

הסתברות למתמטיקאים – הרצאה 3

28.02.11

דיברנו על מספרים נורמליים ועל טור עם סימנים מתחלפים. האם יש קשר בין השניים?
 $x = (0.\beta_1\beta_2\dots)_2$ נורמלי האם בהכרח $\sum_n \frac{(2\beta_n-1)}{n}$? אף אחד לא יודע. כנראה שאפשר לבנות דוגמא נגדית עבור 2-נורמליות אבל הקשר למספרים נורמליים? כנראה שאף אחד לא יודע.
 טענה באנליזה:

טענה. תהי $f: \mathbb{R} \rightarrow \mathbb{R}$ פונקציה מדידה. (מידת לגאס זה לא ברור). ועקיימת

$$f(x + 2^{-n}) = f(x)$$

עבור כמעט כל $x \in \mathbb{R}$ ו- $n = 1, 2, 3, \dots$ אז f קבועה כמעט בכל מקום.
 כלומר

$$\exists a \in \mathbb{R} \quad f(x) = a \quad \text{למעט כל } x$$

הערה. זה לא משנה אם לכל n בנפרד או לכל n יחד כי סה"כ יש קבוצה זניחה של "מספרים רעים". יש לפונקציה מחזורים קטנים כרצוננו.

לא קבועה בכל מקום כי לא דרשנו רציפות. האם בכלל יש פונקציות כאלה? כן, לדוגמא פונקציות דיריכלה. כעת, מדוע הבאנו משפט באנליזה לתורת ההסתברות?

טענה. יהיו $X_1, X_2, \dots$ סימנים מקריים (מקבלים ערכים 0 או 1 בהסתברות חצי) וגם Y משתנה מקרי בצורה לכל $n: Y = f_n(X_n, X_{n+1}, \dots)$ כ"כ.
 אז Y קבוע. כלומר, קיים $a \in \mathbb{R}$ כך ש- $\mathbb{P}(Y = a) = 1$.

איך זה קשור לטענה באנליזה? $\Omega = [0, 1], Y = f: \Omega \rightarrow \mathbb{R}, X_k = \beta_k$ ספרות בינאריות אז התנאי מתקיים מפני שמחזוריות כזאת של פונקציה אומרת שהיא לא תלויה בספרות הבינאריות הראשונות. להוסיף חצי – לשנות ספרה ראשונה, להוסיף 2^{-n} זה כמו לשנות n ספרות בינאריות ראשונות. לכן Y היא פונקציה של זנב.
 שוב מקרה פרטי של חוק 1/0 של קולמוגורוב.

סדרות אינסופיות של מ"מ ב"ת

הגדרה אנחנו כבר מכירים.

משפט. החוק החזק של המספרים הגדולים
 יהיו $X_1, X_2, \dots$ מ"מ ב"ת ש"ה $\mathbb{E}|X_1| < \infty$. אז סכום

$$\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{\text{ב"צ}} \mathbb{E}X_1$$

כמובן שזה הכללה למה שראינו ששיחות מתכנסות להסתברות. זה לא פיקניק להוכיח את זה. מתישהו נצטרך להתמודד עם זה ומזל שעכשיו זה רק סקירה.

משפט. צורה חזקה של גליונקו – קנטלי

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow[n \rightarrow \infty]{\text{ב"ב}} 0$$

כאשר F פונקציית הצטברות ו- F_n פונקציית הצטברות אמפירית.

כדאי להזהר כאן. x לא רץ כאן על סדרה בת מניה. x רץ על רצף. אפילו התכנסות נקודתית כאן זה בעיה. זוכרים את התכנסות טור עם סימנים משתנים? סוף סוף הגענו למשפט המתאים:

משפט. קולמגורוב

יהיו $X_1, X_2, \dots$ מ"מ ב"ת (לאו דווקא ש"ה). $\mathbb{E}X_n = 0$. $\forall n$ נניח ש $\sum_{n=1}^{\infty} \mathbb{E}(X_n^2) < \infty$. אז:

$$\sum_{n=1}^{\infty} X_n$$

מתכנס כ"ב.

