הרצאה 4

1 שיטות המומנט הראשון והשני

יהי X מ"מ אי-שלילי שמחזיר ערכים שלמים (נקרא גם מ"מ **סופר**). נניח שהתוחלת שואפת ל0 כאשר n שואף לאינסוף (עבור פרמטר n כלשהו). אז מי-שוויון מרקוב, נקבל:

$$\mathbb{P}(X \neq 0) = \mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}(X)}{1} = \mathbb{E}(X) \Rightarrow \lim_{n\to\infty} \mathbb{P}(X = 0) = 1$$

בעצם קיבלנו שX הוא "כמעט תמיד" 0.

השימוש הזה באי-שוויון מרקוב נקרא **שיטת המומנט הראשון**. (שימוש בתוחלת כדי להסיק דברים על ההסתברות).

עכשיו, נניח ש $\mathbb{E}(X)$ שואף לאחד או אפילו לאינסוף כאשר n שואף לאינסוף. האם זה אומר ש $\lim_{n\to\infty} \mathbb{P}(X = 0) = 0$? מסתבר שלא בהכרח, לדוגמה:

לכל טבעי n יהי X_n מ"מ כך ש $\mathbb{P}(X_n = n^2) = 1/n$, $\mathbb{P}(X = 0) = 1 - 1/n$. אזי התוחלת היא אינסוף:

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \lim_{n\to\infty} n = \infty$$

אבל בבירור ניתן לראות ש:

$$\lim_{n\to\infty} \mathbb{P}(X = 0) = 1$$

שזה הגיוני, כי ככל ש-n עולה, ההסתברות לקבל את n^2 קטנה.

איזה עוד נתונים נצטרך כדי שכן נוכל להסיק ש $\lim_{n\to\infty} \mathbb{P}(X = 0) = 0$? אנחנו רוצים להסיק משהו על X לפי התוחלת שלו. אבל מסתבר שX לא חייב להיות תמיד קרוב לתוחלת. בדיוק בשביל זה הגדרנו שונות – זה אומר לנו עד כמה X מתרחק מהתוחלת. אם נפעיל את אי-שוויון צ'בישב, נקבל:

$$\mathbb{P}(X = 0) \leq \underbrace{\mathbb{P}\Big(|X - \mathbb{E}(X)| \geq \mathbb{E}(X)\Big)}_{X\geq 2\mathbb{E}(X) \text{ or } x\leq 0} \leq \frac{\mathrm{Var}(X)}{\big(\mathbb{E}(X)\big)^2}$$

אז אם הביטוי הימני שואף לאפס כאשר n שואף לאינסוף, זה יגיד לנו שההסתברות שX הוא אפס גם שואפת לאפס. אפשר לסמן את זה כך:

$$\mathrm{Var}(X) = o\left(\big(\mathbb{E}(X)\big)^2\right)$$

(כלומר, השונות "קטנה בהרבה" מהתוחלת בריבוע).

השימוש הזה באי-שוויון צ'בישב נקרא **שיטת המומנט השני**. (שימוש בתוחלת בריבוע).

באופן כללי, כדי לבדוק אם אובייקט קיים:
נייצר מ"מ X שסופר מופעים של האובייקט הזה,
ונרצה להחליט האם X שווה 0 (כלומר אין אובייקט) או גדול מ0 (כלומר יש).
כדי להראות שהוא שווה 0 נשתמש במומנט הראשון – נראה שהתוחלת שווה 0.
בגלל שX לא יכול להיות שלילי, זב אומר שX תמיד 0.
כדי להראות שזה גדול מאפס, נשתמש במומנט השני – נראה שההסתברות שX שווה 0 קטנה ממשהו ששואף לאפס.

