Felsőbb matematika villamosmérnököknek - Sztochasztika

pótZH megoldások, 2022 ősz

Minden megoldást részletesen indokolni kell. Azon belül minden alkalmazott jelölést be kell vezetni.

Munkaidő: 90 perc

1. A *Badacsony szelet*^{TM} „Minden ötödik nyer!” akciót hirdetett, mely alapján minden csokipapírba „nyert” illetve „nem nyer” feliratok kerülnek 1 : 4 arányban (véletlenszerűen elszórva). A „nyert” feliratú papírokat aztán egy újabb csokira lehet beváltani. (Az akció során olyan sok csoki került a piacra, hogy az Pistike szempontjából végtelennek tekinthető.)

   a.) **(4 pont)** Pistike vett egy Badacsony szeletet, mellyel egy másik szeletet is nyert. Nagy meglepetésére a második csokoládéval egy harmadiknak is a szerencsés birtokosa lett. A harmadikban sajnos már „nem nyert” feliratot talált. Mi ennek az esemény-sornak a valószínűsége?

   b.) **(6 pont)** Átlagosan hány csokihoz jut Pistike az akció során egy csokoládé árából?

   Megoldás:

   a.) $\frac{1}{5}$ valószínűséggel nyer, $\frac{4}{5}$ valószínűséggel nem nyer egy adott csokoládé. Mivel a csokoládék kimenetelei függetlenek, a valószínűségeket összeszorozhatjuk.

   $$\mathbb{P}(\text{nyert, nyert, nem nyert}) = \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} = \frac{4}{125} = 0.032 = 3.2\%$$

   b.) Jelölje X a nyert csokoládék számát plusz egyet, az eredeti (pénzért vett) csokit: ennyit eszik Pistike egy csoki áráért. X (optimista) geometriai eloszlású $p = \frac{4}{5}$ siker-valószínűséggel, hiszen addig kell újabb és újabb csokikat megenni, amíg nem sikerül egy „nem nyert” feliratot találni. A várható érték tehát

   $$\mathbb{E}\left(X\right) = \frac{1}{p} = \frac{5}{4} = 1.2$$

2. Svindlis Sanya körlevelet készít, mely sikeres dolgozatokat és vizsgákat ígér a címzetteknek, amennyiben továbbküldik 3 barátjuknak. (Ellenkező esetben pedig egy kis nyuszi elsírja magát valahol a világban.) Mindenki, aki ezt a levelet megkapja, 40% eséllyel küldi tovább 3 másik embernek (egyébként pedig senkinek), az előzményektől függetlenül. (Az egyszerűség kedvéért tekintsünk el annak a lehetőségétől, hogy valaki olyannak küldjük a levelet, aki már megkapta mástól.) Kezdetben Sanya csak egy embernek, Pistikének küldi el a körlevelet.

   a.) **(5 pont)** Jelölje Y azoknak az embereknek a számát, akik Pistike *barátaitól* kapják meg a körlevelet. (Vagyis nem Pistikétől magától, hanem egy áttételen keresztül.) Mi Y generátorfüggvénye?

   b.) **(5 pont)** Mi annak a valószínűsége, hogy internet-szenzációvá válik Sanya körlevele, ahelyett, hogy a körlevelek láncolata megállna?

   Megoldás:

   Nevezzük *nulladik generációnak* Pistikét egymagát. Nevezzük *első generációnak* azokat, akiknek Pistike a levelet továbbküldi. És így tovább: álljon az $(n+1)$-edik generáció azokból, akiknek az n-edik generáció tagjai a levelet továbbküldik, és legyen Z_n az n-edik generáció tagjainak száma. Így Z_n Galton-Watson elágazó folyamat, az egylépéses utódszám pedig legyen X. Ez 0 vagy 3 értéket vehet fel, az utódszám-eloszlás tehát: $p_0 = 0.6$, $p_3 = 0,4$ és $p_k = 0$ ha $k \neq 0, 3$). Az utódeloszlás generátorfüggvénye tehát:

   $$g(z) = 0.6z^0 + 0.4z^3 = 0.6 + 0.4z^3.$$

a.) A fenti jelölésekkel $Y = Z_2$. Z_n generátorfüggvényét $g(z)$ kompozíció hatványaiként kapjuk meg, vagyis

$$\begin{aligned} g_Y(z) =& g_{Z_2}(z) = g(g(z)) = 0.6 + 0.4\left(0.6 + 0.4z^3\right)^3 = \\ & 0.6 + 0.4\left(0.6^3 + 3\cdot 0.6^2(0.4z^3) + 3\cdot 0.6\cdot (0.4z^3)^2 + (0.4z^3)^3\right) = \\ & 0.6 + 0.4\left(0.216 + 0.144z^3 + 0.288z^6 + 0.064z^9\right) = \\ & 0.6864 + 0.0576z^3 + 0.1152z^6 + 0.0256z^9. \end{aligned}$$

b.) Egy ember átlagosan $m = \mathbb{E}X = 0.6\cdot 0 + 0.4\cdot 3 = 1.2$ embernek adja tovább az üzenetet. Mivel $m > 1$, a folyamat szuperkritikus, így a kihalás esélye kisebb mint 1 és az alábbi fixpontegyenlet egyetlen $0 \le z < 1$ tartományba eső megoldásával lesz egyenlő:

$$\begin{aligned} z =& g(z) \\ z =& 0.6 + 0.4z^3 \\ 0.4z^3 - z + 0.6 =& 0 \\ 2z^3 - 5z + 3 =& 0. \end{aligned}$$

A generátorfüggvénynek $z = 1$ mindig fixpontja, így $z = 1$ biztosan gyök: polinom-osztást hajthatunk végre $z - 1$-gyel.

$$2z^2 + 2z - 3 = 0$$

Elég csupán a pozitív gyökre szorítkoznunk, így

$$z = \frac{-2 \pm \sqrt{4+24}}{4} = \frac{-1 \pm \sqrt{7}}{2}.$$

Az egyetlen $[0,1)$-beli gyök így

$$z = \frac{-1+\sqrt{7}}{2} \approx 0.8229$$

az örökké túlélés valószínűsége pedig $1 - z \approx 0.1771 \approx 18\%$

Ha valaki iterációs módszerrel közelíti z-t, azt is elfogadjuk megoldásként.

3. Piripócs könyvtárában egy régi, elromlásra hajlamos számítógép van. Miden nap háromféle állapotban lehet: „működik", „rossza a monitor", illetve „rossza az internet". Egy adott nap, ha működik, 80% eséllyel fog másnap is működni, 10% eséllyel a monitorja, 10% eséllyel pedig az internet romlik el. Ha az internet rossz, azt könyvtáros 60% eséllyel javítja meg másnapra; ha a monitor rossz, azt csak 20% eséllyel – az előzményektől függetlenül. (A javítás alatt más műszaki probléma nem lép fel.)

   a.) **(4 pont)** Feltéve, hogy hétfőn működik a számítógép, mi az esélye, hogy szerdán nem működik a monitor?

   b.) **(6 pont)** A polgármester elrendelte, hogy a számítógépet 3 év múlva le kell cserélni. Addig kb. a napok hányad részében fog jól működni a számítógép?