כאן עצם העובדה שהסכום קטן מאינסוף לא נותנת שום דבר כי הסימנים מתחלפים. לא לכתוב דברים כאלה במבחן. הרי אף אחד לא אומר שסכום הטור $-1, 1, -1, 1, -1, \dots$ מתכנס רק כי קטן מאינסוף נכון? ואנחנו גם אף פעם לא נשתמש במשפטים מחדו"א על טורים אי-שליליים כאשר לא באמת בדקנו שכל האיברים הם אי-שליליים? נכון?

משפט. חוק 1,0 של קולמגורוב

יהיו $X_1, X_2, \dots$ מ"מ ב"ת (אולי שוב כדאי להזכיר שלא דווקא ש"ה) ו- Y מ"מ מהצורה $Y = f_n(X_n, X_{n+1}, \dots)$ אז Y הוא קבוע כ"ב. כלומר:

$$\exists a \quad \mathbb{P}(Y = a) = 1$$

כמה תוצאות של זה:

תוצאה. יהיו $X_1, X_2, \dots$ מ"מ ב"ת, $S_n = X_1 + \dots + X_n$ אז

1. ההסתברות ש- S_n חסומה היא 0 או 1. לא יכול להיות סיכוי חצי, לא יכול להיות סיכוי שליש או כמעט תמיד חסומה או כמעט תמיד לא חסומה. אם X איכשהו גדולים אז כנראה לא חסומה אם X איכשהו לא גדולים אז כנראה חסומה.

2. ההסתברות ש- S_n מתכנסת (כלומר ש- $\sum X_n$ מתכנס) גם היא 0 או 1. גם פה. בלי שטויות. או כ"ב שמתכנס או כ"ב שמתבדר. תכונה איכותית תמיד מנוונת.

3. $\limsup_n S_n < \infty$ כלומר האם הסדרה חסומה מלמעלה. נחשו מה? ההסתברות היא 0 או 1.

4. $\liminf_n S_n > -\infty$ ברור שגם או 0 או 1.

למה אלה דוגמאות? נגדיר Y להיות אינדיקטור של מאורע (נגדיר 1 אם הטור מתכנס ו-0 כשמתבדר) ואז הטור מתכנס אם הזנב מתכנס ואכן אפשר לבטא Y כפונקציה של הזנב. כל אלה למעשה תכונות זנביות ולכן האינדיקטורים שלהם הם מ"מ זנביים. איך אינדיקטור יכול להיות קבוע? הוא 0 או 1.

סוף התצוגה.

אי תלות – הוכחות

$S_n = X_1 + \dots + X_n$ הילוך מקרי חד ממדי פשוט. $X_1, \dots, X_n$ מ"מ ב"ת ש"ה $(1/2, 1/2) = \pm 1$. ידועים תחת השם סימנים מקריים. נראה:

$$\mathbb{P}\left(\frac{1}{n}|S_n| \leq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 1$$

באופן היסטורי הוכיחו את זה ישירות, עם התפלגות בינומית אבל זאת עבודה קשה אנליטית. קיצור דרך שלא צריך לעבוד קשה איתו הוא אי-שויונות של מרקוב וצ'בישב. עושים גיאומטריה קלה ולא צריך לעבוד אנליטית. D3ERUTCIP D2ERUTCIP מרחק מראשית הצירים עד אמצע משולש שווה שוקיים 1 ישר זווית. המרחק שואף ל-0 כשמימד עולה לאינסוף. איך זה קשור לתורת ההסתברות? יש לפנינו $\Omega = \{-1, 1\}^n$; בסטאפ הזה מ"מ הוא $X : \{-1, 1\}^n \rightarrow \mathbb{R}$. מרחב כל המ"מ הוא למעשה מרחב אוקלידי!

