

הסתברות למתמטיקאים

© ארזים

21 ביוני 2017

1 התכנסות בהתפלגות

תהי (μ_n) סדרת מידות הסתברות על X , ותהי μ מידת הסתברות נוספת על X .

הגדרה 1.1 נאמר כי μ_n מתכנסות בהתפלגות אל μ (ומסמנים $\mu_n \xrightarrow{d} \mu$ או $\mu_n \xrightarrow{w} \mu$) אם לכל פונקציה $f: X \rightarrow \mathbb{R}$ רציפה וחסומה מתקיים

$$\int f(x) d\mu_n(x) \rightarrow \int f(x) d\mu(x)$$

לעיתים נסמן $\mu_n(f) \rightarrow \mu(f)$.

הזכרנו שההגדרה הזו טובה במרחבים פולניים, שהם מרחבים מטריים ספרביליים ושלמים.

דוגמאות $\mathbb{R}, \mathbb{R}^n, \mathbb{R}^\infty, C[0, 1]$ - המטריקה כאן היא התכנסות במידה שווה, משמע

$$d(f, g) = \max_{x \in [0, 1]} |f(x) - g(x)|$$

במרחב זה, למשל, הילוך מקרי פשוט בעל n צעדים מתכנס לתנועת בראון.

מעכשיו, כל המידות שלנו יהיו על $\mathbb{R}$.

הגדרה 1.2 יהיו $\mu, (\mu_n)$ מידות הסתברות על $\mathbb{R}$. אזי $\mu_n \xrightarrow{d} \mu$ אם ורק אם לכל t עם $\mu(\{t\}) = 0$, מתקיים

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

הערה 1.3 אם $\mu(\{t\}) = 0$ נאמר כי t אינה אטום של μ .

נוכיח תיכף שקילות בין שתי ההגדרות, אבל לפני כן ניתן טענה שתסייע בהוכחה.

טענה 1.4 אם $X_n \rightarrow X$ כמעט תמיד, אזי $X_n \xrightarrow{d} X$ - כלומר $\mathcal{L}(X_n) \xrightarrow{d} \mathcal{L}(X)$. למעשה, אפילו אם יש התכנסות בהסתברות, $X_n \xrightarrow{\mathbb{P}} X$, אזי $X_n \xrightarrow{d} X$.

הוכחה: נוכיח את החלק על התכנסות כמעט תמיד - החלק השני יישאר כתרגיל. תהי f פונקצייה רציפה וחסומה. צריך להראות

$$\mathbb{E}(f(X_n)) \rightarrow \mathbb{E}(f(X))$$

■

זוה מתקיים ממשפט ההתכנסות החסומה.

הוכחה: (של שקילות ההגדרות) ראשית, נניח התכנסות לפי ההגדרה הראשונה (עם הפונקציות הרציפות והחסומות). נקבע t , ויהי $\delta > 0$. נגדיר $f_{t,\delta}$ קירוב רציף של האינדקטור של הקטע $(-\infty, t]$, שהוא 1 על הקטע, יורד לינארית אל 0 בנקודה $t + \delta$, וממשיך משם על אפס. מתקיים $\mu_n(f_{t,\delta}) \rightarrow \mu(f_{t,\delta})$ כי זו פונקצייה רציפה. כמו כן,

$$\begin{aligned}\mu_n(f_{t,\delta}) &\geq \mu_n((-\infty, t]) \\ \mu_n(f_{t,\delta}) &\geq \mu_n((-\infty, t + \delta])\end{aligned}$$

כלומר,

$$\limsup_{n \rightarrow \infty} \mu_n((-\infty, t]) \leq \mu((-\infty, t + \delta])$$

כעת, מתקיים

$$\mu((-\infty, t + \delta]) \xrightarrow{\delta \rightarrow 0} \mu((-\infty, t])$$

ולכן נקבל גם

$$\limsup_{n \rightarrow \infty} \mu_n((-\infty, t]) \leq \mu((-\infty, t])$$

בשביל האי שוויון השני, ניקח את $f_{t-\delta,\delta} = g_{t,\delta}$ - קירוב רציף של האינדקטור של הקטע $(-\infty, t - \delta]$, שהוא 1 על הקטע, יורד לינארית לאפס בנקודה t , וממשיך משם לאפס. כעת,

$$\begin{aligned}\mu_n(g_{t,\delta}) &\leq \mu_n((-\infty, t]) \\ \mu(g_{t,\delta}) &\geq \mu((-\infty, t - \delta])\end{aligned}$$

