האוניברסיטה העברית בירושלים
המחלקה לסטטיסטיקה
הסתברות ותהליכים מקריים
מורה הקורס : עופר קלע

אי תלות:
נניח כי $A_\alpha \in \mathcal{F}$ לכל $\alpha \in \Lambda$. אז $\{A_\alpha | \alpha \in \Lambda\}$ נקראים בלתי תלויים אם לכל תת-קבוצה סופית $J \subset \Lambda$ מתקיים כי

$$P\left(\bigcap_{\alpha \in J} A_\alpha\right) = \prod_{\alpha \in J} P(A_\alpha)$$

בקורסים קודמים ראיתם כי אם A_1, A_2 בלתי תלויים אז A_1, A_2^c בלתי תלויים (ולכן גם A_1^c, A_2 או A_1^c, A_2^c למעשה מתקיימת)
טענה: $\{A_\alpha | \alpha \in \Lambda\}$ בלתי תלויים אם ורק אם

$$\{A_\alpha^c | \alpha \in \Lambda_c\} \cup \{A_\alpha | \alpha \in \Lambda \setminus \Lambda_c\}$$

בלתי תלויים לכל $\Lambda_c \subset \Lambda$.
שימו לב כי אם Λ_c ריקה אז מקבלים את אוסף המאורעות המקורי ואם היא שווה ל-Λ אז זה בעצם כמו לקחת את אוסף כל המשלימים.
הוכחה: ניקח $J \subset \Lambda$ סופי עם לפחות שלושה איברים. כאשר יש איבר אחד או שני איברים אז אנחנו כבר יודעים מה לעשות. נניח כי $\beta \in J$. אז שני המאורעות $A = A_\beta$ ו-$B = \cap_{\alpha \in J \setminus \{\beta\}} A_\alpha$ הם בלתי תלויים. זאת מכיון ש-

$$P(A \cap B) = P\left(\bigcap_{\alpha \in J} A_\alpha\right) = \prod_{\alpha \in J} P(A_\alpha) = P(A_\beta) \cdot \prod_{\alpha \in J \setminus \{\beta\}} P(A_\alpha) = P(A)P(B)$$

כאשר השוויון האחרון מתקיים מכיוון ש-$J \setminus \{\beta\}$ היא קבוצה סופית של אינדכסים ולכן מהגדרת אי התלות נובע כי

$$P(B) = P\left(\bigcap_{\alpha \in J \setminus \{\beta\}} A_\alpha\right) = \prod_{\alpha \in J \setminus \{\beta\}} P(A_\alpha)$$

מכיוון ש-$A = A_\beta$ ו-$B = \cap_{\alpha \in J \setminus \{\beta\}} A_\alpha$ בלתי תלויים אז גם $A^c = A_\beta^c$ ו-$B = \cap_{\alpha \in J \setminus \{\beta\}} A_\alpha$ בלתי תלויים ולכן

$$P\left(A_\beta^c \cap \bigcap_{\alpha \in J \setminus \{\beta\}} A_\alpha\right) = P(A_\beta^c) P\left(\bigcap_{\alpha \in J \setminus \{\beta\}} A_\alpha\right) = P(A_\beta^c) \prod_{\alpha \in J \setminus \{\beta\}} P(A_\alpha)$$

קיבלנו איפה כי אם נחליף את A_β ב-A_β^c ואת השאר נשאיר כמו שהם נקבל כי ההסתברות של החיתוך שווה למכפלת ההסתברויות.
המסקנה היא איפה שאם ניקח β כלשהו וניקח קבוצת אינדכסים סופית J שמכילה את β אז החלפת A_β ב- A_β^c שומרת על התכונה שההסתברות של חיתוך המאורעות עם אינדכסים ב-J שווה למכפלת ההסתברויות של מאורעות אילה. ברור כי אם J אינו מכיל את β אז החלפת A_β ב-A_β^c לא

