Felsőbb matematika villamosmérnököknek - Sztochasztika

ZH megoldások, 2020 ősz, 2020.12.02, 18:00

Munkaidő: 90 perc. A megoldásokat fényképezve vagy scannelve kérem vissza. Aki ezt nem tudja megoldani, kérem, jelezze!

Minden írásos segédeszköz illetve számológép/számítógép használható, de a feladatokat **önállóan** kell megoldani, vagyis más embertől kérdezni és segítséget elfogadni nem szabad. Ez alól egyedüli kivétel a tárgy előadója.

Minden megoldást részletesen indokolni kell.

Minden feladat 10 pontot ér.

1. Bergengóciában Móricka nagyon népszerű: minden 3 újságcikkből 2 róla szól (abban az értelemben, hogy minden újságcikk, a többitől függetlenül, $\frac{2}{3}$ valószínűséggel szól róla.

   Amikor este Móricka a neten újságot olvas, válogatás nélkül rákattint minden cikkre, egészen addig, amíg el nem jut a 100-adik róla szóló cikkig. Adjunk nagy eltérés becslést annak valószínűségére, hogy eközben legalább 100 egyéb cikket is talál!

   Megoldás:

1. megoldás:

Legyen $i=1,2,3\ldots$-re $X_i = \begin{cases} 1, & \text{ha az } i\text{-edik kattintásra Móricka magáról szóló cikket talál} \\ 0, & \text{ha nem} \end{cases}$ „siker"

~~legyen $S_n = X_1$~~

Így $\boxed{X_1, X_2, X_3, \ldots \text{ független } \sim B(p), \ p=\frac{2}{3}}$

legyen $S_n = X_1 + X_2 + \ldots + X_n$ a „sikeres" kattintások száma.

A kérdés annak val.sége, hogy 100 sikerhez legalább 200 kattintás kell, avagy: $\boxed{n=199}$ kattintásból még legfeljebb 99 sikeres:

$\boxed{\mathbb{P}(S_n \leq 99)}$ -re kell nagy eltérés becslés. $K := 99$

1/a megoldás: Minden X_i korlátos (és független), így alkalmazható a Hoeffding egyenlőtlenség: $0 = a_i \leq X_i \leq b_i = 1$

$$\Rightarrow \sum_{i=1}^{n} (b_i - a_i)^2 = n = 199 \ ; \quad EX_i = p = \frac{2}{3} \Rightarrow \mathbb{E}S_n = np = 199 \cdot \frac{2}{3} = 132.6\dot{6}$$

$$\Rightarrow \boxed{t := \mathbb{E}S_n - K = 132.6\dot{6} - 99 = 33.6\dot{6}} = \frac{101}{3}$$

választással $K = 99 = \mathbb{E}S_n - t$, így

$$\boxed{\mathbb{P}(S_n \leq 99) = \mathbb{P}(S_n \leq \mathbb{E}S_n - t)} \overset{\text{Hoeffding}}{\leq} \exp\left\{-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right\} = \exp\left\{-\frac{2t^2}{n}\right\}$$

$$= \exp\left\{-\frac{2 \cdot 33.6\dot{6}^2}{199}\right\} \approx e^{-11.39} \approx 1.1 \cdot 10^{-5}$$

1/b megoldás: Minden X_i azonos eloszlású (és független),

így alkalmazható a $\boxed{\text{Cramér nagy eltérés tétel}}$:

$$\mathbb{P}(S_n \le 99) = \mathbb{P}\left(\frac{S_n}{n} \le \frac{99}{199}\right) = \mathbb{P}\left(\frac{S_n}{n} \in (a,b]\right), \text{ ahol } a = -\infty,$$

$$b = \frac{99}{199} \approx 0.4975$$

Mivel ~~$\mathbb{E}X_i$~~ $m := \mathbb{E}X_i = p = \frac{2}{3} = 0.6\dot{6}$, $\boxed{b<m}$,

ezért a Cramér tétel szerint

$$\mathbb{P}(S_n \le 99) = \mathbb{P}\left(\frac{S_n}{n} \in (a,b]\right) \lesssim e^{-nI(b)} = e^{-199 \cdot I\left(\frac{99}{199}\right)},$$

