מבוא להסתברות ח'

חלק 3

אמיר יהודיוף
הפקולטה למתמטיקה, הטכניון

תזכורת

מרחב הסתברות

תכונות נוספות (למשל הכלה-הדחה)

התניה: כיצד מידע משנה תפישת עולם

נוסחת הסתברות שלמה: שימוש בחלוקה של מרחב מדגם

תלות וחוסר תלות בין מאורעות

משתנים מקריים בדידים ורציפים

פונקציית התפלגות

למשתנים מיקריים בדידים יש הסתברות ולרציפים יש צפיפות

ישנה פונקציה שמוגדרת לכל משתנה מיקרי

הגדרה: לכל משתנה מיקרי X נגדיר את פונקציית ההתפלגות שלו כך: לכל a ממשי מתקיים

$$F_X(a) = \mathbb{P}(X \leq a)$$

משתנים בדידים

תזכורת: פונקציית ההתפלגות של X היא $F_X(a) = \mathbb{P}(X \leq a)$

משתנים בדידים: אם X בדיד אז

$$F_X(a) = \sum_{b \leq a} \mathbb{P}_X(b) = \sum_{b \leq a} \mathbb{P}(X = b)$$

ובכיוון השני

$$\mathbb{P}(X = a) = \mathbb{P}(X \leq a) - \mathbb{P}(X < a) = F_X(a) - \lim_{b \to a^-} F_X(b)$$

משתנים רציפים

תזכורת: פונקציית ההתפלגות של X היא $F_X(a) = \mathbb{P}(X \leq a)$

משתנים רציפים: אם X רציף אז

$$F_X(a) = \int_{-\infty}^{a} f_X(y)dy$$

ובכיוון השני: בנקודות a בהן f_X רציפה מתקיים

$$f_X(a) = F_X{}'(a)$$

תכונות

תזכורת: פונקציית ההתפלגות של X היא $F_X(a) = \mathbb{P}(X \leq a)$

תכונות:

א. $\lim_{a\to\infty} F_X(a) = 1$

ב. $\lim_{a\to-\infty} F_X(a) = 0$

ג. מונוטוניות: אם $a \leq b$ אז $F_X(a) \leq F_X(b)$

ד. רציפות מימין: $\lim_{a\to b^+} F_X(a) = F_X(b)$

הערה: כל פונקציה המקיימת א-ד יכולה להיות התפלגות (נדבר בהמשך)

גרפים

בדידים: $\mathbb{P}_X$ נתמכת באופן בדיד ו F_X פונקציית מדרגות

רציפים: f_X מוגדרת על $\mathbb{R}$ ו F_X רציפה

בדידים: דוגמאות

אחיד על קבוצה סופית: בהנתן n נגדיר $X \sim U(\{1,2,\ldots,n\})$ כך

$$\mathbb{P}(X=k) = \begin{cases} \frac{1}{n} & k \in \{1,2,\ldots,n\} \\ 0 & k \notin \{1,2,\ldots,n\} \end{cases}$$

אז

$$F_X(a) = \begin{cases} 0 & a < 1 \\ \frac{k}{n} & 1 \le a \le n,\ a \in [k, k+1) \\ 1 & a > n \end{cases}$$

רציפים: דוגמאות

משתנה מיקרי מעריכי:

$$f_X(a) = \begin{cases} \lambda e^{-\lambda a} & a \geq 0 \\ 0 & a < 0 \end{cases}$$

מתקיים:

$$F_X(a) = \begin{cases} 1 - \mathbb{P}(X > a) = 1 - e^{-\lambda a} & a \geq 0 \\ 0 & a < 0 \end{cases}$$

פונקציית התפלגות של נורמלי

תזכורת: $X \sim N(0,1)$ אם $f_X(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}}$ לכל $y \in \mathbb{R}$

הגדרה:

$$\Phi(a) = \mathbb{P}(X \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-y^2/2} dy$$

פונקציה זו שימושית ביותר, אך אין לה ביטוי פשוט ונחפש את ערכיה בטבלה:
למשל

y	0	1	2	2.5	3
Φ	0.5	0.8413	0.9772	0.9938	0.9987

פונקציית התפלגות של נורמלי

הגדרה:

$$\Phi(a) = \mathbb{P}(X \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-y^2/2} dy$$

