

Mintavizsga feladatok

Sztochasztika

Horváth Illés

2024/12/06

A vizsga témái

- Valószínűesszámítás ismételés, nevezetes eloszlások, teljes valószínűség tétel, teljes várható érték tétel
- Generátorfüggvény
- Elágazó folyamatok
- Poisson folyamat
- Valószínűségi koncentrációs tételek
- Diszkrét idejű Markov láncok
- Folytonos idejű Markov láncok
- Markov-sorok
- Jackson-hálózatok
- Statisztika I: alapok, paraméterbecslés
- Statisztika II: hipotézisvizsgálat, várható érték tesztelése
- Statisztika III: hipotézisvizsgálat, nemparaméteres próbák

Egy számítógépes programban egy véletlen, rekurzív rutin fut: minden egyes részfolyamat véletlen számú, önmagával megegyező alfolyamatot indít. Az indított alfolyamatok száma 0, 1 vagy 2 lehet, melyek valószínűsége rendre 0.4, 0.2 és 0.4, a többi részfolyamattól függetlenül. Minden egyes részfolyamat egységnyi időt vesz igénybe.

- (a) Modellezzük a folyamatot elágazó folyamattól. Mik az egyedek és a leszármazottak? A folyamat szubkritikus, kritikus vagy superkritikus?
- (b) Mennyi a valószínűsége annak, hogy a program előbb-utóbb lefut?
- (c) Mennyi a program futási idejének a várható értéke?

Megoldás.

- (a) Az egyedek a rutinok, és ha egy rutin elindít egy másikat, az a közvetlen leszármazottja. Az utódeloszlás generátorfüggvénye

$$G(z) = \frac{1}{3}z^0 + \frac{1}{3}z^1 + \frac{1}{3}z^2,$$

várható értéke $G'(1) = 1$, így a folyamat kritikus.

- (b) Kritikus folyamatokra a kihalás valószínűsége 1, azaz annak a valószínűsége, hogy a program előbb-utóbb lefut, 1.
- (c) Kritikus folyamat esetén $E(N) = \infty$, ahol N a teljes populáció mérete, azaz a program a futása alatt várhatóan végtelen sok rutint indít, ami mind egységnyi ideig fut, azaz az eredeti program futási idejének várható értéke végtelen.

Egy kis ország elektromos hálózatára 8000 kis fogyasztó és 2000 nagy fogyasztó kapcsolódik. A csúcsidei fogyasztásuk:

- egy kis fogyasztó által felvett teljesítmény legfeljebb 6000W, átlagosan 500W;
- egy nagy fogyasztó által felvett teljesítmény legfeljebb 20000W, átlagosan 4000W.

Mekkora legyen a hálózat C összteljesítménye, ha azt szeretnénk, hogy a csúcsidőbeli fogyasztás legfeljebb 10^{-8} valószínűséggel lépje túl C -t?

Megoldás. Jelölje S azt, hogy mennyi a csúcsidőbeli fogyasztás. Ekkor

$$S = X_1 + \dots + X_n,$$

ahol X_i az i -edik fogyasztó által felvett teljesítmény, és $n = 10000$.

C értékét szeretnénk beállítani úgy, hogy

$$P(S > C) \leq 10^{-8}$$

teljesüljön.

A feltétel szerint $n_1 = 8000$ fogyasztóra

$$a_1 = 0 \leq X_i \leq 6000 = b_1 \quad n_1 = 8000 \text{ fogyasztóra, és}$$

$$a_2 = 0 \leq X_i \leq 20000 = b_2 \quad n_2 = 2000 \text{ fogyasztóra.}$$

Továbbá

$$E(S) = 8000 \times 500 + 2000 \times 4000 = 12000000 \text{ (W).}$$

Legyen $C = E(S) + t$; a Hoeffding-korlát alapján

$$P(S > C) = P(S > E(S) + t) \leq e^{-\frac{2t^2}{n_1(b_1 - a_1)^2 + n_2(b_2 - a_2)^2}}.$$

