תורת ההסתברות 1 — סיכום

7 במרץ 2025

תוכן העניינים

1 שיעור 1 — 29.10.2024

1.1 מבוא הקורס

נלמד לפי ספר שעוד לא יצא לאור שנכתב על־ידי אורי עצמו, הוא עוד לא סופי ויש בו בעיות ואי־דיוקים, תשיג את הספר הזה. כן יש הבדל בין הקורס והספר אז לא לסמוך על הסדר שלו גם כשאתה משיג אותו, אבל זו תוספת מאוד נוחה. יש סימון של כוכביות לחומר מוסף, כדאי לעבור עליו לקראת המבחן כי זה יתן לנו עוד אינטואיציה והעמקה של ההבנה.

נשים לב כי ענף ההסתברות הוא ענף חדש יחסית, שהתפתח הרבה אחרי שאר הענפים הקלאסיים של המתמטיקה, למעשה רק לפני 400 שנה נשאלה על־ידי נזיר במהלך חקר של משחק אקראי השאלה הראשית של העולם הזה, מה ההסתברות של הצלחה במשחק.

נעבור לדבר על פילוסופיה של ההסתברות. מה המשמעות של הטלת מטבע מבחינת הסתברות? ישנה הגישה של השכיחות, שמציגה הסתברות כתוצאה במקרה של חזרה על ניסוי כמות גדולה מאוד של פעמים. יש כמה בעיות בזה, לרבות חוסר היכולת להגדיר במדויק אמירה כזו, הטיות שנובעות מפיזיקה, מטבעות הם לא מאוזנים לדוגמה. הבעיה הראשית היא שלא לכל בעיה אפשר לפנות בצורה כזאת. ישנה גישה נוספת, היא הגישה האוביקטיבית או המתמטית, הגישה הזו בעצם היא תרגום בעיה מהמציאות לבעיה מתמטית פורמלית. לדוגמה נשאל את השאלה מה ההסתברות לקבל 6 בהגרלה של כל המספרים מ־1 עד מיליון. השיטה ההסתברותית קובעת שאם אני רוצה להוכיח קיום של איזשהו אוביקט, לפעמים אפשר לעשות את זה על־ידי הגרלה של אוביקט כזה והוכחה שיש הסתברות חיובית שהוא יוגרל, וזו הוכחה שהוא קיים. מה התחזיות שינבעו מתורת ההסתברות? לדוגמה אי־אפשר לחזות הטלת מטבע בודדת, אבל היא כן נותנת הבנה כללית של הטלת 1000 מטבעות, הסתברויות קטנות מספיק יכולות להיות זניחות ובמקרה זה נוכל להתעלם מהן. לפחות בתחילת הקורס נדבר על תרגום של בעיות מהמציאות לבעיות מתמטיות, זה אומנם חלק פחות ריגורזי, אבל הוא כן חשוב ליצירת קישור בין המציאות לבין החומר הנלמד.

דבר אחרון, ישנה השאלה הפילוסופית של האם באמת יש הסתברות שכן לא בטוח שיש אקראיות בטבע, הגישה לנושא מבחינה פיזיקלית קצת השתנתה בעת האחרונה וקשה לענות על השאלה הזאת. יש לנו תורות פיזיקליות שהן הסתברותיות בעיקרן, כמו תורת הקוונטים, תורה זו לא סתם הסתברותית, אנחנו לא מנסים לפתור בעיות הסתברותיות אלא ממש משתמשים במודלים סטטיסטיים כדי לתאר מצב בעולם. לדוגמה נוכל להסיק ככה מסקנה פשוטה שאם מיכל גז נפתח בחדר, יהיה ערבוב של הגז הפנימי ושל אוויר החדר, זוהי מסקנה הסתברותית. החלק המדהים הוא שתורת הקוונטים מניחה חוסר דטרמניזם כתכונה יסודית ועד כמה שאפשר לראות יש ניסויים שמוכיחים שבאמת יש חוסר ודאות בטבע. דהינו שברמה העקרונית הפשוטה באמת אין תוצאה ודאית בכלל למצבים כאלה במציאות.

1.2 מרחבי מדגם ופונקציית הסתברות

הגדרה 1.1 (מרחב מדגם) מרחב מדגם הוא קבוצה לא ריקה שמהווה העולם להסתברות.
נסמנה Ω. איבר במרחב המדגם נסמן ב־$\omega \in \Omega$ על־פי רוב.

נוכל להגיד שמרחב במדגם הוא הקבוצה של האיברים שעליה אנחנו שואלים בכלל שאלות, זהו הייצוג של האיברים או המצבים שמעניינים אותנו. בהתאם נראה עכשיו מספר דוגמות שמקשרות בין אובייקטים שאנו דנים בהם בהסתברות ובהגדרה פורמלית של מרחבי מדגם עבורם.

דוגמה 1.1 (מרחבי הסתברות שונים) נראה מספר דוגמות למצבים כאלה:

- הטלת מטבע תוגדר על־ידי $\Omega = \{H, T\}$.
- הטלת שלושה מטבעות תהיה באופן דומה $\Omega = \{H, T\}^3$.
- הטלת קוביה היא $\Omega = [6] = \{1, \dots, 6\}$.
- הטלת מטבע ואז אם יוצא עץ (H) אז מטילים קוביה ואם יוצא פלי (T) אז מטילים קוביה עם 8 פאות.
במקרה זה נסמן $\Omega = \{H1, H2, H3, \dots, H6, T1, \dots, T8\} = \{H, T\} \times \{1, \dots, 8\}$ כאשר הכוונה פה היא לזוג סדור $\langle H, 1\rangle$.
- ערבוב חפיסת קלפים, במקרה זה מרחב המדגם שלנו יהיה סימון של הקלפים כרשימה מספרית בלבד, דהינו $\Omega = S_{52}$.
נוכל גם לסמן במקום את $\Omega = \{1, \dots, 52\}^{52}$, זהו סימון זהה.

בדוגמה זו קל במיוחד לראות שכל איבר בקבוצה מתאר מצב סופי כלשהו, ואנו יכולים לשאול שאלות הסתברותיות מהצורה מה הסיכוי שנקבל ω מסוים מתוך Ω, זאת ללא התחשבות בבעיה שממנה אנו מגיעים. נבחן עתה גם דוגמות למקרים שבהם אין לנו מספר סופי של אפשרויות, למעשה מקרים אלה דומים מאוד למקרים שראינו עד כה.

דוגמה 1.2 (מרחבי מדגם לא סופיים) מטילים מטבע עד שיוצא H, אז מרחב המדגם הוא $\Omega = \mathbb{N} \cup \{\infty\}$.
באופן דומה נוכל לבחון מדידת זמן התפרקות חלקיק, היא $\Omega = \mathbb{R}_+ \cup \{\infty\}$.

הגדרה 1.2 (פונקציית הסתברות נקודתית) יהי מרחב מדגם Ω ותהי $p : \Omega \to [0, \infty)$ פונקציה כך שמתקיים

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

אז פונקציה זו נקראת **פונקציית הסתברות**.

למעשה פונקציית הסתברות היא מה שאנחנו נזהה עם הסתברות במובן הפשוט, פונקציה זו מגדירה לנו לכל סיטואציה ממרחב המדגם מה הסיכוי שנגיע אליה, כך לדוגמה אם נאמר שהטלת מטבע תגיע בחצי מהמקרים לעץ ובחצי השני לפלי, אז זו היא פונקציית ההסתברות עצמה, פונקציה שמחזירה חצי עבור עץ וחצי עבור פלי, נראה מספר דוגמות.

דוגמה 1.3 (פונקציית הסתברות להטלת מטבע) נגדיר $\Omega = \{H, T\}$ ויהי $0 \le \alpha \le 1$, נגדיר $p(H) = \alpha, p(T) = 1 - \alpha$.

דוגמה 1.4 (פונקציית הסתברות אינסופית) נגדיר $\Omega = \mathbb{N} \cup \{\infty\}$ ו־ $p(\omega) = \begin{cases} 2^{-\omega} & \omega \in \mathbb{N} \\ 0 & \omega = \infty \end{cases}$. בדוגמה זו נקבל $\sum_{n=1}^{\infty} 2^{-n} = 1$ ולכן זו אכן פונקציית הסתברות.

נבחין כי הדוגמה האחרונה מתארת לנו התפלגות של דעיכה, זאת אומרת שלדוגמה אם קיים חלקיק עם זמן מחצית חיים של יחידה אחת, פונקציית הסתברות זו תניב לנו את הסיכוי שהוא התפרק לאחר כמות יחידות זמן כלשהי.

דוגמה 1.5 נגדיר $\Omega = \mathbb{N}$ ו־ $p(\omega) = \frac{1}{\omega(\omega+1)}$, נבחין כי אכן $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$.

הגדרה 1.3 (תומך) התומך של p הוא $\mathrm{Supp}(p) = \{\omega \in \Omega \mid p(\omega) > 0\}$.
נבחין כי התומך הוא למעשה קבוצת האיברים שאפשרי לקבל לפי פונקציית ההסתברות, כל שאר המצבים מקבלים 0, משמעו הוא שאין אפשרות להגיע אליו.

הערה נבחין כי תמיד $\mathcal{F} \subseteq \mathcal{P}(\Omega)$.

הגדרה 1.4 (מאורע) מאורע הוא תת־קבוצה של מרחב המדגם, קבוצת כל המאורעות תסומן $\mathcal{F}$. עבור מאורע A המאורע המשלים מסומן ב־ $A^C = \Omega \setminus A$.

הגדרה 1.5 (פונקציית הסתברות) נגדיר עתה פונקציית הסתברות שאיננה נקודתית. יהי מרחב מדגם Ω וקבוצת מאורעות $\mathcal{F}$.
תהי $\mathbb{P} : \mathcal{F} \to [0, \infty)$ המקיימת את התכונות הבאות:

1. $\mathbb{P}(\Omega) = 1$
2. לכל $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ סדרת מאורעות שונים מתקיים
$$\sum_{i \in \mathbb{N}} \mathbb{P}(A_i) = \mathbb{P}(\bigcup_{i \in \mathbb{N}} A_i)$$
דהינו, הפונקציה סכימה בתת־קבוצות בנות מניה.

לפונקציה כזו נקרא **פונקציית ההסתברות** על $(\Omega, \mathcal{F})$.

טענה 1.6 *תהי p פונקציית הסתברות נקודתית על Ω אז נגדיר פונקציית הסתברות $\mathbb{P}_p$ על־ידי*

$$\mathbb{P}_p(A) = \sum_{\omega \in A} p(\omega)$$

אז $\mathbb{P}_p$ היא פונקציית הסתברות.

הוכחה. נוכיח ששתי התכונות של פונקציית הסתברות מתקיימות.

$$\mathbb{P}_p(A) = \sum_{\omega \in A} p(\omega) \ge 0$$

שכן זהו סכום אי־שלילי מהגדרת p, בנוסף נקבל מההגדרה של p כי

$$\mathbb{P}_p(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$$

וקיבלנו כי התכונה הראשונה מתקיימת.

תהי $\{A\}_{i=1}^{\infty} \in \mathcal{F}$, אז נקבל

$$\sum_{i\in\mathbb{N}} \mathbb{P}_p(A_i) = \sum_{i\in\mathbb{N}} \left(\sum_{\omega\in A_i} p(\omega) \right) = \sum_{\omega\in\bigcup_{i\in\mathbb{N}} A_i} p(\omega) = \mathbb{P}_p(\bigcup_{i\in\mathbb{N}} A_i)$$

ולכן גם התכונה השנייה מתקיימת וקיבלנו כי $\mathbb{P}_p$ היא אכן פונקציית הסתברות. $\square$

נשים לב כי בעוד פונקציית הסתברות נקודתית מאפשרת לנו לדון בהסתברות של איבר בודד בקבוצות בנות מניה, פונקציית הסתברות למעשה מאפשרת לנו לדון בהסתברות של מאורעות, הם קבוצות של כמה מצבים אפשריים, ובכך להגדיל את מושא הדיון שלנו. מהטענה האחרונה גם נוכל להסיק שבין שתי ההגדרות קיים קשר הדוק, שכן פונקציית הסתברות נקודתית גוררת את קיומה של פונקציית הסתברות כללית.

2 תרגול 1 — 31.10.2024

המתרגל הוא אמיר, amir.behar@mail.huji.ac.il

2.1 מרחבי הסתברות סופיים ובני־מניה

ניזכר בהגדרה למרחב הסתברות, המטרה של הגדרה זו היא לתאר תוצאות אפשריות של מצב נתון.

הגדרה 2.1 (מרחב הסתברות) מרחב הסתברות הוא קבוצה $(\Omega, \mathcal{F}, \mathbb{P})$ כאשר $\mathbb{P} : \mathcal{F} \to [0,1]$, כך שמתקיים

1. חיוביות: $\forall A \in \mathcal{F}, \mathbb{P}(A) \geq 0$
2. נרמול: $\mathbb{P}(\Omega) = 1$
3. סיגמא־אדיטיביות: $\forall \{A_i\}_{i=1}^{\infty} \in \mathcal{F}, (\forall i,j \in \mathbb{N}, i \neq j \implies A_i \cap A_j = \emptyset) \implies \sum_{i\in I} \mathbb{P}(A_i) = \mathbb{P}(\bigcup_{i\in I} A_i)$

תרגיל 2.1 יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, $A, B \in \mathcal{F}$, הוכיחו

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

הוכחה. נבחין כי $\mathbb{P}(A) = \mathbb{P}(A - (A \cap B)) + \mathbb{P}(A \cap B)$ וגם $\mathbb{P}(B) = \mathbb{P}(B - (A \cap B)) + \mathbb{P}(A \cap B)$. נוכל אם כן לסכום ולקבל

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

נבחין כי השוויון האחרון נובע מהזרות של קבוצות אלה. □

לאורך פרק זה נגדיר מעתה שמתקיים Ω סופית, $\mathcal{F} = 2^{\Omega}$ ואף נגדיר כי ההסתברות אחידה, דהינו $\forall A \;\; \mathbb{P}(A) = \frac{|A|}{|\Omega|}$, זה כמובן שקול לטענה

$$\forall \omega, \omega' \in \Omega, \mathbb{P}(\{\omega\}) = \mathbb{P}(\{\omega'\})$$

תרגיל 2.2 מטילים קוביה הוגנת, מה ההסתברות שיצא מספר זוגי?

פתרון נגדיר $\Omega = [6] = \{1, \dots, 6\}$, עם $\mathbb{P}$ אחידה.
נרצה לחשב את $A = \{2,4,6\}$ ולכן נקבל $\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$.

תרגיל 2.3 מטילים מטבע הוגן שלוש פעמים, מה ההסתברות שיצא עץ בדיוק פעמיים, ומה ההסתברות שיצא עץ לפחות פעמיים?

פתרון נגדיר $\Omega = \{TTT, TTP, TPT, PTT, \dots\}$.
עבור המקרה הראשון נגדיר $A = \{TTP, TPT, PTT\}$, ולכן נקבל שההסתברות היא $\mathbb{P}(A) = \frac{3}{8}$.
במקרה השני נקבל $B = A \cup \{TTT\}$ ולכן $P(B) = \frac{1}{2}$.

תרגיל 2.4 מטילים קוביה הוגנת n פעמים.

1. מה ההסתברות שתוצאת ההטלה הראשונה קטנה מ־4?
2. מה ההסתברות שתוצאת ההטלה הראשונה קטנה שווה מתוצאת ההטלה השנייה?
3. מה ההסתברות שיצא 1 לפחות פעם אחת?

פתרון נגדיר $\Omega = [6]^n = \{(x_1, \dots, x_n) \mid x_i \in [6]\}$.

1. נגדיר $A = \{(x_1, \dots, x_n) \in \Omega \mid x_1 < 4\}$ ולכן $\mathbb{P}(A) = \frac{3 \cdot 6^{n-1}}{6^n} = \frac{1}{2}$.
2. נגדיר $B = \{(x_1, \dots, x_n) \in \Omega \mid x_1 \leq x_2\} = \bigcup_{i=1}^{6} \{(x_1, i, x_3, \dots, x_n) \in \Omega \mid x_i \leq i\}$, ולכן נקבל
$$\mathbb{P}(B) = \sum \mathbb{P}(B_i) = \sum \frac{i \cdot 6^{n-2}}{6^n} = \frac{\sum_{i=1}^{6} i}{6^2} = \frac{6 \cdot 7}{6^2 \cdot 2} = \frac{7}{12}$$
3. הפעם $C = \{(x_1, \dots, x_n) \in \Omega \mid \exists i, x_i = 1\}$, בהתאם $C^C = \{(x_1, \dots, x_n) \in \Omega \mid \forall i, x_1 \neq 1\}$.
לכן נקבל $\mathbb{P}(C^C) = \frac{5^n}{6^n} \implies \mathbb{P}(C) = 1 - \frac{5^n}{6^n}$.

תרגיל 2.5 חמישה אנשים בריאים וחמישה אנשים חולי שפעת עומדים בשורה. מה ההסתברות שחולי השפעת נמצאים משמאל לאנשים הבריאים?

פתרון נגדיר Ω ככל הסידורים של $0,1$ כשיש חמישה מכל סוג. לכן נקבל $\binom{10}{5} = |\Omega|$, שכן $\Omega = \{X \subset [10] \mid |X| = 5\}$.
המאורע הפעם הוא $A = \{\{1,2,3,4,5\}\}$ ובהתאם $\mathbb{P}(A) = \frac{5!5!}{10!}$.

נוכל גם להגדיר $\Omega = S_{10}$ כאשר חמשת המספרים הראשונים מייצגים בריאים וחמשת האחרונים מייצגים חולים.
במקרה זה נקבל $A = \{\pi \in \Omega \mid \pi(\{1,2,3,4,5\}) \subseteq \{1,2,3,4,5\}\}$ ולכן $|A| = 5!5!$ וכך נקבל $\mathbb{P}(A) = \frac{5!5!}{10!}$.

3 שיעור 2 — 31.10.2024

3.1 השלמה לטורים דו־מימדיים

נגדיר הגדרה שדרושה לצורך ההרצאה הקודמת כדי להיות מסוגלים לדון בסכומים אינסופיים בני־מניה.

הגדרה 3.1 (סכום קבוצת בת־מניה) אם $\{a_i\}_{i\in I}$ ו־$a_i \ge 0$ לכל $i \in I$ אז נגדיר

$$\sum_{i\in I} a_i = \sup\left\{\sum_{i\in J} \mid J \subseteq I, J \text{ is finite}\right\}$$

3.2 תכונות של פונקציות הסתברות

נעבור עתה לבחון פונקציות הסתברות ואת תכונותיהן, נתחיל מתרגיל שיוצק תוכן לתומך של פונקציית הסתברות:

תרגיל 3.1 הוכיחו כי אם $\sum_{i\in I} a_i < \infty$ ו־$a_i \ge 0$ לכל $i \in I$ אז $|\{i \in I \mid a_i < 0\}| \le \aleph_0$. במילים אחרות הוכיחו כי התומך של a הוא בן־מניה.

בשיעור הקודם ראינו את ההגדרה והטענה הבאות:

הגדרה 3.2 (פונקציית הסתברות מתאימה לנקודתית) בהינתן פונקציית הסתברות נקודתית p נגדיר

$$\mathbb{P}_p(A) = \sum_{\omega\in A} p(\omega)$$

טענה 3.3 *$\mathbb{P}_p$ היא פונקציית הסתברות.*

טענה זו בעצם יוצרת קשר בין פונקציות הסתברות לפונקציות הסתברות נקודתיות, ומאפשרת לנו לחקור את פונקציות ההסתברות לעומק באופן פשוט הרבה יותר. נשתמש עתה בכלי זה.

הגדרה 3.4 (מרחב הסתברות בדיד) אם $\mathbb{P}$ פונקציית הסתברות כך שקיימת פונקציית הסתברות נקודתית p כך ש־$\mathbb{P} = \mathbb{P}_p$, אז נאמר ש־$\mathbb{P}$ היא בדידה ו־$(\Omega, \mathcal{F}, \mathbb{P})$ **מרחב הסתברות בדיד**.

טענה 3.5 *יש פונקציות הסתברות שאינן בדידות. בפרט, עבור מדגם ההסתברות $\Omega = [0, 1]$ קיימת פונקציית הסתברות $\mathbb{P}$ המקיימת*

$$\forall a, b \in \mathbb{R}, 0 \le a \le b \le 1 \implies \mathbb{P}([a, b]) = b - a$$

דוגמה 3.1 ידוע כי $\sum_{n\in\mathbb{N}} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$ ולכן נוכל להגדיר $\Omega = \mathbb{N}$ ו־$p(n) = \frac{1}{\frac{\pi^2}{6} n^2}$, הגדרה זו תניב ש־$\sum_{n\in\mathbb{N}} p(n) = 1$ ולכן זו פונקציית הסתברות. נחשב את $\mathbb{P}_p(A)$ עבור $A = 2\mathbb{N}$:

$$\mathbb{P}_p(A) = \sum_{n\in A} p(n) = \sum_{k\in\mathbb{N}} p(2k) = \frac{1}{\frac{\pi^2}{6}(2k)^2} = \frac{6}{\pi^2}\frac{1}{4}\sum_{k\in\mathbb{N}}\frac{1}{k^2} = \frac{1}{4}$$

נסביר, הגדרנו פונקציית הסתברות של דעיכה, דהינו שככל שהמספר שאנו מבקשים גדול יותר כך הוא פחות סביר באופן מעריכי (לדוגמה זמן מחצית חיים), ואז שאלנו כמה סביר המאורע שבו נקבל מספר זוגי.

משפט 3.6 (תכונות פונקציית הסתברות) *$\mathbb{P}$ פונקציית הסתברות על $(\Omega, \mathcal{F})$, אז*

1. $\mathbb{P}(\emptyset) = 0$
2. *אם I קבוצה סופית ו־$\{A_i\}_{i\in I}$ מאורעות זרים בזוגות, אז* $\mathbb{P}(\bigcup_{i\in I} A_i) = \sum_{i\in I} \mathbb{P}(A_i)$
3. *אם $A \subseteq B$ מאורעות אז* $\mathbb{P}(A) \le \mathbb{P}(B)$
4. $\mathbb{P}(A) \le 1$ *לכל מאורע A*
5. *לכל מאורע A מתקיים* $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$

הוכחה. נוכיח את התכונות

1. נראה כי $\mathbb{P}(\emptyset) = \sum_{i=1}^{\infty} \mathbb{P}(\emptyset)$ שכן כל איחוד של קבוצות ריקות הוא זר, לכן אילו $\mathbb{P}(\emptyset) \ne 0$ נקבל ישר סתירה, נסיק כי $\mathbb{P}(\emptyset) = 0$ בלבד.

2. נגדיר $A_i = \emptyset$ לכל $i > n$ ונשתמש בסיגמא־אדיטיביות ונקבל
$$\mathbb{P}(\bigcup_{i\in I} A_i) = \mathbb{P}(\bigcup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i) = \sum_{i\in I} \mathbb{P}(A_i)$$
3. נשתמש בתכונה 2 על $B, B \setminus A$, אלו הן קבוצות זרות כמובן, אם נגדיר $D = A \cup (B \setminus A)$ נקבל $\mathbb{P}(D) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$.
4. נובע ישירות מתכונה 3 ומ־$A \subseteq \Omega$.
5. ניזכר כי $A^C = \Omega \setminus A$ ולכן $\Omega = A \cup A^C$ ונקבל $1 = \mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A^C)$.

$\square$

נעבור עתה לאפיון של פונקציות הסתברות בדידות, נבין מתי הן כאלה ומתי לא.

משפט 3.7 (תנאים שקולים לפונקציית הסתברות בדידה) *אם $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, התנאים הבאים שקולים:*

1. *$\mathbb{P}$ היא פונקציית הסתברות בדידה*
2. *$\mathbb{P}$ נתמכת על קבוצות בנות־מניה, כלומר קיימת קבוצה $A \in \mathcal{F}$ בת־מניה כך ש־$\mathbb{P}(A) = 1$*
3. *$\sum_{\omega\in\Omega} \mathbb{P}(\{\omega\}) = 1$*
4. *לכל מאורע $A \in \mathcal{F}$ מתקיים $\mathbb{P}(A) = \sum_{\omega\in A} \mathbb{P}(\{\omega\})$*

הוכחה. $2 \implies 1$: נניח ש־$\mathbb{P} = \mathbb{P}_p$ עבור $p : \Omega \to [0, \infty)$ פונקציית הסתברות נקודתית. נסתכל על $\mathrm{Supp}(p) = \{\omega \in \Omega \mid p(\omega) > 0\}$, לפי הגדרת הסכום והתרגיל נובע ש־$A = \mathrm{Supp}(p)$ בת־מניה. נקבל
$$\mathbb{P}(A) = \sum_{\omega\in A} p(\omega) = \sum_{\omega\in\Omega} p(\omega) = \mathbb{P}(\Omega) = 1$$
$4 \implies 2$: נניח ש־$\mathbb{P}(S) = 1$ עבור S בת־מניה. לכן $\mathbb{P}(S^C) = 0$. נראה כי A הוא איחוד זר $A = (A \cap S) \cup (A \cap S^C)$ ולכן נקבל
$$\mathbb{P}(A) = \mathbb{P}(A \cap S) + \mathbb{P}(A \cap S^C) = \mathbb{P}(A \cap S) + 0 = \sum_{\omega\in A\cap S} \mathbb{P}(\{\omega\}) = \sum_{\omega\in A} \mathbb{P}(\{\omega\})$$
$3 \implies 4$: אם נבחר $A = \Omega$ נקבל את טענה 3.

$1 \implies 3$: נגדיר $p : \Omega \to [0, \infty)$ על־ידי $p(\omega) = \mathbb{P}(\{\omega\})$, נקבל $\sum_{\omega\in\Omega} p(\omega) = 1$ ולכן p היא פונקציית הסתברות נקודתית. מהתרגיל והגדרת הסכום נובע ש־$S = \mathrm{Supp}(p)$ היא בת־מניה ומתקיים $\mathbb{P}(S^C) = 0$, אז לכל $A \in \mathcal{F}$ מתקיים
$$\mathbb{P}(A) = \mathbb{P}(A \cap S) + \mathbb{P}(A \cap S^C) = \mathbb{P}(A \cap S) = \sum_{\omega\in A\cap S} \mathbb{P}(\{\omega\}) = \sum_{\omega\in A} \mathbb{P}(\{\omega\}) = \sum_{\omega\in A} p(\omega) = \mathbb{P}_p(A)$$

$\square$

3.3 פרדוקס יום ההולדת

פרדוקס יום ההולדת הוא פרדוקס מוכר הגורס כי גם בקבוצות קטנות יחסית של אנשים, הסיכוי שלשני אנשים שונים יהיה תאריך יום הולדת זהה הוא גבוה במידה משונה. הפרדוקס נקרא כך שכן לכאורה אין קשר בין מספר הימים בשנה לבין הסיכוי הכל־כך גבוה שמצב זה יקרה, נבחן עתה את הפרדוקס בהיבט הסתברותי.

נניח שכל תאריכי יום ההולדת הם סבירים באותה מידה ונבחן את הפרדוקס. נגדיר $\Omega = [365]^k$ עבור k מספר האנשים בקבוצה נתונה כלשהי. $p(\omega) = \frac{1}{365^k}$ לכל $\omega \in \Omega$. נקבל $\mathbb{P}(A) = \mathbb{P}_p(A) = \frac{|A|}{365^k}$. נרצה לחשב את A כמאורע שיש לפחות שני אנשים שיש להם יום הולדת באותו יום, דהינו שיש שני ערכים זהים ברשימת המספרים, נגדיר $A = \{\omega \in \Omega \mid \exists 1 \leq i \neq j \leq k, \omega_i = \omega_j\}$. בשל המורכבות נבחן את המשלים A^C, נקבל $|A^C| = 365 \cdot 364 \cdots (365 - (k-1)) = \frac{365!}{(365-(i-1))!}$. נציב ונחשב:
$$\mathbb{P}(A^C) = \frac{|A^C|}{365^k} = \prod_{i=1}^{k} \frac{365 - (i-1)}{365} = \prod_{i=1}^{k} (1 - \frac{i-1}{365})$$
מהנוסחה שקיבלנו נראה שמההצבה $k = 23$ נקבל שההסתברות היא בערך $\frac{1}{2}$, דהינו בקבוצה של 23 אנשים יש סבירות של חצי שלפחות שניים יחגגו יום הולדת באתו יום.

4 שיעור 3 — 5.11.2024

4.1 מכפלת מרחבי הסתברות בדידים

ניזכר תחילה במרחבי הסתברות אחידים

הגדרה 4.1 (מרחב הסתברות אחיד) מרחב הסתברות אחיד הוא $(\Omega, \mathcal{F}, \mathbb{P}_p)$ המקיים $p(\omega_1) = p(\omega_2)$ לכל $\omega_1, \omega_2 \in \Omega$.

מסקנה 4.2 $\mathbb{P}_p(A) = \frac{|A|}{|\Omega|}$

נבחין כי במקרים מסוימים ההסתברות שלנו מורכבת משני מאורעות בלתי תלויים, במקרים אלה נרצה להגדיר מכפלה של מרחבי ההסתברות.

הגדרה 4.3 (מרחב מכפלת הסתברויות) אם $(\Omega_1, \mathcal{F}_1, \mathbb{P}_{p_1})$ ו־$(\Omega_2, \mathcal{F}_2, \mathbb{P}_{p_2})$ מרחבי הסתברות בדידים נגדיר $q : \Omega_1 \times \Omega_2 \to [0, \infty)$ על־ידי $q(\omega_1, \omega_2) = p(\omega_1) \cdot p(\omega_2)$.

טענה 4.4 *q פונקציית הסתברות נקודתית.*

הוכחה. נשתמש ישירות בהגדרה ונחשב

$$\sum_{(\omega_1,\omega_2)\in\Omega_1\times\Omega_2} q(\omega_1, \omega_2) = \sum_{\omega_1\in\Omega_1,\omega_2\in\Omega_2} q(\omega_1, \omega_2) = \sum_{\omega_1\in\Omega_1} \left(\sum_{\omega_2\in\Omega_2} p_1(\omega_1)p_2(\omega_2) \right) = \sum_{\omega_1\in\Omega_1} p_1(\omega_1) = 1$$

□

עתה כשהוכחנו טענה זו, יש לנו הצדקה אמיתית להגדיר את $(\Omega_1 \times \Omega_2, \mathcal{F}_{1,2}, \mathbb{P}_q)$ כמרחב הסתברות, ונקרא לו מרחב מכפלה.

טענה 4.5 *אם $(\Omega_1, \mathcal{F}_1, \mathbb{P}_{p_1})$ ו־$(\Omega_2, \mathcal{F}_2, \mathbb{P}_{p_2})$ מרחבי הסתברות אחידים, אז מרחב המכפלה $(\Omega_1 \times \Omega_2, \mathcal{F}_{1,2}, \mathbb{P}_q)$ אחיד אף הוא.*

הוכחה.

$$q(\omega_1, \omega_2) = p_1(\omega_1)p_2(\omega_2) = \frac{1}{|\Omega_1|} \cdot \frac{1}{|\Omega_2|} = \frac{1}{|\Omega_1 \times \Omega_2|}$$

□

הגדרה 4.6 (מאורע שוליים ומאורע מכפלה) במרחב מכפלה המאורעות מהצורה $A \times \Omega_2$ או $\Omega_1 \times A$ נקראים שוליים.
מאורע מהצורה $A \times B$ נקרא מאורע מכפלה.

טענה 4.7 *במרחב מכפלה $\mathbb{P}_q(A \times B) = \mathbb{P}_{p_1}(A) \cdot \mathbb{P}_{p_2}(B)$. בפרט $\mathbb{P}_q(A \times \Omega_2) = \mathbb{P}_{p_1}(A)$.*

הוכחה.

$$\sum_{(\omega_1,\omega_2)\in A\times B} q(\omega_1, \omega_2) = \sum_{\omega_1\in A,\omega_2\in B} q(\omega_1, \omega_2) = \sum_{\omega_1\in A} \left(\sum_{\omega_2\in B} p_1(\omega_1)p_2(\omega_2) \right) = \sum_{\omega_1\in A} p_1(\omega_1)\mathbb{P}_{p_2}(B) = \mathbb{P}_{p_1}(A)\mathbb{P}_{p_2}(B)$$

□

דוגמה 4.1 בהינתן n הטלות מטבע כלשהו, מה ההסתברות שיצאו k עצים?

עבור ההטלה הראשונה, $\Omega_1 = \{0, 1\}$. עוד נגדיר $p(1) = \alpha, p(0) = 1 - \alpha$ עבור $0 \le \alpha \le 1$ כלשהו.
בהתאם נקבל $\Omega = \{0, 1\}^n$, וכן

$$q(\omega_1, \dots, \omega_n) = \prod_{i=1}^{n} p(\omega_i) = \prod_{i=1}^{n} \alpha^{\omega_i} \cdot (1 - \alpha)^{1-\omega_i} = \alpha^{\sum_{i=1}^{n} \omega_i}(1 - \alpha)^{n-\sum_{i=1}^{n} \omega_i}$$

נבחין כי היינו יכולים לתאר את המקרה הזה ממש על־ידי $q(\omega) = \alpha^{\omega} \cdot (1 - \alpha)^{1-\omega}$.

נעבור עתה לבחינת המאורע

$$A = \{(\omega_1, \dots, \omega_n) \in \Omega \mid \sum_{i=1}^{n} \omega_i = k\}$$

נקבל מהביטוי שמצאנו כי

$$\mathbb{P}_q(A) = \sum_{(\omega_1,\ldots,\omega_n)\in A} q(\omega_1,\ldots,\omega_n) \sum_{\sum_{i=1}^n \omega_i = k} \alpha^{\sum_{i=1}^n \omega_i}(1-\alpha)^{n-\sum_{i=1}^n \omega_i} = |A|\alpha^k(1-\alpha)^{n-k} = \binom{n}{k}\alpha^k(1-\alpha)^{n-k}$$

דוגמה 4.2 נבחן עתה את המקרה של הטלות הוגנות ובחינת המקרה שחצי מההטלות לפחות יצאו עץ, זאת־אומרת שנבחן את הדוגמה הקודמת כאשר $n = 2m, k = m$, ו־$\alpha = \frac{1}{2}$. מנוסחת סטרלינג שאנחנו לא מכירים $m! \simeq \sqrt{2\pi m}(\frac{m}{e})^m$ ואז נוכל להסיק

$$\mathbb{P}_q(A) = \binom{2m}{m}\frac{1}{2^m} \simeq \frac{\sqrt{4\pi m}(\frac{2m}{e})^{2m}}{(\sqrt{2\pi m}(\frac{k}{e})^m)^2 2^{2m}} = \frac{\sqrt{4\pi m}}{2\pi m} = \frac{1}{\sqrt{\pi m}}$$

4.2 ניסויים דו־שלביים

נניח $(\Omega_1, \mathcal{F}_1, \mathbb{P}_{p_1})$ מרחב הסתברות בדידה עבור הניסוי הראשון, ונניח שיש מרחב הסתברות בדידה עבור הניסוי השני כך שלכל תוצאה בניסוי הראשון, פונקציית ההסתברות תשתנה בהתאם בניסוי השני. לכל $\omega_1 \in \Omega_1$ יש פונקציית הסתברות נקודתית $p_{\omega_1} : \Omega_2 \to [0,\infty)$. נגדיר את מרחב הניסוי הדו־שלבי $(\Omega_1 \times \Omega_2, \mathcal{F}_{1,2}, \mathbb{P}_q)$, כאשר $q(\omega_1,\omega_2) = p_1(\omega_1)\cdot p_{\omega_1}(\omega_2)$.

טענה 4.8 $\mathbb{P}_q$ *פונקציית הסתברות.*

הוכחה.

$$\sum_{(\omega_1,\omega_2)\in\Omega_1\times\Omega_2} q(\omega_1,\omega_2) = \sum_{\omega_1\in\Omega_1}\left(\sum_{\omega_2\in\Omega_2} p_1(\omega_1)p_{\omega_1}(\omega_2)\right) = \sum_{\omega_1\in\Omega_1} p_1(\omega_1)\left(\sum_{\omega_2\in\Omega_2} p_{\omega_1}(\omega_2)\right) = \sum_{\omega_1\in\Omega_1} p_1(\omega_1) = 1$$

□

דוגמה 4.3 $\Omega_1 = \{H,T\}$ ו־$\Omega_2 = \{1,\ldots,8\}$, נגדיר $p_1(H) = p_1(T) = \frac{1}{2}$. עוד נגדיר

$$p_H(\omega_2) = \begin{cases} \frac{1}{6} & 1 \le \omega_2 \le 6 \\ 0 & \text{else} \end{cases}, \qquad p_T(\omega_2) = \frac{1}{8}$$

מהגדרה זו נקבל

$$q(\omega_1,\omega_2) = \begin{cases} \frac{1}{12} & \omega_1 = H, \omega_2 \in [6] \\ 0 & \omega_1 = H, \omega_2 \in \{7,8\} \\ \frac{1}{16} & \omega_1 = T, \omega_2 \in [8] \end{cases}$$

משפט 4.9 (חסם האיחוד) *אם A, B מאורעות אז* $\mathbb{P}(A\cup B) \le \mathbb{P}(A) + \mathbb{P}(B)$.

הוכחה.

$$\mathbb{P}(A\cup B) = \mathbb{P}(A \uplus (B\setminus A)) = \mathbb{P}(A) + \mathbb{P}(B\setminus A) \le \mathbb{P}(A) + \mathbb{P}(B)$$

□

נוכל להשתמש בחסם האיחוד כדי להוכיח גרסה כללית יותר של המשפט:

משפט 4.10 (אי־שוויון בול) *אם $A_1,\ldots,A_k$ מאורעות, אז* $\mathbb{P}(\bigcup_{i=1}^k A_i) \le \sum_{i=1}^k \mathbb{P}(A_i)$.

