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

26 בינואר 2026

תוכן עניינים

1 רציפות פונקציית ההסתברות

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

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

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

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

אם־כך מסכימות נקבל

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

$\square$

2 אי־שיוויון בול

משפט 2.1 (אי־שיוויון בול):

משפט 2.2 (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

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

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

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

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

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

עבור מאורעות יורדים, נשתמש בהיות המשלים שלהם מאורעות עולים. □

משפט 2.3 (אי־שיוויון בול לסדרת מאורעות): לכל סדרת מאורעות $\{A_n\}_{n\in\mathbb{N}}$ במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ מתקיים

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

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

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

□

3 נוסחת ההסתברות השלמה במונחי הסתברות מותנית

משפט 3.1 (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

$$\mathbb{P}(B) = \sum_{A \in \mathcal{A}} \mathbb{P}(A \cap B) = \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(A \cap B) + \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) = 0}} \mathbb{P}(A \cap B) \underset{\text{כלל השרשרת}}{=} \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(B \mid A)\mathbb{P}(A) + \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) = 0}} 0 = \sum_{\substack{A \in \mathcal{A} \\ \mathbb{P}(A) > 0}} \mathbb{P}(B \mid A)\mathbb{P}(A)$$

$\square$

4 כלל בייס

משפט 4.1 (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:

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

□

5 אי־תלות נשמרת תחת הפעלת פונקציות

משפט 5.1 (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

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

$\square$

6 שיוויון כמעט־תמיד גורר שיוויון התפלגויות

משפט 6.1 (שוויון כמעט־תמיד גורר שיוויון התפלגויות): יהיו X,Y משתנים מקריים על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. אם $Y \overset{a.s.}{=} X$ אז $Y \overset{d}{=} X$.

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

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

□

7 שיוויון התפלגויות נשמר תחת הפעלת פונקציה

משפט 7.1 (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

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

$\square$

8 שיוויון כמעט־תמיד נשמר תחת הפעלת פונקציה

משפט 8.1 (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 $\mathbb{P}(N_f) \leq \mathbb{P}(N) = 0$. □

9 הצלחה ראשונה בסדרת ניסויי ברנולי בלתי־תלויים מתפלג גיאומטרי

משפט 9.1 (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

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

□

10 חוסר זיכרון של התפלגות גיאומטרית

משפט 10.1 (חוסר זיכרון של התפלגות גיאומטרית):

הגדרה 10.1 (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 $n\in\mathbb{N}$

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

$\Longleftarrow$ 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 $1-p=\mathbb{P}(X>1)$, אז לכל $k\geq 1$

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

$\square$

11 סכום משתנים ברנולי בלתי־תלויים מתפלג בינומית

משפט 11.1 (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 בידיוק 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

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

$\square$

12 חיבור משתנים מקריים בינומיים בלתי־תלויים

משפט 12.1 (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

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

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

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

□

13 פואסון כגבול של בינומי במובן הנקודתי

משפט 13.1 (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

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

□

14 נוסחת התוחלת השלמה

משפט 14.1 (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 $x \neq 0$

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

$\square$

15 נוסחת סכום לשונות

משפט 15.1 (נוסחת סכום לשונות): לכל אוסף $(X_k)_{k\in[n]}$ של משתנים מקריים מתקיים

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

בכל מקרה בו אגף ימין מוגדר היטב.

תזכורת:

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

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

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

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

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

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

מתקיים

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

ונקבל אם־כך

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

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

16 אי־שיוויון מרקוב

משפט 16.1 (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 $b\mathbb{1}_{X\geq b} \overset{a.s.}{\geq} X\mathbb{1}_{X\geq b}$ 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

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

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

□

17 אי־שיוויון צ'בישב

משפט 17.1 (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 $\{(X - \mathbb{E}(X))^2 \geq a^2\} = \{|X - \mathbb{E}(X)| \geq a\}$ ולכן בבחירת $b = a^2$ נקבל

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

□

18 אי־שיוויון צ'רנוף

משפט 18.1 (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

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

□

19 אי־שיוויון הופדינג

משפט 19.1 (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 19.2 (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: היות ואי־תלות נשמרת תחת הפעלת פונקצייה נובע כי e^{tX}, e^{tY} משתנים מקרים בלתי־תלויים. אז מכפליות התוחלת למשתנים בלתי־תלויים

$$\mathbb{E}\big(e^{t(X+Y)}\big) = \mathbb{E}\big(e^{tX}e^{tY}\big) = \mathbb{E}\big(e^{tX}\big)\mathbb{E}\big(e^{tY}\big)$$

□

משפט 19.3 (הלמה של הופדינג): יהי X משתנה מקרי המקיים $|X| \overset{a.s.}{\leq} 1$ וכן $\mathbb{E}(X) = 1$. אז לכל $t \in \mathbb{R}$ מתקיים

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

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

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

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

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

□

20 הלמה של פאטו

משפט 20.1 (הלמה של פאטו): תהיי $\{A_n\}_{n\in\mathbb{N}}$ סדרת מאורעות. אז

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

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

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

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

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

□

21 הלמה הראשונה של בורל־קנטלי

משפט 21.1 (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 $\lim_{n\to\infty} \sum_{i=n}^{\infty} \mathbb{P}(A_i) = 0$. $\square$

22 הלמה השנייה של בורל־קנטלי

משפט 22.1 (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־$\sum_{i=m}^{\infty} \mathbb{P}(A_i) = \infty \Longleftarrow \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ לכל m. $\square$

23 החוק החלש של המספרים הגדולים

משפט 23.1 (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

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

$\square$

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

24 החוק החזק של המספרים הגדולים

משפט 24.1 (החוק החזק של המספרים הגדולים): תהיי $X_1, X_2, \ldots$ 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־שיוויון 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 $\frac{1}{n}\sum_{i=1}^{n} X_i \overset{a.s.}{\to} \mu$. $\square$