Grafy losowe i sieci złożone

Informatyka algorytmiczna, WIiT PWr
semestr zimowy 2024/2025

Lista 1. momentu

- $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\mathbb{N}^+ = \{1, 2, 3, \ldots\}$
- $[n] = \{1, 2, \ldots, n\}$ dla $n \in \mathbb{N}^+$
- $f(n) \ll g(n)$, jeżeli $f(n)/g(n) \xrightarrow{n\to\infty} 0$
- $f(n) \gg g(n)$, jeżeli $f(n)/g(n) \xrightarrow{n\to\infty} \infty$

Zad. 1 — Niech $G_{n,p}$ będzie dwumianowym grafem losowym. Zdefiniujmy następujące zmienne losowe:
- $|E_{n,p}| = e(G_{n,p})$ - liczba krawędzi w $G_{n,p}$,
- $T_{n,p}$ - liczba trójkątów w $G_{n,p}$,
- $I_{n,p}$ - liczba wierzchołków izolowanych w $G_{n,p}$,
- $D^{\geq 2}_{n,p}$ - liczba wierzchołków stopnia co najmniej 2 w $G_{n,p}$.

Wyznacz ich wartości oczekiwane.

Zad. 2 — Niech $G_{n,p}$ będzie dwumianowym grafem losowym oraz niech $m \in \{0, 1, \ldots, \binom{n}{2}\}$. Wyznacz $q_m = \mathbb{P}[e(G_{n,p}) = m]$ oraz oblicz $\sum_{m=0}^{\binom{n}{2}} q_m$.

Zad. 3 — Ile razy w życiu dowodził(a|e)ś nierówności Markova? Zwiększ tę liczbę o 1.

Nierówność Markova. Niech X będzie nieujemną zmienną losową. Wtedy dla każdego $a > 0$

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

Rozważmy ciąg zmiennych losowych $(X_n)_{n=0}^{\infty}$ przyjmujących wartości w zbiorze liczb naturalnych $\mathbb{N}$.

1. Co można powiedzieć o $\lim_{n\to\infty} \mathbb{P}[X_n = 0]$, jeśli $\lim_{n\to\infty} \mathbb{E}[X_n] = 0$?
2. Podaj przykład ciągu zmiennych losowych takich, że $\mathbb{E}[X_n] \xrightarrow{n\to\infty} \infty$, ale $\mathbb{P}[X_n = 0] \xrightarrow{n\to\infty} 1$.
3. Przemyśl punkt 1. i 2.

Zad. 4 — Niech $G_{n,p}$ będzie dwumianowym grafem losowym, zaś $\mathcal{P}$ własnością „graf posiada co najmniej jedną krawędź". Udowodnij, że dla $p = p(n) \ll 1/n^2$

$$\mathbb{P}[G_{n,p} \in \mathcal{P}] \xrightarrow{n\to\infty} 0.$$

Zad. 5 — Niech $G_{n,p}$ będzie dwumianowym grafem losowym, zaś $\mathcal{P}_\Delta$ własnością „graf posiada co najmniej jeden trójkąt". Zaproponuj sensowną funkcję $f(n)$ taką, że dla $p = p(n) \ll f(n)$

$$\mathbb{P}[G_{n,p} \in \mathcal{P}_\Delta] \xrightarrow{n\to\infty} 0.$$

Zad. 6 — Niech $G_{n,p}$ będzie dwumianowym grafem losowym, zaś $\mathcal{P}_\bullet$ własnością „graf nie posiada wierzchołków izolowanych". Niech $\varepsilon > 0$. Udowodnij, że dla $p = p(n) \geq \frac{(1+\varepsilon)\ln n}{n}$

$$\mathbb{P}[G_{n,p} \in \mathcal{P}_\bullet] \xrightarrow{n\to\infty} 1.$$

(Skorzystaj z nierówności $1 - x \leq e^{-x}$ dla $x \in \mathbb{R}$.)

1. Czy w zadaniu 4 dla $p = p(n) \leq (1-\varepsilon)/n^2$ otrzymalibyśmy $\mathbb{P}[G_{n,p} \in \mathcal{P}] \xrightarrow{n\to\infty} 0$?
2. Czy w zadaniu 5 dla $p = p(n) \leq (1-\varepsilon)f(n)$ otrzymalibyśmy $\mathbb{P}[G_{n,p} \in \mathcal{P}_\Delta] \xrightarrow{n\to\infty} 0$?

