הרצאה 6

1 פונקציות סף בגרפים מקריים

בהרצאה 4 הוכחנו ש:

$$\lim_{n\to\infty} \mathbb{P}(G(n,p) \text{ contains a triangle}) = \begin{cases} 0, & \text{if } p = o(1/n) \\ 1, & \text{if } p = \omega(1/n) \end{cases}$$

כלומר, כל עוד p יותר קטן מ- 1/n (בסדר גודל), ההסתברות למעגל שואפת לאפס. ואם p גדול יותר מ- 1/n, ההסתברות שואפת ל1. אז אם נצייר גרף של ההסתברות למשולש כפונקציה של p, אנחנו יודעים את הצורה הכללית של הגרף: משמאל ל-1/n הוא קרוב לאפס, מימין ל-1/n הוא קרוב ל-1. סוג של מדרגה. בסביבה הקרובה של 1/n אנחנו לא יודעים בדיוק מה מוגדר.

נרחיב את התופעה לכל תכונה מונוטונית עולה:

הגדרה 1.1: פונקציה $p_0(n)$ תיקרא **פונקציית סף** עבור תכונה מונוטונית עולה Q אם מתקיים:

$$\lim_{n\to\infty} \mathbb{P}(G(n,p) \in Q) = \begin{cases} 0, & \text{if } p = o(p_0) \\ 1, & \text{if } p = \omega(p_0) \end{cases}$$

כלומר היא הנקודה שבה הקפיצה מתרחשת.

משפט 1.2: לכל תכונה לא-טריויאלית ומונוטונית עולה יש סף. לא נוכיח את המשפט בקורס הזה.
(תכונה לא טריוויאלית היא תכונה שקיימת בכל גרף או לא קיימת באף גרף)
המשפט הזה נותן לנו את "חוק אפס – אחד": עבור n מספיק גדול, ההסתברות לתכונה היא בפועל 0 או 1 (למעט סביבה כלשהי של הסף). לפעמים נוכל גם להקטין את החלון הזה.

הגדרה 1.3: סף של תכונה מונוטונית עולה Q ייקרא **חד** אם לכל $\varepsilon > 0$ מתקיים:

$$\lim_{n\to\infty} \mathbb{P}(G(n,p) \in Q) = \begin{cases} 0, & \text{if } p \le (1-\varepsilon)p_0 \\ 1, & \text{if } p \ge (1+\varepsilon)p_0 \end{cases}$$

אחרת, הסף ייקרא **גס**.

1.1 דרגה מינימלית וקשירות של גרפים מקריים

משפט 1.4: $\frac{\ln n}{n}$ הוא סף חד עבור דרגה מינימלית חיובית וקשירות של $G(n,p)$.

הוכחה: יהי $\varepsilon > 0$, יהי n מספיק גדול, ויהי $G \sim G(n,p)$. מכיוון שלכל גרף קשיר יש דרגה מינימלית חיובית, מספיק להוכיח:

1) אם $p \le \frac{(1-\varepsilon)\ln n}{n}$, אז $\lim_{n\to\infty} \mathbb{P}(\delta(G) \ge 1) = 0$
2) אם $p \ge \frac{(1+\varepsilon)\ln n}{n}$ אז $\lim_{n\to\infty} \mathbb{P}(G \text{ is connected}) = 1$

כי כל אחד יגרור גם את השני בכיוון שצריך.

הוכחת 1: צריך להוכיח שקיים משהו – זה מתאים לשיטת המומנט השני. נניח ש $p \le \frac{(1-\varepsilon)\ln n}{n}$. לכל $1 \le j \le n$ יהי I_j האינדיקטור למאורע "j הוא קודקוד מבודד". אז $X = \sum_{j=1}^{n} I_j$ סופר את מספר הקודקודים המבודדים. אנחנו נרצה להראות שהתוחלת שואפת לאינסוף. מתקיים:

$$\mathbb{E}(X) \overset{א}{=} \sum_{j=1}^{n} \mathbb{E}(I_j) \overset{ב}{=} \sum_{j=1}^{n} \mathbb{P}(I_j = 1) \overset{ג}{=} n(1-p)^{n-1}$$

א – לינאריות התוחלת
ב – תוחלת של אינדיקטור
ג – ההסתברות שקודקוד מסוים יהיה מבודד, כפול n קודקודים.

