ערבוב קלפים

הקדמה – שרשראות מרקוב

הגדרה 1. שרשרת מרקוב (בזמן בדיד) היא תהליך סטוכסטי $X_1, X_2, \ldots$ של משתנים מקריים המקבלים ערכים בקבוצה נתונה S, כך שלכל t, X_{t+1} תלוי רק בצעד שלפניו X_t. כלומר

$$\mathrm{P}(X_{t+1} = a_{t+1} \mid X_1 = a_1, \ldots, X_t = a_t) = \mathrm{P}(X_{t+1} = a_{t+1} \mid X_t = a_t).$$

אנחנו נניח שהשרשרת הומוגנית (כלומר ההסתברות לעבור ממצב i למצב j לא תלויה ב-t) ועל קבוצת מצבים סופית S מגודל n. לכן היא נתונה על ידי איזושהי מטריצה $P \in \mathbb{R}^{n\times n}$ של הסתברויות.

משפט 2. כל שרשרת מרקוב אי-פריקה (אפשר להגיע מכל מצב ב-S לכל מצב עם איזושהי כמות צעדים) וא-מחזורית (לכל i, $\gcd\left\{k \mid P_{ii}^{(k)} > 0\right\} = 1$) מתכנסת להתפלגות סטציונרית (יחידה) π.

שאלה 3. כמה מהר השרשרת מתכנסת להתפלגות הזו?

הגדרה 4. יהיו μ, π שתי התפלגויות על קבוצה S. נגדיר

$$\|\mu - \pi\|_{\mathrm{TV}} = \frac{1}{2}\sum_{i\in S} |\mu(i) - \pi(i)| = \max_{A\subseteq S} |\mu(A) - \pi(A)| .$$

הגדרה 5. נניח שנתונה שרשרת מרקוב $\{X_t\}$ על קבוצה S. לכל $\varepsilon > 0$, נגדיר

$$\tau_{\mathrm{mix}}(\varepsilon) = \inf\left\{t \mid \|\mathrm{P}(X_t \in \cdot) - \pi\|_{\mathrm{TV}} < \varepsilon\right\}.$$

לרוב בוחרים $\varepsilon = \frac{1}{4}$ ומסמנים $\tau_{\mathrm{mix}} = \tau_{\mathrm{mix}}(\frac{1}{4})$.

כשנרצה לדבר על קצב ההתכנסות של שרשראות מרקוב, נרצה לקחת סדרה שלהן ולהשאיף $|S| \to \infty$. נגדיר את ההגדרות הבאות:

הגדרה 6. לסדרת שרשראות מרקוב $\{X_t^n\}$, נאמר שסדרת זמנים $t = t_n$ היא:

1. **חסם עליון**, אם $\limsup_n \|\mathrm{P}(X_t^n \in \cdot) - \pi_n\|_{\mathrm{TV}} \leq \frac{1}{4}$;
2. **חסם תחתון**, אם $\limsup_n \|\mathrm{P}(X_t^n \in \cdot) - \pi_n\|_{\mathrm{TV}} \geq \frac{1}{4}$;
3. **סף**, אם לכל $\varepsilon > 0$ מתקיים ש-$(1-\varepsilon)t_n$ חסם תחתון ו-$(1+\varepsilon)t_n$ חסם עליון.

הילוכים מקריים על חבורות

אם S היא חבורה, אומרים ששרשרת מרקוב $\{X_t\}$ היא **הילוך מקרי על** S אם יש מידת הסתברות μ על S שעבורה $p_{ij} = \mu(i^{-1}j)$. במקרה הזה, אם חושבים על השרשרת בזמן הפוך $\left\{\hat{X}_t\right\}$, היא מתכנסת בדיוק באותו הקצב של $\{X_t\}$.

ערבוב top-to-random

בערבוב הזה בכל צעד לוקחים את הקלף העליון, ומכניסים אותו למיקום אקראי בחבילה. פורמלית אפשר לתאר זאת כהילוך מקרי על S_n ביחס למידה

$$\mu((k\, k-1\, \cdots\, 1)) = \frac{1}{n}.$$

נטען את הטענה הבאה:

טענה 7. סדרת הזמנים $t_n = n\log n$ היא סף של ערבוב ה-top-to-random.

ערבוב random transpositions

בערבוב הזה בכל צעד בוחרים שני קלפים באקראי i ו-j (מאפשרים $i = j$), ומחליפים בין הקלפים במיקומים ה-i וה-j. זהו הילוך מקרי על S_n ביחס למידה

$$\mu(\mathrm{id}) = \frac{1}{n}, \qquad \mu((i\, j)) = \frac{1}{n^2/2}.$$

נטען:

טענה 8. סדרת הזמנים $t_n = \frac{1}{2}n\log n$ היא סף של ערבוב ה-random transpositions.

ערבוב riffle shuffle

הערבוב הבא שנתעסק איתו דומה יותר לערבוב מציאותי. מחלקים את החפיסה לשני חלקים עם בערך $\frac{n}{2}$ קלפים, ואז משלבים את שני החצאים זה בזה. באופן פורמלי יותר: מגרילים $\mathrm{Bin}(n, \frac{1}{2})$ שאומר לנו כמה קלפים לוקחים מהחפיסה לחצי השמאלי; אחר כך, בהינתן שני חצאים, בוחרים בכל צעד את אחד הקלפים העליונים שלהם להיות הקלף הבא. אם בערימה השמאלית נותרו L קלפים ובימנית R, בוחרים את הקלף הראשון מהשמאלית בסיכוי $\frac{L}{L+R}$.

טענה 9. $(1+o(1))\log n \le \tau_{\mathrm{mix}} \le 2\log_2 n + 1$.

הערה 10. התשובה האמיתית היא $\frac{3}{2}\log_2 n$, הוכחה על ידי Bayer ו-Diaconis.

הערה 11. אבל – האם שבעה ערבובים זה מספיק? בשנת 2004 הציג Peter Doyle משחק שנקרא New Age Solitaire, שבו הסיכוי לזכות בחפיסה מעורבבת באמת הוא $\frac{1}{2}$, אבל הסיכוי לזכות בחפיסה שעורבבה 7 פעמים ב-riffle shuffle הוא בערך 0.8.