$$\begin{aligned} X \perp Y &\Leftrightarrow \langle X, Y \rangle = 0 \Leftrightarrow \int_{\Omega} XY dP = 0 \Leftrightarrow \\ &\Leftrightarrow \frac{1}{2^n} \sum_{\omega \in \Omega} X(\omega)Y(\omega) = 0 \Leftrightarrow \mathbb{E}(XY) = 0 \end{aligned}$$

המ"מ X שלנו הם אורתוגונליים.

$$\mathbb{E}(X_1 X_2) = (+1) \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$$

ובאותו אופן $X_k \perp X_l : 1 \leq k < l \leq n$.
לכן ממשפט פיתגורס:

$$\|X_1 + X_2\|^2 = \|X_1\|^2 + \|X_2\|^2 = 2$$

לכן $\|S_2\|^2 = 2$.
d2 ERUTCIP
באותו חישוב:

$$\|S_n\| = \sqrt{n}$$

$$\left\| \frac{S_n}{n} \right\| = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0$$

אבל זאת התכנסות ב- L_2 ואנחנו רוצים התכנסות בהסתברות. נשתמש באי-שוויון מרקוב:

$$\mathbb{P}\left(\frac{1}{n}|S_n| \geq \epsilon\right) \leq \frac{\mathbb{E}\left(\frac{1}{n}|S_n|^2\right)}{\epsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

$$\frac{\mathbb{E}\left(\frac{1}{n}|S_n|^2\right)}{\epsilon^2} = \frac{\|\frac{1}{n}S_n^2\|_2^2}{\epsilon^2}$$

כלומר ההסתברות להיות מחוץ לסימפלקס שואפת ל-0 לכן ההסתברות להיות בתוך הסימפלקס שואפת ל-1.
מה לגבי אי שוויון מרקוב?

$$\epsilon^2 \mathbb{1}_{Y^2 \geq \epsilon^2} \leq Y^2$$

$$\mathbb{E}(\epsilon^2 \mathbb{1}_{Y^2 \geq \epsilon^2}) \leq \mathbb{E}(Y^2)$$

$$\mathbb{E}(\epsilon^2 \mathbb{1}_{Y^2 \geq \epsilon^2}) = \epsilon^2 \mathbb{P}(Y^2 \geq \epsilon^2) \leq \mathbb{E}(Y^2)$$

$$\mathbb{P}(|Y| \geq \epsilon) = \mathbb{P}(Y^2 \geq \epsilon^2)$$

$$\mathbb{P}(|Y| \geq \epsilon) \leq \frac{\mathbb{E}Y^2}{\epsilon^2}$$

כלומר התכנסות ב- L_2 חזקה מהתכנסות בהסתברות.

קירוב להתפלגות בינומית

המפתח הוא נוסחאת סטירלינג $n! = n^n e^{-n} \sqrt{2\pi n} \beta(n)$, $\beta(n) \xrightarrow{n \rightarrow \infty} 1$. בלי הוכחה. הוכחנו באנליזה. ידוע אפילו $\beta(n) = 1 + O(\frac{1}{n})$ ובזה נשתמש. הערה. אפילו ידוע ש- $\beta(n) = \frac{1}{12n}$ אבל זה יותר מדי חדואיסיטי ולא נשתמש.

$$\ln n! = n \ln n - n + \frac{1}{2} \ln n + \frac{1}{2} \ln(2\pi) + \underbrace{\ln \beta(n)}_{O(\frac{1}{n})}$$

$$\mathbb{P}(S_n = k) = 2^{-n} \binom{n}{\frac{n+k}{2}} = 2^{-n} \frac{n!}{(\frac{n-k}{2})! (\frac{n+k}{2})!}$$

$n+k$ זוגי. אחרת זה 0.

