

הסתברות למתמטיקאים

© ארזים

19 באפריל 2017

1 תוחלת

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות.

הגדרה 1.1 יהי X משתנה מקרי ממשי. מגדירים, כאשר $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ או $X \geq 0$ את התוחלת:

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

סימון

$$\mathbb{E}(X; A) = \mathbb{E}(X \cdot \mathbb{1}_A)$$

משפט 1.2 (אי שוויון מרקוב) יהי $X \geq 0$ משתנה מקרי. אזי לכל $c > 0$ מתקיים:

$$\mathbb{P}(X \geq c) \leq \frac{\mathbb{E}X}{c}$$

הוכחה:

$$\mathbb{E}X \geq \mathbb{E}(X; \{X \geq c\}) = \mathbb{E}(X \cdot \mathbb{1}_{\{X \geq c\}}) \geq c\mathbb{P}(X \geq c)$$

■

וסיימנו.

דוגמאות יהי X משתנה מקרי ממשי כללי, ויהי $c > 0$.

$$\mathbb{P}(X \geq c) \leq \mathbb{P}(X^2 \geq c^2) \leq \frac{\mathbb{E}(X^2)}{c^2}$$

כעת יהי $\theta > 0$.

$$\mathbb{P}(X \geq c) = \mathbb{P}(e^{\theta X} \geq e^{\theta c}) \leq \frac{\mathbb{E}(e^{\theta X})}{e^{\theta c}}$$

תהי $(Z_k)_{k=1}^\infty$ סרת משתנים מקריים אי שליליים. נרצה לדעת האם $\sum Z_k$ מתכנס, כלומר האם הוא סופי? למשל, כמעט תמיד? אפשר לחשב את תוחלת הסכום:

$$\mathbb{E} \left(\sum_{k=1}^{\infty} Z_k \right) = \sum_{k=1}^{\infty} \mathbb{E}(Z_k)$$

נאמר ותוחלת זו סופית. נובע מיד

$$\mathbb{P} \left(\sum_{k=1}^{\infty} Z_k < \infty \right) = 1$$

מסקנה 1.3 (למת בורל קנטלי הראשונה, הוכחה חדשה) יהיו (A_k) מאורעות עם

$$\sum \mathbb{P}(A_k) < \infty$$

אזי

$$\mathbb{P}(\limsup A_k) = 0$$

■

הוכחה: ניקח בחישוב הקודם $Z_k = \mathbb{1}_{A_k}$.

שימוש נוסף: נניח כי $X \geq 0$ ומקבל ערכים שלמים. אזי

$$\mathbb{P}(Z > 0) = \mathbb{P}(X \geq 1) \leq \mathbb{E}(X)$$

משפט 1.4 (אי שוויון ינסן) תהי f פונקציה קמורה על קטע I בתוך $\mathbb{R}$, כלומר לכל $x, y \in I$ ולכל $\lambda \in (0, 1)$ מתקיים

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

כעת יהי X משתנה מקרי ממשי עם

$$\mathbb{E}|X| < \infty$$

$$\mathbb{P}(X \in I) = 1$$

$$\mathbb{E}|f(X)| < \infty$$

אזי

$$\mathbb{E}(f(x)) \geq f(\mathbb{E}X)$$

הוכחה: (רעיון) מקרה קצה - אם $\mathbb{E}X$ היא באחד מקצות הקטע, אזי X שווה בסיכוי 1 לאותו קצה, ואז הטענה ברורה. אחרת, $\mathbb{E}(X)$ בתוך הקטע ממש. מקמירות f , קיים לה משיק בנקודה $\mathbb{E}X$, והפונקציה תמיד מעליו. נסמן את המשיק:

$$g(x) = f(\mathbb{E}X) + a(x - \mathbb{E}X)$$

ואז נקבל

$$f(x) \geq f(\mathbb{E}X) + a(x - \mathbb{E}X)$$

לכל $x \in I$. אי השוויון נובע מהצבת X במקום x , ולקיחת תוכלת בשני האגפים. ■

מסקנה 1.5 (מונוטוניות נורמת L^p) יהי $p > 0$. עבור משתנה מקרי X ממשי, נאמר כי $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ (או $X \in L^p$) אם $\mathbb{E}(|X|^p) < \infty$. אז נסמן

$$\|X\|_p := (\mathbb{E}(|X|^p))^{\frac{1}{p}}$$

אם $0 < p \leq r < \infty$ וכן $X \in L^r$, אזי $X \in L^p$ וכן $\|X\|_p \leq \|X\|_r$.

הוכחה: (במילה) נגדיר את הפונקציה הקמורה:

$$f(x) = x^{\frac{r}{p}}$$

על הקטע $[0, \infty)$. כעת נשתמש באי שוויון ינסן על

$$|X|^p$$

לא ידוע שהתוחלת שלו סופית, ולכן נשתמש בסדרה, למשל $\min(|X|^p, n)$, כדי להשתמש באי שוויון ינסן. ■

1.1 תכונות של מרחבי L^p .

1. אי שוויון הולדר: יהיו $p, q > 1$ עם $\frac{1}{p} + \frac{1}{q} = 1$. יהי X, Y משתנים מקריים עם $X \in L^p, Y \in L^q$. אזי

$$\mathbb{E}(XY) \leq \|X\|_p \|Y\|_q$$

וכן $XY \in L^1$.

2. אי שוויון מינקובסקי: יהי $p \geq 1$ ויהיו $X, Y \in L^p$. אזי $X + Y \in L^p$ וכן

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$$

3. המרחב L^p הוא שלם, כלומר אם (X_n) סדרת משתנים מקריים מתוך L^p המקיימים

$$\sup_{m,n \geq r} \|X_m - X_n\|_p \xrightarrow{r \rightarrow 0} 0$$

אזי קיים $X \in L^p$ כך שמתקיים

$$\|X_n - X\|_p \xrightarrow{n \rightarrow \infty} 0$$

1.2 המרחב L^2

היו $X, Y \in L^2$. בפרט $X, Y \in L^1$, ולכן קיימות $\mathbb{E}X, \mathbb{E}Y$.

הגדרה 1.6 שונות:

$$\text{Var}(X) = \mathbb{E}\left((X - \mathbb{E}X)^2\right) = \mathbb{E}\left(X^2 - 2(\mathbb{E}X)X + (\mathbb{E}X)^2\right) = \mathbb{E}(X^2) - (\mathbb{E}X)^2$$

הגדרה 1.7 שונות משותפת:

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$$

הגדרה 1.8 נניח כי $\|X - \mathbb{E}X\|_2, \|Y - \mathbb{E}Y\|_2 \neq 0$. מקדם מתאם:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \in [-1, 1]$$

תכונות

1. משפט פיתגורס: אם $X, Y \in L^2$, $\text{Cov}(X, Y) = 0$ אזי $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. באופן כללי:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

הוכחה:

$$\begin{aligned} \text{Var}(X + Y) &= \mathbb{E}\left((X - \mathbb{E}X + Y - \mathbb{E}Y)^2\right) = \mathbb{E}\left((X - \mathbb{E}X)^2 + (Y - \mathbb{E}Y)^2 + 2(X - \mathbb{E}X)(Y - \mathbb{E}Y)\right) = \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \end{aligned}$$

■

מסקנה 1.9 אם $X_1, \dots, X_n \in L^2$ וכן לכל $i \neq j$ מתקיים $\text{Cov}(X_i, X_j) = 0$ אזי $\text{Var}(\sum X_k) = \sum \text{Var}(X_k)$.

