פתרונות לתרגיל הראשון

סדר באי־סדר

נוסחת סטירלינג אומרת שמתקיים $n! = (1+o(1))\sqrt{2\pi n}(n/e)^n$. בפרט, לכל n גדול מספיק, יתקיים $n! > (n/e)^n$. נקבע $C = 10$. עבור קבוצה Q מגודל $k = \lceil C\sqrt{n} \rceil$, הסיכוי ש־$\sigma$ תסדר אותה בסדר עולה הוא $1/k!$ בדיוק (ציינו בשיעור שהתפלגות הסדרים ש־σ מקצה על קבוצה $Q \subset \{1,\ldots,n\}$ היא יוניפורמית, בהוכחה של משפט טוראן). כמו כן, מספר הקבוצות האפשריות מגודל זה הוא $\binom{n}{k} < n^k/k!$. לכן, לפי איחוד מאורעות, ההסתברות לקיום קבוצה Q מגודל k שמסודרת באופן עולה חסומה ע"י $n^k/(k!)^2$. עתה נשים לב שאם n גדול דיו אז מתקיים $k! > (k/e)^k > 3^{10\sqrt{n}} n^{5\sqrt{n}}$. מכאן שההסתברות לקיום קבוצה Q כזו חסומה ע"י $n^{10\sqrt{n}+1}/9^{10\sqrt{n}} n^{10\sqrt{n}} = o(1)$, כנדרש.

התאמה זוגית

חישוב ישיר (עם סכומים) היה עובד כאן. בפתרון זה נחסוך את החישובים ונשתמש בשאלות "התפלגויות מותנות" ו"התפלגויות מותנות – הכיוון השני" בחוברת התרגילים הפתורים.

לפי שאלת הכיוון השני, אנו יכולים למצוא התפלגות ξ מעל $S \times S$ כך שהצמצום שלה לקורדינטה הראשונה מתפלג כמו μ_1, הצמצום שלה לקורדינטה השניה מתפלג כמו μ_2, וההסתברות למאורע שהוגרל (i_1, i_2) שעבורו $i_1 \neq i_2$ היא $d(\mu_1, \mu_2)$. השאלה בחוברת התרגילים אומנם מנוסחת במושגים של בניית שני מ"מ X, Y, אבל זה שקול להגדרה $\xi(i_1, i_2) = \Pr[X = i_1 \wedge Y = i_2]$.

באופן דומה נבנה את η מעל $T \times T$ כך שהצמצום שלה לקורדינטה הראשונה מתפלג כמו ν_1, הצמצום שלה לקורדינטה השניה מתפלג כמו ν_2, וההסתברות למאורע שהוגרל (j_1, j_2) שעבורו $j_1 \neq j_2$ היא $d(\nu_1, \nu_2)$. ננתח עתה את $\xi \times \eta$, ההתפלגות של הגרלת זוג ערכים לפי ξ והגרלה (ב"ת בהגרלה לפי ξ) של זוג ערכים לפי η. זוהי התפלגות מעל $(S \times S) \times (T \times T)$, אבל ע"י סידור מחדש של הקורדינטות אפשר להסתכל עליה כהתפלגות מעל $(S \times T) \times (S \times T)$.

עתה ננתח את ההתפלגות על ה"קורדינטה" הראשונה של זוגות ערכים $(i_1, j_1) \in S \times T$. אנחנו יודעים ש־$i_1$ מתפלג כמו μ_1 (זו הקורדינטה הראשונה של בחירת זוג לפי ξ), ו־j_1 מתפלג כמו ν_1 (זו הקורדינטה הראשונה של בחירת זוג לפי η). מכיוון שהגרלנו באופן ב"ת את הערך לפי ξ ואת הערך לפי η, הזוג כולו מתפלג לפי $\mu_1 \times \nu_1$.

