הסתברות וסטטיסטיקה מתמטית - תרגול 9

זמני עצירה

הגדרה:

משתנה מקרי $T:\Omega\to\mathbb{N}\cup\{0,\infty\}$ נקרא <u>זמן עצירה</u> ביחס ל-$\{\mathcal{F}_n\}$ אם לכל $n\in\mathbb{N}_0$, $\{T\le n\}\in\mathcal{F}_n$.

דוגמאות:

אם X_n מרקובי, אז $T=\min\{n\in\mathbb{N}\mid X_n=s\}$ זמן עצירה.

לעומת זאת, $T=\min\{n\in\mathbb{N}\mid X_n=X_{n+1}\}$ לא זמן עצירה.

טענה:

יהיו T_1,T_2 זמני עצירה. אז $T_1\wedge T_2=\min\{T_1,T_2\}$, $T_1\vee T_2=\max\{T_1,T_2\}$, T_1+T_2 זמני עצירה.

הגדרה:

יהי T זמן עצירה ביחס לפילטרציה $\{\mathcal{F}_n\}$. <u>ה-σ-אלגברה העצורה ע"י T</u> (σ-algebra of T-past) היא

$$\mathcal{F}_T=\{A\in\mathcal{F}\mid \forall n\ge 0: A\cap\{T\le n\}\in\mathcal{F}_n\}.$$

איך לחשוב על זה?

$A\in\mathcal{F}_n$ ← כדי לדעת אם $\omega\in A$ או לא מספיק להסתכל עד זמן n.

$A\in\mathcal{F}_T$ ← $A\in\mathcal{F}_T$ אם לכל $\omega\in\Omega$, אם יש לנו את כל המידע עד זמן n וגם $T(\omega)\le n$, אז אפשר לדעת אם $\omega\in A$ או $\omega\notin A$.

טענה:

יהיו T_1,T_2 זמני עצירה.

א. $\mathcal{F}_{T_1}$ היא σ-אלגברה.

ב. T_1 היא $\mathcal{F}_{T_1}$ - מדידה.

ג. אם $T_1 \overset{a.s.}{\leq} T_2$ אז $\mathcal{F}_{T_1} \subseteq \mathcal{F}_{T_2}$.

<u>תרגיל:</u>

אם T_1, T_2 זמני עצירה ביחס ל- $\{\mathcal{F}_n\}$, אז $\mathcal{F}_{T_1} \cap \mathcal{F}_{T_2} = \mathcal{F}_{T_1 \wedge T_2}$.

<u>הוכחה:</u>

$\boxed{\subseteq}$ נניח $A \in \mathcal{F}_{T_1} \cap \mathcal{F}_{T_2}$. לכן לכל $n \geq 0$,

$$A \cap \{T_1 \leq n\} \in \mathcal{F}_n$$
$$A \cap \{T_2 \leq n\} \in \mathcal{F}_n$$

$$A \cap \{T_1 \wedge T_2 \leq n\} = A \cap (\{T_1 \leq n\} \cup \{T_2 \leq n\}) =$$
$$= \underbrace{(A \cap \{T_1 \leq n\})}_{\in \mathcal{F}_n} \cup \underbrace{(A \cap \{T_2 \leq n\})}_{\in \mathcal{F}_n} \in \mathcal{F}_n$$

לכן $A \in \mathcal{F}_{T_1 \wedge T_2}$.

$\boxed{\supseteq}$ נשים לב כי $T_1 \wedge T_2 \leq T_1, T_2$ $\Leftarrow$ $\mathcal{F}_{T_1 \wedge T_2} \subseteq \mathcal{F}_{T_1}, \mathcal{F}_{T_2}$ $\Leftarrow$

$\mathcal{F}_{T_1 \wedge T_2} \subseteq \mathcal{F}_{T_1} \cap \mathcal{F}_{T_2}$.

