יסודות תורת ההסתברות

אוהד נ. פלדהיים
אורי גוראל גורביץ'

מכון איינשטיין למתמטיקה

האוניברסיטה העברית בירושלים

14 בדצמבר 2018

פרק 8

סדרות של משתנים מקריים

"...התורה של ניסויים בלתי-תלויים היא בעת ובעונה אחת החלק הפשוט ביותר מבחינה אנליטית והמתקדם ביותר של תורת ההסתברות."
– וויליאם פלר, מבוא לתורת ההסתברות ושימושיה, 1950.

בפרקים הקודמים הגבלנו את עיסוקנו לקבוצות סופיות של משתנים מקריים. אומנם, באמצעות השאפת אורך הסדרה לאינסוף ניסינו לעמוד על ההתנהגות האסימפטוטית של אוספים כאלו כאשר מספר המשתנים הולך וגדל ולהבין אילו מאורעות נעשים סבירים יותר ויותר ואילו הופכים נדירים יותר ויותר. באמצעות הסתכלות זו ניסחנו והוכחנו את החוק החלש של המספרים הגדולים (משפט 5.32), ותיארנו את התנהגותן של בעיות הסתברותיות שונות. לצורך כך היה עלינו להסתכל על כל אוסף סופי של משתנים מקריים כאילו הוא מוגדר במרחב הסתברות משלו.

בפרק זה נבקש להסתכל על סדרה אינסופית של משתנים מקריים בלתי-תלויים המוגדרת על אותו מרחב הסתברות ולראותה כסדרה מקרית של ערכים. כך נוכל למשל לנסח את השאלה "מה היא ההסתברות שסדרה מסויימת מתכנסת? ". לצורך טיפול בשאלות אלו נידרש למרחב הסתברות שעליו מוגדרים אינסוף משתנים מקריים בלתי-תלויים. אין זה קשה לוודא שאף מרחב הסתברות בדידה אינו מקיים תכונה זו. ואולם, את מרחב ההסתברות האחיד על $[0,1]$ נוכל לראות כמרחב של סדרות אינסופיות של ספרות עשרוניות הנבחרות באופן בלתי-תלוי. מסתבר שניתן להכליל הסתכלות זו לקבלת הטענה הבאה.

טענה 8.1 (קיום סדרה אינסופית של משתנים מקריים בלתי-תלויים)**.** יהי X משתנה מקרי. על מרחב ההסתברות האחיד $([0,1],\mathbb{B}([0,1]),\mathbb{P})$ קיימים משתנים מקריים $\{X_i\}_{i\in\mathbb{N}}$, בלתי-תלויים, כולם שווי-התפלגות ל-X.

בנספח הפניה להלן נדון באופן נרחב יותר בהצדקה לטענה זו, אבל הוכחה פורמלית שלה נסמכת על תוצאות בתורת המידה והיא אינה נכללת במסגרת ספר זה. לאור תכונותיה של $\mathbb{B}([0,1])$ כ-σ-אלגברה, נוכל לשייך הסתברות לכל שאלה הנוגעת לסדרת משתנים מקיים בלתי-תלויים $\{X_i\}_{i\in\mathbb{N}}$ אשר ניתן לנסחה במונחים של רצף סופי של פעולות משלים, חיתוך בן-מניה ואיחוד בן-מניה של מאורעות שכל אחד מהם נוגע למשתנה מקרי בודד (ר' הגדרה 7.1). באופן כללי לא נתעכב על נקודה בהמשך הפרק והעיון בשאלה "לאילו מאורעות ניתן לשייך הסתברות?" ישמר לספרים מתקדמים יותר הנסמכים על תורת המידה.

$\{w \mid \forall n\ \exists m \geq n \Rightarrow w \in A_m\} = \{A_n\ i.o\} = \bigcap_{n=1}^{\infty} \bigcup_{m \geq n} A_m$

לכל $n \in N$ — קיים $m \in N$ כך ש $m \geq n$ — (ומתקיים) $\Rightarrow w \in A_m$

<u>דוגמ':</u> לכל מרחק n שאני מתקדם בסדרת הטלות שלי קיים m רחוק כך שיצא עץ בהטלה ה-m

$\{A_n\ .a.e\} = \bigcup_{n=1}^{\infty} \bigcap_{n=m} A_m$

קיים $n \in N$ — לכל $m \geq n$ מתקיים — $(\Rightarrow)$ יצא רק פלי/מאורע כלשהו

<u>דוגמ':</u> אם יש לי כד גרעינים ואני כל הזמן מוציא את כללי הטילית יש לי זמן שבו כל מעתה

<u>אז</u> יהיו לי תקלות

אז למה: מקובל גם לסמן $\limsup_{i\to\infty} A_n := \{A_n \text{ i.o.}\}$, $\liminf_{i\to\infty} A_n := \{A_n \text{ a.e}\}$.

<u>נזכר בעוד הגדרות:</u>

משפט 2.20 (רציפות פונקציית ההסתברות על מאורעות עולים). יהי $(\Omega, \mathcal{F}, \mathrm{P})$ מרחב הסתברות. ותהי $\{A_n\}_{n\in\mathbb{N}}$ סדרה עולה של מאורעות ב-$\mathcal{F}$. אזי מתקיים,

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathrm{P}(A_n).$$

מסקנה 2.21 (רציפות פונקציית ההסתברות על מאורעות יורדים). אם $(A_n)_{n\in\mathbb{N}}$ היא סדרת מאורעות יורדים אזי

$$\mathrm{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathrm{P}(A_n).$$

<u>עבור המקרה:</u>

$\{A_n\ .a.e\} = \bigcup_{n=1}^{\infty} \bigcap_{n=m}^{\infty} = \bigcup_{n\in N} \bigcap_{i\geq n} A_i$

אני לוקח יותר ויותר מאורעות לכן אני מקבל כל פעם קבוצה קטנה

יותר כלומר יש כאן <u>סדרה מונו' יורדת</u>.

באותו אנלוגי עבור $\{A_n\ i.o\} = \bigcap_{n=1}^{\infty} \bigcup_{m=n} A_i = \bigcap_{n\in N} \bigcup_{i\geq n} A_i$ אני מאחד כל הזמן לכן יש לי סדרה של <u>מאורעות עולים</u>

תמיד אני גדלה

לכן אני מקבל

 # טענה תהיי x_n סדרה ב-$\mathbb{R}$

ויהיו $T_n = \{x_k | n \leq k \in \mathbb{N}\}$

$T_{n+1} \subset T_n \subset T_{n-1} \subset \cdots \subset T_3 \subset T_2 \subset T_1 = \{x_n\} \subseteq [a,b]$

נגדיר $s_n = Sup\, T_n$ ו $i_n = inf\, T_n$

אזי מתקיים כי $i_n \leq i_{n+1} \leq s_{n+1} \leq s_n$

נגדיר

$limSup(x_n) = inf(s_n)$

$liminf(x_n) = Sup(i_n)$

ומכאן הגיע

limsup

liminf

8.1 התכנסות של סדרת מאורעות

בבואנו לדבר על התכנסות של סדרה אינסופית של מאורעות במרחב הסתברות נתעניין בשני מאורעות מרכזיים.

הגדרה 8.2 (התרחשות אסימפטוטית)**.** תהי $(A_n)_{n\in\mathbb{N}}$ סדרת מאורעות במרחב הסתברות. נגדיר את המאורעות הבאים:

(א) $\{A_n \text{ i.o}\} = \{|\{n \ : \ A_n\}| = \infty\}$ המאורע לפיו A_n מתרחשים infinitely often, כלומר, **אינסוף פעמים** (או במילים אחרות אינסוף מתוך A_i מתרחשים). מבחינה פורמלית

$$\{A_n \text{ i.o}\} = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$

(ב) $\{A_n \text{ a.e}\} = \{|\{n \ : \ A_n^c\}| < \infty\}$ המאורע לפיו A_n מתרחשים almost everywhere, או eventually כלומר, **החל ממקום מסויים (המ"מ)** או **כמעט בכל מקום (כב"מ)**. מבחינה פורמלית

$$\{A_n \text{ a.e}\} = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m.$$

מקובל גם לסמן $\liminf_{i\to\infty} A_n := \{A_n \text{ a.e}\}$, $\limsup_{i\to\infty} A_n := \{A_n \text{ i.o.}\}$.