Megoldás:

A számítógép egymást követő állapotait az X_n Markov lánccal modellezhetjük, melynek állapottere $S = \{1, 2, 3\}$, ahol 1 jelöli a működés, 2 a monitor, 3 pedig az internet hiba állapotát. Az átmenetmátrix

$$P = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.6 & 0.4 & 0 \\ 0.2 & 0 & 0.8 \end{pmatrix}.$$

a.) Legyen a 0. nap a hétfő! A feltételünk szerint $X_0 = 1$:

$$\begin{aligned}\mathbb{P}(X_2 = 2 \,|\, X_0 = 1) = \left(P^2\right)_{12} &= P_{11}P_{12} + P_{12}P_{22} + P_{13}P_{32} \\ &= 0.8 \cdot 0.1 + 0.1 \cdot 0.4 + 0.1 \cdot 0 = 0.12 = 12\%\end{aligned}$$

b.) A Markov lánc véges állapoterű és irreducibilis, így az ergodtétel szerint az 1 állapotban hosszú távon az idő π_1 hányadát tölti (a kezdeti állapottól függetlenül), ahol π az egyetlen a stacionárius eloszlás (sorvektor). A meghatározásához meg kell oldanunk a

$$(P^T - \mathbf{1})\pi^t = 0$$

lineáris egyenletrendszert, és annak egy (tetszőleges) megoldását lerormálni, hogy a sorösszeg 1 legyen. Az egyenletrendszer

$$\left(\begin{array}{ccc|c} -0.2 & 0.6 & 0.2 & 0 \\ 0.1 & -0.6 & 0 & 0 \\ 0.1 & 0 & -0.2 & 0 \end{array}\right)$$

, ennek egy megoldása a

$$\tilde{\pi} = \begin{pmatrix} 6 & 1 & 3 \end{pmatrix},$$

amit lenormálva

$$\pi = \begin{pmatrix} \frac{6}{10} & \frac{1}{10} & \frac{3}{10} \end{pmatrix},$$

Vagyis hosszú távon a számítógép az idejének a $\pi_1 = \frac{6}{10}$ részét tölti működőképes állapotban.

4. A Piripócsi Skacok újság átlagosan 10.4 helyreigazítási pert veszít el egy év során. A pereket egymástól független ritka események generálják. (Az egyszerűség kedvéért az év tekinthető 360 naposnak 30 napos hónapokkal.)

a.) **(3 pont)** Feltéve, hogy télen 2 pert veszítettek el, mi az esélye, hogy nyáron legalább 3-at veszítenek?

b.) **(3 pont)** Tiszta lapról indulva átlagosan hány napot kell várni a 4. vesztes helyreigazítási perig?

c.) **(4 pont)** Az elvesztett perek 35%-át a Piripócs a Piripócsiaké (PP) alapítvány kezdeményezi. Mi az esélye, hogy négy hónap alatt az újság egy pert veszít velük szemben, két másikat pedig valaki mással szemben?

Megoldás:

Mivel a sok cikk-beli sok hazugságból külön-külön kis valószínűséggel és függetlenül lesz bukott sajtóper, a bukott pereket Poisson folyamattal modellezhetjük, $\lambda = 10.4$ rátával (az időt években mérve). Ezen belül a PP alapítvány ellen vesztett perek a Poisson folyamat ritkítása 0.35-szeresre.

a.) Mivel a Poisson folyamatban a diszjunkt időintervallumok függetlenek egymástól, így elég annak az esélyét kiszámolnunk, hogy 3 hónap alatt legalább három pert veszítenek el. Ennek számát jelölje X, mely Poisson eloszlású valószínűségi változó $10.4 \cdot \frac{3}{12} = 2.6$ paraméterrel.

$$\begin{aligned}\mathbb{P}(X \geq 3) =& 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2) = 1 - e^{-2.6} + 2.6 \cdot e^{-2.6} + \frac{2.6^2}{2}e^{-2.6} = \\ & 1 - 6.98 \cdot e^{-2.6} \approx 0.4816 \approx 48\%\end{aligned}$$

b.) Jelölje T_k a k. per időpontját, továbbá az egyszerűbb jelölés kedvéért legyen $T_0 := 0$! $\tau_k := T_k - T_{k-1}$ az idő, mely eltelik a k. és a $(k-1)$. per között. A τ_k-k i.i.d. exponenciális eloszlású valószínűségi változók $\lambda = 10.4$ paraméterrel.

