離散数理工学　第 8 回

離散確率論：確率的離散システムの解析 (発展)

岡本 吉央
okamotoy@uec.ac.jp

電気通信大学

2024 年 12 月 24 日

最終更新：2024 年 12 月 16 日　12:18

今日の目標

確率的手法 を用いて，離散数学の定理を証明する

- ラムゼー理論
- 無向グラフの最大カット
- 箱と玉のモデル (負荷分散)

目次

Frank P. Ramsey

イギリスの思想家，数学者，経済学者 (1903–1930)

http://en.wikipedia.org/wiki/File:Frank_Plumpton_Ramsey.JPG

ラムゼー理論とは？

D. West の本 ('01) からの引用 (の試訳)

「ラムゼー理論」とは大きな構造の分割に関する研究を指す．典型的な結果は，分割のある類に特殊な部分構造が必ず生起するというものである．モツキンは「完全な無秩序は不可能である」ということばでこれを表現した．我々が考える対象は単に集合や数であり，...

原文："Ramsey theory" refers to the study of partitions of large structures. Typical results state that a special substructure must occur in some class of the partition. Motzkin described this by saying that "Complete disorder is impossible." The objects we consider are merely sets and numbers, ...

格言

物理学に「物理学的現象」，生物学に「生物学的現象」があるように
数学にも「数学的現象」が存在する

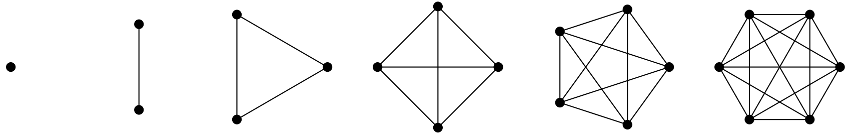
完全グラフ

無向グラフ $G = (V, E)$，正の整数 $n > 0$

定義：完全グラフとは？

G が **完全グラフ** であるとは，V のどの 2 頂点も辺で結ばれていること

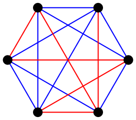
頂点数 n の完全グラフを K_n と表記する



K_1 K_2 K_3 K_4 K_5 K_6

完全グラフの辺集合の分割

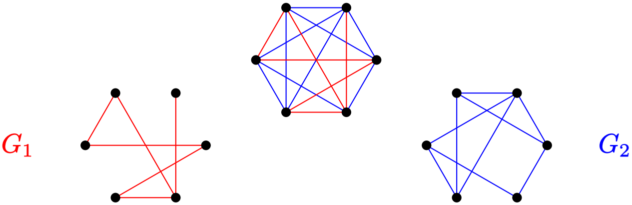
K_6 の辺集合を 2 分割して，2 つのグラフを得る



注：等分割をする必要はない

完全グラフの辺集合の分割

K_6 の辺集合を 2 分割して，2 つのグラフを得る



注：等分割をする必要はない

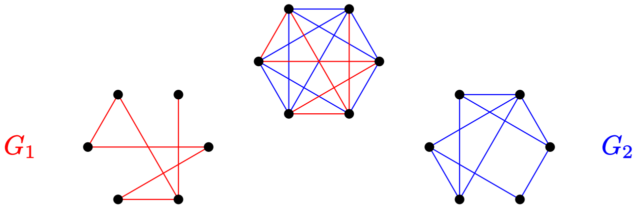
K_6 に対するラムゼー理論

定理：K_6 に対するラムゼー理論

K_6 の辺集合を任意に 2 分割してできたグラフ G_1, G_2 において

$$G_1 \text{ が } K_3 \text{ を含む} \quad \text{または} \quad G_2 \text{ が } K_3 \text{ を含む}$$

が成り立つ

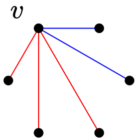


比喩的に，次のような言われ方もする

6 人出席者のいるパーティーでは，互いに知り合いである 3 人組か，
互いに知り合いではない 3 人組が必ず存在する

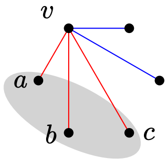
K_6 に対するラムゼー理論：証明 (1)