אבחנה 8.3. לכל סדרת מאורעות A_i במרחב הסתברות מתקיים

$$\{A_i \text{ i.o.}\}^c = \{A_i^c \text{ a.e}\}$$

בכדי להיווכח בנכונות האבחנה, נשים לב כי אם A_n המ"מ – בפרט מתקיים כי A_n אינסוף פעמים. כמו כן המאורע A_n^c כמעט בכל מקום הוא משלימו של המאורע A_n אינסוף פעמים, וכי המאורע A_n^c אינסוף פעמים הוא משלימו של המאורע A_n כמעט בכל מקום.

להמחשה נשווה בנפשנו אינסוף הטלות מטבע המתבצעות בזו אחר זו ונסמן ב-A_n את העובדה שבהטלה מסויימת התקבלה תוצאה של עץ. הטענה "כל המטבעות הראו תוצאה של עץ החל ממקום מסויים" (כלומר A_n המ"מ), שקולה לכך שתוצאת פלי התקבלה לכל היותר במספר סופי של מקומות (כלומר לא A_n^c אינסוף פעמים). העובדה שטענה זו לא מתקיימת, כלומר "לא קיים מקום שממנו ואילך כל המטבעות הראו תוצאה של עץ" שקולה לכך שהתקבלה התוצאה פלי באינסוף מקומות (כלומר A_n^c אינסוף פעמים).

לעיתים קרובות נוח לחשוב על A_n כמאורע המתרחשים בזמן n. מנקודת השקפה זו המאורע $\{A_n \text{ a.e}\}$ מכונה גם A_n **אסימפטוטית כמעט תמיד** (asymptotically almost surely).

העובדה שסדרת מאורעות תתקיים החל ממקום מסויים אינה מבטיחה לנו כמה מן המאורעות לא יתרחשו. כך למשל אם ברשותנו אגרטל ואנחנו מטילים בכל יום מטבע הוגן ובתוצאה של עץ מחליטים לשבור את האגרטל, המאורעות "היום האגרטל שבור" יתרחשו המ"מ ואולם האגרטל עשוי לשרוד כל מספר ימים. כך, גם כאשר מובטח כי אינסוף מתוך סדרת מאורעות יתרחשו אין כל ערבון לכך שתדירותם לא תהיה קטנה להדהים.

המתמטיקאי הצרפתי פייר פאטו (Pierre Fatou), בן זמנו של לבג, הראה את החסמים הפשוטים הבאים

המשמעות

(ד) $1-\mathbb{P}(\{A_i^c \text{ a.e.}\}) \overset{\text{חלק א}}{\geq} 1-\liminf_{n\to\infty}\mathbb{P}(A_n^c)$

$$\Downarrow$$

$$-P(A_i^c, a.e) \geq -\liminf(A_n^c)$$

$$\underbrace{P(A_i^c\ a.e)}_{\text{סיכוי שהחל ממקום מסוים יצא לי רק פלי}} \leq \underbrace{\liminf(A_n^c)}_{\text{הגבול התחתון של כל ההטלות מכבר}}$$

המשל שלי:

[סיכוי שהחל ממקום מסוים יצא לי רק פלי] $\leq$ [הגבול התחתון של כל ההטלות מכבר]

(ה) סדרה של מאורעות יורדים אני כל פעם חותך יותר ויותר (כ

לכן אני מקבל סדרה של מאורעות יורדים

להסתברות שסדרת מאורעות תתרחש אינסוף פעמים ולהסתברות שהיא תתרחש כמעט תמיד.

טענה 8.4 (הלמה של פאטו). תהי $\{A_n\}_{n\in\mathbb{N}}$ סדרה של מאורעות. אזי

$$\mathbb{P}(\{A_i \text{ a.e.}\}) = \mathbb{P}(\liminf_{n\to\infty} A_n) \le \liminf_{n\to\infty} \mathbb{P}(A_n)$$

$$\mathbb{P}(\{A_i \text{ i.o.}\}) = \mathbb{P}(\limsup_{n\to\infty} A_n) \ge \limsup_{n\to\infty} \mathbb{P}(A_n)$$

הוכחה. ראשית נראה כי הטענה השניה נובעת מהפעלת הטענה הראשונה על סדרת המאורעות $\{A_n^c\}$. כלומר -

$$\mathbb{P}(\{A_i \text{ i.o.}\}) \overset{\substack{\text{אבחנה}\\ 8.3}}{=} 1 - \mathbb{P}(\{A_i^c \text{ a.e.}\}) \underset{(\text{א})}{\overset{\text{חלק א}}{\ge}} 1 - \liminf_{n\to\infty} \mathbb{P}(A_n^c) = 1 - \liminf_{n\to\infty}(1-\mathbb{P}(A_n)) = \limsup_{n\to\infty} \mathbb{P}(A_n).$$

להוכחת הטענה הראשונה, ניזכר ברציפות פונקציית ההסתברות למאורעות עולים (משפט 1.20),

$$\liminf_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \inf_{i>n} \mathbb{P}(A_i) \ge \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i>n} A_i\right) \underset{(\text{א})}{\overset{\substack{\text{משפט}\\ 1.20}}{=}} \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \bigcap_{i>n} A_i\right) = \mathbb{P}(\{A_i \text{ a.e.}\})$$

■

המתמטיקאים אמיל בורל (Émile Borel) ופרנצ'סקו קנטלי (Francesco Paolo Cantelli) הבחינו בשנות השלושים של המאה העשרים בקריטריונים הבאים להתרחשות אינסוף מאורעות ולהתרחשותם המ"מ.

משפט 8.5 (הלמה ה-I של בורל-קנטלי). תהי A_i סדרת מאורעות. אם $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty$ אז $\mathbb{P}(A_i \text{ i.o.}) = 0$.

הוכחה. נשים לב כי לפי משפט בול,

$$\mathbb{P}(A_i \text{ i.o.}) \overset{\substack{\text{משפט}\\ 1.20}}{=} \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=n}^{\infty} A_i\right) \overset{\substack{\text{משפט}\\ 1.17}}{\le} \lim_{n\to\infty} \sum_{i=n}^{\infty} \mathbb{P}(A_i) = 0$$

כאשר השוויון האחרון נובע מכך ש- $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty$ אם ורק אם $\lim_{n\to\infty} \sum_{i=n}^{\infty} \mathbb{P}(A_i) = 0$. ■

משפט 8.6 (הלמה ה-II של בורל-קנטלי). תהי A_i סדרת מאורעות בלתי-תלויים. אם $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ אז $\mathbb{P}(A_i \text{ i.o.}) = 1$.

הוכחה. נשים לב כי

$$\mathbb{P}(A_i \text{ i.o.}) = 1 - \mathbb{P}(A_i^c \text{ a.e.}) = 1 - \mathbb{P}\left(\bigcup_{m=1}^{\infty} \bigcap_{i=m}^{\infty} A_i^c\right) \overset{\substack{\text{משפט}\\ 1.20}}{=} 1 - \lim_{m\to\infty} \mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right)$$

על כן, די שנראה כי לכל $m \in \mathbb{N}$ מתקיים:

$$\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right) = 0.$$

נחשב ונקבל באמצעות אי תלות

לשים לב שתמיד מתקיים כי $\mathbb{P}(\bigcap_{i\in\mathbb{N}} A_i) = \mathbb{P}[\bigcap_{n\in\mathbb{N}}[\bigcap_{i=1}^{n} A_i]] = \lim_{n\to\infty}(\bigcap_{i=1}^{n} A_i)$

$$\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right) \overset{\text{משפט } 1.20}{=} \lim_{n\to\infty}\mathbb{P}\left(\bigcap_{i=m}^{n} A_i^c\right) = \lim_{n\to\infty}\prod_{i=m}^{n}(1-\mathbb{P}(A_i)) \le \lim_{n\to\infty}\exp\left(-\sum_{i=m}^{n}\mathbb{P}(A_i)\right) = 0$$

כאשר האי-שיוויון נובע מכך ש-$1+x \le e^x$ לכל x והשוויון האחרון משתמש בהנחה $\sum_{i=1}^{\infty}\mathbb{P}(A_i) = \infty$, ממנה נובע כי $\sum_{i=m}^{\infty}\mathbb{P}(A_i) = \infty$ לכל m. ■

נשים לב שבמקרה של מאורעות בלתי תלויים, שני חלקי הלמה משלימים זה את זה, ובפרט מתקיים שלכל סדרת מאורעות בלתי-תלויים $\mathbb{P}(A_i \text{ i.o}) \in \{0,1\}$. לעומת זאת לא ניתן להכליל את הלמה השניה למאורעות תלויים. כך למשל מפני שאם נטיל מטבע הוגן ונקבע לכל i את A_i להיות המאורע שהתקבלה תוצאה של ראש, אזי נקבל

$$\mathbb{P}(A_i \text{ i.o.}) = \mathbb{P}(A_i \text{ a.e.}) = \frac{1}{2}.$$

דוגמא 8.7. יהיו $X_n \sim \mathrm{Geo}(1/e)$ בלתי-תלויים. נסמן $A_n = \{X_n > \alpha\log n\}$ יש להראות כי $\mathbb{P}(A_n \text{ i.o.}) = 1$ אם ורק אם $\alpha \le 1$.

