MAC5775 - Métodos Probabilísticos em Combinatória e em Teoria da Computação I
Notas de Aula

Prof. Yoshiharu Kohayakawa
yoshi@ime.usp.br
Aluno: Marcel Kenji de Carli Silva
magal@ime.usp.br

29/6/2005

Sumário

1 Teoria de Ramsey

Notação: $n \to (s,t)$ indica que todo grafo com n vértices contém ou um K_s ou um $\overline{K_t}$.

Defina $R(s,t) := \min\{n : n \to (s,t)\}$.

Ramsey provou que $R(s,t)$ é finito para quaisquer s,t.

Teorema 1.1 (Ramsey [17])**.** *Sejam $s,t \geq 1$. Então $R(s,t)$ é finito. Ademais,*

$$R(s,t) \leq \binom{s+t-2}{s-1} \tag{1.1}$$

Demonstração. A prova é por indução em $s+t$. Começamos notando que $R(s,1) = R(1,t) = 1$ para todo $s,t \geq 1$.

Sejam $s > 1$ e $t > 1$. Vamos mostrar que

$$R(s,t) \leq R(s-1,t) + R(s,t-1). \tag{1.2}$$

Seja $n := R(s-1,t) + R(s,t-1)$ e seja G um grafo qualquer sobre n vértices. Seja v um vértice de G. Particione $V_G \setminus \{v\}$ em $\{A,N\}$, onde A é formado pelos vértices de G adjacentes a v e N é formado pelos vértices de G não-adjacentes a v.

Suponha que $|A| < R(s-1,t)$ e que $|N| < R(s,t-1)$. Então $|V_G \setminus \{v\}| = |A| + |N| \leq R(s-1,t) - 1 + R(s,t-1) - 1 < n-1$, um absurdo. Segue que $|A| \geq R(s-1,t)$ ou $|N| \geq R(s,t-1)$.

Se $|A| \geq R(s-1,t)$, então o subgrafo $G[A]$ contém um K_{s-1} ou um $\overline{K_t}$. Se $G[A]$ contiver um $\overline{K_t}$, estamos feitos. Senão, o conjunto de vértices de um K_{s-1} de $G[A]$, junto com v, é um K_s em G.

Se $|N| \geq R(s,t-1)$, então o subgrafo $G[N]$ contém um K_s ou um $\overline{K_{t-1}}$. Se $G[N]$ contiver um K_s, estamos feitos. Senão, o conjunto de vértices de um $\overline{K_{t-1}}$ de $G[N]$, junto com v, é um $\overline{K_t}$ em G.

Resta provarmos (1.1). Note que (1.1) é justo para $s = t = 1$. Suponha então que $s > 1$ ou $t > 1$. Então, por indução em $s+t$ e por (1.2), temos

$$\begin{aligned} R(s,t) &\leq R(s-1,t) + R(s,t-1) \\ &\leq \binom{s+t-2-1}{s-2} + \binom{s+t-2-1}{s-1} \\ nn &= \binom{s+t-2}{s-1}, \end{aligned}$$

onde utilizamos a famosa identidade binomial

$$\binom{a}{b} = \binom{a-1}{b-1} + \binom{a-1}{b}. \tag{1.3}$$

$\square$

Podemos utilizar a cota superior (1.1) para limitar superiormente os números diagonais $R(s) := R(s,s)$:

$$R(s) \leq \binom{2s-2}{s-1} \leq \frac{2^{2s-2}}{\sqrt{s}} \leq 2^{2s-2}. \tag{1.4}$$

Para provarmos uma cota inferior para os números diagonais $R(s)$, vamos utilizar o método probabilístico:

Teorema 1.2 (Erdős [5])**.** *Seja $s \geq 3$. Então*

$$R(s) \geq \lfloor 2^{s/2} \rfloor. \tag{1.5}$$

Demonstração. Seja $n := \lfloor 2^{s/2} \rfloor$ e considere o grafo G com n vértices construído da seguinte maneira. Comece com $E_G := \emptyset$. Para cada $e \in \binom{V_G}{2}$, coloque e em E_G com probabilidade $1/2$. Note que, dado qualquer grafo G_0 com n vértices, temos $\mathbb{P}[\,G = G_0] = 2^{-\binom{n}{2}}$, ou seja, o grafo G pode ser visto como um grafo sorteado aleatoriamente e uniformemente dentre todos os grafos com n vértices.

Seja $W \in \binom{V_G}{s}$ um conjunto de s vértices de G. Defina a variável aleatória

$$X_W := \begin{cases} 1 & \text{se } G[W] \text{ é um } K_s \text{ ou um } \overline{K_s}, \\ 0 & \text{caso contrário.} \end{cases} \tag{1.6}$$

Defina também $X := \sum \left\{ X_W : W \in \binom{V}{s} \right\}$ como o número de cliques e conjuntos independentes de cardinalidade s.

Utilizando a linearidade da esperança, temos

$$\mathbb{E}\,[X] = \mathbb{E}\left[\sum_{W \in \binom{V}{s}} X_W \right] = \sum_{W \in \binom{V}{s}} \mathbb{E}\,[X_W] = \sum_{W \in \binom{V}{s}} \mathbb{P}[X_W = 1]. \tag{1.7}$$

É fácil ver que a probabilidade de $G[W]$ ser um clique é $2^{-\binom{s}{2}}$. E essa também é a probabilidade de $G[W]$ ser um conjunto independente. Assim,

$$\mathbb{P}[X_W = 1] = 2^{1-\binom{s}{2}}. \tag{1.8}$$

Juntando (1.7) e (1.8), obtemos

$$\mathbb{E}\,[X] = \sum_{W \in \binom{V}{s}} 2^{1-\binom{s}{2}} = \binom{n}{s} 2^{1-s^2/2+s/2} \leq \frac{n^s}{s!} \frac{2^{1+s/2}}{2^{s^2/2}}.$$

Como $n \leq 2^{s/2}$, então

$$\mathbb{E}\,[X] \leq \frac{2^{s^2/2}}{s!} \frac{2^{1+s/2}}{2^{s^2/2}} = \frac{2^{1+s/2}}{s!}. \tag{1.9}$$

É fácil ver, através de (1.9), que $\mathbb{E}[X] < 1$ se $s \geq 3$.

Mas isso significa que, dentre todos os grafos com n vértices, ao menos um deles tem $X < 1$, isto é, ao menos um deles não tem nem um clique de cardinalidade s nem um conjunto independente de cardinalidade s. Segue que $R(s) > \lfloor 2^{s/2} \rfloor$ para $s \geq 3$, como queríamos. $\square$

Na verdade, na prova do teorema 1.2, pode-se ver em (1.9) que $\mathbb{E}[X]$ é muito menor que 1 para s grande. Isso significa que, para s grande, quase todo grafo com $\lfloor 2^{s/2} \rfloor$ vértices satisfaz a propriedade desejada, isto é, não contém nem um clique de cardinalidade s e nem um conjunto independente de cardinalidade s. Assim, um grafo sorteado como G prova que $R(s) > \lfloor 2^{s/2} \rfloor$ com alta probabilidade.

2 Grafos aleatórios

O grafo G construído na prova do teorema (1.2) é um grafo aleatório. Podemos generalizar essa definição. Seja $0 \leq p \leq 1$ um real. Denote por $G(n,p)$ o grafo com $V_G := [n]$ construído da seguinte maneira. Comece com $E_G = \emptyset$. Para cada $e \in \binom{V_G}{2}$, coloque e em E_G com probabilidade p.

Podemos definir também $G(n,M)$, onde $0 \leq M \leq \binom{n}{2}$ é um inteiro. Neste caso, o conjunto de arestas é sorteado uniformemente do conjunto

$$\binom{E(K_n)}{M}.$$

Ou seja, nesse caso o número de arestas é pré-determinado.

Seguem alguns exemplos de resultados sobre grafos aleatórios.

Pode-se estudar vários aspectos de grafos aleatórios, como conexidade, hamiltonicidade, etc. É possível provar o seguinte fenômeno curioso. Seja $(1+\varepsilon)\log n$ o grau médio de um grafo aleatório. Se $\varepsilon > 0$, então o grafo é tipicamente conexo. Mais formalmente, a probabilidade de um grafo aleatório com n vértices ser conexo tende a 1, quando $n \to \infty$. Já se $\varepsilon < 0$, então o grafo é tipicamente desconexo. A formalização é semelhante.

Outro resultado que pode ser provado é o número de grafos não-isomorfos com n vértices. Dado um grafo com n vértices, há no máximo $n!$ grafos isomorfos a ele. Assim, o número de grafos não-isomorfos com n vértices é pelo menos $2^{\binom{n}{2}}/n!$. Na verdade, é possível provar, utilizando argumentos probabilísticos, que o número de grafos não-isomorfos com n vértices é, assintoticamente, igual a esse valor.

3 Números cromático e de independência

Evidentemente temos

$$\chi(G) \leq \Delta(G) + 1 \tag{3.1}$$

para qualquer grafo G. Esse fato nos permite obter uma cota inferior para o número de independência de G.

Lembrando que uma coloração dos vértices de um grafo G pode ser vista como uma partição de V_G em conjuntos independentes, seja $\{X_1, \dots, X_{\chi(G)}\}$ uma coloração mínima de G. Suponha, que $|X_i| < n/\chi(G)$ para todo i, onde $n := |V_G|$. Então $|V_G| = \sum_{i=1}^{\chi(G)} |X_i| < n$, o que é um absurdo. Logo, temos $|X_i| \geq n/\chi(G)$ para algum i. Segue que

$$\alpha(G) \geq n/\chi(G). \tag{3.2}$$

Juntando (3.1) e (3.2), obtemos

$$\alpha(G) \geq \frac{n}{\Delta(G)+1}. \tag{3.3}$$

Essa cota pode ser generalizada utilizando o método probabilístico:

Teorema 3.1 (Caro e Wei). *Seja $G = (V, E)$ um grafo. Então*

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{d(v)+1}. \tag{3.4}$$

Note que essa cota de fato generaliza (3.3), pois

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{d(v)+1} \geq \sum_{v \in V} \frac{1}{\Delta(G)+1} = \frac{n}{\Delta(G)+1}.$$

Demonstração. Seja $<$ uma ordem total aleatória sobre V. Considere o conjunto $I := \{v \in V : vw \in E \Rightarrow v < w\}$. Evidentemente I é um conjunto independente.

Seja $v \in V$ e defina a seguinte variável aleatória:

$$X_v := \begin{cases} 1 & \text{se } v \in I, \\ 0 & \text{caso contrário.} \end{cases}$$

Defina também a variável aleatória $X := \sum_{v \in V} X_v$, de modo que $X = |I|$.

Evidentemente $\alpha(G) \geq |I|$. Temos então

$$\alpha(G) \geq \mathbb{E}[X] = \sum_{v \in V} \mathbb{E}[X_v] = \sum_{v \in V} \mathbb{P}[X_v = 1].$$

É fácil ver que $X_v = 1$ se, e somente se, v é o menor elemento de $\{v\} \cup \Gamma(v)$, segundo a ordem total $<$. Assim, a probabilidade de que $X_v = 1$ é $1/(d(v)+1)$. Mas então

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{d(v)+1},$$

como queríamos. □

Essa cota é justa se o grafo for uma união de cliques.

4 Teorema de Turán

Qual o número máximo de arestas que um grafo pode ter sem que contenha necessariamente um triângulo, isto é, um K_3?

Considere a seguinte construção para um grafo de n vértices. Biparticione V_G em $\{A, B\}$, com $|A| = \lceil n/2 \rceil$ e $|B| = \lfloor n/2 \rfloor$. Agora crie arestas ligando todo vértice de A a todo vértice de B. Em outras palavras, tome $G := K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$. Tal grafo tem $\lceil n/2 \rceil \lfloor n/2 \rfloor \leq n^2/4$ arestas. Como G é bipartido, ele não tem circuitos ímpares, de modo que não contém K_3. Além disso, G é maximal, pois a adição de qualquer aresta a G cria um triângulo. Pode-se provar que essa construção é ótima, isto é, que se $|E_G| > n^2/4$, então G contém um triângulo. Esse é o teorema de Turán.

Essa construção pode ser generalizada da seguinte maneira. Queremos proibir que G contenha um K_{k+1}. Particionamos V_G em $\{I_1, \ldots, I_k\}$, de modo que $|I_i| = \lceil n/k \rceil$ para $1 \leq i \leq n \bmod k$ e $|I_i| = \lfloor n/k \rfloor$ para $n \bmod k < i \leq k$. Cada I_i será um conjunto independente. Criamos então arestas ligando cada vértice de I_i a cada vértice de I_j, para todo $i \neq j$. É fácil ver que não existe nenhum K_{k+1} neste grafo, e que a adição de qualquer aresta forma um K_{k+1}. Essa construção é ótima, no sentido de que, se um grafo H com n vértices possui mais que $|E_G|$ arestas, então H contém um K_{k+1}.

Para enunciar o teorema de Turán, vamos inverter a construção acima. Os conjuntos independentes serão cliques e não existirá arestas entre os cliques. Seja n e $k \leq n$. Sejam q e r tais que $n = kq + r$ e $0 \leq r < k$. Considere o grafo $T_{n,t}$ dado por $\bigcup_{i=1}^{r} K_{q+1} \cup \bigcup_{i=r+1}^{k} K_q$, onde

$$t := r\binom{q+1}{2} + (k-r)\binom{q}{2} \tag{4.1}$$

é o número de arestas de $T_{n,t}$.

Teorema 4.1 (Turán [22])**.** *Sejam $n > 0$ e $k \leq n$ e seja t dado por (4.1). Seja H um grafo com $n(H) = n$ e $|E_H| = t$. Então $\alpha(H) \geq k$. Ademais, se $\alpha(H) = k$, então $H \cong T_{n,t}$.*

Demonstração. Começamos observando que $\alpha(T_{n,t}) = k$, pois $T_{n,t}$ é uma união de k cliques. Assim, a cota (3.4) é justa para $T_{n,t}$.

Seja então H um grafo com n vértices e t arestas. Pelo teorema (3.1) de Caro e Wei, temos

$$\alpha(H) \geq \sum_{v \in V_H} \frac{1}{d(v) + 1}.$$

Se minimizarmos $\sum_{v \in V_H} 1/(d(v)+1)$ sujeito a $\sum_{v \in V_H} d(v) = 2t$, a solução que obtivermos deverá satisfazer

$$|d(v) - d(w)| \leq 1 \tag{4.2}$$

para todos $v, w \in V_H$. De fato, suponha que $d(v) - d(w) \geq 2$. Construa agora uma solução d', idêntica a d em todos os vértices, exceto pelo fato de que $d'(v) = d(v) - 1$ e $d'(w) = d(w) + 1$.

Note que

$$\frac{1}{d(v)+1} + \frac{1}{d(w)+1} = \frac{d(v)+d(w)+2}{(d(v)+1)(d(w)+1)} \tag{4.3}$$

e que

$$\frac{1}{d'(v)+1} + \frac{1}{d'(w)+1} = \frac{1}{d(v)} + \frac{1}{d(w)+2} = \frac{d(v)+d(w)+2}{d(v)(d(w)+2)}. \tag{4.4}$$

Suponha, por contradição, que (4.3) é menor ou igual a (4.4). Como elas têm o mesmo numerador, então $(d(v)+1)(d(w)+1) \geq d(v)(d(w)+2)$. Disso segue que $d(v) - d(w) \leq 1$, o que é um absurdo. Segue a validade de (4.2).

Então o d que minimiza $\sum_{v \in V_H} 1/(d(v)+1)$ sujeito a $\sum_{v \in V_H} d(v) = 2t$ é a função grau de $T_{n,t}$. Segue que

$$\alpha(H) \geq \sum_{v \in V_H} \frac{1}{d_H(v)+1} \geq \sum_{v \in V_{T_{n,t}}} \frac{1}{d_{T_{n,t}}(v)+1} = k.$$

Resta mostrar que, se $\alpha(H) = k$, então $H \cong T_{n,t}$. Suponha então que $\alpha(H) = k$. Então todas as desigualdades que apareceram até agora devem ser igualdades, incluindo as da prova do teorema 3.1 de Caro e Wei. Em particular, a variável aleatória X deve ser constante, isto é, ela independe da ordem total $<$.

Suponha que H não é uma união de cliques. Então existem vértices x, y, z com $xy, xz \in E_H$ e $yz \notin E_H$. Considere as seguintes ordens $<$ e $<'$. Considere x o menor elemento de V_H segundo $<$ e tome $x < y < z < \cdots$, onde $\cdots$ é uma ordem total no restante dos vértices. Considere y o menor elemento de V_H segundo $<'$ e tome $y <' z <' x <' \cdots$, onde $\cdots$ é a mesma ordem no restante dos vértices definida para $<$.

Seja $I_<$ o conjunto I referente à ordem $<$ e seja $I'_<$ o referente à ordem $<'$. Então $x \in I_<$, mas $y, z \notin I_<$, enquanto que $x \notin I'_<$, mas $y, z \in I'_<$. Mas então $|I_<| \neq |I'_<|$, de modo que X não é constante.

Segue que H é uma união de cliques da mesma forma como $T_{n,t}$. □

Exercício 1 (Shearer [19])**.** *Defina $d_i(x)$ como o número de vértices à distância i de x. Por exemplo, $d_1(x) = d(x)$. Seja G um grafo livre de triângulos. Prove que*

$$\alpha(G) \geq \sum_{v \in V_G} \frac{d_1(v)}{1 + d_1(v) + d_2(v)}. \tag{4.5}$$

Exercício 2 (Shearer [19])**.** *Seja G um grafo livre de C_3 e C_5. Prove que*

$$\alpha(G) \geq \sqrt{n\overline{d}/2} = \sqrt{|E_G|}, \tag{4.6}$$

onde $\overline{d} := 2|E_G|/n(G)$ é o grau médio de G.

5 Teoria extremal dos conjuntos

Um exemplo de resultado da teoria extremal dos conjuntos é o seguinte.

Teorema 5.1. *Seja $\mathcal{F} \in \mathcal{P}([n])$ uma família de conjuntos intersectantes, isto é, $F \cap F' \neq \emptyset$ para quaisquer $F, F' \in \mathcal{F}$. Então $|\mathcal{F}| \leq 2^{n-1}$.*

Demonstração. Particione $\mathcal{P}([n])$ em 2^{n-1} pares de conjuntos complementares, isto é, cada parte dessa partição é da forma $\{F, [n] \setminus F\}$ para algum $F \in \mathcal{P}([n])$. É claro que toda família intersectante $\mathcal{F}$ pode ter no máximo um dos elementos de cada parte. Mas então $|\mathcal{F}| \leq 2^{n-1}$. □

Além disso, essa cota é justa. Basta tomar a família

$$\mathcal{F}^* := \{F \cup \{n\} : F \in \mathcal{P}([n-1])\}.$$

5.1 Um teorema de Bollobás

Teorema 5.2 (Bollobás [2])**.** *Sejam $(A_1, B_1), \ldots, (A_m, B_m)$ pares de conjuntos finitos satisfazendo $A_i \cap B_j = \emptyset$ se, e somente se, $i = j$. Então*

$$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \leq 1. \tag{5.1}$$

Em particular, se $|A_i| \leq a$ e $|B_i| \leq b$ para todo i, então

$$m \leq \binom{a+b}{a}. \tag{5.2}$$

Demonstração. Suponha, sem perda de generalidade, que $\bigcup_{i=1}^{m}(A_i \cup B_i) \subseteq [n]$.

Considere uma permutação aleatória σ de $[n]$, isto é, $\sigma = \sigma_1 \cdots \sigma_n$. Dizemos que σ é compatível com o par (A_i, B_i) se todos os elementos de A_i aparecem antes de todos os elementos de B_i.

Seja X a variável aleatória que denota o número de pares (A_i, B_i) compatíveis com σ. Suponha que σ é compatível com os pares distintos (A_i, B_i) e (A_j, B_j). Seja k a posição máxima de um elemento de A_i em σ, isto é, $k := \max\{p : \sigma_p \in A_i\}$. Seja l a posição mínima de um elemento de B_i em σ, isto é, $l := \min\{p : \sigma_p \in B\}$. Como σ é compatível com (A_i, B_i), temos $k < l$. Por outro lado, (A_j, B_j) também é compatível com σ. Como $A_j \cap B_i \neq \emptyset$, então a posição máxima de um elemento de A_j em σ é, no mínimo, l. Como $A_i \cap B_j \neq \emptyset$, então a posição mínima de um elemento de B_j em σ é, no máximo, k. Mas então existe um elemento de B_j que aparece antes de um elemento de A_j em σ, contradizendo com a hipótese de compatibilidade. Concluímos que $X \leq 1$.

Seja X_i a variável aleatória definida da seguinte forma:

$$X_i := \begin{cases} 1 & \text{se } \sigma \text{ é compatível com } (A_i, B_i), \\ 0 & \text{caso contrário.} \end{cases}$$

É claro que $X = \sum_{i=1}^{m} X_i$. Então

$$1 \geq \mathbb{E}\left[X\right] = \sum_{i=1}^{m} \mathbb{E}\left[X_i\right] = \sum_{i=1}^{m} \mathbb{P}[X_i = 1].$$

Precisamos então calcular a probabilidade de que a permutação sorteada seja compatível com (A_i, B_i). Dentre todas as $n!$ permutações, temos $\binom{n}{|A_i|+|B_i|}$ formas de escolher as posições ocupadas por A_i e B_i, lembrando que A_i e B_i são disjuntos. Podemos ordenar os elementos de A_i de $|A_i|!$ formas, e os de B_i de $|B_i|!$ formas. Finalmente, há $(n - |A_i| - |B_i|)!$ formas de se ordenar os elementos de $[n]$ que não estão em A_i ou em B_i. Segue que

$$\mathbb{P}[X_i = 1] = \binom{n}{|A_i| + |B_i|} |A_i|!|B_i|!(n - |A_i| - |B_i|)!/n!. \tag{5.3}$$

Assim, temos

$$\begin{aligned}
1 \geq \mathbb{E}\left[X\right] &= \sum_{i=1}^{m} \binom{n}{|A_i| + |B_i|} |A_i|!|B_i|!(n - |A_i| - |B_i|)!/n! \\
&= \sum_{i=1}^{m} \frac{n!}{(|A_i| + |B_i|)!(n - |A_i| - |B_i|)!} |A_i|!|B_i|!(n - |A_i| - |B_i|)!/n! \\
&= \sum_{i=1}^{m} \frac{|A_i|!|B_i|!}{(|A_i| + |B_i|)!} = \sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1},
\end{aligned}$$

como queríamos.

Para (5.2), começamos notando que

$$\binom{|A_i| + |B_i|}{|A_i|} \leq \binom{|A_i| + b}{|A_i|},$$

já que $|B_i| \leq b$ para todo i. Ademais, ao passarmos de

$$\binom{|A_i| + b}{|A_i|}$$

para

$$\binom{a + b}{a},$$

estamos percorrendo uma diagonal do triângulo de Pascal, cujas entradas são não-decrescentes. Segue que

$$\binom{|A_i| + |B_i|}{|A_i|} \leq \binom{|A_i| + b}{|A_i|} \leq \binom{a + b}{a}, \tag{5.4}$$

de modo que

$$m/\binom{a + b}{a} = \sum_{i=1}^{m} \binom{a + b}{a}^{-1} \leq \sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \leq 1.$$

$\square$

É fácil ver que as cotas (5.1) e (5.2) são justas através da seguinte construção. Seja $n := a + b$ e tome como seus A_i todos os subconjuntos de $[n]$ de cardinalidade a. Para cada i, tome $B_i := [n] \setminus A_i$. Com isso temos $\binom{a+b}{a}$ pares de conjuntos (A_i, B_i). Resta verificarmos que $A_i \cap B_j \neq \emptyset$ sempre que $i \neq j$.

Sejam $i \neq j$. Como $|A_i| = |A_j|$, então existe algum $k \in A_i \setminus A_j$. Mas então $k \in B_j$, de modo que $A_i \cap B_j \neq \emptyset$.

Corolário 5.2.1 (Sperner)**.** *Seja $\mathcal{A} \subseteq \mathcal{P}([n])$ uma família de Sperner, isto é, para quaisquer $A, A' \in \mathcal{A}$ com $A \neq A'$, temos $A \not\subseteq A'$ e $A \not\supseteq A'$. Então*

$$\sum_{A \in \mathcal{A}} \binom{n}{|A|}^{-1} \leq 1. \tag{5.5}$$

Em particular,

$$|\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}. \tag{5.6}$$

Demonstração. Seja $\mathcal{A}$ uma família de Sperner.

Vamos construir $|\mathcal{A}|$ pares (A_i, B_i) de conjuntos que satisfazem as hipóteses do teorema 5.2 de Bollobás. Os A_i's serão os elementos de $\mathcal{A}$. Para cada i, tome $B_i := [n] \setminus A_i$. Evidentemente temos $A_i \cap B_i = \emptyset$ para todo i. Sejam $i \neq j$. Como $\mathcal{A}$ é de Sperner, então existe um $k \in A_i \setminus A_j$. Mas então $k \in B_j = [n] \setminus A_j$, de modo que $A_i \cap B_j \neq \emptyset$.

Aplicando o teorema 5.2 de Bollobás aos pares construídos, obtemos

$$\sum_{i=1}^{m} \binom{n}{|A_i|}^{-1} \leq 1.$$

Ademais, como $\binom{n}{\lfloor n/2 \rfloor} \geq \binom{n}{k}$ para todo k, temos

$$|\mathcal{A}| / \binom{n}{\lfloor n/2 \rfloor} = \sum_{A \in \mathcal{A}} \binom{n}{\lfloor n/2 \rfloor}^{-1} \leq \sum_{A \in \mathcal{A}} \binom{n}{|A|}^{-1} \leq 1,$$

de onde (5.6) segue imediatamente. □

É fácil ver que a cota (5.6) é justa: basta tomar $\mathcal{A} := \binom{[n]}{\lfloor n/2 \rfloor}$.

Suponha que as hipóteses do teorema 5.2 de Bollobás sejam enfraquecidas para: $A_i \cap B_i = \emptyset$ para todo i e $A_i \cap B_j \neq \emptyset$ para todo $i < j$. Neste caso, não se conhecem demonstrações puramente combinatórias para cotas. Todas utilizam métodos de álgebra linear.

Exercício 3 (Tuza [23])**.** *Sejam $(A_1, B_1), \ldots, (A_m, B_m)$ pares de conjuntos finitos satisfazendo $A_i \cap B_i = \emptyset$ para todo i e, para todo $i \neq j$, temos $A_i \cap B_j \neq \emptyset$ ou $A_j \cap B_i \neq \emptyset$. Se $|A_i| \leq a$ e $|B_i| \leq b$ para todo i, prove que*

$$m \leq \frac{(a+b)^{a+b}}{a^a b^b}. \tag{5.7}$$

[*Dica:* Prove que, para qualquer $0 < p < 1$, tem-se

$$\sum_{i=1}^{m} p^{|A_i|}(1-p)^{|B_i|} \leq 1. \tag{5.8}$$

Utilize essa cota para derivar o resultado.]

5.2 Pares disjuntos

Seja $\mathcal{F} \subseteq \mathcal{P}([n])$. Defina

$$d(\mathcal{F}) := \big|\big\{\{F, F'\} : F, F' \in \mathcal{F} \text{ e } F \cap F' = \emptyset\big\}\big|. \tag{5.9}$$

Podemos considerar um grafo que tem $\mathcal{F}$ como conjunto de vértices, onde existe uma aresta entre dois vértices F e F' se $F \cap F' = \emptyset$. Então $d(\mathcal{F})$ é o número de arestas desse grafo.

Daykin e Erdős conjecturam o seguinte. Se $m := |\mathcal{F}| \geq 2^{(1/2+\delta)n}$ para algum $\delta > 0$, então $d(\mathcal{F}) = o(m^2)$.

A relevância da conjectura pode ser melhor compreendida através da seguinte construção. Seja n par e particione $[n]$ em $\{U, W\}$ de modo que $|U| = |W|$. Tome $\mathcal{F} := \mathcal{P}(U) \cup \mathcal{P}(W)$. Observe que

$$|\mathcal{F}| = m = 2 \cdot 2^{n/2} = 2^{(1/2+1/n)n} = 2^{(1/2+o(1))n}$$

e que $d(\mathcal{F}) \geq (m/2)^2 = \Omega(m^2)$.

Essa construção, porém, não é um contra-exemplo para a conjectura. De fato, fixe $\delta > 0$. Existe n_0 tal que $1/n < \delta$ para todo $n \geq n_0$. Assim, para todo $n \geq n_0$, essa construção gera uma família $\mathcal{F}$ de cardinalidade menor que a da hipótese da conjectura.

Seja $f(x)$ uma função sobre os reais. A função $f(x)$ é dita *convexa* se

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{5.10}$$

para qualquer $0 \leq \lambda \leq 1$. Utilizando o seguinte lema poderemos provar uma cota.

Lema 5.3 (Jensen). *Seja f uma função convexa. Então*

$$f\left(\sum_{i=1}^{r} \lambda_i x_i\right) \leq \sum_{i=1}^{r} \lambda_i f(x_i), \tag{5.11}$$

sempre que $\sum_{i=1}^{r} \lambda_i = 1$ e $0 \leq \lambda_i \leq 1$ para todo i.

Demonstração. A prova é por indução em r. Para $r = 2$, basta usar a definição de função convexa. Suponha então que $r > 2$. Podemos supor que $\lambda_i > 0$ para todo i, pois podemos remover os λ_i's que são nulos. Temos

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_r x_r =$$
$$(\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2\right) + \cdots + \lambda_r x_r.$$

Observe que o lado direito tem $r-1$ termos. Aplicando a hipótese de indução, obtemos

$$f\left((\lambda_1+\lambda_2)\left(\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\right)+\cdots+\lambda_r x_r\right)\leq \\ (\lambda_1+\lambda_2)f\left(\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\right)+\cdots+\lambda_r f(x_r).$$

Porém, pela convexidade de $f(x)$, temos

$$f\left(\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\right)\leq\frac{\lambda_1}{\lambda_1+\lambda_2}f(x_1)+\frac{\lambda_2}{\lambda_1+\lambda_2}f(x_2),$$

e estamos feitos. $\square$

Agora vamos provar uma cota.

Teorema 5.4 (Alon e Frankl [1]). *Seja $\mathcal{F}\subseteq\mathcal{P}([n])$, com $m:=|\mathcal{F}|=2^{(1/2+\delta)n}$, onde $0<\delta<\delta_0$ para um certo $\delta_0>0$. Então $d(\mathcal{F})<m^{2-\delta^2/2}$.*

Demonstração. A prova é por contradição. Suponha então que $d(\mathcal{F})\geq m^{2-\delta^2/2}$. Começamos fazendo alguns cálculos. Tome

$$t:=\lceil 1+1/\delta\rceil. \tag{5.12}$$

Temos

$$1-t\delta=1-\lceil 1+1/\delta\rceil\delta\leq 1-(1+1/\delta)\delta=-\delta,$$

de modo que

$$2^{(1-t\delta)n}\leq 2^{-\delta n}. \tag{5.13}$$

Ademais, pode-se provar que, tomando $\delta_0:=1/2$, então $0<\delta<\delta_0$ implica que

$$\left(1-\frac{t\delta^2}{2}\right)\left(\frac{1}{2}+\delta\right)>\frac{1}{2}.$$

Se tomarmos cada um dos lados como expoente de 2^n, obtemos

$$m^{1-t\delta^2/2}>2^{n/2}, \tag{5.14}$$

de onde concluímos que

$$m^{-t\delta^2/2}>2^{n/2}/m=2^{-\delta n}. \tag{5.15}$$

Sejam $A_1,\ldots,A_t\in\mathcal{F}$ sorteados uniforme e independentemente, com reposição.

Primeiro queremos calcular a probabilidade de que $|\bigcup_{i=1}^t A_i|\leq n/2$. Podemos limitar superiormente essa probabilidade da seguinte forma. Se esse evento ocorre, então existe um conjunto $S\in\binom{[n]}{\lfloor n/2\rfloor}$ tal que $A_i\subseteq S$ para todo i. Então

$$\mathbb{P}\Big[\big|\textstyle\bigcup_{i=1}^t A_i\big|\leq n/2\Big]\leq\sum_{S\in\binom{[n]}{\lfloor n/2\rfloor}}\mathbb{P}[A_i\subseteq S,\forall i].$$

Dado $S \in \binom{[n]}{\lfloor n/2 \rfloor}$, é fácil calcular $\mathbb{P}[A_i \subseteq S, \forall i]$. Dentre os m elementos de $\mathcal{F}$, no máximo $2^{\lfloor n/2 \rfloor} \leq 2^{n/2}$ são subconjuntos de S. Ademais, todos os A_i são sorteados independentemente e com reposição, de modo que

$$\mathbb{P}\Big[\big|\textstyle\bigcup_{i=1}^{t} A_i\big| \leq n/2\Big] \leq 2^n \left(\frac{2^{n/2}}{m}\right)^t = 2^{(1-t\delta)n} \leq 2^{-\delta n}, \tag{5.16}$$

onde utilizamos (5.13) na última inequação.

Considere agora o grafo relativo a $\mathcal{F}$, como descrevemos no início da subseção. Dado $B \in \mathcal{F}$, seu grau no grafo é

$$v(B) = |\{A \in \mathcal{F} : A \cap B = \emptyset\}|. \tag{5.17}$$

Lembrando que $d(\mathcal{F})$ é o número de arestas do grafo, temos

$$\sum_{F \in \mathcal{F}} v(F) = 2d(\mathcal{F}) \geq 2m^{2-\delta^2/2}, \tag{5.18}$$

onde utilizamos a hipótese de contradição.

Seja Y a variável aleatória que denota o número de conjuntos $B \in \mathcal{F}$ tais que $B \cap A_i = \emptyset$ para todo i. Ou seja, Y é o número de vizinhos de $A_1, \ldots, A_t$ que estão fora desse conjunto. Para cada $B \in \mathcal{F}$, considere a variável aleatória Y_B como valendo 1 se $B \cap A_i = \emptyset$ para todo i e 0 caso contrário. Obviamente temos $Y = \sum_{B \in \mathcal{F}} Y_B$.

Temos

$$\mathbb{E}\,[Y] = \sum_{B \in \mathcal{F}} \mathbb{E}\,[Y_B] = \sum_{B \in \mathcal{F}} \mathbb{P}[B \cap A_i = \emptyset, \forall i].$$

Fixado B, queremos que todos os t conjuntos A_i escolhidos sejam vizinhos de B. Então existem $v(B)$ escolhas dessa maneira, dentre as m escolhas possíveis:

$$\mathbb{E}\,[Y] = \sum_{B \in \mathcal{F}} \left(\frac{v(B)}{m}\right)^t = m \sum_{B \in \mathcal{F}} \frac{1}{m} \left(\frac{v(B)}{m}\right)^t. \tag{5.19}$$

Note que a função $f(x) := (x/m)^t$ é convexa para $x \geq 0$. Podemos então aplicar a desigualdade (5.11) de Jensen a (5.19), obtendo

$$\mathbb{E}\,[Y] \geq m\left(\sum_{B \in \mathcal{F}} \frac{v(B)}{m^2}\right)^t. \tag{5.20}$$

Mas então, por (5.18),

$$\mathbb{E}\,[Y] \geq m\left(\frac{2d(\mathcal{F})}{m^2}\right)^t \geq m\big(2m^{-\delta^2/2}\big)^t \geq 2m^{1-t\delta^2/2}. \tag{5.21}$$

Como $Y \leq m$, temos

$$\begin{aligned} \mathbb{E}\,[Y] &\leq m\mathbb{P}[Y \geq m^{1-t\delta^2/2}] + m^{1-t\delta^2/2}\mathbb{P}[Y < m^{1-t\delta^2/2}] \\ &\leq m\mathbb{P}[Y \geq m^{1-t\delta^2/2}] + m^{1-t\delta^2/2}. \end{aligned} \tag{5.22}$$

Segue de (5.21), (5.22) e (5.15) que

$$\mathbb{P}[Y \geq m^{1-t\delta^2/2}] \geq m^{-t\delta^2/2} > 2^{-\delta n}. \tag{5.23}$$

Evidentemente temos

$$\mathbb{P}\Big[Y \leq 2^{n/2} \text{ ou } \big|\textstyle\bigcup_{i=1}^t A_i\big| \leq n/2\Big] \leq \mathbb{P}[Y \leq 2^{n/2}] + \mathbb{P}\Big[\big|\textstyle\bigcup_{i=1}^t A_i\big| \leq n/2\Big].$$

Então, por (5.14), (5.23) e (5.16), temos

$$\mathbb{P}\Big[Y \leq 2^{n/2} \text{ ou } \big|\textstyle\bigcup_{i=1}^t A_i\big| \leq n/2\Big] < (1 - 2^{-\delta n}) + 2^{-\delta n} = 1, \tag{5.24}$$

ou seja, existem conjuntos $A_1, \ldots, A_t$ tais que $Y > 2^{n/2}$ e $|\bigcup_{i=1}^t A_i| > n/2$.