- K_6 の頂点を 1 つ任意に選んで，v とする
- v に接続する辺は 5 つ存在
- その中の 3 つは G_1 か G_2 に存在　(補足：鳩の巣原理)
- この 3 つが G_1 に存在する場合を考える
  (G_2 に存在する場合も同様)
- その 3 辺に接続する v 以外の頂点を a, b, c とする



K_6 に対するラムゼー理論：証明 (1)

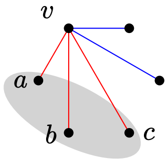
- K_6 の頂点を 1 つ任意に選んで，v とする
- v に接続する辺は 5 つ存在
- その中の 3 つは G_1 か G_2 に存在　(補足：鳩の巣原理)
- この 3 つが G_1 に存在する場合を考える
  (G_2 に存在する場合も同様)
- その 3 辺に接続する v 以外の頂点を a, b, c とする



K_6 に対するラムゼー理論：証明 (2) 場合分け

場合 1：a, b, c を結ぶ辺がすべて G_2 に含まれるとき

場合 2：そうではないとき

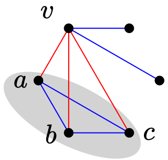


K_6 に対するラムゼー理論：証明 (2) 場合分け

場合 1：a, b, c を結ぶ辺がすべて G_2 に含まれるとき

- a, b, c を頂点集合とする K_3 が G_2 に含まれる

場合 2：そうではないとき



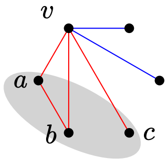
K_6 に対するラムゼー理論：証明 (2) 場合分け

場合 1：a, b, c を結ぶ辺がすべて G_2 に含まれるとき

- a, b, c を頂点集合とする K_3 が G_2 に含まれる

場合 2：そうではないとき

- a, b, c を結ぶ辺の 1 つは G_1 に含まれる
- それを $\{a, b\}$ であるとする (他の場合も同様)
- a, b, v を頂点とする K_3 が G_1 に含まれる $\square$



グラフに対するラムゼー理論とは？

グラフに対するラムゼー理論とは？

- 完全グラフの辺集合を分割して，グラフ $G_1, \ldots, G_r$ を得る
- このとき，その中のどれかがある大きさの完全グラフを含む

先ほどの例

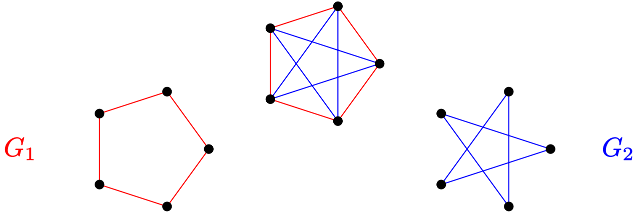
- 頂点数 6 の完全グラフの辺集合を 2 分割
- このとき，どちらかが頂点数 3 の完全グラフを含む

注意

頂点数 6 の完全グラフの辺集合 2 分割でこれが成り立つので，
頂点数 $7, 8, 9, \ldots$ の完全グラフの辺集合を 2 分割しても
どちらかは頂点数 3 の完全グラフを必ず含む

頂点数 5 だとどうか？

K_5 の辺集合を 2 分割しても K_3 は含まれないかもしれない



この意味で,「6」が **極値** になっている

ラムゼー数

正の整数 k, ℓ

定義：ラムゼー数 $R(k, \ell)$ とは？

K_n の辺集合を 2 つに分けてグラフ G_1, G_2 を任意に作ったとき

G_1 が K_k を含む　または　G_2 が K_ℓ を含む

が成り立つような 最小 の n

先ほどの場合に対応するのは：$R(3,3) = 6$

- $\boxed{R(3,3) \le 6}$：
  K_6 の辺集合を 2 つに分けて G_1, G_2 を任意に作ると
  「G_1 が K_3 を含む，または，G_2 が K_3 を含む」が成り立つ
- $\boxed{R(3,3) \ge 6}$：
  K_5 の辺集合を 2 つに分けて G_1, G_2 をうまく作ると
  「G_1 が K_3 を含まず，かつ，G_2 が K_3 を含まない」が成り立つ