עבודה פיזית קשה: נסמן, $k = cn$

$$\begin{aligned}
\ln \mathbb{P}(S_n = k) &= \\
&= -n \ln 2 + n \ln n - \frac{n-k}{2} \ln \frac{n-k}{2} - \frac{n+k}{2} \ln \frac{n+k}{2} + \\
&+ \underbrace{1 - n + \frac{n-k}{2} + \frac{n+k}{2}}_{=0} + \frac{1}{2} \ln n - \frac{1}{2} \ln \frac{n-k}{2} - \frac{1}{2} \ln \frac{n+k}{2} - \frac{1}{2} \ln(2\pi) + \\
&\quad + \ln \beta(n) - \ln \beta\left(\frac{n-k}{2}\right) - \ln \beta\left(\frac{n+k}{2}\right) = \\
&= -n \ln 2 + n \ln n - \frac{n(1-c)}{2} \ln \frac{n(1-c)}{2} - \frac{n(1+c)}{2} \ln \frac{n(1+c)}{2} + \\
&\quad + \frac{1}{2} \ln n - \frac{1}{2} \ln \frac{n(1-c)}{2} - \frac{1}{2} \ln \frac{n(1+c)}{2} - \frac{1}{2} \ln(2\pi) + \\
&\quad + \ln \beta(n) - \ln \beta\left(\frac{n(1-c)}{2}\right) - \ln \beta\left(\frac{n(1+c)}{2}\right) = \\
&= -n \ln 2 + n \ln n - \left(\frac{n(1-c)}{2} + \frac{n(1+c)}{2}\right) \ln n + \\
&\quad - \frac{n(1-c)}{2} \ln \frac{n(1-c)}{2} - \frac{n(1+c)}{2} \ln \frac{n(1+c)}{2} + \\
&\quad - \frac{1}{2} \ln n - \frac{1}{2} \ln \frac{n(1-c)}{2} - \frac{1}{2} \ln \frac{n(1+c)}{2} - \frac{1}{2} \ln(2\pi) + \\
&\quad + \ln \beta(n) - \ln \beta\left(\frac{n(1-c)}{2}\right) - \ln \beta\left(\frac{n(1+c)}{2}\right) = \\
&= -n \ln 2 - n \underbrace{\left(\frac{1-c}{2} \ln(1-c) + \frac{1+c}{2} \ln(1+c)\right)}_{\gamma(c)} + \\
&\quad + n \left(\frac{1-c}{2} \ln 2 - \frac{1+c}{2} \ln 2\right) - \frac{1}{2} \ln n - \frac{1}{2} \ln \frac{1-c^2}{4} - \frac{1}{2} \ln(2\pi) + \\
&\quad + \ln \beta(n) - \ln \beta\left(\frac{n(1-c)}{2}\right) - \ln \beta\left(\frac{n(1+c)}{2}\right) = \\
&= -n\gamma(c) - \frac{1}{2} \ln n - \frac{1}{2} \ln \frac{1-c^2}{4} - \frac{1}{2} \ln(2\pi) + \\
&\quad + \ln \beta(n) - \ln \beta\left(\frac{n(1-c)}{2}\right) - \ln \beta\left(\frac{n(1+c)}{2}\right)
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(S_n = cn) &= e^{-n\gamma(c)} \frac{1}{\sqrt{n}} \frac{2}{\sqrt{1-c^2}} \frac{1}{\sqrt{2\pi}} \cdot e^{\beta(n) - \beta\left(\frac{n(1-c)}{2}\right) - \beta\left(\frac{n(1+c)}{2}\right)} \approx \\
&\approx \frac{2}{\sqrt{2\pi n}} \frac{1}{\sqrt{1-c^2}} e^{-n\gamma(c)}
\end{aligned}$$

מתי זה קירוב טוב? מה עם ה- β שנאשרו? נרצה שהביטוי הבא יהיה קטן:

$$\ln \beta(n) - \ln \beta\left(\frac{n(1-c)}{2}\right) + \ln \beta\left(\frac{n(1+c)}{2}\right)$$

כלומר ש- $n \gg 1$ וגם $n(1-c) \gg 1$ וגם $n(1+c) \gg 1$ כלומר ש- $n(1-|c|) \gg 1$. מצד אחד n צריך להיות גדול ומצד שני ש- c לא יהיה קיצוני מידי. אנחנו יודעים $\ln \beta(n) = O(\frac{1}{n})$ אז $O(\frac{1}{n(1-|c|)})$ תהיה הערכת השגיאה.

