הסתברות וסטטיסטיקה מתמטית - תרגול 2

משתנים מקריים

תזכורת:

$f:(X,\mathcal{F}_X)\to(Y,\mathcal{F}_Y)$ היא מדידה אם לכל $A\in\mathcal{F}_Y$, $f^{-1}(A)\in\mathcal{F}_X$.

הגדרה:

יהי $(\Omega,\mathcal{F})$ מרחב מדיד. משתנה מקרי על Ω הוא פונקציה מדידה $X:(\Omega,\mathcal{F})\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$.

למה דווקא σ-אלגברת בורל?

היינו רוצים למדוד $\{\omega: X(\omega)\le c\}$ $\Leftarrow$ $\mathcal{B}(\mathbb{R})\subseteq\mathcal{F}_{\mathbb{R}}$.

תרגיל:

קיימת פונקציה רציפה $f:\mathbb{R}\to\mathbb{R}$ שאינה מדידה לבורל.

הוכחה:

תהי $g:[0,1]\to[0,1]$ פונקציית קנטור.

$$g(x)=\begin{cases}\sum_{n=1}^{\infty}\frac{a_n}{2^n}, & x=\sum_{n=1}^{\infty}\frac{2a_n}{3^n}\in C,\quad a_n\in\{0,1\}\\ \sup\{f(c)\mid c<x, c\in C\} & \text{אחרת}\end{cases}$$

נגדיר $h:[0,1]\to[0,2]$ לפי $h(x)=x+g(x)$. מסתבר ש-h היא הומיאומורפיזם בין הקטעים. ניקח $f=h^{-1}$.

מתקיים $\mu(h(C))=1$ $\Leftarrow$ קיימת $B\subseteq h(C)$ לא מדידה. אבל $h^{-1}(B)\subseteq C$

לכן היא מדידה לבורל. $h^{-1}(B)$ מדידה, אבל $f^{-1}(h^{-1}(B))=B$ לא מדידה. $\square$

הגדרה:

יהי X משתנה מקרי. ההתפלגות (או החוק) של X היא $\mathcal{L}_X = P\circ X^{-1}$

$$\mathcal{L}_X(A) = P(X(\omega)\in A)$$

פונקציית ההתפלגות המצטברת של X היא $F_X(x)=P(X\le x)$.

תזכורת:

א. F_X רציפה מימין.

ב. F_X מונוטונית לא יורדת.

ג. $\lim_{x\to\infty} F_X(x)=1$, $\lim_{x\to-\infty}F_X(x)=0$.

תרגיל: (inverse cdf)

לכל פונקציה $F:\mathbb{R}\to\mathbb{R}$ שמקיימת את התנאים האלו יש משתנה מקרי

$$X:([0,1],\mathcal{B}([0,1]))\longrightarrow\mathbb{R}$$

כך ש- $F_X=F$.

הוכחה:

רעיון - אם X מ"מ רציף, $F_X(X)\sim U[0,1]$. אם למשל F_X הפיכה,

$$P(F_X(X)\le a) = P(X\le F_X^{-1}(a)) = F_X(F_X^{-1}(a)) = a$$

במקרה הכללי: יהי $U\sim U[0,1)$ מ"מ אחיד על $[0,1]$.

$$F^{-1}(a)=\inf\{t\in\mathbb{R} \mid F(t)\ge a\}$$

- F^{-1} מונוטונית לא יורדת.
- אם $t=F^{-1}(a)$, אז $F(F^{-1}(a))=F(t)\ge a$ (רציפות מימין)
- מההגדרה, $F^{-1}(F(t))\le t$.

נגדיר $X=F^{-1}(U)$.

$$* \quad \{U<F(a)\}\subseteq\{X=F^{-1}(U)\le a\}\subseteq\{U\le F(F^{-1}(U))\le F(a)\}$$

$$P(U<F(a))\le P(X\le a)\le P(U\le F(a))$$

אבל U מ"מ רציף, ולכן $P(X\le a)=P(U\le F(a))=F(a)$. $\square$

תלות ואי-תלות

הגדרה:

מאורעות $A_1,\dots,A_n\in\mathcal{F}$ בלתי-תלויים אם לכל $I\subseteq\{1,\dots,n\}$

$$P\left(\bigcap_{i\in I}A_i\right)=\prod_{i\in I}P(A_i)$$

הגדרה:

תת-σ-אלגברות $\mathcal{F}_1,\dots,\mathcal{F}_n$, $\mathcal{F}_i\subseteq\mathcal{F}$ הן בלתי-תלויות אם לכל $A_i\in\mathcal{F}_i$ המאורעות $A_1,\dots,A_n$ בלתי-תלויים.

טענה:

מספיק לבדוק את זה על קבוצה סגורה לחיתוכים שיוצרת את ה-σ-אלגברה.

בורל-קנטלי

משפט: (למת בורל קנטלי)

תהי $\{A_n\}$ סדרת מאורעות.

א. אם $\sum_{n=1}^{\infty}P(A_n)<\infty$, אז $P(\limsup A_n)=0$.

