מבוא להסתברות וסטטיסטיקה

נכתב ע"י יונתן קלסי, נערך ע"י אוהד היינה

הרצאה 1

הגדרות בסיסיות

מוסכמה: (לא פורמלית) נאמר שניסוי הוא תהליך בדיקה שבסופו יש תוצאה.

דוגמאות: הטלת קוביה, מילוי טופס לוטו, מדידת גובה ומשקל.

סימון: עבור מספר טבעי $n \in \mathbb{N}$ נסמן $[n] = \{1, \ldots, n\}$.

הגדרה: מרחב מדגם הוא קבוצה לא ריקה Ω שמכילה תוצאות אפשריות.

דוגמאות:

א. הטלת קוביה הוגנת: $\Omega = [6] = \{1, \ldots, 6\}$.

ב. הטלת קובית 20 הוגנת: $\Omega = [20] = \{1, \ldots, 20\}$.

ג. הטלת מטבע: $\Omega = \{H, T\}$.

ד. הטלת 3 מטבעות: $\Omega = \{H, T\}^3$.

ה. זמן עד שנורה תשרף: $\Omega = \mathbb{R}_+ = [0, \infty)$.

ו. הטלת מטבע עד שיוצא פלי: $\Omega = \mathbb{N} \cup \{\infty\}$.

הגדרה: פונקצית הסתברות נקודתית היא פונקציה $p : \Omega \longrightarrow [0, \infty)$ אשר מקיימת $\sum\limits_{\omega \in \Omega} p(\omega) = 1$.

הגדרה: התומך של p הוא הקבוצה $\mathrm{Sprt}(p) = \{\omega \in \Omega \mid p(\omega) > 0\}$.

הערה: נשים לב כי בהגדרה מובלעת הדרישה כי $\mathrm{Sprt}(p)$ היא בת מנייה.

דוגמאות:

א. הטלת קוביה הוגנת: $p(\omega) = \frac{1}{6}$.

ב. הטלת קוביה לא הוגנת: $p(i) = p_i$ כלשהם כך ש $\sum p_i = 1$.

ג. הטלת מטבע: $p(H) = p(T) = \frac{1}{2}$.

ד. בחירת מספר טבעי אקראי: $\Omega = \mathbb{N}$, ונוכל להגדיר מספר פונקציות הסתברות נקודתיות:

(1) $p(i) = \frac{1}{i(i+1)}$.

(2) $p(i) = \begin{cases} 1 & i = 7 \\ 0 & i \neq 7 \end{cases}$.

(3) $p(i) = 2^{-i}$.

מאורעות, פונקציית הסתברות ומרחב הסתברות

הגדרה: מאורע $A \subseteq \Omega$ הוא אוסף תוצאות, תת קבוצה של Ω. אוסף כל המאורעות יסומן ב $\mathcal{F}$, כלומר $\mathcal{F} = 2^{\Omega}$.

הגדרה: עבור מאורע A, נגדיר $A^c := \Omega \backslash A$ להיות המאורע המשלים של A.

דוגמאות:

א. עבור הטלת קוביה הוגנת $\Omega = [6]$ והמאורע "התוצא יצאה זוגית" נקבל $A = \{2, 4, 6\} \subseteq \Omega$.

ב. עבור 10 הטלות של קוביה: נבחר $A =$ "התוצאה המקסילית היא של הטלה 4". מתקיים $\Omega = [6]^{10}$, $A = [4]^{10}$.

ג. גם $\emptyset, \Omega$ מאורעות.

הגדרה: פונקצית הסתברות היא פונקציה $\mathbb{P} : \mathcal{F} \longrightarrow [0, \infty)$ המקיימת:

1. $\mathbb{P}(\Omega) = 1$.
2. סכימות בת מניה: לכל סדרת מאורעות זרים $\{A_n\}_{n=1}^{\infty}$ מתקיים $\mathbb{P}\left(\dot{\bigcup}_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

אם $\mathbb{P}(A) = 1$ אז נאמר ש $\mathbb{P}$ נתמכת על A. השלשה $(\Omega, \mathcal{F}, \mathbb{P})$ תקרא מרחב הסתברות.

טענה (תכונות פונקצית ההסתברות):

1. $\mathbb{P}(\emptyset) = 0$.
2. סכימות: לכל אוסף סופי של מאורעות זרים $\{A_n\}_{n=1}^{N}$ מתקיים $\mathbb{P}\left(\dot{\bigcup}_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mathbb{P}(A_n)$.
3. מונוטוניות: $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$.
4. חסימות: $A \in \mathcal{F} \implies \mathbb{P}(A) \in [0, 1]$.
5. משלים: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

הוכחה:

1. נבחר $A_n = \emptyset$ לכל $n \in \mathbb{N}$. מסכימות בת מניה נקבל כי $\mathbb{P}(\emptyset) = 0$.
2. בהינתן אוסף סופי של מאורעות זרים $\{A_n\}_{n=1}^{N}$ נגדיר $\{B_n\}_{n=1}^{\infty}$ על פי $B_n = A_n$ לכל $n \in [N]$ וְ $B_n = \emptyset$ אחרת. כעת השיוויון המבוקש נובע מסכימות בת מניה ומתכונה (1).
3. בהינתן מאורעות $A \subseteq B$ נשתמש בתכונה (2) על המאורעות הזרים A ו $B \backslash A$ ונקבל $\mathbb{P}(A) + \mathbb{P}(B \backslash A) = \mathbb{P}(B)$. על פי הגדרה, $\mathbb{P}$ אי שלילית ונוכל להסיק כי $\mathbb{P}(A) \leq \mathbb{P}(B)$ כנדרש.
4. בהינתן מאורע $A \in \mathcal{F}$ נשים לב כי $\emptyset \subseteq A \subseteq \Omega$. מהמונוטוניות של $\mathbb{P}$ נקבל $0 \leq \mathbb{P}(A) \leq 1$ כנדרש.
5. נפעיל את תכונת הסכימות על A, A^c ונקבל: $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1 \iff \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

טענה: פונקצית הסתברות נקודתית מגדירה פונקציית הסתברות: תהי $p:\Omega\longrightarrow\mathbb{R}_+$ פונקציית הסתברות נקודתית על מרחב מדגם Ω. אזי הפונקציה $\mathbb{P}_p:\mathcal{F}\longrightarrow\mathbb{R}_+$ הנתונה על ידי $\mathbb{P}_p\left(A\right)=\sum\limits_{\omega\in A}p\left(\omega\right)$ לכל מאורע A, היא פונקציית הסתברות הנתמכת על $\mathrm{Sprt}\left(p\right)$.

הוכחה: ראשית נוכיח כי $\mathbb{P}_p\left(\Omega\right)=1$. על פי הגדרת פונקציית ההסתברות הנקודתית מתקיים $\mathbb{P}_p\left(\Omega\right)=\sum\limits_{\omega\in\Omega}p\left(\omega\right)=1$.

כעת נוכיח סכימות בת מניה. יהיו $\left\{A_n\right\}_{n=1}^{\infty}$ מאורעות זרים. נחשב:

$$\mathbb{P}_p\left(\dot{\bigcup_{n\in\mathbb{N}}}A_n\right)=\sum_{\omega\in\dot{\bigcup\limits_{n\in\mathbb{N}}}A_n}p\left(\omega\right)=\sum_{n\in\mathbb{N}}\sum_{\omega\in A_n}p\left(\omega\right)=\sum_{n\in\mathbb{N}}\mathbb{P}_p\left(A_n\right)$$

נותר להוכיח כי $\mathbb{P}_p$ נתמכת על $\mathrm{Sprt}\left(p\right)$. נשים לב כי:

$$\mathbb{P}_p\left(\mathrm{Sprt}\left(p\right)\right)=\sum_{\omega\in\mathrm{Sprt}\left(p\right)}p\left(w\right)=\sum_{\substack{\omega\in\Omega\\p\left(\omega\right)>0}}p\left(\omega\right)=1$$

הגדרה: יהי $\left(\Omega,\mathcal{F},\mathbb{P}\right)$ מרחב הסתברות. אם קיימת פונקציית הסתברות נקודתית p כך ש $\mathbb{P}_p=\mathbb{P}$ אז $\mathbb{P}$ נקראת פונקציית הסתברות בדידה. השלשה $\left(\Omega,\mathcal{F},\mathbb{P}\right)$ מכונה מרחב הסתברות בדידה.

הרצאה 2

עקרונות החיבור והכפל

הגדרה: נאמר כי $|A|=N$ אם קיימת $f:A\longrightarrow[N]$ חח"ע ועל.

עקרון החיבור: אם $A\cap B=\emptyset$ אז $|A\cup B|=|A|+|B|$.

עקרון הכפל: אם A,B קבוצות, נסמן $A\times B=\{(a,b)\mid a\in A\wedge b\in B\}$. אזי $|A\times B|=|A|\cdot|B|$.

הערה: ניתן להוכיח עקרונות אלו בקלות ע"י בניית פונ' חח"ע ועל.

מסקנות:

1. אם $A\subseteq B$ אז $|B\backslash A|=|B|-|A|$.
2. עקרון הכפל נתן להכללה ל k קבוצות.

בחירה עם סדר ועם חזרות

שאלה: כמה פונקציות $f:[N]\longrightarrow[K]$ ישנן?

תשובה: נבחר לאן לשלוח כל איבר מ $[N]$. לכל איבר יש K אפשרויות ויש N איברים ובסה"כ K^N פונקציות.

שאלה: בכמה דרכים ניתן להכניס N כדורים ל K מגירות?

תשובה: זו בדיוק אותה שאלה כמו הקודמת, לכן גם אותה התשובה.

שאלה: בכמה דרכים ניתן להכניס N כדורים ל K מגירות כך שלפחות כדור אחד יהיה במגירה האחרונה?

תשובה: נחשב כמה דרכים יש כך שאין אף כדור במגירה האחרונה. זה שקול לכמה פונ' $f:[N]\longrightarrow[K-1]$ ישנן. ע"פ התרגיל הקודם התשובה היא $(K-1)^N$. כמו כן ראינו כי כמות הדרכים בהן ניתן להכניס N כדורים ל K מגירות היא K^N וסה"כ מעקרון החיסור: $K^N-(K-1)^N$.

שאלה: כמה תת קבוצות יש לקבוצה נתונה A מגודל N?

תשובה: 2^N. שקול לגודל הקבוצה $\{f:[N]\longrightarrow\{0,1\}\}$ - קבוצת הפונ' <u>המציינות</u> לכל איבר אם הוא בתת קבוצה, או שלא.

בחירה עם סדר ובלי חזרות

שאלה: בכמה דרכים אפשר לסדר N ילדים בשורה?

תשובה: בוחרים מיהו הילד הראשון - N אפשרויות, בוחרים מיהו הילד השני - $N-1$ אפשרויות, ..., בוחרים מיהו הילד ה N - אפשרות אחת. בסה"כ מעקרון הכפל: $N\left(N-1\right)\cdot\dots\cdot1=N!$ אפשרויות.

שאלה: בכמה דרכים ניתן לכתוב רשימה סדורה של N ילדים מתוך כיתה של K?

תשובה: בוחרים את הראשון - K אפשרויות, בוחרים את השני - $K-1$ אפשרויות, ..., בוחרים את ה N - $K-N+1$ אפשרויות. בסה"כ מעקרון הכפל: $K\left(K-1\right)\cdot\dots\cdot\left(K-N+1\right)=\frac{K!}{(K-N)!}$ אפשרויות.

שאלה: כמה פונ' חח"ע $f:[N]\longrightarrow[K]$ יש?

תשובה: זאת אותה שאלה, לכן: $K\left(K-1\right)\cdot\dots\cdot\left(K-N+1\right)=\frac{K!}{(K-N)!}$.

שאלה: K כדורים בכד, ממוספרים. מבצעים N שליפות בלי החזרה, עם חשיבות לסדר. בכמה דרכים ניתן לעשות זאת, אם:

א. מובטח שנשלף הכדור ה K?

ב. מובטח שנשלפו הכדור ה K והכדור ה $K-1$?

ג. מובטח שלא נשלפו הכדור ה K והכדור ה $K-1$?

תשובה:

א. לשליפת הכדור ה K היו N אפשרויות זרות (ראשונה, שנייה, ..., N), ועל כל אפשרות כזו כל שאר השליפות הן "רגילות" עבור מקום אחד פחות. לכן: $N(K-1)(K-2)\cdot\dots\cdot(K-N+1)$.

ב. הפעם יש שני כדורים מיוחדים ויש חשיבות לסדר ביניהם: $N(N-1)(K-2)\cdot\dots\cdot(K-N+1)$.

ג. עקרון החיסור: $K(K-1)\cdot\dots\cdot(K-N+1)-N(N-1)(K-2)\cdot\dots\cdot(K-N+1)$.

עקרון ספירת המחלקות

עקרון ספירת המחלקות (ספירה כפולה): אם נחלק את A ל M מחלקות, כל אחת מגודל S, אז: $|A|=M\cdot S$ ומספר המחלקות $M=\frac{|A|}{S}$.

שאלה: כמה סידורים יש ל N אבירים סביב שולחן עגול?

תשובה: למעשה, כל סידור סביב השולחן מתאים למחלקה בת n סידורים בשורה, וכל אחד מהם מתאים לסדר האבירים סביב השולחן כאשר מתחילים למנות אותו מילד מסוים. לכן נוכל לחשוב על השאלה כספירת מספר המחלקות הנ"ל. מחלקות אלו זרות ושוות בגודלן, לכן נוכל להעזר בעקרון, ולקבל כי מספר המחלקות (כמות הסידורים סביב השולחן) הוא $(n-1)!=\frac{n!}{n}$.

שאלה: נתונים $2n$ כדורים, שניים מכל צבע. נסמן הצבעים $\{1,\dots,n\}$. בכמה דרכים ניתן לסדר את הכדורים בשורה? (לא מבחינים בין כדורים שווי צבע).

תשובה: יש $(2n)!$ דרכים לסדר את הכדורים בשורה עם הבחנה בין כדורים שווי צבע. כל סידור תקין מתאים למחלקה בת $2\cdot 2\cdot\dots\cdot 2=2^n$ סידורים, וכל אחד מהם הוא החלפה במיקומים של חלק (או כל) מהכדורים שווי הצבע. מחלקות אלו זרות ושוות בגודלן, לכן נוכל להעזר בעקרון ולקבל כי מספר המחלקות (כמות הדרכים לסדר הכדורים בשורה תקינה) הוא $\frac{(2n)!}{2^n}$.

בחירה בלי סדר ובלי חזרות

שאלה: נתונים n כדורים - k שחורים, $n-k$ לבנים. בכמה דרכים ניתן לסדרם בשורה?

תשובה: יש $n!$ לסדר את הכדורים בשורה (אם מבחינים בן כדורים שווי צבע). יש $(n-k)!$ סידורים פנימיים של לבנים ו $k!$ סידורים פנימיים של שחורים. סה"כ: $\frac{n!}{k!(n-k)!}$.

הגדרה: מקדם בינומי $\binom{n}{k}=\frac{n!}{k!(n-k)!}$ כאשר $k\le n\in\mathbb{N}$. (קרי: "n מעל k", "n בחר k").

פרשנויות:

- מספר התת קבוצות מגודל k של $[n]$.
- מספר הדרכים לבחור k נציגים מתוך n אנשים.
- מספר הדרכים לחלק k כדורים ל n מגירות כך שיש לכל היותר כדור אחד בכל מגירה.
- מספר האפשרויות לשלוף k איברים מתוך קבוצה בגודל n בלי החזרה ובלי חשיבות לסדר.

זהויות קומבינטוריות

שאלה: בכמה דרכים ניתן לבחור k איברים מתוך $[n]$ כך ש:

א. n נבחר.

ב. n לא נבחר.

תשובה: עבור הסעיף הראשון: נשארו $n-1$ איברים מתוכם יש לבחור, לכן $\binom{n-1}{k-1}$. עבור הסעיף השני: נשארו $n-1$ איברים מתוכם יש לבחור, לכן $\binom{n-1}{k}$.

מסקנה (זהות פסקל): $\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k}$.

זהות אבן עזרא: $\binom{n}{k}=\binom{n}{n-k}$.

משפט הבינום של ניוטון: לכל $a,b\in\mathbb{R}$ ו $n\in\mathbb{N}$ מתקיים: $(a+b)^n=\sum\limits_{k=0}^{n}\binom{n}{k}a^kb^{n-k}$.

הוכחה: $\overbrace{(a+b)(a+b)\cdot\cdots\cdot(a+b)}^{n}=(a+b)^n$. הדרגה הכוללת של כל מחובר אחרי פתיחת הסוגריים היא n. איבר כללי יהיה מסוג a^kb^{n-k} כאשר $0\le k\le n$. כדי לבחור את a^kb^{n-k} יש לבחור k קבוצות (סוגריים) שמהן נלקח a ומהשר נלקח b. מספר האפשרויות לכך הוא $\binom{n}{k}$.

הרצאה 3

המשך זהויות קומבינטוריות

משולש פסקל: בהרצאה הקודמת ראינו את זהות פסקל. ננסה לתאר אותה בצורה גיאומטרית.

$$
\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}
\qquad\qquad
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

נכתוב את המקדמים הבינומים בצורה הבאה (כמתואר מימין). נשים לב שלפי זהות פסקל, הסכום של כל שני מקדמים הוא המקדם שנמצא ביניהם בשורה מתחת, למשל $\binom{3}{1} = \binom{2}{1} + \binom{2}{0}$. בדרך שבה כתבנו את המקדמים, שוקי המשולש תמיד יהיו מורכבים מאחדות, לכן, נוכל לסכום אותם בדרך שהצגנו, וכך לחשב במהירות מקדמים בינומים.

טענה: $\sum\limits_{k=0}^{n} \binom{n}{k} = 2^n$.

הוכחה: ידוע כי $|\{A : A \subseteq [n]\}| = 2^n$. ראינו כבר כי $|\{A : A \subseteq [n] \wedge |A| = k\}| = \binom{n}{k}$, לכן $\sum\limits_{k=0}^{n} \binom{n}{k} = 2^n$. במילים: גודל קבוצת החזקה של $[n]$ הוא 2^n, ומספר תתי הקבוצות מגודל k הוא $\binom{n}{k}$. לכן אם נסכום מספר תתי הקבוצות מכל גודל, נקבל מספר תתי הקבוצות - כלומר גודל קבוצת החזקה. דרך נוספת היא להציב $a = b = 1$ במשפט ניוטון: $2^n = (1+1)^n = \sum\limits_{k=0}^{n} \binom{n}{k} 1^k 1^{n-k} = \sum\limits_{k=0}^{n} \binom{n}{k}$.

טענה: $\sum\limits_{k=0}^{n} \binom{n}{k} (-1)^k = 0$.

הוכחה: נציב $a = -1, b = 1$ במשפט ניוטון: $0 = (-1+1)^n = \sum\limits_{k=0}^{n} \binom{n}{k} (-1)^k 1^{n-k} = \sum\limits_{k=0}^{n} \binom{n}{k} (-1)^k$.

טענה: $\sum\limits_{k=0}^{n} k \cdot \binom{n}{k} = n 2^{n-1}$.

הוכחה: לפי זהות ניוטון, לכל x: $(1+x)^n = \sum\limits_{k=0}^{n} \binom{n}{k} x^k$. נגזור ונקבל $n(1+x)^{n-1} = \sum\limits_{k=0}^{n} k \binom{n}{k} x^{k-1}$. נציב $x = 1$ ונקבל את הדרוש.

דרך נוספת: נתבונן במספר הדרכים לבחור תת קבוצה של $[n]$ שבה איבר אחד בצבע אדום. <u>שיטה 1</u> לספור זאת היא לבחור תת קבוצה בגודל k - $\binom{n}{k}$ אפשרויות, ולצבוע איבר אחד באדום - k אפשרויות. כלומר מספר הדרכים הוא $\sum\limits_{k=0}^{n} k \cdot \binom{n}{k}$. <u>שיטה 2</u> לספור זאת היא תחילה לבחור את האיבר האדום - n אפשרויות, ואז לבחור את היתר - 2^{n-1} אפשרויות. לכן מספר הדרכים הוא $n \cdot 2^{n-1}$. שתי הספירות שוות לכן נקבל $\sum\limits_{k=0}^{n} k \cdot \binom{n}{k} = n \cdot 2^{n-1}$.

בחירה בלי סדר ועם חזרה

עלינו לספור כמה קבוצות של איברים מ $[n]$ בגודל k יש - <u>עם חזרות</u>. נסמן קבוצה כזו ע"י: $\left\{ \overbrace{1, \ldots, 1}^{x_1}, \overbrace{2, \ldots, 2}^{x_2}, \ldots, \overbrace{n, \ldots, n}^{x_n} \right\}$.

נקבל בעיה שקולה: כמה פתרונות אי שליליים שלמים יש למשוואה $x_1 + \cdots + x_n = k$? נמיר זאת שוב בבעיה אחרת - בכמה דרכים ניתן להכניס n כוכבים לתוך k תאים, כאשר הכוכבים זהים אך התאים ממוספרים? ראשית, נבחין שנוכל לייצג קונפיגורציה של סידור שכזה ע"י kיה. כעת, נסדר את כל הכוכבים בשורה ונכניס סורגים (או מחיצות) כדי לתאר כמה כדורים יש בכל תא. למשל עבור $n = 7$ ו $k = 5$, הקונפיגורציה $(4, 0, 1, 2, 0)$ תיוצג על ידי:

$$\star \star \star \star \mid \mid \star \mid \star \star \mid$$

סך הכל, יש $n+k-1$ אובייקטים (n כדורים ו $k-1$ סורגים) ולכן יש $(n+k-1)!$ דרכים לסדרם. אבל, צריך להתחשב בסידורים פנימיים, שהרי הסורגים והכוכבים זהים, לכן יש $\frac{(n+k-1)!}{n!\cdot(k-1)!}=\binom{n+k-1}{n-1}$ דרכים לסדרם.

כעת, השלמנו את כל המקרים האפשריים עבור בחירה, ונקבל את הטבלה:

k מתוך n	עם החזרה	בלי החזרה
עם חשיבות לסדר	n^k	$\frac{n!}{(k-n)!}$
בלי חשיבות לסדר	$\binom{n+k-1}{n-1}$	$\binom{n}{k}$

דוגמאות לחישובי הסתברות

הגדרה: מרחב הסתברות בדידה $(\Omega,\mathcal{F},\mathbb{P}_p)$ נקרא אחיד אם לכל $\omega_1,\omega_2\in\Omega$ מתקיים $p(\omega_1)=p(\omega_2)$.

הערה: מאחר ש $\sum_{\omega\in\Omega}p(\omega)=1$, אז במרחב הסתברות אחיד יתקיים כי $p(\omega)=\frac{1}{|\Omega|}$. לכן לכל מאורע A מתקיים כי

$$\mathbb{P}(A)=\sum_{\omega\in A}p(\omega)=\frac{|A|}{|\Omega|}$$

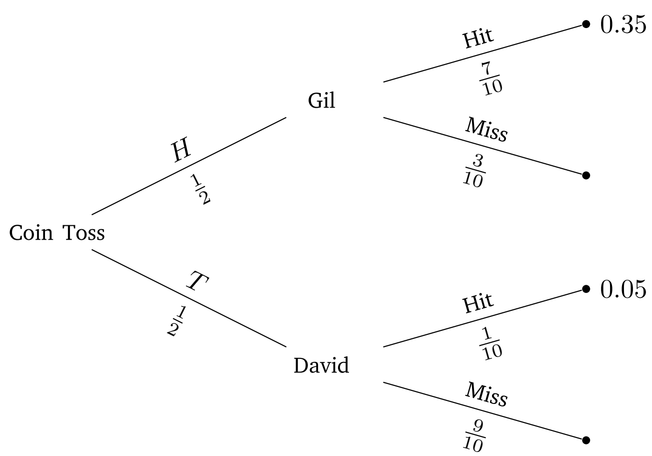
דוגמה: מה הסיכוי שסכום שתי קוביות הוגנות הוא 7?

פתרון: מתקיים כי $\Omega=[6]^2$, נסמן בA את המאורע הרצוי. אזי: $A=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$. לכן:

$$\mathbb{P}(A)=\frac{|A|}{|\Omega|}=\frac{6}{36}=\frac{1}{6}$$

דוגמה: גיל פוגע במטרה בהסתברות 0.7. דוד פוגע במטרה בהסתברות 0.1. מטילים מטבע כדי להכריע מי משניהם ינסה לקלוע. מה הסיכוי שבוצעה פגיעה במטרה?

פתרון: ניעזר בעקרון הכפל ונרכיב עץ עבור בעיה זו, ונקבל כי ההסתברות שגיל יבחר ויפגע היא 0.35 וההסתברות שדוד יבחר ויפגע היא 0.05. נסכום את השתיים ונקבל כי הסיכוי שבוצעה פגיעה במטרה היא 0.4.



דוגמה: בכיתה n ילדים שכל אחד נולד ביום מקרי אחיד בשנה. מה ההסתברות ששני ילדים נולדו באותו היום?

פתרון: $\Omega=[365]^n$ ונסמן $A=\{(x_1,\ldots,x_n)\in\Omega\mid\exists i\neq j: x_i=x_j\}$. נשים לב כי $A^c=\{(x_1,\ldots,x_n)\in\Omega:\forall i\neq j\ x_i\neq x_j\}$, לכן $|A^c|=365\cdot364\cdot\cdots\cdot(365-n+1)=\frac{365!}{(365-n)!}$, כי לילד הראשון יש 365 אפשרויות ליום הולדתו, לילד השני 364 וכן הלאה. כמו כן, $|\Omega|=365^n$. כעת: $\mathbb{P}(A)=1-\mathbb{P}(A^c)=1-\frac{|A^c|}{|\Omega|}=1-\frac{365!}{(365-n)!\cdot365^n}$.

נוסחת ההסתברות השלמה

הגדרה: האוסף $\{A_j\}_{j\in J}$ יקרא חלוקה של Ω אם לכל $i\neq j\in J$ מתקיים $A_i\cap A_j=\emptyset$ ו $\Omega=\dot{\bigcup\limits_{j\in J}}A_j$.

משפט (נוסחת ההסתברות השלמה): תהי $\{A_j\}_{j=1}^{\infty}$ חלוקה בת מניה של Ω. אזי לכל $B\in\mathcal{F}$:

$$\mathbb{P}(B)=\mathbb{P}\left(\bigcup_{j=1}^{\infty}(B\cap A_j)\right)=\sum_{j=1}^{\infty}\mathbb{P}(B\cap A_j)$$

הוכחה: נשים לב כי $B=\dot{\bigcup\limits_{j=1}^{\infty}}(B\cap A_j)$ לכן הטענה נובעת ישירות מסכימות בת מנייה של פונקצית ההסתברות.

דוגמה: מטילים קוביית 6 וקוביית 12. מה הסיכוי שהתקבלה בקוביית ה 12 תוצאה נמוכה יותר?

פתרון: $\Omega=[6]\times[12]$. נסמן $B=\{(x,y)\in\Omega:x>y\}$. נגדיר חלוקה $A_k=\{k\}\times[12]$. מתקיים כי $\mathbb{P}(B\cap A_k)=\frac{|B\cap A_k|}{|\Omega|}=\frac{k-1}{7\cdot 2}$. לכן: $\mathbb{P}(B)=\sum\limits_{k=1}^{6}\frac{k-1}{7\cdot2}=\frac{5}{24}$.

דוגמה: קוביה מוטה בת 6 פאות מוטלת 5 פעמים. הסיכוי לקבל את התוצאה i פורפוציונאלי ל 21. מה הסיכוי שהתוצאה הראשונה התקבלה בדיוק פעם אחת, אם הטלנו 5 פעמים?

פתרון: למעשה נתון $p(i)=\frac{i}{21}$, $\Omega=[6]^5$. נסמן $B=\{(x_1,\ldots,x_5:\forall i>1\ x_1\neq x_i)\}$, ונגדיר $A_k=\{k\}\times[6]^4$. מתקיים כי: $\mathbb{P}(B\cap A_k)=p(k)\cdot(1-p(k))^4$. נקבל: $\mathbb{P}(B)=\sum\limits_{k\in[6]}\frac{k}{21}\left(1-\frac{k}{21}\right)^4\approx 0.417$.

הרצאה 4

חסם האיחוד

משפט (חסם האיחוד): יהיו $\{A_n\}_{n=1}^m$ סדרת מאורעות כלשהם במרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$, אזי: $\mathbb{P}\left(\bigcup\limits_{j=1}^m A_j\right)\leq\sum\limits_{j=1}^m\mathbb{P}\left(A_j\right)$.

הוכחה: באינדוקציה. מקרה בסיס $m=2$: נשים לב כי $A\cup B=A\uplus\left(B\backslash A\right)$ לכן:

$$\mathbb{P}\left(A\cup B\right)=\mathbb{P}\left(A\uplus\left(B\backslash A\right)\right)=\mathbb{P}\left(A\right)+\mathbb{P}\left(B\backslash A\right)\leq\mathbb{P}\left(A\right)+\mathbb{P}\left(B\right)$$

כנדרש. נניח נכונות ל m ונוכיח ל $m+1$:

$$\mathbb{P}\left(\bigcup_{j=1}^{m+1}A_j\right)=\mathbb{P}\left(A_{m+1}\cup\bigcup_{j=1}^{m}A_j\right)\leq\mathbb{P}\left(\bigcup_{j=1}^{m}A_j\right)+\mathbb{P}\left(A_{m+1}\right)\leq\sum_{j=1}^{m}\mathbb{P}\left(A_j\right)+\mathbb{P}\left(A_{m+1}\right)=\sum_{j=1}^{m+1}\mathbb{P}\left(A_j\right)$$

טענה: איחוד סופי של מאורעות זניחים (הסברותם 0) הוא מאורע זניח (הסתברותו 0).