סימטריה:

$$\Phi(-a) = \mathbb{P}(X \leq -a) = \mathbb{P}(X \geq a) = 1 - \Phi(a)$$

בפרט: $\Phi(0) = 1/2$

נורמלי כללי: ראינו שאפשר להפוך נורמלי כללי לתקני. אם $Y \sim N(\mu, \sigma^2)$ אז

$$F_Y(a) = \mathbb{P}(Y \leq a) = \mathbb{P}\left(\frac{Y-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{a-\mu}{\sigma}\right)$$

לסיכום: עם טבלה של $\Phi(a)$ עבור $a > 0$ ניתן לחשב הסתברויות של כל נורמלי

סיכום ביניים

מרחב הסתברות

משתנים מיקריים: בדידים ורציפים

פונקציית התפלגות

משתנים מעורבים

דוגמא: רמזור מתחלף מירוק לאדום כל 30 שניות ולהיפך. מכונית מגיעה לרמזור בזמן אחיד בדקה. יהי X זמן ההמתנה של המכונית בשניות.

מהי ההתפלגות?

$$F_X(a) = \mathbb{P}(X \le a) = \begin{cases} 0 & a < 0 \\ 1 & a \ge 30 \\ \frac{1}{2} \cdot 1 + \frac{1}{2}\frac{a}{30} & 0 \le a \le 30 \end{cases}$$

במילים: בסיכוי חצי מגיעה כשירוק ואז מחכה אפס שניות ובסיכוי חצי באדום ואז סיכוי לחכות לכל היותר a שניות הוא $a/30$

ניתן לרשום: $F_X = \frac{1}{2}F_{green} + \frac{1}{2}F_{red}$

משתנים מעורבים

דוגמא: יהיו $Y \sim U[0,3]$ ו $Z \sim U(\{1,2\})$ בלתי תלויים

נגדיר את X להיות Y בסיכוי $2/3$ ולהיות Z אחרת:

$$\begin{aligned} F_X(a) &= \mathbb{P}(X \leq a) \\ &= \mathbb{P}(X \leq a | X = Y)\mathbb{P}(X = Y) + \mathbb{P}(X \leq a | X = Z)\mathbb{P}(X = Z) \\ &= \frac{2}{3}F_Y(a) + \frac{1}{3}F_Z(a) \end{aligned}$$

ערכים ש X יכול לקבל: בין אפס לשלוש

כמה ערכים לדוגמא:

a	0	1^-	1	2^-	2	3
F_X	0	$\frac{2}{3}\frac{1}{3} + \frac{1}{3}0 = \frac{2}{9}$	$\frac{7}{18}$	$\frac{11}{18}$	$\frac{7}{9}$	1

סכום של בדיד ורציף

יהי X משתנה מיקרי ונרצה לכתוב

$$F_X = \alpha F_Y + (1-\alpha) F_Z$$

כאשר $0 < \alpha < 1$, Y חלק בדיד של X ו Z חלק רציף של X

חלק בדיד: יהיו $a_1, a_2, \dots$ נקודות אי הרציפות של F_X

נסמן $b_i = F_X(a_i) - \lim_{t \to a_i^-} F_X(t)$ (גובה "הקפיצה" ה i ית)

נרצה ש

א. $\mathbb{P}(Y = a_i)$ יהיה ביחס ישר ל b_i , כלומר $\mathbb{P}(Y = a_i) = \frac{b_i}{\sum_{j=1}^{\infty} b_j}$

ב. $F_X - \alpha F_Y$ תהיה פונקציה רציפה

לכן

$$\alpha = \sum_{j=1}^{\infty} b_j \ , \ \mathbb{P}(Y = a_i) = \frac{b_i}{\alpha}$$

ולסיום חלק רציף:

$$F_Z = \frac{1}{1-\alpha} (F_X - \alpha F_Y)$$

קטיעת משתנה רציף

דוגמא: יהי $Z \sim N(0,1)$ ונגדיר

$$X = \begin{cases} -1 & Z < -1 \\ 1 & Z > 1 \\ X & -1 \leq Z \leq 1 \end{cases}$$

אז

$$F_X(a) = \begin{cases} 0 & a < -1 \\ 1 & a > 1 \\ \Phi(a) & -1 \leq a \leq 1 \end{cases}$$

