הכנה למבחן — משפטים והוכחות נבחרים — תורת ההסתברות 1, 80420

31 בינואר 2026

תוכן עניינים

1 שיטות בסיסיות

1.1 רציפות פונקציית ההסתברות (2.15)

משפט 1.1 (רציפות פונקציית ההסתברות): יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ותהיי $\{A_n\}_{n\in\mathbb{N}}$ סדרה עולה של מאורעות. אז מתקיים

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

הוכחה: נקבע $B_1 = A_1$ ולכל $n > 1$ נגדיר $B_n = A_n \setminus A_{n-1}$ ואלו בהכרח מאורעות זרים:
כי אם $m < n$ אז לכל $\omega \in B_n$ מתקיים $\omega \notin A_{n-1}$ ולכן מתקיים $\omega \notin A_m \supset B_m$.
מצד שני, באינדוקציה

$$(\star)\ \ \dot{\bigcup_{k\in[n]}} B_k = \bigcup_{k\in[n]} A_k = A_n$$

עבור $A_1 = B_1$ הטענה מיידית, נניח כי היא מתקיימת עבור $n \geq 1$ ונקבל

$$\dot{\bigcup_{k\in[n+1]}} B_k = \left(\dot{\bigcup_{k\in[n]}} B_k\right) \bigcup B_{n+1} \underset{\text{הנחת האינדוקציה}}{=} A_n \bigcup (A_{n+1} \setminus A_n) \underset{A_n \subset A_{n+1}}{=} A_{n+1}$$

ולכן

$$(\star\star)\ \ \bigcup_{k\in\mathbb{N}} B_k = \bigcup_{k\in\mathbb{N}} A_k$$

אם־כך מסכימות נקבל

$$\mathbb{P}\left(\bigcup_{k\in\mathbb{N}} A_k\right) \underset{(\star\star)}{=} \mathbb{P}\left(\dot{\bigcup_{k\in\mathbb{N}}} B_k\right) \underset{\text{סכימות}}{=} \sum_{k=1}^{\infty} \mathbb{P}(B_k) \underset{\text{הגדרת הטור}}{=} \lim_{n\to\infty} \sum_{k\in[n]} \mathbb{P}(B_k) \underset{\text{סכימות}}{=} \lim_{n\to\infty} \mathbb{P}\left(\dot{\bigcup_{k\in[n]}} B_k\right) \underset{(\star)}{=} \lim_{n\to\infty} \mathbb{P}(A_n)$$

מפתח להוכחה:

1. מוכיחים הזרת מאורעות באינדוקציה

$$B_1 = A_1,\ B_n = A_n \setminus A_{n-1}$$

2. סכימות בת־מנייה של מאורעות זרים
3. הגדרת הגבול

$\square$

1.2 אי־שיוויון בול (2.18)

משפט 1.2 (אי־שיוויון בול למספר מאורעות): לכל $m \in \mathbb{N}$ ולכל סדרה של m מאורעות $\{A_n\}_{n\in[m]}$ במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ מתקיים

$$\mathbb{P}\left(\bigcup_{n\in[m]} A_n\right) \leq \sum_{n\in[m]} \mathbb{P}(A_n)$$

הוכחה: באינדוקציה על m, עבור $m = 2$ בסיס האינדוקציה: יהיו A, B מאורעות כנ"ל אז $A \cup B = A \uplus (B \setminus A)$ ולכן

$$\mathbb{P}(A \cup B) \underset{\substack{\text{סכימות פונקציית ההסתברות} \\ \text{למאורעות זרים}}}{=} \mathbb{P}(A) + \mathbb{P}(B \setminus A) \underset{\text{מונוטוניות פונקציית ההסתברות}}{\leq} \mathbb{P}(A) + \mathbb{P}(B)$$

כעת נניח את נכונות הטענה עבור m ונוכיחה עבור $m+1$: יהיו $A_1, ..., A_{m+1}$ מאורעות ונפעיל את הטענה עבור שני מאורעות $A_{m+1}, \cup_{i=1}^m A_i$ ונקבל

$$\mathbb{P}\left(\bigcup_{i=1}^{m+1} A_i\right) \leq \mathbb{P}\left(\bigcup_{i=1}^{m} A_i\right) + \mathbb{P}(A_{m+1}) \underset{\text{הנחת האינדוקציה}}{\leq} \sum_{i=1}^{m+1} \mathbb{P}(A_i)$$

עבור מאורעות יורדים, נשתמש בהיות המשלים שלהם מאורעות עולים.

מפתח להוכחה: אינדוקציה שבבסיס משתמשים בהזרה ותכונות פונקציית ההסתברות. $\square$

משפט 1.3 (אי־שיוויון בול לסדרת מאורעות): לכל סדרת מאורעות $\{A_n\}_{n\in\mathbb{N}}$ במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ מתקיים

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$$

הוכחה: נגדיר $B_k = \bigcup_{k\in[n]} A_k$ וזו סדרת מאורעות עולה המקיימת $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n$, אז

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} B_n\right) \underset{\text{רציפות פונקציית ההסתברות}}{=} \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\in[n]} A_k\right) \underset{\text{אי־שיוויון בול}}{\leq} \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(A_k) = \sum_{k=1}^{\infty} \mathbb{P}(A_k)$$

מפתח להוכחה: הגדרת $B_k = \bigcup_{k\in[n]} A_k$, שימוש ברציפות פונקציית ההסתברות ובאי־שיוויון בול. $\square$

2 הסתברות מותנית

2.1 נוסחת ההסתברות השלמה במונחי הסתברות מותנית (3.18)

משפט 2.1 (נוסחת ההסתברות השלמה במונחי הסתברות מותנית): תהיי $\mathcal{A}$ חלוקה בת־מנייה של מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. אז לכל מאורע B מתקיים

$$\mathbb{P}(B) = \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(B \mid A)\mathbb{P}(A)$$

הוכחה: נתזכר את כלל השרשרת: יהיו A, B מאורעות במרחב ההסתברות כך שמתקיים $\mathbb{P}(B) > 0$, אז

$$\mathbb{P}(A \mid B)\mathbb{P}(B) = \mathbb{P}(A \cap B)$$

נשתמש בנוסחת ההסתברות השלמה ונקבל

$$\begin{aligned}
\mathbb{P}(B) &= \sum_{A \in \mathcal{A}} \mathbb{P}(A \cap B) \\
&= \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(A \cap B) + \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) = 0}} \mathbb{P}(A \cap B) \\
&\underset{(\star)}{=} \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(B \mid A)\mathbb{P}(A) + \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) = 0}} 0 \\
&= \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(B \mid A)\mathbb{P}(A)
\end{aligned}$$

כאשר $(\star)$ נובע מכלל השרשרת.

