הסתברות למתמטיקאים

ⓒ ארזים

26 באפריל 2017

1 חוק חזק של מספרים גדולים

1.1 אי תלות ותוחלת

טענה 1.1 יהיו X,Y משתנים מקריים בלתי תלויים עם $\mathbb{E}\left|X\right|,\mathbb{E}\left|Y\right|<\infty$. אזי $\mathbb{E}\left|XY\right|<\infty$ ומקתיים $\mathbb{E}\left(XY\right)=\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)$.

הוכחה: ראשית ניזכר כי המשמעות של האי תלות היא שלכל $A\in\sigma\left(X\right),B\in\sigma\left(Y\right)$ מתקיים $\mathbb{P}\left(A\cap B\right)=\mathbb{P}\left(A\right)\mathbb{P}\left(B\right)$.

מספיק להוכיח עבור X,Y אי שליליים (כי $X=X^{+}-X^{-}$, ושני אלה אי שליליים).
כעת, עבור $X\geq0$, קיימת סדרת משתנים מקריים מהצורה

$$X_{n}=\sum_{j}\alpha_{j,n}\mathbb{1}_{A_{j,n}}$$

כאשר $A_{j,n}\in\sigma\left(X\right)$ וכן $X_{n}\uparrow X$ כמעט תמיד. נוכל להגדיר:

$$X_{n}=\sum_{j=0}^{n2^{n}-1}\frac{j}{2^{n}}\mathbb{1}_{\left\{ \frac{j}{2^{n}}\leq X\leq\frac{j+1}{2^{n}}\right\} }+n\mathbb{1}_{\left\{ X\geq n\right\} }$$

נבחין כי $X_{n}\to X$ כמעט תמיד. כמו כן, $X_{n+1}\geq X_{n}$, שכן אם למשל $X_{n}=\frac{j}{2^{n}}$ אזי $\frac{j}{2^{n}}\leq X\leq\frac{j+1}{2^{n}}$, ואז $X_{n+1}\in\left\{ \frac{j}{2^{n}},\frac{2j+1}{2^{n+1}}\right\}$.
כעת, נקבל

$$\begin{aligned}\mathbb{E}\left(X_{n}Y_{n}\right) & =\mathbb{E}\left(\sum_{j,k}\alpha_{j,n}\mathbb{1}_{A_{j,n}}\beta_{k,n}\mathbb{1}_{B_{k,n}}\right)=\\
 & =\sum_{j,k}\alpha_{j,n}\beta_{j,n}\mathbb{P}\left(A_{j,n}\cap B_{k,n}\right)=\sum_{j,k}\alpha_{j,n}\beta_{k,n}\mathbb{P}\left(A_{j,n}\right)\mathbb{P}\left(B_{k,n}\right)=\\
 & =\underbrace{\left(\sum_{j}\alpha_{j,n}\mathbb{P}\left(A_{j,n}\right)\right)}_{\mathbb{E}X_{n}}\underbrace{\left(\sum_{k}\beta_{k,n}\mathbb{P}\left(B_{k,n}\right)\right)}_{\mathbb{E}Y_{n}}
\end{aligned}$$

כעת, $X_{n}Y_{n}\uparrow XY$, ולכן מהתכנסות מונוטונית נקבל את הנדרש. ■

מסקנה 1.2 אם $X,Y\in L^2$ ובלתי תלויים, אזי $\operatorname{Cov}(X,Y)=0$, ולכן $\operatorname{Var}(X+Y)=\operatorname{Var}(X)+\operatorname{Var}(Y)$.

נתאר את החוקים של המספרים הגדולים: יהיו (X_n) סדרת משתנים מקריים בלתי תלויים ושווי התפלגות (לכל m,n מתקיים $\mathcal{L}(X_m)=\mathcal{L}(X_n)$). נניח כי $\mathbb{E}X_1=a$. אזי $\frac{X_1+\cdots+X_n}{n}\xrightarrow{n\to\infty}a$ כמעט תמיד (החוק החזק), וכן לכל $\varepsilon>0$, מתקיים $\mathbb{P}\left(\left|\frac{X_1+\cdots+X_n}{n}-a\right|>\varepsilon\right)\xrightarrow{n\to\infty}0$ (החוק החלש).

משפט 1.3 (חוק חזק תחת הנחת מומנט רביעי) תהי (X_n) סדרת משתנים מקריים בלתי תלויים, לאו דווקא שווי התפלגות. נניח שקיים $K<\infty$ עבורו לכל n מתקיים

$$\mathbb{E}\left(X_n^4\right)\leq K$$

וכן $\mathbb{E}X_n=0$. אזי $\frac{X_1+\cdots+X_n}{n}\xrightarrow{n\to\infty}0$ כמעט תמיד.