נרצה לכתוב $F_X = \alpha F_Y + (1-\alpha) F_Z$
כאשר $0 < \alpha < 1$, Y משתנה בדיד ו Z משתנה רציף

כמו קודם $a_1 = -1, a_2 = 1$ ו $b_1 = b_2 = \Phi(-1)$ ולכן $\alpha = 2\Phi(-1)$

כלומר $Y \sim U(\{-1,1\})$ וגם

$$F_Z = \frac{1}{1-2\Phi(-1)}(F_X - 2\Phi(-1)F_Y)$$

סכום של בדיד ורציף

יהי X משתנה מיקרי ונרצה לכתוב

$$F_X = \alpha F_Y + (1-\alpha) F_Z$$

כאשר $0 < \alpha < 1$, Y חלק בדיד של X ו Z חלק רציף של X

סיכום פתרון: נגדיר את $\{a_i\}_{i=1}^{\infty}$ כנקודות אי הרציפות, את $\{b_i\}_{i=1}^{\infty}$ כגובה ה"קפיצות" ואת

$$\alpha = \sum_{j=1}^{\infty} b_j \ , \quad \mathbb{P}(Y = a_i) = \frac{b_i}{\alpha} \ , \quad F_Z = \frac{1}{1-\alpha}(F_X - \alpha F_Y)$$

הערה: רק הבטחנו ש F_Z רציפה אבל לא בהכרח ש Z רציף ואכן ישנם משתנים מיקריים X שעבורם Z שהגדרנו לא רציף (נקראים סינגולריים) אך לא נרחיב בנושא

סיכום ביניים

מרחב הסתברות

משתנים מיקריים: פונקציית הסתברות וצפיפות

פונקציית התפלגות ופירוק משתנים מיקריים (כלליים)

פונקציה של משתנה מקרי

משתנה מיקרי הוא פונקציה $X : \Omega \to \mathbb{R}$

בהנתן פונקציה "חלקה" $h : \mathbb{R} \to \mathbb{R}$ הפונקציה $h(X) : \Omega \to \mathbb{R}$ המוגדרת על ידי $(h(X))(\omega) = f(X(\omega))$ גם תהיה משתנה מיקרי

דוגמא: גובה של אדם בס"מ X הוא מספר (נגיד) בין 0 ל 300 . כיצד נייצג מספר זה? $\lceil X \rceil$ (כאשר $h(x) = \lceil x \rceil$ הוא השלם הקטן ביותר הגדול מהממשי x)

הערה: יש לוודא ש h חלקה ושתחומי הגדרה מסתדרים (למשל עם לוגריתמים)

בדידות ורציפות

אם X בדיד, אז $h(X)$ בדיד גם כן

אם X רציף, אז $h(X)$ יכול להיות מעורב (למשל קטיעת משתנה מיקרי)

דוגמא

נסמן ב X ריבית בנקאית באחוזים, נניח $X \sim U([0,10])$
ונסמן ב Y שווי פיקדון עם ערך התחלתי של 1000 ש״ח לאחר 3 שנים
נחקור F_Y

ערך מינימלי של Y : מתקיים $Y \geq 1000$
ערך מקסימלי של Y : מתקיים $Y \leq 1000(1+0.1)^3 = 1331$
לכל $1331 \leq a \leq 1000$ מתקיים

$$\mathbb{P}(Y \leq a) = \mathbb{P}(1000(1+X/100)^3 \leq a) = \mathbb{P}\left(X \leq 100\left(\frac{a^{1/3}}{10} - 1\right)\right)$$

$$= \frac{100\left(\frac{a^{1/3}}{10} - 1\right)}{10} = a^{1/3} - 10$$

דוגמא

נסמן ב X ריבית בנקאית באחוזים, נניח $X \sim U([0,10])$
ונסמן ב Y שווי פיקדון עם ערך התחלתי של 1000 ש״ח לאחר 3 שנים
לכן

$$F_Y(a) = \begin{cases} 0 & a < 1000 \\ a^{1/3} - 10 & 1000 \leq a \leq 1331 \\ 1 & a > 1331 \end{cases}$$

המשתנה Y הוא רציף (ציור)

מהי צפיפותו?