1.1 הופעת משולש:

טענה 1.1: נבנה גרף G על $[n]$ בצורה הבאה: לכל $1 \le i < j \le n$ נטיל מטבע **לא הוגן**, כאשר כל ההטלות בת"ל. אם יצא עץ (הסתברות p), נוסיף את הצלע ij ל-G. נגדיר את המאורע $A_G =$ "יש משולש בG". צ"ל:

א) אם $p = o(1/n)$, אז $\lim_{n\to\infty} \mathbb{P}(A_G) = 0$

ב) אם $p = \omega(1/n)$, אז $\lim_{n\to\infty} \mathbb{P}(A_G) = 1$

כלומר, אם p הרבה יותר קטן מ $1/n$, אז ההסתברות למשולש שואפת לאפס כאשר n שואף לאינסוף.
ואם p הרבה יותר גדול מ- $1/n$, אז ההסתברות למשולש שואפת ל-1 כאשר n שואף לאינסוף.

תזכורת: $f = o(g) \Leftrightarrow g = \omega(f)$

הוכחת א: נגדיר $t = \binom{n}{3}$ ויהיו $A_1, \dots, A_t$ כל תתי הקבוצות של $[n]$ בגודל 3.
לכל $1 \le i \le t$, נגדיר את X_i האינדיקטור למאורע "הצלעות של A_i יוצרות משולש בG".
יהי $X = \sum_{i=1}^{t} X_i$, אז X סופר את מספר המשולשים בG. מתקיים:

$$\mathbb{E}(X_i) = \mathbb{P}(X_i = 1) = p^3$$

לכל $1 \le i \le t$. לפי לינאריות התוחלת נקבל:

$$\mathbb{E}(X) = \sum_{i=1}^{t} \mathbb{E}(X_i) = \binom{n}{3} p^3$$

בפרט: אם $p = o(1/n)$ (כלומר $\frac{p}{1/n} \to 0$),
אז $\lim_{n\to\infty} \mathbb{E}(X) = 0$ (כי $\binom{n}{3} \approx n^3$, אז $(np)^3 = o(1)$).

אז לפי המומנט הראשון, ההסתברות שיש בG משולש שואפת לאפס כאשר n שואף לאינסוף. (התוחלת שואפת לאפס, ודרך אי-שוויון מרקוב נקבל שההסתברות שX גדול מ0 שואפת לאפס). מש"ל א.

הוכחת ב: נניח $p = \omega(1/n)$, נובע מכך ש $\lim_{n\to\infty} \mathbb{E}(X) = \infty$.
נשתמש בשיטת המומנט השני: נשתמש באי-שוויון צ'בישב כדי לחסום את ההסתברות ש $\mathbb{P}(X = 0)$.
בשביל זה נצטרך לחסום את השונות של X. נשתמש בנוסחה לשונות של סכום: (2)

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

ראשית, נשים לב שמתקיים: (3) לכל $1 \le i \le t$

$$\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) - \left(\mathbb{E}(X_i)\right)^2 \le \mathbb{E}(X_i) = p^3$$

נבחר $1 \le i < j \le t$ כלשהם ונגדיר $\ell = |A_i \cap A_j|$.
נשים לב שבגלל ש $i \ne j$, מתקיים $\ell \le 2$.
אם $\ell \le 1$, זה אומר ש X_i, X_j נקבעו ע"י קבוצות זרות של הטלות
(כי יש להן לכל היותר קודקוד אחד משותף. אז אין להן צלעות משותפות) ולכן הם בת"ל.
בפרט, $\mathrm{Cov}(X_i, X_j) = 0$.
נניח עכשיו ש $\ell = 2$. אז מתקיים: (4)

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i, X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \mathbb{P}(X_i = 1, X_j = 1) - p^6 = p^5 - p^6 \le p^5$$

נשלב את (2), (3), (4):

$$\mathrm{Var}(X) \leq^{א} n^3p^3 + 2n^4p^5 =^{ב} o(n^6p^6) =^{ג} o\left(\left(\mathbb{E}(X)\right)^2\right)$$

א – חסם עליון לt, כפול השונות של X_i.
וחסם עליון למספר הדרכים לבחור 2 משולשים: אם יש צלע אחת או 2 משותפות צריך לבחור רק 4 קודקודים,
אז זה חסם עליון לזה. ואם אין צלעות משותפות השונות המשותפת היא 0. והשונות המשותפת היא לכל היותר p^5.

ב – בגלל ההנחה ש $p = \omega(1/n)$.

ג – בגלל החסם שמצאנו על התוחלת.