ラムゼー数の上界と下界

正の整数 k, ℓ

定義：ラムゼー数 $R(k,\ell)$ とは？

K_n の辺集合を 2 つに分けてグラフ G_1, G_2 を任意に作ったとき

G_1 が K_k を含む　または　G_2 が K_ℓ を含む

が成り立つような **最小** の n

$R(k,\ell) = N$ を証明するには⋯

- $\boxed{R(k,\ell) \leq N \text{ の証明}}$：
  K_N の辺集合を 2 つに分けて G_1, G_2 を任意に作ると
  「G_1 が K_k を含む，または，G_2 が K_ℓ を含む」が成り立つ
- $\boxed{R(k,\ell) \geq N}$：
  K_{N-1} の辺集合を 2 つに分けて G_1, G_2 をうまく作ると
  「G_1 が K_k を含まず，かつ，G_2 が K_ℓ を含まない」が成り立つ

ラムゼー数に対する疑問

疑問

任意の正整数 k, ℓ に対して，十分に大きな N を考えると
K_N の辺集合を 2 つに分けて G_1, G_2 を任意に作ったとき

$$G_1 \text{ が } K_k \text{ を含む} \quad \text{または} \quad G_2 \text{ が } K_\ell \text{ を含む}$$

は成り立つのか？

疑問：別の言い方

任意の正整数 k, ℓ に対して，$R(k, \ell)$ は存在するのか？

ラムゼー数の存在性

定理：グラフに対するラムゼーの定理

任意の正整数 k, ℓ に対して，ある正整数 N が存在して
K_N の辺集合を 2 分割して G_1, G_2 を任意に作ると

$$G_1 \text{ が } K_k \text{ を含む} \quad \text{または} \quad G_2 \text{ が } K_\ell \text{ を含む}$$

が成り立つ

わざわざ難しく書くと

$\forall$ 正整数 k, ℓ

$\exists$ 正整数 N

$\forall$ K_N の辺集合の 2 分割から作られる G_1, G_2：

G_1 が K_k を含む $\vee$ G_2 が K_ℓ を含む

ラムゼー数の存在性：証明

証明は以下の再帰式に基づいた帰納法

性質：ラムゼー数に対する再帰式

次の式が成立する

- 任意の $k \geq 1$ に対して，$R(k,1) = 1$
- 任意の $\ell \geq 1$ に対して，$R(1,\ell) = 1$
- 任意の $k,\ell > 1$ に対して，$R(k,\ell) \leq R(k-1,\ell) + R(k,\ell-1)$

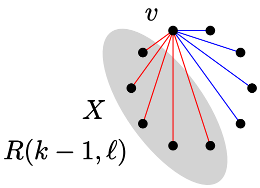
これが証明できれば，グラフに対するラムゼーの定理の証明になる

- $R(k,1) = 1$ と $R(1,\ell) = 1$ は簡単
- 主題は最後の不等式

ラムゼー数に対する再帰式：証明 (1)

$N = R(k, \ell - 1) + R(k - 1, \ell)$ とする

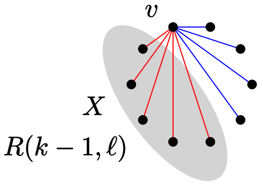
- K_N の頂点を 1 つ任意に選んで，v とする
- v に接続する辺は $N - 1$ 個存在
- その中の $R(k - 1, \ell)$ 個が G_1 に存在するか，または，
  その中の $R(k, \ell - 1)$ 個が G_2 に存在 (補足：鳩の巣原理)



ラムゼー数に対する再帰式：証明 (2)

v に接続する辺の中の $R(k-1,\ell)$ 個が G_1 に存在するとき

- それらの辺に接続する v 以外の頂点の集合を X とする
- $|X| \geq R(k-1,\ell)$
- 帰納法の仮定から，X の頂点を見ると
  1. その中の $k-1$ 頂点を結ぶ辺がすべて G_1 に含まれる，または
  2. その中の ℓ 頂点を結ぶ辺がすべて G_2 に含まれる



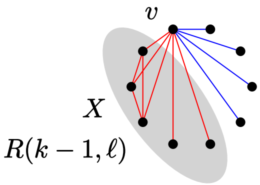
ラムゼー数に対する再帰式：証明 (3)

