פתרונות לתרגיל הראשון

לכו בעקבות החץ

אפשר לחסוך את החישובים אם מגדירים "התפלגות משותפת" מתאימה ומשתמשים בשאלה "התפלגויות מותנות" שעברתם עליה מתוך חוברת התרגילים הפתורים (השאלה בחוברת מנוסחת עבור מ"מ שמקבלים ערכים מתוך $\{1,\ldots,n\}$, אז אצלנו אפשר לקבוע $n=|V|$ ולזהות את V עם קבוצת המספרים הנ"ל). נגדיר את ההתפלגות μ מעל $V \times V$ באופן הבא:

- ראשית נבחר קשת $uv \in E$, באופן יוניפורמי מכל הקשתות. בסימון כאן u הוא צומת היציאה ו־v הוא צומת הכניסה.
- בהסתברות $\frac{1}{2}$, נבחר את $(v,v) \in V \times V$.
- בהסתברות $\frac{1}{2}$, נבחר את $(v,u) \in V \times V$.

כל שנותר לעשות הוא לשים לב לנקודות הבאות.

- ההתפלגות של הקורדינטה הראשונה היא בדיוק לפי τ. עבור צומת $w \in V$ כל שהוא, הקורדינטה הראשונה תהיה שווה לצומת w אם ורק אם נבחרה קשת שנכנסת ל־w (כאשר בחרנו קשת באופן יוניפורמי מכל הקשתות המכוונות האפשריות), ז"א בהסתברות $d_i(w)/|E|$.
- ההתפלגות של הקורדינטה השניה היא בדיוק לפי ν. עבור צומת $w \in V$, בהסתברות $\frac{1}{2}$ בחרנו את צומת היציאה של uv, ואז ההסתברות עבור $v=w$ היא $d_i(w)/|E|$, ובהסתברות $\frac{1}{2}$ בחרנו את צומת הכניסה של uv, ואז ההסתברות עבור $u=w$ היא $d_o(w)/|E|$ (כאשר $d_o(w)$ מסמן את דרגת היציאה של w). סה"כ ההסתברות שלנו היא $\frac{1}{2}(d_i(w)+d_o(w))/|E| = d(w)/2|E|$.
- ההסתברות ששתי הקורדינטות זהות זו לזו היא $\frac{1}{2}$, ההסתברות לבחירת $(v,v) \in V \times V$ (אם מתירים לולאות בגרף אז ההסתברות לשוויון היא עדיין לפחות $\frac{1}{2}$). זה משלים את הדרוש להפעלת תוצאת השאלה "התפלגויות מותנות" מהחוברת להשלמת חסם המרחק.

התבלבלות

עבור היותר פורמליסטים מבינינו, ההתפלגות $\hat{\mu} : S \times S \to [0,1]$ המוגדרת בשאלה נתונה ע"י הפונקציה $\hat{\mu}(i,j) = \frac{1}{2}(\mu_1(i)\mu_1(j) + \mu_2(i)\mu_2(j))$, והמ"מ מוגדרים ע"י הפונקציות $X(i,j)=i$ ו־$Y(i,j)=j$.

עתה נניח שקיים i שעבורו $\mu_1(i) \neq \mu_2(i)$, ונראה שהמ"מ אינם ב"ת זה בזה ע"י כך שנראה שמתקיים $\Pr[X=i \wedge Y=i] \neq \Pr[X=i]\cdot\Pr[Y=i]$. מחישוב ישיר $\Pr[X=i \wedge Y=i] = \frac{1}{2}((\mu_1(i))^2 + (\mu_2(i))^2)$ וכן $\Pr[X=i]\cdot\Pr[Y=i] = (\frac{1}{2}(\mu_1(i)+\mu_2(i)))^2$. להשלמת ההוכחה מחשבים את ההפרש:

$$\Pr[X=i \wedge Y=i] - \Pr[X=i]\cdot\Pr[Y=i] = \frac{(\mu_1(i))^2}{4} + \frac{(\mu_2(i))^2}{4} - \frac{\mu_1(i)\mu_2(i)}{2} = \frac{(\mu_1(i)-\mu_2(i))^2}{4} \neq 0$$

לתפוס את הרכבת

נתחיל מתובנה מאוד שימושית לפתרון שיוצג כאן: עבור כל $i,k \in \mathbb{N}$ ו־$a_0,\ldots,a_k$, ההסתברות המותנה $\Pr[X_{i+1}=a_1,\ldots,X_{i+k}=a_k | X_i=a_0]$ זהה להסתברות $\Pr[X_1=a_1-a_0,\ldots,X_k=a_k-a_0]$

(הסתברות לא מותנה). אם קיים l שעבורו $|a_l - a_{l-1}| \neq 1$ אז ההסתברות היא אפס, ואחרת היא שווה ל־$(\frac{1}{3})^r(\frac{2}{3})^{k-r}$ כאשר r הוא גודל הקבוצה $\{1 \leq l \leq k : a_l = a_{l-1} + 1\}$.