Como $|\bigcup_{i=1}^t A_i| > n/2$, então cada $B \in \mathcal{F}$ disjunto de todos os A_i's deve ser subconjunto de $[n] \setminus \big(\bigcup_{i=1}^t A_i\big)$, que tem $k < n/2$ elementos. Assim, o total de conjuntos $B \in \mathcal{F}$ disjuntos de todos os A_i's é no máximo $2^k < 2^{n/2}$, o que contradiz com $Y > 2^{n/2}$. $\square$

6 O método da alteração

O uso mais simples do método probabilístico consiste em mostrar que um dado evento desejado ocorre com probabilidade positiva num certo espaço de probabilidade. Porém, algumas vezes não se consegue mostrar diretamente que o evento desejado ocorre com probabilidade positiva, mas pode-se alterar um objeto, construído probabilisticamente, de modo a garantir que ele satisfaça certas propriedades. Esse é o método da alteração, que exemplificamos nesta seção.

Antes vamos provar a desigualdade de Markov.

Teorema 6.1 (Markov). *Seja X uma variável aleatória não-negativa. Para qualquer $\lambda \in \mathbb{R}_{>0}$, temos*

$$\mathbb{P}[X \geq \lambda] \leq \frac{\mathbb{E}[X]}{\lambda}. \tag{6.1}$$

Demonstração. Pela definição de esperança,

$$\begin{aligned}\mathbb{E}[X] &= \sum_x x\mathbb{P}[X = x] \geq \sum_{x \geq \lambda} x\mathbb{P}[X = x] \\ &\geq \sum_{x \geq \lambda} \lambda\mathbb{P}[X = x] = \lambda \sum_{x \geq \lambda} \mathbb{P}[X = x] \\ &= \lambda\mathbb{P}[X \geq \lambda],\end{aligned}$$

como queríamos. $\square$

6.1 Grafos com cintura e número cromático grandes

Seja G um grafo. Evidentemente temos $\chi(G) \geq \omega(G)$. Uma questão interessante é a seguinte: existe G com $\chi(G) \geq 2005$ e $\omega(G) = 2$? Em outras palavras, existem grafos com número cromático arbitrariamente maiores que o tamanho de um clique máximo? Essa questão foi inicialmente proposta por B. Descartes, Zykov e Mycielski. Resolvê-la é um bom exercício.

Outro parâmetro sobre um grafo G é $g(G)$, a cintura (= *girth*) de G, definida como

$$g(G) := \min\{l : l \geq 3 \text{ e } C_l \subseteq G\}, \tag{6.2}$$

ou seja, a cintura de G é o comprimento do menor circuito de G.

Suponha que um grafo G tenha cintura grande. Seja $k := \lfloor g(G)/2 \rfloor - 1$ e, dado $v \in V_G$, denote por $D_{\leq k}(v)$ o conjunto de vértices de G cuja distância até v é no máximo k. Então $G[D_{\leq k}(v)]$ é um árvore para todo vértice v. Em outras palavras, o grafo G é localmente uma árvore. Assim, o grafo G pode ser bicolorido localmente. Isso torna o seguinte teorema surpreendente.

Teorema 6.2 (Erdős [6]). *Sejam k, l inteiros positivos. Então existe um grafo G com $\chi(G) > k$ e $g(G) > l$.*

Demonstração. Fixe $\theta < 1/l$ e tome $p := n^{\theta-1}$. Vamos considerar o grafo aleatório $G(n,p)$, para n grande.

Seja $3 \leq j \leq l$. Seja $\mathcal{C}^j$ o conjunto de todos os circuitos de comprimento j em $G(n,1) = K_n$. Para cada $C \in \mathcal{C}^j$, defina X_C como a probabilidade de que $C \subseteq G(n,p)$. Evidentemente temos $\mathbb{E}[X_C] = \mathbb{P}[X_C = 1] = p^j$.

Para $3 \leq j \leq l$, defina uma variável aleatória X_j como o número de circuitos de comprimento j em $G(n,p)$. Note que

$$X_j = \sum_{C \in \mathcal{C}^j} X_C,$$

de modo que

$$\mathbb{E}[X_j] = \sum_{C \in \mathcal{C}^j} \mathbb{E}[X_C] = p^j \sum_{C \in \mathcal{C}^j} 1.$$

Existem $n(n-1)\cdots(n-j+1)$ seqüências de vértices distintos de comprimento j, ou seja, seqüências da forma $v_1 v_2 \cdots v_j v_1$. Dessas, $2j$ representam o mesmo circuito, se considerarmos ambas as direções e o vértice inicial da seqüência. Assim, $|\mathcal{C}^j| = n(n-1)\cdots(n-j+1)/2j$. Logo,

$$\mathbb{E}[X_j] = n(n-1)\cdots(n-j+1)p^j/2j \leq (np)^j/2j \leq (np)^j = n^{j\theta}. \tag{6.3}$$

Seja X uma variável aleatória que denota o número de circuitos "pequenos" em $G(n,p)$, isto é, o números de circuitos de comprimento no máximo l. Temos $X = \sum_{j=3}^{l} X_j$. Assim,

$$\mathbb{E}[X] = \sum_{j=3}^{l} \mathbb{E}[X_j] \leq \sum_{j=3}^{l} n^{j\theta} \leq l n^{l\theta} = o(n), \tag{6.4}$$

pois $l\theta < 1$.

Pela desigualdade (6.1) de Markov, temos

$$\mathbb{P}[X \geq n/2] \leq 2\mathbb{E}\,[X]/n = o(1). \tag{6.5}$$

Tome

$$a := \left\lceil \frac{3}{p} \ln n \right\rceil. \tag{6.6}$$

Para cada $A \in \binom{[n]}{a}$, defina

$$Y_A := \begin{cases} 1, & \text{se o subgrafo de } G(n,p) \text{ induzido por } A \text{ é independente,} \\ 0, & \text{caso contrário.} \end{cases}$$

Seja $Y := \sum_{A \in \binom{[n]}{a}} Y_A$ o número de conjuntos independentes de cardinalidade a em $G(n,p)$. Temos

$$\mathbb{E}\,[Y] = \sum_{A \in \binom{[n]}{a}} \mathbb{E}\,[Y_A] = \sum_{A \in \binom{[n]}{a}} \mathbb{P}[Y_A = 1] = \binom{n}{a}(1-p)^{\binom{a}{2}}. \tag{6.7}$$

Utilizando o fato de que, para todo $x \in \mathbb{R}$, temos

$$1 + x \leq e^x, \tag{6.8}$$

obtemos

$$\mathbb{E}\,[Y] \leq n^a \exp\{-pa(a-1)/2\} = \big(n \exp\{-p(a-1)/2\}\big)^a. \tag{6.9}$$

Pode-se provar que, para $n \geq \exp\{2p\}$, temos $-p(a-1)/2 \leq -5/4 \ln n$, de modo que

$$\mathbb{E}\,[Y] \leq \big(n \exp\{-5 \ln n/4\}\big)^a = n^{-a/4} = o(1), \tag{6.10}$$

pois $a > 0$ para todo n.

Segue de (6.10) e da desigualdade (6.1) de Markov que

$$\mathbb{P}[Y \geq 1] = o(1). \tag{6.11}$$

Utilizando (6.5) e (6.10), com n suficientemente grande, seja G um $G(n,p)$ satisfazendo $X < n/2$ e $Y = 0$. Isto é, o número de circuitos "pequenos" de G é estritamente menor que $n/2$ e $\alpha(G) < a$, ou seja,

$$\alpha(G) \leq 3n^{1-\theta} \ln n. \tag{6.12}$$

Remova de G um vértice de cada um dos circuitos "pequenos", obtendo o grafo G'. Temos $g(G') > l$ e o número de vértices de G' é ao menos $n/2$. Como apenas deletamos alguns vértices, o número de independência não pode ter aumentado, de modo que $\alpha(G') \leq \alpha(G)$. Utilizando (3.2), temos

$$\chi(G') \geq \frac{n(G')}{\alpha(G')} \geq \frac{n/2}{3n^{1-\theta} \ln n} = \frac{n^\theta}{6 \ln n}. \tag{6.13}$$

Tomando n suficientemente grande, teremos $\chi(G') > k$, como queríamos. $\square$

Exercício 4. *Sejam l, k, n inteiros satisfazendo $1 \leq l < k \leq n$. Seja $\mathcal{F} \subseteq \binom{[n]}{k}$. A família $\mathcal{F}$ é dita uma* cobertura *de $\binom{[n]}{l}$ se, para todo $L \in \binom{[n]}{l}$, existe um $F \in \mathcal{F}$ tal que $L \subseteq F$.*

Prove que, se $\mathcal{F}$ é uma cobertura de $\binom{[n]}{l}$, então

$$|\mathcal{F}| \geq \binom{n}{l}\binom{k}{l}^{-1}. \tag{6.14}$$

Prove que existe uma cobertura $\mathcal{F} \subseteq \binom{[n]}{k}$ de $\binom{[n]}{l}$ com

$$|\mathcal{F}| \leq \left(1 + \ln \tbinom{k}{l}\right)\binom{n}{l}\binom{k}{l}^{-1}. \tag{6.15}$$

[*Dica:* Defina uma família $\mathcal{F}$ colocando cada $F \in \binom{[n]}{k}$ em $\mathcal{F}$ com probabilidade $x/\binom{n-l}{k-l}$, onde x deve ser escolhido apropriadamente. Sejam $X := |\mathcal{F}|$ e $Y := |\{L \in \binom{[n]}{l} : L$ não é coberto por $\mathcal{F}\}|$. Considere $X+Y$ e otimize a escolha de x.]

7 O método do segundo momento

A variância de uma variável aleatória X é definida como

$$\mathrm{Var}(X) := \mathbb{E}\left[\left(X - \mathbb{E}[X]\right)^2\right]. \tag{7.1}$$

Em geral, utilizamos as abreviaturas $\mu := \mathbb{E}[X]$ e $\sigma^2 := \mathrm{Var}(X)$.

Vamos provar a desigualdade de Chebyshev:

Teorema 7.1 (Chebyshev)**.** *Para todo $\lambda > 0$, temos*

$$\mathbb{P}\big[|X - \mathbb{E}[X]| \geq \lambda\big] \leq \frac{\mathrm{Var}(X)}{\lambda^2}. \tag{7.2}$$

Demonstração. Defina a variável aleatória $Y := (X - \mathbb{E}[X])^2$ e aplique a desigualdade (6.1) de Markov, obtendo

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq \lambda] \leq \mathbb{P}[Y \geq \lambda^2] \leq \mathbb{E}[Y]/\lambda^2 = \mathrm{Var}(X)/\lambda^2.$$

□

Outra forma de escrever a desigualdade de Chebyshev é, para todo $\lambda > 0$,

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq \lambda\sigma] \leq 1/\lambda^2, \tag{7.3}$$

onde a distância de X à sua média é feita em termos do desvio padrão σ.

Não é possível melhorar significativamente a cota da desigualdade (7.2) sem fortalecer a hipótese. De fato, sejam μ, σ, λ reais, com $\lambda > 0$. Considere uma variável aleatória X que vale μ com probabilidade $1 - 1/\lambda^2$, vale $\mu - \sigma\lambda$ com probabilidade $1/(2\lambda^2)$ e vale $\mu + \sigma\lambda$ com essa mesma probabilidade. Essa distribuição mostra que a cota é justa.

Uma situação comum é definir uma variável aleatória como a soma de outras. Suponha que $X = \sum_{i=1}^m X_i$. Temos $\mathbb{E}[X] = \sum_{i=1}^m \mathbb{E}[X_i]$. No caso da variância, temos

$$\mathrm{Var}(X) = \sum_{i=1}^{m} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j), \tag{7.4}$$

onde

$$\mathrm{Cov}(X, Y) := \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \tag{7.5}$$

7.1 Distribuição do número de divisores de um inteiro

Antes de usarmos o método do segundo momento para provarmos alguns resultados sobre a distribuição do número de divisores de um inteiro, vamos enunciar dois fatos sem prova.

Lema 7.2 (Mertens)**.** *Seja $n > 0$ um inteiro. Então*

$$\sum_{\substack{p \leq n \\ p \text{ primo}}} \frac{1}{p} = \ln \ln n + O(1). \tag{7.6}$$

Teorema 7.3 (dos Números Primos)**.**

$$\pi(n) = (1 + o(1)) \frac{n}{\ln n}, \tag{7.7}$$

isto é,

$$\lim_{n \to \infty} \frac{\pi(n)}{n / \ln n} = 1. \tag{7.8}$$

Dado m inteiro, seja $\nu(m)$ o número de divisores primos distintos de m, ignorando as multiplicidades.

Seja $m \leq n$ um inteiro, e suponha que a fatoração de m em primos é $m = p_1^{a_1} \cdots p_k^{a_k}$. Então $\nu(m) = k$. Pode-se mostrar que,

$$1 \leq \nu(m) \leq \big(1 + o(1)\big) \frac{\ln n}{\ln \ln n}. \tag{7.9}$$

Teorema 7.4 (Hardy e Ramanujan, Turán [21])**.** *Seja $\psi(n)$ uma função tal que*

$$\lim_{n \to \infty} \psi(n) = \infty.$$

Então o número de inteiros $m \in [n]$ tais que

$$|\nu(m) - \ln \ln n| > \psi(n) \sqrt{\ln \ln n} \tag{7.10}$$

é $o(n)$.

A prova que apresentamos é devida a Turán.

Demonstração. Considere $[n]$ um espaço de probabilidade uniforme e sorteie um $m \in [n]$. Defina a variável aleatória X como $X := \nu(m)$. Para cada primo $p \leq n$, defina a variável aleatória X_p como

$$X_p := \begin{cases} 1, & \text{se } p \text{ divide } m \\ 0, & \text{caso contrário.} \end{cases} \tag{7.11}$$

Evidentemente temos $X = \sum_{p \leq n} X_p$. Então

$$\mathbb{E}\left[X\right] = \sum_{p \leq n} \mathbb{E}\left[X_p\right] = \sum_{p \leq n} \mathbb{P}[X_p = 1] = \sum_{p \leq n} \frac{\lfloor n/p \rfloor}{n}.$$

Temos

$$\mathbb{E}\left[X\right] \leq \sum_{p \leq n} \frac{n/p}{n} = \sum_{p \leq n} 1/p$$

e

$$\mathbb{E}\left[X\right] \geq \sum_{p \leq n} \frac{n/p - 1}{n} = \sum_{p \leq n} (1/p - 1/n) \geq \Big[\sum_{p \leq n} 1/p\Big] - 1,$$

de modo que, pelo lema 7.2,

$$\mathbb{E}\left[X\right] = \ln \ln n + O(1). \tag{7.12}$$

Agora vamos calcular a variância de X. Temos

$$\operatorname{Var}(X) = \sum_{p \leq n} \operatorname{Var}(X_p) + \sum_{p \neq q} \operatorname{Cov}(X_p, X_q).$$

Primeiro calcularemos $\operatorname{Var}(X_p)$. Note que, como X_p é uma variável aleatória booleana, então $X_p^2 = X_p$.

$$\begin{aligned} \operatorname{Var}(X_p) &= \mathbb{E}\left[X_p^2\right] - \left(\mathbb{E}\left[X_p\right]\right)^2 = \mathbb{E}\left[X_p\right] - \left(\frac{\lfloor n/p \rfloor}{n}\right)^2 \\ &\leq \frac{1}{p} - \left(\frac{n/p - 1}{n}\right)^2 = \frac{1}{p} - \left(\frac{1}{p} - \frac{1}{n}\right)^2 \\ &= \frac{1}{p} - \frac{1}{p^2} + \frac{2}{pn} - \frac{1}{n^2} = \frac{1}{p}\left(1 - \frac{1}{p}\right) + O(1/n). \end{aligned} \tag{7.13}$$

Portanto, pelo lema 7.2,

$$\begin{aligned} \sum_{p \leq n} \operatorname{Var}(X_p) &= \sum_{p \leq n} \left[\frac{1}{p}\left(1 - \frac{1}{p}\right) + O(1/n)\right] \\ &\leq \sum_{p \leq n} \Big[1/p + O(1/n)\Big] = \ln \ln n + O(1). \end{aligned} \tag{7.14}$$

Agora fixe $p \neq q$ primos. Temos $\mathrm{Cov}(X_p, X_q) = \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q]$. A variável aleatória $X_p X_q$ é booleana e vale 1 se, e somente se, ambos p e q dividem m. Como p e q são primos, isso só ocorre se pq divide m.

$$\begin{aligned}
\mathrm{Cov}(X_p, X_q) &= \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q] \\
&= \frac{\lfloor n/(pq) \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n}. \\
&\leq \frac{1}{pq} - \frac{(n/p-1)}{n}\frac{(n/q-1)}{n} \\
&= \frac{1}{pq} - \left(\frac{1}{p} - \frac{1}{n}\right)\left(\frac{1}{q} - \frac{1}{n}\right) \\
&= \frac{1}{pq} - \left[\frac{1}{pq} - \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right) + \frac{1}{n^2}\right] \\
&\leq \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right).
\end{aligned} \tag{7.15}$$

Utilizando repetidas vezes o lema 7.2, obtemos

$$\begin{aligned}
\sum_{\substack{p,q \leq n \\ p \neq q}} \left(\frac{1}{p} + \frac{1}{q}\right) &\leq \sum_{p \leq n}\sum_{q \leq n} \left(\frac{1}{p} + \frac{1}{q}\right) \\
&= \sum_{p \leq n} \left[\frac{1}{p}\sum_{q \leq n} 1 + \sum_{q \leq n} \frac{1}{q}\right] \\
&= \sum_{p \leq n} \left[\frac{1}{p}\pi(n) + \ln\ln n + O(1)\right] \\
&= \left[\pi(n)\sum_{p \leq n} 1/p\right] + \big(\ln\ln n + O(1)\big)\sum_{p \leq n} 1 \\
&= 2\pi(n)\big(\ln\ln n + O(1)\big).
\end{aligned} \tag{7.16}$$

Juntando (7.15), (7.16) e utilizando o teorema 7.3 dos números primos, obtemos

$$\sum_{\substack{p,q \leq n \\ p \neq q}} \mathrm{Cov}(X_p, X_q) \leq \frac{2\pi(n)}{n}\big(\ln\ln n + O(1)\big) = o(1). \tag{7.17}$$

Isso pode ser interpretado informalmente como uma baixíssima dependência entre as variáveis X_p.

Combinando esse último resultado com (7.14), obtemos

$$\mathrm{Var}(X) \leq \ln\ln n + O(1). \tag{7.18}$$

Estamos prontos para utilizar a desigualdade (7.2) de Chebyshev.

$$\begin{aligned} \mathbb{P}\Big[|X - \ln\ln n| \geq \psi(n)\sqrt{\ln\ln n}\Big] & \\ \leq \mathbb{P}\Big[|X - \mathbb{E}\,[X]| \geq \psi(n)\sqrt{\ln\ln n} + O(1)\Big] & \\ \leq \frac{\ln\ln n + O(1)}{\psi^2(n)\ln\ln n + O(1)} = o(1). & \quad (7.19) \end{aligned}$$

(Note que, da primeira para segunda linha, deslocamos o centro do intervalo por uma constante e expandimos o comprimento do intervalo pela mesma constante.)

□

Na verdade, Erdős e Kac provaram em 1940 que, se m é escolhido uniformemente em $[n]$, então para qualquer $\lambda \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\Big[\nu(m) \geq \ln\ln n + \lambda\sqrt{\ln\ln n}\Big] = \frac{1}{2\pi}\int_\lambda^\infty \exp\{-t^2/2\}\,dt,$$

isto é, $\nu(m)$ tem a mesma distribuição de probabilidade de uma normal.

Podemos evitar a referência a um $n \geq m$ pré-fixado, como fazemos na seguinte definição. Seja $\psi(n)$ uma função tal que $\lim_{n\to\infty}\psi(n) = \infty$. Dizemos que um inteiro m é (ν,ψ)-típico se

$$|\nu(m) - \ln\ln m| \leq \psi(m)\sqrt{\ln\ln m}. \qquad (7.20)$$

Porém, isso não faz muita diferença. Para diminuir $\ln\ln n$ de uma unidade, devemos tomar a raiz e-ésima de n, isto é, para que $\ln\ln n' \leq \ln\ln n - 1$, é preciso que $n' \leq n^{1/e}$. Assim, $\ln\ln n$ é próximo ou igual a $\ln\ln m$ para a grande maioria dos inteiros $m \in [n]$. Segue, do teorema 7.4 que, para qualquer função $\psi(n)$ que tenda ao infinito,

$$\mathbb{P}[\, m \text{ é } (\nu,\psi)\text{-típico}\,] = 1 - o(1). \qquad (7.21)$$

Exercício 5. *Evite o uso do teorema 7.3 dos números primos na prova do teorema 7.4.*

[*Dica:* Considere $X' = \sum_{p\leq\sqrt{n}} X_p$ e note que $X' \leq X \leq X' + 1$.]

Seja $\nu'(m)$ o número de divisores primos de m contando multiplicidades. Por exemplo, $\nu(12) = 2$, mas $\nu'(12) = 3$. É fácil ver que

$$1 \leq \nu'(m) \leq \lg n. \qquad (7.22)$$

De fato, suponha que $m = p_1^{a_1}\cdots p_k^{a_k}$. Tome $m' := 2^{a_1+\cdots+a_k}$ e note que $\lg m' = \nu'(m') = \nu'(m)$ e $m' \leq m \leq n$. Mas então $\nu'(m) = \lg m' \leq \lg n$.

Exercício 6. *Seja $\psi(n)$ uma função tal que $\lim_{n\to\infty}\psi(n) = \infty$. Prove que*

$$\lim_{n\to\infty}\frac{1}{n}\Big|\Big\{m \in [n] : |\nu'(m) - \ln\ln n| \geq \psi(n)\sqrt{\ln\ln n}\Big\}\Big| = 0. \qquad (7.23)$$

[*Dica:* Defina X_{p^k} valendo 1 se p^k divide m e 0 caso contrário, para todo primo p e inteiro positivo k. Note que $\nu'(m) = \sum_{p,k} X_{p^k}$. Prove que

$$\sum_{\substack{k\geq 2\\ p \text{ primo}}} \frac{1}{p^k} = \sum_{p \text{ primo}} \frac{1}{p(p-1)} < \infty.$$

Sorteie m uniformemente em $[n]$. Observe que $\mathbb{E}\big[\nu'(m) - \nu(m)\big] = O(1)$ e conclua que $\mathbb{P}\big[\nu'(m) - \nu(m) \geq \omega(n)\big] = o(1)$.]

Defina $d(m)$ como o número de divisores de m. Se $m = p_1^{a_1} \cdots p_k^{a_k}$, então

$$\begin{aligned}
\nu(m) &= k,\\
\nu'(m) &= a_1 + \cdots a_k,\\
d(m) &= (1+a_1)\cdots(1+a_k).
\end{aligned}$$

Podemos utilizar $\nu(m)$ e $\nu'(m)$ para limitar $d(m)$. Considere o conjunto $P := \{p_1, \ldots, p_k\}$. Para qualquer subconjunto $S \subseteq P$, o produto de todos os elementos de S é um divisor de m. Então $2^{\nu(m)} \leq d(m)$. Considere agora o multiconjunto P' formado por a_1 cópias de p_1, …, a_k cópias de p_k. Qualquer divisor de m equivale a pelo menos um subconjunto de P'. Logo, $d(m) \leq 2^{\nu'(m)}$. Temos assim

$$2^{\nu(m)} \leq d(m) \leq 2^{\nu'(m)}. \tag{7.24}$$

Pelo teorema 7.4, temos, tipicamente,

$$(1-\varepsilon)\ln\ln m \leq \nu(m), \nu'(m) \leq (1+\varepsilon)\ln\ln m. \tag{7.25}$$

Mais formalmente, para qualquer $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n}\Big|\big\{m \in [n] : m \text{ satisfaz } (7.25)\big\}\Big| = 1. \tag{7.26}$$

Para tais m, utilizando (7.24),

$$2^{(1-\varepsilon)\ln\ln m} \leq d(m) \leq 2^{(1+\varepsilon)\ln\ln m},$$

isto é,

$$(\ln m)^{(1-\varepsilon)\ln 2} \leq d(m) \leq (\ln m)^{(1+\varepsilon)\ln 2}. \tag{7.27}$$

Assim, $d(m)$ concentra-se tipicamente em torno de $(\ln m)^{\ln 2}$.

Agora calculemos o valor médio de $d(m)$. Considere a variável aleatória X_d definida como

$$X_d := \begin{cases} 1, & \text{se } d \text{ divide } m,\\ 0, & \text{caso contrário.}\end{cases} \tag{7.28}$$

e seja $X := \sum_{d\leq n} X_d$, de modo que $X = d(m)$. Temos

$$\mathbb{E}[X] = \sum_{d\leq n} \mathbb{E}[X_d] = \sum_{d\leq n} \frac{\lfloor n/d \rfloor}{n} = \Big[\sum_{d\leq n} 1/d\Big] + O(1) = \ln n + O(1). \tag{7.29}$$

Assim, a variável aleatória X não é concentrada em torno de sua média.

7.2 Grafos aleatórios e função limiar

De acordo com a desigualdade (6.1) de Markov, se X é uma variável aleatória não-negativa e $\lambda > 0$ é um real, então $\mathbb{P}[X \geq \lambda] \leq \mathbb{E}[X]/\lambda$.

Para nós, freqüentemente temos X uma variável aleatória inteira, como por exemplo em casos em que X é uma variável de contagem. Então

$$\mathbb{P}[X > 0] = \mathbb{P}[X \geq 1] \leq \mathbb{E}[X]. \tag{7.30}$$

Um exemplo típico do uso dessa desigualdade é no estudo de grafos aleatórios, como vemos a seguir.

Sejam A, B conjuntos e $f : A \to B$ uma função. Utilizamos a notação $f : A \hookrightarrow B$ para indicar que f é uma injeção, isto é, uma função injetora.

Sejam G, H grafos. Defina

$$\#\{H \hookrightarrow G\} := \big|\{f \mid f : V_H \hookrightarrow V_G \text{ e } xy \in E_H \Longrightarrow f(x)f(y) \in E_G\}\big|. \tag{7.31}$$

Isto é, $\#\{H \hookrightarrow G\}$ é o número de injeções de V_H em V_G que preservam arestas. O número $\#\{H \hookrightarrow G\}$ também pode ser visto como o número de "cópias rotuladas" de H em G.

Por exemplo, vamos calcular $\#\{P_2 \hookrightarrow C_4\}$, onde P_2 é um caminho de comprimento 2 e C_4 é um circuito de comprimento 4. Suponha que $P_2 = abc$ e $C_4 = uvwx$. Seja $f : \{a, b, c\} \hookrightarrow \{u, v, w, x\}$ uma injeção que preserve arestas. Existem 4 possibilidades para $f(b)$: podemos escolher qualquer um dos vértices de C_4. Fixado $f(b)$, existem duas possibilidades para $f(a)$: qualquer um dos dois vizinhos de $f(b)$. Depois disso, a escolha de $f(c)$ é fixa. Assim, $\#\{P_2 \hookrightarrow C_4\} = 8$.

No que se segue, vamos utilizar a notação $h := |V_H|$ e $l := |E_H|$. Definimos a densidade de H como $d(H) := l/h$. Note que $d(H)$ é metade do grau médio de H.

Seja $p(n)$ uma função positiva satisfazendo $p(n) = o(n^{-1/d(H)})$, ou seja, temos $p(n) = o(n^{-h/l})$. Tome $p := p(n)$ e considere o grafo $G := G(n, p)$, isto é, G é um grafo aleatório.

Defina a variável aleatória $X := \#\{H \hookrightarrow G\}$. Para cada $f : V_H \hookrightarrow V_G$, defina a variável aleatória

$$X_f := \begin{cases} 1, & \text{se } xy \in E_H \Longrightarrow f(x)f(y) \in E_G, \\ 0, & \text{caso contrário.} \end{cases} \tag{7.32}$$

Assim, X_f é uma variável aleatória booleana que vale 1 se, e somente se, f "preserva arestas". Evidentemente temos

$$X = \sum \{X_f \mid f : V_H \hookrightarrow V_G\}. \tag{7.33}$$

Pela linearidade da esperança, temos $\mathbb{E}[X] = \sum_f \mathbb{E}[X_f] = \sum_f \mathbb{P}[X_f = 1]$. Seja $f : V_H \hookrightarrow V_G$. A probabilidade de que G tenha todas as l arestas $f(x)f(y)$ para todo $xy \in E_H$ é p^l, lembrando que cada aresta de $G = G(n, p)$ é sorteada independentemente de todas as outras.

O número de injeções de V_H em V_G pode ser calculado da seguinte maneira. Suponha que $V_H = \{x_1, \ldots, x_l\}$. Existem n possibilidades para $f(x_1)$. Fixado $f(x_1)$, existem $n-1$ possibilidades para $f(x_2)$, pois f é uma injeção, de modo que não pode ocorrer $f(x_1) = f(x_2)$. Prosseguindo dessa maneira, existem $n-i$ possibilidades para x_{i+1}. Então há $(n)_h = n(n-1)\cdots(n-h+1)$ possíveis injeções f, sendo que esse produto tem h termos. Lembramos que $(m)_k := m!/(m-k)! = m(m-1)\cdots(m-k+1)$.

Mas então

$$\mathbb{E}[X] = \sum_f \mathbb{P}[X_f = 1] = (n)_h p^l \sim n^h p^l = o(n^h n^{-h}) = o(1), \tag{7.34}$$

onde a notação $a \sim b$ abrevia o fato de que $a = (1+o(1))b$. Segue de (7.30) que $\mathbb{P}[X > 0] \leq \mathbb{E}[X] = o(1)$. Assim, a probabilidade de que "exista um H em G" tende a zero.

Dizemos que "quase todo $G(n,p)$ não contém H." Alternativamente, dizemos que "quase sempre $G(n,p)$ não contém H." Como isso vale para n grande, dizemos ainda que "assintoticamente quase sempre $G(n,p)$ não contém H." Na literatura em inglês, a expressão "assintoticamente quase sempre" em geral vem abreviada como *a.a.s.* (= *asymptotically almost surely*).

Se p fosse qualquer coisa maior, então $\mathbb{E}[X]$ tenderia ao infinito. Então é natural ver $p = o(n^{-1/d(H)})$ como um "ponto de corte". É natural desejar que, se $p = \Omega(n^{-1/d(H)})$ (que ocorre se, e somente se, $\mathbb{E}[X] \to \infty$), então $\mathbb{P}[X > 0]$ tenda a 1. Porém, isso não é bem verdade. É necessário fazer alguns ajustes e utilizar a variância de X para provar a afirmação ajustada. É o que fazemos a seguir.

Considere um K_4 com conjunto de vértices $\{1,2,3,4\}$. Tome como H um grafo com $V_H := \{1,2,3,4,5\}$ dado por K_4 mais uma aresta 15. Note que $h = 5$ e $l = 7$, de modo que $d(H) = 7/5$, e tome $p_0(n) := n^{-5/7}$.

Seja $\psi(n)$ uma função tal que $\lim_{n\to\infty} \psi(n) = \infty$ e tome $p := p(n) := \psi(n) p_0(n)$. Observe que $p(n) = \omega(n^{-5/7})$.

Defina a variável aleatória X como acima e note que

$$\mathbb{E}[X] = (n)_5 p^7 \sim n^5 p^7 = (\psi(n))^7 \to \infty. \tag{7.35}$$

Defina agora a variável aleatória Y como $Y := \#\{K_4 \hookrightarrow G\}$ e observe que

$$\mathbb{E}[Y] = (n)_4 p^6 \sim n^4 p^6 = n^4 (\psi(n)(n^{-5/7}))^6 = (\psi(n))^6 n^{-2/7} = o(1), \tag{7.36}$$

se $(\psi(n))^6 = o(n^{2/7})$, isto é, se $\psi(n) = o(n^{1/21})$.

Portanto, por (7.30), temos $\mathbb{P}[Y > 0] \leq \mathbb{E}[Y] = o(1)$, de onde concluímos que $\mathbb{P}[X > 0] = o(1)$, já que $K_4 \subset H$.

Essa anomalia ocorre porque o grafo H tem regiões com alta variação na densidade: o clique $K_4 \subset H$ tem densidade alta, mas nem todo subgrafo de H tem densidade tão alta assim: basta considerar subgrafos contendo o vértice 5.

Prosseguimos agora para fazer um ajuste.

Seja H um grafo com pelo menos uma aresta. Defina

$$m(H) := \max\{d(H') : H' \subseteq H \text{ e } V(H') \neq \emptyset\}. \tag{7.37}$$

Chamamos o valor $m(H)$ de 0-densidade de H. Como $m(H)$ leva em conta a densidade de subgrafos, é razoável pensar que $p_0(n) := n^{-1/m(H)}$ seja um "ponto de corte".

Claramente temos $m(H) \geq d(H)$. Então $1/m(H) \leq 1/d(H)$, de modo que $-1/m(H) \geq -1/d(H)$. Assim,

$$n^{-1/m(H)} \geq n^{-1/d(H)}. \tag{7.38}$$

Observe que quase todo $G(n,p)$ com $p := p(n) := o(n^{-1/m(H)})$ não contém H. De fato, Seja $H^* \subseteq H$ com $V(H^*) \neq \emptyset$ tal que $m(H) = d(H^*)$. Defina a variável aleatória Z como $\#\{H^* \hookrightarrow G\}$. Provamos acima que $\mathbb{P}[Z > 0] = o(1)$. Mas então quase todo $G(n,p)$ não contém H^*. Mas se um grafo não contém H^*, não pode conter H. Então quase todo $G(n,p)$, com p definido dessa forma, não contém H.

Agora consideramos o caso em que $p = \psi(n)n^{-1/m(H)}$, onde $\psi(n)$ é uma função tal que $\lim_{n\to\infty} \psi(n) = \infty$.

Defina a variável aleatória $X := \#\{H \hookrightarrow G\}$ e as variáveis X_f, para cada $f : V_H \hookrightarrow V_G$, como em (7.32). Lembre-se que $G = G(n,p)$.

Seja $\mu := \mathbb{E}[X]$. Se tivéssemos tomado p como $\psi(n)n^{-1/d(H)}$, teríamos $\mu = \mathbb{E}[X] \to \infty$, como já provamos acima. Então, por (7.38), temos

$$\mu = \mathbb{E}[X] \to \infty. \tag{7.39}$$

Seja $\sigma^2 := \operatorname{Var}(X)$. Pela desigualdade (7.2) de Chebyshev, temos

$$\mathbb{P}[X = 0] \leq \mathbb{P}\big[|X - \mu| \geq \mu\big] \leq \frac{\sigma^2}{\mu^2}. \tag{7.40}$$

Se mostrarmos que $\sigma^2 = o(\mu^2)$, teremos então $\mathbb{P}[X = 0] = o(1)$, de modo que $X > 0$ quase sempre, isto é, quase sempre $G(n,p)$ tem uma cópia de H. Prosseguimos então para provar que $\operatorname{Var}(X) = o(\mathbb{E}[X]^2)$.