הוכחה: נגדיר $S_n=X_1+\cdots+X_n$. כעת,

$$\mathbb{E}\left(\frac{S_n}{n}\right)=\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\left(X_j\right)=0$$

כעת, נעריך את המומנט הרביעי:

$$\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right]=\frac{1}{n^4}\mathbb{E}\left[\left(X_1+\cdots+X_n\right)^4\right]$$

בפתיחת הסוגריים, נקבל גורמים מהצורות הבאות:

$$X_n^4,X_mX_n^3,X_m^2X_n^2,X_lX_mX_n^2,X_kX_lX_mX_n$$

כאשר $k\neq l\neq m\neq n$. כל גורם שמכיל אחד מבין X_i בחזקה ראשונה מתאפס בתוחלת, כי הם בלתי תלויים (מהטענה לעיל נוכל לפרק למכפלת התוחלות). לכן יישאר לנו:

$$\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right]=\frac{1}{n^4}\left(\sum_{j=1}^{n}\underbrace{\mathbb{E}\left[X_n^4\right]}_{\leq K}+6\sum_{m<k}\underbrace{\mathbb{E}\left[X_m^2X_k^2\right]}_{\leq\sqrt{\mathbb{E}[X_m^4]\mathbb{E}[X_n^4]}\leq K}\right)\leq\frac{1}{n^4}\left(nK+6\binom{n}{2}K\right)\leq\frac{1}{n^4}3n^2K=\frac{3K}{n^2}$$

נובע שלכל $\varepsilon>0$ מתקיים

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right|>\varepsilon\right)=\mathbb{P}\left(\left|\frac{S_n}{n}\right|^4\geq\varepsilon^4\right)\leq\frac{3K}{n^2\varepsilon^4}$$

הטור הזה סכים, ולכן מלמת בורל קנטלי הראשוני, כמעט תמיד, רק מספר סופי של המאורעות $\left\{\left|\frac{S_n}{n}\right|>\varepsilon\right\}$ קורה. ניקח $\varepsilon_k=\frac{1}{k}$ ואז כמעט תמיד לכל k בו זמנית, רק מספר סופי מבין $\left\{\left|\frac{S_n}{n}>\varepsilon_k\right|\right\}$ קורה, כלומר $\frac{S_n}{n}\to 0$ כמעט תמיד.

סיום מהיר יותר:

$$\mathbb{E}\left(\left(\frac{S_n}{n}\right)^4\right) \leq \sum_n \frac{3K}{n^2} < \infty$$

ולכן $\sum\left(\frac{S_n}{n}\right)^4$ סופי כמעט תמיד, כלומר $\frac{S_n}{n} \to 0$ כמעט תמיד. ■

מסקנה 1.4 תהי (X_n) סדרת משתנים מקריים בתי תלויים עם $\mathbb{E}(X_n) = a$ וכן יהי K עבורו

$$\mathbb{E}\left(X_n^4\right) \leq K$$

אזי $a \xrightarrow{n\to\infty} \frac{X_1+\cdots+X_n}{n}$ כמעט תמיד.

הוכחה: מפעילים את המשפט הקודם על $Y_n = X_n - a$. יש לוודא שהמומנט הרביעי של Y_n חסום במידה אחידה, ואכן

$$||Y_n||_4 \leq ||X_n||_4 + \underbrace{||a||_4}_{|a|} \leq K^{\frac{1}{4}} + |a|$$

ולכן הכל עובד. ■

טענה 1.5 (אי שוויון צ'בישב) יהי $X \in L^2$. אזי לכל $t > 0$ מתקיים

$$\mathbb{P}\left(|X - \mathbb{E}X| \geq t\right) \leq \frac{\mathrm{Var}\left(X\right)}{t^2}$$

הוכחה:

$$\{|X - \mathbb{E}X| \geq t\} = \left\{(X - \mathbb{E}X)^2 \geq t^2\right\}$$

וכעת נפעיל את אי שוויון מרקוב. ■

דוגמא נסמן S_n את אוסף התמורות של $\{1, \ldots, n\}$. למשל, $|S_n| = n!$. תהי π תמורה שנבחרה אחיד מתוך S_n. כמה מעגלים יש לה? נסמן $C\left(\pi\right)$ את כמות המעגלים של π.