באופן דומה, ה"קורדינטה" השניה היא זוג ערכים $(i_2, j_2) \in S \times T$ שמתפלג לפי $\mu_2 \times \nu_2$. לבסוף, נחסום את ההסתברות שההגרלה של $((i_1, j_1), (i_2, j_2)) \in (S \times T) \times (S \times T)$ תקיים $(i_1, j_1) \neq (i_2, j_2)$. נזכיר שבהגדרה של $\xi \times \eta$ בחרנו את (i_1, i_2) לפי ξ, ולכן $\Pr[i_1 \neq i_2] = d(\mu_1, \mu_2)$, ובחרנו את (j_1, j_2) לפי η, ולכן $\Pr[j_1 \neq j_2] = d(\nu_1, \nu_2)$. כמו כן, קבענו את הבחירה של (i_1, i_2) באופן ב"ת בבחירה של (j_1, j_2), ולכן מתקיים:

$$\begin{aligned}\Pr_{\xi\times\eta}[i_1 \neq i_2 \vee j_1 \neq j_2] &= 1 - \Pr_\xi[i_1 = i_2]\Pr_\eta[j_1 = j_2] = 1 - (1 - d(\mu_1, \mu_2))(1 - d(\nu_1, \nu_2)) \\ &= d(\mu_1, \mu_2) + d(\nu_1, \nu_2) - d(\mu_1, \mu_2) \cdot d(\nu_1, \nu_2)\end{aligned}$$

לפי השאלה "התפלגויות מותנות" (הפעם לא הכיוון השני), הבניה הזו נותנת את החסם עליון הנדרש על $d(\mu_1 \times \nu_1, \mu_2 \times \nu_2)$.

סיכויים לרמזי

נסמן ב־$R(k)$ את מספר הצמתים המינימלי שלגביו משפט רמזי מתקיים (כך שבכל גרף בעל $R(k)$ צמתים יש או קליק בעל k צמתים או קבוצה חסרת קשתות בעלת k צמתים), ונסתכל על התהליך הבא לבחירת $X_1, \ldots, X_k$: ראשית נגריל $R(k)$ מ"מ יוניפורמים מתוך $[0,1]$ וב"ת זה בזה, $Y_1, \ldots, Y_{R(k)}$. אחר זאת נגריל סדרה של אינדכסים $1 \leq i_1 < i_2 < \cdots < i_k \leq R(k)$ באופן יוניפורמי מבין $\binom{R(k)}{k}$ האפשרויות (בעצם מגרילים באופן יוניפורמי ת"ק של $\{1, \ldots, R(k)\}$ בגודל k ואז מסדרים אותה לקבלת האינדכסים). לבסוף לכל $1 \leq j \leq k$ נקבע $X_j = Y_{i_j}$.

ראשית, נשים לב שגם בתהליך הזה, $X_1, \ldots, X_k$ מתפלגים בדיוק כמו k מ"מ ב"ת שמוגרלים יוניפורמית מ־$[0,1]$. הסיבה לכך היא שלכל סדרה ספציפית אפשרית $1 \leq i_1 < i_2 < \cdots < i_k \leq R(k)$, סדרת המ"מ $Y_{i_1}, \ldots, Y_{i_k}$ מתפלגת כמו סדרה של k מ"מ ב"ת יוניפורמים, ולכן הדבר נכון גם אם השתמשנו בהגרלה כל שהיא על מנת לבחור את סדרת האינדכסים (חשוב אבל ששיטת הבחירה אינה תלויה בערכים $Y_1, \ldots, Y_{R(k)}$ עצמם, ז"א שהיינו יכולים להגריל את ה־Y_i לאחר קביעת סדרת האינדכסים).

עתה "נעקם" טיפה את הסימונים ונסמן למשל ב־$f(X_1, \ldots, X_k)$ את הגרף מעל $\{1, \ldots, k\}$ שעבורו i, j היא קשת אם ורק אם $f(X_i, X_j) = 1$. עבור גרף G עם קבוצת הצמתים $\{1, \ldots, R(k)\}$, ניתן לנתח את המאורע "$f(Y_1, \ldots, Y_{R(k)}) = G$" וגם לנתח הסתברויות שמותנות עליו, בגלל הנתון ש־f היא מדידה (המאורע הנ"ל הוא חיתוך של מאורעות מהצורה "$f(X_i, X_j) = 1$" ומשלימים שלהם – אם כי לא הייתם חייבים לציין נקודה זו במפורש בפתרונות שלכם).