תשובה: נחשב לפי פונקציית ההתפלגות השיורית של משתנה מקרי גיאומטרי (טענה 3.40).

$$\mathbb{P}(A_n) = \mathbb{P}(X_n > \alpha\log n) = (1-\frac{1}{e})^{\lfloor\alpha\log n\rfloor}$$

מכיוון שמתקיים $e^{-\alpha+1} \ge (1-\frac{1}{e})^{\alpha} \ge e^{-\alpha}$ נסיק כי

$$n^{-\alpha} \le e^{-\lfloor\alpha\log n\rfloor} \le \mathbb{P}(A_n) \le e^{-\lfloor\alpha\log n\rfloor+1} \le e^2 n^{-\alpha}$$

לכן מתקיים

$$\sum_{n=0}^{\infty}\mathbb{P}(A_n) = \infty \Leftrightarrow \alpha \le 1.$$

מכאן שלפי הלמה הראשונה של בורל-קנטלי (משפט 8.5) לכל $\alpha > 1$ מתקיים $\mathbb{P}(A_n \text{ i.o.}) = 0$ ולפי הלמה השניה של בורל-קנטלי (משפט 8.6) לכל $\alpha \le 1$ מתקיים $\mathbb{P}(A_n \text{ i.o.}) = 1$ כנדרש.

דוגמא 8.8. יהיו $\{A_n\}_{n\in\mathbb{N}}$ מאורעות בלתי תלויים בעלי הסתברויות $\mathbb{P}(A_n) = 1/n^{\alpha}$. נגדיר $B_n = A_n \cap A_{n+1}$. עבור אילו α מתקיים $\mathbb{P}(B_n \text{ i.o.}) = 1$?

תשובה:

$$\mathbb{P}(B_n) = (n(n+1))^{-\alpha} \approx n^{-2\alpha}$$

מכאן שלפי הלמה הראשונה של בורל-קנטלי (משפט 8.5) לכל $\alpha > 1/2$ מתקיים $\mathbb{P}(B_n \text{ i.o.}) = 0$.

כאשר $\alpha \le 1/2$ הטור $\sum_{n=1}^{\infty}\mathbb{P}(B_n)$ מתבדר, אולם המאורעות $\{B_n\}_{n\in\mathbb{N}}$ אינם בלתי תלויים ולכן לא ניתן להשתמש בלמה השניה של בורל-קנטלי (משפט 8.6) ישירות. ואולם, נשים לב שהמאורעות $\{B_{2n}\}_{n\in\mathbb{N}}$ הינם בלתי תלויים והטור $\sum_{n=1}^{\infty}\mathbb{P}(B_{2n})$ מתבדר ולפי המה השניה של בורל-קנטלי (משפט 8.6) $\mathbb{P}(B_2n \text{ i.o.}) = 1$ ולפיכך גם $\mathbb{P}(B_n \text{ i.o.}) = 1$ במקרה זה.

דוגמא 8.9 (הילוך שיכור חולף או נשנה)**.** יהיו $\{B_i\}_{i\in\mathbb{N}}$ מ"מ ברנולי עם פרמטר p ב"ת ונסמן $X_n = 2\sum_{i=0}^{n} B_i - n$. תהליך זה מכונה הילוך שיכור (מוטה). כאשר $p < \frac{1}{2}$ ההילוך מוטה שמאלה, כאשר $p > \frac{1}{2}$ הוא מוטה ימינה

וכאשר $p=\frac{1}{2}$ ההילוך מאוזן. נסמן $A_i=\{X_i=0\}$. אם A_i מתרחשים אינסוף פעמים בהסתברות אחת נאמר שההילוך **נשנה** ואם A_i^c החל ממקום מסויים – נאמר שהוא **חולף**. חקור האם ההילוך חולף, נשנה או לא חולף ולא נשנה כתלות בפרמטר p.

תשובה: עבור $p<\frac{1}{2}$ נשים לב ש-$\frac{(2B_n-1+1-2p)}{2}$ מקיימים את תנאי אי-שיוויון הופדינג (משפט 6.10), מפני ש-

$$\mathbb{E}(2B_n-1+1-2p)=2p-1+1-2p=0,$$

וכן $|2B_n-1+1-2p|<2$. נסמן אפוא

$$Y_n=\frac{\sum_{i=1}^{n}(2B_n-1+1-2p)}{2}=\frac{X_n}{2}+\frac{n(1-2p)}{2}$$

. לפי אי-שיוויון הופדינג, נקבל

$$\mathbb{P}(A_i)=\mathbb{P}(X_i=0)\leq\mathbb{P}\left(\frac{X_i}{2}\geq 0\right)=\mathbb{P}\left(Y_i\geq\frac{n(1-2p)}{2}\right)\overset{\text{הופדינג}}{\leq}\exp\left(\frac{-(1-2p)^2n}{8}\right)$$

ולכן $\sum_{i=1}^{\infty}\mathbb{P}(A_i)<\infty$ ולפי הלמה הראשונה של בורל קנטלי (משפט 8.5) נקבל $\mathbb{P}(A_i \text{ i.o.})=0$.

עבור $p>\frac{1}{2}$ נשים לב כי $X_i\overset{d}{=}-X_i'$, כאשר X_i' הוא הילוך שיכור מוטה עם פרמטר $q=1-p$ ומכיוון שהראינו ש-X_i' כזה חולף - הרי שגם X_i חולף.

המקרה $p=\frac{1}{2}$ מורכב יותר, ולא נוכל לטפל בו על ידי שימוש פשוט בלֶמות של בורל-קנטלי. נשים לב כי $\sum_{i=1}^{n}B_i\sim\text{Bin}(1/2,n)$, ולכן

$$\mathbb{P}(X_{2n}=0)=\mathbb{P}\left(\sum_{i=1}^{2n}B_i=n\right)\overset{\substack{\text{דוגמא}\\3.46}}{\geq}\frac{0.4}{\sqrt{n}}$$

כעת נסמן עבור $i\in\mathbb{N}_0$ משתני אינדיקטור לביקור ב-0 בזמן i, על ידי $W_i=\mathbb{1}(A_i)$. ונקבל

$$\mathbb{E}\left(\sum_{i=1}^{\infty}W_i\right)\geq\sum_{i=1}^{\infty}\frac{0.4}{\sqrt{n}}=\infty$$

הערה:

נסמן $p_{\text{last},n}=\mathbb{P}(\{\forall i>n: W_i=0\}\mid W_n=1)$ ונשים לב כי

$$p_{\text{last},n}=\mathbb{P}\left(\forall m>n: \sum_{i=n+1}^{m}X_i\neq 0\right.$$

ולכן הסתברות זו לא תלויה ב-n. נסמנה אפוא ב-

$$p_{\text{last}}=1-\sum_{n=1}^{\infty}\mathbb{P}(W_n=1,W_{n-1}=0,\dots,W_1=0)$$

ונחשב לפי נוסחת התוחלת השלמה (טענה 4.14) עבור $m\in\mathbb{N}$:

$$\mathbb{E}\left(\sum_{n=1}^{m}W_n\right)=\sum_{n=1}^{m}\mathbb{P}(W_n=1,W_{n-1}=0,\dots,W_1=0)\mathbb{E}\left(\sum_{i=n}^{m}W_i\,\middle|\,W_n=1,W_{n-1}=0,\dots,W_1=0\right)$$