מפתח להוכחה: נוסחת ההסתברות השלמה וכלל השרשרת. $\square$

2.2 כלל בייס (3.20)

משפט 2.2 (כלל בייס): יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו A, B שני מאורעות בעלי הסתברות חיובית, אזי

$$\mathbb{P}(A \mid B)\mathbb{P}(B) = \mathbb{P}(B \mid A)\mathbb{P}(A)$$

או בניסוח אחר

$$\mathbb{P}(A \mid B) = \mathbb{P}(B \mid A)\frac{\mathbb{P}(A)}{\mathbb{P}(B)}$$

הוכחה:

$$\mathbb{P}(B)\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B)\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A \cap B) = \frac{\mathbb{P}(A)\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \mathbb{P}(B \mid A)\mathbb{P}(A)$$

מפתח להוכחה: לחשב כל פעם לבד לפי הגדרת ההסתברות המותנית. $\square$

3 יחסים בין משתנים מקריים

3.1 אי־תלות נשמרת תחת הפעלת פונקציות (4.89)

משפט 3.1 (אי־תלות נשמרת תחת הפעלת פונקציות): יהיו $X_1, ..., X_n$ וקטורים מקריים בלתי־תלויים כאשר X_i הוא וקטור d_i־מימדי ותהיינה $f_i \in \mathcal{F}_{\mathbb{R}^{d_i} \to \mathbb{R}^{s_i}}$ עבור s_i כלשהם. אזי $f_1(X_1), ..., f_n(X_n)$ בלתי־תלויים

הוכחה: תהיינה $A_i \in \mathcal{F}_{\mathbb{R}^{s_i}}$ עבור $i \in [n]$, אזי

$$\mathbb{P}(\forall i \in [n],\ f_i(X_i) \in A_i) = \mathbb{P}(\forall i \in [n],\ X_i \in f_i^{-1}(A_i)) \underset{\text{אי־תלות}}{=} \prod_{i \in [n]} \mathbb{P}(X_i \in f_i^{-1}(A_i)) = \prod_{i \in [n]} \mathbb{P}(f_i(X_i) \in A_i)$$

מפתח להוכחה: עדיף להסתכל על המקורות תחת הפונקציה ואז אפשר להשתמש באי־תלות. $\square$

3.2 שיוויון כמעט־תמיד גורר שיוויון התפלגויות (4.29)

משפט 3.2 (שוויון כמעט־תמיד גורר שיוויון התפלגויות): יהיו X, Y משתנים מקריים על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. אם $X \overset{a.s.}{=} Y$ אז $X \overset{d}{=} Y$.
תזכורת:

$$\mathbb{P}(X = Y) = \mathbb{P}(\{\omega \mid X(\omega) = Y(\omega)\}) = 1 \Longrightarrow X \overset{a.s.}{=} Y$$

$$\mathbb{P}_X = \mathbb{P}_Y \Longrightarrow X \overset{d}{=} Y$$

הוכחה: אם $X \overset{a.s.}{=} Y$ אז לכל $S \in \mathcal{F}_{\mathbb{R}}$ מתקיים לפי מונוטוניות $\mathbb{P}(X \in S, Y \notin S) \leq \mathbb{P}(X \neq Y) = 0$ ובדומה $\mathbb{P}(X \notin S, Y \in S) = 0$, אז

$$\mathbb{P}_X(s) = \mathbb{P}(X \in S) = \mathbb{P}(X \in S, Y \in S) + \mathbb{P}(X \in S, Y \notin S) = \mathbb{P}(X \in S, Y \in S) + \mathbb{P}(X \notin S, Y \in S) = \mathbb{P}(Y \in S) = \mathbb{P}_Y(S)$$

מפתח להוכחה: משתמשים בהכלת מאורעות מההנחה. $\square$

3.3 שיוויון התפלגויות נשמר תחת הפעלת פונקציה (4.31)

משפט 3.3 (שיוויון התפלגויות נשמר תחת הפעלת פונקציה): יהיו X, Y משתנים מקריים בדידים ושווי התפלגות (לאו דווקא על אותו מרחב הסתברות) ותהיי $f \in \mathcal{F}_{\mathbb{R}\to\mathbb{R}}$ אזי $f(X) \overset{d}{=} f(Y)$.

הוכחה: תהיי $S \subset \mathbb{R}$ אזי

$$\mathbb{P}_{f(X)}(S) = \mathbb{P}(f(X) = S) = \mathbb{P}\big(X \in f^{-1}(S)\big) \underset{X \overset{d}{=} Y}{=} \mathbb{P}\big(Y \in f^{-1}(S)\big) = \mathbb{P}(f(Y) \in S) = \mathbb{P}_{f(Y)}(S)$$

מפתח להוכחה: עדיף להסתכל על המקורות תחת הפונקציה ואז אפשר להשתמש בשיוויון התפלגויות. □

3.4 שיוויון כמעט־תמיד נשמר תחת הפעלת פונקציה

משפט 3.4 (שיוויון כמעט־תמיד נשמר תחת הפעלת פונקציה): יהיו X,Y משתנים מקריים בדידים המקיימים $X \overset{a.s.}{=} Y$ ותהיי $f \in \mathcal{F}_{\mathbb{R}\to\mathbb{R}}$ אזי $f(X) \overset{a.s.}{=} f(Y)$.

הוכחה:

מכך שמתקיים $X \overset{a.s.}{=} Y$ נובע שמתקיים $\mathbb{P}(X = Y) = \mathbb{P}(\{\omega \in X(\omega) = Y(\omega)\}) = 1$, כלומר $\mathbb{P}(X \neq Y) = 0$ מהגדרת המשלים.
נסמן

$$N := \{\omega \in \Omega \mid X(\omega) \neq Y(\omega)\}$$

נרצה להראות ש־$\mathbb{P}(f(X) \neq f(y)) = 0$, אז נגדיר

$$N_f := \{\omega \in \Omega \mid f(X(\omega)) \neq f(Y(\omega))\}$$

אם $\omega \in N$, מתקיים $X(\omega) \neq Y(\omega)$ ויכול להיות $f(X(\omega)) \neq f(Y(\omega))$ ויכול להיות $f(X(\omega)) \neq f(Y(\omega))$.
אם $\omega \notin N$ מתקיים $X(\omega) = Y(\omega)$ כמספרים ממשיים ולכן מהגדרת הפונקציה נובע שמתקיים בהכרח $f(X(\omega)) = f(Y(\omega))$, כלומר אם $\omega \notin N$ אז בהכרח $\omega \notin N_f$.
כלומר בהכרח מתקיים $N_f \subseteq N$ וממונוטוניות פונקציית ההסתברות מתקיים $\mathbb{P}(N_f) \leq \mathbb{P}(N) = 0$.
מפתח להוכחה: מראים שקבוצת הנקודות שבהם המשתנים המקריים לאחר הפעלת הפונקציה לא זהים מוכלת בקבוצת האיברים שבהם המשתנים המקריים לא זהים ואז ממונוטוניות (מגדירים קבוצה ממידה אפס ומבינים מה נמצא בה אחרי הפעלת הפונקציה). $\square$

4 משתנים מקריים בדידים

4.1 הצלחה ראשונה בסדרת ניסויי ברנולי בלתי־תלויים מתפלג גיאומטרי (4.101)

משפט 4.1 (הצלחה ראשונה בסדרת ניסויי ברנולי בלתי־תלויים מתפלג גיאומטרי): תהיי $\{X_k\}_{k\in\mathbb{N}}$ סדרה אינסופית של משתנים מקריים בלתי־תלויים כאשר $X_k \sim Ber(p)$ לכל $k \in \mathbb{N}$, נסמן

$$X = \min(\{k \mid X_k = 1\})$$

אז $X \sim Geo(p)$

הוכחה: $X(\omega)$ הוא האינדקס של המקום הראשון בסדרה $X_1(\omega), X_2(\omega), ...$ בו מופיע הערך 1 ואם כל איברי הסדרה מתאפסים נסמן $X(\omega) = \infty$.
נשים לב

$$\{X = k\} = \{X_1 = 0, ..., X_{k-1} = 0, X_k = 1\}$$

ולפי האי־תלות נקבל

$$\mathbb{P}(X = k) = \mathbb{P}(X_1 = 0, ..., X_{k-1} = 0, X_k = 1) = (1-p)^{k-1}p$$

כלומר $X \sim Geo(p)$ כנדרש ונשים לב שלפי אבחנה שראינו על מכפלה אינסופית

$$\mathbb{P}(X = \infty) = \prod_{k=1}^{\infty}(1-p) = \lim_{k\to\infty}(1-p)^k = 0$$

מפתח להוכחה: כותבים $\{X = k\} = \{X_1 = 0, ..., X_{k-1} = 0, X_k = 1\}$ ומהאי־תלות ההוכחה כותבת את עצמה. $\square$

4.2 תיאור משתנה גיאומטרי במונחים של התפלגות שיורית (4.105)