עתה נסמן את המאורעות הבאים: A הוא המאורע שקיים i שעבורו $X_i = 1$ (שאת ההסתברות שלו אנחנו מעונינים לחשב), B הוא המאורע שקיים j שעבורו $X_j = 2$, ולכל i נגדיר את המאורע B_i שקיים $j > i$ שעבורו $X_j = 2$ (בפרט המאורעות B_0 ו־B_1 זהים ל־B עצמו). כמו כן נסמן ב־T את המשתנה המקרי שערכו הוא ה־i הקטן ביותר שקיבלנו עבורו $X_i = 1$. חשוב לציין ש־T יכול לקבל גם את הערך "∞", וזאת בדיוק כאשר המאורע A לא מתקיים.

הדבר הראשון לשים לב הוא שמתקיים $\Pr[A|X_1 = -1] = \Pr[B]$. זאת מכיוון שלפי התובנה למעלה, ההתפלגות של $X_2, X_3, \ldots$ בהינתן המאורע $X_1 = -1$ זהה להתפלגות הלא־מותנה של סדרת המ"מ $X_1 - 1, X_2 - 1, \ldots$. אם רוצים להיות יותר פורמלים, התובנה למעלה נכתבה עבור סדרות סופיות, ולכן הקביעה היא שההסתברות שלפחות אחד מ־$X_2, \ldots, X_k$ שווה ל־1 בהינתן המאורע $X_1 = -1$ זהה להסתברות הלא־מותנה שלפחות אחד מ־$X_1, \ldots, X_{k-1}$ שווה ל־2, ומכאן אפשר לקחת את הגבול של סדרת ההסתברויות.

מניתוח ההסתברויות לערכי X_1 נקבל $\Pr[A] = \frac{1}{3}\Pr[A|X_1 = 1] + \frac{2}{3}\Pr[A|X_1 = -1] = \frac{1}{3} + \frac{2}{3}\Pr[B]$. עתה ננסה למצוא בטוי אלטרנטיבי עבור $\Pr[B]$.

לשם כך נפצל את המאורע B לאיחוד זר של מאורעות, לפי החיתוכים עם המאורעות על הערכים האפשריים של T. ההסתברות היא חיבורית מעל מאורעות זרים, ולכן נקבל $\Pr[B] = \sum_{t \in \mathbb{N}} \Pr[B \wedge (T = t)]$ (נשים לב שמתקיים $\Pr[B \wedge (T = \infty)] = 0$ ולכן הוא לא נכנס לסכום). עבור מחובר בודד, הדבר הבא הוא לשים לב שהמאורע $B \wedge (T = t)$ זהה למאורע $B_t \wedge (T = t)$, בגלל שעבור $i \leq T$ לא יכול להתקיים $X_i = 2$ (בשביל "להגיע לערך 2" צריך קודם "לעבור דרך 1", ולפי ההגדרה של T זה לא יכול לקרות לפניו). על כן מתקיים $\Pr[B \wedge (T = t)] = \Pr[B_t \wedge (T = t)] = \Pr[B_t|T = t]\Pr[T = t]$ לכל $t \in \mathbb{N}$.

עתה נשתמש שוב בתובנה השימושית מתחילת הפתרון. הפעם נקבל ממנה שההסתברות המותנה $\Pr[B_t|T = t]$ זהה להסתברות הלא־מותנה $\Pr[A]$. על כן:

$$\Pr[B] = \sum_{t \in \mathbb{N}} \Pr[B_t|T = t]\Pr[T = t] = \Pr[A] \sum_{t \in \mathbb{N}} \Pr[T = t] = (\Pr[A])^2$$

לסיום, נציב את זה במשוואה הקודמת שקיבלנו, ונקבל $\Pr[A] = \frac{1}{3} + \frac{2}{3}(\Pr[A])^2$. פתרונות המשוואה הריבועית הזו הם $\frac{1}{2}$ ו־1. היה נתון לכם שמתקיים $\Pr[A] < 1$, מה שמשאיר לנו את האפשרות $\Pr[A] = \frac{1}{2}$.

