הסתברות 1 - תרגול 8

בדוגמ' בתרגול לא מתקיים $X_1, \dots, X_n \sim X$

אבל כן במעט מתקיים שהרי $X_1, \dots, X_n \sim ber(p_n)$

כאשר $p_n \xrightarrow{n\to\infty} \frac{1}{2}$

החוק החלש של המספרים הגדולים

נזכר במשפט שהוכחתם בכיתה:

משפט 0.1 יהי X מ"מ בעל תוחלת סופית ויהי $\epsilon > 0$. לכל $N \in \mathbb{N}$ יהי

$$p_N = \mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - \mathbb{E}(X)\right| < \epsilon\right)$$

כאשר $\{X_n\}$ מ"מ ב"ת ושווי התפלגות ל X. אזי

$$\lim_{N\to\infty} p_n = 1$$

נסתכל על דוגמא שלא מתאימה לגמרי לתנאי המשפט:

דוגמא

תהי $\{p_n\}$ סדרת מספרים חיוביים כך ש $\lim_n p_n = 1/2$. יהיו $X_n \sim Ber(p_n)$ ב"ת.

טענה 0.2 לכל $\epsilon > 0$ מתקיים

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - \frac{1}{2}\right| < \epsilon\right) = 1$$

הוכחה: מאחר ו X_n ב"ת,

$$Var(\frac{1}{N}\sum_{n=1}^{N} X_n) = \frac{1}{N^2}\sum_{n=1}^{N} Var(X_n) = \frac{1}{N^2}\sum_{n=1}^{N}(p_n - p_n^2)$$

כי בערך $\frac{N \cdot 1/4}{N^2} \xrightarrow{n\to\infty} 0$

מאחר ו $\lim_n p_n = 1/2$ מתקיים

$$\lim_n E(\frac{1}{n}\sum_{i=1}^{n} X_n) = \lim_n \frac{\sum_{i=1}^{n} p_n}{n} = \frac{1}{2}$$

אפשר את החלוקה כדי לשלב את החוק הגדול

נשתמש בזה בנוסף בשיטה בחוק הגדול של מנות כן לפיצ'ח את הממוצע הנ"ל

• הממוצע החשבוני של a_n מתכנס לאותו גבול כמו $\{a_n\}_{n=1}^{\infty}$ (ניתן גם לראות ממוצע זה כמקרה פרטי של ממוצע משוקלל).

אנצל את התועלת כדי
לסגור את הטען הקודם

$$\lim_n E(\frac{1}{n}\sum_{i=1}^{n} X_n) = \lim_n \frac{\sum_{i=1}^{n} p_n}{n} = \frac{1}{2}$$

ומכאן שקיים N_0 כך שלכל $n > N_0$

$$|E(\frac{1}{n}\sum_{i=1}^{n} X_n) - \frac{1}{2}| := r_n < \frac{\epsilon}{2}$$

בנוסף, מאי שוויון המשולש ההפוך,

$$\left|\frac{1}{N}\sum_{n=1}^{N} x_n - \frac{1}{2}\right| = \left|\frac{1}{N}\sum_{n=1}^{N} x_n - \left|E\left(\frac{1}{N}\sum_{n=1}^{N} x_n\right) - r_n\right|\right| \le \left|\frac{1}{N}\sum x_n - E\left(\frac{1}{N}\sum_{n=1}^{N} x_n\right)\right| + |r_n|$$

א״ש משולש

ולכן

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - \frac{1}{2}\right| \ge \epsilon\right) \le \mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - E(\frac{1}{n}\sum_{i=1}^{n} X_n)\right| + \frac{\epsilon}{2} \ge \epsilon\right)$$

$$= \mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - E(\frac{1}{n}\sum_{i=1}^{n} X_n)\right| \ge \frac{\epsilon}{2}\right) \le \frac{4}{\epsilon^2} Var(\frac{1}{N}\sum_{n=1}^{N} X_n)$$

כאשר האי שוויון האחרון הוא אש צ׳בישב. לבסוף,

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - \frac{1}{2}\right| < \epsilon\right) = 1 - \mathbb{P}\left(\left|\frac{1}{N}\sum_{n=1}^{N} X_n - \frac{1}{2}\right| \ge \epsilon\right)$$

$$\ge 1 - \frac{4}{\epsilon^2} Var(\frac{1}{N}\sum_{n=1}^{N} X_n) = 1 - \frac{4}{\epsilon^2}\frac{1}{N^2}\sum_{n=1}^{N}(p_n - p_n^2)$$