לכל גרף G קיימת לפחות קבוצה אחת של צמתים שהיא או קליק או חסרת קשתות, לפי משפט רמזי. ההסתברות שבדיוק האינדכסים של קבוצה זו נבחרו היא $1/\binom{R(k)}{k}$. על כן, בהסתברות לפחות $1/\binom{R(k)}{k}$ יתקיים שהגרף $f(X_1, \ldots, X_k) = f(Y_{i_1}, \ldots, Y_{i_k})$ הוא או חסר קשתות או קליק, לפי נוסחת ההסתברות השלמה (כי החסם מתקיים עבור ההתניה על המאורע "$f(Y_1, \ldots, Y_{R(k)}) = G$" לכל גרף G אפשרי בעל $R(k)$ צמתים). על מנת לסיים את פתרון השאלה נציב $\alpha_k = 1/2\binom{R(k)}{k}$, כי הדבר אומר שלפחות אחת משתי האפשרויות עבור $f(X_1, \ldots, X_k)$ חייבת להתקיים לפחות בהסתברות הזו (אחרת נקבל סתירה לפי אי השוויון על איחוד מאורעות).

פתרונות לתרגיל השני

גלגלים מסתובבים

אנחנו נתייחס לקבוצת הנקודות V כאל קבוצת כל המספרים $\{0,\ldots,kr-1\}$ ב־$\mathbb{Z}_{kr}$, קבוצת השלמים מודולו kr. מעתה, כל החישובים שלנו בפתרון השאלה יהיו מודולו kr. בפרט, אם u,v,w שלוש נקודות ב־$\mathbb{Z}_{kr}$, x,y,z שלוש נקודות אחרות, וכן מתקיים $x-u=y-v=z-w \pmod{kr}$, אז נובע מזה ששני המשולשים המתאימים חופפים (ע"י סיבוב).

עבור $V_1,\ldots,V_r$ נגריל עתה באופן יוניפורמי וב"ת אינדכסים $a_1,\ldots,a_r\in\mathbb{Z}_{kr}$, ולכל i נגדיר את הקבוצה $W_i=a_i+V_i=\{a_i+v : v\in V_i\}$, כאשר החיבור הוא מודולו kr. עבור $w\in\mathbb{Z}_{kr}$ כל שהוא, ההסתברות שמתקיים $w\in\bigcap_{i=1}^r W_i$ היא בדיוק $1/r^r$. זאת מכיוון שההסתברות ל־$w\in W_i$ היא $1/r$, ויש לנו חיתוך של r מאורעות ב"ת כאלה.

נגדיר משתני אינדיקטור X_w עבור המאורעות $w\in\bigcap_{i=1}^r W_i$, ונקבל $|\bigcap_{i=1}^r W_i|=\sum_{w\in\mathbb{Z}_{kr}} X_w$. מלינאריות התוחלת נקבל מכך שמתקיים $\mathrm{E}[|\bigcap_{i=1}^r W_i|]=k/r^{r-1}$. זה אומר שעבור $k>2r^{r-1}$ התוחלת של $|\bigcap_{i=1}^r W_i|$ גדולה מ־2, ולכן קיימת בחירה ספציפית של $i_1,\ldots,i_r$ שעבורה גודל החיתוך הנ"ל גדול מ־2, ז"א שהוא מכיל לפחות שלוש נקודות שונות $\{x,y,z\}$. מתקבל מכך שלכל i הקבוצה V_i מכילה את המשולש $t_i=\{x-a_i,y-a_i,z-a_i\}$, ומשולשים אלו כולם חופפים.

כל אחד את עצמו

נסמן ב־V_i את קבוצת הצמתים בשלב ה־i שיש להם שכן שצבוע באותו צבע כמותם. בפרט, אם הגרף הוא חסר צמתים מבודדים אז $V_0=V$ (כי אז כל הצמתים צבועים באותו צבע), אבל בכל מקרה $|V_0|\le n$. שימו לב ש־$|V_i|$ (לכל $i\ge 0$) הוא משתנה מקרי, כי הוא תלוי בזהות הצמתים החברים ב־V_i, ואלו תלויים בתהליך המקרי של האלגוריתם.