ולכן

$$\liminf_{n \rightarrow \infty} \mu_n((-\infty, t]) \geq \mu((-\infty, t - \delta])$$

ושוב נוכל לקחת גבול ולקבל

$$\liminf_{n \rightarrow \infty} \mu_n((-\infty, t]) \geq \mu((-\infty, t))$$

נובע שאם $\mu(\{t\}) = 0$ אזי $\mu((-\infty, t]) \rightarrow \mu((-\infty, t])$ כמו שרצינו.
 בכיוון השני, נניח את ההגדרה השנייה (עם הקטעים). נוכיח באופן הבא את ההגדרה הראשונה - נבנה סדרת משתנים מקריים $X, (X_n)$ על מרחב ההסתברות $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$ עם $\mathcal{L}(X_n) = \mu_n, \mathcal{L}(X) = \mu$ ועם $X_n \rightarrow X$ כמעט תמיד. נשתמש בבניית סקורוחוד:

$$X(\omega) = \inf \{z \mid \mu((-\infty, z]) > \omega\}$$

$$X_n(\omega) = \inf \{z \mid \mu_n((-\infty, z]) > \omega\}$$

למדנו שמתקיים $\mathcal{L}(X_n) = \mu_n, \mathcal{L}(X) = \mu$. נקבע ω ונדון בקשר בין $X_n(\omega)$ לבין $X(\omega)$. יהי $z > X(\omega)$. מההגדרה, מתקיים $\mu((-\infty, z]) > \omega$. נניח כרגע כי $\mu(\{z\}) = 0$. אזי

$$\mu_n((-\infty, z]) \rightarrow \mu((-\infty, z])$$

ולכן עבור n גדול מספיק נקבל $\mu_n((-\infty, z]) > \omega$, משמע $z \geq X_n(\omega)$. כלומר, לכל $z > X(\omega)$ שאינו אטום של μ , מתקיים

$$\limsup_{n \rightarrow \infty} X_n(\omega) \leq z$$

נובע מכאן כי

$$\limsup_{n \rightarrow \infty} X_n(\omega) \leq X(\omega)$$

אחרת, היה בין שני הערכים האלה קטע, שמתוכו יש נקודה שאינה אטום של μ (כי יש רק כמות בת מניה של אטומים), וזו תהיה סתירה לאי שוויון הקודם. בשביל אי השוויון השני, נשתמש בחלק השני של בניית סקורוחוד:

$$X^-(\omega) = \inf \{z \mid \mu((-\infty, z]) \geq \omega\}$$

$$X_n^-(\omega) = \inf \{z \mid \mu_n((-\infty, z]) \geq \omega\}$$

אזי

$$\mathbb{P}(X_n^- = X_n) = 1 = \mathbb{P}(X^- = X)$$

וכן $X_n(\omega) \geq X_n^-(\omega)$ לכל $(\omega \in [0, 1])$. באותו אופן נקבל את אי השוויון

$$\liminf_{n \rightarrow \infty} X_n^-(\omega) \geq X(\omega)$$

ומכאן

$$X^-(\omega) \leq \liminf_n X_n^-(\omega) \leq \limsup_n X_n(\omega) \leq X(\omega)$$

וזאת לכל $\omega \in [0, 1]$ מסדוויץ' נקבל שלכל ω שבו $X^-(\omega) = X(\omega)$ מתקיים $X_n(\omega) \rightarrow X(\omega)$.
 ■

1.1 קומפקטיות סדרתית של מידות על $\mathbb{R}$

נתונה סדרת מידות הסתברות (μ_n) על $\mathbb{R}$. נרצה לדעת האם בהכרח קיימת מידת הסתברות μ ותת סדרה (n_k) עם $\mu_{n_k} \xrightarrow{d} \mu$.

דוגמא לכך שזה לא תמיד המצב. נבחר $\mu_n = \delta_n$, המידה שנותנת 1 לאטום $\{n\}$. אין לסדרה זו תת סדרה מתכנסת, כי אם הייתה, הגבול היה מקיים $\mu((-\infty, t]) = 0$ לכל t .