דוגמה 4.4 נחזור לבחון את פרדוקס יום ההולדת, הפעם נבחן גרסה כללית יותר של הרעיון. נגדיר $\Omega = [m]^k$ עם הסתברות אחידה. נגדיר גם $A = \{\omega\in\Omega \mid \exists 1 \le i < j \le k, \omega_i = \omega_j\}$. אנו רוצים את ההסתברות $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$, אז נבחן את המשלים

$$A^C = \{\omega\in\Omega \mid \forall 1 \le i,j \le k, i \ne j \implies \omega_i \ne \omega_j\}$$

נחשב

$$|A^C| = m(m-1)\cdots(m-(k-1))$$

בהתאם

$$\mathbb{P}(A^C) = \frac{\prod_{i=0}^{k-1}(m-i)}{m^k} = \prod_{i=0}^{k-1}\frac{m-i}{m^k} = \prod_{i=0}^{k-1}(1-\frac{i}{m})$$

נזכור ש־$\forall x \in \mathbb{R}, 1+x \le e^x$, ונוכל לקבל

$$\prod_{i=0}^{k-1}(1-\frac{i}{m}) \le \prod_{i=0}^{k-1} e^{-\frac{i}{m}} = \exp(-\frac{1}{m}\sum_{i=0}^{k-1} i) = e^{-\frac{k(k-1)}{2m}}$$

כאשר k גדול ביחס ל־$\sqrt{2m}$ מקבלים הסם קרוב ל־0.

נגדיר הפעם $A = \bigcup_{\substack{i \ne j \\ i,j \in [k]}} A_{ij}$ עבור $A_{ij} = \{\omega \in \Omega \mid \omega_i = \omega_j\}$. וגם

$$i \ne j \implies \mathbb{P}(A_{ij}) = \frac{|A_{ij}|}{m^k} = \frac{m \cdot m^{k-2}}{m^k} = \frac{1}{m}$$

ועתה

$$\mathbb{P}(A) \le \sum_{\substack{i \ne j \\ i,j \in [k]}} \mathbb{P}(A_{ij}) = \sum_{\substack{i \ne j \\ i,j \in [k]}} \frac{1}{m} = \binom{k}{2}\frac{1}{m} = \frac{k(k-1)}{2m}$$

לכן אם k קטן ביחס ל־$\sqrt{2m}$ אז ההסתברות ליום־הולדת משותף קטנה.

5 תרגול 2 — 7.11.2024

5.1 פתרון שאלות הסתברותיות

נתחיל בבחינת טענה שימושית לביצוע חישובי הסתברות:

טענה 5.1 (נוסחת ההסתברות השלמה) *יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, $\mathcal{A}$ חלוקה בת־מניה של Ω, לכל $B \in \mathcal{F}$ מתקיים*

$$\mathbb{P}(B) = \sum_{A \in \mathcal{A}} \mathbb{P}(A \cap B)$$

נניח שיש מרחב הסתברות ויש חלוקה בת מניה של המרחב, אז לכל מאורע ההסתברות שלו היא הסכום על החלוקה על החיתוך של החלוקה ו־A.

הוכחה. נשים לב כי $B = \biguplus A \in \mathcal{A}B \cap A$ איחוד זר, ולכן הטענה נובעת מסיגמא־אדיטיביות. $\square$

תרגיל 5.1 קוביה מוטה בעלת 6 פאות עם הסתברות נקודתית $p(i) = \frac{i}{21}$ מוטלת 5 פעמים.
מה ההסתברות שתוצאת ההטלה הראשונה התקבלה פעם אחת ויחידה?

פתרון נגדיר $\Omega = [6]^5$ ונגדיר $\mathbb{P}(x_1, \ldots, x_5) = p(x_1) \cdots p(x_5)$. אנו רוצים לחשב את

$$B = \{(x_1, \ldots, x_5) \in \Omega \mid \forall j \neq 1, x_j \neq x_1\}$$

נגדיר חלוקה $\mathcal{A} = \{A_1, \ldots, A_6\}$ של Ω כך ש־$A_i = \{(i, x_2, \ldots, x_5) \in \Omega \mid 1 \leq x_j \leq 6\}$.
נקבל

$$\mathbb{P}(B \cap A_i) = \frac{i}{21} \cdot (1 - \frac{i}{21})^4$$

על־ידי שימוש בנוסחת ההסתברות השלמה נקבל

$$\mathbb{P}(B) = \sum_{i=1}^{6} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{6} \frac{i}{21}(1 - \frac{i}{21})^4$$

נראה עתה דוגמה לשימוש בחסם האיחוד בן־המניה, אותו נראה בהרצאה הבאה

טענה 5.2 (חסם האיחוד הבן־מניה) *אם $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ו־$\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ אז מתקיים*

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_i)$$

תרגיל 5.2 משלשלים k פתקי הצבעה בין n קלפיות.
מה ההסתברות שאין קלפי עם יותר מפתק אחד?

פתרון נגדיר $\Omega = \{(x_1, \ldots, x_n) \mid 0 \leq x_i, x_1 + \cdots + x_n = k\}$. נחשב ונקבל $|\Omega| = \binom{n+k-1}{k-1}$.
נגדיר את המאורע, $A = \{(x_1, \ldots, x_n) \in \Omega \mid x_i \leq 1\}$.
ננסה לחסום את המשלים,

$$\Omega \setminus A = \{(x_1, \ldots, x_n) \in \Omega \mid \exists i, x_i \geq 2\}$$

אם נגדיר $A_i = \{(x_1, \ldots, x_n) \in \Omega \mid x_i \geq 2\}$ אז נוכל להגדיר

$$\Omega \setminus A = \bigcup_{i \in [n]} A_i$$

נחשב את ההסתברות של כל A_i, מתקבל $|A_i| = \binom{n+k-3}{k-3}$ מהשיקול של סכימת הפתרונות השלמים תוך התעלמות משני פתקים.
לכן

$$\mathbb{P}(A_i) = \frac{|A_i|}{|\Omega|} = \frac{\binom{n+k-3}{k-3}}{\binom{n+k-1}{k-1}} = \frac{k(k-1)}{(k+n-1)(k+n-2)}$$

מחסם האיחוד נובע

$$\mathbb{P}(\Omega - A) \leq \sum_{i=1}^{n} \frac{k(k-1)}{(k+n-1)(k+n-2)} = n \cdot \frac{k(k-1)}{(n+k-1)(n+k-2)}$$

ועל־ידי מעבר למשלים שוב נוכל להסיק $\mathbb{P}(A) \geq 1 - n \cdot \frac{k(k-1)}{(n+k-1)(n+k-2)}$.

נזכור כי אנו מנסים להבין את המגמה כאשר המספרים מאוד גדולים, לכן נבחן את המקרה ש־$n \to \infty$, אז נובע $\mathbb{P}(A) \xrightarrow[n\to\infty]{} 1$, דהינו כאשר יש כמות קלפיות הולכת וגדלה הסיכוי שיהיה פתק יחיד בכל אחת (מספר הפתקים לא משתנה) הולך וגדל ומתקרב לסיכוי מלא.

נראה עתה דוגמה לשימוש במרחבי ניסוי דו־שלביים:

תרגיל 5.3 מה ההסתברות שנגריל מספר m בין 1 ל־n, ואז נגריל עוד מספר והוא יהיה בין 1 לבין m?

פתרון נבנה פונקציית הסתברות עבור הניסוי השני, נניח שבניסוי השני קיבלנו m:

$$p_m(k) = \begin{cases} \frac{1}{m} & k \le m \\ 0 & k > m \end{cases}, \qquad q(m,k) = \begin{cases} \frac{1}{mn} & k \le m \\ 0 & k > m \end{cases}$$

נגדיר A_k המאורע שתוצאת ההגרלה השניה היא k, לכן

$$\mathbb{P}(A_k) = \mathbb{P}(\{(m,k) \in \Omega \mid m \le k\}) = \sum_{m=1}^{n} q(m,k) = \sum_{m=k}^{n} \frac{1}{mn}$$

נבחין כי המעבר האחרון אכן תקין, שכן קיבענו את המשתנה השני, זאת אומרת שעכשיו במקום להסתכל על מספר שיותר קטן ממספר אחר, אנו בוחנים את המספר החוסם מלמעלה, המספר הגדול יותר.

לדוגמה

$$\mathbb{P}(A_n) = \frac{1}{n^2}, \qquad \mathbb{P}(A_1) = \sum_{m=1}^{n} \frac{1}{mn} = \frac{1}{n} \sum_{m=1}^{n} \frac{1}{m} \approx \frac{\log n}{n}$$

נבחן דוגמה ספציפית כהמשך של השאלה הזו, הפעם נגדיר $m = n/2$:

דוגמה 5.1 נגדיר $B_{n/2}$ להיות המאורע בהתחלה השניה ו־B שבהגרלה השניה יצא מספר גדול מ־$n/2$

$$\mathbb{P}(B_{n/2}) = \mathbb{P}(\bigcup_{k\ge n/2}^{n} A_k) = \frac{1}{n} \sum_{k \ge \frac{n}{2}}^{n} \sum_{m=k}^{n} \frac{1}{m} = \frac{1}{n} \sum_{m=\lceil \frac{n}{2} \rceil}^{n} \frac{\frac{n}{2} + 1 - n + m}{m}$$

כמו בשאלה הקודמת, גם הפעם נרצה להבין מגמה כללית, ולכן נבדוק את הביטוי כאשר n שואף לאינסוף, דהינו שהמספרים שאפשר להגדיל הולכים וגדלים בכמותם:

$$\lim_{n\to\infty} \mathbb{P}(B_{n/2}) = \lim_{n\to\infty} \frac{1}{n} \sum_{m=\lceil \frac{n}{2} \rceil}^{n} \frac{1 + m - \frac{n}{2}}{m}$$

נבחין כי $\sum_{m=1}^{n} \frac{1}{m} = \log(n) + e + o(\frac{1}{m})$ ולכן

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \sum_{m=\lceil \frac{n}{2} \rceil}^{n} \frac{1 + m - \frac{n}{2}}{m} &= \lim_{n\to\infty} \frac{1}{2} + \frac{n}{2n}(\log(n) - \log(\frac{n}{2}) + o(\frac{1}{n})) + \frac{1}{n}(\log(n) - \log(\frac{n}{2}) + o(\frac{1}{n})) \\ &= \frac{1}{2} + \frac{1}{n}\log 2 \end{aligned}$$

6 שיעור 4 — 7.11.2024

בשיעור הקודם דיברנו על מרחבי מכפלה וניסויים דו־שלביים. ברור לנו כי על־ידי שרשור דומה לתהליך של ניסוי דו־שלבי נוכל לבנות ניסוי רב־שלבי. עוד דיברנו על חסם האיחוד, הטענה כי $\mathbb{P}(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i)$. השימוש של חסם האיחוד מאפשר לנו לפשט חישובים שבהם אנחנו רוצים הבנה כללית של ההתנהגות של מרחב ההסתברות.

6.1 חסמי איחוד ורציפות

הגדרה 6.1 (סדרת מאורעות עולה) סדרת מאורעות $\{A_n\}_{n=1}^\infty$ נקראת עולה אם $A_n \subseteq A_{n+1}$ לכל $n \in \mathbb{N}$.

סימון 6.2 נסמן $A_\infty = \bigcup_{n\in\mathbb{N}} A_n$.

משפט 6.3 (משפט רציפות פונקציית ההסתברות) *אם $\{A_n\}_{n=1}^\infty$ סדרת מאורעות עולה אז*

$$\mathbb{P}(A_\infty) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

המשפט נקרא כך בשל ההקבלה שלו לקונספט של רציפות בפונקציות רגילות, עבור $f : \mathbb{R} \to \mathbb{R}$ היא רציפה ב־a אם ורק אם לכל סדרה $x_n \to a$ מתקיים $f(x_n) \to f(a)$.

הוכחה. נגדיר $B_n = A_n \setminus A_{n-1}$ כאשר $B_1 = A_1 \setminus \emptyset = A_1$.
נראה כי מתקיים $\biguplus_{n=1}^m B_n = A_m$ איחוד זר:
$\bigcup_{n=1}^m B_n = A_m$ כי לכל $\omega \in A_m$ יש n מינימלי כך ש־$\omega \in A_n$, אבל $\omega \notin A_{n-1}$, לכן נוכל להסיק כי $\omega \in A_n \setminus A_{n-1} = B_n$.
אם $\omega \in B_n = A_n \setminus A_{n-1}$ אז $\omega \notin A_{n-1}$ ולכן $\omega \notin A_k$ לכל $k < n$. מסיגמא־אדיטיביות נסיק

$$\sum_{n=1}^m \mathbb{P}(B_n) = \mathbb{P}(A_m)$$

וגם

$$\sum_{n=1}^\infty \mathbb{P}(B_n) = \mathbb{P}(\biguplus_{n=1}^\infty B_n) = \mathbb{P}(\bigcup_{m=1}^\infty \left(\biguplus_{n=1}^\infty B_n\right)) = \mathbb{P}(\bigcup_{m=1}^\infty A_m)$$

מצד שני מהגדרת הגבול

$$\sum_{n=1}^\infty \mathbb{P}(B_n) = \lim_{m\to\infty} \sum_{n=1}^m \mathbb{P}(B_n) = \lim_{m\to\infty} \mathbb{P}(A_m)$$

□

הגדרה 6.4 (סדרת מאורעות יורדת) נגדיר סדרת מאורעות $\{A_n\}_{n=1}^\infty$ כך שמתקיים $A_{n+1} \subseteq A_n$ לכל $n \in \mathbb{N}$.

נוכל להסיק מהעובדה שמשלים של סדרה עולה הוא סדרה יורדת ונקבל

טענה 6.5

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

טענה 6.6 (חסם האיחוד הבן־מניה) *אם $\{A_n\}_{n=1}^\infty$ סדרת מאורעות אז מתקיים*

$$\mathbb{P}(\bigcup_{n=1}^\infty A_n) \le \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$$

הוכחה. נגדיר $B_m = \bigcup_{n=1}^m A_n$, זוהי סדרה עולה ולכן

$$\mathbb{P}(\bigcup_{n=1}^\infty A_n) = \mathbb{P}(\bigcup_{m=1}^\infty B_m) = \lim_{m\to\infty} \mathbb{P}(B_m) \le \lim_{m\to\infty} \sum_{n=1}^m \mathbb{P}(A_n) = \sum_{n=1}^\infty \mathbb{P}(A_n)$$

□

6.2 עיקרון ההכלה וההדחה

טענה 6.7 *אם A, B מאורעות אז*

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

הוכחה. נגדיר $C = A \setminus B, D = A \cap B, E = B \setminus A$, נקבל

$$A = C \uplus D, \quad B = D \uplus E, \quad A \cup B = C \uplus D \uplus E$$

ונקבל

$$\mathbb{P}(A) = \mathbb{P}(C) + \mathbb{P}(D), \quad \mathbb{P}(D \cup B) = \mathbb{P}(D) + \mathbb{P}(E)$$

ולכן

$$\mathbb{P}(A \cup B) = \mathbb{P}(C) + \mathbb{P}(D) + \mathbb{P}(E)$$

□

משפט 6.8 (הכלה והפרדה לשלושה מאורעות) *עבור שלושה מאורעות A, B, C:*

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

משפט 6.9 (הכלה והפרדה ל־n מאורעות) *יהיו $A_1, \dots, A_n$ מאורעות, אז*

$$\mathbb{P}(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(A_i \cap A_j) + \sum_{i=1}^n \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} \mathbb{P}(A_i \cap A_j \cap A_k) + \dots$$

אם נגדיר $A_I = \bigcap_{i \in I} A_i$ לכל $I \subseteq [n]$ אז נקבל

$$\mathbb{P}(\bigcup_{n=1}^n A_i) = \sum_{k=1}^n (-1)^{k+1} \sum_{\substack{I \subseteq [i] \\ |I|=k}} \mathbb{P}(A_I) = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mathbb{P}(A_I)$$

את משפט זה נוכיח בהמשך הקורס.

נראה דוגמה לבעיה קלאסית במקרים אלה.

תרגיל 6.1 (בעיית ההתאמה) מחלקים n מעטפות ל־n תיבות דואר, אחת לכל תיבה, מה ההסתברות שאף מכתב לא הגיע ליעדו?

פתרון נגדיר $\Omega = S_n$ מרחב אחיד. $A = \{\omega \in \Omega \mid \forall i, \omega(i) \neq i\}$.

נבחן את המשלים, $A^C = \{\omega \in \Omega \mid \exists i, \omega(i) = i\} = \bigcup_{i=1}^n A_i$ עבור $A_i = \{\omega \in \Omega \mid \omega(i) = i\}$. נחשב

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

במקרה של חיתוך $\mathbb{P}(A_i \cap A_j)$ עבור $j < i$ נקבל

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

נוכל להמשיך את התהליך הזה, ונקבל

$$\mathbb{P}(A_I) = \frac{|\bigcap_{i \in I} A_i|}{|\Omega|} = \frac{(n-|I|)!}{n!} = \frac{1}{n(n-1)(n-2)\cdots(n-(I+1))}$$

כעת נותר להשתמש בנוסחה להכלה והדחה, ונקבל

$$\mathbb{P}(\bigcup_{i=1}^n A_i) = \sum_{k=1}^n (-1)^{k+1} \sum_{\substack{I \subseteq [n] \\ |I|=k}} \frac{(n-k)!}{n!} = \sum_{k=1}^n (-1)^{k+1} \binom{n}{k} \frac{(n-k)!}{n!} = \sum_{k=1}^n \frac{(-1)^{k+1}}{k!}$$

נשים לב כי רצינו לחשב את המשלים למאורע, לכן

$$\mathbb{P}(A) = 1 - \mathbb{P}(\bigcup_{i=1}^n A_i) = 1 + \sum_{k=1}^n \frac{(-1)^k}{k!} = \sum_{k=0}^n \frac{(-1)^k}{k!} \xrightarrow[n \to \infty]{} e^{-1}$$

נקבל שאוסף התמורות ללא נקודת שבת הוא

$$|A^n| = n!\sum_{l=0}^{n}\frac{(-1)^l}{l!}$$

נגדיר קבוצה חדשה

$$D_k = \{\omega \in S_n \mid \exists i, \omega(i) = i\} = \biguplus_{\substack{I\subseteq[n]\\|I|=k}} D_I$$

ונבחין כי

$$D_I = \{\omega \in S_n \mid \forall i \in I, \omega(i) = i, \forall i \notin I, \omega(i) \neq i\}$$

ולכן

$$\begin{aligned}
\mathbb{P}(D_k) &= \sum_{\substack{I\subseteq[n]\\|I|=k}} \mathbb{P}(D_I) \\
&= \sum_{\substack{I\subseteq[n]\\|I|=k}} \frac{|D_I|}{n!} \\
&= \sum_{\substack{I\subseteq[n]\\|I|=k}} \frac{(n-k)!\sum_{l=0}^{n}\frac{(-1)^l}{l!}}{n!} \\
&= \binom{n}{k}\frac{(n-k)!}{n!}\sum_{l=0}^{n}\frac{(-1)^l}{l!} \\
&= \frac{1}{k!}\sum_{l=0}^{n}\frac{(-1)^l}{l!} \\
&\xrightarrow[n\to\infty]{} \frac{e^{-1}}{k!}
\end{aligned}$$

7 שיעור 5 — 12.11.2024

7.1 הסתברות מותנית

הגדרה 7.1 (הסתברות מותנית) A, B מאורעות, ההסתברות המותנית של A בהינתן B תוגדר להיות

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

דוגמה 7.1 אם מטילים שתי קוביות מאוזנות, מה ההסתברות שיצא 3 בקוביה הראשונה בהינתן שהסכום הוא 8?

נגדיר כמובן $\Omega = [6]^2$, וכן נגדיר $A = \{(3, i) \in \omega \mid 1 \le i \le 6\}$ וכן $B = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$.

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5}$$

טענה 7.2 *נקבע מאורע B עם הסתברות $\mathbb{P}(B) > 0$, נגדיר $\mathbb{P}_B(A) = \mathbb{P}(A \mid B)$, דהינו $\mathbb{P}_B : \mathcal{F} \to [0, \infty)$.*
אז $\mathbb{P}_B$ היא פונקציית הסתברות.

הוכחה. $\mathbb{P}_B(A)$ היא אי־שלילית.
נראה גם

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

ולבסוף

$$\mathbb{P}_B(\biguplus_{i \in I} A_i) = \frac{(\mathbb{P}_B(\biguplus_{i \in I} A_i)) \cap B}{\mathbb{P}(B)} = \frac{\mathbb{P}_B(\biguplus_{i \in I} A_i \cap B)}{\mathbb{P}(B)} = \sum_{i \in I} \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_{i \in I} \mathbb{P}_B(A_i)$$

□

טענה 7.3 *יהיו C, B מאורעות המקיימים $\mathbb{P}(B \cap C) > 0$, נסמן $\mathbb{P}' = \mathbb{P}_B$ ו־$\mathbb{P}'' = \mathbb{P}'_C$.*
אז לכל מאורע A מתקיים $\mathbb{P}''(A) = \mathbb{P}(A \mid B \cap C)$ או בחילוף סימונים $\mathbb{P}'' = \mathbb{P}_{B \cap C}$.

הוכחה.

$$\mathbb{P}''(A) = \mathbb{P}'_C(A) = \frac{\mathbb{P}'(A \cap C)}{\mathbb{P}'(C)} = \frac{\mathbb{P}_B(A \cap C)}{\mathbb{P}_B(C)} = \frac{\frac{\mathbb{P}(B \cap (A \cap C))}{\mathbb{P}(B)}}{\frac{\mathbb{P}(B \cap C)}{\mathbb{P}(B)}} = \mathbb{P}_{B \cap C}(A)$$

□

מצאנו כי התניה חוזרת היא אסוציאטיבית ולכן נוכל לדבר על הסתברות מותנית בכמה מאורעות ללא התייחסות לסדר שלהם, למעשה התנייה מותנית היא קומוטטיבית כפי שאפשר לראות בהוכחה.

מסקנה 7.4 (נוסחת ההסתברות השלמה בהסתברות מותנית) נניח ש־$\{A_i\}_{i \in \mathbb{N}}$ חלוקה בת־מניה של Ω ו־B מאורע כלשהו, אז

$$\mathbb{P}(B) = \sum_{i \in \mathbb{N}} \mathbb{P}(A_i)\mathbb{P}(B \mid A_i)$$

הוכחה.

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

ולכן

$$\biguplus_{i \in \mathbb{N}} (B \cap A_i) = B \implies \mathbb{P}(B) = \sum_{i \in \mathbb{N}} \mathbb{P}(B \cap A_i)$$

□

למה 7.5 (כלל בייס) *אם A, B מאורעות עם הסתברות חיובית אז*

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

הוכחה. ישירות מהגדרה נסיק

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

□

מסקנה 7.6 (כלל השרשרת)

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

תרגיל 7.1 מטילים מטבע הוגן. אם יוצא עץ נוסעים לתל־אביב ואם יוצא פלי אז נוסעים לחיפה. כשנוסעים לתל־אביב יש הסתברות של אחוז אחד לפנצ'ר, ובנסיעה לחיפה יש הסתברות של 2 אחוז לפנצ'ר.
מה ההסתברות לפנצ'ר ומה ההסתברות שנסעו לתל־אביב?

פתרון נגדיר A הוא עץ או לנסוע לתל־אביב ו־B ההסתרות שיהיה פנצ'ר, בהתאם

$$\mathbb{P}(A^C) = \mathbb{P}(A) = \frac{1}{2}, \qquad \mathbb{P}(B \mid A) = 0.01, \mathbb{P}(B \mid A^C) = 0.02$$

בהתאם

$$\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B \mid A) + \mathbb{P}(A^C) + \mathbb{P}(B \mid A^C) = \frac{1}{2}0.01 + \frac{1}{2}0.02 = 0.015$$

באשר לשאלה השנייה נקבל

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)}\mathbb{P}(B \mid A) = \frac{\frac{1}{2}}{0.015} \cdot 0.01 = \frac{1}{3}$$

נבחין כי התוצאה יצאה מאוד אלגנטית כתוצאה מהמטבע ההוגן, אילו הוא היה לא הוגן היינו מקבלים חישוב שונה במקצת, אך תקף באותה המידה.

דוגמה 7.2 (מונטי הול) *יש שלוש דלתות, בוחרים אחת, מנחה פותח דלת שלא נבחרה ומאחוריה אין כלום, מה שאומר שמאחורי אחת הדלתות הסגורות יש אוצר ובאחרות יש עז. המנחה מציע לכם להחליף את הדלת שבחרתם.*

קשה למדל את הבעיה הזו, שכן חסר תיאור והגדרה, אז נאמר שהגרלנו מספר ב־$[3]$, נניח שבחרנו 1, נניח שהמנחה גם במכוון תמיד בוחר דלת ריקה. נוסיף את ההנחה שאם האוצר מאחורי דלת 1 אז המנחה פותח את 2 או 3, וההסתברויות שוות.

נעבור להגדרה, A_i המאורע שהאוצר ב־i ו־B_i היא שהמנחה פותח את דלת i. מההנחות שלנו נובע $\mathbb{P}(B_3 \mid A_2) = 1, \mathbb{P}(B_2 \mid A_3) = 1, \mathbb{P}(B_3 \mid A_1) = \mathbb{P}(B_2 \mid A_1) = \frac{1}{2}$. לבסוף נניח כי $\mathbb{P}(A_i) = \frac{1}{3}$ לכל שלוש הדלתות.
נרצה לחשב את $\mathbb{P}(A_1 \mid B_2)$:

$$\mathbb{P}(A_1 \mid B_2) = \frac{\mathbb{P}(A_1)}{\mathbb{P}(B_2)} \cdot \mathbb{P}(B_2 \mid A_1) = \frac{\frac{1}{6}}{\mathbb{P}(B_2)}$$

וגם

$$\mathbb{P}(B_2) = \mathbb{P}(A_1)\mathbb{P}(B_2 \mid A_1) + \mathbb{P}(A_2)\mathbb{P}(B_2 \mid A_2) + \mathbb{P}(A_3)\mathbb{P}(B_2 \mid A_3) = \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 1 = \frac{1}{2}$$

8 תרגול 3 — 14.11.2024

8.1 הסתברות מותנית

תרגיל 8.1 מטילים זוג קוביות הוגנות ושונות. נתון שסכום תוצאותיהן גדול מעשר, מה ההסתברות שבהטלה השנייה יצא 6?

פתרון נגדיר $\Omega = [6]^2$ עם $\mathbb{P}$ אחידה.

עוד נגדיר $A = \{(x,y) \in \Omega \mid x+y > 10\}$ וכן $B = \{(x,6) \in \Omega\}$, לכן

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\frac{|A\cap B|}{|\Omega|}}{\frac{|A|}{|\Omega|}} = \frac{|A \cap B|}{|A|} = \frac{2}{3}$$

תרגיל 8.2 אדם מחפש מכתב, זכור לו במעורפל בהסתברות $0 \le p \le 1$ שהניח אותו באחת ממגירות שולחן העבודה.

בשולחן n מגירות והאדם חיפש ב־k המגירות הראשונות ולא מצא את המכתב.

מה ההסתברות שהמכתב בשולחן?

פתרון נגדיר A להיות המאורע שהמכתב בשולחן ו־B_k המכתב לא באף אחת מ־k המגירות הראשונות. אנו מחפשים את $\mathbb{P}(A \mid B_k)$.

לכן

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

עוד אנו יודעים כי

$$\mathbb{P}(A) = p, \mathbb{P}(B_k) = 1 - \frac{kp}{n}$$

אזי

$$\frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(B_k)} = \frac{\frac{(n-k)p}{n}}{\frac{n-kp}{n}} = \frac{(n-k)p}{n-kp}$$

תרגיל 8.3 האדם הוא מתודי והחליט להפסיק את החיפוש אם ההסתברות שהמכתב בשולחן קטנה מ־$\frac{1}{4}$.

נניח ש־$p = \frac{3}{4}$ ושיש 10 מגירות, כמה מגירות תיבדקנה לכל היותר עד שהאדם יפסיק את החיפוש?

פתרון

$$\frac{1}{4} > \mathbb{P}(A \mid B_k) = \frac{(10-k)\frac{3}{4}}{10 - \frac{3k}{4}} \iff k > \frac{89}{11}$$

נבדוק לכל היותר 8 מגירות.

8.2 ניסוי דו־שלבי על־ידי הסתברות מותנית

טענה 8.1 *נניח שנתון ניסוי דו־שלבי על $\Omega_1 \times \Omega_2$ עם פונקציית הסתברות נקודתית p על Ω_1 ולכל $\omega \in \Omega_1$ גם p_ω היא פונקציית הסתברות נקודתית על Ω_2.*

אם $\mathbb{P}$ היא פונקציה על $\Omega_1 \times \Omega_2$ המקיימת

$$\mathbb{P}(\{a,x\}) = p(a), \qquad \mathbb{P}(\{x,b\} \mid \{(a,x)\}) = p_a(b)$$

אז $\mathbb{P}$ היא פונקציית הסתברות יחידה המתאימה לניסוי הדו־שלבי.

הוכחה. יהי $(a,b) \in \Omega_1 \times \Omega_2$, מכלל השרשרת נובע

$$\mathbb{P}(\{(a,b)\}) = \mathbb{P}(\{(a,x)\}) \cdot \mathbb{P}(\{(x,b)\} \mid \{(a,x)\}) = p(a) \cdot p_a(b) = q(a,b)$$

עבור q פונקציית הסתברות נקודתית של $\mathbb{P}$. $\square$

נבחין שוב כי בעוד כל ניסוי דו־שלבי, ניתן לבחון אותו כניסוי מותנה, הכיוון ההפוך לא בהכרח מתקיים; לא כל ניסוי מותנה הוא ניסוי דו־שלבי. נבחן דוגמות לשימוש בקשר זה.

תרגיל 8.4 בשוק ישנם שלושה סוגי מחשבים. חצי מסוג ראשון, 30% מסוג שני ו־20% מסוג שלישי.

הסיכוי שמחשב מסוג ראשון יתקלקל בשנתו הראשונה הוא עשירית, הסיכוי לסוג שני הוא חמישית והסיכוי למחשב מהסוג השלישי הוא $\frac{1}{20}$.

קונים מחשב באקראי מבין מחשבי השוק, מה ההסתברות שהוא יתקלקל בשנתו הראשונה?

פתרון נסמן C_j הוא המאורע שקנינו מחשב מסוג j ו־B המאורע שהמחשב התקלקל בשנתו הראשונה.
עוד נתון $\mathbb{P}(C_1) = \frac{l}{2}, \mathbb{P}(C_2) = \frac{3}{10}, \mathbb{P}(C_3) = \frac{1}{5}$.
נתונים לנו גם $\mathbb{P}(B \mid C_1) = \frac{1}{10}, \mathbb{P}(B \mid C_2) = \frac{1}{5}, \mathbb{P}(B \mid C_3) = \frac{1}{20}$. מנוסחת ההסתברות השלמה נובע

$$\mathbb{P}(B) = \mathbb{P}(B \mid C_1)\mathbb{P}(C_1) + \mathbb{P}(B \mid C_2)\mathbb{P}(C_2) + \mathbb{P}(B \mid C_3)\mathbb{P}(C_3)$$

תרגיל 8.5 במבחן אמריקאי לכל שאלה 4 אפשרויות ובדיוק 1 נכונה. סטודנטית ניגשת למבחן עם האסטרטגיה הבאה:

- אם היא יודעת את התשובה היא עונה נכונה.
- אם היא לא יודעת את התשובה אז היא בוחרת תשובה אקראית.

נתון כי הסטודנטית יודעת את התשובה ל־90% משאלות הבחינה.

בוחרים שאלה באקראי, ונתון שהסטודנטית ענתה עליה נכון, מה ההסתברות שהיא ידעה את התשובה.

פתרון נסמן ב־A את המאורע שהסטודנטית ידעה את התשובה, וב־B את המאורע שהסטודנטית ענתה נכון.
אנו יודעים כי $\mathbb{P}(A) = \frac{9}{10}$ וגם כי $\mathbb{P}(B \mid A) = 1, \mathbb{P}(B \mid A^C) = \frac{1}{4}$.

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A)}{\mathbb{P}(B)}\mathbb{P}(B \mid A) = \frac{\frac{9}{10} \cdot 1}{\mathbb{P}(B \mid A) \cdot \mathbb{P}(A) + \mathbb{P}(B \mid A^C) \cdot \mathbb{P}(A^C)} = \frac{\frac{9}{10}}{\frac{9}{10} + \frac{1}{4} \cdot \frac{1}{10}} = \frac{\frac{9}{10}}{\frac{37}{40}} \approx 0.973$$

9 שיעור 6 — 14.11.2024

9.1 אי־תלות

הגדרה 9.1 (מאורעות בלתי־תלויים) מאורעות A, B המקיימים $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ יקראו **בלתי־תלויים**.

הערה נובע שמתקיים

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

הערה (תזכורת) אם $A \subseteq \Omega_1$ ו־$B \subseteq \Omega_2$ ועובדים עם $\mathbb{P}$ פונקציית הסתברות של מרחב המכפלה $\Omega_1 \times \Omega_2$.
אז ראינו שמתקיים $\mathbb{P}(A \times B) = \mathbb{P}_1(A) \cdot \mathbb{P}_2(B) = \mathbb{P}(A \times \Omega_2) \cdot \mathbb{P}(\Omega_1 \times B)$

דוגמה 9.1 מטילים שתי קוביות, אז $\Omega = [6]^2$.
נגדיר $A = \{4\} \times [6]$ המאורע שיצא 4 בקוביה הראשונה ו־B המאורע שסכום הקוביות הוא 7.

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{1}{6}, \qquad \mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{1}{6}$$

וחישוב היתוך המאורעות יניב

$$\mathbb{P}(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{|\{(4,3)\}|}{36} = \frac{1}{36} = \mathbb{P}(A)\mathbb{P}(B)$$

אז המאורעות בלתי־תלויים.

טענה 9.2 *1. לכל מאורע A, A ו־$\emptyset$ בלתי־תלויים וכן A ו־Ω בלתי־תלויים.*

2. אם A ו־B בלתי־תלויים ו־$\mathbb{P}(B) > 0$ אז $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

3. אם A ו־B בלתי־תלויים אז גם A^C ו־B בלתי תלויים.

הוכחה. נוכיח את הטענה השלישית

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

במעבר הראשון השתמשנו בנוסחת ההסתברות השלמה על החלוקה A, A^C. □

נראה הגדרה לאי־תלות במספר מאורעות, אך לא ההגדרה שאנו רוצים לעבוד איתה.

הגדרה 9.3 (אי־תלות בזוגות) אם $A_1, \dots, A_n$ נקראים בלתי תלויים בזוגות אם

$$\forall 1 \le i < j \le n, \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

הגדרה 9.4 (קבוצה בלתי־תלויה) מאורע A נקרא בלתי־תלוי בקבוצת המאורעות $B_1, \dots, B_n$ אם לכל $I \subseteq [n]$ מתקיים

$$\mathbb{P}(A \mid \bigcap_{i \in I} B_i) = \mathbb{P}(A)$$

דהינו A ו־$\bigcap_{i \in I} B_i$ בלתי־תלוי.

תרגיל 9.1 הביאו דוגמה למאורעות A, B_1, B_2 כך ש־A ו־B_1 בלתי־תלויים וכך גם A ו־B_2 אבל A לא בלתי־תלוי בקבוצה $\{B_1, B_2\}$.
הראו כי אם A, B_1 בלתי־תלויים וגם A, B_2 בלתי־תלויים וגם B_1, B_2 זרים, אז A בלתי־תלוי ב־$\{B_1, B_2\}$.

טענה 9.5 *A בלתי־תלוי ב־$\{B_1, \dots, B_n\}$ אם ורק אם A בלתי תלוי ב־$\{B_1, \dots, B_n, B_1^C, \dots, B_n^C\}$.*

הוכחה. הכיוון השני הוא טריוויאלי, לכן נוכיח את הכיוון הראשון בלבד.
נראה ש־A בלתי־תלוי בקבוצה $\{B_1, \dots, B_n, B_1\}^C$. נרצה להראות שלכל $I \subseteq [n+1]$ מתקיים ש־A ו־$\bigcap_{i \in I} B_i$ בלתי־תלויים.
אם $n+1 \notin I$ אז לפי ההנחה חוסר התלות כבר מתקיים.

אחרת נגדיר $J = I \setminus \{n+1\}$ ולכן $I = J \uplus \{n+1\}$, ומכאן נובע

$$\begin{aligned}
\mathbb{P}((\bigcap_{i\in I} B_i) \cap A) &= \mathbb{P}((\bigcap_{i\in J} B_i) \cap B_1^C \cap A) \\
&= \mathbb{P}(\bigcap_{i\in J} B_i \cap A) - \mathbb{P}(\bigcap_{i\in J} B_i \cap B_1 \cap A) \\
&= \mathbb{P}(\bigcap_{i\in J} B_i)\mathbb{P}(A) - \mathbb{P}(\bigcap_{i\in J} B_i \cap B_1)\mathbb{P}(A) \\
&= \mathbb{P}(\bigcap_{i\in J} B_i \cap B_1^C)\mathbb{P}(A) \\
&= \mathbb{P}(\bigcap_{i\in I} B_i)\mathbb{P}(A)
\end{aligned}$$

ומצאנו כי ניתן להוסיף איבר, בשל כך נוכל לבצע את התהליך איטרטיבית ולקבל את המבוקש. □

הגדרה 9.6 (אי־תלות קבוצת מאורעות) קבוצת מאורעות $\{A_1, \dots, A_n\}$ נקראת בלתי־תלויה אם לכל $I \subseteq [n]$ מתקיים

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

מסקנה 9.7 אם $A_1, \dots, A_n$ בלתי־תלויים, אז גם כל תת־קבוצה של מאורעות היא בלתי־תלויה.

בפרט $A_1, \dots, A_n$ בלתי־תלויים גורר ש־$A_1, \dots, A_n$ בלתי־תלויים בזוגות.