Lista 2. momentu

Zad. 1 — Rozważ funkcję $f(p) = \mathbb{P}[G_{n,p} \in \mathcal{P}]$ dla $p \in [0,1]$. Uzasadnij, że $f(p)$ jest ciągła na $[0,1]$. Ile wynoszą $f(0)$ oraz $f(1)$ dla własności wymienionych w Misji 1 (możesz przyjąć $n > 5$)?

Zad. 2 — Niech $\mathcal{P}$ będzie własnością monotonicznie rosnącą. Uzasadnij, że $\mathcal{P}^C$ jest własnością monotonicznie malejącą.

Zad. 3 — Pokazaliśmy, że jeśli $\mathcal{P}$ jest własnością monotonicznie rosnącą, to

$$\mathbb{P}[G_{n,m} \in \mathcal{P}] \leq \mathbb{P}[G_{n,m'} \in \mathcal{P}] \quad \text{dla} \quad m < m'.$$

Czy nierówność pozostaje prawdziwa, gdy własność $\mathcal{P}$ jest monotonicznie malejąca?

Zad. 4 — Zaproponuj coupling, dzięki któremu uzasadnisz, że jeśli $\mathcal{P}$ jest własnością monotonicznie rosnącą, to

$$\mathbb{P}[G_{n,p} \in \mathcal{P}] \leq \mathbb{P}[G_{n,p'} \in \mathcal{P}] \quad \text{dla} \quad p < p'.$$

Zad. 5 — Ile razy w życiu dowodził(a|e)ś nierówności Czebyszewa? Zwiększ tę liczbę o 1.

Nierówność Czebyszewa. Niech X będzie zmienną losową. Dla każdego $t > 0$ mamy

$$\mathbb{P}[|X - \mathbb{E}X| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}.$$

Rozważmy ciąg dodatnich zmiennych losowych $(X_n)_{n=0}^{\infty}$.

1. Jak wygląda nierówność Czebyszewa dla $t = \mathbb{E}[X_n]$?
2. Wstaw odpowiednią nierówność pomiędzy $\mathbb{P}[X_n = 0]$ i $\mathbb{P}[|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]]$.
3. Co można powiedzieć o $\lim_{n\to\infty} \mathbb{P}[X_n = 0]$, jeśli $\lim_{n\to\infty} \frac{\mathbb{E}[X_n^2]}{E[X_n]^2} = 1$?

Zad. 6 — Niech $G_{n,p}$ będzie dwumianowym grafem losowym. Zdefiniujmy następujące zmienne losowe:

- $|E_{n,p}| = e(G_{n,p})$ - liczba krawędzi w $G_{n,p}$,
- $T_{n,p}$ - liczba trójkątów w $G_{n,p}$.

Spróbuj wyznaczyć ich wariancje.

Lista 3. progów

Zad. 1 — Ponumerujmy wszystkie trójkąty w etykietowanym grafie K_n liczbami naturalnymi od 1 do $\binom{n}{3}$. Niech $T := T_{n,p}$ będzie zmienną losową oznaczającą liczbę trójkątów w grafie losowym $G_{n,p}$. Dla $i \in \left[\binom{n}{3}\right]$ definiujemy zmienne losowe

$$T_i = \begin{cases} 1 & \text{jeżeli trójkąt numer } i \text{ należy do } G_{n,p} \\ 0 & \text{w p.p.} \end{cases}$$

Oczywiście, $T = \sum_{i=1}^{\binom{n}{3}} T_i$.

1. Uzasadnij, że $\mathbb{E}[T^2] = E[T] \sum_{i=1}^{\binom{n}{3}} \mathbb{P}[T_i = 1 | T_1 = 1]$.
2. Wyznacz $\mathrm{Var}[T]$.
3. Czy równość analogiczna do tej z punktu 1. zachodzi dla poniższych zmiennych losowych?
   - $I_{n,p}$ - liczba wierzchołków izolowanych w $G_{n,p}$.
   - $E_{n,p}$ - liczba krawędzi w $G_{n,p}$.
   - $K_{n,p}^{(k)}$ - liczba kopii grafu K_k w $G_{n,p}$.

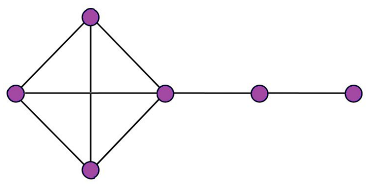
Zad. 2 — Udowodnij, że $p^*(n) = \frac{1}{n}$ jest progiem dla własności $\mathcal{P}_\triangle$ - „graf posiada trójkąt" w $G_{n,p}$.