Állítsuk be t értékét úgy, hogy a jobboldal 10^{-8} legyen, azaz

$$\frac{2t^2}{8000(6000 - 0)^2 + 2000(20000 - 0)^2} = 8 \ln 10,$$

ahonnan

$$t = 3166000,$$

és a hálózat C összteljesítménye

$$C = E(S) + t = 15166000.$$

Az ügyfelek sorbanállnak egy belvárosi ATM-nél. Az ATM-et egyszerre egy ügyfél használhatja, a többiek addig várnak. Minden ügyfél átlagosan 1 percre használja az ATM-et. Átlagosan 2 percenként érkezik egy ügyfél. Ha legalább 2-en állnak az ATM-nél (beleértve azt is, aki éppen használja), akkor a további érkezők egyből távoznak és nem jönnek vissza.

- (a) Modellezzük az ATM-nél állók számát folytonos idejű Markov-lánccal. Mit kell feltenni, hogy teljesüljön a Markov-tulajdonság? Mik a lehetséges állapotok? Írjuk fel a generátort.
- (b) Hosszú távon átlagosan az idő mekkora részében van az ATM használatban?
- (c) A várakozást és a használatot is figyelembe véve átlagosan mennyi időt tölt egy ügyfél az ATM-nél?

Megoldás.

- (a) Ha az ügyfelek érkezési folyamata Poisson-folyamat, és az egy ügyfél által az automata használatával eltöltött idő is exponenciális eloszlású (és független minden mástól), akkor teljesül a Markov-tulajdonság.

Az állapottér $\{0, 1, 2\}$ aszerint, hogy hány ügyfél áll az automatánál.

A generátor

$$Q = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 1 & -3/2 & 1/2 \\ 0 & 1 & -1 \end{bmatrix},$$

mert átlagosan 2 percenként érkezik egy ügyfél $\rightarrow$ az érkezési ráta $1/2$, a kiszolgálási ráta pedig 1 .

- (b) Hosszú távon átlagosan az idő mekkora részében van az ATM használatban? Kiszámítjuk a $v_{\text{st}} = (x_0 \ x_1 \ x_2)$ stacionárius eloszlást. Ez egy Markov-sor, így teljesülnek az egyensúlyi egyenletek:

$$\frac{1}{2}x_0 = 1 \cdot x_1, \quad \frac{1}{2}x_1 = 1 \cdot x_2, \quad x_0 + x_1 + x_2 = 1,$$

ahonnan $x_0 : x_1 : x_2 = 4 : 2 : 1$, így

$$v_{\text{st}} = (4/7 \ 2/7 \ 1/7),$$

és az idő $x_1 + x_2 = 3/7$ részében van az ATM használatban.

- (c) A várakozást és a használatot is figyelembe véve átlagosan mennyi időt tölt egy ügyfél az ATM-nél? Little-formula alapján

$$L = W\lambda_e,$$

ahol L az átlagos sorhossz:

$$L = 0 \cdot x_0 + 1 \cdot x_1 + 2 \cdot x_2 = 4/7,$$

λ_e az effektív érkezési ráta, amibe nem számítanak bele azok az ügyfelek, akik akkor érkeznek, amikor tele van a sor:

$$\lambda_e = (1 - x_2)\lambda = \frac{3}{7},$$

és így az átlagos ATM-nél töltött idő:

$$W = L/\lambda_e = \frac{4}{3} \quad (\text{perc}).$$

Egy úton Poisson-pontfolyamat szerint haladnak el az autók, ismeretlen λ (autó/perc) rátával. 5 diszjunkt 2 perces intervallumban megszámoljuk az elhaladó autókat, és a következő mintát kapjuk: 4, 3, 1, 0, 4. Adjunk a minta alapján maximum likelihood-bebecslést vagy momentum-bebecslést λ értékére. Elég az egyik módszerrel, bármelyik választható. (Figyeljünk a megadott információkra és a kérdésre!)