טענה 9.8 *קבוצת מאורעות $\{A_1, \dots, A_n\}$ בלתי־תלויה אם ורק אם לכל $i \in [n]$ מתקיים ש־A_i בלתי־תלויה ב־$\{A_1, \dots, A_n\} \setminus \{A_i\}$.*

הוכחה. לכיוון הראשון, נרצה להראות ש־A_1 לא תלוי ב־$\{A_2, \dots, A_n\}$, כלומר לכל $I \subseteq \{2, \dots, n\}$ רוצים להראות ש־$\mathbb{P}(\bigcap_{i\in I} A_i \cap A_1) = \mathbb{P}(\bigcap_{i\in I} A_i)\mathbb{P}(A_1)$ על־ידי

$$\mathbb{P}(\bigcap_{i\in I} A_i \cap A_1) = (\prod_{i\in I} \mathbb{P}(A_i))\mathbb{P}(A_1) = \mathbb{P}(\bigcap_{i\in I} A_i)\mathbb{P}(A_1)$$

נעבור לכיוון השני. צריך להראות שלכל $I \subseteq [n]$ מתקיים $\mathbb{P}(\bigcap_{i\in I} A_i) = \prod_{i\in I} \mathbb{P}(A_i)$. תהי $I \subseteq [n] = \{i_1, \dots, i_k\}$ כאשר $|I| = k$.

לפי ההנחה A_{i_1} בלתי־תלוי ב־$\{A_j \mid j \in [n] \setminus \{i_1\}\}$, לכן נקבל באינדוקציה

$$\mathbb{P}(\bigcap_{l=1}^{k} A_{i_l}) = \mathbb{P}(A_{i_1} \cap (\bigcap_{l=2}^{k} A_{i_l})) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(\bigcap_{l=2}^{k} A_{i_l}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2})\mathbb{P}(\bigcap_{l=3}^{k} A_{i_l}) = \cdots = \mathbb{P}(A_{i_1}) \cdots \mathbb{P}(A_{i_k})$$

□

10 שיעור 7 — 19.11.2024

10.1 אי־תלות

נראה הגדרה שקולה לאי־תלות

הגדרה 10.1 (שקולה לאי־תלות) $A_1, \ldots, A_n$ בלתי־תלויים אם ורק אם

$$\forall I \subseteq [n], \mathbb{P}((\bigcap_{i\in I} A_i) \cap (\bigcap_{i\in[n]\setminus I} A_i^C)) = \prod_{i\in I} \mathbb{P}(A_i) \prod_{i\in[n]\setminus I} \mathbb{P}(A_i^C)$$

את השקילות של ההגדרות נראה בתרגיל.

מאורעות $A_1, \ldots, A_n$ בלתי־תלויים בהינתן B אם הם בלתי־תלויים לפי פונקציית ההסתברות המותנית ב־B, $\mathbb{P}_B$.

דוגמה 10.1 בוחרים מטבע באקראי משק ומטילים אותו n פעמים.

A_i יצא עץ בהטלה ה־i בלתי־תלוי בהינתן בחירת המטבע, B המאורע שנבחר מטבע מסוים.

נרצה לנסות לתת הגדרה חדשה עבור מקרים אינסופיים, נראה שיתקיים

$$\forall I \subseteq \mathbb{N}, \mathbb{P}(\bigcap_{i\in I} A_i = \prod_{i\in I} \mathbb{P}(A_i))$$

אבל היא לא מועילה לנו, נגדיר במקום זאת

הגדרה 10.2 (קבוצה בת־מניה בלתי־תלויה) $A_1, A_2, \ldots$ מאורעות הם בלתי־תלויים אם ורק אם לכל קבוצה סופית I מתקיים $\{A_i\}_{i\in I}$ קבוצה בלתי־תלויה.

הערה (מכפלה אינסופית) נגדיר מכפלה אינסופית על־ידי

$$\prod_{i\in\mathbb{N}} a_i = \prod_{i=1}^{\infty} a_i = \lim_{N\to\infty} \prod_{i=1}^{N} a_i$$

טענה 10.3 *אם $A_1, A_2, \ldots$ סדרת מאורעות בלתי־תלויים אז*

$$\mathbb{P}(\bigcap_{i\in\mathbb{N}} A_i) = \prod_{i\in\mathbb{N}} \mathbb{P}(A_i)$$

הוכחה. נגדיר $B_n = \bigcap_{i=1}^n A_i$ סדרה יורדת ולכן מרציפות פונקציית ההסתברות נובע

$$\mathbb{P}(\bigcap_{i\in\mathbb{N}} A_i) = \mathbb{P}(\bigcap_{n\in\mathbb{N}} B_n) = \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{N\to\infty} \prod_{i=1}^{N} \mathbb{P}(A_i) = \prod_{i=1}^{\infty} \mathbb{P}(A_i)$$

□

דוגמה 10.2 אם $A_1, \ldots$ בלתי־תלויים ו־$\mathbb{P}(A_i) = p < 1$ אז $\mathbb{P}(\bigcap_{i\in\mathbb{N}} A_i) = 0$.
לדוגמה בהטלה אינסוף פעמים של מטבע הסיכוי שייצא עץ הוא אפס. דוגמה זו קצת בעייתית שכן כלל לא הראינו כי מרחב זה קיים ומוגדר, אבל המשמעות היא שעבור מרחבי מדגם הולכים וגדלים, אז ההסתברות המבוקשת שואפת להיות אפס.

10.2 משתנים מקריים

עד כה היינו צריכים לבצע ניתוח מלא של הסיטואציה כדי להגיע למסקנה, גם אם בהרבה מקרים שונים הגענו לבדיוק אותה המסקנה, המטרה של משתנים מקריים הוא לבודד את הרעיון הזה ולתקוף אותו.

הגדרה 10.4 (משתנה מקרי) יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, פונקציה מ־Ω ל־$\mathbb{R}$ נקראת **משתנה מקרי**.

סימון 10.5 על־אף שזו פונקציה, נהוג לסמן משתנים מקריים בסימונים שאנו רגילים שמשמשים למשתנים, לדוגמה X, Y, Z.

הערה השם קצת מטעה, אלו הם לא משתנים, ושווה לחשוב עליהם בתור מצבים מקריים יותר.

דוגמה 10.3 נניח $\Omega = \{H, T\}$ הטלת מטבע, ונגדיר את הפונקציה $f : \Omega \to \mathbb{R}$ על־ידי $f(H) = 2, f(T) = -3$, זאת במטרה לייצג שאם יוצא עץ נפסיד שלושה מטבעות ואם מתקבל פלי אז נקבל שני מטבעות.

דוגמה 10.4 נרצה להטיל שתי קוביות ונרצה לדבר על תוצאת אחת ההטלות, נתחיל ונגדיר $\Omega = [6]^2$.

נגדיר $X_1 : \Omega \to \mathbb{R}$ על־ידי $X_1(a,b) = a$, ובאופן דומה נגדיר $X_2(a,b) = b$. יצרנו פונקציות שמהוות משתנה מקרי עבור ההטלה הראשונה ועבור ההטלה השנייה, נגדיר גם עבור הסכום, $Y(a,b) = a+b$.
נקבל עתה $Y = X_1 + X_2$, ונבחין בכוח האמיתי של הגדרה זו, יש לנו איזשהו קישור מורכב במרחב ההסתברות ללא עבודה ישירות מול המרחב.

הגדרה 10.6 (משתנה מקרי מושרה ממאורע) אם A מאורע אז נגדיר 1_A משתנה מקרי על־ידי

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

טענה 10.7 **(תכונות של משתנים מקריים מושרים)**

1. $1_{A^C} = 1 - 1_A$
2. $1_{A \cap B} = 1_A \cdot 1_B$
3. $1_{A \cup B} = \max\{1_A, 1_B\}$

דוגמה 10.5 $\Omega = S_n$, A_i המאורע שיש i נקודות שבת.
נסמן $X_i = 1_{A_i}$ ו־$X = \sum_{i=1}^n X_i$.

בדוגמות הקודמות ההטלה הראשונה זוגית $\{(a,b) \in [6]^2 \mid a \in \{2,4,6\}\}$. נכתוב במקום זאת $X_1 \in \{2,4,6\}$.

הגדרה 10.8 (מאורע מושרה ממשתנה מקרי) אם X משתנה מקרי ו־$S \in \mathcal{F}_\mathbb{R}$, המאורע $X \in S$ מוגדר להיות

$$X^{-1}(S) = \{\omega \in \Omega \mid X(\omega) \in S\}$$

בהתאם נכתוב $\mathbb{P}(\{x \in S\})$ על־ידי $\mathbb{P}(X \in S)$, ובאופן דומה נוכל לציין גם את $\mathbb{P}(X = s), \mathbb{P}(X \le s)$ ודומים.

הגדרה 10.9 (פונקציית הסתברות מושרית ממשתנה מקרי) $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, ויהי X משתנה מקרי.
נגדיר פונקציה $\mathbb{P}_X : \mathcal{F}_\mathbb{R} \to [0, \infty)$ על־ידי

$$\mathbb{P}_X(S) = \mathbb{P}(X \in S) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in S\})$$

$\mathbb{P}_X$ מכונה **ההתפלגות** של X.
אם $\mathbb{P}_X$ נתמכת על S (כלומר $\mathbb{P}_X(S) = 1$) אז אומרים ש־X נתמך על S.

טענה 10.10 *$\mathbb{P}_X$ היא פונקציית הסתברות על $(\mathbb{R}, \mathcal{F}_\mathbb{R})$.*

הוכחה.

$$\forall S, \mathbb{P}_X(S) = \mathbb{P}(X \in S) \ge 0$$

וכן

$$\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(\Omega) = 1$$

ולבסוף סיגמא־אדיטיביות:

$$\begin{aligned} \forall S_1, S_2, \dots, \mathbb{P}_X(\biguplus_{n \in \mathbb{N}} S_n) &= \mathbb{P}(X \in \biguplus_{n \in \mathbb{N}} S_n) \\ &= \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in \biguplus_{n \in \mathbb{N}} S_n\}) \\ &= \mathbb{P}(\biguplus_{n \in \mathbb{N}} \{X \in S_n\}) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(X \in S_n) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}_X(S_n) \end{aligned}$$

□

11 תרגול 4 — 21.11.2024

11.1 אי־תלות

נניח מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$.

תרגיל 11.1 בכד שלושה מטבעות, שניים הוגנים ואחד שמוטבע עץ על שני צדדיו.

שולפים מטבע באקראי ואז מטילים אותו פעמיים.

האם תוצאת ההטלה הראשונה תלויה בתוצאת ההטלה השנייה?

פתרון נסמן ב־A_i את המאורע שבהטלה ה־i יצא עץ.

אנו שואלים אם A_1, A_2 הם תלויים, נסמן גם F המאורע ששלפנו מטבע הוגן.

$$\mathbb{P}(A_1) = \mathbb{P}(A_1 \mid F)\mathbb{P}(F) + \mathbb{P}(A_1 \mid F^C)\mathbb{P}(F^C) = \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}$$

אנו רוצים לבדוק את התלות ולכן נחשב

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_2 \mid F)\mathbb{P}(F) + \mathbb{P}(A_1 \cap A_2 \mid F^C)\mathbb{P}(F^C) = \frac{1}{4} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{2} \neq \frac{4}{9} = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2)$$

ולכן הם תלויים.

תרגיל 11.2 נגדיר $\Omega = \mathbb{N}$ ו־$\mathbb{P}(\{n\}) = \frac{1}{c \cdot n^2}$, כאשר $c = \sum_{n \in \mathbb{N}} \frac{1}{n^2} = \frac{\pi^2}{6}$.

נגדיר $\forall k \in \mathbb{N}, A_k = k\mathbb{N} = \{kn \mid n \in \mathbb{N}\}$.

האם $\{A_k\}_{k \in \mathbb{N}}$ היא תלויה?

פתרון

$$\mathbb{P}(A_k) = \sum_{n \in \mathbb{N}} \mathbb{P}(\{k_n\}) = \sum_{n \in \mathbb{N}} \frac{1}{ck^2n^2} = \frac{1}{ck^2} \sum_{n \in \mathbb{N}} \frac{1}{n^2} = \frac{1}{k^2}$$

ולכן

$$\mathbb{P}(A_2 \mid A_4) = 1 \neq \frac{1}{4} = \mathbb{P}(A_2)$$

ולכן המאורעות תלויים ובכלל הקבוצה לא בלתי־תלויה.

תרגיל 11.3 נגדיר P קבוצת המספרים הראשוניים, האם $\{A_p\}_{p \in P}$ בלתי־תלויה?

פתרון יהיו $p_1, \dots, p_m \in P$ ראשוניים, אז מהמשפט היסודי של האריתמטיקה (או פירוק לגורמים ראשוניים)

$$A_{p_1} \cap \dots \cap A_{p_m} = A_{p_1 \cdots p_m}$$

ולכן

$$\mathbb{P}(A_{p_1} \cap \dots \cap A_{p_m}) = \mathbb{P}(A_{p_1 \cdots p_m}) = \frac{1}{(p_1 \cdots p_m)^2} = \frac{1}{p_1^2} \cdots \frac{1}{p_m^2} = \mathbb{P}(A_{p_1}) \cdots \mathbb{P}(A_{p_m})$$

נגדיר גם $B = \bigcap_{p \in P} A_p^C = \{1\}$, ונחשב

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

מסקנה 11.1 נוכל להסיק מסקנה משמעותית נוספת, לכל $s > 1$ מתקיים

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \prod_{p \in P} (1 - \frac{1}{ps})^{-1} = \zeta(s)$$

הערך ה־s של פונקציית זטא של רימן, וזו זהות אוילר לפונקציית זטא.

מסקנה 11.2 מסקנה נוספת היא שבשל אי־רציונליות π נוכל להסיק כי הטור הוא לא טור סופי, לכן יש אינסוף ראשוניים.

11.2 משתנים מקריים

אנו רוצים להסתכל על משתנה מקרי כדרך להסתכל מחדש על מרחב ההסתברות ובפרט פונקציית ההסתברות באופן נוסף, זה בתורו יאפשר לנו לפתור בעיות בדרך חדשה ואולי אף פשוטה יותר, כפי שנראה בהמשך.

דוגמה 11.1 נגדיר X להיות משתנה מקרי שמתאר סכום הטלת שתי קוביות הוגנות, דהינו נוכל להגדיר $\mathbb{P} = [6]^2$ ו־$\mathbb{P}$ אחידה.
בהתאם נגדיר $X : \Omega \to \mathbb{R}$ על־ידי $a + b \mapsto (a, b)$.
לכן

$$\operatorname{rng}(X) = \{2, \dots, 12\}$$

נעבור לחישוב הסתברויות

$$\begin{aligned}
\mathbb{P}(X = 2) &= \frac{1}{36}, \\
\mathbb{P}(X = 3) &= \frac{2}{36}, \\
\mathbb{P}(X = 4) &= \frac{3}{36}, \\
\mathbb{P}(X = 5) &= \frac{4}{36}, \\
\mathbb{P}(X = 6) &= \frac{5}{36}, \\
\mathbb{P}(X = 7) &= \frac{6}{36}, \\
\mathbb{P}(X = 8) &= \frac{5}{36}
\end{aligned}$$

וכן הלאה, בהתאם נוכל להסיק

$$\forall E \subseteq \mathbb{R}, \mathbb{P}(X \in E) = \mathbb{P}_X(E \cap \operatorname{rng}(X)) = \sum_{i \in E \cap \operatorname{rng}(X)} \mathbb{P}(X = i)$$

נסמן X_i תוצאת ההטלה ה־i, ולכן $X = X_1 + X_2$, נחשב את X ביחס ל־X_i.

$$\begin{aligned}
\forall n \in \{2, \dots, 12\}, \mathbb{P}(X = n) &= \sum_{i=1}^{6} \mathbb{P}(X_1 = i)\mathbb{P}(X_2 = n - i) \\
&= \sum_{i=1}^{6} \frac{1}{6} \min\{6 - i, 0\} \\
&= \frac{1}{36} |\{\{1, \dots, 6\} \cap \{n - 1, \dots, n - 6\}\}|
\end{aligned}$$

אם נגדיר $Y = X(\mod 6)$ אז $\operatorname{rng}(Y) = \{0, \dots, 5\}$ ונחשב

$$\begin{aligned}
\forall n \in \{0, \dots, 5\}, \mathbb{P}(Y = n) &= \mathbb{P}(X = n \lor X = n + 6 \lor X = n + 12) \\
&= \frac{1}{36} \cdot |\{1, \dots, 6\} \cap \{n + 12, \dots, n - 6\}| \\
&= \frac{6}{36} = \frac{1}{6}
\end{aligned}$$

12 שיעור 8 — 21.11.2024

12.1 משתנים מקריים — המשך

הגדרה 12.1 (משתנה מקרי בדיד) X משתנה מקרי נקרא **בדיד** אם $\mathbb{P}_X$ הוא פונקציית הסתברות בדידה.
במקרה זה יש ל־X התפלגות נקודתית $p_X : \mathbb{R} \to [0, \infty)$.

הערה נבחין כי גם אם מרחב ההסתברות הוא לא בדיד, נוכל להגדיר משתנה מקרי בדיד עליו.

דוגמה 12.1 נגדיר $A \in \mathcal{F}$ ו־$X = 1^A$ ונניח $\mathbb{P}(A) = p$.
אם $S \subseteq \mathcal{F}_\mathbb{R}$ אז אם $\{0,1\} \in S$ אז $\Omega = X^{-1}(S)$ ואז $\mathbb{P}_X(S) = \mathbb{P}(\Omega) = 1$.
אם $1 \in S$ אבל $0 \notin S$ אז $A = X^{-1}(S)$ ואז $\mathbb{P}_X(S) = \mathbb{P}(A) = p$.
לבסוף אם $0 \in S$ ו־$1 \notin S$ אז $A^C = X^{-1}(S)$ אז $\emptyset = X^{-1}(S)$ ואז $\mathbb{P}_X(S) = \mathbb{P}(\emptyset) = 0$.
אם נגדיר $p_X : \mathbb{R} \to [0, \infty)$ על־ידי

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

אז מתקיים

$$\mathbb{P}_X(S) = \sum_{s \in S} p_X(s)$$

הגדרה 12.2 (התפלגות ברנולי) משתנה מקרי X מתפלג ברנולי עם פרמטר p אם יש לו פונקציית התפלגות נקודתית

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

במקרה זה נסמן $X \sim \mathrm{Ber}(p)$, סימון לא מאוד מועיל או מתכתב עם השימוש הסטנדרטי, אבל אלה הם החיים.

נשאל את עצמנו את השאלה האם כל משתנה מקרי מתפלג ברנולי הוא מציין של מאורע. אילו נגדיר $A = X^{-1}(1)$ אז מתקבל $X = 1_A$, אנו אומרים ש־X שווה למציין של 1_A כמעט תמיד, נראה זאת בהמשך הפרק.
נמשיך לעוד מקרים.

הגדרה 12.3 (משתנה מקרי קבוע) משתנה מקרי X הוא קבוע אם

$$p_X(s) = \begin{cases} 1 & s = c \\ 0 & \text{else} \end{cases}$$

עבור c קבוע כלשהו.

הגדרה 12.4 (משתנה מקרי אחיד) משתנה מקרי X נקרא אחיד על S תת־קבוצה סופית של $\mathbb{R}$ אם

$$p_X(s) = \begin{cases} \frac{1}{|S|} & s \in S \\ 0 & \text{else} \end{cases}$$

במקרה זה נסמן $X \sim \mathrm{U}(S)$.

הגדרה 12.5 (התפלגות גאומטרית) X מתפלג גאומטרית עם פרמטר p אם

$$p_X(s) = \begin{cases} (1 - p)^{s-1} p & s \in \{1, 2, \dots\} \\ 0 & \text{else} \end{cases}$$

ונסמן $X \sim \mathrm{Geo}(p)$.
לפעמים הגדרה זו תסומן אחרת על־ידי מדידת המקרים שבהם יצאה ההסתברות למאורע הראשון בלבד.
התפלגות זו מתארת את המקרה שניסינו לקבל תוצאה בהסתברות בין שני מקרים וקיבלנו אותה בפעם ה־s.

הגדרה 12.6 (התפלגות בינומית) X מתפלג בינומית עם פרמטרים n ו־p אם

$$p_X(s)=\begin{cases}\binom{n}{s}p^s(1-p)^{n-s} & s\in\{1,2,\dots\}\\ 0 & \text{else}\end{cases}$$

ונסמן $X\sim \mathrm{Bin}(n,p)$.

מאפשר לנו לחשב את מספר המטבעות המוטים שיצאו על צד מסוים. ולבסוף

הגדרה 12.7 (התפלגות פואסונית) X מתפלג פואסונית עם פרמטר λ אם

$$p_X(s)=\begin{cases}e^{-\lambda}\frac{\lambda^s}{s!} & s\in\{0,1,2,\dots\}\\ 0 & \text{else}\end{cases}$$

ונסמן $X\sim \mathrm{Po}(\lambda)$.

בפעם הראשונה ההתפלגות הזו הופיעה בהקשר של מספר החיילים שנהרגו מבעיטה מהסוס שלהם, התפלגות שהייתה מהותית עד מלחמת העולם הראשונה.

12.2 קשרים בין משתנים־מקריים

דוגמה 12.2 $\Omega=[6]^2$ מרחב אחיד להטלת שתי קוביות, ונגדיר שוב $Y=X_1+X_2$ סכום הטלות שתי הקוביות, דהינו

$$X_1(a,b)=a, X_2(a,b)=b, Y(a,b)=a+b$$

בתרגול מצאנו את הערכים של p_Y לכל ערך אפשרי.

נגדיר גם $Z=Y \mod 6$ (ומנוחות נגדיר $Z\in[6]$), ונשאל מה ההתפלגות של Z.

$$p_Z(1)=\mathbb{P}(Z=1)=\mathbb{P}(Y=7)=\frac{1}{6}$$

באופן דומה

$$p_Z(2)=\mathbb{P}(Z=2)=\mathbb{P}(Y=2)+\mathbb{P}(Y=8)=\frac{1}{36}+\frac{5}{36}=\frac{1}{6}$$

באופן כללי מתקיים מחישוב כזה ש־$p_Z(n)=\frac{1}{6}$ לכל $n\in[6]$, נסיק כי $Z\sim \mathrm{U}([6])$.

הגדרה 12.8 (הסתברות כמעט תמיד) אם A מאורע עם הסתברות 1 אז אומרים שהוא מתרחש כמעט תמיד.

הגדרה 12.9 (משתנים שווים שמעט תמיד) אם X ו־Y המקיימים ש־$X=Y$ כמעט תמיד אז נסמן $X\overset{a.s.}{=}Y$.

זה כמובן שקול להגדרה כי $\mathbb{P}(\{\omega\in\Omega\mid X(\omega)=Y(\omega)\})=1$ וזה נכון אם ורק אם $\mathbb{P}(\{\omega\in\Omega\mid X(\omega)\neq Y(\omega)\})=0$.

תרגיל 12.1 הוכיחו כי אם $X\overset{a.s.}{=}Y$ וגם $Y\overset{a.s.}{=}Z$ אז גם $X\overset{a.s.}{=}Z$, דהינו זהו יחס טרנזיטיבי.

הוכחה. נשים לב לעובדה הבאה, אם $X(\omega)=Y(\omega)$ ו־$Y(\omega)=Z(\omega)$ עבור $\omega\in\Omega$ כלשהו, אז $X(\omega)=Z(\omega)$, כלומר

$$\{\omega\in\Omega\mid X(\omega)=Y(\omega)\}\cap\{\omega\in\Omega\mid Y(\omega)=Z(\omega)\}\subseteq\{\omega\in\Omega\mid X(\omega)=Z(\omega)\}$$

ובהתאם גם

$$\{\omega\in\Omega\mid X(\omega)\neq Y(\omega)\}\cup\{\omega\in\Omega\mid Y(\omega)\neq Z(\omega)\}\supseteq\{\omega\in\Omega\mid X(\omega)\neq Z(\omega)\}$$

אז מחסם האיחוד נקבל

$$0\le\{\omega\in\Omega\mid X(\omega)\neq Z(\omega)\}\le 0+0$$

□

טענה 12.10 *$\overset{a.s.}{=}$ הוא יחס שקילות על מרחב כל המשתנים־המקריים על Ω.*

הוכחה. ראינו עתה טרנזיטיביות, וסימטריה ורפלקסיביות נובעות ישירות מההגדרה. □

תרגיל 12.2 האם בדוגמה קודם מתקיים $X_1\overset{a.s.}{=}X_2$?

פתרון מחישוב מתקיים $\mathbb{P}(X_1 = X_2) = \frac{1}{6}$ ולכן התשובה היא שלא.
נבחין כי גם $\mathbb{P}(X_1 \neq Z) \geq \mathbb{P}(X_1 = 2, Z = 3) = \mathbb{P}(\{(2,1)\}) = \frac{1}{36}$.

באופן יותר כללי גם אם יש מאורעות שיש להם אותה ההסתברות, אין הכרח שיהיה קשר לשוויון שלהם כמעט תמיד.

הגדרה 12.11 (משתנים מקריים שווי התפלגות) אם למשתנים מקריים X, Y יש אותה פונקציית התפלגות, דהינו $\mathbb{P}_Y = \mathbb{P}_X$,
דהינו מתקיים $\mathbb{P}(X^{-1}(S)) = \mathbb{P}(Y^{-1}(S)) \iff \mathbb{P}(X \in S) = \mathbb{P}(Y \in S) \iff \forall S \in \mathcal{F}_{\mathbb{R}}, \mathbb{P}_X(S) = \mathbb{P}_Y(S)$,
אז נאמר שהם שווי התפלגות ונסמן $X \overset{d}{=} Y$.

ראינו שיש משתנים מקריים X ו־Y כך ש־$Y \overset{a.s.}{\neq} X$ אבל $X \overset{d}{=} Y$, האם $Y \overset{a.s.}{=} X$ גורר $X \overset{d}{=} Y$? התשובה היא שכן!

טענה 12.12 *אם* $Y \overset{a.s.}{=} X$ *אז גם* $X \overset{d}{=} Y$.

הוכחה. נניח ש־$Y \overset{a.s.}{=} X$ ונרצה להוכיח ש־$\forall S \in \mathcal{F}_{\mathbb{R}}, \mathbb{P}(X \in S) = \mathbb{P}(Y \in S)$.
לכל $S \in \mathcal{F}_{\mathbb{R}}$ מתקיים

$$0 \neq \mathbb{P}(X \neq Y) \geq \mathbb{P}(X \in S, Y \notin S) = 0$$

ובהתאם

$$\mathbb{P}(X \in S) = \mathbb{P}(X \in S, Y \in S) + \overbrace{\mathbb{P}(X \in S, Y \notin S)}^{=0} = \mathbb{P}(X \in S, Y \in S)$$

כמו־כן גם $\mathbb{P}(Y \in S) = \mathbb{P}(X \in S, Y \in S)$. □

13 שיעור 9 — 25.11.2024

13.1 וקטורים מקריים

ניזכר בהגדרה 12.1:

הגדרה (משתנה מקרי בדיד) משתנה מקרי נקרא בדיד אם $\mathbb{P}_X$ פונקציית הסתברות בדידה, כלומר

$$\forall S \in \mathcal{F}_X, \mathbb{P}_X(S) = \sum_{s \in S} p_X(s)$$

כאשר $p_X(s) = \mathbb{P}(X = s) = \mathbb{P}(X^{-1}(s)) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = s\})$.

גם דיברנו על סוגים שונים של התפלגות, לדוגמה

$$\forall i \in [6], p_X(i) = \frac{1}{6} \iff X \sim \mathrm{U}([6])$$

או באופן דומה

$$\forall k \in \{0, \dots, n\}, p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

A מתרחש כמעט תמיד אם $\mathbb{P}(A) = 1$, לכן אם $\mathbb{P}(X = Y) = 1$ אז נאמר ש־$X = Y$ כמעט תמיד, או נסמן $X \overset{a.s.}{=} Y$.
באופן דומה אם $\mathbb{P}_X = \mathbb{P}_Y$ אז נסמן $X \overset{d}{=} Y$ או נאמר ש־X ו־Y שווי התפלגות, וראינו קשר בין שתי ההגדרות.

דוגמה 13.1 נגדיר הטלת מטבע, $\Omega = \{H, T\}$ ו־$X = 1_{\{H\}}, Y = 1_{\{T\}}$, אז $X \overset{d}{=} Y$ שכן

$$p_X(s) = p_Y(s) = \begin{cases} \frac{1}{2} & s = 0 \\ \frac{1}{2} & s = 1 \\ 0 & \text{else} \end{cases}$$

אבל גם $\mathbb{P}(X = Y) = 0$ ולכן $X \overset{a.s.}{\neq} Y$.

טענה 13.1 *אם* $X \overset{d}{=} Y$ ו־$f \in \mathcal{F}_{\mathbb{R} \to \mathbb{R}}$ אז $f(X) \overset{d}{=} f(Y)$.

הוכחה. נגדיר $Z = f(X), W = f(Y)$, צריך להוכיח ש־$\forall S \in \mathcal{F}_{\mathbb{R}}, \mathbb{P}_Z(S) = \mathbb{P}_W(S)$.

$$\begin{aligned}
\mathbb{P}_Z(S) &= \mathbb{P}(Z \in S) \\
&= \mathbb{P}(\{\omega \in \omega \mid Z(\omega) \in S\}) \\
&= \mathbb{P}(\{\omega \in \omega \mid f(X(\omega)) \in S\}) \\
&= \mathbb{P}(\{\omega \in \omega \mid X(\omega) \in f^{-1}(S)\}) \\
&= \mathbb{P}(X \in f^{-1}(S)) \\
&= \mathbb{P}_X(f^{-1}(S)) \\
&= \mathbb{P}_Y(f^{-1}(S)) \\
&= \mathbb{P}(\{\omega \in \omega \mid Y(\omega) \in f^{-1}(S)\}) \\
&= \mathbb{P}(\{\omega \in \omega \mid f(Y(\omega)) \in S\}) \\
&= \mathbb{P}(W \in S) \\
&= \mathbb{P}_W(S)
\end{aligned}$$

□

דוגמה 13.2 נניח ש־$X \sim \mathrm{Ber}(\frac{1}{2})$ וגם $Y \sim \mathrm{Ber}(\frac{1}{2})$, ונרצה לחשב את $\mathbb{P}(X = Y)$.
אין לנו את היכולת לעשות זאת כי אין מספיק מידע.

הגדרה 13.2 (וקטור מקרי) וקטור מקרי הוא משתנה מקרי לתוך $\mathbb{R}^n$ במקום ל־$\mathbb{R}$, $X : \Omega \to \mathbb{R}^n$.

כלל ההגדרות נשארות זהות פרט להגדרה זו, לדוגמה $\mathbb{P}_X(S) = \mathbb{P}(X \in S)$.
המוטיב שלנו הוא היכולת לבנות כמה משתנים מקריים ולעבוד איתם כיציר בודד, לדוגמה $X = (X_1, X_2)$ עבור $X_1, X_2 : \Omega \to \mathbb{R}$ משתנים מקריים.

הגדרה 13.3 (התפלגות משותפת והתפלגויות שוליות) אם $X_1, \dots, X_n$ משתנים מקריים המוגדרים על Ω יחיד אז $X = (X_1, \dots, X_n)$ הוא וקטור מקרי המוגדר על Ω וההתפלגות שלו נקראת ההתפלגות המשותפת של $X_1, \dots, X_n$.
ההתפלגויות של כל אחד מ־$X_1, \dots, X_n$ נקראות ההתפלגויות השוליות.

השם הזה נובע מהגישה שבה נוכל להבין את ההסתברות של משתנה מקרי בודד מתוך הווקטור על־ידי, אם X_1, X_2 מרחבים מקריים אז

$$\mathbb{P}_{X_1}(S) = \mathbb{P}_{(X_1, X_2)}(S \times \mathbb{R}) = \mathbb{P}(\{\omega \in\in \Omega \mid (X_1(\omega), X_2(\omega)) \in S\})$$

דוגמה 13.3 אם $\Omega = [6]^2$ ו־$X_1(a, b) = a, X_2(a, b) = b$ אז $X = (X_1, X_2)$ כאשר $X : \Omega \to \mathbb{R}^2$ פונקציית הזהות.

דוגמה 13.4 נבחן עבור $E = \{(s, y) \in \mathbb{R}^2 \mid s \le t\}$ את

$$\mathbb{P}_{(X,Y)}(E) = \mathbb{P}(X \le Y)$$

הגדרה 13.4 (התפלגות משותפת בדידה) אם לווקטור המקרי התפלגות בדידה, כלומר $\mathbb{P}_{(X_1, \dots, X_n)}$ פונקציית הסתברות בדידה, אז נאמר שההתפלגות המשותפת של $X_1, \dots, X_n$ בדידה.

טענה 13.5 *ההתפלגות המשותפת של $X_1, \dots, X_n$ בדידה אם ורק אם ההתפלגות של כל אחד מ־$X_1, \dots, X_n$ בדידה.*

הוכחה. נוכיח את הכיוון הראשון.
נניח $\mathbb{P}_{X_1, \dots, X_n}$ בדידה, אך זה נכון אם ורק אם היא נתמכת על־ידי קבוצה בת־מניה, נבחר קבוצה $S \in \mathcal{F}_\mathbb{R}$ כזו.
נסמן ב־S_1 את ההטלה של S על הקורדינטה הראשונה, לכן $\mathbb{P}_{X_1}(S_1) = \mathbb{P}_{(X_1, X_2)}(S_1 \times \mathbb{R})$ אבל $S \subseteq S_1 \times \mathbb{R}$.
לכן X_1 נתמך על־ידי קבוצה בת־מניה, S_1, ולכן הוא בדיד.

נעבור לכיוון השני.
נניח ש־X_1, X_2 בדידים, לכן קיימות $S_1, S_2 \in \mathcal{F}_\mathbb{R}$ בנות־מניה,
כך ש־$\mathbb{P}(X_1 \in S_1) = \mathbb{P}(X_2 \in S_2) = 1$.
לכן $\mathbb{P}((X_1, X_2) \in S_1 \times S_2) = \mathbb{P}(X_1 \in S_2, X_2 \in S_2) = 1$.
S_1, S_2 בנות־מניה ולכן נובע ש־$S_1 \times S_2$ בת־מניה.

כמובן לווקטורים בגודל $n > 2$ ההוכחה דומה. □

14 תרגול 5 — 28.11.2024

14.1 משתנים מקריים

בהרצאה זו נניח שכל המשתנים המקריים הם בדידים.

הגדרה 14.1 (התניה במשתנים מקריים בדידים) אם $X : \Omega \to \mathbb{R}^d$ ו־$A \subseteq \Omega$ כך ש־$\mathbb{P}(A) > 0$ אז

$$\forall S \subseteq \mathbb{R}^d, \mathbb{P}_{X|A}(S) = \mathbb{P}(X \in S \mid A) = \mathbb{P}_A(X \in S)$$

הגדרה 14.2 (אי־תלות במשתנים מקריים בדידים) אם $X, Y : \Omega \to \mathbb{R}^d$ בלתי־תלויים אם לכל $S, T \subseteq \mathbb{R}^d$ מתקיים

$$\mathbb{P}(X \in S, Y \in T) = \mathbb{P}(X \in S) \cdot \mathbb{P}(Y \in T)$$

תרגיל 14.1 יהיו $X_1, X_2 \operatorname{Geo}(p)$ בלתי־תלויים ונגדיר גם $Z = X_1 + X_2$.

1. חשבו את ההתפלגות המשותפת של X_1 ו־Z.
2. הראו ש־$X_1 \mid \{Z = 1\}$ מתפלג אחיד על $\{1, \dots, n-1\}$.

פתרון תחילה ניזכר שאם $W \sim \operatorname{Geo}(p)$ אז $\operatorname{Supp}(W) = \mathbb{N}$ ו־$\mathbb{P}(W = k) = (1-p)^{k-1}p$, שכן זוהי ההסתברות שלא הצלחנו $k-1$ פעמים ובניסיון ה־k הצלחנו, עבור איזושהי פעולה.

1. אנו מגדירים $X = (X_1, Z)$ וקטור מקרי ואנו רוצים לחשב את ההתפלגות שלו, נחשב את התומך

$$\operatorname{Supp}(X_1, Z) \subseteq \mathbb{N}^2$$

ישירות מההגדרה של הווקטור, אבל אנו יודעים כי תמיד $X_1 < Z$, וכן גם אם $m < n$

$$\begin{aligned} P_{(X_1,Z)}(m,n) &= \mathbb{P}(X_1 = m, Z = n) \\ &= \mathbb{P}(X_1 = m, X_2 = n - m) \\ &= \mathbb{P}(X_1 = m) \cdot \mathbb{P}(X_2 = m - n) \\ &= (1-p)^{m-1}p(1-p)^{n-m-1}p \\ &= p^2(1-p)^{n-2} \end{aligned}$$

ולכן נסיק

$$P_{(X_1,Z)}(n,n) = \begin{cases} 0 & m \geq n \\ p^2(1-p)^{n-2} & m < n \end{cases}$$

2. נבחן את $X_1 \mid \{Z = n\}$ ונבין מה התומך.

$$\operatorname{Supp}(X_1 \mid \{Z = 1\}) = \{1, \dots, n-1\}$$

שכן Z מייצג סכום ולכן מהווה חסם ל־X_1 יחד עם $\operatorname{Supp}(X_1) = \mathbb{N}$. נעבור לחישוב ההתפלגות

$$\mathbb{P}(X = m \mid Z = n) = \frac{\mathbb{P}(X = m, Z = n)}{\mathbb{P}(Z = n)} = \frac{p^2(1-p)^{n-2}}{\mathbb{P}(Z = n)}$$

אבל

$$\mathbb{P}(Z = n) = \mathbb{P}(X_1 + X_2 = n) = \sum_{i=1}^{n-1} \mathbb{P}(X_1 = i, X_2 = n - i) = \sum_{i=1}^{n-1} p^2(1-p)^{n-2} = (n-1)p^2(1-p)^{n-2}$$

זוהי קונבולוציה, לכן נוכל להסיק

$$\mathbb{P}(X_1 = m \mid Z = n) = \frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}} = \frac{1}{n-1}$$

תרגיל 14.2 מטילים מטבע הוגן, אם יצא 0 מטילים שוב מטבע הוגן ואם יצא 1 מטילים מטבע מוטה עם הסתברות p ל־1.
נתחו את התפלגות ההטלה השנייה.

