

## Chủ đề 11. Đồ thị ngẫu nhiên

Nhận thấy rằng trong các mục trước của bài viết chúng ta đã có giới thiệu qua định lý cơ bản sau của nhà toán học Erdős: “Với mọi số nguyên dương  $k > 0$ , luôn tồn tại một đồ thị  $G$  với  $g(G) > k$  và  $\chi(G) > k$ ”.

Một câu hỏi được đặt ra ở đây đó là: “Làm thế nào để chúng ta có thể chứng minh định lý trên?”.

Cách tiếp cận tiêu chuẩn ở đây sẽ là xây dựng một đồ thị với cả hai thuộc tính nêu trên từng bước một bằng quy nạp toán học theo chỉ số  $k$ . Mặc dù cách tiếp cận nêu trên khá đơn giản nhưng việc trình bày chứng minh trên thực tế khá là dài.

Trong một bài luận tổng hợp của mình vào năm 1959, nhà toán học Erdős đã giới thiệu một cách tiếp cận hoàn toàn khác cho vấn đề trên như sau: “Với mỗi số tự nhiên  $n$ , chúng ta xây dựng một không gian xác suất trên tập các đồ thị với  $n$  đỉnh và chứng minh rằng, với một số phép đo xác suất được lựa chọn cẩn thận, xác suất để một đồ thị  $n$  – đỉnh có cả hai tính chất nêu trên là dương với  $n$  đủ lớn”.

Ngày nay, cách tiếp cận này, **phương pháp xác suất**, đã được phát triển thành một kỹ thuật chứng minh tinh tế và linh hoạt không chỉ trong lý thuyết đồ thị mà còn cho cả các mảng khác của toán rời rạc. Lý thuyết về các đồ thị ngẫu nhiên hiện nay đang là một chủ đề được quan tâm nghiên cứu phát triển mạnh mẽ. Mục đích của bài viết này nhằm cung cấp một bài giới thiệu sơ cấp nhưng chặt chẽ về các đồ thị ngẫu nhiên: không cần thiết phải hiểu rõ các khái niệm, ý tưởng và kỹ thuật chứng minh cơ bản nhưng đủ hiểu sức mạnh và sự tinh tế ẩn sau các phép tính được trình bày.

### 11.1. Khái niệm về đồ thị ngẫu nhiên.

Giả sử cho trước một tập  $V$  với  $n$  phần tử, tạm giả sử  $V = \{0, \dots, n-1\}$ .

Mục đích của chúng ta là biến tập  $\mathcal{G}$  gồm tất cả các đồ thị trên  $V$  thành một không gian xác suất và sau đó xem xét các loại câu hỏi thường gặp về các đối tượng ngẫu nhiên như: “Xác suất để một đồ thị  $G \in \mathcal{G}$  có tính chất này hay tính chất kia là bao nhiêu? Kỳ vọng của một đặc trưng bất biến cho trước của  $G$  là bao nhiêu? (chẳng hạn như girth  $g(G)$  hoặc sắc số  $\chi(G)$ )”.

Một cách trực quan, chúng ta có thể tạo ra  $G$  một cách ngẫu nhiên như sau: “Với mỗi  $e \in [V]^2$ , chúng ta quyết định xem  $e$  có phải là một cạnh của  $G$  hay không bằng một phép thử ngẫu nhiên; các phép thử này được thực hiện độc lập và xác suất thành công của mỗi phép thử này (nói cách khác, xác suất để  $e$  là một cạnh của đồ thị  $G$ ) luôn là một hằng số  $p \in [0,1]$  cố định”.

Khi đó, nếu  $G_0$  là một đồ thị cố định nào đó trên  $V$  với  $m$  cạnh thì biến cố sơ cấp  $\{G_0\}$  sẽ có xác suất là  $p^m q^{\binom{n}{2}-m}$  với  $q := 1 - p$ . Nói cách khác, xác suất để đồ thị được tạo ngẫu nhiên  $G$  của chúng ta là đồ thị  $G_0$  cho trước là  $p^m q^{\binom{n}{2}-m}$ . Nhưng nếu xác suất của tất cả các biến cố sơ cấp đều được xác định như trên thì số đo xác suất của toàn bộ không gian  $\mathcal{G}$  cũng sẽ được xác định. Như vậy, phần việc còn lại của chúng ta là chứng minh rằng thực sự tồn tại một phép đo xác suất như vậy trên  $\mathcal{G}$  – một phép đo mà tất cả các cạnh riêng lẻ xảy ra độc lập với xác suất  $p \in [0,1]$ .

Để xây dựng một phép đo xác suất như vậy trên  $\mathcal{G}$  một cách hình thức, chúng ta sẽ xác định cho mỗi cạnh tiềm năng  $e \in [V]^2$  một không gian xác suất nhỏ hơn của riêng nó như sau:  $\Omega_e := \{0_e, 1_e\}$  với  $P_e(\{1_e\}) := p$  và  $P_e(\{0_e\}) := q$  là các xác suất của các biến cố sơ cấp khác nhau của nó.

Khi đó, không gian xác suất  $\mathcal{G} = \mathcal{G}(n, p)$  cần tìm chính là không gian tích  $\Omega := \prod_{e \in [V]^2} \Omega_e$ .

Như vậy, về mặt hình thức, một phần tử của  $\Omega$  là một ánh xạ  $\omega$  gán cho mỗi cạnh  $e \in [V]^2$  một kí hiệu  $0_e$  hoặc một kí hiệu  $1_e$  còn độ đo xác suất  $P$  trên  $\Omega$  là độ đo tích của tất cả các độ đo  $P_e$ .

Trên thực tế, chúng ta thường đồng nhất ánh xạ  $\omega$  với một đồ thị  $G$  xác định trên  $V$  với tập cạnh là  $E(G) = \{e | w(e) = 1_e\}$  và  $G$  được gọi là một **đồ thị ngẫu nhiên với xác suất cạnh  $p$**  (hoặc đơn giản chỉ là **đồ thị ngẫu nhiên**).

Theo thuật ngữ xác suất tiêu chuẩn, chúng ta có thể xem bất kì một tập hợp đồ thị nào trên  $V$  cũng là một **biến cố** trong  $\mathcal{G}(n, p)$ . Đặc biệt, với mỗi  $e \in [V]^2$ , tập  $A_e := \{\omega | \omega(e) = 1_e\}$  gồm tất cả các đồ thị  $G$  xác định trên  $V$  với  $e \in E(G)$  là biến cố: “ $e$  là một cạnh của  $G$ ”.

Với cách định nghĩa hình thức nêu trên, chúng ta có thể đưa ra một chứng minh hình thức cho những cảm nhận trực quan dẫn đường từ đầu bài viết đến giờ:

### Kết quả 11.1.1.

Các biến cố  $A_e$  là đôi một độc lập và xác suất để xảy ra mỗi biến cố  $A_e$  đều bằng  $p$ .

### Chứng minh kết quả 11.1.1.

Từ định nghĩa không gian xác suất  $\mathcal{G}(n, p)$ ,  $A_e = \{1_e\} \times \prod_{e' \neq e} \Omega_{e'}$ .

Do  $P$  là độ đo tích của các độ đo  $P_e$  nên  $P(A_e) = p \cdot \prod_{e' \neq e} 1 = p$ .

Tương tự, với mọi tập con  $\{e_1, \dots, e_k\}$  bất kì của  $[V]^2$ ,

$$P(A_{e_1} \cap \dots \cap A_{e_k}) = P(\{1_{e_1}\} \times \dots \times \{1_{e_k}\} \times \prod_{e \notin \{e_1, \dots, e_k\}} \Omega_e) = p^k = P(A_{e_1}) \dots P(A_{e_k}). \quad \square$$

Như đã trình bày ở trên, độ đo xác suất  $P$  được xác định duy nhất bởi  $p$  và giả định rằng các biến cố  $A_e$  là đôi một độc lập. Như vậy, khi tính toán xác suất trong  $\mathcal{G}(n, p)$ , chúng ta thường chỉ làm việc với hai giả định nêu trên là đủ: việc trình bày mô hình xác suất trên  $\mathcal{G}(n, p)$  nhằm chứng minh chặt chẽ tính đúng đắn của hai giả định trên nên việc trình bày thêm về nó cũng không còn cần thiết nữa.