משפט 4.2 (תיאור משתנה גיאומטרי במונחים של התפלגות שיורית): משתנה מקרי שנתמך על השלמים מתפלג $Geo(p)$ אם ורק אם לכל $n \in \mathbb{N}$ מתקיים

$$\mathbb{P}(X > n) = (1-p)^n$$

הוכחה: $\Longleftarrow$ $X \sim Geo(p)$ ולכן

$$\mathbb{P}(X > n) = \sum_{k=n+1}^{\infty} \mathbb{P}(X = k) = \sum_{k=n+1}^{\infty} (1-p)^{k-1} p = (1-p)^n p \sum_{\ell=0}^{\infty} (1-p)^{\ell} \underset{\text{טור הנדסי}}{=} (1-p)^n$$

$\Longrightarrow$ לכל $n \in \mathbb{N}$

$$\mathbb{P}(X = n) = \mathbb{P}(\{X > n-1\} \setminus \{X > n\}) \underset{\text{ההנחה}}{=} (1-p)^{n-1} - (1-p)^n = (1-p)^{n-1} p$$

מפתח להוכחה: בכיוון הראשון לסדר אינדקס סכימה לטור הנדסי, בכיוון השני להשתמש בהכלת מאורעות כי זה על השלמים. $\square$

4.3 חוסר זיכרון של התפלגות גיאומטרית (4.107)

משפט 4.3 (חוסר זיכרון של התפלגות גיאומטרית):

הגדרה 4.1 (חוסר זיכרון לכישלון): משתנה מקרי X בדיד שנתמך על $\mathbb{N}$ נקרא **חסר זיכרון לכישלון** אם X ו־$X-1 \mid X>1$ שווי התפלגות.
כלומר, אם לכל $n\in\mathbb{N}$ מתקיים

$$\mathbb{P}(X\in S)=\mathbb{P}(X-1\in S\mid X>1)$$

לכל $S\in\mathcal{F}_{\mathbb{R}}$.

יהי X משתנה מקרי הנתמך על $\mathbb{N}$ המקיים $\mathbb{P}(X=1)<1$, אזי X חסר זיכרון לכשלונות אם ורק אם קיים $p\in(0,1)$ כך ש־$X\sim Geo(p)$

הוכחה: $\Longrightarrow$ נניח כי $X\sim Geo(p)$ ולכן לכל $n\in\mathbb{N}$

$$\mathbb{P}(X-1=n\mid X>1)\underset{\substack{\text{הסתברות מותנית}\\\text{והכלת מאורעות}}}{=}\frac{\mathbb{P}(X=n+1)}{\mathbb{P}(X>1)}=\frac{(1-p)^np}{1-p}=(1-p)^{n-1}p=\mathbb{P}(X=n)$$

$\Longleftarrow$ נניח כי $\mathbb{P}(X=n)=\mathbb{P}(X-1=n\mid X>1)$ לכל $n\in\mathbb{N}$ ונסמן $p:=\mathbb{P}(X=1)$ ולכן $\mathbb{P}(X>1)=1-p$, אז לכל $k\geq 1$

$$\mathbb{P}(X>k+1)\underset{\substack{\text{כלל השרשרת}\\\text{והכלת מאורעות}}}{=}\mathbb{P}(X>k+1\mid X>1)\mathbb{P}(X>1)=\mathbb{P}(X-1>k\mid X>1)\mathbb{P}(X>1)\underset{\text{ההנחה}}{=}\mathbb{P}(X>k)(1-p)$$

נמשיך באינדוקציה ונקבל

$$\mathbb{P}(X>k)=\mathbb{P}(X>k-1)(1-p)=\cdots=\mathbb{P}(X>1)(1-p)^{k-1}=(1-p)^k$$

שזו בידיוק ההגדרה של משתנה גיאומטרי במונחים של התפפלגות שיורית.

מפתח להוכחה:

1. בכיוון הראשון, הסתברות מותנית והכלת מאורעות כותב את ההוכחה
2. בכיוון השני
   1. מפתחים עם כלל השרשרת והכלת מאורעות עם ההנחה
   2. ממשיכים באינדוקציה
   3. הגדרת משתנה גיאומטרי לפי התפלגות שיורית

$\square$

4.4 סכום משתנים ברנולי בלתי־תלויים מתפלג בינומית (4.115)

משפט 4.4 (סכום משתנים ברנולי בלתי־תלויים מתפלג בינומית): יהיו $\{X_i\}_{i\in[n]}$ וקטור של משתני ברנולי עם הסתברות הצלחה p בלתי־תלויים, אזי

$$\sum_{i\in[n]} X_i \sim Bin(n,p)$$

הוכחה: יהי $k\in\{0,...,n\}$ ונסמן $Y=\sum_{i\in[n]} X_i$.
נסמן ב־A_k את אוסף הוקטורים ב־$\{0,1\}^n$ שבהם בדיוק k אחדות ו־$(n-k)$ אפסים, כלומר

$$A_k := \left\{(x_1,...,x_n)\in\{0,1\}^n \mid \sum_{i\in[n]} x_i = k\right\}$$

כך שמתקיים $|A_k|=\binom{n}{k}$ ונחשב

$$\mathbb{P}(Y=k)=\sum_{x\in A_k}\mathbb{P}(X=x)\underset{\text{אי־תלות}}{=}\sum_{x\in A_k}\left(\prod_{i\in[n]}\mathbb{P}(X_i)=x_i\right))=\binom{n}{k}p^k(1-p)^{n-k}$$

מפתח להוכחה:

1. מגדירים

$$A_k := \left\{(x_1,...,x_n)\in\{0,1\}^n \mid \sum_{i\in[n]} x_i = k\right\}$$

2. נוסחת ההסתברות השלמה על A_k כחלוקה של המרחב
3. אי־תלות

□

4.5 חיבור משתנים מקריים בינומיים בלתי־תלויים (4.116)

משפט 4.5 (חיבור משתנים מקריים בינומיים בלתי־תלויים)**:** אם $X \sim Bin(m,p)$ ו־$Y \sim Bin(n,p)$ בלתי־תלויים אזי $X+Y \sim Bin(m+n,p)$.

הוכחה: יהיו $B_1,...,B_{m+n}$ משתנים מקריים בלתי־תלויים המתפלגים ברנולי עם הסתברות הצלחה p, נסמן

$$X' = \sum_{k=1}^{m} B_k \qquad Y' = \sum_{k=m+1}^{m+n} B_k$$

אז לפי הטענה על סכום משתנים מקריים בלתי־תלויים שמתפלגים ברנולי p נקבל

$$X' \sim Bin(m,p), \qquad Y' \sim Bin(n,p), \qquad X'+Y' \sim Bin(m+n,p)$$

כך שמתקיים

$$X' \overset{d}{=} X \qquad Y' \overset{d}{=} Y$$

אלו פונקציות של קבוצות משתנים שונות באוסף של משתנים בלתי־תלויים ולכן X', Y' הם גם בלתי־תלויים כלומר לכל $a,b \in \mathbb{R}$ מתקיים

$$\mathbb{P}(X'=a, Y'=b) = \mathbb{P}(X'=a)\mathbb{P}(Y'=b) = \mathbb{P}(X=a)\mathbb{P}(Y=b) = \mathbb{P}(X=a, Y=b)$$

כלומר ההתפלגות המשותפת של X', Y' זהה לזו של X, Y, אבל שיוויון נשמר תחת הפעלת פונקציות ולכן

$$X'+Y' \overset{d}{=} X+Y$$

מפתח להוכחה:

1. סכום משתנים מקריים ברנולי מתפלג בינומית
2. שיוויון התפלגויות

□

4.6 פואסון כגבול של בינומי במובן הנקודתי (4.126)