מאחר ו $p_n \to \frac{1}{2}$ נקבל שהחל מ n מסוים $p_n - p_n^2 < \frac{1}{2}$ ולכן $\frac{1}{N^2}\sum_{n=1}^{N}(p_n - p_n^2) \to 0$ כאשר $N \to \infty$ ומכאן הטענה. $\blacksquare$

גרף ארדוש רני

יהי $n \in \mathbb{N}$ ו $p \in (0,1)$, גרף ארדוש רני (או גרף $G(n,p)$) הוא גרף (לא מכוון) בעל n קודקודים שכל צלע מופיעה בהסתברות p ובאופן ב״ת בצלעות אחרות.נשלול אפשרות של צלע מקודקוד לעצמו. למרות שאין שום קשר בין הצלעות המופיעות בגרף ניתן לזהות תופעות גלובליות מעניינות (למשל קשירות, קיום מעגל המילטון ועוד). ע״פ רוב $p = p(n)$ והשאלה היא ע״פ איזה חוק חזקה (או כלל אחר) צריך להתנהג p כדי לקבל תופעות מסוימות בגרף. בתרגיל הבית התבקשתם להראות קיום משולשים. בתרגול זה נעסוק בצורה אחרת.

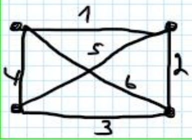
ראשית כדי לקבל קצת אינטואיציה נשים לב לעובדות הבאות:

- עבור X ־מספר הצלעות בגרף מתקיים $\mathbb{E}(X) = p\binom{n}{2}$
- ההסתברות שקודקוד נתון v יהיה מדרגה k מתפלג בינומית

$$P(deg(v) = k) = \binom{n-1}{k} p^k (1-p)^{n-k-1}$$

דרגה של קודקוד = כמות צלעות
שיוצאות ממנו

קיים ריבוע שלם בגרף קרי



נראה תנודה מעניינת

עבור $p(n) = \frac{1}{n^t}$

$t < -\frac{2}{3}$ סיכוי אפסי שיהיה קליק

$t > \frac{2}{3}$ ודאות כמעט מובטחת שיהיה קליק

הגדרה 0.3 4־ קליק בגרף היא קבוצה של 4 קודקודים המקיימת שכל $\binom{4}{2} = 6$ הצלעות שבים הקודקודים קיימות. נסמן ב C_4 את המאורע בו קיים 4־ קליק בגרף.

$f(n) = o(g(n))$ $\lim_{n\to\infty} \frac{|f(n)|}{g(n)} = 0$

טענה 0.4 נניח ש $p = p(n) = o(n^{-2/3})$. אזי

$$\lim_{n\to\infty} \mathbb{P}(C_4) = 0$$

הוכחה: יהי X המ"מ של מספר ה 4־קליקים בגרף. אזי

$$X = \sum_{I\subset[n],|I|=4} X_I$$

כאשר $X_I = 1$ אם קודקודי I יוצרים 4־קליק, ו $X_I = 0$ אחרת. נשים לב שיש $\binom{n}{4}$ מחוברים בסכום.

כמו כן, $P(X_I = 1) = p^6$ (כל הצלעות קיימות) ולכן מליניאריות התוחלת

$$\mathbb{E}(X) = \binom{n}{4} p^6 = n^4 p^6 + \text{lower order terms}$$

ולכן, אם $p = o(n^{-2/3})$ נקבל $\mathbb{E}(X) \to 0$. ולכן מא"ש מרקוב,

$$\mathbb{P}(C_4) = \mathbb{P}(X \geq 1) \leq \mathbb{E}(X) \to 0$$

■

כעת נרצה להראות את ההפך. אם $\lim \frac{p(n)}{n^{-2/3}} = \infty$ אז $\mathbb{P}(C_4) \to 1$. נשים לב שזו טענה קצת יותר עדינה. ע"פ החישוב שמקודם, במקרה זה $\mathbb{E}(X) \to \infty$ זה לא מספיק כדי להסיק ש $\mathbb{P}(C_4) \to 1$ מא"ש מרקוב. יתכן באופן עקרוני שהתוחלת היא גדולה מאוד אך לכל 4־קליק הסתברות מאוד קטנה. כדי להראות זאת צריך מידע נוסף על השונות. אם נראה שהשונות מספיק קטנה נוכל להסיק ש X מספיק קרוב לתוחלת שלו ולכן חיובי. טיעון כזה נקרא second moment method.