נוכיח עתה באינדוקציה שמתקיים $\mathrm{E}[|V_k|]\le(\frac{2}{3})^k n$. הבסיס הוא $k=0$. עבור המעבר מהשלב ה־$k-1$ לשלב ה־k, נחסום מלמעלה את ההסתברות של צומת להיות עם שכן מאותו צבע. נסמן $r=|V_{k-1}|$ (לפי הנחת האינדוקציה מתקיים $\mathrm{E}[r]\le(\frac{2}{3})^{k-1}n$), ונסמן $V_{k-1}=\{u_1,\ldots,u_r\}$. הסדר שבו אנחנו מסמנים את הצמתים יכול להיות שרירותי, זה לא משנה אפילו אם נשתמש בסדר אחר לכל איטרציה של האלגוריתם.

עבור הניתוח נניח שאנחנו בוחרים את הצבעים החדשים של הצמתים לפי הסדר השרירותי הנ"ל, החל מ־u_1 וכלה ב־u_r. נחסום את תוחלת הגודל של W, קבוצת הצמתים שהפכו להיות עם שכן מאותו צבע בשלב כל שהוא של הצביעה של V_{k-1}, שזה בוודאי חוסם את מספר הצמתים ב־$V_k\subseteq W$ (יכול להיות שחלק מהצמתים של W הצליחו "להינצל" כשצבענו מחדש צמתים יותר מאוחרים, אבל בכל מקרה יש הכלה).

כאשר אנחנו צובעים את u_j מחדש, נתון שיש לו לא יותר מ־d שכנים. מכיוון שהצבע של u_j נבחר באופן יוניפורמי (וב"ת בבחירות קודמות) מקבוצה בת $3d$ צבעים, לפי לינאריות התוחלת, תוחלת מספר השכנים שלו שיהיו כתוצאה מזה מאותו צבע חסומה ע"י $\frac{1}{3}$ (כל שכן יהיה בצבע זהה בהסתברות $\frac{1}{3d}$). צריך אבל לזכור שכאשר יש שכן מאותו צבע אז גם u_j עצמו הוא בעל שכן מצבע זהה, ולכן תוחלת מספר הצמתים הכולל שיהיה להם שכן מאותו צבע בגלל הצביעה של u_j חסומה ע"י $\frac{2}{3}$. שוב מלינאריות התוחלת, מקבלים שתוחלת גודל הקבוצה W חסום כולו ע"י $\frac{2}{3}r$. כאשר אנחנו לא מתנים על $r=|V_{k-1}|$, נקבל חסם עבור התוחלת הלא מותנה $\mathrm{E}[|V_k|]\le\mathrm{E}[|W|]\le\frac{2}{3}\mathrm{E}[|V_{k-1}|]$, ומכאן אנחנו משלימים את צעד האינדוקציה עבור $\mathrm{E}[|V_k|]\le(\frac{2}{3})^k n$.

נסמן ב־T את המ"מ של הזמן שבו כל הצמתים היו צבועים בצבעים שונים. לפי אי־שוויון מרקוב מתקבל $\Pr[T>k]=\Pr[|V_k|\ge 1]\le(\frac{2}{3})^k n$. כמו כן, עבור $k\le 2\log n$, בוודאי מתקיים $\Pr[T>k]\le 1$. מכאן, ע"י שימוש בטריק $\sum_{k=0}^{\infty} k\alpha_k=\sum_{k=0}^{\infty}(\sum_{j=k+1}^{\infty}\alpha_j)$ כאשר אצלנו $\alpha_k=\Pr[T=k]$, אפשר להשיג את החסם $\mathrm{E}[T]=\sum_{k=0}^{\infty}\Pr[T>k]\le 2\log n+\sum_{k\ge 2\log n}(\frac{2}{3})^k n\le 2\log n+\sum_{j=0}^{\infty}(\frac{2}{3})^j=O(\log n)$.