Temos

$$\operatorname{Var}(X) = \sum_f \operatorname{Var}(X_f) + \sum_{f\neq g} \operatorname{Cov}(X_f, X_g).$$

Como X_f é uma variável booleana, então $X_f^2 = X_f$, de modo que $\operatorname{Var}(X_f) = \mathbb{E}[X_f^2] - (\mathbb{E}[X_f])^2 \leq \mathbb{E}[X_f]$. Mas então

$$\operatorname{Var}(X) \leq \sum_f \mathbb{E}[X_f] + \sum_{f\neq g} \operatorname{Cov}(X_f, X_g) \leq \mathbb{E}[X] + \sum_{f\neq g} \operatorname{Cov}(X_f, X_g). \tag{7.41}$$

Precisamos calcular então $\sum_{f\neq g} \operatorname{Cov}(X_f, X_g)$. Fixe $f \neq g$ injeções de V_H em V_G. Temos

$$\operatorname{Cov}(X_f, X_g) = \mathbb{E}[X_f X_g] - \mathbb{E}[X_f]\mathbb{E}[X_g] = \mathbb{P}[X_f X_g = 1] - p^{2l}. \tag{7.42}$$

Seja $f(E_H) := \{f(x)f(y) : xy \in E_H\}$ e defina $g(E_H)$ analogamente. Temos que $X_f X_g = 1$ se, e somente se, f e g preservam arestas, ou, equivalentemente, se $f(E_H)$ e $g(E_H)$ são arestas em $G(n,p)$. Equivalentemente, $X_f X_g = 1$ se, e só se, $f(E_H) \cup g(E_H)$ são arestas em $G(n,p)$. Então $\mathbb{P}[X_f X_g = 1] = p^{|f(E_H) \cup g(E_H)|}$. Porém, $|f(E_H)| = |g(E_H)| = 2l$, pois f e g são injeções. Como

$$\begin{aligned} &|f(E_H) \cup g(E_H)| = \\ &\qquad |f(E_H)| + |g(E_H)| - |f(E_H) \cap g(E_H)| = 2l - |f(E_H) \cap g(E_H)|, \end{aligned} \tag{7.43}$$

obtemos

$$\mathrm{Cov}(X_f, X_g) = p^{2l - |f(E_H) \cap g(E_H)|} - p^{2l}. \tag{7.44}$$

Se $f(E_H) \cap g(E_H) = \emptyset$, então $\mathrm{Cov}(X_f, X_g) = 0$. Portanto, basta considerarmos injeções $f \neq g$ com $f(E_H) \cap g(E_H) \neq \emptyset$. Defina o conjunto de arestas

$$F_{f,g} := f^{-1}\big(f(E_H) \cap g(E_H)\big), \tag{7.45}$$

ou seja, $F_{f,g}$ é o conjunto de arestas $f(E_H) \cap g(E_H)$ "trazido de volta para H". Como estamos supondo $f(E_H) \cap g(E_H) \neq \emptyset$, então $F_{f,g} \neq \emptyset$.

Defina $H'_{f,g} := H[F_{f,g}]$, isto é, H' é o subgrafo de H induzido pelas arestas de $F_{f,g}$. Temos $E(H'_{f,g}) = F_{f,g} \neq \emptyset$.

Estamos interessados em pares (f,g) de injeções tais que $f \neq g$ e $F_{f,g} \neq \emptyset$. Podemos particionar a coleção desses pares da seguinte maneira:

$$\begin{aligned} &\big\{(f,g) : f \neq g \text{ e } F_{f,g} \neq \emptyset\big\} = \\ &\qquad \bigcup_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} \big\{(f,g) : f \neq g, F_{f,g} \neq \emptyset \text{ e } H'_{f,g} = H'\big\}. \end{aligned} \tag{7.46}$$

Observe que todo par (f,g) com $f \neq g$ e $F_{f,g} \neq \emptyset$ está em alguma parte dessa partição, pois $F_{f,g}$ é o conjunto de arestas de algum subgrafo de H. Ademais, cada par (f,g) como acima está em uma única parte, pois há exatamente um subgrafo de H com conjunto de arestas $F_{f,g}$ sem vértices isolados.

Como só estamos considerando subgrafos $H' \subseteq H$ com pelo menos uma aresta e $E(H') = E(H'_{f,g}) = F_{f,g}$, a condição $F_{f,g} \neq \emptyset$ pode ser removida:

$$\big\{(f,g) : f \neq g \text{ e } F_{f,g} \neq \emptyset\big\} = \bigcup_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} \big\{(f,g) : f \neq g \text{ e } H'_{f,g} = H'\big\}. \tag{7.47}$$

Queremos agora calcular o tamanho de cada parte. Note que cada parte é totalmente determinada por $H' \subseteq H$. Fixe $H' \subseteq H$ com $E(H') \neq \emptyset$.

Quantos pares (f,g) existem com $H'_{f,g} = H'$? Vamos delimitar superiormente o número de pares (f,g) tais que $H'_{f,g} = H'$. Utilizaremos a notação $h' := |V(H')|$ e $l' = |E(H')|$.

Escolha f de qualquer maneira. Há $(n)_h$ maneiras de se fazer isso. Fixe um tal f.

Queremos contar o número de $g : V_H \hookrightarrow V_G$ tais que $E(H') = F_{f,g} = f^{-1}\big(f(E_H) \cap g(E_H)\big)$ ou, equivalentemente, o número de injeções g tais que $f\big(E(H')\big) = f(E_H) \cap g(E_H)$. Evidentemente f é inversível se restringirmos seu contradomínio à sua imagem. Então f é uma bijeção entre $V(H)$ e $V\big(G[f(V_H)]\big)$. Em particular, f induz uma bijeção entre E_H e $f(E_H)$. Como $f(E_H) \cap g(E_H) \subseteq f(E_H)$, podemos trabalhar sobre o grafo G ao invés de sobre o grafo H. Isto é, basta considerarmos $f\big(E(H')\big)$ como um conjunto de l' arestas de G, sendo que o subgrafo induzido $G' := G\big[f\big(E(H')\big)\big]$ tem h' vértices. Em outras palavras, podemos pensar que o que está fixo não é H', um subgrafo de H, mas sim um subgrafo de G, com mesma estrutura de H'.

Existem $\binom{h}{h'}$ formas de se escolher quais vértices de H serão levados por g a algum vértice de G'. Fixados tais vértices, g pode ordenar estes h' vértices de no máximo $(h')!$ modos diferentes. É claro que nem toda ordenação "funciona", ou seja, pode ser que, para alguma ordenação, g não preserve as arestas. Porém, isso não nos interessa, já que estamos interessados apenas numa delimitação superior, mesmo que "crua", mas que seja o suficiente para provarmos cotas interessantes.

Os vértices restantes de H podem ser levados de qualquer maneira para G: há $(n)_{h-h'}$ maneiras de se fazer isso.

Assim, o número de pares (f, g) com $H'_{f,g} = H'$ é, no máximo,

$$(n)_h (n)_{h-h'} \binom{h}{h'} (h')! \leq n^{2h-h'} (h)_{h'} \leq n^{2h-h'} h^h. \tag{7.48}$$

Vamos utilizar a partição de pares (f, g) com $f \neq g$ e $F_{f,g} \neq \emptyset$ para separar o somatório a seguir:

$$\begin{aligned}
\sum_{f \neq g} \operatorname{Cov}(X_f, X_g) &= \sum_{\substack{f \neq g \\ F_{f,g} \neq \emptyset}} \operatorname{Cov}(X_f, X_g) = \sum_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} \sum_{\substack{f \neq g \\ H'_{f,g} = H'}} p^{2l-l'} - p^{2l} \\
&\leq \sum_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} n^{2h-h'} h^h p^{2l-l'} = h^h n^{2h} p^{2l} \sum_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} \frac{1}{n^{h'} p^{l'}}
\end{aligned} \tag{7.49}$$

Porém, vimos em (7.34) que $\mu = \mathbb{E}[X] \sim n^h p^l$. Então

$$\begin{aligned}
\sum_{f \neq g} \operatorname{Cov}(X_f, X_g) &\leq h^h n^{2h} p^{2l} \sum_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} \left(\frac{1}{n^{h'/l'} p} \right)^{l'} \\
&\leq h^h \mu^2 \sum_{\substack{H' \subseteq H \\ E(H') \neq \emptyset}} \left(\frac{1}{n^{1/m(H)} p} \right)^{l'},
\end{aligned} \tag{7.50}$$

onde utilizamos o fato de que $1/m(H)$ é o valor mínimo de h'/l' para todo subgrafo H' de H.

Como $p = \psi(n)n^{-1/m(H)}$, então

$$\lim_{n\to\infty} \left(\frac{1}{n^{1/m(H)}p}\right)^{l'} = \left(\frac{1}{\psi(n)}\right)^{l'} = 0, \tag{7.51}$$

já que $l' \geq 1$.

Como a somatória em (7.50) tem um tem um número finito de termos, concluímos que $\sum_{f\neq g} \operatorname{Cov}(X_f, X_g) = o(\mu^2)$. Retomando (7.41), obtemos

$$\sigma^2 = \operatorname{Var}(X) = \mathbb{E}\,[X] + \sum_{f\neq g} \operatorname{Cov}(X_f, X_g) = \mu + o(\mu^2) = o(\mu^2), \tag{7.52}$$

pois $\mu = \mathbb{E}\,[X] \to \infty$.

Concluímos então, como mostramos acima utilizando a desigualdade de Chebyshev, que quase sempre $X > 0$, isto é, quase sempre existe H em G.

É um bom exercício provar, utilizando Chebyshev, que, para todo $\varepsilon > 0$, temos

$$\lim_{n\to\infty} \mathbb{P}\Big[\big|X - \mathbb{E}\,[X]\big| \geq \varepsilon\mathbb{E}\,[X]\Big] = 0. \tag{7.53}$$

Enunciamos a seguir o resultado provado acima:

Teorema 7.5 (Erdős-Rényi, Bollobás)**.** *Seja H um grafo com pelo menos uma aresta. Então $p_0 := p_0(n) := n^{-1/m(H)}$ é uma* função limiar *para o evento $H \hookrightarrow G(n,p)$ no sentido que*

1. *se $p = o(p_0)$, então quase sempre $G(n,p)$ não contém H;*
2. *se $p = \psi(n)p_0(n)$, onde $\psi(n)$ é uma função tal que $\lim_{n\to\infty} \psi(n) = \infty$, então quase sempre $G(n,p)$ contém H.*

Exercício 7. *Defina $\mathcal{F} \subseteq \mathcal{P}([n])$ pondo $\mathbb{P}[A \in \mathcal{F}] = p$ para qualquer $A \subseteq [n]$, com esses eventos todos independentes.*

Mostre que $p_0 := p_0(n) = 3^{-n/2}$ é uma função limiar no sentido do teorema 7.5 para o evento "$\mathcal{F}$ é de Sperner".

[*Dica:* estude a existência de um par (A, B) com $A, B \in \mathcal{F}$ e $A \subseteq B$.]

8 A distribuição binomial e algumas cotas

8.1 Um problema geométrico extremal

Começamos com um problema geométrico extremal que pode ser atacado com métodos probabilísticos utilizando a distribuição binomial.

Defina a função $f : \{n : n \in \mathbb{N} \text{ e } n \geq 2\} \to \mathbb{N}$ da seguinte maneira:

$$\begin{split} f(d) := \max\Big\{|S| : S \subseteq \mathbb{R}^d, S \text{ não tem pontos colineares e} \\ \text{cada } T \in \tbinom{S}{3} \text{ determina um triângulo agudo}\Big\}, \end{split} \tag{8.1}$$

sendo que um triângulo é agudo se todos os seus ângulos são agudos.

Para nos familiarizar melhor com o problema, vamos mostrar que $f(2) = 3$. É evidente que $f(2) \geq 3$, pois basta tomar como S o conjunto de vértices de um triângulo equilátero qualquer no plano $\mathbb{R}^2$. Vamos mostrar agora que, se S tem 4 pontos no plano, então existe um subconjunto de S que determina um triângulo que não é agudo.

Sejam $A, B, C, D \in S$. Considere o triângulo T determinado por A, B e C. Suponha primeiro que D está dentro de T. Evidentemente a soma dos ângulos $\widehat{ADB}$, $\widehat{ADC}$ e $\widehat{BDC}$ é 2π. Mas então pelo menos um deles, é no mínimo $2\pi/3 > \pi/2$. Então um dos triângulos formados por D e dois vértices de T é obtuso. Suponha agora que D está fora de T, de modo que $\{A, B, C, D\}$ determina um quadrilátero. Como a soma dos ângulos internos de um equilátero é 2π, pelo menos um dos quatro ângulos internos é no mínimo $\pi/2$. Mas então existe um subconjunto de $\{A, B, C, D\}$ que determina um triângulo que não é agudo. Segue que $f(2) < 4$. Fica provado então que

$$f(2) = 3. \tag{8.2}$$

Croft provou que

$$f(3) = 5 \tag{8.3}$$

e Danzer e Grünbaum conjecturaram que

$$f(d) \leq 2d - 1. \tag{8.4}$$

Porém, Erdős e Füredi [7] provaram que existe um d_0 tal que

$$f(d) \geq (1{,}15)^d \tag{8.5}$$

para todo $d \geq d_0$.

Observe que, se enfraquecermos a exigência de que todos os triângulos sejam agudos, para a de que todos sejam não-obtusos, isto é, se permitirmos ângulos retos, então $f(d) \geq 2^d$ sempre: basta considerar os vértices de um hipercubo.

Exercício 8. *Prove que*

$$f(d) \geq \exp\{\alpha d\} \tag{8.6}$$

para algum $\alpha > 0$.

[*Dica:* Tome $\underline{x}_1, \ldots, \underline{x}_n \in \{\pm 1\}^d \subseteq \mathbb{R}^d$ com cada coordenada de $\underline{x}_i$ valendo ± 1 aleatoriamente, todos independentes. Considere $\langle \underline{x}_i, \underline{x}_j \rangle$.]

8.2 A distribuição binomial

Sejam $X_1, \ldots, X_n$ variáveis aleatórias de Bernoulli, mutuamente independentes, com $\mathbb{E}[X_i] = p$ para todo i. Defina

$$X := \sum_{i=1}^{n} X_i \tag{8.7}$$

como o número de variáveis X_i com $X_i = 1$, isto é, o número de sucessos na repetição de n experimentos aleatórios. Dizemos que X segue a distribuição binomial e denotamos isso por

$$X \sim \mathrm{Bi}(n, p). \tag{8.8}$$

Temos

$$\mu := \mathbb{E}\,[X] = \sum_{i=1}^{n} \mathbb{E}\,[X_i] = \sum_{i=1}^{n} p = np. \tag{8.9}$$

Para calcular a variância de X, comece notando que $\mathbb{E}\,[X_i X_j] = \mathbb{E}\,[X_i]\mathbb{E}\,[X_j]$ para todo $i \neq j$, pois $X_1, \ldots, X_n$ são mutuamente independentes. Mas então $\mathrm{Cov}(X_i, X_j) = 0$ para todo $i \neq j$. Assim,

$$\begin{aligned}\sigma^2 := \mathrm{Var}(X) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} \Big(\mathbb{E}\,[X_i^2] - \big(\mathbb{E}\,[X_i]\big)^2\Big) = \sum_{i=1}^{n} \big(p - p^2\big) \\ &= np(1-p).\end{aligned} \tag{8.10}$$

Observe que, para $0 \leq k \leq n$, temos

$$\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}, \tag{8.11}$$

pois há $\binom{n}{k}$ formas de selecionar k dentre as variáveis $X_1, \ldots, X_n$ que terão valor 1, sendo que as restantes $n-k$ devem ter valor 0.

Considere a variável aleatória

$$Y := \frac{X - \mu}{\sigma}, \tag{8.12}$$

isto é, Y tem média 0 e contamos a distância de Y até sua média em passos do desvio padrão σ. Pode-se provar que, quanto n tende ao infinito a distribuição de Y tende à distribuição normal com média 0 e desvio padrão 1, isto é,

$$\lim_{n\to\infty} \mathbb{P}[Y \leq \lambda] = \frac{1}{\sqrt{2\pi n}} \int_{-\infty}^{\lambda} \exp\{-t^2/2\}\, dt. \tag{8.13}$$

8.3 Cotas para a distribuição binomial

A seguir vamos derivar diversas cotas envolvendo a distribuição binomial.

Aplicando uma versão da desigualdade (7.3) de Chebyshev a uma variável aleatória $X \sim \mathrm{Bi}(n, p)$, obtemos

$$\mathbb{P}\Big[\big|X - np\big| \geq \lambda\sqrt{np(1-p)}\Big] \leq \frac{1}{\lambda^2} \tag{8.14}$$

para todo $\lambda > 0$. Nessa versão da desigualdade de Chebyshev, estamos contando a distância de X até sua média em passos do desvio padrão. Podemos também contar a diferença em passos da média, utilizando a desigualdade (7.2) de Chebyshev: para todo $\varepsilon > 0$, temos

$$\mathbb{P}\Big[\big|X - np\,\big| \geq \varepsilon\mu\Big] \leq \frac{\sigma^2}{\varepsilon^2\mu^2} = \frac{np(1-p)}{\varepsilon^2 n^2 p^2} = \frac{1-p}{\varepsilon^2 np} \leq \frac{1}{\varepsilon^2\mu}. \tag{8.15}$$

Observe então que, se $\lim_{n\to\infty}\mu = \lim_{n\to\infty} np = \infty$, então

$$\mathbb{P}\Big[\big|X - np\,\big| \geq \varepsilon\mu,\Big] = o(1) \tag{8.16}$$

para qualquer $\varepsilon > 0$, isto é, X é concentrada em torno de sua média.

Uma cota exponencial para o evento $X = 0$ pode ser derivado facilmente:

$$\mathbb{P}[X = 0] = (1-p)^n \leq \exp\{-pn\} = \exp\{-\mu\}, \tag{8.17}$$

onde utilizamos (6.8). Note que essa cota é muito melhor que a obtida da desigualdade (8.15) se tomarmos $\varepsilon := 1$: de (8.15), obtemos $\mathbb{P}[X = 0] \leq 1/\mu$. De fato, para μ grande, $\exp\{-\mu\}$ é muito menor que $1/\mu$.

Agora vamos provar que, para $0 \leq k \leq n$, temos

$$\mathbb{P}[X \geq k] \leq \binom{n}{k} p^k. \tag{8.18}$$

Suponha que $X \geq k$. Seja $S \subseteq [n]$ o conjunto dos índices i tais que $X_i = 1$. Temos $|S| \geq k$. Seja S' o conjunto dos k menores elementos de S e seja l o elemento máximo de S'. Então, para $i \leq l$, o evento

$$X_i = \begin{cases} 1, & \text{se } i \in S'; \\ 0, & \text{caso contrário} \end{cases} \tag{8.19}$$

ocorre. Tal evento ocorre com probabilidade $p^k(1-p)^{l-k} \leq p^k$. Assim, o conjunto dos eventos com $X \geq k$ está contido no conjunto dos eventos tais que vale (8.19) para algum conjunto $S' \in \binom{[n]}{k}$. Como existem $\binom{n}{k}$ possibilidades para o conjunto S', obtemos (8.18).

Além disso, é fácil provar que

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k. \tag{8.20}$$

De fato, comece notando que

$$\binom{n}{k} x^k \leq \sum_{l=0}^{n} \binom{n}{l} x^l = (1+x)^n \leq \exp\{xn\}, \tag{8.21}$$

onde utilizamos (6.8) na última passagem. Agora basta tomar $x := k/n$ para obter

$$\binom{n}{k}\left(\frac{k}{n}\right)^k \leq e^k, \tag{8.22}$$

como queríamos.

Combinando (8.18) com (8.20), obtemos

$$\mathbb{P}[X \geq k] \leq \left(\frac{enp}{k}\right)^k. \tag{8.23}$$

Em particular, se tomarmos $k = \lambda\mu = \lambda np$ para algum $\lambda > 0$, obtemos

$$\mathbb{P}[X \geq \lambda\mu] \leq \left(\frac{e}{\lambda}\right)^{\lambda\mu}. \tag{8.24}$$

Se tomarmos $\lambda > e$, temos

$$\mathbb{P}[X \geq \lambda\mu] \leq \left[\left(\frac{e}{\lambda}\right)^{\lambda}\right]^{\mu}. \tag{8.25}$$

Como

$$\left(\frac{e}{\lambda}\right)^{\lambda} = \big(\exp\{1 - \ln\lambda\}\big)^{\lambda} = \exp\{-\lambda(\ln\lambda - 1)\}, \tag{8.26}$$

obtemos

$$\mathbb{P}[X \geq \lambda\mu] \leq \big(\exp\{-\lambda(\ln\lambda - 1)\}\big)^{\mu}. \tag{8.27}$$

Como $\alpha := -\lambda(\ln\lambda - 1)$ é negativo, então $\exp\{\alpha\} < 1$, de modo que o lado direito da cota (8.27) é muito menor que $1/\mu$ quando $\mu \to \infty$.

Passemos agora a limitar a probabilidade de desvios de $\pm\varepsilon\mu$ em relação à média, para $\varepsilon > 0$ pequeno.

Fixe $t > 0$. Para qualquer $s \geq 0$, temos

$$\mathbb{P}[X \geq \mu + t] = \mathbb{P}\big[\exp\{sX\} \geq \exp\{s(\mu + t)\}\big], \tag{8.28}$$

pois a função $f(y) := \exp\{sy\}$ é crescente se $s \geq 0$. Como $\exp\{sX\} \geq 0$, podemos aplicar a desigualdade (6.1) de Markov, obtendo

$$\mathbb{P}[X \geq \mu + t] \leq \exp\{-s(\mu + t)\}\mathbb{E}\big[\exp\{sX\}\big]. \tag{8.29}$$

Temos ainda

$$\mathbb{E}\big[\exp\{sX\}\big] = \mathbb{E}\Big[\exp\big\{s\textstyle\sum_{i=1}^n X_i\big\}\Big] = \mathbb{E}\Big[\textstyle\prod_{i=1}^n \exp\{sX_i\}\Big]. \tag{8.30}$$

Mas como $X_1, \ldots, X_n$ são mutuamente independentes, então as variáveis aleatórias $\exp\{sX_1\}, \ldots, \exp\{sX_n\}$ também o são, de modo que

$$\mathbb{E}\big[\exp\{sX\}\big] = \mathbb{E}\Big[\textstyle\prod_{i=1}^n \exp\{sX_i\}\Big] = \displaystyle\prod_{i=1}^n \mathbb{E}\big[\exp\{sX_i\}\big]. \tag{8.31}$$

Como a variável $\exp\{sX_i\}$ vale 1 com probabilidade $1 - p$ e e^s com probabilidade p, temos

$$\mathbb{E}\big[\exp\{sX\}\big] = \prod_{i=1}^n \mathbb{E}\big[\exp\{sX_i\}\big] = (1 - p + pe^s)^n. \tag{8.32}$$

Concluímos que, para qualquer $s \geq 0$, temos

$$\mathbb{P}[X \geq \mu + t] \leq e^{-s(\mu+t)}(1 - p + pe^s)^n. \tag{8.33}$$

Queremos encontrar o valor de $s \geq 0$ que torna o lado direito de (8.33) o menor possível. Isto é, queremos encontrar $s \geq 0$ que minimiza a função $f(s) := \exp\{-s(\mu+t)\}(1-p+pe^s)^n$. Utilizando ferramentas de cálculo, obtemos que o $s \geq 0$ que minimiza $f(s)$ é tal que

$$e^s = \frac{(\mu+t)(1-p)}{(n-\mu-t)p}, \tag{8.34}$$

onde estamos supondo que $\mu + t < n$. Note que essa restrição não é tão forte, já que $\mathbb{P}[X > n] = 0$, de modo que a única probabilidade que não conseguimos limitar com essa cota é a de que $X = n$. Veremos a seguir que isso poderá ser "recuperado", isto é, ao continuarmos o desenvolvimento da cota, veremos que poderemos usá-la para limitar até mesmo a probabilidade do evento $X = n$.

Utilizando (8.34), obtemos

$$e^{-s(\mu+t)} = \left(\frac{(n-\mu-t)p}{(\mu+t)(1-p)}\right)^{\mu+t} \tag{8.35}$$

e

$$\begin{aligned} 1 - p + pe^s &= 1 - p + \frac{(\mu+t)(1-p)}{n-\mu-t} \\ &= \frac{(n-\mu-t)(1-p) + (\mu+t)(1-p)}{n-\mu-t} \\ &= (1-p)\frac{n}{n-\mu-t}. \end{aligned} \tag{8.36}$$

Utilizando (8.35) e (8.36) em (8.33), obtemos

$$\begin{aligned} \mathbb{P}[X \geq \mu + t] &\leq \left(\frac{(n-\mu-t)p}{(\mu+t)(1-p)}\right)^{\mu+t} \left(\frac{n}{n-\mu-t}\right)^n (1-p)^n \\ &= \frac{(n-\mu-t)^{\mu+t} p^{\mu+t} n^{\mu+t} n^{n-\mu-t} (1-p)^n}{(\mu+t)^{\mu+t}(1-p)^{\mu+t}(n-\mu-t)^n} \\ &= \frac{(np)^{\mu+t}}{(\mu+t)^{\mu+t}} \frac{\big(n(1-p)\big)^{n-\mu-t}}{(n-\mu-t)^{n-\mu-t}} \\ &= \left(\frac{\mu}{\mu+t}\right)^{\mu+t} \left(\frac{n-\mu}{n-\mu-t}\right)^{n-\mu-t}. \end{aligned} \tag{8.37}$$

Observe que, em (8.37), se tivermos $\mu + t = n$, então um denominador vale 0. Entretanto, o expoente da fração cujo denominador vale 0 também é 0. Vamos convencionar então que $(1/0)^0 = 1$. Assim, obtemos que $\mathbb{P}[X \geq n] \leq (np/n)^n = p^n$, o que é verdade. Assim, a cota continua valendo mesmo quando $\mu + t = n$.

Portanto, (8.37) vale para todo $0 \leq t \leq n - \mu$.

Um bom exercício é utilizar a mesma idéia para obter um limitante superior para $\mathbb{P}[X \leq \mu - t]$, com $t \geq 0$.

Teorema 8.1. *Seja*

$$\varphi(x) := (1+x)\ln(1+x) - x \tag{8.38}$$

para $x \geq -1$, *com* $\varphi(x) := \infty$ *se* $x < -1$.

Seja X *uma variável aleatória com* $X \sim \mathrm{Bi}(n,p)$ *e sejam* $\mu := \mathbb{E}[X] = np$ *e* $t \geq 0$. *Então*

$$\begin{aligned}\mathbb{P}[X \geq \mu + t] &\leq \exp\big\{-\mu\varphi(t/\mu)\big\} &(8.39)\\ &\leq \exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\} &(8.40)\end{aligned}$$

e

$$\begin{aligned}\mathbb{P}[X \leq \mu - t] &\leq \exp\big\{-\mu\varphi(-t/\mu)\big\} &(8.41)\\ &\leq \exp\left\{-\frac{t^2}{2\mu}\right\}. &(8.42)\end{aligned}$$

Demonstração. Vamos provar apenas (8.39) e (8.40).

Já provamos em (8.37) que

$$\mathbb{P}[X \geq \mu + t] \leq \left(\frac{\mu}{\mu + t}\right)^{\mu+t}\left(\frac{n-\mu}{n-\mu-t}\right)^{n-\mu-t}. \tag{8.43}$$

Seja A o lado direito de (8.43). Temos

$$\begin{aligned}\ln A &= -(\mu+t)\ln\frac{\mu+t}{\mu} - (n-\mu-t)\ln\frac{n-\mu-t}{n-\mu}\\ &= -\mu(1+t/\mu)\ln\big(1+t/\mu\big)\\ &\qquad - (n-\mu)\big[1-t/(n-\mu)\big]\ln\big[1-t/(n-\mu)\big]\\ &= -\mu\left[\left(1+\frac{t}{\mu}\right)\ln\left(1+\frac{t}{\mu}\right) - \frac{t}{\mu}\right]\\ &\qquad - (n-\mu)\left[\left(1-\frac{t}{n-\mu}\right)\ln\left(1-\frac{t}{n-\mu}\right) + \frac{t}{n-\mu}\right]\\ &= -\mu\varphi(t/\mu) - (n-\mu)\varphi\big(-t/(n-\mu)\big).\end{aligned} \tag{8.44}$$

Queremos mostrar que $-(n-\mu)\varphi\big(-t/(n-\mu)\big) \leq 0$, ou, equivalentemente, que

$$(n-\mu-t)\ln\big((n-\mu-t)/(n-\mu)\big) + t \geq 0 \tag{8.45}$$

para todo $0 \leq t \leq n-\mu$. Tomando cada um dos lados de (8.45) como expoente de e, obtemos que vale (8.45) se, e somente se,

$$\left(\frac{n-\mu-t}{n-\mu}\right)^{n-\mu-t} + e^t \geq 1. \tag{8.46}$$

Mas isso sempre é verdade, pois $e^t \geq 1$ para $t \geq 0$ e a primeira parcela é não-negativa. Concluímos então que

$$\ln A \leq -\mu\varphi(t/\mu). \tag{8.47}$$

Mas então, de (8.43) e (8.47), temos

$$\mathbb{P}[X \geq \mu + t] \leq \exp\{-\mu\varphi(t/\mu)\}, \tag{8.48}$$

como queríamos.

Resta provarmos (8.40). Tome

$$f(x) := \frac{x^2}{2(1 + x/3)}. \tag{8.49}$$

Vamos mostrar que $\varphi(x) \geq f(x)$ para $x \geq 0$. Comece notando que

$$\varphi'(x) = \ln(1 + x), \tag{8.50}$$

$$\varphi''(x) = \frac{1}{1 + x}, \tag{8.51}$$

$$f'(x) = \frac{x(2 + x/3)}{2(1 + x/3)^2} \tag{8.52}$$

e

$$f''(x) = \frac{1}{(1 + x/3)^3}. \tag{8.53}$$

Como $\varphi(0) = f(0)$ e

$$\varphi''(x) = \frac{1}{1 + x} \geq \frac{1}{(1 + x/3)^3} = f''(x) \tag{8.54}$$

para todo $x \geq 0$, segue que

$$\varphi(x) \geq f(x) \tag{8.55}$$

para todo $x \geq 0$.

Agora (8.40) segue imediatamente de (8.39) combinada com o fato de que

$$\mu\varphi(t/\mu) \geq \mu f(t/\mu) = \frac{t^2}{2(\mu + t/3)}. \tag{8.56}$$

Note que podemos estender as cotas (8.39) e (8.40) para todo $t \geq 0$, em contrapartida com o intervalo $0 \leq t \leq n - \mu$ existente na cota (8.37). De fato, para $t > n - \mu$, as probabilidades são nulas, enquanto os lados direitos de (8.39) e (8.40) são positivos. □

Observe que, se tomarmos $t := \varepsilon\mu$ em (8.39), para qualquer $\varepsilon > 0$, obtemos

$$\mathbb{P}[X \geq (1 + \varepsilon)\mu] \leq \left(\frac{(1 + \varepsilon)^{1+\varepsilon}}{e^{\varepsilon}}\right)^{-\mu}. \tag{8.57}$$

O seguinte resultado é útil quando t/μ é pequeno:

Corolário 8.1.1.

$$\mathbb{P}[X \geq \mu + t] \leq \exp\left\{\frac{-t^2}{2\mu} + \frac{t^3}{6\mu^2}\right\}. \tag{8.58}$$

Demonstração. Basta combinar (8.40) com o fato de que

$$\frac{1}{\mu + t/3} \geq \frac{\mu - t/3}{\mu^2}. \tag{8.59}$$

□

Corolário 8.1.2. *Para qualquer $\varepsilon > 0$, temos*

$$\mathbb{P}\Big[\big|X - \mu\big| \geq \varepsilon\mu\Big] \leq 2\exp\{-\mu\varphi(\varepsilon)\}. \tag{8.60}$$

Se $\varepsilon \leq 3/2$, então

$$\mathbb{P}\Big[\big|X - \mu\big| \geq \varepsilon\mu\Big] \leq 2\exp\{-\varepsilon^2\mu/3\}. \tag{8.61}$$

Demonstração. A cota (8.60) segue imediatamente de (8.39) e (8.41).

Agora vamos mostrar (8.61). Usando o fato de que $0 < \varepsilon \leq 3/2$ em (8.55), temos

$$\varphi(\varepsilon) \geq \frac{\varepsilon^2}{2(1 + \varepsilon/3)} \geq \frac{\varepsilon^2}{3}. \tag{8.62}$$

Daí segue que $-\varphi(\varepsilon) \leq -\varepsilon^2/3$, e estamos feitos. □

8.4 Uma modificação da distribuição binomial

Suponha que queiramos definir uma variável aleatória $Y := \sum_{i=1}^n Y_i$, sendo que Y_i vale 1 com probabilidade p e -1 com probabilidade $1 - p$.

Uma maneira de se fazer isso é considerar uma variável aleatória $X := \sum_{i=1}^n X_i \sim \mathrm{Bi}(n, p)$ e tomar

$$Y_i := 2X_i - 1 \tag{8.63}$$

para todo i. Obteremos então

$$Y = 2X - n. \tag{8.64}$$

Assim, a variável aleatória Y pode ser estudada em termos da variável binomial X.

Observe, por exemplo, que $\mathbb{E}[Y] = 0$ se $p = 1/2$.

9 Uniformidade na distribuição de arestas de grafos aleatórios

Seja $G := G(n,p)$ um grafo aleatório. O número de arestas de G é uma variável aleatória X com $X \sim \mathrm{Bi}(\binom{n}{2}, p)$, de forma que o número esperado de arestas de G é $\mathbb{E}[X] = p\binom{n}{2}$.

Dado $U \subseteq V_G$, seja

$$e(U) := |\{e \in E_G : e \subseteq U\}| \tag{9.1}$$

o número de arestas de G com as duas pontas em U. Para grafos em geral, $e(U)$ é uma função do conjunto U. Já para grafos aleatórios, $e(U)$ é uma função apenas da cardinalidade $|U|$, isto é, independente do particular conjunto U escolhido. Pode-se dizer que, em grafos aleatórios, $e(U)$ é um valor muito bem determinado, no sentido de que existe um intervalo pequeno de valores possíveis tais que, quando n tende ao infinito, a probabilidade de que $e(U)$ esteja fora desse intervalo tende a zero.

Por essa razão dizemos que existe uma certa uniformidade na distribuição de arestas de $G(n,p)$.

Cabe observar que, fixado $U \subseteq V_G$ de $G := G(n,p)$, se definirmos uma variável aleatória $X_U := e(U)$, temos $X_U \sim \mathrm{Bi}(\binom{|U|}{2}, p)$. Logo, o valor esperado de X_U é $\mathbb{E}[X_U] = p\binom{|U|}{2}$. Queremos mostrar então que existe um intervalo pequeno, centralizado em torno de $\mathbb{E}[X_U]$, tal que a probabilidade de que X_U esteja fora desse intervalo é $o(1)$.

Vamos formalizar melhor esse conceito de uniformidade.

Seja $0 < \eta < 1$. Dizemos que $G = (V,E)$ é *η-uniforme* se, para qualquer $U \subseteq V$ com $|U| \geq \eta|V|$, temos que

$$\left| e(U) - d\binom{|U|}{2} \right| \leq \eta d\binom{|U|}{2}, \tag{9.2}$$

onde $d := |E|/\binom{n}{2}$ é a densidade do grafo G e $n := |V|$.

Uma forma de enxergar tal definição é a seguinte. Fixado $0 < \eta < 1$, dizemos que G é η-uniforme se, para todo subconjunto $U \subseteq V$ suficientemente grande, temos que $e(U)$ difere de no máximo η (isto é, de 100η por cento) do valor esperado de $e(U)$ no grafo aleatório $G(n,d)$, que é o grafo aleatório cuja densidade esperada é a mesma de G.

Uma outra forma é a seguinte. Para cada $U \subseteq V$ com $|U| \geq \eta|V|$, tome $d_U := e(U)/\binom{|U|}{2}$. Podemos pensar em d como a "densidade global" do grafo G e em d_U como a "densidade local" de G em U. Para definirmos grafos η-uniformes, queremos que d_U não esteja muito distante de d, para todo U. Dividindo (9.2) por $\binom{|U|}{2}$, a exigência passa a ser de que $|d_U - d| \leq \eta d$, isto é, de que

$$(1-\eta)d \leq d_U \leq (1+\eta)d \tag{9.3}$$

para todo U suficientemente grande. Em outras palavras, estabelecemos um intervalo em torno da densidade global e exigimos que a densidade local de todo U suficientemente grande esteja nesse intervalo.

