<u>הסתברות למתמטיקאים</u>

© ארזים

14 ביוני 2017

1 מרטינגלים אינטגרביליים במידה שווה

משפט 1.1 (האי שוויון המקסימלי של דוב) יהי Z תת מרטינגל אי שלילי ויהי $t > 0$. אזי

$$\mathbb{P}\left(\max_{0\le k\le n} Z_k \ge t\right) \le \frac{\mathbb{E}\left(Z_n \mathbb{1}_{\max\limits_{1\le k\le n} Z_k \ge t}\right)}{t} \le \frac{\mathbb{E}\left(Z_n\right)}{t}$$

את המשפט הזה והוכחתו ראינו בתרגול. כעת נשתמש בו כדי להוכיח משפט אחר.

משפט 1.2 (התכנסות המרטינגל בנורמת L^p) יהי $p > 1$. יהי Z תת מרטינגל אי שלילי חסום בנורמת L^p. אזי קיים Z_∞ כך שמתקיים $Z_n \to Z_\infty$ כמעט תמיד ובנורמת L^p. כמו כן,

$$||Z_\infty||_p = \lim_{n\to\infty} ||Z_n||_p = \sup_n ||Z_n||_p$$

בנוסף, אם נגדיר

$$Z^* = \sup_{n\ge 0} Z_n$$

אזי

$$||Z^*||_p \le \frac{p}{p-1} \sup_n ||Z_n||_p$$

הערה 1.3 נובע שהמרטינגל כולו נשלט על ידיד משתנה מקרי אינטגרבילי בנורמת L^p, $Z_n \le Z^*$.

הוכחה: (של המשפט) נסמן

$$Z_N^* = \sup_{0\le n\le N} Z_n$$

מהאי שוויון המקסימלי של דוב נובע לכל $t > 0$ שמתקיים

$$\mathbb{P}\left(Z_N^* \geq t\right) \leq \frac{\mathbb{E}\left(Z_N \mathbb{1}_{\max\limits_{1 \leq n \leq N} Z_n \geq t}\right)}{t}$$

כעת ניעזר בלמה, שאותה נוכיח אחר כך: בתנאים הללו, מתקיים

$$||Z_N^*||_p \leq \frac{p}{p-1} ||Z_N||_p$$

כעוון שהמרטינגל Z חסום בנורמת L^P, נובע שהוא חסום גם בנורמת L^1. לכן קיים Z_∞ כך שמתקיים $Z_n \to Z_\infty$ כמעט תמיד. מהלמה שראינו ומהתכנסות מונוטונית נובע כי

$$||Z^*||_p \leq \frac{p}{p-1} \lim_{N\to\infty} ||Z_N||_p < \infty$$

כעת, נראה שמתקיים $Z_n \to Z_\infty$ גם בנורמת L^p. מתקיים

$$|Z_n - Z_\infty|^p \leq (2Z^*)^p \in L^1$$

ולכן מהתכנסות נשלטת נקבל $Z_n \to Z_\infty$ בנורמת L^p. לבסוף, נרצה להראות שהגבול

$$\lim_n ||Z_n||_p$$

הוא עולה. נשתמש באי שוויון ינסן המותנה:

$$\mathbb{E}\,|Z_n|^p = \mathbb{E}\left(\mathbb{E}\left(|Z_n|^p \mid \mathcal{F}_{n-1}\right)\right) \geq \mathbb{E}\left(|\mathbb{E}\left(Z_n \mid \mathcal{F}_{n-1}\right)|^p\right) \geq \mathbb{E}\,|Z_{n-1}|^p$$

וסיימנו. ■

מסקנה 1.4 יהי M מרטינגל חסום בנורמת L^p. אזי $M_n \to M_\infty$ כמעט תמיד ובנורמת L^p.

הוכחה: נגדיר $Z_n = |M_n|$. אזי Z תת מרטינגל אי שלילי, ועליו נפעיל את המשפט הקודם. ■

נחזור כעת ללמה מההוכחה.