Chẳng hạn, chúng ta hãy xem xét biến cố: “ $G$  có chứa một đồ thị con cố định  $H$  xác định trên một tập con của  $V$  như một đồ thị con”. Kí hiệu  $k := |H|$  và  $l := \|H\|$ . Khi đó, xác suất của biến cố  $H \subseteq G$  bằng tích của tất cả các xác suất  $P_e$  lấy trên tất cả các cạnh  $e \in H$ , nên  $P[H \subseteq G] = p^l$ . Ngược lại, xác suất để  $H$  là một đồ thị con cảm sinh của  $G$  sẽ bằng  $p^l q^{\binom{k}{2} - l}$ : điều này suy ra từ các cạnh bị thiếu trong  $H$  cũng được yêu cầu bị thiếu trong  $G$  và chúng xảy ra độc lập với xác suất  $q$ .

Mặc dù, xác suất  $P_H$  để “ $G$  có chứa một đồ thị con cảm sinh đẳng cấu với  $H$ ” thường sẽ rất khó để tính được chính xác: do các trường hợp có thể có của  $H$  trên các tập con của  $V$  chồng lên nhau và các biến cố này xảy ra trong  $G$  không độc lập với nhau. Tuy nhiên, chúng ta có thể chứng minh được rằng tổng của các xác suất  $P[H \simeq G[U]]$  lấy trên tập tất cả các  $k$  – tập  $U \subseteq V$  đóng vai trò như một chặn trên của  $P_H$  (điều này suy ra từ việc  $P_H$  là xác suất của hợp của các biến cố này).

Kết quả sau đây sẽ cung cấp cho chúng ta một chặn trên cho biến cố “ $G$  có chứa một đồ thị con cảm sinh đẳng cấu với  $H$ ” trong trường hợp đặc biệt  $H = \overline{K_k}$ :

### Kết quả 11.1.2.

Với mọi số nguyên dương  $n, k$  với  $n \geq k \geq 2$ , xác suất để một đồ thị  $G \in \mathcal{G}(n, p)$  có chứa một tập độc lập với  $k$  đỉnh tối đa không vượt quá  $\binom{n}{k} q^{\binom{k}{2}}$ , nghĩa là  $P[\alpha(G) \geq k] \leq \binom{n}{k} q^{\binom{k}{2}}$ .

### Chứng minh kết quả 11.1.2.

Nhận thấy rằng do xác suất để một  $k$  – tập cố định  $U \subseteq V$  là một tập độc lập trong  $G$  là  $q^{\binom{k}{2}}$  nên xác suất để một đồ thị  $G \in \mathcal{G}(n, p)$  có chứa một tập độc lập với  $k$  đỉnh tối đa không vượt quá  $\binom{n}{k} q^{\binom{k}{2}}$  (điều này suy ra từ việc có đúng  $\binom{n}{k}$   $k$  – tập  $U \subseteq V$ ).  $\square$

Lập luận tương tự, xác suất để một đồ thị  $G \in \mathcal{G}(n, p)$  có chứa  $K_k$  tối đa không vượt quá  $\binom{n}{k} p^{\binom{k}{2}}$ , nghĩa là

$$P[\omega(G) \geq k] \leq \binom{n}{k} p^{\binom{k}{2}}.$$

Nhận thấy rằng nếu chúng ta cố định số nguyên dương  $k$  và chọn  $n$  đủ nhỏ sao cho tổng các chặn trên của các xác suất  $P[\alpha(G) \geq k]$  và  $P[\omega(G) \geq k]$  nhỏ hơn 1 thì  $\mathcal{G}$  sẽ chứa các đồ thị  $G$  không thỏa mãn tính chất Ramsey: nghĩa là  $\mathcal{G}$  có chứa các đồ thị  $G$  không có chứa  $K_k$  và  $\overline{K_k}$ . Tuy nhiên, nếu chúng ta chọn  $n$  nhỏ hơn hoặc bằng số Ramsey của  $k$  thì khẳng định trên là hiển nhiên!. (xem lại Kết quả 9.1.1)

Nhận định trên giúp chúng ta đưa ra một chặn dưới tốt cho số Ramsey  $R(k)$  giới thiệu trong “Chủ đề 09. Dạng đồ thị của định lý Ramsey”. Chặn dưới này khá gần với chặn trên  $2^{2k-3}$  nêu Kết quả 9.1.1):

### Kết quả 11.1.3. (Erdős)

Với mọi số nguyên dương  $k \geq 3$ , số Ramsey của  $k$  thỏa mãn đánh giá  $R(k) > 2^{\frac{k}{2}}$ .

### Chứng minh kết quả 11.1.3.

Với  $k = 3$ , tính đúng đắn của khẳng định trên là hiển nhiên, do  $R(3) \geq 3 > 2^{\frac{3}{2}}$ .

Không giảm tính tổng quát của vấn đề, giả sử  $k \geq 4$ .

Ta sẽ chứng minh rằng, với mọi  $n \leq 2^{\frac{k}{2}}$ , với mọi đồ thị  $G \in \mathcal{G}(n, p)$ , các xác suất  $P[\alpha(G) \geq k]$  và  $P[\omega(G) \geq k]$  đều nhỏ hơn  $\frac{1}{2}$ .

Bằng cách chọn  $p = q = \frac{1}{2}$ , Kết quả 11.1.2) và thực hiện một lập luận tương tự Kết quả 11.1.2) cho  $\omega(G)$ , chúng ta nhận thấy rằng, với mọi  $n \leq 2^{\frac{k}{2}}$ ,

$$P[\alpha(G) \geq k], P[\omega(G) \geq k] \leq \binom{n}{k} \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

$$\begin{aligned}
&< \frac{n^k}{k!} \left(\frac{1}{2}\right)^{\binom{k}{2}} \\
&< \frac{n^k}{2^k} 2^{-\frac{k(k-1)}{2}} \text{ (điều này suy ra từ việc } k! \geq 2^k \text{ với mọi } n \geq 4) \\
&\leq \frac{\frac{k^2}{2}}{2^k} 2^{-\frac{k(k-1)}{2}} \\
&= 2^{-\frac{k}{2}} \\
&< \frac{1}{2}.
\end{aligned}$$

□

Theo ngôn ngữ của lý thuyết đồ thị ngẫu nhiên, các đặc trưng bất biến quen thuộc của đồ thị (chẳng hạn như bậc trung bình  $d(G)$ , chỉ số liên thông  $\kappa(G)$ , chỉ số girth  $g(G)$ , sắc số  $\chi(G)$ ,...) có thể được xem như một **đại lượng ngẫu nhiên không âm** xác định trên  $\mathcal{G}(n, p)$ , nghĩa là một hàm  $X: \mathcal{G}(n, p) \rightarrow [0, \infty)$ .

**Kì vọng  $E(X)$  của đại lượng  $X$**  được xác định bởi công thức  $E(X) := \sum_{G \in \mathcal{G}(n, p)} P(\{G\}) \cdot X(G)$ .

Nhận thấy rằng kì vọng  $E(X)$  là một **đại lượng tuyến tính**: nghĩa là, với mọi đại lượng ngẫu nhiên  $X, Y$  xác định trên  $\mathcal{G}(n, p)$ , với mọi số thực  $\lambda \in \mathbb{R}$ ,  $E(X + Y) = E(X) + E(Y)$  và  $E(\lambda X) = \lambda E(X)$ .

Như vậy, việc tính toán kì vọng của một đại lượng ngẫu nhiên  $X$  sẽ là một phương pháp đơn giản và hiệu quả để chứng minh sự tồn tại của một đồ thị  $G$  sao cho  $X(G) < a$  với  $a > 0$  là một số cho trước nào đó.