דוגמא. $n = 200, c = -0.9, \mathbb{P}(S_{200} = -180) = 1.397 \cdot 10^{-44}$. זה חישוב מדויק. בקירוב נקבל: $1.409 \cdot 10^{-44}$ קירוב טוב בתנאים לא כל כך נוחים. נוסחא ממש משהו. (הקירוב הנורמלי יהיה $3.7 \cdot 10^{-37}$).

עכשיו רוצים להתפצל ולהמשיך לשני דברים: אחת לקירוב נורמלי, משפט דה־מואבר לפלאס. השנייה לסטיות גדולות. קירוב נורמלי: כאשר c הוא קטן, קרובים למרכז ההתפלגות. איך γ מתנהגת בסביבה של 0?

$$\gamma : [-1, 1] \rightarrow \mathbb{R}$$

$$\gamma(-c) = \gamma(c)$$

$$\gamma(0) = 0$$

$$\gamma'(0) = 0$$

$$\gamma''(0) = 1$$

לכן,

$$\gamma(c) \approx \frac{1}{2}c^2$$

$$\gamma'''(0) = 0$$

$$\gamma''''(0) = ?$$

אבל זה מספר סופי. אז איבר הבא בפיתוח הוא רק מסדר 4:

$$\gamma(c) = \frac{1}{2}c^2 + O(c^4)$$

$$\begin{aligned}
& \left| \ln \mathbb{P}(S_n = k) - \ln\left(\frac{2}{\sqrt{2\pi n}}\right) e^{-\frac{k^2}{2n}} \right| \leq \\
& \leq \left| \ln \mathbb{P}(S_n = c_n) - \ln\left(\frac{2}{\sqrt{2\pi n}} \frac{1}{\sqrt{1-c^2}} e^{-n\gamma(c)}\right) \right| + \\
& + \left| \ln\left(\frac{2}{\sqrt{2\pi n}} \frac{1}{\sqrt{1-c^2}} e^{-n\gamma(c)}\right) - \ln\left(\frac{2}{\sqrt{2\pi n}}\right) e^{-\frac{c^2 n}{2}} \right| \leq \\
& \leq O\left(\frac{1}{n(1-|c|)}\right) + \underbrace{\left| -n\gamma(c) + n\frac{c^2}{2} - \frac{1}{2} \ln(1-c^2) \right|}_{=nO(c^4)} = \\
& = O\left(\frac{1}{n(1-|c|)}\right) + nO(c^4) + O(c^2)
\end{aligned}$$

לא כל כך ברור כמה זה טוב וכמה זה אז טריק קטן:

$$O\left(\frac{1}{n(1-|c|)}\right) + O(c^2) + O(nc^4) = O\left(nc^4 + \frac{1}{n}\right)$$

מדוע?
עבור $\frac{1}{n} \leq \delta \leq \frac{1}{4}$ מתקיים $nc^4 + \frac{1}{n} \leq \delta \leq \frac{1}{4}$ ואז $\frac{1}{n} \leq \delta$ וכן $nc^4 \leq \delta$ נסיק: $c^2 \leq \delta$ ומבחירת $\delta: |c| \leq \frac{1}{2}$. במילים אחרות

$$\frac{1}{n(1-|c|)} \leq \frac{2}{n} \leq 2\delta$$

לכן,

$$\begin{aligned}
& \mathbb{P}(S_n = k) \frac{2}{\sqrt{2\pi n}} e^{-\frac{k^2}{2n}} O\left(\frac{k^4}{n^3} + \frac{1}{n}\right) \\
& \cdot \frac{k^4}{n^3} = O\left(\frac{1}{n}\right) \text{ נקבל } |k| = O(\sqrt{n}) \text{ כעת אם}
\end{aligned}$$