תהליך המסעדה הסינית: הלקוח מספר k שנכנס למסעדה בוחר באופן אחיד מבין k אפשרויות למקומות ישיבה (בין כל שני לקוחות בשולחן, או פתיחת שולחן חדש). נעצור אחרי שנכנסו n אנשים. נשים לב שיש התאמה חד חד ערכית בין תמורות לבין סידורי ישיבה, ובגלל שכל לקוח מגריל אחיד את מקום הישיבה שלו, התמורה מתפלגת אחיד. לכן התהליך מפיק תמורה אחידה. כעת ננסה לספור מעגלים.
נסמן A_k את המאורע בו הלקוח מספר k פתח שולחן חדש. אזי

$$C\left(\pi\right) = \sum_{k=1}^{n} \mathbb{1}_{A_k}$$

נבחין שהמאורעות A_k בלתי תלויים, וכן $\mathbb{P}(A_k)=\frac{1}{k}$. לכן

$$\mathbb{E}(C)=\sum_{k=1}^{n}\frac{1}{k}\sim\log n$$

$$\mathrm{Var}(C)=\sum_{k=1}^{n}\frac{k-1}{k^2}\sim\log n$$

מאי שוויון צ'בישב נקבל כי לכל $\varepsilon>0$ מתקיים

$$\mathbb{P}(C>(1+\varepsilon)\log n)=\mathbb{P}(C-\mathbb{E}C>(1+\varepsilon)\log n-\mathbb{E}C)\leq$$
$$\leq\frac{\mathrm{Var}C}{((1+\varepsilon)\log n-\mathbb{E}C)^2}\sim\frac{1}{\varepsilon^2\log n}$$

באופן דומה,

$$\mathbb{P}(C<(1-\varepsilon)\log n)\lesssim\frac{1}{\varepsilon^2\log n}$$

2 מידות מכפלה

אם μ_1,μ_2 מידות σ סופיות על (S_1,Σ_1), (S_2,Σ_2) בהתאמה, אזי קיימת מידת מכפלה $\mu_1\times\mu_2$ על מרחב המכפלה $(S_1\times S_2,\Sigma_1\otimes\Sigma_2)$ ומתקיים משפט פוביני: לכל $f:S_1\times S_2\to\mathbb{R}$ מדידה ואי שלילית (או אינטגרבילית) מתקיים

$$\int f(x,y)\,\mathrm{d}(\mu_1\times\mu_2)=\int_{S_2}\int_{S_1}f(x,y)\,\mathrm{d}\mu_1(x)\,\mathrm{d}\mu_2(y)=\int_{S_1}\int_{S_2}f(x,y)\,\mathrm{d}\mu_2(y)\,\mathrm{d}\mu_1(x)$$

2.1 הקשר בין אי תלות לבין מידות מכפלה

יהיו X,Y משתנים מקריים המקבלים ערכים במרחבים מדידים $(S_1,\Sigma_1),(S_2,\Sigma_2)$ בהתאמה.
ניזכר בהתפלגות $\mathcal{L}(X)$ - זו מידת ההסתברות

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

עבור $A\in\Sigma_1$. נוכל לתת גם את ההתפלגות המשותפת $\mathcal{L}(X,Y)$ - זו מידת ההסתברות

$$\mathcal{L}(X,Y)(B)=\mathbb{P}((X,Y)\in B)$$

עבור $B\in\Sigma_1\otimes\Sigma_2$. אזי X,Y בלתי תלויים אם ורק אם

$$\mathcal{L}(X,Y)=\mathcal{L}(X)\times\mathcal{L}(Y)$$

הוכחה: כדי להראות את השוויון, מספיק לבדוק ששתי המידות הללו מזדהות על מערכת π שיוצרת את $\Sigma_1 \otimes \Sigma_2$. נבחר באוסף $\{A_1 \times A_2 \mid A_1 \in \Sigma_1, A_2 \in \Sigma_2\}$. אכן,

$$\begin{aligned}\mathcal{L}(X,Y)(A_1 \times A_2) &= \mathbb{P}((X,Y) \in A_1 \times A_2) = \mathbb{P}(X \in A_1, Y \in A_2) = \\ &= \mathbb{P}(X \in A_1)\mathbb{P}(Y \in A_2) = \mathcal{L}(X)(A_1)\mathcal{L}(X)(A_2) = \\ &= (\mathcal{L}(X) \times \mathcal{L}(Y))(A_1 \times A_2)\end{aligned}$$

■

מקרים פרטיים חשובים

1. נניח כי X,Y ממשיים. אזי X,Y בלתי תלויים אם ורק אם לכל $s,t \in \mathbb{R}$ מתקיים

$$\mathbb{P}(X \le s, Y \le t) = \mathbb{P}(X \le s)\mathbb{P}(Y \le t)$$

2. אם X,Y בעלי צפיפות משותפת $f(x,y)$ אז X,Y בלתי תלויים אם ורק אם קיימות פונקציות $g(x), h(y)$ עם

$$f(x,y) = g(x)h(y)$$

כמעט בכל מקום (אוטומטית נובע כי קיים $c > 0$ קבוע עבורו cg צפיפות של X, $\frac{1}{c}h$ צפיפות של Y).