למה 0.5 יהי X מ"מ, אזי

$$P(X = 0) \leq \frac{Var(X)}{\mathbb{E}(X)^2}$$

בפרט, אם נראה ש $Var(X(n)) = o(\mathbb{E}(X(n))^2)$ כאשר $X(n)$ מספר ה 4־קליקים ב $G(n, p(n))$ נסיק ש $P(X(n) \geq 1) \to 1$. **הוכחה:** (של הלמה) מא"ש צ'בישב,

$$\mathbb{P}(X = 0) \leq \mathbb{P}\left(|X - \mathbb{E}(X)| \geq \mathbb{E}(X)\right) \leq \frac{Var(X)}{\mathbb{E}(X)^2}$$

כאשר הא"ש הראשון נובע מכך ש $\{X = 0\} \subset \{|X - \mathbb{E}(X)| \geq \mathbb{E}(X)\}$ ■

ניגש להוכחת הטענה.

טענה 0.6 אם $\lim \frac{p(n)}{n^{-2/3}} = \infty$ אזי $\mathbb{P}(C_4) \to 1$ כאשר $n \to \infty$.

הוכחה:

כמקודם, נכתוב

$$X = \sum_{I\subset[n],|I|=4} X_I$$

כאשר X_I המציינת של ה 4 ־קליק של הקודקודים שבקבוצה I.
מטענה קודמת, מספיק להוכיח ש $Var(X) = o(E(X)^2)$. נעזר בנוסחה השימושית

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

ראשית, מאחר ו $X_I \sim Ber(p^6)$

$$Var(X_I) = \mathbb{E}(X_I^2) - \mathbb{E}(X_I)^2 = \mathbb{E}(X_I) - \mathbb{E}(X_I)^2 = p^6 - p^{12} = O(p^6)$$

חשב כעת את $Cov(X_I, X_J)$. הערך יהיה תלוי כמובן במידת האי־תלות שתלויה מצדה במספר הקודקודים המשותפים $|I\cap J|$ נחלק למקרים:

- אם $|I\cap J| \leq 1$ אז אין צלעות משותפות ולכן X_I, X_j ב"ת ובפרט $Cov(X_I, X_J) = 0$.
- . נניח ש $|I\cap J| = 2$ אזי $X_I \cdot X_j = 1$ אם קיים 4־קליק ב I וב J מאחר ויש להם צלע אחת משותפת נצטרך רק 11 צלעות ונקבל ש $X_IX_J \sim Ber(p^{11})$ ולכן

$$Cov(X_I, X_j) = \mathbb{E}(X_IX_j) - \mathbb{E}(X_I)\mathbb{E}(X_J) = p^{11} - p^{22} - p^{12} = O(p^{11})$$

  יש $\binom{n}{6}\binom{6}{2} = O(n^6)$ אפשרויות לזוג J, I כנ"ל.
- לבסוף, אם $|I\cap J| = 3$ אז יש להם 3 צלעות משותפות ולכן חישוב דומה יתן $Cov(X_I, X_J) = O(p^9)$
  יש $\binom{n}{5}\binom{5}{3} = O(n^5)$ אפשרויות לזוג J, I כנ"ל.

נקבל אם כן,

$$Var(X) = O(n^4)O(p^6) + O(n^6)O(p^{11}) + O(n^5)O(p^9)$$

אולם, $O(n^5)O(p^9) \to 0$ וכו
כעת, נזכר ש $\mathbb{E}(X) = \binom{n}{4}p^6$ ולכן $\mathbb{E}(X)^2 = O(n^8p^{12})$.
לכן, אם $\lim \frac{p(n)}{n^{-2/3}} = \infty$ אזי

$$\frac{Var(X)}{\mathbb{E}(X)^2} \to 0$$

כנדרש. (להשלים פרטים)