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:

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

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:

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

וסיימנו. ■

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

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

כעת יהי $\theta > 0$.

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

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:

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

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

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

אזי

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

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

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

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

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

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

כעת יהי X משתנה מקרי ממשי עם

$$\begin{aligned} \mathbb{E}\left|X\right| &< \infty \\ \mathbb{P}\left(X \in I\right) &= 1 \\ \mathbb{E}\left|f\left(X\right)\right| &< \infty \end{aligned}$$

אזי

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

הוכחה: (רעיון) מקרה קצה - אם $\mathbb{E}X$ היא באחד מקצות הקטע, אזי X שווה בסיכוי 1 לאותו קצה, ואז הטענה ברורה.

אחרת, $\mathbb{E}\left(X\right)$ בתוך הקטע ממש. מקמירות f, קיים לה משיק בנקודה $\mathbb{E}X$, והפונקציה תמיד מעליו. נסמן את המשיק:

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

ואז נקבל

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

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 $X\in L^p\left(\Omega,\mathcal{F},\mathbb{P}\right)$ (או $X\in L^p$) אם $\mathbb{E}\left(\left|X\right|^p\right)<\infty$. אז נסמן

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

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

הוכחה: (במילה) נגדיר את הפונקציה הקמורה:

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

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

$$\left|X\right|^p$$

לא ידוע שהתוחלת שלו סופית, ולכן נשתמש בסדרה, למשל $\min\left(\left|X\right|^p,n\right)$, 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

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

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

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

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

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

אזי קיים $X\in L^p$ כך שמתקיים

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

1.2 המרחב L^2

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

הגדרה 1.6 שונות:

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

הגדרה 1.7 שונות משותפת:

$$\mathrm{Cov}\left(X,Y\right)=\mathbb{E}\left(\left(X-\mathbb{E}X\right)\left(Y-\mathbb{E}Y\right)\right)=\mathbb{E}\left(XY\right)-\left(\mathbb{E}X\right)\left(\mathbb{E}Y\right)$$

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

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

תכונות

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

$$\mathrm{Var}\left(X+Y\right)=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)+2\mathrm{Cov}\left(X,Y\right)$$

הוכחה:

$$\mathrm{Var}\left(X+Y\right)=\mathbb{E}\left(\left(X-\mathbb{E}X+Y-\mathbb{E}Y\right)^2\right)=\mathbb{E}\left(\left(X-\mathbb{E}X\right)^2+\left(Y-\mathbb{E}Y\right)^2+2\left(X-\mathbb{E}X\right)\left(Y-\mathbb{E}Y\right)\right)=$$
$$=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)+2\mathrm{Cov}\left(X,Y\right)$$

■

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

הוכחה: באינדוקציה על k. נשתמש בדרך בכך שמתקיים

$$\mathrm{Cov}\left(X_1+X_2,Y\right)=\mathrm{Cov}\left(X_1,Y\right)+\mathrm{Cov}\left(X_2,Y\right)$$

■

2. כלל המקבילית: אם $X,Y\in L^2$ אזי

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

הוכחה: נחשב:

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

■

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

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

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 $X-W_n,X-W_m$.

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

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

$$||W_n - W_m||_2^2 = 2\,||X - W_n||_2^2 + 2\,||X - W_m||_2^2 - 4\left|\left|X - \frac{1}{2}(W_n + W_m)\right|\right|_2^2 \le$$
$$\le 2\left(||X - W_n||_2^2 + ||X - W_m||_2^2 - 2\Delta^2\right) \xrightarrow{n,m\to\infty} 0$$

לכן

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

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

$$\Delta \le ||X - Y||_2 \le ||X - W_n||_2 + ||W_n - Y||_2 \to \Delta + 0 = \Delta$$

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

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

שכן לכל $t \in \mathbb{R}$ מתקיים

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

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

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

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 $\mathbb{R}$ כך שלכל קבוצת בורל A מתקיים $\mathcal{L}(X)(A) = \mathbb{P}(X \in A)$. אזי

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

ומתקיים

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

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

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

ומתקיים

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

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

$$\mathbb{E}(\mathbb{1}_A(x)) = \mathbb{P}(X \in A)$$

$$\int \mathbb{1}_A(x)\ \mathrm{d}\mathcal{L}(X) = \mathcal{L}(X)(A)$$

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:

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

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

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

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

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

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

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

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

$$\mu(A) = \int\limits_A f(x)\, \mathrm{d}x$$

לכל A בורל.