Thật vậy, nếu chúng ta chứng minh được kì vọng  $E(X)$  của đại lượng  $X$  nhỏ hơn  $a$  thì  $X(G)$  không thể lớn hơn  $a$  với mọi đồ thị  $G \in \mathcal{G}(n, p)$ , nên  $X$  phải nhỏ hơn  $a$  với một số đồ thị  $G \in \mathcal{G}(n, p)$ .

Ý tưởng đơn giản này đóng một vai trò cốt lõi trong hầu hết các phép chứng minh tồn tại không cần xây dựng cấu hình cụ thể mà sử dụng lý thuyết đồ thị ngẫu nhiên. Về mặt hình thức, nó có thể được phát biểu lại dưới dạng đơn giản sau:

#### Kết quả 11.1.4. (Markov)

Giả sử  $X$  là một đại lượng ngẫu nhiên không âm xác định trên  $\mathcal{G}(n, p)$  và một số thực dương  $a > 0$ .

Khi đó,  $P[X \geq a] \leq \frac{E(X)}{a}$ .

#### Chứng minh kết quả 11.1.4.

$$\begin{aligned}
E(X) &= \sum_{G \in \mathcal{G}(n, p)} P(\{G\}) \cdot X(G) \\
&\geq \sum_{\substack{G \in \mathcal{G}(n, p), \\ X(G) \geq a}} P(\{G\}) \cdot X(G) \\
&\geq \sum_{\substack{G \in \mathcal{G}(n, p), \\ X(G) \geq a}} P(\{G\}) \cdot a \\
&= P[X \geq a] \cdot a.
\end{aligned}$$

□

Do  $\mathcal{G}(n, p)$  là một không gian xác suất hữu hạn và kì vọng  $E[X]$  là một đại lượng tuyến tính nên chúng ta có thể tính  $E[X]$  bằng phương pháp đếm bằng hai cách – một phương pháp tiếp cận tiêu chuẩn trong các bài toán rời rạc. Chẳng hạn, nếu  $X$  là một đại lượng ngẫu nhiên xác định trên  $\mathcal{G}(n, p)$  đếm số các đồ thị con  $H$  của  $G$  nằm trong một tập các đồ thị  $\mathcal{H}$  cố định cho trước. Khi đó, chúng ta tính  $E(X)$  trên các đồ thị  $G \in \mathcal{G}(n, p)$  bằng một “vòng lặp ngoài” và, với mỗi  $G \in \mathcal{G}(n, p)$ , thực hiện một “vòng lặp trong” trên các đồ thị  $H \in \mathcal{H}$ .

Để minh họa phân tích nêu trên một cách chi tiết hơn, chúng ta sẽ thực hiện tính toán mẫu giá trị kì vọng của số các chu trình có độ dài  $k \geq 3$  cho trước trong một đồ thị ngẫu nhiên  $G \in \mathcal{G}(n, p)$ .

Kí hiệu  $X: \mathcal{G}(n, p) \rightarrow \mathbb{N}$  là đại lượng ngẫu nhiên gán mỗi đồ thị ngẫu nhiên  $G \in \mathcal{G}(n, p)$  với số các  $k$  – chu trình – số các đồ thị con đẳng cấu với  $C_k$ . Kí hiệu  $(n)_k := n(n-1)(n-2) \dots (n-k+1)$  là số các dãy gồm  $k$  phần tử phân biệt của một  $n$  – tập cho trước.

### Kết quả 11.1.5.

Kì vọng của số các  $k$  – chu trình chứa trong  $G \in \mathcal{G}(n, p)$  là  $E(X) = \frac{(n)_k}{2k} p^k$ .

### Chứng minh kết quả 11.1.5.

Với mỗi  $k$  – chu trình  $C$  với các đỉnh nằm trong tập  $V = \{0, \dots, n-1\}$  – tập đỉnh của họ các đồ thị ngẫu nhiên  $\mathcal{G}(n, p)$ , kí hiệu  $X_C: \mathcal{G}(n, p) \rightarrow \{0, 1\}$  là **chỉ số đặc trưng** của chu trình  $C$  xác định bởi công thức:

$$X_C: G \mapsto \begin{cases} 1, & \text{nếu } C \subseteq G, \\ 0, & \text{ngược lại.} \end{cases}$$

Do  $X_C$  chỉ có giá trị dương duy nhất là 1 nên kì vọng  $E[X_C]$  chính là xác suất  $P[X_C = 1]$  của tập tất cả các đồ thị  $G$  trong  $\mathcal{G}(n, p)$  có chứa  $C$ .

Tuy nhiên, xác suất  $P[X_C = 1]$  này chính là xác suất để  $C \subseteq G$ , nghĩa là  $E(X_C) = P[C \subseteq G] = p^k$ .

Một vấn đề đặt ra cho chúng ta đó là: “Có bao nhiêu chu trình  $C = v_0 \dots v_{k-1}$ ?”. Nhận thấy rằng do có  $(n)_k$  dãy  $v_0 \dots v_{k-1}$  các đỉnh phân biệt nằm trong  $V$  và mỗi chu trình  $C = v_0 \dots v_{k-1}$  sẽ tương ứng với  $2k$  dãy các đỉnh phân biệt như trên nên có chính xác  $\frac{(n)_k}{2k}$  chu trình  $C = v_0 \dots v_{k-1}$ .

Nhắc lại rằng đại lượng ngẫu nhiên  $X(G)$  đếm số các  $k$  – chu trình nằm trong  $G$  nên hiển nhiên nó sẽ bằng tổng của các đại lượng đặc trưng  $X_C(G)$  trong đó  $C$  là một chu trình nào đó trong số  $\frac{(n)_k}{2k}$  chu trình có chiều dài  $k$  với các đỉnh nằm trong  $V$ :

$$X = \sum_C X_C.$$

Do  $E(X)$  là một đại lượng tuyến tính nên

$$E(X) = E(\sum_C X_C) = \sum_C E(X_C) = \frac{(n)_k}{2k} p^k.$$

□

## 11.2. Phương pháp xác suất.

Một cách khái quát, phương pháp xác suất trong toán học rời rạc được phát triển lên từ ý tưởng đơn giản sau đây. Để chứng minh sự tồn tại của một cấu hình tổ hợp với một số tính chất mong muốn, chúng ta xây dựng một không gian xác suất trên một lớp các đối tượng đủ lớn và chắc chắn không rỗng, và sau đó chỉ ra một phần tử của không gian này có tính chất mong muốn với xác suất dương. Các đối tượng “sinh sống” trong các không gian xác suất này có thể thuộc bất kì loại cấu hình tổ hợp nào: phân hoạch nguyên, đồ thị, ánh xạ, ... Trong mục này của bài viết, chúng ta sẽ minh họa phương pháp xác suất bằng cách đưa ra một phép chứng minh chi tiết cho một trong những kết quả sớm nhất của nó: định lý Erdős.

Định lý Erdős phát biểu rằng: “Với mọi số nguyên dương  $k > 0$ , luôn tồn tại một đồ thị  $G$  với girth  $g(G) > k$  và sắc số  $\chi(G) > k$ ”. Chúng ta sẽ gọi các chu trình có chiều dài  $\leq k$  là **chu trình ngắn** và các tập đỉnh độc lập với  $\geq \frac{|G|}{k}$  đỉnh là **tập đỉnh độc lập lớn**. Để chứng minh định lý Erdős, chúng ta chỉ cần chỉ ra một đồ thị  $G$  không có chứa chu trình ngắn và cũng không có chứa tập đỉnh độc lập lớn. Khi đó, các lớp màu trong các phép tô màu đỉnh bất kì của  $G$  đều  $< k$ , nghĩa là chúng ta sẽ cần nhiều hơn  $k$  để tô màu các đỉnh của  $G$ .