פתרונות לתרגיל השני

לתפוס את המרובה

נחשוב על סדרה לא סופית של מ"מ $X_1, X_2, \ldots$ שכולם ב"ת ומתפלגים לפי μ. לשם הנוחות, לכל $i \in \mathbb{N}$ נגדיר את הקבוצה (המוגרלת מקרית) A_i להיות קבוצת כל הערכים המופיעים ב־$\{X_1, \ldots, X_i\}$, ובאופן תואם לכך נגדיר גם את $A_0 = \emptyset$. בנוסף לכך לכל $j \in \mathbb{N}$ נגדיר את K_j להיות המ"מ שמקבל את ה־i המינימלי שעבורו $|A_i| = j$, ובהתאם נגדיר גם $K_0 = 0$. יכול להיות מצב שיתקיים $K_j = \infty$: זה תמיד קורה אם ל־μ יש תומך סופי שגודלו קטן מ־j, ולא קשה לראות שזה קורה בהסתברות 0 אחרת. עתה נגדיר את המשתנים שננתח, $Z_1, \ldots, Z_n$.

לכל $1 \leq i \leq n$, Z_i יוגדר להיות שווה ל־0 אם קיים $j \leq K_{i-1}$ שעבורו $\mu(A_j) \geq 1 - \epsilon$ (הסיבה היחידה לא להתייחס ישר ל־$A_{K_{i-1}}$ היא מקרה הקצה שמתקיים $K_{i-1} = \infty$), ואחרת הוא יוגדר להיות שווה ל־$K_i - K_{i-1}$, ז"א "מספר הנסיונות להגדיל את $|A_{K_{i-1}}|$ מ־$i - 1$ ל־i עד שמצליחים". עתה נשים לב שהסכום $Z = \sum_{i=1}^{n} Z_i$ הוא המ"מ שערכו שווה ל־j הקטן ביותר שעבורו A_j מקיים את תנאי השאלה "$\mu(A_j) \geq 1 - \epsilon$ או $|A_j| \geq n$". הסיבה לכך היא שהסכום הנ"ל שווה ל־K_n אם לא היה $j < K_n$ שעבורו $\mu(A_j) \geq 1 - \epsilon$, ואחרת הוא שווה ל־$K_i$ שעבורו $j = K_i$ הוא האינדכס שבו האירוע "$\mu(A_j) \geq 1 - \epsilon$" קרה לראשונה. לכן, אם יש לנו חסם $\mathrm{E}[Z] \leq \alpha$, אז אפשר לפי אי־שוויון מרקוב לבחור $k = 2\alpha$ על מנת להבטיח הסתברות של לפחות $\frac{1}{2}$ ש־A_k יקיים את תנאי השאלה.

נוכיח שלכל i מתקיים $\mathrm{E}[Z_i] \leq 1/\epsilon$: נראה שלכל קבוצה סופית $B \subset \mathbb{N}$, מתקיים $\mathrm{E}[Z_i | A_{K_{i-1}} = B] \leq 1/\epsilon$ (אלא אם כן מתקיים $\Pr[A_{K_{i-1}} = B] = 0$). בפרט הדבר יוכיח שכל ה־Z_i יהיו סופיים בהסתברות 1 (שימו לב שהגדרת Z_i לא מבטיחה את זה אוטומטית). אם $\mu(B) \geq 1 - \epsilon$, אז מההגדרה מתקיים $\mathrm{E}[Z_i | A_{K_{i-1}} = B] = 0$. לעומת זאת, אם $\mu(B) = 1 - \eta < 1 - \epsilon$, אז ההתפלגות של Z_i זהה למספר ההטלות של מטבע לא מאוזן עם הסתברות η לתוצאה "1", עד שתוצאה זו מתקבלת (ההתפלגות הגיאומטרית עם פרמטר η), ובפרט מתקיים $\mathrm{E}[Z_i | A_{K_{i-1}} = B] = 1/\eta \leq 1/\epsilon$. מכיוון שהחסם נכון לכל התניה אפשרית על $A_{K_{i-1}}$, מתקיים גם ללא התניה $\mathrm{E}[Z_i] \leq 1/\epsilon$. על כן מלינאריות התוחלת מתקיים $\mathrm{E}[Z] \leq n/\epsilon$, ולכן בחירה של $k = 2n/\epsilon$ תקיים את הנדרש עבור פתרון השאלה.

הערה קטנה: הניתוח של $\mathrm{E}[Z]$ היה צריך להזכיר לכם את השאלה "הגעה מהוססת" שלמדתם בתרגול.