דוגמא

נסמן ב X ריבית בנקאית באחוזים, נניח $X \sim U([0,10])$
ונסמן ב Y שווי פיקדון עם ערך התחלתי של 1000 ש״ח לאחר 3 שנים

תזכורת: $F_Y(a) = a^{1/3} - 10$ עבור $1000 \leq a \leq 1331$

מהי צפיפותו? הנגזרת של F_Y (בנקודות בהן f_Y רציפה)

לכל $1000 < a < 1331$ מתקיים $f_Y(a) = F_Y'(a) = \frac{a^{-2/3}}{3}$

(ציור: $f_Y(1000) = 1/300$ ו $f_Y(1331) = 1/363$)

באופן כללי

תזכורות:

א. אם h חח״ע בתחום הגדרתה, אז קיימת פונקציה הופכית h^{-1}

למשל עבור $h(x) = x^2$ בקטע $[0,1]$ מתקיים $h^{-1}(x) = \sqrt{x}$

ב. כלל שרשרת: בהנתן שתי פונקציות h, g כך שההרכבה $h \circ g$ מוגדרת, מתקיים

$(h \circ g)'(x) = h'(g(x)) \cdot g'(x)$

משפט: יהי X משתנה מקרי רציף, תהי h פונקציה חח״ע וגזירה, ויהי $Y = h(X)$,

אז לכל a בתמונה של h מתקיים

$$F_Y(a) = \mathbb{P}(Y \leq a) = \mathbb{P}(X \leq h^{-1}(a)) = F_X(h^{-1}(a))$$

ומתקיים

$$f_Y(a) = {F_Y}'(a) = {F_X}'(h^{-1}(a)) \cdot |(h^{-1})'|(a) = f_X(h^{-1}(a)) \cdot |(h^{-1})'|(a)$$

הערה: אם h אינה חח״ע, לפעמים ניתן לחלק את תחום הגדרתה כך שבחלוקה מתקיימת חד-חד-ערכיות

דוגמא

יהיו $X \sim U([-1,1])$, $h = x^2$, $Y = h(X) = x^2$

התפלגות: לכל $a \geq 0$ מתקיים

$$F_Y(a) = \mathbb{P}(-\sqrt{a} \leq X \leq \sqrt{a}) = \frac{2\sqrt{a}}{2} = \sqrt{a}$$

צפיפות: לכל $a \in [0,1]$ מתקיים

$$f_Y(a) = F_Y{}'(a) = \frac{a^{-1/2}}{2}$$

הערה: ניתן היה גם להשתמש במשפט עם חלוקת $[-1,1]$ בהתאם לתחומי חח״ע של h
כלומר ל $[0,1]$ ו $[-1,0)$

שימוש

סימולציות במחשב משתמשות לרוב במשתנים מיקריים. נניח שרוצים לדמות התפלגות של Y.

בעיה: מחשב מאפשר גישה רק למשתנים מיקריים מסוג מסוים, למשל $X \sim U([0,1])$ (וגם זה רק בקירוב). כיצד נדמה את Y באמצעות X?

משפט: יהי X משתנה מקרי ותהי F פונקציית צפיפות חח״ע נתונה. נגדיר $h = F^{-1} \circ F_X$ ו $Y = h(X)$. אז

$$F_Y = F$$

שימוש

משפט: יהי X משתנה מקרי ותהי F פונקציית צפיפות נתונה. נניח כי F_X, F חח״ע. נגדיר $h = F^{-1} \circ F_X$ ו $Y = h(X)$. אז $F_Y = F$

הוכחה:

$$F_Y(a) = \mathbb{P}(F^{-1}(F_X(X)) \leq a) = \mathbb{P}(X \leq F_X^{-1}(F(a))) = F_X(F_X^{-1}(F(a)) = F(a)$$

הערה: אם F_X, F אינן הפיכות עדיין אפשר לפתור (לא נרחיב)

דוגמא (מעריכי מאחיד): אם $X \sim U([0,1])$ ונרצה לקבל $Exp(\lambda)$. לכל $a \geq 0$ מתקיים

$$F(a) = 1 - e^{-\lambda a}\ , \quad F^{-1}(a) = -\frac{\ln(1-a)}{\lambda}$$

ולכן המשתנה $Y = -\frac{\ln(1-X)}{\lambda}$ יהיה $Exp(\lambda)$

סיכום ביניים

מרחב הסתברות

משתנים מיקריים: פונקציית הסתברות וצפיפות

פונקציית התפלגות ופירוק משתנים מיקריים (כלליים)

פונקציות של משתנים מיקריים