הערה 2:

$$\overset{\substack{\text{אי}\\\text{תלות}}}{=}\sum_{n=1}^{m}\mathbb{P}(W_n=1,W_{n-1}=0,\dots,W_1=0)\left(1+\mathbb{E}\left(\sum_{i=n+1}^{m}W_i\right)\right)$$

$$\leq 1+\sum_{n=1}^{m}\mathbb{P}(W_n=1,W_{n-1}=0,\dots,W_1=0)\mathbb{E}\left(\sum_{i=1}^{m}W_i\right)=1+(1-p_{\text{last}})\mathbb{E}\left(\sum_{i=1}^{m}W_i\right)$$

הערה:

$P_{last,n}$ מייצג את ההסתברות לא עשה לימודים בזמן ה-n

המלורה של ה׳א לצורך איזשהו שם בית S

נשים לב ש $\sum_{i=n+1}^{\infty} x_i$ מייצגים את הבעיה שלי ו $P_{last,n}$

ונוצר מצב שבו הסיכוי בזמן 100 להיות ב-5 = לסיכוי בזמן 30 להיות בס

הערה 2:

$$E\Big(\sum_{\text{מסביבונים}} w_i \;\Big|\; \underbrace{w_n=1,\ w_{n-1}=0,\ \ldots\ w_1=0}_{\text{הייתי באותו מקום בזמן ה-n}}\Big)$$

נפתור ונסיק כי עבור $p_{\text{last}} > 0$ מתקיים

$$\mathbb{E}\left(\sum_{i=1}^{m} W_i\right) \leq \sum_{n=0}^{\infty}(1-p_{\text{last}})^n = \frac{1}{p_{\text{last}}}$$

בסתירה לכך ש-$\mathbb{E}(\sum_{i=1}^{\infty} W_i) = \infty$. נסיק אפוא כי $p_{\text{last}} = 0$ ולכן לכל n קיים בהסתברות 1 $m > n$ כך ש-$W_i = 1$. מכאן נסיק כי $\mathbb{P}(A_i \text{ i.o.}) = 1$.

8.1.1 הכללת הלמה השנייה של בורל קנטלי למאורעות בלתי-תלויים בזוגות

למעשה ניתן להכליל את הלמה השנייה של בורל-קנטלי, למאורעות בלתי-תלויים בזוגות (ר' הגדרה 2.23).

משפט 8.10 (הלמה ה-II של בורל-קנטלי למשתנים בלתי-תלויים בזוגות). תהי A_i סדרת מאורעות. אם $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ והמאורעות בלתי-תלויים בזוגות אז $\mathbb{P}(A_i \text{ i.o.}) = 1$.

כל זוג מאורעות ב"ת

$A,B \in X$

ב"ת בזוגות ?

הוכחה. כמקודם, נראה כי לכל m

$$\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right) = 0.$$

נגדיר $X_m^n = \sum_{i=m}^{n} \mathbb{I}(A_n)$ ונשים לב כי

$$\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right) \overset{\substack{\text{משפט}\\1.20}}{=} \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i=m}^{n} A_i^c\right) = \lim_{n\to\infty} \mathbb{P}(X_m^n = 0).$$ הערה (1)

נסמן $p_i = \mathbb{P}(A_i)$ ונחשב את תוחלת ושונות X_m^n

$$\mathbb{E}(X_m^n) = \sum_{i=m}^{n} p_i,$$ הערה (2)

$$\text{Var}(X_m^n) = \text{Var}\left(\sum_{i=m}^{n} \mathbb{I}(A_i)\right) \overset{\substack{\text{טענה}\\5.16}}{=} \sum_{i=m}^{n} \text{Var}(\mathbb{I}(A_i)) = \sum_{i=m}^{n} p_i(1-p_i) \leq \sum_{i=m}^{n} p_i = \mathbb{E}(X_m^n).$$

כעת נשתמש באי-שיוויון צ'בישב,

(3)

$$\mathbb{P}\left(\bigcap_{i=m}^{n} A_i^c\right) = \mathbb{P}(X_m^n = 0) \leq \mathbb{P}(|X_m^n - \mathbb{E}(X_m^n)| \geq \mathbb{E}(X_m^n)) \overset{\substack{\text{משפט}\\5.5}}{\leq} \frac{\text{Var}(X_m^n)}{\mathbb{E}(X_m^n)^2} \leq \frac{E(X_m^n)}{[E(X_m^n)]^2} = \frac{1}{E(X_m^n)}$$

נשים לב כי, היות ש-$\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$, הרי שמתקיים $\lim_{n\to\infty} \mathbb{E}(X_m^n) = \sum_{i=m}^{\infty} \mathbb{P}(A_i) = \infty.$ ומכאן נסיק כי

$$\lim_{n\to\infty} \mathbb{P}(X_m^n = 0) \leq \frac{1}{\lim_{n\to\infty} \mathbb{E}(X_m^n)} = 0,$$

כנדרש. ■

8.2 התכנסות של סדרות משתנים מקריים

נציג כעת שני מובנים של התכנסות הנוגעים לסדרות משתנים מקריים המצויים באותו מרחב הסתברות. בכדי להבין את ההבדלים בין שתי צורות התכנסות אלו, נלווה את הצגתן בדוגמא מכוננת.

הערה 1:

$$\lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i=m}^{n} A_i^c\right) = \lim_{n\to\infty} \mathbb{P}\left(X_m^n = 0\right)$$

הסיכוי שלמרות כל המאורעות המשלימים יקרו

זה הסיכוי ש $X_m^n = 0$ כלומר שאף A_i לא קרה כלומר

שאין אף כל בזמן לכל $n \ge i \ge m$ $\quad \mathbb{1}(A_i) = 0$

כי $X_m^n = \sum_{i=m}^{n} \mathbb{1}(A_i) = 0$ ולבסוף נכון $\lim_{n\to\infty}$

הסבר לתחילת השוויון

$$\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i\right) = \mathbb{P}\left[\bigcap_{n\in N}\left(\bigcap_{i=m}^{n} A_i\right)\right] = \mathbb{P}\left(\bigcap_{i=m}^{n} A_i\right)$$

ולבסוף נכון $\lim_{n\to\infty}$

הערה 2: ידוע כי $\mathbb{E}(\mathbb{1}(A_i)) = p_i$ לכל $n \ge i \ge m$

מכיוון שיש לי סכום של אינדיקטורים אז נקבל ש

$$\mathbb{E}(X_m^n) = \sum_{i=m}^{n} p_i$$

(3) **המאורע:**

$$X_m^n = 0 \subset \left\{|X_m^n - \mathbb{E}(X_m^n)| \ge \mathbb{E}(X_m^n)\right\}$$

לכן מתקיים הא"ש

$$\mathbb{P}(X_m^n = 0) \le \mathbb{P}(|X_m^n - \mathbb{E}(X_m^n)| \ge \mathbb{E}(X_m^n))$$

נדמיין שחובב מפות עולם בוחר באקראי מפה מתוך רשימה של מאה מפות הסטוריות ומחליט להתמחות בהעתקתה. לאחר מכן הוא בודק את יצירתו על ידי כך שהוא מסתכל מה אחוז העולם במפה המכוסה במים. נסמן ב-X_n את שטח הים היחסי במפה ה-n-ית וב-X את שטח הים היחסי במפה המקורית. בשנים הראשונות הדגמים שונים מאוד מהמקור, אך ככל שהמעתיק מתמחה X_n נוטה יותר ויותר להתקרב ל-X ואף מזדהה עמו לעתים קרובות.

נאמר ש-X_n מתכנס כמעט תמיד (או כמעט בוודאות) ל-X אם בהסתברות אחת, לאחר אימון מספק, יהיה שטח הים היחסי בכל המפות המועתקות החל מהעתקה מסויימת ואילך, קרוב כרצוננו לשטח הים היחסי במפה המקורית. התכנסות זו נתונה בהגדרה הבאה.