Một vấn đề đặt ra ở đây đó là: “Làm thế nào để chúng ta có thể tìm được một đồ thị  $G$  như vậy?”. Nếu chúng ta chọn  $p$  đủ nhỏ thì các đồ thị ngẫu nhiên trong  $\mathcal{G}(n, p)$  sẽ khó có thể chứa bất kì chu trình ngắn nào. Nếu chúng ta chọn  $p$  đủ lớn thì các đồ thị ngẫu nhiên trong  $\mathcal{G}(n, p)$  sẽ khó có thể chứa bất kì tập đỉnh độc lập lớn nào. Như vậy, câu hỏi đặt ra là: “Liệu có tồn tại cách chọn  $p$  sao cho, với mọi  $n$  cho trước, đủ nhỏ để  $P[g \leq k] < \frac{1}{2}$  nhưng cũng đủ lớn để  $P[\alpha \geq n/k] < \frac{1}{2}$  hay không?”. Nếu có tồn tại một cách chọn  $p$  như trên thì  $\mathcal{G}(n, p)$  sẽ chứa ít nhất một đồ thị  $G$  không có chứa chu trình ngắn và cũng không có tập đỉnh độc lập lớn.

Thật không may mắn, một sự lựa chọn  $p$  như vậy là không thể xảy ra: hai dãy  $p$  này không trùng nhau!. Như chúng ta sẽ thấy trong mục “11.4. Hàm ngưỡng”, chúng ta buộc phải giữ  $p$  ở dưới  $n^{-1}$  để đảm bảo  $G$  không có chứa chu trình ngắn nào nhưng, với các giá trị  $p$  như vậy, rất có thể  $G$  sẽ không chứa bất kì chu trình nào. Điều này chỉ ra  $G$  rất có thể là một đồ thị lưỡng phân nên  $G$  có thể có ít nhất  $\frac{n}{2}$  đỉnh độc lập.

Tuy nhiên, mọi thứ chưa hẳn đã kết thúc. Để đảm bảo  $G$  không có chứa tập đỉnh độc lập lớn nào, chúng ta có thể cố định  $p$  ở trên  $n^{-1}$  và vào khoảng  $n^{\varepsilon-1}$  với một số thực dương  $\varepsilon > 0$  nào đó. May mắn thay, nếu  $\varepsilon$  đủ nhỏ thì việc giả sử như vậy chỉ tạo ra một vài chu trình ngắn trong  $G$  thậm chí có thể so sánh được với  $n$ . Nếu chúng ta xóa đi một đỉnh trong mỗi chu trình này thì đồ thị  $H$  thu được sẽ không có chứa chu trình ngắn nào và chỉ số độc lập  $\alpha(H)$  của nó tối đa không thể vượt quá  $\alpha(G)$ . Do  $H$  không nhỏ hơn quá nhiều so với  $G$  nên sắc số  $\chi(H)$  của  $H$  vẫn sẽ khá lớn nên chúng ta có thể tìm thấy một đồ thị  $H$  có cả chỉ số girth  $g(H)$  và sắc số  $\chi(H)$  đủ lớn.

Để chuẩn bị cho việc trình bày phép chứng minh chính thức của định lý Erdős, chúng ta chỉ cần chỉ ra rằng xác suất cạnh  $p = n^{\varepsilon-1}$  đủ lớn để đảm bảo rằng các đồ thị ngẫu nhiên  $G \in \mathcal{G}(n, p)$  “gần như chắc chắn” là không có chứa tập đỉnh độc lập lớn nào. Chính xác hơn, chúng ta sẽ chứng minh một khẳng định mạnh hơn của nó như sau:

**Kết quả 11.2.1.**

Giả sử cho trước một số nguyên dương  $k > 0$  và  $p = p(n)$  là một hàm của  $n$  sao cho  $p \geq (6k \ln n)n^{-1}$  với  $n$  đủ lớn. Khi đó,  $\lim_{n \rightarrow \infty} P\left[\alpha \geq \frac{1}{2}n/k\right] = 0$ .

**Chứng minh kết quả 11.2.1.**

Với mọi số nguyên dương  $n, r$  với  $n \geq r \geq 2$ , với mọi đồ thị ngẫu nhiên  $G \in \mathcal{G}(n, p)$ , Kết quả 11.1.2) chỉ ra rằng

$$\begin{aligned} P[\alpha \geq r] &\leq \binom{n}{r} q^{\binom{r}{2}} \\ &\leq n^r q^{\binom{r}{2}} \\ &= (nq^{(r-1)/2})^r \\ &\leq (ne^{-p(r-1)/2})^r \text{ (điều này suy ra từ } 1 - p \leq e^{-p} \text{ với mọi } p) \end{aligned}$$

Nhận thấy rằng nếu  $p \geq (6k \ln n)n^{-1}$  và  $r \geq \frac{1}{2}n/k$  thì

$$\begin{aligned} ne^{-p(r-1)/2} &= ne^{-pr/2+p/2} \\ &\leq ne^{-(3/2)\ln n+p/2} \\ &\leq nn^{-3/2}e^{1/2} \\ &= \sqrt{e}/\sqrt{n} \rightarrow_{n \rightarrow \infty} 0. \end{aligned}$$

Như vậy, nếu  $p \geq (6k \ln n)n^{-1}$  với  $n$  đủ lớn và  $r := \left\lceil \frac{1}{2}n/k \right\rceil$ ,

$$\lim_{n \rightarrow \infty} P\left[\alpha \geq \frac{1}{2}n/k\right] = \lim_{n \rightarrow \infty} P[\alpha \geq r] = 0. \quad \square$$

**Kết quả 11.2.2. (Erdős)**

Với mọi số nguyên dương  $k > 0$ , luôn tồn tại một đồ thị  $H$  với  $g(H) > k$  và  $\chi(H) > k$ .

**Chứng minh kết quả 11.2.2.**

Giả sử  $k \geq 3$ , chọn ra một số thực dương  $\varepsilon$  cố định sao cho  $0 < \varepsilon < 1/k$  và kí hiệu  $p := n^{\varepsilon-1}$ . Kí hiệu  $X(G)$  là đại lượng ngẫu nhiên đếm số các chu trình ngắn nằm trong một đồ thị ngẫu nhiên  $G \in \mathcal{G}(n, p)$ , nói cách khác,  $X(G)$  đếm số các chu trình có chiều dài  $\leq k$ .

Từ Kết quả 11.1.5),

$$E(X) = \sum_{i=3}^k \frac{\binom{n}{i}}{2i} p^i \leq \frac{1}{2} \sum_{i=3}^k n^i p^i \leq \frac{1}{2} (k-2)n^k p^k,$$

(điều này suy ra từ việc  $np = n^\varepsilon \geq 1$  và  $(np)^i \leq (np)^k$ ).

Từ Kết quả 11.1.4),

$$\begin{aligned} P[X \geq n/2] &\leq E(X)/(n/2) \\ &\leq (k-2)n^{k-1}p^k \\ &= (k-2)n^{k-1}n^{(\varepsilon-1)k} \\ &= (k-2)n^{k\varepsilon-1}. \end{aligned}$$

Từ cách chọn số thực dương  $\varepsilon > 0$ ,  $k\varepsilon - 1 < 0$ , nên

$$\lim_{n \rightarrow \infty} P[X \geq n/2] = 0.$$

Kí hiệu  $n$  là số nguyên dương đủ lớn sao cho  $P[X \geq n/2] < \frac{1}{2}$  và  $P\left[\alpha \geq \frac{1}{2}n/k\right] < \frac{1}{2}$  (điều này suy ra từ cách chọn xác suất cạnh  $p := n^{\varepsilon-1}$  và Kết quả 11.2.1).

Khi đó, tồn tại một đồ thị ngẫu nhiên  $G \in \mathcal{G}(n, p)$  với ít hơn  $n/2$  chu trình ngắn và  $\alpha(G) < \frac{1}{2}n/k$ .

Bằng cách xóa đi một đỉnh trong mỗi chu trình ngắn này, và kí hiệu  $H$  là đồ thị thu được.

Như vậy,  $|H| \geq n/2$  và  $H$  không có chứa bất kì chu trình ngắn nào, nên  $g(H) > k$ .