פתרון נסמן X תוצאת ההטלה הראשונה ו־Y תוצאת ההטלה השנייה.

מהנתונים נסיק כי $X \sim \operatorname{Ber}(\frac{1}{2})$. אנו גם יודעים כי גם Y הוא בהתפלגות ברנולי כלשהי.

לבסוף אנו גם יודעים שמתקיים $Y \mid \{X=0\} \sim \mathrm{Ber}(\frac{1}{2})$ וגם ש־$Y \mid \{X=1\} \sim \mathrm{Ber}(p)$, לכן

$$\mathbb{P}(Y=1)=\mathbb{P}(Y=1 \mid X=0)\mathbb{P}(X=0)+\mathbb{P}(Y=1 \mid X=1)\mathbb{P}(X=1)=\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot p=\frac{1}{4}+\frac{p}{2}$$

תרגיל 14.3 יהיו $X \sim \mathrm{Ber}(p)$ ו־$Y \sim \mathrm{Ber}(q)$ בלתי־תלויים.
חשבו את ההתפלגות של $X \cdot Y$.

פתרון נתחיל ונראה כי

$$\mathrm{Supp}(XY)=\{0,1\},$$

וכן גם XY בהתפלגות ברנולי כלשהי, אך

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

ולכן $XY \sim \mathrm{Ber}(pq)$.

15 שיעור 10 — 28.11.2024

15.1 התפלגות תחת התניה

הגדרה 15.1 (התפלגות משתנה מקרי בהינתן מאורע) יהי X משתנה מקרי ו־A מאורע כך ש־$\mathbb{P}(A) > 0$ אז אפשר לדבר על התפלגות X בהינתן A. זוהי ההתפלגות של X תחת $\mathbb{P}_A$ במקום $\mathbb{P}$. במקרה זה

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

טענה 15.2 *אם $X \overset{d}{=} Y$ ו־$S \in \mathcal{F}_{\mathbb{R}}$ כך ש־$\mathbb{P}(X \in S) > 0$ וכן $\mathbb{P}(Y \in S) > 0$, אז $X \mid X \in S \overset{d}{=} Y \mid Y \in S$.*

דוגמה 15.1 נניח ש־$X, Y \sim \mathrm{U}([6])$ ו־$S = [3, \infty)$ אז

$$X \mid X \in S \sim \mathrm{U}(\{3,4,5,6\}), \qquad Y \mid Y \in S \sim \mathrm{U}(\{3,4,5,6\})$$

הגדרה 15.3 (אי־תלות משתנים מקריים) X ו־Y נקראים בלתי־תלויים אם לכל $S, T \in \mathcal{F}_{\mathbb{R}}$ המאורעות $X \in S, Y \in T$ בלתי־תלויים.
הגדרה זו שקולה להגדרה שמתקיים

$$\mathbb{P}(X \in S, Y \in T) = \mathbb{P}(X \in S) \cdot \mathbb{P}(Y \in T)$$

טענה 15.4 *אם X ו־Y משתנים מקריים בדידים אז X ו־Y בלתי־תלויים אם ורק אם לכל $s, t \in \mathbb{R}$ מתקיים ש־$X = s$ ו־$Y = t$ בלתי־תלויים.*
טענה זו שקולה לטענה שמתקיים

$$\mathbb{P}(X = s, Y = t) = \mathbb{P}(X = s) \cdot \mathbb{P}(Y = t)$$

הוכחה. הכיוון הראשון הוא טריוויאלי מבחירת יחידונים ושימשו בהגדרה, לכן נניח את הכיוון השני ונראה כי מתקיים לכל $S, T \in \mathcal{F}_{\mathbb{R}}$ מתקיים

$$\mathbb{P}(X \in S, Y \in T) = \mathbb{P}(X \in S)\mathbb{P}(Y \in T)$$

נבחין כי

$$\begin{aligned}
\mathbb{P}(X \in S, Y \in T) &= \mathbb{P}(X \in S \cap \mathrm{Supp}(X), Y \in T \cap \mathrm{Supp}(Y)) \\
&= \sum_{\substack{s \in S \cap \mathrm{Supp}(X) \\ t \in T \cap \mathrm{Supp}(Y)}} \mathbb{P}(X = s, Y = t) \\
&= \sum_{\substack{s \in S \cap \mathrm{Supp}(X) \\ t \in T \cap \mathrm{Supp}(Y)}} \mathbb{P}(X = s)\mathbb{P}(Y = t) \\
&= \sum_{s \in S \cap \mathrm{Supp}(X)} \left(\sum_{t \in T \cap \mathrm{Supp}(Y)} \mathbb{P}(X = s)\mathbb{P}(Y = t) \right) \\
&= \left(\sum_{s \in S \cap \mathrm{Supp}(X)} \mathbb{P}(X = s) \right) \left(\sum_{t \in T \cap \mathrm{Supp}(Y)} \mathbb{P}(Y = t) \right) \\
&= \mathbb{P}(X \in S)\mathbb{P}(Y \in T)
\end{aligned}$$

□

טענה 15.5 *התפלגות X ו־$X + Y$ ו־Y בלתי־תלויים קובע ביחידות את ההתפלגות המשותפת.*

הוכחה עבור בדידים. p_X ו־p_Y בלתי־תלויים קובע את $p_{(X,Y)}(s,t) = p_X(s)p_Y(t)$. □

טענה 15.6 *X, Y, Z משתנים מקריים בדידים ונניח שלכל $s \in \mathrm{Supp}(X)$ מתקיים $Y \mid X = s \overset{d}{=} Z$, אז $Y \overset{d}{=} Z$ ו־X, Y בלתי־תלויים.*

הוכחה. מנוסחת ההסתברות השלמה נובע

$$\begin{aligned}\mathbb{P}(Y=t) &= \sum_{s\in\text{Supp}(X)} \mathbb{P}(X=s)\mathbb{P}(Y=t \mid X=s) \\ &= \sum_{s\in\text{Supp}(X)} \mathbb{P}(X=s)\mathbb{P}(Z=t) \\ &= \mathbb{P}(Z=t)\end{aligned}$$

עבור החלק השני נבחין כי

$$\mathbb{P}(X=s,Y=t)=\mathbb{P}(X=s)\mathbb{P}(Y=t\mid X=s)=\mathbb{P}(X=s)\mathbb{P}(Z=t)=\mathbb{P}(X=s)\mathbb{P}(Y=t)$$

$\square$

טענה 15.7 *אם X,Y משתנים מקריים בלתי־תלויים ו־$f,g\in\mathcal{F}_{\mathbb{R}\to\mathbb{R}}$ אז $f(X),g(Y)$ בלתי־תלויים.*

הוכחה. צריך להראות שלכל $S,T\in\mathcal{F}_{\mathbb{R}}$ מתקיים

$$\mathbb{P}(f(X)\in S,g(Y)\in T)=\mathbb{P}(f(X)\in S)\mathbb{P}(g(Y)\in T)$$

אבל ראינו כבר כי

$$\mathbb{P}(f(X)\in S,g(Y)\in T)=\mathbb{P}(X\in f^{-1}(S),Y\in g^{-1}(T))$$

אבל גם

$$\mathbb{P}(X\in f^{-1}(S),Y\in g^{-1}(T))=\mathbb{P}(f(X)\in S)\mathbb{P}(g(Y)\in T)$$

$\square$

דוגמה 15.2 X ו־Y בלתי־תלויים אז $X^2,\frac{1}{Y}$ בלתי־תלויים, זאת שכן $\mathbb{P}(X=1,Y=1)=\frac{1}{2}\neq\frac{1}{2}\cdot\frac{1}{2}$.
בכיוון ההפוך אם $X=Y\sim\text{Ber}(\frac{1}{2})$ לא בלתי־תלויים אבל אם $f(x)=g(y)=6$ גוררים ש־$f(X)$ ו־$g(Y)$ הם כמעט תמיד 6 ולכן בלתי־תלויים.

הגדרה 15.8 (קבוצת משתנים מקריים בלתי־תלויה) יהיו $X_1,\dots,X_n$ משתנים מקריים, אז הם יקראו בלתי־תלויים אם לכל $S_1,\dots,S_n\in\mathcal{F}_{\mathbb{R}}$ המאורעות $\{X_i\in S_i\}_{i\in[n]}$ הם בלתי־תלויים.

דוגמה 15.3 אם X,Y,Z הם בלתי־תלויים, האם גם $X+Y$ ו־Z בלתי־תלויים? אנו צריכים להראות ש־$\mathbb{P}(X+Y=s,Z=t)=\mathbb{P}(X+Y=s)\mathbb{P}(Z=t)$ עבור כל $s\in\{0,1,2\},t\in\{0,1\}$.
נבחר לדוגמה את $\mathbb{P}(X+Y=1,Z=1)=\mathbb{P}(X=0,Y=1,Z=1)+\mathbb{P}(X=1,Y=0,Z=1)=\frac{1}{8}+\frac{1}{8}$ ונוכל להמשיך כך ולראות שהטענה אכן מתקיימת.

תרגיל 15.1 מאורעות $A_1,\dots,A_n$ הם בלתי־תלויים אם ורק אם $1_{A_1},\dots,1_{A_n}$ הם בלתי־תלויים.

טענה 15.9 *$X_1,\dots,X_n$ משתנים מקריים בלתי־תלויים ונניח שיש אינדקסים $1=i_0<i_1<\dots<i_k=n$.*
נגדיר $Y_0=(X_{i_0},\dots,X_{i-1}),\dots Y_k=(X_{i_{k-1}},\dots,X_{i_k})$.

16 שיעור 11 — 3.12.2024

16.1 אי־תלות משתנים מקריים

נמשיך עם מהלך ההרצאה הקודמת.

טענה 16.1 *יהיו $X_1, \dots, X_n$ משתנים מקריים (יכולים להיות גם וקטורים מקריים ללא השפעה על ההוכחה) בלתי־תלויים, ויהיו $1 = b_0 < b_1 < \cdots < b_k = n$ ונגדיר $Y_1 = (X_{b_0+1}, \dots, X_{b_1}), \dots, Y_k = (X_{b_{k-1}+1}, \dots, X_{b_k})$.*
אז $Y_1, \dots, Y_k$ בלתי־תלויים.

כדוגמה, $X_1, \dots, X_7 \to \overbrace{(X_1, X_2, X_3)}^{Y_1}, \overbrace{(X_4, X_5)}^{Y_2}, \overbrace{(X_6, X_7)}^{Y_3}$.

הוכחה במקרה הבדיד. Y_i הוא וקטור מקרי ממימד $b_i - b_{i-1}$, צריך להוכיח שלכל $s_1, \dots, s_k$ כך ש $s_1 \in \mathbb{R}^{b_i - b_{i-1}}$ מתקיים

$$\mathbb{P}(\forall i \in k, Y_i = s_i) = \prod_{i=1}^{k} \mathbb{P}(Y_i = s_i)$$

נניח ש־$s_i = (a_{i1}, \dots, a_{id_i})$ ולכן נסיק מחוסר התלות של X_i

$$\prod_{i=1}^{k} \mathbb{P}(Y_i = s_i) = \prod_{i=1}^{k} \mathbb{P}(\forall 1 \le j \le d_i, X_{b_{i-1}+j} = a_{ij}) = \prod_{i=1}^{k} \prod_{j=1}^{d_i} \mathbb{P}(X_{b_{i-1}+j} = a_{ij})$$

אבל

$$PP(\forall i \in k, Y_i = s_i) = \mathbb{P}(\forall j = X_j = c_j) = \prod_{j=1}^{h} \mathbb{P}(X_j = c_j) = \prod_{i=1}^{k} \prod_{j=1}^{d_i} \mathbb{P}(X_{b_{i-1}+j} = a_{ij})$$

עבור

$$c = (\overbrace{a_{11}, \dots, a_{1d_1}}^{s_1}, \dots, \overbrace{a_{k1}, \dots, a_{kd_1}}^{s_k})$$

ומצאנו כי השוויון אכן מתקיים ו־$Y_1, \dots, Y_k$ בלתי־תלויים. □

מסקנה 16.2 $X_1, \dots, X_n$ בלתי־תלויים ו־$0 = b_0 < b_1 < \cdots < b_k = n - 1$ ו־$d_i = b_i - b_{i-1}$ כך ש־$Y_i = (X_{b_{i-1}+1}, \dots, X_{b_i})$ ו־$f_1, \dots, f_k$ כאשר $f_i : \mathbb{R}^{d_i} \to \mathbb{R}$, אז $\{f_i(Y_i)\}_{i=1}^k$ בלתי־תלויים.

דוגמה 16.1 אם $X_1, \dots, X_n$ בלתי־תלויים אז גם $X_1 + X_2, \dots, X_3 + X_4, \dots, X_{n-1} + X_n$ כנביעה מהמסקנה,
באופן דומה גם $X_1 + X_2 + X_3, \dots$ בלתי־תלויים.
כרעיון אנו יכולים לחלק משתנים מקריים לווקטורים בלתי־תלויים, ואז להפעיל פונקציה, שלא משנה את חוסר התלות, על כל הקבוצה.

דוגמה 16.2 נניח ש־$A_1, \dots, A_5$ מאורעות בלתי־תלויים, אז המאורעות $(A_5 \cap A_4) \cup (A_1 \cap A_2) \cup A_3^C, A_4 \setminus A_5$ בלתי־תלויים, זאת אנו עושים על־ידי שימוש במשתנים המקריים האופייניים של A_i ושימוש במסקנה.

נבחין כי דרישת סופיות קבוצת המשתנים המקריים היא לא תנאי הכרחי

הגדרה 16.3 (קבוצה בת־מניה של משתנים מקריים בלתי־תלויים) $X_1, X_2, \dots$ משתנים מקריים הם בלתי־תלויים אם לכל $n \in \mathbb{N}$ מתקיים $X_1, \dots, X_n$ בלתי־תלויים.

טענה 16.4 *אם $\{X_n\}_{n \in \mathbb{N}}$ בלתי־תלויים ו־$S_n \in \mathcal{F}_{\mathbb{R}}$ לכל $n \in \mathbb{N}$ אז*

$$\mathbb{P}(\forall n \in \mathbb{N}, X_n \in S_n) = \prod_{n \in \mathbb{N}} \mathbb{P}(X_n \in S_n)$$

נשאל את עצמנו אם מצב זה בכלל אפשרי, נראה טענה ללא הוכחה שעונה על שאלה זו.

טענה 16.5 *קיימת סדרת משתנים מקריים כזאת שכולם* $\mathrm{Ber}(\frac{1}{2})$.

טענה 16.6 *סדרה כזו בהכרח לא מוגדרת על מרחב בדיד.*

הוכחה. נניח ש־$X_1, \dots$ סדרה כזו ונניח ש־$(\Omega, \mathcal{F}, \mathbb{P})$ בדיד.

נניח ש־$\omega_0 \in \Omega$ ו־$\mathbb{P}(\{\omega_0\}) > 0$, נסמן $s_i = X_i(\omega_0)$, אז

$$0 \xleftarrow[n\to 0]{} \left(\frac{1}{2}\right)^n = \mathbb{P}(\forall i \le n, X_i = s_i) \ge \mathbb{P}(\forall i \in \mathbb{N}, X_i = s_i) \ge \mathbb{P}(\{\omega_0\}) > 0$$

וקיבלנו סתירה לקיום ω_0 כזה. □

16.2 התפלגות גאומטרית

ניזכר בהגדרה 12.5, אשר מדברת על ניסוי שאנו עושים שוב ושוב עד שאנו מצליחים.

טענה 16.7 *אם* $X_1, X_2, \dots$ *משתנים מקריים בלתי־תלויים המתפלגים* $\mathrm{Ber}(p)$ *עבור* $0 < p < 1$, *ונסמן* $Y = \min\{k \mid X_k = 1\}$, *אז* $Y \sim \mathrm{Geo}(p)$.

נבחין כי Y מייצג בחירת המופע הראשון של 1 בהתפלגות ברנולי, נזכיר כי היא מייצגת הגרלה יחידה, לדוגמה הטלת מטבע בודד. נעבור להוכחה.

הוכחה. המאורע $Y = l$ הוא המאורע $X_1 = X_2 = \cdots = X_{l-1} = 0$ ו־$X_l = 1$, אבל שלו הם משתנים בלתי־תלויים, לכן

$$\mathbb{P}(X_1 = \cdots = X_{l-1} = 0, X_l = 1) = \mathbb{P}(X_1 = 0) \cdots \mathbb{P}(X_{l-1})\mathbb{P}(X_l = 1) = (1-p)^{l-1}p$$

זוהי התפלגות גאומטרית. □

הערה הסכום הוא

$$\sum_{l=1}^{\infty} (1-p)^{l-1} p = 1$$

ולכן המקרה שבו אין מינימום כפי שהגדרנו לא רלוונטי להגדרה, וניתן להתעלם ממנו.

מה יקרה אם נגדיר ככה $Y = Y_1$ ו־Y_2 המשתנה המקרי שעבורו קיבלנו 1 בפעם השנייה וכן הלאה, אז $Y_2 - Y_1$ וסדרת החיסורים היא בלתי תלויה־אף היא, תוצאה אינטואיטיבית אך לא מובנת מאליו.

17 תרגול 6 — 5.12.2024

17.1 שאלות בנושאי משתנים מקריים בלתי־תלויים

תרגיל 17.1 שתיים מטילות n מטבעות הוגנים כל אחת באופן בלתי־תלוי.
מה ההסתברות שהן קיבלו אותו מספר תוצאות עץ?

פתרון נגדיר $X_i \sim \mathrm{Ber}(\frac{1}{2})$ ההטלה ה־i של הראשונה, ו־$Y \sim \mathrm{Ber}(\frac{1}{2})$ ההטלה ה־i של השנייה, ונגדיר

$$X = \sum_{i=1}^n X_i, \qquad Y = \sum_{i=1}^n Y_i$$

משתנים המייצגים את מספר הטלות העץ של כל אחת מהשתיים.
אבל זאת דרך מורכבת לפתור את השאלה הזאת, נגדיר במקום זה $X, Y \sim \mathrm{Bin}(n, \frac{1}{n})$, את ההוכחה לשקילות נראה בהרצאה הקרובה.
מהגדרה 12.6 נוכל להסיק $\mathrm{Supp}\, X = \{0, \dots, n\}$, ואנו מבקשים לחשב את $\mathbb{P}(X = Y)$, נחשב על־ידי

$$\mathbb{P}(X = Y) = \sum_{k=0}^n \mathbb{P}(X = k = Y) = \sum_{k=0}^n \mathbb{P}(X = k)\mathbb{P}(Y = k) = \sum_{k=0}^n \binom{n}{k}\frac{1}{2^n}\binom{n}{k}\frac{1}{2^n} = \frac{1}{2^{2n}}\sum_{k=0}^n \binom{n}{k}^2$$

טענה 17.1 *יהיו $X \sim \mathrm{Bin}(n, p)$ ו־$Y \sim \mathrm{Bin}(m, p)$ בלתי־תלויים, אז $X + Y \sim \mathrm{Bin}(n + m, p)$.*

הוכחה. נבחין תחילה ש־$\mathrm{Supp}\, X + Y = \{0, \dots, n + m\}$ וכן

$$\begin{aligned}
\mathbb{P}(X + Y = k) &= \sum_{i=0}^{n+m} \mathbb{P}(X = i, Y = k - i) \\
&= \sum_{i=0}^{n+m} \mathbb{P}(X = i)\mathbb{P}(Y = k - i) \\
&= \sum_{i=0}^{n+m} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\
&= p^k (1-p)^{n+m-k} \sum_{i=0}^{n+m} \binom{n}{i}\binom{m}{k-i} \\
&= p^k (1-p)^{n+m-k} \binom{n+m}{k}
\end{aligned}$$

כאשר עלינו להוכיח את השוויון האחרון, זאת נעשה בתרגיל הבא. □

תרגיל 17.2 (זהות ונדרמונדה) מתקיים

$$\sum_{i=0}^{n+m} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$$

פתרון נבחין כי $\binom{n+k}{k}$ הוא קומבינטורית שקול לבחירה k ערכים מתוך קבוצה בגודל m וקבוצה בגודל n.
במקביל $\binom{n}{i}\binom{m}{k-i}$ הוא היכולת לבחור i מתוך m ועוד השארית מ־k מהקבוצה השנייה.
נבחין כי זוהי הוכחה קומבינטורית ואפשרי להוכיח גם אלגברית את השוויון הנתון.

טענה 17.2 *ניזכר בהגדרה 12.7 יהיו $X \sim Pois(\lambda)$ ו־$Y \sim Pois(\lambda + \eta)$ בלתי־תלויים, אז $X + Y \sim Pois(\lambda + \eta)$.*

ניזכר בהתפלגות פואסון עם דוגמה. אם מגיעים לבית־חולים בממוצע $\lambda = 10$ אנשים ביום, אז התפלגות פואסון היא השאלה כמה אנשים הגיעו ביום ספציפי לבית־החולים.
בהתאם השאלה שאנו שואלים מדברת על מקרה שבו יש שני בתי־חולים ואנו שואלים על כמה אנשים הגיעו ביום.
ברור לנו אם כן שטענה זו הגיונית, נעבור להוכחה.

הוכחה. תחילה $\text{Supp}(X+Y)=\mathbb{N}\cup\{0\}$ ונניח $k\in\text{Supp}(X+Y)$, ונחשב

$$\begin{aligned}\mathbb{P}(X+Y=k)&=\sum_{n=0}^{\infty}\mathbb{P}(X=n)\mathbb{P}(Y=k-n)\\&=\sum_{n=0}^{k}\frac{e^{-\lambda}\lambda^n}{n!}\frac{e^{-\eta}\eta^{k-n}}{(k-n)!}\\&=e^{-(\lambda+\eta)}\sum_{n=0}^{k}\frac{\lambda^n\eta^{k-n}}{n!(k-n)!}\\&=\frac{e^{-(\lambda+\eta)}}{k!}\sum_{n=0}^{k}\frac{k!}{n!(k-n)!}\lambda^n\eta^{k-n}\\&=\frac{e^{-(\lambda+\eta)}}{k!}\sum_{n=0}^{k}\binom{k}{n}\lambda^n\eta^{k-n}\\&=\frac{e^{-(\lambda+\eta)}}{k!}(\lambda+\eta)^k\end{aligned}$$

□

טענה 17.3 *נניח* $X\sim Pois(\lambda)$ *ו־*Y *משתנה מקרי המקיים*

$$\forall n\in\mathbb{N}\cup\{0\},\ Y\mid\{X=n\}\sim\text{Bin}(n,p)$$

אז $Y\sim Pois(\lambda p)$.

הוכחה. הפעם $\text{Supp}\,Y=\mathbb{N}\cup\{0\}$.
נניח ש־$k\in\mathbb{N}\cup\{0\}$ ונחשב את $\mathbb{P}(Y=k)$ על־ידי שימוש בנוסחת ההסתברות השלמה.

$$\begin{aligned}\mathbb{P}(Y=k)&=\sum_{n=0}^{\infty}\mathbb{P}(X=n)\mathbb{P}(Y=k\mid X=n)\\&=\sum_{n=0}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k}\\&=\sum_{n=k}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\\&=\frac{e^{-\lambda}p^k}{k!}\sum_{n=k}^{\infty}\frac{\lambda^n(1-p)^{n-k}}{(n-k)!}\\&=\frac{e^{-\lambda}p^k}{k!}\sum_{m=0}^{\infty}\frac{\lambda^{m+k}(1-p)^m}{m!}\\&=\frac{e^{-\lambda}p^k\lambda^k}{k!}\sum_{m=0}^{\infty}\frac{\lambda^m(1-p)^m}{m!}\\&=\frac{e^{-\lambda}p^k\lambda^k}{k!}e^{\lambda(1-p)}\\&=\frac{e^{-\lambda p}p^k(p\lambda)^k}{k!}\end{aligned}$$

□

18 שיעור 12 — 5.12.2024

18.1 התפלגות גאומטרית

טענה 18.1 *אם* $X_i \sim \mathrm{Ber}(p)$ *אז* $Y = \min k \mid X_k = 1$ *אז* $Y \sim \mathrm{Geo}(p)$.

משפט 18.2 (תכונת חוסר הזיכרון) X *משתנה מקרי הנתמך על* $\mathbb{N}$, *ב־*$\mathbb{P}(X > 1) > 0$. *אז התנאים הבאים שקולים:*

1. $X \sim \mathrm{Geo}(p)$ *עבור* $0 < p < 1$ *כלשהו.*
2. *לכל* $l \in \mathbb{N}$ *מתקיים* $X - l \mid X > l \overset{d}{=} X$. *כלומר* $\mathbb{P}(X > l \mid X - l > 0) = \mathbb{P}(X = k)$ *לכל* $l \in \mathbb{N}$.
3. $X - 1 \mid X > 1 \overset{d}{=} X$. *כלומר לכל* $S \in \mathcal{F}_{\mathbb{R}}$ *מתקיים* $\mathbb{P}(X - 1 \in S \mid X > 1) = \mathbb{P}(X \in S)$.

נראה טענה קודמת שתעזור לנו בהוכחת המשפט

טענה 18.3 *אם* X *משתנה מקרי שנתמך על* $\mathbb{N}$ *אז התנאים הבאים שקולים:*

1. $X \sim \mathrm{Geo}(p)$
2. $\mathbb{P}(X > n) = (1-p)^n$

הוכחה. $2 \implies 1$:

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

$1 \implies 2$:

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

□

הוכחת המשפט. $2 \implies 1$: נניח $l, k \in \mathbb{N}$

$$(1-p)^{k-1} p = \mathbb{P}(X = k) = \mathbb{P}(X - l = k \mid X - l > 0) = \frac{\mathbb{P}(X = l + k)}{\mathbb{P}(X > l)} = \frac{(1-p)^{l+k-1} p}{(1-p)^l} = (1-p)^{l-1} p$$

$3 \implies 2$ מיידי.

$1 \implies 3$: נסמן $p = \mathbb{P}(X = 1)$ ונראה ש־$X \sim \mathrm{Geo}(p)$ על־ידי כך שנראה $\forall n \in \mathbb{N},\ \mathbb{P}(X < n) = (1-p)^n$ והטענה.

נוכיח באינדוקציה. עבור $n = 1$ נובע $\mathbb{P}(X > 1) = 1 - \mathbb{P}(X = 1) = 1 - p$.

נניח שהטענה נכונה ל־n ונראה

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

והשלמנו את מהלך האינדוקציה. □

18.2 התפלגות בינומית

נעבור לדבר על התפלגויות בינומיות כפי שהגדרנו בהגדרה 12.6.

טענה 18.4 *אם* $X_1, \ldots, X_n$ *משתנים מקריים בלתי תלויים מתפלגים* $\mathrm{Ber}(p)$, *אז* $Y = \sum_{i=1}^{n} X_i$ *מתפלג* $\mathrm{Bin}(n, p)$.

הוכחה.

$$\begin{aligned}\mathbb{P}(Y=k) &= \sum_{\substack{v\in\{0,1\}^n \\ \sum v_i = k}} \mathbb{P}(X_1=v_1,\ldots,X_n=v_n) \\ &= \sum_{\substack{v\in\{0,1\}^n \\ \sum v_i = k}} \prod_{i=1}^{k} \mathbb{P}(X_i=v_i) \\ &= \sum_{\substack{v\in\{0,1\}^n \\ \sum v_i = k}} p^k(1-p)^{n-k} \\ &= \binom{n}{k} p^k (1-p)^{n-k}\end{aligned}$$

□

ניזכר בטענה 17.1 ונוכיח אותה הפעם בדרך פורמלית ולא על־ידי אינטואיציה.

הוכחה. נניח שיש $Z_1,\ldots,Z_{n+m}$ משתנים מקריים בלתי־תלויים $\mathrm{Ber}(p)$ כך שמתקיים

$$X' = \sum_{i=1}^{n} Z_i, \qquad Y' = \sum_{i=n+1}^{n+m} Z_i$$

אז $X'+Y' = \sum_{i=1}^{n+m} Z_i$, לפי הטענה

$$X' \sim \mathrm{Bin}(n,p), \qquad Y' \sim \mathrm{Bin}(m,p)$$

וכן

$$X'+Y' \sim \mathrm{Bin}(n+m,p)$$

לפי הטענה מההרצאה הקודמת X' ו־Y' בלתי־תלויים.

אבל $X' \overset{d}{=} X$ ו־$Y' \overset{d}{=} Y$ ו־X ו־Y בלתי־תלויים וגם X', Y' בלתי־תלויים, אז $(X,Y) \overset{d}{=} (X',Y')$. לבסוף נובע $X+Y \overset{d}{=} X'+Y'$. □

18.3 התפלגות פואסון

נעבור להתפלגות 12.7 ונבחן אותה

טענה 18.5 *נניח $\lambda > 0$ ונגדיר $X_n \sim \mathrm{Bin}(n, \frac{\lambda}{n})$ עבור $n > \lambda$,*
אז לכל $k \in \mathbb{N}$ מתקיים

$$\mathbb{P}(X_n = k) \xrightarrow[n\to\infty]{} \mathbb{P}(Y=k)$$

עבור $Y \sim Pois(\lambda)$.

הוכחה.

$$\mathbb{P}(X_n=0) = \binom{n}{0}\left(\frac{\lambda}{n}\right)^0 \left(1-\frac{\lambda}{n}\right)^n \xrightarrow[n\to\infty]{} e^{-\lambda}$$

באופן דומה

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

ונעבור למקרה הכללי

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

□

נחזור לטענה שראינו בתרגיל הבית:

טענה 18.6 *אם* $X \sim Pois(\lambda_1)$ *ו־*$Y \sim Pois(\lambda_2)$ *בלתי־תלויים אז* $X + Y \sim Pois(\lambda_1 + \lambda_2)$.

הפעם אפשר יהיה להוכיחה על־ידי הטענה החדשה שראינו.

דוגמה 18.1 נניח $\{X_n\}_{n\in\mathbb{N}}$ משתנים מקריים בלתי־תלויים $\mathrm{U}([26])$ המייצגים את קבלת כל אחת מהאותיות באנגלית, אז

$$\mathbb{P}(X_1 = 1, \dots, X_{1000} = 1) = \frac{1}{26^{1000}} = \mathbb{P}(X_{1001} = 1, \dots, X_{2000} = 1)$$

ההסתברות שיצא טקסט שמורכב מהאות a 1000 פעמים. הרעיון הוא שאין קשר בין המיקום שבו אנו שואלים עם הטקסט הופיע, אלא רק מהו אורך הטקסט, בהתאם

$$\mathbb{P}(\neg\exists k,\ X_{1000k+1} = 1, \dots, X_{1000k+1000} = 1) = (1 - \frac{1}{26^{1000}})^n \to 0$$

ולכן בסופו של דבר הטקסט הזה בהכרח יופיע.

19 שיעור 13 — 10.12.2024

19.1 תוחלת

הגדרה 19.1 (תוחלת במשתנים מקריים בדידים) X משתנה מקרי בדיד. **התוחלת** של X היא

$$\mathbb{E}(X)=\sum_{s\in\mathbb{R}} s\mathbb{P}(X=s)$$

הערה אם הטור $\sum_{s\in\mathbb{R}} s\mathbb{P}(X=s)$ לא מתכנס בהחלט, אז נאמר שאין תוחלת ל־X.

דוגמה 19.1 נניח ש־$\Omega=[100]$ מרחב הסתברות אחיד, מייצג קבוצת תלמידים.

נגדיר $X(\omega)$ הציון של תלמיד ω במבחן.

אז $\mathbb{E}(X)$ יהיה ממוצע הציונים בכיתה.

טענה 19.2 *אם X מוגדר על מרחב הסתברות בדידה, אז* $\mathbb{E}(X)=\sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})$.

הוכחה. מההגדרה של מרחב הסתברות בדידה נוכל להשתמש בתומך ואז נובע

$$\mathbb{E}(X)=\sum_{s\in\mathbb{R}} s\mathbb{P}(X=s)=\sum_{s\in\mathbb{R}} s\sum_{\substack{\omega\in\Omega\\ X(\omega)=s}}\mathbb{P}(\{\omega\})=\sum_{s\in\mathbb{R}}\sum_{\substack{\omega\in\Omega\\ X(\omega)=s}} X(\omega)\mathbb{P}(\{\omega\})=\sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})$$

□

טענה 19.3 *אם $X_1,\dots,X_n$ משתנים מקריים בדידים ו־$f\in\mathcal{F}_{\mathbb{R}^n\to\mathbb{R}}$, $Y=f(X_1,\dots,X_n)$, אז*

$$\mathbb{E}(Y)=\sum_{(s_1,\dots,s_n)\in\mathbb{R}^n} f(s_1,\dots,s_n)\mathbb{P}(X_1=s_1,\dots,X_n=s_n)$$

הוכחה. כמקודם נשתמש בתומך וכך נראה את השוויון.

$$\begin{aligned}\mathbb{E}(Y)&=\sum_{s\in\mathbb{R}} s\mathbb{P}(Y=s)\\&=\sum_{s\in\mathbb{R}} s\sum_{\substack{(s_1,\dots,s_n)\in\mathbb{R}^n\\ f(s_1,\dots,s_n)=s}}\mathbb{P}(X_1=s_1,\dots,X_n=s_n)\\&=\sum_{s\in\mathbb{R}}\sum_{\substack{(s_1,\dots,s_n)\in\mathbb{R}^n\\ f(s_1,\dots,s_n)=s}} f(s_1,\dots,s_n)\mathbb{P}(X_1=s_1,\dots,X_n=s_n)\\&=\sum_{(s_1,\dots,s_n)\in\mathbb{R}^n} f(s_1,\dots,s_n)\mathbb{P}(X_1=s_1,\dots,X_n=s_n)\end{aligned}$$

□

דוגמה 19.2 נניח $X\sim\mathrm{Ber}(p)$ אז

$$\mathbb{E}(X)=0\cdot\mathbb{P}(X=0)+1\cdot\mathbb{P}(X=1)=p$$

דוגמה 19.3 אם $X\sim\mathrm{U}([n])$ אז

$$\mathbb{E}(X)=\sum_{k=1}^n k\cdot\mathbb{P}(X=k)=\frac{1}{n}\sum_{k=1}^n k=\frac{1}{n}\cdot\frac{n(n+1)}{2}=\frac{n+1}{2}$$

דוגמה 19.4 אם $X\sim\mathrm{Bin}(n,p)$ אז

$$\mathbb{E}(X)=\sum_{k=0}^n k\binom{n}{k}p^k(1-p)^{n-k}=np$$

דוגמה 19.5 נניח $X\sim\mathrm{Poi}(\lambda)$, אז

$$\mathbb{E}(X)=\sum_{k=0}^\infty ke^{-\lambda}\frac{\lambda^k}{k!}=\sum_{k=1}^\infty e^{-\lambda}\frac{\lambda^k}{(k-1)!}=\lambda e^{-\lambda}\sum_{k=1}^\infty\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}\sum_{k=0}^\infty\frac{\lambda^k}{k!}=\lambda e^{-\lambda}e^{\lambda}=\lambda$$

ולבסוף גם

דוגמה 19.6 נניח $X \sim \mathrm{Geo}(p)$ ולכן

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} k(1-p)^{k-1} = \frac{1}{p}$$

את החישוב עצמו שמוכיח את הטענה הזאת נעשה בהמשך.

דוגמה 19.7 נניח $X \sim \mathrm{Geo}(\frac{1}{2})$ ונגדיר $Y = 2^X$.
בהתאם $\mathbb{P}(Y=2) = \frac{1}{2}$ וכן $\mathbb{P}(Y=4) = \frac{1}{4}$ וכן הלאה, כך ש־$\mathbb{P}(Y = 2^k) = \frac{1}{2^k}$. נחשב את התוחלת:

$$\mathbb{E}(Y) = \sum_{s \in \{2,4,8,\dots\}} 2^k \mathbb{P}(Y = 2^k) = \sum_{k=1}^{\infty} 2^k \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = \infty$$

ולכן אין תוחלת.

יכולנו להחליף את הגדרת Y להיות $Y = (-2)^X$ והיינו מקבלים טור שלא מתכנס בכלל.

הערה אפשר להרחיב את התוחלת ותכונותיה למקרים אינסופיים, אנו לא נעשה זאת.

19.2 תכונות של תוחלת

טענה 19.4 (תכונות של תוחלת) *אם X, Y משתנים מקריים בעלי תוחלת, אז:*

1. *אם $X \geq 0$ כמעט תמיד אז $\mathbb{E}(X) \geq 0$. אם בנוסף $\mathbb{P}(X > 0) > 0$ אז $\mathbb{E}(X) > 0$.*
2. *לינאריות: אם $a, b \in \mathbb{R}$ אז אם $Z = aX + bY$ יש תוחלת והיא $\mathbb{E}(Z) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.*

הוכחה. 1. אם $X \geq 0$ כמעט תמיד אז כל המחוברים אי־שליליים ולכן $\mathbb{E}(X) \geq 0$.
אם בנוסף $\mathbb{P}(X > 0) > 0$ אז קיים $s > 0$ כך ש־$\mathbb{P}(S = s) > 0$ ולכן $s\mathbb{P}(X = s) > 0$ ולכן הסכום חיובי.