בסך הכל, נקבל שההסתברות שאין משולש שואפת לאפס. כלומר ההסתברות שיש משולש שואפת ל1, כנדרש.

1.2 סכומים זרים:

נאמר שלקבוצה $\{x_1, \dots, x_k\}$ יש **סכומים זרים** אם לכל אחת מ 2^k תתי הקבוצות יש סכום שונה.
עבור מספר טבעי n, יהי $f(n)$ השלם הגדול ביותר k שעבורו קיימים שלמים $1 \leq x_1, \dots, x_k \leq n$ כך שלקבוצה $\{x_1, \dots, x_k\}$ יש
סכומים זרים. (כלומר, עבור כל n, מה הקבוצה הכי גדולה שיש לה סכומים זרים).
לדוגמה $\{2^i : 0 \leq i \leq \lfloor \log_2 n \rfloor\}$. זה מראה שמתקיים $f(n) \geq 1 + \lfloor \log_2 n \rfloor$.
אנחנו נראה שהחסם התחתון הזה הוא הדוק מבחינה אסימפטוטית.

טענה 1.2: $f(n) \leq \log_2 n + \frac{1}{2}\log_2 \log_2 n + O(1)$.

הוכחה: אנחנו רוצים למצוא חסם עליון ל $f(n)$. נעשה את זה כך:
נמציא מ"מ שמייצג סכום תת קבוצה,
נמצא את התוחלת והשונות,
נשתמש בצ'בישב כדי לקבל חסם על ההסתברות של ההפרש בין X לתוחלת,
ונשתמש בתכונת הסכומים הזרים כדי לקבל חסם שתלוי ב $f(n)$:

יהי $k = f(n)$ ויהיו $1 \leq x_1 \dots x_k \leq n$ שלמים כך ש $\{x_1 \cdots x_k\}$ בעלת סכומים זרים.
יהיו $I_1 \dots I_k$ מ"מ בת"ל כך ש $\mathbb{P}(I_i = 0) = \mathbb{P}(I_i = 1) = 1/2$ לכל $1 \leq i \leq k$.
יהי $X = \sum_{i=1}^{k} x_i I_i$. לפי לינאריות התוחלת, נקבל:

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{k} x_i I_i\right) = \sum_{i=1}^{k} x_i \cdot \mathbb{E}(I_i) = \frac{x_1 + \cdots + x_k}{2}$$

האינדיקטורים בעצם נותנים לנו סכום אקראי של תת קבוצה של $\{x_1 \cdots x_k\}$. ורואים שהתוחלת היא בעצם הממוצע.

בנוסף, בגלל שכל האינדיקטורים בת"ל, מתקיים: (5)

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{i=1}^{k} x_i I_i\right) =^{א} \sum_{i=1}^{k} x_i^2 \cdot \mathrm{Var}(I_i) =^{ב} \frac{x_1^2 + \cdots + x_k^2}{4} \leq^{ג} \frac{n^2 k}{4}$$

א – סקלר יוצא מהשונות בחזקת 2. וכל הcovariance הם 0 כי המ"מ בת"ל.
ב - $I_i \sim \mathrm{Ber}(1/2)$, אז השונות היא 1/4.
ג – כל $X_i \leq n$.

עכשיו נוכל להשתמש בצ'בישב: נציב $t = 2\sqrt{\mathrm{Var}(X)}$ (כדי שיצטמצם עם ה Var בצ'בישב) ונקבל:

$$\mathbb{P}(|X - \mathbb{E}(X)| > t) \leq \frac{\mathrm{Var}(X)}{t^2} = \frac{1}{4}$$

ולכן: (6) $\mathbb{P}(|X-\mathbb{E}(X)|\le t)\ge 1-\frac{1}{4}=\frac{3}{4}$.

