הסתברות (1) ־ ד"ר אורי פרזנצ'בסקי

מסכם: יחיאל מרצבך

תוכן העניינים

הקדמה

בשפה היומיומית, אנו אומרים למשל "הטלת מטבע (הוגן) נותנת עץ או פלי, בסיכוי $\frac{1}{2}$". או למשל "חתול נולד לבקן בסיכוי $\frac{1}{1000}$".

שיעור מס' 1:
יום שלישי
20.10.20

מה הכוונה שלנו באמירה הזאת? מבחינה פילוסופית, ניתן לומר שכוונת דברינו כי "ישנה אי וודאות ועדיין לא הוחלט", או מבחינה אחרת ניתן לומר כי הסתכלות על נתוני העבר מוכיחה על העתיד. גישות אחרות הן דטרמיניסטיות יותר, ואין כלל מושג של הסתברות.
בקורס נפתח את התורה המתמטית של הסתברות.
בשונה מאינפי וליניארית, נסתמך על אינטואיציה מסוימת, כיוון שיש לקורס שימושיות גדולה. עלינו לציין כי רק במאה ה־20 המתמטיקה הצליחה לנסח בצורה פורמלית את תורת ההסתברות, כיוון שקיימות נקודות עדינות מבחינת ריגורזיות ופורמליות.

מושגים בסיסיים

ניסוי - הפעולה שנותנת תוצאה "מקרית" (הטלת מטבע, לידת חתול וכו', בדוגמאות דלעיל).
מרחב המדגם - התוצאות האפשריות של הניסוי.
מבחינה פורמלית, **לניסוי** אין הגדרה, ו**מרחב מדגם** הוא קבוצה לא ריקה. לרוב נסמן את **מרחב המדגם** ב־Ω.

הגדרה

פונקצית הסתברות נקודתית (על מרחב מדגם Ω) היא פונקציה
$p:\Omega\to[0,1]$ כך ש־$\sum\limits_{\omega\in\Omega}p(\omega)=1$ (סכום כל ההסתברויות משלים ל־1).

דוגמאות

- $\Omega=\{H,T\}$ ו־$p(H)=p(T)=\frac{1}{2}$. זאת למעשה ההסתברות של מטבע הוגן.
- גם $\Omega=\{H,T,☺\}$:
  עם $p(H)=p(T)=\frac{1}{2}$ ו־$p(☺)=0$, זאת הסתברות של מטבע הוגן.
- $\Omega=[6]=\{1,2\,\ldots,6\}$, כך ש־$\forall j\in\Omega\, p(j)=\frac{1}{6}$ - "קוביה הוגנת".
- כיצד נוכל לתאר הטלת מטבע, עד שיצא H, כך ש**תוצאת הניסוי** היא מס' ההטלות שביצענו? התשובה הינה למעשה $\Omega=\mathbb{N}\cup\{\infty\}$ (איחוד של הטבעיים עם מקרה בו לא קיבלנו את H בכלל):

תוצאה	הטלה 2	הטלה 1
1	H	H
1	T	H
2	H	T
?	T	T

- אם כך, נקבל כי $p(1)=\frac{1}{2},p(2)=\frac{1}{4}$. בסך הכולל, הסיכוי הינו רבע (ישנם ארבע אפשרויות, ולכן זה אחת מהאפשרויות)
  או אם נרצה, ההסתברות של $p(m)=\frac{1}{2^m}$, וההסתברות של $p(\infty)=0$, וכמו כן $\sum\limits_{j=1}^{\infty}i^j=1$.
- נתבונן במרחב המדגם - $\Omega=\{\{1,1\},\{1,2\},\ldots(6,6)\}$ - לכאורה $p(\{1,j\})=\frac{1}{21}$, אך למעשה הסיכוי שיצא $\{1,1\}$ קטן יותר! לכן, למעשה, מדויק יותר לכאורה לומר - $p(\{i,j\})=\begin{cases}\frac{1}{36} & i=j\\ \frac{1}{18} & i\neq j\end{cases}$.
  שתי האפשרויות למעשה נכונות, אלא שהמודל הראשון נקרא **"קוביות שונות"**, ולשני קוראים **"קוביות זהות"**.
  המודל שבו p קבועה נקרא **"מודל של התפלגות אחידה"** והוא נוח לחישובים[1].

[1] אעדיף לצבוע קוביות בצבעים שונים, על מנת לקבל את המודל הראשון.

שיעור מס' 2:
יום חמישי
22.10.20

בעיית ההגרלה האחידה ב־$[0,1]$

כיצד ניתן להגריל מספר ממשי אקראי בקטע $[0,1]$?

ברור כי מרחב המדגם הינו $\Omega=[0,1]$.

באופן עקרוני, נרצה לומר כי בסיכוי $\frac{1}{2}$ נקבל מספר בין 0 ל־$\frac{1}{2}$, ובסיכוי $\frac{1}{4}$ נקבל מספר בין 0 ל־$\frac{1}{4}$.

בהינתן $p:\Omega\to\mathbb{R}_+$ ונדרוש כי $\sum\limits_{\alpha\in\Omega}p(\alpha)=1$ ו־$p(\alpha)=p(\beta)$ (כי אנחנו מחפשים אחידות), לכל $\alpha,\beta\in\Omega$ אין אפשרות ל־p כזו (שהרי עבור $p=0$ מתקבל כי $\sum p(\omega)=0$ ועבור $p=c>0$ מתקבל כי $\sum p(\omega)=\infty$). [2].

הפיתרון הפרקטי בסופו של דבר הוא להטיל קוביה הוגנת עם 10 פאות ולרשום את הספרות מימין לנקודה העשרונית: 0.27056.[3]

אמנם, ננסה לפרמל את המודל בכל זאת:

נשים לב כי לכל $\alpha\in[0,1]$ אכן מתקיים כי $p(\alpha)=0$ (כמו לקבל רק 'פלי' בהטלת מטבע), אבל קיבלנו הסתברות נכונה עבור קטעים.

כך למשל, $p\left(0\le\alpha\le\frac{1}{2}\right)$ אם"ם "הספרה הראשונה היא $0,1,2,3,4$ או שהספרה הראשונה היא חמש וכל השאר הן 0". (דבר זה קורה בסיכוי $5\cdot\frac{1}{10}+0$) .

---

[2] אם נרצה להגדיר סכום שאינו בן מנייה, נאמר: לכל קבוצה X ופונקציה $f:X\to\mathbb{R}_+$ נגדיר $\sum\limits_{x\in X}f(x)=\sup\limits_{A\subseteq X}\sum\limits_{x\in A}f(x)$.

[3] עבור מהנדס\ פיזיקאי הדבר מספיק ־ מטילים עד הדיוק שחפצים בו. עבור מתמטיקאי ־ הדבר לא מספיק, כי לא נוכל לענות על השאלה האם $\alpha\in\mathbb{Q}$.

מרחבי הסתברות - הגדרות ומושגים

הגדרות

הגדרה

מאורע הוא תת קבוצה של Ω.

אוסף כל המאורעות ב־Ω מסומן ב־$\mathcal{F}$.

דוגמאות

עבור $\Omega=[0,1]$: $\left[0,\frac{1}{2}\right]$, $\mathbb{Q}\cap[0,1]$, $C:=\left\{\sum_{i=1}^{\infty}\frac{a_i}{10^i}\mid \begin{array}{c}0\le a_i\le 9\\ a_i\neq 7\end{array}\right\}$.

עבור $\Omega=[6]$: $\{1,3,5\}$(תוצאה אי זוגית), $\{1,\dots,4\}$.

עבור $\omega\in\Omega$ גם $\{\omega\}$ היא מאורע.

הגדרה

פונקציית הסתברות (על מרחב מדגם Ω) היא פונקציה $\mathbb{P}:\mathcal{F}\to\mathbb{R}_+$ שמקיימת:

(1) $\mathbb{P}(\Omega)=1$.[4]

(2) אם $\{A_i\}_{i=1}^{\infty}$ סדרה של מאורעות זרים בזוגות (כלומר $\forall i\neq j\ A_i\cap A_j=\emptyset$) אזי $\mathbb{P}\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}\mathbb{P}(A_i)$.[5]

- לא נרצה לדרוש מעבר לכך (שהרי אם $\{A_i\}_{i\in I}$ אוסף כלשהוא של מאורעות זרים בזוגות, אזי $\mathbb{P}(\cup A_i)=\sum\mathbb{P}(A_i)$ כיוון ש־ $p(\{\omega\})=0$ לכל ω גורר כי $\mathbb{P}(\Omega)=0$).

פונקציית הסתברות **נקודתית** על Ω משרה פונקצית הסתברות על $\mathcal{F}$ של Ω:

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

לדוגמא, קוביה הוגנת:

$\Omega=\{H,T\}$ ו־ $P=\frac{1}{2}$ אזי $\mathcal{F}=\{\{\},\{H\},\{T\},\{H,T\}\}$ כך ש־ $\mathbb{P}(\{\})=0,\mathbb{P}(\{H\}),\mathbb{P}(\{T\})=$1 $\frac{1}{2},\mathbb{P}(\{H,T\})=$

מסקנות מההגדרה

(1) אם $A_1,\dots,A_k\in\mathcal{F}$ מאורעות זרים בזוגות, אזי $\mathbb{P}\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{k}\mathbb{P}(A_i)$.

(2) $\mathbb{P}(\emptyset)=0$.

(א) $A_i=\emptyset$ לכל $i\in\mathbb{N}$ נותן סדרה שהאיחוד שלה הוא הקבוצה הריקה.

(3) $\mathbb{P}(A^c)=1-\mathbb{P}(A)$ כאשר $A^c:\Omega\setminus A$.

(4) לכל $A\in\mathcal{F}$ אזי $\mathbb{P}(A)\le 1$.

(5) אם $A\subseteq B\in\mathcal{F}$ אזי $\mathbb{P}(A)\le\mathbb{P}(B)$.

טענה

אם A_i סדרה עולה של מאורעות (כלומר $A_i\subseteq A_j$ עבור $i\le j$) אזי $\mathbb{P}\left(\bigcup_{i=1}^{\infty}A_i\right)=\lim_{i\to\infty}\mathbb{P}(A_i)$.[6]

הוכחה

---

[4] אם $A,B\in\mathcal{F}$ זרות, אזי $P(A\cup B)=P(A)+P(B)$ (מכך נובע כי $P(\emptyset)=0$ $\Leftarrow$ $P(\emptyset)=P(\emptyset\cup\emptyset)=P(\emptyset)+P(\emptyset)\Rightarrow P(\emptyset)=0$)

[5] הטור הזה מתכנס כיוון שמדובר בטור אי שלילי חסום.

[6] יש שיקראו לזה 'רציפות פונקציית ההסתברות'.

נגדיר $B_1=A_1$ וגם $B_j=A_j\setminus A_{j-1}$ לכל $2\le j$. נשים לב כי $\bigcup\limits_{i=1}^{\infty}B_i=\bigcup\limits_{i=1}^{\infty}A_i$ ו־ $\{B_i\}$ זרים בזוגות. לכן :

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty}A_i\right)=\mathbb{P}\left(\bigcup_{i=1}^{\infty}B_i\right)=\sum_{i=1}^{\infty}\mathbb{P}\left(B_i\right)=\lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{P}\left(B_i\right)\overset{\text{תכונה א'}}{=}\lim_{n\to\infty}\mathbb{P}\left(\bigcup_{i=1}^{n}B_i\right)=\lim_{i\to\infty}\mathbb{P}\left(A_i\right)$$

שיעור מס' 3:
יום שלישי
27.10.20

הגדרה

מרחב הסתברות הוא שלשה $(\Omega,\mathcal{F},\mathbb{P})$ כש־Ω קבוצה לא ריקה ("מרחב המדגם"),$\mathcal{F}=\{A\mid A\subseteq\Omega\}$ [7].
הפונקציה $\mathbb{P}:\mathcal{F}\to\mathbb{R}_+$ הינה פונקצית הסתברות כפי שהגדרנו קודם לכן.

דוגמא

אם p פונקצית הסתברות נקודתית על Ω כלשהוא, אזי $(\Omega,\mathcal{F},P_p)$ מרחב הסתברות כש־$\mathcal{F}=2^{\Omega}$ ו־$\forall A\subseteq$ Ω $\mathbb{P}_p\left(A\right)=\sum\limits_{\omega\in A}p\left(\omega\right)$.

הגדרה

אינטרלוג ־ סכום של מספרים אי שליליים של קבוצה כלשהיא: אם $\{a_i\}_{i\in I}$ כך ש־$a_i\ge 0$ ו־I קבוצה כלשהיא, מגדירים:

$$\sum_{i\in I}a_i=\text{supp}\left\{\sum_{i\in J}a_i\mid\right\}(\in\mathbb{R}\cup\{\infty\})$$

(כש־$I=N$ מקבלים את הגדרת הסכום סכום טור מאינפי 2).

טענה

אם $|I|>\aleph_0$ ו־$a_i>0$ לכל $i\in I$ אזי $\sum\limits_{i\in I}a_i=\infty$.

הוכחה

נגדיר $X_n=\left\{i\in I\mid a_i>\frac{1}{n}\right\}$ אזי $\bigcup\limits_{n\in\mathbb{N}}X_n=I$ ולכן ישנו n_0 כך ש־$\aleph_0<|X_{n_0}|$ (איחוד בן מניה של קבוצות בנות מניה הוא עדיין "בן מנייה").

בפרט X_{n_0} אינסופי ולכן לכל $m\in\mathbb{N}$ יש $J\subseteq X_{n_0}$ בגודל m ולכן $\frac{m}{n_0}=|J|\frac{1}{n_0}\ge\sum\limits_{j\in J}a_i\ge\sum\limits_{i\in I}a_i$ הוא $\sum a_i=\infty$.

לעניינו, דבר זה גורר כי אם p היא פונקצית הסתברות נקודתית על Ω, אזי $\text{supp}\left(p\right)=\{\omega\in\Omega\mid p\left(\omega\right)\neq 0\}$ סופית או בת־מנייה [8].

הגדרה

מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$ נקרא מרחב הסתברות בדיד אם הוא מתקבל מפונקצית הסתברות נקודתית, כמו בדוגמא.

---

[7] יותר מאוחר נדרוש כי $\mathcal{F}$ תכלול רק חלק מתתי הקבוצות של Ω, מסיבות שונות.

[8] קבוצת כל האיברים ב־$\mathcal{F}$ שההסתברות שלהם גדולה מ־0.

מרחבי הסתברות בדידים

שיעור מס' 4:
יום חמישי
29.10.20

טענה

התנאים הבאים שקולים, כש־$(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות:

(1) $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות בדיד (יש p כך ש־$\mathbb{P}=\mathbb{P}_p$).

(2) יש $A\subseteq\Omega$ עם $\mathbb{P}(A)=1$ ו־$|A|\leq\aleph_0$.

(3) $\sum\limits_{\omega\in\Omega}\mathbb{P}(\{\omega\})=1$.

(4) לכל $B\in\mathcal{F}$ מתקיים כי $\mathbb{P}(B)=\sum\limits_{\omega\in\Omega}\mathbb{P}(\{\omega\})$.

בחצי הראשון של הקורס נעסוק במרחבי הסתברות בדידים.

הוכחה

$(4)\Leftrightarrow(1)$: טאוטולוגיה.

$(4)\Rightarrow(3)$: ברור ישירות מההגדרות, כי אם ניקח $B=\Omega$ נקבל כי $\sum\limits_{\omega\in\Omega}\mathbb{P}(\{\omega\})=\mathbb{P}(\Omega)=1$.

$(1)\Rightarrow(2)$: ניקח $A=\text{supp}(p)$. אזי נקבל מהטענה הקודמת כי $|A|\leq\aleph_0$, כי אחרת ראינו ש־$\sum\limits_{\omega\in A}p(\omega)=\infty$.

כעת, נקבל:

$$\mathbb{P}(A)\overset{(*)}{=}P\left(\bigcup_{\omega\in A}\{\omega\}\right)\overset{\text{ס.אדיטיביות}}{=}\sum_{\omega\in A}p(\{\omega\})\overset{(1)}{=}\sum_{\omega\in A}p(\omega)$$
$$\overset{A=\text{supp}(p)}{=}\sum_{\omega\in\Omega}p(\omega)\overset{\text{הסתברות נקודתית}}{=}1$$

$(*)$ איחוד של $\aleph_0$ קבוצות זרות לכל היותר.

$(2)\Rightarrow(1)$: נבחין כי למעשה מתקיים־ $B=(B\cap A)\underset{\text{זר}}{\bigcup}(B\cap A^c)$ ולכן:

$$\mathbb{P}(B)=\mathbb{P}(B\cap A)+\mathbb{P}(B\cap A^c)\underset{(☺)}{=}\mathbb{P}(B\cap A)$$

$$(☺)\,\mathbb{P}(B\cap A^c)\leq\mathbb{P}(A^c)=1-\mathbb{P}(A)\overset{(2)}{=}1-1=0$$

כיוון ש־$|B\cap A|\leq|A|\leq\aleph_0$ מתקבל כי:

$$\mathbb{P}(B)=\mathbb{P}(B\cap A)\overset{\text{ס.תאדיטיביות}}{=}\sum_{\omega\in B\cap A}\mathbb{P}(\{\omega\})\overset{(☻)}{=}\sum_{\omega\in B}\mathbb{P}(\{\omega\})$$
$$(☻)\,\omega\notin A\rightarrow\mathbb{P}(\{\omega\})\leq\mathbb{P}(A^c)=0$$

$(2)\Rightarrow(3)$ ניקח את $A=\{\omega\in\Omega\mid\mathbb{P}(\{\omega\}>0)\}$ (supp עם היחידונים).

ואז נקבל מ־ (3) כי $|A|\leq\aleph_0$ ולכן גם $1\overset{(3)}{=}\sum\limits_{\omega\in\Omega}\mathbb{P}(\{\omega\})=\sum\limits_{\omega\in A}\mathbb{P}(\{\omega\})\overset{\text{ס.אדיטיביות}}{=}\mathbb{P}(A)$.

דוגמא למרחב הסתברות לא בדיד

הגרלת מספר אחד בין 0 ל־1. מרחב המדגם הינו $\Omega = [0,1]$. מיהו $\mathcal{F}$? אם ניקח $\mathcal{F} = 2^{\Omega}$[9], נצטרך להגדיר לכל $A \subseteq [0,1]$ מהו $\mathbb{P}(A)$.

ברור כי אנו מעוניינים ב־ $A = [\alpha, \beta]$ (בפרט כי $0 \le \alpha \le \beta \le 1$) לקבוע $\mathbb{P}(A) = \beta - \alpha$.

מכך, נוכל להסיק ישירות כי לסדרת קטעים זרים $A_i = [\alpha_i, \beta_i]$ עלינו לקבוע כי $\sum_{i=1}^{\infty} \beta_i - \alpha_i = \mathbb{P}\left(\bigcup_{i=1}^{\infty} [\alpha_i, \beta_i]\right)$.

אמנם, ב־$2^{[0,1]}$ ישנן עוד המון קבוצות. התשובה המלאה לבעיה ניתנת בקורס "תורת המידה".

יחד עם זאת, נוכל לראות כי $\mathbb{P}(\mathbb{Q} \cap [0,1]) = 0$

ולכן $\mathbb{P}(\mathbb{R} \setminus \mathbb{Q} \cap [0,1]) = 1$.

בנוסף $\mathbb{P}\left(\sum_{m=1}^{\infty} \frac{c_m}{10^m} \;\middle|\; \begin{matrix} 0 \le c_m \le 9 \\ c_m \ne 7 \end{matrix}\right) = ?$ - מרחב המדגם אינו בן מנייה (אנו מצפים שייצא 0 בגלל מודל הטלת הקוביה).

היינו רוצים למצוא פונקציית הסתברות $\mathbb{P}$ כך שיתקיים כי לכל $A \subseteq [0,1]$ ולכל $0 \le \gamma \le 1$ כך שגם ל־ $A + \gamma = \{\alpha + \gamma \mid \alpha \in A\} \subseteq [0,1]$ מתקיים כי $\mathbb{P}(A) = \mathbb{P}(A + \gamma)$ "אינווריאנטית להזזה".

אך אין כזו.

משפט

ישנה $\mathcal{F} \subseteq 2^{[0,1]}$ המכילה את כל הקטעים ב־$[0,1]$ וסגורה לאיחודים וחיתוכים. בני מניה, ופונקצית הסתברות $\mathbb{P}$ על $\mathcal{F}$ הזו, כך ש־ $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות ו־$\mathbb{P}$ אינווריאנטית להזזה. לפונקציה זו קוראים התפלגות אחידה על $[0,1]$.

הערה

מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$ נקרא רציף אם $\forall \omega \in \Omega \; \mathbb{P}(P\{\omega\}) = 0$.

דוגמא

במגירה זוג גרביים שחורים ושני זוגות גרביים לבנים. מוציאים 2 גרביים, מה הסיכוי שיצאו שני גרביים באותו הצבע?

תשובה: נמספר את הגרביים $[1, \ldots, 6]$ כך ש־$1, 2$ שחורות. ואז נקבל:

$$\Omega = \{(i,j) \mid i, j \in [6], i \ne j\}$$

כעת נוכל להבחין כי $p((i,j)) = \frac{1}{|\Omega|} = \frac{1}{30}$.

היינו גם יכולים לקחת גם זוגות לא סדורים:

$$\Omega' = \{\{i,j\} \mid i \le j \le 6, \}$$

ואז נקבל כי $p'(\{i,j\}) = \frac{1}{|\Omega'|} = \frac{1}{15}$.

(כל זוג לא סדור ב־Ω' מתאים לשני זוגות מ־Ω)

נאבחן מהי הקבוצה הרצויה מבחינתנו:

[9] קבוצת כל הקבוצות החלקיות

$$A=\left\{\{i,j\}\in\Omega'\mid\begin{array}{c}i\leq i<j\leq 2\\ 3\leq i\leq j\leq 6\end{array}\right\}$$

ואז נקבל:

$$\mathbb{P}\left(A\right)=\sum_{\omega\in A}p\left(\omega\right)=\frac{|A|}{|\Omega'|}=\frac{7}{15}$$

בדקו מה היה קורה במודל Ω.

מכפלת מרחבי הסתברות בדידים

אם p_1,p_2 פונקציות הסתברות נקודתיות על מרחבי מדגם Ω_1,Ω_2 בהתאמה, נגדיר פונקצית הסתברות נקודתית p על $\Omega_1\times\Omega_2$:

$$\begin{aligned}&p:\Omega_1\times\Omega_2\to\mathbb{R}_+\\&p\left(\omega_1,\omega_2\right)=p_1\left(\omega_1\right)\cdot p_2\left(\omega_2\right)\end{aligned}$$

נבחין כי זו אכן פונקצית הסתברות נקודתית, שהרי מתקיים:

$$\begin{aligned}&\sum_{\omega\in\Omega_1\times\Omega_2}p\left(\omega_1,\omega_2\right)=\sum_{\omega_1\in\Omega_1}p_1\left(\omega_1\right)\sum_{\omega_2\in\Omega_2}p_2\left(\omega_2\right)=\\&\sum_{\omega_1\in(p_1)\text{supp}}p_1\left(\omega_1\right)\sum_{\omega_2\in(p_2)\text{supp}}p_2\left(\omega_2\right)\overset{\text{פה"נ}}{=}\\&\sum_{\omega_1\in(p_1)\text{supp}}p_1\left(\omega_1\right)\overset{\text{פה"נ}}{=}1\end{aligned}$$

דוגמא

הטלת מטבע הוגן - $\Omega_1=\{H,T\},P_1=\frac{1}{2}$ והטלת קוביה הוגנת - $\Omega_2=[6],P_1=\frac{1}{6}$.
נקבל:

$$\left(p_1\times p_2\right)\left((s,i)\right)=p_1\left(s\right)\cdot p_2\left(i\right)=\frac{1}{2}\cdot\frac{1}{6}=\frac{1}{12}$$

באותו אופן, תמיד כש-p_1,p_2 הסתברויות אחידות על Ω_1,Ω_2 אזי $p_1\times p_2$ ייתן הסתברות אחידה על $\Omega_1\times\Omega_2$. למעשה, $p_1\times p_2$ משקפת ביצוע של הניסוי $\left(\Omega_1,p_1\right)$ והניסוי $\left(\Omega_2,p_2\right)$ באופן בו אין השפעה בין הניסויים.

הגדרה

אם $\left(\Omega_1,2^{\Omega_1},\mathbb{P}_1\right)$ ו-$\left(\Omega_2,2^{\Omega_2},\mathbb{P}_2\right)$ מרחבי הסתברות בדידים,
נגדיר את מרחב המכפלה $\left(\Omega_1\times,\Omega_2,2^{\Omega_1\times\Omega_2},\mathbb{P}_1\times\mathbb{P}_2\right)$ ע"י $\mathbb{P}_1\times\mathbb{P}_2=\mathbb{P}_{p_1\times p_2}$ כאשר p_1 ו-p_2 הינן פונקציות הסתברות שנותנות את $\mathbb{P}_1,\mathbb{P}_2$.

טענה

$$\forall A \subseteq \Omega_1, \quad B \subseteq \Omega_2 : \ (\mathbb{P}_1 \times \mathbb{P}_2)(A \times B) = \mathbb{P}_1(A) \cdot 2(B)$$

הוכחה

נבחין כי מתקיים:

$$\begin{aligned}
(\mathbb{P}_1 \times \mathbb{P}_2)(A \times B) &\overset{\text{הגדרה}}{=} \mathbb{P}_{p_1 \times p_2}(A \times B) \overset{(**)}{=} \sum_{\omega \in A \times B} (p_1 \times p_2)(\omega) \\
&= \sum_{\omega \in A \times B} p_1(\omega) \cdot p_2(\omega) = \sum_{\substack{\omega_1 \in A \\ \omega_2 \in B}} p_1(\omega_1) \cdot p_2(\omega_2) = \\
&= \sum_{\omega_1 \in A} p_1(\omega_1) \sum_{\omega_2 \in A} p_2(\omega_2) = \sum_{\omega_1 \in A} p_1(\omega_1) P_2(B) = \\
&P_1(A) P_2(B)
\end{aligned}$$

מסקנה

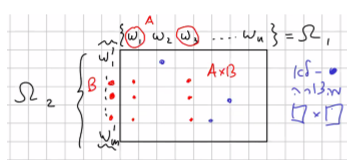
לכל $A \subseteq \Omega_1$: $(\mathbb{P}_1 \times \mathbb{P}_2)(A \times \Omega_2) = \mathbb{P}_1(A)$

שיעור מס' 5:
יום שלישי
3.11.20

לכל $B \subseteq \Omega_1$: $(\mathbb{P}_1 \times \mathbb{P}_2)(\Omega_1 \times B) = \mathbb{P}_2(B)$

מאורעות כאלו $((A \times \Omega_2), (\Omega_1 \times B))$ נקראים **מאורעות שוליים** (מסתכלים על הראשון ולא אכפת לנו מהשני).

נשים לב כי לכל מאורע ב־$\Omega_1 \times \Omega_2$ הוא מהצורה $A \times B$, למשל, לא מדובר על מלבנים מרווחים בדוגמא הבאה:



$A \times B$ הם רק "מלבנים מרווחים" ומאורעות כאלו נקראים **מאורעות מכפלה.**

למשל $\Omega_1 = \Omega_2 = [2]$, אזי $\{(1,1),(2,2)\}$ אינו מאורע מכפלה.

הוכחה

אם $A \times B$ אזי $1,2 \in A$ וגם $1,2 \in B$ ואזי $A \times B = \Omega_1 \times \Omega_2$, שהרי $\Omega_1 \times \Omega_2 = \{(1,1),(1,2),(2,1),(2,2)\}$

מכפלה של הרבה מרחבי הסתברות

אם $p_1, \ldots, p_m$ פה"נ על $\Omega, \ldots, \Omega_1$, נגדיר פה"נ $p_1 \times \ldots \times p_2$ על $\Omega_1 \times \ldots \times \Omega_n$ ע"י

$$(p_1 \times \ldots \times p_n)(\omega_1, \ldots, \omega_n) = p_1(\omega_1) \cdot \ldots \cdot p_n(\omega_n)$$

ואם $(\Omega_i, 2^{\Omega_i}, P_i)$ מרחבי הסתברות (כאשר $i = 1, \ldots, n$), נגדיר את מרחב המכפלה $(\Omega_1 \times \ldots \Omega_n, 2^{\Omega_1 \times \ldots \Omega_n}, \mathbb{P}_1 \times \ldots \times \mathbb{P}_n)$ כש־$\mathbb{P}_1 \times \ldots \times \mathbb{P}_n$ מושרית ממכפלת הפה"נ ולכן בפרט מתקבל הסתברות בדיד.

כלומר, למעשה $(P_1 \times \ldots \times P_n)(A) = \sum_{\omega \in A} p_1(\omega_2) \cdot p_2(\omega_2) \cdot \ldots \cdot p_n(\omega_n)$

מכאן והלאה נסמן $P(\{\omega\}) = P(\omega)$ וניפטר מ־p.