2. נגדיר $f(x, y) = ax + by$ עבור $f \in \mathcal{F}_{\mathbb{R}^2 \to \mathbb{R}}$. אז

$$\begin{aligned}
\mathbb{E}(Z) &= \mathbb{E}(f(x,y)) \\
&= \sum_{(s,t) \in \mathbb{R}^2} f(s,t)\mathbb{P}(X = s, Y = t) \\
&= \sum_{(s,t) \in \mathbb{R}^2} (as + bt)\mathbb{P}(X = s, Y = t) \\
&= a\left(\sum_{(s,t) \in \mathbb{R}^2} s\mathbb{P}(X = s, Y = t)\right) + b\left(\sum_{(s,t) \in \mathbb{R}^2} t\mathbb{P}(X = s, Y = t)\right) \\
&= a\left(\sum_{s \in \mathbb{R}}\sum_{t \in \mathbb{R}} s\mathbb{P}(X = s, Y = t)\right) + b\left(\sum_{t \in \mathbb{R}}\sum_{s \in \mathbb{R}} t\mathbb{P}(X = s, Y = t)\right) \\
&= a\sum_{s \in \mathbb{R}} s\mathbb{P}(X = s) + b\sum_{t \in \mathbb{R}} t\mathbb{P}(Y = t) \\
&= a\mathbb{E}(X) + b\mathbb{E}(Y)
\end{aligned}$$

□

מסקנה 19.5 אם X ו־Y משתנים מקריים בעלי תוחלת ו־$X \leq Y$ כמעט תמיד, אז $\mathbb{E}(Y) \leq \mathbb{E}(X)$.

הוכחה. נגדיר $Z = X - Y$ ואז $Z \geq 0$ כמעט תמיד ולכן $\mathbb{E}(Z) \geq 0$ ואז $X = Y + Z$ ולכן $\mathbb{E}(X) = \mathbb{E}(Y) + \mathbb{E}(Z) \geq \mathbb{E}(Y)$. □

דוגמה 19.8 נעבור לראות את החישוב של תוחלת להתפלגות בינומית: נגדיר $X \sim \mathrm{Bin}(n, p)$ ונגדיר $X_1, \dots, X_n$ משתנים מקריים $\mathrm{Ber}(p)$, ויהי $X = \sum_{i=1}^n X_i$, לכן $X \sim \mathrm{Bin}(n, p)$ וכן

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) = np$$

הגדרה 19.6 (תוחלת מותנית) X משתנה מקרי ו־A מאורע, כך ש־$\mathbb{P}(A) > 0$, אז

$$\mathbb{E}(X \mid A) = \sum_{s \in \mathbb{R}} s\mathbb{P}(X = s \mid A)$$

טענה 19.7

$$\mathbb{E}(X \mid A) = \frac{\mathbb{E}(X \cdot 1_A)}{\mathbb{P}(A)}$$

הוכחה.

$$\mathbb{E}(X \mid A) = \sum_{s \in \mathbb{R}} s\mathbb{P}(X = s \mid A) = \sum_{s \in \mathbb{R}} s\frac{\mathbb{P}(X = s, A)}{\mathbb{P}(A)} = \sum_{s \in \mathbb{R}} s\frac{\mathbb{P}(X \cdot 1_A = s)}{\mathbb{P}(A)}$$

כאשר המעבר האחרון נובע מהגדרת המציין ובדיקה ידנית של המקרים בהם $s \in A, s \notin A$, כלומר

$$\mathbb{P}(X = s, A) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = s, \omega \in A\}) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) = s, 1_A(\omega) = 1\})$$

□

טענה 19.8 *אם $A_1, \ldots, A_n$ חלוקה של X משתנה מקרי בעל תוחלת, אז*

$$\mathbb{E}(X) = \sum_{k=1}^{n} \mathbb{E}(X \cdot q_{A_k})$$

הוכחה. מאותו מעבר כמו בהוכחה הקודמת נסיק

$$X = \sum_{k=1}^{n} X \cdot 1_{A_k}$$

ואז משתמש בתכונת הלינאריות של תוחלות ונקבל את המבוקש. □

טענה 19.9 **(נוסחת התוחלת השלמה)** *אם $A_1, \ldots, A_n$ חלוקה ו־$\mathbb{P}(A_k) > 0$ לכל k אז*

$$\mathbb{E}(X) = \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{E}(X \mid A_k)$$

הוכחה. על־ידי הטענות הקודמות נובע

$$\mathbb{E}(X) = \sum_{k=1}^{n} \mathbb{E}(X \cdot 1_{A_k}) = \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{E}(X \mid A_k)$$

□

דוגמה 19.9 נניח $X \sim \mathrm{Geo}(p)$ ו־$A_1 = \{X = 1\}$ וכן $A_2 = \{X > 1\}$, אז $\{A_1, A_2\}$ הם אכן חלוקה של Ω.
נחשב את התוחלת:

$$\mathbb{E}(X) = \mathbb{P}(A_1)\mathbb{E}(X \mid A_1) + \mathbb{P}(A_2)\mathbb{E}(X \mid A_2) = p \cdot 1 + (1-p) \cdot \mathbb{E}(X \mid X > 1)$$

אבל אז מתכונת חוסר הזיכרון

$$p \cdot 1 + (1-p) \cdot \mathbb{E}(X \mid X > 1) = p + (1-p) \cdot (\mathbb{E}(X - 1 \mid X > 1) + \mathbb{E}(1 \mid X > 1)) = p + (1-p) \cdot (\mathbb{E}(X) + 1)$$

ולכן קיבלנו את השוויון $\mathbb{E}(X) = p + (1-p)(\mathbb{E}(X) + 1)$, ממנו נובע $0 = p + (1-p) - p\mathbb{E}(X)$, לכן $\mathbb{E}(X) = \frac{1}{p}$.

20 תרגול 7 — 12.12.2024

20.1 שאלות ותכונות של תוחלות

באנגלית תוחלת היא Expectancy, מילה שמתארת בצורה יותר נאמנה את מושג התוחלת.

דוגמה 20.1 נניח ש־$X \sim \mathrm{Geo}(p)$, ונחשב את התוחלת של X.

יש להעריך את הטור $\sum_{n=1}^{\infty} n\mathbb{P}(X = n)$.

$$\sum_{k=1}^{\infty} k\mathbb{P}(X = k) = \sum_{n=1}^{\infty} kp(1-p)^{k-1} = p\sum_{n=1}^{\infty} k(1-p)^{k-1}$$

נגדיר $f \in (0,1)$ את $f(q) = \frac{1}{1-q} = \sum_{k=0}^{\infty} q^k$ ולכן f אנליטית ומתכנסת במידה שווה על תת־קבוצות של התחום שלה. אז

$$\frac{1}{(1-q)^2} = f'(q) = \sum_{k=0}^{\infty} kq^{k-1} = \sum_{k=1}^{\infty} kq^{k-1}$$

אם נציב $q = 1 - p$ אז נובע

$$p\sum_{n=1}^{\infty} k(1-p)^{k-1} = p \cdot \frac{1}{(1-(1-p))^2} = \frac{1}{p}$$

דוגמה 20.2 אם $X \sim \mathrm{Bin}(n, p)$ ואם נגדיר $Y \sim \mathrm{Bin}(n-1, p)$ אז

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{k \cdot n!}{(n-k)!k!}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(n-k)!(k-1)!}p^k(1-p)^{n-k} \\
&= \sum_{m=0}^{n-1} \frac{n!}{(n-m-1)!m!}p^{m+1}(1-p)^{n-m-1} \\
&= np\sum_{m=0}^{n-1} \frac{(n-1)!}{(n-m-1)!m!}p^{m}(1-p)^{n-m-1} \\
&= np\mathbb{P}(y \in \operatorname{Supp} Y) \\
&= np
\end{aligned}$$

דוגמה 20.3 (תוחלת של משתנה מקרי היפר־גאומטרי) ניזכר בשאלה: בכד יש a כדורים אדומים ו־b כדורים שחורים ושולפים k כדורים ללא החזרה לכד.

X משתנה מקרי שסופר את מהספר הכדורים האדומים.

נגדיר X_i המשתנה המקרי שבשליפה ה־i יצא כדור אדום, ו־$X = \sum_{i=1}^{k} X_i$. נוכל להגדיר גם $\Omega = \{(y_1, \dots, y_k) \mid i \neq j \implies y_i \neq y_j\}$, כאשר מסמנים את הכדורים האדומים ב־$\{1, \dots, a\}$ וכן את השחורים ב־$\{a+1, \dots, b\}$. אז

$$\mathbb{P}(X_i = 1) = \mathbb{P}(y_i \le a) = \sum_{j=1}^{a} \mathbb{P}(y_i = j) = \sum_{j=1}^{a} \frac{1}{a+b} = \frac{a}{a+b}$$

ולכן $\mathbb{E}(X) = \frac{k \cdot a}{a+b}$.

נעבור לבחינת דוגמה לשימוש בנוסחת התוחלת השלמה, אותה ראינו בהרצאה האחרונה.

תרגיל 20.1 מטילים קובייה הוגנת שוב ושוב עד שיוצא 1.

מה תוחלת סכום ערכי הקובייה?

פתרון נגדיר X משתנה מקרי שסוכם את מספר סיבובי המשחק, לכן $X \sim \mathrm{Geo}(\frac{1}{6})$.

נגדיר גם Y_i להיות תוצאת ההטלה ה־i, אם היא התקיימה.

בנוסף $Y = \sum_{i=1}^{\infty} Y_i$, הערך המעניין אותנו, סכום הקובייה בסוף המשחק.

$$Y_i \mid X = n \sim \begin{cases} \mathrm{U}(2,\dots,6) & i < n \\ 1 & i = n \\ 0 & i > n \end{cases}$$

ולכן נשתמש בנוסחת התוחלת השלמה

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{n=1}^{\infty} \mathbb{E}(Y \mid X = n) \cdot \mathbb{P}(X = n) \\
&= \sum_{n=1}^{\infty} (\sum_{i=1}^{n} \mathbb{E}(X_i \mid X = n)) \cdot \mathbb{P}(X = n) \\
&= \sum_{n=1}^{\infty} (\sum_{i=1}^{n} 4 + 1) \cdot (\frac{5}{6})^{n-1} \\
&= \frac{1}{6} \sum_{n=1}^{\infty} (4n - 3)(\frac{5}{6})^{n-1} \\
&= \frac{1}{6} (4 \sum_{n=1}^{\infty} n(\frac{5}{6})^{n-1} - 3 \sum_{n=1}^{\infty} n(\frac{5}{6})^{n-1}) \\
&= \frac{1}{6} (4 \cdot \frac{1}{(\frac{1}{6})^2} - 3 \cdot 6) \\
&= 21
\end{aligned}$$

21 שיעור 14 — 12.12.2024

21.1 תוחלת — המשך

נבחין כי מתקיימת הטענה הבאה, אך לא נוכיח אותה שכן אין בכך ערך לימודי:

טענה 21.1 (נוסחת התוחלת השלמה הבת־מניה) *אם X, Y משתנים מקריים בדידים, אז*

$$\mathbb{E}(X) = \sum_{t \in \mathbb{R}} \mathbb{P}(Y = t)\mathbb{E}(X \mid Y = t)$$

נבחן תכונה נוספת.

טענה 21.2 (תוחלת מכפלת משתנים מקריים בלתי־תלויים) *אם X, Y משתנים מקריים בלתי־תלויים ובעלי תוחלת, אז*

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

נבחין שבשונה מלינאריות תוחלת, במקרה הזה אנו צריכים את חוסר־התלות.

הוכחה. לפי נוסחת התוחלת השלמה

$$\mathbb{E}(XY) = \sum_{t \in \mathbb{R}} \mathbb{P}(Y = y)\mathbb{E}(XY \mid Y = t)$$

בהינתן $Y = t$ ההתפלגות של Y מתרכזת כולה ב־t, כלומר

$$\mathbb{P}(Y = s \mid Y = t) = \begin{cases} 1 & s = t \\ 0 & \text{else} \end{cases}$$

לכן בהינתן $Y = t$ מתקבל $XY \overset{a.s.}{=} Xt$ ובהתאם

$$XY \mid Y = t \overset{a.s.}{=} Xt \mid Y = t$$

לכן

$$\mathbb{E}(XY \mid Y = t) = \mathbb{E}(Xt \mid Y = t) = t\mathbb{E}(X \mid Y = t) = t\mathbb{E}(X)$$

ומשילוב השוויונות שמצאנו נובע

$$\mathbb{E}(XY) = \sum_{t \in \mathbb{R}} \mathbb{P}(Y = y)t\mathbb{E}(X) = \mathbb{E}(X)\mathbb{E}(Y)$$

□

נעבור לדון במה בכלל המשמעות של תוחלת. עד כה מצאנו תכונות שלה, ואף הגדרות שקולות, אך מה המשמעות של התוחלת בהקשר הסתברותי? באיזה מובן עלינו להתחשב בתוחלת במקרה שבו אנו יודעים את ערכה, כשהיא חיובית? נעבור להליך שנותן לנו מידע בהסתברות מתוך מידע על תוחלות.

משפט 21.3 (אי־שוויון מרקוב) *יהי X משתנה מקרי אי־שלילי (דהינו $X \overset{a.s.}{\geq} 0$) ובעל תוחלת.*
אז לכל $a > 0$ מתקיים

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

הוכחה. נבחן את החלוקה $\{X < 0\}, \{0 \leq X < a\}, \{X \geq a\}$, זוהי חלוקה של Ω, נסמן אותם גם ב־A_0, A_1, A_2 בהתאמה.

$$\mathbb{E}(X) = \overbrace{\mathbb{E}(X1_{A_0})}^{=0} + \mathbb{E}(X1_{A_1}) + \mathbb{E}(X1_{A_2})$$

נוכל לקבל תוצאה דומה עם נוסחת התוחלת השלמה.

המחובר השני הוא אי־שלילי מההגדרות שהנחנו, והמחובר השלישי מקיים

$$\mathbb{P}(X \geq a \mid X \geq a) = 1 \implies \mathbb{E}(X \mid X \geq a) \geq \mathbb{E}(a \mid X \geq a) = a$$

ולכן חסום על־ידי $a\mathbb{P}(a \in X)$. □

21.2 שימושים של אי־שוויון מרקוב

דוגמה 21.1 נניח ש־$X \sim \mathrm{Geo}(\frac{1}{2})$, אז $\mathbb{E}(X) = 2$, ובהתאם $\mathbb{P}(X \geq 4) \leq \frac{\mathbb{E}(X)}{4} = \frac{2}{4} = \frac{1}{2}$.
נוכל לחשב את ההסתברות עצמה על־ידי $\mathbb{P}(X \geq 4) = \sum_{k=4}^{\infty} \mathbb{P}(X = k) = \sum_{k=4}^{\infty} \frac{1}{2^k} = \frac{1}{8}$.
קיבלנו שהחסם שנובע מאי־שוויון ברקוב לא מאוד מועיל לנו.

אם $Y = X - 1$ ואנו מחפשים את $Y \geq 0$ אז $\mathbb{E}(Y) = 1$.
במקרה זה $\mathbb{P}(X \geq 4) = \mathbb{P}(Y \geq 3) \leq \frac{\mathbb{E}(Y)}{3} = \frac{1}{3}$.
אז קיבלנו חסם יותר טוב לערך, זאת אומרת שיש לנו דרך נוספת להשתמש באי־השוויון.

דוגמה 21.2 אם $X \sim Po(\lambda)$ ואנו מחפשים את $\mathbb{P}(X \geq 1)$.
אנו כבר יודעים ש־$\mathbb{E}(X) = \lambda$, ולכן $\mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}(X)}{1} = \lambda$.
אם $\lambda \geq 1$ אז מצאנו שהחסם הוא 1, והוא לא ממש מועיל לנו.
מצד שני $\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-\lambda} \frac{\lambda^k}{k!} = 1 - e^{-\lambda}$.
לכן $1 - e^{-\lambda} \leq \lambda$ ו־$1 - \lambda \leq e^{-\lambda}$. באופן כללי ראינו שהחסם אכן די מדויק עבור מקרים רבים של ערכי λ.

דוגמה 21.3 n מכתבים מגיעים ל־n תיבות דואר ובוחרים תמורה מקרית על $[n]$.
נבחן את X מספר נקודות השבת של התמורה.

$$X = \sum_{i=1}^{n} 1_{A_i}$$

כאשר A_i המאורע של נקודת שבת של התמורה ב־i, כלומר $\sigma(i) = i$.
נחשב גם $\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(1_{A_i}) = \sum_{i=1}^{n} \mathbb{P}(A_i)$ ולכן $\mathbb{P}(X \geq a) \leq \frac{1}{a}$.

דוגמה 21.4 נחזור לפרדוקס יום ההולדת, $X_i \sim \mathrm{U}([n])$ בלתי־תלויים עבור $i \in [k]$, כאשר המשתנה המקרי מייצג את הסיכוי שלאדם ה־i יש יום הולדת.
X הוא מספר ימי ההולדת המשותפים,

$$X = \sum_{1 \leq i < j \leq k} 1_{\{X_i = X_j\}}$$

ונוכל גם לכתוב

$$\mathbb{P}(X_i = X_j) = \sum_{l=1}^{m} \mathbb{P}(X_i = l, X_j = l) = \sum_{l=1}^{m} \mathbb{P}(X_i = l)\mathbb{P}(X_j = l) = \sum_{l=1}^{m} \frac{1}{m} \cdot \frac{1}{m} = \frac{1}{m}$$

ונובע

$$\mathbb{E}(X) = \sum_{1 \leq i < j \leq k} \mathbb{P}(X_i = X_j) = \binom{k}{2} \frac{1}{m}$$

ולבסוף

$$\mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}(X)}{1} = \binom{k}{2} \frac{1}{m}$$

זוהי הכללה של חסם האיחוד.

דוגמה 21.5 יהיו $A_1, \ldots, A_N$ קבוצות, $|A_i| \geq n$ ו־$N < 2^{n-1}$, אז קיימת קבוצה B כך ש־$B \cap A_i \neq \emptyset$ ו־$A_i \not\subseteq B$ $\forall 1 \leq i \leq N$.
נגדיר $A = \bigcup_{i=1}^{N} A_i$. יהיו $X_a \sim \mathrm{Ber}(\frac{1}{2})$ בלתי תלויים לכל $a \in A$ ונגדיר $B = \{a \mid X_a = 1\}$.
נחשב את ההסתברות $\mathbb{P}(A_i \subseteq B) = \mathbb{P}(\forall a \in A_i, X_a = 1) = (\frac{1}{2})^{|A_i|} \leq \frac{1}{2^n}$.
מצד שני $\mathbb{P}(A_i \cap B \neq \emptyset) = \mathbb{P}(\forall a \in A_i, X_a = 0) = (\frac{1}{2})^{|A_i|} \leq \frac{1}{2^n}$. אז

$$\mathbb{P}(\exists 1 \leq i \leq N, A_i \subseteq B \vee A_i \cap B = \emptyset) \leq \sum_{i=1}^{N} \mathbb{P}(A_i \subseteq B) + \mathbb{P}(A_i \cap B = \emptyset) \leq N \cdot (\frac{1}{2^n} + \frac{1}{2^n}) = \frac{N}{2^{n-1}} < 1$$

ולכן קיימת B כזאת.

22 שיעור 15 — 17.12.2024

22.1 נוסחה לתוחלות

נתחיל בנוסחה קטנה שתעזור לנו לפתח אינטואיציה, אך לא נשתמש בה רבות.

טענה 22.1 (נוסחת הזנב לתוחלת) *אם X משתנה מקרי שנתמך על־ידי $\mathbb{N}\cup\{0\}$ אז*

$$\mathbb{E}(X)=\sum_{n\in\mathbb{N}}\mathbb{P}(X\geq n)$$

הוכחה. ממשפט פוביני לסכומים מרובים

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

□

22.2 שונות

באנגלית Variance. אנו רוצים לשאול את השאלה כמה הסתברות רחוקה בעצם מהתוחלת, כך שנוכל לאפיין את שתי התכונות באופן מוצלח יותר אחת על־ידי השנייה. לדוגמה אם $\mathbb{E}(X)=\mathbb{E}(Y)=5$ אבל $X=5$ כמעט תמיד ומצד שני $\mathbb{P}(Y=5000000)=\frac{1}{10^6}$ ו־$\mathbb{P}(Y=0)=1-\frac{1}{10^6}$ בעלות תוחלת זהה אבל בבירור שונות בתכלית.

הגדרה 22.2 (שונות) X משתנה מקרי בעל תוחלת $\mu=\mathbb{E}(X)$, אז **השונות** של X היא

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

דוגמה 22.1 במקרה שראינו זה עתה

$$\operatorname{Var}(X)=\mathbb{E}((X-5)^2)=0$$

בעוד שמתקיים

$$\operatorname{Var}(Y)=\mathbb{E}((Y-5)^2)=(5000000-5)^2\cdot\frac{1}{10^6}+(0-5)^2\cdot(1-\frac{1}{10^6})\approx 25000000$$

כפי שאנו רואים, הפעם השונות מייצגת את ההבדל המשמעותי שבין שני המשתנים המקריים.

נוסיף הגדרה שלא נעסוק בה אך שרבים מאיתנו שמעו בעבר, והוא מושג סטיית התקן, מושג שמשמש רבות בסטטיסטיקה.

הגדרה 22.3 (סטיית תקן) סטיית התקן של X היא $\sigma(X)=\sqrt{\operatorname{Var}(X)}$.

נראה הגדרה נוספת לשונות שמשומשת אף היא, הגדרה זו שקולה להגדרה שראינו

הגדרה 22.4 (הגדרה שקולה לשונות) נגדיר את השונות להיות

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

הוכחת השקילות. נסמן $\mu=\mathbb{E}(X)$ ולכן מתכונות התוחלת

$$\mathbb{E}((X-\mu)^2)=\mathbb{E}(X^2-2X\mu+\mu^2)=\mathbb{E}(X^2)-\mathbb{E}(2X\mu)+\mathbb{E}(\mu^2)=\mathbb{E}(X^2)-2\mu\mathbb{E}(X)+\mu^2=\mathbb{E}(X^2)-2\mu^2+\mu^2$$

ומצאנו כי מתקיים השוויון שחיפשנו. □

טענה 22.5 (תכונות של שונות) *כלל התכונות הבאות מתקיימות עבור X משתנה מקרי בעל תוחלת:*

1. *$\operatorname{Var}(X)\geq 0$ ו־$\operatorname{Var}(X)=0$ אם X קבוע כמעט תמיד. השונות היא חיובית ואם המשתנה המקרי קבוע, אז השינוי שהיא מייצגת הוא אפס סביב התוחלת, היא גם ההסתברות.*
2. *$\operatorname{Var}(X)=\operatorname{Var}(X+a)$ לכל $a\in\mathbb{R}$. השונות לא מושפעת מהזזה, היא תכונה כללית.*
3. *$\operatorname{Var}(aX)=a^2\operatorname{Var}(X)$. מתיחה של המשתנה המקרי מגדילה אפילו יותר את השונות, נבחין כי השונות מייצגת את הטווח סביב התוחלת, ונוכל להסתכל עליה כשטח של איזשהו רדיוס סביב התוחלת, ככה נקבל את הריבוע.*

הוכחה. 1. $(X-\mu)^2 \geq 0$ ולכן גם $\mathbb{E}((X-\mu)^2) \geq 0$. גם

$$X-\mu \overset{a.s.}{=} 0 \iff (X-\mu)^2 \overset{a.s.}{=} 0 \iff \mathbb{E}((X-\mu)^2) = 0$$

2. נגדיר $Y = X + a$ ולכן $\mathbb{E}(Y) = \mu + a$, ואז

$$\mathrm{Var}(Y) = \mathbb{E}((Y-(\mu+a))^2) = \mathbb{E}(((X+a)-(\mu+a))^2) = \mathbb{E}((X-\mu)^2) = \mathrm{Var}(X)$$

3. נגדיר $Y = aX$ ובהתאם $\mathbb{E}(Y) = \mathbb{E}(aX) = a\mathbb{E}(X) = a\mu$, ולכן

$$\mathrm{Var}(Y) = \mathbb{E}((Y-a\mu)^2) = \mathbb{E}((aX-a\mu)^2) = a^2\mathbb{E}((X-\mu)^2) = a^2\,\mathrm{Var}(X)$$

□

נבחין כי בעוד שאנו יודעים כי תוחלות הן לינאריות, זהו לא המקרה עבור שונות, ננסה לחשב שונות של סכום משתנים מקריים. נחשב את $\mathrm{Var}(X+Y)$ עבור $\mu = \mathbb{E}(X), \nu = \mathbb{E}(Y)$, אז $\mathbb{E}(X+Y) = \mu + \nu$. נעבור לחישוב השונות

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathbb{E}(((X+Y)-(\mu+\nu))^2) \\ &= \mathbb{E}(((X-\mu)+(Y-\nu))^2) \\ &= \mathbb{E}((X-\mu)^2 + 2(X-\mu)(Y-\nu) + (Y-\nu)^2) \\ &= \mathbb{E}((X-\mu)^2) + 2\mathbb{E}((X-\mu)(Y-\nu)) + \mathbb{E}((Y-\nu)^2)\end{aligned}$$

ניתן שם לביטוי לחלק הביטוי שיצא לנו, ונגדיר

הגדרה 22.6 (שונות משותפת) נגדיר

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mu)(Y-\nu))$$

עבור $\mathbb{E}(X) = \mu, \mathbb{E}(Y) = \nu$.

כאשר Cov הוא קיצור ל־Covariance, הוא בתורו קיצור למילה Cooperative Variance. ולכן נוכל לקבל

מסקנה 22.7 עבור משתנים מקריים X, Y בעלי תוחלת, מתקיים

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + 2\,\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)$$

טענה 22.8 *אם X ו־Y בלתי־תלויים ובעלי שונות, אז* $\mathrm{Cov}(X,Y) = 0$.

הוכחה. נראה בהמשך שאם ל־X ול־Y יש שונות אז $\mathrm{Cov}(X,Y)$ מוגדר (כלומר הטור מתכנס בהחלט).
נניח כרגע שזה נכון ולכן

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mu)(Y-\nu)) \overset{(1)}{=} \mathbb{E}(X-\mu)\mathbb{E}(Y-\nu) = 0 \cdot 0 = 0$$

כאשר

1. $X-\mu$ ו־$Y-\nu$ בלתי־תלויים, מאי־התלות של X, Y עצמם.

□

דוגמה 22.2 אם $X = c$ כמעט תמיד אז $\mathbb{E}(X) = c$ ו־$\mathrm{Var}(X) = 0$.

דוגמה 22.3 נניח $X \sim \mathrm{Ber}(p)$ ולכן $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1-p)$.
אם $X \sim \mathrm{Ber}(p)$ אז $1-X \sim \mathrm{Ber}(1-p)$ ואז $\mathrm{Var}(1-X) = \mathrm{Var}(-X) = (-1)^2\,\mathrm{Var}(X) = \mathrm{Var}(X)$.
זאת אומרת, לא מפתיע שהשונות היא סימטרית במקרה זה עבור שני המשתנים.

דוגמה 22.4 אם $X \sim \mathrm{Bin}(n,p)$, אז נוכל להשתמש ישירות בהגדרת השונות, אבל נשתמש בעובדה שמשתנה בינומי הוא סכום של משתנים ברנולי, כלומר $X_i \sim \mathrm{Ber}(p)$ בלתי־תלויים עבור $1 \leq i \leq n$ אז אם $X = \sum_{i=1}^n X_i$ גם $X \sim \mathrm{Bin}(n,p)$, ולכן

$$\mathrm{Var}(X) = \mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i) = np(1-p)$$

דוגמה 22.5 נניח עתה $X \sim \mathrm{Poi}(\lambda)$, ונחשב על־ידי ההגדרה השקולה והעובדה שאנו כבר יודעים ש־$\mathbb{E}(X) = \lambda$:

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{k=0}^{\infty} k^2 \mathbb{P}(X = k) \\
&= \sum_{k=1}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=1}^{\infty} \frac{k\lambda^k}{(k-1)!} \\
&= e^{-\lambda} \sum_{k=1}^{\infty} \frac{((k-1)+1)\lambda^k}{(k-1)!} \\
&= e^{-\lambda} \sum_{k=1}^{\infty} \frac{(k-1)\lambda^k}{(k-1)!} + e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda}\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda^2 + \lambda
\end{aligned}$$

ולכן נסיק $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda$.

23 תרגול 8 — 19.12.2024

23.1 שימושים למשפט מרקוב

תרגיל 23.1 מטילים מטבע מוטה עם הסתברות p לעץ, 20 פעמים, חסמו את ההסתברות שהרצף עץ עץ יצא פחות מפעמיים.

פתרון נגדיר $\Omega = \{0,1\}^{20}$ וכן $X = 1_{\{\omega_i = \omega_{i+1} = 1\}}$, וגם $X = \sum X_i$, אנו רוצים לחשב את $\mathbb{P}(X \le 1)$.
נבחין כי $X_i \sim \mathrm{Ber}(p^2)$, וכן $\mathbb{E}(X) = \sum \mathbb{E}(X_i) = 19p^2$, שכן $0 \le X \le 19$.

$$\mathbb{P}(X \le 1) = \mathbb{P}(19 - X \ge 18) \le \frac{\mathbb{E}(19 - X)}{18} = \frac{19 - 19p^2}{18}$$

23.2 שאלות נבחרות בנושא שונות

נתחיל ונבחין ששונות היא תבנית בי־לינארית (תבנית ריבועית), ובשל כך היא מקיימת את הטענה שהיא אי־שלילית, היא אדישה להזזות קבועות ויש לה כיול ריבועי.

דוגמה 23.1 אם $X \sim \mathrm{U}([6])$ משתנה מקרי להטלת קובייה הוגנת, אז $\mathbb{E}(X) = 3\frac{1}{2}$ וכן

$$\mathrm{Var}(X) = \frac{6^2 - 1}{12} = \frac{35}{12} \approx 3$$

זאת־אומרת שהשונות באמת מתכתבת עם המרחק של הערכים מהתוחלת.

תרגיל 23.2 יהי $X \sim \mathrm{Geo}(q)$, חשבו את $\mathrm{Var}(X)$.

פתרון אנו יודעים ש־$\mathbb{E}(X) = \frac{1}{q}$.
עוד אנו יודעים מאנליטיות $\frac{1}{1-x}$ ופיתוח טיילור, מתקיים

$$\sum_{n=0}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$$

נשתמש בנוסחה זו ונובע

$$\mathbb{E}(X(X-1)) = \sum_{n=0}^{\infty} n(n-1)q(1-q)^{n-1} = q(1-q)\sum_{n=0}^{\infty} n(n-1)(1-q)^{n-2} = q(1-q)\frac{2}{q^3}$$

ולכן $\mathbb{E}(X^2) = \frac{2(1-q)}{q^2} + \frac{1}{q}$. נעבור לחישוב השונות

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2(1-q)}{q^2} + \frac{1}{q} - \frac{1}{q^2} = \frac{2(1-q) + q - 1}{q^2} = \frac{1-q}{q^2}$$

ניזכר בשונות משותפת, נבחין כי מבי־לינאריות השונות מתקיים

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{Cov}(X+Y, X+Y) \\
&= \mathrm{Cov}(X,X) + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) + \mathrm{Cov}(Y,Y) \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)
\end{aligned}$$

בהתאם גם

$$\mathrm{Var}(X) = \mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i,j=1}^{n} \mathrm{Cov}(X_i, X_j) = \sum_{i=1}^{n} \mathrm{Cov}(X_i) + 2\sum_{1=i<j=n} \mathrm{Cov}(X_i, X_j)$$

תרגיל 23.3 חשבו את השונות ואת התוחלת של מספר נקודות השבת בתמורה מקרית על $\{1, \dots, n\}$.

פתרון נגדיר X_i נקודת השבת במקום ה־i. $X_i \sim \mathrm{Ber}(\frac{1}{n})$ ו־$X = \sum X_i$. בהתאם

$$\mathbb{E}(X) = n\frac{1}{n} = 1$$

אם $i \ne j$ אז $\mathbb{P}(X_i, X_j = 1) = \frac{1}{n} \cdot \frac{1}{n-1} \ne \frac{1}{n^2} = \mathbb{P}(X_i = 1)\mathbb{P}(X_j = 1)$, כלומר אנו רואים כי משתנים אלה תלויים.

$$\mathrm{Var}(X_i) = \frac{1}{n}(1 - \frac{1}{n})$$

וכן

$$i \neq j \implies \mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \frac{1}{n} \cdot \frac{1}{n-1} - \frac{1}{n^2}$$

לכן

$$\begin{aligned}
\mathrm{Var}(X) &= \sum_{i=1}^{n} \frac{1}{n}(1 - \frac{1}{n}) + 2 \sum_{1=i<j=n} (\frac{1}{n} \cdot \frac{1}{n-1} - \frac{1}{n^2}) \\
&= (1 - \frac{1}{n}) + 2\binom{n}{2}(\frac{1}{n(n-1)} - \frac{1}{n^2}) \\
&= (1 - \frac{1}{n}) + 2\frac{n!}{2!(n-2)!}(\frac{1}{n(n-1)} - \frac{1}{n^2}) \\
&= (1 - \frac{1}{n}) + \frac{n(n-1)}{1}(\frac{1}{n(n-1)} - \frac{1}{n^2}) \\
&= 1 - \frac{1}{n} + 1 - \frac{n-1}{n} \\
&= 1
\end{aligned}$$

24 שיעור 16 — 19.12.2024

24.1 שונות — המשך

נרחיב את הטענה מההרצאה הקודמת.

טענה 24.1 *עבור משתנים מקריים X,Y בעלי תוחלת מתקיים*

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

הוכחה.

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

□

הערה אם X ו־Y בלתי־תלויים אז $\mathrm{Cov}(X,Y)=0$.

הגדרה 24.2 (משתנים מקריים בלתי־מתואמים) אם מתקיים $\mathrm{Cov}(X,Y)=0$ אז נאמר ש־X ו־Y **בלתי־מתואמים**.

נבחין כי זהו תנאי הרבה יותר חלש מאי־תלות.

טענה 24.3 (תכונות של שונות משותפת) *לכל שני משתנים מקריים X,Y בעלי־תוחלת מתקיימות התכונות הבאות*

1. $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
2. $\mathrm{Cov}(a+X,Y)=\mathrm{Cov}(X,Y)$
3. $\mathrm{Cov}(aX,Y)=a\cdot\mathrm{Cov}(X,Y)$
4. $\mathrm{Var}(X)=\mathrm{Cov}(X,X)$
5. $\mathrm{Cov}(X+Y,Z)=\mathrm{Cov}(X,Z)+\mathrm{Cov}(Y,Z)$

הוכחה. התכונה הראשונה טריוויאלית ונובעת מההגדרה, ולכן נוכיח את התכונה השנייה.

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

נוכיח את התכונה החמישית.

$$\begin{aligned}\mathbb{E}(((X+Y)-\mathbb{E}(X+Y))(Z-\mathbb{E}(Z)))&=\mathbb{E}(((X-\mathbb{E}(X))+(Y-\mathbb{E}(Y)))(Z-\mathbb{E}(Z)))\\&=\mathbb{E}((X-\mathbb{E}(X))(Z-\mathbb{E}(Z)))+\mathbb{E}((Y-\mathbb{E}(Y))(Z-\mathbb{E}(Z)))\end{aligned}$$

□

טענה 24.4 *נניח ש־$X_1,\dots,X_n$ משתנים מקריים בעלי תוחלת, אז*

$$\mathrm{Var}(\sum_{i=1}^n X_i)=\sum_{i=1}^n\mathrm{Var}(X_i)+2\sum_{1\le i<j\le n}\mathrm{Cov}(X_i,Y_i)$$

הוכחה.

$$\begin{aligned}\mathrm{Var}(\sum_{i=1}^n X_i)&=\mathrm{Cov}(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i)\\&=\sum_{i=1}^n\mathrm{Cov}(X_i,\sum_{i=1}^n X_i)\\&=\sum_{i=1}^n\sum_{j=1}^n\mathrm{Cov}(X_i,X_j)\\&=\sum_{i=1}^n\mathrm{Var}(X_i)+2\sum_{1\le i<j\le n}\mathrm{Cov}(X_i,Y_i)\end{aligned}$$

□

דוגמה 24.1 מטילים מטבע הוגן n פעמים וסופרים את מספר ה־HH, נחשב את התוחלת ואת השונות.

נגדיר $\{X_i\}_{i=1}^n$ משתנים מקריים $\mathrm{Ber}(\frac{1}{2})$ בלתי־תלויים, ונגדיר $Y_i = X_i + X_{i+1}$ עבור $i = 1, \dots, n-1$. נגדיר גם $Y = \sum_{i=1}^{n-1} Y_i$.

$$\mathbb{E}(Y_i) = \mathbb{E}(X_i)\mathbb{E}(X_{i+1}) = \frac{1}{4}, \qquad \mathbb{E}(Y) = \sum_{i=1}^{n-1} \mathbb{E}(Y_i) = \frac{n-1}{4}$$

וכן

$$\mathrm{Var}(Y) = \mathrm{Var}(\sum_{i=1}^{n-1} Y_i) = \sum_{n=1}^{n-1} \mathrm{Var}(Y_i) + 2 \sum_{1 \le i < j \le n-1} \mathrm{Cov}(Y_i, Y_j)$$

ונעבור לחישוב אלה האחרונים.

$$\mathrm{Var}(Y_i) = \frac{3}{16}$$

נבחן דוגמה ספציפית למקרה שהמשתנים שונים,

$$\mathrm{Cov}(Y_1, Y_7) = \mathrm{Cov}(X_1 X_2, X_7 X_8)$$

לכל $i+1 < j$ המשתנים $Y_i = X_i X_{i+1}$ ו־$Y_j = X_j X_{j+1}$ בלתי־תלויים (בתור פונקציות מעל משתנים מקריים בקבוצות זרות), לכן $\mathrm{Cov}(Y_i, Y_j) = 0$.

נשאר המקרה $j = i+1$.

$$\forall 1 \le i \le n-2, \mathrm{Cov}(Y_i, Y_{i+1}) = \mathrm{Cov}(X_i X_{i+1}, X_{i+1} X_{i+2}) = \mathbb{E}(X_i X_{i+1}^2 X_{i+2}) - \mathbb{E}(Y_i)\mathbb{E}(Y_{i+1}) = \frac{1}{8} - \frac{1}{16} = \frac{1}{16}$$

לבסוף $\mathrm{Var}(Y) = (n-1)\frac{3}{16} + 2(n-2)\frac{1}{16} = o(n)$.