$\square$

<u>משפט העצירה של דוב:</u> (Optional stopping theorem)
sampling

יהי X_n מרטינגל ביחס ל- $\{\mathcal{F}_n\}$, ויהי T זמן עצירה. אם לפחות אחד מהתנאים הבאים מתקיים, אז $X_T(\omega) = X_{T(\omega)}(\omega)$ מוגדר ומתקיים $\mathbb{E}[X_T] = \mathbb{E}[X_0]$:

א. $T \overset{a.s.}{\leq} C < \infty$.

ב. $T \overset{a.s.}{<} \infty$ וגם קיים $C \in \mathbb{R}$ כך ש- $|M_n| \overset{a.s.}{\leq} C$ לכל $n \leq T$.

ג. $\mathbb{E}[T] < \infty$ וגם קיים $C \in \mathbb{R}$ כך ש- $|M_n - M_{n-1}| \overset{a.s.}{\leq} C$ לכל $n \leq T$.

תרגיל:

יהי S_n הילוך מקרי מוטה על $\mathbb{Z}$, כלומר $S_0=0$, $S_n=\sum_{i=1}^{n} X_i$, $X_i=\begin{cases}1, & p\\ -1, & 1-p\end{cases}$,

כך $p\neq\frac{1}{2}$. נגדיר $a\in\mathbb{Z}$

$T_a=\min\{n\geq 0 \mid S_n=a\}$.

עבור $a,b>0$, מהי $P(T_{-a}<T_b)$?

סתרון:

המרטינגל המקובל לפשוט $\leftarrow S_n-(2p-1)n \leftarrow$ לא יעבוד.

ניעזר במרטינגל האקספוננציאלי שראינו בתרגול הקודם

$M_n=\left(\frac{1-p}{p}\right)^{S_n}$.

נגדיר $T=T_{-a}\wedge T_b$. זה זמן עצירה סופי כמעט תמיד, כי הוא חסום על ידי משתנה גיאומטרי. בנוסף M_n חסום לכל $n\leq T$, כי אם $n\leq T$ אז $-a\leq S_n\leq b$ ולכן M_n חסום על ידי ביטוי שתלוי רק ב- a,b,p. לכן, לפי ה- OST,

$$\mathbb{E}[M_T]=\mathbb{E}[M_0]=1$$

$$\mathbb{E}[M_T]=\left(\frac{1-p}{p}\right)^{-a}\cdot P(T_{-a}<T_b)+\left(\frac{1-p}{p}\right)^{b}\cdot P(T_b<T_{-a})$$

אפשר להעביר אגפים ולקבל

$$P(T_{-a}<T_b)=\frac{1-\left(\frac{1-p}{p}\right)^b}{\left(\frac{1-p}{p}\right)^{-a}-\left(\frac{1-p}{p}\right)^b} \qquad \square$$

מסקנה:

אם $p>\frac{1}{2}$ אז $P(T_{-a}=\infty)>0$. $(a>0)$

תרגיל:

במדינה עמומה יש בחירות עם שני מועמדים, א' ו-ב'. מועמד א' קיבל a קולות, ומועמד ב' קיבל b קולות, $b<a$. הקולות נספרו בסדר אקראי. מה הסיכוי שלאורך הספירה מועמד א' הוביל?

פתרון:

נגדיר $n=a+b$, $a_k =$ כמות הקולות ש-א' קיבל עד הקול ה-k (כולל),

$b_k =$ ש-ב' קיבל

$S_k=a_k-b_k =$ כמות ההובלה של א' על ב', $0\le k\le n$.

נייצר מרטינגל מ-S_k הפוך: $S_n, S_{n-1}, \dots, S_0$

$$\mathbb{E}[S_{n-k} \mid S_n,\dots,S_{n-(k-1)}] = ?$$

$$S_{n-k} = \begin{cases} S_{n-k+1}+1, & \text{הקול ה-}(n-k+1)\text{ הוא ל-ב'} \\ S_{n-k+1}-1, & \text{הקול ה-}(n-k+1)\text{ הוא ל-א'} \end{cases} =$$

$$= \begin{cases} S_{n-k+1}+1, & \text{הסתברות } \frac{n-k+1-a_{n-k+1}}{n-k+1} \\ S_{n-k+1}-1, & \text{הסתברות } \frac{a_{n-k+1}}{n-k+1} \end{cases}$$