Zad. 3 — Zaproponuj funkcję $p^*(n)$, która mogłaby być progiem dla własności $\mathcal{P}_{K_4}$ - „graf posiada kopię K_4" w $G_{n,p}$. Udowodnij, że Twoje $p^*(n)$ jest progiem dla $\mathcal{P}_{K_4}$.

Zad. 4 — Zaproponuj próg dla własności $\mathcal{P}_{K_k}$ - „graf posiada kopię K_k" w $G_{n,p}$ dla stałego $k \in \mathbb{N}$. Spróbuj udowodnić, że rzeczywiście jest progiem.

Zad. 5 — Zaproponuj próg dla własności $\mathcal{P}_H$ - „graf posiada kopię H" w $G_{n,p}$, gdzie H jest dowolnym grafem o stałej liczbie wierzchołków v_H i stałej liczbie krawędzi e_H.

1. Jak wyglądałby zaproponowany przez Ciebie próg dla grafu przedstawionego na Rysunku 1?
2. Porównaj go z progiem dla $\mathcal{P}_{K_4}$.
3. Wyciągnij wniosek.



Rysunek 1: Latawiec / Lollipop

Zad. 6 — Udowodnij, że $p^*(n) = \frac{\ln n}{n}$ jest ostrym progiem dla własności $\mathcal{P}_\bullet$ - „graf nie posiada wierzchołków izolowanych" w $G_{n,p}$.

Lista 4. wyzwań

Wszystkie wyzwania tej listy dotyczą grafu losowego $G_{n,p}$. Być może warto sobie przypomnieć, że

- $\binom{n}{k} \leq \frac{n^k}{k!} \leq n^k$, a nawet $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ dla $n, k \in \mathbb{N}$,
- $1 - x \leq e^{-x}$ dla $x \in \mathbb{R}$, a także $1 - f(n) \sim e^{-f(n)}$ dla $f(n) = o(n)$,
- $f(n) = o(g(n)) \iff f(n) \ll g(n)$,
- $f(n) = \omega(g(n)) \iff f(n) \gg g(n)$.

Zad. 1 — Udowodnij, że $p^*(n) = \frac{1}{n^{3/2}}$ jest progiem dla własności $\mathcal{P}$ - „graf jest sumą wierzchołków izolowanych i krawędzi izolowanych”.
Wskazówka: Czego nie ma graf, który jest sumą wierzchołków izolowanych i krawędzi izolowanych?

Zad. 2 — Niech $p = p(n) \ll \frac{1}{n}$ (równoważnie $p = p(n) \leq \frac{1}{\omega(1)n}$). Udowodnij, że

$$\mathbb{P}[G_{n,p} \text{ jest lasem}] \xrightarrow{n\to\infty} 1.$$

Uzasadnij, że $p^*(n) = \frac{1}{n}$ jest progiem dla własności $\mathcal{P}_F$ - „graf jest lasem”.
Wskazówka: Pomyśl o cyklach długości co najmniej 3.

Zad. 3 — Niech $G = (V, E)$ będzie grafem prostym, $|V| = n$. Cięciem w G nazwijmy taką partycję zbioru wierzchołków $\{S, V \setminus S\}$, gdzie $1 \leq |S| \leq n/2$ oraz dla każdej krawędzi $\{v, w\} \in E$ mamy $v, w \in S$ lub $v, w \in V \setminus S$. Niech $\varepsilon > 0$ oraz $p = p(n) \geq \frac{(1+\varepsilon)\ln n}{n}$. Udowodnij, że

$$\mathbb{P}[G_{n,p} \text{ zawiera cięcie}] \xrightarrow{n\to\infty} 0.$$

Wskazówka: Jeśli badasz pewną sumę $\sum_{i=1}^{n/2}(\ldots)$, to rozważ osobno $\sum_{i=1}^{\lfloor\sqrt{n}\rfloor}(\ldots)$ oraz $\sum_{i=\lfloor\sqrt{n}\rfloor+1}^{n/2}(\ldots)$.

Zad. 4 — Udowodnij, że $p^*(n) = \frac{\ln n}{n}$ jest ostrym progiem dla własności $\mathcal{P}_C$ - „graf jest spójny”.

