

הסתברות למתמטיקאים – הרצאה 1

21.02.11

אי-תלות

סדרות ארוכות של מאורעות בלתי תלויים

מטבע הוגן

נפתח בדוגמא. דוגמא טובה היא הטלת מטבע n פעמים.

$$a_1, \dots, a_n \in \{0, 1\}$$

2^n אפשרויות.

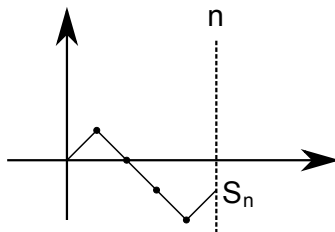
$$p_n(a_1, \dots, a_n) = 2^{-n}$$

ללא שום ספק ההטלות הן בלתי תלויות ונוכל לבחור סדרה ארוכה כרצוננו.
כפי שאולי אנחנו זוכרים ממבוא להסתברות, מדבר בהתפלגות בינומית:

$$\mathbb{P}(a_1 + \dots + a_n = k) = \frac{\binom{n}{k}}{2^n}$$

$$k = 0, 1, \dots, n$$

מטבע הוגן, הטלות בלתי תלויות. נוכל לתאר כהילוך מקרי.



$$S_n = \sum_{k=1}^n (2a_k - 1)$$

בפועל כמובן ההילוך הוא מקבל ערכים 0 או 1. אבל נרצה לקבל הילוך יפה, לדוגמא סביב ציר ה- x . בגלל זה נכתוב קצת אחרת אבל חשוב לזכור שהמשמעות לא משתנה! כעת, עבור $k = -n, -n+2, \dots, n-2, n$:

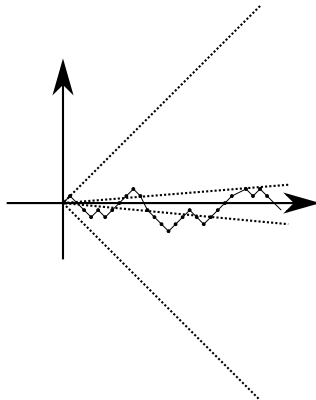
$$\mathbb{P}(S_n = k) = P\left(a_1 + \dots + a_n = \frac{n+k}{2}\right) = 2^{-n} \binom{n}{\frac{n+k}{2}}$$

אמנם אנחנו חכמים ומכירים התפלגות בינומית, נרצה להסתכל איך אפשר לקרב אותה.

טענה. לכל $\varepsilon > 0$:

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

מה זה אומר אם נסתכל על ההילוך המקרי שלנו? זה אומר שככל שנלך יותר זמן, ונסתכל על הישר שמחבר בין הראשית לערך האחרון של S אז הזווית עם הציר האופקי תהיה קטנה, כלומר אחרי הרבה זמן נהיה קרובים לציר.



וזה מקרה פרטי של החוק החלש של המספרים הגדולים.

טענה. כאשר $k+n$ זוגי:

$$\mathbb{P}(S_n = k) = \frac{2}{\sqrt{2\pi n}} e^{-\frac{k^2}{2n}} \left(1 + \alpha_n\left(\frac{k}{\sqrt{n}}\right)\right)$$

$$\alpha_n(\cdot) \xrightarrow{n \rightarrow \infty} 0$$

במידה שווה בכל קטע חסום.

כלומר,

$$\forall M \sup_{[-M, M]} |\alpha_n(\cdot)| \rightarrow 0$$

היה אפשר לכתוב גם מקסימום כי P מוגדר רק על סריג. זה מקרה פרטי של משפט הגבול המקומי. באופן ברור ומובהק $P(S_n = k) = 0$ עבור $k+n$ איזוגי.