למה 1.5 (הלייברי, Helly – Bray) תהי (μ_n) סדרת מידות הסתברות על $\mathbb{R}$. אזי קיימת פונקציה $F : \mathbb{R} \rightarrow [0, 1]$ שאינה יורדת ורציפה מימין, וקיימת תת סדרה (n_k) כך שלכל t שבו F רציפה מתקיים

$$\mu_{n_k}((-\infty, t]) \rightarrow F(t)$$

הוכחה: נבחר תת קבוצה צפופה ובת מניה בתוך $\mathbb{R}$ - נקרא לה $C = \{c_i\}_{i=1}^\infty$. כעת, עבור c_1 , תהי (n_k^1) תת סדרה שעליה $\mu_{n_k^1}((-\infty, c_1])$ מתכנס. נבצע אותו דבר עבור $c_2, c_3, \dots$ ניקח תת סדרה (n_k^2) של (n_k^1) . נמשיך כך - לכל c_j תהיה לנו תת סדרה (n_k^j) של התת סדרה הקודמת. נתבונן כעת בתת הסדרה $(n_k^k) = (n_k)$. על תת סדרה זו,

$$\mu_{n_k^k}((-\infty, c_j]) \rightarrow H(c_j)$$

לכל j , כאשר $H : C \rightarrow [0, 1]$. נבחין גם כי H לא יורדת. כעת, נגדיר לכל $t \in \mathbb{R}$ נגדיר

$$F(t) = \lim_{c \downarrow t} H(c)$$

משמעות הסימון היא שהגבול נלקח על פני $c \in C$ כן $t < c$. נובע כי F לא יורדת, מקבלת ערכים בתוך $[0, 1]$ ורציפה מימין. כמו כן, אם F רציפה בנקודה t אזי

$$\mu_{n_k^k}((-\infty, t]) \rightarrow F(t)$$

■

הערה 1.6 אם במקרה

$$\begin{aligned} \lim_{t \downarrow -\infty} F(t) &= 0 \\ \lim_{t \uparrow \infty} F(t) &= 1 \end{aligned}$$

אזי קיימת מידת הסתברות μ על $\mathbb{R}$ עם $F(t) = \mu((-\infty, t])$.

1.1.1 הדיקות

הגדרה 1.7 נאמר שסדרת מידות הסתברות (μ_n) על $\mathbb{R}$ היא הדוקה לכל $\varepsilon > 0$ קיים $K < \infty$ עם

$$\mu_n([-K, K]) \geq 1 - \varepsilon$$

לכל n .

מסקנה 1.8 תהי (μ_n) סדרת מידות הסתברות.

1. אם (μ_n) הדוקה אזי קיימת (n_k) ומידת הסתברות μ עם $\mu_{n_k} \xrightarrow{d} \mu$.

2. אם קיימת מידת הסתברות μ עם $\mu_n \xrightarrow{d} \mu$ אזי (μ_n) .

הוכחה:

1. ניקח את F ואת (n_k) מלמת הלי-בריי. ההדיקות מבטיחה שתנאי ההערה האחרונה מתקיימים, ולכן יש μ שכזו.

2. יהי $\varepsilon > 0$, ויהי $K < \infty$ עם

$$\mu((-K, K]) \geq 1 - \varepsilon$$

ונניח גם כי $K, -K$ אינם אטומים של μ . כעת, ידוע כי

$$\begin{aligned}\mu_n((-\infty, K]) &\rightarrow \mu((-\infty, K]) \\ \mu_n((-\infty, -K]) &\rightarrow \mu((-\infty, -K])\end{aligned}$$

ולכן

$$\mu_n((-K, K]) \rightarrow \mu((-K, K]) \geq 1 - \varepsilon$$

ולכן עבור n גדול מספיק נקבל $\mu_n \geq 1 - 2\varepsilon$. נותרו מספר סופי של ערכי n , ונוכל להגדיל את K כדי לטפל בהם.

■

2 משפט הגבול המרכזי

ניזכר בפונקציה האופיינית φ_X של משתנה מקרי X :

$$\begin{aligned}\varphi_X : \mathbb{R} &\rightarrow \mathbb{C} \\ \varphi_X(\theta) &= \mathbb{E}(e^{i\theta X})\end{aligned}$$

הפונקציות $\operatorname{Re}(e^{i\theta X}), \operatorname{Im}(e^{i\theta X})$ הן פונקציות רציפות וחסומות על $\mathbb{R}$, ולכן נובע שאם $X_n \xrightarrow{d} X$ אזי $\varphi_{X_n}(\theta) \rightarrow \varphi_X(\theta)$ לכל θ .