דוגמא

ניקח $\{H,T\} = \Omega_1 \times \ldots \Omega_n$ ולכל i $P_i(\{H\}) = \alpha$ (ולכן $P(T) = 1-\alpha$).
מרחב המכפלה מתאים ל־n הטלות שלא משפיעות זו על זו.
איברי $\Omega_1 \times \ldots \Omega_n$ הם סדרות ב־$\{H,T\}^n$.
מהי ההסתברות על סדרה כזו, למשל: $(\mathbb{P}_1 \times \ldots \times \mathbb{P}_6)((H,H,H,T,H,T))$?
נקבל כי:

$$\mathbb{P}_1(H)\,\mathbb{P}_2(H)\,\mathbb{P}_3(H)\,\mathbb{P}_4(T)\,\mathbb{P}_5(H)\,\mathbb{P}_6(T) = \alpha^4(1-\alpha)^2$$

נסמן: $\Omega = \Omega_1 \times \ldots \Omega_n$ וגם $P = \mathbb{P}_1 \times \ldots \times \mathbb{P}_n$, אזי נקבל כי:

$$\mathbb{P}\left(\underbrace{\omega}_{\text{סדרת הטלות}}\right) = \alpha^{\text{מספר H ב־}\omega} \cdot (1-\alpha)^{\text{מספר T ב־}\omega}$$

אם נרצה לדבר על מאורעות, נקבל כי:

$$A = \{\omega \in \Omega \mid \text{יש H בדיוק k פעמים ב־}\omega\}$$

למרות שהמרחב אינו אחיד:

$$\begin{aligned} P(A) = \sum_{\omega \in A} p(\omega) = \sum_{\omega \in A} \alpha^k (1-\alpha)^{n-k} = \\ |A|\,\alpha^k(1-\alpha)^{n-k} = \binom{n}{k}\alpha^k(1-\alpha)^{n-k} \end{aligned}$$

אם $(\Omega_1, \mathbb{P}_1)$ ו־$(\Omega_2, \mathbb{P}_2)$ מרחבי הסתברות, הגדרנו מרחב הסתברות על מרחב המדגם $\Omega_1 \times \Omega_2$.
נבחין כי יש עוד מרחבי הסתברות. למשל, אם הטלנו מטבע הוגן פעמיים ו־$(\Omega_1, \mathbb{P}_1)$ מייצג את מספר הראשים:

$$\Omega_1 = \{0,1,2,3\}, \qquad \mathbb{P}_1 : \begin{array}{c} 0 \to \frac{1}{4} \\ 1 \to \frac{1}{2} \\ 2 \to \frac{1}{4} \end{array}$$

ולעומת זאת $(\Omega_2, \mathbb{P}_2)$ מייצג את מספר הזנבות:

$$\Omega_1 = \{0,1,2,3\}, \qquad \mathbb{P}_2 : \begin{array}{c} 0 \to \frac{1}{4} \\ 1 \to \frac{1}{2} \\ 2 \to \frac{1}{4} \end{array}$$

כעת נבחין כי $\mathbb{P}((2,2))=\mathbb{P}((0,1))=\mathbb{P}((0,0))=0$

ומאידך נקבל כי $\mathbb{P}((1,1))=\frac{1}{2}$ $\mathbb{P}((0,2))=\frac{1}{4}$.

אם כך, זהו לא מרחב המכפלה!

שיעור מס' 6:
יום חמישי
5.11.20

שאלה:

נתונה הסתברות $\mathbb{P}$ על $\Omega_1\times\Omega_2$. האם $\mathbb{P}$ היא הסתברות מכפלה של הסתברויות כלשהן על $\mathbb{P}_1,\mathbb{P}_2$ על Ω_1,Ω_2.

נתבונן בטבלה הבאה, המייצגת את $\mathbb{P}$:

	1	2	3	4
A	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{3}{24}$	$\frac{4}{24}$
B	$\frac{4}{24}$	$\frac{3}{24}$	$\frac{2}{24}$	$\frac{1}{24}$
C	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{24}$

האם בהכרח $\mathbb{P}$ הינה מכפלת פונקציות $\mathbb{P}_1,\mathbb{P}_2$ כלשהן? [10]

ננסה לשחזר מי היו $\mathbb{P}_1,\mathbb{P}_2$. ראינו כי במכפלה של 2 מרחבי הסתברות מתקיים:

$$\mathbb{P}(\{\omega\}\times\Omega_2)=\mathbb{P}_1(\omega)$$

כעת, נגדיר $A=\omega$ ונקבל:

$$\mathbb{P}(\{A\}\times\Omega_2)=\mathbb{P}_1(A)$$

אם נסכום את הטבלה מקודם, נקבל כי מאורע השוליים הספציפי הינו $\{A\}\times\Omega_2$.

כלומר סכומי הטורים והשורות משחזרים את $\mathbb{P}_1,\mathbb{P}_2$. כעת ננסה לבדוק האם אכן כך מתקבל:

$$\mathbb{P}((\omega_1,\omega_2))=\mathbb{P}_1(\omega_1)\cdot\mathbb{P}_2(\omega_2)=\mathbb{P}(\omega_1\times\Omega_2)\,\mathbb{P}(\Omega_1\times\omega_2)$$

ואם אכן (ורק אם) מתקבל, אז אכן $(\Omega_1\times\Omega_2,P)$ הוא מרחב מכפלה.

נבדוק זאת על הדוגמא שלנו, ואכן, אצלנו מתקיים כי $\mathbb{P}(A,2)=\frac{2}{24}=\frac{10}{24}\cdot\frac{6}{24}=\mathbb{P}_1(A)\cdot\mathbb{P}_2(2)$

כלומר, זהו לא מרחב הסתברות של מכפלה (דהיינו, הניסויים כן מושפעים זה מזה).

איך ניתן לתאר מקרה שבו שני ניסויים כן משפיעים זה על זה? ניסוי דו שלבי.

ניסוי דו שלבי

ניסוי דו שלבי מורכב ממרחב הסתברות $(\Omega_1,\mathbb{P})$ - "הניסוי הראשון", ומרחב מדגם Ω_2 ולכל $\omega_1\in\Omega_1$ פונקצית הסתברות $\mathbb{P}_{\omega_1}$ על Ω_2.

[10] ניתן לומר גם כי אם"ם כל שתי עמודות וכל שתי שורות הן תלויות לינארית זו בזו (ולכן בכל שתי שורות מימד השורות שווה למימד העמודות)

הרעיון: $\mathbb{P}_{\omega_1}$ מתארת את ההתפלגות בניסוי השני אחרי שיצא ω_1 בניסוי הראשון.

דוגמא

נבחר קובית $D\&D$ אקראית, ומטיל אותה. הניסוי הראשון: איזו קוביה בחרתי. הניסוי השני: מה יצא בהטלה.

למשל, $\Omega_1=\{D_4,D_6,D_8,D_{10},D_{12},D_{20}\}$

נקבל $\mathbb{P}\left(D_m\right)=\frac{1}{6}$.

בנוסף, $\Omega_2=\{1,\ldots,20\}$.

כעת נקבל כי $\mathbb{P}_{D_4}\left(i\right)=\begin{cases}\frac{1}{4} & 1\le i\le 4\\ 0 & \text{אחרת}\end{cases}$.

ובאופן כללי, נקבל כי:

$$\mathbb{P}_{D_m}\left(i\right)=\begin{cases}\frac{1}{m} & 1\le i\le m\\ 0 & \text{אחרת}\end{cases}$$

ניסוי דו שלבי למעשה מגדיר התפלגויות על $\Omega_1\times\Omega_2$ ע"י $\mathbb{P}_x\left(\left(\omega_1,\omega_2\right)\right)=\mathbb{P}\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2\right)$.

בדוגמא למשל, נוכל לראות $\mathbb{P}_x\left(\left(D_8,17\right)\right)=0$, ומאידך $\mathbb{P}_x\left(\left(D_8,5\right)\right)=\frac{1}{6}\cdot\frac{1}{8}$.

ובאופן כללי, נקבל $\mathbb{P}_x\left(\left(D_m,i\right)\right)=\begin{cases}\frac{1}{6m} & i\le i\le m\\ 0 & \text{Otherwise}\end{cases}$.

נוכל גם לבדוק מה ההסתברויות לקבל מספר ספציפי (למשל, מה הסיכוי שנקבל 5), באמצעות $\mathbb{P}\left(\Omega_1\times\{i\}\right)=\sum_m\mathbb{P}\left(\left(D_m,i\right)\right)$.

הערה

נשים לב כי אם $\mathbb{P}_{\omega_1}=\mathbb{P}_{\omega_1'}$ לכל ω_1,ω_1' אזי ההסתברות המושרית ע"י הניסוי הדו שלבי היא הסתברות מכפלה - "הניסוי הראשון לא משפיע על השני". כלומר, למעשה הניסוי הדו שלבי הוא הכללה של הסתברויות מכפלה.

שאלה

נתונה הסתברות $\mathbb{P}$ על $\Omega_1\Omega_2\times$. כיצד ניתן להכריע האם היא ניתנת להצגה כניסוי דו שלבי?

תשובה

בתרגיל.

הגדרה

נאמר כי ניסוי הינו ניסוי רב שלבי, אם $\Omega=\Omega_1\times\Omega_2\times\ldots\times\Omega_m$.

ו-$\mathbb{P}$ שהינה פונקציית הסתברות על Ω_1.

כאשר מתקיים - $\forall\omega_1\in\Omega_1\mathbb{P}_{\omega_1}$ פונקציית הסתברות על Ω_2.

כמו כן, מתקיים $\forall\left(\omega_1,\omega_2\right)\in\Omega_1\times\Omega_2\mathbb{P}_{\omega_1},\omega_{\omega_1,\omega_2}$ פונקציית הסתברות על Ω_3.

ובמקרה כללי, נקבל כי $\forall\left(\omega_1,\ldots,\omega_{m-1}\right)\in\Omega_1\mathbb{P}_{\omega_1,\ldots,\omega_{m-1}}$

הינה פונקציית הסתברות על Ω_m.

לסיכום, ניסוי רב שלבי משרה פונקציית הסתברות על $\Omega_1\times\ldots\times\Omega_m$ ע"י:

$$\mathbb{P}_x\left(\left(\omega_1,\ldots\omega_m\right)\right)=\mathbb{P}\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2\right)\cdot\ldots\mathbb{P}_{\omega_1,\ldots,\omega_{m-1}}\left(\omega_m\right)$$

נוסחת ההסתברות השלימה, תלות וחוק בייס

נוסחת ההסתברות השלימה

טענה

אם נתון מרחב הסתברות $(\Omega, \mathbb{P})$ ומאורע $A \in \mathcal{F}$ אזי לכל למאורע B מתקיים:

$$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)$$

הוכחה

הם זרים ואיחודם הוא B.

טענה

אם $A_i, \ldots, A_n$ זרים בזוגות ו־$\bigcup\limits_{i=1}^{m} A_i = \Omega$ אזי מתקיים כי $\mathbb{P}(B) = \sum\limits_{i=1}^{m} \mathbb{P}(B \cap A_i)$.

הוכחה

באינדוקציה.

טענה

עבור $(\Omega, \mathcal{F}, \mathbb{P})$, אם $\{A_i\}_{i=1}^{\infty}$ סדרה של מאורעות זרים בזוגות, כאשר $A_i \in \mathcal{F}$ ו־$\bigcup\limits_{i=1}^{\infty} A_i = \Omega$ ו($\{A_i\}$ היא חלוקה של Ω") אזי לכל $B \in \mathcal{F}$ מתקיים כי $\mathbb{P}(B) = \sum\limits_{i=1}^{\infty} \mathbb{P}(B \cap A_i)$.

הוכחה

$\{B \cap A_i\}$ הם זרים בזוגות, ואיחודם הוא B.

הטענות האלו נקראות "נוסחת ההסתברות השלימה".

דוגמא

נבחר קובית $D\&D$ אקראית, כאשר $\Omega = \{1, \ldots, 20\}$.

מהו ההסתברות לתוצאה שמתחלקת ב־3? יהיה לנו קשה לחשב. לכן, ניקח מרחב מדגם גדול יותר, בו נציין גם איזו קוביה בחרתי, כלומר $\Omega = \{D_4, \ldots, D_{20}\} \times [20]$.

כעת, נבחר $\mathbb{P}(A)$ כאשר A הינה התוצאה המתחלקת ב־3.

לכל $m \in \{4, 6, 8, 10, 12, 20\}$ נגדיר $A_m = \{D_m\} \times [20]$. זו חלוקה של איחודים זרים $A_4 \uplus \ldots \uplus A_{20} = \Omega$.

סך הכל נקבל:

$$\mathbb{P}(B) = \mathbb{P}\underbrace{(B \cap A_4)}_{\frac{1}{6} \cdot \frac{1}{4}} + \mathbb{P}\underbrace{(B \cap A_6)}_{\frac{1}{6} \cdot \frac{1}{3}} + \mathbb{P}(B \cap A_8) + \underbrace{\mathbb{P}(B \cap A_{10})}_{\frac{1}{6} \cdot \frac{3}{10}} + \ldots$$

כלומר, למעשה נוסחת ההסתברות השלימה הינה פירוק של המאורע לניסויים רב שלביים.

שאלה

במידה ונתונה חלוקה כללית - $\Omega = \biguplus\limits_{i \in I} A_i$, האם לכל B יתקיים $\mathbb{P}(B) = \sum\limits_{i \in I} \mathbb{P}(B \cap A_i)$.

שיעור מס' 7:
יום שלישי
10.11.20

תשובה:

לא. אם $(\Omega, \mathbb{P})$ הוא מ"ה רציף, $(\mathbb{P}\{\omega\} = 0)$ אזי $I = \Omega$ ו־$A_i = \{i\}$ ואז לכל B נקבל $\sum_{i \in I} \mathbb{P}(B \cap A_i) = 0$ (כי $B \cap A_i$ יחידון או ריק), אבל $\mathbb{P}(\Omega) = 1$.

תרגיל

הוכיחו שדבר זה נכון אם $(\Omega, \mathbb{P})$ מרחב הסתברות בדיד.

חסם האיחוד

טענה:

אם $\{A_i\}_{i=1}^{\infty}$ סדרה של מאורעות במ"ה, אזי $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

הוכחה

נגדיר $B_1 = A_1$ וגם $B_j = A_j \setminus (A_1 \cup \ldots \cup A_{j-1})$ לכל $j \geq 2$.

אזי נקבל כי $\{B_j\}_{j=1}^{\infty}$ סדרה של מאורעות זרים.

כמו כן, $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ ולכן:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \overset{\text{הגדרה}}{=} \mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) \overset{\text{ס.אדיטיביות}}{=} \sum_{i=1}^{\infty} \mathbb{P}(B_i) \overset{\text{מונטוניות}}{\leq} \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

תרגיל

הסיקו חסם איחוד עבור סדרה סופית של מאורעות.

הסתברות מותנית

אם נרצה, הסתברות מותנית משקפת לנו, מציאותית, מצב בו יש לנו מידע מסוים. למשל "הטלתי קוביה ויצא מספר זוגי. מה הסיכוי שהוא 4". אנחנו יודעים שזה $\frac{1}{3}$.

הגדרה

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מ"ה, ויהי $B \in \mathcal{F}$ מאורע עם $\mathbb{P}(B) > 0$. אזי לכל $A \in \mathcal{F}$, ההסתברות של A בהינתן B היא:

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

דוגמאות

1. בדוגמה שהבאנו מקודם, $B = \{2, 4, 6\}$ ו־$\{A\} = 4$ אזי נקבל כי:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(\{4\})}{\mathbb{P}(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

הרעיון הינו כי תמיד יתקיים בהכרח ־ $\mathbb{P}(A\cap B)=\mathbb{P}(B)\,\mathbb{P}(A\mid B)$ [11]

2. נוכל להתבונן גם בשאלה הבאה ־ מהי ההסתברות לקבל מספר זוגי בקוביה, בהינתן שיצא מספר קטן מ־4?

$$\mathbb{P}(\{2,4,6\}\mid\{1,2,3\})=\frac{\mathbb{P}(\{2\})}{\mathbb{P}(\{1,2,3\})}=\frac{1}{3}$$

3. אם $\mathbb{P}$ הסתברות אחידה על Ω אזי:

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

הגדרה

אם $(\Omega,\mathcal{F},\mathbb{P})$ מ"ה, לא דווקא בדיד ו־$B\in\mathcal{F}$ עם $\mathbb{P}(B)>0$, אזי נגדיר פונקצית הסתברות מותנית על $\Omega,\mathcal{F}$ ע"י B:

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

למשל אם $(\Omega,\mathbb{P})$ הטלת קוביה, ו־B תוצאה זוגית, אזי:

$$\mathbb{P}_B(\{i\})=\begin{cases}0 & \text{odd}\\ \frac{1}{3} & \text{even}\end{cases}$$

טענה

$\mathbb{P}_B$ היא פונקצית הסתברות.

הוכחה

נבחין כי מתקיים, מההגדרה:

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

בנוסף, אם $\{A_i\}_{i=1}^{\infty}$ זרים בזוגות, אזי נקבל:

$$\mathbb{P}_B\left(\bigcup A_i\right)=\frac{\mathbb{P}\left(\bigcup\limits_{i=1}^{\infty}A_i\cap B\right)}{\mathbb{P}(B)}=\frac{\mathbb{P}\left(\bigcup\limits_{i=1}^{\infty}(A_i\cap B)\right)}{\mathbb{P}(B)}\overset{\text{ס.אדיטיביות}}{=}\frac{\sum\limits_{i=1}^{\infty}\mathbb{P}(A_i\cap B)}{\mathbb{P}(B)}=\sum_{i=1}^{\infty}\frac{\mathbb{P}(A_i\cap B)}{\mathbb{P}(B)}=\sum_{i=1}^{\infty}\mathbb{P}_B(A_i)$$

מה הקשר לניסוי דו שלבי? נוכל להבחין

[11] למעט זה ש־$\mathbb{P}(A\mid B)$ איננו מוגדר, אך $\mathbb{P}(B)=0$ ואז נקבל כי לא מוגדר $0=0*$

$$\mathbb{P}_x\left(\left(\omega_1,\omega_2\right)\right)=\mathbb{P}_1\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2\right)$$
$$\mathbb{P}\left(A\cap B\right)=\mathbb{P}\left(A\right)\cdot\mathbb{P}\left(A\mid B\right)$$

שיעור מס' 8:

יום חמישי

12.11.20

מבחינה פורמלית, נסמן מאורעות שוליים ע"י $\left(\cdot,\,\omega_2\right)=\Omega_1\times\left\{\omega_{n\omega}\right\}$ ו־$\left(\omega_1,\cdot\right)=\left\{\omega_1\right\}\times\Omega_2$.
במקום לכתוב הסתברות במרחב מכפלה, נוכל לכתוב בצורה שונה:

$$\mathbb{P}\left(\Omega_1\times\left\{\omega_2\right\}\right)=\mathbb{P}\left(\left(\cdot,\omega_2\right)\right)=\mathbb{P}\left(\left(\cdot,\omega_2\right)\right)$$

אם נתון ניסוי דו שלבי ־ $\left(\Omega,\mathbb{P}\right)$, $\left\{\left(\Omega_2,\mathbb{P}_{\omega_1}\right)\right\}_{\omega_1\in\Omega_1}$ ו־$\mathbb{P}_x\left(\omega_1,\omega_2\right)=\mathbb{P}\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2\right)$.
דבר זה מתאים להסתברות מותנה במאורעות שוליים:

$$\mathbb{P}_x\left(\left(\cdot,\,\omega_2\right)\mid\left(\omega_1,\cdot\right)\right)=\frac{\mathbb{P}_x\left(\left(\left(\cdot,\,\omega_2\right)\cap\left(\omega_1,\cdot\right)\right)\right)}{\mathbb{P}_x\left(\left(\omega_1,\cdot\right)\right)}=\frac{\mathbb{P}_x\left(\omega_1,\omega_2\right)}{\mathbb{P}_x\left(\omega_1,\cdot\right)}\overset{\text{הגדרות}}{=}$$
$$\frac{\mathbb{P}\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2\right)}{\sum\limits_{\omega_2'\in\Omega}\mathbb{P}_x\left(\omega_1\cdot\omega_2'\right)}=$$
$$\frac{\mathbb{P}\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2\right)}{\sum\limits_{\omega_2'\in\Omega}\mathbb{P}\left(\omega_1\right)\cdot\mathbb{P}_{\omega_1}\left(\omega_2'\right)}\overset{\text{פונקציית הסתברות}}{=}\mathbb{P}_{\omega_1}\left(\omega_2\right)$$

לא כל ניסוי הוא מרחב מכפלה, אבל דבר זה אכן דומה למתרחש בניסוי דו־שלבי.

חיתוך של 3 מאורעות

נניח שנרצה לבחון את ההסתברות של חיתוך 3 מאורעות:

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

דוגמא

בוחרים אדם אקראי ־ מהי ההסתברות שהוא קרח, מכריס ושעיר:

$$\mathbb{P}\left(\text{קרח, מכריס ושעיר}\right)=\mathbb{P}\left(\text{קרח}\right)\mathbb{P}\left(\text{מכריס}\mid\text{קרח}\right)\mathbb{P}\left(\text{שעיר}\mid\text{קרח ומכריס}\right)$$

תרגיל

הכלילו ל־$\mathbb{P}\left(A_1\cap\ldots\cap A_n\right)$

התנייה כפולה

נזכיר כי $\mathbb{P}_B(A)=\mathbb{P}(A\mid B)$. אם כך, מהו $(\mathbb{P}_B)_{B'}$? כלומר, למשל מקרה בו קיבלנו לפי הגדרה, מתקיים:

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

כל זה עד כי העובדה שדברים עשויים להיות לא מוגדרים אם חילקנו באפס (ואם $\mathbb{P}(B\cap B')>0$ אין חלוקות ב־0).

חוק בייס

נבחין כי בהסתברות אין משמעות לזמן, ולכן בהכרח מתקיים כי:

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

ולכן, אם $\mathbb{P}(A)>0$ אזי:

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

דוגמא

הטלתי קוביית $D\&D$ אקראית, ויצא 11. מהי ההסתברות שבחרתי את D_{12} ?

$$\mathbb{P}(\text{בחרתי}D_{12}\mid\text{הטלתי }11)=\mathbb{P}(\text{הטלתי}11\mid\text{בחרתי}D_{12})\cdot\frac{\mathbb{P}(\text{בחרתי}D_{12})}{\mathbb{P}(\text{הטלתי}11)}=$$
$$\frac{\frac{1}{6}}{\mathbb{P}(\text{הטלתי}11)}=\frac{\frac{1}{6}}{\frac{1}{45}}=7.5$$

כי מתקיים:

$$\mathbb{P}(\text{הטלתי}11)\overset{\text{הסתברות השלימה}}{=}\sum_{m=4,6\ldots20}\mathbb{P}(\text{הטלתי}1\cap\text{בחרתי}D_m)=$$
$$0+0+0+0+\frac{1}{6}\cdot\frac{1}{12}+\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{1}{20}=\frac{1}{45}$$

נוסחת הסתברות שלימה נוספת

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מ"ה לא דווקא בדיד ו־$\{A_i\}_{i=1}^{\infty}$ או $\{A_i\}_{i=1}^{n}$ חלוקה של Ω.
(כלומר $A_i\cap A_j=\emptyset\quad\forall i\neq j$ וגם $\bigcup\limits_i A_i=\Omega$).
אזי לכל $B\in\mathcal{F}$ מתקיים כי $\mathbb{P}(B)=\sum\limits_i\mathbb{P}(B\mid A_i)\mathbb{P}(A_i)$

הוכחה

$$\mathbb{P}(B) \overset{\text{נה"ש}}{=} \sum_i \mathbb{P}(B\cap A_i) = \sum_i \mathbb{P}(B\mid A_i)\,\mathbb{P}(A_i)$$

דוגמא

נתבונן בסיכוי לקבלת תוצאה בהטלת קובייה אקראית:

$$\mathbb{P}(\text{תוצאה } j \text{ בהטלת קובייה אקראית}) = \sum_{m\in\{4,6,8,10,12,20\}} \mathbb{P}(\text{הטלת } j \mid \text{בחרתי את } D_m)\cdot \mathbb{P}(\text{בחרתי את } D_m) = \frac{1}{6}\sum_m \mathbb{P}(\text{הטלת } j \mid \text{בחרתי } D_m) = \frac{1}{6}\cdot \sum_{\substack{m\in\{4,6,\dots,20\}\\ j\le m}} \frac{1}{m}$$

נשים לב כי דילגנו על בחירת $\mathbb{P},\Omega$. אין בעיה בכך, כל עוד החישובים נכונים.

בנוסף, ניתן לבחור גם $\mathbb{P},\Omega$ שונים בכל פעם (ניתן לבחור אפשרות לזוגות סדורים או לא סדורים).

ניסוח אחר לבייס

פעמים רבות מנסחים את בייס כך:

בהינתן חלוקה $\{A_i\}_i$ (חלוקה סופית או בת מנייה או מלכתחילה מ"ה בדיד), לכל מאורע B מתקיים:

$$\mathbb{P}(A_i\mid B) \overset{\text{בייס}}{=} \frac{\mathbb{P}(B\mid A_i)\,\mathbb{P}(A_i)}{\underbrace{\sum_j \mathbb{P}(B\mid A_j)\,\mathbb{P}(A_j)}}$$

הסקה בייסאנית

נתבונן במקרה הבא:

שיעור מס' 9:

יום שלישי

17.11.20

$$\mathbb{P}(\text{דני חולה קורונה} \mid \text{דני קיבל תוצאה חיובית בבדיקה}) = \frac{\mathbb{P}(\text{דני קיבל תוצאה חיובית בבדיקה} \mid \text{דני חולה קורונה})\,\mathbb{P}(\text{דניחולהקורונה})}{\mathbb{P}(\text{דניקיבלתוצאהחיוביתבבדיקה})}$$

נוכל להבחין שהמקרה האחרון למעשה שווה למקרה הבא:

$$\overset{\text{נה"ש}}{=} \frac{\overbrace{\mathbb{P}\left(\text{אדם חולה}\right)}^{\text{שיעור החולה באוכלוסיה}}\overbrace{\mathbb{P}\left(\text{תוצאה חיובית} \mid \text{אדם חולה}\right)}^{\text{מהימנות הבדיקה}}}{\mathbb{P}\left(\text{תוצאה חיובית}\right)}$$

$$\frac{\mathbb{P}\left(\text{חולה}\right)\mathbb{P}\left(\text{חיובי}\mid\text{חולה}\right)}{\mathbb{P}\left(\text{חולה}\right)\mathbb{P}\left(\text{חיובי} \mid \text{חולה}\right)+\underbrace{\mathbb{P}\left(\text{חיובי}\mid\text{בריא}\right)}_{\text{FalsePositive}}\mathbb{P}\left(\text{בריא}\right)}$$

אם נציב את הנתונים האמיתיים, נקבל:

$$\frac{0.9\cdot 0.01}{0.9\cdot 9.0.1+0.05+0.99}=0.15$$

נניח ודני לא מרוצה מהתוצאה, והולך לבדיקה נוספת. עכשיו ב־$\mathbb{P}\left(\text{דני חולה}\right)$ כבר לא נציב 0.01 אלא 0.15.

שאלה

אם נמשיך לעשות עוד ועוד בדיקות, לאיזה סיכוי נוכל להגיע?

נוסחת בייס לשלושה מאורעות

נוכל לרשום את הנוסחה הבאה:

$$\begin{aligned}\mathbb{P}\left(A\mid B\cap C\right) &= \frac{\mathbb{P}\left(A\cap B\cap C\right)}{\mathbb{P}\left(B\cap C\right)}=\frac{\mathbb{P}\left(B\mid A\cap B\right)\mathbb{P}\left(A\cap C\right)}{\mathbb{P}\left(B\mid C\right)\mathbb{P}\left(C\right)}\\ &= \frac{\mathbb{P}\left(B\mid A\cap C\right)\mathbb{P}\left(A\mid C\right)\mathbb{P}\left(C\right)}{\mathbb{P}\left(B\mid C\right)\mathbb{P}\left(C\right)}=\frac{\mathbb{P}\left(B\mid A\cap C\right)\mathbb{P}\left(A\mid C\right)}{\mathbb{P}\left(B\mid C\right)}\end{aligned}$$

שיעור מס' 10:

יום חמישי

$19.11.20$

תלות ואי תלות

כיצד ניתן לפרמל את הרעיון כי תופעות אקראיות אינן משפיעות זו על זו?

הרעיון הוא כזה: נשאל תחילה, מתי ב־$(\Omega,\mathbb{P})$ המאורע B לא משפיע על המאורע A?

למשל, אם $\Omega=\{H,T\}^2$ ו־$\mathbb{P}=\frac{1}{4}$, אזי מתי A ־ המאורע בו יצא H בהטלה הראשונה, ו־B ־ המאורע בו יצא H בהטלה השנייה, בלתי תלויים?

כאשר $\mathbb{P}\left(A\right)=\mathbb{P}\left(A\mid B\right)$!

הגדרה

נאמר כי A בלתי תלוי ב־B (או ש־A,B מאורעות בלתי תלויים) אם $\mathbb{P}\left(A\cap B\right)=\mathbb{P}\left(A\right)\mathbb{P}\left(B\right)$.[12]

הערה

אם A בלתי תלוי ב־B, אזי B בלתי תלוי ב־A.

בנוסף, אם A,B בלתי תלויים ו־$\mathbb{P}\left(B\right)>0$ אזי $\mathbb{P}\left(A\right)=\mathbb{P}\left(A\mid B\right)$.

טענה

[12] למעלה למעשה מדובר על מקרה בו $\mathbb{P}\left(B\right)\neq 0$, שהרי $\mathbb{P}\left(A\right)=\mathbb{P}\left(A\mid B\right)=\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$ וכן הלאה.

