Sztochasztikus folyamatok
8. feladatsor
Martingálok II.

8.1 Lássa be, hogy ha T_1 és T_2 megállási idők az $\big(\mathcal{F}_n\big)_{n=1}^{\infty}$ filtrációra nézve, akkor $T_1 \wedge T_2$ is megállási idő.

8.2 Legyenek X_i, $i = 1, 2, \ldots$ független és azonos eloszlású valószínűségi változók, melyeknek közös eloszlása

$$\mathbf{P}(X_i = 1) = p, \qquad \mathbf{P}(X_i = -1) = q = 1 - p.$$

Legyen $S_n := X_1 + X_2 + \cdots + X_n$.

(a) Mutassuk meg, hogy $Z_n := (q/p)^{S_n}$ martingál, melynek várható értéke 1.

(b) $a \in \mathbb{N}$ esetén jelölje $T_a = \min\{n : S_n = a\}$ az a elérési idejét. A martingál-megállítási tétel segítségével számolja ki a $\mathbf{P}(T_a < T_b)$ valószínűséget, ahol $a < 0$, $b > 0$.

(c) Mely $C > 0$ és $\lambda > 0$ választások mellett lesz

$$Z_n := C^n \lambda^{S_n}$$

martingál? Melyik korábbi feladat átfogalmazása ez?

8.3 Legyenek $X_1, X_2, \ldots$ független, azonos eloszlású, véges μ várhatóértékű valószínűségi változók és ν tőlük független, véges várható értékű nemnegatív egész értékű valószínűségi változó. Legyen $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n, \nu \cdot \mathbb{1}[\nu \leq n])$. Lássa be, hogy $M_n = S_n - n\mu$ martingál és ν megállási idő az $\mathcal{F}_n$ filtrációra nézve, és hogy $\mathbf{E}(\sum_{n=1}^{\nu} X_n) = \mathbf{E}(\nu)\mu$.

8.4 Legyenek $X_1, X_2, \ldots$ független, azonos eloszlású, véges μ várhatóértékű, pozitív valószínűségi változók, $T(n) = \sum_{k=1}^{n} X_k$ az n-edik felújítási időpont és $N(t) = \max\{n : T(n) \leq t\}$ a $(0, t]$ intervallumba eső felújítások száma. $\mathcal{F}_n := \sigma(X_1, X_2, \ldots, X_n) = \sigma(T(1), T(2), \ldots, T(n))$.

(a) Mutassa meg, hogy tetszőleges t-re $N(t) + 1$ megállási idő, és hogy

$$\mathbf{E}\big(T(N(t) + 1)\big) = \mathbf{E}\big(N(t)\big)\mu + \mu$$

Segítség: Igaz, hogy semelyik órán tanult martingál-megállítási tétel nem alkalmazható, de a dominált konvergencia-tétel feltételeit kézzel le lehet ellenőrizni...

(b) Igaz-e, hogy $\mathbf{E}\big(T(N(t))\big) = \mathbf{E}\big(N(t)\big)\mu$?

8.5 Tekintsük az irreducibilis, szubsztochasztikus P mátrixot, és a hozzá tartozó Markov láncot (amit úgy kapunk, hogy hozzáveszünk egy elnyelő állapotot, és a maradék valószínűségekkel mindenhonnan az elnyelő állapotba lépünk). A *Perron-Frobenius*-tétel miatt P legnagyobb abszolútértékű λ sajátértéke $0 < \lambda < 1$, a sajátérték egyszeres, a hozzá tartozó jobb oldali $\mathbf{v}$ sajátvektor csupa pozitív elemből áll, és a többi sajátártékhez tartozó sajátvektor nem ilyen. Lássa be, hogy annak a valószínűsége, hogy X_n még nem az elnyelő állapot, legfeljebb $\frac{\mathbf{v}(X_0)}{\mathbf{v}_{min}}\lambda^n$, ahol $\mathbf{v}_{min}$ a $\mathbf{v}$ vektor legkisebb elemét jelöli.

8.6 *A Wright-Fisher modell*

Van m golyó egy dobozban, az n-edik lépésben X_n darab piros és $m - X_n$ darab kék. Az $n+1$-edik golyó-leosztást úgy kapjuk, hogy m-szer húzunk visszatevéssel az n-edik dobozból, és olyan színű golyókat rakunk az új dobozba, amilyeneket kihúztunk.