הערה על האפשרות לפתרון של 6 נקודות: כאשר ההתפלגות μ מוגדרת מעל $\{1, \ldots, n\}$ בלבד, ניתן להגדיר את המשפחה $\mathcal{B} = \{B \subseteq \{1, \ldots, n\} : \mu(B) > \epsilon\}$. עתה אפשר לפתור את השאלה למשל עבור $k = 2n/\epsilon$: לכל $B \in \mathcal{B}$ ההסתברות שקבוצת ערכי $X_1, \ldots, X_k$ לא תכיל איבר של B תהיה חסומה ע"י $(1-\epsilon)^k < e^{-\epsilon k} < 2^{-1-n}$, וכל שנותר הוא להשתמש בחסימת איחוד מאורעות עבור כל האיברים של $\mathcal{B}$ (שמספרם חסום ע"י 2^n).

רוקדים על שתי החתונות

נשתמש בשיטה של הגרלה עם תיקונים (בסגנון דומה לזה של השאלה "קבוצות ב"ת בהיפרגרפים" שהועברה בתרגול). נגדיר תהליך מקרי בן מספר שלבים לבנית הקבוצה A שתמיד תקיים את הנדרש מבחינת השכנים של כל צומת ב־$V \setminus A$, ונחסום את התוחלת של הגודל $|A|$.

- בשלב הראשון, מגדירים קבוצה A_0 ע"י כך שבאופן ב"ת, לכל צומת $v \in V$, נבחר אותו להיות ב־A_0 בהסתברות $2\log(d)/d$, באופן ב"ת לכל הצמתים. נשים לב שמתקיים $\mathrm{E}[|A_0|] = 2n\log(d)/d$.
- בשלב השני, נסמן ב־B_0 את קבוצת הצמתים ב־$V \setminus A_0$ שכל שכניהם ב־$V \setminus A_0$. מלינאריות התוחלת מתקיים $\mathrm{E}[|B_0|] \leq n \cdot (1 - 2\log(d)/d)^{d+1} < ne^{-2\log(d)} < n\log(d)/d$ עבור d גדול דיו (הסבר קצר – לכל צומת $v \in V$ יש לפחות d שכנים, ולכל אחד מהם כולל v עצמו יש הסתברות ב"ת של $1 - 2\log(d)/d$ להשאר ב־$V \setminus A_0$). נגדיר $A_1 = A_0 \cup B_0$. בפרט אין אף צומת ב־$V \setminus A_1$ שכל שכניהם ב־$V \setminus A_1$.

- בשלב השלישי, נסמן ב־B_1 את קבוצת הצמתים ב־$V \setminus A_1$ שכל שכניהם ב־A_1. על מנת לחסום את התוחלת של B_1, נחשוב על צומת $v \in V$ מדרגה $d' \geq d$. לכל צומת w שהוא שכן של v, יש לו הסתברות $2\log(d)/d$ להיות ב־A_0, והסתברות חסומה ע"י $\log(d)/d$ להיות ב־B_0 (ראו את החישוב של הסעיף הקודם). על כן תוחלת מספר השכנים של v שנמצאים ב־A_1 חסומה ע"י $3d' \log(d)/d$, ולכן לפי אי־שוויון מרקוב הסיכוי שכל d' השכנים יהיו ב־A_1 (הדבר שיכניס את v ל־B_1) חסום ע"י $3\log(d)/d$. על כן מתקיים $\mathrm{E}[|B_1|] \leq 3n\log(d)/d$. לבסוף נגדיר $A = A_1 \cup B_1$. נשים לב שהפעולה הזו לא השפיעה כלל על מיקום השכנים של הצמתים שנשארו ב־$V \setminus A$ (יחסית למצב שהיה עם A_1), ולכן עדיין לא יהיו צמתים ב־$V \setminus A$ שכל שכניהם ב־$V \setminus A$, ז"א ש־A מקיימת את הנדרש בשאלה.

נפעיל שוב את לינאריות התוחלת, ונקבל $\mathrm{E}[|A|] = \mathrm{E}[|A_0|] + \mathrm{E}[|B_0|] + \mathrm{E}[|B_1|] \leq 6n\log(d)/d$, ולכן יש בחירה ספציפית של A עם חסם הגודל הזה שמקיימת את הנדרש.