אם $\mathbb{P}(B)\in\{0,1\}$ אזי A,B בלתי תלוי לכל מאורע A.[13]

הוכחה

אם $\mathbb{P}(B)=0$, אזי מתקיים:

$$\mathbb{P}(A\cap B)\overset{\text{מונטוניות}}{\leq}\mathbb{P}(B)=0=\mathbb{P}(A)\mathbb{P}(B)$$

אם $\mathbb{P}(B)=1$, אזי $\mathbb{P}(B^c)=0$ ואז נקבל:

$$\mathbb{P}(A\cap B)\overset{\text{ס.אדיטביו}}{=}\mathbb{P}(A)-\mathbb{P}(A\cap B^c)=\mathbb{P}(A)-0=\mathbb{P}(A)\mathbb{P}(B)$$

דוגמאות

בהטלת קוביה הוגנת, ניקח את A להיות הטלה זוגית, ואת B להיות הטלה שמתחלקת ב־3.
נשים לב כי:

$$\mathbb{P}(A\cap B)=\mathbb{P}(\{6\})=\frac{1}{6}=\frac{1}{2}\cdot\frac{1}{3}=\mathbb{P}(\{2,4,6\})\mathbb{P}(\{3,6\})=\mathbb{P}(A)\mathbb{P}(B)$$

למעשהת דבר זה רומז על "משפט השאריות הסיני", שהרי מודולו 2 ו־מודולו 3 אינם מתקשרים. אם ניקח את C להיות הקבוצה שמתחלקת ב־5, נקבל כי A וC תלויים זה בזה. ־ אמנם ב־D_{10} הם יהיו בלתי תלויים.
אם ניקח את D_{10}, האם המאורעות A,B שראינו בלתי תלויים?
תחילה נחשב־ $\mathbb{P}(A)=\frac{1}{2}$. בנוסף מתקיים:

$$\mathbb{P}(A\mid B)=\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}=\frac{\mathbb{P}(\{6\})}{\mathbb{P}(\{3,6,9\})}=\frac{\frac{1}{10}}{\frac{3}{10}}=\frac{1}{3}$$

כלומר, הידיעה שהתוצאה התחלקה ב־3, השפיעה על תוצאת הניסוי.

טענה

אם A,B בלתי תלויים אזי גם A,B^c בלתי תלויים.

הוכחה

נוכל לשים לב שמתקיים:

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

מה הקשר של דבר זה לניסוי דו שלבי? נבחין כי מרחב המכפלה של הניסוי הדו שלבי הינו, $\Omega=\Omega_1\times\Omega_2$.
ראינו גם שמתקיים:

[13] "מאורע שתמיד קורה או לא קורה, לא יכול ללמד כלום".

$$\mathbb{P}_x(\omega_1,\omega_2)=\mathbb{P}_x(\omega,\cdot)\,\mathbb{P}_x((\cdot,\omega_2)\mid(\omega_1,\cdot))$$

אם הניסוים לא משפיעים זה על זה, אנחנו מניחים ש־

$$\mathbb{P}_x((\cdot,\omega_2)\mid(\omega_1,\cdot))=\mathbb{P}_x((\cdot,\omega_2))$$

ולכן בפרט מתקיים כי:

$$\mathbb{P}_x(\omega_1,\omega_2)=\mathbb{P}_x((\omega_1,\cdot))\cdot\mathbb{P}_x(\cdot,\omega_2)$$

כלומר, למעשה $(\omega_1,\cdot)$ ו־$(\cdot,\omega_2)$ הם מאורעות בלתי תלויים. כי הרי $\mathbb{P}_x(\omega_1,\omega_2)$ וזה בדיוק המקרה של הסתברות מכפלה.
ב'חיים האמיתיים', ניקח את Ω להיות ה'אנושות', ואת B להיות מעשן, ואת A להיות מת מסרטן. כלומר, נרצה לשאול האם הם משפיעים זה על זה? כלומר $\mathbb{P}(A\mid B)\overset{?}{=}\mathbb{P}(A)$.(אבל אמנם, קורלציה לא מוכיחה סיבתיות)

הערה
אם $\mathbb{P}(A\mid B)>\mathbb{P}(A)$ נאמר ש־B מאשש את A.
אם $\mathbb{P}(A\mid B)=\mathbb{P}(A)$ בלתי תלויים.
אם $\mathbb{P}(A\mid B)<\mathbb{P}(A)$ נאמר ש־B מחליש את A.

מה לגבי 3 מאורעות? ניקח את הדוגמה מקודם, כאשר $\Omega=\{H,T\}^2$ ו־$\mathbb{P}=\frac{1}{4}$.
ניקח את $A=(H,\cdot)$ ו־$B=(\cdot,H)$ וגם $C=\{\omega\mid\omega_1=\omega_2\}$.
אזי נקבל כי $\mathbb{P}(A)=\mathbb{P}(B)=\mathbb{C}=\frac{1}{2}$.
בסך הכל נקבל כי:

$$\mathbb{P}(A\cap B)=\frac{1}{4}=\mathbb{P}(A)\cdot\mathbb{P}(B)$$
$$\mathbb{P}(A\cap C)=\frac{1}{4}=\mathbb{P}(A)\cdot\mathbb{P}(C)$$
$$\mathbb{P}(B\cap C)=\frac{1}{4}=\mathbb{P}(B)\cdot\mathbb{P}(C)$$

האם אלו מאורעות בלתי תלויים? מאידך:

$$\mathbb{P}(A\cap B\cap C)=\frac{1}{4}\neq\frac{1}{8}=\mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C)$$

כמו כן, מתקיים:

$$\mathbb{P}(C \mid A \cap B) = 1 \neq \frac{1}{2} = \mathbb{P}(C)$$

לכן נאמר שהם תלויים. נאמר שהם בלתי תלויים בזוגות, כי כל זוג הוא בלתי זוג.

הצעה

נדרוש $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.
אבל זה לא עובד כי אם ניקח $A, B = (H, \cdot)$ ו־C בהטלה השנייה יצא לימון.
אז מתקיים כי $\mathbb{P}(A \cap B \cap C) = 0 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ אבל מאידך $\mathbb{P}(A \mid B) = 1 \neq \frac{1}{2} = \mathbb{P}(A)$.
לכן נצטרך הגדרה חזקה יותר.

הגדרה

מאורעות $A_1, \ldots, A_n$ נקראים בלתי תלויים אם:

$$\forall J \subseteq [n] \quad \mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$$

כלומר כל תתי הקבוצות צריכים לקיים את התנאי שראינו קודם לכן.

הגדרה נוספת

אם $B_1, \ldots, B_n$ מאורעות, וגם A, נאמר כי A ב״ת ב־$B_1, \ldots, B_n$ אם:

$$\forall J \subseteq [n] \quad \mathbb{P}\left(A \mid \bigcap_{i \in J} B_i\right) = \mathbb{P}(A)$$

תרגיל

אם $0 < \mathbb{P}(A_1 \cap \ldots \cap A_n)$ אזי $A_1, \ldots, A_n$ בלתי תלויים אם״ם לכל $1 \leq i \leq n$, המאורע A_i ב״ת בשאר $A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n$.

הגדרה

אוסף $\{A_i\}_{i=1}$ נקראת ב״ת אם $A_1, \ldots, A_n$ ב״ת לכל $J \leq I$ סופית.

תרגיל

סדרה $\{A_i\}_{i=1}^{\infty}$ נקראת ב״ת אם $A_1, \ldots, A_n$ ב״ת לכל $n \in \mathbb{N}$.

דוגמא

אם $(\Omega_1, \mathbb{P}_i)_{i=1}^{n}$ מרחבי הסתברות ־ $(\Omega, \mathbb{P})$ מרחב המכפלה, אזי לכל $\omega_1 \in \Omega_1, \ldots, \omega_n \in \Omega_n$ המאורעות: $(\omega_1, \ldots, \ldots), (\cdot, A_2, \cdot, \ldots), \ldots (\cdot, \ldots, A_n)$ הם בלתי תלויים.

טענה

אם A ב״ת ב־$B_1, \ldots B_n$ אזי הוא גם ב״ת ב־$B_1, \ldots B_n, B_1^c, \ldots B_n^c$.

הוכחה

מספיק להראות כי הוא ב"ת ב־ $B_1, \dots B_n, B_1^c$.

צ"ל למעשה ־ $\mathbb{P}(A) = \mathbb{P}\left(A \mid \bigcap_{i\in J} B_i\right)$ או $\mathbb{P}\left(\bigcap_J B_i\right) = 0$ לכל $J \subseteq [n]$ או

וגם $\mathbb{P}(A) = \mathbb{P}\left(A \mid B_1^c \cap \left(\bigcap_{i\in J} B_i\right)\right)$ או $\mathbb{P}\left(B_1^c \cap \bigcap_J B_i\right) = 0$ לכל $J \subseteq [n]$.

הראשון כבר נתון, לכן נותר לנו רק להוכיח את השני.

אם $1 \in J$ אזי $B_1^c \cap \bigcap_j B_i = \emptyset$.

ואז $\mathbb{P}\left(B_1^c \cap \bigcap_J B_i\right) = 0$.

אם $1 \notin J$, אזי:

$$\mathbb{P}\left(A \cap B_1^c \cap \bigcap_J B_i\right) = \mathbb{P}\left(A \cap \dots \bigcap_J B_i\right) - \mathbb{P}\left(A \cap B_i \cap \bigcap_J B_i\right)$$
$$\overset{(*)}{=} \mathbb{P}(A)\,\mathbb{P}\left(\bigcap_{i\in J} B_i\right) - \mathbb{P}(A)\,\mathbb{P}\left(B_1 \bigcap_{i\in A} B_i\right)$$
$$= \mathbb{P}(A)\left(\mathbb{P}\left(\bigcap_{i\in J} B_i\right)\right) - \mathbb{P}\left(B_1 \cap \bigcap_{i\in J} B_i\right)$$
$$= \mathbb{P}(A)\,\mathbb{P}\left(B_1^c \cap \bigcap_{i\in J} B_i\right)$$

ואז למעשה נקבל כי $\mathbb{P}\left(B_1^c \cap \bigcap_{i\in J} B_i\right) = 0$ או

$(*)$ A ב"ת ב־ $B_1, \dots B_n$.

משתנים מקריים, תוחלת והתפלגויות

שיעור מס' 11:
יום שלישי
24.11.20

משתנה מקרי X על מ"ה $\left(\Omega,2^{\Omega},\mathbb{P}\right)$ הוא פונקציה $X:\Omega\to\mathbb{R}$.

דוגמאות

גובה של ירושלמי אקראי. מס' הטלות H ב־n הטלות מטבע (הוגן). נקבל:

$$\Omega=\{H,T\}^{2}\quad\mathbb{P}=\frac{1}{2^{n}}\quad X\left(\omega\right)=\left|\left\{ 1\leq i\leq n\mid\omega_{i}=H\right\} \right|$$

סכום 2 קוביות (הוגנות). נשים לב כי ישנם שני Ω הגיוניים:

$$\Omega=[6]^{2}\quad\mathbb{P}=\frac{1}{36}\quad X\left(\omega\right)=\omega_{1}+\omega_{2}$$

משתנים מקריים (מ"מ) מגדירים מאורעות:
אם X מ"מ על מ"ה $(\Omega,\mathbb{P})$ אזי אפשר להסתכל על ־ $\{X=7\}=\{\omega\in\Omega\mid X\left(\omega\right)=7\}$
או למשל $\{X\leq10\}$ וגם $\{x\in\{2,5,14\}\}$ וכן הלאה.
נוכל גם לקחת את $\{X=Y\}:=\{\omega\in\Omega\mid X\left(\omega\right)=Y\left(\omega\right)\}$.
למשל, אם נטיל 2 קוביות (הוגנות). יהי X סכומן ו־Y מכפלתן, נקבל:

$$\{X=4\}=\{\left(1,3\right),\left(2,2\right),\left(3,1\right)\}$$
$$\{X=Y\}=\{\left(2,2\right)\}$$

יתרה מכך, נוכל לקצר זאת עוד יותר ולסמן:

$$\mathbb{P}\left(X=7\right)=\mathbb{P}\left(\{X=7\}\right)=\mathbb{P}\left(\{\omega\in\Omega\mid X\left(\omega\right)=7\}\right)=\mathbb{P}\left(X^{-1}\left(7\right)\right)$$

(כי הרי אם $f:S\to T$ פונקציה, אזי $f^{-1}\left(t\right)=\{s\in S\mid f\left(s\right)=t\}$. וגם $f^{-1}\left(A\right)=\{s\in S\mid f\left(s\right)\in A\}$.

דוגמא

יהי X סכום קוביות הוגנות. $\mathbb{P}\left(X=4\right)=\frac{3}{36}$ (בין אם כ־$\frac{3}{36}$ או כ־$\frac{1}{8}+\frac{1}{36}$).

ניתן לעשות גם פעולות על מ"מ.
אם $X,Y:\Omega\to\mathbb{R}$ מ"מ על אותו מ"ה, אפשר לעשות עליהם אריתמטיקה, למשל $X+Y,X\cdot Y,X-Y,3X+5Y$.
למשל, $Z=X+Y$ הוא המ"מ המוגדר על ידי $Z\left(\omega\right)=X\left(\omega\right)+Y\left(\omega\right)$

דוגמא

המזכיר המבולבל מכניס n מכתבים ל־n מעטפות בסדר אקראי.
יהי X מס' המכתבים שמועננו נכון. לכל $1\leq i\leq n$ נגדיר:

$$Y_i\left(\omega\right)=\begin{cases}1 & \omega\left(i\right)=i\\0 & else\end{cases}$$

כלומר, Y_i לפי הוא 1 או 0 האם המכתב ה־i מוען נכונה.
אזי־ $X=Y_1+Y_2+\ldots+Y_n$.
אבל נשים לב כי כל Y_i הוא די פשוט:

$$\mathbb{P}\left(Y_i=1\right)=\frac{1}{n}\quad\mathbb{P}\left(Y_i=0\right)=1-\frac{1}{n}=\frac{n-1}{n}$$

בנוסף, נשים לב כי:

$$X\left(\omega\right)=Y_1\left(\omega\right)+\ldots+Y_n\left(\omega\right)\quad\forall\omega\in S_n$$

כאשר $X=\left(\omega\right)$ הוא מספר ה־i כך ש־$\omega\left(i\right)=i$, ו־$Y_1\left(\omega\right)$ הוא האינדיקטור שבודק האם $\omega\left(i\right)=i$ ו־0 אחרת, וכן הלאה.

אינדיקטורים

אם $\left(\Omega,\mathbb{P}\right)$ מ"ה ו־A מאורע. האינדיקטור של A הוא המ"מ $\mathbb{1}_A:\Omega\to\mathbb{R}$ המוגדר על ידי $\mathbb{1}_A=\begin{cases}1 & \omega\in A\\0 & else\end{cases}$.
למשל, בבעיית המזכיר המבולבל:
$Y_i=\mathbb{1}_{A_i}$ כאשר A_i הוא המאורע שמכתב i נשלח נכון ־ $A_i=\left\{\omega\in S_n\mid\omega\left(i\right)=i\right\}$.

נבחין כי כל מ"מ משרה מרחב הסתברות על $\mathbb{R}$ (כמרחב מדגם). אם X מ"מ על $\left(\Omega,\mathbb{P}\right)$, נגדיר מ"ה חדש: $\left(\mathbb{R},2^{\mathbb{R}},\mathbb{P}_X\right)$, לדוגמא:

$$\mathbb{P}_X\left(7\right)=\mathbb{P}\left(X=7\right)=\mathbb{P}\left(\left\{\omega\in\Omega\mid X\left(\omega\right)=7\right\}\right)=\mathbb{P}\left(X^{-1}\left(7\right)\right)$$

ובאופן כללי מתקיים:

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

נוודא ש־ $\left(\mathbb{R},\mathbb{P}_X\right)$ הוא אכן מרחב הסתברות:

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

ואם $A_i\subset\mathbb{R}$ זרות בזוגות אזי:

$$\mathbb{P}_X\left(\bigcup_{i=1}^{\infty}A_i\right)=\mathbb{P}\left(X^{-1}\bigcup_{i=1}^{\infty}A_i\right)\overset{\text{מקורות זרים}}{=}\mathbb{P}\left(\bigcup_{i=1}^{\infty}X^{-1}\left(A_i\right)\right)\overset{\text{ס.אדיטיביות}}{=}$$
$$\sum_{i=1}^{\infty}\mathbb{P}\left(X^{-1}\left(A_i\right)\right)=\sum\mathbb{P}_X\left(A_i\right)$$

$\mathbb{P}_X$ נקרא ההתפלגות של X.

הגדרה

X נקרא משתנה מקרה בדיד אם $\mathbb{P}_X$ היא פונקציית הסתברות בדידה (ראינו 4 קריטריונים שקולים לזה).

דוגמא

אם $(\Omega,\mathcal{F},\mathbb{P})$ מתאר הגרלת מספר אחיד בקטע $[0,1]$ ו־X הוא הספרה השלישית אחרי הנקודה אזי $(\mathbb{P},\mathcal{F},\mathbb{P})$ לא מ"ה בדיד, אבל X כן מ"מ בדיד.

ההתפלגות הנקודתית של X היא $\mathbb{P}_X\left(\alpha\right)=\begin{cases}\frac{1}{10} & \alpha=0,\ldots,9\\ 0 & else\end{cases}$

למשל

יהי $(\Omega,\mathbb{P})$ הניסוי של 2 הטלות מטבע הוגן. יהי X מספר הראשים ו־Y מס' הזנבות, אזי $\Omega=\{H,T\}^2$, ולכן $X\left(\omega\right)=\left|\{1\leq i\leq 2\mid\omega_i=H\}\right|$.

נשים לב כי $Y=2-X$.

ממילא נקבל כי:

$$\mathbb{P}_X\left(\alpha\right)=\begin{cases}\frac{1}{4} & \alpha=0,\ldots,2\\ \frac{1}{2} & \alpha=1\\ 0 & else\end{cases}$$

אם כך, פונקציית ההתפלגות הינה - $\mathbb{P}_X\left(s\right)=\sum_{\alpha\in S}p_X\left(\alpha\right)$.

נוכל גם להבחין כי אם היינו מחשבים את ההתפלגות של Y, היא הייתה אותו הדבר. אם כך, נאמר ש־X ו־Y שווי התפלגות, אבל הם לא שווים.

<u>סימונים:</u>

ל־2 מ"מ X,Y על אותו מ"ה $(\Omega,\mathbb{P})$ מסמנים:

$X=Y$: $\forall\omega\in\Omega:X\left(\omega\right)=Y\left(\omega\right)$

$X\overset{d}{=}Y$: X,Y **שווי התפלגות:** $\forall s\subseteq\mathbb{R}:\mathbb{P}_X\left(s\right)=\mathbb{P}_Y\left(s\right)$

$X\overset{a,s}{=}Y$: X,Y **שווים כמעט תמיד** $\mathbb{P}\left(X=Y\right)=1$.

למשל, אם $\Omega=\{H,T\}$ ו־$p\left(H\right)=1,p\left(T\right)=0$.

ניקח $X\left(H\right)=Y\left(H\right)$, אזי קיבלנו כי $X\left(T\right)=9\neq 8=Y\left(T\right)$.

אזי $X\overset{a,s}{=}Y$.

הרבה פעמים נתעניין רק בהתפלגות של מ"מ ולא באיזה ω הוא מקבל מה. לכן התפלגות זה מושג מרכזי.

תוחלת

שיעור מס' 12:
יום חמישי
26.11.20

תוחלת היא הגרסה ההסתברותית לממוצע. הממוצע של המון הטלות קובייה יוצא בערך 3.5. "במטבע מוטה של $\left(\frac{2}{3},\frac{1}{3}\right)$, מס' הראשים ב־$n$ הטלות יהיה בערך $\frac{2}{3}n$ כש־n מאוד גדול."

הגדרה

התוחלת של מ"מ מקרי בדיד X היא:

$$\mathbb{E}\left(X\right)=\sum_{\alpha\in\mathbb{R}}\alpha\cdot\mathbb{P}_{X}\left(\alpha\right)$$

אם הטור מתכנס בהחלט (כולל התכנסות ל־∞ או $-\infty$).

אם הטור לא מתכנס בהחלט, נאמר ש־X חסר תוחלת.[14]

אבחנה

אם X מ"מ על $\left(\Omega,\mathbb{P}\right)$ מ"ה בדיד, אזי:

$$\mathbb{E}\left(X\right)=\sum_{\alpha\in\mathbb{R}}\alpha\mathbb{P}_{X}\left(\alpha\right)=\sum_{\alpha\in\mathbb{R}}\alpha\sum_{\substack{\omega\in\Omega,\\ X\left(\omega\right)=\alpha}}\mathbb{P}\left(\omega\right)=$$

$$\overset{\text{שינוי סדר הסכימה}}{=}\sum_{\omega\in\Omega}X\left(\omega\right)\mathbb{P}\left(\omega\right)$$

לפעמים דבר שימושי אבל לעיתים יותר קרובות נוח לשכוח את Ω ולדעת שאפשר להסיק את $\mathbb{E}\left(X\right)$ מהתפלגות X.

טענה

אם X ו־Y שווי התפלגות, אזי $\mathbb{E}\left(X\right)=\mathbb{E}\left(Y\right)$.

(X ו־Y גם יכולים להיות על מ"ה שונים)

הוכחה

ההגדרה לא השתמשה ב־$\left(\Omega,\mathbb{P}\right)$ אלא רק ב־$\mathbb{P}_{X}$ (ההתפלגות).

התפלגויות נפוצות

ברנולי

נאמר שמ"מ X מתפלג "ברנולי עם סיכוי p" אם:

$$\mathbb{P}_{X}\left(1\right)=0\quad\mathbb{P}_{X}\left(0\right)=1-p$$

נסמן זאת בתור $X\sim\text{Ber}\left(p\right)$

[14] כיוון ש־X בדיד $0<\mathbb{P}_{X}\left(\alpha\right)$ רק עבור $\alpha\leq\aleph_{0}$ שונות.

"מספר ההצלחות ב־n חזרות".

טענה

אם $X\sim\text{Ber}\left(p\right)$ אזי $\mathbb{E}\left(X\right)=p$

הוכחה

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

גיאומטרית

נאמר ש־X מתפלגת "גיאומטרית עם סיכוי p" ונסמן $X\sim Geo\left(p\right)$ אם:

$$\mathbb{P}_{X}=\begin{cases}p\left(1-p\right)^{n-1} & n\in\mathbb{N}\\ 0 & else\end{cases}$$

דוגמא

נתבונן בניסוי של סדרת הטלות מטבע מוטה p שאינן משפיעות זו על זו (סדרה אינסופית), ויהי X־ באיזה הטלת H הראשונה ו־27 אם לא התקבל H.

פורמלית, נקבל כי $\Omega=\left\{H,T\right\}^{\mathbb{N}}$. בנוסף $\forall i\ \ \mathbb{P}\left(\underbrace{\left\{\omega\mid\omega_{i}=H\right\}}_{A_{i}}\right)=p$, והמאורעות A_{i} בלתי תלויים.

בנוסף, נקבל:

$$X\left(\omega\right)=\begin{cases}\min\left\{i\mid\omega_{i}=H_{i}\right\} & H\in\omega\\ 27 & H\notin\omega\end{cases}$$

אזי $X\sim\text{Geo}\left(p\right)$:

$$\begin{aligned}\mathbb{P}_{X}\left(n\right)=\mathbb{P}\left(A_{1}^{c}\cap A_{2}^{c}\cap\ldots\cap A_{n-1}^{c}\cap A_{n}\right)=\\ \mathbb{P}\left(A_{1}^{c}\right)\mathbb{P}\left(A_{2}^{c}\right)\ldots\mathbb{P}\left(A_{n-1}^{c}\right)\mathbb{P}\left(A_{n}^{c}\right)=\\ \left(1-p\right)\left(1-p\right)\cdot\ldots\cdot\left(1-p\right)=p\left(1-p\right)^{n-1}\end{aligned}$$

טענה

אם $X\sim\text{Geo}\left(p\right)$ אזי $\mathbb{E}\left(X\right)=\frac{1}{p}$.

"בממוצע צריך לחכות 6 הטלות בכדי לקבל 5 בקובייה".(מיקום ההצלחה הראשונה)

הוכחה

$$\mathbb{E}\left(X\right)=\sum_{n=1}^{\infty}n\mathbb{P}_{X}\left(n\right)=p\sum_{n\geq1}n\left(1-p\right)^{n-1}$$

כעת, נסמן $q=1-p$ ונקבל:

$$p\sum_{n\geq1}n\left(1-p\right)^{n-1}=p\sum_{n\geq1}nq^{n-1}\overset{\text{כפל לסכום}}{=}p\sum_{n=1}^{\infty}\sum_{m=1}^{n}q^{n-1}\overset{\text{שס"ס}}{=}$$
$$p\sum_{m=1}^{\infty}\sum_{n\geq m}^{\infty}q^{n-1}\overset{(*)}{=}p\sum_{m=1}^{\infty}q^{m-1}\frac{1}{1-q}=\sum_{m=1}^{\infty}q^{m-1}=\frac{1}{1-q}=\frac{1}{p}$$

$(*)$ סכום סדרה הנדסית.

נאמר ש־X מתפלג בינומית (n,p) (כש־$n\in\mathbb{N}$ ו־$0\leq p\leq1$)

אם $\mathbb{P}_X\left(k\right)=\binom{n}{k}p^k\left(1-p\right)^{n-k}$. נסמן $X\sim\text{Bin}\left(n,p\right)$.

דוגמא

מס' הטלות H ב־n הטלות מטבע מוטה p.

נאמר ש־X מתפלג פואסון עם תוחלת λ אם:

$$\forall n\in\mathbb{N}_0\ \ \mathbb{P}_X\left(n\right)=\frac{e^{-\lambda}\lambda^n}{n!}$$

נסמן $X\sim\text{Poi}\left(\lambda\right)$.

"זמן עד ההצלחה הראשונה".

טענה

אם $X\sim$ **Poi** $\left(\lambda\right)$ **אזי** $\mathbb{E}\left(X\right)=\lambda$.

הוכחה

$$\mathbb{E}\left(X\right)=\sum_{n=0}^{\infty}\frac{ne^{-\lambda}\lambda^n}{n!}\overset{(*)}{=}e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}\lambda}{(n-1)!}\overset{(**)}{=}$$
$$\lambda e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\overset{(***)}{=}\lambda e^{-\lambda}e^{\lambda}=\lambda$$

$(*)$ ביטול של n והוצאת קבוע החוצה.

$(**)$ הוצאת λ החוצה על מנת להגיע לסכימה מ־0.

$(***)$ נוסחה כלשהיא.

שיעור מס' 13:
יום שלישי
1.12.20

טענה (פואסון כגבול של בינומי במובן נקודתי) - "בינומי שואף לפואסון"

אם $Y\sim\text{Poi}\left(\lambda\right)$ עבור $\lambda\geq0$. ו־$\{X_n\}$ משתנים מקריים, כך שלכל $n>\lambda$ מתקיים כי $X_n\sim\left(n,\frac{\lambda}{n}\right)$. אזי לכל $k\in\mathbb{N}_0$ מתקיים:

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

הוכחה

נשים לב כי מתקיים, עבור k קבוע, ו־n שואף לאינסוף:

$$\left(\begin{array}{c} n \\ k \end{array}\right)\left(\frac{\lambda}{n}\right)^{k}=\frac{n\cdot(n-1)\cdots(n-k+1)}{n^{k}}\frac{\lambda^{k}}{k}\cdot\underbrace{\left(1-\frac{1}{n}\right)^{n}}_{e^{-\lambda}}\underbrace{\left(1-\frac{1}{n}\right)^{-k}}_{\to 1}=$$

$$\underbrace{\frac{n\cdot(n-1)\cdots(n-k+1)}{k!}}_{1}\frac{\lambda^{k}}{n^{k}}\cdot\underbrace{\left(1-\frac{1}{n}\right)^{n}}_{e^{-\lambda}}\underbrace{\left(1-\frac{1}{n}\right)^{-k}}_{\to 1}\underset{n\to\infty}{\to}\frac{\lambda^{k}\cdot e^{-\lambda}}{k!}$$

משתנים מותנים

אם X מ"מ על $(\Omega,\mathbb{P})$ ו־A מאורע עם $\mathbb{P}(A)\geq 0$, אזי אפשרי להתנות את X ב־B.

נסמן:

$$\begin{aligned}
&\mathbb{P}_{X|B}(S)=\mathbb{P}(x\in S\mid B)=\\
&\mathbb{P}(\{\omega\in\Omega\mid X(\omega\in S)\}\mid B)=\\
&\frac{\mathbb{P}(\{x\in S\}\cap B)}{\mathbb{P}(B)}
\end{aligned}$$

דבר זה מתלכד עם $\mathbb{P}_B(x\in S)$.

ניתן גם לחשוב על $(X\mid B)$ כהתפלגות. למשל $(X\mid B)\sim Bin(n,p)$ פירושו כי $\mathbb{P}_{X|B}(k)=\binom{n}{k}p^kq^{n-k}$.

דוגמא או טענה ־ חוסר הזיכרון של משתנה גיאומטרי [15]

אם $X\sim \text{Gep}(p)$ אזי $\mathbb{P}(X=n\mid X>m)=\mathbb{P}(X=n-m)$.

במילים אחרות, $(X-m\mid X>m)\sim \text{Gep}(p)$.

"אם עשיתי אלף ניסויים, ההסתברות שאצטרך 1007 היא כמו ההסתברות שאצטרך 7 ניסויים בלתי תלויים".