Từ cách định nghĩa  $G$ ,

$$\chi(H) \geq \frac{|H|}{\alpha(G)} \geq \frac{n/2}{\alpha(G)} \geq k. \quad \square$$

### Kết quả 11.2.3.

Luôn tồn tại các đồ thị với chỉ số girth có giá trị lớn tùy ý và giá trị lớn tùy ý của các đặc trưng bất biến  $\kappa, \varepsilon$  và  $\delta$ .

### Chứng minh kết quả 11.2.3.

Tính đúng đắn của khẳng định này suy ra từ Kết quả 1.4.2) và Kết quả 5.1.3). □

### 11.3. Tính chất hầu hết của các đồ thị ngẫu nhiên.

Giả sử cho trước một hàm cố định  $p = p(n)$  (có thể là một hằng số) và một họ các đồ thị  $\mathcal{P}$  đóng kín qua phép đẳng cấu.

Một vấn đề được đặt ra ở đây đó là:

“Xác định xác suất  $P[G \in \mathcal{P}]$  với  $G \in \mathcal{G}(n, p)$  khi  $n \rightarrow \infty$ ?”.

Nếu xác suất này tiến đến 1 khi  $n \rightarrow \infty$  thì chúng ta nói rằng ***hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có thuộc tính  $\mathcal{P}$*** ; nếu xác suất này tiến đến 0 khi  $n \rightarrow \infty$  thì chúng ta nói rằng ***hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều không có thuộc tính  $\mathcal{P}$*** .

Để minh họa khái niệm mới này, chúng ta sẽ chứng minh rằng, với mọi hằng số thực  $p \in (0,1)$ , mọi đồ thị  $H$  đều là một đồ thị con cảm sinh của hầu hết các đồ thị:

### Kết quả 11.3.1.

Với mọi hằng số thực  $p \in (0,1)$ , với mọi đồ thị  $H$ , hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có chứa một bản sao cảm sinh của  $H$ .

### Chứng minh kết quả 11.3.1.

Giả sử cho trước một đồ thị  $H$  bất kì và kí hiệu  $k := |H|$ . Nhận thấy rằng nếu  $n \geq k$  và  $U \subseteq \{0, \dots, n-1\}$  là một  $k$ -tập nào đó của  $G$  thì biến cố  $G[U]$  đẳng cấu  $H$  sẽ xảy ra với một xác suất  $r > 0$  nào đó. Xác suất  $r$  này chỉ phụ thuộc vào  $p$  chứ không phụ thuộc vào  $n$ .

Nhận thấy rằng do  $G$  chứa một họ gồm  $\lfloor n/k \rfloor$   $k$ -tập  $U$  như trên nên xác suất để không có bất kì đồ thị  $G[U]$  nào trong số chúng đẳng cấu với  $H$  sẽ bằng  $(1-r)^{\lfloor n/k \rfloor}$  (tính độc lập của các biến cố này được suy ra từ tính phân biệt của tập các cạnh  $[U]^2$ ).

Như vậy,  $P[H \not\subseteq_{\text{cảm sinh}} G] \leq (1-r)^{\lfloor n/k \rfloor} \rightarrow_{n \rightarrow \infty} 0$ . □

Giả sử cho trước các số tự nhiên  $i, j \in \mathbb{N}$ , kí hiệu  $\mathcal{P}_{i,j}$  là họ các đồ thị nào đó sao cho, với hai tập đỉnh phân biệt  $U, W$  bất kì với  $|U| \leq i$  và  $|W| \leq j$ , luôn tồn tại một đỉnh  $v \notin U \cup W$  liên hợp với tất cả các đỉnh nằm trong  $U$  nhưng không liên hợp với bất kì đỉnh nào nằm trong  $W$ .

Kết quả sau đây sẽ đóng vai trò như một công cụ đơn giản cho phép chúng ta chứng minh được khá nhiều tính chất thú vị xoay quanh khái niệm “hầu hết các đồ thị”:

### Kết quả 11.3.2.

Với mọi hằng số thực  $p \in (0,1)$ , với mọi số tự nhiên  $i, j \in \mathbb{N}$ , hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có thuộc tính  $\mathcal{P}_{i,j}$ .

### Chứng minh kết quả 11.3.2.

Nhận thấy rằng, với hai tập phân biệt  $U, W$  cố định với  $|U| \leq i$  và  $|W| \leq j$  và  $v \in G - (U \cup W)$ , xác suất để có một đỉnh  $v$  liên hợp với tất cả các đỉnh nằm trong  $U$  nhưng không liên hợp với bất kì đỉnh nào nằm trong  $W$  là  $p^{|U|} q^{|W|} \geq p^i q^j$ .

Như vậy, xác suất để không tồn tại đỉnh  $v$  nào thỏa mãn tính chất như trên cho hai tập đỉnh phân biệt  $U, W$  cố định là  $(1 - p^{|U|} q^{|W|})^{n - |U| - |W|} \leq (1 - p^i q^j)^{n - i - j}$  (với  $n \geq i + j$ ) (điều này suy ra việc các biến cố tương ứng với các đỉnh  $v$  phân biệt là đôi một độc lập).

Nhận thấy rằng có tối đa không quá  $n^{i+j}$  cặp  $(U, W)$  như trên nằm trong  $V(G)$  nên xác suất để không tồn tại đỉnh  $v$  nào thỏa mãn tính chất như trên tối đa không vượt quá  $n^{i+j} (1 - p^i q^j)^{n - i - j}$ .

Như vậy, xác suất này sẽ tiến đến 0 khi  $n \rightarrow \infty$  (điều này suy ra từ việc  $0 < 1 - p^i q^j < 1$ ). □

### Kết quả 11.3.3.

Với mọi hằng số thực  $p \in (0,1)$ , với mọi số tự nhiên  $k \in \mathbb{N}$ , hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều là  $k$  – liên thông.

### Chứng minh kết quả 11.3.3.

Từ Kết quả 11.3.2), để chứng minh hầu như tất cả các đồ thị  $G \in \mathcal{G}(n, p)$  đều là  $k$  – liên thông, chúng ta chỉ cần chứng minh rằng mọi đồ thị nằm trong  $\mathcal{P}_{2,k-1}$  đều là  $k$  – liên thông. Tuy nhiên, khẳng định trên là hiển nhiên: mọi đồ thị nằm trong  $\mathcal{P}_{2,k-1}$  đều có bậc  $\geq k+2$  và nếu  $W$  là một tập với ít hơn  $k$  đỉnh thì hai đỉnh  $x, y$  bất kì đều có một đỉnh láng giềng chung  $v \notin W$  (điều này suy ra từ định nghĩa của  $\mathcal{P}_{2,k-1}$ ), nghĩa là tập  $W$  không tách  $x$  ra khỏi  $y$ .  $\square$

### Kết quả 11.3.4.

Với mọi hằng số thực  $p \in (0,1)$ , với mọi số thực dương  $\varepsilon > 0$ , hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có sắc số  $\chi(G) > \frac{\log(1/q)}{2+\varepsilon} \cdot \frac{n}{\log n}$ .

### Chứng minh kết quả 11.3.4.

Với mọi số nguyên dương  $n, k$  cố định với  $n \geq k \geq 2$ , Kết quả 11.1.2) chỉ ra rằng

$$\begin{aligned} P[\alpha \geq k] &\leq \binom{n}{k} q^{\binom{k}{2}} \\ &\leq n^k q^{\binom{k}{2}} \\ &= q^{k \frac{\log n}{\log q} + \frac{1}{2} k(k-1)} \\ &= q^{\frac{k}{2} \left( -\frac{2 \log n}{\log(1/q)} + k - 1 \right)}. \end{aligned}$$

Bằng cách chọn  $k := (2 + \varepsilon) \frac{\log n}{\log(1/q)}$ , do số mũ của biểu thức trên sẽ tiến đến  $\infty$  khi  $n \rightarrow \infty$  nên biểu thức chính của nó sẽ tiến đến 0 khi  $n \rightarrow \infty$ . Như vậy, hầu như tất cả các đồ thị  $G \in \mathcal{G}(n, p)$ , với mọi phép tô màu đỉnh của  $G$ , đều không có  $k$  đỉnh nào được tô bởi cùng một màu, nghĩa là mọi phép tô màu đỉnh của  $G$  đều sử dụng tối thiểu  $\frac{n}{k} = \frac{\log(1/q)}{2+\varepsilon} \cdot \frac{n}{\log n}$  màu.  $\square$

Năm 1988, nhà toán học Bollobás đã chỉ ra rằng đánh giá thu được ở trên là chặn theo nghĩa nếu chúng ta thay  $\varepsilon$  bởi  $-\varepsilon$  thì chặn dưới nêu trong đánh giá trên sẽ quay đầu thành một chặn trên cho sắc số  $\chi$ .