משנה כלום לגבי המאורעות עם אינדכסים ב-J ולכן גם אז ההסתברות של החיתוך שווה למכפלת ההסתברויות.
המסקנה מכך היא שאם נחליף את אחד מהמאורעות הבלתי תלויים $\{A_\alpha|\alpha\in\Lambda\}$ במשלים שלו אז שוב נקבל אוסף של מאורעות בלתי תלויים. זה נכון עבור Λ ולכן זה גם נכון עבור כל תת קבוצה של Λ (למשל J כפי שנציין עוד מעט). מכאן שאפשר להחליף עוד מאורע ששונה מהמאורע הראשון במשלים שלו ושוב לקבל אוסף של מאורעות בלתי תלויים. ככה אפשר להמשיך ולקבל כי אם נחליף איזשהו אוסף סופי של מאורעות מתוך $\{A_\alpha|\alpha\in\Lambda\}$ (או כל תת אוסף) במשלימים שלהם אז שוב נקבל אוסף מאורעות בלתי תלויים. מה לגבי החלפת אוסף כלשהו של מאורעות (לא בהכרח סופי)?
אם כן, נסתכל על האוסף

$$\{A_\alpha^c|\alpha\in\Lambda_c\}\cup\{A_\alpha|\alpha\in\Lambda\setminus A_c\}$$

עכשיו ניקח אוסף סופי $\Lambda \subset J$ (שאינו ריק) כלשהו. אז

$$\{A_\alpha|\alpha\in J\}$$

הם מאורעות בלתי תלויים. לכן גם אם נחליף אחד מהם במשלים שלו נקבל כי הם (אוסף סופי של) מאורעות בלתי תלויים. נמשיך להחליף עוד אחד ועוד אחד עד שהחלפנו את כל המאורעות עם אינדכסים ב-$J\cap\Lambda_c$ במשלימים שלהם (אם $J\cap A_c$ ריקה אז אין צורך לעשות דבר) ונקבל כי ההסתברות של המכפלה של המאורעות שהתקבלו בסוף התהליך היא מכפלת ההסתברויות.
כלומר,

$$P\left(\left(\bigcap_{\alpha\in J\cap\Lambda_c}A_\alpha^c\right)\cap\left(\bigcap_{\alpha\in J\cap(\Lambda\setminus\Lambda_c)}A_\alpha\right)\right)=\left(\prod_{\alpha\in J\cap\Lambda_c}P(A_\alpha^c)\right)\cdot\left(\prod_{\alpha\in J\cap(\Lambda\setminus\Lambda_c)}P(A_\alpha)\right)$$

המסקנה היא שגם אוסף המאורעות

$$\{A_\alpha^c|\alpha\in\Lambda_c\}\cup\{A_\alpha|\alpha\in\Lambda\setminus A_c\}$$

הם בלתי תלויים וסיימנו.
נשים לב כי עבור אוסף מאורעות בלתי תלויים כזה, אם ניקח $\alpha_1,\alpha_2,\ldots\in\Lambda$ (אוסף בן מניה של אברים שונים) אז מכיון ש-$B_n=\bigcap_{i=1}^n A_{\alpha_i}$ היא סדרה לא עולה של מאורעות נובע מרציפות ההסתברות כי

$$P\left(\bigcap_{i=1}^\infty A_i\right)=\lim_{n\to\infty}P\left(\bigcap_{i=1}^n A_i\right)=\lim_{n\to\infty}\prod_{i=1}^n P\left(A_{\alpha_i}\right)=\prod_{i=1}^\infty P\left(A_{\alpha_i}\right)$$

מכאן שתכונת המכפלה שהגדרנו עבור כל אוסף סופי $\Lambda\subset J$ מתקיימת גם עבור אוסף בן-מניה כלשהו.
נאמר כי $\{X_\alpha|\alpha\in\Lambda\}$ בלתי תלויים אם ורק אם לכל אוסף של מספרים $\{t_\alpha|\alpha\in\Lambda\}$ מתקיים כי $\{\{X_\alpha\le t_\alpha|\alpha\in\Lambda\}$ הם מאורעות בלתי תלויים. יותר מאוחר בקורס נראה כי זה שקול לכך כי לכל אוסף של קבוצות בורל $\{B_\alpha|\alpha\in\Lambda\}$ מתקיים כי $\{\{X_\alpha\in B_\alpha\}|\alpha\in\Lambda\}$ הם מאורעות בלתי תלויים.
מההגדרה נובע כי כל תת-אוסף של מאורעות בלתי תלויים הוא בעצמו אוסף של מאורעות בלתי תלויים וכי כל תת-אוסף של משתנים מקריים בלתי תלויים הוא אוסף של משתנים מקריים בלתי תלויים.
הלמה של בורל קנטלי: $A_1,A_2,\ldots\in\mathcal{F}$.