למה 1.5 יהי $p > 1$. יהיו X, Y משתנים מקריים אי שליליים, ונניח שלכל $t > 0$ מתקיים

$$\mathbb{P}\left(X \geq t\right) \leq \frac{\mathbb{E}\left(Y \mathbb{1}_{X \geq t}\right)}{t}$$

אזי

$$||X||_p \leq \frac{p}{p-1} ||Y||_p$$

הוכחה: נפתח נוסחא עבור $\mathbb{E}\left(X^p\right)$ באמצעות $\mathbb{P}\left(X \geq t\right)$. מתקיים

$$\mathbb{E}\left(X^p\right) = \mathbb{E}\left(p\int\limits_0^X t^{p-1}\,\mathrm{d}t\right) = \mathbb{E}\left(p\int\limits_0^\infty t^{p-1}\mathbb{1}_{t\leq X}\,\mathrm{d}t\right)$$

כעת, הכל כאן אי שלילי, ולכן ממשפט פוביני נקבל

$$\begin{aligned}\mathbb{E}\left(X^p\right) &= p\int\limits_0^\infty t^{p-1}\mathbb{E}\left(\mathbb{1}_{t\leq X}\right)\,\mathrm{d}t = p\int\limits_0^\infty t^{p-1}\mathbb{P}\left(X \geq t\right)\,\mathrm{d}t \leq \\ &\leq p\int\limits_0^\infty t^{p-2}\mathbb{E}\left(Y\mathbb{1}_{X\geq t}\right)\,\mathrm{d}t = p\mathbb{E}\left(Y\underbrace{\int\limits_0^\infty t^{p-1}\mathbb{1}_{t\leq X}\,\mathrm{d}t}_{\int_0^X t^{p-2}\,\mathrm{d}t=\frac{X^{p-1}}{p-1}}\right) = \\ &= \frac{p}{p-1}\mathbb{E}\left(YX^{p-1}\right)\end{aligned}$$

נפעיל את אי שוויון הלדר על חזקת p (כי זה מה שרוצים לקבל בסוף) ונקבל

$$\mathbb{E}\left(X^p\right) \leq \frac{p}{p-1}\mathbb{E}\left(Y^p\right)^{\frac{1}{p}}\mathbb{E}\left(X^p\right)^{\frac{p-1}{p}}$$

אם הגורם הימני ביותר סופי, נעביר אגפים ונקבל

$$\begin{aligned}\mathbb{E}\left(X^p\right)^{\frac{1}{p}} &\leq \frac{p}{p-1}\mathbb{E}\left(Y^p\right)^{\frac{1}{p}} \\ ||X||_p &\leq \frac{p}{p-1}\,||Y||_p\end{aligned}$$

אחרת, כדי לאשפר את הצקרה שבו $\mathbb{E}\left(X^p\right) = \infty$ נביט במשתנה $\min\left(X,n\right)$ ונבחין שגם הוא מקיים

$$\mathbb{P}\left(\min\left(X,n\right) \geq t\right) \leq \frac{\mathbb{E}\left(Y\mathbb{1}_{\min(X,n)\geq t}\right)}{t}$$

לכל $t > 0$. לכן נקבל מהשוויונות שלמעלה

$$||\min\left(X,n\right)||_p \leq \frac{p}{p-1}\,||Y||_p$$

ונסיים מהתכנסות מונוטונית. ■

1.1 דיכוטומיית קקוטני

ראינט שלוש דוגמאות בסיסיות למרטינגלים:

1. סכום של משתנים בלתי תלויים בתוחלת אפס.
2. חשיפה של משתנה מקרי אינטגרבילי לפי פילטרציה $\mathcal{F}_n$.
3. מכפלה של משתנים בלתי תלויים אי שליליים בתוחלת 1.

על שני המקרים הראשונים דנו בהרחבה. כעת נדבר על השלישי.

משפט 1.6 (דיכוטומיית קקוטני) תהי (X_n) סדרת משתנים מקריים בלתי תלויים ואי שליליים עם $\mathbb{E}X_n = 1$ לכל n. נגדיר

$$M_n = \prod_{i=1}^{n} X_i$$

אזי התכונות הבאות שקולות:

1. $\mathbb{E}\left(M_\infty\right) = 1$.
2. $M_n \to M_\infty$ בנורמת L^1.
3. (M_n) אינטגרבילי במידה שווה.
4. נגדיר

$$0 \leq a_n = \mathbb{E}\left(\left(X_n\right)^{\frac{1}{2}}\right) \leq 1$$

אזי

$$\prod_n a_n > 0$$

5. עם אותה הגדרה כמו בסעיף הקודם,

$$\sum_n 1 - a_n < \infty$$

הוכחה: השקילות בין $1,2,3$ ידועה לנו כבר (בעזרת למת שפה מראים $2 \Rightarrow 1$). השקילות בין $4,5$ נכונה לכל סדרת מספרים $a_n \in [0,1]$. נותר להראות, למשל, כי $2,4$ שקולים.
נראה בסופו של דבר שתחת תנאי 4 מתקיים $\sup_n M_n \in L^1$, ונסיים מהתכנסות נשלטת.
נגדיר מרטינגל

$$N_n = \prod_{i=1}^{n} \frac{X_i^{\frac{1}{2}}}{a_i}$$

תחת הנחה 4, נקבל כי

$$\mathbb{E}\left(N_n^2\right)=\left(\prod_{j=1}^{n}a_j^2\right)^{-1}\leq\left(\prod_n a_n^2\right)^{-1}<\infty$$

