

הסתברות למתמטיקאים – הרצאה 21

30.5.11

נחשוב על הילוך מקרי חד ממדי רגיל. $S_n = X_1 + \dots$. $B = \{S_n \xrightarrow{n \rightarrow \infty} +\infty\}$. זה מאורע מהסתברות 0 כי אנחנו יודעים שהילוך מקרי לא חסום לא מלמטה ולא מלמעלה. נכתוב את התנאי B באופן שקול: $X_{11} + X_{12} + \dots \rightarrow +\infty$. זה לא תלוי ב- $X_1, \dots, X_{10}$ לכן

$$\mathbb{P}((X_1, \dots, X_{10}) \in A|B) = \mathbb{P}((X_1, \dots, X_{10}) \in A)$$

לכל A . זה נכון גם לכל מספר סופי של X ים. וזה אומר שהתפלגות משותפת של כל הסדרה האינסופית בהינן B זהה להתפלגות הלא מותנית. בפרט, $\mathbb{P}(B|B) = \mathbb{P}(B) = 0$ זה בעיה. וזה אומר שאין התנאיה מהסוג הזה וכנראה שגם לעולם לא תהיה. דוגמא נוספת מהסוג הזה היא $U \sim U(0, 1)$ ונתבונן ב- $X = e^{2\pi i u}$ מתפלג באופן אחיד במעגך יחידה במישור. עכשיו ניקח קבוצה: $B = \{e^{2\pi i u} : u \in \mathbb{Q}\}$ זאת קבוצה צפופה של מספרים על המעגל. קל לראות שהקבוצה אינוורטנטית לסיבובים אבל התנאי $X|_{X \in B}$ אינווריאנטי רק לסיבובים הרציונליים. וזה בעיה כי הרציונליים הם בני מניה ואי אפשר להגדיר על קבוצה כזאת הסתברות. נניח יש לנו צפיפות משותפת: $f_{X,Y}(x,y)$. אז הצפיפות המותנית היא $f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$. אז:

$$\mathbb{P}(Y = y|X = x) = \frac{\mathbb{P}(X = x, Y = 0)}{\mathbb{P}(X = x)}$$

נניח ש- X רציף ו- Y בדיד. עכשיו לא ברור איך עושים התפלגות משותפת. הבעיה שאין תיאוריה אחת שמכסה את כל המקרים שבהם אפשר לעשות התנאיה. החדשות הטובות הן שאם $X, Y : \Omega \rightarrow \mathbb{R}^m$ מ"מ אז יש תיאוריה שנותנת $Y|_{X=x}$. נסגור קודם מקרה בדיד. $(\Omega, \mathcal{F}, P)$ סופית או בת מניה, $\mathcal{F} = 2^{-\Omega}$ ו- P מידת הסתברות כלשהי. תת-שדה σ $\mathcal{F}_1 \subseteq \mathcal{F}$ אז תמיד קיים X כך ש- $\mathcal{F}_1 = \sigma(X)$. אם $A_i \subseteq \Omega$ לכל $i \in I$. Ω סופי לכן I_0 ואם Ω בת מניה אז I_0 בת מניה כאשר

$$\bigcup_{i \in I} A_i = \bigcup_{i \in I_0} A_i$$

כעת לכל ω נגדיר $\bigcap_{A \ni \omega, A \in \mathcal{F}_1} A \in \mathcal{F}_1$ וע"י חלוקה כזאת אפשר ליצור את השדה σ . עכשיו נכתוב נוסחאות להסתברות מותנית.

$$\mathbb{P}(A|X = x) = \frac{\mathbb{P}(A, X = x)}{\mathbb{P}(X = x)} = \frac{\sum_{\omega \in A: X(\omega)=x} \mathbb{P}(\omega)}{\sum_{\omega: X(\omega)=x} \mathbb{P}(\omega)} = P_x(A)$$

כאשר

$$P_x(\omega) = \begin{cases} \frac{\mathbb{P}(\omega)}{\mathbb{P}(X=x)} & X(\omega) = x \\ 0 & \text{אחרת} \end{cases}$$

עבור A נתון $P_x(A) = f(x)$ פונקציה ב- x בלבד. נוכל לכתוב:

$$\mathbb{P}(A|X) = f(X)$$

$$\mathbb{E}Y = \int_{\Omega} Y(\omega) P(d\omega)$$

$$\mathbb{E}(Y|X = x) = \int_{\Omega} Y(\omega) P_x(d\omega) = \int_{X^{-1}(x)} Y(\omega) P_x(d\omega)$$