場合 1：X の中の $k-1$ 頂点を結ぶ辺がすべて G_1 に含まれる

- G_1 において，それら $k-1$ 個の頂点と v が K_k を作る

場合 2：X の中の ℓ 頂点を結ぶ辺がすべて G_2 に含まれる

- G_2 において，それら ℓ 個の頂点が K_ℓ を作る



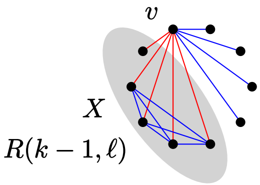
ラムゼー数に対する再帰式：証明 (3)

場合 1：X の中の $k-1$ 頂点を結ぶ辺がすべて G_1 に含まれる

- G_1 において，それら $k-1$ 個の頂点と v が K_k を作る

場合 2：X の中の ℓ 頂点を結ぶ辺がすべて G_2 に含まれる

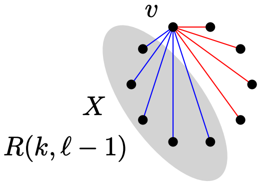
- G_2 において，それら ℓ 個の頂点が K_ℓ を作る



ラムゼー数に対する再帰式：証明 (4)

v に接続する辺の中の $R(k,\ell-1)$ 個が G_2 に存在するとき

- それらの辺に接続する v 以外の頂点の集合を X とする
- $|X| \geq R(k,\ell-1)$
- 帰納法の仮定から，X の頂点を見ると
  1. その中の k 頂点を結ぶ辺がすべて G_1 に含まれる，または
  2. その中の $\ell-1$ 頂点を結ぶ辺がすべて G_2 に含まれる



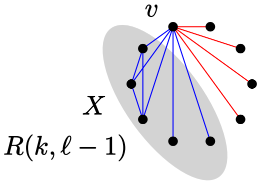
ラムゼー数に対する再帰式：証明 (5)

場合 1：X の中の k 頂点を結ぶ辺がすべて G_1 に含まれる

- G_1 において，それら k 個の頂点が K_k を作る

場合 2：X の中の $\ell - 1$ 頂点を結ぶ辺がすべて G_2 に含まれる

- G_2 において，それら $\ell - 1$ 個の頂点と v が K_ℓ を作る □



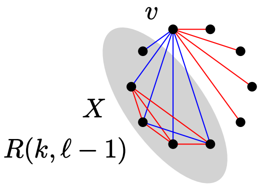
ラムゼー数に対する再帰式：証明 (5)

場合 1：X の中の k 頂点を結ぶ辺がすべて G_1 に含まれる

- G_1 において，それら k 個の頂点が K_k を作る

場合 2：X の中の $\ell - 1$ 頂点を結ぶ辺がすべて G_2 に含まれる

- G_2 において，それら $\ell - 1$ 個の頂点と v が K_ℓ を作る □



ラムゼー数の上界

ラムゼー数に対する再帰式 (再掲)

次の式が成立する

- 任意の $k \geq 1$ に対して，$R(k, 1) = 1$
- 任意の $\ell \geq 1$ に対して，$R(1, \ell) = 1$
- 任意の $k, \ell > 1$ に対して，$R(k, \ell) \leq R(k-1, \ell) + R(k, \ell-1)$

この式から次の上界が得られる (演習問題)

$$R(k, \ell) \leq \binom{k+\ell-2}{k-1}$$

ラムゼー数の上界：$k=\ell$ の場合

特に，$k=\ell$ の場合

$$R(k,k) \leq \binom{k+k-2}{k-1} = \binom{2k-2}{k-1} \leq \left(\frac{\mathrm{e}(2k-2)}{k-1}\right)^{k-1} \leq (2\mathrm{e})^k$$

二項係数の性質：簡単な評価 （第 1 回講義の復習）

任意の正整数 $a \geq b$ に対して，

$$\left(\frac{a}{b}\right)^b \leq \binom{a}{b} \leq \left(\frac{\mathrm{e}a}{b}\right)^b$$