הוכחה

$$\begin{aligned}
&\mathbb{P}(X=n\mid X>m)=\frac{\mathbb{P}(\{X=n\}\cap\{X>m\})}{\mathbb{P}(X>m)}\overset{n\geq m}{=}\\
&\frac{\mathbb{P}(X=n)}{1-\mathbb{P}(X\leq m)}=\frac{p\cdot q^{n-1}}{1-(p+pq+\ldots+pq^{n-1})}=\\
&\frac{pq^{n-1}}{1-p\frac{1-q^n}{1-p}}=pq^{n-m-1}=\mathbb{P}(X=n-m)
\end{aligned}$$

ולכן בפרט מתקיים כי:

$$\begin{aligned}
&\mathbb{P}(X-m=n\mid X>m)=\\
&\mathbb{P}(X=m+n\mid X>m)=pq^{n-1}=\mathbb{P}(X=n)
\end{aligned}$$

[15] עמוד 100 בספר

גם ההיפך נכון - "אם X נתמך על $\mathbb{N}$ וחסר זיכרון, אזי $X \sim \text{Geo}\left(p\right)$"

שיעור מס' 14:
יום חמישי
3.12.20

טענה
אם X מ"מ שנתמך (ההסתברות ש-X שייך לטבעיים היא 1) על $\mathbb{N}$ וחסר זיכרון, אזי X מתפלג גיאומטרית.
הוכחה
נבחין כי:

$$\mathbb{P}\left(X>n\right)=\mathbb{P}\left(X>1\right)\cdot\mathbb{P}\left(X>2\mid X>1\right)\ldots\mathbb{P}\left(X>n\mid X>n-1\right)$$

(דבר זה נכון כי $\mathbb{P}\left(\bigcap_{i=1}^{n}A_i\right)=\mathbb{P}\left(A_1\right)\cdot\mathbb{P}\left(A_2\mid A_1\right)\ldots\cdot\mathbb{P}\left(A_3\mid A_1\cap A_2\right)$)

אבל:

$$\mathbb{P}\left(X>2\mid X>1\right)=\sum_{n=3}^{\infty}\mathbb{P}\left(X=n\mid X>1\right)\overset{\text{חוסר זיכרון}}{=}\sum_{n=3}^{\infty}\mathbb{P}\left(X=n-1\right)=\mathbb{P}\left(X>1\right)$$

ולכן בפרט:

$$\mathbb{P}\left(X>1\right)\cdot\mathbb{P}\left(X>1\right)\cdot\ldots\mathbb{P}\left(X>1\right)=\mathbb{P}\left(X>1\right)^{n}$$

נגדיר כעת $q=\mathbb{P}\left(X>1\right)$ ו-$p=1-q$ ונקבל:

$$\mathbb{P}\left(X=n\right)=\mathbb{P}\left(X>n-1\right)-\mathbb{P}\left(X>n\right)=q^{n-1}-q^{n}=pq^{n-1}=\mathbb{P}$$

הגדרה
התוחלת המותנה של מ"מ X בדיד בהינתן מאורע A עם $\mathbb{P}\left(A\right)>0$ היא:

$$\mathbb{E}\left(X\mid A\right)=\sum_{\alpha\in\mathbb{R}}\alpha\cdot\mathbb{P}\left(X=\alpha\mid A\right)=\sum_{\alpha\in\mathbb{R}}\alpha\cdot\mathbb{P}_{X|A}\left(\alpha\right)$$

נוסחת התוחלת השלימה

אם $\{A_i\}$ חלוקה (אוסף של מאורעות זרים בזוגות שאיחודם הוא Ω) בת מנייה של Ω אזי:

$$\mathbb{E}\left(X\right)=\sum_{i}\mathbb{P}\left(A_{i}\right)\mathbb{E}\left(X\mid A_{i}\right)$$

(כאשר לא מוגדר כפול 0 שווה ל־0)

הוכחה

נבחין כי:

$$\begin{aligned}&\mathbb{E}\left(X\right)=\sum_{\alpha\in\mathbb{R}}\alpha\cdot\mathbb{P}_{X}\left(\alpha\right)\overset{\text{נה"ש}}{=}\\&\sum_{\alpha}\alpha\sum_{i}\mathbb{P}\left(A_{i}\right)\mathbb{P}\left(X=\alpha\mid A_{i}\right)\overset{(*)}{=}\\&\sum_{i}\mathbb{P}\left(A_{i}\right)\sum_{\alpha}\alpha\mathbb{P}\left(X=\alpha\mid A_{i}\right)=\\&\sum_{i}\mathbb{P}\left(A_{i}\right)\cdot\mathbb{E}\left(X\mid A_{i}\right)\end{aligned}$$

$(*)$ מותר לשנות סדר סכימה כי הטור מתכנס בהחלט לפי הגדרת התוחלת.

דוגמא

X= תוצאת הטלת קוביות $D\&D$ אקראית.

A_m הוא המאורע "בחרנו בקוביה D_m" ואז $A_4,A_6,A_8,A_{10},A_{12},A_{20}$ היא חלוקה של Ω ולכן:

$$\mathbb{E}\left(X\right)=\sum_{m}\mathbb{P}\left(A_{m}\right)\mathbb{E}\left(X\mid A_{m}\right)=\sum_{m}\frac{1}{6}\cdot\frac{1+m}{2}$$

(בדקו והבינו)

כמה משתנים מקריים

אם X,Y מ"מ על מ"ה $(\Omega,\mathbb{P})$ לא מספיק להבין את $\mathbb{P}_X,\mathbb{P}_Y$ כדי להבין דברים שתלויים בשני המשתנים (למשל $Z=X+Y$).

לדוגמא, אם X מספר הראשים בשתי הטלות מטבע ו־Y מספר הזנבות, אזי $\mathbb{P}_X=\mathbb{P}_Y=\begin{cases}\frac{1}{2} & 1\\ \frac{1}{4} & 0,2\end{cases}$ ו־$\mathbb{P}_Z\left(2\right)=\mathbb{P}_{X+Y}\left(2\right)=1$.

לעומת זאת, אם ניקח 4 הטלות המטבע, ו־X יהיה מספר הראשים מבין 2 ההטלות הראשונות, ו־Y מספר הזנבות מבין 2 ההטלות האחרונות, אזי עדיין $\mathbb{P}_X=\mathbb{P}_Y=\begin{cases}\frac{1}{2} & 1\\ \frac{1}{4} & 0,2\end{cases}$ אבל עכשיו $Z\sim\text{Bin}\left(4,\frac{1}{2}\right)$.

הגדרה

עבור מ"מ X,Y על אותו מ"ה $(\Omega,\mathbb{P})$, **ההתפלגות המשותפת** שלהם היא:

$$\mathbb{P}_{X,Y}\left(a,b\right)=\mathbb{P}\left(X=a\wedge Y=b\right)=\mathbb{P}\left(\left\{\omega\in\Omega\mid X\left(\omega\right)\right\},Y\left(\omega\right)=b\right)$$

באופן כללי יותר, ניתן לכתוב:

$$\mathbb{P}_{X,Y}\left(S,T\right)=\mathbb{P}\left(X\in S\wedge Y\in T\right)$$

תרגיל: $P_{X,Y}$ היא פונקציית הסתברות על $\mathbb{R}^{2}$.

דוגמא

נטיל 2 קוביות וניקח את X להיות ההטלה הגדולה מביניהן ו־Y הקטנה מביניהן.

מתוך ההתפלגות המשותפת $\mathbb{P}_{X,Y}$ אפשר לשחזר את $\mathbb{P}_{X}$ ואת $\mathbb{P}_{Y}$, כי מתקיים: $\mathbb{P}_{X}\left(S\right)=\mathbb{P}_{X,Y}\left(S,\mathbb{R}\right)$.
למעשה, $\mathbb{P}_{X},\mathbb{P}_{Y}$ נקראות ההתפלגויות השוליות של ההתפלגות המשותפת $\mathbb{P}_{X,Y}$.
באופן כללי יותר, אם $X_{1},\ldots X_{m}$ מ"מ על $\left(\Omega,\mathbb{P}\right)$, נגדיר את ההתפלגות המשותפת שלהם:

$$\mathbb{P}_{X_{1},\ldots X_{n}}\left(S_{1},\ldots S_{n}\right):=\mathbb{P}\left(\forall i:X_{i}\in S_{i}\right)$$

זאת פונקצית הסתברות על $\mathbb{R}^{n}$.
ניתן לחשוב על $X_{1},\ldots X_{n}$ כמשתנה מקרי אחד שמקבל ערכים ב־$\mathbb{R}^{n}$.
כלומר ניקח $X:\Omega\to\mathbb{R}^{n}$ על ידי $X\left(\omega\right)=\left(X_{1}\left(\omega\right),X_{2}\left(\omega\right),\ldots,X_{n}\left(\omega\right)\right)$.
X נקרא מ"מ וקטורי, ואז:

$$\mathbb{P}_{X}\left(\overrightarrow{v}\right)=\mathbb{P}_{X_{1},\ldots X_{n}}\left(v_{1}\ldots v_{n}\right)$$

ונקבל כי:

$$\mathbb{P}_{X}\left(S_{1}\times S_{2}\times\ldots\times S_{n}\right)=\mathbb{P}_{X_{1},X_{2},\ldots X_{n}}\left(S_{1},\ldots,S_{n}\right)=$$
$$\mathbb{P}\left(X^{-1}\left(S_{1}\times S_{2}\times\ldots\times S_{n}\right)\right)$$

מה לגבי $\mathbb{P}_{X}\left(A\right)$ כללית?

טענה

$X=\left(X_{1},\ldots,X_{n}\right)$ הוא בדיד (כלומר $\mathbb{P}_{X}$ היא פונקצית הסתברות בדידה על $\mathbb{R}^{n}$) אם"ם כל X_{i} בדיד.
מסקנה
לכל $\left(\sum_{\overrightarrow{v}\in A}\mathbb{P}_{X_{1},\ldots,X_{n}}\left(v_{1},\ldots,v_{n}\right)\right)=\sum_{\overrightarrow{v}\in A}\mathbb{P}_{X}\left(\overrightarrow{v}\right)=\mathbb{P}_{X}\left(A\right)$.
(תסתכלו בדוגמא 80 בספר)

נחזור לבעיית הסכום:
אם X,Y מ"מ בדידי. וניקח $Z=X+Y$. מתקיים כי:

$$\mathbb{P}_Z(\gamma) = \sum_{\substack{\alpha,\beta\in\mathbb{R}^2 \\ \alpha+\beta=\gamma}} \underbrace{\mathbb{P}_{X,Y}(\alpha,\beta)}_{(*)}$$

$(*)$ אומר כי דבר זה שווה ל־0 מלבד מספר בן מנייה של $(\alpha,\beta)\in\mathbb{R}^2$.

תרגיל

אם $X_1,\dots,X_n$ ו־$Y_1,\dots,Y_n$ מ"מ, אזי $X \overset{a.s}{=} Y$ (כמ"מ וקטוריים) אם"ם לכל i יתקיים כי $X_i \overset{a.s}{=} Y_i$. מה לגבי $X \overset{d}{=} Y$ ו־$X_i \overset{d}{=} Y_i$?

אי תלות של משתנים מקריים

נאמר שמ"מ X,Y (על אותו מ"ה) הם ב"ת. כלומר $\forall S,T\subseteq\mathbb{R}$, המאורעות $\{X\in S\},\{Y\in T\}$ בלתי תלויים.
למעשה, מתקיים:

$$\begin{aligned}&\mathbb{P}_{X,Y}(S,T) = \mathbb{P}(\{X\in S\}\cap\{Y\in T\}) = \\ &\mathbb{P}(X\in S)\,\mathbb{P}(Y\in T) = \mathbb{P}_X(S)\,\mathbb{P}_Y(T)\end{aligned}$$

כלומר $\mathbb{P}_{X,Y}$ היא בדיוק הסתברות המכפלה של $\mathbb{P}_X$ו־$\mathbb{P}_Y$.
(לסיכום, X,Y ב"ת אם ההתפלגות שלהם היא מכפלת ההסתברויות השוליות שלהם).

הגדרה

$X_1,\dots X_n$ הם ב"ת אם לכל $S_1,\dots,S_n\subseteq\mathbb{R}$ המאורעות $\{X\in S_1\},\dots\{X_n\in S_n\}$ ב"ת.

דבר זה שקול לכך ש־$\mathbb{P}_{X_1,\dots X_n}$ היא מכפלת ההתפלגויות $\mathbb{P}_{X_1}\times\dots\mathbb{P}_{X_n}$.
הכיוון הראשון קל, כיוון שאם $X_1,\dots X_n$ ב"ת, אזי:

$$\begin{aligned}&\mathbb{P}_X(S_1\times\dots\times S_n) = \mathbb{P}(X_1\in S\wedge\dots X_n\in S_n) = \\ &\prod \mathbb{P}(X_i\in S_i) \overset{\text{הגדרה}}{=} (\mathbb{P}_{X_1}\times\dots\mathbb{P}_{X_n})(S_1\times\dots\times S_n)\end{aligned}$$

אם ההתפלגות המשותפת היא מכפלת ההסתברויות השוליות, אזי:

$$\forall I \subseteq [n] \qquad \mathbb{P}\left(\bigcap_{i\in I}\{X_i \in S_i\}\right) =$$
$$\mathbb{P}\left(\bigcap_{i\in I}\{X_i \in S_i\} \cap \bigcap_{i\notin I}\{X_i \in \mathbb{R}\}\right) =$$
$$\mathbb{P}_X\left(S_1 \times S_2 \times \mathbb{R} \times \ldots \times S_n\right) =$$
$$\prod_{i\in I}\mathbb{P}\left(X_i \in S_i\right)\underbrace{\prod_{i\notin I}\mathbb{P}\left(X_i \in S_i\right)}_{1} =$$
$$\prod_{i\in I}\mathbb{P}\left(X_i \in S_i\right)$$

כלומר, $\{X_i \in S_i\}$ אם ב"ת ומכיון שזה נכון לכל $S_1, \ldots S_n \subseteq \mathbb{R}$, אזי $X_1, \ldots X_n$ ב"ת לפי ההגדרה.

טענות על תוחלת

שיעור מס' 16: **יום חמישי** 8.12.20

טענה

יהיו X, Y מ"מ ב"ת (על אותו מ"ה) אזי:

$$\mathbb{E}\left(X+Y\right) = \mathbb{E}\left(X\right) + \mathbb{E}\left(Y\right)$$
$$\mathbb{E}\left(XY\right) = \mathbb{E}\left(X\right) \cdot \mathbb{E}\left(Y\right)$$

הוכחה

נוכיח עבור מ"מ בדידים.

נשים לב כי מתקיים:

$$\mathbb{E}\left(X+Y\right) = \sum_{\alpha\in\mathbb{R}}\alpha\left(X+Y=\alpha\right) =$$
$$\sum_{\alpha}\alpha\sum_{\beta}\mathbb{P}_{X,Y}\left(\beta, \alpha-\beta\right) \overset{(*)}{=}$$
$$\sum_{\substack{\beta\in\mathbb{R}\\ \gamma\in\mathbb{R}}}\left(\beta+\gamma\right)_{X,Y}\left(\beta,\gamma\right) \overset{\text{ב"ת}}{=}$$
$$\sum_{\beta,\gamma}\left(\beta+\gamma\right)_X\left(\beta\right)\mathbb{P}_Y\left(\gamma\right) =$$
$$\sum_{\beta}\beta\mathbb{P}_X\underbrace{\sum_{\gamma}\mathbb{P}_Y\left(\gamma\right)}_{1} + \sum_{\gamma}\gamma\mathbb{P}_Y\left(\gamma\right)\cdot\underbrace{\sum_{\beta}\mathbb{P}_X\left(\beta\right)}_{1} =$$
$$\mathbb{E}\left(X\right) + \mathbb{E}\left(Y\right)$$

$(*)$ נגדיר $\gamma = \alpha - \beta$ (סכום בן מנייה מתכנס בהחלט)

בצורה דומה לגבי המכפלה:

$$\mathbb{E}\left(XY\right) = \sum_{\beta,\gamma} \beta\gamma\mathbb{P}_{X,Y}\left(\beta,\gamma\right) \overset{\text{ב"ת}}{=}$$
$$\sum_{\beta,\gamma} \beta\gamma\mathbb{P}_{X}\left(\beta\right)\mathbb{P}_{Y}\left(\gamma\right) = \sum_{\beta} \beta\mathbb{P}_{X}\left(\beta\right)\sum_{\gamma}\gamma\mathbb{P}_{Y}\left(\gamma\right) =$$
$$\sum_{\beta} \beta\mathbb{P}_{X}\left(\beta\right)\mathbb{E}\left(Y\right) = \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)$$

דוגמא

נוכיח שאם $X \sim \text{Bin}\left(n,p\right)$ אזי $\mathbb{E}\left(X\right) = np$.

הוכחה

נתבונן ב־n הטלות מטבע מוטה p־ב"ת, ויהא A_i ־ הצלחה בהטלה ה־i ו־Y מספר ההצלחות.

אזי $Y = 1_{A_i} + \ldots + 1_{A_n}$ ו־ $1_{A_i}\left(\omega\right) = \begin{cases} 1 & \omega_i = H \\ 0 & \omega_i = T \end{cases}$.

הוכחנו כבר ש־$Y \sim \text{Bin}\left(n,p\right)$ ועכשיו מתקבל כי:

$$\mathbb{E}\left(Y\right) = \mathbb{E}\left(1_{A_i}\right) + \ldots + \mathbb{E}\left(1_{A_n}\right) = np$$

(ניזכר כי התוחלת של $\text{Ber}\left(p\right)$ היא p, ובנוסף ־ 1_{A_i} ב"ת כי A_i ב"ת).

לסיום, $\mathbb{E}\left(X\right) = \mathbb{E}\left(Y\right)$ כי $X \overset{d}{=} Y \Rightarrow \mathbb{E}\left(X\right) = \mathbb{E}\left(Y\right)$.

למעשה, תוחלת היא אדיטיבית גם עבור מ"מ תלויים (אבל לא כפלית).

טענה

אם X,Y מ"מ על אותו מ"ה, אזי $\mathbb{E}\left(X+Y\right) = \mathbb{E}\left(X\right) + \mathbb{E}\left(Y\right)$ (כשיש להם תוחלת).

הוכחה (עבור בדידים)

$$\mathbb{E}\left(X+Y\right) = \sum_{\beta,\gamma}\left(\beta+\gamma\right)\mathbb{P}_{X,Y}\left(\beta,\gamma\right) =$$
$$\sum_{\beta,\gamma}\left(\beta+\gamma\right)\mathbb{P}\left(X=\beta \mid Y=\gamma\right)\mathbb{P}\left(Y=\gamma\right)$$
$$= \sum_{\gamma}\gamma\mathbb{P}\left(Y=\gamma\right)\underbrace{\sum_{\beta}\mathbb{P}\left(X=\beta \mid Y=\gamma\right)}_{1} +$$
$$\sum_{\beta}\beta\mathbb{P}\sum_{\gamma}\mathbb{P}\left(X=\beta \mid Y=\gamma\right)\mathbb{P}\left(Y=\gamma\right) =$$
$$\mathbb{E}\left(Y\right) + \sum_{\gamma}\mathbb{P}\left(Y=\gamma\right)\mathbb{E}\left(X \mid Y=\gamma\right) \overset{\text{משפט התוחלת השלימה}}{=}$$
$$\mathbb{E}\left(X\right) + \mathbb{E}\left(Y\right)$$

טענה

אם ל־X יש תוחלת $\mathbb{E}\left(\alpha X\right)=\alpha\mathbb{E}\left(X\right)$ לכל $\alpha\in\mathbb{R}$.(הומוגניות)

מכך שהתוחלת מקיימת הומוגניות וגם אדיטיביות, עולה כי היא ליניארית.

דוגמא

בבעיית המזכיר המבולבל, תוחלת מספר המעטפות, שיגיעו ליעדן היא 1.
(גם כאן אם A_i - המאורע שהמכתב ה־i הגיע ליעדו, אזי $\mathbb{1}_{A_1}+\ldots+\mathbb{1}_{A_n}$ שווה למספר ההצלחות, ולכן $\mathbb{1}_{A_i}$ היא ברנולי - $\frac{1}{n}$. מצד שני, הפעם הם תלויים!
(לכאורה, יש כאן הטייה חיובית, כי אם מכתב הגיע ליעד מסוים, זה מעלה את הסיכוי שהמכתב הבא יגיע ליעדו). השווה ל־n הטלות מטבע מוטה $\frac{1}{n}$.

שיעור מס' 16:
יום חמישי
11.12.20

טענה (חיוביות התוחלת)

יהי X מ"מ בדיד ובעלת תוחלת סופית.
(1) אם $\mathbb{P}\left(X\geq 0\right)=1$ אזי $\mathbb{E}\left(X\right)\geq 0$.
(2) אם בנוסף $\mathbb{P}\left(X>0\right)>0$ (במילים אחרות X הוא לא המ"מ הקבוע 0)
אזי $\mathbb{E}\left(X\right)>0$.

הוכחה

נבחין כי מתקיים:

$$\mathbb{E}\left(X\right)=\sum_{x\in\text{Supp}(X)}x\cdot\mathbb{P}_X\left(x\right)\Rightarrow\text{Supp}\left(x\right)\subset[0,\infty)\Rightarrow$$
$$\sum_{x\in\text{Supp}(X)}x\cdot\mathbb{P}_X\left(x\right)\geq\sum_{x\in\text{Supp}(X)}0\cdot\mathbb{P}_X\left(x\right)=0$$

מאידך, עבור חלק (2) , מההנחה קיים $x_0>0$ כך ש־ $\mathbb{P}_X\left(x_0\right)>0$.
אם כך, מתקיים:

$$\mathbb{E}\left(X\right)\geq x_0\cdot\mathbb{P}_X\left(x_0\right)>0$$

כנדרש.

טענה (מונוטוניות התוחלת)

אם X,Y מ"מ בעל תוחלת סופית ואם $\mathbb{P}\left(X\geq Y\right)=1$ אזי $\mathbb{E}\left(X\right)\geq\mathbb{E}\left(Y\right)$

הוכחה

נגדיר $Z=X-Y$.
מתקיים כי $\mathbb{P}\left(Z\geq 0\right)=1$, ולכן מחיוביות התוחלת, $\mathbb{E}\left(Z\right)\geq 0$.
כלומר $\mathbb{E}\left(X-Y\right)=\mathbb{E}\left(X\right)-\mathbb{E}\left(Y\right)\geq 0$ (מליניאריות התוחלת) ובפרט נקבל כי $\mathbb{E}\left(X\right)\geq\mathbb{E}\left(Y\right)$.

הערה

אם $\mathbb{P}\left(X>Y\right)>0$ וגם $\mathbb{P}\left(X\geq Y\right)=1$, אזי $\mathbb{E}\left(X\right)>\mathbb{E}\left(Y\right)$.

אי שוויון ינסן

נניח כי $(a,b) \subset \mathbb{R}$ קטע.
פונקציה $f:(a,b) \to \mathbb{R}$ נקראת "קמורה" אם לכל $0 \le t \le 1$ ולכל $x_1, x_2 \in (a,b)$ מתקיים כי

$$f\left(tx_1 + (1-t)\,x_2\right) \le t\left(f\left(x_1\right) + (1-t)\right) f\left(x_2\right)$$

(הן פונקציות גזירות למעט מספר בן־מניה של נקודות)

כעת, תהי $f:(a,b) \to \mathbb{R}$ פונקציה קמורה.
יהי X מ"מ המקיים $\mathrm{Supp}\left(X\right) \subset (a,b)$. כלומר $\{x \in \mathbb{R} \mid \mathbb{P}_X\left(x\right) > 0\} \subset (a,b)$.
אם למ"מ $X, f\left(X\right)$ יש תוחלת סופית אזי:

$$f\left(\mathbb{E}\left(X\right)\right) \le \mathbb{E}\left(f\left(X\right)\right)$$

הערה

נניח כי $\mathrm{Supp}\left(X\right) = \{x_1, \ldots x_m\}$ ונסמן $p_i = \mathbb{P}_X\left(x_i\right)$.
נשים לב כי $\sum\limits_{i=1}^{m} p_i = 1$.

הוכחה

נסמן $\gamma = \mathbb{E}\left(X\right)$. נשים לב כי $\mu \in (a,b)$. מכך ש־f קמורה, נובע כי קיים ישר משיק לגרף של f בנקודה γ.
החותך של הישר הוא $f\left(\mu\right)$. נסמן את השיפוע ב־z.
נכתוב:

$$\ell\left(x\right) = f\left(\mu\right) + t\left(x - \mu\right)$$

לכל $x \in (a,b)$ מתקיים:

$$f\left(x\right) \ge f\left(\mu\right) + t\left(x - \mu\right) = \ell\left(x\right)$$

ומכאן עולה גם כ $\ell\left(0\right) = f\left(\mu\right) + t\mu$.
בפרט נובע שלכל $\omega \in \Omega$ מתקיים כי $f\left(X\left(\omega\right)\right) \ge f\left(\mu\right) + t\left(X\left(\omega\right) - \mu\right)$ וממילא מתקיים:

$$\mathbb{E}\left(f\left(X\right)\right) \ge \mathbb{E}\left(f\left(\mu\right)\right) + t\left(X - \mu\right) = f\left(\mu\right) + t\left(\mathbb{E}\left(X\right) - \mu\right) \overset{\mu = \mathbb{E}(X)}{=} f\left(\mathbb{E}\left(X\right)\right)$$

מסקנה

אם X, X^2 בעלי תוחלת סופית, אזי $\left(\mathbb{E}\left(X\right)\right)^2 \le \mathbb{E}\left(X^2\right)$ ($f\left(t\right) = t^2$ קמורה), ולכן $\mathbb{E}\left(X\right)^2 \le \mathbb{E}\left(X^2\right)$.

דוגמא

אם $x_1, \ldots x_m \in \mathbb{R}$ אי שליליים, אזי:

$$\underbrace{\left(\frac{x_1+\ldots+x_n}{n}\right)}_{\mathbb{E}(X)} \le \underbrace{\frac{x_1^2+\ldots+x_n^2}{n}}_{\mathbb{E}(X^2)}$$

כיוון ש־$X \sim \text{Unif}\left(x_1, \ldots x_n\right)$

דוגמא נוספת

אם $x_1, \ldots x_n \in \mathbb{R}$ אי שלילים, אזי $\frac{x_1+\ldots+x_n}{n} \le \left(x_1, \ldots x_n\right)^{\frac{1}{n}}$.

דבר זה מתקיים כי :

$$-\ln\left(\frac{x_1+\ldots+}{n}\right) \le \frac{1}{n}\sum_{i=1}^{n} -\ln\left(x_i\right)$$

כיוון שהפונקציה $f\left(t\right) = \ln\left(t\right)$ קעורה (ובמילים אחרות $-f$ קמורה).

אם f קעורה, אזי $-f$ קמורה, ולכן בפרט $-f\left(\mathbb{E}\left(X\right)\right) \le \mathbb{E}\left(-f\left(X\right)\right)$ גורר כי $f\left(\mathbb{E}\left(X\right)\right) \ge \mathbb{E}\left(f\left(X\right)\right)$.

ניקח כעת $X \sim \text{Unif}\left(\{x_1, \ldots, x_n\}\right)$ומא"ש ינסן לפונקציות נקבל כי:

$$f\left(\mathbb{E}\left(X\right)\right) \ge \mathbb{E}\left(f\left(X\right)\right) \Rightarrow \ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \ge \frac{1}{n}\sum_{i=1}^{n}\ln\left(x_i\right) = \ln\left(\left(x_1, \ldots, x_n\right)^{\frac{1}{n}}\right)$$

ומכיון ש־e^x מונטונית עולה אז נקבל בצורה דומה כי :

$$\frac{1}{n}\sum_{i=1}^{n} x_i \ge \left(x_1, \ldots, x_n\right)^{\frac{1}{n}}$$

אי שוויון מרקוב

אם $x_1, \ldots x_m$ מספרים חיוביים שהממוצע שלהם הוא $\overline{x}$ אזי לא ייתכן כי $x_i > \overline{x}$ לכל i, ומאידך לא ייתכן כי $x_i < \overline{x}$ לכל i.

כעת ננסח טענה הקשורה לזה.

טענה

יהי X מ"מ בעל תוחלת סופית, אזי $\mathbb{P}\left(X < \mathbb{E}\left(X\right)\right) < 1$ וכן $\mathbb{P}\left(X > \mathbb{E}\left(X\right)\right) < 1$

הוכחה

נניח בשלילה כי $\mathbb{P}\left(X < \mathbb{E}\left(X\right)\right) = 1$.

אזי מתקיים כי:

$$\mathbb{E}\left(X\right) < \mathbb{E}\left(\mathbb{E}\left(X\right)\right) = \mathbb{E}\left(X\right)$$

בסתירה.

אינטואטיבית, אי שוויון מרקוב מרחיב זאת. לא ייתכן שיותר מחצי מאיברי $x_1,\dots,x_n$ גדולים מ־$2\overline{x}$.

משפט (א"ש מרקוב)

יהי X מ"מ אי שלילי ($\text{Supp}\left(X\right)\subset\left[a,\infty\right)$) אזי לכל $a>0$ מתקיים כי:

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

באופן שקול:

$$\mathbb{P}\left(X\geq a\mathbb{E}\left(X\right)\right)\leq\frac{1}{a}$$

הוכחה

נגדיר $Y=a\cdot\mathbb{1}_A\left\{X\geq a\right\}$. המוגדר על ידי $X\left(\omega\right)\leq$ $Y\left(\omega\right)=\begin{cases}a & X\left(\omega\right)\geq a\\0 & X\left(\omega\right)<a\end{cases}$.