משפט 2.1 (ההתכנסות של לוי) תהי (μ_n) סדרת מידות הסתברות,

$$\varphi_{\mu_n}(\theta) = \mu_n(e^{i\theta X}) = \int_{-\infty}^{\infty} e^{i\theta x} d\mu_n(x)$$

נניח שקיימת $g: \mathbb{R} \rightarrow \mathbb{C}$ עם $\varphi_{\mu_n}(\theta) \rightarrow g(\theta)$ לכל θ שהיא גם רציפה בנקודה $\theta = 0$, אזי קיימת מידת הסתברות μ עם $g = \varphi_\mu$ וכן $\mu_n \xrightarrow{d} \mu$.

נוכיח את המשפט בעתיד (בשיעור הבא). בינתיים נרצה להראות למה צריך רציפות באפס.

דוגמא ניקח $\mu_n = U[-n, n]$. אזי

$$\varphi_{\mu_n}(\theta) = \int_{-n}^n \frac{1}{2n} e^{i\theta x} dx = \frac{e^{i\theta n} - e^{-i\theta n}}{2ni\theta} = \frac{\sin(\theta n)}{\theta n}, \theta \neq 0$$

$$\varphi_{\mu_n}(0) = 1$$

מתקיים

$$\varphi_{\mu_n}(\theta) \xrightarrow{n \rightarrow \infty} \begin{cases} 1 & \theta = 0 \\ 0 & \theta \neq 0 \end{cases}$$

נבחין שהגבול אינו רציף בנקודה 0, ואכן (μ_n) לא מתכנסת בהתפלגות.

הערה 2.2 רציפות של הגבול בנקודה $\theta = 0$ גוררת הדיקות של μ_n - נראה בהוכחה.

2.1 הקשר בין המומנטים של μ והפונקציה האופיינית ליד 0

נזכר כי

$$e^{i\theta x} = \sum_{k=0}^{\infty} \frac{(i\theta x)^k}{k!}$$

לכן עבור משתנה מקרי X אשר לקוות שיתקיים

$$\varphi_X(\theta) = \mathbb{E}(e^{i\theta X}) = \sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!} \mathbb{E}(X^k)$$

כדי להצדיק פיתוח מסוג זה, נעריך קירובים של $e^{i\theta x}$. נגדיר

$$R_n(x) = e^{ix} - \sum_{k=0}^n \frac{(ix)^k}{k!}$$

נרצה לבדוק שזו קטנה. עבור $n = 0$ נקבל

$$R_0(x) = e^{ix} - 1 = \int_0^x ie^{it} dt$$

ולכן

$$|R_0(x)| \leq \min(2, |x|)$$

עבור $n \geq 1$, נבין כי

$$R_n(x) = i \int_0^x R_{n-1}(t) dt$$

אפשר לבדוק זאת בקלות. לכן נקבל

$$|R_n(x)| \leq \int_0^x |R_{n-1}(t)| dt$$

באינדוקציה נקבל

$$|R_n(x)| \leq \min\left(\frac{2|x|^n}{n!}, \frac{|x|^{n+1}}{(n+1)!}\right)$$

בפרט,

$$\left| e^{i\theta x} - \sum_{k=0}^n \frac{(i\theta x)^k}{k!} \right| \leq \min\left(\frac{2|\theta x|^n}{n!}, \frac{|\theta x|^{n+1}}{(n+1)!}\right)$$

עבור $n = 2$:

$$R_2(\theta x) \leq |\theta x|^2 \min\left(1, \frac{|\theta x|}{6}\right)$$

כעת, אם X משתנה מקרי עם $\mathbb{E}(X) = 0$, $\text{Var}(X) < \infty$, אזי

$$\begin{aligned} |\mathbb{E}(R_2(\theta X))| &= \left| \mathbb{E}\left(e^{i\theta X} - \left(1 + i\theta X - \frac{\theta^2 X^2}{2}\right)\right) \right| = \\ &= \left| \mathbb{E}\left(e^{i\theta X}\right) - \left(1 - \frac{\theta^2 \text{Var}(X)}{2}\right) \right| \leq \\ &\leq \mathbb{E}\left(\left|e^{i\theta X} - \left(1 - \frac{\theta^2 \text{Var}(X)}{2}\right)\right|\right) \leq \\ &\leq \theta^2 \min\left(\mathbb{E}X^2, \frac{\theta \mathbb{E}|X|^3}{6}\right) \end{aligned}$$