המכנה המשותף

ראשית נמצא $\alpha > \frac{1}{2}$, שעבורו נוכיח שתוחלת אורך תת־המחרוזת הגדולה ביותר היא לפחות αn, עבור n זוגי. תוצאת השאלה נובעת מטענה זו ע"י התהליך הבא: נקבע $\alpha > \alpha' > \frac{1}{2}$ כל שהוא (למשל $\alpha' = \frac{1}{2}\alpha + \frac{1}{4}$), ואז עבור $n > 1/(\alpha - \alpha')$ שהוא אי־זוגי אפשר פשוט להתעלם בניתוח מהאיברים v_{n+1} ו־w_{n+1}, וזה חוסם מלמטה (ע"י $\alpha' n$) את תוחלת האורך של תת־מחרוזת משותפת שהיא לא בהכרח הארוכה ביותר. מכאן שזה חוסם מלמטה גם את התוחלת של תת־המחרוזת הארוכה ביותר, כי עבור מ"מ שתמיד מקיימים $X \geq Y$ יתקיים גם $\mathrm{E}[X] \geq \mathrm{E}[Y]$.

אחרי ההנחה ש־n זוגי, נגביל אפילו עוד יותר את תתי־המחרוזות המשותפות של v ו־w שעבורן נבדוק מי הארוכה ביותר מביניהן. שוב, אם אנחנו חוסמים את תוחלת האורך עבור תת־מחרוזת משותפת שאולי אינה הארוכה ביותר, זה נותן גם חסם תחתון עבור התוחלת של $k(v, w)$.

אנחנו נתמקד בסדרות אינדכסים $1 \leq i_1 < \cdots < i_k \leq n$ ו־$1 \leq j_1 < \cdots < j_k \leq n$ כך ש־$v_{i_l} = w_{j_l}$ לכל $1 \leq l \leq k$ (ז"א שיש לנו תת־מחרוזת משותפת), ובנוסף מקיימות $\lceil i_l/2 \rceil = \lceil j_l/2 \rceil$. המשמעות היא שאנחנו מחלקים את האינדכסים $\{1, \ldots, n\}$ לזוגות $\{(1,2), (3,4), \ldots, (n-1, n)\}$, ודורשים שכל ההתאמות של תת־המחרוזות המשותפת שלנו יהיו דרך אינדכסים שבאים מאותו זוג.

עבור זוג ספציפי $(2r-1, 2r)$ יש לנו את ההסתברויות הבאות:

- בסיכוי $\frac{1}{4}$ יהיו לנו שתי התאמות (ז"א שגם $2r-1$ וגם $2r$ נמצאים בשתי סדרות האינדכסים של תת־המחרוזות המשותפת). זה קורה כאשר $v_{2r-1} = w_{2r-1}$ וגם $v_{2r} = w_{2r}$.
- בסיכוי $\frac{1}{8}$ לא תהיה לנו אף התאמה. זה קורה רק כאשר מתקיים $v_{2r-1} = v_{2r} \neq w_{2r-1} = w_{2r}$.
- בסיכוי הנותר $\frac{5}{8}$ תהיה לנו התאמה בודדת (כאשר יש שתי אפשרויות לבחור את ההתאמה הבודדת).

סה"כ, תוחלת מספר ההתאמות בזוג $(2r-1, 2r)$ היא $\frac{9}{8}$. מכיוון שיש לנו $n/2$ זוגות סה"כ, לפי לינאריות התוחלת קיבלנו שעבור תתי־מחרוזות מהטיפוס המסויים הזה, תוחלת האורך המקסימלי היא αn עבור $\alpha = \frac{9}{16} > \frac{1}{2}$. מכאן שזה חוסם מלמטה את התוחלת של האורך המקסימלי הכללי $k(v, w)$ עבור n זוגי.

פתרונות לתרגיל השלישי

סף למעגלים

למען פישוט הסימונים, נסמן p במקום $p(n)$. כמו כן נשתמש בסימון $\mathcal{C}$ עבור קבוצת כל המעגלים מגודל k בגרף השלם עם n צמתים (זוהי "קבוצת כל המעגלים האפשריים ב־G"), כפי שנעשה בהדרכה לשאלה.