3 תורת המרטינגלים

3.1 תוחלת מותנה

נזכרת בתוחלת מותנה על מרחב הסתברות בדיד.
נאמר כי X,Z משתנים מקריים המקבלים מספר סופי של ערכים:

$$\begin{aligned}X &: x_1, \dots, x_n \in \mathbb{R} \\ Z &: z_1, \dots, z_n \in \mathbb{R}\end{aligned}$$

הסתברות מותנה: נתנה בכך שמתקיים $Z = z_1$:

$$\mathbb{P}(A \mid Z = z_1) = \frac{\mathbb{P}(A \cap \{Z = z_1\})}{\mathbb{P}(Z = z_1)}$$

ואז

$$\mathbb{E}(X \mid Z = z_1) = \sum_i x_i \mathbb{P}(X = x_i \mid Z = z_1)$$

אפשר גם לדבר על $\mathbb{E}\left(X \mid Z\right)$. זהו משתנה מקרי, שהוא פונקציה של Z. הערך שלו בנקודה מסויימת z_i הוא

$$f\left(z_i\right) = \mathbb{E}\left(X \mid Z = z_i\right)$$

נשים לב שהמשתנה $\mathbb{E}\left(X \mid Z\right)$ מדיד לפי $\sigma\left(Z\right)$. בנוסף,

$$\begin{aligned}\mathbb{E}\left(\mathbb{E}\left(X \mid Z\right)\mathbb{1}_{Z=z_1}\right) &= \mathbb{E}\left(\mathbb{E}\left(X \mid Z = z_1\right)\cdot \mathbb{1}_{Z=z_1}\right) = \\ &= \mathbb{E}\left(X \mid Z = z_1\right)\mathbb{P}\left(Z = z_1\right) = \\ &= \sum_i x_i \mathbb{P}\left(X = x_i, Z = z_1\right) = \mathbb{E}\left(X \cdot \mathbb{1}_{Z=z_1}\right)\end{aligned}$$

למעשה, לכל $A \in \sigma\left(Z\right)$ מתקיים, מלינאריות,

$$\mathbb{E}\left(\mathbb{E}\left(X \mid Z\right)\cdot \mathbb{1}_A\right) = \mathbb{E}\left(X\mathbb{1}_A\right)$$

משפט 3.1 (עם הגדרה) יהי $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$ מרחב הסתברות, ויהי X משתנה מקרי אינטגרבילי. תהי $\mathcal{G} \subseteq \mathcal{F}$ תת סיגמא אלגברה. אזי קיים משתנה מקרי Y עם

1. Y מדיד לפי $\mathcal{G}$.
2. Y אינטגרבילי.
3. לכל קבוצה $A \in \mathcal{G}$ מתקיים

$$\mathbb{E}\left(Y\mathbb{1}_A\right) = \mathbb{E}\left(X\mathbb{1}_A\right)$$

אם $\tilde{Y}$ משתנה מקרי שמקיים תכונות אלה אזי $Y = \tilde{Y}$ כמעט תמיד. כל משתנה מקרי Y שמקיים את התכונות הללו נקרא <u>גרסה</u> של $\mathbb{E}\left(X \mid \mathcal{G}\right)$. נרשום זאת בצורה $Y = \mathbb{E}\left(X \mid \mathcal{G}\right)$ כמעט תמיד.

הגדרה 3.2 נגדיר גם

$$\mathbb{E}\left(X \mid Z\right) = \mathbb{E}\left(X \mid \sigma\left(Z\right)\right)$$

כאשר Z משתנה מקרי. באופן דומה גם

$$\mathbb{E}\left(X \mid Z_1, Z_2, \ldots\right) = \mathbb{E}\left(X \mid \sigma\left(Z_1, Z_2, \ldots\right)\right)$$

נאמר לרגע כי X,Z ממשיים, $Z\sim U\left([0,1]\right)$. אזי $\mathbb{P}\left(Z=z\right)=0$ לכל $z\in\mathbb{R}$, ולכן לא נוכל להגדיר ישירות את $\mathbb{E}\left(X\mid Z=\frac{1}{2}\right)$. עדיין נוכל להגדיר את $\mathbb{E}\left(X\mid Z\right)$ והוא מוגדר לכמעט כל ערך של Z.