(a) A Markov láncnak van két elnyelő állapota. Mekkora valószínűséggel melyikben nyelődik el?

(b) Lássa be, hogy $\frac{X_n(m-X_n)}{(1-\frac{1}{m})^n}$ martingál. A 5. feladat segítségével mondjon egy olyan n-et, hogy a fele piros-fele kék állapotból indított Markov lánc az n-edik lépésben már legalább $\frac{1}{2}$ valószínűséggel valamelyik egyszínű állapotban lesz.

8.7 Legyen X_n, $n = 0, 1, 2, \ldots$ irreducibilis Markov lánc a *véges* S állapottéren, melynek átmenetmátrixát jelöljük P-vel. Legyen

$$T := \min\{n > 0 : X_n = X_0\}$$

a kezdő állapotba való első visszatérés ideje. Tegyük fel, hogy a kezdeti állapot $X_0 = z$ és minden $y \in S$-re legyen

$$r(y) := \mathbf{E}\Big(\sum_{i=0}^{T-1} \mathbb{1}_{\{X_i = y\}}\Big),$$

a kezdő állapotba való első visszatérés előtt az y állapotban töltött idő várható értéke. Vegyük észre, hogy ha $X_0 = z$, akkor $r(z) = 1$. Bizonyítsuk be, hogy

$$\sum_{x \in S} r(x) P_{x,y} = r(y),$$

azaz az $r(x)$ számokból képezett sorvektor az 1 sajátértékhez tartozó baloldali sajátvektora a Markov lánc P átmenetmátrixának.

Segítség: $\sum_{i=1}^{n}\big(\mathbb{1}_{\{X_i=y\}} - \mathbf{E}(\mathbb{1}_{\{X_i=y\}}|\mathcal{F}_{i-1})\big)$ korlátos növekményű martingál és T véges várható értékű megállási idő a természetes filtrációra nézve.

8.8 Legyen S diszkrét (véges vagy megszámlálható) állapottér és felette a $\big(P_{x,y}\big)_{x,y\in S}$ sztochasztikus mátrix. Az $f : S \to \mathbb{R}$ függvényről azt mondjuk, hogy (a P átmenet mátrix szerint) *szuperharmonikus az $x \in S$ pontban*, ha

$$\sum_{y \in S} P_{x,y} f(y) \le f(x).$$

Rögzítsünk egy $z \in S$ állapotot. Legyen $\mathcal{A}_z$ azon $f : S \to \mathbb{R}$ függvények halmaza, melyekre

(i) $f(z) = 1$,

(ii) minden $y \in S$-re $0 \le f(y) \le 1$,

(iii) f szuperharmonikus minden $x \neq z$ pontban.

(a) Legyen X_n Markov lánc az S-en, melynek átmenetmátrixa P, de tegyük a z állapotot elnyelővé. Lássa be, hogy $f \in \mathcal{A}_z$ esetén $f(X_n)$ szupermartingál.

(b) Értelmezzük a $g_z : S \to \mathbb{R}$ függvényt a következőképpen:

$$g_z(x) = \inf_{f \in \mathcal{A}_z} f(x).$$

Mutassuk meg, hogy $g_z \in \mathcal{A}_z$.

(c) Mutassuk meg, hogy bármely $x \neq z$ pontban

$$\sum_{y \in S} P_{x,y} g_z(y) = g_z(x),$$

azaz a g_z függvény *harmonikus* az $S \setminus \{z\}$ halmazon.

Útmutatás: Feltéve, hogy $\sum_{y \in S} P_{x,y} g_z(y) < g_z(x)$ valamely $x \neq z$ pontban, mutassuk meg, hogy g_z csökkenthető lenne egy picivel az x pontban, úgy, hogy szuperharmonikus tulajdonsága megmaradna minden $y \neq z$ pontban.