כמו כן, נקבל

$$\mathbb{E}e^{i\theta x} = 1 - \frac{\theta^2}{2} \text{Var}(X) + o(\theta^2)$$

נראה זאת:

$$\left| \frac{e^{i\theta x} - \left(1 - \frac{\theta^2}{2} \text{Var}(X)\right)}{\theta^2} \right| \leq \min \left(|X|^2, \frac{\theta |X|^3}{6} \right)$$

נובע שאגף שמאל שואף לאפס נקודתית כאשר $\theta \rightarrow 0$. בנוסף, אגף שמאל חסום על ידי X^2 , שהוא אינטגרבילי, ולכן נסיק מהתכנסות נשלטת:

$$\mathbb{E} \left(\left| \frac{e^{i\theta X} - \left(1 - \frac{\theta^2}{2} \text{Var}(X)\right)}{\theta^2} \right| \right) \xrightarrow{\theta \rightarrow 0} 0$$

למה 2.3 (טכנית) עבור $|z| < \frac{1}{2}$ מתקיים

$$|\log(1+z) - z| \leq z^2$$

הוכחה: מתקיים

$$\log(1+z) - z = \int_0^z \frac{-w}{1+w} dw = -z^2 \int_0^1 \frac{t}{1+tz} dt$$

■

את הסוף נשאיר כתרגיל.

משפט 2.4 (הגבול המרכזי) תהי (X_n) סדרת משתנים מקריים בלתי תלויים, שווי התפלגות ובעלי $\text{Var}(X_1) = 1$, $\mathbb{E}X_1 = 0$. אזי

$$\hat{S}_n = \frac{X_1 + \dots + X_n}{\sqrt{n}} \xrightarrow{d} N(0, 1)$$

באופן שקול, לכל $a < b$ מתקיים

$$\mathbb{P} \left(a < \frac{X_1 + \dots + X_n}{\sqrt{n}} < b \right) \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

הוכחה: נסתמך על משפט ההתכנסות של לוי.

$$\begin{aligned} \varphi_{\hat{S}_n}(\theta) &= \mathbb{E} \left(e^{i\theta \hat{S}_n} \right) = \mathbb{E} \left(e^{i\theta \frac{X_1 + \dots + X_n}{\sqrt{n}}} \right) = \prod_{j=1}^n \mathbb{E} \left(e^{i \frac{\theta}{\sqrt{n}} X_j} \right) = \\ &= \prod_{j=1}^n \varphi_{X_j} \left(\frac{\theta}{\sqrt{n}} \right) = \varphi_{X_1} \left(\frac{\theta}{\sqrt{n}} \right)^n = \left(1 - \frac{\theta^2}{2n} + \frac{1}{n} \cdot o(\theta^2) \right)^n \end{aligned}$$

נותר לנו להסיק שהביטוי בסוף מתכנס אל $e^{-\frac{\theta^2}{2}}$, כי זו רציפה באפס, ונקבל את הנדרש ממשפט ההתכנסות של לוי, כי

$$\varphi_{N[0,1]}(\theta) = e^{-\frac{\theta^2}{2}}$$

כדי להראות את ההתכנסות ניקח לוגריתם:

$$\log [\varphi_{\hat{S}_n}(\theta)] = n \log \left(1 - \frac{\theta^2}{2n} + o\left(\frac{\theta^2}{n}\right) \right)$$

כאשר n גדול מספיק, מתקיים

$$\left| -\frac{\theta^2}{2n} + o\left(\frac{\theta^2}{n}\right) \right| \leq \frac{1}{2}$$

לכן

$$\left| \log [\varphi_{\hat{S}_n}(\theta)] - n \left(-\frac{\theta^2}{2n} + o\left(\frac{\theta^2}{n}\right) \right) \right| \leq n \cdot O\left(\frac{C(\theta)}{n}\right)^2$$

כאשר $C(\theta)$ היא החזקה המתאימה של θ (אפשר לבדוק ולראות מה היא). כאשר $n \rightarrow \infty$, נקבל שאגף ימין מתכנס לאפס, וסיימנו. ■