הוכחה: יהיו $\{A_j\}_{j=1}^m$ מאורעות זניחים, כלומר $\mathbb{P}\left(A_j\right)=0$ לכל j. אזי מתקיים: $0\leq\mathbb{P}\left(\bigcup\limits_{j=1}^m A_j\right)\leq\sum\limits_{j=1}^m\mathbb{P}\left(A_j\right)=0$.

מסקנה: חיתוך סופי של מאורעות שכל אחד מהם מתרחש כמעט תמיד (התסברותו 1) הוא מאורע שמתרחש כמעט תמיד (הסתברותו 1).

דוגמה: ידוע כי ההסתברות שירד גשם מחר היא 0.5, ואילו ההסתברות שטל ילך לעבודה מחר היא 0.7. יש להראות כי ההסתברות שירד גשם מחר ולמרות זאת טל ילך לעבודה היא לפחות 0.2.

פתרון: נסמן ב A את המאורע שירד גשם וב B את המאורע שטל ילך לעבודה. אזי $\mathbb{P}\left(A\cap B\right)=1-\mathbb{P}\left(A^c\cup B^c\right)\geq 1-\mathbb{P}\left(A^c\right)-\mathbb{P}\left(B^c\right)=0.2$ מחסם האיחוד.

דוגמה: בוחרים באקראי ובאופן אחיד n מספרים מתוך $[m]$. נבקש לקבל חסם עליון על ההסתברות שנבחר אותו מספר פעמיים.

פתרון: נסמן $A=\{(x_1,\dots,x_n)\in\Omega\mid\exists i\neq j\in[n]:x_i=x_j\}$ ולכל $1\leq i<j\leq n$ נגדיר $B_{ij}=\{(x_1,\dots,x_n)\in\Omega\mid x_i=x_j\}$. אזי מתקיים $A=\bigcup\limits_{1\leq i<j\leq n}B_{ij}$. נשים לב כי $|B_{ij}|=m^{n-1}$ ולכן $\mathbb{P}\left(B_{ij}\right)=\frac{m^{n-1}}{m^n}=\frac{1}{m}$. נשתמש בחסם האיחוד ונקבל

$$\mathbb{P}\left(A\right)=\mathbb{P}\left(\bigcup_{1\leq i<j\leq n}B_{ij}\right)\leq\sum_{1\leq i<j\leq n}\mathbb{P}\left(B_{ij}\right)=\sum_{1\leq i<j\leq n}\frac{1}{m}=\frac{n\left(n-1\right)}{2m}$$

רציפות פונקציית ההסתברות

הגדרה: יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. סדרת מאורעות $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ נקראת עולה אם לכל j מתקיים $A_j \subseteq A_{j+1}$. הסדרה נקראת יורדת אם $A_{j+1} \subseteq A_j$ לכל j.

משפט (רציפות פונקציית ההסתברות): יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ותהי $\{A_n\}_{n=1}^{\infty}$ סדרה עולה של מאורעות. אזי

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

הוכחה: נגדיר סדרת מאורעות $\{B_n\}_{n=1}^{\infty}$ באופן הבא: נקבע $B_1 = A_1$ ולכל $n > 1$ נגדיר $B_n = A_n \backslash A_{n-1}$. מתקיים כי מאורעות אלו זרים בזוגות. כדי להוכיח זאת, מספיק להראות כי עבור $m > n$, לכל $\omega \in B_n$ מתקיים $\omega \notin B_m$. אכן, עבור $\omega \in B_n$ מתקיים $\omega \notin A_{m-1}$ ולכן $\omega \notin A_m$ כי הסדרה $\{A_n\}_{n=1}^{\infty}$ עולה. מכאן: $\omega \notin B_m$ כי $B_m \subseteq A_m$.

מצד שני, נראה באינדוקציה כי: $\bigcup_{k\in[n]} A_k = A_n$ $\dot\bigcup_{k\in[n]} B_k =$. בסיס $n = 1$ מתקיים באופן מידי כי $A_1 = B_1$ ע"פ הגדרה. נניח נכונות עבור n ונוכיח עבור $n+1$:

$$\dot\bigcup_{k\in[n+1]} B_k = B_{n+1} \cup \dot\bigcup_{k\in[n]} B_k = B_{n+1} \cup A_n = (A_{n+1} \backslash A_n) \cup A_n = A_n$$

כעת נובע נובע באופן מידי כי $\bigcup_{n=1}^{\infty} A_n = \dot\bigcup_{k=1}^{\infty} B_k$. נותר רק להשלים את ההוכחה:

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\dot\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} \mathbb{P}\left(B_k\right) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}\left(B_k\right) = \lim_{n\to\infty} \mathbb{P}\left(\dot\bigcup_{k=1}^{n} B_k\right) = \lim_{n\to\infty} \mathbb{P}\left(A_n\right)$$

מסקנה: יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ותהי $\{A_n\}_{n=1}^{\infty}$ סדרה יורדת של מאורעות. אזי $\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}\left(A_n\right)$.

משפט (חסם איחוד בן מנייה): לכל סדרת מאורעות $\{A_n\}_{n=1}^{\infty}$ במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ מתקיים $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(A_n\right)$.

הוכחה: נגדיר $B_n = \bigcup_{k=1}^{n} A_k$. זו סדרת מאורעות עולה המקיימת $\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n} B_k$. כעת נשתמש בחסם האיחוד למספר סופי של מאורעות וברציפות פונקציית ההסתברות ונקבל:

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mathbb{P}\left(B_n\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \leq \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}\left(A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}\left(A_k\right)$$

עקרון ההכלה וההפרדה

טענה: יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות.

1. יהיו A, B מאורעות, אזי מתקיים $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
2. יהיו A, B, C מאורעות, אזי מתקיים $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) - \mathbb{P}(A \cap C) + \mathbb{P}(A \cap B \cap C)$.

הוכחה: (1) נשים לב כי $A \cup B = (A \backslash B) \uplus (B \backslash A) \uplus (A \cap B)$, לכן מתכונת הסכימות:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \backslash B) + \mathbb{P}(B \backslash A) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

(2) נשים לב כי:

$$\mathbb{P}((A \cup B) \cup C) = \mathbb{P}(A \cup B) + \mathbb{P}(C) - \mathbb{P}((A \cup B) \cap C)$$

אבל $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ לכן:

$$= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - (\mathbb{P}(A \cup C) + \mathbb{P}(B \cup C) - \mathbb{P}(A \cap B \cap C))$$

$$= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cup C) - \mathbb{P}(B \cup C) + \mathbb{P}(A \cap B \cap C)$$

משפט (עקרון ההכלה וההפרדה): יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו $A_1, \ldots, A_n$ מאורעות. אזי מתקיים:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} \mathbb{P}(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} \mathbb{P}(A_1 \cap \cdots \cap A_n)$$

או בצורה מפורשת יותר: נסמן $A_I = \bigcap_{j \in I} A_j$. אזי:

$$\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}(A_I)$$

הוכחה: באינדוקציה. את בסיס האינדוקציה $n = 2$ הוכחנו כבר בטענה הקודמת, ואף הוכחנו את המקרה $n = 3$. כעת נניח נכונות עבור n ונוכיח עבור $n + 1$:

$$\mathbb{P}\left(\bigcup_{j=1}^{n+1} A_j\right) = \mathbb{P}\left(A_{n+1} \cup \left(\bigcup_{j=1}^{n} A_j\right)\right) = \mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(A_{n+1} \cap \left(\bigcup_{j=1}^{n} A_j\right)\right)$$

$$= \mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\bigcup_{j=1}^{n} (A_j \cap A_{n+1})\right) \stackrel{*}{=} \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}(A_I) + \mathbb{P}(A_{n+1}) - \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}\left(A_{I \cup \{n+1\}}\right)$$

$$= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}(A_I) + \sum_{I \subseteq [n]} (-1)^{|I \cup \{n+1\}|+1} \mathbb{P}\left(A_{I \cup \{n+1\}}\right)$$

כאשר במעבר האחרון כללנו $\mathbb{P}\left(A_{n+1}\right)$ בסכום הימני על ידי הוספת המקרה של קבוצה ריקה, וכן החלפנו פלוס במינוס ושינינו בהתאם את חזקת (-1). כעת נסמן $J = I \cup \{n+1\}$ ונמשיך:

$$= \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}\left(A_I\right) + \sum_{n+1 \in J \subseteq [n+1]} (-1)^{|J|+1} \mathbb{P}\left(A_J\right) = \sum_{\emptyset \neq I \subseteq [n+1]} (-1)^{|I|+1} \mathbb{P}\left(A_I\right)$$

כאשר השתמשנו בכך שאוסף התת קבוצות הלא ריקות של $[n+1]$ ניתן לתיאור כאיחוד זר של קבוצות שמכילות את $n+1$ ותת קבוצות לא ריקות של $[n]$.

* ניתן להשתכנע במעבר על ידי הצבת $B_j = A_{I\cup\{n+1\}}$ ופיתוח הביטוי:

$$\mathbb{P}\left(\bigcup_{j=1}^{n}\left(\overbrace{A_j \cap A_{n+1}}^{B_j}\right)\right) = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}\left(B_I\right)$$

אכן:

$$B_I = \bigcap_{j\in I} B_j = \bigcap_{j\in I}\left(A_j \cap A_{n+1}\right) = A_{n+1} \cap \bigcap_{j\in I} A_j = \bigcap_{j\in I\cup\{n+1\}} A_j = A_{I\cup\{n+1\}}$$

הרצאה 5

עקרון ההכלה והפרדה - המשך

דוגמא: מדפיסים n מכתבים המיועדים לנמענים שונים ומכניסים כל אחד למעטפה שונה באקראי - ושולחים.

א. מהי ההסתברות שכל המכתבים ישלחו לנמען הנכון?

נחשוב על המספרים ועל המעטפות כממוספרים מ 1 עד n בהתאם לנמען שאליו הם מכוונים. כל תוצאה של הניסוי תייצוג ע"י וקטור $(x_1,\dots,x_n)$. בייצוג זה, מרחב המדגם הוא מרחב כל התמורות על $[n]$. כלומר:

$$\Omega=\{(x_1,\dots,x_n)\mid\forall i,j\in[n]:i\neq j\implies x_i\neq x_j\}$$

ומתקיים $|\Omega|=n!$. פונקציית ההסתברות המתאימה לניסוי היא אחידה ולכן לכל A אוסף של תמורות על $[n]$ מתקיים $\mathbb{P}(A)=\frac{|A|}{n!}$.

כעת, המאורע בו כל המכתבים נשלחים ליעדם מכיל רק את תמורת הזהות כי רק היא מקיימת $\forall i:f(i)=i$. לכן ההסתברות $\frac{1}{n!}$.

ב. מה ההסתברות שאף מכתב לא הגיע ליעדו?

נגדיר את המאורע שאף מכתב לא הגיע ליעדו $A=\{(x_1,\dots,x_n)\mid\forall i\in[n]:x_i\neq i\}$. כדי לנתח המאורע, נגדיר B_i להיות המאורע שהמכתב ה i הוכנס למעטפה הנכונה, ובאופן דומה את B_I להיות המאורע שכל המכתבים $i\in I$ הוכנסו למעטפות הנכונות:

$$B_i=\{(x_1,\dots,x_n)\in\Omega\mid x_i=i\}$$

$$B_I=\{(x_1,\dots,x_n)\in\Omega\mid\forall i\in I:x_i=i\}$$

נשים לב כי:

$$A^c=\{(x_1,\dots,x_n)\mid\exists i\in[n]:x_i=i\}$$

לכן $A^c=\bigcup_{i=1}^{n}B_i$. נרצה להשתמש בעיקרון ההכלה וההדחה. נשים לב כי לכל $i\in[n]$ מתקיים:

$$\mathbb{P}(B_i)=\frac{|B_i|}{|\Omega|}=\frac{(n-1)!}{n!}=\frac{1}{n}$$

כי מספר האיברים ב B_i הוא מספר התמורות על $[n]\setminus\{i\}$. באופן דומה, עבור $i<j$ מתקיים:

$$\mathbb{P}(B_i\cap B_j)=\frac{|B_i\cap B_j|}{|\Omega|}=\frac{(n-2)!}{n!}=\frac{1}{n(n-1)}$$

ועבור קבוצת אינדקסים כללית $I\subseteq[n]$:

$$\mathbb{P}(B_I)=\frac{\left|\bigcap_{i\in I}B_i\right|}{|\Omega|}=\frac{(n-|I|)!}{n!}$$

כעת נשתמש בעקרון ההכלה וההדחה ונקבל:

$$\mathbb{P}(A^c)=\mathbb{P}\left(\bigcup_{i=1}^{n}B_i\right)=\sum_{i=1}^{n}\mathbb{P}(B_i)-\sum_{1\leq i<j\leq n}\mathbb{P}(B_i\cap B_j)+\cdots+(-1)^{n+1}\,\mathbb{P}\left(B_{[n]}\right)$$

$$=\binom{n}{1}\cdot\frac{(n-1)!}{n!}-\binom{n}{2}\cdot\frac{(n-2)!}{n!}+\binom{n}{3}\cdot\frac{(n-3)!}{n!}+\cdots+(-1)^{n+1}\binom{n}{n}\cdot\frac{(n-n)!}{n!}$$

$$=\sum_{\ell=1}^{n}(-1)^{\ell+1}\binom{n}{\ell}\cdot\frac{(n-\ell)!}{n!}=\sum_{\ell=1}^{n}(-1)^{\ell+1}\frac{\cancel{n!}}{\cancel{(n-\ell)!}\cdot\ell!}\cdot\frac{\cancel{(n-\ell)!}}{\cancel{n!}}=\sum_{\ell=1}^{n}\frac{(-1)^{\ell+1}}{\ell!}$$

לכן:

$$\mathbb{P}\left(A\right)=1-\mathbb{P}\left(A^{c}\right)=1-\sum_{\ell=1}^{n}\frac{\left(-1\right)^{\ell+1}}{\ell!}=\sum_{\ell=0}^{n}\frac{\left(-1\right)^{\ell}}{\ell!}$$

זהו התחלתו של פיתוח טיילור של e^{-x} סביב $x_0=0$ עבור $x=-1$, ולכן $\mathbb{P}\left(A\right)\underset{n\to\infty}{\longrightarrow}\frac{1}{e}$.

כמו כן, מאחר שהמרחב אחיד, הרי שמתקיים $\mathbb{P}\left(A\right)=\frac{|A|}{|\Omega|}$. נוכל להסיק אם כן את הנוסחא הבאה למספר התמורות אשר אין להן נקודת שבת:

$$|A|=n!\cdot\sum_{\ell=0}^{n}\frac{\left(-1\right)^{\ell}}{\ell!}$$

ג. מה ההסתברות שבדיוק k מכתבים הגיעו ליעדם?

נגדיר מאורע C_n^k המתאר את אוסף התמורות על $[n]$ עם בדיוק k נקודות שבת. מאורע זה מתאים לכך שבדיוק k מכתבים הגיעו ליעדם.

נסמן, עבור $I\subseteq[n]$:

$$D_I=\left(\bigcap_{i\in I}B_i\right)\cap\left(\bigcap_{i\in I}B_i^c\right)$$

המאורע שכל המכתבים בקבוצה I הגיעו ליעדם ואילו השאר לא. כעת מתקיים:

$$C_n^k=\bigcupdot_{|I|=k}D_I$$

נרצה אם כן, לחשב את הסתברות D_I עבור $|I|=k$. מספר האיברים ב I הוא כמספר התמורות על $[n]\setminus I$ אשר אין להן נקודות שבת. לפי החישוב הקודם:

$$|D_I|=(n-k)!\cdot\sum_{\ell=0}^{n-k}\frac{\left(-1\right)^{\ell}}{\ell!}$$

לכן:

$$\mathbb{P}\left(D_I\right)=\frac{|D_I|}{n!}=\frac{(n-k)!}{n!}\cdot\sum_{\ell=0}^{n-k}\frac{\left(-1\right)^{\ell}}{\ell!}$$

על פי סכימות בת מניה:

$$\mathbb{P}\left(C_n^k\right)=\mathbb{P}\left(\bigcupdot_{|I|=k}D_I\right)=\sum_{|I|=k}\mathbb{P}\left(D_I\right)=\sum_{|I|=k}\left(\frac{(n-k)!}{n!}\cdot\sum_{\ell=0}^{n-k}\frac{\left(-1\right)^{\ell}}{\ell!}\right)$$

ראינו כבר שכמות תת הקבוצות מגודל k של $[n]$ היא $\binom{n}{k}$ לכן:

$$=\binom{n}{k}\cdot\frac{(n-k)!}{n!}\cdot\sum_{\ell=0}^{n-k}\frac{\left(-1\right)^{\ell}}{\ell!}=\frac{1}{k!}\cdot\frac{(n-k)!}{n!}\cdot\sum_{\ell=0}^{n-k}\frac{\left(-1\right)^{\ell}}{\ell!}$$

עבור k קבוע ו n שואף לאינסוף נקבל:

$$\lim_{n\to\infty}\mathbb{P}\left(C_n^k\right)=\frac{1}{e\cdot k!}$$

הסתברות מותנית

הגדרה: יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. יהיו $A, B \in \mathcal{F}$ כך ש $\mathbb{P}(B) > 0$. נגדיר את ההסתברות המותנית של A בהינתן B על ידי:

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

טענה: $(\Omega, \mathcal{F}, \mathbb{P}_B)$ הוא מרחב הסתברות.

הוכחה: עלינו לוודא כי $\mathbb{P}_B$ אכן פונקצית הסתברות כשרה. ראשית:

$$\mathbb{P}_B(\Omega) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

נותר להראות כי $\mathbb{P}_B$ מקיימת סכימות בת מנייה. תהי $\{A_n\}_{n=1}^{\infty}$ סדרת מאורעות זרים, לפי סכימות בת מניה של $\mathbb{P}$ נקבל:

$$\begin{aligned}\mathbb{P}_B\left(\dot{\bigcup_{n=1}^{\infty}} A_n\right) &= \frac{\mathbb{P}\left(\left(\dot{\bigcup_{n=1}^{\infty}} A_n\right) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}\left(\dot{\bigcup_{n=1}^{\infty}} (A_n \cap B)\right)}{\mathbb{P}(B)} \\ &= \frac{1}{\mathbb{P}(B)} \sum_{n=1}^{\infty} \mathbb{P}(A_n \cap B) = \sum_{n=1}^{\infty} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \mathbb{P}_B(A_n)\end{aligned}$$

דוגמה: בכד 5 כדורים ירוקים, 10 שחורים ו-10 לבנים. כדור נשלף באקראי.

א. מה ההסתברות שצבעו ירוק?

נשתמש במרחב ההסתברות האחיד על $\Omega = [25]$, כאשר הכדורים הירוקים מיוצגים ע"י המספרים $\{1, \ldots, 5\}$ הלבנים ע"י $\{6, \ldots, 15\}$ והשחורים ע"י $\{15, \ldots, 25\}$. המאורע המבוקש הוא $A = [5]$. מאחר שהמרחב אחיד נקבל שהסתברותו $5/25 = 1/5$.

ב. מה ההסתבות שצבעו ירוק, אם ידוע שלא יצא כדור שחור?

נתנה אם כן במאורע $B = [15]$ - הכדורים הירוקים והלבנים. נקבל:

$$\mathbb{P}\left(\begin{array}{c}\text{כדור} \\ \text{ירוק}\end{array} \;\middle|\; \begin{array}{c}\text{לא} \\ \text{שחור}\end{array}\right) = \mathbb{P}(A \mid B) = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{|\{1, \ldots, 5\}|}{|\{1, \ldots, 15\}|} = \frac{5}{15} = \frac{1}{3}$$

דוגמה: מוטלות שתי קוביות הוגנות.

א. מה ההסתברות שסכום התוצאות יהיה 8?

מרחב המדגם שלנו הוא $\Omega = [6]^2$. את המאורע "סכום הקוביות הוא 8" נסמן ב A, כלומר: $A = \{(x,y) \in \Omega : x + y = 8\}$. כעת:

$$\mathbb{P}(A) = \frac{|\{(2,6),(3,5),(4,4),(5,3),(6,2)\}|}{|\Omega|} = \frac{5}{36}$$

ב. כיצד משתנה ההסתברות אם ידוע שתוצאת הקוביה הראשונה היא 3?

ההסתברות החדשה היא:

$$\mathbb{P}(A \mid \{(x,y) \in \Omega : x = 3\}) = \frac{\mathbb{P}(A \cap \{(x,y) \in \Omega : x = 3\})}{\mathbb{P}(\{(x,y) \in \Omega : x = 3\})} = \frac{\mathbb{P}(\{(3,5)\})}{\mathbb{P}(\{(3,1),\ldots,(3,6)\})} = \frac{1/36}{6/36} = \frac{1}{6}$$

ג. מה אם נודע לנו רק שהתוצאה של לפחות אחת מהקוביות היא 3?

ההתסברות החדשה היא:

$$\mathbb{P}(A \mid \{(x,y) \in \Omega : x = 3 \vee y = 3\}) = \frac{\mathbb{P}(A \cap \{(x,y) \in \Omega : x = 3 \vee y = 3\})}{\mathbb{P}(\{(x,y) \in \Omega : x = 3 \vee y = 3\})}$$

$$= \frac{\mathbb{P}(\{(3,5),(5,3)\})}{\mathbb{P}(\{(3,1),\ldots,(3,6),(1,3),\ldots,(1,6)\})} = \frac{2/36}{11/36} = \frac{2}{11}$$

אבחנה (כלל השרשרת): יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ויהיו A, B מאורעות כך ש $\mathbb{P}(B) > 0$. אזי:

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

דוגמה: נזכר שוב בדוגמה של גיל ודוד: גיל פוגע בהסתברות 7.0 .דוד פוגע בהסתברות 1.0. מטילים מטבע כדי להכריע מי משניהם ינסה לקלוע. לעומת זאת, הפעם ידוע שהקליע פגע - מה ההסתברות שהיורה היה גיל?

ראשית, נבחר $\Omega = \{\text{גיל}, \text{דוד}\} \times \{\text{פגע}, \text{החטיא}\}$ והרי שנתוני השאלה הם: $\mathbb{P}((\text{דוד}, \cdot)) = \mathbb{P}((\text{גיל}, \cdot)) = 0.5$, $\mathbb{P}((\cdot, \text{פגע}) \mid (\text{גיל}, \cdot)) = 0.7$ ו- $\mathbb{P}((\cdot, \text{פגע}) \mid (\text{דוד}, \cdot)) = 0.1$.

כעת:

$$\mathbb{P}((\text{גיל}, \cdot) \mid (\cdot, \text{פגע})) = \frac{\mathbb{P}((\text{גיל}, \cdot) \cap (\cdot, \text{פגע}))}{\mathbb{P}((\cdot, \text{פגע}))} = \frac{\mathbb{P}((\text{גיל}, \text{פגע}))}{\mathbb{P}((\cdot, \text{פגע}))} = \frac{0.5 \cdot 0.7}{0.5 \cdot 0.7 + 0.5 \cdot 0.1} = \frac{0.35}{0.4} = \frac{7}{8}$$

נוסחת ההסתברות השלמה וכלל בייס

טענה (נוסחת ההתסברות השלמה עם התניה): תהי $\mathcal{A}$ חלוקה בת מנייה של מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. אזי לכל מאורע B מתקיים

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

הוכחה:

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

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

כאשר המעבר הראשון נובע מנוסחת ההתסברות השלמה והמעבר הראשון בשורה השנייה נובע מכלל השרשרת.

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

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

הוכחה: ההוכחה נובעת ישירות מהגדרת ההתסברות המותנית:

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

דוגמה: שוב נחזור לדוגמה של גיל ודוד, אך הפעם נציג דרך פשוטה יותר לחישוב ההסתברות שהיורה היה גיל, אילו היה ידוע שהקליע פגע:

$$\mathbb{P}((\text{גיל},\cdot)\mid(\cdot,\text{פגע}))=\mathbb{P}((\cdot,\text{פגע})\mid(\text{גיל},\cdot))\cdot\frac{\mathbb{P}((\text{גיל},\cdot))}{\mathbb{P}((\cdot,\text{פגע}))}=0.7\cdot\frac{0.5}{0.4}=\frac{7}{8}$$

הרצאה 6

הגישה הבייסיאנית להפרדת השערות פשוטות

נניח שנתונות לנו מדידות שהתקבלו מאחד משני מרחבי הסתברות - כל מרחב הסתברות מתאים לתיאוריה אחרת בנוגע לתהליך הגרלת הדגימות. עלינו להכריע מהי הסבירות שכל אחת משתי התיאוריות מתארת נכונה את המרחב ממנו הופקו הדגימות. במצב זה מכנים את ההשערה הראשונה $\mathcal{H}_0$ - השערת האפס, ואת השנייה $\mathcal{H}_1$ - ההשערה החלופית.

בכל אחד מהמקרים אנו מניחים שאחת התיאוריות מתארת נכונה את דרך הפקת הדגימות, ושתחת כל אחת מהתיאוריות יש ברשותנו מודל מדוייק לדרך הפקת הדגימות. אלה נתונים על ידי $p_{\mathcal{H}_0}$ ו- $p_{\mathcal{H}_1}$.

הגישה הבייסיאנית היא גישה סטטיסטית שבה רואים את הפקת הדגימות כניסוי דו-שלבי: קודם נבחרה את התאוריה שעל פיה פועלת המציאות, ואחר כך הוגרלו הדגימות לפי תיארויה זו. בגישה בייסיאנית אנו מניחים מידה התחלתית של סבירות שאנו מייחסים לכל אחת משתי התיאוריות, כלומר ההסתברות שאנו מייחסים לכך שהמציאות פועלת לפי תיאוריה זו. ההסתברות שייחסנו לכך שתיאוריה נכונה לפני עריכת הניסוי נקראת ההסתברות אפריורי (Prior) שלה. נסמן הסתברויות אלה על ידי $p(\mathcal{H}_0), p(\mathcal{H}_1)$. מטרתנו תהיה לעדכן את ההסתברות הזו לאור תוצאת הניסוי ω, כלומר לחשב את ההתסברות אפוסטריורי (Postprior) - לאחר ביצועו. באופן פורמלי, נרצה לחשב את:

$$\mathbb{P}\left(\{(\mathcal{H}_0,\cdot)\} \mid \{(\cdot,\omega)\}\right)$$

ניעזר בכלל בייס ובנוסחת ההסתברות השלמה:

$$= \mathbb{P}\left(\{(\cdot,\omega)\} \mid \{(\mathcal{H}_0,\cdot)\}\right)\cdot\frac{\mathbb{P}\left(\{(\mathcal{H}_0,\cdot)\}\right)}{\mathbb{P}\left(\{(\cdot,\omega)\}\right)} = \frac{p_{\mathcal{H}_0}(\omega)\cdot p(\mathcal{H}_0)}{p_{\mathcal{H}_0}(\omega)\,p(\mathcal{H}_0)+p_{\mathcal{H}_1}(\omega)\,p(\mathcal{H}_1)}$$

דוגמה (רוזנקרנץ וגילדנשטרן): כאשר רוזנקרנץ מוציא מטבע מכיסו, גילדנשטרן מאמין שמדובר במטבע הוגן, אך מוכן לקבל כי בהסתברות של אחד למיליון מדובר במטבע מכושף התמיד מוציא "עץ". רוזנקרנץ מטיל את המטבע שלושים פעם, ופעם אחר פעם מתקבל "עץ". כיצד על גילדנשטרן לעדכן את הסבירות שהמטבע מכושף?

נשתמש במרחב המדגם $\Omega=\{\mathcal{H}_0,\mathcal{H}_1\}\times\{H,T\}^{30}$ כאשר $\mathcal{H}_1$ היא ההשערה שהמטבע מכושף ואילו $\mathcal{H}_0$ היא ההשערה שהוא הוגן. כמו כן, נסמן $A=\{(\mathcal{H}_1,\cdot)\}$ - "המטבע מכושף", $B=\left\{\left(\cdot,H^{30}\right)\right\}$ - "שלושים פעמים רצוף קיבלנו עץ".

לפי גילדנשטרן, $\mathbb{P}(A)=p(\mathcal{H}_1)=10^{-6}$ ולכן $\mathbb{P}(A^c)=p(\mathcal{H}_0)=1-10^{-6}$. כמו כן, $\mathbb{P}(B\mid A)=p_{\mathcal{H}_1}\left(H^{30}\right)=1$ ומפני שבשלושים הטלות מטבע הוגן יש 2^{30} סדרות תוצאות אפשריות שוות-הסתברות, הרי ש $\mathbb{P}(B\mid A^c)=p_{\mathcal{H}_0}\left(H^{30}\right)=2^{-30}$. נחשב:

$$\mathbb{P}(A\mid B)=\mathbb{P}(B\mid A)\cdot\frac{\mathbb{P}(A)}{\mathbb{P}(B)}=\frac{\mathbb{P}(B\mid A)\cdot\mathbb{P}(A)}{\mathbb{P}(B\mid A)\cdot\mathbb{P}(A)+\mathbb{P}(B\mid A^c)\,\mathbb{P}(A^c)}$$

$$=\frac{10^{-6}}{10^{-6}+(1-10^{-6})\cdot 2^{-30}}\approx 1-\frac{1}{1073}$$

כאשר המעבר הראשון הוא לפי כלל בייס והמעבר השני הוא לפי נוסחת ההסתברות השלמה.