1. megoldás. A 2 perc alatt érkező autók száma $\text{POI}(2\lambda)$ eloszlású, ez a háttéreloszlás. A momentum-becsléshez szükségünk van a várható értékére: $X_1 \sim \text{POI}(2\lambda)$ esetén

$$g(\lambda) = \mathbb{E}_\lambda(X_1) = 2\lambda,$$

és így

$$g^{-1}(x) = x/2,$$

és a momentum-becslés az $\bar{x} = \frac{1}{5}(4 + 3 + 1 + 0 + 4) = 2.4$ mintaátlag alapján

$$\hat{\lambda} = g^{-1}(\bar{x}) = \bar{x}/2 = 1.2.$$

2. megoldás. A maximum likelihood becsléshez szükség van a likelihood függvényre:

$$\begin{aligned} L(\lambda) &= \frac{(2\lambda)^4}{4!} e^{-2\lambda} \cdot \frac{(2\lambda)^3}{3!} e^{-2\lambda} \cdot \frac{(2\lambda)^1}{1!} e^{-2\lambda} \cdot \frac{(2\lambda)^0}{0!} e^{-2\lambda} \cdot \frac{(2\lambda)^4}{4!} e^{-2\lambda} \\ &= \frac{\lambda^{12}}{C} e^{-10\lambda}, \end{aligned}$$

a log-likelihood függvény pedig

$$\ell(\lambda) = -\log(C) + 12 \log(\lambda) - 10\lambda.$$

(A C egy pozitív konstans jelöl, a pontos értéke lényegtelen.)

$$\frac{d}{d\lambda} \ell(\lambda) = \frac{12}{\lambda} - 10 = 0,$$

aminek egyetlen megoldása $\lambda = 1.2$, ami a $\lambda > 0$ lehetséges paramétertartomány belsejében van, úgyhogy ez a maximum-likelihood becslés.

Amikor valaki elvégez egy IQ tesztet, az eredmény véletlenszerű, melynek várható értéke a valódi IQ, szórása 4. Jónás azt állítja, hogy az IQ-ja 120. Ennek bizonyítására elvégez 5 különböző IQ tesztet, melyek eredménye rendre 118, 116, 119, 119, 118.

Teszteljük 95%-os szinten azt a hipotézist, hogy Jónás IQ-ja 120, azon hipotézis ellenében, hogy az IQ-ja nem 120.

Megoldás. A várható értéket teszteljük; a próba, amit végzünk:

- μ -próba, mivel ismert a szórás;
- egymintás, mivel a minta várható értékét teszteljük ismert érték ellen;
- kétoldali, mivel az ellenhipotézis az, hogy Jónás IQ-ja nem 120.

Az ismert szórás $\sigma = 4$, a várható érték m ismeretlen. A hipotézis és ellenhipotézis:

- $H_0: m = 120$;
- $H_1: m \neq 120$.

Kiszámítjuk a statisztikát:

$$u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{118 - 120}{4} \sqrt{5} \approx 1.118.$$

Kiszámítjuk a kvantilist. A konfidenciaszint 95%, így $\varepsilon = 0.05$, és

$$u_{\varepsilon/2} = \phi^{-1}(0.975) = 1.96.$$

A próba:

$$u = 1.118 \in [-u_{\varepsilon/2}, u_{\varepsilon/2}] = [-1.96, 1.96]$$

teljesül, így elfogadjuk H_0 -t, miszerint Jónás IQ-ja 120.

Béla egy videójátékkal játszik, amelyben átlagosan 10 percenként zsákmányol fegyvert és 5 percenként aranyat.

- (a) Mekkora az esélye, hogy előbb zsákmányol fegyvert, mint aranyat?
- (b) Minden egyes zsákmányolt fegyver 10% eséllyel jobb, mint a jelenlegi fegyvere. Mekkora a valószínűsége, hogy a következő 3 óra játék alatt talál jobb fegyvert, mint a jelenlegi?
- (c) Mekkora az esélye, hogy a következő 30 perc alatt csak aranyat zsákmányol (beleértve azt az esetet is, hogy semmit sem zsákmányol)?