כדאי לחשוב על $\mathbb{E}\left(X\mid Z\right)$ כגבול -

$$\mathbb{E}\left(X\mid Z=z_0\right)=\lim_{\varepsilon\downarrow0}\mathbb{E}\left(X\mid Z\in\left(z_0-\varepsilon,z_0+\varepsilon\right)\right)$$

ככה אנו מגדירים תוחלת מותנה, ואפשר גם להגדיר הסתברות מותנה

$$\mathbb{P}\left(A\mid\mathcal{G}\right)\coloneqq\mathbb{E}\left(\mathbb{1}_A\mid\mathcal{G}\right)$$

מקרה פרטי אם $X\in L^2$, אזי $\mathbb{E}\left(X\mid\mathcal{G}\right)$ היא הטלה אורתוגונלית לתת הרחב $L^2\left(\Omega,\mathcal{G},\mathbb{P}\right)$. כלומר

$$\inf_{Y\in L^2\left(\Omega,\mathcal{G},\mathbb{P}\right)}\mathbb{E}\left[\left(X-Y\right)^2\right]=\mathbb{E}\left[\left(X-\mathbb{E}\left(X\mid\mathcal{G}\right)\right)^2\right]$$

הוכחה: (של המשפט) יהיו $Y,\tilde{Y}$ זוג משתנים מקריים שמקיימים את $1,2,3$. אם $\mathbb{P}\left(Y>\tilde{Y}\right)>0$, אז קיים n עבורו

$$\mathbb{P}\left(Y>\tilde{Y}+\frac{1}{n}\right)$$

יהי $A=\left\{Y>\tilde{Y}+\frac{1}{n}\right\}$. אזי $A\in\mathcal{G}$ לפי תכונה 1. לכן, לפי תכונה 3, מתקיים

$$\mathbb{E}\left(Y\mathbb{1}_A\right)=\mathbb{E}\left(X\mathbb{1}_A\right)=\mathbb{E}\left(\tilde{Y}\mathbb{1}_A\right)$$

ולכן נקבל כי

$$\mathbb{E}\left(\left(Y-\tilde{Y}\right)\mathbb{1}_A\right)=0$$

אבל

$$\mathbb{E}\left(\left(Y-\tilde{Y}\right)\mathbb{1}_A\right)=\mathbb{E}\left(Y-\tilde{Y}\right)\mathbb{1}_{\left\{Y>\tilde{Y}+\frac{1}{k}\right\}}\geq\frac{1}{n}\mathbb{P}\left(Y>\tilde{Y}+\frac{1}{n}\right)>0$$

בסתירה. לכן קיבלנו יחידות.

נוכיח קיום במקרה הפרטי $X\in L^2$. תהי Y הטלה אורתוגונלית של X למרחב $L^2\left(\Omega,\mathcal{G},\mathbb{P}\right)$. אזי Y מדיד לפי $\mathcal{G}$, וכן $Y\in L^2$. כעת, ראינו כי לכל $Z\in L^2$ מתקיים

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

נבחר $Z = \mathbb{1}_A$ עבור $A \in \mathcal{G}$ ונקבל

$$\mathbb{E}\left(\mathbb{1}_A\left(X-Y\right)\right) = \mathbb{E}\left(X\mathbb{1}_A\right) - \mathbb{E}\left(Y\mathbb{1}_A\right) = 0$$

כנדרש.

כעת נדון במקרה הכללי. מספיק לבדוק קיום כאשר $X \geq 0$. קיימת סדרה $X_n \geq 0$ כך שכל X_n חסום וכן $X_n \uparrow X$ כמעט תמיד. לכל X_n קיימת $\mathbb{E}\left(X_n \mid \mathcal{G}\right)$ כי $X_n \in L^2$ מחסימות. נותר להוכיח כי $\mathbb{E}\left(X_n \mid \mathcal{G}\right) \leq \mathbb{E}\left(X_{n+1} \mid \mathcal{G}\right)$. בהנחה זו ניקח את הגבול

$$Y = \limsup_n \mathbb{E}\left(X_n \mid \mathcal{G}\right)$$

ומהתכנסות מונוטונית, $Y = \mathbb{E}\left(X \mid \mathcal{G}\right)$. בשביל זה ניעזר בלמה (בשיעור הבא):

למה 3.3 אם U משתנה מקרי חסום ואי שלילי, אזי $\mathbb{E}\left(U \mid \mathcal{G}\right) \geq 0$ כמעט תמיד.

מלמה זו נובע הכל, על ידי לקיחת $U = X_{n+1} - X_n$. ∎