הגדרה 8.11 (התכנסות כמעט תמיד)**.** תהי $(X_n)_{n\in\mathbb{N}}$ סדרת משתנים מקריים במרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. נאמר כי סדרה זו מתכנסת למשתנה המקרי X **כמעט-תמיד**, ונסמן $X_n \xrightarrow{\text{a.s.}} X$, אם מתקיים

$$\mathbb{P}\left(\lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1.$$

נאמר ש-X_n מתכנס בהסתברות ל-X אם הסיכוי לטעות בכל גודל סופי שטח הים היחסי במפה ה-n שואף לאפס כאשר n שואף לאינסוף. דוגמא למקרה כזה היא כאשר שכיחות הטעויות של המעתיק יורד והולך, אך בכל זאת מפעם בפעם הוא ממשיך ומבצע טעויות העתקה.

הגדרה 8.12 (התכנסות בהסתברות)**.** תהי $(X_n)_{n\in\mathbb{N}}$ סדרת משתנים מקריים במרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. נאמר כי סדרה זו מתכנסת למשתנה המקרי X **בהסתברות**, ונסמן $X_n \xrightarrow{\text{p}} X$, אם לכל $\epsilon > 0$ מתקיים

$$\lim_{n\to\infty} \mathbb{P}(|X_n(\omega) - X(\omega)| \le \epsilon) = 1.$$

בהתקלות ראשונה, שתי ההגדרות נדמות דומות להפליא. בכדי לאמוד על ההבדלים שביניהם, נבהיר את האתגר הטכני בהוכחת כל אחד מסוגי ההתכנסות במושגים של חשבון אינפיניטיסימלי.

טענה 8.13. תהי $(X_n)_{n\in\mathbb{N}}$ סדרת משתנים מקריים במרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. אזי

$$X_n \xrightarrow{\text{p}} X \iff \left(\forall\epsilon > 0 \;:\; \limsup_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0\right)$$

$$X_n \xrightarrow{\text{a.s.}} X \iff \left(\forall\epsilon > 0 \;:\; \mathbb{P}(\limsup_{n\to\infty} |X_n - X| > \epsilon) = 0\right)$$

הוכחה. להוכחת השקילות הראשונה, נשים לב כי, לפי הגדרה, $X_n \xrightarrow{\text{p}} X$ שקול לכך ש-

$$\lim_{n\to\infty} \mathbb{P}(|X_n(\omega) - X(\omega)| > \epsilon) = 0,$$

לכל $\epsilon > 0$. כעת מכיוון ש $\mathbb{P}(|X_n(\omega) - X(\omega)| > \epsilon) \ge 0$ טענה זו שקולה לכך ש-

$$\limsup_{n\to\infty} \mathbb{P}(|X_n(\omega) - X(\omega)| > \epsilon) = 0,$$

כנדרש.

להוכחת השקילות השנייה, נשים לב כי, לפי הגדרת הגבול, התכנסות כמעט תמיד, $X_n \xrightarrow{\text{a.s.}} X$, מתקיימת אם ורק אם

$$\mathbb{P}\left(\forall \epsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n > n_0\ :\ |X_n - X| \le \epsilon\right) = 1,$$

נסתכל על המאורע המשלים ונקבל שטענו זו שקולה לכך ש-

$$\mathbb{P}\left(\exists k \in \mathbb{N} \forall n_0 \in \mathbb{N}\ \exists n > n_0\ :\ |X_n - X| > k^{-1}\right) = 0,$$

נשים לב כי

$$\mathbb{P}\left(\exists k \in \mathbb{N} \forall n_0 \in \mathbb{N}\ \exists n > n_0\ :\ |X_n - X| > k^{-1}\right) \le \sum_{k \in \mathbb{N}} \mathbb{P}\left(\forall n_0 \in \mathbb{N}\ \exists n > n_0\ :\ |X_n - X| > k^{-1}\right),$$

וכי אגף שמאל שווה ל-0 אם ורק אם כל אברי הסכום מתאפסים. והרי

$$\forall_{k\in\mathbb{N}} \mathbb{P}\left(\forall n_0 \in \mathbb{N}\ \exists n > n_0\ :\ |X_n - X| > k^{-1}\right) \iff \forall_{\epsilon>0} \mathbb{P}\left(\forall n_0 \in \mathbb{N}\ \exists n > n_0\ :\ |X_n - X| > \epsilon\right).$$

קיבלנו ש-$X_n \xrightarrow{\text{a.s.}} X$ אם ורק אם,

$$\forall \epsilon > 0\ :\ \mathbb{P}(\limsup_{n\to\infty} |X_n - X| > \epsilon) = 0,$$

כנדרש. ■

משפט 8.14 (התכנסות כמעט תמיד גוררת התכנסות בהסתברות)**.** אם $X_n \xrightarrow{\text{a.s.}} X$ אז $X_n \xrightarrow{\text{p}} X$.

הוכחה. נסמן $A_n^k = \{\omega : |X_n(\omega) - X(\omega)| \le 1/k\}$. כעת נוכל לכתוב את המאורע $\{X_n(\omega) \to X(\omega)\}$ בתור

$$\bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^k.$$

כלומר התנאי $X_n \xrightarrow{\text{a.s.}} X$ שקול לכך שלכל k מתקיים

$$\mathbb{P}\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^k\right) = \mathbb{P}\left(\liminf_{n\to\infty} A_n^k\right) = 1.$$

לעומת זאת, התנאי $X_n \xrightarrow{\text{p}} X$ שקול לכך שלכל k מתקיים

$$\lim_{n\to\infty} \mathbb{P}(A_n^k) = 1.$$

ואולם לפי הלמה של פאטו

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n^k\right) \le \liminf_{n\to\infty} \mathbb{P}\left(A_n^k\right)$$

ולכן אם צד שמאל שווה ל-1 אז גם צד ימין. ■

הדוגמאות הבאות ממחישות כי התכנסות בהסתברות אינה גוררת התכנסות כמעט-תמיד.

דוגמא 8.15 (התכנסות כמעט תמיד לעומת התכנסות בהסתברות). יהיו $\{A_n\}_{n\in\mathbb{N}}$ מאורעות בלתי תלויים בעלי הסתברויות $p_n=\mathbb{P}(A_n)$ ונגדיר $X_n=\mathbb{1}(A_n)$. מתי $X_n \xrightarrow{\text{p}} 0$ ומתי $X_n \xrightarrow{\text{a.s.}} 0$?

תשובה: אם $p_n \to 0$ אז לכל $\epsilon>0$ מתקיים $\mathbb{P}(|X_n-0|>\epsilon)=\mathbb{P}(X_n=1)=p_n\to 0$ ולכן $X_n \xrightarrow{\text{p}} 0$ במקרה זה.

לעומת זאת, הסדרה $X_n(\omega)$ מקבלת ערכים שהם אפס או אחד בלבד ולכן היא מתכנסת לאפס אם ורק אם היא שווה לאפס החל ממקום מסוים, או במילים אחרות, בדיוק כאשר מתקיים $\{A_n^c \text{ a.e.}\}$. זה שקול למאורע $\{A_n \text{ i.o.}\}^c$ ולפי הלמות של בורל-קנטלי מאורע זה קורה בהסתברות 1 אם ורק אם $\sum_{n\in\mathbb{N}} p_n<\infty$.

דוגמא 8.16. יהי $Y\sim \mathrm{Unif}([0,1])$ ויהיו $X_{m,n}=\mathbb{1}(Y\in[\frac{m-1}{2^n},\frac{m}{2^n}])$ עבור $m,n\in\mathbb{N}$ המקיימים $m\le 2^n$. נראה כי

$$X_{1,n} \xrightarrow{\text{a.s.}} 0,\qquad X_{1,n} \xrightarrow{\text{p}} 0,\qquad X_{n,\lceil\log_2 n\rceil} \xrightarrow{\text{p}} 0,\qquad X_{n,\lceil\log_2 n\rceil} \not\xrightarrow{\text{a.s.}} 0.$$

תשובה: נראה כי $X_{1,n} \xrightarrow{\text{a.s.}} 0$. לפי טענה 8.13, עלינו להראות כי לכל $\epsilon>0$ מתקיים

$$\mathbb{P}\left(\forall n_0\in\mathbb{N}\ \exists n>n_0\ :\ |X_{1,n}|>\epsilon\right)=0.$$

נשים לב כי

$$\mathbb{P}\left(\exists n>n_0\ :\ |X_n|>\epsilon\right)\le\sum_{n=n_0+1}^{\infty}\mathbb{P}\left(|X_n|>\epsilon\right)=\sum_{n=n_0+1}^{\infty}\mathbb{P}\left(Y\in\left[0,\frac{1}{2^n}\right]\right)=\sum_{n=n_0+1}^{\infty}2^{-n}=2^{-n_0}.$$

ולכן

$$\mathbb{P}\left(\forall n_0\in\mathbb{N}\ \exists n>n_0\ :\ |X_{1,n}|>\epsilon\right)\overset{\substack{\text{משפט}\\ 1.20}}{=}\lim_{n_0\to\infty}\mathbb{P}\left(\exists n>n_0\ :\ |X_{1,n}|>\epsilon\right)\le\lim_{n_0\to\infty}\left(2^{-n_0}\right)=0,$$

כנדרש.