Lista 5. ćwiczeń

Zad. 1 — Niech $X = \sum_{i=1}^{M} X_i$, gdzie X_i są zmiennymi losowymi. Pokaż, że

$$\mathrm{Var}[X] = \sum_{i=1}^{M}\sum_{i=1}^{M} \mathrm{Cov}[X_i, X_j].$$

Dodatkowo, jeżeli X_i są zmiennymi losowymi przyjmującymi wartości w zbiorze $\{0, 1\}$, to

$$\mathrm{Var}[X] = \sum_{i=1}^{M}\sum_{i=1}^{M} \left(\mathbb{P}[X_i = 1, X_j = 1] - \mathbb{P}[X_i = 1]\mathbb{P}[X_j = 1]\right).$$

Zad. 2 — Niech X, Y będą zmiennymi losowymi o dodatnich wariancjach, przyjmującymi wartości w zbiorze $\{0, 1\}$, przy czym $\mathbb{P}[Y = 1] > 0$. Pokaż, że jeżeli X i Y są pozytywnie skorelowane (czyli $\mathrm{Cor}[X, Y] \geq 0$), to

$$\mathbb{P}[X = 1|Y = 1] \geq \mathbb{P}[X = 1].$$

Zad. 3 — Niech $X_1, X_2, X_3, \ldots$ oraz X będą zmiennymi losowymi przyjmującymi wartości w zbiorze liczb naturalnych. Pokaż, że

$$X_n \xrightarrow{d} X \quad \iff \quad \forall k \in \mathbb{N} \quad \mathbb{P}[X_n = k] \xrightarrow{n\to\infty} \mathbb{P}[X = k].$$

Zad. 4 — Rozważmy ciąg zmiennych losowych $(X_n)_{n\geq 1}$ takich, że X_n pochodzi z rozkładu dwumianowego z parametrami n oraz p_n, przy czym $np_n \xrightarrow{n\to\infty} \lambda \in \mathbb{R}^+$. Pokaż, że $X_n \xrightarrow{d} Po(\lambda)$.

Zad. 5 — Niech $\{X_n\}_{n\geq 1}$ będzie ciągiem zmiennych losowych przyjmujących wartości w zbiorze liczb naturalnych. Niech $\{\lambda_n\}_{n\geq 1}$ będzie ciągiem dodatnich liczb rzeczywistych takich, że $\lambda_n \xrightarrow{n\to\infty} \lambda$. Pokaż, że

$$d_{TV}(X_n, Po(\lambda_n)) \xrightarrow{n\to\infty} 0 \quad \iff \quad X_n \xrightarrow{d} Po(\lambda).$$

Przypomnienie: Dla całkowitoliczbowych zmiennych losowych X i Y

$$d_{TV}(X, Y) = \frac{1}{2}\sum_{k\in\mathbb{Z}} |\mathbb{P}[X = k] - \mathbb{P}[Y = k]|,$$

zaś $X_n \xrightarrow{d} Po(\lambda)$ oznacza zbieżność według rozkładu.

Lista 6. - po prostu lista

Zad. 1 — Wyznacz próg dla własności $\mathcal{P}_\bullet$ - „graf nie posiada wierzchołków izolowanych" w dwudzielnym grafie losowym $G_{n,n,p}$.

Zad. 2 — Niech $G = (V, E)$ będzie grafem dwudzielnym z partycją wierzchołków $V = U \cup W$, gdzie $|U| = |W| = n$. Niech $S \subseteq U$ będzie najmniejszym (w sensie mocy zbioru) zbiorem naruszającym warunek Halla (czyli $|S| > |N(S)|$, gdzie $N(S)$ to zbiór sąsiadów wierzchołków z S).

1. Pokaż, że $|S| = |N(S)| + 1$.
2. Pokaż, że $|S| \leq \lceil n/2 \rceil$.
3. Pokaż, że każdy wierzchołek z $N(S)$ jest sąsiadem co najmniej dwóch wierzchołków z S.

Zad. 3 — Niech $I_{n,p}$ oznacza liczbę wierzchołków izolowanych w grafie losowym $G_{n,p}$. Niech $p = \frac{\ln n + c}{n}$, gdzie $c \in \mathbb{R}$ jest stałą. Jeżeli $I_{n,p} \xrightarrow{d} Y$, to jaki rozkład ma Y? Wyznacz $\lim_{n\to\infty} \mathbb{P}[I_{n,p} = 0]$.

Zad. 4 — Graf H jest zbalansowany, jeżeli $\mu(H) = d(H)$. Graf H jest ściśle zbalansowany, jeżeli $d(H) > d(K)$ dla każdego K będącego właściwym podgrafem H. Podaj przykład grafu, który

- nie jest zbalansowany,
- jest zbalansowany, ale nie jest ściśle zbalansowany,
- jest ściśle zbalansowany.