- אם $\sum_{n=1}^{\infty} P(A_n) < \infty$ אז $P(A_n,\, i.o.) = 0$.
- אם $\sum_{n=1}^{\infty} P(A_n) = \infty$ וגם $A_1, A_2, \ldots$ בלתי תלויים, אז $P(A_n,\, i.o.) = 1$.

הוכחה: אם $\sum_{n=1}^{\infty} P(A_n) < \infty$ אז

$$P(A_n,\, i.o.) = P\left(\cap_{n\geq 1} \cup_{m\geq n} A_m\right) = \lim_{n\to\infty} P\left(\cup_{m\geq n} A_m\right) \leq \boxed{\lim_{n\to\infty} \sum_{m=n}^{\infty} P(A_m) = 0}$$

כי זנב של טור מתכנס שואף לאפס.

אם $A_1, A_2, \ldots$ בלתי תלויים אז גם $A_1^c, A_2^c, \ldots$ בלתי תלויים ולכן

כאן אנחנו מסתכלים על k מאורעות ב"ת ואז נשאיף את k ל∞

$$P\left(\cap_{m=n}^{k} A_m^c\right) = \prod_{m=n}^{k} P(A_m^c) = \prod_{m=n}^{k} (1-P(A_m)) \leq \prod_{m=n}^{k} e^{-P(A_i)} = e^{-\sum_{m=n}^{k} P(A_i)}$$

$1 - x \leq e^{-x}$ לכל $x \geq 0$

אם נשאיף את k לאינסוף נקבל

$$P\left(\cap_{m=n}^{\infty} A_m^c\right) = \lim_{k\to\infty} P\left(\cap_{m=n}^{k} A_m^c\right) = \lim_{k\to\infty} e^{-\sum_{m=n}^{k} P(A_m)} = 0$$

מכיוון ש-$\sum_{m=n}^{\infty} P(A_m) = \infty$. עכשיו נשאיף את $n \to \infty$ ונקבל

$$\mathbb{P}([A_n\, i.o]^C) = \boxed{P\left(\cup_{n\geq 1} \cap_{m\geq n} A_m^c\right)} = \lim_{n\to\infty} P\left(\cap_{m=n}^{\infty} A_m^c\right) = 0$$

לבסוף

$$\begin{aligned} P(A_n,\, i.o.) &= P\left(\cap_{n\geq 1} \cup_{m\geq n} A_m\right) = P\left(\left(\cup_{n\geq 1} \cap_{m\geq n} A_m^c\right)^c\right) \\ &= 1 - \boxed{P\left(\cup_{n\geq 1} \cap_{m\geq n} A_m^c\right)} = 1 - 0 = 1 \end{aligned}$$

אפשר בשלב זה להגדיר למה אנו מתכוונים באי תלות של משתנים מקריים, אך יהיה יותר טבעי לעשות זאת יותר מאוחר.

נניח כי $X, X_1, X_2, \ldots$ הם משתנים מקריים סופיים. נאמר כי X_n שואף ל-X בהסתברות ונסמן $X_n \xrightarrow{p} X$ אם לכל $\epsilon > 0$ מתקיים כי

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0$$

נאמר כי X_n שואף ל-X בהסתברות אחת ונסמן $X_n \xrightarrow{1} X$ אם

$$P\left(\exists \lim_{n\to\infty} X_n = X\right) = P\left(\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{|X_n - X| < \frac{1}{k}\right\}\right) = 1$$

כלומר, אם

$$P\left(\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{|X_n - X| \geq \frac{1}{k}\right\}\right) = P\left(\bigcup_{k=1}^{\infty} \left\{|X_n - X| \geq \frac{1}{k},\, i.o.\right\}\right) = 0$$

ככל ש k יותר גדול האינטרוול $(X_n - X, X_n + X)$ הולך וקטן לכן $A_k \subset A_{k+1}$