נראה כי $X_{1,n} \xrightarrow{\text{p}} 0$. עלינו להראות כי לכל $\epsilon,\eta>0$, נוכל לבחור $n_0\in\mathbb{N}$ כך שלכל $n>n_0$ יתקיים $\mathbb{P}\left(|X_{1,n}|>\epsilon\right)<\eta$. נחשב

$$\mathbb{P}\left(|X_{1,n}|>\epsilon\right)=\mathbb{P}\left(Y\in\left[\frac{n-1}{2^n},\frac{n}{2^n}\right]\right)=2^{-n}$$

ולכן עבור $n>-\log_2(\eta)$, נקבל $\mathbb{P}\left(|X_{1,n}|>\epsilon\right)<\eta$ כנדרש.

נראה כי $X_{n,\lceil\log n\rceil} \xrightarrow{\text{p}} 0$. כמקודם נווכח כי $X_{n,\lceil\log n\rceil} \xrightarrow{\text{p}} 0$. נחשב כמקודם

$$\mathbb{P}\left(|X_{n,\lceil\log n\rceil}|>\epsilon\right)=\mathbb{P}\left(Y\in\left[\frac{n-1}{2^{\lceil\log n\rceil}},\frac{n}{2^{\lceil\log n\rceil}}\right]\right)=2^{-\lceil\log n\rceil}\le\frac{1}{n}$$

ולכן עבור $n>\frac{1}{\eta}$, נקבל $\mathbb{P}\left(|X_{n,\lceil\log n\rceil}|>\epsilon\right)<\eta$ כנדרש.

לסיום, **נראה כי** $X_{n,\lceil\log n\rceil} \not\xrightarrow{\text{a.s.}} 0$. לצורך כך עלינו להראות כי

$$\mathbb{P}\left(\forall n_0\in\mathbb{N}\ \exists n>n_0\ :\ |X_{n,\lceil\log n\rceil}|>\epsilon\right)>0.$$

נסמן $E_{n_0}=\{\exists n>n_0\ :\ |X_{n,\lceil\log n\rceil}|>\epsilon\}$ ו- $A_{n,m}=\{Y\in\left[\frac{m-1}{2^{\lceil\log n\rceil+1}},\frac{m}{2^{\lceil\log n\rceil+1}}\right]\}$. יהי $n_0\in\mathbb{N}$ נשים לב כי

$$\mathbb{P}\left(E_{n_0}\right) = \sum_{m=1}^{2^{\lceil \log n \rceil+1}} \mathbb{P}\left(A_{n,m}\right) \mathbb{P}\left(E_{n_0} \,\middle|\, A_{n,m}\right) \geq \sum_{m=2^{\lceil \log n \rceil}+1}^{2^{\lceil \log n \rceil+1}} \mathbb{P}\left(A_{n,m}\right) \mathbb{P}\left(E_{n_0} \,\middle|\, A_{n,m}\right)$$

והרי עבור $m \in [2^n + 1, 2^{n+1}]$ מתקיים

$$\mathbb{P}\left(E_{n_0} \,\middle|\, A_{n,m}\right) \geq \mathbb{P}\left(\left|X_{m,\lceil \log m \rceil}\right| = 1 \,\middle|\, A_{n,m}\right) = \mathbb{P}\left(\left|X_{m,\lceil \log n \rceil+1}\right| = 1 \,\middle|\, A_{n,m}\right) = 1$$

ולכן

$$\mathbb{P}\left(\mathbb{E}_{n_0}\right) \geq \sum_{m=2^{\lceil \log n \rceil}+1}^{2^{\lceil \log n \rceil+1}} \mathbb{P}\left(Y \in \left[\frac{m-1}{2^{n+1}}, \frac{m}{2^{n+1}}\right]\right) = \frac{1}{2},$$

ולכן $X_{n,\lceil \log n \rceil} \overset{\text{a.s.}}{\not\to} 0$, כנדרש.

את החוק החלש של המספרים הגדולים (משפט 5.8) נוכל לנסח מחדש במונחים של התכנסות בהסתברות.

8.3 חוקי המספרים הגדולים

משפט 8.17 (החוק החלש של המספרים הגדולים). תהי $\{X_n\}_{n\in\mathbb{N}}$ סדרת משתנים מקריים בלתי-תלויים ושווי התפלגות, שתוחלתם μ. אזי

$$\frac{1}{n}\sum_{i=1}^{n} X_i \overset{\text{p}}{\to} \mu$$

כאשר אנו מזהים את μ עם המשתנה המקרי המקבל את הערך μ בהסתברות 1.

כזכור, הוכחנו את החוק תחת התנאי הנוסף שלמשתנים המקריים שונות סופית בטענה 5.8 ולמשתנים בעלי תוחלת בטענה 5.32. כעת נוכל לתאר גם את החוק החזק של המספרים הגדולים. להלן נביא את טענת החוק תחת ההנחה המגבילה שהמשתנים המקריים חסומים, ובפרק 8.4 נציג את החוק בגרסתו הכללית.

משפט 8.18 (החוק החזק של המספרים הגדולים). יהי $M > 0$. תהא $\{X_n\}_{n\in\mathbb{N}}$ סדרת משתנים מקריים בלתי-תלויים ושווי התפלגות כך ש-$|X_i| \overset{\text{a.s.}}{\leq} M$ ו-$\mathbb{E}(X_i) = \mu$ עבור $\mu \in \mathbb{R}$. אזי

גרסא מוגבלת
של החוק החזק

$$\frac{1}{n}\sum_{i=1}^{n} X_i \overset{\text{a.s.}}{\to} \mu$$

כאשר אנו מזהים את μ עם המשתנה המקרי המקבל את הערך μ בהסתברות 1.

הוכחה. נגדיר משתנים מקריים חדשים

$$Y_n = \frac{X_n - \mu}{2M}$$

ונשים לב שהם מקיימים את תנאי אי-שיוויון הופדינג (משפט 6.9 ??). לכן, לכל n ולכל $a > 0$ נקבל

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} Y_i\right| \geq a\right) \leq 2\exp\left(-\frac{a^2}{2n}\right).$$

עבור $\epsilon > 0$ כלשהו, אם נציב $a = \epsilon n$ נקבל

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Y_i\right| \geq \epsilon\right) \leq 2\exp\left(-\frac{\epsilon^2}{2}n\right).$$

נסמן

$$A_n^k = \left\{\left|\frac{1}{n}\sum_{i=1}^{n} Y_i\right| \geq 1/k\right\}$$

ולפי הלמה הראשונה של בורל-קנטלי (משפט 8.5), נקבל

$$\mathbb{P}(\limsup_{n\to\infty} A_n^k) = 0$$

. ניקח איחוד על כל ערכי k ונקבל

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} \limsup_{n\to\infty} A_n^k\right) = 0$$

או, באופן שקול

$$\mathbb{P}\left(\bigcap_{k=1}^{\infty} \liminf_{n\to\infty} (A_n^k)^c\right) = 1$$

וזו ההגדרה של $\frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{\text{a.s.}} 0$ וכמובן ש-$\frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{\text{a.s.}} 0$ אם ורק אם $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} \mu$. ■

*8.4 החוק החזק של המספרים הגדולים

הוכחה לא למבחן

בפרק זה נוכיח את גרסתו המלאה של החוק החזק של המספרים הגדולים.