(d) Mutassuk meg, hogy ha $g_z(x) < 1$ valamely $x \neq z$ pontban, akkor

$$\inf_{y \in S} g_z(y) = 0.$$

Útmutatás: Feltéve, hogy $\inf_{y \in S} g_z(y) =: \varepsilon > 0$ tekintsük a $h_z(x) = (g_z(x) - \varepsilon)/(1 - \varepsilon)$ függvényt.

(e) Legyen $\hat{g}_z(x)$, $x \in S$, annak a valószínűsége, hogy az x állapotból indított lánc valaha is eléri a z elnyelő állapotot, azaz $\hat{g}_z(x) = \mathbf{P}_x(T_z < +\infty)$. Bizonyítandó, hogy $\hat{g}_z(x)$ harmonikus az $S \setminus \{z\}$ halmazon, és hogy $\hat{g}_z \in \mathcal{A}_z$.

(f) Legyen $M_n = g_z(X_n)$. A martingál konvergenciatétel és a dominált konvergenciatétel segítségével lássa be, hogy $g_z(x) \ge \hat{g}_z(x)$.

(g) A fentiek segítségével lássuk be a következő tételt:

Tétel. Legyen egy S diszkrét állapotterű irreducibilis Markov lánc átmenet mátrixa P. Ha létezik egy $z \in S$ állapot és egy $f : S \to \mathbb{R}$ függvény, melyre igaz, hogy

(i) $\forall y \in S :\ 0 \le f(y) \le 1$ és $f(z) = 1$,

(ii) f szuperharmonikus az $S \setminus \{z\}$ halmazon,

(iii) valamely $y \neq z$-re $f(y) < 1$.

Ekkor a Markov lánc tranziens.

8.9 Egy szerencsejátékos minden fogadási menetben egyenlő valószínűséggel nyer vagy veszít 1 petákot. Azzal a szilárd elhatározással kezd el játszani, hogy amikor n petáknyi nyereséget vagy m petáknyi veszteséget ér el, abbahagyja a játékot.

(a) Számoljuk ki annak a valószínűségét, hogy nyereséggel távozik.

(b) Számoljuk ki fogadásainak várható számát.

8.10 A "most piros" játék szabályai a következőek: a játékvezetőnél van egy jól megkevert francia-kártya-pakli, aminek a lapjait egyenként megmutatja a játékosnak, akinek az utólsó lap megmutatása előtt valamikor azt kell mondania, hogy "most piros", és ha a következő lap színe tényleg piros, akkor a játékos nyert. Mi a legjobb stratégiája?

8.11 Legyen X_n diszkrét idejű születési-halálozási folyamat, melynek átmenet valószínűségei:

$$P_{i,i+1} = p_i, \qquad P_{i,i-1} = 1 - p_i =: q_i, \qquad p_0 = 1.$$

Értelmezzük a

$$g : \mathbb{N} \to \mathbb{R}_+, \qquad g(k) := 1 + \sum_{j=1}^{k-1} \prod_{i=1}^{j} \frac{q_i}{p_i}$$

függvényt.

(a) Bizonyítandó, hogy

$$Z_n := g(X_n)$$

martingál a természetes filtrációra nézve.

(b) Rögzítsük a $0 < i < n$ számokat. Számoljuk ki annak a valószínűségét, hogy az i állapotból indított folyamat előbb éri el az n állapotot, mint a 0-t.

8.12 *Elcserélt kalapok — turbo változat.*

n személy egy kupacba dobja a kalapját és véletlenszerűen kihúznak a kupacból egyet-egyet. Azok, akik a véletlenül éppen a saját kalapjukat húzták, boldogan hazamennek. A maradók újból egy kupacba dobják a náluk lévő kalapokat és újból véletlenszerűen kihúznak egyet-egyet. Akik most véletlenül pont a saját kalapjukat húzzák, hazamennek. Ez folytatódik, mindaddig, amíg mindenki meg nem találja a saját kalapját. Legyen R a menetek száma (pl. $R = 1$, ha az elsű húzásnál mindenki éppen a saját kalapját húzza). Számoljuk ki $\mathbf{E}(R)$-t.

8.13 Két (korrekt) kockával dobunk és feljegyezzük az számok összegét. Mindaddig dobunk ameddig a $7, 2, 12, 7, 2$ eredménysorzat meg nem jelenik. Mennyi a dobások számának várhatú értéke?