מכיוון ש-$A_k = \left\{|X_n - X| \geq \frac{1}{k},\ i.o.\right\}$ מקיימות כי $A_{k+1} \subset A_k$ (מדוע?) נובע כי לכל $\ell \geq 1$ מתקיים

$$P\left(|X_n - X| \geq \frac{1}{\ell},\ i.o.\right) \leq P\left(\bigcup_{k=1}^{\infty}\left\{|X_n - X| \geq \frac{1}{k},\ i.o.\right\}\right) = \lim_{k\to\infty} P\left(|X_n - X| \geq \frac{1}{k},\ i.o.\right)$$

ולכן תנאי שקול לכך ש-$X \xrightarrow{1} X_n$ הוא שלכל $k \geq 1$ מתקיים כי

$$P\left(|X_n - X| \geq \frac{1}{k},\ i.o.\right) = 0$$

ומכיוון שלכל $\epsilon > 0$ אפשר לקחת $k \geq 1/\epsilon$ ואז $1/k \leq \epsilon$ נקבל כי

$$P\left(|X_n - X| \geq \epsilon,\ i.o.\right) \leq P\left(|X_n - X| \geq \frac{1}{k},\ i.o.\right) = 0$$

מכאן ש-$X \xrightarrow{1} X_n$ אם ורק אם לכל $\epsilon > 0$ מתקיים כי

$$P\left(|X_n - X| \geq \epsilon,\ i.o.\right) = 0$$

נשים לב כי

$$\lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty}\{|X_n - X| \geq \epsilon\}\right) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\{|X_n - X| \geq \epsilon\}\right) = P\left(|X_n - X| \geq \epsilon,\ i.o.\right)$$

ומכיוון ש-

$$P\left(|X_n - X| \geq \epsilon\right) \leq P\left(\bigcup_{m=n}^{\infty}\{|X_n - X| \geq \epsilon\}\right)$$

אז אם הגבול של צד ימין הוא אפס אז גם הגבול של צד שמאל חייב להיות אפס. מכאן נובע כי אם $X \xrightarrow{1} X_n$ אז גם $X \xrightarrow{p} X_n$.

עכשיו, נניח כי $X \xrightarrow{p} X_n$. אז מכיוון ש-

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0$$

בפרט נובע כי לכל $k \geq 1$ קיים n_k עבורו

$$P\left(|X_{n_k} - X| \geq \frac{1}{k}\right) \leq \frac{1}{2^k}$$

ולכן

$$\sum_{k=1}^{\infty} P\left(|X_{n_k} - X| \geq \frac{1}{k}\right) \leq \sum_{k=1}^{\infty}\frac{1}{2^k} = 1 < \infty$$

ומהלמה של בורל-קנטלי נובע כי

$$P\left(|X_{n_k}-X|\geq\frac{1}{k},\ i.o.\right)=0$$

עכשיו, לכל $k\geq 1/\epsilon$ מתקיים כי

$$\{|X_{n_k}-X|\geq\epsilon\}\subset\left\{|X_{n_k}-X|\geq\frac{1}{k}\right\}$$

ומכאן ש-ולכן לכל $\epsilon>0$ מתקיים כי

$$P\left(|X_{n_k}-X|\geq\epsilon,\ i.o.\right)\leq P\left(|X_{n_k}-X|\geq\frac{1}{k},\ i.o.\right)=0$$

ולכן

$$X_{n_k}\xrightarrow{1}X$$

עכשיו, כאשר $X_n\xrightarrow{p}X$ אז אם ניקח תת סדרה $\{n_k\}$ אז גם $X_{n_k}\xrightarrow{p}X$ (מדוע?) ולכן קיימת לתת סדרה זו תת-תת סדרה $\{n_{k_\ell}\}$ עבורה מתקיים כי

$$X_{n_{k_\ell}}\xrightarrow{1}X$$

בכיוון ההפוך, אם לכל תת סדרה $\{n_k\}$ קיימת תת -תת סדרה $\{n_{k_\ell}\}$ עבורה

$$X_{n_{k_\ell}}\xrightarrow{1}X$$

אז גם

$$X_{n_{k_\ell}}\xrightarrow{p}X$$

(כי שאיפה בהסתברות אחת גוררת שאיפה בהסתברות). מכאן שעבור $\epsilon>0$ כלשהו מתקיים כי

$$a_n=P\left(|X_n-X|>\epsilon\right)$$

היא סדרה המקיימת שלכל תת סדרה $\{n_k\}$ קיימת תת-תת סדרה $\{n_{k_\ell}\}$ עבורה

$$\lim_{\ell\to\infty}a_{n_{k_\ell}}=0$$

ולכן גם

אחרת היה ל $a_{n_{k_l}}$ תת סדרה שלא מתכנסת ל0 וזה סתירה לכך שעבור סדרה מתכנסת כל תת הסדרות שלה מתכנסות לאותו הגבול.