Teorema 9.1. *Suponha que $p = p(n)$ com*

$$\lim_{n\to\infty} \frac{pn}{\ln n} = \infty \tag{9.4}$$

Então $G := G(n,p)$ é η-uniforme para qualquer $\eta > 0$ fixo, quase sempre. Isto é, para qualquer $\eta > 0$,

$$\lim_{n\to\infty} \mathbb{P}[G \text{ é } \eta\text{-uniforme}] = 1. \tag{9.5}$$

Demonstração. Fixe $\eta > 0$. Seja $\gamma := \eta/2$.

Seja $d := \big|E(G(n,p))\big|/\binom{n}{2}$ a densidade de $G(n,p)$, de modo que d é uma variável aleatória com distribuição

$$\frac{\mathrm{Bi}\left(\binom{n}{2}, p\right)}{\binom{n}{2}}. \tag{9.6}$$

Fixe $U \subseteq V$, com $|U| \geq \eta n$. Temos $e(U) \sim \mathrm{Bi}\left(\binom{|U|}{2}, p\right)$. Pela cota (8.61), obtemos

$$\mathbb{P}\left[\left|e(U) - p\tbinom{|U|}{2}\right| \geq \gamma p\tbinom{|U|}{2}\right] \leq 2\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tbinom{|U|}{2}\Big\}. \tag{9.7}$$

Assim,

$$\begin{aligned}\mathbb{P}\Big[\exists U \subseteq V \text{ com } |U| \geq \eta n \text{ e } \left|e(U) - p\tbinom{|U|}{2}\right| \geq \gamma p\tbinom{|U|}{2}\Big] \\ \leq \sum_{\eta n \leq u \leq n} \binom{n}{u} 2\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tbinom{u}{2}\Big\}.\end{aligned} \tag{9.8}$$

Observe agora que

$$\begin{aligned}\sum_{\eta n \leq u \leq n} \binom{n}{u} 2\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tbinom{u}{2}\Big\} &\leq 2\sum_{u \geq \eta n} n^u \exp\Big\{-\tfrac{1}{3}\gamma^2 p\tbinom{u}{2}\Big\} \\ &= 2\sum_{u \geq \eta n} \left(n\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tfrac{u-1}{2}\Big\}\right)^u.\end{aligned} \tag{9.9}$$

Vamos mostrar que, para n suficientemente grande, temos

$$n\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tfrac{u-1}{2}\Big\} \leq \frac{1}{2}. \tag{9.10}$$

Ou seja, queremos mostrar que existe n_0 tal que (9.10) vale para todo $n \geq n_0$ e $u \geq \eta n$.

Suponha então que $u \geq \eta n$. Então $u \geq \gamma n$. Temos

$$\begin{aligned}n\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tfrac{u-1}{2}\Big\} &= \exp\Big\{\ln n - \tfrac{1}{6}\gamma^2 p(u-1)\Big\} \\ &\leq \exp\Big\{\ln n - \tfrac{1}{6}\gamma^2 p(\gamma n - 1)\Big\} \\ &= \exp\Big\{\ln n - \tfrac{1}{6}\gamma^2(\gamma pn - p)\Big\}.\end{aligned} \tag{9.11}$$

De (9.4), sabemos que, para qualquer $C \in \mathbb{R}$, existe $n_0' = n_0'(C)$ tal que $pn \geq C \ln n$ para todo $n \geq n_0'$. Então, para todo $n \geq n_0'(C)$, temos

$$\exp\Big\{ \ln n - \tfrac{1}{6}\gamma^2(\gamma pn - p)\Big\} \leq \exp\Big\{ \ln n - \tfrac{1}{6}\gamma^2(\gamma C \ln n - p)\Big\}. \tag{9.12}$$

Tome

$$C := \frac{1 + \ln 2 + \gamma^2 p/2}{\gamma^3/6}. \tag{9.13}$$

É fácil verificar que

$$1 - \gamma^3 C/6 = -\ln 2 - \gamma^2 p/6.$$

Como $\ln n \geq 1$ para todo $n \geq e$ e $-\ln 2 - \gamma^2 p/6 < 0$, temos

$$(\ln n)(1 - \gamma^3 C/6) \leq -\ln 2 - \gamma^2 p/6.$$

Mas então

$$\ln n - \tfrac{1}{6}\gamma^2(\gamma C \ln n - p) \leq -\ln 2, \tag{9.14}$$

de modo que, por (9.12), temos

$$\exp\Big\{ \ln n - \tfrac{1}{6}\gamma^2(\gamma pn - p)\Big\} \leq \exp\{-\ln 2\} = \tfrac{1}{2}. \tag{9.15}$$

Combinando (9.11) com (9.15), obtemos que (9.10) vale para todo $n \geq n_0 := \max\{n_0'(C), e\}$.

Retomando (9.9) e utilizando (9.10), temos, para todo $n \geq n_0$,

$$\begin{aligned} 2\sum_{u \geq \eta n}\left(n\exp\Big\{-\tfrac{1}{3}\gamma^2 p\tfrac{u-1}{2}\Big\}\right)^u &\leq 2\sum_{u \geq \lceil \eta n\rceil}\left(\frac{1}{2}\right)^u \\ &= 2\frac{1/2^{\lceil \eta n \rceil}}{1 - 1/2} \leq \frac{4}{(2^\eta)^n} \\ &= o(1), \end{aligned} \tag{9.16}$$

pois $\eta > 0$ implica em $2^\eta > 1$.

Concluímos, por (9.8), (9.9) e (9.16), que

$$\mathbb{P}\left[\exists U \subseteq V \text{ com } |U| \geq \eta n \text{ e } \left|e(U) - p\tbinom{|U|}{2}\right| \geq \gamma p\tbinom{|U|}{2}\right] = o(1). \tag{9.17}$$

Ou seja, quase todo $G(n,p)$ satisfaz a seguinte propriedade: para todo $U \subseteq V$ com $|U| \geq \eta n$, temos $\left|e(U) - p\binom{|U|}{2}\right| < \frac{\eta}{2}p\binom{|U|}{2}$.

Considere agora a variável aleatória $m := |E(G(n,p))|$, isto é, $m = d\binom{n}{2}$. Claramente $m \sim \mathrm{Bi}\left(\binom{n}{2}, p\right)$. Observe que

$$\frac{\mathrm{Var}(m)}{(\mathbb{E}[m])^2} = \frac{p(1-p)\binom{n}{2}}{p^2\binom{n}{2}^2} = \frac{1-p}{p\binom{n}{2}} \leq \frac{1}{\mathbb{E}[m]} = o(1), \tag{9.18}$$

pois $\mathbb{E}[m] = p\binom{n}{2} \to \infty$ quando $n \to \infty$. Então, pela desigualdade (7.2) de Chebyshev, temos

$$\mathbb{P}\Big[\big|m - p\tbinom{n}{2}\big| \geq \gamma p\tbinom{n}{2}\Big] \leq \frac{1}{\gamma^2}\frac{\mathrm{Var}(m)}{(\mathbb{E}[m])^2} = o(1). \tag{9.19}$$

Dividindo tudo que esta dentro da probabilidade em (9.19) por $\binom{n}{2}$, obtemos

$$\mathbb{P}\big[|d - p| \geq \gamma p\big] = o(1). \tag{9.20}$$

Combinando (9.17) e (9.20), concluímos que quase todo $G(n,p)$ satisfaz as seguintes propriedades: para todo $U \subseteq V$ com $|U| \geq \eta n$, temos $\big|e(U) - p\binom{|U|}{2}\big| < \frac{\eta}{2}p\binom{|U|}{2}$ e, além disso, $|d-p| < \frac{\eta}{2}p$.

Seja G um $G(n,p)$ satisfazendo essas propriedades. Fixe um conjunto $U \subseteq V$ com $|U| \geq \eta n$. Pela desigualdade triangular, temos

$$\begin{aligned}\big|e(U) - d\tbinom{|U|}{2}\big| &= \big|e(U) - p\tbinom{|U|}{2} + p\tbinom{|U|}{2} - d\tbinom{|U|}{2}\big| \\ &\leq \big|e(U) - p\tbinom{|U|}{2}\big| + \big|p\tbinom{|U|}{2} - d\tbinom{|U|}{2}\big| \\ &< \frac{\eta}{2}p\tbinom{|U|}{2} + \frac{\eta}{2}p\tbinom{|U|}{2} = \eta p\tbinom{|U|}{2}.\end{aligned} \tag{9.21}$$

Mas então G é η-uniforme. Concluímos que quase todo $G(n,p)$ é η-uniforme, como queríamos. □

Podemos formalizar o conceito de uniformidade na distribuição de arestas de outra forma: sejam G um grafo com n vértices e $A > 0$ uma constante. Ponha $p := |E(G)|/\binom{n}{2}$. Dizemos que G é (p,A)-*uniforme* se, para quaisquer $U, W \subseteq V_G$, com $U \cap W = \emptyset$ e $1 \leq |U| \leq |W| \leq pn|U|$, temos

$$\Big|e(U,W) - p \cdot |U| \cdot |W|\Big| \leq A\sqrt{pn \cdot |U| \cdot |W|}, \tag{9.22}$$

onde $e(U,W)$ denota o número de arestas que têm uma ponta em U e outra em W.

Teorema 9.2. *Para qualquer $0 < p = p(n) < 1$, temos*

$$\lim_{n\to\infty} \mathbb{P}[G(n,p) \text{ é } (p,5)\text{-uniforme}] = 1. \tag{9.23}$$

Exercício 9. *Seja G um grafo com $V_G := \mathcal{P}([n])$ e*

$$E_G := \big\{\{f,g\} : f,g \subseteq [n] \text{ com } |f \cap g| \text{ par e } f \neq g\big\}. \tag{9.24}$$

Mostre que G é η-uniforme se $n \geq n_0 = n_0(\eta)$, isto é, mostre que, para todo η, existe um n_0 tal que G é η-uniforme para todo $n \geq n_0$.

9.1 Mais formalizações do conceito de uniformidade

Sejam $0 < \eta \leq 1/2$, $0 < p \leq 1$ e G um grafo com n vértices. Dizemos que G é *η-uniforme com densidade* p se, para todo $U, W \subseteq V_G$ com $U \cap W = \emptyset$ e $|U|, |W| \geq \eta n$, vale que

$$\Big| e(U,W) - p \cdot |U| \cdot |W| \Big| \leq \eta p \cdot |U| \cdot |W|. \tag{9.25}$$

É conveniente utilizarmos a seguinte notação: $O_1(x)$ denota um termo y com $|y| \leq x$. Por exemplo, na definição acima, é preciso que, para U, W dessa forma, tenhamos

$$e(U,W) = p \cdot |U| \cdot |W| + O_1\big(\eta p \cdot |U| \cdot |W|\big). \tag{9.26}$$

Na condição acima, podemos interpretar $O_1\big(\eta p \cdot |U| \cdot |W|\big)$ como um "termo de erro". Assim, na definição de η-uniforme com densidade p, estamos exigindo que $e(U,W)$ não fique muito distante de seu valor esperado num $G(n,p)$, que é justamente $p \cdot |U| \cdot |W|$. Mais especificamente, estamos exigindo que o "termo de erro" seja, em valor absoluto, no máximo uma fração do valor esperado de $e(U,W)$.

Agora vamos redefinir o conceito de (p, A)-uniformidade para facilitar nosso trabalho.

Sejam $A > 0$ uma constante, $0 < p \leq 1$ e G um grafo com n vértices. Ponha $d := pn$. Dizemos que G é (p, A)-*uniforme* se, para todo $U, W \subseteq V_G$ com $U \cap W = \emptyset$ e $1 \leq |U| \leq |W| \leq d|U|$, temos

$$e(U,W) = p \cdot |U| \cdot |W| + O_1\big(A\sqrt{d \cdot |U| \cdot |W|}\big). \tag{9.27}$$

Podemos comparar (9.26) e (9.27) para ver em que condições uma definição é mais forte que a outra. Por exemplo, para que seja mais forte dizer que G é (p, A)-uniforme do que dizer que ele é η-uniforme com densidade p, é preciso que

$$\frac{A\sqrt{d \cdot |U| \cdot |W|}}{\eta p \cdot |U| \cdot |W|} < 1, \tag{9.28}$$

isto é, que $A^2 n < \eta^2 p \cdot |U| \cdot |W|$. Por exemplo, se A e η forem constantes e $p \cdot |U| \cdot |W|$ for muito maior que n, então é muito mais forte dizer que G é (p, A)-uniforme.

Vamos agora ver o que as condições de uniformidade de $e(U,W)$ implicam sobre uniformidade de $e(U)$.

Lema 9.3. *Seja G um grafo com n vértices e suponha que G é η-uniforme com densidade p. Então, para qualquer $U \subseteq V_G$ com $|U| \geq 2\eta n$, vale que*

$$e(U) = p\binom{|U|}{2} + O_1\left(\eta p\binom{|U|}{2}\right). \tag{9.29}$$

Demonstração. Exercício. □

Lema 9.4. *Seja G um grafo (p, A)-uniforme. Suponha que $d := pn \geq 2$. Então, para todo $U \subseteq V_G$, temos*

$$e(U) = p\tbinom{|U|}{2} + O_1\big(A|U|\sqrt{d}\big). \tag{9.30}$$

Demonstração. Fixe $U \subseteq V_G$ e seja $u := |U|$. Podemos supor que $u \geq 2$.

Para contar as arestas de $G[U]$, vamos fixar $1 \leq s < u$ e considerar a soma

$$\sum_{S \in \binom{U}{s}} e(S, U \setminus S). \tag{9.31}$$

É evidente que cada aresta de $G[U]$ é contada ao menos uma vez no somatório (9.31) acima. Entretanto, cada aresta é contada repetidas vezes. Precisamos contar o número de repetições para determinar exatamente $e(U)$ em função desse somatório.

Fixe $1 \leq s < u$ e seja $e = vw$ uma aresta de $G[U]$. Precisamos contar quantos $S \in \binom{U}{s}$ satisfazem $v \in S$ e $w \notin S$ ou então $w \in S$ e $v \notin S$, isto é, quantas vezes a aresta e será contabilizada em (9.31). Por simetria, basta calcular quantos $S \in \binom{U}{s}$ satisfazem $v \in S$ e $w \notin S$ e multiplicar esse número por 2.

Precisamos que v esteja em S e w fora de S, de modo que, dentre os u elementos de U, dois deles já têm sua "posição" (dentro ou fora de S) determinada. Dentre os $u - 2$ elementos restantes de U, precisamos escolher $s - 1$ deles para colocá-los em S (lembre-se que v já faz parte de S e que S precisa ter cardinalidade s). Portanto, existem $\binom{u-2}{s-1}$ conjuntos $S \in \binom{U}{s}$ tais que $v \in S$ e $w \notin S$. Mas então cada aresta de $G[U]$ é contabilizada exatamente $2\binom{u-2}{s-1}$ vezes em (9.31):

$$2\binom{u-2}{s-1}e(U) = \sum_{S \in \binom{U}{s}} e(S, U \setminus S). \tag{9.32}$$

Pela definição de (p, A)-uniformidade, se $1 \leq s \leq u - s \leq ds$, então

$$\begin{aligned} e(U) &= \frac{1}{2}\binom{u-2}{s-1}^{-1} \sum_{S \in \binom{U}{s}} \Big\{p \cdot |S| \cdot |U \setminus S| + O_1\big(A\sqrt{d \cdot |S| \cdot |U \setminus S|}\big)\Big\} \\ &= \frac{1}{2}\binom{u-2}{s-1}^{-1}\binom{u}{s}\Big\{ps(u-s) + O_1\big(A\sqrt{ds(u-s)}\big)\Big\}. \end{aligned} \tag{9.33}$$

Tome $s := \lfloor u/2 \rfloor$, de modo que $1 \leq s \leq u - s \leq ds$ vale se, e somente se, $\lceil u/2 \rceil \leq pn\lfloor u/2 \rfloor$. Mas $d = pn \geq 2$. Como $\lceil u/2 \rceil \leq 2\lfloor u/2 \rfloor$ para todo $u \geq 2$, então vale $1 \leq s \leq u - s \leq ds$. Portanto (9.33) vale para $s = \lfloor u/2 \rfloor$.

Temos

$$\begin{aligned}
\binom{u}{s}\binom{u-2}{s-1}^{-1} &= \binom{u}{\lfloor u/2\rfloor}\binom{u-2}{\lfloor u/2\rfloor-1}^{-1} \\
&= \frac{u!}{\lfloor u/2\rfloor!\lceil u/2\rceil!}\frac{(\lfloor u/2\rfloor-1)!(\lceil u/2\rceil-1)!}{(u-2)!} \\
&= \frac{u(u-1)}{\lfloor u/2\rfloor\lceil u/2\rceil} \leq \frac{u(u-1)}{(u/2-1/2)(u/2)} = 4.
\end{aligned} \tag{9.34}$$

Combinando (9.33) com (9.34), obtemos

$$\begin{aligned}
e(U) &= \frac{1}{2}\frac{u(u-1)}{\lfloor u/2\rfloor\lceil u/2\rceil}p\lfloor u/2\rfloor\lceil u/2\rceil + O_1\left(2A\sqrt{d\lfloor u/2\rfloor\lceil u/2\rceil}\right) \\
&= p\binom{u}{2} + O_1\left(2A\sqrt{d\lfloor u/2\rfloor\lceil u/2\rceil}\right).
\end{aligned} \tag{9.35}$$

Ademais, como $\lfloor u/2\rfloor\lceil u/2\rceil \leq (u/2-1/2)(u/2+1/2) \leq u^2/4$, obtemos

$$e(U) = p\binom{|U|}{2} + O_1\big(A|U|\sqrt{d}\big), \tag{9.36}$$

como queríamos. □

Note que, nessa última demonstração, mostramos que $e(U)$ pode ser escrito em termos de $e(S,W)$, onde $S, W \subseteq U$. Podemos também escrever $e(U)$ em função de $e(S)$, onde $S \subseteq U$, da seguinte forma.

Fixe $2 \leq s \leq u := |U|$ e considere a soma $\sum_{S\in\binom{U}{s}} e(S)$. Cada aresta de $G[U]$ é contada repetidas vezes. Vamos contar quantas vezes contamos cada aresta de $G[U]$ nesse somatório.

Seja $vw \in E(G[U])$. Precisamos contar quantos $S \in \binom{U}{s}$ existem tais que $vw \in E(G[S])$. Para que isso ocorra, é preciso que $v, w \in S$. É preciso ainda escolher outros $s-2$ elementos para fazerem parte de S, e há $u-2 = |U \setminus \{v,w\}|$ possíveis elementos dentre os quais podemos escolher. Logo, cada aresta é contada $\binom{u-2}{s-2}$ vezes. Concluímos que

$$e(U) = \binom{u-2}{s-2}^{-1} \sum_{S\in\binom{U}{s}} e(S). \tag{9.37}$$

Antes de provar mais um lema, considere $G := G(n,p)$ com $0 < p := p(n) < 1$ e note que $d := d(n) := pn$ é, "basicamente", o grau médio de G. De fato, para todo $v \in V_G$, o grau esperado de v é $p(n-1)$, de modo que o grau médio esperado de G é $p(n-1)$, que é muito próximo de pn para n grande.

Lema 9.5. *Sejam $0 < \eta < 1$, $0 < p := p(n) < 1$ e $d := d(n) := pn$. Então existe $d_0 = d_0(\eta)$ tal que quase todo $G(n,p)$ com $d \geq d_0$ é η-uniforme com densidade p.*

Demonstração. Seja $G := (V, E) := G(n, p)$.

Sejam $U, W \subseteq V$ com $U \cap W = \emptyset$, $u := |U| \geq \eta n$ e $w := |W| \geq \eta n$. Observe que

$$\mathbb{E}\left[e(U, W)\right] = puw \geq p\eta nu = \eta du. \tag{9.38}$$

Então, por (8.61), temos

$$\begin{aligned} P_{U,W} &:= \mathbb{P}\Big[\big|e(U, W) - puw\big| > \eta puw\Big] \\ &\leq 2\exp\left\{-\tfrac{1}{3}\eta^2 puw\right\} \leq 2\exp\left\{-\tfrac{1}{3}\eta^3 du\right\}, \end{aligned} \tag{9.39}$$

onde utilizamos (9.38) na última passagem.

Sejam $u \geq w \geq \lceil \eta n \rceil \geq \eta n$. Tome

$$\begin{aligned} P_{u,w} &:= \mathbb{P}\Big[\exists U, W \subseteq V \text{ com } U \cap W = \emptyset, \\ &\qquad\qquad |U| = u, |W| = w \text{ e } \big|e(U, W) - puw\big| > \eta puw\Big] \\ &\leq \sum_{U \in \binom{V}{u}, W \in \binom{V}{w}} P_{U,W}. \end{aligned} \tag{9.40}$$

Vamos limitar superiormente o número de pares (U, W) com $U \in \binom{V}{u}$ e $W \in \binom{V}{w}$. Utilizando (8.20), temos

$$\binom{n}{u}\binom{n}{w} \leq \left(\frac{en}{u}\right)^u \left(\frac{en}{w}\right)^w \leq \left(\frac{en}{\eta n}\right)^u \left(\frac{en}{\eta n}\right)^w \leq \left(\frac{e}{\eta}\right)^{2u}. \tag{9.41}$$

Combinando (9.40) com (9.41) e (9.39), obtemos

$$P_{u,w} \leq \left(\frac{e}{\eta}\right)^{2u} 2\exp\left\{-\tfrac{1}{3}\eta^3 du\right\} = 2\left[\left(\frac{e}{\eta}\right)^2 \exp\left\{-\tfrac{1}{3}\eta^3 d\right\}\right]^u. \tag{9.42}$$

Tome

$$d_0(\eta) := \frac{3}{\eta^3}\Big(2 - 2\ln\eta + \ln 2\Big). \tag{9.43}$$

Se $d \geq d_0(\eta)$, temos

$$\tfrac{1}{3}\eta^3 d \geq 2(1 - \ln\eta) + \ln 2,$$

de modo que

$$2(1 - \ln\eta) - \tfrac{1}{3}\eta^3 d \leq -\ln 2,$$

ou seja,

$$\left(\frac{e}{\eta}\right)^2 \exp\left\{-\tfrac{1}{3}\eta^3 d\right\} \leq \frac{1}{2}. \tag{9.44}$$

Usando (9.42) e (9.44), temos que, se $d \geq d_0(\eta)$, então

$$P_{u,w} \leq 2\left(\frac{1}{2}\right)^u. \tag{9.45}$$

Portanto, se $d \geq d_0(\eta)$, então

$$
\begin{aligned}
P := \mathbb{P}\Big[&\exists U, W \subseteq V \text{ com } U \cap W = \emptyset, |U| = u \geq \eta n, \\
&|W| = w \geq \eta n \text{ e } \big|e(U,W) - puw\big| > \eta puw\Big] \\
&\leq \sum_{u=\lceil \eta n \rceil}^{n} \sum_{w=\lceil \eta n \rceil}^{n} P_{u,w} \leq \sum_{u=\lceil \eta n \rceil}^{n} \sum_{w=\lceil \eta n \rceil}^{n} 2\left(\frac{1}{2}\right)^{u} \\
&\leq 2 \sum_{u \geq \lceil \eta n \rceil} \big((1-\eta)n + 1\big)\left(\frac{1}{2}\right)^{u} \leq (2n+1) \sum_{u \geq \lceil \eta n \rceil} \left(\frac{1}{2}\right)^{u} \\
&= (2n+1)\frac{1/2^{\lceil \eta n \rceil}}{1/2} = \frac{4n+2}{2^{\lceil \eta n \rceil}} \leq \frac{4n+2}{2^{\eta n}} = \frac{4n+2}{(2^{\eta})^{n}} = o(1),
\end{aligned}
\tag{9.46}
$$

pois $\eta > 0$ implica $2^{\eta} > 1$. Mas isso é justamente o que queríamos. □

Lema 9.6. *Seja* $0 < p := p(n) < 1$. *Quase todo* $G(n,p)$ *é* (p,A)*-uniforme para* $A = e^{3/2} < 5$.

Demonstração. Seja $G := (V,E) := G(n,p)$.

Seja $d := d(n) := pn$ e suponha, sem perda de generalidade, que $d = pn \geq 1$. Podemos fazer essa suposição pois, se $d < 1$, então a definição de (p,A)-uniformidade é vazia, isto é, todo grafo com $d < 1$ é (p,A)-uniforme para qualquer $A > 0$.

Tome $A := e^{3/2}$. Seja

$$\mathcal{F} := \big\{(U,W) : U, W \subseteq V \text{ com } U \cap W = \emptyset \text{ e } 1 \leq |U| \leq |W| \leq d|U|\big\}. \tag{9.47}$$

Para cada $(U,W) \in \mathcal{F}$, defina a variável aleatória

$$X_{U,W} := \begin{cases} 1, & \text{se } \big|e(U,W) - puw\big| > A\sqrt{duw}; \\ 0, & \text{caso contrário,} \end{cases} \tag{9.48}$$

onde u denota $|U|$ e w denota $|W|$. Temos

$$P_{U,W} := \mathbb{E}\left[X_{U,W}\right] = \mathbb{P}\Big[\big|e(U,W) - puw\big| > A\sqrt{duw}\Big]. \tag{9.49}$$

Observe que $P_{U,W}$ depende unicamente de u e w, isto é, independe dos particulares U e W escolhidos. Logo, podemos escrever $P_{u,w}$ no lugar de $P_{U,W}$ quando conveniente.

Defina a variável aleatória $X := \sum_{(U,W) \in \mathcal{F}} X_{U,W}$, que conta o número de pares "ruins", isto é, pares de conjuntos que violam a condição que queremos satisfeita. Temos

$$\mathbb{E}\left[X\right] = \sum_{(U,W) \in \mathcal{F}} \mathbb{E}\left[X_{U,W}\right] = \sum_{(U,W) \in \mathcal{F}} P_{U,W}. \tag{9.50}$$

Vamos mostrar que $\mathbb{E}[X] = o(1)$. Para tanto, vamos particionar convenientemente a coleção $\mathcal{F}$.

Seja $(U, W) \in \mathcal{F}$. Denote $|U|$ por u e $|W|$ por w. Defina

$$\mu := \mu(U, W) := puw, \tag{9.51}$$

$$t := t(U, W) := A\sqrt{duw} \tag{9.52}$$

e

$$\eta := \eta(U, W) := \frac{t(U, W)}{\mu(U, W)}. \tag{9.53}$$

Observe que $\mu(U, W)$ é o valor esperado de $e(U, W)$, que $t(U, W)$ é o desvio que podemos tolerar de $e(U, W)$ com relação a $\mu(U, W)$ e que $\eta(U, W)$ é o desvio relativo que podemos tolerar. Não vamos nos referir a μ como $\mu(U, W)$ sempre, mas é importante lembrar que μ é uma função de U e W. Evidentemente o mesmo vale para t e η.

Note que

$$\eta(U, W) = \frac{A\sqrt{duw}}{puw} = \frac{An\sqrt{duw}}{duw} = \frac{An}{\sqrt{duw}}. \tag{9.54}$$

Tome $\mathcal{F}_1 := \{(U, W) \in \mathcal{F} : \eta(U, W) \leq e^2\}$ e $\mathcal{F}_2 := \mathcal{F} \setminus \mathcal{F}_1$. Defina a variável aleatória $X_i := \sum_{(U,W)\in\mathcal{F}_i} X_{U,W}$ para $i = 1, 2$ e observe que $X = X_1 + X_2$. Vamos mostrar separadamente que $\mathbb{E}[X_1] = o(1)$ e $\mathbb{E}[X_2] = o(1)$.

Primeiro vamos mostrar que

$$\mathbb{E}[X_1] = \sum_{(U,W)\in\mathcal{F}_1} P_{U,W} = o(1). \tag{9.55}$$

Fixe $(U, W) \in \mathcal{F}_1$. Usando (8.40) e (8.42), temos

$$\begin{aligned} P_{U,W} &= \mathbb{P}\Big[\big|X_{U,W} - \mu\big| > t\Big] \\ &\leq \exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\} + \exp\left\{-\frac{t^2}{2\mu}\right\} \\ &\leq 2\exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\}. \end{aligned} \tag{9.56}$$

Como $(U, W) \in \mathcal{F}_1$, temos $t = t(U, W) \leq \mu(U, W)e^2 = \mu e^2$, de modo que

$$\frac{t^2}{2(\mu + t/3)} \geq \frac{t^2}{2(\mu + \mu e^2/3)} = \frac{A^2 duw}{2puw(1 + e^2/3)} = \frac{e^3 n}{2(1 + e^2/3)}. \tag{9.57}$$

Portanto

$$P_{U,W} \leq 2\exp\left\{-\frac{e^3 n}{2(1 + e^2/3)}\right\}. \tag{9.58}$$

Mas então

$$\begin{aligned}
\mathbb{E}\left[X_1\right] &= \sum_{(U,W)\in\mathcal{F}_1} P_{U,W} \le (2^n)^2 2\exp\left\{-\frac{e^3 n}{2(1+e^2/3)}\right\} \\
&= 2\exp\left\{n\ln 4\right\}\exp\left\{-\frac{e^3 n}{2(1+e^2/3)}\right\} \\
&= 2\exp\left\{-n\frac{e^3-2(1+e^2/3)\ln 4}{2(1+e^2/3)}\right\} = o(1),
\end{aligned} \tag{9.59}$$

pois $e^3 > 2(1+e^2/3)\ln 4$ e $2(1+e^2/3) > 0$.

Resta mostrarmos que $\mathbb{E}\left[X_2\right] = o(1)$.

Seja $(U,W) \in \mathcal{F}_2$. Temos $\eta(U,W) > e^2$, de modo que $t = t(U,W) > \mu(U,W)e^2 = \mu e^2$. Logo, $\mu - t < (1-e^2)\mu < 0$, de maneira que $P_{U,W} = \mathbb{P}[X_{U,W} > \mu + t]$.

Continue denotando $|U|$ por u e $|W|$ por w. Usando (8.23), temos

$$\begin{aligned}
P_{U,W} = \mathbb{P}[X_{U,W} > \mu + t] &\le \mathbb{P}\big[X_{U,W} \ge \lceil t\rceil\big] \le \left(\frac{euw}{\lceil t\rceil}\right)^{\lceil t\rceil} p^{\lceil t\rceil} \\
= \left(\frac{epuw}{\lceil t\rceil}\right)^{\lceil t\rceil} = \left(\frac{e\mu}{\lceil t\rceil}\right)^{\lceil t\rceil} &\le \left(\frac{e\mu}{t}\right)^{\lceil t\rceil} = \left(\frac{e}{\eta}\right)^{\lceil t\rceil} \le \left(\frac{e}{\eta}\right)^{t},
\end{aligned} \tag{9.60}$$

onde a última passagem é justificada pelo fato de que $\eta > e^2$, de modo que $e/\eta < 1$.

Como $1 \le u \le w \le du$, então

$$\sqrt{duw} \ge \sqrt{w^2} = w \ge \frac{1}{2}(u+w). \tag{9.61}$$

Seja $r := r(U,W) := u + w$. Da equação acima, temos

$$\sqrt{duw} \ge \frac{r(U,W)}{2}. \tag{9.62}$$

Defina também $y(U,W) := e/\eta(U,W)$, de modo que

$$y(U,W) = \frac{e}{\eta(U,W)} = \frac{e\sqrt{duw}}{An} \ge \frac{er(U,W)}{2An}. \tag{9.63}$$

Observe ainda que

$$y(U,W) < \frac{1}{e}, \tag{9.64}$$

já que $e^2 < \eta(U,W)$.

As equações (9.60), (9.62), (9.63) e (9.64) valem para todo $(U,W) \in \mathcal{F}_2$.

Tome $f(x) := x^{Bx} = (x^x)^B$, onde $B := A^2n/e$, e note que, para todo $(U,W) \in \mathcal{F}_2$, temos

$$f\big(y(U,W)\big) = \left(\frac{e}{\eta}\right)^{(e/\eta)(A^2n/e)} = \left(\frac{e}{\eta}\right)^{A^2n/\eta} = \left(\frac{e}{\eta}\right)^{t}, \tag{9.65}$$

onde utilizamos (9.54) na última passagem.

Como $f'(x) = Bf(x)(1 + \ln x)$ e $B > 0$, temos que $f'(y) < 0$ para todo $0 < y < 1/e$, isto é, $f(x)$ é decrescente nesse intervalo. Utilizando esse fato em conjunto com (9.63), obtemos que, para todo $(U, W) \in \mathcal{F}_2$,

$$\left(\frac{e}{\eta}\right)^t = f\big(y(U,W)\big) \leq f\left(\frac{er(U,W)}{2An}\right), \tag{9.66}$$

ou seja,

$$\left(\frac{e}{\eta}\right)^t \leq \left(\frac{er}{2An}\right)^{[A^2n/e][er/(2An)]} = \left(\frac{er}{2An}\right)^{Ar/2}. \tag{9.67}$$

Note ainda que $r(U, W)$, assim como todas as outras funções de (U, W), dependem unicamente de u e w, assim como $P_{U,W}$ acima. Então, para todo $1 \leq u \leq w \leq du$, usamos (9.60), (8.20) e (9.67) para obter

$$\binom{n}{r} P_{u,w} \leq \left(\frac{en}{r}\right)^r \left(\frac{er}{2An}\right)^{Ar/2} = \left[\frac{en}{r}\left(\frac{er}{2An}\right)^{A/2}\right]^r. \tag{9.68}$$

Como $r \leq 2n$, então $\frac{er}{2An} < 1$. Ademais, como $A > 4$, temos

$$\binom{n}{r} P_{u,w} \leq \left[\frac{en}{r}\left(\frac{er}{2An}\right)^{A/2}\right]^r \leq \left[\frac{en}{r}\left(\frac{er}{2An}\right)^{2}\right]^r = \left(\frac{r}{4n}\right)^r. \tag{9.69}$$

Podemos, finalmente, limitar superiormente $\mathbb{E}[X_2]$, usando (9.69):

$$\begin{aligned}
\mathbb{E}[X_2] = \sum_{(U,W)\in\mathcal{F}_2} P_{U,W} &\leq \sum_{1\leq u<r\leq n} \sum_{\substack{U\in\binom{V}{u} \\ W\in\binom{V}{r-u} \\ \eta(U,W)>e^2}} P_{u,w} \\
&\leq \sum_{1\leq u<r\leq n} \binom{n}{r}\binom{r}{u} P_{u,w} \leq \sum_{1\leq u<r\leq n} \binom{r}{u}\left(\frac{r}{4n}\right)^r \\
&\leq \sum_{2\leq r\leq n} \left(\frac{r}{4n}\right)^r \sum_{1\leq u<r} \binom{r}{u} \leq \sum_{2\leq r\leq n} \left(\frac{r}{4n}\right)^r 2^r \\
&= \sum_{2\leq r\leq n} \left(\frac{r}{2n}\right)^r.
\end{aligned} \tag{9.70}$$

Mas

$$
\begin{aligned}
\sum_{r=2}^{n}\left(\frac{r}{2n}\right)^r &\le \sum_{r=2}^{\lfloor 2\lg n\rfloor}\left(\frac{r}{2n}\right)^r + \sum_{r=\lceil 2\lg n\rceil}^{n}\left(\frac{r}{2n}\right)^r \\
&\le \sum_{r=2}^{\lfloor 2\lg n\rfloor}\left(\frac{\lg n}{n}\right)^2 + \sum_{r=\lceil 2\lg n\rceil}^{\infty}\left(\frac{1}{2}\right)^r \\
&\le \big(\lfloor 2\lg n\rfloor - 2 + 1\big)\left(\frac{\lg n}{n}\right)^2 + \frac{1/2^{\lceil 2\lg n\rceil}}{1-1/2} \\
&\le (2\lg n)\left(\frac{\lg n}{n}\right)^2 + \frac{2}{2^{2\lg n}} = 2\frac{\lg^3 n}{n^2} + \frac{2}{n^2} \\
&= o(1).
\end{aligned}
\tag{9.71}
$$

Mostramos assim que $\mathbb{E}[X] = o(1)$. Utilizando agora a desigualdade (6.1) de Markov, obtemos

$$\mathbb{P}[X > 0] = \mathbb{P}[X \ge 1] \le \mathbb{E}[X] = o(1), \tag{9.72}$$

isto é, quase todo $G(n,p)$ é (p,A)-uniforme para $A = e^{3/2}$, como queríamos. $\square$

9.2 (n,d,λ)-grafos

Vamos estudar agora um pouco de um tópico favorito de Noga Alon[1].