כלומר, לאחר שחזה בתופעה המופלאה, שינה גילדנשטרן את הערכתו וכעת הסבירות שהוא מייחס לכך שהמטבע הוגן היא כאחד לאלף.

דוגמה: אדם אקראי ניגש לבדוק אם הוא נשא קורונה. ידוע ש 1% מהאוכלוסיה נושאים את המחלה. כאשר נבדק אדם שאינו נשא, ההסתברות לתוצאה חיובית כוזבת היא α וההסתברות לתוצאה שלילית כוזבת היא β. אדם נבדק וקיבל תוצאה חיובית. מה ההסתברות שהוא באמת נושא את הנגיף? יש להעריך בפרט עבור $\alpha,\beta\in\{0,0.05\}$.

נשתמש במרחב המדגם $\Omega=\{\mathcal{H}_0,\mathcal{H}_1\}\times\{+,-\}$ כאשר $\mathcal{H}_0$ מייצגת את ההשערה שהנבדק אינו נשא, ו- $\mathcal{H}_1$ את ההשערה שהנבדק נשא. נסמן $A=\{(\mathcal{H}_1,\cdot)\}$ - המאורע שהנבדק נשא, $B=\{(\cdot,+)\}$ - המאורע שהתוצאה חיובית.

נתון כי: $\mathbb{P}(B^c\mid A)=p_{\mathcal{H}_1}(-)=\beta$ (ההתסברות לתוצאה שלילית כוזבת), $\mathbb{P}(B\mid A^c)=p_{\mathcal{H}_0}(+)=\alpha$ (ההסתברות לתוצאה חיובית כוזבת) ו- $\mathbb{P}(A)=p(\mathcal{H}_1)=0.01$ (1% מהאוכלוסיה נשאים).

לפי נוסחת ההסתברות השלמה:

$$\mathbb{P}(B)=\mathbb{P}(B\mid A)\cdot\mathbb{P}(A)+\mathbb{P}(B\mid A^{c})\cdot\mathbb{P}(A^{c})=(1-\beta)\cdot 0.01+\alpha\cdot 0.99$$

כעת מכלל בייס:

$$\mathbb{P}(A\mid B)=\frac{\mathbb{P}(B\mid A)\cdot\mathbb{P}(A)}{\mathbb{P}(B)}=\frac{(1-\beta)\cdot 0.01}{(1-\beta)\cdot 0.01+\alpha\cdot 0.99}$$

עבור $\alpha=\beta=0$ נקבל $\mathbb{P}(A\mid B)=1$, כלומר אמינות אידיאלית. לעומת זאת עבור $\alpha=\beta=0.05$ נקבל $\mathbb{P}(A\mid B)\approx\frac{1}{6}$.

אי תלות

הגדרה: יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהיו A,B שני מאורעות בו. נאמר כי A ו- B הם בלתי תלויים (ב"ת) ונסמן $A\perp\!\!\!\perp B$ אם $\mathbb{P}(A\cap B)=\mathbb{P}(A)\cdot\mathbb{P}(B)$. אחרת, המאורעות יקראו תלויים.

אבחנה: עבור מאורעות בלתי תלויים A,B מתקיים כי $\mathbb{P}(B\mid A)=\mathbb{P}(B)$, $\mathbb{P}(A\mid B)=\mathbb{P}(A)$.

דוגמה: מטילים שתי קוביות הוגנות. נסמן ב A את המאורע שבקוביה הראשונה יצא 6, וב B את המאורע שבקוביה השנייה יצא 6. יש להראות ששני המאורעות בלתי תלוים.

מרחב ההסתברות הוא המרחב האחיד על $\Omega=[6]^2$, המאורעות $A=\{(6,\cdot)\}$ ו- $B=\{(\cdot,6)\}$. מצד אחד: $\mathbb{P}(A)=\mathbb{P}(B)=\frac{6}{36}=\frac{1}{6}$, מצד שני: $\mathbb{P}(A\cap B)=\mathbb{P}(\{(6,6)\})=\frac{1}{36}$. אכן מתקיים ש $\mathbb{P}(A\cap B)=\mathbb{P}(A)\cdot\mathbb{P}(B)$ לכן המאורעות בלתי תלויים.

דוגמה: מוטלת קובית 6 פעם אחת. מרחב ההסתברות הוא המרחב האחיד על $\Omega=[6]^2$. נגדיר את המאורעות $A=\{2,4,6\}, B=\{1,2\}, C=\{1,2,3\}$ ומתקיים כי:

$$\mathbb{P}(A)=\frac{1}{2},\mathbb{P}(B)=\frac{1}{3},\mathbb{P}(C)=\frac{1}{2}$$

$A\perp\!\!\!\perp B$: אכן $\mathbb{P}(A\cap B)=\mathbb{P}(\{2\})=\frac{1}{6}=\mathbb{P}(A)\cdot\mathbb{P}(B)$.

$A\not\perp\!\!\!\perp C$: $\mathbb{P}(A\cap C)=\frac{1}{6}\neq\mathbb{P}(A)\cdot\mathbb{P}(C)$. באופן דומה, $B\not\perp\!\!\!\perp C$.

טענה (תכונות של אי תלות): יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהיו $A,B\in\mathcal{F}$ מאורעות כך ש $\mathbb{P}(A),\mathbb{P}(B)>0$. אזי:

1. אם $A\cap B=\emptyset$ או $A\subseteq B\neq\Omega$ אז $A\not\perp\!\!\!\perp B$.
2. A בלתי תלוי ב $\emptyset$ וב Ω.
3. אם $A\perp\!\!\!\perp B$ אז $A^{c}\perp\!\!\!\perp B$.
4. אם $A\perp\!\!\!\perp A$ אז $\mathbb{P}(A)\in\{0,1\}$.

הוכחה:

1. אם $A\cap B=\emptyset$ אז $\mathbb{P}(A\cap B)=0\neq\mathbb{P}(A)\cdot\mathbb{P}(B)$ כי $\mathbb{P}(A),\mathbb{P}(B)>0$. בדומה עבור $A\subseteq B\neq\Omega$.
2. מיידי מהגדרת אי תלות.
3. נשים לב כי:

$$\mathbb{P}(A^{c}\cap B)=\mathbb{P}(B)-\mathbb{P}(A\cap B)=\mathbb{P}(B)-\mathbb{P}(A)\mathbb{P}(B)=\mathbb{P}(B)(1-\mathbb{P}(A))=\mathbb{P}(B)\cdot\mathbb{P}(A^{c})$$

4. מיידי מהגדרת אי תלות.

הגדרה: יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהיו $A,B_1,\ldots,B_k\in\mathcal{F}$ מאורעות בו. נאמר כי A בלתי תלוי באוסף $\{B_1,\ldots,B_k\}$ אם לכל $I\subseteq[k]$ מתקיים כי $A\perp\!\!\!\perp B_I$.

דוגמה: מטילים שתי קוביות הוגנות. נסמן ב A את המאורע שבקוביה הראשונה יצא 4, ב B_1 את המאורע שתוצאת הקוביה השניה היא זוגית וב B_2 את המאורע שתוצאת הקוביה השניה גדולה מ 3. יש להראות כי A בלתי תלוי ב $\{B_1,B_2\}$ אך B_1 תלוי ב B_2.

מרחב ההסתברות הוא מרחב ההסתברות האחיד על $\Omega=[6]^2$, אזי $A=\{4\}\times[6]$, $B_1=[6]\times\{2,4,6\}$, $B_2=[6]\times\{4,5,6\}$ וקל לראות כי הדרוש מתקיים.

הגדרה: יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות. מאורעות $A_1,\ldots,A_n\in\mathcal{F}$ יקראו בלתי תלויים אם לכל $I\subseteq[k]$ מתקיים כי:

$$\mathbb{P}\left(\bigcap_{i\in I}A_i\right)=\prod_{i\in I}\mathbb{P}\left(A_i\right)$$

טענה: יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהיו $A_1,\ldots,A_n\in\mathcal{F}$ מאורעות בו. מאורעות אלה בלתי תלויים $\iff$ לכל $j\in[n]$ מתקיים $A_j\perp\!\!\!\perp\{A_1,\ldots,A_{j-1},A_{j+1},\ldots,A_n\}$.

הוכחה: ($\Longleftarrow$): נניח כי $\{A_i\}_{i=1}^n$ בלתי תלויים. יהי $j\in[n]$ ותהי $I\subseteq[n]\setminus\{j\}$. לפי הגדרת אי תלות:

$$\mathbb{P}\left(\bigcap_{i\in I}A_i\cap A_j\right)=\prod_{i\in I}\mathbb{P}\left(A_i\right)\cdot\mathbb{P}\left(A_j\right)=\mathbb{P}\left(\bigcap_{i\in I}A_i\right)\cdot\mathbb{P}\left(A_j\right)$$

($\Longrightarrow$): נניח שלכל $j\in[n]$ ו- $I\subseteq[n]\setminus\{j\}$ מתקיים:

$$\mathbb{P}\left(\bigcap_{i\in I}A_i\cap A_j\right)=\mathbb{P}\left(\bigcap_{i\in I}A_i\right)\cdot\mathbb{P}\left(A_j\right)$$

כעת, תהי $J\subseteq[n]$ ונסמנה $J=\{j_1,\ldots,j_k\}$. נקבל:

$$\mathbb{P}\left(A_J\right)=\mathbb{P}\left(\bigcap_{m=1}^{k}A_{j_m}\right)=\mathbb{P}\left(A_{j_k}\right)\cdot\mathbb{P}\left(\bigcap_{m=1}^{k-1}A_{j_m}\right)=\cdots=\prod_{j\in J}\mathbb{P}\left(A_j\right)$$

הרצאה 7

אי תלות - המשך

הגדרה: יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. מאורעות $A_1, \dots, A_n$ יקראו בלתי תלויים בזוגות אם לכל $i \neq j$ מתקיים כי A_i בלתי תלוי ב A_j.

דוגמה: מוטלים שני מטבעות הוגנים, באופן בלתי תלוי, כלומר מרחב המדגם הוא המרחב האחיד $\Omega = \{0,1\}^2$. נגדיר את המאורעות הבאים: $A = \{(x,y) \in \Omega : x = y\}$ - בשתי ההטלות יצאה אותה תוצאה, $B_1 = \{(x,y) \in \Omega : x = 1\}$ - בהטלה הראשונה יצא עץ, $B_2 = \{(x,y) \in \Omega : y = 0\}$ - בהטלה השניה יצא פלי. נטען כי המאורעות הם בלתי תלויים בזוגות.

ראשית, נשים לב כי $\mathbb{P}(A) = \mathbb{P}(B_1) = \mathbb{P}(B_2) = \frac{1}{2}$. כעת:

$$\text{(I) } \mathbb{P}(A \cap B_1) = \mathbb{P}(\{(1,1)\}) = \frac{1}{4}$$

$$\text{(II) } \mathbb{P}(A \cap B_2) = \mathbb{P}(\{(1,1)\}) = \frac{1}{4}$$

$$\text{(III) } \mathbb{P}(A \cap B_1) = \mathbb{P}(\{(1,1)\}) = \frac{1}{4}$$

לכן המארועות בלתי תלויים בזוגות. אבל $\mathbb{P}(A \cap B_1 \cap B_2) = \mathbb{P}(\emptyset) = 0$ לכן הם לא בלתי תלויים.

הגדרה: אוסף מאורעות $\{A_i\}_{i \in J}$ נקרא בלתי תלוי אם המאורעות בכל תת אוסף סופי שלו הם בלתי תלויים.

טענה: תהי $\{A_n\}_{n=1}^{\infty}$ סדרה של מאורעות בלתי תלויים אזי:

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \prod_{n=1}^{\infty} \mathbb{P}(A_n)$$

הוכחה: נגדיר $B_m = \bigcap_{n=1}^{m} A_n$, אזי $\bigcap_{n=1}^{\infty} A_n = \bigcap_{m=1}^{\infty} B_m$ ולכל m מתקיים $B_{m+1} \subseteq B_m$. מרציפות פונקציית ההסתברות על סדרה יורדת של מאורעות:

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\bigcap_{m=1}^{\infty} B_m\right) = \lim_{m \to \infty} \mathbb{P}(B_m) = \lim_{m \to \infty} \prod_{n=1}^{m} \mathbb{P}(A_n) = \prod_{n=1}^{\infty} \mathbb{P}(A_n)$$

דוגמה: נסתכל על מרחב ההסתברות על $\mathbb{N}$ עם פונקצית ההסתברות הנקודתית $p(n) = \frac{6}{\pi^2} \cdot \frac{1}{n^2}$. לכל $k \in \mathbb{N}$, נגדיר את המאורע $A_k = k\mathbb{N}$.

א. הראו כי $\{A_p\}_{p \in P}$ הוא אוסף בלתי תלוי, כאשר P היא קבוצת כל הראשוניים.

לפי ההגדרה, צריך להראות שלכל קבוצת מספרים ראשוניים $p_1, \dots, p_k$ מתקיים $\mathbb{P}(A_{p_1} \cap \dots \cap A_{p_k}) = \prod_{j=1}^{k} \mathbb{P}(A_{p_j})$.

אכן, לכל p מתקיים כי:

$$\mathbb{P}(A_p) = \mathbb{P}(p \cdot \mathbb{N}) = \sum_{n=1}^{\infty} \frac{6}{\pi^2} \cdot \frac{1}{(pn)^2} = \frac{1}{p^2} \cdot \sum_{n=1}^{\infty} \frac{6}{\pi^2} \cdot \frac{1}{n^2} = \frac{1}{p^2}$$

עבור שני ראשוניים $p_1 \neq p_2$:

$$\mathbb{P}(A_{p_1} \cap A_{p_2}) = \mathbb{P}(A_{p_1 \cdot p_2}) = \frac{1}{(p_1 p_2)^2}$$

כי p_1 ו- p_2 זרים. כעת, עבור $p_1, \ldots, p_k$ ראשוניים:

$$\mathbb{P}\left(A_{p_1} \cap \cdots \cap A_{p_k}\right) = \frac{1}{\left(p_1 \cdots p_k\right)^2} = \prod_{j=1}^{k} \frac{1}{p_j^2} = \prod_{j=1}^{k} \mathbb{P}\left(A_{p_j}\right)$$

ב. הוכיחו כי $\frac{6}{\pi^2} = \prod_{p \in P}\left(1 - \frac{1}{p^2}\right)$.

בתרגיל הבית ראינו שאם $\{A_n\}$ אוסף בלתי תלוי, אז $\{A_n^c\}$ גם אוסף בלתי תלוי. מתקיים כי:

$$\frac{6}{\pi^2} = \mathbb{P}\left(\{1\}\right) = \mathbb{P}\left(\bigcap_{p \in P} A_p^c\right) = \prod_{p \in P}\left(1 - \frac{1}{p^2}\right)$$

בעיית מונטי הול

הבעיה קרויה על שם שעשועון הטלוויזיה האמריקאי המצליח "Let's Make a Deal" בהנחיית מונטי הול. בסוף התוכנית מציגים בפני השחקן שלוש דלתות. מאחורי אחת מהן מצויה מכונית, ומאחורי שתי הדלתות האחרות עזים. השחקן אינו יודע איזו דלת מסתירה את המכונית, ונאלץ לבחור באקראי. אם יבחר בדלת הנכונה - יזכה בה. אחרת, לא יקבל דבר.

לאחר שהשחקן הצביע על אחת הדלתות, המנחה, היודע מהי הדלת הנכונה, פותח אחת משתי הדלתות האחרות, ומגלה מאחוריה עז. כעת מאפשרים לשחקן להחליט האם לדבוק בבחירה המקורית שלו, או לשנות את ההחלטה ולהעדיף במקומה את הדלת האחרונה שנותרה סגורה. היכן סביר יותר למצוא את המכונית?

ננתח את ההסתברויות בהנחה שהמשתתף בחר את דלת מספר 1 (עבור דלתות 2 ו- 3 הניתוח זהה).

נסמן ב- C_i את המאורע שמהכונית נמצאת מאחורי הדלת ה i, אזי $\Omega = C_1 \uplus C_2 \uplus C_3$. מאחר שמיקום המכונית מוגרל באופן אחיד, מתקיים ש $\mathbb{P}\left(C_i\right) = \frac{1}{3}$.

נסמן ב- A את המאורע שהמשתתף מנצח לאחר שהוא <u>מחליף</u> את בחירתו בדלת האחרונה. כעת:

$$\mathbb{P}\left(A \mid C_1\right) = 0,\ \mathbb{P}\left(A \mid C_2\right) = 1,\ \mathbb{P}\left(A \mid C_3\right) = 1$$

מנוסחת ההסתברות השלמה:

$$\mathbb{P}\left(A\right) = \sum_{i=1}^{3} \mathbb{P}\left(A \mid C_i\right) \cdot \mathbb{P}\left(C_i\right) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}$$

כלומר, בכל מקרה משתלם למשתתף להחליף את בחירתו הראשונית.

משתנים מקריים

הגדרה: יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. משתנה מקרי (מ"מ) הוא פונקציה $X : \Omega \longrightarrow \mathbb{R}$. לכל $S \subseteq \mathbb{R}$ נגדיר: $X^{-1}\left(S\right) = \{X \in S\} := \{\omega \in \Omega \mid X\left(\omega\right) \in S\}$.

הערה: כפי שראינו, המשתנה המקרי נותן לנו להגדיר מאורעות במונחים של הערכים אותם הוא מקבל, ולשייך להם הסתברות. מאורעות אלו נכתבים בכתיב מקוצר. למשל:

$$\{X \notin S\} = \{\omega \in \Omega \mid X\left(\omega\right) \notin S\}$$

$$\{X \geq a\} = \{\omega \in \Omega \mid X\left(\omega\right) \geq a\}$$

דוגמה: מטילים קוביה הוגנת עשר פעמים, כלומר מרחב המדגם הוא המרחב האחיד $\Omega=[6]^{10}$. לכן, לכל $\omega\in\Omega$ נוכל לומר $\omega=(\omega_1,\ldots,\omega_{10})$. נגדיר את המשתנה המקרי $X(\omega)=\sum\limits_{i=1}^{10}\omega_i$. נרצה לחשב את ההסתברות שסכום ההטלות קטן מ 11 - כלומר $\mathbb{P}(\{X\leq 11\})$. לכך יש 11 אפשרויות - או שכל התוצאות הן 1, שכולן מלבד אחת הן 1 והאחרת היא 2. מאחר שהמרחב אחיד, נקבל כי: $\mathbb{P}(\{X\leq 11\})=\frac{11}{6^{10}}$.

הגדרה: יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהי מאורע $A\in\mathcal{F}$. המשתנה מקרי

$$\mathbb{1}_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$$

יקרא אינדיקטור של A או משתנה מקרי מציין של A.

הרצאה 8

משתנים מקריים - המשך

הגדרה: יהי X משתנה מקרי על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. התפלגותו של X היא הפונקציה $\mathbb{P}_X:\mathcal{F}_{\mathbb{R}}\longrightarrow[0,1]$ כאשר $\mathcal{F}_{\mathbb{R}}=2^{\mathbb{R}}$, הנתונה על ידי:

$$\mathbb{P}_X\left(S\right)=\mathbb{P}\left(X^{-1}\left(S\right)\right)=\mathbb{P}\left(\left\{\omega\in\Omega:X\left(\omega\right)\in S\right\}\right)$$

דוגמה: מטילים קוביה הוגנת שלושים פעמים, כלומר מרחב המדגם הוא המרחב האחיד $\Omega=[6]^{30}$. נגדיר את המשתנה המקרי:

$$X\left(\left(\omega_1,\ldots,\omega_{30}\right)\right)=\max_{1\leq i\leq 30}\omega_i$$

אזי:

$$\mathbb{P}_X\left(\left(-\infty,5\right]\right)=\mathbb{P}_X\left(\left[1,5\right]\right)=\mathbb{P}\left(X\leq 5\right)$$

המעבר הראשון נכון כי תוצאת קוביה היא מספר שלם בין 1 ל 6.

טענה: יהי X משתנה מקרי על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. אזי $\left(\mathbb{R},\mathcal{F}_{\mathbb{R}},\mathbb{P}_X\right)$ הוא מרחב הסתברות.

הוכחה: $\mathbb{P}_X\left(S\right)$ אי שלילי לכל $S\in\mathbb{R}$ ומתקיים כי $\mathbb{P}_X\left(\mathbb{R}\right)=\mathbb{P}\left(X^{-1}\left(\mathbb{R}\right)\right)=\mathbb{P}\left(\Omega\right)=1$. נותר לבדוק שמתקיימת סכימות בת מניה. תהי $\left\{S_k\right\}_{k=1}^{\infty}\subseteq\mathcal{F}_{\mathbb{R}}$ סדרה של קבוצות זרות. אזי:

$$\mathbb{P}\left(\dot{\bigcup_{k=1}^{\infty}}S_k\right)=\mathbb{P}\left(X^{-1}\left(\dot{\bigcup_{k=1}^{\infty}}S_k\right)\right)=\mathbb{P}\left(\dot{\bigcup_{k=1}^{\infty}}\left\{\omega\in\Omega:X\left(\omega\right)\in S_k\right\}\right)$$

$$=\sum_{k=1}^{\infty}\mathbb{P}\left(X^{-1}\left(S_k\right)\right)=\sum_{k=1}^{\infty}\mathbb{P}_X\left(S_k\right)$$

הגדרה: משתנה מקרי X יקרא בדיד, והתפלגותו תקרא בדידה אם $\mathbb{P}_X$ היא פונקציית הסתברות בדידה. כלומר, קיימת פונקציית הסתברות נקודתית $p_X:\mathbb{R}\longrightarrow[0,1]$ כך ש $\sum_{x\in\mathbb{R}}p_X\left(x\right)=1$ ולכל $S\subseteq\mathbb{R}$ מתקיים $\mathbb{P}_X\left(S\right)=\sum_{x\in S}p_X\left(x\right)$.

דוגמה: מוטל מטבע מוטה שעל צדדיו המספרים 0 ו- 1 פעם אחת, כלומר מרחב המדגם הוא $\Omega=\{0,1\}$, ההסתברות שיצא 0 היא p וההסתברות שיצא 1 היא $1-p$. נגדיר את המשתנה המקרי $X\left(\omega\right)=\omega$. נשים לב כי X הוא משתנה מקרי בדיד, עבור

$$p_X\left(x\right)=\begin{cases}1-p & x=1\\ p & x=0\\ 0 & \text{else}\end{cases}$$

הגדרה: נאמר כי משתנה מקרי X מתפלג ברנולי עם פרמטר p ונכתוב $X\sim\text{Ber}\left(p\right)$ אם $\mathbb{P}_X\left(\{0\}\right)=p$ ו- $\mathbb{P}_X\left(\{1\}\right)=1-p$.

אבחנה: יהי $A\in\mathcal{F}$ מאורע במרחב $(\Omega,\mathcal{F},\mathbb{P})$ אז $\mathbb{1}_A\sim\text{Ber}\left(\mathbb{P}\left(A\right)\right)$.

הגדרה: נאמר שמשתנה מקרי X מתפלג באופן אחיד על קבוצה סופית $S\subseteq\mathbb{R}$ ונכתוב $X\sim\text{Unif}\left(S\right)$ אם $p_X\left(i\right)=\frac{1}{|S|}$ לכל $i\in S$.

הגדרה: נאמר ש משתנה מקרי X הוא משתנה מקרי קבוע אם קיים $c\in\mathbb{R}$ כך ש $\mathbb{P}\left(X=c\right)=1$.

פעולות על משתנים מקריים

הערה: לכל פונקציה $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ ולכל אוסף משתנים מקריים $X_1,\ldots,X_n$ על אותו מרחב הסתברות, נוכל להגדיר משתנה מקרי $Y=f\left(X_1,\ldots,X_n\right)$ כלומר משתנה מקרי הנתון ע"י:

$$Y\left(\omega\right)=f\left(X_1\left(\omega\right),\ldots,X_n\left(\omega\right)\right)$$

כך למשל, מתוך משתנים מקריים X,Y,Z נוכל להגדיר את $\left(XY\right)^Z,Ye^{X+Z},\sqrt[X]{Y+Z}$ וכו'.

דוגמה: מוטל מטבע מוטה שנופל על עץ בסיכוי $\frac{2}{3}$, 30 פעמים. נרצה לספור כמה מופעים של עץ היו. נסמן את מרחב המדגם $\Omega=\left\{0,1\right\}^{30}$, ונגדיר לכל $\omega=\left(\omega_1,\ldots,\omega_{30}\right)\in\Omega$ נגדיר $X_i\left(\omega\right)=\omega_i$. כעת, המ"מ הרצוי הוא $X=\sum\limits_{i=1}^{30}X_i$.

סוגי שיוויון בין משתנים מקריים

הגדרה: יהיו X,Y שני משתנים מקריים בדידים על מרחב הסתברות $\left(\Omega,\mathcal{F},\mathbb{P}\right)$. נאמר כי X ו- Y <u>שווים כמעט תמיד</u>, ונסמן $X\overset{\text{a.s}}{=}Y$ אם מתקיים:

$$\mathbb{P}\left(X=Y\right)=\mathbb{P}\left(\left\{\omega\in\Omega:X\left(\omega\right)=Y\left(\omega\right)\right\}\right)=1$$

הגדרה: אם לשני משתנים מקריים שונים X ו- Y (שעשויים להיות מוגדרים על שני מרחבי הסתברות שונים) יש אותה התפלגות (כלומר מתקיים $\mathbb{P}_X\equiv\mathbb{P}_Y$), נאמר כי הם <u>שווי התפלגות</u> ונכתוב $X\overset{\text{d}}{=}Y$.

דוגמאות:

1. קוביה הוגנת מוטלת פעמיים באופן בלתי תלוי. נסמן את מרחב המדגם $\Omega=\left[6\right]^2$ ונגדיר שני משתניים מקריים על המרחב: $X_1\left(\left(\omega_1,\omega_2\right)\right)=\omega_1$ ו- $X_2\left(\left(\omega_1,\omega_2\right)\right)=\omega_2$. אזי, $\mathbb{P}_{X_1}\equiv\mathbb{P}_{X_2}$ ולכן $X_1\overset{d}{=}X_2$ אבל $\mathbb{P}\left(\left\{\omega\in\Omega:X_1\left(\omega\right)=X_2\left(\omega\right)\right\}\right)\neq1$ ולכן $X_1\overset{\text{a.s}}{\neq}X_2$.
2. עבור אותם משתנים מקריים כמקודם, $X_1-1\overset{\text{d}}{=}\left(X_1+X_2\right)\left(\mod 6\right)$ (תרגיל).

טענה: יהיו X,Y משתנים מקריים על מרחב הסתברות $\left(\Omega,\mathcal{F},\mathbb{P}\right)$. אם $X\overset{\text{a.s}}{=}Y$ אז $X\overset{\text{d}}{=}Y$.

הוכחה: תהי $S\subseteq\mathbb{R}$, צריך להוכיח כי $\mathbb{P}_X\left(S\right)=\mathbb{P}_Y\left(S\right)$. מנוסחת ההסתברות השלמה:

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

נשים לב כי $\mathbb{P}\left(X\in S,Y\notin S\right)=\mathbb{P}\left(X\notin S,Y\in S\right)=0$ כי המשתנים שווים כמעט תמיד. נוכל להציב זאת ולהעזר שוב בנוסחת ההסתברות השלמה:

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

טענה (תרגיל): יהיו X,Y משתנים מקריים בדידים, ותהי $f:\mathbb{R}\longrightarrow\mathbb{R}$. אזי:

1. אם $X\overset{\text{a.s}}{=}Y$ אז $f\left(X\right)\overset{\text{a.s}}{=}f\left(Y\right)$.
2. אם $X\overset{\text{d}}{=}Y$ אז $f\left(X\right)\overset{\text{d}}{=}f\left(Y\right)$.
3. אם $X\overset{d}{=}X^2$ אז X משתנה מקרי מציין.

הרצאה 9

וקטור מקרי והתפלגות משותפת

דוגמה: נתונים שני משתנים מקריים X,Y המתפלגים ברנולי $\frac{1}{2}$. האם הנתון מספיק כדי לקבוע את $\mathbb{P}\left(X=Y\right)$?

התשובה - לא ניתן. נבחר למשל את מרחב ההסתברות האחיד $\Omega=\left\{0,1\right\}^{2}$ ונגדיר $X\left(\left(\omega_1,\omega_2\right)\right)=\omega_1$, $Y\left(\left(\omega_1,\omega_2\right)\right)=\omega_2$. אכן $X,Y\sim\text{Ber}\left(\frac{1}{2}\right)$ ו- $\mathbb{P}\left(X=Y\right)=\frac{2}{4}=\frac{1}{2}$. אבל עבור $\tilde{X}\left(\left(\omega_1,\omega_2\right)\right)=\tilde{Y}\left(\left(\omega_1,\omega_2\right)\right)=\omega_1$ מתקיים כי $\mathbb{P}\left(\tilde{X}=\tilde{Y}\right)=1$, ואכן $\tilde{X},\tilde{Y}\sim\text{Ber}\left(\frac{1}{2}\right)$.

הגדרה: אוסף סופי של משתנים מקריים $X=\left(X_1,\ldots,X_n\right)$ המוגדרים על אותו מרחב הסתברות $\left(\Omega,\mathcal{F},\mathbb{P}\right)$ יכונה וקטור מקרי.