ראשית נוכיח במהירות את החסם התחתון: עבור מעגל ספציפי C מגודל k, הסיכוי שהוא יהיה מוכל ב־G הוא בדיוק p^k. מספר המעגלים האפשריים הוא $|\mathcal{C}| = O(n^k)$; לאלו הזוכרים את בניית הגרפים עם מותן גבוהה מההרצאה על הגרלה עם תיקונים, המספר המדויק הוא $\frac{k!}{2k}\binom{n}{k}$. מכאן, לפי החסם על איחוד מאורעות, הסיכוי לקיום מעגל מתוך $\mathcal{C}$ ב־G חסום ע"י $O(n^k p^k) = O((np)^k)$. מכיוון ש־k קבוע, אם $p = o(1/n)$ אז הביטוי הזה שואף ל־0 כאשר n שואף ל־∞.

עתה נעבור לחסם העליון. נניח אם כן שמתקיים $p = \omega(1/n)$. לכל $C \in \mathcal{C}$ נסמן ב־X_C את מ"מ האינדיקטור עבור הקיום של C ב־G, ונסמן את מספר המעגלים הכולל ב־G ב־$X = \sum_{C\in\mathcal{C}} X_C$. בדומה למה שנעשה בהרצאה עבור פונקצית סף לקיום K_4, נראה אצלנו שעבור k קבוע מתקיים $\mathrm{V}[X] = o((\mathrm{E}[X])^2)$, אשר לפי אי־שוויון צ'בישף ייתן לנו את החסם המבוקש על ההסתברות לאי־קיום מעגל. קודם כל נשים לב שחישוב מהיר נותן $\mathrm{E}[X] = |\mathcal{C}|p^k = \Omega((np)^k)$.

עתה, עבור $k \le l \le 0$, נסמן ב־$\mathcal{I}_l$ את קבוצת זוגות המעגלים (C, C') שמקיימים $C \neq C'$ ויש להם בדיוק l צמתים משותפים (הדרישה $C \neq C'$ היא על מנת שלא להכליל זוגות מהצורה (C, C) ב־$\mathcal{I}_k$). מתקיים אם כן $\mathrm{V}[X] = \sum_{C,C'\in\mathcal{C}} \mathrm{Cov}[X_C, X_{C'}] = \sum_{C\in\mathcal{C}} \mathrm{V}[X_C] + \sum_{l=0}^{k} \sum_{(C,C')\in\mathcal{I}_l} \mathrm{Cov}[X_C, X_{C'}]$. עתה נחסום כל אחד מהסכומים הנ"ל לחוד.

- חישוב ישיר עבור סכום השונויות נותן $\sum_{C\in\mathcal{C}} \mathrm{V}[X_C] = |\mathcal{C}|(p^k - p^{2k}) = O(n^k p^k) = O((np)^k)$.
- עבור $(C, C') \in \mathcal{I}_0 \cup \mathcal{I}_1$ יש אי־תלות של X_C ב־$X_{C'}$, ולכן הסכומים עבור $l = 0, 1$ מתאפסים.
- עבור $1 < l \le k$, נשים לב שאם $(C, C') \in \mathcal{I}_l$, אז יש ל־$C$ ו־C' לכל היותר $l - 1$ קשתות משותפות (זה נובע מהטענה שהוזכרה בהדרכה לשאלה). על כן $\mathrm{Cov}[X_C, X_{C'}] \le \Pr[C, C' \subset G] \le p^{2k+1-l}$ (בדומה לחסימת הקוואריאנס עבור מ"מ אינדיקטור מההרצאה). כמו כן מתקיים $|\mathcal{I}_l| = O(n^{2k-l})$, מכיוון שאפשר להגדיר כל זוג מעגלים כזה ע"י סידרה של $2k - l$ צמתים (תהיה יותר מסידרה אחת שתגדיר כל זוג $(C, C') \in \mathcal{I}_l$), אבל בכל מקרה אנחנו מעונינים בחסם עליון). מכל אלו מתקיים $\sum_{(C,C')\in\mathcal{I}_l} \mathrm{Cov}[X_C, X_{C'}] = O(n^{2k-l}p^{2k+1-l}) = O(p(np)^{2k-l}) = O((np)^{2k-l})$.