מה קורה במקרה $k > n$?

$$0 = \mathbb{P}(S_n = k) = \frac{2}{\sqrt{2\pi n}} e^{-\frac{k^2}{2n}} \left(1 + \alpha_n \left(\frac{k}{\sqrt{n}} \right) \right)$$

ולכן,

$$\alpha_n \left(\frac{k}{\sqrt{n}} \right) = -1$$

אז למה בכלל הרשנו את זה? כי אפילו ל- $k = n$

$$\underbrace{2^{-n}}_{\left(\frac{1}{2}\right)^n} = \underbrace{\frac{2}{\sqrt{2\pi n}} e^{-\frac{n}{2}} (1 + \alpha_n)}_{\left(\frac{1}{\sqrt{e}}\right)^n}$$

וברור שאין שוויון. פה יש בעיה כי ההפרש אקספוננציאלי. לכן הקירוב הנורמלי עובד אבל אי אפשר לדרוש ממנו יותר מדי.

משפט. DeMoivre—Laplace

לכל a, b כך ש- $-\infty \leq a < b \leq +\infty$:

$$\mathbb{P}(a\sqrt{n} < S_n < b\sqrt{n}) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{u^2}{2}} du$$

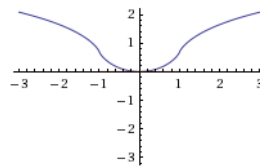
כנראה שזה גם קשור לטענה הקודמת אבל נראה בהמשך.
נדבר על סטיות גדולות:

טענה. לכל $c \in (0, 1)$

$$\frac{1}{n} \ln \mathbb{P}(S_n > cn) \xrightarrow{n \rightarrow \infty} -\gamma(c)$$

כאשר:

$$\gamma(c) = \frac{1}{2}(1+c) \ln(1+c) + \frac{1}{2}(1-c) \ln(1-c)$$



עכשיו אם רוצים לחשב הסתברות שהסכום גבוה יש לנו שני קירובים שונים. זה קצת בעייתי כי סטיות גדולות זה קירוב של זנב ונרמלית זה קירוב של כל הטור. אבל בכל זאת נרצה להשוות.
ניקח דוגמא מספרית ונסתכל.

דוגמא. מטילים מטבע הוגן $n = 200$ פעמים. מספר הפעמים שיצא עץ c .

		$c = 0$	$c = 30$	$c = 90$	$c = 150$
בינום (החישוב המדויק)	$2^{-n} \binom{n}{\frac{n+c}{2}}$	$6 \cdot 10^{-2}$	$6 \cdot 10^{-3}$	$5 \cdot 10^{-11}$	$3 \cdot 10^{-29}$
קירוב נורמלי	$\frac{1}{\sqrt{2\pi n}} e^{-\frac{c^2}{2n}}$	$6 \cdot 10^{-2}$	$6 \cdot 10^{-3}$	$9 \cdot 10^{-11}$	$2 \cdot 10^{-26}$
סטיות גדולות	$e^{n\gamma(\frac{c}{n})}$	1	$1 \cdot 10^{-1}$	$8 \cdot 10^{-10}$	$3 \cdot 10^{-28}$

קירוב נורמלי טוב לסטיות קטנות וסטיות גדולות כשמה עדיפה לסטיות גדולות.

מטבע לא הוגן

נעבור למטבע לא הוגן. יש לו פרמטר p לקבל עץ והסתברות $1 - p$ לקבל פלי.

$$p_n(a_1, \dots, a_n) = p_1(a_1) \cdot \dots \cdot p_1(a_n) = p^{a_1 + \dots + a_n} \cdot (1 - p)^{n - (a_1 + \dots + a_n)}$$

זה מביא להתפלגות בינומית.

$$\mathbb{P}(a_1 + \dots + a_n = k) = \binom{n}{k} p^k (1 - p)^{(n-k)}$$

עכשיו בכל מקרה לא נוכל לקבל הילוך סימטרי אז במקרה זה לא נשנה את הנוסחא.

$$S_{n,p} = a_1 + \dots + a_n$$

p מופיע רק באופן סמוי.

