2025 年 5 月 19 日

神戸大学 理学研究科数学専攻/理学部数学科

特別講義 統計学 B 演習課題 (2025 年 7 月 31 日締め切り)

滋賀大学 データサイエンス学部
来嶋 秀治 (Shuji Kijima)

注意 1: 各問題の最後の [ ] 内の数字は配点である．得点が 100 点を超えた場合は 100 点とする．
注意 2: Dropbox に提出する `https://www.dropbox.com/request/GpdPVong39CRzfjrpJxM`
複数ファイル（写真など）は zip 等で圧縮して一つのファイルにすること．50MB 以内.
学籍番号，氏名を明記すること．わからない場合は成績評価を行わない．メールアドレスも任意で記載.
注意 3: 文献やインターネット，生成 AI などで調べてよいが，自分の理解の範囲で証明を記せ.（生成 AI を用いた場合でも）書籍，論文，URL など，出典あるいは利用した定理の根拠となる文献を明記のこと.
計算機（プログラム，数値計算ソフトなど）も用いてよい.
注意 4: レポートの返却は行わない.
講義資料: `https://shuji-kijima.com/kobe25/index.html`
連絡先: shuji-kijima@biwako.shiga-u.ac.jp

1．講義の話題から

1.1. Hoeffding の不等式「$X_1, \ldots, X_n$ は相互に独立で，確率 1 で $a_i \leq X_i \leq b_i$ を満たす．$Y_n := X_1 + \cdots + X_n$ の期待値は μ とする．任意の $t > 0$ に対して，

$$\Pr[|Y_n - \mu| > t] < 2\exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)}\right)$$

が成り立つ.」を示せ．[20]

1.2. $2 \times n$ 分割表の ϵ-近似一様サンプリングが期待計算時間 $\mathrm{O}(n^2(\log n - \log \epsilon))$ で得られるとき，$2 \times n$ 分割表の FPRAS を設計し，その期待計算時間の上界を求めよ．[20]

1.3. 簡単のため，遷移確率行列 P は既約とする．以下の各問いに答えよ．[10]
(1) $p_{ii} > 0$ を満たす状態 $i \in \{1, \ldots, n\}$ がひとつでも存在すれば非周期的であることを示せ.
(2) $p_{ij} > 0$ かつ $p_{ji} > 0$ を満たす状態対 i, j が一組でも存在すれば，周期は 2 以下であることを示せ.
(3) 上記の (1),(2) を踏まえ，周期 3 の遷移確率行列の例をひとつあげよ.

1.4. Metropolis-Hastings 法を記せ．Metropolis-Hastings 法が reversible であることを示せ．[10]

1.5. 有限格子上の lazy RW の mixing time：グラフ $G = (\Omega, E)$ は $\Omega = \{0, 1, \ldots, k\}^n$，$E = \{\{x, y\} \mid |x \oplus y| = 1\}$ とし，$N(x) = \{y \mid \{x, y\} \in E\}$ とする．ランダムウォークの推移確率を

$$P(x, y) = \begin{cases} \frac{1}{4n} & (\{x, y\} \in E \text{ の場合}) \\ \frac{1}{2} + \frac{2n - N(x)}{4n} & (x = y \text{ の場合}) \\ 0 & (\text{それ以外}) \end{cases}$$

と定める．以下の各問いに答えよ．[100]
(1†) coupling 法を用いて mixing time の上界を求めよ.
(2) path coupling 法を用いて mixing time の上界を求めよ.
(3†) 固有値を用いて mixing time の上界を求めよ.
(4) multi-commodity flow を用いて mixing time の上界を求めよ.
(5) conductance を用いて mixing time の**下界**を求めよ.

1.6. $G = (U, V; E)$ は $E \subseteq U \times V$ を枝集合とする二部グラフとし，$|U| = |V| = n$，$|E| = m$ とする．$\mathcal{M}$ は G 上のマッチング全体の集合とし，$\mathcal{M}_k$ は G 上のサイズ k のマッチング全体の集合とする．[100]
(1) $|\mathcal{M}|$ に対する FPRAS を設計せよ．
(2) $\mathcal{M}_n \cup \mathcal{M}_{n-1}$ 上の一様分布に従うマルコフ連鎖を設計せよ．
(3) (2) で設計したマルコフ連鎖の mixing time の上界を求めよ．
(4) $|\mathcal{M}_n \cup \mathcal{M}_{n-1}|$ に対する FPRAS を設計せよ．
(5) $\frac{|\mathcal{M}_n|}{|\mathcal{M}_{n-1}|} \leq 2^{-\frac{n-2}{4}}$ となる例を見つけよ．
1.7. おねえさん問題のマルコフ連鎖について以下の問いに答えよ．[100]
(1) 既約性を示せ．
(2†) mixing time の上界を求めよ．

2．調査課題（サーベイ論文）

下記のテーマについて文献/論文を調べ，モデルを定義し，主要な定理を示せ．レポート本文は原則 8 ページ以下とし，証明や追試実験などで必要があれば 8 ページを超えて付録 (appendix) をつけてよい．言語は日本語または英語とする．**引用の方法は通常のサーベイ論文の基準に準拠すること**．

2.1. ページランク [100]

2.2. Azuma-Hoeffding 不等式 [100]

2.3. Probabilistic counting[100]

2.4. Expander グラフ [100]

2.5. Hitting time と cover time[100]

2.6. Balabasi-Albert モデル [100]

2.7. Galton-Watson 木 [100]

3．自由課題（サーベイ論文）

広い意味で MCMC 法に関する話題について，論文を調査し報告せよ．レポートでは，研究の動機，関連研究，その研究の結果，成果のカギとなるアイデアを述べよ．レポート本文は原則 8 ページ以下とし，証明や追試実験などで必要があれば 8 ページを超えて付録 (appendix) をつけてよい．言語は日本語または英語とする．**引用の方法は通常のサーベイ論文の基準に準拠すること**．[100]

参考文献

[1] Olle Häggström, Finite Markov Chains and Algorithmic Applications (London Mathematical Society Student Texts, Series Number 52), Cambridge University Press, 2002.
[2] Alistair Sinclair, Algorithms for Random Generation and Counting: A Markov Chain Approach, Springer, 1993
[3] David A. Levin and Yuval Peres, Markov Chains and Mixing Times, second edition, American Mathematical Society, 2017. `https://pages.uoregon.edu/dlevin/MARKOV/`
[4] Mark Jerrum, Counting, Sampling and Integrating: Algorithms and Complexity, Lectures in Mathematics. ETH Zuerich, Springer, 2013.
[5] Bernhard Korte and Jens Vygen, Combinatorial Optimization: Theory and Algorithms, 6th ed., Springer, 2018.
[6] Michael Sipser, Introduction to the theory of computation, 3rd ed., Cengage Learning, 2012.