כלומר, בסיכוי 1, רק מספר סופי של A_n-ים מתרחש ביחד.

ב. אם $\{A_n\}$ בלתי-תלויים וגם $\sum_{n=1}^{\infty}P(A_n)=\infty$ אז $P(\limsup A_n)=1$.

במילים, בסיכוי 1, אינסוף מה-A_n-ים קורים במקביל.

דוגמה:

בסדרה של אינסוף הטלות מטבע, $A_n=\{\text{בהטלה ה-}n\text{ יצא } \mathcal{H}\}$, $P(A_n)=\frac{1}{2}$,

A_n-ים ב"ת, $\sum P(A_n)=\infty \Leftarrow$ נקבל אינסוף פעמים $\mathcal{H}$.

תרגיל:

יהיו $X_1, X_2, \dots$ ממתשב"ה עם $X_i \sim N(0,1)$. נגדיר את המאורע

$$A_n = \{\max\{X_{n^2}, \dots, X_{n^2+2n}\} > 5\}.$$

הראו כי $P(\{A_n \text{ a.s.}\}) = 1$.

הוכחה:

$$P(A_n^c) = P(\max\{X_{n^2}, \dots, X_{n^2+2n}\} \le 5) = \prod_{i=n^2}^{n^2+2n} \underbrace{P(X_i \le 5)}_{\substack{\text{מספר קבוע } p \\ 0<p<1}} = p^{2n+1}$$

לכן $\sum P(A_n^c) = \sum p^{2n+1} < \infty \Leftarrow$ A_n^c קורים מספר סופי של פעמים $\Leftarrow P(A_n \text{ a.s.}) = 1$. $\square$

תרגיל:

יהי $X \sim U[0,1]$. הראו כי בפיתוח העשרוני של X, כל רצף ספרות יופיע אינסוף פעמים כמעט תמיד.

הוכחה:

הספרה ה-n-ית אחרי הנקודה היא $X_n = \lfloor X \cdot 10^n \rfloor \pmod{10}$. אלה ממתשב"ה, $X_i \sim U[0,9]$ (בדיד)

יהיו $d_1, \dots, d_k$ סדרת ספרות. נגדיר

$$A_n = \{(X_{kn+1}, \dots, X_{kn+k}) = (d_1, \dots, d_k)\}$$

$$P(A_n) = \frac{1}{10^k}$$

A_n בת"ל, $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{10^k} = \infty \overset{\text{בורל קנטלי}}{\Leftarrow} A_n \text{ i.o.} \Leftarrow$ בסיכוי 1 הרצף $d_1, \dots, d_k$ יופיע אינסוף פעמים. $\square$

$\mathcal{F}_n = \sigma(X_n, X_{n+1}, \dots)$

חוק ה-0-1 של קולמוגורוב

משפט:

יהיו $Y, X_1, X_2, \dots$ מ"מ ב"ת כך ש-Y הוא $\sigma(\{X_n\})$-מדיד.
אז Y הוא קבוע כ"ת.

למאורעות – כל מאורע שאינו תלוי במספר סופי של ה-X_n-ים הוא בהסתברות 0 או 1.

תרגיל:

יהיו $X_1, X_2, \dots$ מ"מ ב"ת, ויהי $S_n = \sum_{i=1}^{n} X_i$. תהי $\{b_n\}$ סדרת מספרים
כך ש-$b_n \to \infty$. הראו כי $P\left(\limsup \frac{S_n}{b_n} = c\right) = 1$ לאיזשהו קבוע
$c \in \mathbb{R} \cup \{\pm\infty\}$.

הוכחה:

$A_c = \left\{\limsup \frac{S_n}{b_n} \le c\right\}$ מאורע זנב ביחס ל-$X_1, X_2, \dots$. לכל k קבוע,

$$\frac{S_{n+k}}{b_{n+k}} = \underbrace{\frac{X_1 + \dots + X_k}{b_{n+k}}}_{\xrightarrow{n\to\infty} 0} + \frac{X_{k+1} + \dots + X_{n+k}}{b_{n+k}} = \frac{X_1 + \dots + X_k}{b_{n+k}} + \frac{S_{n+k} - S_k}{b_{n+k}}$$

$\Leftarrow \limsup_{n\to\infty} \frac{S_{n+k}}{b_{n+k}} = \underbrace{\limsup_{n\to\infty} \frac{S_{n+k}-S_k}{b_{n+k}}}$. זה מראה של A_c מאורע זנב, לכן
לפי קולמוגורוב $P(A_c) \in \{0,1\}$.

ניקח $F(c) = P(A_c)$, $c_0 = \sup\{c \in \mathbb{R} \cup \{\pm\infty\} \mid F(c) = 0\}$. ממנו נקבל

$$P\left(\limsup \frac{S_n}{b_n} \le c_0\right) = \lim_{c \to c_0^+} P\left(\limsup \frac{S_n}{b_n} \le c\right) \underset{\text{רציפות}}{=} 1$$

$\Leftarrow P\left(\limsup \frac{S_n}{b_n} = c_0\right) = 1$.

$(F(\infty) = 1)$

□