Một điều khá đặc biệt là hầu hết tất cả các kết quả được trình bày ở đây đều có một điểm chung đó là các giá trị của  $p$  không đóng một vai trò quan trọng nào cả: nếu hầu hết các đồ thị trong  $\mathcal{G}\left(n, \frac{1}{2}\right)$  có tính chất  $\mathcal{P}$  thì hầu hết các đồ thị  $\mathcal{G}\left(n, \frac{1}{1000}\right)$  cũng sẽ có tính chất  $\mathcal{P}$ . Một câu hỏi được đặt ra ở đây đó là: “Nguyên nhân nào khiến cho mô hình đồ thị ngẫu nhiên của chúng ta vô cảm với những thay đổi của xác suất cạnh  $p$ ?”.

Sự vô cảm như vậy của mô hình đồ thị ngẫu nhiên mà chúng ta đã xây dựng đối với những thay đổi của  $p$  chắc chắn là không được trù tính từ trước. Với hầu hết các thuộc tính, bậc tới hạn của  $p$  để xung quanh đó thuộc tính sẽ chỉ xảy ra hoặc không xảy ra nằm thấp hơn nhiều so với bất kì giá trị không đổi nào của  $p$ : nó thường là một hàm của  $n$  tiến đến 0 khi  $n \rightarrow \infty$ . Điều này giải thích cho hiện tượng nêu trên.

Tiếp theo, chúng ta sẽ xem xét xem điều gì sẽ xảy ra nếu  $p$  có thể thay đổi được theo  $n$ . Ngay lập tức, một bức tranh hấp dẫn mở ra cho chúng ta. Đối với xác suất cạnh  $p$  có bậc lớn nằm dưới  $n^{-2}$ , hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  không có chứa cạnh nào cả. Khi xác suất cạnh  $p$  phát triển,  $G$  sẽ ngày càng có nhiều cấu trúc hơn: từ khoảng  $p = \sqrt{n}n^{-2}$  trở đi, nó gần như chắc chắn có chứa một thành phần liên thông với nhiều hơn hai đỉnh, các thành phần liên thông này phát triển thành các cấu trúc cây; xung quanh  $p = n^{-1}$ , các chu trình đầu tiên được sinh ra, ngay sau đó, một trong số các chu trình này có sẽ các dây cung bắt chéo nhau khiến cho đồ thị không còn khả phẳng, đồng thời, một thành phần liên thông lớn hơn những thành phần liên thông khác sẽ nuốt chửng chúng, cho đến khi  $p$  nằm xung quanh  $p = (\log n)n^{-1}$ , để tạo ra một đồ thị liên thông; hầu như sau này, với  $p = (1 + \varepsilon)(\log n)n^{-1}$ , hầu hết các đồ thị đều có chứa một chu trình Hamilton!

#### 11.4. Hàm ngưỡng và phương sai.

Từ bức tranh mà chúng ta phác thảo ở trên, chúng ta sẽ so sánh sự phát triển của các đồ thị ngẫu nhiên khi hàm  $p(n)$  thay đổi với sự tiến của các loài sinh vật: với mỗi hàm  $p = p(n)$ , chúng ta có thể xem xét sự phát triển của các đồ thị  $G \in \mathcal{G}(n, p)$  bằng cách nghiên cứu cách mỗi đồ thị  $G$  thay đổi các đặc điểm của nó với hàm  $p(n)$  đóng vai trò như tốc độ tăng trưởng của một loài sinh vật. Tương tự như sự tiến hóa của các loài sinh vật, sự phát triển của đồ thị ngẫu nhiên xảy ra với những bước nhảy tương đối đột ngột: các xác suất cạnh tới hạn được giới thiệu ở trên đóng vai trò như các ngưỡng mà bên dưới ngưỡng này hầu như không có đồ thị nào thỏa mãn tính chất mong muốn nhưng trên ngưỡng này thì hầu như mọi đồ thị đều có tính chất đang xem xét. Chính xác hơn, chúng ta sẽ gọi một hàm số thực  $t = t(n)$  (với  $t(n) \neq 0$  với mọi  $n$ ) được gọi là một **hàm ngưỡng của thuộc tính  $\mathcal{P}$**  nếu với mọi hàm  $p = p(n)$ , với mọi  $G \in \mathcal{G}(n, p)$ :

$$\lim_{n \rightarrow \infty} P[G \in \mathcal{P}] = \begin{cases} 0, & \text{nếu } \frac{p}{t} \rightarrow 0 \text{ khi } n \rightarrow \infty, \\ 1, & \text{nếu } \frac{p}{t} \rightarrow \infty \text{ khi } n \rightarrow \infty. \end{cases}$$

Nhận thấy rằng nếu  $\mathcal{P}$  có một hàm ngưỡng  $t$  thì mọi hàm  $ct$  ( $c > 0$  là một hằng số thực dương nào đó) cũng là một hàm ngưỡng của  $\mathcal{P}$ , nghĩa là hàm ngưỡng  $t$  tồn tại sai khác nhau đến một hằng số nhân.

Một câu hỏi được đặt ra cho chúng ta đó là: “Một thuộc tính  $\mathcal{P}$  như thế nào thì sẽ có hàm ngưỡng?”.

Năm 1987, nhà toán học Bollobás và Thomason chứng minh được rằng mọi thuộc tính đơn điệu  $\mathcal{P}$  (ngoại trừ một số trường hợp đặc biệt) đều có một hàm ngưỡng.

Trong mục tiếp theo của bài viết, chúng ta sẽ khảo sát một phương pháp tổng quát thường dùng để xác định các hàm ngưỡng của một thuộc tính  $\mathcal{P}$ :

Nhận thấy rằng có vô số các thuộc tính  $\mathcal{P}$  có thể biểu diễn lại dưới dạng  $\mathcal{P} = \{G | X(G) > 0\}$  với  $X \geq 0$  là một đại lượng ngẫu nhiên xác định trên không gian xác suất  $\mathcal{G}(n, p)$ . Chẳng hạn, nếu kí hiệu  $X$  là đại lượng ngẫu nhiên đếm số các cây khung thì  $\mathcal{P}$  chính là thuộc tính liên thông của các đồ thị.

Một vấn đề tự nhiên được đặt ra cho chúng ta đó là: “Làm thế nào để chúng ta có thể chứng minh thuộc tính  $\mathcal{P}$  có một hàm ngưỡng  $t$ ?”. Khi đó, phép chứng minh của chúng ta sẽ bao gồm hai phần: phần đầu chúng ta sẽ chứng minh rằng hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều không có tính chất  $\mathcal{P}$  khi  $p$  nhỏ hơn nhiều so với  $t$ ; phần thứ hai chúng ta sẽ chứng minh hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có tính chất  $\mathcal{P}$  khi  $p$  lớn hơn nhiều so với  $t$ .

Nếu  $X$  nhận giá trị nguyên thì chúng ta có thể sử dụng bất đẳng thức Markov để chứng minh phần đầu và xác định một chặn trên cho  $E(X)$ . Bên cạnh đó, kì vọng dễ tính hơn nhiều so với xác suất: chẳng hạn như chúng ta không cần phải lo lắng về tính độc lập và tính không tương thích của đại lượng ngẫu nhiên này mà chúng ta chỉ cần tính tổng của các giá trị kì vọng riêng lẻ - chẳng hạn như tổng của các chỉ số đặc trưng.