עכשיו, בגלל ה-p שבחרנו, מתקיים: (1)

$$\lim_{n\to\infty}\mathbb{E}(X) =^{\text{א}} \lim_{n\to\infty} n(1-p)^{n-1} \geq^{\text{ב}} \lim_{n\to\infty} ne^{-(p+p^2)(n-1)} \geq^{\text{ג}}$$
$$\lim_{n\to\infty} ne^{-\frac{(1-\varepsilon/2)\ln n}{n}\cdot n} =^{\text{ד}} \lim_{n\to\infty} n^{1-(1-\varepsilon/2)} =^{\text{ה}} \lim_{n\to\infty} n^{\varepsilon/2} = \infty$$

א – נציב את מה שמצאנו

ב – אנחנו רגילים להשתמש באי-השוויון $1-p \leq e^{-p}$, אבל פה אנחנו צריכים את הכיוון ההפוך. אז נשתמש בזה שהם קרובים: $1-p \approx e^{-p}$, ונתקן. פיתוח טור טיילור של e^{-x+x^2} מראה ש $(1-p)^{n-1} \geq e^{-(p+p^2)(n-1)}$.

ג – היינו רוצים להגיד שמה שיש בחזקה זה np. זה לא בדיוק מה שכתוב אבל זה קרוב, כי n שואף לאינסוף אז מינוס 1 זה זניח, ו-p שואף לאפס אז להוסיף p^2 זה זניח. אם היה כתוב את זה, אז היה מתקיים:

$$np \leq n\cdot\frac{(1-\varepsilon)\ln n}{n} = (1-\varepsilon)\ln n$$

אבל יש שם משהו קצת יותר גדול מ- np, אז כדי לתקן נכפול את הכל ב $\left(1+\frac{\varepsilon}{10}\right)$ ואז זה כן גדול מ- np, אבל עדיין קטן מ- $(1-\varepsilon/2)\ln n$:

$$\left(1+\frac{\varepsilon}{10}\right)np \leq n\cdot\frac{\left(1+\frac{\varepsilon}{10}\right)(1-\varepsilon)\ln n}{n} = \left(1+\frac{\varepsilon}{10}\right)(1-\varepsilon)\ln n \leq \left(1-\frac{\varepsilon}{2}\right)\ln n$$

אז רצינו להקטין את הגבול, ובגלל שהחזקה במינוס זה אומר שצריך להגדיל את מה שיש בתוך הביטוי. במקום $p+p^2$ לקחנו את p שזה קטן יותר, אז במקום לכפול ב $(1-\varepsilon)$ כפלנו ב $(1-\varepsilon/2)$.

ד – ה-n מצטמצמים, e בחזקת לן מצטמצם.

ה – פתיחה של הסוגריים.

נחשב את השונות: לכל $1 \leq j \leq n$ מתקיים:

$$\mathrm{Var}(I_j) = \mathbb{E}(I_j^2) - \left(\mathbb{E}(I_j)\right)^2 \leq^{\text{א}} \mathbb{E}(I_j) = (1-p)^{n-1}$$

א – תוחלת של אינדיקטור, $\mathbb{E}(I_j) = \mathbb{E}(I_j^2)$

ומתקיים לכל $1 \leq i < j \leq n$:

$$\mathrm{Cov}(I_i, I_j) = \mathbb{E}(I_iI_j) - \mathbb{E}(I_i)\mathbb{E}(I_j) =^{\text{א}} \mathbb{P}(I_i = 1, I_j = 1) - (1-p)^{2n-2} =^{\text{ב}}$$
$$(1-p)^{2n-3} - (1-p)^{2n-2} =^{\text{ג}} p(1-p)^{2n-3}$$

א – אנחנו יודעים את התוחלת של כל אחד.

ב – מה ההסתברות של שני האינדיקטורים ביחד? עבור שני קודקודים, ההסתברות ששניהם מבודדים.
ההסתברות שהראשון מבודד: עבור $n-1$ קודקודים יצא $1-p$, ועבור הקודקוד השני צריך לבדוק רק $n-2$ קודקודים.

ג – גורם משותף.

ולכן,

$$\mathrm{Var}(X) = \sum_{j=1}^{n}\mathrm{Var}(I_j) + 2\sum_{1\leq i<j\leq n}\mathrm{Cov}(I_i, I_j) \leq^{\text{א}} n(1-p)^{n-1} + n^2p(1-p)^{2n-3}$$

א – לפי מה שמצאנו לעיל.

אז מתקיים:

$$\frac{\mathrm{Var}(X)}{\left(\mathbb{E}(X)\right)^2} \leq^{א} \frac{\mathbb{E}(X)+n^2p(1-p)^{2n-3}}{\left(\mathbb{E}(X)\right)^2} =^{ב} \frac{1}{\mathbb{E}(X)}+\frac{p}{1-p} =^{ג} o(1)$$

א – לפי מה שמצאנו
ב – נפרק את השבר, ונציב את מה שמצאנו וזה מצטמצם.
ג – כי התוחלת שואפת לאינסוף, ו-p שואף לאפס.