ניזכר באי־שוויון מרקוב 21.3 ונגדיר אי־שוויון חדש

משפט 24.5 (אי־שוויון צ'בישב) *נניח ש־X משתנה מקרי בעל שונות (ולכן בעל תוחלת), אז*

$$\forall \lambda > 0, \ \mathbb{P}(|X - \mathbb{E}(X)| \ge \lambda) \le \frac{\mathrm{Var}(X)}{\lambda^2}$$

הוכחה. נשים לב ש־$\{|X - \mathbb{E}(X)| \ge \lambda\} = \{(X - \mathbb{E}(X))^2 \ge \lambda^2\}$.

נגדיר $Y = (X - \mathbb{E}(X))^2$ ואז $0 \le Y$ וכן $\mathbb{E}(Y) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathrm{Var}(X)$, ולכן מ־21.3

$$\mathbb{P}(Y \ge \lambda^2) \le \frac{\mathrm{Var}(X)}{\lambda^2}$$

□

דוגמה 24.2 מטילים מטבע הוגן 10^6 פעמים. X מספר העצים ואנו רוצים לדעת מה ההסתברות ש־$495000 < X < 505000$. אנו יודעים ש־$X \sim \mathrm{Bin}(10^6, \frac{1}{2})$.

$$\begin{aligned}\mathbb{P}(495000 < X < 505000) &= 1 - \mathbb{P}(X \le 495000 \vee X \ge 505000) \\ &= 1 - \mathbb{P}(|X - 500000| \ge 5000) \\ &\ge \frac{\mathrm{Var}(X)}{5000^2}\end{aligned}$$

אנו גם יודעים שמתקיים

$$X = \sum_{i=1}^{10^6} X_i \implies \mathrm{Var}(X) = \sum_{i=1}^{10^6} \mathrm{Var}(X_i) = 10^6 \cdot \frac{1}{4}$$

ולכן

$$\mathbb{P}(|X - 500000| \ge 5000) \le \frac{\mathrm{Var}(X)}{5000^2} = \frac{10^6 \cdot \frac{1}{4}}{5000^2} = \frac{1}{100}$$

ולכן ההסתברות הזאת גדולה מ־0.99, זאת־אומרת שמצאנו חסם מאוד טוב למספר ההטלות שקיבלו עץ באופן יחסי.

דוגמה 24.3 אם נחזור לדוגמה איתה פתחנו את ההרצאה, אז נוכל לקבוע

$$\mathrm{Var}(Y_n) < 100n \qquad \mathbb{E}(Y_n) = \frac{n-1}{4}$$

אז

$$\mathbb{P}\left(\left|\frac{Y_n}{\frac{n-1}{4}}-1\right|\geq\epsilon\right)\leq\frac{\mathrm{Var}(\frac{Y_n}{\frac{n-1}{4}})}{\epsilon^2}=\frac{\mathrm{Var}(Y_n)}{\epsilon^2(\frac{n-1}{4})^2}\leq\frac{o(n)}{\frac{\epsilon^2}{16}(n-1)^2}=o(\frac{1}{n})$$

משפט 24.6 (החוק החלש של המספרים הגדולים) *תהי* $X_1, X_2, \ldots$ *סדרת משתנים מקריים בלתי־תלויים שווי התפלגות ובעלי תוחלת* μ*.*
אם $Y_n = \frac{\sum_{i=1}^n X_i}{n}$*, אז לכל* $\epsilon > 0$

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

הוכחה. נוכיח בהנחת קיום שונות, כאשר ניתן להוכיח גם ללא הנחה זו.

$$\mathbb{E}(Y_n)=\frac{\mathbb{E}(\sum_{i=1}^n X_i)}{n}=\frac{\sum_{i=1}^n\mathbb{E}(X_i)}{n}=\mu$$

ולכן

$$\mathbb{P}(|Y_n-\mu|\geq\epsilon)\leq\frac{\mathrm{Var}(Y_n)}{\epsilon^2}=\frac{\mathrm{Var}(\frac{\sum_{i=1}^n X_i}{n})}{\epsilon^2}=\frac{\mathrm{Var}(\sum_{i=1}^n X_i)}{n^2\epsilon^2}=\frac{n\,\mathrm{Var}(X_1)}{n^2\epsilon^2}\xrightarrow[n\to\infty]{}0$$

□

25 שיעור 17 — 31.12.2024

25.1 בעיית אספן הקופונים

תרגיל 25.1 (אספן הקופונים) יהי אספן קופונים אשר מקבל כל יום קופון כלשהו מבין מספר קופונים אפשריים, מה החסם שמעיד שהאספן השיג את כל הקופונים?

פתרון נגדיר $X_1, \ldots, X_m$ משתנים מקריים בלתי־תלויים המתפלגים אחיד $U([n])$, כלומר m מספר סוגי הקופונים ו־n מספר השליפות, ואנו רוצים לחסום את $\mathbb{P}(\forall k \in [n], \exists i \in [m], X_i = k)$. נבחן את המאורע המשלים, $\mathbb{P}(\exists k \in [m], \forall i \in [n], X_i \neq k) = \mathbb{P}(\bigcup_{k=1}^n A_k)$ כאשר האיחוד איננו זר ו־$A_k = \forall i \in [m], X_i \neq k$, אפשר לחסום זאת עם חסם האיחוד,

$$\mathbb{P}(\bigcup_{k=1}^n A_k) \le \sum_{k=1}^n \mathbb{P}(A_k) = \sum_{k=1}^n (1-\frac{1}{n})^m = n(1-\frac{1}{n})^m$$

אנו רוצים להבין מתי הביטוי שהתקבל שואף ל־0, נשים לב שידוע כי $\forall x \in \mathbb{R}, 1+x \le e^x$, ולכן

$$n(1-\frac{1}{n})^m \le n(e^{-\frac{1}{n}})^m = ne^{-\frac{m}{n}}$$

אז אם $m = \lceil cn \log n \rceil$ ו־$c > 1$ אז

$$ne^{-\frac{m}{n}} = ne^{-c\log n} = nn^{-c} = n^{1-c} \xrightarrow[n\to\infty]{} 0$$

נבחין כי חסם האיחוד הוא מקרה פרטי של אי־שוויון מרקוב. נעבור לחישוב של תוחלת ושונות של כמה קופונים האספן לא השיג, נסמן ב־Y את מספר הקופונים החסרים, וכן $Y = \sum_{k=1}^n Y_n$ כאשר $Y_k = 1_{A_k}$. מהחישובים שעשינו עד כה נוכל להסיק

$$\mathbb{E}(Y) = n(1-\frac{1}{n})^m$$

שכן מאי־שוויון מרקוב $\mathbb{P}(Y \ge 1) \le \frac{\mathbb{E}(Y)}{1}$. נראה שבהצבה של m שקטן ממה שמצאנו (יחד עם $c < 1$) חסמנו מלמעלה את $\mathbb{P}(Y=0)$. נעבור לחישוב השונות של Y,

$$\mathrm{Var}(Y_k) = (1-\frac{1}{n})^m(1-(1-\frac{1}{n})^m) \le (1-\frac{1}{n})^m = \mathbb{E}(Y_k)$$

וכן

$$\begin{aligned} k \le l, \mathrm{Cov}(Y_k, Y_l) &= \mathbb{E}(Y_k \cdot Y_l) - \mathbb{E}(Y_k)\mathbb{E}(Y_l) \\ &= (1-\frac{2}{n})^m - (1-\frac{1}{n})^{2m} \\ &= (1-\frac{2}{n})^m - ((1-\frac{1}{n})^2)m \\ &= (1-\frac{2}{n})^m - ((1-\frac{2}{n}+\frac{1}{n^2})^2)m \\ &\le 0 \end{aligned}$$

ולכן $\mathrm{Var}(Y) \le n(1-\frac{1}{n})^m = \mathbb{E}(Y)$ (מחקנו את האיברים השליליים) ובהתאם

$$\frac{\mathrm{Var}(Y)}{(\mathbb{E}(Y))^2} \le \frac{1}{\mathbb{E}(Y)}$$

נשאר למצוא מתי $EE(Y) \to \infty$, עבור $m = cn\log n$ כאשר $1 > c$ נובע

$$\log(n(1-\frac{1}{n})^m) = \log(n) + m\log(1+\frac{1}{n}) = \log(n) + cn\log(1-\frac{1}{n}) \to \infty$$

כאשר את המקרה $c = 1$ אין לנו היכולת להראות.

25.2 בין הסתברות ללינארית

טענה 25.1 *X משתנה מקרי בעל שונות ו־$f(a) = \mathbb{E}((X-a)^2)$, אז f מקבלת מינימום ב־$a = \mathbb{E}(X)$.*

הוכחה.

$$f(a) = \mathbb{E}((X-a)^2) = \mathbb{E}(X^2 02aX + a^2) = \mathbb{E}(X^2) - 2a\mathbb{E}(X) + a^2$$

ולכן
$$f'(a) = -2\mathbb{E}(X) + 2a$$
ונובע ש־$f'(\mathbb{E}(X)) = 0$. □

אנו נתקלים בקושי של הקשר בין משתנים מקריים, תוחלת ושונות עם אלגברה לינארית.

הגדרה 25.2 (קבוצת כל המשתנים המקריים) נגדיר את L_2 להיות קבוצת כל המשתנים המקריים (במרחב הסתברות כלשהו) בעלי שונות, כאשר אנו מזהים משתנים מקריים ששווים כמעט תמיד.

טענה 25.3 *L_2 הוא מרחב מכפלה פנימית עם* $\langle X, Y\rangle = \mathbb{E}(XY)$.

לפני שניגש להוכחה נבחן דוגמה שתבהיר לנו את הטענה.

דוגמה 25.1 נגדיר $\Omega = [n]$ עם הסתברות אחידה, אז אם X משתנה מקרי אז $X : \Omega \to \Omega$, כלומר $L_2 = \mathbb{R}^\Omega$. מתקיים
$$\langle X, Y\rangle = \sum_{\omega\in\Omega} \mathbb{P}(\{\omega\})X(\omega)Y(\omega) = \sum_{k=1}^{n} \frac{1}{n} X(k)Y(k)$$
וזו אכן מכפלה פנימית.

נעבור להוכחה.

הוכחה. מהגדרת התוחלת של $\mathbb{E}(XY)$ קל להוכיח כי $\langle aX, Y\rangle = a\langle X, Y\rangle$ וכן ש־$\langle X + Y, Z\rangle = \langle X, Z\rangle + \langle Y, Z\rangle$, ואף $\mathbb{E}(X^2) = \langle X, X\rangle \iff X = 0$. אם מצטמצמים לתת־מרחב של המשתנים המקריים של $\mathbb{E}(X) = 0$ אז $\langle X, Y\rangle = \operatorname{Cov}(X, Y)$ ו־$\|X\| = \operatorname{Var}(X)$.
נשים לב כי הפעולה של התוחלת היא פונקציה ולכן $\mathbb{E} : L_2 \to \mathbb{R}$ לינארית ונחשוב על $\mathbb{R}$ כ־K כאשר K כל המשתנים המקריים הקבועים. נזהה את התוחלת על־ידי הטלה אורתוגונלית. כלומר $\mathbb{E}(X)$ אורתוגונלי ל־$X - \mathbb{E}(X)$ וזה ברור כי $\langle X, X - \mathbb{E}(X)\rangle = \mathbb{E}((X - \mathbb{E}(X))\mathbb{E}(X)) = 0$.
את ההוכחה הפורמלית נעשה עם משפט קושי־שוורץ. לכל $X, Y \in L_2$, נראה
$$|\mathbb{E}(XY)| \le \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$$
נגדיר
$$\overline{X} = \frac{X}{\sqrt{\mathbb{E}(X^2)}}, \qquad \overline{Y} = \frac{Y}{\sqrt{\mathbb{E}(X^2)}}$$
ואז $\mathbb{E}(\overline{X}^2) = \mathbb{E}(\overline{Y}^2) = 0$. ניזכר שאם X, Y אי־שליליים אז $0 \le XY \le \frac{X^2+Y^2}{2}$ ונניח שהמשתנים המקריים שלנו אי־שליליים. אז
$$\mathbb{E}(\overline{XY}) \le \mathbb{E}(\frac{\overline{X}^2 + \overline{Y}^2}{2}) = \frac{1}{2}(\mathbb{E}(\overline{X}^2) + \mathbb{E}(\overline{Y}^2)) = 1$$
□

26 תרגול 9 — 2.1.2025

26.1 תרגילים שונים בנושא שונות

דוגמה 26.1 סופר שואל סטודנטים בקמפוס אם הם קראו ספר שלו או לא.

אחוז הסטודנטים שקראו את הספר הוא q. אנו רוצים למצוא את q על־ידי שאילת רק חלק מהסטודנטים.

נניח שדגמנו n סטודנטים (מספר הסטודנטים הכללי לא ידוע), נסמן ב־X את מספר הסטודנטים שענו שהם קראו. נמצא כמה סטודנטים אנו צריכים לשאול כדי שיתקיים

$$\mathbb{P}(|\frac{X}{n} - q| \geq 0.1) < 0.05$$

נניח ש־$X \sim \mathrm{Bin}(n, q)$ (זוהי הנחה מסטטיסטיקה שלא נעסוק בה). לכן

$$\mathbb{E}(X) = nq, \mathbb{E}(\frac{X}{n}) = q, \mathrm{Var}(X) = nq(1-q) \leq n\frac{1}{4}$$

בהתאם

$$\mathrm{Var}(\frac{X}{n}) \leq \frac{1}{4n}$$

מאי־שוויון צ'בישב

$$\mathbb{P}(|\frac{X}{n} - q| \geq 0.1) < \frac{4n}{0.01} = \frac{25}{n} < \frac{5}{100} \iff 500 < n$$

תרגיל 26.1 הזמן (בחודשים) עד שתנור חדש מתקלקל מתפלג X עם תוחלת μ ושונות σ^2.

חברת ביטוח רוצה להציע אחריות לתנורים, היא חישבה שעל־מנת להרוויח היא צריכה שיחס מימושי הביטוח תהיה לכל היותר p.

כמה חודשי ביטוח על החברה להציע?

פתרון אנו מחפשים את p עבור $\mathbb{P}(X < M) < p$ מספר חודשים. נעשה שינוי ונכתוב $\{X < \mu - k\} \subseteq \{|X - \mu| > k\}$. צ'בישב

$$\mathbb{P}(X < \mu - k) \leq \mathbb{P}(|X - \mu| \geq k) < \frac{\mathrm{Var}(X)}{k^2} = \frac{\sigma^2}{k^2} < p$$

ולכן

$$k > \frac{\sigma}{\sqrt{p}}$$

ועל החברה להציע ביטוח לכל היותר ל־$\mu - \frac{\sigma}{\sqrt{p}}$ חודשים.

תרגיל 26.2 מגרילים מטריצה $n \times n$ ממקדמים ב־$\mathbb{Z}_{/2}$ על־ידי זה שמגרילים מטבע בכל כניסה באופן בלתי תלוי.

חשבו את התוחלת והשונות של מספר תתי־המטריצות 2×2 מהצורה $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

פתרון נסמן $X_{i,j}$ משתנה מקרי ברנולי לפי האם הכניסה ה־i, j היא 1 ו־$Y_{i,j}$ ברנולי של האם התת־מטריצה בגודל שתיים במקום ה־i, j היא מטריצת יחידות כרצוי. נבחין שמתקיים

$$Y_{i,j} = X_{i,j} X_{i+1,j} X_{i,j+1} X_{i+1,j+1}$$

נסמן

$$Y = \sum_{1 \leq i,j \leq n-1} Y_{i,j}$$

ונעבור לחישוב הערכים,

$$\mathbb{E}(Y) = (n-1)^2 \frac{1}{2^2}$$

וכן עבור השונות

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{Cov}(Y,Y)\\ &= \sum_{1\le i,j\le n-1}\mathrm{Cov}(Y_{i,j},Y_{i,j})\\ &= \sum_{1\le i,j\le n-1}\mathrm{Var}(Y_{i,j}) + \sum_{\substack{1\le i\le n-1\\ 1\le j\le n-2}}\mathrm{Cov}(Y_{i,j},Y_{i+1,j}) + \cdots + \sum_{1\le j\le n-2}\mathrm{Cov}(Y_{i,j},Y_{i+1,j+1}) + \sum_{\substack{1\le i,j\le n-1\\ |i-j|\ge 2}} 0\end{aligned}$$

ויש לנו שלושה סוגי חפיפה שנוספים אף הם ולא כתבנו, עתה נוכל לעבור לחישוב החלקים השונים ביתר קלות. לדוגמה

$$\mathrm{Cov}(Y_{i,j},Y_{i,j+1}) = \mathbb{E}(X_{i,j}X_{i+1,j}X_{i,j+1}X_{i+1,j+1}X_{i+1,j+1}X_{i+2,j+1}X_{i+1,j+2}X_{i+2,j+2}) - \mathbb{E}(Y_{i,j})\mathbb{E}(Y_{i,j+1}) = \frac{1}{2^6} - \frac{1}{2^8}$$

ולכן

$$\mathrm{Var}(Y) = (n-1)^2(\frac{1}{2^4} - \frac{1}{2^8}) + 4(n-1)(n-2)(\frac{1}{2^6} - \frac{1}{2^8}) + 4(n-2)^2(n-1)^2(\frac{1}{2^7} - \frac{1}{2^8})$$

27 שיעור 18 — 2.1.2025

27.1 מומנטים גבוהים

הגדרה 27.1 (מומנט) המומנט ה־k של X הוא $\mathbb{E}(X^k)$.

הגדרה 27.2 (מומנט מרכזי) המומנט המרכזי הוא $\mathbb{E}((X-\mathbb{E}(X))^k)$.

טענה 27.3 (צ'בישב מוכלל)

$$\mathbb{P}(|X-\mathbb{E}(X)|\geq\lambda)\leq\frac{\mathbb{E}((X-\mathbb{E}(X)))^k}{\lambda^k}$$

לכל k זוגי.

ההוכחה מאוד דומה להוכחה של אי־השוויון במקרה הרגיל.

דוגמה 27.1 מטילים מטבע הוגן n פעמים ואנו רוצים לחסום את ההסתברות שקיבלנו יותר מ $\frac{3}{4}n$ עצים.
מגדירים $X_1,\dots,X_n\sim\mathrm{Ber}(\frac{1}{2})$ וכן $X=\sum_{i=1}^n X_i$ ולכן

$$\mathbb{P}(X\geq\frac{3}{4}n)\leq\mathbb{P}(|X-\frac{n}{2}|\geq\frac{n}{4})\leq\frac{\mathrm{Var}(X)}{(\frac{n}{4})^2}$$

אבל מ־$X\sim Bin$ נסיק

$$\mathrm{Var}(X)=\frac{n}{4}$$

ולכן

$$\frac{\mathrm{Var}(X)}{(\frac{n}{4})^2}=\frac{\frac{n}{4}}{(\frac{n}{4})^2}=\frac{4}{n}$$

ננסה להשתמש בנוסחה החדשה.

$$\mathbb{E}((X-\frac{n}{2})^4)=\mathbb{E}((\sum_{i=1}^n(X_i-\frac{1}{2}))^4)$$

נגדיר $Y_i=X_i-\frac{1}{2}$ ולכן

$$\mathbb{E}((X-\frac{n}{2})^4)=\mathbb{E}((\sum_{i=1}^n Y_i)^4)=\mathbb{E}(\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n Y_iY_jY_kY_l)=\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n\mathbb{E}(Y_iY_jY_kY_l)$$

במצב הרגיל אנו יכולים לחלק למקרים עבור אינקסים זהים ושונים, הפעם יש לנו סוגי התלכדות שונים, נחשב לדוגמה את המקרה הזר, נניח שi,j,k,l שונים ולכן

$$\mathbb{E}(Y_iY_jY_kY_l)=0$$

למעשה מספיק שאחד מהם יהיה שונה מכל השאר, לדוגמה $i\neq j,k,l$ אז

$$\mathbb{E}(Y_iY_jY_kY_l)=\mathbb{E}(Y_i)\mathbb{E}(Y_jY_kY_l)=0$$

ולכן

$$\mathbb{E}((X-\frac{n}{2})^4)=\sum_{i=1}^n\mathbb{E}(Y_i^4)+\binom{4}{2}\sum_{i=1}^n\sum_{j=1,j\neq i}^n\mathbb{E}(Y_i^2)\mathbb{E}(Y_j^2)\leq Kn^2$$

ומאי־שוויון צבישב המוכלל נקבל

$$\mathbb{P}(|X-\frac{n}{2}|\geq\frac{3n}{4})\leq o(\frac{1}{n^2})$$

במקום לעבוד עם פולינומים נעבוד עם משתנים מערכיים על־ידי ההגדרה

$$Z=2^X$$

ולכן גם $Z-2^X=2^{\sum_{i=1}^n X_i}=\prod_{i=1}^n 2^{X_i}$, נגדיר את מכפלות אלה כ־$Z_i$ ואז

$$\mathbb{P}(X\geq\frac{3n}{4})=\mathbb{P}(Z\geq 2^{\frac{3n}{4}})\leq\frac{\mathbb{E}(Z)}{2^{\frac{3n}{4}}}=\frac{\prod_{i=1}^n\mathbb{E}(Z_i)}{2^{\frac{3n}{4}}}$$

וגם
$$\mathbb{E}(Z_i) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 2 = \frac{3}{2}$$
ולכן
$$\mathbb{E}(Z) = \left(\frac{3}{2}\right)^n$$
וכן
$$\frac{\prod_{i=1}^n \mathbb{E}(Z_i)}{2^{\frac{3n}{4}}} = \frac{\frac{3^n}{2^n}}{2^{\frac{3n}{4}}} = \left(\frac{\frac{3}{2}}{2^{\frac{3}{4}}}\right)^n \xrightarrow[n\to\infty]{} 0$$

הגדרה 27.4 (פונקציה יוצרת מומנטים) X משתנה מקרי, אז הפונקציה יוצרת המומנטים שלו היא
$$M_X(t) = \mathbb{E}(e^{tX})$$
והיא מוגדרת עבור חלק מערכי $t \in \mathbb{R}$.

טענה 27.5 (אי־שוויון צ'רנוף) *נניח X משתנה מקרי ו־$\lambda \in \mathbb{R}$ אז*
$$\mathbb{P}(X \geq \lambda) \leq \frac{M_X(t)}{e^{t\lambda}}$$
לכל $t > 0$ עבורו $M_X(t)$ מוגדרת.

הוכחה.
$$\mathbb{P}(X \geq \lambda) = \mathbb{P}(e^{tX} \geq e^{t\lambda}) \leq \frac{\mathbb{E}(e^{tX})}{e^{t\lambda}}$$
□

טענה 27.6 (כפליות פונקציה יוצרת מומנטים) *אם X, Y בלתי־תלויים אז*
$$M_{X+Y}(t) = M_X(t)\cdot M_Y(t)$$
לכל t בתחום ההגדרה של X, Y.

הוכחה.
$$\mathbb{E}(e^{t(X+Y)}) = \mathbb{E}(e^{tX}\cdot e^{tY}) = \mathbb{E}(e^{tX})\mathbb{E}(e^{tY})$$
□

דוגמה 27.2 נניח ש־$X \sim \mathrm{Ber}(p)$ אז
$$M_X(t) = \mathbb{E}(e^{tX}) = (1-p)e^{t\cdot 0} + pe^{t\cdot 1} = 1 + p(e^t - 1)$$
ועל־ידי הטענה האחרונה אם $X \sim \mathrm{Bin}(n,p)$ אז
$$M_X(t) = (1 + p(e^t-1))^n$$

דוגמה 27.3 נניח ש־$X \sim \mathrm{Poi}(\lambda)$ אז
$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{k=0}^\infty e^{tk}e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^\infty \frac{(\lambda\cdot e^t)^k}{k!} = e^{-\lambda}e^{e^t\lambda} = e^{\lambda(e^t-1)}$$

הערה אם ניקח $X_n \sim \mathrm{Bin}(n, \frac{\lambda}{n})$ אז $M_{X_n}(t) = (1 + \frac{\lambda(e^t-1)}{n})^n \xrightarrow[n\to\infty]{} e^{\lambda(e^t-1)} = M_X(t)$.

דוגמה 27.4 נניח ש־$X \sim \mathrm{Geo}(p)$ אז
$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{k=1}^\infty e^{tk}(1-p)^{k-1}p = pe^t\sum_{k=1}^\infty (e^t(1-p))^{k-1} = \frac{pe^t}{1-e^t(1-p)}$$
וזה מוגדר רק כאשר $e^t(1-p) < 1$.

משפט 27.7 (אי־שוויון הופדינג) *יהיו $X_1, \dots, X_n$ משתנים מקריים בלתי־תלויים כך ש־$\mathbb{E}(X_i) = 0$, ו־$|X_i| \leq 1$ כמעט תמיד, אז*
$$\mathbb{P}\left(\sum_{i=1}^n X_i \geq \lambda\right) \leq e^{-\frac{\lambda^2}{2n}}$$

את ההוכחה נראה בהרצאה הבאה, אבל כן נראה דוגמה

דוגמה 27.5 מטילים $n = 10^6$ מטבעות הוגנים באופן בלתי־תלוי,

$$0.99 \leq \mathbb{P}(495000 \leq X \leq 505000)$$

ונמצא חסם נוסף להסתברות זו. נגדיר $Y_i = 2(X_i - \frac{1}{2})$ כדי למרכז אותם, ואז

$$\mathbb{E}(Y_i) = 0, |Y_i| \leq 1$$

אז אפשר להשתמש באי־שוויון הופדינג על $Y = \sum_{i=1}^n Y_i$.

$$\mathbb{P}(Y \geq \lambda) \leq e^{-\frac{\lambda^2}{2n}}$$

רוצים לחסום את $|X - \frac{n}{2}| \geq 5000$ אז

$$Y = \sum_{i=1}^n Y_i = \sum_{i=1}^n 2X_i - 1 = 2X - n = 2(X - \frac{n}{2})$$

ולכן במקרה של Y אנו מחפשים את

$$|Y| \geq 10000$$

ועתה נוכל מטעמי סימטריה לבחון רק את המקרה החיובי,

$$\mathbb{P}(|X - \frac{n}{2}| \geq 5000) \leq 2\mathbb{P}(Y \geq 10000) \leq e^{\frac{10000^2}{2n}} = e^{-50}$$

28 שיעור 19 — 7.1.2024

28.1 פונקציה יוצרת מומנטים — המשך

בהרצאה הקודמת ראינו את אי־שוויון הופדינג 27.7, אי־שוויון שימושי במיוחד עבור חסמים, עתה נראה את ההוכחה שלו ודוגמות נוספות.

דוגמה 28.1 אם $X_i \sim \mathrm{U}(\{-1,1\})$ מקיימים את תנאי 27.7.

מאי־שוויון צ'בישב נקבל את החסם

$$\mathbb{P}(|\sum_{i=1}^n X_i| \geq \lambda) \leq \frac{\mathrm{Var}(\sum_{i=1}^n X_i)}{\lambda^2} = \frac{n}{\lambda^2}$$

בזמן שמאי־שוויון הופדינג נובע

$$\mathbb{P}(|\sum_{i=1}^n X_i| \geq \lambda) \leq 2\exp(-\frac{\lambda^2}{2n})$$

נראה עתה למה שנצטרך להוכחת אי־השוויון.

למה 28.1 (הלמה של הופדינג) *אם X משתנה מקרי כך ש־$\mathbb{E}(X) = 0$ ו־$|X| \leq 1$ אז*

$$M_X(t) = \mathbb{E}(e^{tX}) \leq e^{\frac{t^2}{2}}$$

לכל $t \in \mathbb{R}$.

הוכחה. יהי $t \in \mathbb{R}$. נגדיר את הפונקציה הקווית $L(x) = \frac{e^t - e^{-t}}{2}x + \frac{e^t + e^{-t}}{2}$ על־ידי שימוש בשיפוע על גרף הפונקציה e^{tX}. לכל $x \in [-1,1]$ נובע $e^{tx} \leq L(x)$. עבור X שמקיים $|X| \leq 1$ מתקיים $e^{tX} \leq L(X)$. לכן מתכונות התוחלת נובע $\mathbb{E}(e^{tX}) \leq \mathbb{E}(L(X))$.

נותר אם כן לחשב את הערך האחרון,

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

ומצאנו חסם, אך לא האחד שרצינו, נמשיך ונראה כי החסם המבוקש מתקיים אף הוא. מפיתוח טיילור נקבל

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

מהצד השני

$$e^{\frac{t^2}{2}} = \sum_{l=0}^{\infty}\frac{\left(\frac{t^2}{2}\right)^l}{l!} = \sum_{l=0}^{\infty}\frac{t^{2l}}{2^l l!}$$

אבל לכל $l \geq 0$ מתקיים

$$2^l l! = 2l(2l-2)\cdots 2 \leq 2l(2l-1)\cdots 1 = (2l)!$$

ולכן מצאנו את החסם הרצוי בדיוק. □

נעבור להוכחת אי־השוויון 27.7.

הוכחה. לפי הלמה לכל i,

$$M_{X_i}(t) \leq e^{\frac{t^2}{2}}$$

אנו גם יודעים ש־X_i בלתי־תלויים ולפי כפליות

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t) \leq (e^{\frac{t^2}{2}})^n = e^{\frac{nt^2}{2}}$$

ולכן לפי צ'רנוף

$$\mathbb{P}(\sum_{i=1}^n X_i \geq \lambda) \leq \frac{M_{\sum_{i=1}^n X_i}(t)}{e^{\lambda t}} \leq e^{\frac{nt^2}{2} - \lambda t}$$

נמצא את הערך הקן ביותר עבור חסם זה על־ידי מציאת ערך קטן ביותר לביטוי $\lambda t - \frac{nt^2}{2}$, נגזור את הביטוי ונקבל $nt - \lambda$ ולכן נבחר $t = \frac{\lambda}{n}$ ונקבל

$$e^{\frac{n(\frac{\lambda}{n})^2}{2} - \lambda\frac{\lambda}{n}} = e^{\frac{\lambda^2}{2n} - 2\frac{1}{2}\frac{\lambda^2}{n}} = e^{-\frac{\lambda^2}{2n}}$$

כפי שרצינו להראות. □

דוגמה 28.2 $X_i \sim \mathrm{U}([6])$ עבור $i \in [100]$ ונרצה לחשב את ההסתברות שממוצע מאה ההטלות הוא לפחות 4.

$$\mathbb{P}(\frac{\sum_{i=1}^{100} X_i}{100} \ge 4)$$

נמרכז את המשתנים המקריים על־ידי הגדרת $Y_i = \frac{X_i - \frac{7}{2}}{\frac{5}{2}}$ כדי לעמוד בדרישות אי־השוויון 27.7.
מההגדרה אכן $\mathbb{E}(Y_i) = 0$ וגם $|Y_i| \le 1$. לכן נובע מאי־שוויון הופדינג

$$\mathbb{P}(\sum_{i=1}^{100} Y_i \ge \lambda) \le e^{-\frac{\lambda^2}{100}}$$

ואנו רוצים למצוא את ערך λ המתאים כדי לתרגם את אי־השוויון הזה לאחד ששואל על $\frac{\sum_{i=1}^{100} X_i}{100} \ge 4$. מהצבה

$$X_i = \frac{5}{2}Y_i + \frac{7}{2}$$

ולכן

$$\frac{\sum_{i=1}^{100} \frac{5}{2}Y_i + \frac{7}{2}}{100} \ge 4 \iff \frac{\frac{5}{2}\sum_{i=1}^{100} Y_i}{100} + \frac{7}{2} \ge 4 \iff \frac{5}{2}\sum_{i=1}^{100} Y_i \ge 50 \iff \sum_{i=1}^{100} Y_i \ge 20$$

ולכן נקבל

$$\mathbb{P}(\frac{\sum_{i=1}^{100} X_i}{100} \ge 4) \le e^{-\frac{20^2}{200}} = e^{-2}$$

לבסוף נעיר הערה על השם פונקציה יוצרת מומנטים, הסיבה שאנו קוראים לה ככה הוא שהנגזרות שלה הם המומנטים, כלומר $\mathbb{E}(X^2)$ יהיה ערך הנגזרת השנייה של $M_X(t)$ וכן הלאה, כך נראה עתה.

טענה 28.2 *אם $M_X(t)$ מוגדרת בסביבת 0 אז היא חלקה בסביבת 0 ונגזרותיה הן המומנטים של X.*

הוכחה במקרה של תומך סופי.

$$M_X(t) = \sum_{s \in \mathrm{Supp}\, X} e^{ts}\mathbb{P}(X = s)$$

ולכן

$$M'_X(t) = \sum_{s \in \mathrm{Supp}\, X} se^{ts}\mathbb{P}(X = s)$$

ובאופן כללי

$$M_X^{(k)}(t) = \sum_{s \in \mathrm{Supp}\, X} s^k e^{ts}\mathbb{P}(X = s)$$

ולכן

$$M_X^{(k)}(0) = \sum_{s \in \mathrm{Supp}\, X} s^k \mathbb{P}(X = s) = \mathbb{E}(X^k)$$

□

28.2 מבוא למרחבי הסתברות רציפים

ניזכר שהגדרנו $(\Omega, \mathcal{F}, \mathbb{P})$ עבור מרחב הסתברות, אבל לא דיברנו על המשמעות של $\mathcal{F}$ כדי להבין מה היכולות האמיתיות של ההגדרה שלנו. ננסה עתה להגדיר מרחב הסתברות לא בדיד בצורה נאיבית. למען הפשטות נבחן את $\Omega = [0, 1]$. אנו רוצים ש־$\mathbb{P}([0, \frac{1}{2}]) = \frac{1}{2}$ וכן ש־$\mathbb{P}([\frac{1}{2}, 1]) = \frac{1}{2}$ אבל אז נובע ישירות ש־$\mathbb{P}(\{\frac{1}{2}\}) = 0$. ככלל נגדיר ש־$\mathbb{P}([a, b]) = b - a$. אבל בהגדרה זו אנו נתקלים בפרדוקס בשם בנך־טרסקי. אפשר לחלק את כדור היחידה ב־$\mathbb{R}^3$ למספר סופי של חלקים זרים ולהזיז ולסובב את החלקים ולהרכיב מהם שני כדורי יחידה. כדי לא להיתקל בפרדוקס הזה אנו הולכים להשתמש במידה ולהגביל את $\mathcal{F}$.

29 תרגול 10 — 9.1.2025

29.1 פונקציות יוצרות מומנטים

ניזכר כי המומנט ה־k של X הוא $\mathbb{E}(X^k)$.

דוגמה 29.1 נניח ש־$X \sim \mathrm{Ber}(p)$, אז $\forall t \in \mathbb{R}$,

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{x \in \mathrm{Supp}\, X} e^{tx}\mathbb{P}(X = x) = e^0\mathbb{P}(X = 0) + e^t\mathbb{P}(X = 1) = (1 - p) + pe^t$$

דוגמה 29.2 נניח ש־$X \sim \mathrm{Bin}(n, p)$, אז נשתמש בקשר להתפלגות ברנולי ונקבל

$$M_X(t) = \mathbb{E}(e^{tX}) = \mathbb{E}(e^{t\sum X_i}) = \mathbb{E}(\prod_{i=1}^{n} e^{tX_i}) = \prod_{i=1}^{n} M_{X_i}(t) = ((1 - p) + pe^t)^n$$

דוגמה 29.3 נגיד ש־$X \sim \zeta(2)$ אם $\mathrm{Supp}\, X = \mathbb{N}$ וכן

$$\mathbb{P}(X = n) = \frac{1}{cn^2}$$

כאשר

$$c = \sum \frac{1}{n^2} = \frac{\pi^2}{6}$$

ולכן

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{n=1}^{\infty} \frac{e^{tn}}{cn^2} < \infty \iff t \le 0$$

אז $\mathbb{E}(X) \sim \sum \frac{1}{n}$ ולכן X חסר תוחלת.

הערה נראה סימון נוסף למה שכבר ראינו על מומנטים,

$$\left.\frac{\partial^k M_X(t)}{\partial^k t}\right|_{t=0} = \mathbb{E}(X^k)$$

29.2 אי־שוויון צ׳רנוף

דוגמה 29.4 נניח ש־$X \sim \mathrm{Poi}(\lambda)$ ואנו רוצים לחסום את $\mathbb{P}(|X - \lambda| > \lambda)$.

נתחיל בחסימה על־ידי צ׳בישב, נובע

$$\mathbb{P}(|X - \lambda| > \lambda) < \frac{\mathrm{Var}(X)}{\lambda^2} = \frac{1}{\lambda}$$

ועתה נשתמש באי־שוויון צ׳רנוף,

$$\{|X - \lambda| > \lambda\} = \{X - \lambda > \lambda\} \cup \overbrace{\{X - \lambda < -\lambda\}}^{=0}$$

ולכן

$$\mathbb{P}(|X - \lambda| > \lambda) = \mathbb{P}(X > 2\lambda) < M_\lambda(t)e^{-t2\lambda} = e^{\lambda(e^t - 1)}e^{-t2\lambda} = \exp(\lambda e^t - \lambda - 2\lambda t)$$

נגזור את $f(t) = \lambda e^t - \lambda - 2\lambda t$ כדי למצוא את החסם היעיל ביותר,

$$f'(t) = \lambda e^t - 2\lambda = 0 \iff t = \log 2$$

בנוסף

$$f''(t) = \lambda e^t > 0$$

ולכן זוהי נקודת מינימום, וקיבלנו מאי־השוויון את החסם

$$\mathbb{P}(|X - \lambda| > \lambda) < \exp(\lambda e^{\log 2} - \lambda - 2\lambda \log 2) = \exp(2\lambda - \lambda - 2\lambda \log 2) = \frac{e^\lambda}{e^{\log 2 \cdot 2\lambda}} = \left(\frac{e}{4}\right)^\lambda$$

זהו כמובן חסם הרבה יותר הדוק, ואחד שנותן לנו מידע נוסף על התנהגות ההתפלגות.

תרגיל 29.1 מטילים מטבע הוגן n פעמים באופן בלתי־תלוי. נגדיר את X להיות מספר הפעמים שיצא הרצף $1, 1$. חסמו את $\mathbb{P}(X < \frac{n-1}{4} - c)$ לכל $c > 0$.