הוכחה: באינדוקציה על k . נשתמש בדרך בכך שמתקיים

$$\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$$

■

2. כלל המקבילית: אם $X, Y \in L^2$ אזי

$$\|X\|_2^2 + \|Y\|_2^2 = \frac{1}{2} (\|X + Y\|_2^2 + \|X - Y\|_2^2)$$

הוכחה: נחשב:

$$\begin{aligned} \frac{1}{2} (\|X + Y\|_2^2 + \|X - Y\|_2^2) &= \frac{1}{2} (\mathbb{E}((X + Y)^2) + \mathbb{E}((X - Y)^2)) = \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) = \|X\|_2^2 + \|Y\|_2^2 \end{aligned}$$

■

1.3 הטלות אורתוגונליות

משפט 1.10 יהי V תת מרחב סגור של L^2 (כלומר V אוסף משתנים מקריים מתוך L^2 שהוא תת מרחב לינארי וכן אם $(X_n) \subseteq V$ סדרת קושי, אזי גם גבול הסדרה בתוך V). יהי $X \in L^2$. אזי קיים $Y \in V$ המקיים:

1.

$$\|X - Y\|_2 = \inf_{W \in V} \|X - W\|_2 =: \Delta$$

2. לכל $W \in V$ מתקיים

$$\mathbb{E}((X - Y)W) = 0$$

יתר על כן, אם Y מקיים אחת משתי התכונות, הוא מקיים גם את האחרת, והוא יחיד במובן שאם $\tilde{Y} \in V$ מקיים את אחת התכונות אזי $Y = \tilde{Y}$ כמעט תמיד.

הוכחה: מהגדרת אינפימום, קיימת סדרה $(W_n) \subseteq V$ המקיימת

$$\|X - W_n\|_2 \downarrow \Delta$$

נוכיח כי (W_n) קושי. נפעיל את כלל המקבילית על $X - W_n, X - W_m$

$$\begin{aligned} \|X - W_n\|_2^2 + \|X - W_m\|_2^2 &= \frac{1}{2} (\|W_n - W_m\|_2^2 + \|2X - W_n - W_m\|_2^2) \\ \|W_n - W_m\|_2^2 &= 2\|X - W_n\|_2^2 + 2\|X - W_m\|_2^2 - 4 \left\| X - \frac{1}{2}(W_n + W_m) \right\|_2^2 \end{aligned}$$

כעת, $\frac{1}{2}(W_n - W_m) \in V$ ולכן $\|X - \frac{1}{2}(W_n - W_m)\| \geq \Delta$ אם כן:

$$\begin{aligned} \|W_n - W_m\|_2^2 &= 2\|X - W_n\|_2^2 + 2\|X - W_m\|_2^2 - 4\left\|X - \frac{1}{2}(W_n + W_m)\right\|_2^2 \leq \\ &\leq 2\left(\|X - W_n\|_2^2 + \|X - W_m\|_2^2 - 2\Delta^2\right) \xrightarrow{n,m \rightarrow \infty} 0 \end{aligned}$$

לכן

$$\|W_n - W_m\|_2 \rightarrow 0$$

וזו אכן סדרת קושי. V סגור, ולכן קיים לה גבול $Y \in V$. מתקיים

$$\Delta \leq \|X - Y\|_2 \leq \|X - W_n\|_2 + \|W_n - Y\|_2 \rightarrow \Delta + 0 = \Delta$$

ולכן $\|X - Y\|_2 = \Delta$.
כעת, נבחין שלכל $Z \in V$ מתקיים

$$\mathbb{E}((X - Y)Z) = 0$$

שכן לכל $t \in \mathbb{R}$ מתקיים

$$\begin{aligned} \Delta^2 &\leq \|X - (Y + tZ)\|_2^2 = \mathbb{E}\left((X - Y - tZ)^2\right) = \mathbb{E}\left((X - Y)^2\right) + t^2\mathbb{E}(Z^2) - 2t\mathbb{E}((X - Y)Z) \\ 0 &\leq t^2\mathbb{E}(Z^2) - 2t\mathbb{E}((X - Y)Z) \end{aligned}$$

הגורם הלינארי חייב להיות 0, כי אחרת עבור t קטן מאוד נקבל שאגף ימין שלילי.
באופן דומה, התכונה השנייה גוררת את הראשונה, כיוון שכל $W \in V$ ניתן לרשום בתור $Y + tZ$, ואז

$$\|X - W\|_2^2 = \|X - (Y + tZ)\|_2^2 = \|X - Y\|_2^2 + t^2\|Z\|_2^2 - 2t\mathbb{E}((X - Y)Z)$$

■

ומכאן גם יחידות Y .