ラムゼー数に対する上界の表

$\binom{k+\ell-2}{k-1}$ の表

$k \backslash \ell$	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7
3	1	3	6	10	15	21	28
4	1	4	10	20	35	56	84
5	1	5	15	35	70	126	210
6	1	6	21	56	126	252	462
7	1	7	28	84	210	462	924

小さなラムゼー数の表

小さなラムゼー数の表

$k \backslash \ell$	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7
3	1	3	6	9	14	18	23
4	1	4	9	18	25	36–41	49–61
5	1	5	14	25	43–48	59–87	80–143
6	1	6	18	36–41	59–87	102–165	115–298
7	1	7	23	49–61	80–143	115–298	205–540

(Radziszowski '24 によるまとめ)

未解決問題

この表にあるギャップを埋めよ
($R(5,5) = 43$ であると予想されている)

ラムゼー数の下界

疑問？

ラムゼー数 $R(k,k)$ を下から評価できるか？

復習；$R(3,3) \geq 6$ を示すときに，どうしたか

- K_5 の辺集合を 2 つに分けて G_1, G_2 をうまく作ると
「G_1 が K_3 を含まず，かつ，G_2 が K_3 を含まない」が成り立つ

ラムゼー数の下界：定理

定理：ラムゼー数の下界

任意の正整数 k に対して

$$R(k, k) > 2^{k/2-1}$$

証明で行いたいこと

- $K_{2^{k/2-1}}$ の辺集合を 2 つに分けて G_1, G_2 をうまく作ると「G_1 が K_k を含まず，かつ，G_2 が K_k を含まない」が成り立つ

任意の k に対してうまく作るのは難しい

$\rightsquigarrow$ **確率を使って**，うまい作り方の **存在だけを証明する**

ラムゼー数の下界：確率的証明 (1)

$N = 2^{k/2-1}$ とする

- K_N の各辺を G_1 か G_2 に確率 $1/2$ で独立に振り分けて，G_1 と G_2 を構成する (G_1, G_2 は確率変数)
- このとき，k 個の頂点 $v_1, v_2, \ldots, v_k$ を固定したとき，

$$\Pr(G_1 \text{ で } v_1, v_2, \ldots, v_k \text{ が } K_k \text{ になる}) = \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

$$\Pr(G_2 \text{ で } v_1, v_2, \ldots, v_k \text{ が } K_k \text{ になる}) = \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

- したがって，

$$\Pr(G_1 \text{ か } G_2 \text{ が } K_k \text{ を含む}) \leq \binom{N}{k} \cdot 2 \cdot \left(\frac{1}{2}\right)^{\binom{k}{2}}$$

ラムゼー数の下界：確率的証明 (2)

- したがって，

$$\begin{aligned}
\Pr(G_1 \text{ も } G_2 \text{ も } K_k \text{ を含まない}) &\geq 1 - 2\binom{N}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}} \\
&= 1 - 2\binom{2^{k/2-1}}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}} \\
&\geq 1 - 2\left(2^{k/2-1}\right)^{k}\left(\frac{1}{2}\right)^{\binom{k}{2}} \\
&= 1 - 2^{1+(k/2-1)k-\binom{k}{2}} \\
&= 1 - 2^{1-3k/2} > 0
\end{aligned}$$

ただし，$1 + \left(\frac{k}{2} - 1\right)k - \binom{k}{2} = 1 + \frac{k^2}{2} - k - \frac{k^2}{2} - \frac{k}{2} = 1 - \frac{3}{2}k$

ラムゼー数の下界：確率的証明 (2)

したがって,

$$\Pr(G_1 \text{ も } G_2 \text{ も } K_k \text{ を含まない}) > 0$$

- つまり, G_1 も G_2 も K_k を含まないように,
  辺を振り分ける方法が存在する
- $\therefore R(k,k) > N = 2^{k/2-1}$ $\square$

注

確率空間 (標本空間) が有限で, $\Pr(\text{事象 } A) > 0 \Rightarrow$
事象 A に属する要素が存在する

補足：ラムゼー数の上界と下界

ここまで証明したこと

$$\sqrt{2}^k \lesssim R(k,k) \lesssim (2\mathrm{e})^k$$

上界にスターリングの公式を使うと証明できること

$$\sqrt{2}^k \lesssim R(k,k) \lesssim \frac{1}{\sqrt{k}} 4^k$$

知られている中でもっともよい下界と上界

$$k\sqrt{2}^k \lesssim R(k,k) \lesssim \frac{1}{k^{c \log k / \log\log k}} 4^k$$

(下界は Spencer '75, 上界は Conlon '09, c はある定数)