ahol I a $p = \frac{2}{3}$ paraméterű Bernoulli eloszlás ráta-függvénye. Ezt tudjuk pl. az előadásról:

$$I(x) = x \ln \frac{x}{p} + (1-x) \ln \frac{1-x}{1-p},$$

ami $p = \frac{2}{3}$, $x = b = \frac{99}{199}$ helyettesítéssel $I(b) \approx 0.060646$,

$$nI(b) \approx 12.0685$$

$$\Rightarrow \boxed{\mathbb{P}(S_n \le 99) \lesssim e^{-12.0685} \approx 5.7 \cdot 10^{-6}}$$

2. megoldás:

Legyen „sikeres" egy kattintás, ha Móricka önmagáról szóló cikket talál, és legyen $i = 1, 2, 3 \ldots$ -re

Y_i az i-edik sikerhez szükséges kattintások száma (az előző sikertől számítva).

Így $\boxed{Y_1, Y_2, Y_3 \ldots \text{ független } \sim \mathrm{Geom}(p), \; p = \frac{2}{3}}$

A kérdés ~~annak~~ annak val-sége, hogy $\boxed{n = 100}$ sikerhez legalább 200 kattintás kell. Vagyis

$S_n := Y_1 + Y_2 + \ldots + Y_n$ jelöléssel

$\boxed{\mathbb{P}(S_n \geq 200)\text{-ra kell nagy eltérés becslés.}}$

Mivel az Y_i-k függetlenek és azonos eloszlásúak, a $\boxed{\text{Cramér nagy eltérés tétel}}$ használható:

$m := \mathbb{E}Y_i = \frac{1}{p} = \frac{3}{2} = 1.5, \quad a = 2, \quad b = \infty$ jelöléssel

$$\mathbb{P}(S_n \geq 200) = \mathbb{P}\left(\frac{S_n}{n} \geq 2\right) = \mathbb{P}\left(\frac{S_n}{n} \in [a, b)\right) \underset{a > m}{\overset{\text{Cramér}}{\lessapprox}} e^{-nI(a)} = e^{-100 \cdot I(2)},$$

ahol I a $p = \frac{2}{3}$ paraméterű geometriai eloszlás ráta-függvénye: $I(x) = x \ln\left(\frac{x-1}{x(1-p)}\right) + \ln\left(\frac{1-p}{p(x-1)}\right)$ (mondjuk a gyakorló feladatsorból).

Ez $p=\frac{2}{3}$, $x=a=2$ helyettesítéssel $I(a)\approx 0.1178$,

$nI(a)\approx 11.78$

$$\Rightarrow \boxed{\mathbb{P}(S_n \geq 200) \lessapprox e^{-11.78} \approx 7.7\cdot 10^{-6}}$$

2. Bergengóciában járvány terjed. Aki megfertőződik, az pontosan egy hétig fertőző (mielőtt meggyógyulna vagy karanténba kerülne). A bergengócok hetente egyszer tartanak házibulit, így mindenki legfeljebb 1 házibuliba mehet el úgy, hogy éppen fertőző. Ha egy fertőző ember kihagyja ezt az egy házibulit, akkor nem fertőz meg senkit. Ha viszont elmegy, akkor ott megfertőz 24 embert. Márpedig Bergengóciában mindenki minden héten az előzményektől függetlenül 5% valószínűséggel megy el a buliba.

   Kezdetben Móricka az egyetlen fertőzött.

   a.) Várhatóan hányan lesznek fertőzöttek egy év (52 hét) elteltével?

   b.) Mennyi annak a valószínűsége, hogy 3 hét elteltével senki sem fertőzött?

   c.) **Bónusz kérdés +4 pontért:** Mennyi a valószínűsége, hogy a járvány előbb-utóbb eltűnik, vagyis előbb-utóbb senki sem lesz fertőző?

   Megoldás:

Legyen $n=0,1,2,\ldots$-re Z_n a fertőző bergengócok száma n hét elteltével.