לכן מתקיים כי:

$$\mathbb{E}\left(Y\right)\leq\mathbb{E}\left(X\right)\Rightarrow a\cdot\mathbb{1}_A\left\{X\geq a\right\}\leq\mathbb{E}\left(X\right)\Rightarrow\mathbb{P}\left(X\geq a\right)\leq\frac{\mathbb{E}\left(X\right)}{a}$$

דוגמא

יהי $x_1,\dots x_n$ מ"מ ב"ת ב־$\text{Unif}\left(\left[m\right]\right)$, נחסום את ההסתברות שיש לפחות k מהם שקיבלו את אותו ערך.
לכל $J\subset[n]$ עם $|J|=k$. יהי A_j המאורע שהמ"מ $\left\{X_j\right\}_{j\in J}$ קיבלו את אותו הערך. אזי:

$$\mathbb{P}\left(A_j\right)=\sum_{i=1}^{m}\frac{1}{m^k}=\frac{1}{m^{k-1}}$$

לכן עבור המ"מ $Y_J=\mathbb{1}_{A_J}\sim\text{Ber}\left(\frac{1}{m^{k-1}}\right)$ מתקיים $\mathbb{E}\left(Y_J\right)=\frac{1}{m^{k-1}}$. כעת, נתבונן במשתנה המקרי X שבודק לנו כמה קבוצות בגודל k קיימות. אנחנו רוצים שיתקיים כי $X\geq 1$:

$$\mathbb{P}\left(X\geq1\right)=\mathbb{P}\left(\sum_{J\subset[n],J\neq k}Y_J\geq1\right)\leq\frac{\mathbb{E}\left(X\right)}{1}=\mathbb{E}\left(\sum_{\substack{J\subset[n]\\|J|=k}}Y_J\right)=\sum_{\substack{J\subset[n]\\|J|=k}}\mathbb{E}\left(Y_J\right)=\begin{pmatrix}n\\k\end{pmatrix}\frac{1}{m^{k-1}}$$

דוגמא נוספת (השיטה ההסתברותית)

נקבע $N\in\mathbb{N}$. יהיו $A_1\dots A$ קבוצות שכולן בנות n איברים. נניח כי $N<2^{n-1}$. אזי קיימת קבוצה B ש־:

א. חותכת כל אחת מ־$A_1,\dots A_n$.

ב. לא מכילה אף אחת מ־$A_1,\dots A_n$.

התנאי $N<2^{n-1}$ הכרחי, כי אם נניח $A_1=\left\{1,2\right\},A_2=\left\{2,3\right\},A_3=\left\{1,3\right\}$ אזי $N=3$ ו־$2^{n-1}=2$, כלומר התנאי לא מתקיים ולא קיימת B כזאת.

תהי $A=\bigcup_{i=1}^{N}A_i$. נחפש $B\subset A$. לכל $a\in A$ יהי $X_a\sim \text{Ber}\left(\frac{1}{2}\right)$ שמתאר האם $a\in B$.
נגדיר עוד מ"מ:

$$X=\left|\left\{k\in[n]\mid X_a=X_b\quad \forall a,b\in A_k\right\}\right|$$

נבחין כי אם $B\cap A_i\neq 0$ וגם $A_i\not\subseteq B$ אזי A_i אינה מקיימת את התנאי כי לכל $a,b\in A_i$ נקבל כי $X_a=X_b$.
כלומר, המאורע $\{X=0\}$ הוא המאורע המבוקש. אם $\{X=0\}\neq\emptyset$ 'ניצחנו'.
נבחין כי מתקיים:

$$X=\sum_{k=1}^{N}\sum_{i\in\{0,1\}}\mathbb{1}_{\{X_a|i\quad \forall a\in A_k\}}$$

כי התנאי $\mathbb{1}_{\{X_a|i\quad \forall a\in A_k\}}$ הוא 1 אם A_k מקיימת את התנאי $X_a=X_b$ לכל $a,b\in A_i$,ו־0 אחרת.

$$\mathbb{E}\left(X\right)=\sum_{k=1}^{N}\sum_{i\in\{0,1\}}\mathbb{P}\left(X_a=i\ \ \forall a\in A_k\right)=\sum_{k=1}^{N}\sum_{i\in\{0,1\}}\frac{1}{2^n}=N\cdot 2\cdot\frac{1}{2^n}=\frac{N}{2^{n-1}}<1$$

דבר זה גורר כי:

$$\mathbb{P}\left(X=0\right)=1-\mathbb{P}\left(X\geq 1\right)\geq 1-\mathbb{E}\left(X\right)=1-\mathbb{E}\left(X\right)>0$$

שונות

שיעור מס' 17:
יום שלישי
15.12.20

נרצה לבדוק כמה X נוטה להיות שונה מתוחלתו.למשל, קוביה הוגנת ומ"מ קבוע עם ערך 3.5, הם בעלי אותה תוחלת.
הרעיון של $X-\mathbb{E}\left(X\right)$ לא עובד כי :

$$\mathbb{E}\left(X-\mathbb{E}\left(X\right)\right)=\mathbb{E}\left(X\right)-\mathbb{E}\left(\mathbb{E}\left(X\right)\right)=0$$

והסטיות של X מעל ומתחת ל־$\mathbb{E}\left(X\right)$ מאזנת זו את זו.
אמנם, $\mathbb{E}\left(\left|X-\mathbb{E}\left(X\right)\right|\right)$ יכול להתאים. איננו אוהבים ערכים מוחלטים ולכן נתבונן בריבוע.

הגדרה
למ"מ בדיד X עם תוחלת $\mathbb{E}\left(X\right)$, השונות של X היא $\text{Var}\left(X\right)=\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^2\right)$ או בקיצור $V\left(X\right)$.
($V\left(X\right)$ בהכרח קיים אם מרשים שיתקיים ∞ ־ סכום טור אי שלילי).

סטיית התקן של X היא $\sigma\left(X\right)=\sqrt{V\left(X\right)}$.

דוגמא

אם $X\sim \text{Ber}\left(p\right)$ אזי $\mathbb{E}\left(X\right)=p$ ולכן:

$$V\left(X\right)=\mathbb{E}\left(\left(X-p\right)^{2}\right)=p\left(1-p\right)^{2}+\left(1-p\right)\left(-p\right)^{2}=p\left(1-p\right)$$

טענה

אם ל־X יש שונות, אזי $V\left(X\right)=\mathbb{E}\left(X^{2}\right)-\mathbb{E}\left(X\right)^{2}$.

הוכחה

נסמן $\mu=\mathbb{E}\left(X\right)$ ואז יתקיים כי:

$$\begin{aligned}V\left(X\right)=\mathbb{E}\left(\left(X-\mu\right)^{2}\right)=\mathbb{E}\left(X^{2}-2\mu X+\mu^{2}\right)= \\ \mathbb{E}\left(X^{2}\right)-2\mu\mathbb{E}\left(X\right)+\mu^{2}=\mathbb{E}\left(X^{2}\right)-2\mu^{2}+\mu^{2}= \\ \mathbb{E}\left(X^{2}\right)-\mathbb{E}\left(X\right)^{2}\end{aligned}$$

תכונות

(א) $V\left(X\right)\geq 0$ ו־$V\left(X\right)=0$ אם"ם $X\overset{\text{a.s}}{=}c$ ל־$c\in\mathbb{R}$ (מדוע? הוכחנו את חיוביות התוחלת, ולכן כיוון ש־$\left(X-\mathbb{E}\left(X\right)\right)^{2}$ אי שלילי, ולמעשה בעל תוחלת 0 אם ורק אם הוא 0 כמעט תמיד)

(ב) לכל $a\in\mathbb{R}$ מתקיים כי $V\left(aX\right)=\mathbb{E}\left(\left(aX-\mathbb{E}\left(aX\right)\right)^{2}\right)=a^{2}V\left(X\right)$.

(ג) מתקיים כי $V\left(X+a\right)=\mathbb{E}\left(\left(X+a-\mathbb{E}\left(X+a\right)\right)^{2}\right)=\mathbb{E}\left(\left(\mathbb{E}-\mathbb{E}X\right)^{2}\right)=V\left(X\right)$

האם $V\left(X+Y\right)=V\left(X\right)+V\left(Y\right)$ ־ לא תמיד.

נסמן $\mu=\mathbb{E}\left(X\right),\nu=\mathbb{E}\left(Y\right)$ ואז נקבל:

$$\begin{aligned}&V\left(X+Y\right)-V\left(X\right)-V\left(Y\right)= \\ &\mathbb{E}\left(\left(X+Y-\mu-\nu\right)^{2}\right)-\mathbb{E}\left(\left(X-\mu\right)^{2}\right)-\mathbb{E}\left(\left(Y-\nu\right)^{2}\right)= \\ &2\mathbb{E}\left(XY-X\nu-Y\mu+\mu\nu\right)=2\left(\mathbb{E}\left(XY\right)-\mu\nu-\nu\mu+\mu\nu\right)- \\ &2\left(\mathbb{E}\left(XY\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)\right)\end{aligned}$$

מסקנה

$V\left(X+Y\right)=V\left(X\right)+V\left(Y\right)$ אם"ם $\mathbb{E}\left(XY\right)=\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)$ ובפרט אם X,Y בלתי תלויים. (המושג עצמו של מכפלת התוחלת נקרא **בלתי מתואמים**).

דוגמא

אם $X\sim \text{Bin}\left(n,p\right)$ אזי $V\left(X\right)=np\left(1-p\right)$

הוכחה

יהיו $Y_{1},\ldots,Y_{n}$ משתני $\text{Ber}\left(P\right)$ ב"ת ו־$Y=\sum\limits_{i=1}^{n}Y_{i}$ אזי:

$$V\left(X\right)=V\left(Y\right)=V\left(\sum_{i=1}^{n}Y_{i}\right)=\sum_{i=1}^{n}V\left(Y_{i}\right)=np\left(1-p\right)$$

תרגיל

חשבו שונות של משתנה פואסון, אחיד, גיאומטרי.

אי שוויון צ'בישב

יהיו X מ"מ בעל שונות סופית, אזי לכל $a \in \mathbb{R}$ מתקיים כי:

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

בניסוח אחר:

$$\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq b\cdot\sigma\left(X\right)\right)<\frac{1}{b^{2}}$$

"הסיכוי ש־X רחוק לפחות 3 סטיות תקן מהתוחלת שלו, הוא לכל היותר $\frac{1}{9}$"

הוכחה

$Y=\left(X-\mathbb{E}\left(X\right)\right)^{2}$ הוא מ"מ אי שלילי, ולכן ממרקוב מתקיים:

$$\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq\sqrt{c}\right)=\mathbb{P}\left(Y\geq c\right)=\frac{\mathbb{E}\left(Y\right)}{c}=\frac{V\left(X\right)}{c}$$

ולכן נציב $a=\sqrt{c}$ ונסיים.

דוגמא

נניח כי $X\sim\text{Bin}\left(1000,\frac{1}{2}\right)$ אזי:

$$\mathbb{P}\left(451\leq X\leq549\right)=1-\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq50\right)\geq$$
$$1-\frac{V\left(X\right)}{2500}=1-\frac{250}{2500}-\frac{9}{10}$$

כי $\mathbb{E}\left(X\right)=500$ והמרחק מהתוחלת (של המשלים) גדול מ־50. כל השאר זה פשוט הצבה בנוסחה.

שיעור מס' 18:
יום חמישי
17.12.20

משפט (החוק החלש של המספרים הגדולים)

תהי $\left\{X_{n}\right\}_{n=1}^{\infty}$ סדרה אינסופית של **משתנים מקריים, בלתי תלויים ושווי התפלגות** (נסמנם iid). עם תוחלת סופית μ ושונות סופית, אזי לכל $\varepsilon>0$ מתקיים כי:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_{i}-\mu\right|\geq\varepsilon\right)\underset{n\to\infty}{\rightarrow}0$$

(ככל שאעשה יותר ניסויים, הסיכוי שאתרחק מהממוצע, הולך וקטן).

הוכחה

נסמן את הממוצע ־ $A_{n}=\frac{1}{n}\sum\limits_{i=1}^{n}X_{i}$ וכמו כן, נסמן את השונות בתור $V=V\left(X_{i}\right)$.

מליניארית התוחלת מתקיים כי:

$$\mathbb{E}\left(A_n\right)=\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i\right)=\frac{1}{n}\cdot n\cdot\mu=\mu$$

ואם כך, מתקיים כי:

$$V\left(A_n\right)=V\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)\overset{(*)}{=}\frac{1}{n^2}V\left(\sum_{n=1}^{n}X_i\right)\overset{(**)}{=}$$
$$\frac{1}{n^2}n\cdot V=\frac{V}{n}$$

$(*)$ מתכונות השונות.

$(**)$ הם בלתי תלויים ולכן $V\left(\sum X_i\right)=\sum V\left(X_i\right)$ (

ניזכר באי שוויון צ'בשיב -

$\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq b\cdot\sigma\left(X\right)\right)<\frac{1}{b^2}$.

וכעת נסמן $\varepsilon=b\cdot\sigma\left(X\right)$, ואז מתקיים כי $\frac{1}{b^2}=\frac{V\left(A_n\right)}{\varepsilon^2}$)

$$\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq\underbrace{b\cdot\sigma\left(X\right)}_{\varepsilon}\right)\leq\frac{V\left(A_n\right)}{\varepsilon^2}=\frac{\frac{V}{n}}{\varepsilon^2}=\frac{V}{n\left(\varepsilon^2\right)}$$

ואכן כאשר $n\to\infty$ מתקיים כי $\frac{V}{n\left(\varepsilon^2\right)}\to 0$.

הערה

המשפט נכון גם אם השונות של X לא סופית, אבל ההוכחה שונה (ראו פרק 6.5 בספר).

הגדרה

יהיו X,Y משתנים מקריים על אותו מרחב הסתברות. אזי **השונות המשותפת** (Cov) של X ו־Y מוגדרת על ידי:

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

אם נציב בנוסחה לעיל, מתקיים כי:

$$V\left(X+Y\right)=V\left(X\right)+V\left(Y\right)+2\left(\mathbb{E}\left(XY\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)\right)=$$
$$V\left(X\right)+V\left(Y\right)+2\text{Cov}\left(X,Y\right)$$

אם $\text{Cov}\left(X,Y\right)=0$, אזי נאמר כי X,Y **בלתי מתואמים** (בפרט כש־X,Y ב"ת).

הערה

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

הגדרה

נאמר כי ל־X,Y **מתאם שלילי** אם $\text{Cov}\left(X,Y\right)<0$.

נאמר כי ל־X,Y **מתאם חיובי** אם $\text{Cov}\left(X,Y\right)>0$.

אינטואיטיבית, מתאם שלילי אומר כי השונות $V\left(X+Y\right)<V\left(X\right)+V\left(Y\right)$ - כלומר כאשר נחבר את שני המשתנים המקריים, **נתקרב יותר לממוצע** מאשר אם נסתכל על כל אחד בנפרד - ערכים "גדולים מהרגיל" של X נוטים לבוא עם ערכים "קטנים מהרגיל" של Y.

מתאם חיובי יכול להתקבל כי ערכים גדולים של X מטים את Y לכיוון ערכים גדולים, ולכן מרחק הסכום מהתוחלת גדול מהצפוי.

דוגמה (קיצונית)

$$\text{Cov}\left(X,X\right)=\mathbb{E}\left(XX\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(X\right)=\mathbb{E}\left(X^{2}\right)-\mathbb{E}\left(X\right)^{2}=V\left(X\right)$$

ואכן, נקבל כי:

$$V\left(X+X\right)=V\left(2X\right)=4V\left(X\right)$$

הערה

$\text{Cov}\left(X,Y\right)$ היא **תבנית בילינארית סימטרית** על מ"ו של המשתנים המקריים מעל מרחב הסתברות מסוים.

כלומר, $\text{Cov}\left(X,Y\right)=\text{Cov}\left(Y,X\right)$ וגם $\text{Cov}\left(aX,Y\right)=a\text{Cov}\left(X,Y\right)$ וגם $\text{Cov}\left(X+Z,Y\right)=\text{Cov}\left(X,Y\right)+\text{Cov}\left(Z,Y\right)$

הוכחה

תרגיל.

זהירות! זו לא מכפלה פנימית. כי ישנה חיוביות, אבל אין **חיוביות בהחלט**. כי אם X קבוע שונה מאפס, אזי $V\left(X\right)=0$ למרות ש־$X\neq0$.

כעת, נפעיל את קושי שוורץ (עובד גם על טענות בילינאריות סימטריות):

$$\begin{aligned}\left|\text{Cov}\left(X,Y\right)\right| & \leq\sqrt{\text{Cov}\left(X,X\right)}\cdot\sqrt{\text{Cov}\left(Y,Y\right)}=\\ & \sqrt{V\left(X\right)}\sqrt{V\left(Y\right)}=\\ & \sigma\left(X\right)\sigma\left(Y\right)\end{aligned}$$

הגדרה

המקדם המתאם של X ו־Y מוגדר על ידי:

$$r_{X,Y}=\rho_{X,Y}=\frac{\text{Cov}\left(X,Y\right)}{\sigma\left(X\right)\sigma\left(Y\right)}$$

הערה

כאשר $X=Y$ ראינו כי $\text{Cov}\left(X,X\right)=V\left(X\right)$ ולכן בפרט $\left|\text{Cov}\left(X,Y\right)\right|=\sigma\left(X\right)\sigma\left(X\right)$.
נבחין כי $r_{X,Y}\in\left[-1,1\right]$

תרגיל

מתי $r_{X,Y}=\pm1$?

טענה

אם $X_1,\dots X_n$ משתנים מקריים על אותו מרחב הסתברות, אזי השונות המשותפת הינה:

$$V\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}V\left(X_i\right)+2\sum_{1\leq i<j\leq n}\text{Cov}\left(X_i,Y_j\right)$$

הוכחה

נשים לב שמתקיים:

$$V\left(\sum_{i=1}^{n}X_i\right)=\text{Cov}\left(\sum_{i=1}^{n}X_i,\sum_{j=1}^{n}X_j\right)\overset{\text{תבה"ס}}{=}$$
$$\sum_{i,j}\text{Cov}\left(X_i,X_j\right)\overset{(*)}{=}$$
$$\sum_{i=1}^{n}V\left(X_i\right)+2\sum_{1\leq i<j\leq n}\text{Cov}\left(X_i,X_j\right)$$

$(*)$ חילקנו למצב של שונות משותפת עם עצמו, וכאשר הם שונים.

דוגמה

ננסה לחשב את השונות המשותפת בבעיית המזכיר המבולבל.
נסמן X = מספר המכתבים שנכנסו למעטפה הנכונה. כעת, נרצה לחשב את $V\left(X\right)$.
נגדיר אינדיקטור X_i המציין האם המכתב ה־i נכנס למעטפה הנכונה או לא.
מתקיים כי $X_i\sim\text{Ber}\left(\frac{1}{n}\right)$ וגם $X=\sum\limits_{i=1}^{n}X_i$.
כמו כן, נקבל:

$$V\left(X_i\right)=\mathbb{E}\left(X_i^2\right)-\mathbb{E}\left(X_i\right)^2=$$
$$\frac{1}{n}-\frac{1}{n^2}=\frac{n-1}{n^2}$$

אם כך, מהי השונות המשותפת?

$$\text{Cov}\left(X_i,X_j\right)=\mathbb{E}\left(X_i,X_j\right)-\mathbb{E}\left(X_i\right)\left(\mathbb{E}\left(X_j\right)\right)$$

נבחין כי $X_iX_j=1$ אם"ם $X_i=X_j=1$. בנוסף, $\mathbb{P}\left(X_j=X_i=1\right)=\frac{1}{n}\cdot\frac{1}{n-1}$.

אם כך, $\mathbb{E}\left(X_i, X_j\right) = \frac{1}{n(n-1)}$.
כלומר קיבלנו בסך הכל:

$$\text{Cov}\left(X_i, X_j\right) = \frac{1}{n\left(n-1\right)} - \frac{1}{n^2} = \frac{1}{n^2\left(n-1\right)}$$

ולסיכום:

$$\begin{aligned} V\left(X\right) = \sum_{i=1}^{n} V\left(X_i\right) + 2\sum_{1\le i<j\le n}\text{Cov}\left(X_i, X_j\right) = \\ n\cdot\frac{n-1}{n^2} + 2\binom{n}{2}\frac{1}{n^2\left(n-1\right)} = \\ \frac{n-1}{n} + 2\cdot\frac{n\left(n-1\right)}{2}\frac{1}{n^2\left(n-1\right)} \\ \frac{n-1}{n} + \frac{1}{n} = 1 \end{aligned}$$

שיעור מס' 19:
יום שלישי
22.12.20

דוגמא:
בקורס כלשהו יש 2 מסלולים של סטודנטים שלוקחים את הקורס. $\frac{1}{3}$ מהתלמידים במסלול א' ו$\frac{2}{3}$ מהמסלולים במסלול ב'.
לתלמידי מסלול א' יש תוחלת ציון 80 ושונות 84.
לתלמידי מסלול ב' יש תוחלת ציון 83 ושונות 49.
מה התוחלת והשונות של ציון תלמיד מקרי?
נגדיר W להיות תוחלת ושונת של ציון תלמיד מקרי. נקבל[16] $W = ZX + \left(1-Z\right)Y$ כאשר $Z\sim\text{Ber}\left(\frac{1}{3}\right)$ בוחר מאיזה מסלול הסטודנט שלנו, X ציון תלמיד מקרי ממסלול א' וY ציון תלמיד מקרי ממסלול ב'.
נשים לב ש3 המ"מ X, Y, Z הם ב"ת. לכן

$$\mathbb{E}\left(W\right) \overset{\overbrace{\text{ליניאריות}}^{}}{\downarrow}{=} \mathbb{E}\left(ZX\right) + \mathbb{E}\left(\left(1-Z\right)Y\right) \overset{\overbrace{\text{אי תלות}}^{}}{\downarrow}{=} \mathbb{E}\left(Z\right)\mathbb{E}\left(X\right) + \mathbb{E}\left(1-Z\right)\mathbb{E}\left(Y\right) = \\ = \frac{1}{3}\cdot 80 + \frac{2}{3}\cdot 83 = 82$$

נחשב את השונות באופן כללי, כדי שנוכל לקבל נוסחה כללית ו"יפה":

[16] צריך להוכיח עד הסוף שזה באמת W, אבל מה שמעניין אותנו הוא המבנה הזה, ולא באמת סיפור הרקע.

$$V\left(ZX+(1-Z)Y\right)=\mathbb{E}\left(\left(ZX+(1-Z)Y\right)^2\right)-\left(\mathbb{E}\left(ZX+(1-Z)Y\right)\right)^2=$$

$$=\mathbb{E}\left(Z^2\right)\mathbb{E}\left(X^2\right)+\mathbb{E}\left((1-Z)^2\right)\mathbb{E}\left(Y^2\right)+\underbrace{2\mathbb{E}\left(ZX(1-Z)Y\right)}_{Z(1-Z)=0}-\left(\mathbb{E}\left(ZX+(1-Z)Y\right)\right)^2$$

$$=\mathbb{E}(Z)\mathbb{E}\left(X^2\right)+\mathbb{E}(1-Z)\mathbb{E}\left(Y^2\right)+0-\left(\mathbb{E}\left(ZX+(1-Z)Y\right)\right)^2=$$

$$=\mathbb{E}(Z)\mathbb{E}\left(X^2\right)+\mathbb{E}(1-Z)\mathbb{E}\left(Y^2\right)-\mathbb{E}(Z)^2\mathbb{E}(X)^2-\mathbb{E}(1-Z)^2\mathbb{E}(Y)^2-2\mathbb{E}(Z)\mathbb{E}(X)\mathbb{E}(1-Z)\mathbb{E}(Y)=$$

$$=\mathbb{E}\left(Z\mathbb{E}\left(X^2\right)-Z\mathbb{E}(X)^2+(1-Z)\mathbb{E}\left(Y^2\right)-(1-Z)\mathbb{E}(Y)^2+Z\mathbb{E}(X)^2+(1-Z)\mathbb{E}(Y)^2\right)-$$

$$-\mathbb{E}(Z)^2\mathbb{E}(X)^2-\mathbb{E}(1-Z)^2\mathbb{E}(Y)^2-2\mathbb{E}(Z)\mathbb{E}(X)\mathbb{E}(1-Z)\mathbb{E}(Y)=$$

$$=\mathbb{E}\left(Z\cdot V(a)+(1-Z)V(Y)\right)+\mathbb{E}(Z)\mathbb{E}(X)^2+\mathbb{E}(1-Z)\mathbb{E}(Y)^2-\mathbb{E}(Z)^2\mathbb{E}(X)^2-$$

$$-\mathbb{E}(1-Z)^2\mathbb{E}(Y)^2-2\mathbb{E}(Z)\mathbb{E}(X)\mathbb{E}(1-Z)\mathbb{E}(Y)=$$

$$=\mathbb{E}\left(Z\cdot V(a)+(1-Z)V(Y)\right)+\mathbb{E}\left(\underbrace{\left(Z\mathbb{E}(X)+(1-Z)\mathbb{E}(Y)\right)^2}_{\substack{Z^2\mathbb{E}\left(X^2\right)+(1-Z)^2\mathbb{E}(Y)^2+2\cdot 0=\\ =Z\mathbb{E}(X)^2+(1-Z)\mathbb{E}(Y)^2}}\right)-$$

$$\mathbb{E}\left(Z\mathbb{E}(X)+(1-Z)\mathbb{E}(Y)\right)^2=$$

$$=\mathbb{E}\left(Z\cdot V(a)+(1-Z)V(Y)\right)+V\left(Z\mathbb{E}(X)+(1-Z)\mathbb{E}(Y)\right)$$

קבלנו שהשונות של ציון מקרי, הוא [התוחלת של השונויות] ועוד [השונות של ממוצעים]. אם הערך הראשון קטן והשני גדול, אזי יש לנו ערך לפצל את הקורס ל2, כי רואים שהציון של האנשים משתנה לפי השייכות שלו לקבוצה.

טענה (שלא נוכיח)

אוסף המשתנים המקריים בעלי שונות ותוחלת 0 מהווים מרחב מכפלה פנימית, כאשר השונות המשותפת היא המכפלה הפנימית.

טענה (קושי שוורץ הסתברותי)

לכל X,Y בעלי שונות סופית, אזי מתקיים כי $\mathbb{E}(XY)\leq\sqrt{\mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)}$ ושוויון מתקיים אם"ם $X=aY$.

עבור משתנים עם תוחלת שונה מ-0 מתקיים כי $\text{Cov}(XY)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)=\mathbb{E}(XY)$.

אם התוחלת שווה ל-0 אזי $\mathbb{E}\left(X^2\right)=\text{Var}(X)$.

הוכחה

אם $\mathbb{E}\left(X^2\right)=0$ או $\mathbb{E}\left(Y^2\right)=0$ הטענה טרוויאלית.

נגדיר $\overline{Y}=\frac{Y}{\sqrt{\mathbb{E}(Y^2)}},\overline{X}=\frac{X}{\sqrt{\mathbb{E}(X^2)}}$ - הגדרות אלו נקראת "תקנון".

נשים לב שמתקיים:

$$\mathbb{E}\left(\overline{X^2}\right)=\frac{\mathbb{E}\left(X^2\right)}{\mathbb{E}\left(X^2\right)}=1=\mathbb{E}\left(\overline{Y^2}\right)$$

וגם מתקיים כי $|XY|\leq\frac{X^2+Y^2}{2}$. ולכן, סך הכל קיבלנו כי:

$$\mathbb{E}\left(\left|\overline{X}\cdot\overline{Y}\right|\right)\leq\frac{\mathbb{E}\left(\overline{X^2}\right)+\mathbb{E}\left(\overline{Y^2}\right)}{2}=1$$

אם נציב, יתקיים:

$$\mathbb{E}\left(\frac{X}{\sqrt{\mathbb{E}\left(X^2\right)}}\cdot\frac{Y}{\sqrt{\mathbb{E}\left(Y^2\right)}}\right)\leq\frac{\mathbb{E}\left(\left(\frac{X}{\sqrt{\mathbb{E}(X^2)}}\right)^2\right)+\mathbb{E}\left(\left(\frac{Y}{\sqrt{\mathbb{E}(Y^2)}}\right)^2\right)}{2}\Rightarrow$$
$$\mathbb{E}\left(XY\right)\leq\sqrt{\mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right)}$$

מומנטים

שיעור מס' 20:
יום חמישי
24.12.20

טענה
יהי X מ"מ כך ש־$\mathbb{E}\left(X\right)<\infty$.
אזי:

$$\left(\mathbb{E}\left(X-a\right)^2\right)\geq\mathbb{E}\left(\left(x-\mathbb{E}\left(X\right)\right)^2\right)=V\left(X\right)$$

ה"שונות היא מומנט מוזז מינימלי".

הוכחה
ניקח $\mathbb{E}\left(\left(X-\mathbb{E}-b\right)^2\right)$ ונראה שהינו מינימלי כאשר $b=0$:

$$\mathbb{E}\left(\left(X-\mathbb{E}-b\right)^2\right)=\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^2\right)-2\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)b\right)\right)+\mathbb{E}\left(b^2\right)=$$
$$V\left(X\right)-0+b^2\geq V\left(X\right)$$

מסקנה - עבור מ"מ בעלי שונות, תוחלת היא הערך שהמומנט השני המוזז סביבו, הוא מינימלי.