$$\mathbb{E}[S_{n-k} \mid S_n,\dots,S_{n-(k-1)}] = (S_{n-k+1}+1)\cdot\frac{n-k+1-a_{n-k+1}}{n-k+1} +$$

$$+ (S_{n-k+1}-1)\cdot\frac{a_{n-k+1}}{n-k+1} =$$

$= -S_{n-k+1}$

$$= S_{n-k+1} + \frac{n-k+1-2a_{n-k+1}}{n-k+1} =$$

$$= S_{n-k+1}\cdot\frac{n-k}{n-k+1}$$

במילים אחרות, $M_k=\frac{S_{n-k}}{n-k}$ מרטינגל.

נגדיר זמן עצירה

$$T = \begin{cases} \min\{k \mid M_k = 0\} & \text{אם קיים } k \text{ כזה,} \\ n-1 & \text{אחרת.} \end{cases}$$

$T \le n-1$ זמן עצירה חסום, לכן אפשר להשתמש ב- OST:

$$\mathbb{E}[M_T] = \mathbb{E}[M_0] = \mathbb{E}\left[\frac{S_n}{n}\right] = \frac{a-b}{a+b}$$

$$\parallel$$

$$P\left(\text{א' מוביל לכל אורך הדרך}\right)\cdot \overset{1}{\overset{\parallel}{M_{n-1}}} + P\left(\text{ב' מוביל מתישהו או היה תיקו}\right)\cdot 0 = P\left(\text{א' מוביל לכל אורך הדרך}\right)$$

כשא' מוביל לכל אורך הדרך $M_{n-1} = S_1 = 1$

מרטינגלים וריכוז מידה

טענה: (אי-שוויון אזומה - הופדינג)

יהי X_n מרטינגל עם $|X_n - X_{n-1}| \le C$, אז

$$P(|X_n - X_0| \ge \lambda) \le 2e^{-\frac{\lambda^2}{2\sum_{k=1}^{n} c_k^2}}$$

טענה: (Doob's Maximal inequality)

יהי X_n תת-מרטינגל. אז לכל $\lambda>0$,

$$P\left(\max_{0\le k\le n} X_k \ge \lambda\right) \le \frac{\mathbb{E}[\max\{X_n, 0\}]}{\lambda}$$

דוגמה:

מחלקים m כדורים ל-n תאים. בלי חשיבות לסדר, עם החזרה.

$Z_{n,m}$ = כמה תאים ריקים יש? $Z_{n,m} = \sum_{i=1}^{n} B_i$, B_i = האם התא ה-i ריק

$$\mathbb{E}[Z_{n,m}] = \sum_{i=1}^{n} \mathbb{E}[B_i] = n\cdot\left(1-\frac{1}{n}\right)^m$$

כמה $Z_{n,m}$ מרוכז סביב התוחלת?

נגדיר לכל $1 \le j \le m$, X_j = האינדקס של התא שבו שמים את כדור j.

$f(X_1, \dots, X_m)$ = כמות התאים הריקים בחלוקה הזו.

נגדיר $M_k = \mathbb{E}[f(X_1, \dots, X_m) | X_1, \dots, X_k]$.

$M_m = Z_{n,m}$, $M_0 = \mathbb{E}[f(X_1, \dots, X_m)] = \mathbb{E}[Z_{n,m}]$.

נשים לב כי $|M_k - M_{k-1}| \le 1$, כי כשמוסיפים כדור אנחנו משנים את כמות התאים הריקים לכל היותר ב-1.

M_k מרטינגל, לכן לפי אי-שוויון הופדינג

$$P(|Z_{n,m} - \mathbb{E}[Z_{n,m}]| \ge \lambda) \le 2e^{-\frac{\lambda^2}{2m}}$$

אם ניקח $\lambda = b\sqrt{m}$, אז

$$P(|Z_{n,m} - \mathbb{E}[Z_{n,m}]| \ge b\sqrt{m}) \le 2e^{-\frac{b^2}{2}}.$$