$$\mathbb{E}(T_4) = \mathbb{E}\left(\sum_{k=1}^{4} \tau_k\right) = \sum_{k=1}^{4} \mathbb{E}(\tau_k) = 4\mathbb{E}(\tau_1) = 4 \cdot \frac{1}{10.4} \approx 0,3846 \ (\acute{e}v) \approx 139 \text{ nap}$$

c.) Legyen X a PP ellen vesztett perek számát, Y pedig a mások ellen vesztett perek számát 4 hónap alatt! Ezek függetlenek és Poisson eloszlásúak rendre $0.35 \cdot 10.4 \cdot \frac{4}{12} \approx 1.213$ illetve $0.65 \cdot 10.4 \cdot \frac{4}{12} \approx 2.253$ paraméterrel. Így

$$\mathbb{P}\left(X = 1.Y = 2\right) = \mathbb{P}(X = 1)\mathbb{P}(Y = 2) \approx 1.213 \cdot e^{-1.213} \cdot \frac{2.253^2}{2} \cdot e^{-2.253} \approx 0.096 = 9.6\%$$

5. **(10 pont)** Egy szemétszállítási vállalat 100 literes és 400 literes kukákra köt szerződést. Az előbbit 7000, az utóbbit 3000 ügyfél veszi igénybe. A kis kukákban átlagosan 80, a nagyokban 300 liter szemetet halmoznak fel egy adott héten, egymástól függetlenül.

A cég (annyi és) akkora kukásautót kíván kiküldeni, ami legfeljebb 1% eséllyel nem lesz képes mindenki szemetét elvinni. Hány literesre tervezzük a kukásautó(k) teljes kapacitását?

Megoldás:

Jelölje X_i $i = 1, 2, \ldots, 10000$ az egy adott ügyfél szemétmennyiségét. Tartozzon az első 7000 index a 100 literes, a maradék 3000 pedig a 400 literes szerződésekhez. Tudjuk, hogy

$$a_i \leq X_i \leq b_i$$

ahol $a_i = 0$, $b_i = 100$, ha $1 \leq i \leq 7000$, és $b_i = 400$, ha $7001 \leq i \leq 1000$.

Az összes szemét mennyisége: $S_n := \sum_{i=1}^n X_i$, ahol $n = 10000$. Legyen t egy tetszőleges nem negatív szám! A Höeffding-egyenlőtlenség alapján

$$\mathbb{P}\left(S_n \geq \mathbb{E}(S_n) + t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

t értékét úgy kell meghatároznunk, hogy a jobboldal egyenlő legyen 0.01-gyel. Vegyük észre, hogy

$$\sum_{i=1}^n (b_i - a_i)^2 = 7000 \cdot 100^2 + 3000 \cdot 400^2 = 1.9 \cdot 10^8.$$

t-re tehát azt kapjuk, hogy

$$\begin{aligned}
&\exp\left(-\frac{2t^2}{1.9 \cdot 10^8}\right) = 10^{-3} \\
&\frac{2t^2}{1.9 \cdot 10^8} = 3\ln(10) \\
&t = \sqrt{\ln(10) \cdot 2.85 \cdot 10^8} \approx 25617.
\end{aligned}$$

A várható érték:

$$\begin{aligned}
\mathbb{E}(S_n) =& \mathbb{E}\left(\sum_{i=1}^{10000} X_i\right) = \sum_{i=1}^{7000} \mathbb{E}(X_i) + \sum_{i=7001}^{10000} \mathbb{E}(X_i) = 7000 \cdot E(X_1) + 3000 \cdot \mathbb{E}(X_{7001}) = \\
& 7000 \cdot 80 + 3000 \cdot 300 = 1460000.
\end{aligned}$$

Így tehát a kapacitás, amire fel kell készülnünk $\mathbb{E}(S_n) + t \approx 1485617$ liter.