Đối với phần thứ hai của phép chứng minh, mọi thứ trở nên phức tạp hơn rất nhiều so với những gì chúng ta vẫn nghĩ. Để chứng minh  $P[X > 0]$  có giá trị đủ lớn, chúng ta không chỉ cần đảm bảo  $E(X)$  có giá trị đủ lớn mà còn phải đảm bảo  $X$  không bị lệch quá nhiều so với giá trị kì vọng  $E(X)$  của nó quá thường xuyên. Thật vậy, kì vọng  $E(X)$  có thể nhận giá trị rất lớn trong khi  $X$  chỉ nhận giá trị đủ lớn trên một vài đồ thị  $G \in \mathcal{G}(n, p)$  và triệt tiêu với hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  (chẳng hạn, với  $p$  nằm giữa  $n^{-1}$  và  $(\log n)n^{-1}$ , hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có chứa một đỉnh cô lập (nên  $X = 0$ ) trong khi kì vọng  $E(X)$  của nó tiến đến  $\infty$  khi  $n \rightarrow \infty$ ).

Công cụ sơ cấp sau đây trong lý thuyết xác suất sẽ giúp chúng ta đạt được mục đích nêu trên.

Theo thông lệ, chúng ta kí hiệu  $\mu := E(X)$  và xác định  $\sigma^2 \geq 0$  bởi công thức  $\sigma^2 := E((X - \mu)^2)$ .

Đại lượng  $\sigma^2$  được gọi là **phương sai của  $X$** . Từ cách định nghĩa, phương sai  $\sigma^2$  đóng vai trò như số đo độ lệch của  $X$  so với kì vọng  $\mu := E(X)$ . Do  $E$  là một đại lượng tuyến tính nên chúng ta có thể biểu diễn lại phương sai  $\sigma^2$  dưới dạng

$$\sigma^2 = E(X^2 - 2\mu X + \mu^2) = E(X^2) - \mu^2.$$

Kết quả sau đây chính xác là những gì chúng ta cần:  $X$  không lệch quá nhiều so với kì vọng  $\mu := E(X)$  quá thường xuyên.

#### Kết quả 11.4.1. (Chebyshev)

Với mọi số thực dương  $\lambda > 0$ ,  $P[|X - \mu| \geq \lambda] \leq \sigma^2 / \lambda^2$ .

#### Chứng minh kết quả 11.4.1.

Từ Kết quả 11.1.4) và định nghĩa của phương sai  $\sigma^2$ ,

$$P[|X - \mu| \geq \lambda] = P[(X - \mu)^2 \geq \lambda^2] \leq \sigma^2 / \lambda^2. \quad \square$$

Để chứng minh  $X(G) > 0$  với hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$ , chúng ta có thể sử dụng bất đẳng thức Chebyshev theo cách như sau:

#### Kết quả 11.4.2.

Nếu  $\mu > 0$  với mọi  $n$  đủ lớn và  $\sigma^2 / \mu^2 \rightarrow 0$  khi  $n \rightarrow \infty$  thì  $X(G) > 0$  với hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$ .

### Chứng minh kết quả 11.4.2.

Nhận thấy rằng nếu  $G$  là một đồ thị nào đó thỏa mãn  $X(G) = 0$  thì  $|X(G) - \mu| = \mu$ .

Từ Kết quả 11.4.1),

$$P[X = 0] \leq P[|X - \mu| \geq \mu] \leq \sigma^2 / \lambda^2 \rightarrow_{n \rightarrow \infty} 0.$$

Do  $X \geq 0$  nên điều này chỉ ra rằng  $X(G) > 0$  với hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$ . □

Kết quả sau đây là nội dung chính của bài viết. Nó đóng vai trò như công cụ chính yếu giúp chúng ta xác định hàm ngưỡng của một lớp khá rộng các thuộc tính đồ thị khác nhau:

Giả sử cho trước một đồ thị  $H$ , kí hiệu  $\mathcal{P}_H$  là họ các đồ thị có chứa một bản sao của  $H$  như một đồ thị con.

Đồ thị  $H$  được gọi là **cân bằng** nếu  $\varepsilon(H') \leq \varepsilon(H)$  với mọi đồ thị con  $H'$  của  $H$ .

### Kết quả 11.4.3. (Erdős – Rényi)

Nếu  $H$  là một đồ thị cân bằng với  $k$  đỉnh và  $l \geq 1$  cạnh thì  $t(n) := n^{-k/l}$  chính là hàm ngưỡng của thuộc tính  $\mathcal{P}_H$ .

### Chứng minh kết quả 11.4.3.

Kí hiệu  $X(G)$  là đại lượng ngẫu nhiên đếm số các đồ thị con của  $G$  đẳng cấu với  $H$ .

Giả sử cho trước một số tự nhiên  $n \in \mathbb{N}$ , kí hiệu  $\mathcal{H}$  là tập tất cả các đồ thị đẳng cấu với  $H$  với các đỉnh đều nằm trong  $\{0, \dots, n-1\}$  – tập đỉnh của các đồ thị  $G \in \mathcal{G}(n, p)$ :

$$\mathcal{H} := \{H' \mid H' \simeq H, V(H') \subseteq \{0, \dots, n-1\}\}.$$

Giả sử cho trước  $H' \in \mathcal{H}$  và  $G \in \mathcal{G}(n, p)$ , chúng ta sẽ kí hiệu  $H' \subseteq G$  để chỉ bản thân  $H'$  – chứ không phải là một bản sao đẳng cấu của  $H'$  – là một đồ thị con của  $G$ .

Kí hiệu  $h$  là số các bản sao đẳng cấu của  $H$  trên một  $k$ -tập cố định. Hiển nhiên,  $h \leq k!$ . Do có  $\binom{n}{k}$  cách để chọn các tập đỉnh cho các đồ thị nằm trong  $\mathcal{H}$  nên chúng ta sẽ có

$$|\mathcal{H}| = \binom{n}{k} h \leq \binom{n}{k} k! \leq n^k. \quad (1)$$

Giả sử cho trước một hàm cố định  $p = p(n)$ . Kí hiệu  $\gamma := p/t$ . Khi đó,  $p = \gamma n^{-k/l}$ . (2)

Ta sẽ chứng minh rằng nếu  $\gamma \rightarrow 0$  khi  $n \rightarrow \infty$  thì hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều không có thuộc tính  $\mathcal{P}_H$ ; và nếu  $\gamma \rightarrow \infty$  khi  $n \rightarrow \infty$  thì hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có thuộc tính  $\mathcal{P}_H$ .

Để chứng minh phần đầu của khẳng định trên, chúng ta cần xác định một chặn trên đủ tốt cho kì vọng  $E(X)$  – giá trị kì vọng của số các đồ thị con của  $G$  đẳng cấu với  $H$ . Lập luận tương tự như khi chứng minh Kết quả 11.1.5), phương pháp đếm bằng hai cách chỉ ra rằng

$$E(X) = \sum_{H' \in \mathcal{H}} P[H' \subseteq G]. \quad (3)$$

Nhận thấy rằng, với mỗi  $H' \in \mathcal{H}$  cố định,  $P[H' \subseteq G] = p^l$  (điều này suy ra từ việc  $\|H\| = l$ ). (4)

Khi đó,  $E(X) =_{(3),(4)} |\mathcal{H}| p^l \leq_{(1),(2)} n^k (\gamma n^{-k/l})^l = \gamma^l$ . (5)

Như vậy, nếu  $\gamma \rightarrow 0$  khi  $n \rightarrow \infty$  thì  $P[G \in \mathcal{P}_H] = P[X \geq 1] \leq E(X) \leq \gamma^l \rightarrow_{n \rightarrow \infty} 0$  (điều này suy ra từ bất đẳng thức Markov), nên hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều không có thuộc tính  $\mathcal{P}_H$ .