Ez Galton-Watson elágazó folyamat, $Z_0=1$ (Móricka maga).

Az 1-lépéses utódszám-eloszlás:

k	0	24
$p_k=\mathbb{P}(k \text{ utód})$	0.95	0.05

Ennek várható értéke $m=0.95\cdot 0+0.05\cdot 24=1.2$,

generátorfüggvénye $g(z)=\sum_k p_k z^k = 0.95\cdot z^0+0.05\,z^{24}$

$$=0.95+0.05\,z^{24}=\frac{19+z^{24}}{20}.$$

a.) ~~$n=52$-vel~~

~~$\mathbb{E}Z_n=m^n=1.2^{52}\approx 13105$~~

$\boxed{\text{Bocs}}$, a kérdésfeltevés rossz: nem tudhatjuk, hogy az 52 hét alatt megfertőződő emberek közül hány gyógyul meg (és hány kerül karanténba) az 52. hét végére.

Amit tudunk:

a1.) Várhatóan hányan lesznek fertőzőek $n=52$ hét elteltével? Válasz: $\boxed{\mathbb{E}Z_n=m^n=1.2^{52}\approx 13105}$

a2.) Várhatóan hányan fertőződnek meg összesen az $n=52$ hét alatt (Mórickát is beleértve)?

Válasz: $\boxed{\mathbb{E}(Z_0+Z_1+\ldots+Z_n)}=1+m+m^2+\ldots+m^n=\frac{m^{n+1}-1}{m-1}=\frac{1.2^{53}-1}{0.2}$

$\approx \boxed{78623}$

b) $r_n = \mathbb{P}(Z_n = 0)$ jelöléssel

$$r_0 = 0$$

$$r_1 = g(r_0) = 0.95$$

$$r_2 = g(r_1) \approx 0.9646$$

$$\boxed{r_3 = g(r_2) \approx 0.9711 \approx 97\%}$$

c. <u>Bónusz:</u> $r_\infty = \mathbb{P}(\text{kihalás}) = \lim_{n\to\infty} r_n$.

Ez a $g(z) = z$ fixpont-egyenlet megoldása, de ezt megoldani csak numerikusan tudjuk.

Numerikusan viszont könnyű, hiszen $r_n \to r_\infty$, és r_n rekurzívan számolható.

Matlab / Octave vagy egy táblázat-kezelő század-másodperc alatt kiszámolja, hogy $n = 1000$-re

$$\left.\begin{array}{l} r_n \approx 0.98372, \\ r_{2n} \text{ szintén} \end{array}\right\} \Longrightarrow$$ hihető, hogy elértük a határértéket: $\boxed{r_\infty \approx 0.98372}$

Megj: A jó kérdés persze nem ez, hanem a kihalás val.sége, FELTÉVE, hogy Móricka elmegy a buliba. Ekkor $Z_1 = 24$, vagyis a 2. lépésben 24 független G-W folyamat indul, és <u>mindnek</u> ki kell halni:

$$\boxed{\mathbb{P}(\text{kihalás} \mid \text{Móricka elmegy a buliba}) = r_\infty^{24} \approx 0.67 = 67\%}.$$

3. Egy ereszcsatorna sok szegeccsel van összeszerelve. Ezek mindegyike, a többitől függetlenül, valamilyen kis valószínűséggel rosszul van beverve, és kiáll: méterenként átlagosan 2. Ha Móricka sötétben mászik végig az ereszcsatornán, akkor minden kiálló szegecsbe – a többitől függetlenül – $\frac{1}{5}$ valószínűséggel belenyúl, és az elvágja a kezét.

   Móricka 5 métert mászik az ereszen a sötétben. Mennyi a valószínűsége, hogy legalább kétszer elvágja a kezét?

   Megoldás:

   Legyen X a kiálló szegecsek száma 5 méteren,

   Y pedig, hogy Móricka ebből hányba nyúl bele.

   A kérdés $\mathbb{P}(Y \geq 2)$.