Seja $G = (V,E)$ um grafo com n vértices. Suponha que $V = [n]$, sem perda de generalidade. A matriz de adjacência $A = (a_{ij})_{1\le i,j\le n}$ de G, é definida da seguinte maneira:

$$a_{ij} := \begin{cases} 1, & \text{se } ij \in E; \\ 0, & \text{caso contrário.} \end{cases} \tag{9.73}$$

Note que a soma de elementos de cada linha i é o grau do vértice i. Obviamente o mesmo vale para cada coluna j.

Como a matriz de adjacência é simétrica, então ela é diagonalizável e todos seus autovalores são reais. Sejam $\lambda_0 \ge \cdots \ge \lambda_{n-1}$ seus autovalores. Dizemos que $\{\lambda_0, \ldots, \lambda_{n-1}\}$ são os *autovalores do grafo G*.

Sejam $\underline{x}_0, \ldots, \underline{x}_{n-1}$ os autovetores de A, sendo que o autovalor associado ao autovetor $\underline{x}_i$ é λ_i. Como A é simétrica, podemos supor que $\{\underline{x}_0, \ldots, \underline{x}_{n-1}\}$ é uma base ortonormal de $\mathbb{R}^n$ (não é difícil provar que autovetores associados a autovalores diferentes de uma matriz simétrica são ortogonais).

Podemos decompor A da seguinte maneira:

$$A = \sum_{i=0}^{n-1} \lambda_i \big(\underline{x}_i \underline{x}_i^{\,T}\big), \tag{9.74}$$

[1] Visite `http://www.pims.math.ca/science/2002/aga/phenomenavideos/alon/`.

onde é fácil ver que $\underline{x}_i\underline{x}_i{}^T$ é uma matriz de mesmas dimensões que A e com posto 1 para todo i. Para ver que (9.74) de fato vale, observe que

$$\Big(\sum_{i=0}^{n-1}\lambda_i\big(\underline{x}_i\underline{x}_i{}^T\big)\Big)\underline{x}_j=\sum_{i=0}^{n-1}\lambda_i\underline{x}_i\big(\underline{x}_i{}^T\underline{x}_j\big)=\lambda_j\underline{x}_j=A\underline{x}_j. \tag{9.75}$$

Isso é suficiente para mostrar a validade de (9.74) pois um operador linear é totalmente determinado pela sua ação em elementos de uma base.

Suponha, de agora em diante, que G é d-regular. É fácil ver que $(1,\dots,1)^T$ é um autovetor de A associado ao autovalor d. Ademais, $|\lambda_i|\le d$ para todo i. De fato, seja x_{ip} a coordenada de $\underline{x}_i$ com o maior valor absoluto, isto é, $|x_{ip}|\ge|x_{ij}|$ para todo j. Vamos limitar superiormente o valor absoluto da p-ésima coordenada do vetor $A\underline{x}_i=\lambda_i\underline{x}_i$:

$$\begin{aligned}|\lambda_i x_{ip}|&=\big|(\lambda_i\underline{x}_i)_p\big|=\big|(A\underline{x}_i)_p\big|=\big|\textstyle\sum_{j=0}^{n-1}a_{pj}x_{ij}\big|\le\sum_{j=0}^{n-1}|a_{pj}x_{ij}|\\&=\textstyle\sum_{j=0}^{n-1}a_{pj}|x_{ij}|\le\sum_{j=0}^{n-1}a_{pj}|x_{ip}|=d|x_{ip}|.\end{aligned} \tag{9.76}$$

Então $|\lambda_i x_{ip}|=|\lambda_i||x_{ip}|\le d|x_{ip}|$. Como $\underline{x}_i\ne 0$, então $|x_{ip}|>0$, de modo que $|\lambda_i|\le d$, como queríamos.

Concluímos que

$$\lambda_0=d \tag{9.77}$$

e podemos supor que $\underline{x}_0=(1,\dots,1)^T/\sqrt{n}$, já que $\|\underline{x}_0\|=1$.

Podemos agora definir um (n,d,λ)-grafo. Seja G um grafo como acima, isto é, com n vértices, e com autovalores $\lambda_0\ge\cdots\ge\lambda_{n-1}$. Dizemos que G é um *(n,d,λ)-grafo* se G é d-regular e $\lambda\ge|\lambda_i|$ para todo $i>0$.

Por exemplo, seja G um grafo d-regular com n vértices e seja λ tal que $|\lambda_i|\le\lambda$ para todo $i>0$, isto é, tome

$$\lambda:=\max\{|\lambda_1|,\dots,|\lambda_{n-1}|\}=\max\{\lambda_1,-\lambda_{n-1}\}. \tag{9.78}$$

Então G é um (n,d,λ)-grafo.

Estamos interessados em (n,d,λ)-grafos com λ pequeno. Porém, como mostramos a seguir, λ não pode ser muito pequeno.

Proposição 9.7. *Seja G um (n,d,λ)-grafo. Se $d\le(1-\varepsilon)n$ para algum $\varepsilon>0$, então $\lambda=\Omega(\sqrt{d})$.*

Demonstração. Suponha que $G=([n],E)$. Seja A a matriz de adjacência de G e tome $B:=A^2$. É fácil ver que b_{ii} é o grau de i em G. Mas então

$$nd=\mathrm{tr}(A^2), \tag{9.79}$$

onde $\mathrm{tr}(M)$ denota o traço de uma matriz M, ou seja, a soma dos elementos da diagonal de M.

Sejam $\lambda_0,\dots,\lambda_{n-1}$ os autovalores de G, com os respectivos autovetores $\underline{x}_0,\dots,\underline{x}_{n-1}$. É fácil ver que $\lambda_0^2,\dots,\lambda_{n-1}^2$ são os autovalores de A^2. De fato, temos $A^2\underline{x}_i=A(A\underline{x}_i)=A(\lambda_i\underline{x}_i)=\lambda_i(A\underline{x}_i)=\lambda_i^2\underline{x}_i$.

Agora usamos o fato de que o traço de uma matriz é a soma de seus autovalores para obter

$$nd = \mathrm{tr}(A^2) = \sum_{i=0}^{n-1} \lambda_i^2 = d^2 + \sum_{i=1}^{n-1} \lambda_i^2 \leq (1-\varepsilon)nd + (n-1)\lambda^2. \tag{9.80}$$

Mas então $(n-1)\lambda^2 \geq \varepsilon nd$, de modo que $\lambda^2 \geq \varepsilon d$, e portanto $\lambda \geq \sqrt{\varepsilon}\sqrt{d}$. Concluímos que $\lambda = \Omega(\sqrt{d})$, como queríamos. $\square$

(n, d, λ)-grafos podem ser úteis em algoritmos probabilísticos que fazem passeios aleatórios em certos grafos. Existem construções determinísticas de (n, d, λ)-grafos, o que permite que eles sejam utilizados em tais algoritmos. Por exemplo, Lubotzky, Phillips e Sarnak [13] e Margulis [14] mostraram como construir deterministicamente os chamados grafos de Ramanujan, que têm $d = p+1$, onde p é um primo e $\lambda = 2\sqrt{d-1}$.

9.3 Um lema de Lindsey

Seja $H = (h_{ij})_{1\leq i,j\leq n}$ uma matriz com $h_{ij} \in \{-1, +1\}$ para todo i, j. Se as linhas de H são mutuamente ortogonais, então dizemos que H é uma *matriz de Hadamard*. Note então que $HH^T = nI$. Segue que as colunas de H também são mutuamente ortogonais. De fato, como $HH^T = nI$, então $H^{-1} = \frac{1}{n}H^T$. Logo, $I = H^{-1}H = \frac{1}{n}H^T H$, isto é, $H^T H = nI$, como queríamos.

Dados inteiros $1 \leq a, b \leq n$, defina

$$\mathrm{disc}(H; a, b) := \max\left\{\left|\textstyle\sum_{i\in I, j\in J} h_{ij}\right| : I \in \binom{[n]}{a}, J \in \binom{[n]}{b}\right\}. \tag{9.81}$$

É evidente que $0 \leq \mathrm{disc}(H; a, b) \leq ab$.

Teorema 9.8 (Lindsey)**.** *Seja H uma matriz de Hadamard $n \times n$ e sejam $1 \leq a, b \leq n$ inteiros. Então*

$$\mathrm{disc}(H; a, b) \leq \sqrt{abn}. \tag{9.82}$$

Demonstração. Sejam $\underline{v}_1, \ldots, \underline{v}_n$ as linhas de H e $\underline{\chi}_J \in \mathbb{R}^n$ o vetor característico de J. Dado $i \in I$, temos

$$\sum_{j\in J} h_{ij} = \langle \underline{v}_i, \underline{\chi}_J\rangle, \tag{9.83}$$

de modo que

$$\sum_{i\in I, j\in J} h_{ij} = \sum_{i\in I}\sum_{j\in J} h_{ij} = \sum_{i\in I}\langle \underline{v}_i, \underline{\chi}_J\rangle = \left\langle \textstyle\sum_{i\in I}\underline{v}_i, \underline{\chi}_J\right\rangle. \tag{9.84}$$

Por Cauchy-Schwarz, temos

$$\left|\textstyle\sum_{i\in I, j\in J} h_{ij}\right| = \left|\left\langle \textstyle\sum_{i\in I}\underline{v}_i, \underline{\chi}_J\right\rangle\right| \leq \left\|\textstyle\sum_{i\in I}\underline{v}_i\right\| \cdot \|\underline{\chi}_J\| = \left\|\textstyle\sum_{i\in I}\underline{v}_i\right\|\sqrt{b}. \tag{9.85}$$

Ademais, como as linhas de H são mutuamente ortogonais,

$$\left\|\sum_{i\in I}\underline{v}_i\right\| = \sqrt{\left\langle \sum_{i\in I}\underline{v}_i, \sum_{i\in I}\underline{v}_i\right\rangle} = \sqrt{\sum_{i\in I}\langle \underline{v}_i, \underline{v}_i\rangle} = \sqrt{an}, \tag{9.86}$$

já que $\langle \underline{v}_i, \underline{v}_i\rangle = n$ para todo i.

Assim,

$$\left|\sum_{i\in I, j\in J} h_{ij}\right| \leq \left\|\sum_{i\in I}\underline{v}_i\right\|\sqrt{b} = \sqrt{abn}, \tag{9.87}$$

como queríamos. □

Seja H uma matriz de Hadamard $n \times n$. Defina a *discrepância de H* como

$$\operatorname{disc}(H) := \max_{1\leq a,b\leq n} \operatorname{disc}(H; a, b). \tag{9.88}$$

Intuitivamente, a discrepância de H é uma medida de "quão desbalanceada" pode estar uma submatriz de H, sendo que o nível de "desequilíbrio" de uma matriz aumenta se ela tem muito mais 1's do que -1's, ou vice-versa.

Corolário 9.8.1. *Seja H uma matriz de Hadamard $n \times n$. Então*

$$\operatorname{disc}(H) \leq n^{3/2}. \tag{9.89}$$

Demonstração. Imediato de (9.82). □

9.4 Uniformidade em (n, d, λ)-grafos

Seja $G = (V, E)$ um grafo e sejam $U, W \subseteq V$. Até agora, consideramos apenas $e(U, W)$ para $U \cap W = \emptyset$. Vamos estender essa notação. Mesmo que $U \cap W \neq \emptyset$, o valor $e(U, W)$ continuará bem definido. Porém, arestas com as duas pontas em $U \cap W$ são contadas duas vezes. Isto é,

$$e(U, W) = e(U \setminus W, W \setminus U) + 2e(U \cap W). \tag{9.90}$$

Observe ainda que, se G é d-regular com n vértices, então a densidade de G é $[nd/2][\binom{n}{2}/2] = \frac{d}{n-1}$, isto é, $\frac{d}{n}$ é "basicamente" a densidade de G, para n grande.

Teorema 9.9. *Seja $G = (V, E)$ um grafo d-regular com n vértices. Sejam $\lambda_0 \geq \cdots \geq \lambda_{n-1}$ os autovalores de G e $\lambda := \max_{1\leq i\leq n} |\lambda_i|$. Para quaisquer $U, W \subseteq V$, temos*

$$\left|e(U, W) - \tfrac{d}{n}|U| \cdot |W|\right| \leq \lambda\sqrt{|U| \cdot |W| \left(1 - \tfrac{|U|}{n}\right)\left(1 - \tfrac{|W|}{n}\right)}. \tag{9.91}$$

Demonstração. Sejam $\underline{x}_0, \ldots, \underline{x}_{n-1}$ os autovetores de G, sendo que o autovalor associado a $\underline{x}_i$ é λ_i para todo i. Como já discutido anteriormente, podemos supor que $B := \{\underline{x}_0, \ldots, \underline{x}_{n-1}\}$ é uma base ortonormal de $\mathbb{R}^n$.

Fixe $U, W \subseteq V$. Como sempre, tome $u := |U|$ e $w := |W|$.

Seja A a matriz de adjacência de G. Como já mostramos em (9.74), temos $A = \sum_{i=0}^{n-1} \lambda_i \underline{x}_i \underline{x}_i^T$. Tome

$$A_0 := \lambda_0 \underline{x}_0 \underline{x}_0^T \tag{9.92}$$

e

$$E := \sum_{i=1}^{n-1} \lambda_i \underline{x}_i \underline{x}_i^T, \tag{9.93}$$

de modo que $A = A_0 + E$.

Como λ_0 é o maior autovalor, podemos pensar que A_0 é o "termo principal" e que E é o "termo de erro" de A. Da mesma forma, podemos pensar que o "termo principal" de $e(U, W)$ é $\frac{d}{n} uw$ e que seu "termo de erro" é $\lambda\sqrt{uw(1 - u/n)(1 - w/n)}$.

Sejam $\underline{\chi}_U$ e $\underline{\chi}_W$ os vetores característicos de U e W, respectivamente. Podemos escrever $\underline{\chi}_U$ e $\underline{\chi}_W$ na base B da seguinte maneira:

$$\underline{\chi}_U = \sum_{i=0}^{n-1} \alpha_i \underline{x}_i, \tag{9.94}$$

onde, para todo i,

$$\alpha_i := \langle \underline{\chi}_U, \underline{x}_i \rangle = \underline{\chi}_U^T \underline{x}_i, \tag{9.95}$$

e

$$\underline{\chi}_W = \sum_{i=0}^{n-1} \beta_i \underline{x}_i, \tag{9.96}$$

onde, para todo i,

$$\beta_i := \langle \underline{\chi}_W, \underline{x}_i \rangle = \underline{\chi}_W^T \underline{x}_i. \tag{9.97}$$

Não é difícil observar que

$$e(U, W) = \underline{\chi}_U^T A \underline{\chi}_W. \tag{9.98}$$

De fato, o produto $A\underline{\chi}_W$ é um vetor em $\mathbb{R}^n$ cujo i-ésimo componente conta o número de arestas iw, para algum $w \in W$. Multiplicando este vetor por $\underline{\chi}_U^T$ à esquerda, obteremos o número de arestas que ligam um vértice de U a um vértice de W. Note ainda que, como requer a definição de $e(U, W)$, as arestas de $G[U \cap W]$ são contadas duas vezes: se $uw \in E$ e $u, w \in U \cap W$, então a aresta uw é contabilizada tanto no u-ésimo componente de $A\underline{\chi}_W$ quanto no w-ésimo.

Temos então

$$e(U, W) = \underline{\chi}_U^T A_0 \underline{\chi}_W + \underline{\chi}_U^T E \underline{\chi}_W. \tag{9.99}$$

Primeiro vamos mostrar o "termo principal", isto é, que $\underline{\chi}_U^T A_0 \underline{\chi}_W = \frac{d}{n} uw$. Combinando (9.92) com (9.94) e (9.96), e usando o fato de que B é uma base ortonormal, temos

$$\begin{aligned} \underline{\chi}_U^T A_0 \underline{\chi}_W &= \Big(\sum_{i=0}^{n-1} \alpha_i \underline{x}_i^T \Big) \lambda_0 \underline{x}_0 \underline{x}_0^T \Big(\sum_{j=0}^{n-1} \beta_j \underline{x}_j \Big) \\ &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \lambda_0 \alpha_i \beta_j \big(\underline{x}_i^T \underline{x}_0 \big) \big(\underline{x}_0^T \underline{x}_j \big) \\ &= \lambda_0 \alpha_0 \beta_0. \end{aligned} \tag{9.100}$$

Como G é d-regular, temos $\lambda_0 = d$ e $\underline{x}_0^T = (1, \ldots, 1)/\sqrt{n}$, como já discutimos em (9.77). Assim, $\alpha_0 = \underline{\chi}_U^T \underline{x}_0 = u/\sqrt{n}$ e $\beta_0 = \underline{\chi}_W^T \underline{x}_0 = w/\sqrt{n}$. Obtemos então o "termo principal":

$$\underline{\chi}_U^T A_0 \underline{\chi}_W = \lambda_0 \alpha_0 \beta_0 = \frac{d}{n} uw. \tag{9.101}$$

Vamos agora ao "termo de erro". Novamente, podemos combinar (9.92) com (9.94) e (9.96) e usar o fato de que B é uma base ortonormal para obter

$$\begin{aligned} \underline{\chi}_U^T E \underline{\chi}_W &= \Big(\sum_{i=0}^{n-1} \alpha_i \underline{x}_i^T \Big) \Big(\sum_{j=1}^{n-1} \lambda_j \underline{x}_j \underline{x}_j^T \Big) \Big(\sum_{k=0}^{n-1} \beta_k \underline{x}_k \Big) \\ &= \sum_{i=0}^{n-1} \sum_{j=1}^{n-1} \sum_{k=0}^{n-1} \lambda_j \alpha_i \beta_k \big(\underline{x}_i^T \underline{x}_j \big) \big(\underline{x}_j^T \underline{x}_k \big) \\ &= \sum_{j=1}^{n-1} \lambda_j \alpha_j \beta_j. \end{aligned} \tag{9.102}$$

Utilizando a desigualdade triangular, obtemos

$$\big| \underline{\chi}_U^T E \underline{\chi}_W \big| \le \sum_{j=1}^{n-1} |\lambda_j| \cdot |\alpha_j \beta_j| \le \lambda \sum_{j=1}^{n-1} |\alpha_j| \cdot |\beta_j|. \tag{9.103}$$

Considere os vetores $\underline{\alpha}' = (\alpha_j')_{j=1}^{n-1}$ e $\underline{\beta}' = (\beta_j')_{j=1}^{n-1}$, onde $\alpha_j' := |\alpha_j|$ e $\beta_j' := |\beta_j|$ para todo j, e aplique a desigualdade de Cauchy-Schwarz para obter

$$\sum_{j=1}^{n-1} |\alpha_j| \cdot |\beta_j| = \big| \langle \underline{\alpha}', \underline{\beta}' \rangle \big| \le \|\underline{\alpha}'\| \cdot \|\underline{\beta}'\| = \sqrt{\textstyle\sum_{j=1}^{n-1} |\alpha_j|^2} \cdot \sqrt{\textstyle\sum_{j=1}^{n-1} |\beta_j|^2}. \tag{9.104}$$

Prosseguindo (9.103), temos

$$\big| \underline{\chi}_U^T E \underline{\chi}_W \big| \le \lambda \sum_{j=1}^{n-1} |\alpha_j| \cdot |\beta_j| \le \lambda \sqrt{\Big(\textstyle\sum_{j=1}^{n-1} \alpha_j^2 \Big) \Big(\textstyle\sum_{j=1}^{n-1} \beta_j^2 \Big)}. \tag{9.105}$$

Por outro lado, temos

$$\sum_{j=1}^{n-1} \alpha_j^2 = \|\underline{\chi}_U\|^2 - \alpha_0^2 = u - u^2/n = u(1 - u/n) \tag{9.106}$$

e

$$\sum_{j=1}^{n-1} \beta_j^2 = \|\underline{\chi}_W\|^2 - \beta_0^2 = w - w^2/n = w(1 - w/n). \tag{9.107}$$

Então

$$\begin{aligned} \left|\underline{\chi}_U^T E \underline{\chi}_W\right| &\leq \lambda \sqrt{\left(\sum_{j=1}^{n-1} \alpha_j^2\right)\left(\sum_{j=1}^{n-1} \beta_j^2\right)} \\ &= \lambda \sqrt{uw\left(1 - \tfrac{u}{n}\right)\left(1 - \tfrac{w}{n}\right)}. \end{aligned} \tag{9.108}$$

Combinando (9.99) com (9.101) e (9.108), obtemos que

$$e(U, W) = \tfrac{d}{n} uw + O_1\left(\lambda \sqrt{uw\left(1 - \tfrac{u}{n}\right)\left(1 - \tfrac{w}{n}\right)}\right), \tag{9.109}$$

como queríamos. □

Corolário 9.9.1. *Seja $G = (V, E)$ um (n, d, λ)-grafo. Sejam $U, W \subseteq V$ com $U \cap W = \emptyset$. Tome $u := |U|$ e $w := |W|$. Então*

$$e(U, W) = \frac{d}{n} uw + O_1\big(\lambda \sqrt{uw}\big). \tag{9.110}$$

Demonstração. Imediato do teorema 9.9. □

Se combinarmos o corolário 9.9.1 acima com a proposição 9.7, obtemos que qualquer (n, d, λ)-grafo com $\lambda = O\big(\sqrt{d}\big)$ é (p, A)-uniforme com $p = d/n$ e $A = O(1)$.

9.5 Conexidade, aresta-conexidade e emparelhamentos

Sejam $k \geq 2$ um inteiro e G um grafo com n vértices. Dizemos que G é *k-conexo* se $n \geq k+1$ e a remoção de quaisquer $k-1$ vértices não desconecta G.

Por exemplo, tome $n := k + 1$. Queremos que G seja k-conexo. Então a remoção de quaisquer $k - 1$ vértices deve deixar G conexo. Porém, tal remoção deixa G com apenas 2 vértices. Concluímos que deve existir uma aresta entre cada par de vértices de G, ou seja, $G = K_n$.

O teorema de Menger [16] afirma que um grafo G é k-conexo se, e somente se, para quaisquer pares de vértices u e w de G, existem k caminhos internamente disjuntos ligando u a w, onde dizemos que dois caminhos P e P' de u a w são internamente disjuntos se $V(P) \cap V(P') = \{u, w\}$.

A definição de 1-conexidade é a mesma de conexidade. Definimos ainda que todo grafo é 0-conexo.

Pomos

$$\kappa(G) := \max\{k : G \text{ é } k\text{-conexo}\}. \tag{9.111}$$

Teorema 9.10. *Seja G um (n, d, λ)-grafo com $1 \leq d \leq n/2$. Então*

$$\kappa(G) \geq d - \left\lceil 36\frac{\lambda^2}{d} \right\rceil. \tag{9.112}$$

Demonstração. Queremos que $n \geq d - \lceil 36\lambda^2/d \rceil + 1$. Como G tem pelo menos dois vértices e $d \geq 1$, então $n \geq 2d \geq d+1 \geq d+1-\lceil 36\lambda^2/d \rceil$, como precisamos. Ademais, podemos supor, sem perda de generalidade, que

$$\lambda \leq d/6, \tag{9.113}$$

pois caso contrário $d - \lceil 36\lambda^2/d \rceil \leq d - 36\lambda^2/d < 0$.

Suponha que existe $S \subseteq V_G$ com

$$|S| < d - \lceil 36\lambda^2/d \rceil \leq d - 36\lambda^2/d \leq d \tag{9.114}$$

tal que $G - S$ é desconexo e vamos ver que isso é impossível. Para tanto, vamos aplicar o teorema 9.9 repetidas vezes.

Sejam U um conjunto de vértices de uma menor componente de $G - S$ e $W := V_G \setminus (U \cup S)$. Como U é uma menor componente de $G - S$, temos

$$|W| \geq \frac{n - |S|}{2} \geq \frac{n-d}{2} \geq \frac{n}{4}, \tag{9.115}$$

pois $d \leq n/2$.

Seja $v \in U$. Como U é um componente de $G - S$, todos os vizinhos de v estão em $U \cup S$, de modo que $|U| + |S| = |U \cup S| \geq d + 1$. Assim, por (9.114),

$$|U| \geq d + 1 - |S| > 36\lambda^2/d. \tag{9.116}$$

Sejam $u := |U|$ e $w := |W|$. Como $e(U, W) = 0$, pelo teorema 9.9, temos $0 \geq e(U, W) \geq \frac{d}{n}uw - \lambda\sqrt{uw}$, de modo que $\frac{d}{n}uw \leq \lambda\sqrt{uw}$. Mas então

$$u \leq \frac{\lambda^2 n^2}{d^2 w} = \frac{\lambda}{d} \cdot \frac{n}{w} \cdot \frac{\lambda n}{d} \leq \frac{1}{6} \cdot 4 \cdot \frac{\lambda n}{d} < \frac{\lambda n}{d}, \tag{9.117}$$

onde utilizamos (9.113) e (9.115).

Novamente pelo teorema 9.9, temos

$$e(U) = \frac{1}{2}e(U, U) \leq \frac{1}{2}\frac{d}{n}u^2 + \frac{\lambda u}{2} < \frac{d}{2n}\frac{\lambda n}{d}u + \frac{\lambda u}{2} = \lambda u, \tag{9.118}$$

onde utilizamos (9.117).

Observe que $du = \sum_{v\in U} d(v) = 2e(U) + e(U,S)$, de modo que

$$e(U,S) = du - 2e(U) > du - 2\lambda u = (d-2\lambda)u \geq \frac{2}{3}du, \tag{9.119}$$

onde utilizamos (9.118) e (9.113).

Finalmente, tomando $s := |S|$ e usando (9.116), (9.114) e o fato de que $d \leq n/2$, temos, pelo teorema 9.9, que

$$\begin{aligned} e(U,S) &\leq \frac{d}{n}us + \lambda\sqrt{us} < \frac{d}{n}ud + \lambda\sqrt{ud} \\ &\leq \frac{d}{2}u + \frac{\lambda\sqrt{d}}{\sqrt{u}}u \leq \frac{d}{2}u + \frac{\lambda\sqrt{d}}{6\lambda/\sqrt{d}}u \\ &= \frac{d}{2}u + \frac{d}{6}u = \frac{2}{3}du. \end{aligned} \tag{9.120}$$

Obtemos então que $e(U,S) \geq \frac{2}{3}du$ em (9.119) e $e(U,S) < \frac{2}{3}du$ em (9.120). Isso é um absurdo! Portanto é impossível que tal conjunto S exista, e estamos feitos. □

Note que, se G é um (n,d,λ)-grafo, então $\kappa(G) \leq d$, pois a remoção de d vizinhos de qualquer vértice desconecta G. O teorema 9.10 acima implica que, se λ é da ordem de $\sqrt{d}$, isto é, se $\lambda = O(\sqrt{d})$, então $\kappa(G) = d - O(1)$:

Corolário 9.10.1. *Seja G um (n,d,λ)-grafo com $1 \leq d \leq n/2$. Se $\lambda = O(\sqrt{d})$, então $\kappa(G) = d - O(1) \sim d$.*

Demonstração. Imediato do teorema 9.10. □

Exercício 10. *Fixe $\varepsilon > 0$. Seja $p := p(n) := (1+\varepsilon)\frac{\ln n}{n}$. Prove que quase todo $G(n,p)$ é $\lceil \ln n \rceil$-conexo, isto é, que*

$$\mathbb{P}\Big[\kappa\big(G(n,p)\big) \geq \lceil \ln n \rceil\Big] = 1 - o(1). \tag{9.121}$$

É interessante notar que, quase sempre, $\kappa\big(G(n,p)\big) = \delta\big(G(n,p)\big)$, como ilustra o seguinte fenômeno.

Fixe n. Defina $N := \binom{n}{2}$ e seja $(G_i)_{i=0}^N$ uma seqüência de grafos tais que G_0 é o $\overline{K}_n$ e $G_{i+1} = G_i + e$ para alguma aresta $e \notin E(G_i)$. Se tomarmos a aresta e equiprovavelmente dentre todas as arestas fora de $E(G_i)$, obtemos um experimento aleatório cujo espaço amostral é o conjunto de todas as seqüências $(G_i)_{i=0}^N$, todas equiprováveis. Note que o espaço amostral tem cardinalidade $N!$.

Defina as variáveis aleatórias

$$\tau(\text{conexidade}) := \min\{i : G_i \text{ é conexo}\} \tag{9.122}$$

e

$$\tau(\delta > 0) := \min\{i : \delta(G_i) > 0\}. \tag{9.123}$$

Note que τ(conexidade) é o primeiro instante em que G_i é conexo e $\tau(\delta > 0)$ é o primeiro instante em que G_i não tem mais vértices isolados. É evidente que

$$\tau(\text{conexidade}) \geq \tau(\delta > 0). \tag{9.124}$$

O seguinte resultado mostra que, quase sempre, o instante em que G_i passa a ser conexo é justamente o instante em que G_i deixa de ter vértices isolados:

Teorema 9.11. *Temos*

$$\lim_{n\to\infty} \mathbb{P}\big[\tau(\textit{conexidade}) = \tau(\delta > 0)\big] = 1. \tag{9.125}$$

Em geral, para qualquer k fixo, temos

$$\lim_{n\to\infty} \mathbb{P}\big[\tau(k\textit{-conexidade}) = \tau(\delta \geq k)\big] = 1. \tag{9.126}$$

Exercício 11. *Prove o seguinte teorema:*

Seja G um (n, d, λ)-grafo com $d - \lambda \geq 2$. Então G é d-aresta-conexo, isto é, $G - F$ é conexo para qualquer $F \subseteq E(G)$ com $|F| < d$. Conseqüentemente, se n é par, então G tem um emparelhamento perfeito.

[*Dica:* Use o teorema de Tutte e a d-aresta-conexidade de G.]

10 O método das diferenças limitadas

Vamos estudar agora o método das diferenças limitadas (= *bounded differences method*). Nesta discussão inicial, vamos apenas enunciar alguns resultados. Suas provas virão mais adiante.

Lema 10.1 (Chernoff)**.** *Sejam $X_1, \dots, X_n$ variáveis aleatórias de Bernoulli, mutuamente independentes, com $\mathbb{P}[X_k = 1] = p$ para todo k. Tome $X := \sum_{k=1}^n X_k$. Para qualquer $t > 0$, temos*

$$\mathbb{P}\Big[\big|X - \mathbb{E}\,[X]\big| \geq t\Big] \leq 2\exp\{-2t^2/n\}. \tag{10.1}$$

Observe que as cotas (8.40) e (8.42) são melhores que a cota (10.1) de Chernoff se $p = o(1)$. Porém, para p constante, a cota (10.1) de Chernoff é melhor.

No que se segue, vamos trabalhar no seguinte contexto. Sejam $X_1, \dots, X_n$ variáveis aleatórias mutuamente independentes, com $X_k \in A_k$ para todo k, onde supomos por conveniência que A_k é finito. Seja $f : \prod_{k=1}^n A_k \to \mathbb{R}$ uma função que associa um número real a cada possível resultado dos experimentos aleatórios relativos a $X_1, \dots, X_n$. Dizemos que f é (c_k)-*Lipschitz* se existe um vetor $(c_k)_{k=1}^n$ tal que, para quaisquer $\underline{x}, \underline{y} \in \prod_{k=1}^n A_k$, vale que

$$\big|f(\underline{x}) - f(\underline{y})\big| \leq c_k \tag{10.2}$$

sempre que $\underline{x}$ e $\underline{y}$ diferirem apenas na k-ésima coordenada.

Lema 10.2 (McDiarmid [15]). *Sejam $X_1, \ldots, X_n$ e f como acima. Defina a variável aleatória $Y := f(X_1, \ldots, X_n)$. Para todo $t \geq 0$, temos*

$$\mathbb{P}\big[Y \geq \mathbb{E}\,[Y] + t\big] \leq \exp\{-2t^2/\textstyle\sum_{k=1}^{n} c_k^2\} \tag{10.3}$$

e

$$\mathbb{P}\big[Y \leq \mathbb{E}\,[Y] - t\big] \leq] \exp\{-2t^2/\textstyle\sum_{k=1}^{n} c_k^2\}. \tag{10.4}$$

Deixamos para mais adiante a demonstração desse lema. Observe que o lema 10.1 de Chernoff é um corolário desse lema 10.2 de McDiarmid: basta tomar $c_k := 1$ para todo k e $f(X_1, \ldots, X_n) = \sum_{k=1}^{n} X_k$.

10.1 O número cromático de quase todo $G(n, p)$

Vamos aplicar o método das diferenças limitadas para provar um resultado sobre o número cromático de quase todo $G(n, p)$, no caso em que p é constante.

Sejam $0 < p < 1$, $q := 1 - p$ e $b := 1/q$. Pode-se mostrar que existe uma função $r := r_p(n)$ tal que

$$\lim_{n\to\infty} \mathbb{P}\Big[r - 1 \leq \alpha\big(G(n, p)\big) \leq r\Big] = 1. \tag{10.5}$$

Ademais, sabe-se que $r_p(n) \sim 2\log_b n$. Utilizando (3.2), obtemos que quase sempre temos

$$\chi\big(G(n, p)\big) \geq \frac{n}{2\log_b n} \tag{10.6}$$

Para grafos em geral, a cota (3.2) pode ser muito fraca, isto é, $\chi(G)$ pode ser arbitrariamente superior a $n/\alpha(G)$. Porém, em 1975, provou-se que esse não é o caso para grafos aleatórios: utilizando o algoritmo guloso, obtemos que quase sempre

$$\frac{\big(1 + o(1)\big)n}{2\log_b n} \leq \chi\big(G(n, p)\big) \leq \frac{\big(1 + o(1)\big)n}{\log_b n}. \tag{10.7}$$

Porém, nada se sabia a respeito da distribuição de $\chi\big(G(n, p)\big)$ dentro desse intervalo: podia ser que $\chi\big(G(n, p)\big)$ fosse "altamente concentrada" em torno de um valor desse intervalo, ou podia ser que $\chi\big(G(n, p)\big)$ fosse equiprovável dentro desse intervalo. Utilizando-se o método das diferenças limitadas, pode-se mostrar que o que ocorre é o primeiro caso, como veremos a seguir.

Vamos dizer que uma variável aleatória $X := X\big(G(n, p)\big)$ é *concentrada com largura* $s = s(n, p)$ se existe $u = u(n, p)$ tal que

$$\lim_{n\to\infty} \mathbb{P}\Big[u \leq X\big(G(n, p)\big) \leq u + s\Big] = 1. \tag{10.8}$$

Por exemplo, $\alpha\big(G(n, p)\big)$ é concentrada com largura 1 e já vimos que $\chi\big(G(n, p)\big)$ é concentrada com largura $n/(2\log_b n)$.

Teorema 10.3 (Shamir e Spencer [18])**.** *Para qualquer* $0 < p = p(n) < 1$*, temos que* $\chi\big(G(n,p)\big)$ *é concentrado com largura* $\psi\sqrt{n}$ *para qualquer função* $\psi = \psi(n)$ *tal que* $\lim_{n\to\infty} \psi(n) = \infty$*.*

Note que esse resultado é muito mais forte: para n grande, o intervalo de comprimento $2\psi\sqrt{n}$, comparado ao intervalo de comprimento $n/(2\log_b n)$, é tão pequeno que é um ponto dentro desse intervalo maior. Na verdade, provou-se que esse intervalo é centrado em torno da cota inferior dada por (10.7):

Teorema 10.4 (Bollobás [3])**.** *Quase sempre*

$$\chi\big(G(n,p)\big) = \big(1 + o(1)\big)\frac{n}{2\log_b n}. \tag{10.9}$$

Antes de provar o teorema 10.3, vamos estabelecer alguns lemas.