משפט 4.6 (פואסון כגבול של בינומי במובן הנקודתי): יהי $Y \sim Poi(\lambda)$ עבור $\lambda \geq 0$ ויהיו $\{X_n\}_{n\in\mathbb{N}}$ משתנים מקריים כך שלכל $n > \lambda$ מתקיים $X_n \sim Bin\big(n, \frac{\lambda}{n}\big)$.
אזי לכל $k \in \mathbb{N}_0$ מתקיים

$$\lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(Y = k)$$

הוכחה: עבור k קבוע ו־n שואף לאינסוף מתקיים

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k)}{k!} = \frac{n^k(1+o(1))}{k!}$$

וכן

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)\left(1 - \frac{\lambda}{n}\right)^{-k} = e^{-\lambda} \cdot 1$$

ונובע אם כך

$$\lim_{n\to\infty} \mathbb{P}(X_n = k) = \lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \frac{n^k(1+o(1))}{k!} \cdot \frac{\lambda^k}{n^k} \cdot e^{-\lambda} = \frac{e^{-\lambda}\lambda^k}{k!} \lim_{n\to\infty} \frac{n^k(1+o(1))}{n^k} = \mathbb{P}(Y = k)$$

מפתח להוכחה:

1. חישוב גבול הכשלונות של X_n
2. הצבה בגבול המלא של ההסתברות
3. סידור טור יפה

$\square$

4.7 סכום של משתנים מקריים פואסונים בלתי־תלויים (4.127)

משפט 4.7 (סכום של משתנים מקריים פואסונים בלתי־תלויים)**:** יהיו $X \sim Poi(\lambda), Y \sim Poi(\eta)$ בלתי־תלויים, אזי $X+Y \sim Poi(\lambda+\eta)$.

הוכחה: עבור $n \in \mathbb{N}_0$, מנוסחת ההסתברות השלמה

$$\mathbb{P}(X+Y=n) \underset{\text{קונבולוציה}}{=} \sum_{i\in\mathbb{Z}} \mathbb{P}(X=i)\mathbb{P}(Y=n-i) \underset{(\star)}{=} \sum_{i=0}^{n} \frac{e^{-\lambda}}{i!}\frac{e^{-\eta}\eta^{n-i}}{(n-i)!} = \frac{e^{-\lambda-\eta}}{n!}\sum_{i=0}^{n}\binom{n}{i}\lambda^i\eta^{n-i} \underset{(\star\star)}{=} \frac{(\lambda+\eta)^n e^{-\lambda-\eta}}{n!}$$

כאשר $(\star)$ נובע מכך ששאר המחוברים מתאפסים ו־$(\star\star)$ זה נוסחת הבינום מה שמופיע בסכום.

מפתח להוכחה:

1. קונבולוציה
2. שינוי טור
3. בינום

$\square$

5 תוחלת ושונות

5.1 נוסחת התוחלת השלמה (5.26)

משפט 5.1 (נוסחת התוחלת השלמה): תהיי $\mathcal{A}$ חלוקה בת־מנייה של מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ ויהי X משתנה מקרי בעל תוחלת סופית על מרחב זה. אז

$$\mathbb{E}(X) = \sum_{A\in\mathcal{A}} \mathbb{E}(X\mathbb{1}_A)$$

הוכחה: נוכיח עבור X בדיד: $\mathcal{A}$ חלוקה ולכן $\sum_{A\in\mathcal{A}} \mathbb{1}_A = \mathbb{1}_\Omega = 1$ ולכן גם $\sum_{A\in\mathcal{A}} X\mathbb{1}_A = X$ ונחשב

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{A\in\mathcal{A}} X\mathbb{1}_A\right) \underset{\text{הגדרת התוחלת}}{=} \sum_{x\in\mathbb{R}} x\mathbb{P}\left(\sum_{A\in\mathcal{A}} X\mathbb{1}_A\right) \underset{\text{הסתברות שלמה}}{=} \sum_{x\in\mathbb{R}} x \sum_{A\in\mathcal{A}} \mathbb{P}(X\mathbb{1}_A = x)$$

$$\underset{\substack{\text{שינוי סדר סכימה}\\ \text{בטור מתכנס בהחלט}}}{=} \sum_{A\in\mathcal{A}} \sum_{x\in\mathbb{R}} x\mathbb{P}(X\mathbb{1}_A = x) \underset{\text{הגדרת התוחלת}}{=} \sum_{A\in\mathcal{A}} \mathbb{E}(X\mathbb{1}_A)$$

כאשר השיוויון של הסתברות שלמה נובע מכך שלכל $x \neq 0$

$$\bigcupdot_{A\in\mathcal{A}} \{X\mathbb{1}_A = x\} = \bigcupdot_{A\in\mathcal{A}} (\{X = x\} \cap A) = \{X = x\}$$

מפתח להוכחה:

1. בגלל שזוהי חלוקה, $X = \sum_{A\in\mathcal{A}} X\mathbb{1}_A$
2. לשחק עם השיוויונות לפי הגדרת התוחלת והסתברות שלמה

□

5.2 נוסחת הזנב לחישוב תוחלת משתנה מקרי על הטבעיים (5.19)

משפט 5.2 (נוסחת הזנב לחישוב תוחלת משתנה מקרי על הטבעיים): יהי X משתנה מקרי הנתמך על $\mathbb{N}\cup\{0\}$, אזי $\mathbb{E}(X)=\sum_{n=1}^{\infty}\mathbb{P}(X\geq n)$.

הוכחה: נשים לב שכל המחוברים בסכום הבא הם אי־שליליים ולכן ניתן להפעיל עליהם את משפט פוביני, אז מהגדרת התוחלת

$$\mathbb{E}(X)=\sum_{n=1}^{\infty}n\mathbb{P}(X=n)=\sum_{n=1}^{\infty}\sum_{k=1}^{n}\mathbb{P}(X=n)=\sum_{\substack{k,n\in\mathbb{N}\\ k\leq n}}\mathbb{P}(X=n)\underset{\text{פוביני}}{=}\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathbb{P}(X=n)=\sum_{n=1}^{\infty}\mathbb{P}(X\geq n)$$

מפתח להוכחה: להשתמש בהגדרת התוחלת ולהגיע לטור כפול כדי להשתמש במשפט פוביני. □

5.3 תוחלת של פונקציה על וקטור מקרי (5.3)

משפט 5.3 (תוחלת של פונקציה על וקטור מקרי)**:** יהי $X = (X_1, ..., X_d)$ וקטור מקרי דיד ותהיי $f \in \mathcal{F}_{\mathbb{R}^d \to \mathbb{R}}$ פונקציה. אז המשתנה המקרי $Y = f(X)$ מקיים

$$\mathbb{E}(Y) = \sum_{x \in \mathbb{R}^d} f(x)\mathbb{P}(X = x)$$

אם טור זה מתכנס בהחלט ואחרת ל-Y אין תוחלת סופית.

הוכחה: ראינו כי התפלגותו של X היא פונקציית הסתברות בדידה על $\mathbb{R}^d$.
נגדיר $Z(x) = f(x)$ משתנה מקרי חדש ומתקיים $Y \overset{d}{=} Z$ ונוכל להפעיל את תוחלת של משתנה מקרי על מרחב הסתברות בדידה על המרחב $\left(\mathbb{R}^d, \mathcal{F}_{\mathbb{R}}, \mathbb{P}_X\right)$ ולקבל

$$\mathbb{E}(Z) = \sum_{x \in \mathbb{R}^d} Z(x)\mathbb{P}_X(\{x\}) = \sum_{x \in \mathbb{R}^d} f(x)\mathbb{P}(X = x)$$

בגלל שהתוחלת נקבעת לפי ההתפלגות, אז מכך ש-$Z \overset{d}{=} Y$ נובע כי $\mathbb{E}(Y) = \mathbb{E}(Z)$.