פתרונות לתרגיל השלישי

שוויון הנפש

נבדוק תהליך שבו מגרילים קבוצה B באופן יוניפורמי מבין 2^n תתי־הקבוצה האפשריים, וננתח את הערכים האפשריים עבור המ"מ $X_1,\ldots,X_k$, כאשר מגדירים $X_i = |A_i \cap B|$. בסופו של דבר הרעיון יהיה להשתמש בטיעון "שובך יונים" עבור סדרות הערכים האפשריות של המ"מ הנ"ל, אבל אי אפשר לעשות את זה עבור X_i שבאופן תאורטי יכולים לקבל כל ערך בין 0 ל־n.

ננתח עבור $1 \leq i \leq k$ קבוע את התפלגות X_i: תוחלת המשתנה היא $|A_i|/2$. מחסימת סטיות גדולות, מתקיים $\Pr[X_i > |A_i|/2 + \sqrt{|A_i|\log(n)}] < e^{-2|A_i|\log(n)/|A_i|} < \frac{1}{4k}$ וכן $\Pr[X_i < |A_i|/2 - \sqrt{|A_i|\log(n)}] < \frac{1}{4k}$ (עבור ההשוואה ל־$\frac{1}{4k}$ אנחנו משתמשים בכך שבפרט מתקיים $k < 2n$, ומניחים ש־N_α נקבע כך שיתקיים $e^{-2\log(n)} < n^{-2} < \frac{1}{8n}$). על כן, בהסתברות לפחות $\frac{1}{2}$, אף אחד מהמאורעות המתוארים למעלה לא מתקיים לאף $1 \leq i \leq k$. זה אומר שלפחות $2^n/2$ מהבחירות האפשריות עבור B הן כאלה שעבורן לכל X_i יש לכל היותר $1 + 2\sqrt{|A_i|\log(n)} \leq 4\sqrt{n\log(n)}$ ערכים (שוב בהנחה ש־n הוא גדול דיו).

עכשיו נשתמש בעקרון שובך היונים עבור $2^n/2$ הבחירות הנ"ל. מספר סדרות הערכים האפשריות עבור $X_1,\ldots,X_k$ חסום כאן ע"י $(4\sqrt{n\log(n)})^k = 2^{k(\log(n)/2+\log\log(n)/2+2)} \leq 2^{(\alpha/2+o(1))n}$. מכיוון שנתון $\alpha < 2$, עבור n גדול דיו (כתוצאה מבחירה מתאימה של N_α) הביטוי הנ"ל יהיה קטן מ־$2^n/2$, ולכן יהיו לפחות שתי בחירות של B עם אותה סדרת ערכים עבור $|A_1 \cap B|,\ldots,|A_k \cap B|$, כנדרש.

צריכים לחסוך

נראה איך מתרגמים אלגוריתם הסתברותי שמבצע q קריאות (אבל לא מוגבל בכמות האקראיות שלו) לכזה שמבצע $3q$ קריאות ומקיים את חסם האקראיות הדרוש. ראשית דבר נתרגם אותו לאלגוריתם שמבצע $3q$ קריאות, עדיין לא מציית לחסם אקראיות כל שהוא, אבל נותן את התשובה הנכונה בהסתברות לפחות $\frac{20}{27}$ (במקום $\frac{2}{3}$). המעבר פשוט: מריצים את האלגוריתם המקורי שלוש פעמים, כל פעם באופן ב"ת בקודמות (מבצעים את כל ההגרלות באופן ב"ת בהרצות קודמות), ומחזירים את התשובה שלפחות שתיים מההרצות הנ"ל החזירו. אם האלגוריתם המקורי טעה בהסתברות $\alpha \leq \frac{1}{3}$, אז האלגוריתם החדש יטעה בהסתברות $\alpha^3 + 3(1-\alpha)\alpha^2 \leq \frac{7}{27}$.

עתה נראה תרגום של האלגוריתם החדש לאלגוריתם עם אותו מספר שאילתות, שמשתמש ב־$\log(n) + O(1)$ הטלות מטבע בלבד, ונותן את התשובה הנכונה בהסתברות לפחות $\frac{2}{3}$. לשם כך נסתכל על האלגוריתם ההסתברותי שאנחנו הולכים לתרגם באופן הבא: נניח שכל ההגרלות שהאלגוריתם יכול לעשות במהלך ההרצה נעשות מראש, ובמהלך הריצה של האלגוריתם הוא רק "קורא" אותן יחד עם הקריאות שלו מהקלט. במילים אחרות, מסתכלים על האלגוריתם ההסתברותי כמרחב הסתברות מעל אלגוריתמים דטרמיניסטים (שימו לב אבל שהאלגוריתמים הדטרמיניסטים הנ"ל אינם בד"כ בעלי "תיאור קצר", במקרה הכללי אלגוריתם כזה מתואר כעץ החלטות בגובה חסום ע"י $3q$).