   Mivel sok szegecs próbál kiállni, és egymástól függetlenül kis val.séggel sikerül nekik, a sikeres próbálkozók száma jó közelítéssel $X \sim \text{Poi}(\lambda)$, ahol ~~$\lambda = \mathbb{E}X = 2$~~

   $$\lambda = \mathbb{E}X = 2\,\frac{(\text{darab})}{m} \cdot 5\,m = 10 \ (\text{darab}).$$

   Az Y ennek ritkítása $p = \frac{1}{5}$ megmaradási valószínűséggel, így $\boxed{Y \sim \text{Poi}(p\lambda) = \text{Poi}(2)}$. $\mathbb{P}(Y=k) = e^{-2}\frac{2^k}{k!}$ $k = 0, 1, 2 \ldots$

   Vagyis $\mathbb{P}(Y \geq 2) = 1 - \mathbb{P}(Y=0) - \mathbb{P}(Y=1) = 1 - e^{-2}(1+2)$

   $$= 1 - 3e^{-2} \approx 0{,}59 = 59\%$$

4. Mórickát megkérik a barátai, hogy hozza el nekik a tablettáikat a doktor bácsitól. 24 barátja van, ezek mindegyikének, a többitől függetlenül, $\frac{2}{3}$ valószínűséggel van szüksége tablettára. Akinek szüksége van rá, annak a doktor bácsi véletlen számú tablettát küld, egyenletes eloszlással az $\{1, 2, 3\}$ halmazon, a többiektől függetlenül, és beteszi Móricka hátizsákjába.

a.) Írjuk fel a Móricka hátizsákjába kerülő tabletták számának generátorfüggvényét, **VAGY**

b.) **VAGY** számoljuk ki a hátizsákba kerülő tabletták számának várható értékét és szórását!

(Mindenki választhat, hogy az a.) vagy b.) feladatot oldja meg.)

Megoldás:

Legyen N azon barátok száma, akiknek tabletta kell:

$N \sim Bin(24, \frac{2}{3})$. (vagyis $N \sim Bin(n,p)$, $n=24$, $p=\frac{2}{3}$)

Legyen X_i az i-edik ilyen barát számára küldött tabletták száma: $X_i \sim Uni(\{1,2,3\})$.

Mivel ezek mind függetlenek, a hátizsákba kerülő tabletták száma

$$S_N = \sum_{i=1}^{N} X_i \quad \text{véletlen tagszámú összeg.}$$

a.) ~~p~~ $q := 1-p = \frac{1}{3}$ jelöléssel

N generátorfv-e $g_N(z) = \cancel{p}(q+pz)^n = \left(\frac{1}{3} + \frac{2}{3}z\right)^{24}$

X_i generátorfv-e $g_X(z) = \frac{1}{3}z^1 + \frac{1}{3}z^2 + \frac{1}{3}z^3 = \frac{z+z^2+z^3}{3}$

$$\xRightarrow[\text{véletlen tagszámú összegről}]{\text{Tétel a}} \boxed{g_{S_N}(z) = g_N(g_X(z)) = \left(\frac{1}{3} + \frac{2}{3}\,\frac{z+z^2+z^3}{3}\right)^{24}}$$

b.) $\mathbb{E}N = np = 24 \cdot \frac{2}{3} = 16$, $Var\,N = npq = 16 \cdot \frac{1}{3} = \frac{16}{3}$

$\mathbb{E}X = \frac{1+2+3}{3} = 2$, $\mathbb{E}X^2 = \frac{1+4+9}{3} = \frac{14}{3} \Rightarrow Var\,X = \frac{14}{3} - 2^2 = \frac{2}{3}$

$$\xRightarrow[\text{tagszámú összegről}]{\text{tétel a véletlen}} \boxed{\mathbb{E}S_N = \mathbb{E}N\,\mathbb{E}X = \frac{16}{\cancel{3}} \cdot \frac{2}{\cancel{3}} = \boxed{\frac{32}{\cancel{3}}}} \ \cancel{10.6}$$

(4. jegyzet)

$$Var\,S_N = \mathbb{E}N\,Var\,X + (\mathbb{E}X)^2\,Var\,N = 16 \cdot \frac{2}{3} + 2^2 \cdot \frac{16}{3} = 32 \Rightarrow \boxed{DS_N = \sqrt{32} \approx 5.7}$$