פתרון $X \sim ber(\frac{1}{4})$ שכן X מציין שיצא $1,1$ בהטלות $i, i+1$, כאשר $X = \sum_{i=1}^{n-1} X_i$.
נגדיר $Y_i = \frac{1}{4} - X_i$ כדי למרכז ולהתמודד עם הכיוון של אי־השוויון שאנו מחפשים.
נחשב ונקבל ש־$\operatorname{Supp} Y_i = \{\frac{1}{4}, -\frac{3}{4}\} \subseteq [-1,1]$ כפי שאנו רוצים.

$$\begin{aligned}\mathbb{P}(X < \frac{n-1}{3} - c) &= \mathbb{P}(\sum X_i < \frac{n-1}{4} - c) \\ &= \mathbb{P}(\sum(\frac{1}{4} - Y_i) < \frac{n-1}{4} - c) \\ &= \mathbb{P}(\frac{n-1}{4} - \sum Y_i < \frac{n-1}{4} - c) \\ &= \mathbb{P}(\sum Y_i > c)\end{aligned}$$

ולכן נותר לחשב את $\mathbb{P}(\sum Y_i > c)$, אבל בין המשתנים הללו יש תלות (שלא כמו X_i), כדי להשתמש באי־שוויון צ׳רנוף עלינו להתמודד עם בעיה זו. לשם כך נגדיר שני סכומים נפרדים,

$$S_1 = \sum_{\substack{i=1 \\ i \equiv 0 \mod 2}}^{n-1} Y_i, \qquad S_2 = \sum_{\substack{i=1 \\ i \equiv 1 \mod 2}}^{n-1} Y_i$$

עדיין יש תלות בין S_1 ו־S_2, אז נשתמש בעובדה ש־$\{S_1, S_2\} \subseteq \{S_1 > \frac{c}{2}\} \cup \{S_2 > \frac{c}{2}\}$, לכן מאי־שוויון הופדינג,

$$\mathbb{P}(\sum Y_i > c) = \mathbb{P}(S_1 + S_2 > c) \le \mathbb{P}(S_1 > \frac{c}{2}) + \mathbb{P}(S_2 > \frac{c}{2}) < \begin{cases} 2\exp(\frac{-(\frac{c}{2})^2}{n-1}) & n-1 \equiv 0 \mod 2 \\ \exp(\frac{-(\frac{c}{2})^2}{n-2}) + \exp(\frac{-(\frac{c}{2})^2}{n}) & n-1 \equiv 1 \mod 2 \end{cases}$$

30 שיעור 20 — 9.1.2025

30.1 משתנים מקריים לא בדידים

לפני שאנחנו מתחילים לעבוד עם משתנים כאלה, חשוב שנבין קודם איפה הם בכלל מופיעים, ומה המשמעות שלהם. במקרים בדידים ראינו מספר גדול מאוד של דוגמות לשאלות הסתברותיות על מקרים סופיים או בדידים, ועתה נראה דוגמות עבור המקרים הלא בדידים. ככלל, נדבר פה על מקרים שבהם יש לנו שאלות שהרזולוציה שלהן היא לא טבעית, כשלדוגמה ראשונה נוכל לדבר על משקלים, אלו הם מספרים שתנים שניתנים לדיוק כרצוננו, ואנו יכולים לדבר על התפלגות המשקל של אדם ברזולוציות שונות. המשמעות היא שמשתנה מקרי לא בדיד הוא משתנה מקרי שכשנמדוד אותו בכל אמת מידה נקבל התפלגות יחסית לאמת המידה, כך לדוגמה נוכל למדוד משקל בקילוגרמים, בגרמים, במיקרוגרמים וכן הלאה, בכל פעם נקבל אמות מידה מדויקות יותר ויותר. לכן הפעם במקום לשאול למה שווה משתנה מקרי, נשאל את השאלה מתי המשתנה המקרי נמצא בתחומי קטע מסוים, אך במקום זה נתאר את המקרים שבהם המשתנה המקרי נמצא בקרן, מתוך היכולת לחשב התפלגות בקטעים על־ידי חיסור קרניים.

הגדרה 30.1 (פונקציית התפלגות צוברת) X משתנה מקרי. פונקציית ההתפלגות המצטברת של X,

$$F_X : \mathbb{R} \to [0, 1], \qquad F_X(a) = \mathbb{P}(X \le a)$$

מייצגת את ההתפלגות של קרניים כפי שרצינו.

הערה F_X תלויה רק בהתפלגות X.

נעבור לתכונות פונקציות מעין אלה.

טענה 30.2 (תכונות של פונקציית התפלגות מצטברת) *אם X משתנה מקרי אז*

1. *F_X מונוטונית עולה (במובן החלש), $\forall a < b, F_X(a) \le F_X(b)$*
2. *$\lim_{a\to\infty} F_X(a) = 1$ וכן $\lim_{a\to-\infty} F_X(a) = 0$*
3. *F_X רציפה מימין, $\lim_{a\to b^+} F_X(a) = F_x(b)$*

הוכחה. נוכיח את כלל התכונות.

1. $\forall a < b, F_X(a) = \mathbb{P}(X \le a) \le \mathbb{P}(X \le b) = F_X(b)$
2. $\lim_{a\to\infty} F_X(a) = \lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} \mathbb{P}(X \le n)$, עברנו מהמקרה הרציף לגבול בדיד, ואז נקבל סדרה עולה של מאורעות מכילים ואז המסקנה נובעת ממשפט 6.3,
$$\lim_{n\to\infty} \mathbb{P}(X \le n) = \mathbb{P}(\Omega) = 1$$
באופן דומה גם
$$\lim_{a\to-\infty} F_X(a) = \lim_{n\to-\infty} F_X(a) = \lim_{n\to-\infty} \mathbb{P}(X \le -n) = \mathbb{P}(\emptyset) = 0$$
3. נשתמש בעובדה שפונקציה מונוטונית וחסומה היא בעלת גבול, ולכן
$$\lim_{a\to b^+} F_X(a) = \lim_{n\to\infty} F_X(b + \frac{1}{n}) = \lim_{n\to\infty} \mathbb{P}(X \le b + \frac{1}{n})$$
ומרציפות פונקציית ההסתברות ביטוי זה שווה ל־$\mathbb{P}(X \le b) = F_X(b)$.

□

דוגמה 30.1 נניח ש־$X \sim \mathrm{Ber}(p)$ ונבין את התנהגות F_X.
כאשר $a < 0$ אז $F_X(a) = 0$, כאשר $0 \le a < 1$ אז $F_X(a) = 1 - p$ ולבסוף בשאר התחום $F_X(a) = 1$.

טענה 30.3 *אם X משתנה מקרי, אז* $\mathbb{P}(X = a) = F_X(a) - \lim_{b\to a^-} F_X(b)$.

הוכחה.

$$\begin{aligned}F_X(a) - \lim_{b\to a^-} F_X(b) &= F_X(a) - \lim_{n\to\infty} F_X(a-\frac{1}{n})\\ &= \lim_{n\to\infty} F_X(a) - F_X(a-\frac{1}{n})\\ &= \lim_{n\to\infty} \mathbb{P}(X\le a) - \mathbb{P}(X \le a-\frac{1}{n})\\ &= \lim_{n\to\infty} \mathbb{P}(a-\frac{1}{n} < X \le a)\\ &\overset{(1)}{=} \mathbb{P}(X=a)\end{aligned}$$

כאשר (1) נובע מ־6.3. ☐

טענה 30.4 *אם X, Y משתנים מקריים כך ש־$F_X = F_Y$ אז $X \overset{d}{=} Y$, כלומר המשתנים שווי התפלגות.*

הוכחה (עבור משתנים מקריים בדידים). הראינו שניתן לחשב את $\mathbb{P}(X=a)$ מתוך F_X. ☐

את ההוכחה למקרה הכללי לא נוכל להראות בקורס זה שכן היא מתבססת על תורת המידה.

טענה 30.5 *אם F מקיימת את שלוש התכונות שראינו זה עתה, אז קיים משתנה מקרי X (על מרחב כלשהו) כך ש־$F \equiv F_X$.*

גם את הטענה הזו לא נוכל להוכיח בתחומי קורס זה. טענה זו כמובן חזקה במיוחד, שכן היא מספקת אפיון מלא למשתנים מקריים על־ידי הפונקציות המצטברות שלהם.

הגדרה 30.6 (משתנה מקרי רציף) משתנה מקרי X הוא **רציף** אם F_X רציפה.
טענה זו שקולה למקרה ש־$\mathbb{P}(X=a)=0$ לכל a.

הגדרה זו היא הגדרה בעייתית במקצת, היא לא מתכתבת עם ההגדרה הנפוצה בספרות המתמטית, והגדרנו אותה כך לצורך היכולת להבין בין כמה מקרים שנראה בקרוב.

דוגמה 30.2 קיים משתנה מקרי X עבורו F_X מוגדרת על־ידי

$$F_X(a) = \begin{cases} 0 & a \le 0 \\ a & 0 < a \le 1 \\ 1 & 1 < a \end{cases}$$

עבור X זה מתקיים $\mathbb{P}(X\in[a,b]) = b-a$ לכל $0 \le a \le b \le 1$. זאת שכן

$$\mathbb{P}(X\in[a,b]) = \mathbb{P}(a \le X \le b) = \mathbb{P}(X \le b) - \mathbb{P}(X < a) \overset{(1)}{=} \mathbb{P}(X\le b) - \mathbb{P}(X \le a) = F_X(b) - F_X(a) = b-a$$

כאשר (1) נובע מרציפות.

הפונקציה שראינו זה עתה היא פונקציה רציפה, ולכן היא אינטגרל של איזושהי פונקציה אחרת, רעיון זה נותן לנו השראה להגדרה נוספת ושימושית מאוד,

הגדרה 30.7 (משתנה מקרי רציף בהחלט) משתנה מקרי נקרא **רציף בהחלט** אם קיימת פונקציה $f_X : \mathbb{R} \to [0,\infty)$ אינטגרבילית כך שמתקיים

$$F_X(a) = \int_{-\infty}^a f_X(s)\,ds$$

פונקציה f_X כזו היא למעשה המקבילה של פונקציית ההסתברות הבדידה $p(\omega)$ שראינו כבר, והיא מספקת אפיון נוסף להתנהגות ההתפלגות.
הערה רציפות בהחלט גוררת רציפות, זאת שכן האינטגרל של פונקציה הוא פונקציה רציפה.

הגדרה 30.8 (פונקציית הצפיפות) ל־f_X קוראים פונקציית **הצפיפות** של X.

דוגמה 30.3 בדוגמה שראינו קודם מתקיים

$$f_X(s) = \begin{cases} 1 & 0 \le s \le 1 \\ 0 & \text{else} \end{cases}$$

טענה 30.9 *אם $f : \mathbb{R} \to [0,\infty)$ מקיימת*
$$\int_{-\infty}^{\infty} f(s)\,ds = 1$$
אז קיים משתנה מקרי X (על מרחב כלשהו) עבורו f היא פונקציית צפיפות.

הוכחה. נגדיר
$$F(a) = \int_{-\infty}^{a} f(s)\,ds$$
ונראה ששלוש התכונות הדרושות מתקיימות, כך שהטענה חלה. □

נראה טענה שמקבילה אף יותר את פונקציית הצפיפות לפונקציית ההסתברות הבדידה,

טענה 30.10 *נניח ש־X משתנה מקרי עם צפיפות f_X אז* $\mathbb{P}(a \le X \le b) = \int_a^b f_X(s)\,ds$.

הוכחה.
$$\begin{aligned}\mathbb{P}(a \le X \le b) &= \mathbb{P}(X \le b) - \mathbb{P}(X < a)\\ &= \mathbb{P}(X \le b) - \mathbb{P}(X \le a)\\ &= F_X(b) - F_X(a)\\ &= \int_{-\infty}^{b} f_X(s)\,ds - \int_{-\infty}^{a} f_X(s)\,ds\\ &= \int_{a}^{b} f_X(s)\,ds\end{aligned}$$
□

נבחן עתה מספר התפלגויות חשובות.

הגדרה 30.11 (התפלגות רציפה אחידה) X מתפלג אחיד על $[a,b]$, ונסמן $X \sim \mathrm{Unif}([a,b])$, אם
$$f_X(s) = \begin{cases}\frac{1}{b-a} & a \le s \le b\\ 0 & \text{else}\end{cases}$$

הגדרה 30.12 (התפלגות מעריכית) X מתפלג מעריכית אם פרמטר λ, ונסמן $X \sim \mathrm{Exp}(\lambda)$, אם
$$f_X(s) = \begin{cases}\lambda e^{-\lambda s} & 0 \le s\\ 0 & s < 0\end{cases}$$
נצטרך להראות שהגדרה זו בכלל תקפה על־ידי בדיקת האינטגרל לפי התנאי שמצאנו קודם,
$$\int_{-\infty}^{\infty} f_X(s)\,ds = \int_0^{\infty} \lambda e^{-\lambda s}\,ds = -e^{\lambda s}\mid_{s=0}^{s=\infty} = 0 - (-1) = 1$$

הגדרה 30.13 (התפלגות נורמלית סטנדרטית) X מתפלג נורמלי סטנדרטי, אם
$$f_X(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$$
כאשר ראינו באינפי 3 את ערך אינטגרל זה ובהתאם את ההצדקה להגדרה זו.

דוגמה 30.4 נניח ש־$X \sim \mathrm{Exp}(\lambda)$, ואנו רוצים למצוא a עבורו $\mathbb{P}(X \le a) = \frac{1}{2}$, ערך זה נקרא גם הציון.
$$\int_{-\infty}^{\infty} f_X(s)\,ds = \int_0^{\infty} \lambda e^{-\lambda s}\,ds = -e^{-\lambda a} - (-e^0) = 1 - e^{-\lambda a} = \frac{1}{2} \implies -\lambda a = -\ln 2$$
ולכן $a = \frac{\ln 2}{\lambda}$.

דוגמה 30.5 נניח ש־$X \sim \mathrm{Unif}([0,1])$ ואנו רוצים למצוא צפיפות של $Y = X^2$.

נחשב את F_Y ונגזור,

$$F_Y(a)=\mathbb{P}(Y\le a)=\mathbb{P}(X^2\le a)=\mathbb{P}(-\sqrt{a}\le X\le\sqrt{a})=\int_{-\sqrt{a}}^{\sqrt{a}} f_X(s)\,ds=\int_0^{\sqrt{a}} 1\,ds=\sqrt{a}$$

ולכן

$$F_Y(a)=\begin{cases}1 & a\ge 1\\ \sqrt{a} & 0\le a\le 1\\ 0 & a\le 0\end{cases}$$

אם נבחר $f_Y=F_y'$ נקבל צפיפות עבור Y,

$$f_Y(a)=\begin{cases}\frac{1}{2\sqrt{a}} & 0\le a\le 1\\ 0 & \text{else}\end{cases}$$

31 שיעור 21 — 14.1.2025

31.1 מעבר לעולם הרציף

הגדרה 31.1 (תוחלת רציפה) X משתנה מקרי רציף בעל פונקציית צפיפות f_X אז

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} s f_X(s)\, ds$$

ונבחין שעל האינטגרל להתכנס בהחלט כדי שנוכל לומר שהתוחלת מתכנסת.

נבחין שבעולם של תורת המידה הגדרה זו, יחד עם ההגדרה הבדידה, הן מקרים פרטיים של הגדרה רחבה יותר. לא נראה אותה במסגרת הקורס.

טענה 31.2 *אם $Y = g(X)$ אז*

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} g(s) f_X(s)\, ds$$

טענה זו נכונה גם כאשר Y לא רציף בהחלט (רציף).

כל התכונות שראינו עד היום על תוחלות, שונות, שונות משותפת וכן הלאה שראינו למשתנים מקריים בדידים נכונים גם עבורו המקרה הרציף. כולל אי־שוויון מרקוב, אי־שוויון צ'בישב, אי־שוויון הופדינג ואף צ'רנוף.

דוגמה 31.1 $X \sim \mathrm{Unif}([0,1])$, אז $f_X(s) = \begin{cases} 1 & 0 \le s \le 1 \\ 0 & \text{else} \end{cases}$, ולכן

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} s f_X(s)\, ds = \int_0^1 s\, ds\, \frac{s^2}{2}\Big|_{s=0}^{s=1} = \frac{1}{2}$$

וכן גם

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} s^2 f_X(s)\, ds = \int_0^1 s^2\, ds = \frac{s^3}{3}\Big|_{s=0}^{s=1} = \frac{1}{3}$$

ונסיק

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{1}{3} - \frac{1}{2^2} = \frac{1}{12}$$

באופן דומה כאשר $t \ne 0$ אז,

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{ts} f_X(s)\, ds = \int_0^1 e^{ts}\, ds = \frac{e^{ts}}{t}\Big|_{s=0}^{s=1} = \frac{e^t - 1}{t}$$

וכן $t = 0$ גורר ש־$M_X(t) = 1$.

31.2 צפיפות משותפת

הגדרה 31.3 (וקטור מקרי רציף) (X,Y) וקטור מקרי עם צפיפות f_{XY} אם

$$\mathbb{P}((X,Y) \in D) = \int_D f_{XY}(s,t)\, dt\, ds$$

וניזכר במשפט פוביני שיאפשר לנו לחשב את האינטגרלים הללו:

משפט 31.4 (משפט פוביני) *תהי $f : D \to \mathbb{R}$ עבור $D[a,b] \times [c,d] \subseteq \mathbb{R}^2$ אז*

$$\int_D f(s,t)\, ds\, dt = \int_a^b \left(\int_c^d f(s,t)\, dt \right) ds = \int_c^d \left(\int_a^b f(s,t)\, ds \right) dt$$

אם D היא קטע מורכב יותר, לדוגמה מהצורה $D = \{(x,y) \in \mathbb{R}^2 \mid a \le x \le b, c(x) \le b \le d(x)$ עבור פונקציות $c, d : [a,b] \to \mathbb{R}$ מתאימות, אז נוכל לחשב

$$\int_D f(x,y)\, ds\, dt = \int_a^b \int_{c(s)}^{d(s)} f(s,t)\, dt\, ds$$

דוגמה 31.2 נגדיר
$$f_{XY}(s,t)=\begin{cases}6s & 0\le s,0\le t,s+t\le 1\\ 0 & \text{else}\end{cases}$$
ונרצה לראות אם היא מגדירה פונקציית צפיפות,
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{XY}(s,t)\,dt\,ds=\int_0^1\int_0^{1-t}6s\,ds\,dt=\int_0^1 3s^2\Big|_{s=0}^{s=1-t}\,dt=\int_0^1 3(1-t)^2\,dt=-(1-t)^3\Big|_{t=0}^{t=1}=0-(-1)^3=1$$
והתנאי ההכרחי לפונקציית צפיפות אכן מתקיים, וזוהי אכן פונקציית צפיפות.

הערה התוחלת של וקטור מקרי באופן מאוד דומה תהיה
$$\mathbb{E}(g(X,Y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(s,t)f_{XY}(s,t)\,dt\,dt$$

טענה 31.5 (צפיפות שולית) *אם X ו־Y הם משתנים מקריים רציפים בעלי צפיפות משותפת f_{XY}, אז*
$$f_X(s)=\int_{-\infty}^{\infty}f_{XY}(s,t)\,dt$$

הוכחה.
$$\mathbb{P}(X\le a)=\int_{-\infty}^{a}\int_{-\infty}^{\infty}f_{XY}(s,t)\,ds\,dt$$
אבל מההגדרה יש התכנסות ולכן יש הצדקה להגדרה
$$f_X(s)=\int_{-\infty}^{\infty}f_{XY}(s,t)\,dt$$
ומתקבל
$$\mathbb{P}(X\le a)=\int_{-\infty}^{a}f_X(s)\,ds$$
כפי שרצינו. □

הגדרה 31.6 (חוסר תלות במשתנים מקריים רציפים) X ו־Y בלתי־תלויים אם
$$\mathbb{P}(X\in S,Y\in T)=\mathbb{P}(X\in S)\mathbb{P}(Y\in T)$$
זוהי ההגדרה המלאה גם עבור המקרה הבדיד, אך כמו במקרה הבדיד ישנה הגדרה שקולה שתהיה לנו לעזר,

הגדרה 31.7 (הגדרה שקולה לאי־תלות משתנים מקריים רציפים) אם X ו־Y משתנים מקריים רציפים בעלי צפיפות משותפת f_{XY} המקיימת
$$f_{XY}(s,t)=f_X(s)f_Y(t)$$
אז X ו־Y בלתי־תלויים.

דוגמה 31.3 נגדיר
$$f_{XY}(s,t)=\begin{cases}1 & 0\le s,t\le 1\\ 0 & \text{else}\end{cases}$$
אז
$$f_X(s)=\int_{-\infty}^{\infty}f_{XY}(s,t)=\begin{cases}\int_0^1 1\,dt=1 & 0\le s\le 1\\ 0 & \text{else}\end{cases}$$
ומצאנו שאכן תכונת חוסר התלות חלה.

32 תרגול 11 — 16.1.2025

32.1 משתנים מקריים רציפים בהחלט

תרגיל 32.1 יהי X משתנה מקרי רציף בהחלט עם פונקציית צפיפות

$$f_X(t) = \begin{cases} c(2t-t^2) & t \in [0,2] \\ 0 & \text{else} \end{cases}$$

חשבו את c וחשבו את ההתפלגות המצטברת.

פתרון נשתמש בתנאי ההכרחי לפונקציית צפיפות,

$$1 = \mathbb{P}(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f_X(t)\,dt = \int_0^2 c(2t-t^2)\,dt = c(t^2 - \frac{1}{3}t^3)\Big|_{t=0}^{t=2} = c \cdot \frac{4}{3}$$

ולכן $c = \frac{3}{4}$ בלבד.

נעבור לחישוב פונקציית ההתפלגות המצטברת,

$$F_X(s) = \mathbb{P}(X \le s) = \int_{-\infty}^{s} f_X(t)\,dt$$

כשאר $s \in [0,2]$,

$$F_X(s) = \int_0^s \frac{3}{4}(2t-t^2)\,dt = \frac{3}{4}(s^2 - \frac{1}{3}s^3)$$

אם $s < 0$ אז $F_X(s) = 0$ ואם $s > 2$ אז $F_X(s) = 1$.

דוגמה 32.1 נניח ש־$X \sim \text{Exp}(\lambda)$, אז

$$F_X(t) = \int_{-\infty}^{t} f_X(x)\,dx = \int_0^t \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\big|_{x=0}^{x=t} = 1 - e^{-\lambda t}$$

דוגמה 32.2 זמן שיחת טלפון ממוצעת מתפלגת מעריכית $\lambda = \frac{1}{10}$, מה ההסתברות ששיחה נמשכה יותר מעשר דקות?

אנו מחפשים את $\mathbb{P}(X > 10) = 1 - \mathbb{P}(X \le 10) = 1 - 1 - e^{-\frac{1}{10}\cdot 10} = e^{-1}$.

נניח שהשיחה נמשכה יותר מ־10 דקות ונחשב את ההסתברות שהיא נמשכה יותר מ־20 דקות,

$$\mathbb{P}(X > 20 \mid X > 10) = \frac{\mathbb{P}(X > 20, X > 10)}{\mathbb{P}(X > 10)} = \frac{e^{-2}}{e^{-1}} = e^{-1}$$

אנו רוצים לבדוק אם תכונת חוסר הזיכרון חלה על התפלגות זו, נבדוק.

$$\mathbb{P}(X > s+t \mid X > t) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbb{P}(X > s)$$

דוגמה 32.3 נגדיר $Y = [x]$ וכן $X \sim \text{Exp}(\lambda)$, אז $\operatorname{Supp} Y = \mathbb{N} \cup \{0\}$,

$$\mathbb{P}(Y = n) = \mathbb{P}([x] = n) = \mathbb{P}(n \le X \le n+1) = \int_n^{n+1} \lambda e^{-\lambda t}\,dt = e^{-\lambda n} - e^{-\lambda(n+1)} = e^{-\lambda n}(1-e^{-\lambda}) - 1 \sim \text{Geo}(1-e^{-\lambda})$$

משתנה מקרי גאומטרי ומשתנה מקרי מעריכי הם אם כך בעלי קשר, לא במקרה שניהם בעלי ההתפלגויות חסרות הזיכרון היחידות.

תרגיל 32.2 נניח ש־$X \sim \text{Exp}(\lambda)$ ונגדיר $Z = \log X$, חשבו את ההתפלגות של Z.

פתרון יהי $t \in \mathbb{R}$, אז

$$F_X(t) = \mathbb{P}(Z < t) = \mathbb{P}(\log X < t) = \mathbb{P}(X < e^t) = F_X(e^t) = 1 - e^{-\lambda e^t}$$

וכן נחשב את פונקציית הצפיפות, $f_X(t) = F_X'(t) = e^{-\lambda e^t} \cdot \lambda e^t$.

דוגמה 32.4 נחשב את התוחלת של $X \sim \text{Exp}(\lambda)$,

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t)\,dt = \int_0^{\infty} t\lambda e^{-\lambda t}\,dt = \frac{1}{\lambda}\int_0^{\infty} ue^{-u}\,du = \frac{1}{\lambda}\big((-ue^{-u})_{u=0}^{u=\infty} + \int_0^{\infty} e^{-u}\,du\big) = \frac{1}{\lambda}(0 + (1-0)) = \frac{1}{\lambda}$$

באופן דומה גם

$$\mathbb{E}(X^2) = \int_0^{\infty} t^2 e^{-\lambda t}\,dt = \frac{1}{\lambda^2}\int_0^{\infty} u^2 e^{-u}\,du = \frac{1}{\lambda^2}\big((-u^2e^{-u})_{u=0}^{u=\infty} + 2\int_0^{\infty} ue^{-u}\,duoo\big) = \frac{2}{\lambda^2} \implies \text{Var}(X) = \frac{1}{\lambda^2}$$

33 שיעור 22 — 16.1.2025

33.1 אי־תלות במשתנים מקריים רציפים

בשיעור הקודם דיברנו ואפיינו אי־תלות משתנים מקריים רציפים, נראה עתה דוגמה למקרים כאלה.

דוגמה 33.1 נניח ש־$X, Y \sim \mathrm{Exp}(1)$ בלתי־תלויים ונרצה לחשב את $\mathbb{P}(Y \ge 2X)$. נובע מההגדרה של התפלגות מעריכית,

$$f_{XY}(s,t) = f_X(s)f_Y(t) = \begin{cases} e^{-s-t} & t, s \ge 0 \\ 0 & \text{else} \end{cases}$$

נעבור לחישוב ההתפלגות שאנו מחפשים,

$$\mathbb{P}(Y \ge 2X) = \int_0^\infty \int_{2s}^\infty e^{-s-t}\,dt\,ds = \int_0^\infty e^{-s} \int_{2s}^\infty e^{-t}\,dt\,ds$$

וכן

$$e^{-s}\int_{2s}^\infty e^{-t}\,dt = -e^{-t}\big|_{t=2s}^{t=\infty} = 0 - (-e^{-2s}) = e^{-2s}$$

ובהתאם

$$\int_0^\infty e^{-s}\int_{2s}^\infty e^{-t}\,dt\,ds = \int_0^\infty e^{-s}e^{-2s}\,ds = \int_0^\infty e^{-3s}\,ds = -\frac{1}{3}e^{-3s}\Big|_{s=0}^{s=\infty} = -\frac{1}{3}(0-1) = \frac{1}{3}$$

33.2 התפלגות נורמלית

טענה 33.1 (נוסחת הקונבולוציה) *אם X ו־Y בעלי צפיפות f_X ו־f_Y בלתי־תלויים בהתאמה אז למשתנה המקרי $Z = X + Y$ יש צפיפות*

$$f_Z(t) = \int_{-\infty}^\infty f_X(s)f_Y(t-s)\,ds$$

ניזכר שבמקרה הבדיד,

$$\mathbb{P}(Z = t) = \sum_{s \in \mathbb{R}} \mathbb{P}(X = s)\mathbb{P}(Y = t - s)$$

את הטענה הזו לא נוכל להוכיח למקרה הרציף.

משפט 33.2 (משפט הגבול המרכזי) *אם $\{X_i \mid i < \mathbb{N}\}$ משתנים מקריים בלתי־תלויים שווי־התפלגות עם $\mathbb{E}(X_i) = 0, \mathrm{Var}(X_i) = 1$ לכל $i \in \mathbb{N}$ ונגדיר*

$$Z_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$$

(אז $\mathbb{E}(Z_n) = 0, \mathrm{Var}(Z_n) = 1$) אז

$$\mathbb{P}(Z_n \le a) \xrightarrow[n \to \infty]{} \int_{-\infty}^a \frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}\,ds$$

כלומר התפלגות זו שווה עבור $\mathbb{P}(Z \le a)$ עבור $Z \sim \mathrm{N}(0,1)$

הכוונה היא שהמשפט הוא המרכזי, לא הגבול. נראה את ההגדרה המלאה להתפלגות נורמלית,

הגדרה 33.3 (התפלגות נורמלית) $Z \sim \mathrm{N}(\mu, \sigma^2)$ אם הצפיפות שלו מקיימת

$$f_Z(s) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(s-\mu)^2}{2\sigma^2}\right)$$

כאשר $\mathrm{N}(0,1)$ נקראת ההתפלגות הנורמלית הסטנדרטית.

אם μ משתנה, אז נקבל הזזה של ההתפלגות, אם נשנה את σ^2 אז נשנה את השונות, הרדיוס, הריווח של ההתפלגות.

דוגמה 33.2 נניח ש־$a > 0$ וש־$b \in \mathbb{R}$, נניח ש־$X \sim \mathrm{N}(0,1)$ ו־$Y = aX + b$, אנו יודעים שמתקיים

$$f_X(s) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{s^2}{2})$$

ונרצה לחשב את $f_Y(t)$,

$$F_Y(t) = \mathbb{P}(Y \le t) = \mathbb{P}(aX + b \le t) = \mathbb{P}(X \le \frac{t-b}{a}) = \int_{-\infty}^{\frac{t-b}{a}} \frac{1}{\sqrt{2\pi}} \exp(-\frac{s^2}{2})\, ds = F_X(\frac{t-b}{a})$$

אבל

$$F_Y'(t) = F_X'\left(\frac{t-b}{a}\right) \cdot \frac{1}{a} = \frac{1}{\sqrt{2\pi a^2}} \exp\left(-\frac{(t-b)^2}{2a^2}\right)$$

דוגמה 33.3 נחשב את התוחלת של התפלגות נורמלית, נשתמש באי־זוגיות ונקבל,

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} s \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right)\, ds = 0$$

באופן דומה על־ידי אינטגרציה בחלקים נובע

$$\mathrm{Var}(X) = \mathbb{E}(X^2) = \int_{-\infty}^{\infty} s^2 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right)\, ds = 1$$

מסקנה 33.4 אם $Y \sim \mathrm{N}(\mu, \sigma^2)$ אז $\mathbb{E}(Y) = \mu$ ו־$\mathrm{Var}(Y) = \sigma^2$.

טענה 33.5 *אם $X_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ עבור $i \in [2]$ כך שהם בלתי־תלויים, אז*

$$X_1 + X_2 \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

ההוכחה מושארת כתרגיל ומופיעה בספר.

33.3 התפלגות סדרת התפלגויות

נדבר על התכנסות בהתפלגות של משתנים מקריים, נבחין שהשם מוזר קצת כי הוא קופץ מעל המשתנים המקריים ועובר ישר למשתנים המקריים, זאת שכן במקרים אלה הם לא מעניינים אותנו באופן ישיר.

הגדרה 33.6 (התכנסות משתנים מקריים בהתפלגות) $\{X_i\}_{i\in\mathbb{N}}$ סדרת משתנים מקריים, Z משתנה מקרי רציף.
נאמר ש־$X_n \xrightarrow{d} Z$ (X_n מתכנסת בהתפלגות ל־Z) אם לכל $a \in \mathbb{R}$,

$$F_{X_n}(a) \xrightarrow[n\to\infty]{} F_Z(a)$$

דוגמה 33.4 נניח ש־$X_n \sim \mathrm{U}([n])$ אז אם $Y_n = \frac{X_n}{n}$ אז $Y_n \xrightarrow{d} Z$ כאשר $Z \sim ([0,1])$.
נבחין שמההגדרה $Y_n \sim \mathrm{U}(\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\})$. נגדיר

$$f_Z(s) = \begin{cases} 1 & 0 \le s \le 1 \\ 0 & \text{else} \end{cases}, \qquad F_Z(a) = \begin{cases} 0 & a \le 0 \\ a & 0 \le a \le 1 \\ 1 & \text{else} \end{cases}$$

אם $a \le 0$ אז $F_Z(a) = 0 \to F_Y(a) = 0$. אם $a \ge 1$ אז $F_Z(a) = 1 \to F_Y(a) = 1$. אם $0 < a < 1$ אז

$$F_{Y_n}(a) = \mathbb{P}(Y_n \le a) = \mathbb{P}(X_n \le n \cdot a) \le \frac{\lceil n \cdot a \rceil}{n} \xrightarrow{n\to\infty} a$$

נעבור לבחון את הקשר שבין התפלגות גאומטרית ומעריכית.

דוגמה 33.5 נניח ש־$X_n \sim \mathrm{Geo}(\frac{\lambda}{n})$ עבור $\lambda > 0$. עבור $a > 0$,

$$\begin{aligned}
F_{Y_n}(a) &= \mathbb{P}(Y_n \leq a) \\
&= \mathbb{P}(X_n \leq na) \\
&= 1 - \mathbb{P}(X_n > nk) \\
&= 1 - \sum_{k=\lfloor nk \rfloor + 1}^{\infty} \mathbb{P}(X_n = k) \\
&= 1 - \sum_{k=\lfloor nk \rfloor + 1}^{\infty} (1 - \frac{\lambda}{n})^{k-1} \frac{\lambda}{n} \\
&= 1 - (1 - \frac{\lambda}{n})^{k-1} \frac{\lambda}{n} \sum_{l=1}^{\infty} (1 - \frac{\lambda}{n})^{l} \\
&= 1 - (1 - \frac{\lambda}{n})^{\lfloor na \rfloor} \\
&= 1 - ((1 - \frac{\lambda}{n})^{n})^{\lfloor a \rfloor} \\
&\to 1 - e^{-\lambda a}
\end{aligned}$$

זאת אומרת ש־$F_Z(a) = 1 - e^{-\lambda a}$ ו־$Z \sim \mathrm{Exp}(\lambda)$.

34 שיעור 23 — 21.1.2025

34.1 התכנסות משתנים מקריים

טענה 34.1 *X_n סדרת משתנים מקריים כך ש־$\mathbb{E}(X_n)=0$ לכל $n\in\mathbb{N}$, ו־$\operatorname{Var}(X_n)\xrightarrow{n\to\infty}0$, אז $X_n\xrightarrow{d}0$.*

הוכחה. תרגיל, כאשר זוהי תוצאה של צ'בישב. □

טענה 34.2 *אם X_n משתנים מקריים בדידים הנתמכים על $\mathbb{N}$, ו־X גם נתמך על הטבעיים, אז $X_n\xrightarrow{d}X$ אם ורק אם $\mathbb{P}(X_n=k)\xrightarrow{n\to\infty}\mathbb{P}(X=k)$ לכל $k\in\mathbb{N}$.*

ראינו שאם $X_n\sim\operatorname{Bin}(n,\frac{\lambda}{n})$ ו־$X\sim\operatorname{Poi}(\lambda)$ אז מתקיימת הטענה ו־$X_n\xrightarrow{d}X$.

הוכחה. נניח ש־$\operatorname{Supp}X=\mathbb{N}$. עבור $k<a<k+1$ צריך לבדוק שמתקיים $F_{X_k}(a)=\mathbb{P}(X_n\le a)=\sum_{i=1}^k\mathbb{P}(X_n=i)$, ושבאותו אופן $F_X(a)=\sum_{i=1}^k\mathbb{P}(X=i)$, לכן צריך להראות שלכל k, $\sum_{i=1}^k\mathbb{P}(X_n=i)\xrightarrow{n\to\infty}\sum_{i=1}^k\mathbb{P}(X=i)$ אם ורק אם $\mathbb{P}(X_n=i)\xrightarrow{n\to\infty}\mathbb{P}(X=i)$.

הכיוון השני מיידי, ולכן נעבור להוכחת הכיוון הראשון, באינדוקציה על k. עבור $k=1$ הטענה נובעת מיד, ואז נניח שהטענה נכונה עבור $k\in\mathbb{N}$ ונבדוק את $k+1$, אנו כבר יודעים שמתקיים

$$\sum_{i=1}^k\mathbb{P}(X_n=i)+\mathbb{P}(X_n=k+1)\xrightarrow{m\to\infty}\sum_{i=1}^k\mathbb{P}(X=i)+\mathbb{P}(X=k+1)$$

נראה הוכחה פשוטה אף יותר,

$$\sum_{i=1}^{k+1}\mathbb{P}(X_n=i)-\sum_{i=1}^{k}\mathbb{P}(X_n=i)\xrightarrow{n\to\infty}\sum_{i=1}^{k+1}\mathbb{P}(X=i)-\sum_{i=1}^{k}\mathbb{P}(X=i)$$

אבל למעשה מכאן נובע ישירות $\mathbb{P}(X_n=k+1)\to\mathbb{P}(X=k+1)$. □

ניזכר במשפט 33.2 ונראה לו דוגמה.