מפתח להוכחה:

1. $Z(x) = f(x)$ ו-$Z \overset{d}{=} Y$
2. תוחלת של משתנה מקרי על מרחב הסתברות בדידה
3. שימוש בשיוויון התפלגויות

□

5.4 נוסחת סכום לשונות (6.35)

משפט 5.4 (נוסחת סכום לשונות): לכל אוסף $(X_k)_{k\in[n]}$ של משתנים מקריים מתקיים

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_k\right) = \sum_{\ell,k\le n} \mathrm{Cov}(X_k, X_\ell) = \sum_{k\le n} \mathrm{Var}(X_k) + 2\sum_{k<\ell\le n} \mathrm{Cov}(X_k, X_\ell)$$

בכל מקרה בו אגף ימין מוגדר היטב.

תזכורת:

$$\mathrm{Var}(X) := \mathbb{E}((X-\mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

$$\mathrm{Cov}(X,Y) := \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

הוכחה: נמרכז את המשתנים המקריים $\{X_k\}$ על־ידי $\overline{X_k} = X_k - \mathbb{E}(X_k)$ ולכן

$$\mathbb{E}\left(\overline{X_k}\right) = 0$$

$$\mathrm{Var}\left(\overline{X_k}\right) = \mathbb{E}\left(\overline{X_k}^2\right)$$

$$\mathrm{Cov}\left(\overline{X_k}, \overline{X_\ell}\right) = \mathbb{E}\left(\overline{X_k X_\ell}\right)$$

מתקיים

$$\mathrm{Var}\left(\sum_{k=1}^{n} \overline{X_k}\right) \underset{\text{אדישות להזזות}}{=} \mathrm{Var}\left(\sum_{k=1}^{n} \overline{X_k} - \sum_{k=1}^{n} \mathbb{E}(X_k)\right) = \mathrm{Var}\left(\sum_{k=1}^{n} X_k\right)$$

ולכן

$$\mathrm{Var}\left(\sum_{k=1}^{n} \overline{X_k}\right) = \mathbb{E}\left(\left(\sum_{k=1}^{n} \overline{X_k}\right)^2\right) = \mathbb{E}\left(\sum_{k=1}^{n}\sum_{\ell=1}^{n} \overline{X_k X_\ell}\right) \underset{\text{לינאריות}}{=} \sum_{k=1}^{n}\sum_{\ell=1}^{n} \mathbb{E}\left(\overline{X_k X_\ell}\right) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} \mathrm{Cov}(X_k, X_\ell)$$

$$= \sum_{k,\ell\le n} \mathrm{Cov}(X_k, X_\ell)$$

והשיוויון הימני נובע מהיות $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$ והכנסה של ערכים אלו בסכום.

מפתח להוכחה:

1. מרכוז על־ידי התוחלת
2. לרשום את **כל** מה שנובע מהמרכוז בהקשרי תוחלת ושונות
3. אדישות להזזות של השונות כדי להראות שהמשתנה המנורמל מספק אותנו
4. הגדרת השונות על המשתנה הממורכז עם הממצאים שלנו

□

6 אי־שיוויונות הסתברותיים

6.1 אי־שיוויון מרקוב (5.38)

משפט 6.1 (אי־שיוויון מרקוב): יהי X משתנה מקרי אי־שלילי (כלומר $X \overset{a.s.}{\geq} 0$) בעל תוחלת סופית. אזי לכל $a > 0$ מתקיים

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

הוכחה: נפעיל את נוסחת התוחלת השלמה על החלוקה $\{\{X<0\},\{X\in[0,a)\},\{a\leq X\}\}$ ונקבל

$$\mathbb{E}(X) = \mathbb{E}(X\mathbb{1}_{X<0}) + \mathbb{E}\Big(X\mathbb{1}_{X\in[0,a)}\Big) + \mathbb{E}(X\mathbb{1}_{X\geq a})$$

X הוא אי־שלילי ולכל $b \in \mathbb{R}$ מתקיים $X\mathbb{1}_{X\geq b} \overset{a.s.}{\geq} b\mathbb{1}_{X\geq b}$ והרי

$$X\mathbb{1}_{X<0} \overset{a.s.}{=} 0 \qquad X\mathbb{1}_{X\in[0,a)} \overset{a.s.}{\geq} 0 \qquad X\mathbb{1}_{X\geq a} \overset{a.s.}{\geq} a\mathbb{1}_{X\geq a}$$

וממונוטוניות התוחלת נקבל

$$\mathbb{E}(X) = \mathbb{E}(X\mathbb{1}_{X<0}) + \mathbb{E}\Big(X\mathbb{1}_{X\in[0,a)}\Big) + \mathbb{E}(X\mathbb{1}_{X\geq a}) \geq 0 + 0 + a\mathbb{E}(\mathbb{1}_{X\geq a}) = a\mathbb{P}(X \geq a)$$
$$\Longrightarrow \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

מפתח להוכחה:

1. מסתכלים על החלוקה $\{\{a\leq X\},\{X<0\},\{X\in[0,a)\}\}$
2. נוסחת התוחלת השלמה
3. חסימה איבר איבר
4. מונוטוניות התוחלת

$\square$

6.2 אי־שיוויון צ'בישב (6.9)

משפט 6.2 (אי־שיוויון צ'בישב): יהי X משתנה מקרי בעל שונות סופית. אז לכל $a > 0$ מתקיים

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}$$

הוכחה: נגדיר משתנה חדש $Y = (X - \mathbb{E}(X))^2$ וזה משתנה מקרי אי־שלילי המקיים $\mathbb{E}(Y) = \mathrm{Var}(X)$.
לכן לפי אי־שיוויון מרקוב לכל $b > 0$ מתקיים

$$\mathbb{P}(Y \geq b) \leq \frac{\mathbb{E}(Y)}{b} = \frac{\mathrm{Var}(X)}{b}$$

נשים לב $\{|X - \mathbb{E}(X)| \geq a\} = \{(X - \mathbb{E}(X))^2 \geq a^2\}$ ולכן בבחירת $b = a^2$ נקבל

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) = \mathbb{P}((X - \mathbb{E}(X))^2 \geq a^2) = \mathbb{P}(Y \geq a^2) \leq \frac{\mathrm{Var}(X)}{a^2}$$

מפתח להוכחה:

1. הגדרת משתנה מקרי חדש $Y = (X - \mathbb{E}(X))^2$
2. אי־שיוויון מרקוב
3. הכלת מאורעות $\{|X - \mathbb{E}(X)| \geq a\} = \{(X - \mathbb{E}(X))^2 \geq a^2\}$
4. שוב אי־שיוויון מרקוב

□

6.3 אי־שיוויון צ'רנוף (7.9)

משפט 6.3 (אי־שיוויון צ'רנוף): *יהי X משתנה מקרי בעל מומנט מעריכי. אזי לכל $t > 0$ עבורו $M_X(t)$ מוגדרת ולכל $a \in \mathbb{R}$ מתקיים*

$$\mathbb{P}(X \geq a) \leq M_X(t)e^{-ta}$$

תזכורת: יהי X משתנה מקרי. הפונקציה הממשית $M_X(t)$ הנתונה על־ידי

$$M_X(t) := \mathbb{E}(e^{tX})$$

לכל t עבורו התוחלת מוגדרת נקרא הפונקציה היוצרת מומנטים של X.

