mivel $1 - \varepsilon = 90\%$-os szinten kell döntenünk, $\varepsilon = 0.1$, és az elfogadási küszöb

$$K_\varepsilon = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right) = \Phi^{-1}(0.95) \approx 1.65$$

(itt Φ a standard normális eloszlásfüggvény).

Döntés: mivel $|u| \leq K_\varepsilon$, a nullhipotézist elfogadjuk.