8.14 Legyenek $X_1, \ldots, X_n$ független valószínűségi változók. Legyen φ olyan n változós függvény, hogy tetszőleges $x_1, \ldots, x_n$-re és $1 \le k \le n$-re és $\hat{x}_k$-ra

$$|\varphi(x_1, \ldots, x_k, \ldots, x_n) - \varphi(x_1, \ldots, \hat{x}_k, \ldots, x_n)| \le 1$$

Legyen $Y := \varphi(X_1, \ldots, X_n)$.

Bizonyítandó, hogy tetszőleges $a \ge 0$-ra

$$\mathbf{P}\left(Y - \mathbf{E}(Y) \ge a\right) \le e^{-a^2/2n},$$
$$\mathbf{P}\left(Y - \mathbf{E}(Y) \le -a\right) \le e^{-a^2/2n}.$$

8.15 Tekintsük az $\{1, 2, \ldots, n\}$ halmaz egy véletlen f leképezését önmagára, azaz $f(1), \ldots, f(n)$ függetlenek és egyenletes eloszlásúak $\{1, 2, \ldots, n\}$-en. Bizonyítandó, hogy

$$\mathbf{P}\left(\left|\frac{|\mathcal{R}(f)|}{n} - \left(1 - (1 - \frac{1}{n})^n\right)\right| \ge \frac{\lambda}{\sqrt{n}}\right) \le 2\exp(-\frac{\lambda^2}{2})$$

8.16 n hosszú $0-1$ sorozatok Hamming távolsága egyenlő azon koordináták számával, ahol a két sorozat különbözik:

$$\rho(\mathbf{x}, \mathbf{y}) := \sum_{i=1}^{n} |x_i - y_i|.$$

Legyen $\mathcal{A}$ tetszőleges részhalmaza az n hosszú $0-1$ sorozatok $\{0,1\}^n$ halmazának és $X_1, X_2, \ldots, X_n$ független, azonos eloszlású valószínűségi változók, melyeknek közös eloszlása $\mathbf{P}(X_i = 0) = 1/2 = \mathbf{P}(X_i = 1)$. Legyen

$$D_{\mathcal{A}} := \min_{\mathbf{x} \in \mathcal{A}} \rho(\mathbf{X}, \mathbf{x}).$$

Tetszőleges $a > 0$-ra találjunk felső korlátot a $\mathbf{P}\left(|D_{\mathcal{A}} - \mathbf{E}(D_{\mathcal{A}})| \geq a\right)$ valószínűségre.

8.17 Vegyünk egy valószínűségi eloszlást a A, C, T, G (adenin, citozin, timin és guanin) betűk halmazán, és képezzünk két n hosszú i.i.d. sorozatot (DNS-molekulát) ilyen eloszlású betükből. Jelöljük a sorozatokat $u_1, \ldots, u_n$-el és $v_1, \ldots, v_n$-el. Azt mondjuk, hogy a két sorozatban van m hosszú közös részsorozat, ha van olyan $i_1 \leq \cdots \leq i_m$ és $j_1 \leq \cdots \leq j_m$, hogy $u_{i_1} = v_{j_1}, \ldots, u_{i_m} = v_{j_m}$. Jelölje M_n a leghosszabb közös részsorozat hosszát.

(a) Lássa be, hogy $\mathbf{E}(M_n)$ sorozatra teljesül a $\mathbf{E}(M_{n+m}) \geq \mathbf{E}(M_n) + \mathbf{E}(M_m)$ egyenlőtlenség, és ebből vezesse le, hogy $\mathbf{E}(\frac{M_n}{n})$ konvergál $\sup_n \mathbf{E}(\frac{M_n}{n})$-hez.

(b) Bizonyítsa be az alábbi becslést:

$$\mathbf{P}(|M_n - \mathbf{E}(M_n)| \geq \lambda\sqrt{n}) \leq 2e^{-\frac{\lambda^2}{8}}$$

8.18 A 7.13 feladat feltételei mellett (Pólya urna, kezdetben egy piros és egy kék golyó) Mutassuk meg, hogy annak a valószínűsége, hogy a piros golyók aránya valaha is eléri a 3/4-et, kisebb mint 2/3.