אז לפי המומנט השני, נקבל:

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

הוכחה 2: נניח ש $p \geq (1+\varepsilon)\ln n/n$. אם G לא קשיר, אז רכיב הקשירות הקטן ביותר שלו הוא בגודל k עבור: $1 \leq k \leq \lfloor n/2 \rfloor$. ולכן: (2)

$$\mathbb{P}(\text{G is disconnected}) \leq^{א} \sum_{k=1}^{n\backslash 2}\binom{n}{k}(1-p)^{k(n-k)} \leq^{ב} \sum_{k=1}^{\sqrt{n}}\binom{n}{k}e^{-pk(n-k)}+\sum_{k=\sqrt{n}}^{n\backslash 2}\binom{n}{k}e^{-pk(n-k)}$$

א – עבור כל גודל אפשרי של הרכיב הקשירות הקטן, נבחר שאין צלעות בין הרכיב לשאר הגרף.

ב – נפצל את הסכום לשניים. וגם $1-p \leq e^{-p}$

נתבונן בכל חלק של הסכום בנפרד. החלק הראשון: (3)

$$\sum_{k=1}^{\sqrt{n}}\binom{n}{k}e^{-pk(n-k)} \leq^{א} \sum_{k=1}^{\sqrt{n}}n^k e^{-pk(n-k)}=\sum_{k=1}^{\sqrt{n}}\left(ne^{-p(n-k)}\right)^k$$
$$\leq^{ב} \sum_{k=1}^{\sqrt{n}}\left(ne^{-\frac{(1+\varepsilon)\ln n}{n}(n-\sqrt{n})}\right)^k \leq^{ג} \sum_{k=1}^{\sqrt{n}}\left(ne^{-\left(1+\frac{\varepsilon}{2}\right)\ln n}\right)^k=\sum_{k=1}^{\sqrt{n}}n^{-\frac{\varepsilon k}{2}} \leq \sum_{k=1}^{\infty}n^{-\frac{\varepsilon k}{2}}=\frac{n^{-\varepsilon\backslash 2}}{1-n^{-\varepsilon\backslash 2}}$$

א – חסם עליון על $\binom{n}{k}$

ב – נציב במקום p. זה קטן יותר מp, ולכן אי השוויון מתקיים. וגם, k חסום ב $\sqrt{n}$, אז אם נרשום את השורש במקום k זה המקסימום, ובגלל שהחזקה שלילית זה בעצת מגדיל את הסכום.

ג – בגלל ש $\sqrt{n}=o(n)$, כאשר n שואף לאינסוף ההפרש שואף ל-n, אז אפשר "לשלם" בזה שנקטין את אפסילון.

החלק השני: (4)

$$\sum_{k=\sqrt{n}}^{n\backslash 2}\binom{n}{k}e^{-pk(n-k)} \leq^{א} \sum_{k=\sqrt{n}}^{n\backslash 2}\left(\frac{en}{k}\right)^k e^{-pk(n-k)}=\sum_{k=\sqrt{n}}^{n\backslash 2}\left(\frac{en}{k}\cdot e^{-p(n-k)}\right)^k \leq^{ב} \sum_{k=\sqrt{n}}^{n\backslash 2}\left(\frac{en}{\sqrt{n}}\cdot e^{-\frac{(1+\varepsilon)\ln n}{n}\cdot\left(n-\frac{n}{2}\right)}\right)^k$$
$$=^{ג} \sum_{k=\sqrt{n}}^{n\backslash 2}\left(en^{1\backslash 2}\cdot e^{-(1\backslash 2+\varepsilon\backslash 2)\ln n}\right)^k=\sum_{k=\sqrt{n}}^{n\backslash 2}\left(en^{-\varepsilon\backslash 2}\right)^k \leq \sum_{k=1}^{\infty}\left(en^{-\varepsilon\backslash 2}\right)^k=\frac{en^{-\varepsilon\backslash 2}}{1-en^{-\varepsilon\backslash 2}}$$

א - $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$

ב – כמו בסכום הקודם, נציב במקום p ונציב את ה-k המקסימלי. במכנה נציב את ה-k המינימלי.
ג – צמצום.

נחבר את 2,3,4 ונקבל:

$$\lim_{n\to\infty}\mathbb{P}(\text{G is disconnected}) \leq \lim_{n\to\infty}\frac{n^{-\varepsilon\backslash 2}}{1-n^{-\varepsilon\backslash 2}}+\lim_{n\to\infty}\frac{en^{-\varepsilon\backslash 2}}{1-en^{-\varepsilon\backslash 2}}=0+0=0$$

כנדרש.