5. Mórickáék egy házibuliban háromféleképpen táncolnak: egyénileg, párban, illetve körben állva. Mindegyiket az előzményektől független exponenciális eloszlású véletlen idő után unják meg: az egyéni tánc átlagosan 2 percig, a páros átlagosan 3 percig, a körtánc átlagosan 5 percig tart. Ha valamelyiket megunták, váltanak: az egyéni tánc és a körtánc után érmedobással döntenek a másik két lehetőség közül, viszont a páros után mindig körtánc következik.

   Modellezzük a táncot folytonos idejű Markov lánccal! Az időt mérjük percben!

   a.) Írjuk fel az állapotteret és az infinitezimális generátort!

   b.) Keressük meg a Markov lánc stacionárius eloszlásait!

   c.) Hosszú távon az idő mekkora hányadát töltik körtánccal? Miért?

   d.) **Bónusz kérdés +4 pontért**: Átlagosan hány percenként váltanak formációt hosszú távon?

   Megoldás:

a.) Az állapottér legyen

$$\boxed{S=\{1,2,3\}=\{\text{egyéni, páros, gruppen}\}}$$

A tartózkodási idő várható értéke i-ben $\mathbb{E}(Exp(\lambda_i))=\frac{1}{\lambda_i}$,

ahol λ_i a tartózkodási idő paraméter, vagyis

$$\frac{1}{\lambda_1}=2,\quad \frac{1}{\lambda_2}=3,\quad \frac{1}{\lambda_3}=5 \quad\Rightarrow\quad \underline{\lambda}=\begin{pmatrix}\frac{1}{2} & \frac{1}{3} & \frac{1}{5}\end{pmatrix}.$$

A beépített diszkrét idejű Markov lánc átmenetmátrixa a szöveg szerint

érmedobás

$$Q=\begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

páros után csoportos

Ebből a generátor: $G=(G_{ij})$, $G_{ij}=\begin{cases}-\lambda_i, & \text{ha } j=i\\ \lambda_i Q_{ij}, & \text{ha } j\neq i\end{cases}$

$$\boxed{G=\begin{pmatrix} -\frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{10} & \frac{1}{10} & -\frac{1}{5} \end{pmatrix}}$$

b.) A $G^T\pi^T=0$ lineáris egyenletrendszert oldjuk meg:

$$\left(\begin{array}{ccc|c} -\frac{1}{2} & 0 & 1/10 & 0 \\ 1/4 & -1/3 & 1/10 & 0 \\ 1/4 & 1/3 & -1/5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & 0 & 1/5 & 0 \\ 3/4 & -1 & 3/10 & 0 \\ 3/4 & 1 & -3/5 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} -1 & 0 & 1/5 & 0 \\ 0 & -1 & 9/20 & 0 \\ \cancel{0} & \cancel{1} & \cancel{-9/20} & \cancel{0} \end{array}\right) \Rightarrow \begin{array}{l} \pi_1=\frac{1}{5}\pi_3 \\ \pi_2=\frac{9}{20}\pi_3 \end{array}$$

Pl $\pi_3=20$ választással $\pi=(4 \;\; 9 \;\; 20)$, lenormálva

$\boxed{\pi=\begin{pmatrix}\frac{4}{33} & \frac{9}{33} & \frac{20}{33}\end{pmatrix}}$ az egyetlen stac. eloszlás.

c.) A ~~z~~ Markov lánc véges állapotterű és irreducibilis, ezért az ergodtétel szerint a 3-as állapotban töltött idő aránya hosszú távon $\boxed{\pi_3 = \frac{20}{33} \approx 60.6\%}$.

Avagy: ! $f(i) = \begin{cases} 1, & \text{ha } i=3 \\ 0, & \text{ha nem} \end{cases}$, $f: S \to \mathbb{R}$ a 3-as állapot indikátora.