מהאי שוויון במשפט הקודם נקבל

$$\mathbb{E}\left(\sup_n N_n^2\right)\leq 4\sup_n\mathbb{E}\left(N_n^2\right)<\infty$$

אבל

$$\sup_n N_n^2=\sup_n\frac{M_n}{a_1\cdots a_n}\geq\sup_n M_n$$

ולכן

$$\sup_n M_n\in L^1$$

מכאן יש התכנסות נשלטת כמו שרצינו.
בכיוון השני, נניח כי

$$\prod_n a_n=0$$

ונראה שאין התכנסות בנורמת L^1. בפרט נראה כי $M_\infty=0$ כמעט תמיד. נבחין שגם $N_n\to N_\infty$ כמעט תמיד, כי $\mathbb{E}N_n=1$ לכל n, וכן $N_n\geq 0$. לכן, תחת ההנחה שהמכפלה מתאפסת נובע שגם

$$\prod_n X_n^{\frac{1}{2}}=0$$

כמעט תמיד. לכן נסיים. ■

2 משפט הגבול המרכזי

2.1 פונקציות אופייניות

אנחנו מתחילים כעת תחום חדש לחלוטין שילווה אותנו עד סוף הקורס. ננסח כבר עכשיו את משפט הגבול המרכזי, שהוא אחת המטרות שלנו בפרק.

משפט 2.1 (הגבול המרכזי) תהי (X_n) סדרת משתנים מקריים בלתי תלויים ושווי התפלגות עם $\mathbb{E}X_n=0$ ועם $\mathrm{Var}\left(X_n\right)=1$. אזי לכל $a<b$ מתקיים

$$\mathbb{P}\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\in[a,b]\right)\to\mathbb{P}\left(Z\in[a,b]\right)$$

כאשר $Z \sim N[0,1]$. משמע,

$$\mathbb{P}\left(\frac{X_1+\cdots+X_n}{\sqrt{n}} \in [a,b]\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\,\mathrm{d}x$$

המטרה האחרת של הפרק היא לפתח כלים להוכחת המשפט. נתחיל לדון בפונקציות אופייניות (טרנספורם פורייה).

הגדרה 2.2 הפונקצייה האופיינית של X היא $\varphi = \varphi_X : \mathbb{R} \to \mathbb{C}$ המוגדרת על ידי

$$\varphi(\theta) = \mathbb{E}\left(e^{i\theta X}\right) = \mathbb{E}(\cos(\theta X)) + i\mathbb{E}(\sin(\theta X))$$

תכונות

1. $\varphi(0) = 1$.
2. $|\varphi(\theta)| \leq 1$ לכל θ.
3. φ פונקציה רציפה (אם $\theta_n \to \theta$ אזי $\mathbb{E}\left(e^{i\theta_n X}\right) \to \mathbb{E}\left(e^{i\theta X}\right)$ מהתכנסות נשלטת).
4. $\varphi_{-X}(\theta) = \overline{\varphi_X(\theta)}$.
5. $\varphi_{aX+b}(\theta) = e^{i\theta b}\varphi_X(a\theta)$.
6. תכונה מרכזית לדיון שלנו: אם X, Y בלתי תלויים אזי $\varphi_{X+Y}(\theta) = \varphi_X(\theta)\varphi_Y(\theta)$.

הערה 2.3 נובע שכאשר (X_n) בלתי תלויים ושווי התפלגות מתקיים

$$\varphi_{\frac{X_1+\cdots+X_n}{\sqrt{n}}}(\theta) = \left(\varphi_{X_1}\left(\frac{\theta}{\sqrt{n}}\right)\right)^n$$

ומכאן, כאשר $\mathbb{E}X_1 = 0$, $\mathrm{Var}(X_1) = 1$, נראה בעתיד שאגף ימין שואף אל $e^{-\frac{\theta^2}{2}} = \varphi_{N(0,1)}$, היות ומתקיים

$$\varphi_{X_1}(\theta) = 1 - \frac{\theta^2}{2n} + o\left(\frac{1}{n}\right)$$

את כל זה נראה בעתיד.