נחשוב עתה על קלט ספציפי $x \in \{0,1\}^n$, וננתח מה קורה כאשר מבצעים l הרצות בלתי־תלויות של האלגוריתם ההסתברותי עליו. באופן מדויק, נרצה לדעת את ההסתברות שפחות מ־$\frac{2}{3}l$ מההרצות הנ"ל יתנו תשובה נכונה. לפי חסימת סטיות גדולות (החסם השני מההרצאה), ההסתברות שזה יקרה חסומה ע"י $e^{-2(2/3-20/27)^2 l} < 2^{-l/100}$. על כן, לפי איחוד מאורעות, עבור $l \geq 100n$ יש סיכוי חיובי ש־l ההרצות הנ"ל יקיימו בו זמנית את התנאי שלכל $x \in \{0,1\}^n$ לפחות $\frac{2}{3}l$ מהן נותנות תשובה נכונה.

שימו לב שלכל $1 \leq i \leq l$, מה שאנחנו מכנים "ההרצה ה־i" משתמשת בסידרה אחת של "הגרלות הנעשות מראש" (ראו את הפסקה הלפני־קודמת), ולא מבצעת שום הגרלות נוספות תוך כדי קריאת $x \in \{0,1\}^n$ עצמו – בהינתן ההגרלות הנ"ל זהו אלגוריתם דטרמיניסטי. ההגרלות כן נעשות מחדש (באופן ב"ת) למשל עבור "ההרצה ה־j" כאשר $j \neq i$.

נבחר את ה־k המינימלי שעבורו $2^k \geq 100n$ ונבחר $l = 2^k$. שימו לב שבפרט מתקיים $k = \log(n) + O(1)$.
עתה נרשום סדרה ספציפית של l אלגוריתמים דטרמיניסטים שמקיימת שלכל $x \in \{0,1\}^n$, לפחות $\frac{2}{3}l$ מתוכם נותנים תשובה נכונה. סדרה כזו קיימת לפי הפסקה הקודמת: כזכור כל הרצה של אלגוריתם הסתברותי ניתנת לתיאור כבחירה הסתברותית של אלגוריתם דטרמיניסטי שאותו מריצים, ומכיוון שיש חסם הסתברות חיובי על כך שלכל $x \in \{0,1\}^n$ לפחות $\frac{2}{3}l$ מההרצות מהן נותנות תשובה נכונה, יש בחירה ספציפית של l הרצות שמקיימות את המבוקש.

לאחר שרשמנו את הרשימה הארוכה הנ"ל, אפשר לנסח את האלגוריתם הסופי שלנו: האלגוריתם יגריל יוניפורמית ערך i מתוך $\{1,\ldots,l\}$, שאפשר ע"י ייצוג בינארי לחשוב עליו כעל התוצאה של הגרלה יוניפורמית מתוך $\{0,1\}^k$, ואז יבצע מהרשימה את ההרצה ה־i.

פתרונות לתרגיל הרביעי

אי אפשר להתחמק

על מנת "לחסוך מקום", נסמן $U_i = U\cap\{1,\ldots,i\}$ עבור $0\le i\le n$, ובפרט $U_0=\emptyset$ ו־$U_n=U$. נגדיר עכשיו את המ"מ $X_0,\ldots,X_n$ מעל מרחב ההסתברות של בחירת הקבוצה U (אלו פונקציות ממשיות $X_i:\mathcal{P}(\{1,\ldots,n\})\to\mathbb{R}$), לפי $X_i(U)=|A\cap U_i|-p|A\cap\{1,\ldots,i\}|$. ברור שמתקיים $X_0=0$ בהסתברות 1, וכן מתקיים $X_n=|A\cap U|-p|A|$. עתה נוכיח שהסדרה הנ"ל היא מרטינגל, ושמתקיים $|X_i(U)-X_{i-1}(U)|\le 1$ לכל U ו־$1\le i\le n$.

