VillamosMSc Sztoch pótZH Megoldások 2021 ősz ; 2021.12.06

(1)

a.) Legyen X a verekedések száma 1 év alatt.

$X \sim Bin(n,p)$ ahol $n = 365$, $p = \frac{1}{100}$

$$\Rightarrow \mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X \leq 2) = 1 - \left[\mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2)\right] =$$

$$= 1 - \left[\binom{365}{0}\left(\frac{1}{100}\right)^0\left(\frac{99}{100}\right)^{365} + \binom{365}{1}\left(\frac{1}{100}\right)^1\left(\frac{99}{100}\right)^{364} + \binom{365}{2}\left(\frac{1}{100}\right)^2\left(\frac{99}{100}\right)^{363}\right]$$

Így is kiszámolható.

Avagy: n nagy, p kicsi $\Rightarrow$ X jó közelítéssel $\sim Poi(\lambda)$

ahol $\lambda = np = 3.65$

$$\Rightarrow \mathbb{P}(X \geq 3) = 1 - \sum_{k=0}^{2} \mathbb{P}(X=k) \approx 1 - \sum_{k=0}^{2} e^{-\lambda}\frac{\lambda^k}{k!} =$$

$$= 1 - e^{-\lambda}\left[1 + \lambda + \frac{\lambda^2}{2}\right] = 1 - e^{-3.65}\left[1 + 3.65 + \frac{3.65^2}{2}\right]$$

$$\approx 0.706 = 70.6\%$$

b.) Legyen $X(t)$ a verekedések száma t nap alatt. Ha t nagy, akkor $X(t)$ jó közelítéssel $\sim Poi\left(t \cdot \frac{1}{100}\right)$

$$\Rightarrow \mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X=0) \approx 1 - e^{-t \cdot \frac{1}{100}} := \frac{1}{2},$$

vagyis $t = -100 \ln\frac{1}{2} = 100 \ln 2 \approx 69.3$ (nap)

Avagy: Y jelölje azt, hogy hanyadik napon van az első verekedés. Így $Y \sim Geom\left(\frac{1}{100}\right)$ $\Rightarrow$ $\forall k \in \mathbb{N}$-re

$$\mathbb{P}(Y > k) = \mathbb{P}(k \text{ napig nincs verekedés}) = \left(\frac{99}{100}\right)^k := \frac{1}{2}$$

$$\Rightarrow \text{kb } k = \log_{\frac{99}{100}} \frac{1}{2} = \frac{\ln 1/2}{\ln 99/100} \approx \underline{\underline{69.0}} \text{ napra van szükség.}$$

②

a) A képletgyűjteményből X és Y közös generátorfüggvénye

$$\cancel{p_x}\ g_X(z) = g_Y(z) = \frac{pz}{1-qz} \quad \text{ahol } q = 1-p$$

~~+~~ és függetlenek $\Rightarrow$ $Z = X+Y$ generátorfüggvénye

$$g_Z(z) = g_X(z) \cdot g_Y(z) = \left(\frac{pz}{1-qz}\right)^2$$

b) A képletgyűjteményből $Var X = Var Y = \frac{q}{p^2}$, így a

függetlenség miatt $Var(Z) = Var(X+Y) = Var(X) + Var(Y)$

$$= 2\frac{q}{p^2} = 2\frac{1-p}{p^2}$$

(3)

a.) Legyen $X(t)$ a bent lévő vendégek száma t óra eltel-tével. Ez folytonos idejű időben homogén Markov lánc az $S=\{0,1,2\}$ állapottéren, és az ugrási ráták

$$\begin{tikzcd} 0 \arrow[r, bend left, "6"] & 1 \arrow[l, bend left, "3"] \arrow[r, bend left, "6"] & 2 \arrow[l, bend left, "6"] \end{tikzcd}$$

óránként átlag 6 vendég érkezik

Ha 1 vendég van, a távozás rátája 3, mert várhatóan $\frac{1}{3}$ órát marad

Ha 2 vendég van, a távozási ráták összeadódnak, így 6 rátával megy el valamelyik.