- (a) Mekkora az esélye, hogy előbb zsákmányol fegyvert, mint aranyat?

Az aranyak folyamata $PPP(1/5)$, a fegyverek folyamata $PPP(1/10)$ (a ráták mértékegysége 1/perc). A két folyamat független, így az uniójuk, vagyis az összes zsákmány folyamata $PPP(3/10)$, és minden egyes zsákmány $\frac{1/5}{1/5+1/10} = 2/3$ valószínűséggel arany és $\frac{1/10}{1/5+1/10} = 1/3$ valószínűséggel fegyver (a többitől függetlenül).

Az, hogy előbb zsákmányol fegyvert, mint aranyat, pont ugyanaz, mint az, hogy a következő zsákmány fegyver, így a valószínűsége $1/3$.

- (b) Minden egyes zsákmányolt fegyver 10% eséllyel jobb, mint a jelenlegi fegyvere. Mekkora a valószínűsége, hogy a következő 3 óra játék alatt talál jobb fegyvert, mint a jelenlegi?

A jelenleginél jobb fegyverek folyamata ritkítással kapható meg az összes fegyverek folyamatából, így az egy $\text{PPP}(1/10 \cdot 1/10 = 1/100)$. Ha X a 3 óra (180 perc) alatt talált fegyverek száma, akkor $X \sim \text{POI}(180 \cdot 1/100 = 1.8)$, és

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{1.8^0}{0!} e^{-1.8} \approx 0.8347.$$

- (c) Mekkora az esélye, hogy a következő 30 perc alatt csak aranyat zsákmányol (beleértve azt az esetet is, hogy semmit sem zsákmányol)?

Az, hogy csak aranyat zsákmányol, pont ugyanaz, mint az, hogy nulla fegyvert zsákmányol. Legyen Y a fegyverek száma 30 perc alatt; $Y \sim \text{POI}(30 \cdot 1/10 = 3)$, és

$$P(Y = 0) = \frac{3^0}{0!} e^{-3} = 0.0498.$$

Egy éttermi fogadásra 400 vendéget várnak, és kétféle menü lesz: A és B. Az A menüből 220 adagot tudnak készíteni, a B menüből 500 adagot. Minden vendég (egymástól függetlenül) 50 – 50% eséllyel választ A illetve B menüt. Mekkora az esélye, hogy nem jut mindenkinek abból, amit szeretne?

Megoldás. Legyen $X_i = 1$, ha az i -edik vendég A menüt kér, 0, ha B menüt. Ekkor $S = X_1 + \dots + X_n$ vendég kér összesen A menüt, ahol $n = 400$.

Pontosan akkor kaphat mindenki olyan menüt, amit szeretne, ha $S \leq 220$.

B/2. feladat

Megbecsüljük $P(S \leq 220)$ értékét. Ehhez szükségünk lesz a következő paraméterekre:

$$m = E(X_1) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = 0.5,$$

$$\sigma = D(X_1) = \sqrt{E(X_1^2) - (E(X_1))^2} = \sqrt{\frac{1}{2} \cdot 0^2 + \frac{1}{2} \cdot 1^2 - 0.5^2} = 0.5.$$

Végül a CHT alapján

$$\begin{aligned} P(S \leq 220) &= P\left(\frac{S - nm}{\sigma\sqrt{n}} \leq \frac{220 - nm}{\sigma\sqrt{n}}\right) \\ &\approx \phi\left(\frac{220 - nm}{\sigma\sqrt{n}}\right) = \phi\left(\frac{220 - 400 \cdot 0.5}{0.5 \cdot \sqrt{400}}\right) = \phi(2) = 0.9772, \end{aligned}$$

és így annak a valószínűsége, hogy *nem* jut mindenkinek abból, amit szeretne,

$$1 - 0.9772 = 0.0228.$$

Egy vásáron az arcfestő művész kétféle mintát fest: tigrist vagy pillangót. Egy tigrist 10 perc alatt fest meg, egy pillangót 20 perc alatt. Minden egyes gyerek $2/3$ eséllyel azt a mintát kéri, amit az előtte lévő, $1/3$ eséllyel a másikat. Hosszú sorban állnak a gyerekek, az arcfestő folyamatosan dolgozik.