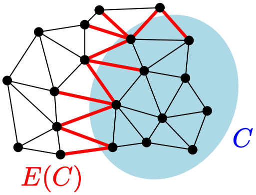
目次

無向グラフの最大カット問題

無向グラフ $G = (V, E)$

定義：無向グラフのカットとは？

- G の **カット** とは頂点部分集合 $C \subseteq V$ で $C \neq \emptyset, V$ であるもの
- カット C の辺集合 $E(C)$ とは,
  C の頂点と $V - C$ の頂点を結ぶ辺をすべて集めてできるもの



最大カット問題：$|E(C)|$ ができる限り大きいカット C を見つけたい
(量子アニーリングでよく出てくる問題)

無向グラフの最大カット問題

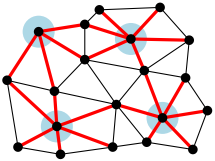
定理

辺を持つ任意の無向グラフ $G = (V, E)$ に対して,

$$|E(C)| \geq \frac{1}{2}|E|$$

を満たすカット C が存在する

例： $|E| = 44$, $|E(C)| = 22$



無向グラフの最大カット問題：証明 (1)

次のように頂点部分集合 $S \subseteq V$ を確率的に構成する (S は確率変数)

- 各頂点を独立に確率 $\frac{1}{2}$ で S に含める

このとき，任意の辺 $e \in E$ に対して，

$$\Pr(e \in E(S)) = \frac{1}{2}$$

無向グラフの最大カット問題：証明 (1)

次のように頂点部分集合 $S \subseteq V$ を確率的に構成する (S は確率変数)

- 各頂点を独立に確率 $\frac{1}{2}$ で S に含める

このとき，任意の辺 $e \in E$ に対して，

$$\Pr(e \in E(S)) = \frac{1}{2}$$

ここで，任意の辺 $e \in E$ に対して，次の確率変数 X_e を考える

$$X_e = \begin{cases} 1 & (e \in E(S)), \\ 0 & (e \notin E(S)) \end{cases}$$

このとき，$\mathrm{E}[X_e] = \frac{1}{2}$

無向グラフの最大カット問題：証明 (2)

したがって，

$$\mathsf{E}[|E(S)|] = \mathsf{E}\left[\sum_{e \in E} X_e\right] = \sum_{e \in E} \mathsf{E}[X_e] = \frac{1}{2}|E|$$

無向グラフの最大カット問題：証明 (2)

したがって，

$$\mathsf{E}[|E(S)|] = \mathsf{E}\left[\sum_{e \in E} X_e\right] = \sum_{e \in E} \mathsf{E}[X_e] = \frac{1}{2}|E|$$

つまり，ある部分集合 $C \subseteq V$ に対して，

$$|E(C)| \geq \frac{1}{2}|E|$$

が成り立ち，そのような C は空集合でも V でもないので，
G のカットである $\square$

証明からアルゴリズムへ

S をランダムに作ると，$\mathrm{E}[|E(S)|] = \dfrac{1}{2}|E|$ になる

アルゴリズム

各頂点を確率 $\dfrac{1}{2}$ で独立に S に入れて，S を出力する

問題点：$|E(S)| \geq \dfrac{1}{2}|E|$ が保証されるわけではない

証明からアルゴリズムへ

S をランダムに作ると，$\mathsf{E}[|E(S)|] = \frac{1}{2}|E|$ になる

アルゴリズム

各頂点を確率 $\frac{1}{2}$ で独立に S に入れて，S を出力する

問題点：$|E(S)| \geq \frac{1}{2}|E|$ が保証されるわけではない

解決策

S を繰り返し選んで，$|E(S)|$ を最大とするものを出力する

どれだけ繰り返すとよいのか？

確率増幅 (1)