גם כאן $\int_{X^{-1}(x)} Y(\omega) P_x(d\omega) = g(x)$ ואז $\mathbb{E}(Y|X) = g(X)$. יתרון גדול של הגדרה כזאת היא ששום דבר לא תלוי יותר בערכים של x .

$$\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$$

$$P = \sum_x (\mathbb{P}(X = x) P_x)$$

במיוחד אם $Y = \mathbb{1}_A$ מקבלים $\mathbb{E}(\mathbb{P}(A|X)) = \mathbb{P}(A)$ אפשר לגשת תוחלת גם ככה:

$$\min_{a \in \mathbb{R}} \mathbb{E}(Y - a)^2$$

המינימום יתקבל $a = \mathbb{E}Y = \int_{\Omega} Y dP$ ואז $\min_{a \in \mathbb{R}} \mathbb{E}(Y - a)^2 = \text{Var}(Y)$ עכשיו,

$$\min_g \mathbb{E}(Y - g(x))^2$$

$g(x) = \mathbb{E}(Y|X = x) = \int Y dP_x$.
 וזה רעיון טוב שיעזור לנו לעבור לתוחלת מותנית רציפה: התוחלת המותנית היא הטלה אורתוגונלית. $L_2(\Omega, \mathcal{F}, P) = L_2(\mathcal{F})$ ויש $L_2(\mathcal{F}_1)$ (פונקציות קבועות על כל חלק מהחלוקה) ו- $L_2(\mathcal{F}_1) \subseteq L_2(\mathcal{F})$. אז $\mathbb{E}(\cdot|\mathcal{F}_1)$ היא הטלה ללדת המחרב.
 אז נדבר על מרחב הסתברות כללי $(\Omega, \mathcal{F}, P)$, $\mathcal{F}_1 \subseteq \mathcal{F}$ תת שדה σ : $L_2(\mathcal{F}_1) \subseteq L_2(\mathcal{F})$. אז יש הטלה שנסמן $Y \mapsto (Y|F_1)$.
 $L_2(\mathcal{F})$ הוא למעשה המישור האוקלידי ו- $L_2(\mathcal{F}_1)$ הוא ישר. צריך לראות מה זה הטלה אורתוגונלית.

$$\langle Y - \mathbb{E}(Y|\mathcal{F}_1), z \rangle = 0 \quad \forall z \in L_2(\mathcal{F}_1)$$

$$\langle Y, z \rangle = \langle \mathbb{E}(Y|\mathcal{F}_1), z \rangle$$

$$\mathbb{E}(Yz) = \mathbb{E}(\mathbb{E}(Y|\mathcal{F}_1) \cdot z)$$

במיוחד עבור $z = 1$:

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|\mathcal{F}_1))$$

אפשר לקחת גם $z = \mathbb{1}_A$, כאשר $A \in \mathcal{F}_1$. אז:

$$\mathbb{E}(Y; A) = \mathbb{E}(\mathbb{E}(Y|\mathcal{F}_1); A)$$

הטלה היא אופרטור לינארי לכן

$$\mathbb{E}(aY|\mathcal{F}_1) = a\mathbb{E}(Y|\mathcal{F}_1)$$

וגם

$$\mathbb{E}(X + Y|\mathcal{F}_1) = \mathbb{E}(X|\mathcal{F}_1) + \mathbb{E}(Y|\mathcal{F}_1)$$

בפרט רציף:

$$X_n \xrightarrow{L_2} X \Rightarrow \mathbb{E}(X_n|\mathcal{F}_1) \xrightarrow{L_2} \mathbb{E}(X|\mathcal{F}_1)$$

לכן,

$$\|\mathbb{E}(X|\mathcal{F}_1)\|_{L_2} \leq \|X\|_{L_2}$$

יש גם תכונה מיוחדת: עבור $L_2(\mathcal{F}_1)$:

$$x \in L_2(\mathcal{F}_1) \Rightarrow X^+ \in L_2(\mathcal{F}_1)$$

לכן אם $x \geq 0$ כ"ב אז $\mathbb{E}(x|\mathcal{F}_1) \geq 0$ כ"ב. כי $\|Y - X\| \geq \|Y^+ - X\|$ לכל $Y \in L_2(\mathcal{F}_1)$.
כי הקירוב האופטימלי הוא חיובי. $(L_2(\mathcal{F}))$ הוא צפוף ב- $L_1(\mathcal{F})$.
נראה כי

$$\|\mathbb{E}(X|\mathcal{F}_1)\|_{L_1} \leq \|X\|_{L_1}$$

נניח $X \in L_2(\mathcal{F})$. זה מתקבל מחיוביות בצירוף עם

$$\mathbb{E}x = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_1))$$

זה כי $X = X^+ - X^-$.