Lema 10.5. *Seja* $\{F_1, \dots, F_m\}$ *uma partição de* $E(K_n)$ *em* m *partes. Suponha que* f *é uma função que associa um número real a cada grafo e que, para quaisquer* $G, G' \subseteq K_n$ *tais que*

$$E(G) \bigtriangleup E(G') \subseteq F_k \tag{10.10}$$

para algum k*, vale que*

$$\big|f(G) - f(G')\big| \leq 1. \tag{10.11}$$

Então a variável aleatória $Y := f\big(G(n,p)\big)$ *satisfaz*

$$\mathbb{P}\Big[\big|Y - \mathbb{E}\,[Y]\big| \geq t\Big] \leq 2\exp\{-2t^2/m\} \tag{10.12}$$

para todo $t > 0$*.*

Demonstração. É fácil provar esse lema utilizando o lema 10.2 de McDiarmid. Basta tomar $A_k := \mathcal{P}(F_k)$, $X_k := E\big(G(n,p)\big) \cap F_k$ e $c_k := 1$ para todo k. □

Corolário 10.5.1. *Se um parâmetro* f *de grafos é tal que*

$$\big|f(G) - f(G')\big| \leq 1 \tag{10.13}$$

se G *e* G' *diferem apenas nas arestas incidentes a um único vértice, então a variável aleatória* $Y := f\big(G(n,p)\big)$ *satisfaz*

$$\mathbb{P}\Big[\big|Y - \mathbb{E}\,[Y]\big| \geq t\Big] \leq 2\exp\{-2t^2/n\} \tag{10.14}$$

para todo $t > 0$*.*

Demonstração. Considere que $V(K_n) = [n]$. Tome a seguinte partição de $E(K_n)$ em $n-1$ partes. Para cada $2 \leq i \leq n$, tome $B_i := \{ji : j < i\}$. Nossa partição será $\{B_2, \dots, B_n\}$. Precisamos mostrar que tal partição e o parâmetro f satisfazem a hipótese do lema 10.5.

Sejam $G, G' \subseteq K_n$ tais que $E(G) \bigtriangleup E(G') \subseteq B_k$ para algum k. Então G e G' diferem apenas nas arestas incidentes a um único vértice. Mas então, por hipótese, temos $\big|f(G) - f(G')\big| \leq 1$, que é justamente o que é exigido na hipótese do lema 10.5 em (10.11). Agora o corolário segue imediatamente do lema 10.5. □

Podemos finalmente provar o teorema 10.3 de Shamir e Spencer.

Demonstração do teorema 10.3. Sejam $G, G' \subseteq K_n$ grafos tais que G e G' diferem apenas nas arestas incidentes a um único vértice. Não é difícil ver que $|\chi(G) - \chi(G')| \leq 1$. Podemos então aplicar o corolário 10.5.1. Seja $\psi = \psi(n)$ uma função tal que $\lim_{n\to\infty} \psi(n) = \infty$. Tome $t := \psi\sqrt{n}/2$ e $Y := \chi\big(G(n,p)\big)$. Obtemos então

$$\mathbb{P}\big[|Y - \mathbb{E}\,[Y]| \geq \psi\sqrt{n}/2\big] \leq 2\exp\{-\psi^2/2\} = o(1). \tag{10.15}$$

Mas então Y é concentrada com largura $\psi\sqrt{n}$, como queríamos.

□

Exercício 12. *Sejam $0 \leq x_1, \ldots, x_n \leq 1$ número reais. Queremos colocar estes itens em bins de tamanho 1. Seja $B := B(x_1, \ldots, x_n)$ o número de bins necessários. Suponha que $X_1, \ldots, X_n$ são variáveis aleatórias independentes assumindo valores em $[0,1]$. Seja $Y := B(X_1, \ldots, X_n)$. Prove que, para todo $t > 0$,*

$$\mathbb{P}\Big[\big|Y - \mathbb{E}\,[Y]\big| \geq t\Big] \leq 2\exp\{-2t^2/n\}. \tag{10.16}$$

e que Y é concentrada com largura $\psi\sqrt{n}$ para qualquer função $\psi = \psi(n)$ tal que $\lim_{n\to\infty} \psi(n) = \infty$.

10.2 Desigualdades isoperimétricas discretas

Seja S uma região no plano tal que $\mathrm{area}(S) = 1$. Queremos limitar inferiormente $\mathrm{per}(S)$, o perímetro de S. Por exemplo, se S é um círculo de raio r, então $r = 1/\sqrt{\pi}$, de modo que $\mathrm{per}(S) = 2\sqrt{\pi}$. Pode-se mostrar que o formato de S que minimiza sua fronteira é justamente o círculo, de modo que $\mathrm{per}(S) \geq 2\sqrt{\pi}$.

Uma forma de se obter o perímetro de S a partir de sua área é a seguinte. Seja $\varepsilon > 0$. Seja S_ε a região do plano formada por todos os pontos do plano distantes no máximo ε de algum ponto de S. Então

$$\mathrm{per}(S) = \lim_{\varepsilon\to 0} \frac{\mathrm{area}(S_\varepsilon) - \mathrm{area}(S)}{\varepsilon}. \tag{10.17}$$

Por exemplo, se S é um círculo de raio r, temos

$$\begin{aligned} \lim_{\varepsilon\to 0} \frac{\mathrm{area}(S_\varepsilon) - \mathrm{area}(S)}{\varepsilon} &= \lim_{\varepsilon\to 0} \frac{\pi(r+\varepsilon)^2 - \pi r^2}{\varepsilon} \\ &= \lim_{\varepsilon\to 0} 2\pi r + \pi\varepsilon = 2\pi r. \end{aligned} \tag{10.18}$$

Se S é um quadrado de lado l, então

$$\begin{aligned} \lim_{\varepsilon\to 0} \frac{\mathrm{area}(S_\varepsilon) - \mathrm{area}(S)}{\varepsilon} &= \lim_{\varepsilon\to 0} \frac{l^2 + 4l\varepsilon + 4(\pi\varepsilon^2/4) - l^2}{\varepsilon} \\ &= \lim_{\varepsilon\to 0} 4l + \pi\varepsilon = 4l. \end{aligned} \tag{10.19}$$

Seja (V, d) um espaço métrico finito. Isto é, seja V um conjunto finito e d uma função distância. Mais formalmente, seja $d : V^2 \to \mathbb{R}_{\geq 0}$ uma função tal que

$$d(x,x) = 0 \text{ para todo } x \in V, \tag{10.20a}$$

$$d(x,y) = d(y,x) \text{ para todos } x, y \in V, \tag{10.20b}$$

e

$$d(x,z) \leq d(x,y) + d(y,z) \text{ para todos } x, y, z \in V. \tag{10.20c}$$

Sejam $A \subseteq V$ e $t > 0$. Analogamente a S_ε, defina

$$A_{(t)} := \{v \in V : d(v, A) \leq t\}, \tag{10.21}$$

onde $d(v, A) := \min\{d(v, a) : a \in A\}$. É evidente que $A \subseteq A_{(t)}$. Estamos interessados em cotas inferiores para $|A_{(t)}|$, se $|A|$ é fixo.

Seja $(\mathcal{P}_k, c_k)_{k=0}^n$ uma seqüência de pares, onde $\mathcal{P}_k$ é uma partição de V e $c_k \in \mathbb{R}$ para todo k. Suponha que $\mathcal{P}_0 = \{V\}$ é uma partição com um único bloco, que $\mathcal{P}_n = \big\{\{v\} : v \in V\big\}$ é uma partição com blocos unitários, e que $\mathcal{P}_0 \leq \mathcal{P}_1 \leq \cdots \leq \mathcal{P}_n$, onde a relação $\leq$ é definida a seguir.

Dadas $\mathcal{P}, \mathcal{P}'$ partições de V, dizemos que

$$\mathcal{P} \leq \mathcal{P}' \tag{10.22}$$

se, para todo bloco $P' \in \mathcal{P}'$, existe um bloco $P \in \mathcal{P}$ tal que $P' \subseteq P$. Intuitivamente, $\mathcal{P}$ é uma partição mais grossa do que $\mathcal{P}'$, ou seja, $\mathcal{P}'$ é uma partição mais refinada do que $\mathcal{P}$. Ou seja, estamos chamando de "refinar" a operação de quebrar um bloco em vários.

Suponha ainda que os valores c_k são tais que, se $A, B \in \mathcal{P}_k$ e $A, B \subseteq C$ para algum bloco $C \in \mathcal{P}_{k-1}$, então existe uma bijeção $\varphi : A \to B$ tal que

$$d\big(x, \varphi(x)\big) \leq c_k \tag{10.23}$$

para todo $x \in A$.

Neste caso, dizemos que (V, d) tem tamanho de partição $\leq \sum_{k=1}^n c_k^2$. Observe que c_0 não tem nenhum papel nessa definição.

Considere o seguinte exemplo.

Seja Q^n o grafo n-cubo, isto é, $V(Q^n) = \mathcal{P}([n]) = \{0,1\}^n$ e, dados $\underline{x}, \underline{y} \in V(Q^n)$, temos $\{\underline{x}, \underline{y}\} \in E(Q^n)$ se $\underline{x}, \underline{y}$ coincidem em todas as coordenadas exceto uma.

Dados $\underline{x}, \underline{y} \in V(Q^n)$, defina $d(\underline{x}, \underline{y})$ como a distância de Hamming entre $\underline{x}$ e $\underline{y}$, isto é, o número de coordenadas em que $\underline{x}$ e $\underline{y}$ diferem. Note que d é justamente a função distância em Q^n.

Vamos construir partições $\mathcal{P}_k$ para cada k. Para cada $\underline{b} \in \{0,1\}^k$, tome

$$B(\underline{b}) := \{\underline{x} \in V(Q^n) : (x_1, \ldots, x_k) = \underline{b}\}. \tag{10.24}$$

Tome

$$\mathcal{P}_k := \big\{B(\underline{b}) : \underline{b} \in \{0,1\}^k\big\}. \tag{10.25}$$

Observe que $\mathcal{P}_0$ tem um único bloco e $\mathcal{P}_n$ tem blocos unitários. Ademais, é fácil ver que $\mathcal{P}_k$ refina $\mathcal{P}_{k-1}$, isto é, que $\mathcal{P}_{k-1} \leq \mathcal{P}_k$ para $1 \leq k \leq n$.

Sejam $X, Y \in \mathcal{P}_k$ com $X, Y \subseteq Z \in \mathcal{P}_{k-1}$. Então temos $Z = B(\underline{b})$ para algum $b \in \{0,1\}^{k-1}$ e, sem perda de generalidade, $X = B(\underline{b}^{(0)})$ e $Y = B(\underline{b}^{(1)})$, onde $\underline{b}^{(j)} := (b_1, \dots, b_{k-1}, j)$ para $j = 0, 1$. Tome $\varphi : X \to Y$ como sendo

$$\varphi(x_1, \dots, x_{k-1}, 0, x_{k+1}, \dots, x_n) := (x_1, \dots, x_{k-1}, 1, x_{k+1}, \dots, x_n) \tag{10.26}$$

e note que φ satisfaz a propriedade (10.23) desejada se tomarmos $c_k := 1$ para $k = 1, \dots, n$.

Concluímos que o tamanho da partição de Q^n é $\leq \sum_{k=1}^{n} c_k^2 = n$.

Teorema 10.6. *Seja (V, d) um espaço métrico e seja $(\mathcal{P}_k, c_k)_{k=1}^{n}$ como acima. Seja $f : V \to \mathbb{R}$ tal que*

$$\big|f(x) - f(y)\big| \leq d(x, y) \tag{10.27}$$

para quaisquer $x, y \in V$. Seja X uma variável aleatória uniformemente distribuída sobre V. Então, para todo $t > 0$, temos

$$\mathbb{P}\big[f(X) - \mathbb{E}\left[f(X)\right] \geq t\big] \leq \exp\left\{\frac{-2t^2}{\sum_{k=1}^{n} c_k^2}\right\} \tag{10.28}$$

e

$$\mathbb{P}\big[f(X) + \mathbb{E}\left[f(X)\right] \leq -t\big] \leq \exp\left\{\frac{-2t^2}{\sum_{k=1}^{n} c_k^2}\right\}. \tag{10.29}$$

Não vamos provar esse teorema.

Corolário 10.6.1. *Sejam (V, d) um espaço métrico e $(\mathcal{P}_k, c_k)_{k=1}^{n}$ como acima.*

Seja $A \subseteq V$, com $|A|/|V| = \alpha$ para algum $0 < \alpha < 1$. Então, para qualquer $t \geq t_0$, onde

$$t_0 := \sqrt{\tfrac{1}{2}\left(\ln \tfrac{1}{\alpha}\right) \textstyle\sum_{k=1}^{n} c_k^2}, \tag{10.30}$$

temos

$$\frac{|A_{(t)}|}{|V|} \geq 1 - \exp\left\{\frac{-2(t - t_0)^2}{\sum_{k=1}^{n} c_k^2}\right\}. \tag{10.31}$$

Demonstração. Tomamos $f(x) := d(x, A)$ para todo $x \in V$.

Sejam $x, y \in V$. Seja $a \in A$ tal que $d(y, A) = d(y, a)$. Por (10.20c), temos

$$d(x, A) \leq d(x, a) \leq d(x, y) + d(y, a) = d(x, y) + d(y, A), \tag{10.32}$$

de modo que $f(x) - f(y) \leq d(x, y)$. Analogamente, temos $f(y) - f(x) \leq d(y, x) = d(x, y)$, onde utilizamos (10.20b) na última igualdade. Concluímos que

$$|f(x) - f(y)| \leq d(x, y) \tag{10.33}$$

para todo $x, y \in V$.

Seja X uma variável aleatória uniformemente distribuída sobre V. Tome $t_1 := \mathbb{E}[X]$. Pelo teorema 10.6, temos

$$\begin{aligned}\alpha &= \mathbb{P}[f(X) = 0] = \mathbb{P}[f(X) \le t_1 - t_1] \\ &= \mathbb{P}[f(X) - t_1 \le -t_1] \le \exp\left\{\frac{-2t_1^2}{\sum_{k=1}^n c_k^2}\right\}.\end{aligned} \tag{10.34}$$

Tomando o logaritmo natural em (10.34), obtemos que $t_1 \le t_0$.

Pelo teorema 10.6, para todo $t \ge t_0$, temos

$$\begin{aligned}1 - \frac{|A_{(t)}|}{|V|} &= \mathbb{P}[f(X) > t] \le \mathbb{P}[f(X) > t_1 + t - t_0] \\ &\le \mathbb{P}[f(X) - t_1 \ge t - t_0] \le \exp\left\{\frac{-2(t-t_0)^2}{\sum_{k=1}^n c_k^2}\right\},\end{aligned} \tag{10.35}$$

onde utilizamos o fato de que $t_1 - t_0 \le 0$.

Mas isso é justamente o que queríamos. □

Corolário 10.6.2. *Considere o espaço métrico $(V(Q^n), d)$ sobre o n-cubo, como definido acima. Seja $A \subseteq V(Q^n)$, com $|A| = \alpha 2^n$ para algum $0 < \alpha < 1$. Para todo $t > t_0$, onde*

$$t_0 := \sqrt{\tfrac{1}{2}\Big(\ln \tfrac{1}{\alpha}\Big)n}, \tag{10.36}$$

temos

$$\frac{|A_{(t)}|}{2^n} \ge 1 - \exp\Big\{-2(t-t_0)^2/n\Big\}. \tag{10.37}$$

Exercício 13. *Seja $S_n := \{\pi : [n] \to [n] : \pi$ é bijeção$\}$, de modo que $|S_n| = n!$. Crie um grafo $G = (S_n, E)$, onde $\{\pi, \sigma\} \in E$ se $\pi^{-1}\sigma$ é uma transposição, isto é, uma troca de dois elementos (note que $\pi^{-1}\sigma$ é uma transposição se, e somente se, $(\pi^{-1}\sigma)^{-1} = \sigma^{-1}\pi$ é uma transposição).*

Prove que S_n tem diâmetro $n-1$ e tamanho de partição $\le n-1$.

Seja $A \subseteq S_n$ com $|A| = \alpha n!$, onde $0 < \alpha < 1$. Prove que, para qualquer $t \ge t_0$, onde

$$t_0 := \sqrt{\tfrac{1}{2}\Big(\ln \tfrac{1}{\alpha}\Big)(n-1)}, \tag{10.38}$$

temos

$$\frac{|A_{(t)}|}{n!} \ge 1 - \exp\Big\{-2(t-t_0)^2/(n-1)\Big\}. \tag{10.39}$$

10.3 Martingais

Seja $(\Omega, \mathbb{P})$ um espaço de probabilidade finito, ou seja, Ω é um conjunto finito e $\mathbb{P} : \Omega \to [0,1]$ é uma função tal que $\sum_{\omega \in \Omega} \mathbb{P}[w] = 1$.

Seja $\mathcal{P} = \{B_i : i \in I\}$ uma partição de Ω. Seja $X : \Omega \to \mathbb{R}$ uma variável aleatória. Defina a função $\mathbb{E}(X|\mathcal{P}) : \Omega \to \mathbb{R}$, chamada de esperança de X em relação à partição $\mathcal{P}$, como $\omega \in \Omega$

$$\mathbb{E}\left[X|\mathcal{P}\right](\omega) := \mathbb{E}\left[X|B_i\right] \tag{10.40}$$

onde B_i é o bloco de $\mathcal{P}$ que contém ω e

$$\mathbb{E}\left[X|B_i\right] := \sum_{\omega \in B_i} X(\omega)\mathbb{P}[w \mid B_i] \tag{10.41}$$

é a média de X condicionada a B_i. Note que $\mathbb{E}\left[X|\mathcal{P}\right]$ é uma variável aleatória que é constante nos blocos de $\mathcal{P}$. Em geral, se uma variável aleatória $Z : \Omega \to \mathbb{R}$ é constante nos blocos de uma partição $\mathcal{P}$, dizemos que Z é $\mathcal{P}$-*mensurável*.

Sejam $\mathcal{P}_0, \ldots, \mathcal{P}_n$ partições de Ω tais que $\mathcal{P}_0 = \{\Omega\}$ é uma partição com um único bloco, $\mathcal{P}_n = \big\{\{\omega\} : \omega \in \Omega\big\}$ é uma partição com blocos unitários e, ademais, $\mathcal{P}_0 \leq \cdots \leq \mathcal{P}_n$, onde a relação $\leq$ foi definida em (10.22). Uma tal seqüência de partições é chamada de *filtro*.

Fixe um filtro $\mathcal{P}_0, \ldots, \mathcal{P}_n$. Se $X_0, \ldots, X_n$ são variáveis aleatórias tais que

$$\mathbb{E}\left[X_k|\mathcal{P}_{k-1}\right] = X_{k-1} \tag{10.42}$$

para todo $1 \leq k \leq n$, então $X_0, \ldots, X_n$ é chamado de *martingal*.

Um jeito de entender melhor a definição de martingal é a seguinte. Fixado um filtro e uma variável aleatória X_n, obtém-se as outras variáveis aleatórias para que elas satisfaçam (10.42).
[a continuar…]

11 Desigualdade de Janson

Seja $\{B_i : i \in I\}$ uma família de eventos num espaço de probabilidade. Para nós, esses eventos $B_i's$ são os eventos bons. Queremos estimar a probabilidade de que nenhum evento bom ocorra, isto é, $\mathbb{P}\big[\bigcap_{i\in I} B_i^c\big]$. Equivalentemente, queremos estimar a probabilidade do complemento desse evento anterior, ou seja, $\mathbb{P}\big[\bigcup_{i\in I} B_i\big]$.

Se os eventos $\{B_i : i \in I\}$ são mutuamente independentes, então é claro que $\mathbb{P}\big[\bigcap_{i\in I} B_i^c\big] = \prod_{i\in I} \mathbb{P}[B_i^c]$. Porém, isso não vale se os eventos não forem mutuamente independentes. Vamos ver que, se eles são "basicamente" independentes, então a desigualdade de Janson nos garante que $\mathbb{P}\big[\bigcap_{i\in I} B_i^c\big]$ é muito próximo de $\prod_{i\in I} \mathbb{P}[B_i^c]$.

Vamos especificar melhor a situação em que a desigualdade de Janson se aplica. Fixado um conjunto U, sorteamos um $R \subseteq U$, com $\mathbb{P}[u \in R] = p_u$ para todo $u \in U$, e com todos esses eventos independentes. Fixamos também uma família $\mathcal{A} = \{A_i \subseteq U : i \in I\}$. Para cada $i \in I$, tome $B_i := \{R : A_i \subseteq R\}$ como o evento $A_i \subseteq R$. Seja X_i uma variável aleatória indicadora do evento B_i, isto é, vale 1 se B_i ocorre e 0 caso contrário. Tome $X := \sum_{i\in I} X_i$ como a variável aleatória que conta o número de conjuntos de $\mathcal{A}$ englobados pelo subconjunto R sorteado. Note que o evento $\bigcap_{i\in I} B_i^c$ é o mesmo que o evento $X = 0$.

Por exemplo, seja $U := \binom{[n]}{2}$, com $p_e = p$ para todo $e \in U$, e p fixo. Então estamos trabalhando com $G(n,p)$, isto é, o conjunto R sorteado será o conjunto de arestas do grafo aleatório com n vértices. Tome $I := \binom{[n]}{3}$ e, para cada conjunto $i \in I$, tome $A_i := \binom{i}{2}$. Note que $\mathcal{A} := \{A_i : i \in I\}$ é uma família formada pelo conjunto de arestas de cada triângulo do K_n.

Seja $i = \{u,v,w\} \in I$. Então o evento $B_i := \{R : A_i \subseteq R\}$ é o mesmo que o evento "$G(n,p)$ contém o triângulo de vértices $\{u,v,w\}$." Assim, o evento $\bigcap_{i\in I} B_i^c$ é o mesmo que o evento "$G(n,p)$ não contém um triângulo." Observe que os eventos $\{B_i : i \in I\}$ não são mutuamente independentes. De fato, cada evento B_i é independente de vários outros B_j's, mas B_i e B_j são dependentes sempre que $|i \cap j| = 2$.

Voltando ao contexto genérico, sejam $i,j \in I$. Vamos escrever $i \sim j$ se $i \neq j$ e $A_i \cap A_j \neq \emptyset$. Isso define um grafo sobre I. Observe que o evento B_i é dependente de $\{B_j : i \sim j\}$ e independente de $\{B_j : i \not\sim j\}$. Isto é,

$$\mathbb{P}\big[B_i \,\big|\, \text{expressão booleana envolvendo } B_j\text{'s}, i \not\sim j\big] = \mathbb{P}[B_i]. \tag{11.1}$$

Pomos

$$\begin{aligned}\Delta &:= \sum \Big\{\mathbb{P}[B_i \cap B_j] : (i,j), i \sim j\Big\} \\ &= 2\sum \Big\{\mathbb{P}[B_i \cap B_j] : (i,j), i \sim j \text{ e } i < j\Big\},\end{aligned} \tag{11.2}$$

onde estamos supondo a existência de uma ordem total $<$ arbitrária definida sobre I. Podemos ver Δ como uma certa "medida de independência." Ademais, pomos

$$M := \prod_{i\in I} \mathbb{P}[B_i^c] \tag{11.3}$$

e

$$\mu := \mathbb{E}\,[X] = \sum_{i\in I} \mathbb{P}[B_i]. \tag{11.4}$$

Observe que M é o valor de $\mathbb{P}\big[\bigcap_{i\in I} B_i^c\big]$ quando os B_i's são independentes.

Teorema 11.1 (Janson [10])**.** *Sejam $\{B_i : i \in I\}, \Delta, M, \mu$ como acima. Suponha ainda que, para um certo $0 < \varepsilon < 1$, temos $\mathbb{P}[B_i] \leq \varepsilon$ para todo $i \in I$. Então*

$$M \leq \mathbb{P}\Big[\textstyle\bigcap_{i\in I} B_i^c\Big] \leq M \exp\Big\{\frac{1}{1-\varepsilon}\frac{\Delta}{2}\Big\}. \tag{11.5}$$

Ademais,

$$\mathbb{P}\Big[\textstyle\bigcap_{i\in I} B_i^c\Big] \leq \exp\Big\{-\mu + \Delta/2\Big\}. \tag{11.6}$$

Demonstração. Na prova desse teorema, vamos usar as seguintes desigualdades de correlação, que deixaremos sem prova. Para todo $i \in I$ e $J \subseteq I$, temos

$$\mathbb{P}\big[B_i \,\big|\, \textstyle\bigcap_{j\in J} B_j^c\big] \le \mathbb{P}[B_i] \tag{11.7}$$

A desigualdade acima faz sentido: se temos informações de que certos A_j não foram englobados pelo subconjunto R sorteado, isso não pode aumentar a probabilidade de que A_i seja englobado. De maneira semelhante, temos

$$\mathbb{P}\big[B_i \,\big|\, B_k \cap \textstyle\bigcap_{j\in J} B_j^c\big] \le \mathbb{P}[B_i \,|\, B_k], \tag{11.8}$$

para todos $i, k \in I$ e $J \subseteq I$.

Suponha que $I = [m]$.

Da definição de probabilidade condicional, dada por

$$\mathbb{P}[A \,|\, B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}, \tag{11.9}$$

para quaisquer eventos A e B, não é difícil ver que,

$$\mathbb{P}\Big[\textstyle\bigcap_{i\in I} B_i^c\Big] = \mathbb{P}\big[B_1^c\big]\mathbb{P}\big[B_2^c \,\big|\, B_1^c\big]\mathbb{P}\big[B_3^c \,\big|\, B_1^c \cap B_2^c\big] \cdots \mathbb{P}\big[B_m^c \,\big|\, \textstyle\bigcap_{1\le j<m} B_j^c\big]. \tag{11.10}$$

Agora, por (11.7), temos

$$\mathbb{P}\big[B_i^c \,\big|\, \textstyle\bigcap_{1\le j<i} B_j^c\big] = 1 - \mathbb{P}\big[B_i \,\big|\, \textstyle\bigcap_{1\le j<i} B_j^c\big] \ge 1 - \mathbb{P}[B_i] = \mathbb{P}[B_i^c]. \tag{11.11}$$

Mas então, por (11.10),

$$\mathbb{P}\Big[\textstyle\bigcap_{i\in I} B_i^c\Big] \ge \mathbb{P}[B_1^c] \cdots \mathbb{P}[B_m^c] = M, \tag{11.12}$$

que é justamente a primeira desigualdade de (11.5).

Vamos agora mostrar a segunda desigualdade de (11.5). Fixe $1 \le i \le m$. Primeiro vamos mostrar que

$$\mathbb{P}\big[B_i \,\big|\, \textstyle\bigcap_{1\le j<i} B_j^c\big] \ge \mathbb{P}[B_i] - \displaystyle\sum_{j<i,\, i\sim j} \mathbb{P}[B_i \cap B_j]. \tag{11.13}$$

Por simplicidade, suponha que $i \sim j$ para $1 \le j \le d$ e que $i \not\sim j$ para $d < j < i$.

Observe que, para quaisquer eventos A, B e C,

$$\mathbb{P}[A \,|\, B \cap C] = \frac{\mathbb{P}[A \cap B \cap C]}{\mathbb{P}[B \cap C]} \ge \frac{\mathbb{P}[A \cap B \cap C]}{\mathbb{P}[C]} = \mathbb{P}[A \cap B \,|\, C]. \tag{11.14}$$

Tomando $A := B_i$, $B := \bigcap_{j=1}^{d} B_j^c$ e $C := \bigcap_{d<j<i} B_j^c$, temos

$$\begin{aligned}
\mathbb{P}[B_i \,|\, \textstyle\bigcap_{1\le j<i} B_j^c] &= \mathbb{P}[A \,|\, B \cap C] \ge \mathbb{P}[A \cap B \,|\, C] \\
&= \mathbb{P}[A \,|\, C]\mathbb{P}[B \,|\, A \cap C] \\
&= \mathbb{P}\big[B_i \,\big|\, \textstyle\bigcap_{d<j<i} B_j^c\big]\mathbb{P}[B \,|\, A \cap C] \\
&= \mathbb{P}[B_i]\mathbb{P}[B \,|\, A \cap C],
\end{aligned} \tag{11.15}$$

onde usamos, na última igualdade, a equação (11.1) e o fato de que $i \not\sim j$ para todo $d < j < i$.

Mas, usando (11.8),

$$\begin{aligned}
\mathbb{P}[B \mid A \cap C] &= \mathbb{P}\big[\textstyle\bigcap_{1 \le j \le d} B_j^c \,\big|\, B_i \cap \bigcap_{d<j<i} B_j^c\big] \\
&= \mathbb{P}\Big[\big(\textstyle\bigcup_{1 \le j \le d} B_j\big)^c \,\Big|\, B_i \cap \bigcap_{d<j<i} B_j^c\Big] \\
&= 1 - \mathbb{P}\Big[\big(\textstyle\bigcup_{1 \le j \le d} B_j\big) \,\Big|\, B_i \cap \bigcap_{d<j<i} B_j^c\Big] \\
&\ge 1 - \sum_{j=1}^{d} \mathbb{P}\big[B_j \,\big|\, B_i \cap \textstyle\bigcap_{d<j<i} B_j^c\big] \\
&\ge 1 - \sum_{j=1}^{d} \mathbb{P}[B_j \mid B_i].
\end{aligned} \tag{11.16}$$

Assim, de (11.15),

$$\begin{aligned}
\mathbb{P}\big[B_i \,\big|\, \textstyle\bigcap_{1 \le j < i} B_j^c\big] &\ge \mathbb{P}[B_i]\Big(1 - \sum_{j=1}^{d} \mathbb{P}[B_j \mid B_i]\Big) \\
&= \mathbb{P}[B_i] - \sum_{j=1}^{d} \mathbb{P}[B_i \cap B_j].
\end{aligned} \tag{11.17}$$

Mas isso é justamente (11.13).

Complementando (11.13), obtemos

$$\begin{aligned}
\mathbb{P}\big[B_i^c \,\big|\, \textstyle\bigcap_{1 \le j < i} B_j^c\big] &= 1 - \mathbb{P}\big[B_i \,\big|\, \textstyle\bigcap_{1 \le j < i} B_j^c\big] \\
&\le 1 - \mathbb{P}[B_i] + \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j] \\
&= \mathbb{P}[B_i^c] + \frac{\mathbb{P}[B_i^c]}{1 - \mathbb{P}[B_i]} \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j] \\
&\le \mathbb{P}[B_i^c]\Big(1 + \frac{1}{1-\varepsilon} \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j]\Big),
\end{aligned} \tag{11.18}$$

onde usamos o fato de que $\mathbb{P}[B_i] \le \varepsilon$ para todo $i \in I$ na última passagem. Mas então

$$\mathbb{P}\big[B_i^c \,\big|\, \textstyle\bigcap_{1 \le j < i} B_j^c\big] \le \mathbb{P}[B_i^c] \exp\Big\{\frac{1}{1-\varepsilon} \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j]\Big\}, \tag{11.19}$$

Utilizando (11.19) em (11.10), temos

$$\mathbb{P}\Big[\textstyle\bigcap_{i=1}^{m} B_i^c\Big] \le \Big(\prod_{i=1}^{m} \mathbb{P}[B_i^c]\Big) \exp\Big\{\frac{1}{1-\varepsilon} \sum_{i=1}^{m} \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j]\Big\} \tag{11.20}$$

Observe em (11.2) que

$$\sum_{i=1}^{m} \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j] = \Delta/2. \tag{11.21}$$

Mas então

$$\mathbb{P}\Big[\textstyle\bigcap_{i=1}^{m} B_i^c\Big] \le M \exp\Big\{\dfrac{1}{1-\varepsilon}\dfrac{\Delta}{2}\Big\}, \tag{11.22}$$

que é justamente a segunda desigualdade de (11.5).

Resta apenas mostrarmos (11.6). Repetindo algumas passagens de (11.18), obtemos

$$\begin{aligned}\mathbb{P}\big[B_i^c \,\big|\, \textstyle\bigcap_{1\le j<i} B_j^c\big] &\le 1 - \mathbb{P}[B_i] + \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j] \\ &\le \exp\Big\{ - \mathbb{P}[B_i] + \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j]\Big\}.\end{aligned} \tag{11.23}$$

Usando novamente (11.10),

$$\begin{aligned}\mathbb{P}\Big[\textstyle\bigcap_{i=1}^{m} B_i^c\Big] &\le \exp\Big\{ - \sum_{i=1}^{m} \mathbb{P}[B_i] + \sum_{i=1}^{m} \sum_{j<i, i\sim j} \mathbb{P}[B_i \cap B_j]\Big\} \\ &= \exp\Big\{ - \mu + \Delta/2\Big\},\end{aligned} \tag{11.24}$$

como queríamos. □

Façamos algumas observações.

Primeiro note que, se ε é constante e Δ tende a zero, então o lado direito de (11.5) tende a M.

Observe ainda que, para qualquer $i \in I$, temos $\mathbb{P}[B_i^c] = 1 - \mathbb{P}[B_i] \le e^{-\mathbb{P}[B_i]}$, de modo que $M = \prod_{i\in I} \mathbb{P}[B_i^c] \le \exp\big\{ - \sum_{i\in I} \mathbb{P}[B_i]\big\} = e^{-\mu}$.

Suponha que estamos numa situação em que $\varepsilon = o(1)$ e $\varepsilon\mu = o(1)$, para algum n tal que $n \to \infty$. Então $M \sim e^{-\mu}$. De fato, temos que existe um $x_0 > 0$ tal que

$$\exp\{-x - x^2\} \le 1 - x \le \exp\{-x\} \tag{11.25}$$

para todo $0 < x \le x_0$.

Como $\mathbb{P}[B_i] \le \varepsilon \to 0$, existe um n_0 a partir do qual $\varepsilon \le x_0$ para todo $n \ge n_0$. Tomando $\mathbb{P}[B_i]$ como x em (11.25) para todo $i \in I$ e multiplicando tudo, obtemos

$$\prod_{i\in I} \exp\big\{ - \mathbb{P}[B_i] - \mathbb{P}[B_i]^2\big\} \le \prod_{i\in I} \mathbb{P}[B_i^c] \le e^{-\mu}. \tag{11.26}$$

Mas $\sum_{i\in I} \mathbb{P}[B_i]^2 \le \varepsilon \sum_{i\in I} \mathbb{P}[B_i] = \varepsilon\mu \to 0$. Como essa somatória é o segundo termo do expoente de e na primeira inequação da desigualdade sanduíche (11.26), obtemos que $M \sim e^{-\mu}$.

Se, além de $\varepsilon = o(1)$ e $\varepsilon\mu = o(1)$, tivermos que $\Delta = o(1)$, então (11.5) implica que

$$\mathbb{P}\big[\bigcap_{i\in I} B_i^c\big] \sim M. \tag{11.27}$$

Retomando o exemplo dado acima, em que $U = \binom{[n]}{2}$, isto é, estamos trabalhando no $G(n,p)$, suponha que $p = c/n$ para alguma constante c. Temos $\mathbb{P}[B_i] = p^3 = c^3/n^3 = o(1)$, $\mu = \binom{n}{3}p^3 \sim (n^3/3!)(c^3/n^3) = c^3/6$, de modo que $\varepsilon\mu \sim c^6/(6n^3) = o(1)$. Note ainda que

$$\Delta = \binom{n}{3} 3(n-3)p^5 \leq n^4 p^5/2 = c^5/2n = o(1). \tag{11.28}$$

Por (11.27), temos

$$\mathbb{P}\big[\bigcap_{i\in I} B_i^c\big] \sim M \sim e^{-\mu} \sim e^{-c^3/6}. \tag{11.29}$$

Em outras palavras, a probabilidade de um $G(n,p)$, com $p = c/n$, ser livre de triângulos é, assintoticamente, $e^{-c^3/6}$.

Exercício 14. *Seja G um grafo. Denote por $\chi_3(G)$ o número cromático para triângulos de G, ou seja, o menor número de cores necessário para se colorir os vértices de G evitando triângulos monocromáticos. Prove que existe um grafo G com $\omega(G) = 3$ e $\chi_3(G) \geq 2005$.*

Observe que, se $\Delta \geq 2\mu$, então a cota (11.6) do teorema 11.1 é vazia.