הגדרה: יהי $X=\left(X_1,\ldots,X_n\right)$ וקטור מקרי על מרחב הסתברות $\left(\Omega,\mathcal{F},\mathbb{P}\right)$. הפונקציה $\mathbb{P}_X=\mathbb{P}_{X_1,\ldots,X_n}:\mathcal{F}_{\mathbb{R}^n}\longrightarrow\left[0,1\right]$ הנתונה על ידי:

$$\mathbb{P}_X\left(S\right)=\mathbb{P}\left(\left(X_1,\ldots,X_n\right)\in S\right)=\mathbb{P}\left(\left\{\omega\in\Omega:X\left(\omega\right)\in S\right\}\right)$$

מכונה ההתפלגות המשותפת של $X=\left(X_1,\ldots,X_n\right)$. ההתפלגות של כל אחד מהמשתנים המקריים $X_1,\ldots,X_n$ מכונה התפלגות שולית.

אבחנה: יהי $\left(X_1,\ldots,X_n\right)$ וקטור מקרי, אזי $\mathbb{P}_{X_1,\ldots,X_n}$ היא פונקציית הסתברות על $\mathbb{R}^n$.

הגדרה: וקטור מקרי $X=\left(X_1,\ldots,X_n\right)$ יקרא בדיד, והתפלגותו תקרא בדידה אם $\mathbb{P}_X$ היא פונקציית הסתברות בדידה. פונקצית ההסתברות הנקודתית p_X המתאימה ל $\mathbb{P}_X$ תכונה פונקציית ההתפלגות הנקודתית של X.

עבור קונפיגורציה $x=\left(x_1,\ldots,x_n\right)$ מסויימת, נסמן:

$$p_{X_1,\ldots,X_n}\left(x_1,\ldots,x_n\right)=p_X\left(\left(x_1,\ldots,x_n\right)\right)=p_X\left(x\right)=\mathbb{P}\left(X_1=x_1,\ldots,X_n=x_n\right)$$

דוגמה: עבור מרחב המדגם $\Omega=\left\{0,1\right\}$, לכל $a\in\left[0,\frac{1}{2}\right]$ נגדיר התפלגות משותפת של שני משתנים מקריים X,Y באמצעות פונקציית התפלגות נקודתית הנתונה על ידי:

$$p_{X,Y}\left(\left(0,0\right)\right)=a,\ p_{X,Y}\left(\left(0,1\right)\right)=\frac{1}{2}-a,\ p_{X,Y}\left(\left(1,0\right)\right)=\frac{1}{2}-a,\ p_{X,Y}\left(\left(1,1\right)\right)=a$$

נראה כי לכל בחירה של a, התפלגויותיהם השוליות של X ו- Y הן התפלגות ברנולי $\frac{1}{2}$. עבור X:

$$\mathbb{P}\left(X=0\right)=\mathbb{P}\left(X=0,Y\in\mathbb{R}\right)=\sum_{y\in\mathbb{R}}p_{X,Y}\left(\left(0,y\right)\right)=p_{X,Y}\left(0,0\right)+p_{X,Y}\left(0,1\right)=\frac{1}{2}$$

$$\mathbb{P}\left(X=1\right)=\mathbb{P}\left(X=1,Y\in\mathbb{R}\right)=\sum_{y\in\mathbb{R}}p_{X,Y}\left(\left(1,y\right)\right)=p_{X,Y}\left(1,0\right)+p_{X,Y}\left(1,1\right)=\frac{1}{2}$$

ונקבל שאכן $X\sim\text{Ber}\left(\frac{1}{2}\right)$. באופן דומה גם Y.

דוגמה: מוטלות שתי קוביות משחק סטנדרטיות. נסמן ב X_1, X_2 את תוצאות הקוביה. כמו כן, נסמן $Y = \min(X_1, Y_2)$ ו- $Z = \max(X_1, X_2)$. נרצה לחשב את ההתפלגות המשותפת של Y, Z והתפלגותם השולית.

מהנתון, $X_1, Y_1 \sim \text{Unif}\left([6]^2\right)$. נשים לב שלכל $y < z \in [6]$ מתקיים:

$$\mathbb{P}(Y = y, Z = z) = \mathbb{P}_{X_1,X_2}(\{(y,z),(z,y)\}) = \mathbb{P}(X_1 = y, X_2 = z) + \mathbb{P}(X_1 = z, X_2 = y) = \frac{2}{36}$$

לעומת זאת:

$$\mathbb{P}(Y = y, Z = y) = \mathbb{P}(X_1 = y, X_2 = y) = \frac{1}{36}$$

ועבור כל צירוף אחר ההסתברות היא 0. אם כן:

$Y \backslash Z$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$
6	0	0	0	0	0	$\frac{1}{36}$

הגדרה: נאמר שוקטור מקרי $X = (X_1, \ldots, X_n)$ מתפלג אחיד על קבוצה סופית $S \subseteq \mathbb{R}^n$ ונכתוב $X \sim \text{Unif}(S)$ אם לכל $a \in S$ מתקיים $p_X(a) = \frac{1}{|a|}$.

הגדרה: יהי X וקטור מקרי בדיד ממימד d על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ ויהי $A \in \mathcal{F}$ מאורע המקיים $\mathbb{P}(A) > 0$. ההתפלגות המותנית של X ב- A היא:

$$\mathbb{P}_{X|A}(S) = \mathbb{P}(X \in S \mid A) = \mathbb{P}_A(X \in S)$$

למעשה התפלגותו של X על מרחב ההסתברות $(\Omega, \mathcal{F}, \mathbb{P}_A)$. כמו כן, נסמן $(X \mid A) \sim \mathcal{D}$ כדי לציין שבהינתן A, המשתנה X מתפלג לפי $\mathcal{D}$.

דוגמה: יהי $X \sim \text{Unif}(\{0,1,2\})$ ונסמן $S = \{0,1\}$. נראה כי $(X \mid X \in S) \sim \text{Ber}\left(\frac{1}{2}\right)$. אכן:

$$\mathbb{P}(X = 0 \mid X \in S) = \frac{\mathbb{P}(X = 0, X \in S)}{\mathbb{P}(X \in S)} = \frac{p_X(0)}{p_X(0) + p_X(1)} = \frac{1/3}{1/3 + 1/3} = \frac{1}{2}$$

ובאופן דומה, $\mathbb{P}(X = 1 \mid X \in S) = \frac{1}{2}$. משתנה שמקיים הסתברויות כאלה הוא משתנה ברנולי חצי.

אי תלות של משתנים מקריים ווקטורים מקריים

הגדרה: נאמר שמשתנים מקריים X ו- Y המוגדרים על אותו מרחב הסתברות הנם בלתי-תלויים, אם לכל שתי קבוצות $S, T \subseteq \mathbb{R}$ מתקיים שהמאורעות $\{X \in A\}$ ו- $\{Y \in B\}$ בלתי תלויים, כלומר:

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)$$

בניסוח אחר, X ו- Y בלתי תלויים אם לכל $S \subseteq \mathbb{R}$ עבורה $\mathbb{P}(X \in S) > 0$ מתקיים $\mathbb{P}_{Y|X\in S} = \mathbb{P}_Y$.

הרצאה 10

המשך אי תלות של משתנים מקריים ווקטורים מקריים

טענה: יהיו X,Y משתנים מקריים בדידים. X ו- Y בלתי תלויים $\iff$ לכל $x,y\in\mathbb{R}$ המאורעות $\{X=x\}$ ו- $\{Y=y\}$ בלתי תלויים. כלומר:

$$\mathbb{P}\left(X=x,Y=y\right)=\mathbb{P}\left(X=x\right)\cdot\mathbb{P}\left(Y=y\right)$$

הוכחה: ($\Leftarrow$): בהינתן $x,y\in\mathbb{R}$, נבחר את המאורעות $A=\{x\}$ ו- $B=\{y\}$ ונקבל:

$$\mathbb{P}\left(X=x,Y=y\right)=\mathbb{P}\left(X\in A,Y\in B\right)=\mathbb{P}\left(X\in A\right)\mathbb{P}\left(Y\in B\right)=\mathbb{P}\left(X=x\right)\mathbb{P}\left(Y=y\right)$$

($\Longrightarrow$): נניח ש $\{X=x\}$ ו- $\{Y=y\}$ בלתי תלויים לכל $x,y\in\mathbb{R}$. מאחר ש X ו- Y בדידים, גם ההתפלגות המשותפת שלהם בדידה, ומאי התלות: $p_{X,Y}=p_X\cdot p_Y$. כעת:

$$\mathbb{P}_{X,Y}\left(A\times B\right)=\sum_{(x,y)\in A\times B}p_{X,Y}\left(x,y\right)=\sum_{x\in A}\sum_{y\in B}p_X\left(x\right)\cdot p_Y\left(y\right)=\left(\sum_{x\in A}p_X\left(x\right)\right)\left(\sum_{y\in B}p_Y\left(y\right)\right)=\mathbb{P}\left(X\in A\right)\mathbb{P}\left(Y\in B\right)$$

כנדרש.

אבחנה: יהיו X,Y משתנים מקריים בלתי תלויים ובדידים. אזי לכל $x\in\mathbb{R}$ עבורו $\mathbb{P}\left(X=x\right)>0$ מתקיים:

$$\left(Y\mid X=x\right)\overset{\mathrm{d}}{=}\left(Y\right)$$

אכן, לכל $y\in\mathbb{R}$:

$$\mathbb{P}\left(Y=y\mid X=x\right)=\frac{\mathbb{P}\left(Y=y,X=x\right)}{\mathbb{P}\left(X=x\right)}=\frac{\mathbb{P}\left(Y=y\right)\cdot\mathbb{P}\left(X=x\right)}{\mathbb{P}\left(X=x\right)}=\mathbb{P}\left(Y=y\right)$$

דוגמה: יהיו $X\sim\text{Ber}\left(p\right)$, $Y\sim\text{Ber}\left(q\right)$ בלתי תלויים. נרצה להראות כי $XY\sim\text{Ber}\left(pq\right)$. מהגדרה, מתקיים כי $X,Y:\Omega\longrightarrow\{0,1\}$ ולכן גם $XY:\Omega\longrightarrow\{0,1\}$. כעת:

$$\mathbb{P}\left(XY=1\right)=\mathbb{P}\left(X=1,Y=1\right)=\mathbb{P}\left(X=1\right)\mathbb{P}\left(Y=1\right)=pq$$

$$\mathbb{P}\left(XY=0\right)=\mathbb{P}\left(X=0,Y=1\right)+\mathbb{P}\left(X=1,Y=0\right)+\mathbb{P}\left(X=0,Y=0\right)=1-\mathbb{P}\left(X=1,y=1\right)=1-pq$$

לכן $XY\sim\text{Ber}\left(pq\right)$, כנדרש.

הגדרה: וקטורים מקריים $X=\left(X_1,\ldots,X_n\right)$, $Y=\left(Y_1,\ldots,Y_m\right)$ המוגדרים על אותו מרחב הסתברות יקראו בלתי תלויים אם לכל $A\subseteq\mathbb{R}^n$ ו- $B\subseteq\mathbb{R}^m$ מתקיים כי $\{X\in A\}$ ו- $\{Y\in B\}$ בלתי תלויים, כלומר:

$$\mathbb{P}\left(X\in A,Y\in B\right)=\mathbb{P}\left(X\in A\right)\cdot\mathbb{P}\left(Y\in B\right)$$

אם X,Y בדידים, אז התנאי שקול ל:

$$\forall x\in\mathbb{R}^n,y\in\mathbb{R}^m:\mathbb{P}\left(X=x,Y=y\right)=\mathbb{P}\left(X=x\right)\cdot\mathbb{P}\left(Y=y\right)$$

טענה (אדישות ההתפלגות להתניה שקולה לאי תלות): יהיו X, Y, Z וקטורים מקריים בדידים ונניח כי לכל $x \in \mathbb{R}$ עבורו $\mathbb{P}(X = x) > 0$ מתקיים:

$$(Y \mid X = x) \overset{\mathrm{d}}{=} (Z)$$

אזי X ו- Y בלתי תלויים וכן $Y \overset{\mathrm{d}}{=} Z$.

הוכחה: לפי נוסחת ההסתברות השלמה:

$$\mathbb{P}(Y = y) = \sum_{x \in \mathrm{Sprt}(X)} \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y \mid X = x) = \sum_{x \in \mathrm{Sprt}(X)} \mathbb{P}(X = x) \cdot \mathbb{P}(Z = y)$$

$$= \mathbb{P}(Z = y) \cdot \sum_{x \in \mathrm{Sprt}(X)} \mathbb{P}(X = x) = \mathbb{P}(Z = y)$$

ולכן $Y \overset{\mathrm{d}}{=} Z$. כעת:

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y \mid X = x) = \mathbb{P}(X = x) \cdot \mathbb{P}(Z = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

לכן הוקטורים המקריים בלתי תלויים.

הגדרה: יהיו $X_1, \ldots, X_n$ משתנים מקריים המוגדרים על אותו מרחב הסתברות. נאמר שזהו אוסף משתנים מקריים בלתי תלוי אם לכל $E_1, \ldots, E_n \subseteq \mathbb{R}$ המאורעות $\{X_i \in E_i\}$ בלתי תלויים, כלומר:

$$\mathbb{P}(X_1 \in E_1, \ldots, X_n \in E_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in E_i)$$

הגדרה: יהיו $X_1, \ldots, X_n$ וקטורים מקריים d-מימדיים המוגדרים על אותו מרחב הסתברות, כאשר $X_i = (X_{i,1}, \ldots, X_{i,d})$. נאמר שזהו אוסף וקטורים מקריים בלתי תלויים אם לכל $E_1, \ldots, E_n \subseteq \mathbb{R}^d$ המאורעות $\{X_i \in E_i\}$ בלתי תלויים.

טענה (תרגיל): יהיו $X_1, \ldots, X_n$ משתנים מקריים בלתי תלויים, יהיו מספרים שלמים $\{b_i\}_{i \in [n] \cup \{0\}}$ המקיימים $0 = b_0 < b_1 < \cdots < b_{k-1} < b_k = n$ ונגדיר k וקטורים מקריים $Y_i = \left(X_{b_{i-1}+1}, \ldots, X_{b_i}\right)$ אז $Y_1, \ldots, Y_k$ בלתי תלויים.

הגדרה: סדרה של משתנים מקריים $\{X_n\}_{n=1}^{\infty}$ נקראת סדרת משתנים בלתי תלויים אם לכל $N \in \mathbb{N}$ המשתנים $X_1, \ldots, X_N$ בלתי תלויים.

אבחנה: תהי $\{X_n\}_{n=1}^{\infty}$ סדרה של משתנים מקריים בלתי תלויים ותהיינה $\{E_n\}_{n=1}^{\infty}$ סדרה של קבוצות $E_i \subseteq \mathbb{R}$. אזי:

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \{X_n \in E_n\}\right) = \prod_{n=1}^{\infty} \mathbb{P}(X_n \in E_n)$$

התפלגות גיאומטרית

לכאורה, ניתן להמציא מספר עצום של ניסויים מקריים כדי לסייע למדל את היבטיה המקריים של המציאות. אולם, ברבים משימושיה של תורת ההסתברות, אנו נתקלים במספר מצומצם של התפלגויות. אנו נקדיש את מירב תשומת ליבנו לחמש משפחות של התפלגויות כאלה. הראשונות - התפלגות ברנולי והתפלגות אחידה, אותן כבר ראינו. כעת נוכל להציג את שלוש ההתפלגויות הנוספות.

חשוב לציין כי ניתן לתאר התפלגות בדידה על ידי תרשים עמודות (היסטוגרמה). הציר האופקי של התרשים יתאר את ערכי המשתנה המקרי והציר האנכי - את ההסתברות הנקודתית לקבלתו של כל ערך.

הגדרה: יהי $p \in (0,1]$. נאמר שמשתנה מקרי X מתפלג לפי התפלגות גיאומטרית עם הסתברות p להצלחה, ונכתוב $X \sim \text{Geo}(p)$ אם לכל $n \in \mathbb{N}$ מתקיים $p_X(n) = (1-p)^{n-1} p$.

טענה: תהי $\{X_k\}_{k=1}^{\infty}$ סדרה אינסופית של משתנים מקריים בלתי תלויים כאשר $X_k \sim \text{Ber}(p)$, ונסמן $X = \min(\{k : X_k = 1\})$. אזי $X \sim \text{Ber}(p)$.

הוכחה: נשים לב כי $\{X = k\} = \{X_1 = 0, \dots, X_{k-1} = 0, X_k = 1\}$ ונחשב באמצעות האי תלות בין המשתנים:

$$\mathbb{P}(X = k) = \mathbb{P}(X_1 = 0, \dots, X_{k-1} = 0, X_k = 1) = \overbrace{(1-p) \cdot \dots \cdot (1-p)}^{k-1 \text{ times}} \cdot p = (1-p)^{k-1} \cdot p$$

כלומר $X \sim \text{Geo}(p)$ כנדרש. כמו כן, נשים לב כי לפי האבחנה האחרונה:

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

הרצאה 11

המשך התפלגות גיאומטרית

טענה: משתנה מקרי שנתמך על הטבעיים מתפלג $\mathrm{Geo}(p)$ $\iff$ לכל $n \in \mathbb{N}$ מתקיים $\mathbb{P}(X > n) = (1-p)^n$.

הוכחה: ($\Longleftarrow$): אם X מתפלג גיאומטרית, הרי שמתקיים:=

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

$$= (1-p)^n \cancel{p} \cdot \frac{1}{\cancel{1-(1-p)}} = (1-p)^n$$

($\Longrightarrow$): אם $\mathbb{P}(X > n) = (1-p)^n$ ו- X נתמך על $\mathbb{N}$, הרי שלכל $n \in \mathbb{N}$:

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

משפט (תכונת חוסר זכרון): יהי X משתנה מקרי הנתמך על $\mathbb{N}$ אשר מקיים $\mathbb{P}(X = 1) < 1$. שלושת הבאים שקולים:

1. X מתפלג גיאומטרית.
2. $X \overset{\mathrm{d}}{=} (X - 1 \mid X > 1)$.
3. לכל $m \in \mathbb{N}$, $X \overset{\mathrm{d}}{=} (X - m \mid X > m)$.

הוכחה: $(1 \Longleftarrow 2)$: לכל $n \in \mathbb{N}$ מתקיים:

$$\mathbb{P}(X - 1 = n \mid X > 1) = \frac{\mathbb{P}(X = n+1)}{\mathbb{P}(X > 1)} \overset{(1)}{=} \frac{(1-p)^n \cdot p}{1-p} = (1-p)^{n-1} \cdot p = \mathbb{P}(X = n)$$

$(2 \Longleftarrow 3)$: נשאר כתרגיל. $(3 \Longleftarrow 1)$: יהי X משתנה מקרי הנתמך על $\mathbb{N}$ ומקיים את תכונה 3. ניעזר בכלל השרשרת, ונקבל שלכל $k \in \mathbb{N}_0$:

$$\mathbb{P}(X > k) = \mathbb{P}(X > m \mid X > m-1) \cdot \mathbb{P}(X > m-1 \mid X > m-2) \cdots \mathbb{P}(X > 1) \overset{(3)}{=} \mathbb{P}(X > 1)^m$$

לכן מהטענה האחרונה, X מתפלג גיאומטרית.

תרגיל:

1. יהיו $X_1 \sim \mathrm{Geo}(p)$, $X_2 \sim \mathrm{Geo}(q)$ בלתי תלויים. אזי $\min(X_1, X_2) \sim \mathrm{Geo}(1 - (1-p)(1-q))$.
2. יהיו $X_1, X_2 \sim \mathrm{Geo}(p)$ בלתי תלויים. אזי $(X \mid X + Y = n) \sim \mathrm{Unif}([n-1])$.

התפלגות בינומית

הגדרה: נאמר שמשתנה מקרי X מתפלג לפי התפלגות בינומית עם n ניסיונות והסתברות הצלחה p ונכתוב $X\sim\text{Bin}\left(n,p\right)$ אם לכל $k\in\{0,\ldots,n\}$ מתקיים $p_{X}\left(k\right)=\binom{n}{k}p^{k}\left(1-p\right)^{n-k}$.

אבחנה: אכן התפלגות בינומית הנה התפלגות מוגדרת היטב:

$$\mathbb{P}\left(X\in\{0,\ldots,n\}\right)=\sum_{k=0}^{n}\binom{n}{k}p^{k}\left(1-p\right)^{n-k}=\left(p+\left(1-p\right)\right)^{n}=1$$

טענה: יהיו $X=\{X_{i}\}_{i=1}^{n}$ וקטור של משתני ברנולי בלתי תלויים עם הסתברות הצלחה p. אזי: $\sum\limits_{i=1}^{n}X_{i}\sim\text{Bin}\left(n,p\right)$. כלומר, $\text{Bin}\left(n,p\right)$ מייצג את מספר ההצלחות בסדרה של n ניסוים בלתי תלויים עם סיכוי הצלחה p.

הוכחה: יהי $k\in\{0,\ldots,n\}$ ונסמן $Y=\sum\limits_{i=1}^{n}X_{i}$. נסמן ב A_{k} את אוסף הוקטורים ב - $\{0,1\}^{n}$ בהם בדיוק k אחדות ו - $\left(n-k\right)$ אפסים. מפורשות:

$$A_{k}=\left\{ \left(x_{1},\ldots,x_{n}\right)\in\{0,1\}^{n}:\sum x_{i}=k\right\}$$

מטעמי קומבינטוריקה, מתקיים כי $\left|A_{k}\right|=\binom{n}{k}$. נחשב:

$$\mathbb{P}\left(Y=k\right)=\sum_{x\in A_{k}}\mathbb{P}\left(X=x\right)=\sum_{x\in A_{k}}\left(\prod_{i=1}^{n}\mathbb{P}\left(X_{i}=x_{i}\right)\right)=\binom{n}{k}p^{k}\left(1-p\right)^{n-k}$$

כאשר המעבר השני נכון מאי תלות.

תרגיל: אם $X\sim\text{Bin}\left(m,p\right)$ ו- $Y\sim\text{Bin}\left(n,p\right)$ בלתי תלויים, אז $X+Y\sim\text{Bin}\left(m+n,p\right)$.

הערה (נוסחת סטרלינג): המתמטיקאי הצרפתי אברהם דה-מואבר הוכיח לראשונה את האומדן המדוייק הבא לקצב הגידול של פונקציית העצרת, אשר זכה, שלא בצדק, להיקרא על שם המתמטיקאי הסקוטי ג'יימס סטרלינג:

$$n!\approx\sqrt{2\pi n}\cdot\left(\frac{n}{e}\right)^{n}$$

דוגמה: מוטלים $2n$ מטבעות הוגנים באופן בלתי תלוי. נרצה להעריך את ההסתברות שבדיוק ב n מטבעות תתקבל תוצאה של עץ.

לפי הטענה הקודמת, מספר המטבעות שתוצאת הטלתם היא עץ מתפלגת בינומית עם מספר ניסויים $2n$ והסתברות הצלחה $\frac{1}{2}$. על כן, ההסתברות לשיוויון היא בדיוק $\binom{2n}{n}2^{-2n}$. נשתמש בנוסחת סטרלינג ונקבל:

$$\binom{2n}{n}2^{-2n}=\frac{\left(2n\right)!}{2^{2n}\cdot\left(n!\right)^{2}}\approx\frac{\sqrt{4\pi n}\cdot\left(\frac{2n}{e}\right)^{2n}}{2^{2n}\cdot2\pi n\cdot\left(\frac{n}{e}\right)^{2n}}=\frac{1}{\sqrt{2\pi n}}\cdot\frac{2\cdot2^{2n}\cdot\left(\frac{n}{e}\right)^{2n}}{2^{2n}\cdot2\cdot\left(\frac{n}{e}\right)^{2n}}=\frac{1}{\sqrt{2\pi n}}$$

הרצאה 12

התפלגות פואסונית

הגדרה: נאמר שמשתנה מקרי X מתפלג לפי התפלגות פואסון (או פואסונית) עם שכיחות $\lambda \geq 0$ ונכתוב $X \sim \text{Poi}(\lambda)$ אם לכל $n \in \mathbb{N}_0$ מתקיים:

$$\mathbb{P}(X = n) = \frac{e^{-\lambda}\lambda^n}{n!}$$

המשמעות של התפלגות זו היא ש X סופר את כמות המאורעות שהתרחשו ביחידת זמן, אם בממוצע מתרחשים λ מאורעות ביחידת זמן.

אבחנה: התפלגות זו מוגדרת היטב:

$$\sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda} \cdot \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

משפט: יהי $Y \sim \text{Poi}(\lambda)$ עבור $\lambda \geq 0$, ויהיו $\{X_n\}_{n=1}^{\infty}$ משתנים מקריים כך שלכל $n > \lambda$ מתקיים $X_n \sim \text{Bin}\left(n, \frac{\lambda}{n}\right)$. אזי $X_n \xrightarrow[n\to\infty]{d} Y$ במובן שלכל $k \in \mathbb{N}_0$ מתקיים $\mathbb{P}(X_n = k) \xrightarrow[n\to\infty]{} \mathbb{P}(Y = k)$.

הוכחה: יהי $k \in \mathbb{N}_0$, נחשב:

$$\mathbb{P}(X_n = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} \xrightarrow[n\to\infty]{} \frac{n(n-1)\cdots(n-k)}{k!} \cdot \frac{\lambda^k}{n^k} \cdot e^{-\lambda}$$

$$= \frac{n^k(1 + o(1))}{k!} \cdot \frac{\lambda^k}{n^k} \cdot e^{-\lambda} = \frac{e^{-\lambda}\lambda^k}{k!} \cdot \frac{n^k(1 + o(1))}{n^k} \xrightarrow[n\to\infty]{} \frac{e^{-\lambda}\lambda^k}{k!} = \mathbb{P}(Y = k)$$

טענה (תכונות התפלגות פואוסונית): יהי $X \sim \text{Poi}(\lambda)$.

1. אם $Y \sim \text{Poi}(\mu)$ בלתי תלוי ב X אז $X + Y \sim \text{Poi}(\lambda + \mu)$.
2. אם לכל $n \in \mathbb{N}_0$ מתקיים $(Y \mid X = n) \sim \text{Bin}(n, p)$ אז $Y \sim \text{Poi}(\lambda p)$.

הוכחה:

1. נחשב לפי נוסחת ההסתברות השלמה עבור $n \in \mathbb{N}_0$:

$$\mathbb{P}(X + Y = n) = \sum_{i \in \mathbb{Z}} \mathbb{P}(X = i)\,\mathbb{P}(Y = n - i) = \sum_{i=0}^{n} \frac{e^{-\lambda}\lambda^i}{i!} \cdot \frac{e^{-\mu}\mu^{n-i}}{(n-i)!}$$

$$= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{i=0}^{n}\binom{n}{i}\lambda^i\mu^{n-i} = \frac{e^{-(\lambda+\mu)}}{n!}(\lambda + \mu)^n$$

2. יהי $k \in \mathbb{N}_0$, נחשב:

$$\mathbb{P}(Y = k) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)\,\mathbb{P}(Y = k \mid X = n) = \sum_{n=k}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k}$$

$$\overset{m=n-k}{=} (p\lambda)^k \sum_{m=0}^{\infty} \frac{\lambda^m e^{-\lambda}}{(k+m)!}\binom{k+m}{k}(1-p)^m = \frac{(p\lambda)^k e^{-\lambda}}{k!}\sum_{m=0}^{\infty}\frac{(\lambda - \lambda p)^m}{m!} = \frac{(p\lambda)^k e^{-\lambda p}}{k!}$$

תוחלת

הגדרה: יהי X משתנה מקרי בדיד. התוחלת של X מוגדרת על ידי

$$\mathbb{E}[X]=\sum_{x\in\mathbb{R}}x\cdot\mathbb{P}(X=x)$$

כאשר טור זה מתכנס בהחלט. במקרה כזה נאמר כי X בעל תוחלת סופית ואחרת נאמר שהוא חסר תוחלת סופית. למעשה, התוחלת מוגדרת להיות ממוצע ערכיו של משתנה מקרי, משוקלל לפי פונקציית ההסתברות הנקודתית.

דוגמה: במשחק מזל כל משתתף מטיל קובייה וזוכה בכמות כסף מסוימת בהתאם לתוצאת הקוביה, כמתואר בטבלה:

תוצאת הקובייה	1	2	3	4	5	6
פרס	5-	3-	0	2	3	4

כאשר פרס שלילי הוא "קנס". נרצה לדעת האם כדאי לשחק במשחק. אם כן, נחשב תוחלת הזכייה:

$$\frac{1}{6}(4+3+2+0-3-5)=\frac{1}{6}>0$$

מאחר שהתוחלת חיובית, אכן כדאי לשחק במשחק.

דוגמה: בהגרלה שבה הפרס $5\cdot10^6$ ש"ח משתתפים 10^6 אנשים, ועלות ההשתתפות היא 5 ש"ח לכל משתתף. יהי X מ"מ שסופר את רווח משתתף בהגרלה, אזי:

$$X=\begin{cases}5\cdot10^6-5 & \text{בהסתברות } \frac{1}{10^6}\\ -5 & \text{בהסתברות } 1-\frac{1}{10^6}\end{cases}$$

לכן תוחלת הרווח היא:

$$\mathbb{E}[X]=\left(5\cdot10^6-5\right)\cdot\frac{1}{10^6}-5\left(1-\frac{1}{10^6}\right)=0$$

טענה: יהי X משתנה מקרי בעל תוחלת, ותהי $f:\mathbb{R}\longrightarrow\mathbb{R}$ פונקציה. אזי:

$$\mathbb{E}[f(X)]=\sum_{x\in\mathbb{R}}f(x)\cdot\mathbb{P}(X=x)$$

אם טור זה מתכנס בהחלט. אחרת ל $f(X)$ אין תוחלת סופית.