עד כה ראינו את אי שוויוני מרקוב וצ'בישב.
נוכל להבחין כי לפעמים אי השוויון $\mathbb{P}\left(X\geq a\right)\leq\frac{\mathbb{E}\left(X^2\right)}{a^2}$ טוב יותר, כאשר $a\to\infty$, שהרי $\mathbb{E}\left(X^2\right)$ הינו קבוע.
אם כך, א"ש צ'בישיב הן "הערכות זנב" - הסתברות ש־X רחוק מ־0 או מהתוחלת שלו[17].
למעשה, אי שוויון צ'בישיב טוב יותר בהערכות "זנב", יותר מאשר מרקוב (צ'בישב הוא שיפור של האי שוויון האחרון שהראינו).
באותה מידה, נוכל גם לקחת X^3. כלומר נקבל $\mathbb{P}\left(X^3>a^3\right)\leq\frac{\mathbb{E}\left(X^3\right)}{a^3}$ שהינה הערכה טובה יותר.
נשים לב כי אנחנו בכל פעם מניחים כי ישנה תוחלת למשתנה המקרי של X^2, אמנם ייתכן כי לא יהיה תוחלת למשתנה המקרי בריבוע.

הגדרה

[17] מצא f כך ש־$\infty>\lim_{a\to\infty}\frac{\mathbb{P}(X>a)}{f(a)}$. או מצא f כך ש־$\mathbb{P}\left(X>a\right)=O\left(f\left(a\right)\right)$.

מומנט k הינו $\mathbb{E}\left(X^k\right)$.
בנוסף מומנט ממורכז k הינו $\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^k\right)$.
בפרט, תוחלת היא מומנט ראשון, שונות היא מומנט ממורכז שני.

אם מומנט k מוגדר (סופי), אזי לכל $l<k$ המומנט l מוגדר וסופי.

טענה - א"ש צ'בשיב מוכלל
נבחין כי $\mathbb{P}\left(X\geq a\right)\leq\frac{\mathbb{E}\left(X^k\right)}{a^k}$ וגם $\mathbb{P}\left(\left|X-\mathbb{E}\left(X\right)\right|\geq a\right)\leq\frac{\mathbb{E}\left(\left|X-\mathbb{E}(X)\right|^k\right)}{a^k}$.

למה שניקח דווקא X^k? אכן, אין סיבה לכך. נוכל לקחת פונקציה כללית יותר.
כלומר:

$$\mathbb{P}\left(X\geq a\right)=\mathbb{P}\left(e^{tX}\geq e^{ta}\right)\leq\frac{\mathbb{E}\left(e^{tX}\right)}{e^{ta}}=\mathbb{E}\left(E^{tX}\right)e^{-ta}$$

כאשר המעבר לתוחלת נובע ממרקוב והמעבר הראשון נובע ממונוטוניות של e^X. נשים לב כי מדובר בהערכה טובה יותר (כי e^{-ta} יורדת ל-0 הרבה יותר מהר).

הגדרה
עבור מ"מ X הפונקציה $M_X\left(t\right)=\mathbb{E}\left(e^{tX}\right)$ מכונה פונקציה יוצרת מומנטים של X.

א"ש צ'רנוף

$$\mathbb{P}\left(X>a\right)\leq M_X\left(t\right)\cdot e^{-at}$$

כאשר אגף ימין מוגדר.

דוגמא
מטילים p כדורים ל-n תאים עם חזרות באקראי. נרצה חסם עליון לעומס בתא העמוס ביותר.
נסמן S_i להיות העומס בתא ה-i. (כלומר $S_i=\left|\sum_j X_j=i\right|$ כאשר X_j הינו מיקום הכדור ה-j).
נרצה כי יתקיים (נרצה אגף שמאל גדול יותר ,ולכן שאגף ימין יהיה קטן יותר):

$$\begin{aligned}\mathbb{P}\left(\max_{i\in[n]}S_i<a\right)&=1-\mathbb{P}\left(\max_{i\in[n]}S_i\geq a\right)\\&\overset{\text{חסם האיחוד}}{\geq}1-\sum_{i\in[n]}\mathbb{P}\left(S_i\geq a\right)=\\&1-n\mathbb{P}\left(S_1\geq a\right)\end{aligned}$$

אנחנו מחפשים a כך ש- $\mathbb{P}\left(S_1\geq a\right)\leq\frac{1}{n}$.
נחשב את M_{S_1}. נבחין כי $S_1\sim\text{Bin}\left(n,\frac{1}{n}\right)$.
באופן כללי, כאשר נרצה לחשב פונקציה יוצרת של $X\sim\text{Bin}\left(n,p\right)$, נבחין כי $X\overset{d}{=}\sum_{i=1}^{n}B_i$, כאשר $B_i\sim\text{Ber}\left(p\right)$.

נחשב:

$$M_{B_i}(t)=\mathbb{E}\left(e^{tB_i}\right)=pe^t+1-p$$

ולכן נקבל:

$$M_X(t)=\mathbb{E}\left(e^{\sum\limits_{i=1}^{n}B_it}\right)=\mathbb{E}\left(\prod e^{B_it}\right)=$$
$$\mathbb{E}\left(e^{B_1t}\right)^n=\left(pe^t+1-p\right)^n$$

ולכן נקבל עבור מה שעשינו קודם לכן כי $M_{S_i}=\left(1+\frac{e^t-1}{n}\right)^n\leq e^{e^{t-1}}$ (כאשר המעבר נובע מכך ש־$\left(\frac{1+x}{n}^n\right)\leq$ e^x).
קודם לכן אמרנו שהטענה נכונה לכל t ולכן עלינו למצוא את ה־t האופטימלי:

$$\mathbb{P}\left(S_1\geq a\right)\leq e^{-ta}M_{S_1}(t)=\exp\left(e^t-1-ta\right)<\frac{1}{n}$$

נגזור t מיטיבי לחיפוש מינימום:

$$\left(e^t-1-ta\right)'=0\Rightarrow e^t=a\Rightarrow t=\log a$$

(עלינו להוכיח פורמלית שמדובר בנקודת מינימום ולא מקסימום, ולא נקודות קצה).
אם כך, קיבלנו בסך הכל:

$$\mathbb{P}\left(S_1\geq a\right)\leq\exp\left(e^{\log a}-1-a\log a\right)\leq\exp\left(-a\left(\log a-1\right)\right)$$

אנחנו רוצים כי $\exp\left(-a\left(\log a-1\right)\right)<\frac{1}{n}$ ולכן בפרט נוציא log ונקבל כי $\left(-a\log\left(a-1\right)\right)<\log\frac{1}{n}=-\log n$.
נחפש כי $a\log\left(a-1\right)-\log n\to\infty$.
נציב $a=\frac{C\log n}{\log\log n}$ עבור $C>1$ ונקבל:

$$\frac{C\log n}{\log\log n}\log\left(\frac{C\log n}{\log\log n}-1\right)-\log n\leq$$
$$\frac{C\log n}{\log\log n}\cdot\left(\log\log n+\log C-\log\log n\right)-\log n\to\infty$$

עבור $a=\frac{C\log n}{\log\log n}$ עבור $c>1$ יתקיים כי $\mathbb{P}\left(S_1\geq a\right)\leq\frac{1}{n}$.

הפעלת צ'רנוף לא תמיד קלה, כי קשה לחשב את $M_X(t)$ וקשה לבחור t מיטיבי. אמנם, ישנו מקום אחד שבו ניתן להשתמש ברעיונות אלו בדרך שטובה מאוד וקלה להפעלה.

א"ש הופטינג

יהיו X_k משתנים מקריים בעלי התכונות הבאות:

(1) ב"ת.
(2) בעלי תוחלת 0.
(3) $|X_k| \overset{a.s}{\leq} 1$

אזי:

$$\mathbb{P}\left(\sum_{k\in[n]} X_k \geq a\right) \leq \exp\left(\frac{-a^2}{2n}\right)$$

עבור $a > 0$.

טענת עזר

אם X משתנה מקרי כך שהוא מקיים את התנאים שהראינו לעיל, אזי $\mathbb{E}\left(e^{tX}\right) \leq \exp\left(\frac{t^2}{2}\right)$

הוכחה

נגדיר את $S = \sum\limits_{i=1}^{n} X_i$

$$M_S(t) = \mathbb{E}\left(e^{t\sum\limits_{i=1}^{n} X_i}\right) \overset{\text{אי תלות}}{=} \prod \mathbb{E}\left(e^{tX_i}\right) \overset{\text{טענת עזר}}{\leq} \prod \exp\left(\frac{t^2}{2}\right) = \exp\left(\frac{nt^2}{2}\right)$$

כעת נפעיל את צ'רנוף ונקבל:

$$\mathbb{P}(S \geq a) \leq \exp\left(\frac{nt^2}{2} - ta\right)$$

נבצע להם אופטימיציה, על ידי גזירה והשוואה ל-0:

$$\left(\frac{nt^2}{2} - ta\right)' = nt - a = 0 \Rightarrow t = \frac{a}{n}$$

ולכן נקבל:

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

שיעור מס' 21:
יום שלישי
29.12.20

מרחבי הסתברות כלליים

הגדרות ומושגים

אם נרצה להגדיר (Ω, P) כאשר p הסתברות נקודתית, ניתקל בבעיה - היחידונים יכולים להיות בעלי הסתברות 0.
הפתרון הינו $(\Omega, \mathbb{P})$ כאשר $\mathbb{P}$ פונקציית הסתברות על Ω.
אמנם, ניתקל בבעיה אחרת - רצינו למצוא $\mathbb{P}: 2^{[0,1]} \to \mathbb{R}_+$ פונקצית הסתברות, אינווריאנטית להזזת $\mathbb{P}(A+c) = \mathbb{P}(A)$, והגענו לסתירה: $A \subseteq [0,1]$ ספציפית כך ש-$\aleph_0$ הזזות זרות של A מהוות חלוקה של $[0,1]$.
הפתרון הוא לוותר על חלק מהקבוצות: $\mathbb{P}(A)$ לא מוגדר לכל $A \subseteq \Omega$.

הגדרה

σ- אלגברה על קבוצה Ω היא אוסף של תתי קבוצות של Ω.
$\mathcal{F} \subseteq 2^{\Omega}$ (כלומר, תת קבוצה של כל תתי הקבוצות החלקיות) שמקיימת:
(I) $\emptyset \in \mathcal{F}$.
(II) **סגירות למשלים** ($A \in \mathcal{F}$ גורר כי $\Omega \setminus A = A^c \in \mathcal{F}$)
(III) **סגירות לאיחוד בן-מנייה** ($A_i \in \mathcal{F}$ $i \in \mathbb{N}$ גורר כי $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$)

הגדרה

מ"ה הוא $(\Omega, \mathcal{F}, \mathbb{P})$ כש-Ω קבוצה, $\mathcal{F}$ היא σ-אלגברה על Ω ו-$\mathbb{P}: \mathcal{F} \to \mathbb{R}_+$ פונקציית הסתברות.

דוגמאות ל-σ- אלגברה על Ω:
(I) $\mathcal{F} = \{\emptyset, \Omega\}$.
(II) $\mathcal{F} = 2^{\Omega}$.
(III) אם $\Omega = \biguplus_{i \in \mathbb{N}} A_i$ (איחוד זר של קבוצות) אזי $\mathcal{F} = \left\{ \bigcup_{i \in I} A_i \mid I \subseteq \mathbb{N} \right\}$.
(IV) $\mathcal{F} = \{A \subseteq \Omega \mid |A^c| \le \aleph_0 \vee |A| \le \aleph_0\}$ (כש-$|\Omega| > \aleph_0$, אחרת מדובר ב-(II)).

דוגמאות למרחב הסתברות עם $\mathcal{F} \neq 2^{\Omega}$: ניקח את $\mathcal{F}$ מ-(IV) וניקח:

$$\mathbb{P}(A) = \begin{cases} 0 & |A| \le \aleph_0 \\ 1 & |A^c| \le \aleph_0 \end{cases}$$

אזי $(\Omega, \mathcal{F}, \mathbb{P})$ מ"ה. צריך כאן $|\Omega| > \aleph_0$ על מנת שהחלוקה למקרים זרים תהיה מוגדרת היטב.

עבור $\mathscr{S} \subseteq 2^{\Omega}$, נחפש האם יש σ- אלגברה מינימלית $\mathcal{F}$ שמכילה את $\mathscr{S}$.
מתקיים:

$$\sigma(\mathscr{S}) = \bigcap_{\substack{\mathscr{S} \subseteq \mathcal{F} \subseteq 2^{\Omega} \\ \mathcal{F}\ \sigma\text{-אלגברה}}} \mathcal{F}$$

ה־σ־ אלגברה על $\mathbb{R}$ שנוצרת ע"י $\mathscr{S}=\{[a,b]\mid a\leq b\}$ נקראת **σ־ אלגברה בורל** של $\mathbb{R}$.
כלומר $\sigma\left(\mathscr{S}\right)=\mathcal{B}=\mathcal{B}_{\mathbb{R}}$.

בדומה, σ־ אלגברה־ בורל של $I=[0,1]$ היא $\mathcal{B}_I=\sigma\left(\{[a,b]\mid 0\leq a\leq b\leq 1\}\right)$.
מי זו $\mathcal{B}_{\mathbb{R}}$?
כל קטע סגור וכל קטע פתוח נמצא בתוכה (באמצעות המשלימים).
אולי $\mathcal{B}_{\mathbb{R}}=2^{\mathbb{R}}$־ לא, A שבנינו בקובץ במודל על התפלגות אחידה איננה ב־$\mathcal{B}_{\mathbb{R}}$ (וזה מצוין עבורנו כי זו לא קבוצה משהו).
יש גם דרך איטרינזית לבנות את $\mathcal{B}_{\mathbb{R}}$ ־ היא אינה נעימה.

משפט (לבג)
ישנה פונקציית הסתברות $\mathbb{P}:\mathcal{B}_I\to\mathbb{R}_+$ אינווריאנטית להזזות.
היא מקיימת $\mathbb{P}\left([a,b]\right)=b-a$.
$\mathbb{P}_\lambda$ הזאת נקראת מידת בורל.
מתקיים כי $\mathbb{P}_\lambda\left(A\right)=\inf\left\{\sum\limits_{i=1}^{\infty}b_i-a_i\mid A\subseteq\bigcup\limits_{i=1}^{\infty}\left(a_i,b_i\right)\right\}$.
הוכחה
לא במסגרת הקורס.

המרחב $\left(I,\mathcal{B}_I,\mathbb{P}_\lambda\right)$ נקרא "הסתברות אחידה על I". גם נקרא "מרחב ההסתברות הסטנדרטי".
גם ההגדרה של משתנה מקרי תשתמש ב־$\mathcal{B}_{\mathbb{R}}$.
נרצה שלכל $A\in\mathcal{B}_{\mathbb{R}}$ נוכל לשאול מהי $\mathbb{P}\left(X\in A\right)$.

הגדרה
משתנה מקרי על מ"ה $\left(\Omega,\mathcal{F},\mathbb{P}\right)$ הוא פונקציית $X:\Omega\to\mathbb{R}$ שמקיימת: $\forall A\in B_I,\ \ X^{-1}\left(A\right)\in\mathcal{F}$.

שיעור מס' 22:
יום חמישי
31.12.20

ההגדרה הזה נותנת לנו לחשב כי $\mathbb{P}\left(X\in A\right)$ מוגדר:

$$\mathbb{P}\left(X\in A\right)=\mathbb{P}\left(\{\omega\in\Omega\mid X\left(\omega\right)\in A\}\right)=\mathbb{P}\left(\underbrace{X^{-1}\left(A\right)}_{\in\mathcal{F}}\right)$$

דוגמה נגדית
ניקח $A\notin\mathcal{B}_I$ ו־$\left(\Omega,\mathcal{F}=\mathcal{B}_I,\mathbb{P}_\lambda\right)$ ו־ $X==\{\mathbb{1}\}_{\mathbb{A}}$.
אזי $\{1\}\in\mathcal{B}_I$ אבל $X^{-1}\left(\{1\}\right)=A\in\mathcal{F}$ לכן X איננו מ"מ.

עוד דוגמה נגדית
ניקח $\left(\Omega=\{0,1\},\mathcal{F}=\{\emptyset,\Omega\},\mathbb{P}\right)$ ו־$X=\mathrm{Id}$, אזי $X^{-1}\left(\{1\}\right)=\{1\}\notin\mathcal{F}$. כל מ"מ על מרחב זה הוא קבוע.
אם $\mathcal{F}=2^{\Omega}$ אזי מ"מ הוא פשוט פונקציה $X:\Omega\to\mathbb{R}$.

עובדה
אם $X^{-1}\left([a,b]\right)\in\mathcal{F}$ לכל $a\leq b$, כבר נובע כי X מ"מ.
(כי $\mathcal{B}_I$ נוצרת על ידי הקטעים).

כמו קודם, אם X מ"מ על מ"ה, נקבל ממנו מ"ה על $\mathbb{R}$:

$$(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mathbb{P}_X)$$

כיצד ניתן לקודד את $\mathbb{P}_X$ הזו? מסתבר ששוב מספיק לבדוק קורה בקטעים - להבין מהו $\mathbb{P}_X([a,b])$ לכל $a \le b$.

דוגמא למ"מ לא בדיד

ניקח $X(\omega) = \omega^2$ ו-$(\Omega = I, \mathcal{F} = \mathcal{B}_I, \mathbb{P}_\lambda)$.

(ולמעשה כל פונקציה רציפה $X : I \to \mathbb{R}$ מקיימת $X^{-1}(A) \in \mathcal{B}_I$ לכל $A \in \mathcal{B}_{\mathbb{R}}$)

מעכשיו, פשוט נחקור את X בהינתן התפלגותו, ולא נחשוב כמעט על Ω.

פונקצית ההתפלגות המצטברת

הגדרה

אם X מ"מ, ה-CDF- פונקציית ההתפלגות המצטברת, היא:

$$F_X(a) = \mathbb{P}(X \le a) = \mathbb{P}_X((-\infty, a])$$

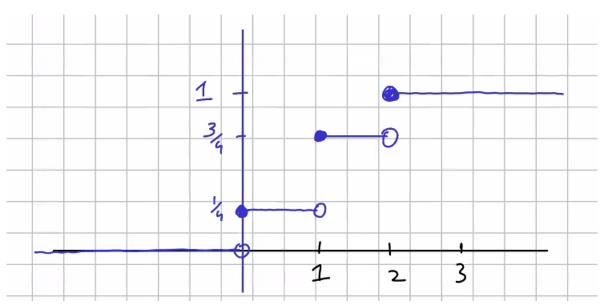
לכל $A \in \mathbb{R}$.

דוגמא

אם X מתפלג אחיד על $[a,b]$, נסמן $X \sim \text{Unif}([a,b])$.

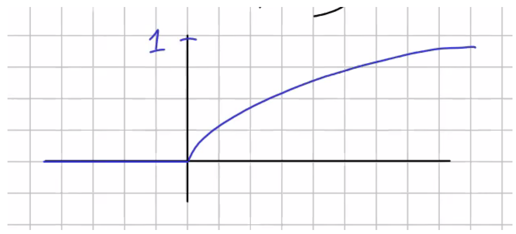
כלומר, $F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & b < x \end{cases}$.

אם ניקח $X \sim \text{Bin}\left(2, \frac{1}{2}\right)$, מתקבלת התמונה הבאה (הסיכוי להיות קטן מ-1, ואז הסיכוי להיות קטן מ-2 ואז קטן מ-3...):



התפלגות מעריכית

נאמר כי X מתפלג מעריכית עם פרמטר λ אם $F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{\lambda X} & x \ge 0 \end{cases}$, וזה נראה כך:



כמו שקיבלנו גבול $\text{Poi}(\lambda)$ $\overset{n\to\infty}{\to}$ $\text{Bin}\left(n,\frac{\lambda}{n}\right)$, נקבל גבול $\exp(\lambda)$ $\overset{n\to\infty}{\to}$ $\text{Geo}\left(\frac{\lambda}{n}\right)$ולכן ההתפלגות מתארת "זמן עד תקרית ראשונה" בתהליך פואסוני (בכל רגע יכולה להתרחש תקרית).

תכונות של פונקציית הסתברות מצטברת

(I) אם X מ"מ, אזי F_X מונטוניות עולה חלש.

(II) $\lim\limits_{x\to\infty}F_X=1$, $\lim\limits_{x\to-\infty}F_X=0$.

בנוסף:

$$F_X\left(b\right)-F_X\left(a\right)=\mathbb{P}\left(X\le b\right)-\mathbb{P}\left(X\le a\right)=\mathbb{P}\left(X\in\left(a,b\right]\right)$$

(מתקבל א"ש חזק בצד שמאל).

טענה

$\mathbb{P}\left(X<a\right)=\lim\limits_{t\to a^-}F_x\left(t\right)$ לכל $a\in\mathbb{R}$.

הוכחה

לכל סדרה $\{t_i\}_{i=1}^{\infty}$ שעולה ממש עם גבול a, נקבל:

$$\lim_{i\to\infty}F_x\left(t_i\right)=\lim_{i\to\infty}\mathbb{P}\left(\left\{X\le t_i\right\}\right)\overset{\text{מונטוניות}}{=}\mathbb{P}\left(\bigcup_{i=1}^{\infty}\left\{X\le t_i\right\}\right)=\mathbb{P}\left(X<a\right)$$

בפרט הגבול $\lim\limits_{t\to a^-}F_X\left(t\right)$ קיים תמיד (מהיינה).

מסקנה

$$\mathbb{P}_X\left(\left[a,b\right]\right)=F_X\left(b\right)-\lim_{t\to a^-}F_X\left(a\right)$$

ובפרט F_X מספיקה לגלות את $\mathbb{P}_X$ על קטעים.

נובע מכך (לא נוכיח) ש־F_X מספיקה לגלות את $\mathbb{P}_X$ לגמרי (כי $\mathcal{B}_I$ נוצרת מהקטעים). במילים אחרות, אם $F_X=F_Y$, אז $\mathbb{P}_X=\mathbb{P}_Y$ (או $X\overset{d}{=}Y$).

עוד מסקנה

מתקיים כי :

$$\mathbb{P}\left(X=a\right)=F_X\left(a\right)-\lim_{t\to a^-}F_X\left(a\right)$$

לכן יש ל־a הסתברות חיובית בדיוק כשל־F_X יש קפיצה ב־a.

תרגיל

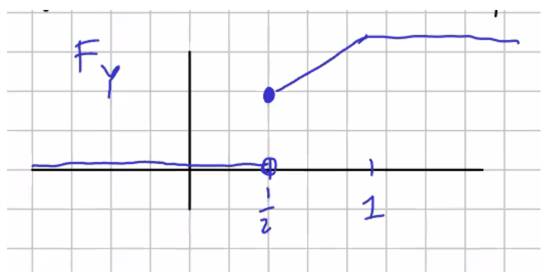
X מ"מ רציף ($\mathbb{P}\left(X=a\right)=0$ לכל a) אם"ם F_Xרציפה.

תזכורת

X נקרא בדיד אם"ם $\sum\limits_{a\in\mathbb{R}}\mathbb{P}\left(X=a\right)=1$.

אם $X\sim\text{Unif}\left(I\right)$ ו־$Y=\max\left(X,\frac{1}{2}\right)$ אזי Y לא בדיד ולא רציף.

כלומר, למשל:



מכאן נצטמצם למשפחה מיוחדת של מ"מ.

פונקציית צפיפות

הגדרה

מ"מ X ייקרא **רציף בהחלט**, אם קיימת פונקציית $f_X:\mathbb{R}\to\mathbb{R}_{\geq0}$, כך ש־ $F_X\left(a\right)=\mathbb{P}\left(X\leq a\right)=\int\limits_{-\infty}^{a}f_X\left(x\right)dx$

(עבורנו זהו אינטגרל רימן, אבל בעולם המבוגרים לוקחים אינטגרל לבג)

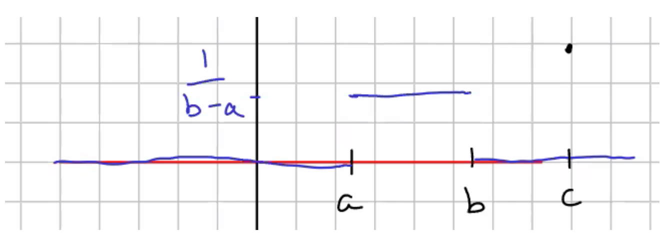
f_X נקראת פונקציית הצפיפות של X וזה נקרא באנגלית PDF.

אם X רציף בהחלט, אזי X רציף (פונקציה קדומה היא תמיד רציפה).

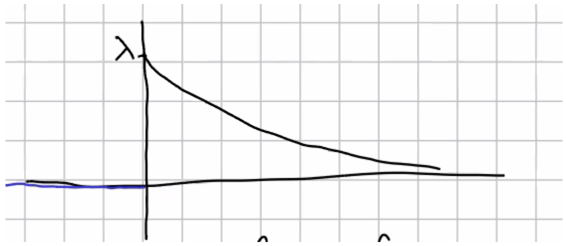
ההיפך איננו נכון, אבל הדוגמאות הן פתולוגיות (הגרלת מספר מקבוצת קנטור, למשל).

דוגמאות

(א) ניקח $X\sim\text{Unif}\left(\left[a,b\right]\right)$ וגם $f_X=\frac{1}{b-a}\cdot\mathbb{1}_{\left[a,b\right]}$ (נבחין כי תמיד אפשר לשנות את f במספר סופי של נקודות בלי שזה ישפיע על $\int f$.



(ב) ניקח את $X\sim\exp\left(\lambda\right)$ כאשר $f_{X}\left(x\right)=\begin{cases} 0 & x<0\\ \lambda e^{-\lambda X} & 0\leq x \end{cases}$



תמיד יתקיים כי $\int\limits_{-\infty}^{\infty}f_{X}\left(x\right)dx=1$.

עובדה

לכל f אי שלילית עם $\int\limits_{-\infty}^{\infty}f=1$, אפשר לבנות מ"ה ועליו מ"מ X עם צפיפות $f_{X}=f$.

(זכרו שעבור $f=1_{I}$ מקטעים $\mathbb{P}_{\lambda}$ התפלגות לבג).

אינטואטיבית, $f_{X}\left(a\right)$ הוא **השינוי** ב־F_{X} סביב a (למשל אם F_{X} גזירה אם $f_{X}\left(a\right)=F_{X}'\left(a\right)$. דבר זה מתאר את ההבדל בין $\mathbb{P}\left(X\leq a-\varepsilon\right)$ ל־$\mathbb{P}\left(X\leq a+\varepsilon\right)$.

כלומר, מהו $\mathbb{P}\left(a-\varepsilon\leq X\leq a+\varepsilon\right)$, כלומר מה ההסתברות ש־$X$ הוא "**בערך**" a.

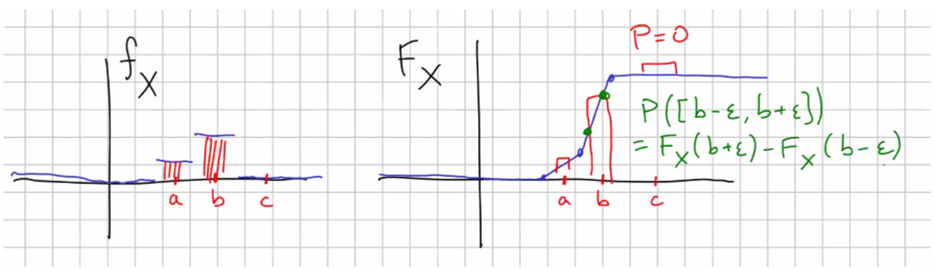
לכן f_{X} **אנאלוגית** ל־p_{X} - פונקצית ההסתברות הנקודתית.

דוגמה

שיעור מס' 23:

יום שלישי

05.01.21



$$P([b-\varepsilon,b+\varepsilon])=F_{X}(b+\varepsilon)-F_{X}(b-\varepsilon)=\int_{b-\varepsilon}^{b+\varepsilon}f_{X}(x)\text{dx}$$

"הצפיפות מתארת את הסבירות להימצא בסביבת הנקודה. כלומר $[b-\varepsilon,b+\varepsilon]$ כאשר $f_{X}\left(b\right)$."

תוחלת ושונות של מ"מ רציפים

תזכורת

תוחלת של X בדיד היא $\mathbb{E}\left(X\right)=\sum\limits_{a\in\mathbb{R}}p_{X}\left(a\right)\cdot a$.

הגדרה

אם X רציף בהחלט, נגדיר:

$$\mathbb{E}\left(X\right)=\int\limits_{-\infty}^{\infty}x\cdot f_{X}\left(x\right)dx$$

(המקבילה של סכום בצורה אינסופית).

בתנאי שהאינטגרל מתכנס בהחלט. $\left(\int\limits_{-\infty}^{\infty}\left|x\right|f_{X}\left(x\right)dx<\infty\right)$

(אפשר גם כאן להגדיר $\mathbb{E}\left(X\right)=\pm\infty$, אך נוותר על זה).

הערה

מיד אפשר להגדיר שונות - $V\left(X\right)=\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)^{2}\right)$

טענה

אם $g:\mathbb{R}\to\mathbb{R}$ מונטונית עולה ממש, ו-X מ"מ עם צפיפות, אזי $Y=g\left(x\right)$ בעל צפיפות $f_{Y}\left(g\right)=\frac{f_{X}\left(g^{-1}\left(a\right)\right)}{g'\left(g^{-1}\left(a\right)\right)}$

הוכחה

ראינו בתרגול.