- (a) Modellezzük az arcfestő tevékenységét diszkrét idejű Markov-lánccal! Mik az állapotok? Írjuk fel az átmenet-valószínűség mátrixot.
- (b) Hosszú távon az idő mekkora részét tölti tigris festésével?
- (c) Az arcfestő egy tigrisért 1800 forintot kér, egy pillangóért 1200 forintot. Mennyi az átlagos bevétele 8 óra alatt?

Megoldás.

- (a) Modellezzük az arcfestő tevékenységét diszkrét idejű Markov-lánccal! Mik az állapotok? Írjuk fel az átmenet-valószínűség mátrixot.

Osszuk fel az időt 10 perces időegységekre. 3 állapot lesz: tigris, vagy pillangó első fele, vagy pillangó második fele. Az átmenetvalószínűség-mátrix:

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

(b) Hosszú távon az idő mekkora részét tölti tigris festésével?

Kiszámítjuk a $v_{\text{st}} = (x_1 \ x_2 \ x_3)$ stacionárius eloszlást. A $v_{\text{st}} P = v_{\text{st}}$ egyenlet kifejtve:

$$2/3x_1 + 1/3x_3 = x_1$$

$$1/3x_1 + 2/3x_3 = x_2$$

$$x_2 = x_3$$

$$x_1 + x_2 + x_3 = 1,$$

aminek a megoldása

$$v_{\text{st}} = \left(\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3} \right).$$

Hosszú távon az idő $1/3$ részét tölti tigris festésével.

- (c) Az arcfestő egy tigrisért 1800 forintot kér, egy pillangóért 1200 forintot. Mennyi az átlagos bevétele 8 óra alatt?

Ergodtételt alkalmazunk; a tigris állapothoz rendeljünk 1800 forintot, a pillangó első feléhez 0-t, a második feléhez 1200 forintot. Ekkor az ergodtétel miatt a hosszú távú átlagos bevétele időegységenként

$$\frac{1}{3} \cdot 1800 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1200 = 1000 \quad (\text{forint}).$$

Viszont a kérdés az, mennyi az átlagos bevétele 8 óra alatt. Az időegység 10 perc, 8 órában összesen 48 időegység van, így alatta az átlagos bevétele 48000 forint.

2. megoldás.

- (a) Lehet másképp is modellezni diszkrét idejű Markov-lánccal. Legyen egy lépés egy minta megfestése (ekkor ugyan a lépések nem egyforma ideig tartanak, de se baj, ezt majd figyelembe vesszük később).

2 állapot van: tigris, vagy pillangó. Az átmenetvalószínűség-mátrix:

$$P = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}.$$

(b) Hosszú távon az idő mekkora részét tölti tigris festésével?

A stacionárius eloszlás ezúttal

$$v_{\text{st}} = (1/2 \ 1/2),$$

ami azt jelenti, hogy hosszú távon átlagosan ugyanannyi pillangót fest, mint tigrist.

v_{st} -t súlyozzuk azzal, mennyi időt tölt el egy tigrissel illetve pillangóval:

$$u_{\text{st}} = \left(\frac{1/2 \cdot 10}{1/2 \cdot 10 + 1/2 \cdot 20} \ \frac{1/2 \cdot 20}{1/2 \cdot 10 + 1/2 \cdot 20} \right) = (1/3 \ 2/3),$$

azaz átlagosan az idő $1/3$ részét tölti tigris festésével.

- (c) Az arcfestő egy tigrisért 1800 forintot kér, egy pillangóért 1200 forintot. Mennyi az átlagos bevétele 8 óra alatt?