הוכחה: נסמן $Y=f(X)$, אזי:

$$\mathbb{E}[Y]=\sum_{y\in\mathbb{R}}y\cdot\mathbb{P}(Y=y)=\sum_{y\in\mathbb{R}}y\cdot\sum_{x\in f^{-1}(\{y\})}\mathbb{P}(X=x)$$

$$=\sum_{y\in\mathbb{R}}\sum_{x\in f^{-1}(\{y\})}f(x)\cdot\mathbb{P}(X=x)=\sum_{x\in\mathbb{R}}f(x)\cdot\mathbb{P}(X=x)$$

המעבר האחרון נכון כי סכמנו על כל התמונות ההפוכות לכל $y\in\mathbb{R}$, כלומר סכמנו על כל התחום.

הערה: ניתן להכליל את הטענה הקודמה עבור וקטורים מקריים: יהי $X=(X_1,\dots,X_d)$ וקטור מקרי בדיד ותהי $f\in\mathbb{R}^d\longrightarrow\mathbb{R}$ פונקציה. אזי המשתנה המקרי $Y=f(X)$ מקיים

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

אם טור זה מתכנס בהחלט, אחרת ל Y אין תוחלת סופית.

דוגמה (תוחלות משתנים מקריים מוכרים):

א. תוחלת משתנה מקרי בנומי $X \sim \text{Bin}(n,p)$ היא:

$$\mathbb{E}[X] = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \overset{m=k-1}{=} \sum_{k=0}^{n-1} \frac{n!}{m!(n-m-1)!} p^{m+1} (1-p)^{n-m-1}$$

$$= np \sum_{k=0}^{n-1} \binom{n-1}{m} p^m (1-p)^{n-m-1} = np \cdot (p + (1-p))^{n-1} = np$$

ב. תוחלת משתנה מקרי פואסוני $X \sim \text{Poi}(\lambda)$ היא:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \overset{m=k-1}{=} \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

ג. תוחלת משתנה מקרי גיאומטרי $X \sim \text{Geo}(p)$ היא:

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} n \cdot \mathbb{P}(X=n) = \sum_{n=0}^{\infty} n \cdot p (1-p)^{n-1} = p \sum_{n=0}^{\infty} n \cdot (1-p)^{n-1} = p \sum_{n=0}^{\infty} \frac{d}{dp} \left(-(1-p)^n\right)$$

$$= -p \cdot \frac{d}{dp} \left(\sum_{n=0}^{\infty} (1-p)^n \right) = -p \cdot \frac{d}{dp} \left(\frac{1-p}{1-(1-p)} \right) = -p \cdot \frac{d}{dp} \left(\frac{1}{p} - 1 \right) = -p \cdot -\frac{1}{p^2} = \frac{1}{p}$$

הרצאה 13

המשך תוחלת

טענה (תכונות התוחלת): יהיו X,Y משתנים מקריים בעלי תוחלת סופית, המוגדרים על אותו מרחב הסתברות. אזי:

1. חיוביות: אם $X \overset{\text{a.s}}{\geq} 0$ אז $\mathbb{E}\left[X\right] \geq 0$. אם בנוסף $\mathbb{P}\left(X>0\right)>0$ אז $\mathbb{E}\left[X\right]>0$.
2. לינאריות: לכל $a,b\in\mathbb{R}$ מתקיים $\mathbb{E}\left[aX+bY\right]=a\mathbb{E}\left[X\right]+b\mathbb{E}\left[Y\right]$.
3. מונוטוניות: אם $X \overset{\text{a.s}}{\geq} Y$ אז $\mathbb{E}\left[X\right]\geq\mathbb{E}\left[Y\right]$. אם בנוסף $\mathbb{P}\left(X>Y\right)>0$ אז $\mathbb{E}\left[X\right]>\mathbb{E}\left[Y\right]$.

הוכחה: נוכיח את המקרה הבדיד.

1. על פי הגדרה:

$$\mathbb{E}\left[X\right]=\sum_{x\in\mathbb{R}}x\cdot\mathbb{P}\left(X=x\right)=\sum_{x\in\mathbb{R}_{-}}x\cdot\mathbb{P}\left(X=x\right)+\sum_{x\in\mathbb{R}_{+}}x\cdot\mathbb{P}\left(X=x\right)=\sum_{x\in\mathbb{R}_{+}}x\cdot\mathbb{P}\left(X=x\right)\geq0$$

כאשר המעבר האחד לפני האחרון נובע מהנחת הסעיף. אי שיוויון חזק מתקבל אם $\mathbb{P}\left(X=x\right)>0$ עבור x חיובי כלשהו.

2. נחשב, על פי הגדרה:

$$\mathbb{E}\left[aX+bY\right]=\sum_{(x,y)\in\mathbb{R}^{2}}\left(ax+by\right)\mathbb{P}\left(X=x,Y=y\right)$$

$$=a\sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}}x\cdot\mathbb{P}\left(X=x,Y=y\right)+b\sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}}y\cdot\mathbb{P}\left(X=x,Y=y\right)$$

שינוי סדר סכימה מותר בטור מתכנס בהחלט, לכן:

$$=a\sum_{x\in\mathbb{R}}\left(x\sum_{y\in\mathbb{R}}\mathbb{P}\left(X=x,Y=y\right)\right)+b\sum_{y\in\mathbb{R}}\left(y\sum_{x\in\mathbb{R}}\mathbb{P}\left(X=x,Y=y\right)\right)$$

מנוסחת ההסתברות השלמה:

$$=a\sum_{x\in\mathbb{R}}x\cdot\mathbb{P}\left(X=x\right)+b\sum_{y\in\mathbb{R}}y\cdot\mathbb{P}\left(Y=y\right)=a\cdot\mathbb{E}\left[X\right]+b\cdot\mathbb{E}\left[Y\right]$$

נותר לוודא את קיום $\mathbb{E}\left[aX+bY\right]$. נוכיח התכנסות בהחלט באופן הבא:

$$\mathbb{E}\left[\left|aX+bY\right|\right]=\sum_{(x,y)\in\mathbb{R}^{2}}\left|aX+bY\right|\cdot\mathbb{P}\left(X=x,Y=y\right)\leq\sum_{(x,y)\in\mathbb{R}^{2}}\left(\left|a\right|\left|X\right|+\left|b\right|\left|Y\right|\right)\cdot\mathbb{P}\left(X=x,Y=y\right)$$

$$=\mathbb{E}\left[\left|a\right|\left|X\right|+\left|b\right|\left|Y\right|\right]=\left|a\right|\mathbb{E}\left[\left|X\right|\right]+\left|b\right|\mathbb{E}\left[\left|Y\right|\right]<\infty$$

3. נחשב על פי הסעיפים הקודמים:

$$\mathbb{E}\left[X\right]=\mathbb{E}\left[X+Y-Y\right]=\mathbb{E}\left[X-Y\right]+\mathbb{E}\left[Y\right]\geq\mathbb{E}\left[X-Y\right]$$

נשים לב כי השתמשנו כאן בכך שלמשתנה המקרי $X-Y$ יש תוחלת סופית (לפי סעיף 2). האי שיוויון החזק נובע מאותו אי שיוויון חזק בסעיף א'.

טענה: יהיו X,Y משתנים מקריים בלתי תלויים ובעלי תוחלת סופית. אזי התוחלת של XY קיימת ומקיימת

$$\mathbb{E}\left[XY\right]=\mathbb{E}\left[X\right]\cdot\mathbb{E}\left[Y\right]$$

הוכחה: נוכיח את המקרה הבדיד. נחשב:

$$\mathbb{E}\left[XY\right]=\sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}}xy\cdot\mathbb{P}\left(X=x,Y=y\right)=\sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}}xy\cdot\mathbb{P}\left(X=x\right)\cdot\mathbb{P}\left(Y=y\right)$$

$$=\sum_{x\in\mathbb{R}}x\cdot\mathbb{P}\left(X=x\right)\sum_{y\in\mathbb{R}}y\cdot\mathbb{P}\left(Y=y\right)=\mathbb{E}\left[X\right]\cdot\mathbb{E}\left[Y\right]$$

כאשר המעבר האחרון נכון משינוי סדר סכימה על טורים מתכנסים בהחלט. נציין כי הוכחת קיום התוחלת דומה להוכחה בטענה הקודמת.

טענה (תרגיל): יהיו X,Y משתנים מקריים על מרחב הסתברות. אזי X,Y בלתי תלויים אם ורק אם לכל שתי פונקציות $f,g:\mathbb{R}\longrightarrow\mathbb{R}$ מתקיים

$$\mathbb{E}\left[f\left(X\right)g\left(Y\right)\right]=\mathbb{E}\left[f\left(X\right)\right]\mathbb{E}\left[g\left(Y\right)\right]$$

בכל מקרה בו f,g חסומות או $g\left(X\right),f\left(Y\right)$ בעלות תוחלת.

הערה (משפט פוביני לטורים חיוביים): לכל אוסף מספרים חיוביים $\left\{a_{ij}\right\}_{i,j\in\mathbb{N}}$ מתקיים

$$\sum_{i\in\mathbb{N}}\sum_{j\in\mathbb{N}}a_{ij}=\sum_{j\in\mathbb{N}}\sum_{i\in\mathbb{N}}a_{ij}$$

טענה (נוסחת הזנב לחישוב תוחלת משתנה מקרי על הטבעיים): יהי X משתנה מקרי המקיים $\text{Sprt}\left(X\right)\subseteq\mathbb{N}$. אזי:

$$\mathbb{E}\left[X\right]=\sum_{n=1}^{\infty}\mathbb{P}\left(X\geq n\right)$$

הוכחה: נחשב את התוחלת לפי ההגדרה, ונחליף את סדר הסכימה באמצעות משפט פוביני:

$$\mathbb{E}\left[X\right]=\sum_{n=1}^{\infty}n\cdot\mathbb{P}\left(X=n\right)=\sum_{n=1}^{\infty}\left(\mathbb{P}\left(X=n\right)+\mathbb{P}\left(X=n\right)+\cdots+\mathbb{P}\left(X=n\right)\right)$$

$$=\sum_{n=1}^{\infty}\sum_{k=1}^{n}\mathbb{P}\left(X=n\right)=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathbb{P}\left(X=n\right)=\sum_{k=1}^{\infty}\mathbb{P}\left(X\geq n\right)$$

כדי להבהיר את שינוי סדר הסכימה, להלן המחשה קצרה:

$$\begin{array}{ccccc}
\mathbb{P}\left(X=1\right) & \mathbb{P}\left(X=2\right) & \mathbb{P}\left(X=3\right) & \mathbb{P}\left(X=4\right)\\
 & \mathbb{P}\left(X=2\right) & \mathbb{P}\left(X=3\right) & \mathbb{P}\left(X=4\right)\\
 & & \mathbb{P}\left(X=3\right) & \mathbb{P}\left(X=4\right)\\
 & & & \mathbb{P}\left(X=4\right)\\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

בסכום $\sum_{n=1}^{\infty}\sum_{k=1}^{n}\mathbb{P}\left(X=n\right)$ אנו סוכמים את עמודות הטבלה, בסכום $\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\mathbb{P}\left(X=n\right)$ אנו סוכמים את שורותיה.

הרצאה 14

שונות

הגדרה: יהי X משתנה מקרי בעל תוחלת סופית μ. השונות של X, מוגדרת

$$\text{Var}\left(X\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)^{2}\right]=\mathbb{E}\left[\left(X-\mu\right)^{2}\right]$$

אם תוחלת זו סופית. אחרת נאמר של- X אין שונות או ששונות אינסופית. הרעיון של מושג השונות פשוט למדי - נשתמש בתוחלת עצמה להגדרת פיזור, על ידי כך שנסתכל על תוחלת ריבוע הסטיה של X מתוחלתו. כלומר, השונות בודקת כמה מפוזר המשתנה המקרי סביב התוחלת שלו. הסיבה בבחירה בריבוע הסטיה דווקא, ולא למשל ערכה המוחלט נובעת מזיקה לאלגברה לינארית - אותה נראה בהמשך.

טענה: לכל משתנה מקרי X בעל תוחלת סופית מתקיים

$$\text{Var}\left(X\right)=\mathbb{E}\left[X^{2}\right]-\mathbb{E}\left[X\right]^{2}$$

כאשר $\mathbb{E}\left[X^{2}\right]$ אינה מוגדרת אם ורק אם השונות אינה מוגדרת. כמו כן, מעתה נתייחס לטענה זו כהגדרה שקולה.

הוכחה: נסמן $\mu=\mathbb{E}\left[X\right]$ ונחשב:

$$\text{Var}\left(X\right)=\mathbb{E}\left[\left(X-\mu\right)^{2}\right]=\mathbb{E}\left[X^{2}-2\mu X+\mu^{2}\right]=\mathbb{E}\left[X^{2}\right]-2\mu\mathbb{E}\left[X\right]+\mu^{2}$$

$$=\mathbb{E}\left[X^{2}\right]-2\mu\cdot\mu+\mu^{2}=\mathbb{E}\left[X^{2}\right]-\mu^{2}=\mathbb{E}\left[X^{2}\right]-\mathbb{E}\left[X\right]^{2}$$

המעבר השלישי נכון מלינאריות התוחלת.

טענה (תכונות השונות): יהי X משתנה מקרי בעל שונות סופית ויהי $a\in\mathbb{R}$, אזי:

1. $\text{Var}\left(X\right)\geq0$ ושיוויון $\text{Var}\left(X\right)=0$ מתקיים אם ורק אם $X\overset{\text{a.s}}{=}c$ עבור $c\in\mathbb{R}$ כלשהו.
2. $\text{Var}\left(X+a\right)=\text{Var}\left(X\right)$.
3. $\text{Var}\left(aX\right)=a^{2}\text{Var}\left(X\right)$.
4. אם Y משתנה מקרי בעל שונות סופית שהנו בלתי תלוי ב-X אז $\text{Var}\left(X+Y\right)=\text{Var}\left(X\right)+\text{Var}\left(Y\right)$.

הוכחה:

1. נשים לב כי $\left(X-\mu\right)^{2}\overset{\text{a.s}}{\geq}0$ לכן מחיוביות התוחלת $\text{Var}\left(X\right)=\mathbb{E}\left[\left(X-\mu\right)^{2}\right]\geq0$.
2. לפי לינאריות התוחלת מתקיים כי $\mathbb{E}\left[X+a\right]=\mathbb{E}\left[X\right]+a$. לכן:

$$\text{Var}\left(X+a\right)=\mathbb{E}\left[\left(X+a-\mathbb{E}\left[X+a\right]\right)^{2}\right]=\mathbb{E}\left[\left(X+a-\mathbb{E}\left[X\right]-a\right)^{2}\right]=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]^{2}\right)\right]=\text{Var}\left(X\right)$$

3. שוב, מליאריות התוחלת ומהטענה הקודמת:

$$\text{Var}\left(aX\right)=\mathbb{E}\left[a^{2}X^{2}\right]-\mathbb{E}\left[aX\right]^{2}=a^{2}\mathbb{E}\left[X^{2}\right]-a^{2}\mathbb{E}\left[X^{2}\right]=a^{2}\text{Var}\left(X\right)$$

4. ראשית, נבחין כי $\mathbb{E}[X+Y]=\mathbb{E}[X]+\mathbb{E}[Y]$ לכן $X+Y$ בעלת תוחלת סופית ונוכל לחשב את שונותו. נשתמש בלינאריות התוחלת ונחשב:

$$\text{Var}(X+Y)=\mathbb{E}\left[(X+Y-\mathbb{E}[X+Y])^2\right]=\mathbb{E}\left[(X+Y-\mathbb{E}[X]-\mathbb{E}[Y])^2\right]$$

$$=\mathbb{E}\left[(X-\mathbb{E}[X]+Y-\mathbb{E}[Y])^2\right]=\mathbb{E}\left[(X-\mathbb{E}[X])^2\right]-2\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]+\mathbb{E}\left[(Y-\mathbb{E}[Y])^2\right]$$

$$=\text{Var}(X)+\text{Var}(Y)-2\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$$

נותר להוכיח כי $\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=0$. ראשית, מאחר ש X,Y בלתי תלויים גם $X-\mathbb{E}[X]$ ו- $Y-\mathbb{E}[Y]$ בלתי תלויים. כעת מכפליות התוחלת למשתנים בלתי תלויים:

$$\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=\mathbb{E}[X-\mathbb{E}[X]]\cdot\mathbb{E}[Y-\mathbb{E}[Y]]=0\cdot 0=0$$

דוגמאות (שונות משתנים מקריים מוכרים):

א. שונות משתנה מקרי ברנולי $X\sim\text{Ber}(p)$ היא:

$$\text{Var}(X)=\mathbb{E}\left[X^2\right]-\mathbb{E}[X]^2=p-p^2=p(1-p)$$

ב. שונות משתנה מקרי בעל התפלגות אחידה $X\sim\text{Unif}([n])$ היא:

$$\text{Var}(X)=\mathbb{E}\left[X^2\right]-\mathbb{E}[X]^2=\sum_{k=1}^{n}k^2\cdot\frac{1}{n}-\left(\frac{n+1}{2}\right)^2=\frac{n(n+1)(2n+1)}{6n}-\frac{(n+1)^2}{4}=\frac{n^2-1}{12}$$

ג. שונות משתנה מקרי בינומי $X\sim\text{Bin}(n,p)$ ניתן לחישוב באופן הבא: ניזכר כי $X\overset{\text{d}}{=}\sum_{k=1}^{n}Y_k$ כאשר $Y_i\sim\text{Ber}(p)$ בלתי תלויים. מתכונה 4 של השונות והחישוב בסעיף א' נקבל:

$$\text{Var}(X)=\text{Var}\left(\sum_{k=1}^{n}Y_k\right)=\sum_{k=1}^{n}\text{Var}(Y_k)=np(1-p)$$

ד. שונות משתנה מקרי פואסוני $X\sim\text{Poi}(\lambda)$: ראינו כבר כי $\mathbb{E}[X]=\lambda$. נחשב את $\mathbb{E}\left[X^2\right]$:

$$\mathbb{E}\left[X^2\right]=\sum_{n=0}^{\infty}n^2\cdot\frac{\lambda^n}{n!}\cdot e^{-\lambda}=\lambda\sum_{n=1}^{\infty}n\cdot\frac{\lambda^{n-1}}{(n-1)!}\cdot e^{-\lambda}\overset{m=n-1}{=}\lambda\sum_{n=1}^{\infty}(m+1)\cdot\frac{\lambda^m}{m!}\cdot e^{-\lambda}=\lambda\mathbb{E}[X+1]=\lambda(\lambda+1)=\lambda^2+\lambda$$

לכן:

$$\text{Var}(X)=\mathbb{E}\left[X^2\right]-\mathbb{E}[X]^2=\lambda^2+\lambda-(\lambda)^2=\lambda$$

ה. שונות משתנה מקרי גיאומטרי $X \sim \text{Geo}(p)$ היא:

$$\text{Var}(X) = \sum_{n=1}^{\infty} n^2 p (1-p)^{n-1} - \frac{1}{p^2} = p(1-p) \sum_{n=1}^{\infty} n^2 (1-p)^{n-2} - \frac{1}{p^2}$$

$$= p(1-p) \sum_{n=1}^{\infty} \left(n^2 - n\right)(1-p)^{n-2} + \sum_{n=1}^{\infty} np(1-p)^{n-1} - \frac{1}{p^2} = p(1-p) \sum_{n=1}^{\infty} \left(n^2 - n\right)(1-p)^{n-2} + \mathbb{E}[X] - \frac{1}{p^2}$$

$$= p(1-p) \sum_{n=1}^{\infty} \left(n^2 - n\right)(1-p)^{n-2} + \frac{1}{p} - \frac{1}{p^2} = p(1-p) \sum_{n=1}^{\infty} \frac{d^2}{dp^2}\left((1-p)^n\right) + \frac{1}{p} - \frac{1}{p^2}$$

$$= p(1-p) \frac{d^2}{dp^2}\left(\sum_{n=1}^{\infty} (1-p^n)\right) + \frac{1}{p} - \frac{1}{p^2} = p(1-p) \frac{d^2}{dp^2}\left(\frac{1-p}{p}\right) + \frac{1}{p} - \frac{1}{p^2}$$

$$= p(1-p) \cdot \frac{2}{p^3} + \frac{1}{p} - \frac{1}{p^2} = \frac{2-2p}{p^2} + \frac{p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

טענה: יהי X משתנה מקרי בעל שונות סופית, אזי:

$$\min_{a \in \mathbb{R}} \left\{ \mathbb{E}\left[(X-a)^2\right] \right\} = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \text{Var}(X)$$

הוכחה: נסמן $Y = X - \mathbb{E}[X]$ ומתקיים $\mathbb{E}[Y] = 0$. לכל $\Delta \in \mathbb{R}$ נחשב:

$$\mathbb{E}\left[(X - \mathbb{E}[X] + \Delta)^2\right] = \mathbb{E}\left[(Y+\Delta)^2\right] = \mathbb{E}\left[Y^2\right] + 2\mathbb{E}\left[\Delta^2\right] + 2\mathbb{E}[\Delta Y] = (X)\text{Var} + \Delta^2 + 2\Delta\mathbb{E}[Y] = (X)\text{Var} + \Delta^2$$

לכן המינימום מתקבל עבור $\Delta = 0$.

סטיית תקן ותיקנון

הגדרה: יהי X משתנה מקרי בעל תוחלת סופית μ. את השורש הריבועי של השונות

$$\sigma_X = \sqrt{\text{Var}(X)}$$

נכנה בשם סטיית התקן של X.

הגדרה: יהי X משתנה מקרי בעל שונות סופית. התיקנון של X הוא המשתנה המקרי

$$\hat{X} = \frac{X - \mathbb{E}[X]}{\sqrt{\text{Var}(X)}} = \frac{X - \mathbb{E}[X]}{\sigma_X}$$

הערה: יהי X משתנה מקרי בעל שונות סופית. אזי:

1. $\mathbb{E}\left[\hat{X}\right] = 0$.
2. $\text{Var}\left(\hat{X}\right) = 1$.

שונות משותפת

הגדרה: יהיו X,Y משתנים מקריים על אותו מרחב הסתברות, בעלי תוחלת סופית. השונות המשותפת של X ו- Y היא:

$$\text{Cov}\left(X,Y\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)\left(Y-\mathbb{E}\left[Y\right]\right)\right]=\mathbb{E}\left[XY\right]-\mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$$

כאשר תוחלת זו מוגדרת היטב. שני משתנים אשר שונותם המשותפת מתאפסת - נקראים בלתי מתואמים.

אבחנות: יהיו X,Y משתנים מקריים.

1. אם X,Y בלתי תלויים אז $\text{Cov}\left(X,Y\right)=0$.
2. $\text{Var}\left(X+Y\right)=\text{Var}\left(X\right)+\text{Var}\left(Y\right)+2\text{Cov}\left(X,Y\right)$ בכל מקרה בו אגף ימין של המשוואה מוגדר היטב.
3. $\text{Cov}\left(X,X\right)=\text{Var}\left(X\right)$.

דוגמה: יהיו $X_1,X_2\sim\text{Unif}\left([6]\right)$ משתנים מקריים המתארים תוצאותיהן של הטלות שתי קוביות משחק הוגנות. נסמן $Y=\min\left(X_1,X_2\right)$ ו- $Z=\max\left(X_1,X_2\right)$ נרצה לחשב את $\text{Var}\left(Y+Z\right)$. נוכל לחשב את השונות באמצעות הנוסחה לעיל, אך חישוב זה יהיה ארוך מאוד. דרך אחרת לחשב אותה, היא להבחין כי $Y+Z\overset{\text{d}}{=}X_1+X_2$, ומאחר ש X_1,X_2 בלתי תלויים:

$$\text{Var}\left(Y+Z\right)=\text{Var}\left(X_1+X_2\right)=\text{Var}\left(X_1\right)+\text{Var}\left(X_2\right)=\frac{35}{6}$$

טענה (תכונות השונות המשותפת, תרגיל): יהיו X,Y,Z משתנים מקריים בעלי שונות סופית, ויהיו $a,b\in\mathbb{R}$. אזי בכל מקרה בו אגף שמאל מוגדר היטב מתקיים:

1. סימטריות: $\text{Cov}\left(X,Y\right)=\text{Cov}\left(Y,X\right)$.
2. בי-לינאריות: $\text{Cov}\left(aX+bZ,Y\right)=a\cdot\text{Cov}\left(X,Y\right)+b\cdot\text{Cov}\left(Z,Y\right)$.
3. $\text{Cov}\left(X+a,Y\right)=\text{Cov}\left(X,Y\right)$.

טענה (תרגיל): יהי $\{X_k\}_{k=1}^{n}$ אוסף של משתנים מקריים. אזי מתקיים:

$$\text{Var}\left(\sum_{k=1}^{n}X_k\right)=\sum_{k=1}^{n}\text{Var}\left(X_k\right)+2\sum_{1\le i<j\le n}\text{Cov}\left(X_i,X_j\right)$$

הרצאה 15

חסם התוחלת - אי שיוויון מרקוב

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

$$\mathbb{P}\left(X \geq a\right) \leq \frac{\mathbb{E}\left[X\right]}{a}$$

הוכחה: נגדיר משתנה מקרי חדש $Y = a \cdot \mathbb{1}_{\{X \geq a\}}$ ונשים לב שמתקיים $X \overset{\text{a.s}}{\geq} Y$. לכן, ממונוטוניות ולינאריות התוחלחת:

$$\mathbb{E}\left[X\right] \geq \mathbb{E}\left[Y\right] = a \cdot \mathbb{E}\left[\mathbb{1}_{\{X \geq a\}}\right] = a \cdot \mathbb{P}\left(X \geq a\right)$$

נחלק ב a לקבלת האי שיוויון המבוקש.

מסקנה: יהי X משתנה מקרי אי שלילי בעל תוחלת סופי. אזי לכל $b > 0$ מתקיים

$$\mathbb{P}\left(X \geq b \cdot \mathbb{E}\left[X\right]\right) \leq \frac{1}{b}$$

דוגמה (פרדוקס יום ההולדת): יהיו $\{X_j\}_{j=1}^{n}$ משתנים מקריים בלתי תלויים, המתפלגים אחידה על $[m]$ - ימי ההולדת של n אנשים. נרצה לחסום מלמעלה את ההסתברות שיש יותר מ ℓ "התנגשויות" - אנשים בעלי אותו יום הולדת.

לכל זוג $1 \leq i < j \leq n$ נגדיר משתנה מקרי $Y_{ij} = \mathbb{1}_{\{X_i = X_j\}}$, ונסמן $Y = \sum\limits_{1 \leq i < j \leq n} Y_{ij}$. כלומר, Y סופר את כמות ההתנגשויות. על מנת שנוכל להפעיל את אי שיוויון מרקוב על Y, עלינו לחשב את תוחלתו. לשם כך, נחשב את תוחלת Y_{ij} ונשתמש בלינאריות התוחלת:

$$\mathbb{E}\left[Y_{ij}\right] = \mathbb{P}\left(X_i = X_j\right) \overset{\substack{\text{הסתברות} \\ \text{שלמה}}}{=} \sum_{k=1}^{m} \mathbb{P}\left(X_i = X_j \mid X_j = k\right) \cdot \mathbb{P}\left(X_j = k\right) = \sum_{k=1}^{m} \mathbb{P}\left(X_i = k\right) \cdot \mathbb{P}\left(X_j = k\right)$$

$$= \sum_{k=1}^{m} \frac{1}{m} \cdot \frac{1}{m} = \frac{m}{m^2} = \frac{1}{m}$$

כעת:

$$\mathbb{E}\left[Y\right] = \mathbb{E}\left[\sum_{1 \leq i < j \leq n} Y_{ij}\right] = \sum_{1 \leq i < j \leq n} \mathbb{E}\left[Y_{ij}\right] = \sum_{1 \leq i < j \leq n} \frac{1}{m} = \binom{n}{2} \cdot \frac{1}{m} = \frac{n\left(n-1\right)}{2m}$$

לכן, מאי שיוויון מרקוב:

$$\mathbb{P}\left(Y > \ell\right) \leq \frac{\mathbb{E}\left[Y\right]}{\ell} = \frac{2n\left(n-1\right)}{2m\ell}$$

דוגמה (בעיית האספן): אספן רוכש בזו אחר זו ביצי הפתעה שכל אחת מהן מכילה אחד מבין n סוגי הפתעות, שנבחרות באופן אחיד ובלתי תלוי.

א. כמה ביצים על האספן לרכוש, כדי להשיג עותק אחד לפחות מכל סוג, בהסתברות גדולה מ $\frac{1}{2}$?

ראשית, לכל i נגדיר X_i משתנה מקרי שיתאר את סוג ההפתעה בביצה ה i. מתקיים כי $X_i \sim \text{Unif}\left(\left[n\right]\right)$ והמשתנים הללו בלתי תלויים. נקבע $Y_0 = 0$ ולכל $k \in [n]$ נגדיר $Y_k = \min\left\{t \in \mathbb{N} : \left|\{X_1, \ldots, X_t\}\right| \geq k\right\}$ - מספר הרכישות עד שימצאו ברשות האספן k סוגי הפתעות שונות. כמו כן, נגדיר $Z_k = Y_k - Y_{k-1}$ - מספר הרכישות מאז שנמצאו ברשות האספן לראשונה $k-1$ סוגי הפתעות עד להשגת הפתעה מסוג חדש. נבחין כי $Z_k \sim \text{Geo}\left(\frac{n-k+1}{n}\right)$ כי בכל רכישה, לאחר שלאספן יש k סוגי הפתעות שונות, הוא מקבל הפתעה מהאוסף בהסתברות $\frac{k-1}{n}$ והפתעה חדשה בהתסברות $\frac{n-k+1}{n}$.