Teorema 11.2. *Sejam $\{B_i : i \in I\}, \Delta, \mu, \varepsilon$ como na hipótese do teorema 11.1. Suponha, adicionalmente, que $\Delta \geq \mu$. Então*

$$\mathbb{P}\Big[\bigcap_{i\in I} B_i^c\Big] \leq \exp\Big\{-\frac{\mu^2}{2\Delta}\Big\}. \tag{11.30}$$

Demonstração. De (11.6), temos

$$-\ln \mathbb{P}\Big[\bigcap_{i\in I} B_i^c\Big] \geq \mu - \Delta/2 = \sum_{i\in I}\mathbb{P}[B_i] - \frac{1}{2}\sum_{i\sim j}\mathbb{P}[B_i \cap B_j]. \tag{11.31}$$

Claramente tal desigualdade vale para todo $S \subseteq I$:

$$-\ln \mathbb{P}\Big[\bigcap_{i\in S} B_i^c\Big] \geq \sum_{i\in S}\mathbb{P}[B_i] - \frac{1}{2}\sum_{i,j\in S: i\sim j}\mathbb{P}[B_i \cap B_j]. \tag{11.32}$$

Queremos encontrar um subconjunto $S \subseteq I$ que maximiza o lado direito de (11.32). Vamos sortear tal S: para cada $i \in I$, independentemente, colocamos i em S com probabilidade $p := \mu/\Delta \leq 1$. Vamos denotar a esperança nesse espaço de probabilidade por $\mathbb{E}_S$. De (11.32), temos

$$\mathbb{E}_S\left[-\ln \mathbb{P}\Big[\bigcap_{i\in S} B_i^c\Big]\right] \geq \mathbb{E}_S\left[\sum_{i\in S}\mathbb{P}[B_i] - \frac{1}{2}\sum_{i,j\in S: i\sim j}\mathbb{P}[B_i \cap B_j]\right]. \tag{11.33}$$

Mas $\mathbb{E}_S\left[\sum_{i\in S}\mathbb{P}[B_i]\right] = p\mu$ e $\mathbb{E}_S\left[\sum_{i\sim j,i,j\in S}\mathbb{P}[B_i\cap B_j]\right] = p^2\Delta$, de modo que

$$\mathbb{E}_S\left[-\ln\mathbb{P}\Big[\bigcap_{i\in S}B_i^c\Big]\right] \geq p\mu - p^2\Delta/2 = \frac{\mu^2}{\Delta} - \frac{\mu^2\Delta}{2\Delta^2} = \frac{\mu^2}{2\Delta}. \tag{11.34}$$

Mas então existe $S\subseteq I$ tal que

$$-\ln\mathbb{P}\Big[\bigcap_{i\in S}B_i^c\Big] \geq \frac{\mu^2}{2\Delta}. \tag{11.35}$$

Assim,

$$\mathbb{P}\Big[\bigcap_{i\in I}B_i^c\Big] \leq \mathbb{P}\Big[\bigcap_{i\in S}B_i^c\Big] \leq \exp\Big\{-\frac{\mu^2}{2\Delta}\Big\}, \tag{11.36}$$

como queríamos. □

A seguinte cota é forte se $\Delta = o(\mu)$.

Teorema 11.3. *Sejam $\{B_i : i\in I\},\Delta,\mu,\varepsilon$ como na hipótese do teorema 11.1. Para cada $i\in I$, seja X_i a variável aleatória indicadora do evento B_i. Tome $X := \sum_{i\in I}X_i$. Para todo $\gamma > 0$, temos*

$$\mathbb{P}\big[X \leq (1-\gamma)\mu\big] \leq \exp\Big\{-\frac{\gamma^2\mu}{2+\Delta/\mu}\Big\}. \tag{11.37}$$

12 O lema local

Vamos agora estudar o lema local de Erdős e Lovász [8], também conhecido como lema local de Lovász, ou crivo local de Lovász.

Sejam $A_1,\dots,A_n$ eventos num espaço de probabilidade arbitrário. Para nós, esses são os eventos ruins: estamos interessados em provar que $\mathbb{P}\big[\bigcap_{i=1}^k A_i^c\big] > 0$.

Se $A_1,\dots,A_n$ são mutuamente independentes, então basta mostrar que $\mathbb{P}[A_i] < 1$ para todo i. Mas como não estamos supondo independência, não é suficiente provarmos isso. No nosso caso, precisamos provar que a dependência existente é limitada de certa forma.

Um *grafo de dependência* para os eventos $A_1,\dots,A_n$ é um grafo dirigido $D = ([n],A)$ tal que, para todo i, o evento A_i é mutuamente independente de $\{A_j : j\neq i \text{ e } (i,j)\notin A\}$. Isto é, se tomarmos $J := [n]\setminus\Gamma^+(i)$, onde $\Gamma^+(i) := \{j : (i,j)\in A\}$, então

$$\mathbb{P}\big[A_i \,\big|\, \text{expressão booleana envolvendo } A_j\text{'s}, j\in J\big] = \mathbb{P}[A_i]. \tag{12.1}$$

Teorema 12.1 (Lema Local)**.** *Sejam $A_1,\dots,A_n$ eventos num espaço de probabilidade e seja D um grafo de dependências desses eventos. Suponha que $d = \Delta^+(D) := \max\big\{|\Gamma^+(i)| : i\in[n]\big\}$ e, ademais, $\mathbb{P}[A_i]\leq p$ para todo i, com $ep(d+1)\leq 1$. Então*

$$\mathbb{P}\Big[\bigcap_{i=1}^k A_i^c\Big] > 0. \tag{12.2}$$

12.1 Hipergrafos e a propriedade B

Vamos ver alguns exemplos de aplicação deste teorema.

Dado um hipergrafo $H = (V, E)$, dizemos que H *tem a propriedade B* se existe uma bipartição $\{V_1, V_2\}$ de V tal que $e \cap V_1 \neq \emptyset$ e $e \cap V_2 \neq \emptyset$ para toda aresta $e \in E$. Isto é, existe uma coloração de V com duas cores tais que nenhuma aresta é monocromática.

Teorema 12.2. *Seja $H = (V, E)$ um hipergrafo tal que, para todo $f \in E$, vale que $|f| \geq k$ e $f \cap g \neq \emptyset$ para no máximo d arestas $g \in E$, $g \neq f$. Se $e(d+1) \leq 2^{k-1}$, então H tem a propriedade B.*

Demonstração. Colorimos V com duas cores aleatoriamente.

Para cada $f \in E$, seja A_f o evento "a aresta f é monocromática". Claramente, podemos definir um grafo de dependência $D = (E, A)$ com $\Delta^+(D) \leq d$: para cada $f \in E$, há no máximo d outras arestas que a intersectam. Todas as outras arestas são mutuamente independentes de A_f.

Ademais, para todo $f \in E$,

$$\mathbb{P}[A_f] = \frac{1}{2^{|f|-1}} \leq \frac{1}{2^{k-1}}. \tag{12.3}$$

E a hipótese nos garante que $e(d+1)/2^{k-1} \leq 1$. Mas então podemos aplicar o teorema 12.1 para obtermos $\mathbb{P}\big[\bigcap_{f \in E} A_f^c\big] > 0$, isto é, existe uma coloração que não deixa nenhuma aresta monocromática, que é justamente o que queríamos. □

Seja $H = (V, E)$ um hipergrafo. Dizemos que H é *k-uniforme* se, para todo $f \in E$, temos $|f| = k$. Dizemos que H é *k-regular* se, para todo $x \in V$, temos $d(x) = k$, onde $d(x) = |\{f \in E : x \in f\}|$.

Corolário 12.2.1. *Seja $H = (V, E)$ um hipergrafo k-uniforme k-regular. Se $k \geq 9$, então H tem a propriedade B.*

Demonstração. Vamos utilizar o teorema 12.2.

Seja $f \in E$. Seja $x \in f$. Como H é k-regular, então $d(x) = k$, ou seja, existem exatamente $k-1$ outras arestas que contém x e, portanto, intersectam f. Como H é k-uniforme, então $|f| = k$. Mas então existem no máximo $k(k-1)$ outras arestas que intersectam f. Logo, podemos tomar $d = k(k-1)$.

Para aplicarmos o teorema 12.2, precisamos que $e[k(k-1)+1] \leq 2^{k-1}$. É fácil verificar que isso vale para $k \geq 9$. □

O corolário acima vale para $k = 8$: isso foi provado opr Alon e Bregman. É um problema em aberto para $4 \leq k < 8$.

Seja $G = (V, E)$ um grafo. Para cada $v \in V$, seja L_v um conjunto de cores. Uma *$(L_v)_{v \in V}$-coloração de G* é uma atribuição de cores para os vértices de G de forma que a cor atribuída a um vértice v pertence a L_v.

Exercício 15. *Seja $G = (V, E)$ um grafo e, para cada $v \in V$, seja L_v um conjunto de cores. Suponha que $|L_v| \geq l$ para todo $v \in V$ e, para todo $v \in V$, toda cor c ocorre no máximo $l/8$ vezes nas listas $L_{u_1}, \ldots, L_{u_k}$, onde $\Gamma(v) = \{u_1, \ldots, u_k\}$. Mostre que G admite uma L_v-coloração própria.*

[*Dica:* Para cada cor i e aresta $e = xy$ tal que $i \in L_x \cap L_y$, chame de $A_{i,e}$ o evento em que x e y são coloridos com a cor i.]

12.2 Demonstração do lema

Primeiro vamos provar um resultado mais forte, do qual o teorema 12.1 seguirá como corolário.

Teorema 12.3. *Sejam $A_1, \ldots, A_n$ eventos num espaço de probabilidade e D um grafo de dependências desses eventos. Suponha que $x_1, \ldots, x_n \in [0, 1)$ são tais que*

$$\mathbb{P}[A_i] \leq x_i \prod_{j \in \Gamma_D^+(i)} (1 - x_j) \tag{12.4}$$

para todo i. Então

$$\mathbb{P}\Big[\textstyle\bigcap_{i=1}^n A_i^c\Big] \geq \displaystyle\prod_{i=1}^n (1 - x_i) > 0. \tag{12.5}$$

Demonstração. Vamos mostrar, por indução em s, a seguinte afirmativa I_s: para qualquer $S \subseteq [n]$ com $|S| = s$, temos

$$\mathbb{P}\Big[\textstyle\bigcap_{j \in S} A_j^c\Big] > 0 \tag{12.6}$$

e, para todo $i \in [n] \setminus S$,

$$\mathbb{P}\Big[A_i \,\Big|\, \textstyle\bigcap_{j \in S} A_j^c\Big] \leq x_i. \tag{12.7}$$

Para a base da indução, temos $s = 0$, ou seja, $S = \emptyset$. Note que $\mathbb{P}\big[\bigcap_{j \in \emptyset} A_j^c\big] = \mathbb{P}[\Omega] = 1 > 0$, de modo que (12.6) é satisfeita. Ademais, a condição (12.4) da hipótese nos garante que (12.7) é satisfeita, pois $1 - x_i \leq 1$ para todo i.

Para o passo da indução, suponha que $s > 0$ e que a afirmativa $I_{s'}$ vale para todo $0 \leq s' < s$.

Sejam $S \subseteq [n]$ com $|S| = s$ e $i \in [n] \setminus S$.

Seja $k \in S$ e $S' := S \setminus \{k\}$. Temos

$$\mathbb{P}\Big[A_k^c \cap \textstyle\bigcap_{j \in S'} A_j^c\Big] = \mathbb{P}\Big[A_k^c \,\Big|\, \textstyle\bigcap_{j \in S'} A_j^c\Big] \mathbb{P}\Big[\textstyle\bigcap_{j \in S'} A_j^c\Big] \tag{12.8}$$

Como $|S'| < s$, a hipótese de indução nos garante que $\mathbb{P}\big[\bigcap_{j \in S'} A_j^c\big] > 0$ e que $\mathbb{P}\big[A_k \mid \bigcap_{j \in S'} A_j^c\big] \leq x_i < 1$, de modo que $\mathbb{P}\big[A_k^c \mid \bigcap_{j \in S'} A_j^c\big] > 0$. Ou seja, os dois fatores do lado direito de (12.8) são positivos. Mas então provamos (12.6) para S.

Resta provarmos (12.7). Ponha

$$S_1 := S \cap \Gamma_D^+(i) \tag{12.9}$$

e $S_2 := S \setminus S_1$.

Suponha que $S_1 = \emptyset$, de modo que $S = S_2$. Por (12.1), temos

$$\mathbb{P}\Big[A_i \,\Big|\, \bigcap_{j\in S} A_j^c\Big] = \mathbb{P}[A_i] \le x_i. \tag{12.10}$$

Podemos supor então que $S_1 \neq \emptyset$, e portanto $|S_2| < s$. Pela hipótese de indução,

$$\mathbb{P}\Big[\bigcap_{l\in S_2} A_l^c\Big] > 0. \tag{12.11}$$

Suponha que $S_1 = \{j_1, \dots, j_r\}$. Usando (11.9), não é difícil ver que

$$\begin{aligned}\mathbb{P}\Big[\bigcap_{i=1}^r A_{j_i}^c \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big] = \mathbb{P}\Big[A_{j_1}^c \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big] \times \\ \mathbb{P}\Big[A_{j_2}^c \,\Big|\, A_{j_1}^c \cap \bigcap_{l\in S_2} A_l^c\Big] \times \cdots \times \\ \mathbb{P}\Big[A_{j_r}^c \,\Big|\, \bigcap_{i=1}^{r-1} A_{j_i}^c \cap \bigcap_{l\in S_2} A_l^c\Big].\end{aligned} \tag{12.12}$$

Mas então, pela hipótese de indução, temos

$$\begin{aligned}\mathbb{P}\Big[\bigcap_{i=1}^r A_{j_i}^c \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big] &= \Big(1 - \mathbb{P}\Big[A_{j_1} \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big]\Big) \times \\ &\quad \Big(1 - \mathbb{P}\Big[A_{j_2} \,\Big|\, A_{j_1}^c \cap \bigcap_{l\in S_2} A_l^c\Big]\Big) \times \cdots \times \\ &\quad \Big(1 - \mathbb{P}\Big[A_{j_r} \,\Big|\, \bigcap_{i=1}^{r-1} A_{j_i}^c \cap \bigcap_{l\in S_2} A_l^c\Big]\Big) \\ &\ge (1 - x_{j_1})(1 - x_{j_2}) \cdots (1 - x_{j_r}) \\ &= \prod \Big\{(1 - x_j) : j \in S \cap \Gamma_D^+(i)\Big\}.\end{aligned} \tag{12.13}$$

Por outro lado, por (12.1) e (12.4), temos

$$\begin{aligned}\mathbb{P}\Big[A_i \cap \bigcap_{j\in S_1} A_j^c \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big] &\le \mathbb{P}\Big[A_i \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big] = \mathbb{P}[A_i] \\ &\le x_i \prod_{j\in\Gamma_D^+(i)} (1 - x_j) \\ &\le x_i \prod \Big\{(1 - x_j) : j \in S \cap \Gamma_D^+(i)\Big\}.\end{aligned} \tag{12.14}$$

De (11.9), não é difícil verificar que

$$\mathbb{P}\Big[A_i \,\Big|\, \bigcap_{j\in S} A_j^c\Big] = \frac{\mathbb{P}\Big[A_i \cap \bigcap_{j\in S_1} A_j^c \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big]}{\mathbb{P}\Big[\bigcap_{j\in S_1}^r A_j^c \,\Big|\, \bigcap_{l\in S_2} A_l^c\Big]}. \tag{12.15}$$

Usando (12.13) e (12.14) em (12.15), obtemos então

$$\mathbb{P}\Big[A_i \,\Big|\, \bigcap_{j\in S} A_j^c\Big] \le x_i, \tag{12.16}$$

o que completa a prova do passo de indução.

Para provar (12.5), basta ver que

$$\begin{aligned}
\mathbb{P}\Big[\bigcap_{i=1}^n A_i^c\Big] &= \mathbb{P}\big[A_1^c\big] \times \mathbb{P}\big[A_2^c \,\big|\, A_1^c\big] \times \cdots \times \mathbb{P}\big[A_n^c \,\big|\, \bigcap_{1\le i<n} A_i^c\big] \\
&= \Big(1 - \mathbb{P}\big[A_1^c\big]\Big) \times \Big(1 - \mathbb{P}\big[A_2^c \,\big|\, A_1^c\big]\Big) \times \cdots \times \\
&\qquad\qquad \times \Big(1 - \mathbb{P}\big[A_n^c \,\big|\, \bigcap_{1\le i<n} A_i^c\big]\Big) \\
&\ge (1-x_1)\cdots(1-x_n) > 0.
\end{aligned} \tag{12.17}$$

□

Podemos agora provar o teorema 12.1:

Demonstração do teorema 12.1. Se $d = 0$, então os eventos são mutuamente independentes, e o teorema segue imediatamente.

Suponha então que $d \ge 1$. Tome $x_i := 1/(d+1) < 1$ para todo i. Queremos que x_i satisfaça (12.4). Temos

$$x_i \prod_{j\in\Gamma_D^+(i)} (1-x_j) \ge \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d. \tag{12.18}$$

Mas

$$\left(1 - \frac{1}{d+1}\right)^d = \left(\frac{d}{d+1}\right)^d = \left(\frac{d+1}{d}\right)^{-d} \ge e^{-1}, \tag{12.19}$$

pois, por (6.8),

$$e \ge \left(1 + \frac{1}{d}\right)^d = \left(\frac{d+1}{d}\right)^d. \tag{12.20}$$

Assim, utilizando a hipótese de que $ep(d+1) \le 1$, obtemos

$$x_i \prod_{j\in\Gamma_D^+(i)} (1-x_j) \ge \frac{1}{e(d+1)} \ge p \ge \mathbb{P}[A_i], \tag{12.21}$$

ou seja, os x_i's satisfazem (12.4). Agora basta aplicar o teorema 12.3 para obter

$$\mathbb{P}\Big[\bigcap_{i=1}^k A_i^c\Big] > 0. \tag{12.22}$$

□

12.3 Dimensão de ordens parciais

Seja $\mathcal{P} = (P, \leq_P)$ uma ordem parcial com P finito. Uma *extensão linear* L de P é uma ordem total $\leq_L$ tal que $x \leq_P y \Longrightarrow x \leq_L y$. A *dimensão* de $\mathcal{P}$, definida por Dushnik, é $\dim(\mathcal{P}) := \min|\mathcal{L}|$, onde $\mathcal{L}$ varia sobre a família de extensões lineares de $\mathcal{P}$ com $\bigcap_{L \in \mathcal{L}} L = \mathcal{P}$.

Equivalentemente, temos que $\dim(\mathcal{P}) \leq d$ se, e somente se, existe uma injeção $\varphi : P \to \mathbb{R}^d$ com $x \leq_p y \iff \varphi(x) \leq_d \varphi(y)$, onde $(x_1, \ldots, x_d) \leq_d (y_1, \ldots, y_d) \iff x_i \leq y_i, \forall i$. Note que isso é uma interpretação geométrica.

Exercício 16. *Seja* $P := \{a_1, \ldots, a_{k+1}, b_1, \ldots, b_{k+1}\}$*. A relação* $\leq_P$ *é definida da seguinte forma. Temos* $a_i <_P b_j$ *para todo* $j \neq i$*. Isso define uma ordem parcial* $\mathcal{P} := (P, \leq_P)$*. Prove que* $\dim(\mathcal{P}) = k + 1$*.*

Dado $x \in P$, defina

$$c(x) := c_{\mathcal{P}}(x) := \big|\{y \in P : x <_P y \text{ ou } y <_P x\}\big| \tag{12.23}$$

como o número de elementos de P comparáveis com x. Tome

$$k := k(\mathcal{P}) := \max_{x \in P} \big|c(x)\big|. \tag{12.24}$$

Trotter [20] conjecturou que

$$\dim(\mathcal{P}) \leq f(k) \tag{12.25}$$

para alguma função f. Na verdade, Rödl e Trotter (cf. [20]) provaram que

$$f(k) \leq 2k^2 + 2. \tag{12.26}$$

Já vimos no exercício 16 acima que $f(k) \geq k + 1$.

Füredi e Kahn [9] melhoraram significantemente o resultado de Rödl e Trotter, provando que

$$\dim(\mathcal{P}) \leq ck(\ln k)^2 \tag{12.27}$$

para alguma constante $c > 0$.

Tome $N := |P|$ e seja $\pi : [N] \to P$ uma bijeção. Chamaremos π de permutação de P. Dada uma permutação π de P, denote por $\leq_\pi$ a ordem total sobre P definida por π. Note que π é uma extensão linear de $\mathcal{P}$ se $x \leq_P y \Longrightarrow x \leq_\pi y$.

Seja π uma permutação qualquer de P. Denote $\pi = x_1 \cdots x_N$, isto é, $x_i = \pi(i)$ para $i = 1, \ldots, N$. Sejam $i, j \in [N]$ com $i < j$. Temos

$$\pi = x_1 \cdots x_i \cdots x_j \cdots x_N. \tag{12.28}$$

Uma rotação de π é uma permutação da forma

$$\pi = x_1 \cdots x_{i-1} x_j x_i \cdots x_{j-1} x_{j+1} \cdots x_N. \tag{12.29}$$

Considerarmos apenas tais rotações para i e j satisfazendo $x_i >_P x_j$ e $i < j$.

Lema 12.4. *Seja π uma permutação de P. Podemos transformar π em uma extensão linear de $\mathcal{P}$ fazendo rotações como acima.*

Demonstração. Seja $\pi = x_1 \cdots x_N$ uma permutação de P e seja k o maior inteiro entre 0 e N tal que $x_j <_P x_i \Longrightarrow x_j <_\pi x_i$ para todo $i \leq k$ e $j \in [n]$. Se $k = N$ nada temos a demonstrar, pois π é uma extensão linear de $\mathcal{P}$.

Suponha então que $k < N$. Pela escolha de k, sabemos que existem $j \in [n]$ tais que $x_j <_P x_{k+1}$ e $x_j >_\pi x_{k+1}$. Em outras palavras, o conjunto

$$J := \{j : x_j <_P x_{k+1} \text{ e } j > k+1\} \tag{12.30}$$

é não vazio. Mas então a rotação que leva x_j imediatamente antes de x_{k+1} é aplicável para todo $j \in J$. Escolha j em J tal que x_j é minimal. É fácil ver que, após a rotação que leva x_j imediatamente antes de x_{k+1}, a permutação resultante π' satisfaz $x_j <_P x_i \Longrightarrow x_j <_{\pi'} x_i$ para todo $i \leq k+1$ e $j \in [n]$. O resultado agora segue por indução em k. □

Lema 12.5. *A dimensão de $\mathcal{P}$ é o menor α tal que existem permutações $\pi_1, \ldots, \pi_\alpha$ de P tais que*

$$y \not\leq_P x \Longrightarrow \exists i : x <_{\pi_i} U(y), \tag{12.31}$$

onde

$$U(y) := \{z \in P : y \leq_P z\}. \tag{12.32}$$

Demonstração. Primeiro vamos mostrar que $\dim(\mathcal{P}) \geq \alpha$.

Dada uma permutação π de P, chame de L_π a ordem total induzida por π.

Sejam $\pi_1, \ldots, \pi_d$ permutações de P tais que $\mathcal{P} = \bigcap_{i=1}^d L_{\pi_i}$ e $d = \dim(\mathcal{P})$. Afirmamos que (12.31) vale para estas permutações π_i e, portanto, $\alpha \leq \dim(\mathcal{P})$.

Sejam $x, y \in P$ com $y \not\leq_P x$. Como $\mathcal{P} = \bigcap_{i=1}^d L_{\pi_i}$, então existe um índice i tal que $x <_{\pi_i} y$. Como π_i é extensão linear de $\mathcal{P}$, temos que $y \leq_{\pi_i} U(y)$ e assim $x <_{\pi_i} U(y)$, por transitividade. Provamos assim que vale (12.31). Logo, $\dim(\mathcal{P}) \geq \alpha$.

Agora vamos mostrar que $\dim(\mathcal{P}) \leq \alpha$. Sejam $\pi_1, \ldots, \pi_\alpha$ como no enunciado do lema, com α mínimo.

Fixe i. Transforme π_i em uma extensão linear de $\mathcal{P}$ usando rotações (como no lema 12.4). Afirmamos que, ao fazermos isso para todo i, obtemos permutações $\pi'_1, \ldots, \pi'_\alpha$ que são extensões lineares de $\mathcal{P}$ e que também satisfazem (12.31).

Sejam $x, y \in P$ com $y \not\leq_P x$. Por hipótese, existe um índice i tal que $x <_{\pi_i} U(y)$. Suponha que, numa rotação, algum elemento $u \in U(y)$ seja movido para imediatamente antes de um elemento $v \in P$, e acabe ficando menor do que x nesta nova permutação. Pela definição de rotação, temos então $v >_P u$. Ademais, como u é menor que x nesta nova permutação (após a rotação), temos $v \leq_{\pi_i} x$. Por outro lado, como $u \in U(y)$, então $y \leq_P u <_P v$, de modo que $v \in U(y)$. Mas então $x <_{\pi_i} v$. Temos assim $v \leq_{\pi_i} x$ e $x <_{\pi_i} v$, um absurdo! Logo, (12.31) vale para $\pi'_1, \ldots, \pi'_\alpha$.

Afirmamos agora que $\mathcal{P} = \bigcap_{i=1}^{\alpha} L_{\pi_i'}$. Como $L_{\pi_i'}$ é extensão linear de $\mathcal{P}$ para cada i, é evidente que $\mathcal{P} \subseteq \bigcap_{i=1}^{\alpha} L_{\pi_i'}$. Suponha que $y \not\leq_P x$. De (12.31) temos $x <_{\pi_i'} U(y)$ para algum i. Em particular, $x <_{\pi_i'} y$ e, portanto, $(y, x) \notin \bigcap_{i=1}^{\alpha} L_{\pi_i'}$. Assim, concluímos que $\alpha \geq \dim(\mathcal{P})$, como queríamos. $\square$

Proposição 12.6. *Seja*

$$u := u(\mathcal{P}) := \max_{y \in P} \big|U(y)\big|. \tag{12.33}$$

Temos

$$\dim(\mathcal{P}) \leq \big\lceil 2(u+1) \ln |P| \big\rceil. \tag{12.34}$$

Demonstração. Sejam $\pi_1, \ldots, \pi_d$ permutações aleatórias de P, com

$$d := \big\lceil 2(u+1) \ln |P| \big\rceil. \tag{12.35}$$

Afirmamos que essas permutações satisfazem (12.31) com probabilidade positiva e, portanto, o resultado segue do lema 12.5.

Fixe $x, y \in P$ distintos e $1 \leq i \leq d$. Então

$$\mathbb{P}\big[x \leq_{\pi_i} U(y)\big] = \frac{1}{1 + |U(y)|} \geq \frac{1}{1+u}. \tag{12.36}$$

Assim,

$$\mathbb{P}\big[\forall i, x \not<_{\pi_i} U(y)\big] \leq \left(1 - \frac{1}{1+u}\right)^d \leq \exp\left\{-\frac{d}{u+1}\right\} \leq \frac{1}{|P|^2}, \tag{12.37}$$

onde a última desigualdade segue da definição de d. Portanto,

$$\begin{aligned} &\mathbb{P}\Big[\exists x, y \in P, y \not\leq_P x \text{ e } \forall i x \not<_{\pi_i} U(y)\Big] \\ &\qquad \leq \mathbb{P}\Big[\exists x \neq y \in P, \forall i x \not<_{\pi_i} U(y)\Big] < |P|^2 \frac{1}{|P|^2} = 1. \end{aligned} \tag{12.38}$$

Assim, existem $\pi_1, \ldots, \pi_d$ que satisfazem as hipóteses do lema 12.5 e a proposição está provada. $\square$

12.4 Famílias misturadoras de permutações

Sejam $\pi_1, \ldots, \pi_d$ permutações de $[n]$. Dizemos que $\{\pi_1, \ldots, \pi_d\}$ é uma família *t-misturadora*, onde $2 \leq t \leq n$, se, para qualquer $X \in \binom{[n]}{t}$ e $x \in X$, existe um índice i tal que $x <_{\pi_i} X \setminus \{x\}$. Seja

$$\begin{aligned} d(n,t) := \min\big\{d : \ &\text{existe uma família } t\text{-misturadora} \\ &\qquad\qquad \text{de } [n] \text{ com } d \text{ permutações}\big\}. \end{aligned} \tag{12.39}$$

Por exemplo, $d(n, 2) = 2$.

Exercício 17. *Mostre que, para quaisquer n e $2 \leq t \leq n$, temos*

$$d(n,t) \leq \left\lceil t^2\big(1+\ln(n/t)\big) + t\ln t \right\rceil. \tag{12.40}$$

Considere um hipergrafo $(V, \mathcal{H})$. Para cada $v \in V$, defina o grau de v, denotado por $d(v)$ ou $d_{\mathcal{H}}(v)$, como o número de hiperarestas de $\mathcal{H}$ contendo v:

$$d(v) := d_{\mathcal{H}}(v) := |\{H \in \mathcal{H} : v \in H\}|. \tag{12.41}$$

Defina o grau máximo do hipergrafo como

$$\Delta(\mathcal{H}) := \max_{v \in V} \big\{d_{\mathcal{H}}(v)\big\}. \tag{12.42}$$

Lema 12.7. *Seja $(X, \mathcal{H})$ um hipergrafo, com $\Delta(\mathcal{H}) \leq b$ e $|H| \leq b$ para toda hiperaresta $H \in \mathcal{H}$ e para algum $b \geq b_0$. Ponha $s := \lceil b/\ln b\rceil$ e $v := \lceil 2e\ln b\rceil$. Então existe uma "coloração" $X = X_1 \cup \cdots \cup X_s$ tal que $|H \cap X_i| \leq v$ para todo $H \in \mathcal{H}$ e $1 \leq i \leq s$.*

Demonstração. Considere o seguinte experimento aleatório: para cada vértice $x \in X$, escolha aleatoriamente uma dentre as s possíveis e atribua esta cor a x. Vamos usar o lema local para mostrar que, com probabilidade positiva, existe uma coloração satisfazendo as propriedades desejadas.

Para cada hiperaresta $H \in \mathcal{H}$ e cor $i \in [s]$, defina o evento $A_{H,i}$ como a ocorrência de $|H \cap X_i| > v$. Ou seja, os eventos $A_{H,i}$ são os "eventos ruins". Considere o grafo D sobre estes $A_{H,i}$ pondo um arco ligando $A_{H,i}$ a $A_{H',i'}$ se, e somente se, $H \cap H' \neq \emptyset$. Então é fácil ver que D é um grafo de dependências para os eventos $A_{H,i}$, já que os eventos $A_{H,i}$ e $A_{H',i'}$ são independentes se H e H' forem disjuntos.

Seja d o grau máximo de D. Seja H uma hiperaresta de $\mathcal{H}$ e $i \in [s]$ uma cor. Seja $x \in H$. Sabemos que x está em no máximo $d_{\mathcal{H}}(x) - 1 \leq \Delta(\mathcal{H}) - 1 \leq b - 1$ outras hiperarestas. Ademais, H tem no máximo b elementos. Então há no máximo $b(b-1)$ hiperarestas $H' \in \mathcal{H}$ que intersectam H. Além disso, contando a própria aresta H, e todas as s cores, temos que $d \leq \big(b(b-1)+1\big)s$. Note ainda que estamos contando o próprio evento $A_{H,i}$ nessa contagem, de modo que $d + 1 \leq \big(b(b-1)+1\big)s$. Concluímos que $d + 1 \leq b^3$.

Antes de aplicar o lema local, precisamos ainda limitar a probabilidade do evento $A_{H,i}$. Observe que a variável aleatória $|H \cap X_i|$ segue a distribuição binomial com parâmetros $|H|$ e $1/s$. Mas então, usando (8.18) e (8.20), temos

$$\begin{aligned}
\mathbb{P}[A_{H,i}] = \mathbb{P}\big[|H \cap X_i| > v\big] &\leq \binom{b}{v}\left(\frac{1}{s}\right)^v \leq \left(\frac{eb}{vs}\right)^v \\
&\leq \left(\frac{eb}{(2e\ln b)(b/\ln b)}\right)^v = \left(\frac{1}{2}\right)^v \leq 2^{-2e\ln b} \\
&= b^{-2e\ln 2} \leq b^{-3,7}.
\end{aligned} \tag{12.43}$$

Logo, $\mathbb{P}[A_{H,i}] \leq p$ para $p := b^{-3,7}$.

Como

$$ep(d+1) \leq eb^{-3,7}b^3 < 1 \tag{12.44}$$

para todo $b \geq b_0 := 5$, podemos aplicar o lema local para mostrar que, com probabilidade positiva, nenhum dos eventos ruins acontece, como queríamos. □

Observe que, no resultado acima, não foi feita nenhuma suposição sobre $|X|$. Tudo é feito de "forma local".

Lema 12.8. *Sejam $a, b \geq 1$ inteiros e $(Y, \mathcal{H})$ um hipergrafo tal que $\Delta(\mathcal{H}) \leq b$ e $|H| \leq a$ para todo $H \in \mathcal{H}$. Então existe uma "coloração" $Y = Y_1 \cup \cdots \cup Y_n$, com $n := (a-1)b+1$, tal que $|H \cap Y_i| \leq 1$ para toda hiperaresta $H \in \mathcal{H}$ e cor $i \in [n]$.*

Demonstração. Podemos aplicar um algoritmo guloso para provar este lema: para cada elemento $y \in Y$ ainda não colorido, consideramos as "cores proibidas" para y como sendo as cores que, se atribuídas a y, violariam a propriedade do lema. As "cores proibidas" para y são as cores já atribuídas a algum elemento x que está numa mesma hiperaresta contendo y. Mas y está em no máximo $\Delta(\mathcal{H}) \leq b$ hiperarestas, e cada uma delas tem no máximo $a-1$ outros elementos além de y. Mas então sempre existe uma cor, dentre as $(a-1)b+1$ cores possíveis, que não é proibida para y, e podemos colorir y com tal cor. □

Exercício 18. *Seja $\mathcal{P}$ uma ordem parcial e seja $k := k(\mathcal{P})$, como definido em (12.24). Mostre que, para todo $k \geq 2$,*

$$\dim(\mathcal{P}) \leq ck^2 \ln k \tag{12.45}$$

para alguma constante $c > 0$.

Podemos agora provar o teorema de Füredi e Kahn citado acima:

Teorema 12.9 (Füredi e Kahn [9]). *Seja $\mathcal{P}$ uma ordem parcial e seja $k := k(\mathcal{P})$, como definido em (12.24). Então*

$$\dim(\mathcal{P}) \leq ck(\ln k)^2 \tag{12.46}$$

para alguma constante $c > 0$.

Demonstração. Podemos supor, sem perda de generalidade, que $k \geq k_0$ para algum k_0 grande. De fato, se (12.46) não valer para algum $k < k_0$, basta trocar c por uma outra constante, suficientemente grande, e (12.46) passará a valer para todo k.

Tome $\mathcal{H} := \{U(x) : x \in P\}$, usando a definição de $U(x)$ dada em (12.32). Então $\mathcal{H}$ é o conjunto de hiperarestas de um hipergrafo sobre P, sendo que $\Delta(\mathcal{H}) \leq k+1$ e $|H| \leq k+1$ para todo $H \in \mathcal{H}$.

Pelo lema 12.7 acima, existe uma "coloração" $P = X_1 \cup \cdots \cup X_s$, com $s := \lceil (k+1)/\ln(k+1) \rceil$, tal que $|U(x) \cap X_i| \leq v$ para todo $x \in P$ e $i \in [s]$, onde $v := \lceil 2e \ln(k+1) \rceil$.

Para cada $i \in [s]$, ponha $\mathcal{H}_i := \{U(x) \cap X_i : x \in P\}$. Então $\mathcal{H}_i$ é o conjunto de hiperarestas de um hipergrafo sobre X_i, com $\Delta(\mathcal{H}_i) \le k+1$ e $|H| \le v$ para toda hiperaresta $H \in \mathcal{H}_i$.

Pelo lema 12.8 acima, existe uma "coloração" $X_i = X_{i1} \cup \cdots \cup X_{in}$, com $n := (v-1)(k+1)+1$, tal que $|U(x) \cap X_{ij}| \le 1$ para todo $x \in X_i$ e $j \in [n]$.

Seja $\pi_1, \ldots, \pi_d$ uma família de permutações $(v+1)$-misturantes, onde $d := d(n, v+1)$ como definido em (12.39). Isto é, para todo $T \subseteq [n]$ com $|T| \le v$ e qualquer $\alpha \in [n] \setminus T$, existe um índice $i \in [d]$ tal que $\alpha \le_{\pi_i} T$. Por (12.40), temos

$$d \le \left\lceil (v+1)^2 \Big(1 + \ln \frac{(v-1)(k+1)+1}{v+1}\Big) + (v+1)\ln(v+1) \right\rceil. \qquad (12.47)$$

Para quaisquer i, j, considere R_{ij} uma ordem total arbitrária sobre X_{ij} e seja R_{ij}^{-1} a ordem reversa, isto é, $x \le_{R_{ij}} y \iff y \le_{R_{ij}^{-1}} x$. Para cada $i \in [s]$ e cada $l \in [d]$, considere

$$\pi_{i,l} := (R_{i,\pi_l(1)}, \ldots, R_{i,\pi_l(n)}, R_{P \setminus X_i}) \qquad (12.48)$$

$$\pi'_{i,l} := (R^{-1}_{i,\pi_l(1)}, \ldots, R^{-1}_{i,\pi_l(n)}, R'_{P \setminus X_i}), \qquad (12.49)$$

onde $R_{P \setminus X_i}$ e $R'_{P \setminus X_i}$ são ordens arbitrárias sobre $P \setminus X_i$.