1.4 דרך נוספת לחישוב תוחלת

משפט 1.11 יהי X משתנה מקרי ממשי, ותהי $\mathcal{L}(X)$ ההתפלגות שלו - כלומר מידת הסתברות על $\mathbb{R}$ כל שלכל קבוצת בורל A מתקיים $\mathcal{L}(X)(A) = \mathbb{P}(X \in A)$. אזי

$$X \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \iff \int |x| d\mathcal{L}(X) < \infty$$

ומתקיים

$$\mathbb{E}X = \int X(\omega) \, d\mathbb{P}(\omega) = \int x \, d\mathcal{L}(X)$$

באופן כללי יותר, יהי X משתנה מקרי שמקבל ערכים במרחב מדיד $(S, \mathcal{S})$ ותהי $h : S \rightarrow \mathbb{R}$ מדידה. אזי

$$h(X) \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \iff \int |h(x)| \, d\mathcal{L}(X) < \infty$$

ומתקיים

$$\mathbb{E}h(X) = \int h(x) \, d\mathcal{L}(X)$$

הוכחה: נבחין כי הגדרת $\mathcal{L}(X)$ היא השוויון האחרון עבור $h(x) = \mathbb{1}_A(x)$, עבור A מדידה בורל. אכן,

$$\begin{aligned} \mathbb{E}(\mathbb{1}_A(X)) &= \mathbb{P}(X \in A) \\ \int \mathbb{1}_A(x) \, d\mathcal{L}(X) &= \mathcal{L}(X)(A) \end{aligned}$$

מכאן השוויון נובע עבור h פשוטה. לכן אם $h \geq 0$ נקבל את השוויון מהתכנסות מונוטונית על ידי קירוב מלמטה עם פונקציות פשוטות, ומנסיים עבור h כללית מלינאריות ופירוק לחלק שלילי וחיובי. ■

שני מקרים פרטיים חשובים:

1. אם X מקבל מספר בן מנייה של ערכים:

$$\int h(X) \, d\mathcal{L}(X) = \sum_x h(x) \mathcal{L}(X)(x)$$

2. X ממשי (או מקבל ערכים מתוך $\mathbb{R}^n$) בעל התפלגות רציפה בהחלט עם צפיפות f ביחס למידת לבג. למשל, במרחב $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$, המשתנה המקרי $X(\omega) = \omega$ מקיים $X \sim U([0, 1])$, כלומר

$$\mathcal{L}(X) = \text{Leb}|_{[0,1]}$$

על אותו מרחב, $Y(\omega) = -\log(\omega)$ מקיים $Y \sim \text{Exp}(1)$, כלומר $\mathcal{L}(Y)$ רציפה בהחלט עם צפיפות $e^{-x} \mathbb{1}_{[0, \infty)}$, שכן

$$\mathbb{P}(-\log(\omega) \in [t, \infty)) = \mathbb{P}(\omega \leq e^{-t}) = e^{-t}$$

אם X ממש, $\mathcal{L}(X)$ רציפה בהחלט עם צפיפות f , אזי

$$\mathbb{E}h(X) = \int h(x) \, d\mathcal{L}(X) = \int h(x) f(x) \, dx$$

הגדרה 1.12 מידה μ על $\mathbb{R}$ נקראת רציפה בהחלט אם קיימת פונקציה $f : \mathbb{R} \rightarrow [0, \infty]$ כך שמתקיים

$$\mu(A) = \int_A f(x) \, dx$$

לכל A בורל.