Az ergodtétel alapján az egy mintára eső átlagos bevétele

$$1/2 \cdot 1800 + 1/2 \cdot 1200 = 1500 \quad (\text{forint}).$$

Szintén az ergodtétel alapján egy minta átlagosan

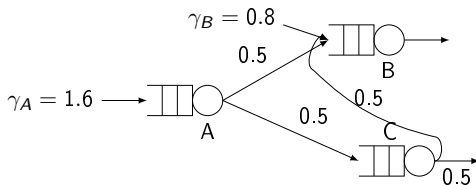
$$1/2 \cdot 20 + 1/2 \cdot 10 = 15$$

percig tart, így 8 órába átlagosan 32 minta elkészítése fér bele, és a 8 óra alatt az átlagos bevétele

$$32 \times 1500 = 48000 \quad (\text{forint}).$$

B/4. kérdés

Az alábbi hálózatban mindhárom szerver FIFO, a kiszolgálási rátájuk rendre $\mu_A = 2, \mu_B = 3, \mu_C = 1$.



- (a) Stabil-e a hálózat? Aciklikus-e a hálózat?
- (b) Mekkora az egyes szerverek terheltsége?
- (c) Egy kívülről B-be beérkező igény átlagosan mennyi időt tölt a rendszerben?

Megoldás.

- (a) A hálózat aciklikus, A, C, B sorrendben véve a szervereket. Az egyes szerverek beérkezési rátája így sorban expliciten számolható:

$$\lambda_A = \gamma_A = 1.6,$$

$$\lambda_C = 0.5\lambda_A = 0.8,$$

$$\lambda_B = \gamma_B + 0.5\lambda_A + 0.5\lambda_C = 0.8 + 0.8 + 0.4 = 2.0.$$

A stabilitás feltétele:

$$\lambda_A = 1.6 < \mu_A = 2.0,$$

$$\lambda_C = 0.8 < \mu_C = 1.0,$$

$$\lambda_B = 2.0 < \mu_B = 3.0$$

teljesül, így a hálózat stabil.

- (b) Mekkora az egyes szerverek terheltsége?

Megoldás.

$$\rho_A = \lambda_A / \mu_A = 1.6 / 2.0 = 0.8,$$

$$\rho_C = \lambda_C / \mu_C = 0.8 / 1.0 = 0.8,$$

$$\rho_B = \lambda_B / \mu_C = 2.0 / 3.0 = 0.667.$$

- (c) Egy kívülről B-be beérkező igény átlagosan mennyi időt tölt a rendszerben?

Megoldás. Egy kívülről B-be érkező igény csak a B szerverben tölt el időt. A B szerver sora M/M/1 típusú, így az átlagos sorban töltött idő

$$\frac{1}{\mu_B - \lambda_B} = \frac{1}{3.0 - 2.0} = 1.$$

Egy irodaházban a tűzriadók Poisson-folyamat szerint történnek, ismeretlen λ (tűzriadó per hónap) rátával. Feljegyeztük az egymást követő tűzriadók között eltelt napok számát, amire a következő minta adódott: 7, 12, 44, 29, 28. A minta alapján adjunk maximum-likelihood becslést λ értékére. (Egy hónapot tekintsünk 30 napnak.)

Megoldás. A háttéreloszlás az érkezési időköz, ami $\text{EXP}(\lambda)$. A minta likelihood-függvénye

$$\begin{aligned} L(\lambda) &= \lambda e^{-\lambda x_1} \dots \lambda e^{-\lambda x_5} = \\ &= \lambda^5 e^{-\lambda(x_1 + \dots + x_5)} = \lambda^5 e^{-120\lambda}. \end{aligned}$$

A log-likelihood függvény

$$\ell(\lambda) = 5 \log \lambda - \frac{120}{\lambda}.$$

A

$$\frac{d}{d\lambda} \ell(\lambda) = \frac{5}{\lambda} - 120 = 0$$

egyenlet egyetlen megoldása $\hat{\lambda} = 1/24$ (tűzriadó/nap), vagy az eredeti kérdésre átváltva 30/24 tűzriadó/hónap, ez a maximum likelihood becslés.

0/8