מצד שני, בגלל שלפי ההנחה כל הסכומים של תתי הקבוצות זרים,
ההסתברות ש $X-\mathbb{E}(X)=s$ עבור s ממשי כלשהו היא או 0 או 2^{-k}
(כי יש בדיוק 2^k סכומים שונים, אז יש בדיוק את המספר הזה של הפרשים אפשריים.
אז אם s הוא לא אחד ההפרשים, אין סיכוי שנגיע אליו, לא משנה מה X יוצא).
בפרט, מתקיים $\mathbb{P}(X-\mathbb{E}(X)=s)\le 2^{-k}$ לכל $s\in[-t,t]$.
יש לכל היותר $2t+1$ ערכים בטווח הזה שעבורם $\mathbb{P}(x-\mathbb{E}(X)=s)>0$,
(ערכים שיכולים להיות ההפרש. אם הסכום של כל האיקסים היה זוגי, אז התוחלת שלמה וזה מספר הערכים השלמים בטווח. ואם הסכום היה אי-זוגי, התוחלת היא משהו ועוד חצי אבל עדיין שיש את אותו מספר של הפרשים אפשריים). אז חסם איחוד ייתן לנו: (7)

$$\mathbb{P}(|X-\mathbb{E}(X)|\le t)\le^{\text{א}} 2^{-k}(2t+1)\le^{\text{ב}} 2^{-k}(2n\sqrt{k}+1)$$

א – חסם על ההסתברות כפול מספר הערכים האפשריים.

ב – חסם עליון על t. ניזכר שt מוגדר לפי השונות, ויש את החסם העליון על השונות (5).

מצאנו חסם עליון ותחתון להסתברות. אם נשווה את (6), (7) נקבל אי שוויון של k: (8)

$$\frac{3}{4}\le^{\text{א}} 2^{-k}\left(2n\sqrt{k}+1\right)\Rightarrow^{\text{ב}}\frac{3}{4}\cdot 2^k\le 2n\sqrt{k}+1\Rightarrow^{\text{ג}}\frac{2^k}{\sqrt{k}}\le Cn$$

א – שילוב של 6, 7.

ב – נכפול את שני האגפים ב 2^k

ג – נחלק ב $\sqrt{k}$, ו"נבלע" את ה 1+ ע"י הקבוע C.

עבור $C>0$ כלשהו. עכשיו אנחנו רוצים לטעון שk קטן יותר ממשהו. איך נעשה את זה?
אנחנו רוצים לפתוח את הצד השמאלי של האי-שוויון האחרון, אבל ה2 בחזקת k והשורש מקשים.
אנחנו יודעים שזה כנראה ייצא משהו עם לוג.
הטריק הוא להציב במקום k משהו דומה אבל יותר גדול ממה שרוצים להוכיח, ונראה שאי השוויון לא מתקיים.
וזה מוכיח שk צריך להיות קטן יותר ממה שהצבנו.
אז נציב את אותה הנוסחה, אבל במקום ה $O(1)$ נשים משהו גדול.
נציב את $k=L>O(1)$. במונה, נפרק את החזקה כדי שיהיה יותר נוח:

$$\frac{2^k}{\sqrt{k}}=^{\text{א}}\frac{\overbrace{2^{\log_2 n}}^{=n}\cdot 2^{\overbrace{\frac{1}{2}\log_2\log_2 n}^{=\log_2(\log_2 n)^{\frac{1}{2}}}}\cdot 2^L}{\sqrt{\log_2 n+\frac{1}{2}\log_2\log_2 n+L}}\ge^{\text{ב}}\frac{n\cdot\sqrt{\log_2 n}\cdot 2^L}{\sqrt{2\log_2 n}}=^{\text{ג}} L'\cdot n>Cn$$

א – נציב את מה שרוצים להוכיח.

ב – המונה שווה. המכנה גדול יותר (כי חצי לוג של לוג זה הרבה יותר קטן, עבור n גדול).

ג – נבחר 'L מתאים – חישוב של ה2^L ביחד עם השורש 2 שנשאר במכנה. עבור L גדול, ברור שזה יותר גדול מC, כי C קבוע קטן כלשהו. גם אם L לא הרבה יותר גדול מC, 2 בחזקת L זה עדיין יותר גדול.

כלומר כדי שאי השוויון 8 יתקיים, צריך שk יהיה קטן ממה שהצבנו.

מכאן נובע ש $k=f(n)\le\log_2 n+\frac{1}{2}\log_2\log_2 n+O(1)$, כנדרש.