Ebből az infinitezimális generátor $G=\begin{pmatrix} -6 & 6 & 0 \\ 3 & -9 & 6 \\ 0 & 6 & -6 \end{pmatrix}$

$\Rightarrow$ a $\pi(t)$ időfejlődését leíró diff. egyenletrendszer

$$\boxed{\dot{\pi}(t)=\pi(t)\,G=\pi(t)\begin{pmatrix} -6 & 0 & 0 \\ 3 & -9 & 6 \\ 0 & 6 & -6 \end{pmatrix}}$$

③

b.) $t=12$ óra hosszú idő, $X(t)$ pedig véges állapotterű, irreducibilis és folytonos idejű, ezért a Markov láncok alaptétele szerint $\Pi(t) \approx \Pi$, ahol Π az egyetlen stacionárius eloszlás. Mivel $X(t)$ születési-halálozási folyamat, Π leolvasható a gráf-reprezentációból:

$$\begin{cases} \Pi_0 \cdot 6 = \Pi_1 \cdot 3 \\ \Pi_1 \cdot 6 = \Pi_2 \cdot 6 \end{cases} \iff \Pi_2 = \Pi_1 = 2\Pi_0$$

vagyis $\Pi = \text{konstans} \cdot (1 \quad 2 \quad 2)$.

Ahhoz hogy ez eloszlásvektor legyen, kell $\Pi_0 + \Pi_1 + \Pi_2 = 1$,

vagyis $\Pi = \left(\frac{1}{5} \quad \frac{2}{5} \quad \frac{2}{5}\right) = (0.2 \quad 0.4 \quad 0.4)$

$$\Rightarrow \mathbb{P}(X(t) \geq 1) \overset{t=12}{\approx} \Pi_1 + \Pi_2 = 0.8 = 80\%$$

(4)

Legyen $n=80$, és $i=1,2,\dots,n$-re legyen

$$X_i = \begin{cases} 1, & \text{ha az } i\text{-edik hallgató megbukik} \\ 0, & \text{ha nem.}\end{cases}$$

Így $X_i \sim B(p)$ ahol $p=\frac{1}{10}$, és függetlenek.

$S_n := X_1+X_2+\dots+X_n$ a megbukó — vagyis pótZH-t író — hallgatók száma. A keresett esemény

$$A=\{S_n\ge 21\}$$

a.) Az X_i-k korlátosak: $0=a_i\le X_i\le b_i=1$,

továbbá $\mathbb{E}S_n = np = 80\cdot\frac{1}{10}=8 \implies \underline{t:=13}$

választással a Hoeffding egyenlőtlenség szerint

$$\mathbb{P}(S_n\ge21)=\mathbb{P}(S_n\ge \mathbb{E}S_n+t)\le \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i-a_i)^2}\right)=$$

$$=\exp\left(-\frac{2\cdot13^2}{80\cdot(1-0)^2}\right)=\exp\left(-\frac{169}{40}\right)=e^{-4.225}\approx 0.015=1.5\%$$

b.) $m:=\mathbb{E}X_i=p=\frac{1}{10}$ a Cramér tétele szerint ahol $\boxed{a:=\frac{21}{80}>m}$, $b:=\infty$

$$\mathbb{P}(S_n\ge21)=\mathbb{P}\left(\frac{S_n}{n}\ge\frac{21}{80}\right)=\mathbb{P}\left(\frac{S_n}{n}\in[a,b]\right)\lessapprox$$

$$\lessapprox e^{-nI(a)}=e^{-80\cdot I(\frac{21}{80})},$$ ahol I a $p=\frac{1}{10}$ paraméterű