Az ergodtétel miatt

$$\cancel{f(X_0)+f(X_1)+\dots+f} \quad \frac{1}{T}\int_0^T f(X(t))\,dt = \sum_{i \in S} f(i)\pi_i = \pi f$$

$$f = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{4}{33} & \frac{9}{33} & \frac{20}{33} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \underline{\underline{\frac{20}{33}}}$$

ahol $X(t)$ a Markov lánc.

d.) 1. megoldás: Keressük meg a beépített diszkrét idejű M.L. stac. eloszlását: $(Q^T - \mathbb{1})\pi^{d^T} = 0$ alapján $\pi^d = \begin{pmatrix} \frac{2}{9} & \frac{3}{9} & \frac{4}{9} \end{pmatrix}$. Vagyis hosszú távon (az ergodtétel miatt)

- a menetek $\frac{2}{9}$-e egyéni – ezek átlag 2 percesek
- a menetek $\frac{3}{9}$-e páros – ezek átlag 3 percesek
- $\frac{4}{9}$-e csoportos – ezek átlag 5 percesek

$\Longrightarrow$ a négy átlag $\frac{2}{9}\cdot 2 + \frac{3}{9}\cdot 3 + \frac{4}{9}\cdot 5 = \frac{33}{9} = \frac{11}{3} = 3.6\dot{6}$ perc, vagyis $\boxed{\text{3 perc 40 másodperc}}$.

d.) <u>2. megoldás</u>: Legyen $\Delta t \ll 1$ <u>rövid</u> idő!

Hány ugrás történik a $[t, t+\Delta t)$ idő-intervallumban?

Jelöljük ezt $N_{t,\Delta t}$ -vel!

Mivel Δt kicsi, $\mathbb{P}(N_{t,\Delta t} \geq 2) \approx 0$ elhanyagolható, így

$$\mathbb{E} N_{t,\Delta t} \approx \mathbb{P}(N_{t,\Delta t} = 1) \approx \mathbb{P}(N_{t,\Delta t} \geq 1) :$$

ez annak a val.sége, hogy <u>történik ugrás.</u>

Az exponenciális eloszlás természete miatt

$$\left.\begin{array}{l} \mathbb{P}(\text{történik ugrás} \mid X(t) = 1) \approx \lambda_1 \Delta t \\ \mathbb{P}(\quad -\!\!/\!/\!\!- \quad \mid X(t) = 2) \approx \lambda_2 \Delta t \\ \mathbb{P}(\quad -\!\!/\!/\!\!- \quad \mid X(t) = 3) \approx \lambda_3 \Delta t \end{array}\right\} \text{ahol } X(t) \in S \text{ a Markov lánc.}$$

Ha a lánc (már) stacionárius, akkor $\mathbb{P}(X(t) = i) = \pi_i$

$$\xRightarrow[\text{tétel}]{\text{teljes valsés}} \mathbb{E} N_{t,\Delta t} \approx \mathbb{P}(\text{történik ugrás}) \approx \sum_{i \in S} \pi_i \lambda_i \Delta t$$

<u>Köv.</u>: $T > 0$ idő alatt

ha $T = n\Delta t$,

$$\text{akkor } \mathbb{E} N_{0,T} = \sum_{i=0}^{n-1} \mathbb{E} N_{t_i,\Delta t} \approx n \sum_{i \in S} \pi_i \lambda_i \Delta t = T \sum_{i \in S} \pi_i \lambda_i$$

$$\text{Esetünkben } \sum_{i \in S} \pi_i \lambda_i = \frac{4}{33} \cdot \frac{1}{2} + \frac{9}{33} \cdot \frac{1}{3} + \frac{20}{33} \cdot \frac{1}{5} = \frac{2+3+4}{33} = \frac{9}{33} = \frac{3}{11}$$

$\Longrightarrow$ egy <u>hosszú</u> T idő alatt kb $\frac{3}{11} \cdot T$ ugrás történik

$\Longrightarrow$ az átlagos tartózkodási idő $\dfrac{T}{\frac{3}{11} T} = \underline{\underline{\dfrac{11}{3}}} = \underline{\underline{3 \text{ min } 40 \text{ sec}}}$