דוגמה 34.1 נניח $X_n\sim\operatorname{U}([\sigma])$ בלתי־תלויים ואנו רוצים להעריך את

$$\mathbb{P}(3500\le\sum_{i=1}^{1000}X_i\le 3700)$$

נתחיל בנרמול, נגדיר $Y_i=\frac{X_i-\frac{7}{2}}{\sqrt{\frac{35}{12}}}$, לכן נובע $\mathbb{E}(Y_i)=0,\operatorname{Var}(Y_i)=1$. נציין שגם $X_i=\sqrt{\frac{35}{12}}Y_i+\frac{7}{2}$, ובהתאם

$$\begin{aligned}\mathbb{P}(3500\le\sqrt{\frac{35}{12}}\sum_{i=1}^{1000}Y_i+3500\le 3700)&=\mathbb{P}(0\le\sqrt{\frac{35}{12}}\sum_{i=1}^{1000}Y_i\le 200)\\&=\mathbb{P}(0\le\sum_{i=1}^{1000}Y_i\le\sqrt{\frac{12}{35}}200)\\&=\mathbb{P}(0\le\frac{1}{\sqrt{1000}}\sum_{i=1}^{1000}Y_i\le\frac{200}{\sqrt{\frac{35}{12}}\sqrt{1000}})\\&\approx\int_0^1\frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}\,ds\end{aligned}$$

דוגמה 34.2 נרצה לחשב את הגבול של $e^{-n}\sum_{k=0}^n\frac{n^k}{k!}$. נוכל כמובן להשתמש בטור טיילור ולנסות לחשב את הערך הזה ישירות, אך נראה דרך נוספת. למעשה שאלה זו שקולה לחלוטין ל־$\mathbb{P}(X\le n)$ עבור $X\sim\operatorname{Poi}(n)$. ניזכר שגם אם $X\sim\operatorname{Poi}(\lambda_1)$ וכן $Y\sim\operatorname{Poi}(\lambda_2)$ אז $X+Y\sim\operatorname{Poi}(\lambda_1+\lambda_2)$ אז נסיק שאם $X_i\sim\operatorname{Poi}(1)$ עבור $1\le i\le n$ נקבל

$$\operatorname{Poi}(n)\sim\sum_{i=1}^n X_i$$

ולכן נוכל לנסח מחדש את השאלה כגבול $\mathbb{P}(\sum_{i=1}^{n} X_i \le n)$ עבור X_i בלתי־תלויים המתפלגים פואסונית 1. נבחין שגם $\mathbb{E}(X_i) = 1$ ולכן ערך זה שקול לערך

$$\mathbb{P}(\sum_{i=1}^{n}(X_i - 1) \le 0)$$

נגדיר $Y_i = X_i - 1$ בלתי־תלויים כך ש־$\mathbb{E}(Y_i) = 0$ ו־$\mathrm{Var}(Y_i) = 1$. אז ממשפט הגבול המרכזי נסיק

$$\mathbb{P}(\sum_{i=1}^{n}(X_i - 1) \le 0) = \mathbb{P}(\frac{\sum_{i=1}^{n}(X_i - 1)}{\sqrt{n}} \le 0) \to \mathbb{P}(Z \le 0) = \frac{1}{2}$$

34.2 סדרות של מאורעות

נעבור לדבר על העולם של סדרות של מאורעות, במטרה לאפיין את ההתנהגות של המאורעות האלה ולבנות כלים שיאפשרו לנו לשאול שאלות מוכללות על הסדרות, נעבור עוד רגע להגדיר שאלות כאלה באופן פורמלי ומדויק.

הגדרה 34.3 (מאורעות מיוחדים עבור סדרת מאורעות) תהי $\{A_n\}_{n=1}^{\infty}$ סדרת מאורעות.

נגדיר

$$\limsup_{n\to\infty} A_n = \{A_n \text{ infinitely often}\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\omega \in \Omega \mid \forall n \in \mathbb{N}, \exists k \ge n, \omega \in A_k\}$$

עבור המאורע שקורה בסדרה אינסוף פעמים. כלומר המקרה שאינסוף פעמים הם התרחשו, לדוגמה שאינסוף פעמים מטבע נפל על עץ.

נגדיר גם

$$\liminf_{n\to\infty} A_n = \{A_n \text{ eventually}\} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \{\omega \in \Omega \mid \exists n \in \mathbb{N}, \forall k \ge n, \omega \in A_k\}$$

עבור מאורע שיקרה מתישהו.

נסתכל על משתנים מקריים שמייצגים את המאורעות האלה, במטרה לנסות להבין את ההגדרות האלה.

תרגיל 34.1 נגדיר 1_{A_n}, המשתנה המקרי המציין, אז נובע

$$1_{\limsup_{n\to\infty} A_n} = \limsup_{n\to\infty} 1_{A_n}$$

ובאופן דומה

$$1_{\liminf_{n\to\infty} A_n} = \liminf_{n\to\infty} 1_{A_n}$$

הערה מדה־מורגן נובע

$$\limsup_{n\to\infty} A_n^C = (\liminf_{n\to\infty} A_n)^C$$

למה 34.4 (הלמה של פאטו) *עבור סדרה מאורעות* $\{A_n\}$ *מתקיים,*

$$\mathbb{P}(\liminf_{n\to\infty} A_n) \le \liminf_{n\to\infty} \mathbb{P}(A_n)$$

וכן באופן דומה,

$$\mathbb{P}(\limsup_{n\to\infty} A_n) \ge \limsup_{n\to\infty} \mathbb{P}(A_n)$$

הוכחה. נוכיח את אחד מהחסמים ואת השני נקבל מההערה.

$$\mathbb{P}(\liminf_{n\to\infty} A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k)$$

ולכן נגדיר $B_n = \bigcap_{k=n}^{\infty} A_k$. זוהי כמובן סדרה עולה, זאת שכן ככל שאנו מתקדמים בסדרה אנו חותכים, כלומר $A_n \cap B_{n+1} = B_n$. לכן נובע מ־6.3 כי

$$\mathbb{P}(\bigcup_{n=1}^{\infty} B_n) = \lim_{n\to\infty} \mathbb{P}(B_n) \le \liminf_{n\to\infty} \mathbb{P}(A_n)$$

כאשר המעבר האחרון נובע מהתכנסות הגבול והגבול התחתון. □

35 תרגול 12 — 23.1.2025

35.1 צפיפות משותפת

תרגיל 35.1 נניח ש־$X, Y \sim \text{Exp}(1)$ בלתי־תלויים (כלומר מתקיים $f_X \cdot f_Y = f_{X,Y}$ לכל פונקציות צפיפות שנבחר).
חשבו את ההתפלגות והצפיפות של המשתנה המקרי $X - Y$.

פתרון נוכל לפתור את השאלה על־ידי שימוש בקונבולוציה, אך ננסה לפתור ישירות. אנו רוצים לחשב את $F_{X-Y}(t)$ עבור $t \in \mathbb{R}$, לפי ההגדרה נובע

$$\begin{aligned} F_{X-Y}(t) &= \iint_{\{X-Y<t\}} f_{X,Y}(x,y)\,dx\,dy \\ &= \iint_{\{X-Y<t\}} f_X(x) f_Y(y)\,dx\,dy \\ &= \iint_{\{X-Y<t\}} e^{-x}e^{-y}\,dx\,dy \\ &= \int_{-\infty}^{\infty} e^{-x} \int_{-x-t}^{\infty} e^{-y}\,dy\,dx \\ &= \int_{-\infty}^{\infty} e^{-y} \int_{-\infty}^{y+t} e^{-x}\,dx\,dy \end{aligned}$$

נבחין כי ראינו פה שתי דרכים שונות להשתמש במשפט פוביני על התחום הנתון. נוכל להשתמש בכל אחד מהביטויים לפי נוחות, במקרה הזה לפי חיוביות ושליליות של הביטויים $y + t$ ו־$-x - t$. כאשר $t < 0$ נחשב,

$$\int_{-\infty}^{\infty} e^{-x} \int_{-x-t}^{\infty} e^{-y}\,dy\,dx = \int_{-\infty}^{\infty} e^{-x}(e^{-x+t})\,dx = \int_{-\infty}^{\infty} e^{-2x+t}\,dx = -\frac{1}{2}e^{-2x+t}\Big|_{x=0}^{x=\infty} = \frac{1}{2}e^t$$

כאשר $t > 0$ נקבל,

$$\int_{-\infty}^{\infty} e^{-y} \int_{-\infty}^{y+t} e^{-x}\,dx\,dy = \int_{0}^{\infty} e^{-y} \int_{0}^{y+t} e^{-x}\,dx\,dy = \int_{0}^{\infty} e^{-y}(1 - e^{-y-t})\,dy = \frac{1}{2}e^{-2y-t} - e^{-y}\Big|_{y=0}^{y=\infty} = 1 - \frac{1}{2}e^{-t}$$

ולכן נובע

$$F_{X-Y}(t) = \begin{cases} 1 - \frac{1}{2}e^{-t} & t \ge 0 \\ \frac{1}{2}e^t & t < 0 \end{cases}$$

ובהתאם לזה נגזור ונקבל גם את פונקציית הצפיפות,

$$f_{X-Y}(t) = F'_{X-Y}(t) = \begin{cases} \frac{1}{2}e^{-t} & t \ge 0 \\ \frac{1}{2}e^t & t < 0 \end{cases}$$

תרגיל 35.2 יהי $i \in [n]$ ונגדיר $X_i \sim \text{Unif}([0,1])$ ונגדיר $Y_i = X_i \cdot X_{i+1}$, כאשר $Y = \sum Y_i$, חשבו את $\mathbb{E}(Y)$ ואת $\text{Var}(Y)$.

פתרון נשתמש בלינאריות התוחלת,

$$\mathbb{E}(Y_i) = \mathbb{E}(X_i) \cdot \mathbb{E}(X_{i+1}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

ובהתאם

$$\mathbb{E}(Y) = \frac{n-1}{4}$$

נעבור לשונות,

$$\text{Var}(Y) = \text{Cov}(Y,Y) = \sum \text{Var}(Y_i) + 2\sum_{i<j} \text{Cov}(Y_i, Y_j)$$

אנו גם יודעים ש־$\mathbb{E}(Y_i) = \mathbb{E}(Y_i^2) - \mathbb{E}(Y_i)^2$, וכן

$$\mathbb{E}(Y_i^2) = \iint_{\mathbb{R}^2} x^2y^2\,dx\,dy = \int_0^1 \int_0^1 x^2y^2\,dx\,dy = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9} \implies \text{Var}(Y_i) = \frac{7}{144}$$

ונותר לחשב את $\mathrm{Cov}(Y_i, Y_j)$. אם $|i-j| \geq 2$ אז $\mathrm{Cov}(Y_i, Y_j) = 0$. נחשב גם את

$$\mathrm{Cov}(Y_i, Y_{i+1}) = \mathbb{E}(Y_i Y_{i+1}) - \mathbb{E}(Y_i)\mathbb{E}(Y_{i+1})$$

וכן

$$\mathbb{E}(Y_i Y_{i+1}) = \mathbb{E}(X_i X_{i+1}^2 X_{i+2}) = \iiint_{[0,1]^3} xy^2 z \, dx \, dy \, dz = \left(\int_0^1 x \, dx\right)\left(\int_0^1 y^2 \, dy\right)\left(\int_0^1 z \, dz\right) = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{12}$$

נסיק אם כך ש-$\mathrm{Cov}(Y_i, Y_{i+1}) = \frac{1}{12} - \frac{1}{16} = \frac{1}{48}$. קיבלנו ש-$\mathrm{Var}(Y) = (n-1)\frac{7}{144} + 2(n-2)\frac{1}{48}$.

36 שיעור 24 — 23.1.2025

36.1 סדרות מאורעות — המשך

דוגמה 36.1 נניח ש־$\mathbb{P}(A_n)=\frac{1}{2}$ בלתי־תלויים. אז

$$\mathbb{P}(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k)=\lim_{n\to\infty}\mathbb{P}(\bigcap_{k=n}^{\infty}A_k)=0\le\liminf_{n\to\infty}\mathbb{P}(A_n)=\frac{1}{2}$$

כלומר הלמה של פאטו לא מועילה לנו במיוחד במקרה זה. באופן דומה

$$\mathbb{P}(\limsup_{n\to\infty}A_n)=1\ge\limsup_{n\to\infty}\mathbb{P}(A_n)=\frac{1}{2}$$

נניח שכל A_n זהים, $\forall n, A_n=A_1$.

למה 36.1 (הלמה הראשונה של בורל־קנטלי) *אם $\{A_n\}$ סדרת מאורעות ומתקיים $\sum_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$ אז $\mathbb{P}(\limsup_{n\to\infty}A_n)=0$.*

נבחן לא הוכחה רשמית, אבל את הרעיון שלה, נסתכל על $X=\sum_{n=1}^{\infty}1_{A_n}$, א

$$\mathbb{E}(X)=\sum_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$$

וכן

$$\mathbb{P}(\limsup_{n\to\infty}A_n)=\mathbb{P}(X=\infty)\le\frac{\mathbb{E}(X)}{\infty}=0$$

ועתה נעבור להוכחה אמיתית.

הוכחה.

$$\mathbb{P}(\limsup_{n\to\infty}A_n)=\mathbb{P}(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k)=\lim_{n\to\infty}\mathbb{P}(\bigcup_{k=n}^{\infty}A_k)\le\lim_{n\to\infty}\sum_{k=n}^{\infty}\mathbb{P}(A_k)=0$$

כאשר אי־השוויון נובע מחסם האיחוד, והשוויון האחרון נובע מהנתון $\sum_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$. □

נרצה למצוא גרסה הפוכה ללמה, אך ניסיון להפוך את התנאים יוביל לטענה שאיננה נכונה, במקום זאת נראה את הלמה הבאה.

למה 36.2 (הלמה השנייה של בורל־קנטלי) *אם A_n סדרת מאורעות בלתי־תלויים ו־$\sum_{n=1}^{\infty}\mathbb{P}(A_n)=\infty$ אז $\mathbb{P}(\limsup_{n\to\infty}A_n)=1$.*

הוכחה. אנו רוצים לחסום מלמטה ולכן נשתמש במשלים, נרצה למצוא את איפוס המשלים, לכן,

$$\mathbb{P}(\liminf_{n\to\infty}A_n^C)=\mathbb{P}(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k^C)=\lim_{n\to\infty}\mathbb{P}(\bigcap_{k=n}^{\infty}A_k^C)$$

אבל לכל $N\in\mathbb{N}$,

$$\mathbb{P}(\bigcap_{k=n}^{\infty}A_k^C)\ge\mathbb{P}(\bigcap_{k=n}^{N}A_k^C)=\prod_{k=n}^{N}(1-\mathbb{P}(A_k))\le\prod_{k=n}^{N}e^{-\mathbb{P}(A_k)}=e^{-\sum_{k=n}^{N}\mathbb{P}(A_k)}\xrightarrow{N\to\infty}0$$

□

דוגמה 36.2 אם A_n בלתי־תלויים ו־$\mathbb{P}(A_n)=p_n$ אז

$$\mathbb{P}(\limsup_{n\to\infty}A_n)=0\iff\sum_{n=1}^{\infty}p_n<\infty$$

וכן

$$\mathbb{P}(\limsup_{n\to\infty}A_n)=1\iff\sum_{n=1}^{\infty}p_n=\infty$$

נגדיר $B_n=A_n\cap A_{n+1}$, אם $\sum_{n=1}^{\infty}p_np_{n+1}<\infty$ אז $\mathbb{P}(\limsup_{n\to\infty}B_n)=0$. אם $\sum_{n=1}^{\infty}p_np_{n+1}=\infty$ אז $\mathbb{P}(\limsup_{n\to\infty}B_n)=$

1. אם $\{B_{2n}\}_{n\in\mathbb{N}}$ בלתי־תלויים ו־$\sum_{n\in2\mathbb{N}}p_np_{n+1}=\infty$ או $\sum_{n\in2\mathbb{N}+1}p_np_{n+1}=\infty$,
אז $\mathbb{P}(\limsup_{n\to\infty}B_{2n})=1$ או $\mathbb{P}(\limsup_{n\to\infty}B_{2n+1})=1$.

דוגמה 36.3 נגדיר $X_n \sim \mathrm{Unif}(\{-1,1\})$ וכן $Y_n = \sum_{k=1}^n X_k$. מצאנו בתרגיל שמתקיים $\mathbb{P}(Y_{3n}=0) \le e^{-cn}$ עבור איזשהו $c>0$. נסיק אם כך,

$$\sum_{n=1}^{\infty} \mathbb{P}(Y_n = 0) < \infty$$

ומהלמה הראשונה נובע $\mathbb{P}(\limsup_{n\to\infty}\{Y_n=0\}) = 0$. זה נכון אם ורק אם $\mathbb{P}(\exists n \forall k \ge n Y_k \ne 0) = 1$.

36.2 התכנסות של סדרות משתנים מקריים

בהגדרה 33.6 דיברנו על התכנסות של משתנים מקריים בהקשר ההתפלגות שלהם, עתה נראה התכנסות במובן חזק יותר.

הגדרה 36.3 (התכנסות סדרת משתנים מקריים כמעט תמיד) נאמר ש־$X_n \xrightarrow{a.s.} X$ אם

$$\mathbb{P}(\{\omega \in \Omega \mid X_n(\omega) \to X(\omega)\}) = 1$$

נבחן דוגמה למקרה זה.

דוגמה 36.4 נבחר $X_n = 1_{A_n}$ עבור A_n בלתי־תלויים. נגדיר גם $\mathbb{P}(A_n) = p_n$, כך ש־$p_n \le \frac{1}{2}$ לכל $n \in \mathbb{N}$. באילו תנאים $X_m \xrightarrow{a.s.} X$ ולאיזה X? מתי בסדרה $1_{A_n}(\omega)$ יש 0 החל ממקום מסוים? זהו למעשה $(\limsup_{n\to\infty} A_n)^C$. אם $\sum_{n=1}^{\infty} p_n < \infty$ אז $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$ ולכן $X_m \xrightarrow{a.s.} 0$. אם $\sum p_n = \infty$ אז $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$ ואז X_n לא מתכנסת כמעט תמיד.

משפט 36.4 **(החוק החזק של המספרים הגדולים)** *אם X_n משתנים מקריים בלתי־תלויים שווי־התפלגות עם $\mathbb{E}(X_n) = \mu$ אז*

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

הוכחה. (תחת ההנחה $|X_n| \le M$ כמעט תמיד) יהי $\epsilon > 0$, נרצה לבדוק את ההסתברויות,

$$\mathbb{P}(\frac{\sum_{i=1}^n X_i}{n} \ge \mu + \epsilon), \qquad \mathbb{P}(\frac{\sum_{i=1}^n X_i}{n} \le \mu - \epsilon)$$

אבל קיבלנו ששני אלה חסומים על־ידי e^{-cn}. על־ידי שימוש בטענה זו,

$$\{\frac{\sum_{i=1}^n X_i}{n} \to \mu\} = \{\forall k \in \mathbb{N}, \exists N, \forall n \ge N, \left|\frac{\sum_{i=1}^n X_i}{n} - \mu\right| < \frac{1}{k}\}$$

ואנו רוצים להראות שלכל k שנבחר מתקיים $\mathbb{P}(\exists N \forall n \ge N \left|\frac{\sum_{i=1}^n X_i}{n} - \mu\right| < \frac{1}{k}) = 1$. באופן שקול

$$\mathbb{P}(\bigcup_{N=1}^{\infty}\bigcap_{n=N}^{\infty} A_k^n) = 1 \iff \mathbb{P}(\limsup_{n\to\infty}(A_k^n)^C) \overset{(1)}{=} 0$$

כאשר (1) נובע מהלמה הראשונה. $\square$

מסקנה 36.5 אם X_n בלתי־תלויים שווי־התפלגות עם תוחלת $\mathbb{E}(X_n) = \mu \ne 0$ אז $Y_n = \sum_{i=1}^n X_i$ הוא הילוך חולף. זאת שכן מהחוק החזק $\frac{Y_n}{n} \to \mu \ne 0$.

37 תרגול 13 — 30.1.2025

37.1 המשפט הגדול המרכזי

נדבר על התפלגות נורמלית סטנדרטית, 30.13, ונתחיל בסימון,

סימון 37.1 אם $X \sim \mathrm{N}(0,1)$, אז נסמן $\Phi(t) = F_X(t)$.

נבהיר שמשפט הגבול המרכזי, 33.2, הוא משפט מרכזי שמסביר למה התפלגות נורמלית מופיעה בכל כך הרבה מקומות. המשפט למעשה גורס שתקנון של התפלגויות ישאף להתנהג כמו התפלגות נורמלית. נעבור לדוגמה,

תרגיל 37.1 מטילים n קוביות הוגנות באופן בלתי־תלוי. העריכו את ההסתברות שסכום הקוביות נמצא בטווח $I = [\frac{7}{2}n - \sqrt{n}, \frac{7}{2} + \sqrt{n}]$.

פתרון נגדיר $X_i \sim \mathrm{U}([6])$ וכן $X = \sum X$, אז

$$\mathbb{E}(X_i) = \frac{7}{2}, \quad \sigma^2 = \mathrm{Var}(X_i) = \frac{35}{12}$$

נגדיר $Y_i = \frac{X_i - \frac{7}{2}}{\sigma}$. Y_i עומדים בתנאי משפט הגבול המרכזי ולכן אם $n \to \infty$ אז

$$\frac{\sum Y_i}{\sqrt{n}} \xrightarrow{d} \mathrm{N}(0,1)$$

אנו רוצים להעריך את $\mathbb{P}(X \in I) = \mathbb{P}(-\frac{\sqrt{n}}{\sigma} \le \sum Y_i \le \frac{\sqrt{n}}{\sigma})$. ממשפט הגבול המרכזי נוכל להסיק,

$$\mathbb{P}(X \in I) = \mathbb{P}(\frac{-1}{\sigma} \le \frac{\sum Y_i}{\sqrt{n}} \le \frac{1}{\sigma}) \approx \Phi(\frac{1}{\sigma}) - \Phi(\frac{-1}{\sigma}) \approx 0.44$$

נעבור לדוגמה שהופיעה במבחן של אמיר,

תרגיל 37.2 נניח ש־$X \sim \mathrm{Poi}(10000)$, תנו הערכה ל־$\mathbb{P}(X \le 9800)$.

רמז: תנו הערכתכם בצורת אינטגרל.

פתרון נגדיר $X_i \sim \mathrm{Poi}(1)$ אז אנו יודעים ש־$X \overset{d}{=} \sum_{i=1}^{10000} X_i$. אנו יודעים גם ש־$\mathbb{E}(X_i) = 1$ וכן ש־$\mathrm{Var}(X_i) = 1$ ולכן נגדיר $Y_i = X_i - 1$. מ־33.2 וההנחה ש־$n = 10000$ מספיק גדול נסיק,

$$\frac{\sum Y_i}{\sqrt{10000}} \overset{d}{\approx} \mathrm{N}(0,1)$$

לכן

$$\mathbb{P}(X \le 9800) = \mathbb{P}(\sum X_i \le 9800) = \mathbb{P}(\frac{\sum Y_i}{100} \le -2) = \Phi(-2) = \frac{1}{\sqrt{2\pi}} - \int_{-\infty}^{-2} \exp(\frac{-t^2}{2})\, dt \approx 0.0228$$

37.2 התכנסויות

נדבר על התכנסות 36.2, התכנסות שמאוד מזכירה את ההתכנסות של אינפי. יש סוג שני של התכנסות,

$$X_n \xrightarrow{P} X \implies \forall \epsilon > 0, \lim_{n\to\infty} \mathbb{P}(|X_n - X| < \epsilon) = 1$$

לעומתה יש את ההתכנסות 33.6, היא מדברת על התכנסות נקודתית באופן נאיבי. כל התכנסות גוררת את הבאות אחריה, אך לא הפוך. נשתכנע בזה עכשיו.

נסכם ונגיד שהתכנסות כמעט תמיד גוררת התכנסות בהסתברות גוררת התכנסות בהתפלגות.

דוגמה 37.1 נניח ש־$X_n \sim \mathrm{Ber}(\frac{1}{2})$. בוודאי שסדרה זו מתכנסת בהתפלגות ל־$\mathrm{Ber}(\frac{1}{2})$.

נראה שהיא לא מתכנסת בהסתברות (הגרסה השנייה),

$$\mathbb{P}(|X_n - X_m| \ge \epsilon) \le \mathbb{P}(|X_n - X_m| \ge \frac{\epsilon}{2}) + \mathbb{P}(|X_m - X| \ge \frac{\epsilon}{2})$$

לכל ϵ קטן כרצוננו. מהצד השני עבור $\epsilon < 1$,

$$\mathbb{P}(|X_n - X_m| \ge \epsilon) = \mathbb{P}(X_n = 0, X_m = 1) + \mathbb{P}(X_n = 1, X_m = 0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

על־ידי שימוש באי־תלות, וזו כמובן סתירה, בשורה הראשונה ראינו שההסתברות קטנה כרצוננו, ואז הצלחנו לחסום אותה על־ידי $\frac{1}{2}$.

דוגמה 37.2

$$\begin{aligned}\{\lim_{n\to\infty} X_n = X\} &= \{\omega\in\Omega \mid \forall\epsilon\exists N\forall n>N, |X_n(\omega)-X(\omega)|<\epsilon\} \\ &= \bigcap_{\epsilon>0}\bigcup_N A_{N,\epsilon} = \{\omega\in\Omega \mid \forall n>N, |X_n(\omega)-X(\omega)|<\epsilon\} \\ &= \liminf_{n\to\infty} A_{N,\epsilon}\end{aligned}$$

ובהתאם

$$\mathbb{P}(\liminf_{n\to\infty} A_{N,\epsilon}) = 1 \iff X_n \xrightarrow{a.s.} X$$

וכן,

$$\liminf_{n\to\infty} \mathbb{P}(A_{N,\epsilon}) = 1 \iff X_i \xrightarrow{P} X$$

זוהי הלמה של פאטו.

דוגמה 37.3 נרצה למצוא (X_n) מתכנסת בהסתברות ולא מתכנסת כמעט תמיד.

נגדיר X_n בלתי־תלויים עם התפלגות,

$$\mathbb{P}(X_n = n) = \frac{1}{n}, \qquad \mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$$

ולכן עבור $\epsilon > 0$,

$$\mathbb{P}(X_n > \epsilon) \to \frac{1}{n} \to 0$$

ולכן $X_n \xrightarrow{D} 0$, כלומר הסדרה מתכנסת בהסתברות למשתנה המקרי הקבוע 0.

נגדיר $B_{n,\epsilon} = \{X_n > \epsilon\}$ ולכן $\mathbb{P}(B_{n,\epsilon}) = \frac{1}{n}$ ובהתאם $\sum_{n=1}^{\infty} \mathbb{P}(B_{n,\epsilon}) = \infty$. מהלמה השנייה של בורל־קנטלי נסיק,

$$\mathbb{P}(\limsup_{n\to\infty} B_{n,\epsilon}) = 1$$

ולכן גם

$$\mathbb{P}(\limsup_{n\to\infty} B_{n,\epsilon}^C) = 0$$

ונובע ש־X_n לא מתכנס כמעט תמיד למשתנה מקרי קבוע 0. לכן גם לא מתכנס כמעט תמיד.

38 שיעור 25 — 30.1.2025

38.1 התכנסויות של סדרות משתנים מקריים

ניזכר שראינו שלושה סוגים של התכנסויות, התכנסות כמעט תמיד, 36.2, את הסוג השני נגדיר עתה,

הגדרה 38.1 (התכנסות בהסתברות) תהי סדרת משתנים מקריים X_n, אז נאמר ש־$X_n \xrightarrow{P} X$, או שסדרת המשתנים המקריים מתכנסת בהסתברות ל־X, אם,

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

הסוג השלישי הוא התכנסות בהתפלגות, 33.6.

דוגמה 38.1 נניח ש־$(X_n)_{n=1}^\infty$ בלתי־תלויים בעלי תוחלת μ ושונות σ^2.
מהחוק החזק של המספרים הגדולים נובע $\mu \xrightarrow{a.s.} \frac{\sum i=1^n X_i}{n}$.
החוק החלש גורר ש־$\mu \xrightarrow{P} \frac{\sum_{i=1}^n X_i}{n}$.
משפט הגבול המרכזי גורר ש־$Z \xrightarrow{d} \frac{(\sum_{i=1}^n X_i)-n\mu}{\sqrt{n}\sigma}$, כאשר $Z \sim \mathrm{N}(0,1)$.

משפט 38.2 (יחס סוגי ההתכנסות) *אם $X_n \xrightarrow{a.s.} X$ אז $X_n \xrightarrow{P} X$, אם $X_n \xrightarrow{P} X$ אז $X_n \xrightarrow{d} X$, כאשר הכיוון ההפוך איננו נכון.*
אם $X_n \xrightarrow{d} C$ עבור משתנה מקרי קבוע C, אז $X_n \xrightarrow{P} C$ (אבל לא מעבר).

הוכחה גרירה של כמעט תמיד להסתברות. יהי $\epsilon > 0$. נבחין כי $\lim_{n\to\infty} X_n = X$ אז קיים $k \in \mathbb{N}$ כך שלכל $n > k$, $|X_n - X| < \epsilon$. לכן

$$\{\lim_{n\to\infty} X_n = X\} \subseteq \bigcup_{k=1}^\infty \bigcap_{n=k}^\infty \{|X_n - X| < \epsilon\} = \liminf_{n\to\infty}\{|X_n - X| \le \epsilon\}$$

ממונוטוניות נובע,

$$1 \overset{(1)}{=} \mathbb{P}(\{\lim_{n\to\infty} X_n = X\}) \le \mathbb{P}(\liminf_{n\to\infty}\{|X_n - X| \le \epsilon\}) \overset{(2)}{=} \liminf_{n\to\infty} \mathbb{P}(|X_n - X| \le \epsilon)$$

כאשר

1. $X_n \xrightarrow{a.s.} X$
2. הלמה של פאטו

ולכן $\lim_{n\to\infty} \mathbb{P}(|X_n - X| \ge \epsilon) = 0$. □

דוגמה 38.2 נראה שאכן לא בהכרח $X \xrightarrow{P} X_n$ מעיד ש־$X \xrightarrow{a.s.} X_n$. גם כאן וגם בדוגמה הטובה יותר מהתרגול X הוא קבוע. תהי $(Y_n)_{n=1}^\infty$ בלתי־תלויים כך ש־$Y_n \sim \mathrm{Geo}(\frac{1}{2})$, ונסמן $X_n = \frac{Y_n}{\log_2 n}$, ונגדיר $X = 0$.
נראה ש־$X \xrightarrow{P} X_n$, יהי $\epsilon > 0$,

$$\mathbb{P}(|X_n - X| > \epsilon) = \mathbb{P}(Y_n > \epsilon \log_2 n) = \left(\frac{1}{2}\right)^{\lfloor \epsilon \log_2 n\rfloor} \xrightarrow{n\to\infty} 0$$

כאשר המהלך האחרון נובע מהתפלגות גאומטרית ישירות.
נעבור להראות שלא מתקיים $X \xrightarrow{a.s.} X_n$. נטען ש־$X_n > 1$ אינסוף פעמים בהסתרות 1,

$$\mathbb{P}(X_n > 1) = \frac{1}{2^{\lfloor \log_2 n\rfloor}} \ge \frac{1}{2^{\log_2 n}} = \frac{1}{n}$$

ולכן

$$\sum_{n=1}^\infty \mathbb{P}(X_n > 1) = \infty$$

ומהלמה השנייה של בורל־קנטלי נובע $\mathbb{P}(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty X_n > 1) = 1$, לכן $\mathbb{P}(\lim_{n\to\infty} X_n = X) = 0 \ne 1$.

נעבור לחלק הבא של הוכחת המשפט.

הוכחת הסתברות גורר בהתפלגות. תהי t נקודת רציפות של F_X, לכל $n \in \mathbb{N}$ ולכל $\epsilon > 0$ נרצה להראות שמתקיים

$$F_X(t-\epsilon) - \mathbb{P}(|X_n - X| > \epsilon) \le F_X(t) \le F_X(t+\epsilon) + \mathbb{P}(|X_n - X| > \epsilon)$$

עבור החלק הראשון של אי־השוויון,

$$\begin{aligned}F_X(t) &= \mathbb{P}(X_n \le t)\\ &= \mathbb{P}(X_n \le t, X \le t+\epsilon) + \mathbb{P}(X_n \le t, X > t+\epsilon)\\ &\le \mathbb{P}(X \le t+\epsilon) + \mathbb{P}(|X_n - X| > \epsilon)\\ &= F_X(t+\epsilon) + \mathbb{P}(|X_n - X| > \epsilon)\end{aligned}$$

ממונוטוניות והסתברות שלמה. עבור החלק השני של אי־השוויון,

$$\begin{aligned}F_X(t-\epsilon) &= \mathbb{P}(X \le t-\epsilon)\\ &= \mathbb{P}(X \le t-\epsilon, X_n \le t) + \mathbb{P}(X \le t-\epsilon, X_n > t)\\ &\le \mathbb{P}(X_n \le t) + \mathbb{P}(|X_n - X| > \epsilon)\\ &= F_X(t) - \mathbb{P}(|X_n - X| > \epsilon)\end{aligned}$$

ולכן נובע ש־$F_X(t) \ge F_X(t-\epsilon) - \mathbb{P}(|X_n - X| > \epsilon)$. מצאנו אי־שוויונות החוסמים את $F_X(t)$ ולכן כאשר $n \to \infty$ נובע

$$F_X(t-\epsilon) \le \liminf_{n\to\infty} F_{X_n}(t) \le \limsup_{n\to\infty} F_{X_n}(t) \le F_X(t+\epsilon)$$

וכאשר $\epsilon \to 0$ ובשימוש ברציפות של t ב־F_X נובע ש־$\lim_{n\to\infty} F_{X_n}(t) = F_X(t)$. □

נראה עתה דוגמה למקרה שיש התכנסות בהתפלגות אבל לא בהסתברות. הסיבה הטכנית היא שייתכן ש־X_n ו־X לא מוגדרים על אותו מרחב הסתברות.

דוגמה 38.3 נניח ש־X_n, X כולם בלתי־תלויים ומתפלגים $\mathrm{Unif}([0,1])$, אז $X_n \xrightarrow{d} X$ שכן $F_{X_n}(t) = F_X(t)$ לכל t, אבל $\mathbb{P}(|X_n - X| > \frac{1}{4}) \ge \mathbb{P}(X_n \le \frac{1}{2}, X \ge \frac{3}{4}) = \frac{1}{8} \not\to 0$.

טענה 38.3 *אם $X = C$ משתנה מקרי קבוע ו־$X_n \xrightarrow{d} C$ אז $X_n \xrightarrow{P} C$.*

הוכחה. יהי $\epsilon > 0$,

$$\mathbb{P}(|X_n - X| \le \epsilon) = \mathbb{P}(X_n \le C+\epsilon) - \mathbb{P}(X_n < C-\epsilon) \ge \mathbb{P}(X_n \le C+\epsilon) - \mathbb{P}(X_n \le C-\epsilon) = F_{X_n}(C+\epsilon) - F_{X_n}(C-\epsilon)$$

אבל $F_X(t) = F_C(t) = \begin{cases} 0 & t < C \\ 1 & t \ge C\end{cases}$ לכן שגם $C+\epsilon$ וגם $C-\epsilon$ נקודות רציפות של F_X, ולכן

$$\mathbb{P}(|X_n - X| \le \epsilon) \ge F_{X_n}(C+\epsilon) - F_{X_n}(C-\epsilon) \xrightarrow{n\to\infty} F_X(C+\epsilon) - F_X(C-\epsilon) = 1 - 0$$

כפי שרצינו. □

38.2 תכונות של התכנסות

טענה 38.4 *הטענות הבאות נכונות עבור התכנסות כמעט תמיד,*

1. *אם $X_n \xrightarrow{a.s.} X$ וכן $Y_n \xrightarrow{a.s.} Y$ אז $X_n + Y_n \xrightarrow{a.s.} X + Y$*
2. *אם $X_n \xrightarrow{a.s.} X$ ו־$f \in C(\mathbb{R})$ אז $f(X_n) \xrightarrow{a.s.} f(X)$.*

הוכחה. 1. מתקיים

$$\{\lim_{n\to\infty} X_n + Y_n = X + Y\} \supseteq \{\lim_{n\to\infty} X_n = X\} \cap \{\lim_{n\to\infty} Y_n = Y\}$$

צריך להוכיח שמתקיים $\mathbb{P}(\lim_{n\to\infty} X_n = X, \lim_{n\to\infty} Y_n = Y) = 1$, נראה שמתקיים

$$\begin{aligned}\mathbb{P}(\{\lim_{n\to\infty} X_n = X\} \cap \{\lim_{n\to\infty} Y_n = Y\}^C) &= \mathbb{P}(\{\lim_{n\to\infty} X_n = X\}^C \cup \{\lim_{n\to\infty} Y_n = Y\}^C)\\ &\le \mathbb{P}(\{\lim_{n\to\infty} X_n = X\}^C) + \mathbb{P}(\{\lim_{n\to\infty} Y_n = Y\}^C)\\ &= 0 + 0\end{aligned}$$

2. $\{\lim_{n\to\infty} X_n = X\} \subseteq \{\lim_{n\to\infty} f(X_n) = f(X)\}$ ולכן

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

וסיימנו.

□

דוגמה 38.4 נניח ש־$(X_n)_{n=1}^{\infty}$ בלתי־תלויים ומתפלגים $\mathrm{Unif}([0,1])$, נראה שמתקיים

$$(X_1 \cdots X_n)^{\frac{1}{n}} \xrightarrow{a.s.} \frac{1}{e}$$

נבחין כי $(X_1 \cdots X_n)^{\frac{1}{n}} = (e^{\log X_1} \cdots e^{\log X_n})^{\frac{1}{n}} = \exp(\frac{1}{n}\sum_{i=1}^n \log X_i)$, אם מתקיים $\frac{1}{n}\sum_{i=1}^n \log X_i \xrightarrow{a.s.} -1$ אז מהטענה שהוכחנו עתה הטענה נובעת. נוכיח ש־$-\log X_i \sim \mathrm{Exp}(1)$,

$$\mathbb{P}(-\log X_i > t) = \mathbb{P}(X_i < e^{-t}) = e^{-t}$$

זאת שכן $X_i \sim \mathrm{U}([0,1])$ ולכן,

$$\mathbb{E}(\log X_i) = -1$$

ומהחוק החזק נובע

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

39 סיכום תוצאות

39.1 התפלגויות בדידות

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

39.2 התפלגויות רציפות

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

הגדרות ומשפטים