$B(p)$ eloszlás Cramér rátafv-e: $I(x)=x\ln\frac{x}{p}+(1-x)\ln\frac{1-x}{1-p}$

$$\Rightarrow I\left(\frac{21}{80}\right)=\frac{21}{80}\ln\frac{21/80}{1/10}+\frac{59}{80}\ln\frac{59/80}{9/10}\approx 0.10647634$$

(4) b.) folyt.

$$\Rightarrow 80 \cdot I\left(\frac{21}{80}\right) \approx 8.5181$$

$$\Rightarrow P(S_n \geq 21) \lesssim e^{-8.5181} \approx 0.00020 = 0.02\%.$$

Mint látható, a Cramér tétel sokkal erősebb becslést ad.

(5)

a.) $\mathbb{P}(X_i=0)=\mathbb{P}(X_i=1000)=\frac{1}{2} \Rightarrow \boxed{m:=\mathbb{E}X_i=\frac{0+1000}{2}=500}$

Ugyanis $\mathbb{E}(X_i^2)=\frac{0^2+1000^2}{2}=500000$

$\Rightarrow \sigma^2:=\operatorname{Var}X_i=\mathbb{E}(X_i^2)-(\mathbb{E}X_i)^2=500000-500^2$

$=250000$

$\Rightarrow \boxed{\sigma=\sqrt{\operatorname{Var}}=\sqrt{\sigma^2}=500}=\mathbb{D}X_i$

[Hát persze: $X_i=1000\cdot Y_i$ ahol $Y_i\sim B\left(\frac{1}{2}\right)$, amiből

$\mathbb{E}Y_i=\frac{1}{2}$, $\operatorname{Var}Y_i=\frac{1}{2}\left(1-\frac{1}{2}\right)=\frac{1}{4}$, $\mathbb{D}Y_i=\frac{1}{2}$

$\Rightarrow \mathbb{E}X_i=1000\cdot\frac{1}{2}$, $\mathbb{D}X_i=1000\cdot\frac{1}{2}$]

Mivel X_i csak 0 vagy 1000 lehet, és ~~ezért~~ $m=500$,

ezért $|X_i-m|^3$ lehetséges értékei $|0-500|^3 = 500^3$ illetve $|1000-500|^3 = 500^3$,

vagyis $|X_i-m|^3$ biztosan 500^3

$\Rightarrow \boxed{\delta:=\mathbb{E}(|X_i-m|^3)=500^3}$

(5)

b.) Legyen $n = 2500$ és
$S_n = X_1 + X_2 + \dots + X_n$ a Télapó össz-költsége, petákban.

Mivel az X_i-k függetlenek és azonos eloszlásúak,

$$\mathbb{E}S_n = n\mathbb{E}X_i = 2500 \cdot 500 \text{, és}$$

$$\mathbb{D}S_n = \sqrt{n}\,\mathbb{D}X_i = 50 \cdot 500$$

A CHT szerint S_n normális eloszlással közelíthető,

Így $K := 1.3 \cdot 10^6$-nal

$$\mathbb{P}(S_n \le K) \approx \phi\left(\frac{K - \mathbb{E}S_n}{\mathbb{D}S_n}\right) = \phi\left(\frac{1300000 - 1250000}{25000}\right)$$

$$= \phi\left(\frac{50000}{25000}\right) = \underline{\phi(2)} \approx 0.98 = \underline{98\%}$$

c.) A Berry-Esseen tétel szerint a CHT közelítés hibája

$$\text{hiba} \le \frac{C\,\delta}{\sqrt{n}\,\sigma^3} = \frac{0.4748 \cdot \cancel{500^3}}{\sqrt{2500}\,\cancel{500^3}} = \frac{0.4748}{50} = 0.009496$$
$$\approx \underline{1\%}$$

[Megjegyzés: Persze minden számolás könnyebb / szebb lett volna, ha végig peták helyett $\boxed{\text{ePeták}}$-ban (vagyis „ezer peták”-ban) számolunk.]