מסקנה

מתקיים כי $\mathbb{E}\left(Y\right)=\int\limits_{-\infty}^{\infty}g\left(x\right)f_{X}\left(x\right)dx$ (לבדיד- $\mathbb{E}\left(g\left(x\right)\right)=\sum\limits_{a\in\mathbb{R}}g\left(a\right)p_{X}\left(a\right)$).

גם במקרה הלא בדיד דבר זה נכון, לכל g (רציפה למקוטעין). נוכיח זאת רק עבור g גזירה מונטונית, ונשתמש בזה עבור g כללית.

הוכחה

$$\mathbb{E}\left(Y\right)=\int\limits_{-\infty}^{\infty}yf_{Y}\left(y\right)dy\overset{\text{טענת עזר}}{=}\int\limits_{-\infty}^{\infty}\frac{f_{X}\left(g^{-1}\left(y\right)\right)}{g'\left(g^{-1}\left(y\right)\right)}$$

נציב כעת $y=g\left(x\right)$ ונקבל:

$$\int\limits_{\infty}^{\infty}g\left(x\right)\frac{f_{X}\left(x\right)}{g'\left(x\right)}g'\left(x\right)dx=\int\limits_{-\infty}^{\infty}g\left(x\right)f_{X}\left(x\right)dx$$

דוגמאות

(א) $X\sim\exp\left(1\right)$ ו-$Y=e^{X}$, נקבל:

$$\mathbb{E}\left(Y\right)=\int\limits_{-\infty}^{\infty}e^{x}f_{X}\left(x\right)dx=\int\limits_{-\infty}^{\infty}e^{x}e^{-x}dx=-\infty$$

(ב) $X\sim\text{Unif}\left([a,b]\right)$:

$$\mathbb{E}\left(X\right)=\int\limits_{-\infty}^{\infty}x\frac{1}{b-a}\mathbb{1}_{[a,b]}dx=$$

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

אם נרצה לחשב את השונות, נקבל:

$$V\left(X\right)=\mathbb{E}\left(X^{2}\right)-\left(\mathbb{E}\left(X\right)\right)^{2}=\int\limits_{-\infty}^{\infty}x^{2}\frac{\mathbb{1}_{[a,b]}}{b-a}dx-\left(\frac{a+b}{2}\right)^{2}=$$

$$\frac{b^{3}-a^{3}}{3\left(b-a\right)}-\left(\frac{a+b}{2}\right)^{2}=\frac{\left(b-a\right)^{2}}{12}$$

(ג) $X\sim\exp\left(\lambda\right)$. נקבל:

$$\mathbb{E}\left(X\right)=\int\limits_{-\infty}^{\infty}xf_{X}\left(x\right)dx=\int\limits_{0}^{\infty}x\lambda e^{-\lambda x}x$$

$$\overset{\text{אינטגרציה בחלקים}}{=}\left[xe^{-\lambda x}\right]_{0}^{\infty}+\int\limits_{0}^{\infty}e^{-\lambda x}dx=$$

$$\left[-\frac{e^{-\lambda x}}{\lambda}\right]_{0}^{\infty}=\frac{1}{\lambda}$$

אם נרצה לחשב את הפונקציה יוצרת המומנטים, נקבל:

$$M_{X}\left(t\right)=\mathbb{E}\left(e^{tX}\right)=\int\limits_{0}^{\infty}e^{tX}\lambda e^{-\lambda x}dx=$$

$$\lambda\int\limits_{0}^{\infty}e^{(t-\lambda)x}dx=\left[\frac{\lambda e^{(t-\lambda)x}}{t-\lambda}\right]_{0}^{\infty}$$

$$\overset{(*)}{=}\frac{\lambda}{\lambda-t}$$

$(*)$ התעניינו בדרך כלל בדברים שקורים בסביבת 0 (כי זאת הדרך לקבל את המומנטים, על ידי גזירה). ולכן נתבונן ב־$t<\lambda$ על מנת שדבר זה יתקיים.
אם כך, מתקיים (תזכורת בהמשך)

$$V(X)=M_{X}''(0)-\mathbb{E}(X)^{2}=\frac{2}{\lambda^{2}}-\frac{1}{\lambda^{2}}=\frac{1}{\lambda^{2}}$$

תזכורת

אם $M_X(t)$ מוגדרת בסביבה של 0, אזי $\mathbb{E}(X^n) = M_X^{(n)}(0)$.

לתוחלת ושונות של מ"מ **כלליים** יש אותן תכונות כמו בבדידים. אם $X \geq 0$ אזי $\mathbb{E}(X) \geq 0$.
כמו כן, אם $X \geq Y$ אזי $\mathbb{E}(X) \geq \mathbb{E}(Y)$ וגם $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ כשאגף ימין מוגדר וסופי.
נוסחת השונות הינה $V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
בנוסף, $V(X) \geq 0$ ו־$V(X) = 0$ אם"ם $X \overset{\text{a.s}}{=} c$.
יתר על כן, $V(aX+b) = a^2 V(X)$ ו־$\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$.
אם X, Y ב"ת אזי $\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ וגם $V(X+Y) = V(X) + V(Y)$.
בנוסף, כל הא"שים - מרקוב, צ'בישב וצ'רנוף עובדים כאן. גם שונות משותפת עובדת כאן.

הטענה $V(X) = 0$ אם"ם $X \overset{\text{a.s}}{=} c$ קשה, שכן היא מתקבלת מהטענה הבאה:
אם f אינטגרבילית אי שלילית ו־$\int_{-\infty}^{\infty} f(x)\,dx$ אזי יש $A \in \mathcal{B}_{\mathbb{R}}$ עם $\mathbb{P}_\lambda(A) = 0$ כך ש־$f|_{\mathbb{R}\setminus A} = 0$ (כאשר $\mathbb{P}_\lambda(A) = 0$ מידת לבג)

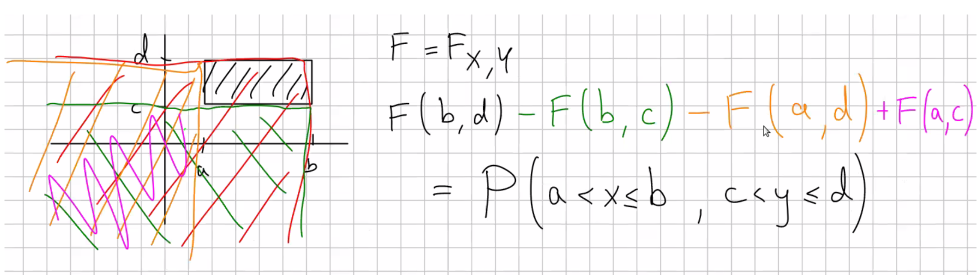
מספר משתנים רציפים

הגדרה

אם X, Y מ"מ על מ"ה $(\Omega, \mathcal{F}, \mathbb{P})$ נגדיר את פונקציית ההתפלגות המצטברת המשותפת שלהם:

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

היינו רוצים למצוא עבור $A \subseteq \mathbb{R}^2$ כללית, מהי $\mathbb{P}((X,Y) \in A)$.
קצת מורכב מדי, ולכן נצטמצם ל־$\left(\left\{\text{מלבנים מקבילים לצירים ב־}\mathbb{R}^2\right\}\right)$ $\mathcal{B}_{\mathbb{R}} = \sigma$.
$F_{X,Y}$ מספיקה להבין הסתברות של מלבנים (ולכן גם של כל $\mathcal{B}_{\mathbb{R}^2}$).
ולכן מתקיים:



על מנת לקבל את $\mathbb{P}(a < x \leq b, c < y \leq d)$, עלינו לקחת גבולות.
כעת, אפשר למצוא את פונקצית ההתפלגות המשותפת של X או של Y:

$$F_X(a) = \mathbb{P}(X \leq a) = \mathbb{P}(X \leq a \wedge Y \leq \infty) = \lim_{b\to\infty} F_{X,Y}(a,b)$$

מעתה, נתעסק בעיקר במ"מ עם צפיפות משותפת.

X,Y **בעלי צפיפות משותפת** $f_{X,Y}$ אם:

$$F_{X,Y}\left(a,b\right)=\in\int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}f_{X,Y}\left(x,y\right)dydx$$

לכל $a,b\in\mathbb{R}$.
מתקבל ש־X,Y רציפים ולכן:

$$\mathbb{P}\left(a\leq X\leq b,c\leq Y\leq d\right)=\int\limits_{a}^{b}\int\limits_{c}^{d}f_{X,Y}\left(x,y\right)dydx$$

טענה
אם ל־X,Y צפיפות משותפת, אזי **הצפיפות השולית** של X (f_X) מתקבלת על ידי $f_X\left(a\right)=\int\limits_{-\infty}^{\infty}f_{X,Y}\left(a,y\right)dy$
(במקרה הבדיד, מתקיים כי $\mathbb{P}_X\left(a\right)=\sum\limits_{b\in\mathbb{R}}p_{X,Y}\left(a,b\right)$).
הוכחה

$$\int\limits_{-\infty}^{a}\int\limits_{-\infty}^{\infty}f_{X,Y}\left(a,y\right)dy=\mathbb{P}\left(-\infty\leq x\leq a,-\infty\leq y\leq\infty\right)=F_X\left(a\right)$$

ולכן $\int\limits_{-\infty}^{\infty}f_{X,Y}\left(a,y\right)dy$ היא פונקציית צפיפות של X.

דוגמא
ניקח $f_{X,Y}\left(x,y\right)=\mathbb{1}_{I\times I}$.
אם $a<0$ או $b<0$ ברור כי $F_{X,Y}\left(0,b\right)=0$
ואז מתקיים:

$$\begin{aligned}F_{X,Y}\left(a,b\right)&=\int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}\mathbb{1}_{I\times I}\left(x,y\right)dy\ dx=\\&\int\limits_{-\infty}^{a}\int\limits_{0}^{\min(b,1)}\mathbb{1}_{I}\left(x\right)dydx\overset{\text{אינטגרל של קבוע}}{=}\\&\int\limits_{-\infty}^{a}\min\left(b,1\right)\cdot\mathbb{1}_{I}\left(x\right)dx=\\&\int\limits_{-\infty}^{\min(a,1)}\min\left(b,1\right)\cdot\mathbb{1}_{I}\left(x\right)dx=\\&\min\left(b,1\right)\min\left(a,1\right)\end{aligned}$$

בפרט עבור $a,b\in I$ מתקבל: $\mathbb{P}\left(0\le x\le a,0\le y\le b\right)=ab$ שהינו למעשה שטח המלבן.
ולכן לכל מלבן ב־$I\times I$ ההסתברות ליפול בו היא שטחו. לכן נקרא להתפלגות זו **אחידה** על $I\times I$.

נניח את משפט פוביני: סדר האינטגרציה לא משנה (על פי אינטגרביליות רימן):

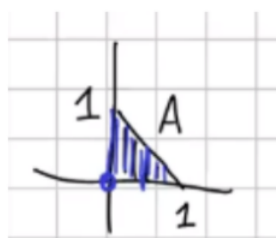
$$\int_{a}^{b}\int_{c}^{d}f\left(x,y\right)dydx=\int_{c}^{d}\int_{a}^{b}f\left(x,y\right)dxdy$$

עוד **עובדה** מתורת האינטגרציה - ל־$A\subseteq\mathbb{R}^{2}$ טובה מספיק (למשל קבוצה פתוחה וכן הלאה). יתקיים:

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

אינטגרל על קבוצה זה למעשה "הנפח מתחת לפונקציה $f_{X,Y}$ ב־A (בין גרף הפונקציה למישור $X-Y$).
אפשר לחשב את זה ע"י פרמטריזציה של A, אם למשל A היא המשולש מ־1 עד 1:

$$\int_{A}f\left(x,y\right)d\left(x,y\right)=\int_{0}^{1}\int_{0}^{1-x}f\left(x,y\right)dydx$$



אי תלות של משתנים מקריים
X,Y ב"ת אם לכל $A,B\in\mathcal{B}_{\mathbb{R}}$ המאורעות $\{X\in A\}$ ו־$\{Y\in B\}$ ב"ת.
כלומר:

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

טענה
אם ל־X,Y יש צפיפות, אזי הם ב"ת אם"ם $f_{X,Y}=F_{X}\left(x\right)f_{Y}\left(y\right)$.
(כמו במקרה הבדיד עם $p_{X,Y}$).
הוכחה
נניח כי $f_{X,Y}=F_{X}\cdot F_{Y}$. אזי לכל $a,b\in\mathbb{R}$ מתקיים:

$$\mathbb{P}\left(X\le a,Y\le b\right)=\int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}f_{X,Y}\left(x,y\right)dydx\overset{\text{אי תלות}}{=}$$
$$\int\limits_{-\infty}^{a}f_{X}\left(x\right)\int\limits_{-\infty}^{b}f_{Y}\left(y\right)dydx=$$
$$\int\limits_{-\infty}^{a}f_{X}\left(x\right)\mathbb{P}\left(Y\le b\right)dx=$$
$$\mathbb{P}\left(X\le a\right)\mathbb{P}\left(Y\le b\right)$$

כלומר, המאורעות $\{X\in(-\infty,a)\}$ ו־$\{Y\in(-\infty,b)\}$ ב"ת. מזה אפשר להגיע לכך שלכל $A,B\in\mathcal{B}_{\mathbb{R}}$ המאורעות $X\in A$ ו־$Y\in B$ ב"ת (על ידי איחודים בני מניה ומשלימים).

בכיוון השני, אם X,Y ב"ת, אנחנו צריכים להוכיח כי $f_X\left(x\right)f_Y\left(y\right)$ היא פונקציית צפיפות משותפת עבור X,Y:

$$\int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}f_{X,Y}\left(x,y\right)dydx=$$
$$\int\limits_{-\infty}^{a}f_{X}\left(x\right)\int\limits_{-\infty}^{b}f_{Y}\left(y\right)dydx=$$
$$\int\limits_{-\infty}^{a}f_{X}\left(x\right)\mathbb{P}\left(Y\le b\right)dx=\mathbb{P}\left(X\le a\right)\mathbb{P}\left(Y\le b\right)=$$
$$F_{X,Y}\left(a,b\right)$$

כאשר המעבר האחרון נובע מהעובדה כי X,Y ב"ת.

ליניאריות התוחלת

אם X,Y בעלי תוחלת (לאו דווקא ב"ת):

$$\mathbb{E}\left(X+Y\right)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left(x+y\right)f_{X,Y}\left(x,y\right)dydx=$$
$$\int\limits_{-\infty}^{\infty}x\int\limits_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)dydx+\int\limits_{-\infty}^{\infty}y\int\limits_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)dxdy=$$
$$\int\limits_{-\infty}^{\infty}xf_{X}\left(x\right)dx+\int\limits_{-\infty}^{\infty}yf_{Y}\left(y\right)dx=$$
$$\mathbb{E}\left(X\right)+\mathbb{E}\left(Y\right)$$

ובאופן כללי יותר:

$$\mathbb{E}\left(g\left(X,Y\right)\right)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}g\left(x,y\right)f_{X,Y}\left(x,y\right)dydx$$

עבור g מספיק טובה (רציפה למקוטעין למשל).

התפלגות נורמלית ומשפט הגבול המרכזי

הגדרה

אם f,g פונקציות, אזי הקונבולוציה שלהן הינה:

$$\left(f*g\right)\left(z\right)=\int\limits_{-\infty}^{\infty}f\left(x\right)g\left(z-x\right)dx$$

טענה

אם X,Y ב"ת בעלי צפיפות אזי:

$$f_{X+Y}\left(z\right)=\left(f_{X}*f_{Y}\right)\left(z\right)=\int\limits_{-\infty}^{\infty}f_{X}\left(x\right)f_{Y}\left(z-x\right)dx$$

הוכחה

אנחנו צריכים להוכיח כי:

$$F_{X+Y}\left(z\right)=\int\limits_{-\infty}^{z}\int\limits_{-\infty}^{\infty}f_{X}\left(x\right)f_{Y}\left(t-x\right)dxdt$$

נשתמש בפוביני:

$$\int_{-\infty}^{z}\int_{-\infty}^{\infty} f_X(x)\, f_Y(t-x)\, dxdt \overset{\text{פוביני ואי תלות}}{=}$$

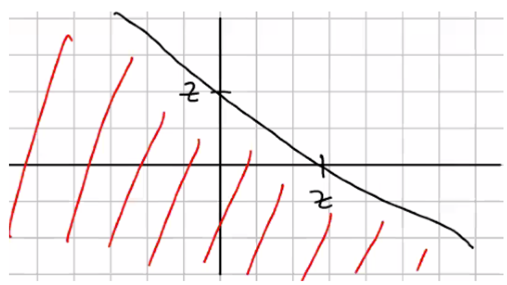
$$\int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_{X,Y}(x, t-x)\, dtdx \overset{y=t-x}{=}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_{X,Y}(x,y)\, dyd =$$

$$\int_{\substack{\text{משולש}\\ \{y\le z-x\}}} f_{X,Y}(x,y)\, d(x,y) = \mathbb{P}(Y \le z - X) =$$

$$\mathbb{P}(X+Y\le z)$$

כשבמשולש אנחנו מתכוונים ל:



שיעור מס' 25:

יום שלישי

12.01.21

דוגמא

אם $X, Y \sim \text{Unif}([0,1])$ ב"ת, מהי צפיפות $X+Y$?

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_0^1 1\text{גסא} f_Y(z-x)dx = \int_0^1 \mathbb{1}_I(z-x)\, dx$$

אם $0 \le z \le 1$ אזי נקבל $\int_0^z 1dx = z$, ואם $1 \le z \le 2$ נקבל $\int_{z-1}^{1} 1dx = 1-(z-1) = 2-z$.

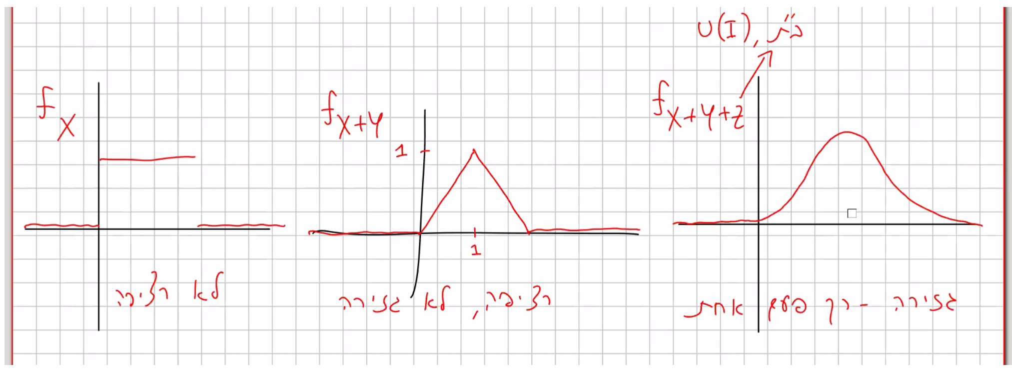
כלומר, קיבלנו סך הכל:

$$f_{X+Y}(z) = \begin{cases} z & 0\le z\le 1\\ 2-z & 1\le z\le 2\\ 0 & else\end{cases}$$

תמונתה הינה:

קיבלנו כי כעת פונקציית הצפיפיות רציפה ולא גזירה, לעומת המשתנה הבודד בו הפונקציה לא הייתה רציפה. האמת היא שכל שנפעיל יותר קונבולוציות, אזי נתקרב יותר ויותר לפונקציה גזירות (הוספת קונבולוציה נוספת תגרום לגזירות פעם אחת וכן הלאה).

כלומר, סך הכל:



מסתבר שכשלוקחים $A_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ כאשר $X_i \sim \text{Unif}\left(I\right)$ וב"ת - יש גבול לפונקציות ההתפלגות - $F_{A_n} \to F_Y$ למ"מ כלשהוא.

מה שבאמת מדהים שזה נכון לכל X_i שווי התפלגות, ב"ת ובעלי שונות סופית.

לא סתם כי F_{A_i} מתכנסות, אלא שהן תמיד מתכנסות לאותה התפלגות.

הגדרה

נאמר ש- X מתפלג נורמלית סטנדרטית - $X \sim N\left(0,1\right)$, אם:

$$f_X\left(x\right) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

(הוכחנו בתרגול כי $\int\limits_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ ונובע מכך ש-$\int\limits_{-\infty}^{\infty} f_x\left(x\right) dx = 1$).

מהי הפונקציה המצטברת $F_X\left(x\right)$?

$$\Phi\left(x\right) = F_X\left(x\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{\frac{-t^2}{2}} dt$$

אי אפשר לבטא את Φ ע"י פולינום, פונקציות טריגונומטריות, לוגריתמים, אקספוננט..

יש לה טור טיילור:

$$\Phi(x)=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j x^{2j+1}}{2^j(2j+1)j!}$$

משפט הגבול המרכזי (בהקדמה ראשונית בלבד)

סכום של משתנים מקריים בלתי תלויים שווי התפלגות מתכנס לנורמלית.

תוחלת של משתנה נורמלי

התוחלת הינה (אם קיימת):

$$\mathbb{E}\left(|X|\right)=\int\limits_{-\infty}^{\infty}|x|\,f_X\left(x\right)dx=$$
$$\frac{2}{\sqrt{2\pi}}\int\limits_{0}^{\infty}xe^{-\frac{x^2}{2}}dx\overset{t=\frac{x^2}{2}}{=}\frac{2}{\sqrt{2\pi}}\int\limits_{0}^{\infty}e^{-t}dt=\sqrt{\frac{2}{\pi}}<\infty$$

מאי הזוגיות (של $x\cdot f_x$) נקבל כי התוחלת הינה 0 על כל התחום.

שונות של משתנה נורמלי

השונות הינה:

$$\mathbb{E}\left(X^2\right)=\sqrt{\frac{2}{\pi}}\int\limits_{0}^{\infty}x^2e^{-\frac{x^2}{2}}=\sqrt{\frac{2}{\pi}}\int\limits_{0}^{\infty}x\cdot\left(xe^{-\frac{x^2}{2}}\right)dx=$$
$$\sqrt{\frac{2}{\pi}}\left[-xe^{\frac{-x^2}{2}}\right]_0^{\infty}-\underbrace{\int\limits_{0}^{\infty}xe^{-\frac{x^2}{2}}d}_{\text{חישבנו}}=\ldots=1$$

ולסיום נקבל כי $V\left(X\right)=\mathbb{E}\left(X^2\right)-\mathbb{E}\left(X\right)^2=1$

לכן מסמנים $N\left(0,1\right)$ - ביחס לתוחלת ולשונות.

הגדרה

נאמר ש-X מתפלג $N\left(\mu,\sigma^2\right)$ (נורמלי עם תוחלת μ ושונות σ^2) אם $\frac{X-\mu}{\sigma}\sim N\left(0,1\right)$.

(לכל משתנה עם תוחלת μ ושונות σ^2, $\frac{X-\mu}{\sigma}$ בעל תוחלת 0 ושונות 1. $\frac{X-\mu}{\sigma}$ נקרא התקנון (נורמליזציה) של X).

ראינו את פונקציית הצפיפות של $f_{g(x)}$ כש-g מונוטונית גזירה.

במקרה שלנו, $g\left(x\right)=\sigma x+\mu$.

$X\sim N\left(0,1\right)$ ו-$Y=g\left(X\right)$ - אזי $X=\frac{Y-\mu}{\sigma}$ באמת מתפלג $N\left(0,1\right)$.

נקבל:

$$f_Y\left(y\right)=\frac{f_X\left(g^{-1}\left(y\right)\right)}{g'\left(g^{-1}\left(y\right)\right)}$$

ולכן, נגיע להגדרה שקולה.

הגדרה שקולה

נאמר כי $Y\sim N\left(\mu,\sigma^{2}\right)$

$$f_{Y}(y)=\frac{f_{X}\left(\frac{y\cdot\mu}{\sigma}\right)}{\sigma}=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^{2}}$$

משפט הגבול המרכזי

אם $\{X_i\}_{i=1}^{\infty}$ משתנים מקריים בלתי תלויים ושווי התפלגות בעלי תוחלת μ ובעלי שונות σ^2, אזי נגדיר $A_n=\frac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$ ומתקיים כי $A_n\underset{n\to\infty}{\overset{d}{\to}}N\left(0,1\right)$.

הוכחה

בהמשך.

הגדרה

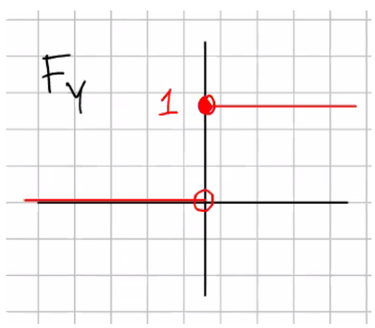
נאמר ש־Y $X_n\overset{d}{\to}Y$ (מתכנסת בהתפלגות מסוימת), אם $F_{X_{n}(t)}\overset{n\to\infty}{\to}F_{Y}\left(t\right)$ לכל t שהיא נקודת רציפות של F_Y.

שיעור מס' 26:
יום חמישי
14.01.21

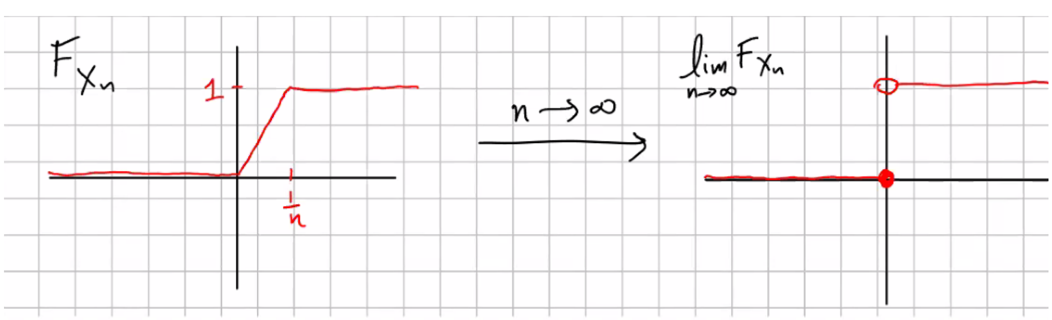
דוגמא (1)

יהיו $X_n\sim\text{Unif}\left(\left[0,\frac{1}{n}\right]\right)$ ו־$Y=0$.

מתקיים:

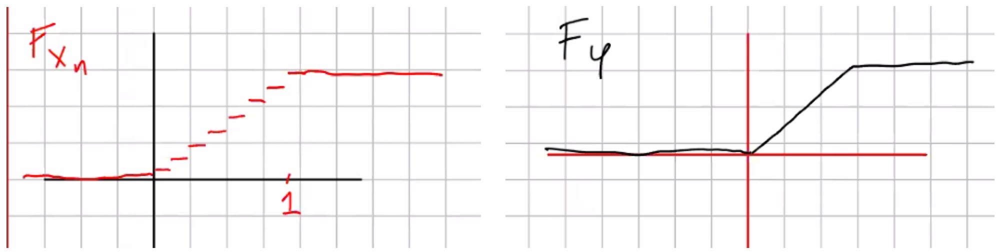


וגם:



כלומר, $F_{X_n}\left(t\right)\to F_{Y}\left(t\right)$ חוץ מאשר ב־$t=0$. עדיין $X_n\overset{d}{\to}Y$ כי 0 היא נקודת אי רציפות של F_Y.

דוגמא (2)
ניקח $X_n = \text{Unif}\left(\left\{\frac{k}{n} \mid k=0,\ldots,n\right\}\right)$ ו- $Y = \text{Unif}\left([0,1]\right)$.
כלומר:



לכל $0 \leq t \leq 1$ מתקיים כי:

$$F_{X_n}(t) = \frac{\lfloor nt \rfloor}{n} \overset{n\to\infty}{\to} t = F_Y(t)$$

מכיוון שהתכנסות בהתפלגות עוסקת רק בהתפלגות של X_n ו-Y (ולא בערכים שלהם על $\omega \in \Omega$) אפשר לכתוב גם:

$$\text{Unif}\left(\left\{\frac{k}{n} \mid k=0,\ldots,n\right\}\right) \overset{d}{\to} \text{Unif}\left([0,1]\right)$$

דוגמא (3)
$\text{Bin}\left(n, \frac{\lambda}{n}\right) \underset{n\to\infty}{\overset{d}{\to}} \text{Poi}\left(\lambda\right)$.

דוגמא (4)
אם $X_n \sim \text{Geo}\left(\frac{\lambda}{n}\right)$ אזי $\frac{1}{n}X_n \overset{d}{\to} \text{Exp}\left(\lambda\right)$

סוגים של התכנסות של מ"מ:
אם X_n, Y מ"מ על אותו מ"ה $(\Omega, \mathbb{R}, p)$, אזי:
$X_n \overset{d}{\to} Y$: התכנסות בהתפלגות - הגדרנו.
$X_n \overset{p}{\to} Y$: התכנסות בהסתברות:

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\left(\{\omega \in \Omega |\, |X_n(\omega) - Y(\omega)| \leqslant \varepsilon\}\right) = 1$$

$X_n \overset{a.s}{\to} Y$: התכנסות כמעט תמיד:

$$\mathbb{P}\left(\omega \in \Omega \mid X_n(\omega) \overset{n\to\infty}{\to} Y(\omega)\right) = 1$$

"התכנסות ממש תמיד":

$$\forall \omega \in \Omega : \quad X_n(\omega) \overset{n\to\infty}{\to} Y(\omega)$$

.

טענה

$X_n \overset{a.s}{\to} Y \Rightarrow X_n \overset{p}{\to} Y \Rightarrow X_n \overset{d}{\to} Y$ אבל $\not\Leftarrow$ בשני הכיוונים.