Temos assim $2sd$ permutações $\pi_{i,l}$ e $\pi'_{i,l}$. Afirmamos que estas permutações de P satisfazem (12.31) do lema 12.5, e portanto $\dim(\mathcal{P}) \le 2sd$.

Primeiro verificamos que $2sd = O\big(k(\ln k)^2\big)$, como precisamos. Para tanto, note que $s = O(k/\ln k)$ e que $v = O(\ln k)$, de modo que $d = O\big((\ln k)^3\big)$. Mas então $2sd = O\big(k(\ln k)^2\big)$, como queríamos. Assim, resta apenas provarmos (12.31).

Fixe $x, y \in P$ com $y \not\le_P x$. Queremos que $x <_{\pi_{i,l}} U(y)$ ou $x <_{\pi_{i,l}} U(y)$ para algum $i \in [s], l \in [d]$. Temos $x \in X_{i\alpha}$ para algum $i \in [s], \alpha \in [n]$. Em qualquer $\pi_{i,l}, \pi'_{i,l}$, temos $x <_{\pi_{i,l}} U(y) \setminus X_i$ e $x <_{\pi'_{i,l}} U(y) \setminus X_i$, pela construção de $\pi_{i,l}$ e $\pi'_{i,l}$, onde os elementos fora de X_i são os últimos.

Consideremos agora $U(y) \cap X_i \in \mathcal{H}_i$. Ponha $T := \{j \in [n] : U(y) \cap X_{ij} \ne \emptyset\}$ como o conjunto de cores j que ocorrem em $U(y) \cap X_i$. Temos então $|T| \le |U(y) \cap X_i| \le v$. Pela escolha das permutações $\pi_1, \ldots, \pi_d$, existe π_l tal que $\alpha <_{\pi_l} T \setminus \{\alpha\}$. Temos assim que $R_{i\alpha}$ aparece antes de $R_{i\beta}$ em $\pi_{i,l}$ e em $\pi'_{i,l}$ para todo $\beta \in T \setminus \alpha$. Se $\alpha \notin T$, estamos feitos, pois aí vale que $x <_{\pi_{i,l}} U(y) \cap X_i$. Suponha então que $\alpha \in T$. Então $|U(y) \cap X_{i\alpha}| = 1$. Seja z o único elemento de $U(y) \cap X_{i\alpha}$. Então $x <_{R_{i\alpha}} z$ ou $x <_{R_{i\alpha}^{-1}} z$. Mas aí temos ou $x <_{\pi_{i,l}} z$ ou $x <_{\pi'_{i,l}} z$. Concluímos que $x <_{\pi_{i,l}} U(y) \cap X_i$ ou $x <_{\pi_{i,l}} U(y) \cap X_i$. Ou seja, vale (12.31), como queríamos. □

13 Entropia

Seja X uma variável aleatória tomando valores em um conjunto finito S. Para todo $x \in S$, ponha $p_x := \mathbb{P}[X = x]$. Temos então um vetor $(p_x)_{x \in S}$ indexado por S, com $0 \le p_x \le 1$ para todo $x \in S$ e $\sum_{x \in S} p_x = 1$. Defina a *entropia de* X, também chamada de *entropia binária de* X e denotada por $H(X)$ ou $H_2(X)$, como

$$H(X) := H_2(X) := \mathbb{E}\left[\lg(1/p_x)\right] = \sum_{x\in S} p_x \lg \frac{1}{p_x} = \sum_{x\in S} -p_x \lg p_x. \qquad (13.1)$$

Convencionamos ainda que $0\lg 0 = 0\lg(1/0) = 0$.

Se Y é uma variável aleatória tomando valores em um conjunto T finito, definimos a *entropia de* X *dado* Y, denotada por $H(X|Y)$, como

$$H(X|Y) := H(X,Y) - H(Y). \qquad (13.2)$$

Observe que neste caso estamos considerando $Z := (X,Y)$ como uma variável aleatória tomando valores em $S\times T$. Desta forma,

$$\begin{aligned} H(X|Y) &= H(X,Y) - H(Y) \\ &= \left(\sum_{(x,y)\in S\times T} p_{xy} \lg \frac{1}{p_{xy}}\right) - \left(\sum_{y\in T} p_y \lg \frac{1}{p_y}\right). \end{aligned} \qquad (13.3)$$

Intuitivamente, $H(X)$ mede a informação em X e $H(X|Y)$ a informação em X, dado que sabemos Y.

Observe ainda que

$$\frac{\mathrm{d}^2(\lg x)}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x\ln 2}\right) = -\frac{1}{x^2\ln 2} \qquad (13.4)$$

é sempre negativo para $x>0$, de modo que $\lg x$ é côncava para $x>0$, e que

$$\frac{\mathrm{d}^2(x\lg x)}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\ln 2} + \lg x\right) = \frac{1}{x\ln 2} \qquad (13.5)$$

é sempre positivo para $x>0$, de modo que $x\lg x$ é convexa para $x>0$.

Podemos verificar no seguinte lema que a afirmação de que $H(X)$ mede a informação em X faz sentido.

Lema 13.1. *Sejam X,Y,Z variáveis aleatórias tomando valores nos conjuntos finitos S,T,U, respectivamente. Então*

$$\begin{aligned} H(X) &\le \lg|S|, && (13.6)\\ H(X,Y) &\ge H(X), && (13.7)\\ H(X,Y) &\le H(X)+H(Y) && (13.8) \end{aligned}$$

com igualdade sempre que X e Y forem independentes, e

$$H(X|Y,Z) \le H(X|Y). \qquad (13.9)$$

Demonstração. Para provar (13.6), basta utilizar a desigualdade (5.11) de Jensen, com $f(x) := -\lg x$ e $\lambda_x := 1/p_x$ e para todo $x \in S$:

$$H(X) = -\sum_{x\in S} p_x f(1/p_x) \leq -f\left(\sum_{x\in S} p_x \frac{1}{p_x}\right) = \lg |S|$$

Para provar (13.7), vamos usar o fato de que $\lg x$ é uma função crescente. Assim, como $p_{xy} = \mathbb{P}[X = x, Y = y] \leq p_x$, temos

$$\begin{aligned} H(X,Y) &= \sum_{x\in S}\sum_{y\in T} p_{xy} \lg \frac{1}{p_{xy}} \geq \sum_{x\in S}\sum_{y\in T} p_{xy} \lg \frac{1}{p_x} \\ &= \sum_{x\in S}\left(\sum_{y\in T} p_{xy}\right) \lg \frac{1}{p_x} = \sum_{x\in S} p_x \lg \frac{1}{p_x} = H(X). \end{aligned} \tag{13.10}$$

Vamos agora mostrar (13.8). Temos

$$\begin{aligned} &H(X) + H(Y) - H(X,Y) = \\ &\quad \sum_{x\in S}\left(\sum_{y\in T} p_{xy}\right) \lg \frac{1}{p_x} + \sum_{y\in T}\left(\sum_{x\in S} p_{xy}\right) \lg \frac{1}{p_y} - \sum_{x\in S}\sum_{y\in T} p_{xy} \lg \frac{1}{p_{xy}}. \end{aligned} \tag{13.11}$$

Assim,

$$\begin{aligned} H(X) + H(Y) - H(X,Y) &= \sum_{x\in S}\sum_{y\in T} p_{xy}\left(\lg \frac{1}{p_x} + \lg \frac{1}{p_y} - \lg \frac{1}{p_{xy}}\right) \\ &= \sum_{x\in S}\sum_{y\in T} p_{xy} \lg\left(\frac{p_{xy}}{p_x p_y}\right). \end{aligned} \tag{13.12}$$

Considere a função $f(w) := w \lg w$ e ponha $w_{xy} := p_{xy}/(p_x p_y)$ para todo $(x, y) \in S \times T$. Então

$$H(X) + H(Y) - H(X,Y) = \sum_{x\in S}\sum_{y\in T} p_x p_y f(w_{xy}). \tag{13.13}$$

Mas $\sum_{x\in S}\sum_{y\in T} p_x p_y = 1$ e assim, pela convexidade de f e pela desigualdade (5.11) de Jensen, temos

$$\begin{aligned} \sum_{x\in S}\sum_{y\in T} p_x p_y f(w_{xy}) &\geq f\left(\sum_{x\in S}\sum_{y\in T} p_x p_y w_{xy}\right) \\ &= f\left(\sum_{x\in S}\sum_{y\in T} p_{xy}\right) \\ &= f(1) = 0, \end{aligned} \tag{13.14}$$

de modo que $H(X) + H(Y) - H(X,Y) \geq 0$, como queríamos.

Note que, se X e Y são independentes, então $w_{xy} = 1$ para todo $(x, y) \in S \times T$, de modo que, por (13.13), temos $H(X) + H(Y) - H(X, Y) = 0$, ou seja, vale a igualdade em (13.8).

Resta apenas provarmos (13.9). Temos

$$\begin{aligned} H(X|Y) &= H(X,Y) - H(Y) \\ &= \left(\sum_{x\in S}\sum_{y\in T} p_{xy} \lg \frac{1}{p_{xy}}\right) - \left(\sum_{y\in T}\left(\sum_{x\in S} p_{xy}\right) \lg \frac{1}{p_y}\right) \\ &= \sum_{x\in S}\sum_{y\in T} p_{xy} \lg \frac{p_y}{p_{xy}} = \sum_{x\in S}\sum_{y\in T}\sum_{z\in U} p_{xyz} \lg \frac{p_y}{p_{xy}}. \end{aligned} \tag{13.15}$$

Ademais,

$$\begin{aligned} H(X|Y,Z) &= H(X,Y,Z) - H(Y,Z) \\ &= \left(\sum_{x\in S}\sum_{y\in T}\sum_{z\in U} p_{xyz} \lg \frac{1}{p_{xyz}}\right) - \left(\sum_{y\in T}\sum_{z\in U}\left(\sum_{x\in S} p_{xyz}\right) \lg \frac{1}{p_{yz}}\right) \\ &= \sum_{x\in S}\sum_{y\in T}\sum_{z\in U} p_{xyz} \lg \frac{p_{yz}}{p_{xyz}}. \end{aligned} \tag{13.16}$$

Assim,

$$H(X|Y) - H(X|Y,Z) = \sum_{x\in S}\sum_{y\in T}\sum_{z\in U} p_{xyz} \lg \left(\frac{p_y}{p_{xy}} \frac{p_{xyz}}{p_{yz}}\right). \tag{13.17}$$

Ponha $f(w) := w \lg w$ e

$$w_{xyz} := \frac{p_y}{p_{xy}} \frac{p_{xyz}}{p_{yz}} \tag{13.18}$$

para todo $(x, y, z) \in S \times T \times U$. Então

$$H(X|Y) - H(X|Y,Z) = \sum_{x\in S}\sum_{y\in T}\sum_{z\in U} \frac{p_{xy}p_{yz}}{p_y} f(w_{xyz}). \tag{13.19}$$

Temos

$$\sum_{x\in S}\sum_{y\in T}\sum_{z\in U} \frac{p_{xy}p_{yz}}{p_y} = \sum_{y\in T}\sum_{z\in U}\left(\sum_{x\in S} p_{xy}\right)\frac{p_{yz}}{p_y} = \sum_{y\in T}\sum_{z\in U} p_{yz} = 1. \tag{13.20}$$

Pela convexidade da função f e pela desigualdade (5.11) de Jensen, temos

$$\begin{aligned} H(X|Y) - H(X|Y,Z) &= \sum_{x\in S}\sum_{y\in T}\sum_{z\in U} \frac{p_{xy}p_{yz}}{p_y} f(w_{xyz}) \\ &\geq f\left(\sum_{x\in S}\sum_{y\in T}\sum_{z\in U} \frac{p_{xy}p_{yz}}{p_y} w_{xyz}\right) \\ &= f\left(\sum_{x\in S}\sum_{y\in T}\sum_{z\in U} p_{xyz}\right) = f(1) = 0, \end{aligned} \tag{13.21}$$

como queríamos. $\square$

A propriedade (13.8) é chamada de *subaditividade*.

É trivial obtermos o seguinte corolário usando indução e a propriedade (13.8) de subaditividade:

Corolário 13.1.1. *Sejam $X_1, \dots, X_n$ variáveis aleatórias. Então*

$$H(X_1, \dots, X_n) \leq \sum_{i=1}^{n} H(X_i). \tag{13.22}$$

Dado um número real $0 \leq p \leq 1$, defina

$$H(p) := p \lg \frac{1}{p} + (1-p) \lg \frac{1}{1-p}, \tag{13.23}$$

isto é, $H(p)$ é a entropia de uma variável aleatória de Bernoulli que tem probabilidade de sucesso p.

Corolário 13.1.2 (Kleitman, Shearer e Sturtevant [11])**.** *Seja $\emptyset \subset \mathcal{F} \subseteq \mathcal{P}([n])$ o conjunto de hiperarestas de um hipergrafo sobre $[n]$. Ponha*

$$p_i := \frac{d_{\mathcal{F}}(i)}{|\mathcal{F}|} = \frac{\big|\{F \in \mathcal{F} : i \in F\}\big|}{|\mathcal{F}|} \tag{13.24}$$

para todo $i \in [n]$. Então

$$|\mathcal{F}| \leq 2^{\sum_{i=1}^{n} H(p_i)}, \tag{13.25}$$

com igualdade se $\mathcal{F} = \mathcal{P}([n])$.

Demonstração. Considere o seguinte experimento aleatório. Sorteie um elemento $F \in \mathcal{F}$ uniformemente. Seja $X = (X_1, \dots, X_n)$ o vetor característico do F sorteado. Então

$$H(X_i) = \mathbb{P}[X_i = 1] \lg \frac{1}{\mathbb{P}[X_i = 1]} + \mathbb{P}[X_i = 0] \lg \frac{1}{\mathbb{P}[X_i = 0]} = H(p_i). \tag{13.26}$$

Pelo corolário 13.1.1 de subaditividade, temos

$$H(X) \leq \sum_{i=1}^{n} H(X_i) = \sum_{i=1}^{n} H(p_i). \tag{13.27}$$

Mas

$$\begin{aligned} H(X) &= \sum_{F_0 \in \mathcal{F}} \mathbb{P}[F = F_0] \lg \frac{1}{\mathbb{P}[F = F_0]} \\ &= \frac{1}{|\mathcal{F}|} \sum_{F \in \mathcal{F}} \lg |\mathcal{F}| = \lg |\mathcal{F}|. \end{aligned} \tag{13.28}$$

Assim,

$$|\mathcal{F}| \leq 2^{\sum_{i=1}^{n} H(p_i)}, \tag{13.29}$$

como queríamos.

É fácil ver que (13.25) vale com igualdade se $\mathcal{F} = \mathcal{P}([n])$, pois neste caso $p_i = 1/2$ e $H(p_i) = 1$ para todo $i \in [n]$. □

Corolário 13.1.3 (Frankl)**.** *Sejam $n \geq 1$ e $0 < p \leq 1/2$. Então*

$$\sum_{k \leq pn} \binom{n}{k} \leq 2^{H(p)n}. \tag{13.30}$$

Demonstração. Seja $\mathcal{F} := \{F \subseteq [n] : |F| \leq pn\}$. É fácil ver que $p_i := d_{\mathcal{F}}(i)/|\mathcal{F}|$ independe de $i \in [n]$, isto é, $p_i = p_j$ para todos $i, j \in [n]$.

Temos

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{d_{\mathcal{F}}(i)}{|\mathcal{F}|} = \frac{1}{|\mathcal{F}|} \sum_{i=1}^{n} d_{\mathcal{F}}(i). \tag{13.31}$$

Por contagem dupla, temos

$$\sum_{i=1}^{n} d_{\mathcal{F}}(i) = \sum_{F \in \mathcal{F}} |F| \leq |\mathcal{F}| pn. \tag{13.32}$$

Mas então

$$\sum_{i=1}^{n} p_i \leq pn. \tag{13.33}$$

Como todos os p_i's são iguais, temos $p_i \leq p$ para todo $i \in [n]$.

Por (13.25), temos então

$$\sum_{k \leq pn} \binom{n}{k} = |\mathcal{F}| \leq 2^{\sum_{i=1}^{n} H(p_i)}. \tag{13.34}$$

Observe ainda que

$$\frac{\mathrm{d}H}{\mathrm{d}x} = \lg \frac{1-x}{x}. \tag{13.35}$$

Assim, para $x > 0$, temos $\frac{\mathrm{d}H}{\mathrm{d}x} \geq 0$ se, e somente se, $(1-x)/x \geq 1$, o que vale para $0 < x \leq 1/2$. Assim, $H(x)$ é crescente no intervalo $0 < x \leq 1/2$.

Concluímos que

$$\sum_{k \leq pn} \binom{n}{k} = |\mathcal{F}| \leq 2^{\sum_{i=1}^{n} H(p_i)} \leq 2^{nH(p)}. \tag{13.36}$$

□

13.1 Uma generalização de subaditividade

Seja $X := (X_1, \ldots, X_n)$ uma variável aleatória tomando valores em $S := S_1 \times \cdots \times S_n$, onde S_i é um conjunto finito para todo i. Para cada $I \subseteq [n]$, defina $X(I) := (X_i)_{i \in I}$ como a projeção de X em I.

O seguinte lema é uma generalização de (13.22), isto é, de subaditividade:

Lema 13.2. *Sejam X e S como acima. Seja $\mathcal{G} := (G_\lambda)_{\lambda \in \Lambda}$, onde $G_\lambda \subseteq [n]$ para todo $\lambda \in \Lambda$ e Λ é um conjunto finito. Suponha que*

$$\deg_{\mathcal{G}}(i) := \big|\{\lambda \in \Lambda : i \in G_\lambda\}\big| \geq k \geq 1 \tag{13.37}$$

para todo $i \in [n]$. Então

$$H(X) \leq \frac{1}{k} \sum_{\lambda \in \Lambda} H\big(X(G_\lambda)\big). \tag{13.38}$$

Demonstração. A prova é por indução em k.

Para a base, suponha que $k = 1$. Podemos transformar $\mathcal{G}$ em uma partição $\mathcal{G}'$ de $[n]$ substituindo cada G_λ por um $G'_\lambda \subseteq G_\lambda$ adequado. Utilizando (13.7) e (13.22), obtemos

$$\sum_{\lambda \in \Lambda} H\big(X(G_\lambda)\big) \geq \sum_{\lambda \in \Lambda} H\big(X(G'_\lambda)\big) \geq H(X). \tag{13.39}$$

Suponha agora $k \geq 2$ e que a asserção é verdadeira para valores menores de k. Suponha primeiro que $[n] = G_{\lambda_0}$ para algum $\lambda_0 \in \Lambda$. Então não é difícil ver que $\mathcal{G}' := (G_\lambda)_{\lambda \in \Lambda'}$, onde $\Lambda' := \Lambda \setminus \{\lambda_0\}$, satisfaz $\deg_{\mathcal{G}'}(i) \geq k - 1$ para todo $i \in [n]$. Pela hipótese de indução, temos

$$\begin{aligned} \sum_{\lambda \in \Lambda} H\big(X(G_\lambda)\big) &= H(X) + \sum_{\lambda \in \Lambda'} H\big(X(G'_\lambda)\big) \\ &\geq H(X) + (k-1)H(X) = kH(X), \end{aligned} \tag{13.40}$$

e o resultado segue.

Suponha então que $G_\lambda \neq [n]$ para todo $\lambda \in \Lambda$. Se todos os G_λ são iguais, então temos necessariamente $G_\lambda = [n]$ para todo $\lambda \in \Lambda$, já que $\deg_{\mathcal{G}}(i) \geq 1$ para todo $i \in [n]$. Logo, existem $\lambda, \lambda' \in \Lambda$ com $G_\lambda \neq G_{\lambda'}$.

Por (13.9), temos

$$\begin{aligned} &H\Big(X(G_\lambda \setminus G_{\lambda'}) \Big| X(G_\lambda \cap G_{\lambda'}), X(G_{\lambda'} \setminus G_\lambda)\Big) \\ &\qquad\qquad \leq H\Big(X(G_\lambda \setminus G_{\lambda'}) \Big| X(G_\lambda \cap G_{\lambda'})\Big) \end{aligned} \tag{13.41}$$

O lado esquerdo de (13.41) é

$$\begin{aligned} &H\Big(X(G_\lambda \setminus G_{\lambda'}), X(G_\lambda \cap G_{\lambda'}), X(G_{\lambda'} \setminus G_\lambda)\Big) \\ &\qquad\qquad - H\Big(X(G_\lambda \cap G_{\lambda'}), X(G_{\lambda'} \setminus G_\lambda)\Big), \end{aligned} \tag{13.42}$$

que é igual a

$$H\Big(X(G_\lambda \cup G_{\lambda'})\Big) - H\Big(X(G_{\lambda'})\Big). \tag{13.43}$$

O lado direito de (13.41) é

$$H\Big(X(G_\lambda \setminus G_{\lambda'}), X(G_\lambda \cap G_{\lambda'})\Big) - H\Big(X(G_\lambda \cap G_{\lambda'})\Big), \tag{13.44}$$

que é igual a

$$H\Big(X(G_\lambda)\Big) - H\Big(X(G_\lambda \cap G_{\lambda'})\Big). \tag{13.45}$$

Logo,

$$H\Big(X(G_\lambda \cup G_{\lambda'})\Big) - H\Big(X(G_{\lambda'})\Big) \leq H\Big(X(G_\lambda)\Big) - H\Big(X(G_\lambda \cap G_{\lambda'})\Big), \tag{13.46}$$

ou seja,

$$H\Big(X(G_\lambda \cup G_{\lambda'})\Big) + H\Big(X(G_\lambda \cap G_{\lambda'})\Big) \leq H\Big(X(G_\lambda)\Big) + H\Big(X(G_{\lambda'})\Big), \tag{13.47}$$

isto é, vale a submodularidade.

Seja $\mathcal{G}' := (G'_\lambda)_{\lambda \in \Lambda}$ o sistema de conjuntos que obtemos ao substituir o par $G_\lambda, G_{\lambda'}$ por $G_\lambda \cup G_{\lambda'}, G_\lambda \cap G_{\lambda'}$. É evidente que $\deg_{\mathcal{G}'}(i) = \deg_{\mathcal{G}}(i) \geq k$ para todo $i \in [n]$. Ademais, por (13.47), temos

$$\sum_{\lambda \in \Lambda} H\big(X(G'_\lambda)\big) \leq \sum_{\lambda \in \Lambda} H\big(X(G_\lambda)\big). \tag{13.48}$$

Repetimos esse processo de substituição até obtermos um $\mathcal{G}'$ contendo $[n]$, o que eventualmente deve acontecer, já que $\deg_{\mathcal{G}}(i) \geq 1$ para todo $i \in [n]$. O resultado segue do caso já tratado anteriormente. □

Observe que este lema de fato generaliza (13.22): basta tomar $\Lambda := [n]$ e $G_\lambda := \{\lambda\}$ para todo $\lambda \in \Lambda$.

Corolário 13.2.1. *Sejam S e $\mathcal{G}$ como no lema 13.2. Em particular, suponha que $\deg_{\mathcal{G}}(i) \geq k \geq 1$ para todo $i \in [n]$. Seja $\mathcal{F} \subseteq S$. Para cada $\lambda \in \Lambda$, defina*

$$F_\lambda := \Big\{\underline{x}|_{G_\lambda} := (x_i)_{i \in G_\lambda} : \underline{x} \in \mathcal{F}\Big\} \tag{13.49}$$

como a projeção de $\mathcal{F}$ nas coordenadas de G_λ. Então

$$|\mathcal{F}|^k \leq \prod_{\lambda \in \Lambda} |F_\lambda|. \tag{13.50}$$

Demonstração. Escolha uniformemente um elemento X em $\mathcal{F}$. Pelo lema (13.2), temos

$$H(X) \leq \frac{1}{k} \sum_{\lambda \in \Lambda} H\big(X(G_\lambda)\big). \tag{13.51}$$

Como X tem distribuição uniforme, é fácil ver que $H(X) = \lg |\mathcal{F}|$. Ademais, por (13.6), temos

$$H\big(X(G_\lambda)\big) \leq \lg |F_\lambda|, \tag{13.52}$$

pois $X(G_\lambda)$ é uma variável aleatória tomando valores em F_λ. Logo,

$$k \lg |\mathcal{F}| \leq \sum_{\lambda \in \Lambda} \lg |F_\lambda|, \tag{13.53}$$

de onde o resultado segue imediatamente. □

Esse resultado é equivalente ao seguinte, que vamos provar informalmente.

Corolário 13.2.2 (Loomis e Whitney [12]). *Defina* corpo *como um compacto que é igual ao fecho de seu interior. Seja $B \subseteq \mathbb{R}^n$ um corpo com volume n-dimensional* $\mathrm{vol}_n(B)$. *Seja B_j a projeção de B no hiperplano gerado pelos vetores $\{e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n\}$, onde $e_1, \ldots, e_n$ são os vetores canônicos de $\mathbb{R}^n$. Seja* $\mathrm{vol}_{n-1}(B_j)$ *o volume $(n-1)$-dimensional de B_j. Então*

$$\big(\mathrm{vol}_n(B)\big)^{n-1} \leq \prod_{1 \leq j \leq n} \mathrm{vol}_{n-1}(B_j). \tag{13.54}$$

Demonstração. Tome $\mathcal{G} := (G_j)_{1 \leq j \leq n}$ com $G_j := [n] \setminus \{j\}$ para todo $j \in [n]$. Então $\deg_{\mathcal{G}}(j) = n - 1$ para todo $j \in [n]$.

A prova segue observando que podemos aproximar volumes de corpos por volumes da união de caixas, onde uma *caixa* é um conjunto da forma

$$[x_1 - \varepsilon/2, x_1 + \varepsilon/2] \times \cdots \times [x_n - \varepsilon/2, x_n + \varepsilon/2],$$

com os x_j em $\mathbb{Z}$. □

Corolário 13.2.3 (Chung, Frankl, Graham e Shearer [4]).
Seja $\mathcal{F} \subseteq \mathcal{P}([n])$. Seja $\mathcal{G} = (G_\lambda)_{\lambda \in \Lambda}$ como no lema 13.2. Em particular, $\deg_{\mathcal{G}}(i) \geq k \geq 1$ *para todo $i \in [n]$. Para cada $\lambda \in \Lambda$, ponha*

$$\mathcal{F}_\lambda := \mathrm{Tr}(\mathcal{F}, G_\lambda) := \{F \cap G_\lambda : F \in \mathcal{F}\}. \tag{13.55}$$

Então

$$|\mathcal{F}|^k \leq \prod_{\lambda \in \Lambda} |\mathcal{F}_\lambda|. \tag{13.56}$$

Demonstração. Seja $S := S_1 \times \cdots \times S_n$ com $S_i := \{0, 1\}$ para todo i. Seja $\mathcal{F}' \subseteq S$. Então cada elemento de $\mathcal{F}'$ é um vetor de n bits e existe uma relação biunívoca entre estes e subconjuntos de $[n]$. Assim, podemos considerar $\mathcal{F}$ como um subconjunto de S e aplicar o corolário 13.2.1, de onde este resultado segue imediatamente. □

Suponha que $F_1, \dots, F_M$ são grafos sobre $[t]$ tais que $F_i \cap F_j$ contém um triângulo para quaisquer i, j. Uma construção ingênua é a seguinte: fixamos um triângulo, que todos os grafos terão, e variamos todo o restante do conjunto de arestas. Obtemos assim uma construção com $M = 2^{\binom{t}{2}-3} = 2^{\binom{t}{2}}/8$. O seguinte corolário mostra uma cota superior para M.

Corolário 13.2.4 (Chung, Frankl, Graham e Shearer [4]**).**
Seja $\mathcal{F} \subseteq \mathcal{P}\left(\binom{[t]}{2}\right)$ tal que, se $F, F' \in \mathcal{F}$, então $F \cap F'$ contém um triângulo. Então

$$|\mathcal{F}| < \frac{1}{4} 2^{\binom{t}{2}}. \tag{13.57}$$

Demonstração. Seja $\mathcal{S} := \binom{[t]}{\lfloor t/2 \rfloor}$. Para cada $S \in \mathcal{S}$, tome

$$G_S := \binom{S}{2} \cup \binom{[t] \setminus S}{2}, \tag{13.58}$$

isto é, G_S é o conjunto de todas as arestas em $\binom{[t]}{2}$, exceto as que têm uma ponta em S e a outra fora de S. Note que

$$s := |G_S| = \binom{\lfloor t/2 \rfloor}{2} + \binom{\lceil t/2 \rceil}{2} \tag{13.59}$$

e

$$m := |\mathcal{S}| = \binom{t}{\lfloor t/2 \rfloor}. \tag{13.60}$$

Afirmamos que cada aresta $xy \in \binom{[t]}{2}$ pertence a exatamente $k := ms/\binom{t}{2}$ dos G'_S. Vamos usar contagem dupla. Considere um grafo bipartido H com as seguintes classes de cores: uma delas é $\mathcal{G} := (G_S)_{S \in \mathcal{S}}$, isto é, o conjunto dos G_S's, e a outra é $\binom{[t]}{2}$, isto é, o conjunto das arestas. Dado um $G_S \in \mathcal{G}$ e uma aresta $xy \in \binom{[t]}{2}$, ligamos estes dois vértices em H se $xy \in G_S$. Queremos mostrar então que $d_H(xy) = k$.

É fácil ver que $d_H(e) = d_H(f)$ para quaisquer arestas $e, f \in \binom{[t]}{2}$, pois nenhuma aresta é "privilegiada" em relação às demais. Seja d o valor comum de todos os $d_H(e)$, com $e \in \binom{[t]}{2}$. Já notamos acima que $d_H(G_S) = s$ para todo $G_S \in \mathcal{G}$. Como H é bipartido, temos

$$\binom{t}{2} d = \sum_{e \in \binom{[t]}{2}} d_H(e) = \sum_{G_S \in \mathcal{G}} d_H(G_S) = ms. \tag{13.61}$$

Segue que $d = ms/\binom{t}{2} = k$, como queríamos.

Temos assim $\deg_{\mathcal{G}}(xy) \geq k$ para todo $xy \in \binom{[t]}{2}$, ou seja, podemos aplicar o corolário 13.2.3.

Fixe $S \in \mathcal{S}$ e considere

$$\mathrm{Tr}(\mathcal{F}, G_S) = \mathcal{F}_S = \{F \cap G_S : F \in \mathcal{F}\}. \tag{13.62}$$

Afirmamos que $\mathcal{F}_S$ é intersectante. De fato, sejam $F, F' \in \mathcal{F}$. Por hipótese, $F \cap F'$ contém um triângulo. Seja T o conjunto de vértices de um tal triângulo. Se $T \subseteq S$ ou $T \subseteq [t] \setminus S$, então é claro que $(F \cap G_S) \cap (F' \cap G_S)$ contém o mesmo triângulo que $F \cap F'$, de modo que $F \cap G_S$ e $F' \cap G_S$ têm intersecção não-vazia. Suponha então que exatamente um dos vértices de T está em S, e os outros dois em $[t] \setminus S$. Neste caso, a aresta do triângulo que liga os dois vértices em $[t] \setminus S$ está tanto em $F \cap G_S$ quanto em $F' \cap G_S$, ou seja, a intersecção destes dois conjuntos é não-vazia. O caso em que T tem exatamente dois vértices em S é análogo. Segue que $\mathcal{F}_S$ é intersectante. Como $\mathcal{F}_S \subseteq \mathcal{P}(G_S)$, então

$$|\mathcal{F}_S| \leq 2^{|G_S|-1} = 2^{s-1}. \tag{13.63}$$

Aplicando finalmente o corolario 13.2.3, obtemos

$$|\mathcal{F}|^{ms/\binom{t}{2}} \leq \left(2^{s-1}\right)^m. \tag{13.64}$$

Portanto,

$$|\mathcal{F}| \leq \left(2^{ms-m}\right)^{\binom{t}{2}/ms} = 2^{\binom{t}{2} - \frac{1}{s}\binom{t}{2}}. \tag{13.65}$$

É fácil verificar que

$$\frac{1}{s}\binom{t}{2} = \frac{t(t-1)}{\lfloor t/2 \rfloor \left(\lfloor t/2 \rfloor - 1\right) + \lceil t/2 \rceil \left(\lceil t/2 \rceil - 1\right)} > 2. \tag{13.66}$$

Basta separar os casos t par e t ímpar. Assim,

$$|\mathcal{F}| \leq 2^{\binom{t}{2} - \frac{1}{s}\binom{t}{2}} < 2^{\binom{t}{2}-2} = \frac{1}{4} 2^{\binom{t}{2}}, \tag{13.67}$$

como queríamos. □

Exercício 19. *Seja $\mathcal{F} \subseteq \mathcal{P}\left(\binom{[t]}{2}\right)$ tal que, para quaisquer $F, F' \in \mathcal{F}$, não há vértice isolado em $F \cap F'$. Mostre que*

$$|\mathcal{F}| \leq 2^{\binom{t}{2} - t/2}. \tag{13.68}$$

Prove que a cota (13.68) é justa para todo t par.

[*Dica:* Para todo $j \in [t]$, considere $G_j := \left\{ jl : l \in [t] \setminus \{j\} \right\}$, a estrela de j em K_t.]

13.2 Entropia e programação linear

14 Conjuntos livres de somas

Referências

[1] N. Alon and P. Frankl. The maximum number of disjoint pairs in a family of subsets. *Graphs Combin.*, 1(1):13–21, 1985.

[2] B. Bollobás. On generalized graphs. *Acta Math. Acad. Sci. Hungar.*, 16:447–452, 1965.

[3] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988.

[4] F. R. K. Chung, R. L. Graham, P. Frankl, and J. B. Shearer. Some intersection theorems for ordered sets and graphs. *J. Combin. Theory Ser. A*, 43(1):23–37, 1986.

[5] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947.

[6] P. Erdős. Graph theory and probability. *Canad. J. Math.*, 11:34–38, 1959.

[7] P. Erdős and Z. Füredi. The greatest angle among n points in the d-dimensional Euclidean space. In *Combinatorial Mathematics (Marseille-Luminy, 1981)*, volume 75 of *North-Holland Math. Stud.*, pages 275–283. North-Holland, Amsterdam, 1983.

[8] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II*, pages 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10. North-Holland, Amsterdam, 1975.

[9] Z. Füredi and J. Kahn. On the dimensions of ordered sets of bounded degree. *Order*, 3(1):15–20, 1986.

[10] S. Janson. Poisson approximation for large deviations. *Random Structures Algorithms*, 1(2):221–229, 1990.

[11] D. J. Kleitman, J. Shearer, and D. Sturtevant. Intersections of k-element sets. *Combinatorica*, 1(4):381–384, 1981.

[12] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc.*, 55:961–962, 1949.

[13] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[14] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[15] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.

[16] K. Menger. Zur allgemeinen Kurventheorie. *Fundamenta Mathematicae*, 10:96–115, 1927.

[17] F. P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc.*, 30:264–286, 1930.

[18] E. Shamir and J. Spencer. Sharp concentration of the chromatic number on random graphs $G_{n,p}$. *Combinatorica*, 7(1):121–129, 1987.

[19] J. B. Shearer. The independence number of dense graphs with large odd girth. *Electron. J. Combin.*, 2:Note 2, approx. 3 pp. (electronic), 1995.

[20] W. T. Trotter. *Combinatorics and partially ordered sets*. Johns Hopkins Series in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, 1992. Dimension theory.

[21] P. Turán. On a theorem of Hardy and Ramanujan. *J. London Math. Soc.*, 9:274–276, 1934.

[22] P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

[23] Z. Tuza. Critical hypergraphs and intersecting set-pair systems. *J. Combin. Theory Ser. B*, 39(2):134–145, 1985.