נשים לב ש־$X_0(U),\ldots,X_{i-1}(U)$ תלויים אך ורק ב־U_{i-1}, ישירות מההגדרות של המ"מ יחד עם זה ש־$A\cap\{1,\ldots,i-1\}$ תלוי אך ורק ב־$U\cap\{1,\ldots,i-2\}$ (שימו לב להגדרות של $\mathcal{A}_1,\ldots,\mathcal{A}_{i-1}$ בהבהרות לתרגיל). נראה שלכל ערך אפשרי V של U_{i-1} מתקיים $\mathrm{E}[X_i(U)|U_{i-1}=V]=X_{i-1}(V)$, כאשר $X_{i-1}(V)$ מסמן את הערך המשותף של $X_{i-1}(U)$ לכל ה־U המקיימים $U_{i-1}=V$. זה מוכיח שהמדובר במרטינגל, כי התוחלת המותנה על סדרת הערכים של $X_0,\ldots,X_{i-1}$ תהיה ממוצע משוקלל לפי כל ה־V שנותנים את הסדרה הזו. את השלמת ההוכחה מבצעים לפי פיצול למקרים: אם $\mathcal{A}_i(V)=0$, אז $X_i=X_{i-1}$ (כי A לא יכיל את i). אם $\mathcal{A}_i(V)=1$, אז בהסתברות p יתקיים $X_i=X_{i-1}+1-p$ (אם i נבחר להיות ב־U), ובהסתברות $1-p$ יתקיים $X_i=X_{i-1}-p$ (אם i לא נבחר להיות ב־U). שני אלו שוב נותנים תוחלת מותנה ששווה ל־X_{i-1}.

לחסימת $|X_{i-1}(U)-X_i(U)|$, פשוט מסתכלים על כל האפשרויות שנזכרו כבר בפסקה הקודמת: בפרט תמיד מתקיים $X_i(U)\in\{X_{i-1}(U),X_{i-1}(U)+1-p,X_{i-1}(U)-p\}$, ולכן הם מקיימים את החסם הנדרש.

אחרי שהוכחנו את שתי התכונות של $X_0,\ldots,X_n$, כל שנותר הוא להפעיל את משפט אזומה לקבלת $\Pr[|A\cap U|-p|A|>a]=\Pr[X_n>a]<e^{-a^2/2n}$.

לסחוט את הלימון

ראשית נתייחס להוכחה של משפט הריכוז על משתנים מוטים מהחוברת. אנחנו נוכיח שלכל $\lambda>0$ מתקיים $\mathrm{E}[e^{\lambda X_m}]<e^{\lambda^2 m/8}$. אחרי שמוכיחים זאת מציבים $\lambda=\frac{4a}{m}$ ומשתמשים באי־שוויון מרקוב, כפי שנעשה שם.

הוכחת הטענה הזו משתמשת בשיטה דומה לחסם על תוחלת $e^{\alpha X_m}$ בהוכחה של משפט אזומה. הטענה הראשונה שנוכיח היא שלכל מ"מ Z שמקיים $\mathrm{E}[Z]=0$ ו־$\Pr[-p\le Z\le 1-p]=1$ עבור $p\in[0,1]$ מתאים, מתקיים $\mathrm{E}[e^{\lambda Z}]<e^{\lambda^2/8}$.

לשם כך נשתמש בקמירות של הפונקציה $f(z)=e^{\lambda z}$ על התחום $[-p,1-p]$. כזכור מתקיים בתחום הזה $f(z)\le g(z)$, כאשר $g(z)$ היא הפונקציה הלינארית העוברת דרך $(-p,f(-p))$ ו־$(1-p,f(1-p))$. אם נרשום את הפונקציות במפורש, נקבל שעבור כל $-p\le z\le 1-p$ מתקיים

$$e^{\lambda z}\le\left(pe^{\lambda(1-p)}+(1-p)e^{-\lambda p}\right)+\left(e^{\lambda(1-p)}-e^{-\lambda p}\right)z$$

ע"י שימוש בלינאריות התוחלת, נובע מזה $\mathrm{E}[f(Z)]\le\mathrm{E}[g(Z)]=pe^{\lambda(1-p)}+(1-p)e^{-\lambda p}\le e^{\lambda^2/8}$, כאשר אי־השוויון מימין מופיע (אומנם בלי הוכחה) בהוכחת משפט הריכוז המקורית. עתה אנחנו יכולים להוכיח באינדוקציה את הטענה הדרושה על $\mathrm{E}[e^{\lambda X_m}]$: נגדיר $Y_i=X_i-X_{i-1}$ לכל $1\le i\le m$, ונכתוב

$$\mathrm{E}\big[e^{\lambda X_m}\big]=\mathrm{E}\big[e^{\lambda Y_m}e^{\lambda X_{m-1}}\big]\le e^{\lambda^2/8}\mathrm{E}\big[e^{\lambda X_{m-1}}\big]\le\cdots\le e^{\lambda^2 m/8}$$

זוהי בדיוק האינדוקציה שנעשתה בחסימת התוחלת בהוכחה של משפט אזומה, וגם כאן בשביל המעבר משתמשים בכך שתכונות Y_m, ולכן החסם על התוחלת של $e^{\lambda Y_m}$, מתקיימות לכל התניה ספציפית מהצורה $X_0=\alpha_0,\ldots,X_{i-1}=\alpha_{i-1}$.