מספר הרכישות עד שימצאו ברשות האספן כל סוגי ההפתעות מתפלג כמו $Z=\sum_{k=1}^{n}Z_k$, ומלינאריות התוחלת:

$$\mathbb{E}\left[Z\right]=\sum_{k=1}^{n}\mathbb{E}\left[Z_k\right]=\sum_{k=1}^{n}\frac{n}{n-k+1}=n\cdot\sum_{k=1}^{n}\frac{1}{n-k+1}\overset{j=n-(k-1)}{=}n\cdot\sum_{j=1}^{n}\frac{1}{j}\leq n\cdot\left(\log n+1\right)$$

כעת, על פי אי שיוויון מרקוב:

$$\mathbb{P}\left(Z\geq 2n\left(\log n+1\right)\right)\leq\frac{\mathbb{E}\left[Z\right]}{2n\left(\log n+1\right)}\leq\frac{n\left(\log n+1\right)}{2n\left(\log n+1\right)}=\frac{1}{2}$$

לכן מספיקות $2n\left(\log n+1\right)$ ביצים כדי שהסיכוי לא לקבל ביצה אחת מכל סוג יהיה לכל היותר $\frac{1}{2}$.

ב. הראו כי הסיכוי שיהיה צורך ברכישת יותר $2n\log n$ ביצים כדי לאסוף את כל סוגי ההפתעות קטן מ $\frac{1}{n}$.

נסמן ב A_j^k את המאורע שלאחר קניית k ביצים, אף ביצה לא הכילה את ההפתעה ה j, כלומר $A_j^k=\{X_1\neq j,\dots,X_k\neq j\}$.
נרצה לחסום את הסתברות $\bigcup_{j=1}^{n}A_j^k$ - שקיים סוג הפתעה j שלא התקבל באף אחת מ k הביצים הראשונות שהאספן קנה. ההסתברות שהביצה i אינה מכילה את ההפתעה ה j היא $1-\frac{1}{n}$ לכן $\mathbb{P}\left(A_j^k\right)=\left(1-\frac{1}{n}\right)^k$. נשתמש בחסם האיחוד ונקבל:

$$\mathbb{P}\left(\bigcup_{j=1}^{n}A_j^k\right)\leq\sum_{j=1}^{n}\mathbb{P}\left(A_j^k\right)=\sum_{j=1}^{n}\left(1-\frac{1}{n}\right)^k=n\left(1-\frac{1}{n}\right)^k\overset{1+x\leq e^x}{\leq}n\cdot e^{-\frac{k}{n}}\overset{k=2n\log n}{=}n\cdot\frac{1}{n^2}=\frac{1}{n}$$

כלומר, הסיכוי שלאחר קניית $2n\log n$ ביצים, האספן לא יקבל את כל סוגי ההפתעות קטנה מ $\frac{1}{n}$ כנדרש.

שונות כמדד פיזור

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

$$\mathbb{P}\left(\left|X-\mathbb{E}\left[X\right]\right|\geq a\right)\leq\frac{\text{Var}\left(X\right)}{a^2}$$

הוכחה: נגדיר משתנה מקרי $Y=\left(X-\mathbb{E}\left[X\right]\right)^2$ ונשים לב כי הוא אי שלילי ומקיים $\mathbb{E}\left[Y\right]=\text{Var}\left(X\right)$. לכן, מאי שיוויון מרקוב, לכל $b\geq 0$ מתקיים:

$$\mathbb{P}\left(Y\geq b\right)\leq\frac{\mathbb{E}\left[Y\right]}{b}=\frac{\text{Var}\left(X\right)}{b}$$

נציב $b=a^2$ ונקבל:

$$\mathbb{P}\left(\left|X-\mathbb{E}\left[X\right]\right|\geq a\right)=\mathbb{P}\left(\left(X-\mathbb{E}\left[X\right]\right)^{\cancel{2}}\geq a^{\cancel{2}}\right)=\mathbb{P}\left(Y\geq a^2\right)\leq\frac{\text{Var}\left(X\right)}{a^2}$$

דוגמה: נחזור לבעיית האספן. נרצה להשתמש באי שיווין צ'בישב כדי למצוא טווח בו מספר הרכישות הנחוץ להשגת כל הפרסים השונים ימצא בהסתברות $\frac{3}{4}$. נשתמש באותם הסימונים: מספר הרכישות מאז שנמצאו ברשות האספן לראשונה $k-1$ סוגי הפתעות עד להשגת הפתעה מסוג חדש היא $Z_k\sim\text{Geo}\left(\frac{n-k+1}{n}\right)$ וכמות הרכישות עד לקבלת כל סוגי ההפתעות היא $Z=\sum_{k=1}^{n}Z_k$. נחשב:

$$\text{Var}\left(Z\right)=\text{Var}\left(\sum_{k=1}^{n}Z_k\right)\overset{\text{המשתנים בלתי תלוים}}{=}\sum_{k=1}^{n}\text{Var}\left(Z_k\right)=\sum_{k=1}^{n}\frac{\frac{k-1}{n}}{\left(\frac{n-k+1}{n}\right)^2}=\sum_{k=1}^{n}\frac{n\left(k-1\right)}{\left(n-k+1\right)^2}\overset{j=n-k+1}{=}\sum_{k=1}^{n}\frac{n\left(n-j\right)}{j^2}$$

$$=n^2\sum_{k=1}^{n}\frac{1}{j^2}-n\sum_{k=1}^{n}\frac{1}{j}\leq n^2\sum_{k=1}^{\infty}\frac{1}{j^2}=n^2\cdot\frac{\pi^2}{6}$$

לכן, מאי שיוויון צ'בישב:

$$\mathbb{P}\left(|Z-\mathbb{E}[Z]|\geq\frac{2\pi n}{\sqrt{6}}\right)\leq\frac{\text{Var}(Z)}{\frac{4\pi^2n^2}{6}}\leq\frac{\frac{\pi^2n^2}{6}}{\frac{4\pi^2n^2}{6}}=\frac{1}{4}$$

ניזכר כי $n\log n\leq n\sum\limits_{i=1}^{n}\frac{1}{i}\leq n(\log n+1)$ לכן בהסתברות לפחות $\frac{3}{4}$, מספר הביצים שנדרש האספן לקנות עד לקבלת עותק מכל סוג הפתעה יהיה בין $n\log n+\frac{2\pi n}{\sqrt{6}}$ לבין $n(\log n+1)+\frac{2\pi n}{\sqrt{6}}$.

דוגמה (ריכוז מידה של משתנה בינומי): מוטלים מיליון מטבעות הוגנים. נציג חסם מלמטה להסתברות שכמות תוצאות העץ שתתקבלנה תהיה בין $495,000$ ל $505,000$.

נסמן ב X את כמות העצים שהתקבלו בניסוי, אזי $X\sim\text{Bin}\left(10^6,\frac{1}{2}\right)$. ראשית, על פי תוחלת ושונות של משתנה בינומי, אנו יודעים כי $\mathbb{E}[X]=10^6\cdot\frac{1}{2}=500,000$ ו- $\text{Var}(X)=10^6\cdot\frac{1}{2}\cdot\left(1-\frac{1}{2}\right)=250,000$. נפעיל את אי שיוויון צ'בישב ונקבל:

$$\mathbb{P}(495,000<X<505,000)=1-\mathbb{P}(|X-\mathbb{E}[X]|\geq 5000)=1-\frac{250,000}{5000^2}=\frac{99}{100}$$

כלומר הסיכוי לסטות מהתוחלת ביותר מ $5,000$ הוא לכל היותר אחוז אחד.

משפט (החוק החלש של המספרים הגדולים למשתנים בעלי שונות): יהי X משתנה מקרי בעל שונות סופית, ויהי $\epsilon>0$. אזי:

$$\lim_{n\to\infty}\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mathbb{E}[X]\right|<\epsilon\right)=1$$

כאשר $\{X_k\}_{k=1}^{n}$ משתנים מקריים בלתי תלויים שווי התפלגות ל X. כלומר, בהינתן n "תצפיות" של המשתנה המקרי X, ממוצען שואף לתוחלת של המשתנה המקרי.

הוכחה: יהי $n\in\mathbb{N}$ ונסמן $S_n=\sum\limits_{k=1}^{n}X_k$. מלינאריות התוחלת מתקיים:

$$\mathbb{E}[S_n]=\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[X_k]=\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[X]=\mathbb{E}[X]$$

מאחר ש $\{X_k\}_{k=1}^{n}$ בלתי תלויים, מתקיים כי:

$$\text{Var}(S_n)=\text{Var}\left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\text{Var}\left(\sum_{k=1}^{n}X_k\right)=\frac{1}{n^2}\sum_{k=1}^{n}\text{Var}(X_k)=\frac{1}{n^2}\sum_{k=1}^{n}\text{Var}(X)=\frac{\text{Var}(X)}{n}$$

נפעיל את אי שיוויון צ'בישב ונקבל:

$$1\geq\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mathbb{E}[X]\right|<\epsilon\right)=1-\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mathbb{E}[X]\right|\geq\epsilon\right)\geq 1-\frac{\text{Var}(S_n)}{\epsilon^2}=1-\frac{\text{Var}(X)}{n\epsilon^2}\xrightarrow[n\to\infty]{}1$$

וממשפט הסנדוויץ' סיימנו.

הרצאה 16

נוסחת התוחלת השלמה

הגדרה: יהיו X משתנה מקרי בעל תוחלת סופית, ו- A מאורע בעל הסתברות חיובית. תוחלתו המותנית של X בהינתן מאורע A מוגדר להיות תוחלתו של X תחת פונקציית ההסתברות המותנית $\mathbb{P}_A$, כלומר:

$$\mathbb{E}\left[X \mid A\right] = \sum_{x\in\mathbb{R}} x \cdot \mathbb{P}\left(X = x \mid A\right)$$

אבחנה: נוכל לכתוב את התוחלת המותנית גם באופן הבא:

$$\mathbb{E}\left[X \mid A\right] = \sum_{x\in\mathbb{R}} x \cdot \mathbb{P}\left(X = x \mid A\right) = \sum_{x\in\mathbb{R}} x \cdot \frac{\mathbb{P}\left(X = x, A\right)}{\mathbb{P}\left(A\right)} = \frac{1}{\mathbb{P}\left(A\right)} \sum_{x\in\mathbb{R}} x \cdot \mathbb{P}\left(\mathbb{1}_A \cdot X = x\right) = \frac{\mathbb{E}\left[\mathbb{1}_A \cdot X\right]}{\mathbb{P}\left(A\right)}$$

כיוון שמתקיים $|\mathbb{1}_A \cdot X| = \mathbb{1}_A \cdot |X| \leq |X|$ נוכל להסיק כי התוחלת המותנית מוגדרת היטב.

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

$$\mathbb{E}\left[X\right] = \sum_{\substack{A\in\mathcal{A} \\ \mathbb{P}(A)>0}} \mathbb{E}\left[X \mid A\right] \cdot \mathbb{P}\left(A\right)$$

הוכחה: עבור X בדיד, על פי הגדרה:

$$\mathbb{E}\left[X\right] = \sum_{s\in\mathbb{R}} s \cdot \mathbb{P}\left(X = s\right) = \sum_{s\in\mathbb{R}} s \cdot \sum_{\substack{A\in\mathcal{A} \\ \mathbb{P}(A)>0}} \mathbb{P}\left(X = s \mid A\right) \cdot \mathbb{P}\left(A\right)$$

$$= \sum_{s\in\mathbb{R}} \mathbb{P}\left(A\right) \sum_{\substack{A\in\mathcal{A} \\ \mathbb{P}(A)>0}} s \cdot \mathbb{P}\left(X = s \mid A\right) = \sum_{\substack{A\in\mathcal{A} \\ \mathbb{P}(A)>0}} \mathbb{E}\left[X \mid A\right] \cdot \mathbb{P}\left(A\right)$$

כאשר המעבר השני נובע מנוסחת ההסתברות השלמה.

דוגמה: יהי N משתנה מקרי בעל תוחלת הנתמך על $\mathbb{N}$. בהינתן ערך של N, נסמנו k, מטילים קוביה k פעמים וסוכמים את התוצאות. מהי תוחלת הסכום?

לכל i, נגדיר משתנה מקרי $X_i \sim \text{Unif}\left([6]\right)$ תוצאת ההטלה ה i. כמו כן, נגדיר $Y = \sum\limits_{i=1}^{N} X_i$ סכום ההטלות. מנוסחת התוחלת השלמה:

$$\mathbb{E}\left[Y\right] = \mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \sum_{k=1}^{\infty} \mathbb{E}\left[Y \mid N = k\right] \cdot \mathbb{P}\left(N = k\right) = \sum_{k=1}^{\infty} \mathbb{E}\left[\sum_{i=1}^{k} X_i\right] \cdot \mathbb{P}\left(N = k\right)$$

מלינאריות התוחלת $\mathbb{E}\left[\sum\limits_{i=1}^{k} X_i\right] = \sum\limits_{i=1}^{k} \mathbb{E}\left[X_i\right]$ אך מאחר שהמשתנים $\{X_i\}$ שווי התפלגות, לכל i מתקיים $\mathbb{E}\left[X_i\right] = \mathbb{E}\left[X_1\right]$. לכן:

$$\mathbb{E}\left[Y\right] = \sum_{k=1}^{\infty} k \cdot \mathbb{E}\left[X_1\right] \cdot \mathbb{P}\left(N = k\right) = \mathbb{E}\left[X_1\right] \cdot \mathbb{E}\left[N\right] = 3.5 \cdot \mathbb{E}\left[Y\right]$$

תוחלת מותנית כמשתנה מקרי

הגדרה: יהיו Y משתנה מקרי ו- X משתנה מקרי בעל תוחלת על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. התוחלת המותנית של X בהינתן Y שנסמנה $\mathbb{E}[X \mid Y]$ היא המשתנה המקרי Z המוגדר על ידי:

$$Z(\omega) = \mathbb{E}[X \mid Y](\omega) = \mathbb{E}[X \mid Y = Y(\omega)]$$

משפט (נוסחת התוחלת השלמה): יהיו Y משתנה מקרי ו- X משתנה מקרי בעל תוחלת על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. אזי:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]]$$

נעיר כי למשפט קוראים גם "חוק התוחלת החוזרת" ו- "משפט ההחלקה".

הוכחה: על פי משפט התוחלת השלמה:

$$\mathbb{E}[X] = \sum_{y \in \text{Sprt}(Y)} \mathbb{E}[X \mid Y = y] \cdot \mathbb{P}(Y = y) = \mathbb{E}[\mathbb{E}[X \mid Y]]$$

דוגמה: נשוב לדוגמה הקודמת, והפעם נחשב את תוחלת הסכום בעזרת חוק התוחלת החוזרת:

$$\mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{N} X_i \mid N\right]\right] = \mathbb{E}[N \cdot \mathbb{E}[X_1]] = \mathbb{E}[X_1] \cdot \mathbb{E}[N]$$

טענה (תכונות התוחלת המותנית): יהיו X, Y, Z משתנים מקריים בעלי תוחלת ותהי $f : \mathbb{R} \longrightarrow \mathbb{R}$. אזי:

1. הוצאת החלק הידוע: $\mathbb{E}[f(Y) \cdot X \mid Y] = f(Y) \cdot \mathbb{E}[X \mid Y]$.
2. השמטת החלק הבלתי-תלוי: אם X בלתי תלוי ב Y אז $\mathbb{E}[X \mid Y] = \mathbb{E}[X]$.

הוכחה:

1. תרגיל.
2. נחשב:

$$\mathbb{E}[X \mid Y] = \sum_{y \in \text{Sprt}(Y)} \mathbb{E}[X \mid Y = y] \cdot \mathbb{P}(Y = y) = \sum_{y \in \text{Sprt}(Y)} \mathbb{E}[X] \cdot \mathbb{P}(Y = y) = \mathbb{E}[X] \cdot \sum_{y \in \text{Sprt}(Y)} \mathbb{P}(Y = y) = \mathbb{E}[X]$$

דוגמה: יהיו $\{X_i\}_{i=1}^{n}$ משתני ברנולי p בלתי תלוים ונסמן את סכומם $Z = \sum_{i=1}^{n} X_i$ מהי $\mathbb{E}[X_1 \mid Z]$?

נשים לב כי $Z \sim \text{Bin}(n, p)$ לכן, לכל $k \in [n]$:

$$\mathbb{E}[X_1 \mid Z = k] = \sum_{x \in \mathbb{R}} x \cdot \mathbb{P}(X_1 = x \mid Z = k) = \mathbb{P}(X_1 = 1, Z = k) = \frac{\mathbb{P}(X_1 = 1, Z = k)}{\mathbb{P}(Z = k)}$$

$$= \frac{\mathbb{P}(X_1 = 1, Z = k)}{\mathbb{P}(Z = k)} = \frac{\cancel{p} \cdot \binom{n-1}{k-1} \cdot \cancel{p^{k-1}} \cdot \cancel{(1-p)^{n-k}}}{\binom{n}{k} \cdot \cancel{p^k} \cdot \cancel{(1-p)^{n-k}}} = \frac{k}{n}$$

כלומר:

$$\mathbb{E}[X_1 \mid Z] = \frac{Z}{n}$$

שונות מותנית

הגדרה: יהיו X ו-Y משתנים מקריים בדידים על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$ כאשר X בעל שונות סופית. השונות המותנית של X בהינתן Y היא המשתנה המקרי

$$\text{Var}\left(X\mid Y\right)\left(\omega\right)=\begin{cases}\text{Var}\left(X\mid Y=Y\left(\omega\right)\right) & \mathbb{P}\left(Y=Y\left(\omega\right)\right)>0\\ 0 & \mathbb{P}\left(Y=Y\left(\omega\right)\right)=0\end{cases}$$

אבחנה: אם למשתנה מקרי X קיימת שונות מותנית, אז מתקיים:

$$\text{Var}\left(X\mid Y\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\right]\right)^{2}-Y\right]=\mathbb{E}\left[X^{2}\mid Y\right]-\mathbb{E}\left[X\mid Y\right]^{2}$$

ההוכחה דומה לפתיחת הסוגרים בהוכחה ללא ההתניה.

טענה: יהי X משתנה מקרי בעל שונות סופית על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$, ויהי Y משתנה מקרי על אותו מרחב הסתברות. אזי:

$$\mathbb{E}\left[\text{Var}\left(X\mid Y\right)\right]=\text{Var}\left(X-\mathbb{E}\left[X\mid Y\right]\right)$$

הוכחה: על פי הגדרה:

$$\text{Var}\left(X-\mathbb{E}\left[X\mid Y\right]\right)=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\mid Y\right]-\mathbb{E}\left[X-\mathbb{E}\left[X\mid Y\right]\right]\right)^{2}\right]$$

נשים לב כי $\mathbb{E}\left[X-\mathbb{E}\left[X\mid Y\right]\right]=\mathbb{E}\left[X\right]-\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]\right]=0$ מחוק התוחלת החוזרת. לכן:

$$=\mathbb{E}\left[\left(X-\mathbb{E}\left[X\mid Y\right]\right)^{2}\right]=\mathbb{E}\left[X^{2}\right]-2\cdot\mathbb{E}\left[X\cdot\mathbb{E}\left[X\mid Y\right]\right]+\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]\right]^{2}$$

נפתח את המחובר האמצעי, לפי חוק התוחלת החוזרת:

$$\mathbb{E}\left[X\cdot\mathbb{E}\left[X\mid Y\right]\right]=\mathbb{E}\left[\mathbb{E}\left[X\cdot\mathbb{E}\left[X\mid Y\right]\mid Y\right]\right]$$

בהינתן Y, $\mathbb{E}\left[X\mid Y\right]$ ידוע לכן:

$$=\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]\cdot\mathbb{E}\left[X\mid Y\right]\right]=\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]^{2}\right]$$

סה"כ:

$$\text{Var}\left(X-\mathbb{E}\left[X\mid Y\right]\right)=\mathbb{E}\left[X^{2}\right]-2\cdot\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]^{2}\right]+\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]\right]^{2}$$

שוב, מחוק התוחלת החוזרת:

$$=\mathbb{E}\left[X^{2}\right]-\mathbb{E}\left[\mathbb{E}\left[X\mid Y\right]^{2}\right]=\mathbb{E}\left[\mathbb{E}\left[X^{2}\mid Y\right]-\mathbb{E}\left[X\mid Y\right]^{2}\right]=\mathbb{E}\left[\text{Var}\left(X\mid Y\right)\right]$$

משפט (נוסחת השונות השלמה): יהי X משתנה מקרי בעל שונות סופית על מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$, ויהי Y משתנה מקרי על אותו מרחב הסתברות. אזי:

$$\text{Var}\left(X\right)=\mathbb{E}\left[\text{Var}\left(X\mid Y\right)\right]+\text{Var}\left(\mathbb{E}\left[X\mid Y\right]\right)$$

הגורם $\text{Var}\left(\mathbb{E}\left[X\mid Y\right]\right)$ נקרא "השונות המוסברת" והגורם $\mathbb{E}\left[\text{Var}\left(X\mid Y\right)\right]$ נקרא "השונות הבלתי מוסברת".

הוכחה: נחשב:

$$\text{Var}\left(X\right)=\text{Var}\left(X-\mathbb{E}\left[X\mid Y\right]+\mathbb{E}\left[X\mid Y\right]\right)$$

$$=\text{Var}\left(X-\mathbb{E}\left[X\mid Y\right]\right)+\text{Var}\left(\mathbb{E}\left[X\mid Y\right]\right)+2\cdot\text{Cov}\left(X-\mathbb{E}\left[X\mid Y\right],\mathbb{E}\left[X\mid Y\right]\right)$$

לפי האבחנה הקודמת:

$$=\left(\mathbb{E}\left[X\mid Y\right]\right)\text{Var}+\mathbb{E}\left[\text{Var}\left(X\mid Y\right)\right]$$

נותר להסביר מדוע $2\cdot\text{Cov}\left(X-\mathbb{E}\left[X\mid Y\right],\mathbb{E}\left[X\mid Y\right]\right)=0$, נשאר כתרגיל בפתיחת סוגריים.

הרצאה 17

מרחבי הסתברות כללים

הגדרה: יהי Ω מרחב מדגם כלשהו. קבוצה $\mathcal{F} \subseteq 2^{\Omega}$ נקראת σ-אלגברה, אם היא מקיימת את התכונות הבאות:

1. $\emptyset \in \mathcal{F}$.
2. סגירות לאיחוד בן מנייה: אם $\{U_i\}_{i=1}^{\infty} \in \mathcal{F}$ אז $\bigcup_{i=1}^{\infty} U_i \in \mathcal{F}$.
3. סגירות למשלים: אם $U \in \mathcal{F}$ אז $\Omega \backslash U = U^c \in \mathcal{F}$.

אבחנות: תהי $\mathcal{F}$ σ-אלגברה על מרחב מדגם Ω, אזי מתקיים:

1. $\Omega \in \mathcal{F}$.
2. סגירות לחיתוך בן מנייה: אם $\{U_i\}_{i=1}^{\infty} \in \mathcal{F}$ אז $\bigcap_{i=1}^{\infty} U_i \in \mathcal{F}$.

הגדרה: תהי $\mathcal{F}$ σ-אלגברה על מרחב מדגם Ω. פונקצית הסתברות $\mathbb{P} : \mathcal{F} \longrightarrow [0, \infty)$ על $(\Omega, \mathcal{F})$ היא פונקציה המקיימת את שתי התכונות הבאות:

1. $\mathbb{P}(\Omega) = 1$.
2. סכימות בת מניה ($\sigma -$ additivity): לכל אוסף בן מנייה של מאורעות זרים $\{A_i\}_{i=1}^{\infty}$ ב $\mathcal{F}$ מתקיים

$$\mathbb{P}\left(\dot{\bigcup_{i=1}^{\infty}} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

הגדרה: תהי $\mathcal{F}$ σ-אלגברה על מרחב מדגם Ω ותהי $\mathbb{P}$ פונקצית הסתברות על $(\Omega, \mathcal{F})$. השלשה $(\Omega, \mathcal{F}, \mathbb{P})$ מכונה מרחב הסתברות.

אלגברת בורל ומרחב ההסתברות התקני

טענה: קיימת $\mathbb{B}(\mathbb{R})$, ה-σ-אלגברה של בורל, שמכילה את כל הקטעים $[a, b] \subseteq \mathbb{R}$ והיא מינימלית במובן שהיא מוכלת בכל σ-אלגברה אחרת שמכילה קבוצות אלו. לכל קטע $I \subseteq \mathbb{R}$ נגדיר σ-אלגברה של בורל על ידי:

$$\mathbb{B}(I) = \{A \cap I : A \in \mathbb{B}(\mathbb{R})\}$$

אבחנה: לכל $a < b \in \mathbb{R}$ מתקיים כי $(a, b) \in \mathbb{B}(\mathbb{R})$, מאחר שכל קטע פתוח שכזה ניתן להצגה על ידי איחוד בן מנייה של קטעים סגורים:

$$(a, b) = \bigcup_{n=k}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n}\right]$$

כאשר k המספר הקטן ביותר עבורו $\left[a + \frac{1}{n}, b - \frac{1}{n}\right] \subseteq (a, b)$. לכן מתכונת הסכימות בת מניה של ה-σאלגברה, $(a, b) \in \mathbb{B}(\mathbb{R})$. באופן דומה, ניתן להראות כי:

$$[a, b) . (a, b], \{a\}, [a, \infty) . (a, \infty), (-\infty, b], (-\infty, b) \in \mathbb{B}(\mathbb{R})$$

הערה: יהיו $\{\mathcal{F}_\alpha\}_{\alpha \in I}$ σ-אלגבראות על אותו מרחב Ω. אזי $\bigcap_{\alpha \in I} \mathcal{F}_\alpha$ היא σ-אלגברה.

משפט (מידת לבג): קיימת פונקצית הסתברות על $\mathbb{B}([0,1])$ כך שלכל $0 \leq a \leq b \leq 1$ מתקיים:

$$\mathbb{P}([a,b]) = b - a$$

הגדרה: המרחב $([0,1], \mathbb{B}([0,1]), \mathbb{P})$ כאשר $\mathbb{P}$ היא מידת לבג יכונה מרחב ההסתברות התקני, או מרחב הסתברות אחיד על $[0,1]$.

הערה: נשים לב שזהו המרחב הראשון אותו אנו חוקרים, בו ההסתברות אינה בדידה ולמעשה הוא מקיים את התכונה המעניינת $\mathbb{P}(\{a\}) = \mathbb{P}([a,a]) = a - a = 0$ לכל $a \in \Omega$. כלומר, הסתברות כל נקודון היא 0 ולכן גם מתקיים $\mathbb{P}((a,b)) = b - a$ לכל $0 \leq a \leq b \leq 1$, שהרי:

$$\mathbb{P}((a,b)) = \mathbb{P}([a,b]) - \mathbb{P}(\{a\}) - \mathbb{P}(\{b\}) = b - a$$

למעשה, הסתברותה של כל קבוצה בת מנייה $A \subset [0,1]$ היא 0, מתכונת סכימות בת מניה של פונקציות הסתברות.

משתנים מקריים מעל מרחבי הסתברות כללים

הגדרה: יהי Ω מרחב מדגם ותהי $\mathcal{F}$ σ-אלגברה על Ω. משתנה מקרי על $(\Omega, \mathcal{F})$ הוא פונקציה $X : \Omega \longrightarrow \mathbb{R}$ המקיימת $X^{-1}(B) \in \mathcal{F}$ לכל $B \in \mathbb{B}(\mathbb{R})$.

הגדרה: יהי X משתנה מקרי על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. התפלגותו של X היא הפונקציה $\mathbb{P}_X : \mathbb{B}(\mathbb{R}) \longrightarrow [0,1]$ הנתונה על ידי:

$$\mathbb{P}_X(A) = \mathbb{P}\left(X^{-1}(A)\right)$$

כמקודם, נכתוב $X \sim \mathbb{P}_X$ ונסמן $\mathbb{P}(X \in A) = \mathbb{P}_X(A)$ ונקרא כי זו ההסתברות שערכו של X נמצא ב A.

הגדרה: יהי X משתנה מקרי על מרחב הסתברות. הפונקציה $F_X : \mathbb{R} \longrightarrow \mathbb{R}$ הנתונה על ידי $F_X(s) = \mathbb{P}(X \leq s)$ מכונה פונקציית ההתפלגות המצטברת של X. באופן דומה, הפונקציה $\overline{F}_X : \mathbb{R} \longrightarrow \mathbb{R}$ הנתונה על ידי $\overline{F}_X(s) = \mathbb{P}(X > s) = 1 - F_X(s)$ מכונה פונקציית ההתפלגות השיורית של X.

טענה (תכונות פונקציית ההתפלגות המצטברת): יהי X משתנה מקרי על מרחב הסתברות, אזי:

1. F_X מונוטונית עולה.
2. F_X מקיימת $\lim_{s \to \infty} F_X(s) = 1$, $\lim_{s \to -\infty} F_X(s) = 0$.
3. F_X רציפה מימין ובעלת גבול משמאל.