הוכחה: נשתמש באי־שיוויון מרקוב בשביל המשתנה המקרי החיובי e^{tX} ונקבל

$$\mathbb{P}(X \geq a) = \mathbb{P}(e^{tX} \geq e^{ta}) \underset{\text{אי־שיוויון מרקוב}}{\leq} \frac{\mathbb{E}(e^{tX})}{e^{ta}} = M_X(t)e^{-ta}$$

מפתח להוכחה: אי־שיוויון מרקוב (לציין שהמשתנה אי־שלילי ולכן הכיוון של אי־השיוויון נשמר). □

6.4 אי־שיוויון הופדינג (7.17)

משפט 6.4 (אי־שיוויון הופדינג): יהיו $\{X_k\}_{k\in[n]}$ משתנים מקריים בלתי־תלויים ובעלי תוחלת אפס אשר מקיימים $|X_k| \overset{a.s.}{\leq} 1$ לכל $k \in [n]$ אז

$$\forall d > 0, \ \left(\sum_{k\in[n]} X_k \geq d\right) \leq \exp\left(-\frac{d^2}{2n}\right)$$

משפט 6.5 (כפליות פונקציה יוצרת מומנטים עבור סכום משתנים מקריים בלתי־תלויים): יהיו X ו־Y משתנים מקריים בלתי־תלויים, אז

$$M_{X+Y}(t) = M_X(t)M_Y(t) \qquad \text{(לכל t עבורו שתיהן מוגדרות)}$$

הוכחה: נובע מכך שאי־תלות נשמרת תחת הפעלת פונקצייה ומכפליות התוחלת למשתנים בלתי־תלויים. □

משפט 6.6 (הלמה של הופדינג): יהי X משתנה מקרי המקיים $|X| \overset{a.s.}{\leq} 1$ וכן $\mathbb{E}(X) = 0$. אז לכל $t \in \mathbb{R}$ מתקיים

$$\mathbb{E}(e^{tX}) \leq \exp\left(\frac{t^2}{2}\right)$$

הוכחה: נקבע את t ונסמן ב־$L(x)$ את הפונקציה

$$L(x) = \frac{e^t + e^{-t}}{2} + x\frac{e^t - e^{-t}}{2}$$

הפונקציה e^{tx} היא בעלת נגזרת שנייה חיובית ולכן קמורה, אז לכל $x \in [-1, 1]$ מתקיים $e^{tx} \leq L(x)$, ממונוטוניות ולינאריות התוחלת נקבל

$$\mathbb{E}(e^{tX}) \leq \mathbb{E}(L(X)) = \frac{e^t + e^{-t}}{2} + \mathbb{E}(X)\frac{e^t - e^{-t}}{2} \underset{\mathbb{E}(X)=0}{=} \frac{e^t + e^{-t}}{2}$$

לכל $t \in \mathbb{R}$ מתקיים $\frac{e^t+e^{-t}}{2} \leq e^{\frac{t^2}{2}}$ וזה נובע מטור טיילור

$$\frac{e^t + e(-t)}{2} = \frac{1}{2}\sum_{n=0}^{\infty} \frac{t^n + (-t)^n}{n!} = \sum_{m=0}^{\infty} \frac{t^{2m}}{(2m)!} \leq \sum_{m=0}^{\infty} \frac{t^{2m}}{(2^m m)!} = \sum_{m=0}^{\infty} \frac{\left(\frac{t^2}{2}\right)^m}{m!} = e^{\frac{t^2}{2}}$$

□

הוכחה: אם־כך, נסמן $X = \sum_{k\in[n]} X_k$ ומתקיים מהטענות לעיל

$$M_X(t) = \prod_{k\in[n]} M_{X_k}(t) \leq \prod_{k\in[n]} \exp\left(\frac{t^2}{2}\right) = \exp\left(\frac{nt^2}{2}\right)$$

מאי־שיוויון צ'רנוף לכל $t > 0$

$$\mathbb{P}(X \geq d) \leq \exp\left(\frac{nt^2}{2} - td\right)$$

כדי למצוא t שימעזר את החסם נגזור את המעריך ונשווה לאפס

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{nt^2}{2} - td\right) = nt - d = 0 \Longrightarrow t = \frac{d}{n}$$

נקבל

$$\mathbb{P}(X \geq d) \leq \exp\left(\frac{n\left(\frac{d}{n}\right)^2}{2} - \left(\frac{d}{n}\right)d\right) = \exp\left(-\frac{d^2}{2n}\right)$$

מפתח להוכחה:

1. כפליות הפונקציה יוצרת מומנטים למשתנים מקריים בלתי־תלויים
2. הלמה של הופדינג
3. אי־שיוויון צ'רנוף + גזירה למזעור של המעריך

□

7 סדרות והתכנסויות

7.1 תנאי תוחלת ושונות להתכנסות לקבוע (6.19)

משפט 7.1 (תנאי תוחלת ושונות להתכנסות לקבוע): תהיי $(X_n)_{n=1}^{\infty}$ סדרת משתנים מקריים המקיימת עבור $\mu \in \mathbb{R}$ כי $\mu \to \mathbb{E}(X_n)$ וכן $\mathrm{Var}(X_n) \to 0$. אזי

$$X_n \xrightarrow{d} \mu$$

הוכחה: יהי $\varepsilon > 0$, נראה כי $\lim_{n\to\infty} \mathbb{P}(|X_n - \mu| < \varepsilon) = 1$ או באופן שקול $\lim_{n\to.,r\infty} \mathbb{P}(|X_n - \mu| \geq \varepsilon) = 0$:
מהיות $\mathbb{E}(X_n) \to \mu$, נבחר n_0 גדול מספיק כך שיתקיים לכל $n > n_0$ כי $|\mathbb{E}(X_n) - \mu| \leq \frac{\varepsilon}{2}$ וכן מאי־שוויון המשולש

$$|X_n - \mu| \leq |X_n - \mathbb{E}(X_n)| + |\mathbb{E}(X_n) - \mu| \leq |X_n - \mathbb{E}(X_n)| + \frac{\varepsilon}{2}$$

ולכן

$$\{|X_n - \mu| \geq \varepsilon\} \subset \left\{|X_n - \mathbb{E}(X_n)| + \frac{\varepsilon}{2} \geq \varepsilon\right\} = \left\{|X_n - \mathbb{E}(X_n)| \geq \frac{\varepsilon}{2}\right\}$$

ומאי־שוויון צ'בישב נקבל

$$\lim_{n\to\infty} \mathbb{P}(|X_n - \mu| \geq \varepsilon) \leq \lim_{n\to\infty} \mathbb{P}\Big(|X_n - \mathbb{E}(X_n)| \geq \frac{\varepsilon}{2}\Big) \underset{\text{אי־שוויון צ'בישב}}{\leq} \lim_{n\to\infty} \frac{\mathrm{Var}(X_n)}{\varepsilon^2 4} = 0$$

מפתח להוכחה:

1. משתמשים בהתכנסות התוחלת
2. אי־שיויון המשולש והכלת מאורעות
3. אי־שוויון צ'בישב

□

7.2 הלמה של פאטו (10.4)

משפט 7.2 (הלמה של פאטו): תהיי $\{A_n\}_{n\in\mathbb{N}}$ סדרת מאורעות. אז

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

$$\mathbb{P}(\{A_i, i.o.\}) = \mathbb{P}\left(\limsup_{n\to\infty} A_n\right) \ge \limsup_{n\to,r\infty} \mathbb{P}(A_n)$$

הוכחה: ראשית נראה כי הטענה השנייה נובעת מנכונות הטענה הראשונה:

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

$$\liminf_{n\to\infty}\mathbb{P}(A_n) = \lim_{n\to\infty}\inf_{i>n}\mathbb{P}(A_i) \ge \lim_{n\to\infty}\mathbb{P}\left(\bigcap_{i>n} A_i\right) \underset{\substack{\text{רציפות פונקציית ההסתברות} \\ \text{למאורעות עולים}}}{=} \mathbb{P}\left(\bigcup_{n\in\mathbb{N}}\bigcap_{i>n} A_i\right) = \mathbb{P}(\{A_i, a.e.\})$$

מפתח להוכחה: רציפות פונקציית ההסתברות למאורעות עולים. $\square$

7.3 הלמה הראשונה של בורל־קנטלי (10.5)

משפט 7.3 (הלמה הראשונה של בורל־קנטלי): תהיי A_i סדרת מאורעות. אם $\sum_{i=1}^{\infty}\mathbb{P}(A_i) < \infty$ אז $\mathbb{P}(A_i, i.o.) = 0$.