Tiếp theo, chúng ta sẽ chứng minh phần thứ hai của khẳng định trên: nếu  $\gamma \rightarrow \infty$  khi  $n \rightarrow \infty$  thì hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có thuộc tính  $\mathcal{P}_H$ . Ý tưởng cho phép minh này nằm ở Kết quả 11.4.2), chúng ta cần chỉ ra rằng  $\sigma^2/\mu^2 = (E(X^2) - \mu^2)/\mu^2 \rightarrow_{n \rightarrow \infty} 0$ .

Lập luận tương tự như khi chứng minh Kết quả 11.1.5), phương pháp đếm bằng hai cách chỉ ra rằng

$$E(X^2) = \sum_{(H', H'') \in \mathcal{H}^2} P[H' \cup H'' \subseteq G]. \quad (6)$$

Đầu tiên, chúng ta sẽ thực hiện một số tính toán đơn giản để đánh giá giá trị của  $P[H' \cup H'' \subseteq G]$ .

Nhận thấy rằng, với mọi tập  $H', H'' \in \mathcal{H}$  cố định,  $P[H' \cup H'' \subseteq G] = p^{2l - \|H' \cap H''\|}$ .

Do  $H$  là một đồ thị cân bằng nên  $\varepsilon(H' \cap H'') \leq \varepsilon(H) = l/k$ .

Bằng cách kí hiệu  $i := |H' \cap H''|$ , chúng ta suy ra rằng  $\|H' \cap H''\| \leq il/k$ , nên

$$P[H' \cup H'' \subseteq G] \leq p^{2l - il/k}. \quad (7)$$

(điều này suy ra từ việc  $0 \leq p \leq 1$ )

Tiếp theo, chúng ta sẽ ước tính giá trị của tổng hữu hạn nêu trong khẳng định (6). Do khẳng định (7) chỉ ra rằng các số hạng chỉ phụ thuộc vào  $i = |H' \cap H''|$  nên chúng ta sẽ phân hoạch tập  $\mathcal{H}^2$  thành các tập con  $\mathcal{H}_i^2$  xác định bởi công thức  $\mathcal{H}_i^2 := \{(H', H'') \in \mathcal{H}^2 \mid |H' \cap H''| = i\}$ ,  $i = 0, 1, \dots, k$ .

Như vậy, chúng ta cần đánh giá các  $A_i := \sum_{(H', H'') \in \mathcal{H}_i^2} P[H' \cup H'' \subseteq G]$  là đủ.

Đầu tiên, nếu  $i = 0$  thì  $H'$  và  $H''$  là các tập phân biệt, nên  $H' \subseteq G$  và  $H'' \subseteq G$  là các biến cố độc lập.

Như vậy,

$$\begin{aligned} A_0 &= \sum_{(H', H'') \in \mathcal{H}_0^2} P[H' \cup H'' \subseteq G] \\ &\leq \sum_{(H', H'') \in \mathcal{H}_0^2} P[H' \subseteq G] \cdot P[H'' \subseteq G] \\ &= (\sum_{H' \in \mathcal{H}} P[H' \subseteq G]) \cdot (\sum_{H'' \in \mathcal{H}} P[H'' \subseteq G]) \\ &=_{(3)} \mu^2. \end{aligned} \quad (8)$$

Tiếp theo, với  $i \geq 1$ , chúng ta sẽ tiến hành đánh giá các  $A_i = \sum_{(H', H'') \in \mathcal{H}_i^2} P[H' \cup H'' \subseteq G]$ .

Kí hiệu lại  $\Sigma'$  thay cho  $\sum_{H' \in \mathcal{H}}$  và  $\Sigma''$  thay cho  $\sum_{H'' \in \mathcal{H}}$ .

Khi đó, tổng  $\sum_{(H', H'') \in \mathcal{H}_i^2}$  có thể được biểu diễn lại dưới dạng  $\Sigma' \Sigma''|_{|H' \cap H''|=i}$ .

Nhận thấy rằng, với mỗi tập  $H'$  cố định (tương ứng với tổng  $\Sigma'$  thứ nhất), tổng  $\Sigma''$  thứ hai sẽ chạy trên  $\binom{k}{i}\binom{n-k}{k-i}h$  số hạng tương ứng: số các đồ thị  $H'' \in \mathcal{H}$  với  $|H'' \cap H'| = i$ .

Như vậy, với mọi  $i \geq 1$ , với mọi  $\gamma \geq 1$ ,

$$\begin{aligned} A_i &= \sum_{(H', H'') \in \mathcal{H}_i^2} P[H' \cup H'' \subseteq G] \\ &\leq_{(7)} \Sigma' \binom{k}{i} \binom{n-k}{k-i} h p^{2l} p^{-il/k} \\ &=_{(2)} |\mathcal{H}| \binom{k}{i} \binom{n-k}{k-i} h p^{2l} (\gamma n^{-k/l})^{-il/k} \\ &= |\mathcal{H}| p^l \binom{k}{i} \binom{n-k}{k-i} h p^l \gamma^{-il/k} n^i. \end{aligned}$$

Nhận thấy rằng, với  $n \geq k$ ,  $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$ , với  $e$  là hằng số siêu việt.

Như vậy, với mọi  $i \geq 1$ , với mọi  $\gamma \geq 1$ ,

$$\begin{aligned} A_i &\leq |\mathcal{H}| p^l c_1 n^{k-i} h p^l \gamma^{-il/k} n^i \quad (c_1 \text{ là một hằng số nào đó không phụ thuộc } n) \\ &=_{(5)} \mu c_1 n^k h p^l \gamma^{-il/k}. \end{aligned}$$

Nhận thấy rằng, với  $n \geq k$ ,

$$\binom{n}{k} n^{-k} = \frac{1}{k!} \left(\frac{n}{n} \cdots \frac{n-k+1}{n}\right) \geq \frac{1}{k!} \left(\frac{n-k+1}{n}\right)^k \geq \frac{1}{k!} \left(1 - \frac{k-1}{k}\right)^k.$$

Như vậy, với mọi  $i \geq 1$ , với mọi  $\gamma \geq 1$ ,

$$\begin{aligned} A_i &\leq \mu c_2 \binom{n}{k} h p^l \gamma^{-il/k} \quad (c_2 \text{ là một hằng số không phụ thuộc } n) \\ &=_{(1),(5)} \mu^2 c_2 \gamma^{-il/k} \\ &\leq \mu^2 c_2 \gamma^{-l/k}. \end{aligned} \tag{9}$$

Từ định nghĩa của các  $A_i$ ,

$$E(X^2)/\mu^2 =_{(6)} \left(A_0/\mu^2 + \sum_{i=1}^k A_i/\mu^2\right) \leq_{(8),(9)} 1 + c_3 \gamma^{-l/k} \quad (c_3 := k c_2).$$

Như vậy,

$$\sigma^2/\mu^2 = (E(X^2) - \mu^2)/\mu^2 = c_3 \gamma^{-l/k} \rightarrow_{n \rightarrow \infty} 0.$$

(điều này suy ra từ việc  $\gamma \rightarrow \infty$  khi  $n \rightarrow \infty$ )

Từ Kết quả 11.4.2),  $X > 0$  với hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$ , nghĩa là hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có thuộc tính  $\mathcal{P}_H$ .  $\square$

Như đã đề cập, Kết quả 11.4.3) đóng vai trò như công cụ chính yếu giúp chúng ta xác định hàm ngưỡng của một lớp khá rộng các thuộc tính đồ thị khác nhau:

**Kết quả 11.4.4.**

Nếu  $k \geq 3$  thì  $t(n) = n^{-1}$  là hàm ngưỡng đảm bảo hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có chứa một  $k$  – chu trình.

**Kết quả 11.4.5.**

Nếu  $T$  là một cây với bậc  $k \geq 2$  thì  $t(n) = n^{-k/(k-1)}$  là hàm ngưỡng đảm bảo hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có chứa một bảo sao của  $T$ .

**Kết quả 11.4.6.**

Nếu  $k \geq 2$  thì  $t(n) = n^{-2/(k-1)}$  là hàm ngưỡng đảm bảo hầu hết các đồ thị  $G \in \mathcal{G}(n, p)$  đều có chứa một bảo sao của  $K_k$ .