דוגמא לשימוש במשפט הגבול המרכזי

אם משתנה מסוים מתקבל כסכום של הרבה משתנים מקריים שווי התפלגות ובלתי תלויים, כל מדען פשוט יניח שהוא מתפלג נורמלית:

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \approx N(0,1) \Rightarrow$$
$$\sum_{i=1}^{n} X_i \approx N\left(\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = n\mu, V\left(\sum_{i=1}^{n} X_i\right) = n\sigma^2\right)$$

למשל, אם $X \sim \text{Bin}\left(1000, \frac{1}{2}\right)$ אזי $X \approx N(500, 250)$.

נסמן $Z \sim N(0,1)$ ומתקיים (השונות היא 250 כיוון ששונות של משתנה בינומי זה ריבוע ההסתברות כפול כמות ההטלות):

$$\mathbb{P}(450 < X < 550) = \mathbb{P}\left(\frac{-50}{\sqrt{250}} < \frac{X-500}{\sqrt{250}} < \frac{50}{\sqrt{250}}\right) \approx$$
$$\mathbb{P}\left(\frac{-50}{\sqrt{250}} < Z < \frac{50}{\sqrt{250}}\right) =$$
$$F_Z\left(\frac{50}{\sqrt{250}}\right) - F_Z\left(\frac{-50}{\sqrt{250}}\right) =$$
$$\Phi\left(\frac{50}{\sqrt{250}}\right) - \Phi\left(\frac{-50}{\sqrt{250}}\right) = 0.99843$$

מצ'בישב קיבלנו:

$$\mathbb{P}(450 < X < 550) \geq 0.9$$

והאמת היא:

$$\mathbb{P}(450 < X < 550) = \sum_{k=451}^{549} \left(\binom{1000}{k}\right) 2^{-1000} = 0.99827$$

משימה

מה נותנים צ'רנוף והופטינג?

משפט

$X_n \overset{d}{\to} Y$ אם (ורק אם) לכל $g \in C_c^{\infty}(\mathbb{R})$ ‏(c - "תומך קומפקטי" - $\text{Supp}(g)$ חסום ו־∞ גזירה אינסוף פעמים) מתקיים:

$$\mathbb{E}\left(g\left(X_n\right)\right) \overset{n\to\infty}{\rightarrow} \mathbb{E}\left(g\left(Y\right)\right)$$

הערה

אותו משפט נכון עם $C_c^0\left(\mathbb{R}\right)$, $C_c^7\left(\mathbb{R}\right)$ וכו'.
למעשה, אנחנו נצטרך את $C_c^3\left(\mathbb{R}\right)$.

הוכחה

אם ניקח $g=\mathbb{1}_{(-\infty,t]}$ אזי $g\left(X_n\right)$ הוא ברנולי ולכן:

$$\mathbb{E}\left(g\left(X_n\right)\right)=\mathbb{P}\left(g\left(X_n\right)=1\right)=\mathbb{P}\left(X_n\leqslant t\right)=F_{X_n}(t)$$

לכן אם g הייתה C_c^∞ (חלקה עם תומך חסום) היינו מסיימים:

$$F_{X_n}\left(t\right)=\mathbb{E}\left(g\left(X_n\right)\right)\underset{\text{נתון}}{\overset{n\to\infty}{\rightarrow}}\mathbb{E}\left(g\left(Y\right)\right)=F_Y\left(t\right)$$

ולכן $X_n\overset{d}{\rightarrow}Y$ אבל לצערנו $g\notin C_c^\infty\left(\mathbb{R}\right)$.

יהי $0<\varepsilon$. נבחר $m\in\mathbb{R}$ עם $F_Y\left(m\right)<\varepsilon$ וניקח $g=\mathbb{1}_{(m,t]}$, כעת התומך חסום.
נקבל:

$$\begin{aligned}
&F_Y\left(t\right)-\varepsilon<F_Y\left(t\right)-F_Y\left(m\right)=\\
&\mathbb{E}\left(g\left(Y\right)\right)=\\
&\lim_{n\to\infty}\mathbb{E}\left(g\left(X_n\right)\right)\overset{\text{היינו רוצים}}{=}\\
&\lim_{n\to\infty}F_{X_n}\left(t\right)-F_{X_n}\left(m\right)\leq\\
&\liminf_{n\to\infty}F_{X_n}\left(t\right)
\end{aligned}$$

($\liminf$ - הגבול החלקי התחתון).
כתוב שם היינו רוצים, כי g לא חלקה ואפילו לא רציפה.
עכשיו ניקח M עם $F_Y\left(M\right)>1-\varepsilon$ ואז נקבל, עבור $g=\mathbb{1}_{(t,M]}$:

$$\begin{aligned}
&\left(1-\varepsilon\right)-F_Y\left(t\right)<F_Y\left(m\right)-F_Y\left(t\right)=\\
&\mathbb{E}\left(g\left(Y\right)\right)\overset{\text{כמעט}}{=}\\
&\lim_{n\to\infty}\mathbb{E}\left(g\left(X_n\right)\right)\overset{\text{חלקה}}{=}\\
&\lim_{n\to\infty}F_{X_n}\left(M\right)-F_{X_n}\left(t\right)\leq\\
&\liminf_{n\to\infty}1-F_{X_n}\left(t\right)
\end{aligned}$$

לכן (נפחית 1 מכל אגף) $\limsup\limits_{n\to\infty} F_{X_n} \leq F_Y(t)+\varepsilon$. כלומר, סך הכל:

$$F_Y(t)-\varepsilon < \liminf_{n\to\infty} F_{X_n}(t) \leq \limsup_{n\to\infty} F_{X_n}(t) < F_Y(t)+\varepsilon$$

לכל $\varepsilon > 0$ ולכן $\lim\limits_{n\to\infty} F_{X_n}(t) = F_Y(t)$.

אבל g לא חלקה, וגם לא השתמשנו בכך ש־F_Y רציפה ב־t.

תרגיל: הפכו את g לחלקה (או רציפה) שם תצטרכו את ההנחה ־ש־F_Y רציפה ב־t.

משפט

סכום של מ"מ נורמליים ב"ת הוא נורמלי.

אם $X\sim N\left(\mu,\sigma^2\right)$ ו־$Y\sim\left(\nu,\tau^2\right)$ ו־X,Y אזי $X+Y\sim N\left(\mu+\nu,\sigma^2+\tau^2\right)$.

הוכחה

תמיד אפשר להפעיל פונקציה אפינית ($x\to ax+b$) על X,Y.

ולכן נקבל כי $X+Y$ הזזה אפינית של נורמלי היא עדיין נורמלית.

נשתמש בזה על מנת להניח כי $\mu=\nu=0$ ו־$\sigma^2+\tau^2=1$ (חילקנו את X,Y בשורש סכום שונותיהם).

עכשיו אנחנו צריכים להוכיח כי $X+Y\sim N(0,1)$ ־ כלומר $f_{X+Y}(z)=\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$.

נקבל כי $f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}}$(ואותו דבר עבור y). כיוון שהם ב"ת נקבל:

$$\begin{aligned}
&f_{X+Y}(z)=\\
&\int_{-\infty}^{\infty} f_X(x)\,f_Y(z-x)\,dx=\\
&\frac{1}{2\pi\sigma\tau}\int_{-\infty}^{\infty} e^{-\frac{x^2}{\sigma^2}-\frac{(z-x)^2}{\tau}}dx=\\
&\frac{1}{2\pi}e^{-\frac{z^2}{2}}\cdot\frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}-\frac{(z-x)^2}{2(1-\sigma^2)}+\frac{z^2}{2}}dx=\\
&\underbrace{\frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{\left(x-z\sigma^2\right)^2}{2\sigma^2(1-\sigma^2)}}dx}_{(*)}\frac{1}{2\pi}e^{-\frac{z^2}{2}}.
\end{aligned}$$

למעשה $(*)$ הינה צפיפות של משתנה נורמלי $1=\left(N\left(Z\sigma^2,\sigma^2\left(1-\sigma^2\right)\right)\right)$, ולכן קיבלנו מה שרצינו.

הוכחת משפט הגבול המרכזי (עבור המקרה בו $\mathbb{E}\left(\left|X_i^3\right|\right)<\infty$)

שיעור מס' 27:
יום שלישי
19.01.21

אפשר להניח מראש כי $\mathbb{E}(X)=0$ ו־$V(X)=1$ (כש־X הוא כל אחד מה־X_i . זה אפשרי כי ניתן לתקנן) ואז צ"ל כי $\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}X_i\xrightarrow{d}N(0,1)$.

ראינו שדבר זה ינבע אם נראה שלכל $g\in C_c^{\infty}\left(\mathbb{R}\right)$ מתקיים כי $\mathbb{E}\left(g\left(N\right)\right) \overset{n\to\infty}{\to} \mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}X_i\right)\right)$ כש־N משתנה נורמלי סטנדרטי.

לינדברג הציע לקחת סדרה של מ"מ נורמליים מתוקנים $\left\{N_i\right\}_{i=1}^{\infty}$ שהם ב"ת ביניהם וגם ב"ת ב־X_i.

נשים לב כי $\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}N_i\sim N\left(0,1\right)$ (כי סכום של משתנים נורמליים הוא נורמלי, השונות של כל אחד היא 1, ולכן יש לחלק ב־$\sqrt{n}$ על מנת "לנרמל").

אם כך, קיבלנו כי $\mathbb{E}\left(g\left(N\right)\right)=\mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}N_i\right)\right)$. כלומר, מספיק להראות כי:

$$\mathbb{E}\left(g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i\right)-g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}N_i\right)\right)\overset{n\to\infty}{\to}0$$

נעשה את זה טלסקופית:

$$\begin{gathered}=\mathbb{E}\left(g\left(\tfrac{1}{\sqrt{n}}\left(\textstyle\sum_{i=1}^{n}X_i\right)\right)-g\left(\tfrac{1}{\sqrt{n}}\left(\left(\textstyle\sum_{i=1}^{n-1}X_i\right)+N_n\right)\right)\right)+\\ \mathbb{E}\left(g\left(\tfrac{1}{\sqrt{n}}\left(\left(\textstyle\sum_{i=1}^{n-1}X_i\right)+N_n\right)\right)-g\left(\tfrac{1}{\sqrt{n}}\left(\left(\textstyle\sum_{i=1}^{n-2}X_i\right)+N_{n-1}+N_n\right)\right)\right)+\\ +\ldots+\\ \mathbb{E}\left(g\left(\tfrac{1}{\sqrt{n}}\left(X_1+\textstyle\sum_{i=2}^{n}N_i\right)\right)-g\left(\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^{n}N_i\right)\right)\end{gathered}$$

פרקנו את הביטוי ל־n מחוברים. צריך להראות שגם סכומם שואף לאפס. נראה שכל אחד מהם חסום ב־$\frac{C}{n^{\frac{3}{2}}}$ וזה ייתן מה שרצינו.

כל הביטויים דומים ומתנהגים אותו הדבר, ולכן נתמקד בראשון.

נגדיר $Z=\frac{X_1+\ldots+X_{n-1}}{\sqrt{n}}$ ואז הביטוי יהיה:

$$\mathbb{E}\left(g\left(Z+\frac{\sum\limits_{i=1}^{n}X_i}{\sqrt{n}}\right)-g\left(\frac{\sum\limits_{i=1}^{n-1}X_i}{\sqrt{n}}+\frac{N_n}{\sqrt{n}}\right)\right)$$

ועל ידי טיילור מתקיים:

$$g\left(Z+\frac{X_n}{\sqrt{n}}\right)=g\left(Z\right)+g'\left(Z\right)\frac{X_n}{\sqrt{n}}+\frac{g''\left(Z\right)}{2}\frac{X_n^2}{n}+R$$

כאשר:

$$\left|R\right|=\left|\frac{g'''\left(3\right)}{6}\cdot\frac{X_n^3}{n^{\frac{3}{2}}}\right|\le C\frac{X_n^3}{n^{\frac{3}{2}}}$$

האי שוויון האחרון מגיע מכך ש־$g\in C_c^3\left(\mathbb{R}\right)$ ולכן g''' חסומה (רציפה על קטע קומפקטי).

באותו אופן:

$$g\left(Z+\frac{N_n}{\sqrt{n}}\right)=g(Z)+g'(Z)\frac{N_n}{\sqrt{n}}+\frac{g''(Z)}{2}\frac{N_n^2}{n}+R'$$

כאשר $|R'| \le C\frac{N_n^3}{n^{\frac{3}{2}}}$.
אבל, כיוון ש־X_n ו־Z ב"ת, מתקבל:

$$\mathbb{E}\left(g\left(Z+\frac{X_n}{\sqrt{n}}\right)\right)=\mathbb{E}(g(Z))+\mathbb{E}\left(g'(Z)\right)\frac{\mathbb{E}X}{\sqrt{n}}+\frac{\mathbb{E}\left(g''(z)\right)}{2}\frac{\mathbb{E}X^2}{n}+\mathbb{E}R$$

ואותו הדבר:

$$\mathbb{E}\left(g\left(Z+\frac{X_n}{\sqrt{n}}\right)\right)=\mathbb{E}(g(Z))+\mathbb{E}\left(g'(Z)\right)\frac{\mathbb{E}N}{\sqrt{n}}+\frac{\mathbb{E}\left(g''(z)\right)}{2}\frac{\mathbb{E}N^2}{n}+\mathbb{E}R$$

ולכן, תוחלת ההפרש שלהם:

$$\left|\mathbb{E}\left(g\left(Z+\frac{X_n}{\sqrt{n}}\right)-g\left(Z+\frac{N_n}{\sqrt{n}}\right)\right)\right|\le\mathbb{E}\left|R\right|+\mathbb{E}\left|R'\right|\le C\frac{\mathbb{E}\left(\left|X^3\right|\right)+\mathbb{E}\left(\left|N^3\right|\right)}{n^{\frac{3}{2}}}$$

אותו ניתוח יעבוד לכל גורם בסכום הטלסקופי, ולכן נקבל:

$$\left|\mathbb{E}\left(g\left(\frac{\sum\limits_{i=1}^{n}X_i}{\sqrt{n}}\right)-g\left(\frac{\sum\limits_{i=1}^{n}N_i}{\sqrt{n}}\right)\right)\right|\le C\frac{\mathbb{E}\left(\left|X^3\right|\right)+\mathbb{E}\left(\left|N^3\right|\right)}{\sqrt{n}}$$

כי יש n מחוברים בסכום הנ"ל.
נוכל להבחין כי $\infty > \frac{\sqrt{8}}{\sqrt{\pi}} = \mathbb{E}\left(\left|N^3\right|\right)$ ו־$\mathbb{E}\left(\left|X^3\right|\right)$ סופי. סך הכל קיבלנו כי:

$$\mathbb{E}\left(g\left(\frac{\sum^n X_i}{n}\right)\right)\overset{n\to\infty}{\to}\mathbb{E}(g(N))$$

ולכן $\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}X_i\overset{d}{\to}N\left(0,1\right)$.

שיעור מס' 28:
יום חמישי
21.01.21

התכנסויות בהסתברות והתכנסות כמעט תמיד

$X_n\overset{p}{\to}X$ - "התכנסות בהסתברות" אם:

$$\forall\varepsilon>0:\mathbb{P}\underbrace{\left(\left|X_n-X\right|>\varepsilon\right)}_{\{\omega\in\Omega||X_n(\omega)-X(\omega)|>\varepsilon\}}\overset{n\to\infty}{\to}D$$

דוגמאות:

(א) יהיו $X_n\sim\text{Ber}\left(p_n\right)$ ב"ת. מתי $X_n\overset{p}{\to}0$?
אם $\varepsilon>1$ אזי תמיד $\mathbb{P}\left(\left|X_n-0\right|>\varepsilon\right)=0$.
אם $0<\varepsilon<1$:

$$\mathbb{P}\left(|X_n-o|>\varepsilon\right)=\mathbb{P}\left(X_n=1\right)=p_n$$

לכן $X_n\xrightarrow{p}0$ אם ורק אם $\lim\limits_{n\to\infty}p_n=0$, ובדומה $X_n\xrightarrow{p}1$ אם ורק אם $\lim\limits_{n\to\infty}p_n=1$.

(ב) החוק החלש של המספרים הגדולים: אם $\{x_n\}_{n=1}^{\infty}$ iid בעלי תוחלת סופית μ אזי:

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{p}\mu$$

הוכחנו זאת (באמצעות צ'בישב) תחת ההנחה ש־$\mathbb{E}\left(X^2\right)<\infty$.

(ג) דוגמא דומה ־ אם x_n סדרה עם $\mathbb{E}\left(X_n\right)\overset{n\to\infty}{\to}\alpha$ ו־$V\left(X_n\right)\overset{n\to\infty}{\to}0$ אזי $X_n\xrightarrow{p}\alpha$ (הוכיחו באמצעות צ'בישב).

(ד) ניקח $\left(\Omega=I,\mathcal{B}_I,\mathbb{P}_\lambda\right)$ (התפלגות אחידה על $[0,1]$) ונגדיר $X_n=\frac{\lfloor\omega\cdot10^n\rfloor}{10^n}$ (כלומר n הספרות הראשונות של ω).
אזי $X_n\xrightarrow{p}\text{Id}$ כי:

$$\forall\varepsilon>0:\quad|X_n(n)-\omega|\le\frac{1}{10^n}<\varepsilon$$

החל מ־n מסוים, לכל ω. ובפרט $\mathbb{P}\left(|X_n-X|>\varepsilon\right)\overset{n\to\infty}{\longrightarrow}0$.

נזכיר:$X_n\xrightarrow{a.s}X$ ־ "התכנסות כמעט תמיד":

$$\mathbb{P}(\underbrace{X_n\to X}_{\{\omega|X_n(\omega)\overset{n\to\infty}{\to}X(\omega)\}}=1$$

כעת, נתבונן בדוגמאות לעיל:
(א) $X_n\sim\text{Ber}\left(p_n\right)$ ב"ת:
$X_n\xrightarrow{a.s}0$ אם"ם $\sum\limits_{n=1}^{\infty}p_n<\infty$ (אולי נוכיח בהמשך).
(ב) החוק החזק של המספרים הגדולים: אם X_n iid בעלי תוחלת סופית μ אזי:

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

(לא נוכיח־ אפשר לקרוא בספר).

(ג) שוב $X_n=\frac{\lfloor\omega\cdot10^n\rfloor}{10^n}\xrightarrow{a.s}X$ כי:

$$\forall\omega\in\Omega:\frac{\lfloor\omega\cdot10^n\rfloor}{10^n}\overset{n\to\infty}{\longrightarrow}\omega$$

(נבחין כי $a.s$ היא התכנסות נקודתית מלבד בפונקציות ממידה 0)

טענה

$X_n \overset{p}{\to} X$ לא גורר $X_n \overset{a.s}{\to} X$.

הוכחה

נגדיר סדרה של קטעים:

$$\begin{aligned}
&I_1=[0,1]\\
&I_2=\left[0,\tfrac{1}{2}\right),I_3=\left(\tfrac{1}{2},1\right]\\
&I_4=\left[0,\tfrac{1}{4}\right],\ldots\ldots,I_7=\left[\tfrac{3}{4},1\right]
\end{aligned}$$

ניקח $(\Omega=I,\mathcal{B}_I,\mathbb{P}_\lambda)$ ואז $X_n(\omega)$ נראית:

$$1,(0,1)\text{ or }(1,0),0,1,0,0,0,1,0\ldots$$

לכן לכל ω מתקיים כי $X_n(\omega)$ אינה מתכנסת. אם כך $X_n \overset{a.s}{\not\to} X$ עבור שום X.

אבל $X_n \overset{p}{\to} 0$, כי מתקיים:

$$\forall 1>\varepsilon>0:\quad P\left(|X_n-0|>\varepsilon\right)=\text{len}\left(I_n\right)\overset{n\to\infty}{\to}0$$

(מכך שהדוגמא היא לא פשוטה במיוחד, עולה שההסתברויות האלו די דומות בסך הכל).

נגדיר כמה מושגים. אם $\{A_n\}_{n=1}^{\infty}$ סדרה של מאורעות, אזי:

$$\begin{aligned}
\overline{\lim_{n\to\infty}}A_n&:=\{\omega\mid(*)\}=\{\omega\mid\{**\}\}=\bigcap_{i=1}^{\infty}\bigcup_{j\geq i}A_j\\
\underline{\lim_{n\to\infty}}A_n&:=\{\omega\mid(***)\}=\{\omega\mid\{****\}\}=\bigcup_{i=1}^{\infty}\bigcap_{j\geq i}A_j
\end{aligned}$$

כאשר $(*)$ מוגדר על ידי "ω שייך ל־A_n עבור ∞ n־ים". ו$(**)$ מוגדר על ידי "בכל זנב $\{A_j\}_{j\geq i}$ עדיין מופיע ω". ו־$(***)$ מוגדר על ידי ω שייך לכל A_n למעט מספר סופי" ו־$(****)$ מוגדר על ידי "יש זנב ש־ω מופיע בכל איבר".

אבחנות: $\underline{\lim}A_n\subseteq\overline{\lim}A_n$.

כמו כן, דה מורגן:

$$\left(\overline{\lim}A_n\right)^c=\underline{\lim}\left(A_n^c\right)$$

הלמה של פאטו

$\mathbb{P}\left(\overline{\lim}A_n\right) \leq \overline{\lim}\mathbb{P}\left(A_n\right)$ (מדה מורגן נובע גם כי $\mathbb{P}\left(\underline{\lim}A_n\right) \geq \underline{\lim}\mathbb{P}\left(A_n\right)$.

הוכחה

$$\mathbb{P}\left(\underline{\lim}A_n\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty}\bigcap_{j\geq i}A_j\right) \overset{(*)}{=} \lim_{i\to\infty}\mathbb{P}\left(\bigcap_{j\geq i}A_j\right) \leq \lim_{i\to\infty}\left(\inf_{j\geq i}\mathbb{P}\left(A_j\right)\right) \overset{\text{הגדרה}}{=} \underline{\lim}_{n\to\infty}\mathbb{P}\left(A_n\right)$$

$(*)$ אם $i \geq k$ אזי $\bigcap_{j\geq k}A_j \supseteq \bigcap_{j\geq i}A_j$ ולכן זו סדרה מונטונית ב-i. הוכחנו שהסתברות של איחוד על סדרה מונטונית הוא גבול הסתברויות איבריה.

טענה

$X_n \overset{a.s}{\to} X$ גורר $X_n \overset{p}{\to} X$.

טענת עזר

$X_n \overset{a.s}{\to} X$ אם"ם $\forall\varepsilon\mathbb{P}\left(\underline{\lim}A_n^{\varepsilon}\right) = 1$.

הוכחה

נקבע $\omega \in \Omega$. אזי $X_n\left(\omega\right) \overset{n\to\infty}{\to} X\left(\omega\right)$. דבר זה שקול לכך:

$$\forall\varepsilon\exists N\ \ \forall n \geq N : |X_n(\omega) - X(\omega)| \leqslant \varepsilon \Leftrightarrow \omega \in A_n^{\varepsilon}$$

ומכך עולה כי:

$$\omega \in \bigcap_{n\geq N}A_n^{\varepsilon} \Rightarrow \exists N \Rightarrow \omega \in \bigcup_{N=1}^{\infty}\bigcap_{n\geq N}A_n^{\varepsilon}$$

ומכך עולה כי ההגדרה שקולה ל-$\omega \in \underline{\lim}A_n^{\varepsilon}$.

הוכחה

נגדיר $A_n^{\varepsilon} = \left\{\omega \mid \left|X_n\left(\omega\right) - X\left(\omega\right)\right| \leq \varepsilon\right\}$.

כעת, מתקיים:

$$X_n \overset{p}{\to} X \ \ \Leftrightarrow\ \ \forall\varepsilon : \mathbb{P}\left(A_n^{\varepsilon}\right) \overset{n\to\infty}{\longrightarrow} 1 \ \ \Leftrightarrow\ \ \ \forall\varepsilon : \underline{\lim}\mathbb{P}\left(A_n^{\varepsilon}\right) = 1$$

(המעבר האחרון נובע מכך שאם $\mathbb{P}\left(A_n^{\varepsilon}\right) = 1$ אזי גם הלימאינפ הוא 1. בכיוון השני, זו סדרה שחסומה על ידי 1, ואם זהו הגבול החלקי הקטן ביותר, סימן שזהו הגבול.

כלומר, $X_n\left(\omega\right) \overset{n\to\infty}{\to} X\left(\omega\right)$ אם"ם לכל $\omega \in \underline{\lim}A_n^{\varepsilon}$.

אם $X_n \overset{a.s}{\to} X$ אזי לכל $\varepsilon > 0$ מתקיים, מטענת העזר כי:

$$X_n \overset{a.s}{\to} X \Rightarrow \forall\varepsilon\ \ \mathbb{P}\left(\underline{\lim}A_n^{\varepsilon}\right) = 1 \overset{\text{פאטו}}{\Rightarrow} \underline{\lim}\mathbb{P}\left(A_n^{\varepsilon}\right) = 1 \Rightarrow X_n \overset{p}{\to} X$$

האם $X_n \xrightarrow{d} X$ גורר כי $X_n \xrightarrow{p} X$? וודאי שלא. (כמו שבהטלת 7 מטבעות מס' הראשים מתפלג כמו מס' הזנבות, אבל הם תמיד שונים).
אמנם, ההפך כן נכון וראינו בתרגול.

טענה

$X_n \xrightarrow{p} X$ גורר כי $X_n \xrightarrow{d} X$.

הוכחה

נניח כי a נקודת רציפות של F.
אנחנו צ"ל כי $F_{X_n}(a) \to F_X(a)$.
נקבע ε. מרציפות, קיימת $\delta > 0$ כך ש:

$$|F_X(t) - F_X(a)| < \varepsilon \Leftarrow |t-a| < \delta$$

ואז:

$$\begin{aligned} F_{X_n}(a) = \mathbb{P}(x_n \leqslant a) &\overset{(*)}{\leq} \mathbb{P}(x \leqslant a+\delta) \text{or } \mathbb{P}(|x_n - x| > \delta) \\ &\overset{\text{חסם האיחוד}}{\leq} F_X(a+\delta) + \mathbb{P}(|X_n - X| > \delta) \\ &\leq F_X(a) + \varepsilon + \mathbb{P}\left(\left(A_n^\delta\right)^c\right) \underset{X_n \xrightarrow{p} X}{\overset{n\to\infty}{\to}} F_X(a) + \varepsilon \end{aligned}$$

$(*)$ אם $X > a+\delta$ וגם $|X_n - X| < \delta$ אזי $X > a$.
ניקח $\overline{\lim}$ משני האגפים, ונקבל:

$$\overline{\lim} F_{X_n}(a) \leqslant F_X(a) + \varepsilon$$

זה היה נכון לכל $\varepsilon > 0$ ולכן $\overline{\lim} F_{X_{n(a)}} \leq F_X(a)$.
באופן אנלוגי (ע"י $a-\delta$) מראים ש-$\underline{\lim} F_{X_n}(a) \geq F_X(a)$, ולכן $F_X(a) \overset{n\to\infty}{\to} F_{X_n}(a)$.
סך הכל הראינו כי $\xrightarrow{a,s} \Rightarrow \xrightarrow{p} \Rightarrow \xrightarrow{d}$.
הכיוונים ההפוכים אינם נכונים.

משפטי בורל קנטלי

תהי A_n סדרת מאורעות, אזי:
(I) אם $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$ אזי $\mathbb{P}\left(\overline{\lim} A_n\right) = 0$.
(II) אם $\{A_n\}$ ב"ת (מספיק ב"ת בזוגות), אזי $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \Rightarrow \mathbb{P}\left(\overline{\lim} A_n\right) = 1$.

הוכחה

(I):

$$P\left(\lim A_n\right) \overset{\text{מונוטוניות}}{=} \lim_{i\to\infty} \mathbb{P}\left(\bigcup_{j\geq i} A_j\right) \overset{\text{חסם האיחוד}}{\leq} \lim_{i\to\infty} \sum_{j\geq i} \mathbb{P}\left(A_j\right) \overset{\text{סכום על זנב של טור מתכנס}}{=} 0$$

(II) מתקיים:

$$\mathbb{P}\left(\overline{\lim} A_n\right) \overset{\text{דה מורגן}}{=} 1-\mathbb{P}\left(\underline{\lim} A_n^c\right) \overset{\text{מונוטוניות}}{=} 1-\lim_{i\to\infty} \mathbb{P}\left(\bigcap_{j\geq i} A_j^c\right)$$

מספיק שנראה כי אפילו לכל i מתקיים כי $\mathbb{P}\left(\bigcap\limits_{j\geq i} A_j^c\right)=0$ ואכן:

$$\mathbb{P}\left(\bigcap_{j\geq i} A_j\right) = \prod_{j=i}^{\infty} \mathbb{P}\left(A_j^c\right) = \prod_{j=1}^{\infty}\left(1-\mathbb{P}\left(A_j\right)\right) \overset{1-X\leq e^{-X}}{\leq}$$

$$\prod_{j=i}^{\infty} e^{-\mathbb{P}\left(A_j\right)} = e^{-\sum\limits_{j=1}^{\infty}\mathbb{P}\left(A_j\right)} = e^{-\infty} = 0$$

עכשיו בדקו כי אם $X_n \sim \text{Ber}\left(p_n\right)$ ב"ת אזי $X_n \overset{a.s}{\to} X$ אמ"ם $\sum\limits_{n=1}^{\infty} p_n < \infty$.