הוכחה:

$$\mathbb{P}(A_i, i.o.) \underset{\substack{\text{רציפות פונקציית ההסתברות} \\ \text{למאורעות עולים}}}{=} \lim_{n\to\infty}\mathbb{P}\left(\bigcup_{i=n}^{\infty} A_i\right) \underset{\text{אי־שיוויון בול}}{\leq} \lim_{n\to\infty}\sum_{i=n}^{\infty}\mathbb{P}(A_i) = 0$$

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

מפתח להוכחה: רציפות פונקציית ההסתברות וחסם האיחוד (והניסוח ממידה יותר יפה/ברור). $\square$

7.4 הלמה השנייה של בורל־קנטלי (10.6)

משפט 7.4 (הלמה השנייה של בורל־קנטלי): תהיי A_i סדרת מאורעות בלתי־תלויים. אם $\sum_{i=1}^{\infty} A_i = \infty$ אז $\mathbb{P}(A_i, i.o.) = 1$.

הוכחה:

$$\mathbb{P}(A_i, i.o.) = 1 - \mathbb{P}(A_i^c, a.e.) = 1 - \mathbb{P}\left(\bigcup_{m=1}^{\infty} \bigcap_{i=m}^{\infty} A_i^c\right) \underset{\substack{\text{רציפות פונקציית ההסתברות} \\ \text{למאורעות עולים}}}{=} 1 - \lim_{m\to\infty} \mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right)$$

אז מספיק שנראה שלכל $m \in \mathbb{N}$ מתקיים $\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right) = 0$ ואכן מהאי־תלות

$$\mathbb{P}\left(\bigcap_{i=m}^{\infty} A_i^c\right) \underset{\substack{\text{רציפות פונקציית ההסתברות} \\ \text{למאורעות עולים}}}{=} \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{i=m}^{n} A_i^c\right) = \lim_{n\to\infty} \prod_{i=m}^{n} (1 - \mathbb{P}(A_i)) \leq \lim_{n\to\infty} \exp\left(-\sum_{i=m}^{n} \mathbb{P}(A_i)\right) = 0$$

כאשר האי־שיוויון נובע מכך ש־$1 + x \leq e^x$ לכל x והשיוויון נובע מכך ש־$\sum_{i=m}^{\infty} \mathbb{P}(A_i) = \infty \Longrightarrow \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ לכל m.

מפתח להוכחה:

1. משלים
2. רציפות פונקציית ההסתברות
3. לכל x מתקיים $1 + x \leq e^x$

$\square$

7.5 החוק החלש של המספרים הגדולים (6.21)

משפט 7.5 (החוק החלש של המספרים הגדולים): תהיי $X_1, X_2, \ldots$ סדרת משתנים מקריים בלתי־תלויים, שווי התפלגות ובעלי תוחלת μ.
אם $Y_n = \frac{\sum_{i=1}^{n} X_i}{n}$ אזי לכל $\varepsilon > 0$

$$\mathbb{P}(|Y_n - \mu| \geq \varepsilon) \xrightarrow[n\to\infty]{} 0$$

הוכחה: **הוכחה תחת הנחת קיום שונות:**

$$\mathbb{E}(Y_n) = \frac{\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)}{n} \underset{\text{לינאריות התוחלת}}{=} \frac{\sum_{i=1}^{n} \mathbb{E}(X_i)}{n} = \frac{n \cdot \mu}{n} = \mu$$

ולכן

$$\mathbb{P}(|Y_n - \mu| \geq \varepsilon) \underset{\text{צ'בישב}}{\leq} \frac{\operatorname{Var}(Y_n)}{\varepsilon^2} = \frac{\operatorname{Var}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)}{\varepsilon^2} \underset{\text{כיול ריבועי}}{=} \frac{\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right)}{n^2\varepsilon^2} \underset{\text{סכום שונות בלתי־תלויות}}{=} \frac{n \operatorname{Var}(X_1)}{n^2\varepsilon^2} \xrightarrow[n\to\infty]{} 0$$

מפתח להוכחה:

1. חישוב תוחלת של Y_n
2. אי־שיוויון צ'בישב

$\square$

הערה: במילים אחרות, החוק החלש של המספרים הגדולים אומר $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{p} \mu$.

7.6 החוק החזק של המספרים הגדולים (10.20)

משפט 7.6 (החוק החזק של המספרים הגדולים): תהיי $X_1, X_2, ...$ סדרת משתנים מקריים בלתי־תלויים, שווי התפלגות עם $\mathbb{E}(X_n) = \mu$ ו־$|X_i| \overset{a.s.}{\leq} M$ אזי

$$\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow[n\to\infty]{a.s.} \mu$$

הוכחה: נגדיר משתנים מקריים חדשים

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

תנאי אי־שיווין הופדינג מתקיימים ולכן 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 $\varepsilon > 0$ אם נציב $a = \varepsilon n$ נקבל

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

נסמן

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

ולכל $k \in \mathbb{N}$ מתקיים $\sum_{n=1}^{\infty} \mathbb{P}\left(A_n^k\right) < \infty$ ולכן לפי הלמה הראשונה של בורל־קנטלי נקבל

$$\mathbb{P}\left(A_n^k n\ i.o.\right) = 1 - \mathbb{P}\left(\left(A_n^k\right)^c n\ a.e.\right) = 0$$

אז

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_n^k n\ i.o.\right) = 0$$

ובאופן שקול

$$\mathbb{P}\left(\bigcap_{k=1}^{\infty} \left(A_n^k\right)^c n\ a.e.\right) = 1$$

וזו ההגדרה של $\frac{1}{n}\sum_{i=1}^{n} Y_i \overset{a.s.}{\to} 0$ אבל $\frac{1}{n}\sum_{i=1}^{n} Y_i \overset{a.s.}{\to}$ אם ורק אם $\frac{1}{n}\sum_{i=1}^{n} X_i \overset{a.s.}{\to} \mu$.

מפתח להוכחה:

1. מגדירים משתנה מקרי ממורכז
2. משתמשים באי־שיווין הופדינג
3. $a = \varepsilon n$ עבור $\varepsilon > 0$
4. $A_n^k := \left\{\left|\frac{1}{n}\sum_{i=1}^{n} Y_i\right| \geq \frac{1}{k}\right\}$
5. הלמה הראשונה של בורל־קנטלי

□

8 מיפוי התכנסויות

8.1 הגדרות

הגדרה 8.1 (התכנסות כמעט־תמיד): תהיי $(X_n)_{n=1}^{\infty}$ סדרת משתנים מקריים במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$.
נאמר כי סדרה זו מתכנסת למשתנה המקרי X **כמעט־תמיד** ונסמן $X_n \overset{a.s.}{\to} X$ אם מתקיים

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

באופן שקול

$$X_n \overset{a.s.}{\to} X \iff \forall\varepsilon > 0,\ \mathbb{P}\Big(\lim_{n\to\infty}|X_n - X| > \varepsilon\Big) = 0 \iff \forall\varepsilon > 0,\ \mathbb{P}(|X_n - X| \le \varepsilon\ a.e.) = 1$$

הגדרה 8.2 (התכנסות בהסתברות): תהיי $(X_n)_{n=1}^{\infty}$ סדרת משתנים מקריים במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$.
נאמר כי סדרה זו מתכנסת למשתנה המקרי X **בהסתברות** ונסמן $X_n \overset{p}{\to} X$ אם לכל $\varepsilon > 0$ מתקיים

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

באופן שקול

$$X_n \overset{p}{\to} X \iff \forall\varepsilon > 0,\ \mathbb{P}\left(\limsup_{n\to\infty}|X_n - X| > \varepsilon\right) = 0$$