עוד מקרה פרטי של חוק המספרים הגדולים:

טענה. לכל $p \in [0, 1]$ ולכל $\varepsilon > 0$:

$$\mathbb{P}\left(\left|\frac{1}{n} S_{n,p} - p\right| \leq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 1$$

כלומר השכיחות קרובה להסתברות.

אמנם כמשפט בפני עצמו זה אולי פחות מעניין אבל מתוך זה אפשר לקבל את משפט וירשטראס שכל פונקציה רציפה אפשר לקרב במידה שווה ע"י פולינום בקטע.

טענה. לכל $\lambda \in (0, \infty)$ ולכל $k = 0, 1, 2, \dots$:

$$\mathbb{P}(S_{n, \frac{\lambda}{n}} = k) \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

עכשיו זה לא הילוך מקרי. כי הסיכוי לעץ ופלי היה קבוע ולא תלוי ב- n . כאן כבר כמובן תלויה ב- n . זאת התפלגות פואסון.

משתנים מקריים

הגדרה. מרחב הסתברות הוא כמעט אותו דבר כמו מרחב מידה. ההבדל היחיד הוא שכאן $P(\Omega) = 1$. נסמן מרחב הסתברות בדרך כלל $(\Omega, \mathcal{F}, P)$.

הגדרה. משתנה מקרי הוא פונקציה מדידה:

$$X : \Omega \rightarrow \mathbb{R}$$

הגדרה. תוחלת של משתנה מקרי X :

$$\mathbb{E}(x) = \int_{\Omega} X dP$$

בדוגמא של מטבע מרחב ההסתברות היה

$$\Omega = \{0, 1\}^n$$

$$\mathcal{F} = 2^{\Omega}$$

$$P(A) = \frac{|A|}{|\Omega|}$$

לכן,

$$P(\{\omega\}) = \frac{1}{|\Omega|} = 2^{-n}$$

היה גם משתנה מקרי:

$$S_n : \Omega \rightarrow \mathbb{R}$$

$$S_n(a_1, \dots, a_n) = \sum_{k=1}^n (2a_k - 1)$$

כאשר סימנו $S_{n,p}$ ואמרנו שאין בהגדרה p למעשה התכוונה ש- p קובע את מרחב ההסתברות.

סימון:

$$\mathbb{P}(X > a) = P(\{\omega \in \Omega : X(\omega) > a\})$$

לא מקובל לכתוב על מה פועל המשתנה המקרי. כי זה ברור מאוד. אז לא נכתוב. $\mathbb{P}$ קיצור למילה הסתברות. P פונקציית הסתברות. ננסה לא להתבלבל אבל אין כאן הבטחה לשום דבר.

לעיתים נרצה לכתוב גם:

$$\mathbb{P}(X > a, Y > b) = P(\{\omega : X(\omega) > a, Y(\omega) > b\}) = P(\underbrace{X^{-1}(a, \infty)}_{\in \mathcal{F}} \cap \underbrace{Y^{-1}(b, \infty)}_{\in \mathcal{F}})$$

פעולות על משתנים מקריים:

$$(cX)(\omega) = cX(\omega)$$

$$(X + Y)(\omega) = X(\omega) + Y(\omega)$$

$$(X^3)(\omega) = (X(\omega))^3$$

$$(\sin X)(\omega) = \sin(X(\omega))$$

תוחלת:

$$\mathbb{E}(X + Y) = (\mathbb{E}X + \mathbb{E}Y)$$

מדוע?

$$\int_{\Omega} (X + Y) dP = \int_{\Omega} X dP + \int_{\Omega} Y dP$$

זוה משפט מאוד מפורסם שאומר שתוחלת של סכום היא לא אחרת מאשר סכום התוחלות.