הערה: יהי X משתנה מקרי על מרחב הסתברות. נאמר ש a היא נקודת קפיצה של F_X אם ורק אם $\mathbb{P}(X = a) > 0$.

הגדרה: יהי X משתנה מקרי על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. נאמר כי ל X התפלגות <u>בדידה</u> אם קיימת קבוצה A בת מניה, כך שמתקיים $\mathbb{P}(X \in A) = 1$.

הגדרה: יהי X משתנה מקרי על מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. נאמר כי ל X התפלגות <u>רציפה</u> אם לכל $x \in \mathbb{R}$ מתקיים $\mathbb{P}(X = x) = 0$.

דוגמה: יהי $([0,1], \mathbb{B}([0,1]), \mathbb{P})$ מרחב ההסתברות התקני ונגדיר משתנה מקרי $X(\omega) = \min\left(\omega, \frac{1}{2}\right)$. נשים לב כי מצד אחד $\mathbb{P}\left(X = \frac{1}{2}\right) = \mathbb{P}\left(\omega \in \left[\frac{1}{2}, 1\right]\right) = \frac{1}{2}$ ולכן X אינו רציף. אבל מצד שני, לכל $x \neq \frac{1}{2}$ מתקיים כי $\mathbb{P}(X = x) = 0$ לכן לכל קבוצה בת מניה A מתקיים $\mathbb{P}(X \in A) \leq \mathbb{P}\left(X = \frac{1}{2}\right) = \frac{1}{2}$ ולכן X אינו בדיד. כלומר, המשתנה המקרי X הוא "מעורב".

הערה: העיסוק במשתנים מקריים כללים דורש שימוש נרחב בתורת המידה. לכן, נגביל את עצמנו למשפחה שימושית במיוחד ונגישה של משתנים מקריים המכונים משתנים רציפים בהחלט.

תלות של משתנים מקריים כלליים

הגדרה: נאמר שמשתנים מקריים X ו- Y הנם בלתי תלויים ונסמן $X \perp\!\!\!\perp Y$, אם לכל שתי קבוצות $A, B \subset \mathbb{B}(\mathbb{R})$ מתקיים:

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B)$$

משתנים מקריים רציפים בהחלט

הגדרה: יהי X משתנה מקרי על מרחב הסתברות. נאמר ש X רציף בהחלט אם קיימת פונקציה אינטגרבילית $f : \mathbb{R} \longrightarrow [0, \infty)$ כך שלכל $-\infty \leq a \leq b \leq \infty$ מתקיים:

$$\mathbb{P}(X \in (a, b)) = \int_a^b f(x)\, dx$$

הפונקציה f נקראת הצפיפות של X.

אבחנה: יהי X משתנה מקרי רציף בהחלט, אזי הוא רציף. כלומר, לכל $a \in \mathbb{R}$ מתקיים $\mathbb{P}(X = a) = 0$. ניתן להוכיח אבחנה זו באמצעות רציפות פונקצית ההסתברות:

$$\mathbb{P}(X = a) = \lim_{a' \longrightarrow a^-} \int_a^{a'} f(x)\, dx = 0$$

אבחנה: יהי X משתנה רציף בהחלט על מרחב הסתברות, אזי מתקיים כי:

$$F_X(s) = \mathbb{P}(X \leq s) = \int_{-\infty}^{s} f(x)\, dx$$

$$\overline{F}_X(s) = \mathbb{P}(X > s) = \int_s^{\infty} f(x)\, dx$$

הרצאה 18

המשך משתנים מקריים רציפים בהחלט

אבחנה: עבור כל פונקציה $f:\mathbb{R}\longrightarrow[0,\infty)$ אינטגרבילית המקיימת $\int\limits_{-\infty}^{\infty}f(x)\,dx=1$ קיים משתנה מקרי X כך ש $f_X=f$.

דוגמה (חישוב הסתברות וצפיפות): נתון משתנה מקרי X בעל צפיפות $f_X(x)=\frac{x+1}{2}\cdot\mathbb{1}_{[-1,1]}(x)$. נחשב את ההסתברות $\mathbb{P}\left(X\geq\frac{1}{2}\right)$ ונמצא את צפיפותם של $Y=5X$ ושל $Z=X^2$.

ראשית, נבחין שאכן מדובר בפונקציית צפיפות, שכן:

$$\int\limits_{-\infty}^{\infty}f_X(x)\,dx=\int\limits_{-1}^{1}\frac{x+1}{2}dx=\left.\frac{x^2+2x}{4}\right|_{-1}^{1}=1$$

כעת:

$$\mathbb{P}\left(X\geq\frac{1}{2}\right)=\int\limits_{\frac{1}{2}}^{\infty}f_X(x)\,dx=\int\limits_{\frac{1}{2}}^{1}\frac{x+1}{2}dx=\left.\frac{x^2+2x}{4}\right|_{\frac{1}{2}}^{1}=\frac{3}{4}-\frac{5}{16}=\frac{7}{16}$$

לחישוב צפיפות Y, נחשב את פונקציית ההתפלגות המצטברת F_Y. עבור $y<-5$ מתקיים $\mathbb{P}(Y\leq y)=0$ ועבור $y>5$ מתקיים $\mathbb{P}(Y\leq y)=1$. עבור $y\in[-5,5]$ נחשב:

$$F_Y(y)=\mathbb{P}(Y\leq y)=\mathbb{P}(5X\leq y)=\mathbb{P}\left(X\leq\frac{y}{5}\right)=F_X\left(\frac{y}{5}\right)$$

$$=\int\limits_{-\infty}^{\frac{y}{5}}f_X(x)\,dx=\int\limits_{-1}^{\frac{y}{5}}\frac{x+1}{2}dx=\left.\frac{x^2+2x}{4}\right|_{-1}^{\frac{y}{5}}=\frac{y^2+10y}{100}+\frac{1}{4}$$

נגזור ונקבל $f_Y(y)=\left(\frac{y}{50}+\frac{1}{10}\right)\cdot\mathbb{1}_{[-5,5]}(y)$. כעת נחשב את צפיפות Z, באופן דומה. עבור $z<0$ מתקיים $\mathbb{P}(Z\leq z)=0$ ועבור $z>1$ מתקיים $\mathbb{P}(Z\leq z)=1$. עבור $z\in[0,1]$ נחשב:

$$F_Z(z)=\mathbb{P}\left(X^2\leq z\right)=\mathbb{P}\left(X\in\left[-\sqrt{z},\sqrt{z}\right]\right)=\int\limits_{-\sqrt{z}}^{\sqrt{z}}f_X(x)\,dx=\int\limits_{-\sqrt{z}}^{\sqrt{z}}\frac{x+1}{2}dx=\sqrt{z}$$

נגזור ונקבל $f_Z(z)=\frac{1}{2\sqrt{z}}\cdot\mathbb{1}_{[0,1]}(z)$.

דוגמה: נתון משתנה מקרי X בעל צפיפות $f_X(x)=\frac{1}{\pi(1+x^2)}$. נחשב את ההסתברות $\mathbb{P}(X\in[a,b])$.

ראשית, נבחין שאכן מדובר בפונקציית צפיפות, שכן:

$$\int\limits_{-\infty}^{\infty}f_X(x)\,dx=\int\limits_{-\infty}^{\infty}\frac{dx}{\pi\left(1+x^2\right)}=\frac{1}{\pi}\cdot\left.\arctan(x)\right|_{-\infty}^{\infty}=\frac{1}{\pi}\left(\frac{\pi}{2}-\left(-\frac{\pi}{2}\right)\right)=1$$

כעת:

$$\mathbb{P}(X\in[a,b])=\int\limits_{a}^{b}f_X(x)\,dx=\frac{1}{\pi}\int\limits_{a}^{b}\frac{dx}{1+x^2}=\frac{1}{\pi}\left(\arctan(b)-\arctan(a)\right)$$

נעיר שהתפלגות זו מכונה התפלגות קושי.

תוחלת ושונות של משתנה מקרי רציף בהחלט

הגדרה (תוחלת משתנה מקרי רציף בהחלט): יהי X משתנה מקרי רציף בהחלט בעל פונקצית צפיפות $f_x(X)$, אזי:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx$$

בכל מקרה בו האינטגרל מתכנס בהחלט. אחרת אין ל- X תוחלת.

טענה: לכל משתנה מקרי רציף בהחלט מתקיים:

$$\mathbb{E}[X] = \int_{0}^{\infty} \overline{F}(t)\,dt - \int_{-\infty}^{0} F(t)\,dt$$

הוכחה: נחשב באמצעות שינוי סדר אינטגרציה:

$$\int_{0}^{\infty} \overline{F}(t)\,dt = \int_{0}^{\infty} \left(\int_{t}^{\infty} f(x)\,dx\right) dt = \int_{0}^{\infty} \left(\int_{0}^{x} dt\right) \cdot f(x)\,dx = \int_{0}^{\infty} x \cdot f(x)\,dx$$

$$\int_{-\infty}^{0} F(t)\,dt = \int_{0}^{\infty} \left(\int_{t}^{\infty} f(-x)\,dx\right) dt = \int_{0}^{\infty} \left(\int_{0}^{x} dt\right) \cdot f(-x)\,dx = \int_{0}^{\infty} x \cdot f(-x)\,dx = \left\{ \begin{array}{c} y=-x \\ dy=-dx \end{array} \right\} = -\int_{-\infty}^{0} y \cdot f(y)\,dy$$

אכן:

$$\int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_{-\infty}^{0} yf(y)\,dy + \int_{0}^{\infty} xf(x)\,dx = \int_{0}^{\infty} \overline{F}(t)\,dt - \int_{-\infty}^{0} F(t)\,dt$$

טענה: יהי X משתנה מקרי בעל צפיפות f_X ותהי $g:\mathbb{R}\longrightarrow\mathbb{R}$ פונקציה. אזי $Y=g(X)$ הוא משתנה מקרי המקיים:

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$$

אם ורק אם האינטגרל מתכנס בהחלט.

הגדרה: יהי X משתנה מקרי רציף בהחלט בעל תוחלת. השונות של X, מוגדרת על ידי:

$$\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - \left(\int_{-\infty}^{\infty} x f_X(x)\,dx\right)^2$$

אם תוחלת זו סופית. אחרת נאמר של- X אין שונות או ששונות אינסופית.

טענה (תכונות התוחלת, תרגיל): יהיו X,Y משתנים מקריים רציפים בהחלט על מרחב הסתברות המקיימים $\mathbb{E}[Y], \mathbb{E}[X] \in \mathbb{R}\cup\{\infty\}$.
אזי:

1. חיוביות: אם $X \overset{\text{a.s}}{\geq} 0$ אז $\mathbb{E}[X] \overset{\text{a.s}}{\geq} 0$, ואם $X \overset{\text{a.s}}{>} 0$ אז $\mathbb{E}[X] > 0$.
2. לינאריות: $\mathbb{E}[aX+bY] = a\cdot\mathbb{E}[X] + b\cdot\mathbb{E}[Y]$ לכל $a,b\in\mathbb{R}$.
3. מונוטוניות: אם $\mathbb{P}(X\geq Y)=1$ אז $\mathbb{E}[X]\geq\mathbb{E}[Y]$.

טענה (תכונות השונות, תרגיל): יהי X משתנה מקרי רציף בהחלט בעל שונות סופית, ויהי $a\in\mathbb{R}$. אזי:

1. $\text{Var}\left(X\right)\geq0$ ושוויון מתקיים אם ורק אם X קבוע.
2. $\text{Var}\left(X+a\right)=\text{Var}\left(X\right)$.
3. $\text{Var}\left(aX\right)=a^{2}\text{Var}\left(X\right)$.
4. אם Y בעל שונות סופית ובלתי תלוי ב X אז $\text{Var}\left(X+Y\right)=\text{Var}\left(X\right)+\text{Var}\left(Y\right)$.

דוגמה: שוברים מקל באורך מטר לשניים בנקודה מקרית X אשר נבחרת באופן אחיד כלומר לפי הצפיפות $f_{X}=\mathbb{1}_{[0,1]}$. לאחר מכן בונים עפיפון בצורת דלתון שאלכסוניו הם שני שברי המקל. מהי תוחלת שטח העפיפון (במטרים רבועים)? מהי שונותו?

שטח דלתון הוא מחצית מכפלת אלכסוניו. שטח העפיפון אם כן הוא $Y=\frac{X(1-X)}{2}$. נקבל:

$$\mathbb{E}\left[Y\right]=\int\limits_{-\infty}^{\infty}\frac{x\left(1-x\right)}{2}\cdot\mathbb{1}_{[0,1]}\left(x\right)dx=\int\limits_{0}^{1}\frac{x\left(1-x\right)}{2}dx=\left.\frac{x^{2}}{4}-\frac{x^{3}}{6}\right|_{0}^{1}=\frac{1}{12}$$

כעת:

$$\text{Var}\left(Y\right)=\mathbb{E}\left[Y^{2}\right]-\mathbb{E}\left[Y\right]^{2}=\frac{1}{4}\int\limits_{-\infty}^{\infty}x^{2}\left(1-x\right)^{2}\cdot\mathbb{1}_{[0,1]}\left(x\right)dx-\frac{1}{144}$$

$$=\frac{1}{4}\int\limits_{0}^{1}x^{2}-2x^{3}+x^{4}dx-\frac{1}{144}=\frac{1}{720}$$

התפלגויות רציפות חשובות

הגדרה (התפלגות אחידה): יהי $[a,b]\subseteq\mathbb{R}$ קטע. נאמר שלמשתנה מקרי X התפלגות אחידה על $[a,b]$ ונכתוב $X\sim\text{Unif}\left([a,b]\right)$ אם צפיפותו היא:

$$f_{X}\left(x\right)=\frac{\mathbb{1}_{[a,b]}}{b-a}$$

להלן מספר דוגמאות מחיי היום-יום: (1) הזווית בה יפול סביבון מתפלגת אחיד על $[0,2\pi]$ (2) בבדיקת אדם מקרי, אחוז האנשים הגבוהים ממנו מתפלג אחיד על $[0,1]*$ (3) הזמן בין הגעתנו לצומת לזמן חילוף הרמזור הבא מתפלג בקירוב אחיד.

אבחנה: יהי $X\sim\text{Unif}\left([a,b]\right)$ משתנה מקרי בעל התפלגות אחידה, אזי מתקיים:

$$F_{X}\left(t\right)=\begin{cases}0 & t<a\\ \frac{t-a}{b-a} & a\leq t\leq b\\ 1 & t>b\end{cases},\mathbb{E}\left[X\right]=\frac{a+b}{2},\text{Var}\left(X\right)=\frac{\left(b-a\right)^{2}}{12}$$

הוכחה: נחשב:

$$F_{X}\left(t\right)=\int\limits_{-\infty}^{t}\frac{\mathbb{1}_{[a,b]}}{b-a}dx=\begin{cases}0 & t<a\\ \frac{t-a}{b-a} & a\leq t\leq b\\ 1 & t>b\end{cases}$$

$$\mathbb{E}\left[X\right]=\int\limits_{a}^{b}\frac{x}{b-a}dx=\left.\frac{x^{2}}{2\left(b-a\right)}\right|_{a}^{b}=\frac{a+b}{2}$$

$$\text{Var}\left(X\right)=\mathbb{E}\left[X^{2}\right]-\mathbb{E}\left[X^{2}\right]=\int\limits_{a}^{b}\frac{x^{2}dx}{b-a}-\left(\frac{a+b}{2}\right)^{2}=\frac{b^{3}-a^{3}}{3\left(b-a\right)}-\frac{\left(a+b\right)^{2}}{2}=\frac{\left(b-a\right)^{2}}{12}$$

הגדרה (התפלגות מעריכית): נאמר שלמשתנה מקרי X התפלגות מעריכית עם פרמטר λ עבור $\lambda > 0$, ונכתוב $X \sim \text{Exp}\left(\lambda\right)$, אם צפיפותו היא:

$$f_X\left(x\right) = \mathbb{1}_{[0,\infty)}\left(x\right) \cdot \lambda e^{-\lambda x}$$

ההתפלגות המעריכית היא המקבילה הרציפה של ההתפלגות הגיאומטרית. היא מתארת את הזמן שחולף להתרחשות בתנאי חוסר זכרון, אך גם התפלגויות פיזיקליות נוספות. דוגמאות: (1) הזמן עד לכניסת לקוח לחנות (2) הזמן בין לידות של תינוקות במדינה (3) הזמן עד להתפרקותו של אטום אורניום.

אבחנה: יהי $X \sim \text{Exp}\left(\lambda\right)$. אזי:

$$F_X\left(t\right) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases}, \mathbb{E}\left[X\right] = \frac{1}{\lambda}, \text{Var}\left(x\right) = \frac{1}{\lambda^2}$$

הוכחה: נחשב:

$$F_X\left(t\right) = \int_0^t \lambda e^{-\lambda x} \mathbb{1}_{[0,\infty)}\left(x\right) dx = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

$$\mathbb{E}\left[X\right] = \int_0^\infty \lambda x e^{-\lambda x} dx = \lim_{s \to \infty} -xe^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x}\Big|_0^s = \frac{1}{\lambda}$$

$$\text{Var}\left(X\right) = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = \int_0^\infty \lambda x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} = \left\{ \begin{array}{cc} u = x^2 & dv = \lambda e^{-\lambda x} \\ du = 2xdx & v = -e^{-\lambda x} \end{array} \right\}$$

$$= \lim_{s \to \infty} \left(-x^2 e^{-\lambda x}\Big|_0^s + 2\int_0^s xe^{-\lambda x} dx \right) - \frac{1}{\lambda^2} = \lim_{s \to \infty} -x^2 e^{-\lambda x} - \frac{2}{\lambda} xe^{-\lambda x} - \frac{2}{\lambda^2} e^{-\lambda x}\Big|_0^s - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

טענה (חוסר זכרון): יהי $X \sim \text{Exp}\left(\lambda\right)$. אזי המשתנה המקרי $Y = \left(X - x_0 \mid X > x_0\right)$ מתפלג $\text{Exp}\left(\lambda\right)$.

הוכחה: נחשב:

$$\overline{F}_Y\left(t\right) = \mathbb{P}\left(X - x_0 > t \mid X > x_0\right) = \frac{\overline{F}_X\left(t + x_0\right)}{\overline{F}_X\left(x_0\right)} = \frac{e^{-\lambda(t+x_0)}}{e^{-\lambda x_0}} = e^{-\lambda t}$$

זו פונקצית התפלגות שיורית של משתנה מעריכי עם פרמטר λ והיא קובעת את התפלגותו.

הרצאה 19

המשך התפלגויות רציפות חשובות

הגדרה (התפלגות נורמלית): נאמר שלמשתנה מקרי X התפלגות נורמלית עם תוחלת μ ושונות σ^2 ונכתוב $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$, אם צפיפותו היא:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

כאשר $X \sim \mathrm{N}(0,1)$ נאמר כי X נורמלי תקני. ההתפלגות הנורמלית דומה בצורתה להתפלגות בינומית עבור מספר רב של ניסויים והיא מתארת משתנים מקריים אשר נקבעים כסכום של מספר רב של גורמים שהנם בלתי תלויים בקירוב. דוגמאות למשתנים מקריים המתפלגים בקירוב נורמלי הנן: (1) גובהו של אדם מקרי באוכלוסיה (2) משך דקה בשעונים אנלוגים שונים (3) כמות הגשם במ"מ שירדה בשנה מסוימת.

הגדרה (פונקצית התפלגות מצטברת של התפלגות מקרית): נסמן מעתה את פונקציית ההתפלגות המצטברת של משתנה מקרי $X \sim \mathrm{N}(0,1)$ על ידי $\Phi(x) = \mathbb{P}(X \leq x)$.

הערה: החישוב $\int\limits_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$, מכונה האינטגרל הגאוסי או אינטגרל אוילר-פואסון. שיטת החישוב פותחה על ידי דה-מואבר, והחישוב עצמו חושב לראשונה על ידי לפלאס ב-1774 אך זכה להתפרסם ב-1809 על ידי גאוס שפיתח אותו באופן בלתי תלוי. כאן נביא שיטה לחשב אינטגרל זה באמצעות העברת הבעיה למישור ושימוש בקואורדינטות קוטביות. נסמן $I = \int\limits_{-\infty}^{\infty} e^{-x^2} dx$. אזי:

$$I^2 = \left(\int\limits_{-\infty}^{\infty} e^{-x^2} dx\right)^2 = \left(\int\limits_{-\infty}^{\infty} e^{-x^2} dx\right)\left(\int\limits_{-\infty}^{\infty} e^{-y^2} dy\right) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-\left(x^2+y^2\right)} dxdy$$

$$= \left\{\begin{array}{c} x = r\cos\theta \quad y = r\sin\theta \\ dxdy = rdrd\theta \end{array}\right\} = \int\limits_{0}^{2\pi}\int\limits_{0}^{\infty} e^{-r^2} rdrd\theta = \int\limits_{0}^{2\pi} \frac{d\theta}{2} = \pi$$

אבחנה: יהי $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$. אזי, לכל $\alpha > 0, \beta \in \mathbb{R}$ מתקיים $\alpha X + \beta \sim \mathrm{N}\left(\alpha\mu + \beta, \alpha^2\sigma^2\right)$.

אבחנה: יהי $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$. אזי:

$$F_X(t) = \Phi\left(\frac{t-\mu}{\sigma}\right), \mathbb{E}[X] = \mu, \mathrm{Var}(X) = \sigma$$

הוכחה: מספיק להוכיח שלמשתנה נורמלי תקני אכן יש תוחלת 0 ושונות 1, מפני שכל משתנה נורמלי $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$ נוכל לבטא כהעתקה לינארית של משתנה מקרי נורמלי תקני $Y \sim \mathrm{N}(0,1)$, על ידי $X = \sigma Y + \mu$.

כדי לוודא שהתוחלת מוגדרת היטב, נחשב ונקבל:

$$\mathbb{E}[|Y|] = \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} |s| e^{-\frac{1}{2}s^2} ds = 2\int\limits_{0}^{\infty} \frac{1}{\sqrt{2\pi}} |s| e^{-\frac{1}{2}s^2} ds = \left\{\begin{array}{c} t = \frac{s^2}{2} \\ dt = sds \end{array}\right\} = \int\limits_{0}^{\infty} \frac{e^{-t}}{\sqrt{2\pi}} dt = \frac{1}{\sqrt{2\pi}}$$

כעת:

$$\mathbb{E}[Y] = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} t \cdot e^{-\frac{t^2}{2}} dt = 0$$

כי זו פונקציה אי זוגית. לבסוף נרצה לחשב את השונות:

$$\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \cancel{\mathbb{E}[X]} = \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx = \left\{\begin{array}{cc} u = x & dv = xe^{-\frac{x^2}{2}} \\ du = dx & v = -e^{-\frac{x^2}{2}} \end{array}\right\}$$

$$= -\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}} = 0 + \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1$$

התפלגות משותפת של משתנים מקריים רציפים

הגדרה(התפלגות משותפת): נאמר כי לשני משתנים מקריים X, Y מעל מרחב הסתברות משותף $(\Omega, \mathcal{F}, \mathbb{P})$ יש צפיפות משותפת אם קיימת פונקציה אינטגרבילית $f_{X,Y}(x,y) : \mathbb{R}^2 \longrightarrow [0,\infty)$ המקיימת לכל שני קטעים $[a,b], [c,d]$ כי:

$$\mathbb{P}(X \in [a,b], Y \in [c,d]) = \int\limits_a^b \int\limits_c^d f_{X,Y}(x,y)\, dydx$$

הפונקציה $f_{X,Y}$ נקראת צפיפות משותפת של X ו- Y. מהגדרת ההתפלגות המשותפת נובע:

$$\mathbb{P}(X \le s, Y \le t) = \int\limits_{-\infty}^{s} \int\limits_{-\infty}^{t} f_{X,Y}(x,y)\, dydx$$

אבחנה: יהיו X, Y משתנים מקריים רציפים בהחלט בעלי צפיפויות f_X ו- f_Y בהתאמה. אזי X ו- Y בלתי תלויים אם ורק אם קיימת להם צפיפות משותפת המקיימת

$$f_{X,Y} = f_X \cdot f_Y$$

הוכחה: $(\Longrightarrow)$: נניח כי $f_{X,Y} = f_X \cdot f_Y$ ונקבל לפי הגדרה כי:

$$\mathbb{P}(X \in I, Y \in J) = \int_I \int_J f_{X,Y}(x,y)\, dxdy = \int_I \int_J f_X(x) f_Y(y)\, dxdy$$

$$= \int_J f_X(x)\, dx \cdot \int_I f_Y(y)\, dy = \mathbb{P}(X \in I) \cdot \mathbb{P}(Y \in J)$$

לכן, X ו- Y בלתי תלויים. $(\Longleftarrow)$: נניח ש X ו- Y בלתי תלויים ונקבל לפי אותה משוואה כי $f_{X,Y} = f_X \cdot f_Y$ היא צפיפות משותפת שלהם.

דוגמה: נתונים משתנים מקריים X ו- Y בעלי צפיפות משותפת $f_{X,Y}(x,y) = (x+y) \cdot \mathbb{1}_{[0,1]^2}(x,y)$. נחשב את ההסתברויות $\mathbb{P}\left(X \ge \frac{1}{2}, Y \ge \frac{1}{2}\right)$ ו- $\mathbb{P}(X + Y \le 1)$.

ראשית, נבחין שאכן מדובר בפונקציית צפיפות:

$$\iint_{\mathbb{R}^2} f(x,y)\, dydx = \iint_{[0,1]^2} f(x,y)\, dydx = \int\limits_0^1 \left(\int\limits_0^1 (x+y)\, dy \right) dx$$

$$= \int\limits_0^1 xy + \frac{y^2}{2} \Big|_{y=0}^{y=1} dx = \int\limits_0^1 \left(x + \frac{1}{2} \right) dx = \frac{x^2 + x}{2} \Big|_{x=0}^{x=1} = 1$$

כעת:

$$\mathbb{P}\left(X \ge \frac{1}{2}, Y \ge \frac{1}{2}\right) = \iint_{\left[\frac{1}{2},\infty\right)^2} f(x,y)\, dtydx = \int\limits_{\frac{1}{2}}^{\infty} \left(\int\limits_{\frac{1}{2}}^{\infty} f_{X,Y}(x,y)\, dy \right) dx = \int\limits_{\frac{1}{2}}^{1} \left(\int\limits_{\frac{1}{2}}^{1} (x+y)\, dy \right) dx$$

$$= \int_{\frac{1}{2}}^{1} xy + \frac{y^2}{2}\Big|_{y=\frac{1}{2}}^{y=1} dx = \int_{\frac{1}{2}}^{1} \left(x + \frac{1}{2}\right) - \left(\frac{x}{2} + \frac{1}{8}\right) dx = \int_{\frac{1}{2}}^{1} \frac{x}{2} + \frac{3}{8} dx = \frac{x^2}{4} + \frac{3x}{8}\Big|_{x=\frac{1}{2}}^{x=1} = \frac{3}{8}$$

אם נסמן $A = \{(x,y) \mid x,y \geq 0, x+y \leq 1\}$ נקבל:

$$\mathbb{P}(X+Y \leq 1) = \iint_A (x+y)\, dydx = \int_0^1 \left(\int_0^{1-x} (x+y)\, dy\right) dx = \int_0^1 xy + \frac{y^2}{2}\Big|_{y=0}^{y=1-x} dx$$

$$= \int_0^1 x(1-x) + \frac{(1-x)^2}{2} dx = \int_0^1 \frac{1}{2} - \frac{x^2}{2} dx = \frac{x}{2} - \frac{x^3}{6}\Big|_{x=0}^{x=1} = \frac{1}{3}$$

צפיפות שולית

אבחנה: יהיו X,Y משתנים מקריים בעלי צפיפות משותפת $f_{X,Y}$ אז X ו- Y רציפים בהחלט ומתקיים כי

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy, \;\; f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$

הן צפיפויות של X ו- Y בהתאמה. התפלגויותיהם של X ו- Y נקראות ההתפלגויות השוליות של X ו- Y.

דוגמה: נתונים משתנים מקריים X ו- Y בעלי צפיפות משותפת $f_{X,Y}(x,y) = (x+y)\cdot \mathbb{1}_{[0,1]^2}(x,y)$. נחשב את התפלגות X.
עבור $0 \leq x \leq 1$:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy = \int_0^1 (x+y)\, dy = xy + \frac{y^2}{2}\Big|_{y=0}^{y=1} = x + \frac{1}{2}$$

עבור ערכי x אחרים, $f_X(x) = 0$.