$$\lim_{n\to\infty}a_n=0$$

זה נכון כי אם לא היה קיים גבול, או שהגבול אינו 0, אז קיים $\delta>0$ ותת סדרה $\{n_k\}$ המקיימת לכל כי $a_{n_k}\geq\delta$ ולכן לא יתכן שקיימת תת-תת סדרה ששואפת לאפס. נסכם בתוצאה הבאה:

משפט:

- אם $X_n\xrightarrow{1}X$ אז גם $X_n\xrightarrow{p}X$.

- $X \xrightarrow{p} X_n$ אם ורק אם לכל תת סדרה $\{n_k\}$ קיימת תת-תת סדרה $\{n_{k_\ell}\}$ עבורה $X_{n_{k_\ell}} \xrightarrow{1} X$ (בפרט אם תת הסדרה הראשונה היא הסדרה כולה).

הערה: כפי שתראו בתרגילי בית, יש דוגמאות לכך ש-$X \xrightarrow{p} X_n$ אך לא מתקיים כי $X_n \xrightarrow{1} X$.

תכונות:

- אם $X_n \xrightarrow{1} X$ אז גם $g(X_n) \xrightarrow{1} g(X)$ לכל פונקציה רציפה g.

- אם $X_n \xrightarrow{1} X_n$ וגם $Y_n \xrightarrow{1} Y$ אז גם

$$\|(X_n, Y_n) - (X, Y)\| = \sqrt{(X_n - X)^2 + (Y_n - Y)^2} \xrightarrow{1} 0$$

ולכן $g(X_n, Y_n) \xrightarrow{1} g(X, Y)$ לכל פונקציה רציפה. באותו אופן זה מתקיים עבור מספר סופי כלשהו של משתנים מקריים.

- התכונות הנ״ל נכונות אם נחליף התכנסות בהסתברות אחת בהתכנסות בהסתברות. בפרט, התכונה הראשונה נובעת מכך שלכל תת סדרה יש תת-תת סדרה שמתכנסת בהסתברות אחת. התכונה השניה נובעת מכך שלכל תת סדרה ניתן לקחת תת-תת סדרה כך ש-$X_{n_{k_\ell}} \xrightarrow{1} X$ ואז לקחת מתוך תת-תת הסדרה הזו תת סדרה נוספת עבורה גם Y_n מתכנס בהסתברות אחת (לאורך התת סדרה החדשה). אז נקבל הזוג (X_n, Y_n) מתכנס בהסתברות אחת ל-(X, Y) לאורך תת סדרה זו.

תחילה נבנה תת סדרה משותפת:

$$X_n \xrightarrow{p} X \Longrightarrow X_n \trianglerighteq X_{n_k} \Longrightarrow X \xrightarrow{1} X_{n_k}$$

$$Y_n \xrightarrow{p} Y \Longrightarrow Y_n \trianglerighteq Y_{n_k} \Longrightarrow Y \xrightarrow{p} Y_{n_k} \Longrightarrow Y_{n_k} \trianglerighteq Y_{n_{k_l}} \Longrightarrow Y \xrightarrow{1} Y_{n_{k_l}}$$

כעת נשים לב כי $X \xrightarrow{1} X_{m_{k_l}}$ לכן לסיכום $\begin{cases} X_{n_{k_l}} \xrightarrow{1} X \\ Y_{n_{k_l}} \xrightarrow{1} Y \end{cases}$ הגענו לכך שלכל תת סדרה n_k קיימת תת תת תת סדרה כך ש $\begin{cases} g(X_{n_{k_l}}) \xrightarrow{1} g(X) \\ g(Y_{n_{k_l}}) \xrightarrow{1} g(Y) \end{cases}$

ולכן $\begin{cases} g(X_n) \xrightarrow{p} g(X) \\ g(Y_n) \xrightarrow{p} g(Y) \end{cases}$