Sztochasztika félvizsga megoldások

Felsőbb matematika villamosmérnököknek A, B vizsgakurzus

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

1. Az X nemnegatív egész értékű valószínűségi változó generátorfüggvénye

$$g(z) = \frac{1}{4} + \frac{1}{4}z + \frac{3}{8}z^2 + cz^3, \quad \text{ahol } c \in \mathbb{R}.$$

(a) Mennyi c értéke?

(b) Mennyi X várható értéke?

(c) Mennyi a $\mathbb{P}(X = 2)$ valószínűség?

Megoldás:

(a) Minden genarátorfüggvényre igaz, hogy $g(1) = 1$, ezért

$$g(1) = \frac{1}{4} + \frac{1}{4} \cdot 1 + \frac{3}{8} \cdot 1^2 + c \cdot 1^3 = \frac{1}{4} + \frac{1}{4} + \frac{3}{8} + c = 1,$$

amiből $c = \frac{1}{8}$. Ebből

$$g(z) = \frac{1}{4} + \frac{1}{4}z + \frac{3}{8}z^2 + \frac{1}{8}z^3$$

(b) Mint mindig, $\mathbb{E}X = g'(1)$. Ehhez

$$g'(z) = \frac{1}{4} + \frac{3}{8} \cdot 2z + \frac{1}{8} \cdot 3z^2,$$

amiből

$$\mathbb{E}X = g'(1) = \frac{1}{4} + \frac{3}{8} \cdot 2 + \frac{1}{8} \cdot 3 = \frac{11}{8} = 1.375.$$

(c) Mint mindig, $\mathbb{P}(X = 2) = \frac{g''(0)}{2!}$. Esetünkben $g''(0) = \frac{3}{4}$, amiből

$$\mathbb{P}(X = 2) = \frac{\frac{3}{4}}{2!} = \frac{3}{8}.$$

Alternatív megoldás: A generátorfüggvényből kiolvasható X eloszlása: $p_k := \mathbb{P}(X = k)$ jelöléssel

$$g(z) = \sum_{k=0}^{\infty} p_k z^k = \frac{1}{4} + \frac{1}{4}z + \frac{3}{8}z^2 + cz^3,$$

amiből $p_0 = \frac{1}{4}$, $p_1 = \frac{1}{4}$, $p_2 = \frac{3}{8}$, $p_3 = c$, és a többi k-ra $p_k = 0$. Így persze

(a) $\sum_{k=0}^{\infty} p_k = 1$ miatt $c = \frac{1}{8}$.

(b) $\mathbb{E}X = \sum_{k=0}^{\infty} kp_k = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{11}{8}$.

(c) $\mathbb{P}(X = 2) = p_2 = \frac{3}{8}$.

2. 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.

3. 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.

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

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

A λ paraméterű Poisson eloszlás Cramér féle rátafüggvénye

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

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)

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

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. Ádám és Éva is kitöltött néhány független IQ-tesztet, egymástól is függetlenül. Ádám pontszámai: 111, 108, 111, 112. Éva pontszámai: 114, 112, 114, 119, 113. Döntsünk 95%-os szinten arról a hipotézisről, hogy Ádám és Éva intelligencia-hányadosa azonos.

   Megoldás: Ismert szórású normális eloszlásokból vettünk mintát, a cél pedig a várható értékek összehasonlítása. Jelöljük x_i-vel az Ádám, y_i-vel pedig az Éva pontszámait. A null-hipotézis $m_x = m_y$. Ezért kétmintás kétoldali u-próbát végzünk. A szórások $\sigma_1 = \sigma_2 = 3$, a minta-elemszámok $n_1 = 4$ és $n_2 = 5$. Az átlagok $\bar{x} = \frac{111+108+111+112}{4} = 110.5$ és $\bar{y} = \frac{114+112+114+119+113}{5} = 114.4$. A teszt-statisztika a képletgyűjtemény szerint

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} = \frac{110.5 - 114.4}{\sqrt{\frac{3^2}{4} + \frac{3^2}{5}}} = \frac{-3.9}{3\sqrt{\frac{1}{4} + \frac{1}{5}}} \approx -1.938.$$

   $1 - \varepsilon = 95\%$-os szinten akarunk döneni, vagyis $\varepsilon = 0.05$. A küszöbérték a képletgyűjtemény szerint

$$K = u_{\frac{\varepsilon}{2}} = \Phi^{-1}(1 - \frac{\varepsilon}{2}) = \phi^{-1}(0.975) \approx 1.96.$$

   Döntés: $|u| \leq K$, azért a hipotézist *elfogadjuk*.