משפט 8.19 (החוק החזק של המספרים הגדולים)**.** יהי $M > 0$ תהי $\{X_n\}_{n\in\mathbb{N}}$ סדרת משתנים מקריים בלתי-תלויים ושווי התפלגות. אזי $\mathbb{E}(X_i) = \mu$ עבור $\mu \in \mathbb{R}$ אם ורק אם

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} \mu.$$

לצורך הוכחת המשפט ניעזר בהגדרה הבאה ובשתי טענות הנוגעות אליה.

הגדרה 8.20 (זהות המ"מ)**.**שתי סדרות מ"מ $\{X_n\}_{n\in\mathbb{N}}$ ו-$\{Y_n\}_{n\in\mathbb{N}}$ נקראות זהות המ"מ (Tail Equivalent) אם

$$\mathbb{P}(X_n = Y_n \text{ a.e.}) = 1.$$

טענה 8.21. אם $\{X_n\}_{n\in\mathbb{N}}$ ו-$\{Y_n\}_{n\in\mathbb{N}}$ זהות המ"מ אז לכל סדרה $a_n \to \infty$ ולכל משתנה מקרי Z מתקיים

$$\lim_{n\to\infty}\sum_{m=1}^{n}\frac{X_m}{a_n} \xrightarrow{\text{a.s.}} Z \iff \lim_{n\to\infty}\sum_{m=1}^{n}\frac{Y_m}{a_n} \xrightarrow{\text{a.s.}} Z.$$

הוכחה. לאור הסימטריה בין שני צידי השקילות, די שנניח כי

$$\lim_{n\to\infty}\sum_{m=1}^{n}\frac{X_m}{a_n}\xrightarrow{\text{a.s.}} Z$$

ונוכיח כי

$$\lim_{n\to\infty}\sum_{m=1}^{n}\frac{Y_m}{a_n}\xrightarrow{\text{a.s.}} Z.$$

יהי $n_0=\min\{n\ :\ \forall m>n\ X_m=Y_m\}$. נרשום

$$\lim_{n\to\infty}\sum_{m=1}^{n}\frac{Y_m}{a_n}\overset{\text{a.s.}}{=}\lim_{n\to\infty}\sum_{m=1}^{n_0-1}\frac{Y_n}{a_n}+\lim_{n\to\infty}\sum_{m=n_0}^{\infty}\frac{X_m}{a_n}=0+\lim_{n\to\infty}\sum_{m=n_0}^{\infty}\frac{X_m}{a_n}=\lim_{n\to\infty}\sum_{m=1}^{n}\frac{X_m}{a_n}\overset{\text{a.s.}}{=}Z,$$

כנדרש. ■

טענה 8.22. תהי $\{X_n\}_{n\in\mathbb{N}}$ סדרת משתנים מקריים בלתי-תלויים ותהי $\{x_n\}_{n\in\mathbb{N}}$ סדרה של מספרים ממשיים. אזי מתקיים $\sum_{n=1}^{\infty}\mathbb{P}(|X_n|>x_n)<\infty$ אם ורק אם $\{X_n\}_{n\in\mathbb{N}}$ ו- $\{X_n\cdot\mathbb{1}_{X_n\in[-x_n,x_n]}\}_{n\in\mathbb{N}}$ זהות המ"מ.

הוכחה. המאורעות $A_n=\{|X_n|>x_n\}$ הנם בלתי-תלויים, ולכן לפי שילוב שתי הלמות של בורל-קנטלי (משפטים 8.5 ו-8.6), $\mathbb{P}(\{A_n \text{ i.o.}\})=1$ אם ורק אם $\sum_{n=1}^{\infty}\mathbb{P}(|X_n|>x_n)=\infty$. ולפי הגדרה טענות אלה שקולות לכך ש-$\mathbb{P}(\{A_n^c \text{ a.e.}\})=0$, וזהו בדיוק המאורע ש-$\{X_n\}_{n\in\mathbb{N}}$ ו- $\{X_n\cdot\mathbb{1}_{X_n\in[-x_n,x_n]}\}_{n\in\mathbb{N}}$ זהות המ"מ. ■

הוכחה של משפט 8.19: קיום תוחלת גורר התכנסות כ"ת של סדרת הממוצעים. נסמן לאורך ההוכחה $S_n=\frac{1}{n}(X_1+\cdots+X_n)$ להיות סדרת הממוצעים האמפיריים (שאת התכנסותה אנו רוצים להוכיח). ההוכחה נחלקת למספר שלבים.

שלב 1: אי-שליליות. ניתן להניח כי $X\geq 0$ כמעט תמיד. זאת מכיוון שבאופן כללי $X=X_+-X_-$ כאשר X_+ ו-X_- הם משתנים מקריים אי-שליליים בעלי תוחלת, לכן מספיק להוכיח את המשפט לכל אחד מהם.

שלב 2: דילול. נשים לב שסדרת הממוצעים האמפריים אינה יכולה להשתנות מהר מדי. באופן פורמלי, עבור אינדקסים $n,m\in\mathbb{N}$ ומספר קטן $\epsilon>0$ כך ש- $n\leq m\leq(1+\epsilon)n$ מתקיים

$$S_m=\frac{1}{m}(X_1+\cdots+X_m)\leq\frac{n}{m}\cdot\frac{1}{n}(X_1+\cdots+X_m+\cdots+X_n)\leq(1+\epsilon)S_n.$$

ככל ש- ϵ קטן יותר (כלומר האינדקסים n ו-m נבדלים בקבוע כפלי קרוב יותר ל-1) הרי שהפער הכפלי בין S_m ל-S_n שואף ל-1. לפיכך, די להוכיח את ההתכנסות של הסדרה $\{S_{n_j}\}_{j\in\mathbb{N}}$ עבור כל $c>1$ וכל סדרת אינדקסים עולה n_j המקיימת $\frac{n_{j+1}}{n_j}\geq c$. סדרת אינדקסים כזו נקראת לקונרית והיא מתאפיינת בכך ש- $\sum_{j=1}^{\infty}\frac{1}{n_j}$.

שלב 3: בורל-קנטלי. לפי למת בורל-קנטלי הראשונה, משפט 8.5, די להראות כי:

$$\sum_{j=1}^{\infty}\mathbb{P}\left(|S_{n_j}-\mathbb{E}X|>\epsilon\right)<\infty, \tag{*}$$

לכל $\epsilon>0$ ולכל סדרה לקונרית $\{n_j\}$. (או אז ינבע כי החל ממקום מסויים לכל j, $|S_{n_j}-\mathbb{E}X|\leq\epsilon$ ולפי שלב 2, ינבע כי החל ממקום מסויים גם לכל m, $|S_m-\mathbb{E}X|\leq\epsilon$).

שלב 4: השוואה למשתנים קטומים זהים כב"מ. לו היה X בעל שונות סופית, אז בעזרת אי-שיוויון צ'בישב (משפט 5.5) היינו מקבלים כי

$$\mathbb{P}\left(|S_{n_j} - \mathbb{E}X| > \epsilon\right) \le \frac{\mathrm{Var}(S_{n_j})}{\epsilon^2} = \frac{n_j \mathrm{Var}(X)}{n_j^2\epsilon^2} = \frac{\mathrm{Var}(X)}{n_j\epsilon^2}.$$

מכיוון ש- n_j סדרה לקונרית, $\sum_{j=1}^{\infty}\frac{1}{n_j} < \infty$, ולכן היינו מקבלים כי (*) מתקיים.