$$\mathbb{E}(X|\mathcal{F}_1) = \underbrace{\mathbb{E}(X^+|\mathcal{F}_1)}_{\geq 0} - \underbrace{\mathbb{E}(X^-|\mathcal{F}_1)}_{\geq 0}$$

אבל:

$$\|\mathbb{E}(X^+|\mathcal{F}_1)\|_{L_1} = \mathbb{E}(\mathbb{E}X^+|\mathcal{F}_1) = \mathbb{E}(\mathbb{E}X^+) = \mathbb{E}X^+$$

באותו אופן:

$$\|\mathbb{E}(X^-|\mathcal{F}_1)\|_{L_1} = \mathbb{E}X^-$$

לכן,

$$\|\mathbb{E}(X|\mathcal{F}_1)\|_{L_1} \leq \mathbb{E}(X^+) + \mathbb{E}(X^-) = \mathbb{E}|x| = \|X\|_{L_1}$$

אז לכל $x \in L_1(\mathcal{F})$ יש $\mathbb{E}(X|\mathcal{F}_1) \in L_1(\mathcal{F}_1)$.
תכונה נוספת:

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}$$

אז

$$\mathbb{E}(X|\mathcal{F}_1) = \mathbb{E}(\mathbb{E}(x|\mathcal{F}_2)|\mathcal{F}_1)$$

נחזור למרטינגלים. אמרנו שמרטינגל שחסום ב- L_2 מתכנס ב- L_2 כ"ב.
ראינו שתי הגדרות למרטינגלים:

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n$$

למקרה הבדיד, ובמקרה הרציף:

$$\mathbb{E}(M_{n+1}Y) = \mathbb{E}(M_nY)$$

לכל $Y \in L_\infty(\mathcal{F}_n)$. עכשיו אם $Y = \mathbb{1}_A, A \in \mathcal{F}_n$ זה בדיוק כמו המקרה הראשון ועכשיו שתי ההגדרות שקולות.

יהי (M_n) מרטינגל חסום חיובי ב- L_2 אז הוא מתכנס ב- L_2 : $M_n \xrightarrow{L_2} M_\infty$. הבעיה היא האם זה מתכנס כ"ב. כשהיה לנו מרטינגל חיובי השתמשנו באי שוויון דובינס. כעת $M_\infty = M_\infty^+ - M_\infty^-$ אז

$$\underbrace{\mathbb{E}(M_\infty|\mathcal{F}_n)}_{=M_n} = \underbrace{\mathbb{E}(M_\infty^+|\mathcal{F}_n)}_{\text{מרטינגל חיובי}} - \underbrace{\mathbb{E}(M_\infty^-|\mathcal{F}_n)}_{\text{מרטינגל}}$$

כלומר מרטינגל חסום ב- L_2 הוא הפרש של מרטינגלים חיוביים ב- L_2 ולכן יש התכנסות כ"ב.

עכשיו נניח שיש לנו פילטרציה $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$ כך ש- $\mathcal{F}_n \uparrow \mathcal{F}_\infty$. אמרנו שבמקרה כזה נובע ש- $\bigcup_n \mathcal{F}_n$ הוא צפוף ב- $\mathcal{F}_\infty$. כאשר $\text{dist}(A, B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A)$. למה חדשה אומרת ש- $\bigcup_n L_2(\mathcal{F}_n)$ צפוף ב- $L_2(\mathcal{F}_\infty)$ כלומר שהפונקציות צפופות בין הפונקציות. כלומר צריך להראות ש- $L_2(\mathcal{F}_\infty) = \overline{\bigcup_n L_2(\mathcal{F}_n)}$ יהי $A \in \mathcal{F}_\infty$ אז $\mathbb{1}_A \in \overline{\bigcup_n L_2(\mathcal{F}_n)}$ כי יש $A_n \in \mathcal{F}_n$ כך ש- $A_n \rightarrow A$ ולכן $\mathbb{1}_{A_n} \rightarrow \mathbb{1}_A$. וזה סוגר כי תת המרחב שלנו מכיל את האינדיקטורים לכן את הקומבינציות הלינאריות שלהם ואלה צפופים.