$\varepsilon > 0$ を十分小さい正の実数とする

- マルコフの不等式から，S を 1 回選んだとき

$$
\begin{aligned}
&\Pr\left(|E(S)| \leq \left(\frac{1}{2} - \varepsilon\right)|E|\right) \\
&= \Pr\left(|E| - |E(S)| \geq |E| - \left(\frac{1}{2} - \varepsilon\right)|E|\right) \\
&\leq \frac{\mathrm{E}[|E| - |E(S)|]}{\left(\frac{1}{2} + \varepsilon\right)|E|} = \frac{|E| - \mathrm{E}[|E(S)|]}{\left(\frac{1}{2} + \varepsilon\right)|E|} = \frac{|E| - \frac{1}{2}|E|}{\left(\frac{1}{2} + \varepsilon\right)|E|} \\
&= \frac{\frac{1}{2}|E|}{\left(\frac{1}{2} + \varepsilon\right)|E|} = \frac{1}{1 + 2\varepsilon} \leq \mathrm{e}^{-\varepsilon}
\end{aligned}
$$

性質 (演習問題)

$0 \leq x \leq 1$ のとき，$\dfrac{1}{1+2x} \leq \mathrm{e}^{-x}$

確率増幅 (2)

$\varepsilon > 0$ を十分小さい正の実数とする

- k 回続けて S を作るとき
$$\Pr\left(k \text{ 回続けて } |E(S)| \le \left(\frac{1}{2} - \varepsilon\right)|E|\right) \le (\mathrm{e}^{-\varepsilon})^k = \mathrm{e}^{-\varepsilon k}$$
- つまり，k 回続けて S を作るとき
$$\Pr\left(1 \text{ 回は } |E(S)| > \left(\frac{1}{2} - \varepsilon\right)|E|\right) \ge 1 - \mathrm{e}^{-\varepsilon k}$$

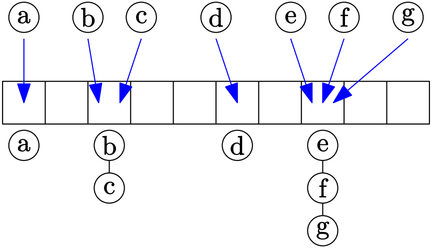
一般論：確率増幅

失敗確率が $p < 1$ のアルゴリズムを k 回繰り返せば，
失敗確率を p^k にできる

目次

ハッシュ：連鎖法 (チェイニング，オープンハッシング)

同じハッシュ値を持つ要素は，連結リストで保持する



疑問

一番長い連結リストの長さはどれほどか？

- 一番長い連結リストの長さの期待値

箱と玉のモデル

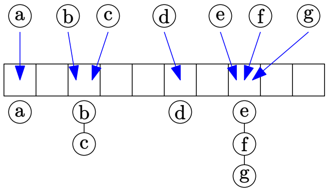
設定

- 箱の集合 $K = \{1, \ldots, k\}$
- 玉の集合 $N = \{1, \ldots, n\}$
- 玉を箱へランダムに入れる (独立に)
- $\Pr(\text{玉 } i \text{ を箱 } j \text{ に入れる}) = \dfrac{1}{k}$
- 単純化として，$n = k$ で，n は十分大きいと仮定する

問題

- 箱に入る玉の数の最大値は？

注意：「箱に入る玉の数の最大値」は確率変数



箱と玉のモデル

箱 j に玉が ℓ 個入る確率は？

$$\begin{aligned}\Pr(\text{箱 } j \text{ に玉が } \ell \text{ 個入る}) &\le \binom{n}{\ell}\left(\frac{1}{k}\right)^{\ell} \\ &\le \left(\frac{\mathrm{e}n}{\ell}\right)^{\ell}\left(\frac{1}{k}\right)^{\ell} \\ &= \left(\frac{\mathrm{e}n}{\ell k}\right)^{\ell} = \left(\frac{\mathrm{e}}{\ell}\right)^{\ell}\end{aligned}$$

二項係数の性質：簡単な評価 (第 1 回講義の復習)

任意の正整数 $a \ge b$ に対して，

$$\left(\frac{a}{b}\right)^{b} \le \binom{a}{b} \le \left(\frac{\mathrm{e}a}{b}\right)^{b}$$