הגדרה. משתנים מקריים $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ ב"ת אם

$$\mathbb{P}(X_1 \in I_1, \dots, X_n \in I_n) = \mathbb{P}(X_1 \in I_1) \cdot \dots \cdot \mathbb{P}(X_n \in I_n)$$

לכל קטעים $I_1, \dots, I_n \subseteq \mathbb{R}$.

הגדרה. משתנים מקריים $X_1 : \Omega_1 \rightarrow \mathbb{R}, \dots, X_n : \Omega_n \rightarrow \mathbb{R}$ הם שויי התפלגות אם:

$$\mathbb{P}(X_1 \in I) = \dots = \mathbb{P}(X_n \in I)$$

לכל קטע $I \subseteq \mathbb{R}$.

כאן מספיק היה להגדיר לזוג ובהגדרה הקודמת אפילו אם כל זוג ב"ת עדיין זה לא מספיק כדי שכולם יהיו ב"ת.
לדוגמא:

$$X, Y, Z = \pm 1$$

$$XYZ = +1$$

יש 4 אפשרויות של סיכוי $\frac{1}{4}$:

$$X = Y = -1, Z = 1$$

$$X = Z = -1, Y = 1$$

$$Y = Z = -1, X = 1$$

$$X = Y = Z = 1$$

הבדל חשוב נוסף הוא שבמקרה הראשון כולם באותו מרחב הסתברות ובהגדרה השנייה אפשר להגדיר מעל מרחבי הסתברות שונים. נחזור למטבע הוגן.

$$\Omega = \{0, 1\}^n$$

$$S_m = \sum_{i=1}^m (2a_i - 1)$$

$$S_0, S_1, \dots, S_n : \Omega \rightarrow \mathbb{R}$$

אלה משתנים מקריים תלויים. ואגב גם לא שויי התפלגות. איפה יש כאן משתנים מקריים בלתי תלויים?

$$A_m(a_1, \dots, a_n) = a_m$$

$$A_1, \dots, A_n : \Omega \rightarrow \mathbb{R}$$

הם בלתי תלויים. האם הם שויי התפלגות? כן.

משפט. החוק החלש של המספרים הגדולים

יהיו $X_1, X_2, \dots$ משתנים מקריים שויי התפלגות אינטגרליים. אז לכל $\varepsilon > 0$

$$\mathbb{P} \left(\left| \frac{x_1 + \dots + x_n}{n} - \mathbb{E}X_1 \right| \leq \varepsilon \right) \xrightarrow{n \rightarrow \infty} 1$$

שויי התפלגות לכן שויי תוחלת לכן שויי תוחלת. אז בחרנו את $\mathbb{E}X_1$ שם באופן שרירותית לגמרי, ויכולנו באותו אופן בדיוק לבחור תוחלת של כל משתנה מקרי אחר.

בעיה פילוסופית: האם קיים Ω עבורו אפשר לקבל את התוצאה שאמרנו עבור מטבע הוגן? כן אבל לא ברור. אפשר למשל איכשהו להגדר "משולש" של משתנים בדידים כך שההתפלגות

של כל משתנה שתהיה שווה להתפלגות של המשתנה שמתחתיו, ההתפלגות של כל זוג שווה להתפלגות של כל זוג מתחתיו וכו'. למשל:

$$x_1^{(1)}$$

$$x_1^{(2)} x_2^{(2)}$$

$$x_1^{(3)} x_2^{(3)} x_3^{(3)}$$

⋮

ואז ל- $x_1^{(1)}$ ול- $x_1^{(2)}$ ול- $x_1^{(3)}$ אותה התפלגות, ל- $x_2^{(2)}$ ול- $x_2^{(3)}$ אותה התפלגות ואפילו ל- $x_1^{(2)} x_2^{(2)}$ ול- $x_1^{(3)} x_2^{(3)}$ אותה התפלגות. הבעיה שזה די מסובך ובסופו של דבר נצטרך להתמודד עם הסתברויות לא בדידות.