הרצאה 20

צפיפות של סכום משתנים מקריים רציפים בהחלט

טענה (נוסחת הקונבולוציה): יהיו X ו- Y משתנים מקריים בעלי צפיפות משותפת ויהי סכומם $Z = X + Y$. אזי Z רציף בהחלט וצפיפותו:

$$f_Z\left(z\right) = \int\limits_{-\infty}^{\infty} f_X\left(x\right) \cdot f_Y\left(z-x\right) dx$$

הוכחה: תחילה, נבחין כי:

$$\mathbb{P}\left(X \leq a, Z \leq b\right) = \mathbb{P}\left(X \leq a, X+Y \leq b\right) = \int\limits_{-\infty}^{a} \int\limits_{-\infty}^{b-x} f_{X,Y}\left(x,y\right) dydx = \int\limits_{-\infty}^{a} \int\limits_{-\infty}^{b} f_{X,Y}\left(x,z-x\right) dzdx$$

על כן, קיבלנו ש $f_{X,Z}\left(x,z\right) = f_{X,Y}\left(x,z-x\right) = f_X\left(x\right) \cdot f_Y\left(z-x\right)$, כאשר המעבר האחרון נובע מהאי-תלות. כעת:

$$f_Z\left(z\right) = \int\limits_{-\infty}^{\infty} f_{X,Z}\left(x,z\right) dx = \int\limits_{-\infty}^{\infty} f_X\left(x\right) \cdot f_Y\left(z-x\right) dx$$

דוגמה: נתונים משתנים X ו- Y בלתי תלויים המתפלגים מעריכית עם פרמטר 1. נחשב את התפלגות $Z = X + Y$. עבור $z \geq 0$:

$$f_Z\left(z\right) = \int\limits_{-\infty}^{\infty} f_X\left(x\right) \cdot f_Y\left(z-x\right) dx = \int\limits_{0}^{z} e^{-x} \cdot e^{x-z} dx = z \cdot e^{-z}$$

עבור $z < 0$ הצפיפות f_Z מתאפסת.

צפיפות מותנית

הגדרה(צפיפות מותנית): יהיו X,Y משתנים מקריים רציפים בהחלט בעלי צפיפות משותפת. נסמן את הצפיפות של X בהינתן Y על ידי:

$$f_{X|Y=y}\left(x\right) = \frac{f_{X,Y}\left(x,y\right)}{f_Y\left(y\right)}$$

טענה (נוסחת ההסתברות השלמה הרציפה): לכל X,Y בעלי צפיפות משותפת מתקיים:

$$f_X\left(x\right) = \int\limits_{-\infty}^{\infty} f_Y\left(y\right) \cdot f_{X|Y=y}\left(x\right) dy$$

הוכחה: נבטא את $f_X\left(x\right)$ כצפיפות שולית, ונציב את הגדרת הצפיפות המותנית:

$$f_X\left(x\right) = \int\limits_{-\infty}^{\infty} f_{X,Y}\left(x,y\right) dy = \int\limits_{-\infty}^{\infty} f_Y\left(y\right) \cdot f_{X|Y=y}\left(x\right) dy$$

דוגמה: יהי X משתנה מקרי המתפלג אחיד בקטע $[0,1]$. יהי Y משתנה מקרי שבהנתן ש $X=x$ מתפלג אחיד בקטע $[0,x]$. נרצה לחשב צפיפות Y.

לכל $0<y\le 1$ נקבל:

$$f_Y(y)=\int\limits_{-\infty}^{\infty} f_X(x)\cdot f_{Y|X=x}(y)\,dx=\int\limits_0^1 \frac{1}{x}\cdot \mathbb{1}_{[0,x]}(y)\,dx=\int\limits_0^1 \frac{1}{x}\cdot \mathbb{1}_{[0,x]}(y)\,dx=\int\limits_0^y \frac{1}{x}dx=-\log(y)$$

אבחנה (כלל בייס רציף): לכל X,Y בעלי צפיפות משותפת מתקיים:

$$f_{X|Y=y}(x)\cdot f_Y(y)=f_{Y|X=x}(y)\cdot f_X(x)$$

אבחנה (תנאי לאי תלות): מתשנים מקריים רציפים בהחלט X ו- Y בלתי תלויים אם ורק אם:

$$f_{X|Y=y}(x)=f_X(x)$$

לכל y עבורו $f_Y(y)\neq 0$.

מומנטים גבוהים

הגדרה: יהי X משתנה מקרי. המומנט מסדר k של X מוגדר בתור $m_k(X)=\mathbb{E}\left[X^k\right]$, כאשר תוחלת זו מוגדרת היטב.

הגדרה (קמירות): פונקציה $f:D\longrightarrow\mathbb{R}$ תקרא קמורה, אם לכל $x_0\in D$ יש $\beta\in\mathbb{R}$ כך שמתקיים $f(x)\ge f(x_0)+\beta\cdot(x-x_0)$ לכל $x\in D$.

משפט (אי שיוויון ינסן): יהי X משתנה מקרי בעל תוחלת ותהי f פונקציה קמורה, אזי $\mathbb{E}[f(X)]\ge f(\mathbb{E}[X])$.

הוכחה: נסמן $\mu=\mathbb{E}[X]$, ולפי הגדרת הקמירות קיים β כך ש $f(x)\ge f(\mu)+\beta\cdot(x-\mu)$. כעת, לפי מונוטוניות ולינאריות התוחלת:

$$\mathbb{E}[f(X)]\ge\mathbb{E}[f(\mu)]+\beta\cdot(\mathbb{E}[X]-\mathbb{E}[\mu])=f(\mu)+\beta\cdot(\mu-\mu)=f(\mathbb{E}[X])$$

טענה: יהי X משתנה מקרי. אם $m_k(X)$ סופי אז גם $m_{k-1}(X)$ סופי.

הוכחה: לפי הגדרה, לכל $k\in\mathbb{N}$ קיום $m_k(X)$ שקול לקיום $\mathbb{E}\left[|X|^k\right]$. נפעיל את אי שיוויון ינסן על הפונקציה הקמורה $f(x)=x^{\frac{k}{k-1}}$ והמשתנה המקרי $|X|^{k-1}$ ונקבל:

$$\mathbb{E}\left[|X|^k\right]\ge\left(\mathbb{E}\left[|X|^k\right]\right)^{\frac{k}{k-1}}$$

לכן, אם $m_k(X)$ סופי, גם $m_{k-1}(X)$ סופי, כנדרש.

פונקציה יוצרת מומנטים

הגדרה: יהי X משתנה מקרי. הפונקציה $M_X\left(t\right)$ הנתונה על ידי $M_X\left(t\right)=\mathbb{E}\left[e^{tX}\right]$ מכונה הפונקציה יוצרת מומנטים של X.

טענה: יהיו X ו- Y משתנים מקריים בלתי תלויים ויהי $Z=X+Y$. אזי:

$$M_Z\left(t\right)=M_X\left(t\right)\cdot M_Y\left(t\right)$$

הוכחה: נחשב, ונעזר בכפליות התוחלת של משתנים בלתי תלויים:

$$M_{X+Y}\left(y\right)=\mathbb{E}\left[e^{t\left(X+Y\right)}\right]=\mathbb{E}\left[e^{tX}\cdot e^{tY}\right]=\mathbb{E}\left[e^{tX}\right]\cdot\mathbb{E}\left[e^{tY}\right]=M_X\left(t\right)\cdot M_Y\left(t\right)$$

דוגמה: נחשב את הפונקציה יוצרת מומנטים של מספר משתנים מוכרים.

(1) פונקציה יוצרת מומנטים של משתנה ברנולי $X\sim\text{Ber}\left(p\right)$:

$$M_X\left(t\right)=\mathbb{E}\left[e^{tX}\right]=e^t\cdot p+\left(1-p\right)$$

(2) פונקציה יוצרת מומנטים של משתנה בינומי $X\sim\text{Bin}\left(n,p\right)$:

$$M_X\left(t\right)=\mathbb{E}\left[e^{tX}\right]=\sum_{k=0}^{n}e^{tk}\cdot\binom{n}{k}\cdot p^k\cdot\left(1-p\right)^{n-k}=\sum_{k=0}^{n}\binom{n}{k}\cdot\left(e^t\cdot p\right)^k\cdot\left(1-p\right)^{n-k}=\left(e^t\cdot p+\left(1-p\right)\right)^n$$

נציין שניתן להגיע לאותה תוצאה על ידי הצגת X כסכום משתני ברנולי בלתי תלויים, ושימוש בטענה הקודמת.

(3) פונקציה יוצרת מומנטים של משתנה אקספוננציאלי $X\sim\text{Exp}\left(\lambda\right)$:

$$M_X\left(t\right)=\int_0^\infty e^{t\cdot x}\cdot\lambda e^{-\lambda x}dx=\lambda\cdot\int_0^\infty e^{\left(t-\lambda\right)\cdot x}dx=\left.\frac{\lambda\cdot e^{\left(t-\lambda\right)\cdot x}}{t-\lambda}\right|_0^\infty=\frac{\lambda}{t-\lambda}$$

והיא מוגדרת רק עבור $t<\lambda$.

(4) פונקציה יוצרת מומנטים של משתנה גיאומטרי $X\sim\text{Geo}\left(p\right)$:

$$M_X\left(t\right)=\sum_{k=1}^{\infty}e^{t\cdot k}\cdot\left(1-p\right)^{k-1}\cdot p=pe^t\cdot\sum_{k=1}^{\infty}\left(e^t-pe^t\right)^{k-1}=\frac{p\cdot e^t}{1-\left(e^t-p\cdot e^t\right)}=\frac{p}{e^{-t}+p-1}$$

והיא מוגדרת רק עבור $t<-\log\left(1-p\right)$.

הרצאה 21

אי שיוויון צ'רנוף

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

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

הוכחה: נציב באי שיוויון מרקוב את המשתנה המקרי החיובי e^{tX} ונקבל:

$$\mathbb{P}\left(X\geq a\right)=\mathbb{P}\left(e^{tX}\geq e^{ta}\right)\leq\frac{\mathbb{E}\left[e^{tX}\right]}{e^{ta}}=M_X\left(t\right)\cdot e^{-ta}$$

דוגמה: מחלקים באקראי n כדורים ל-n תאים באופן אחיד ובלתי תלוי. נחסום מלמעלה את כמות הכדורים בתא העמוס ביותר. נגדיר משתנים S_i מספר הכדורים בתא ה-i. כעת, לכל $a>0$ מתקיים:

$$\mathbb{P}\left(\max_{i\in[n]}S_i<a\right)=1-\mathbb{P}\left(\max_{i\in[n]}S_i\geq a\right)\overset{\text{חסם איחוד}}{\geq}1-\sum_{i=1}^{n}\mathbb{P}\left(S_i\geq a\right)\overset{\text{סימטריה}}{=}1-n\cdot\mathbb{P}\left(S_1\geq a\right)$$

נסמן ב- X_i אינדקטור שהכדור ה-i נפל בתא הראשון. אלו משתני ברנולי בלתי-תלוים, לכן $S_1\sim\text{Bin}\left(n,\frac{1}{n}\right)$. נחשב ונקבל:

$$\mathbb{P}\left(S_1\geq a\right)\leq e^{-ta}\cdot M_{S_1}\left(t\right)=e^{-ta}\cdot\left(1+\frac{e^t-1}{n}\right)^n\leq e^{e^t-1}\cdot e^{-ta}=e^{e^t-a\cdot t}$$

המעבר נובע מכך ש $\left(1+\frac{x}{n}\right)^n\underset{n\to\infty}{\longrightarrow}e^x$, וזו פונקציה עולה. נגזור לפי t, ונקבל ש $t=\log a$ ממזער את החסם. כלומר:

$$\mathbb{P}\left(S_1\geq a\right)\leq e^{a-a\log(a)-1}<e^{-a(\log(a)-1)}$$

נרצה למצוא a עבורו $n\cdot e^{-a(\log(a)-1)}\underset{n\to\infty}{\longrightarrow}0$. למשל, $a=c\cdot\frac{\log n}{\log\log n}$ עובד, עבור $c>1$. אכן, נציב במשוואה הראשונה ונקבל:

$$\mathbb{P}\left(\max_{i\in[n]}S_i<c\cdot\frac{\log n}{\log\log n}\right)\geq1-\frac{n}{\omega\left(n\right)}\longrightarrow1$$

כלומר, לכל $c>1$ כמות הכדורים בתא העמוס ביותר הוא לכל היותר $c\cdot\frac{\log n}{\log\log n}$, בהסתברות השואפת ל-1.

אי שיוויון הופדינג

טענה (הלמה של הופדינג): יהי X משתנה מקרי המקיים $|X|\overset{\text{a.s}}{\leq}1$ וגם $\mathbb{E}\left[X\right]=0$. אזי לכל $t\in\mathbb{R}$ עבורו $M_X\left(t\right)$ קיים מתקיים:

$$M_X\left(t\right)\leq e^{\frac{t^2}{2}}$$

הוכחה: יהי נתון t. הפונקציה e^{tx} (כפונקציה של x), היא בעלת נגזרת שנייה חיובית ולכן קמורה. בפרט, אם נסמן ב $L\left(x\right)$ את הישר העובר דרך $\left(-1,e^{-t}\right)$ ו- $\left(1,e^t\right)$, אז לכל $x\in\left[-1,1\right]$ מתקיים:

$$e^{tx}\leq L\left(x\right)$$

משוואת הישר היא $L\left(x\right)=\frac{e^t+e^{-t}}{2}+x\cdot\frac{e^t-e^{-t}}{2}$. כעת, ממונוטוניות התוחלת מתקיים:

$$M_X\left(t\right)=\mathbb{E}\left[e^{tX}\right]\leq\mathbb{E}\left[L\left(X\right)\right]=\frac{e^t+e^{-t}}{2}+\frac{e^t-e^{-t}}{2}\cdot\mathbb{E}\left[X\right]$$

ניעזר בעובדה ש $\frac{e^t+e^{-t}}{2} \leq e^{\frac{t^2}{2}}$ (ניתן להשתכנע בה ע"י חישוב טורי הטיילור של שני האגפים), ונקבל:

$$M_X\left(t\right) \leq \frac{e^t+e^{-t}}{2} + \frac{e^t-e^{-t}}{2} \cdot 0 \leq e^{\frac{t^2}{2}}$$

משפט (אי שיוויון הופדינג): תהי $\left\{X_k\right\}_{k=1}^n$ סדרת משתנים מקריים בלתי תלויים, ובעלי תוחלת אפס, שמקיימים $\left|X_k\right| \overset{\text{a.s}}{\leq} 1$ לכל $k \in [n]$. אזי לכל $a > 0$ מתקיים:

$$\mathbb{P}\left(\sum_{k=1}^n X_k \geq a\right) \leq e^{-\frac{a^2}{2n}}$$

הוכחה: נסמן $X = \sum\limits_{k=1}^n X_k$. מאי תלות המשתנים, והלמה הקודמת:

$$M_X\left(t\right) = \prod_{k=1}^n M_{X_k}\left(t\right) \leq \prod_{k=1}^n e^{\frac{t^2}{2}} = e^{\frac{n}{2} \cdot t^2}$$

כעת, מאי שיוויון צ'רנוף, לכל $t > 0$ מתקיים:

$$\mathbb{P}\left(X \geq a\right) \leq e^{\frac{n}{2} \cdot t^2 - a \cdot t}$$

נגזור לפי t, ונקבל ש $t = \frac{a}{n}$ ממזער את החסם. נקבל:

$$\mathbb{P}\left(X \geq a\right) \leq e^{\frac{n}{2} \cdot \left(\frac{a}{n}\right)^2 - a \cdot \left(\frac{a}{n}\right)} = e^{-\frac{a^2}{2n}}$$

דוגמה: מיליון מטבעות הוגנים מוטלים. נחסום את הסיכוי שכמות המטבעות שתוצאתם עץ תהיינה בין $495,000$ ל- $505,000$.

נסמן ב- X_i משתנה מקרי המקבל את הערך 1 אם בהטלה ה-i התקבלה תוצאה של עץ ו- -1 אחרת. בנוסף, נגדיר $X = \sum\limits_{i=1}^{10^6} X_i$.

מטרתנו לחסום ההסתברות ש $\left|X\right| < 10^4$. נפעיל את אי שיוויון הופדינג ונקבל:

$$\mathbb{P}\left(-10^4 < X < 10^4\right) \overset{\text{סימטריה}}{=} 1 - 2 \cdot \mathbb{P}\left(X \geq 10^4\right) \geq 1 - 2 \cdot e^{-\frac{10^8}{2 \cdot 10^6}} = 1 - 2e^{-50}$$

הערה (הכללה של אי שיוויון הופדינג): תהי $\left\{X_k\right\}_{k=1}^n$ סדרת משתנים מקריים בלתי תלויים, שמקיימים $\left|X_k - \mathbb{E}\left[X_k\right]\right| \overset{\text{a.s}}{\leq} M$ לכל $k \in [n]$. נסמן $X = \sum\limits_{k=1}^n X_k$. אזי לכל $a > 0$ מתקיים:

$$\mathbb{P}\left(\sum_{k=1}^n X_k - \mathbb{E}\left[X\right] \geq a\right) \leq e^{-\frac{a^2}{2nM^2}},\ \mathbb{P}\left(\left|\sum_{k=1}^n X_k - \mathbb{E}\left[X\right]\right| \geq a\right) \leq 2 \cdot e^{-\frac{a^2}{2nM^2}}$$

דוגמה: יהי $X \sim \text{Bin}\left(N, p\right)$, ויהיו $Y_k \sim \text{Ber}\left(p\right)$ עבור $k \in [N]$. נשים לב כי $\mathbb{E}\left[Y_k\right] = p$ וכי $\left|Y_k - p\right| \leq \max\left\{p, 1-p\right\}$, לכן $\left\{Y_k\right\}$ מקיימים תנאי אי שיוויון הופדינג הכללי. נקבל:

$$\mathbb{P}\left(X - Np \geq a\right) \leq e^{\frac{a^2}{N \cdot \left(\max\left\{1, 1-p\right\}\right)^2}}$$

הרצאה 22

התכנסות התפלגויות

הגדרה (התכנסות בהסתברות): תהי $\{X_n\}_{n=1}^{\infty}$ סדרת משתנים מקריים, לא בהכרח על אותו מרחב הסתברות, ויהי X משתנה מקרי. נאמר כי $\{X_n\}$ מתכנסת בהתפלגות ל-X, ונסמן $X_n \xrightarrow{d} X$, אם לכל a שהיא נקודת רציפות של F_X מתקיים:

$$F_{X_n}(a) \underset{n\to\infty}{\longrightarrow} F_X(a)$$

דוגמה: תהי $X_n \overset{\text{a.s}}{=} c_n$ סדרת משתנים מקריים קבועים, כאשר $c_n \longrightarrow 0$, ויהי $X \overset{\text{a.s}}{=} 0$. פונקציות ההתפלגות המצטברות של הסדרה הן:

$$F_{X_n}(a) = \begin{cases} 0 & a < c_n \\ 1 & a \geq c_n \end{cases}$$

ואומנם, $\lim_{n\to\infty} F_{X_n}(a) = F_X(a)$ בכל נקודה, מלבד 0 - נקודת אי הרציפות היחידה של F_X. ההגדרה שנתנו להתכנסות נועדה לאפשר גם להתכנסות כזו להקרא התכנסות בהתפלגות, ואכן לפי הגדרה זו $X_n \xrightarrow{d} X$.

דוגמה: תהי $X_n \sim \text{Unif}\left(\left[-\frac{1}{n}, \frac{1}{n}\right]\right)$ סדרת משתנים מקריים, אז $X_n \xrightarrow{\text{d}} 0$. פונקציות ההתפלגות המצטברת של הסדרה הן:

$$F_{X_n}(t) = \begin{cases} 0 & t < -\frac{1}{n} \\ \frac{t+\frac{1}{n}}{\frac{2}{n}} & -\frac{1}{n} \leq t \leq \frac{1}{n} \\ 1 & t > \frac{1}{n} \end{cases}$$

ואמנם,

$$\lim_{n\to\infty} F_{X_n} = \begin{cases} 0 & t < 0 \\ \frac{1}{2} & t = 0 \\ 1 & t > 0 \end{cases}$$

ואכן, לכל נקודת רציפות של F_X - כלומר $t \neq 0$, מתקיים $F_{X_n}(a) \underset{n\to\infty}{\longrightarrow} F_X(a)$. באופן דומה, גם $X_n \sim \text{Unif}\left(\left[0, \frac{1}{n}\right]\right)$ וגם $X_n \sim \text{Unif}\left(\left[\frac{1}{n}, \frac{2}{n}\right]\right)$ מתכנסות בהתפלגות ל- $X = 0$.

טענה: תהי $\{X_n\}_{n=1}^{\infty}$ סדרת משתנים מקריים על $\mathbb{N}_0$. אז X_n מתכנסת בהתפלגות ל-X אם ורק אם לכל $k \in \mathbb{N}_0$ מתקיים:

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

הוכחה: ($\Longleftarrow$): תהי $\{X_n\}$ סדרה כנ"ל, ונניח שהיא מתכנסת בהתפלגות ל-X. אזי:

$$\mathbb{P}(X_n = k) = F_{X_n}\left(k + \frac{1}{2}\right) - F_{X_n}\left(k - \frac{1}{2}\right) \underset{n\to\infty}{\longrightarrow} F_X\left(k + \frac{1}{2}\right) - F_X\left(k - \frac{1}{2}\right) = \mathbb{P}(X = k)$$

($\Longrightarrow$): תהי $t \notin \mathbb{N}_0$ נקודת רציפות של F_X, מתקיים: מאחר שהמשתנים נתמכים על הטבעיים.

$$F_{X_n}(t) = \sum_{k=0}^{\lfloor t \rfloor} \mathbb{P}(X_n = k) \underset{n\to\infty}{\longrightarrow} \sum_{k=0}^{\lfloor t \rfloor} \mathbb{P}(X = k) = F_X(t)$$

כנדרש.

דוגמה: תהי $X_n \sim \text{Bin}\left(n, \frac{1}{n}\right)$ סדרת משתנים מקריים. נרצה להוכיח כי $X_n \underset{n\to\infty}{\longrightarrow} X \sim \text{Poi}\left(1\right)$. נעזר בטענה הקודמת: לכל $k \in \mathbb{N}_0$:

$$\mathbb{P}\left(X_n = k\right) = \binom{n}{k} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-k} = \frac{n\left(n-1\right)\cdots\left(n-k+1\right)}{n^k} \cdot \frac{1}{k!} \cdot \left(1 - \frac{1}{n}\right)^{n-k} \underset{n\to\infty}{\longrightarrow} \frac{e^{-1}}{k!} = \mathbb{P}\left(X = k\right)$$

כנדרש.

דוגמה: תהי $X_n \sim \text{Geo}\left(\frac{1}{n}\right)$ סדרת משתנים מקריים, ויהי $X \sim \text{Exp}\left(1\right)$. אזי $Y_n = \frac{1}{n} \cdot X_n \xrightarrow{\text{d}} X$. אכן, יהי נתון $a > 0$:

$$\mathbb{P}\left(Y_n \le a\right) = \sum_{k=1}^{\lfloor n \cdot a \rfloor} \left(1 - \frac{1}{n}\right)^{k-1} \cdot \frac{1}{n} = \frac{1 - \left(1 - \frac{1}{n}\right)^{\lfloor n \cdot a \rfloor}}{1 - \left(1 - \frac{1}{n}\right)} \cdot \frac{1}{n} \underset{n\to\infty}{\longrightarrow} 1 - e^{-a} = F_X\left(a\right)$$

ולפי הגדרה, סיימנו.

דוגמה: תהי סדרת משתנים מקריים אחידים $X_n \sim \text{Unif}\left(\left[-n, n\right]\right)$. אזי $F_{X_n}\left(s\right) = \min\left(1, \max\left(0, \frac{s+n}{2n}\right)\right)$. לכל a מתקיים כי $\lim\limits_{n\to\infty} F_{X_n}\left(a\right) = \frac{1}{2}$ לכן $F_X \equiv \frac{1}{2}$. זו אינה פונקצית התפלגות מצטברת, ולכן מיחידות הגבול נסיק כי X_n אינם מתכנסים בהתפלגות לשום משתנה מקרי.

משפט הגבול המרכזי

משפט הגבול המרכזי קובע שכאשר מספר גדול של גורמים בלתי-תלויים בעלי שונות נסכמים, לאחר נרמול בסטיית התקן של הסכום - סטייתו מהתוחלת שואפת להתפלגות נורמלית סטנדרטית. תופעה זו אוניברסלית - כלומר בחירת ההתפלגות הפרטנית של המודל אינה משפיעה על ההתפלגות הגבולית. תוצאות כאלו זוכות לשימושים רבים בכל תחומי המדע, מפני שהן מאפשרות לנו לדעת את ההתפלגות הגבולית גם מבלי שבנינו מודל מדויק לכל אחד מהמשתנים.

משפט (משפט הגבול המרכזי): תהי $\left\{X_n\right\}_{n=1}^{\infty}$ סדרת משתנים מקריים בלתי תלויים ושווי התפלגות, בעלי תוחלת 0 ושונות 1. אזי:

$$\frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} X_i \xrightarrow{\text{d}} Z$$

כאשר Z הוא משתנה מקרי נורמלי סטנדרטי. באופן שקול, לכל $a \in \mathbb{R}$ מתקיים:

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \cdot \sum_{i=1}^{n} X_i \le a\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{a} e^{-\frac{x^2}{2}}\, dx$$

ניתן להכליל את המשפט לסדרה $\left\{X_n\right\}_{n=1}^{\infty}$ של משתנים מקריים בלתי תלויים ושווי התפלגות, בעלי תוחלת μ ושונות σ^2. עבורה:

$$\frac{1}{\sigma\sqrt{n}} \cdot \sum_{i=1}^{n} \left(X_i - \mu\right) \xrightarrow{\text{d}} Z$$

ההכללה נובע מהצבת $Y_n = \frac{1}{\sigma}\left(X_n - \mu\right)$ במשפט המקורי.

משפט: יהיו X ו-Y משתנים מקריים. אם $M_X\left(t\right) = M_Y\left(t\right)$ בקטע פתוח סביב 0, אז $X \overset{\text{d}}{=} Y$ (לא נוכיח).

הוכחה: נוכיח את המקרה הפרטי. נגדיר $Z_n = \frac{1}{\sqrt{n}}\sum\limits_{j=1}^{n} X_j$, צריך להוכיח כי $Z_n \xrightarrow{\text{d}} Z$. תחילה, נבחין כי לכל k מתקיים $M_X^{(k)}(0) = \mathbb{E}\left[X^k\right]$. אם נצליח להראות כי $M_{Z_n}(t) \underset{n\to\infty}{\longrightarrow} M_Z(t)$, סיימנו, לפי המשפט האחרון. נחשב את $M_{Z_n}(t)$:

$$M_{Z_n}(t) = M_{\frac{1}{\sqrt{n}}\sum\limits_{j=1}^{n} X_j}(t) = \prod_{i=1}^{n} M_{\frac{1}{\sqrt{n}}X_i}(t) = \left(M_{\frac{1}{\sqrt{n}}X_1}(t)\right)^n = \left(\mathbb{E}\left[e^{\frac{t}{\sqrt{n}}X_1}\right]\right)^n = \left(M_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

המעבר השני נובע מאי-תלות והמעבר השלישי נובע משיוויון ההתפלגויות. נחשב את $M_Z(t)$:

$$M_Z(t) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{tx}\cdot e^{-\frac{x^2}{2}}dx = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\left(\frac{t^2}{2}-tx+\frac{x^2}{2}\right)}dx = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}}dx = e^{\frac{t^2}{2}}$$

על כן, עלינו להראות כי $\left(M_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n \underset{n\to\infty}{\longrightarrow} e^{\frac{t^2}{2}}$. יהי $s\in[-1,1]$, נפתח טור טיילור סביב אפס:

$$M_{X_1}(s) = M_{X_1}(0) + M'_{X_1}(0)\cdot s + M''_{X_1}(0)\cdot\frac{s^2}{2} + o\left(s^2\right)$$

כעת, מהנחות המקרה הפרטי:

$$= \mathbb{E}[1] + \mathbb{E}[X_1]\cdot s + \mathbb{E}\left[X_1^2\right]\cdot\frac{s^2}{2} + o\left(s^2\right) = 1 + \frac{s^2}{2} + o\left(s^2\right)$$

קיבלנו כי $M_{Z_n}(t) = \left(1+\frac{t^2}{2}+\epsilon_n(t)\right)^n$ כאשר $\frac{\epsilon_n}{\frac{t^2}{2n}} \underset{n\to\infty}{\longrightarrow} 0$. לבסוף, מאחר ש $\left(1+\frac{a_n}{n}\right)^n \underset{n\to\infty}{\longrightarrow} e^a$ כאשר $a_n \underset{n\to\infty}{\longrightarrow} a$, נקבל ש $M_{Z_n}(t) \underset{n\to\infty}{\longrightarrow} e^{\frac{t^2}{2}}$, כנדרש.

דוגמה: תהי $X_n \sim \text{Bin}(n,p)$ סדרת משתנים מקרים בינומים ויהי $X \sim \text{N}(0,1)$. נראה כי $\frac{X_n - pn}{\sqrt{p(1-p)n}} \xrightarrow{\text{d}} X$. אכן, יהיו $\{Y_k\}_{k=1}^{\infty}$ משתנים מקריים בלתי תלוים שמתפלגים ברנולי p. אזי $X_n = \sum\limits_{i=1}^{n} Y_i$ ולכן:

$$\frac{X_n - pn}{\sqrt{p(1-p)n}} = \frac{1}{\sqrt{n}}\cdot\sum_{i=1}^{n}\frac{Y_i - \mathbb{E}[Y_i]}{\sqrt{\text{Var}(Y_i)}}$$

נסמן $Z_i = \frac{Y_i - \mathbb{E}[Y_i]}{\sqrt{\text{Var}(Y_i)}}$ ונשים לב כי אלו משתנים מקריים בלתי תלויים בעלי תוחלת 0 ושונות 1. נקבל לפי משפט הגבול המרכזי, שאכן:

$$\frac{X_n - pn}{\sqrt{p(1-p)n}} = \frac{1}{\sqrt{n}}\cdot\sum_{i=1}^{n}\frac{Y_i - \mathbb{E}[Y_i]}{\sqrt{\text{Var}(Y_i)}} \xrightarrow{\text{d}} X$$