באופן כללי, אי אפשר להניח כי X חסום או בעל שונות סופית, אך ניתן לקרבו על-ידי משתנים כאלה. נשתמש בסימון הבא: לכל מספר $L > 0$ נגדיר $X_{\le L} = X \cdot \mathbb{1}_{X\le L}$. נשים לב כי מתקיים $\sum_{j=1}^{\infty} n_j\mathbb{P}(X \ge n_j) \le \mathbb{E}[X] < \infty$. ולכן אם נגדיר $j(i) = \min\{j > i \ : \ n_j \ge i\}$ נקבל לפי טענה 8.22 כי הסדרה $Y_n = X_n \cdot \mathbb{1}_{X_n \le n_{j(n)}}$ זהה כב"מ לסדרה X_n. נגדיר $\overline{S}_n = \frac{1}{n}(Y_1 + \cdots + Y_n)$ ונקבל, לפי טענה 8.21, כדי להוכיח כי סדרת ממוצעים זו מתכנסת בהסתברות לתוחלת של X. כלומר, לפי בורל-קנטלי, די להראות

$$\sum_{j=1}^{\infty}\mathbb{P}\left(|\overline{S}_{n_j} - \mathbb{E}X| > \epsilon\right) < \infty, \qquad (**)$$

קיים j_0 כך ש- $|\mathbb{E}\overline{S}_{n_j} - \mathbb{E}X| = |\mathbb{E}X\mathbb{1}_{X\le n_j}| < \epsilon/2$. לפי צ'בישב נקבל כי לכל $j \ge j_0$,

$$\mathbb{P}\left(|\overline{S}_{n_j} - \mathbb{E}X| > \epsilon\right) \le \mathbb{P}\left(|\overline{S}_{n_j} - \mathbb{E}\overline{S}_{n_j}| > \epsilon/2\right) \le \frac{4\,\mathrm{Var}[|X_{\le n_j}|^2]}{n_j\epsilon^2} \le \frac{4\mathbb{E}[|X_{\le n_j}|^2]}{n_j\epsilon^2}$$

להוכחת (**), די אפוא אם נוכיח כי $\sum_{j=1}^{\infty}\mathbb{E}[X_{\le n_j}^2/n_j] < \infty$. לפי משפט ההתכנסות המונוטונית ניתן להחליף סכום ואינטגרל בטור החיובי הבא:

$$\begin{aligned}\sum_{j=1}^{\infty}\frac{1}{n_j}\mathbb{E}[X_{\le n_j}^2] &= \sum_{j=1}^{\infty}\int_{\Omega} X^2(\omega)\frac{1}{n_j}\mathbb{1}\left\{X(\omega) \le n_j\right\}d\mathbb{P}(\omega) \\ &= \int_{\Omega} X^2(\omega)\sum_{j=1}^{\infty}\frac{1}{n_j}\mathbb{1}\left\{X(\omega) \le n_j\right\}d\mathbb{P}(\omega)\end{aligned}$$

נסדר מחדש את הסכימה ונשתמש כעת בתנאי הלקונריות של הסדרה $\{n_j\}$:

$$\begin{aligned}\sum_{j=1}^{\infty}\frac{1}{n_j}\mathbb{E}[X_{\le n_j}^2] &= \int_{\Omega} X^2(\omega)\sum_{j=1}^{\infty}\left(\frac{1}{n_j} + \frac{1}{n_{j+1}} + \dots\right)\mathbb{1}\left\{n_{j-1} < X(\omega) \le n_j\right\}d\mathbb{P}(\omega) \\ &\le \int_{\Omega} X(\omega)\left(1 + c^{-1} + \cdots + c^{-\ell} + \dots\right)\sum_{j=1}^{\infty}\frac{X(\omega)}{n_j}\mathbb{1}\left\{n_{j-1} < X(\omega) \le n_j\right\}d\mathbb{P}(\omega) \\ &\le \frac{c}{c-1}\int_{\Omega} X(\omega)\sum_{j=1}^{\infty}\mathbb{1}\left\{n_{j-1} \le X(\omega) \le n_j\right\}d\mathbb{P}(\omega) \le \frac{c}{c-1}\mathbb{E}[X] < \infty,\end{aligned}$$

כנדרש. ∎

הוכחה. הוכחה של משפט 8.19 : התכנסות כ"ת של ממוצעים גוררת קיום תוחלת] נסמן כמקודם $S_n = \frac{1}{n}(X_1 + \cdots + X_n)$. נניח כי קיים $c \in \mathbb{R}$ כך ש- $S_n \xrightarrow{\text{a.s.}} c$. מתקיים

$$\frac{X_n}{n} = \frac{X_1 + \cdots + X_{n-1} + X_n}{n} - \frac{X_1 + \cdots + X_{n-1}}{n-1} \cdot \frac{n-1}{n} = S_n - \frac{n-1}{n} S_{n-1} \xrightarrow{\text{a.s.}} c - 1 \cdot c = 0.$$

בפרט $\mathbb{P}\left(\left\{\left|\frac{X_n}{n}\right| \geq 1\right\} \text{ i.o.}\right) = 0$. לפי הלמה השניה של בורל-קנטלי, משפט 8.6, נגיע למסקנה כי

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\frac{X_n}{n}\right| \geq 1\right) = \sum_{n=1}^{\infty} \mathbb{P}\Big(|X| \geq n\Big) < \infty.$$

ואולם, לפי הגדרת התוחלת,

$$\mathbb{E}|X| \leq \sum_{n=1}^{\infty} n\mathbb{P}\Big(n-1 \leq |X| \leq n\Big) = \sum_{k=0}^{\infty} \mathbb{P}\Big(|X| \geq k\Big) < \infty,$$

כלומר X בעל תוחלת כנדרש. ■

בעיות הרחבה והעשרה

✎ **בעיה 8.1.** יהיו $\{A_n\}_n \in \mathbb{N}$, $\{B_n\}_n \in \mathbb{N}$ מאורעות על מרחב הסתברות.

(א) יש להראות כי אם $\mathbb{P}(A_n \text{ a.e.}) = \mathbb{P}(B_n \text{ a.e.}) = 1$ אז $\mathbb{P}(A_n \cap B_n \text{ a.e.}) = 1$.

(ב) יש למצוא דוגמא שבה $\mathbb{P}(A_n \text{ a.e.}) = \mathbb{P}(B_n \text{ a.e.}) = 0$ אך $\mathbb{P}(A_n \cup B_n \text{ a.e.}) = 1$.

✎ **בעיה 8.2.** תהי $\{A_n\}_n \in \mathbb{N}$ סדרת מאורעות על מרחב הסתברות. נשתמש ברישום $A \oplus B = (A \setminus B) \cup (B \setminus A)$. יש להוכיח כי לכל $\epsilon > 0$ ו-$N \in \mathbb{N}$ קיימים $n \geq m > N$ המקיימים

$$\mathbb{P}\left(\{A_n \text{ a.e.}\} \oplus \left(\bigcap_m^n A_n\right)\right) < \epsilon$$

✎ **בעיה 8.3.** יהיו $\{X_n\}_n \in \mathbb{N}$, X משתנים מקריים על מרחב הסתברות. יש להראות כי אם $X_n \xrightarrow{\text{p}} X$ אז קיימת תת סדרה $\{m_i\}_{i \in \mathbb{N}}$ כך ש-$X_{m_i} \xrightarrow{\text{a.s.}} X$.

✎ **בעיה 8.4.** יהיו $\{X_n\}_n \in \mathbb{N}$, X, $\{Y_n\}_n \in \mathbb{N}$, Y משתנים מקריים על מרחב הסתברות.

(א) יש להראות כי אם $X_n \xrightarrow{\text{a.s.}} X$ ו-$Y_n \xrightarrow{\text{a.s.}} Y$ אז $X_n + Y_n \xrightarrow{\text{a.s.}} X_n + Y_n$.

(ב) יש להראות כי אם $X_n \xrightarrow{\text{p}} X$ ו-$Y_n \xrightarrow{\text{p}} Y$ אז $X_n + Y_n \xrightarrow{\text{p}} X_n + Y_n$.

✎ **בעיה 8.5 (*).** [הלמה המותנית של בורל-קנטלי] יהיו $\{A_n\}_n \in \mathbb{N}$ מאורעות על מרחב הסתברות ונניח כי לכל $B_1, \ldots, B_n$ כך ש-$B_i \in \{A_i, A_i^c\}$ מתקיים $\mathbb{P}(A_n \mid B_1, \ldots, B_{n-1}) \geq a_n$ עבור $\{a_n\}_n \in \mathbb{N}$ המקיימים $\sum_{n \in \mathbb{N}} a_n = \infty$. אז $\mathbb{P}(A_n \text{ i.o.}) = 1$.

✎ **בעיה 8.6 (*).** סיזיפוסה, התרנגולת בת-האלמוות, נכנסת ללול בגודל אינסופי. בכל בוקר היא מטילה שלוש ביצים. בכל ערב היא בוחרת באקראי את אחת הביצים השלמות בלול, דוגרת עליה ומבקיעה אותה. מה הסיכוי שבסופו של דבר תוטלנה בלול ביצים שתשארנה שלמות עד קץ-הימים.