箱と玉のモデル

ここで，$\ell \geq 3\dfrac{\ln n}{\ln\ln n}$ とすると，$\mathrm{e} < 3$ なので，

$$
\begin{aligned}
\left(\frac{\mathrm{e}}{\ell}\right)^{\ell} &\leq \left(\frac{\mathrm{e}}{3\frac{\ln n}{\ln\ln n}}\right)^{3\frac{\ln n}{\ln\ln n}} = \left(\frac{\mathrm{e}\ln\ln n}{3\ln n}\right)^{3\frac{\ln n}{\ln\ln n}} \leq \left(\frac{\ln\ln n}{\ln n}\right)^{3\frac{\ln n}{\ln\ln n}} \\
&= \left(\mathrm{e}^{\ln\left(\frac{\ln\ln n}{\ln n}\right)}\right)^{3\frac{\ln n}{\ln\ln n}} = \left(\mathrm{e}^{\ln\ln\ln n - \ln\ln n}\right)^{3\frac{\ln n}{\ln\ln n}} \\
&= \mathrm{e}^{3\frac{\ln\ln\ln n}{\ln\ln n}\ln n - 3\ln n}
\end{aligned}
$$

ここで，n が十分大きいので，$3\dfrac{\ln\ln\ln n}{\ln\ln n} \leq 1$ となり，

$$
\left(\frac{\mathrm{e}}{\ell}\right)^{\ell} \leq \mathrm{e}^{\ln n - 3\ln n} = \frac{1}{n^2}
$$

補足

関数 $x \mapsto \frac{1}{x^x}$ は $x \geq 1$ において単調減少

箱と玉のモデル

したがって，$\ell \geq 3\dfrac{\ln n}{\ln\ln n}$ のとき，

$$\Pr(\text{箱 } j \text{ に玉が } \ell \text{ 個入る}) \leq \left(\frac{e}{\ell}\right)^{\ell} \leq \frac{1}{n^2}$$

したがって，

$$\begin{aligned}
&\Pr(\text{玉が } \ell \text{ 個入る箱が存在})\\
&= \Pr(\text{箱 } 1 \text{ に玉が } \ell \text{ 個入る，または，}\ldots\text{，または箱 } n \text{ に玉が } \ell \text{ 個入る})\\
&\leq \sum_{j=1}^{n} \Pr(\text{箱 } j \text{ に玉が } \ell \text{ 個入る}) \leq n \cdot \frac{1}{n^2} = \frac{1}{n}
\end{aligned}$$

したがって，

$$\Pr\left(\text{どの箱にも高々 } 3\frac{\ln n}{\ln\ln n} \text{ 個しか玉がない}\right) \geq 1 - \frac{1}{n}$$

箱と玉のモデル：まとめ

設定

- 箱の集合 $K = \{1, \ldots, k\}$
- 玉の集合 $N = \{1, \ldots, n\}$
- 玉を箱へランダムに入れる (独立に)
- $\Pr(\text{玉 } i \text{ を箱 } j \text{ に入れる}) = \frac{1}{k}$
- 単純化として，$n = k$ で，n は十分大きいと仮定する

問題

- 箱に入る玉の数の最大値は？

回答

- 箱に入る玉の数の最大値は高い確率で $3\frac{\ln n}{\ln \ln n}$ 以下
  (高い確率で $= n \to \infty$ のとき確率 1 で)

箱と玉のモデル：他の問題との関係

「箱と玉のモデル」における文脈で言い直すと…

クーポン収集問題

すべての箱に少なくとも 1 つ玉が入るようになるまで，
入れる玉の総数は何か？

誕生日のパラドックス

ある箱に玉が複数入る確率は何か？

その他の様々な問題も考えることができ，考えられている

目次

1. ラムゼー理論
2. 無向グラフのカット
3. 箱と玉のモデル：負荷分散
4. 今日のまとめ

今日の目標

今日の目標

確率的手法 を用いて，離散数学の定理を証明する

- ラムゼー理論
- 無向グラフの最大カット
- 箱と玉のモデル (負荷分散)

格言

確率は離散数学において **当たり前の道具** として使われている

目次