הגדרה 8.3 (התכנסות בהתפלגות): תהיי $(X_n)_{n=1}^{\infty}$ סדרת משתנים מקריים לא בהכרח על אותו מרחב הסתברות ויהי X משתנה מקרי.
נאמר כי סדרה זו מתכנסת למשתנה המקרי X **בהתפלגות** ונסמן $X_n \overset{d}{\to} X$ אם לכל a שהיא נקודת רציפות של F_X מתקיים

$$\lim_{n\to\infty} F_{X_n}(a) = F_X(a)$$

הגדרה 8.4 (התכנסות בהתפלגות לקבוע): תהיי $(X_n)_{n=1}^{\infty}$ סדרת משתנים מקריים במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$.
נאמר כי סדרה זו **מתכנסת לקבוע** ונסמן $X_n \overset{d}{\to} a$ אם לכל $\varepsilon > 0$ מתקיים

$$\lim_{n\to\infty} \mathbb{P}(|X_n - a| < \varepsilon) = 1$$

מסקנה 8.1: אם נסמן

$$A_{n,\varepsilon} := \{\omega \mid |X_n(\omega) - X(\omega)| < \varepsilon\}$$

מההגדרות

$$X_n \overset{a.s.}{\to} X \iff \forall\varepsilon > 0,\ \mathbb{P}\Big(\liminf_{n\to\infty} A_{n,\varepsilon}\Big) = 1$$

$$X_n \overset{p}{\to} X \iff \forall\varepsilon > 0,\ \liminf_{n\to\infty} \mathbb{P}\big(A_{n,\varepsilon}\big) = 1$$

מסקנה 8.2: אם לכל $\varepsilon > 0$ מתקיים $\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty$ אזי $X_n \overset{a.s.}{\to} X$.

8.2 גרירות

משפט 8.1 (גרירות):

1. התכנסות כמעט־תמיד גורר התכנסות בהסתברות
2. התכנסות בהסתברות גוררת התכנסות בהתפלגות
3. התכנסות בהתפלגות לקבוע גוררת התכנסות בהסתברות (ואז נוח לעבוד עם משפט 6.19)

8.3 כלים שימושים

1. הלמה השנייה של בורל־קנטלי טובה להפרכת התכנסות כמעט־תמיד
2. הלמה הראשונה של בורל־קנטלי טובה להוכחת התכנסות כמעט־תמיד

9 סיכום תוצאות

9.1 התפלגויות בדידות

$X \sim$	Parameters	$\text{supp}(X)$	$\mathbb{P}(X=k)$	$\mathbb{E}(X)$	$\text{Var}(X)$	$M_X(t)$
$Unif([n])$	$n \in \mathbb{N}$	$[n]$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{e^{nt}-e^{2t}}{n(1-e^t)}$
$Ber(p)$	$0 \leq p \leq 1$	$\{0,1\}$	$\begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	p	$p(1-p)$	$pe^t+(1-p)$
$Bin(n,p)$	$n \in \mathbb{N},\ 0 \leq p \leq 1$	$\mathbb{N} \cup \{0\}$	$\binom{n}{k}(1-p)^{n-k}p^k$	np	$np(1-p)$	$\left(pe^t+(1-p)\right)^n$
$Geo(p)$	$0 \leq p \leq 1$	$\mathbb{N}$	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
$Poi(\lambda)$	$0 < \lambda$	$\mathbb{N} \cup \{0\}$	$e^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ	$\exp(\lambda(e^t-1))$

9.2 התפלגויות רציפות

$X \sim$	Parameters	$\text{supp}(X)$	$f_X(t)$	$F_X(t)$	$\mathbb{E}(X)$	$\text{Var}(X)$	$M_X(t)$
$Unif([a,b])$	$a \leq b$	$t \in [a,b]$	$\begin{cases} \frac{1}{b-a} & t \in [a,b] \\ 0 & \text{else} \end{cases}$	$\begin{cases} 0 & t < a \\ \frac{t-a}{t-b} & a \leq t < b \\ 1 & t > b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & t \neq 0 \\ 1 & t=0 \end{cases}$
$Exp(\lambda)$	$\lambda > 0$	$0 \leq t$	$-\lambda e^{\lambda t}$	$1-e^{-\lambda}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$
$\mathcal{N}(0,1)$	—	$\mathbb{R}$	$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)$	$\Phi(t)$	0	1	—
$\mathcal{N}(\mu,\sigma)$	$\sigma^2 \geq 0$	$\mathbb{R}$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$	$\Phi\left(\frac{s-\mu}{\sigma}\right)$	μ	σ^2	—

10 הוכחות ממבחני עבר של אוהד

מבחן	משפט
הסתברות למתמטיקאים 2019 סמסטר א' מועד א'	1. שאלה 1 1. אי־שיוויון מרקוב 2. אי־שיוויון צ'רנוף 2. שאלה 2 1. הלמה הראשונה של בורל־קנטלי 2. הלמה השנייה של בורל־קנטלי
הסתברות למתמטיקאים 2019 סמסטר א' מועד ב'	1. שאלה 1 1. אי־שיוויון בול 2. הכלה והדחה לשלושה מאורעות 2. שאלה 2 1. להגדיר מרחב מדגם, פונקציית הסתברות בדידה, משתנה מקרי ותוחלת 2. חסימות השונות
הסתברות למתמטיקאים 2018 סמסטר א' מועד א'	1. אי־שיוויון בול 2. משהו מוזר
הסתברות למתמטיקאים 2018 סמסטר א' מועד ב'	1. להגדיר שונות משותפת ולהוכיח סכום שנויות
הסתברות למדמ"ח 2025 סמסטר א' מועד א'	1. הגדרת שיוויון התפלגויות 2. הגדרת שיוויון כמעט־תמיד 3. שיוויון כמעט־תמיד גורר שיוויון התפלגויות 4. שיוויון בהתפלגויות נשמר תחת הפעלת פונקציה
הסתברות למדמ"ח 2025 סמסטר א' מועד ב'	1. תכונות של נוסחת התוחלת השלמה עם הסתברות מותנית
הסתברות למדמ"ח 2025 סמסטר א' מועד ג'	1. נוסחת התוחלת השלמה עם הסתברות מותנית 2. נוסחת השונות לסכום
הסתברות למדמ"ח 2024 סמסטר א' מועד א'	1. סכום משתני ברנולי בלתי־תלויים מתפלג בינומית 2. תנאי תוחלת ושונות להתכנסות לקבוע 3. הגדרת התכנסות לקבוע 4. הוכחת החוק החלש של המספרים הגדולים
הסתברות למדמ"ח 2024 סמסטר א' מועד ב'	1. ניסוח והוכחה של אי־שיוויון מרקוב 2. ניסוח והוכחה של אי־שיוויון הופדינג ללא הלמה
הסתברות למדמ"ח 2023 סמסטר א' מועד א'	1. תוחלת של משתנה מקרי שנתמך על הטבעיים (עם פוביני)
הסתברות למדמ"ח 2023 סמסטר א' מועד ב'	1. אי־שיוויון מרקוב (בניסוח מוזר) 2. אי־שיוויון צ'בישב 3. אי־שיוויון צ'נרוף
הסתברות למדמ"ח 2022 סמסטר א' מועד א'	1. לינאריות התוחלת 2. חסימות השונות?
הסתברות למדמ"ח 2022 סמסטר א' מועד ב'	1. אי־שיוויון צ'בישב 2. אי־שיוויון צ'נרוף
הסתברות למדמ"ח 2022 סמסטר א' מועד ג'	1. אי־שיוויון בול עבור מספר סופי של מאורעות
הסתברות למדמ"ח 2018 סמסטר א' מועד א'	1. להגדיר שונות משותפת 2. להוכיח נוסחת סכום שנויות לשני משתנים מקריים