2.1.1 משפט ההיפוך של לוי

משפט זה אומר לנו שהפונקצייה האופיינית קובעת את ההתפלגות.

משפט 2.4 (ההיפוך של לוי) יהי X משתנה מקרי, ותהי $\varphi = \varphi_X$ הפונקציה האופיינית שלו. אזי לכל $a \le b$ מתקיים

$$\lim_{T\uparrow\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-i\theta a} - e^{-i\theta b}}{i\theta} \varphi\left(\theta\right) \,\mathrm{d}\theta = \frac{1}{2}\mathbb{P}\left(X=a\right) + \mathbb{P}\left(X\in\left(a,b\right)\right) + \frac{1}{2}\mathbb{P}\left(X=b\right)$$

לפעמים יש נוסחא יותר פשוטה. אם $\int \left|\varphi\left(x\right)\right| \,\mathrm{d}x < \infty$ אזי ההתפלגות של X רציפה בהחלט ביחס למידת לבג והצפיפות שלה fנתונה על ידי

$$f\left(x\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\theta x} \varphi\left(\theta\right) \,\mathrm{d}x$$

הערה 2.5 כאשר להתפלגות של X יש צפיפות f אזי

$$\varphi\left(\theta\right) = \mathbb{E}\left(e^{i\theta X}\right) = \int_{-\infty}^{\infty} e^{i\theta x} f\left(x\right) \,\mathrm{d}x$$

מספר דוגמאות לפונקציות אופייניות

1. התפלגות נורמלית - $X \sim N\left(a,\sigma^2\right)$, הצפיפות היא $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$ והפונקצייה האופיינית היא

$$e^{ia\theta - \frac{1}{2}\sigma^2\theta^2}$$

2. התפלגות אחידה - $X \sim U\left[0,1\right]$, הצפיפות היא $\mathbb{1}_{[0,1]}\left(x\right)$, והפונקצייה האופיינית היא

$$\frac{e^{i\theta} - 1}{i\theta}$$

3. התפלגות קושי - הצפיפות היא $\frac{1}{\pi(1+x^2)}$, והפונקצייה האופיינית היא

$$e^{-|\theta|}$$

2.2 התכנסות בהתפלגות (התכנסות חלשה, התכנסות חלשה*)

נרצה לדבר על התכנסות של סדרת מידות הסתברות μ_n על $\mathbb{R}$ למידה גבולית μ על $\mathbb{R}$.

ניסיון לכל A בורל, $\mu_n(A) \to \mu(A)$ (התכנסות חזקה - שם לא קנוני).

התכנסות זו מאוד מגבילה את אילו סדרות יתכנסו ולכן לא תהיה שימושית.

דוגמא ניקח (X_n) משתנה מקרי, $\mathbb{P}\left(X_n = \frac{1}{n}\right) = 1$. טבעי שהסדרה הזו תתכנס למשתנה מקרי X_∞ שמקיים $\mathbb{P}(X_n = 0) = 1$. זה לא מקיים את ההגדרה הניסיונית שלנו מלמעלה - למשל כי

$$\mathbb{P}(X_n = 0) = 0 \not\to \mathbb{P}(X_\infty = 0) = 1$$

ניסיון נרצה שלכל $t \in \mathbb{R}$ יתקיים $\mathbb{P}(X_n \le t) \to \mathbb{P}(X \le t)$.

רעיון זה עדיין לא תופס את הדוגמא כי $\mathbb{P}(X_n \le 0) = 0 \not\to 1 = \mathbb{P}(X_\infty \le 0)$.

הגדרה 2.6 סדרת מידות הסתברות μ_n מתכנסת בהתפלגות למידת הסתברות μ אם לכל פונקציה רציפה וחסומה f מתקיים

$$\int f(x)\,\mathrm{d}\mu_n(x) \to \int f(x)\,\mathrm{d}\mu(x)$$

הערה 2.7 זוהי הגדרה טובה להתכנסות של מידות הסתברות לא רק על $\mathbb{R}$ אלא גם על כל מרחב פולני (מטרי, ספרבילי ושלם) - למשל $\mathbb{R}^n$, $\mathbb{R}^{\aleph_0}$, $[0,1]^\infty$.

הגדרה 2.8 (ששקולה לקודמת במקרה של הישר הממשי) סדרת μ_n מ תכנסת בהתפלגות אל μ אם לכל $t \in \mathbb{R}$ עם $\mu(t) = 0$ מתקיים

$$\mu_n((-\infty, t]) \to \mu((-\infty, t])$$