לא מכילים

נקבע מספר ממשי $0<\alpha<\frac{1}{2}$ כל שהוא. נגריל את הקבוצה B באופן יוניפורמי מכל תתי־הקבוצה של $\{1,\ldots,n\}$ (ז"א שכל $i\in\{1,\ldots,n\}$ יבחר להיות ב־B בהסתברות $\frac{1}{2}$ באופן ב"ת בבחירות האחרות), ונראה, עבור n גדול דיו, ש־B לא תכיל אף אחד מהאיברים של $\mathcal{F}$ וגם תקיים $|B|\geq\alpha n$.

לכל $F\in\mathcal{F}$ נסמן ב־A_F את המאורע ש־B אינה מכילה את F. נשים לב שהמאורע הנ"ל (אם מסתכלים עליו כעל משפחת כל תתי־הקבוצה של $\{1,\ldots,n\}$ שאינם מכילים את F) הוא מונוטוני־לא עולה (כמשפחה של קבוצות). כמו כן מתקיים $\Pr[A_F]=1-2^{-|F|}$ ישירות מההגדרות. על כן ניתן להשתמש במשפט FKG (או לחילופין במשפט קלייטמן), ולקבל $\Pr[\bigwedge_{F\in\mathcal{F}}A_F]\geq\prod_{F\in\mathcal{F}}\Pr[A_F]=\prod_{F\in\mathcal{F}}(1-2^{-|F|})$ (כרגיל השתמשנו באינדוקציה עבור אי השוויון לסדרות של יותר מ־2 מאורעות).

עבור המשך הפיתוח, נשתמש בכך שמתקיים $(1-x)>e^{-2x}$ עבור $0<x<\frac{1}{2}$ (אפשרות הוכחה – הפונקציה e^{-2x} קמורה ממש ולכן חותכת את הישר $y=1-x$ בלא יותר משתי נקודות; אחת הנקודות הנ"ל היא $x=0$, והשניה נמצאת מימין ל־$x=\frac{1}{2}$). על כן מתקיים $\Pr[\bigwedge_{F\in\mathcal{F}}A_F]>e^{-2\sum_{F\in\mathcal{F}}2^{-|F|}}\geq e^{-2n/\log(n)}$.

לפי חסימת סטיות גדולות מתקיים $\Pr[|B|<\alpha n]<e^{-2(1/2-\alpha)^2n}$. מכיוון שעבור α קבוע מתקיים $e^{-2(1/2-\alpha)^2n}=o\big(e^{-2n/\log(n)}\big)$, עבור n גדול דיו הסתברות איחוד המאורעות ה"רעים" $|B|<\alpha n$ ו־$\neg\bigwedge_{F\in\mathcal{F}}A_F$ תהיה קטנה מ־1, ולכן בהסתברות חיובית יתקיים גם $|B|\geq\alpha n$ וגם $\bigwedge_{F\in\mathcal{F}}A_F$.

דומה בדומה

נראה כיוון אחד, $\mathrm{H}[\mu]-\mathrm{H}[\nu]\leq d(\mu,\nu)\log(n-1)+H(d(\mu,\nu))$. החסם מהצד השני סימטרי לחלוטין.

נשתמש בשאלה "התפלגויות מותנות – הכיוון השני" מחוברת התרגילים ונבנה מרחב הסתברות עם שני משתנים מקריים X ו־Y, כאשר X מתפלג כמו μ, Y מתפלג כמו ν, ומתקיים $\Pr[X\neq Y]=d(\mu,\nu)$. עבור אלו מתקיים $\mathrm{H}[\mu]-\mathrm{H}[\nu]=\mathrm{H}[X]-\mathrm{H}[Y]\leq\mathrm{H}[X,Y]-\mathrm{H}[Y]=\mathrm{H}[X|Y]$. כל שנותר הוא להשתמש באי־שוויון פאנו (בגרסה ה"מקוצרת"):

$$\mathrm{H}[X|Y]\leq\Pr[X\neq Y]\log(n-1)+H(\Pr[X\neq Y])=d(\mu,\nu)\log(n-1)+H(d(\mu,\nu))$$