לסיום, נשווה כל אחד מהמחוברים (יש לנו $k - 1$ מחוברים שלא מתאפסים) לחוד מול $(\mathrm{E}[X])^2$. מכיוון ש־np שואף ל־∞ (כזכור הנחנו שמתקיים $p = \omega(1/n)$), מתקיים $(np)^k = o((np)^{2k}) = o((\mathrm{E}[X])^2)$, ובאופן דומה לכל $1 < l \le k$ מתקיים $(np)^{2k-l} = o((\mathrm{E}[X])^2)$. לכן זה נכון גם עבור הסכום של המחוברים (שמספרם קבוע), ז"א שמתקיים $\mathrm{V}[X] = o((\mathrm{E}[X])^2)$.

מרטינגל בריבוע

לשם הדגמה, הפתרון כאן יותר מפורט ממה שאתם הייתם צריכים לכתוב בשביל לקבל את מלוא הנקודות.

על מנת להראות את תכונת חוסר הזיכרון, מחשבים:

$$\begin{aligned} Z_m = (X_m)^2 - m &= (X_{m-1} + Y_m)^2 - m \\ &= (X_{m-1})^2 - (m-1) + 2X_{m-1}Y_m + Y_m^2 - 1 = Z_{m-1} + 2X_{m-1}Y_m \end{aligned}$$

שימו לב שהשתמשנו בכך שתמיד מתקיים $(Y_m)^2 = 1$, כי הוא נבחר מתוך $\{1, -1\}$.

על מנת להשלים את הטיעון, נשים לב שגם בהתניה מהסוג $Z_0 = a_0, \ldots, Z_{m-1} = a_{m-1}$ יתקיים ש־Y_m הוא ב"ת ב־$Y_1, \ldots, Y_{m-1}$. הסיבה לכך היא ש־$Z_0, \ldots, Z_{m-1}$ נקבעים ע"י $Y_1, \ldots, Y_{m-1}$ בלבד, ואנחנו כבר יודעים שהתפלגות Y_m תישאר יוניפורמית בכל התניה על קומבינציה של תנאים שקשורים בערכי $Y_1, \ldots, Y_{m-1}$ (במקרה הזה, דרך תנאים על ערכי X_{m-1} ו־$Z_0, \ldots, Z_{m-1}$).

בנוסף לכך (גם משימור ההתפלגות של Y_m) מתקיים $\mathrm{E}[Y_m | Z_0 = a_0, \ldots, Z_{m-1} = a_{m-1}] = 0$, ולכן

$$\begin{aligned}\mathrm{E}[Z_m | Z_0 = a_0, \ldots, Z_{m-1} = a_{m-1}] &= \mathrm{E}[Z_{m-1} | a_0, \ldots, a_{m-1}] + 2\mathrm{E}[X_{m-1} Y_m | a_0, \ldots, a_{m-1}] \\ &= a_{m-1} + 2\mathrm{E}[X_{m-1} | a_0, \ldots, a_{m-1}]\mathrm{E}[Y_m | a_0, \ldots, a_{m-1}] = a_{m-1}\end{aligned}$$

הערה: הסיבה שאי אפשר להשתמש בזה עבור חסימת $\Pr[X_m > 2\sqrt{m}]$ היא שאין חסם טוב על $|Z_i - Z_{i-1}|$.

מתמקדים

נניח שמתקיים $\delta < \frac{1}{3}$, אחרת אפשר פשוט לקחת הרצה בודדת של התהליך ההסתברותי ולפלוט את התוצאה. עבור k אי־זוגי שנבחר בקרוב, נבצע k פעמים את התהליך ההסתברותי, ומתוך סדרת התוצאות $c_1, \ldots, c_k$ נפלוט את החציון, ז"א את הערך c_j שעבורו $|\{i : c_i \leq c_j\}| \geq \frac{k}{2}$ וגם $|\{i : c_i \geq c_j\}| \geq \frac{k}{2}$ (צורת הכתיבה כאן מתאימה גם למקרה שיש שוויונות בסדרת הערכים).

על מנת שהפלט יהיה קטן מ־a, צריך שיהיו לפחות $\frac{k}{2}$ קריאות לתהליך ההסתברותי שנתנו ערך קטן מ־a, ועל מנת שהפלט יהיה גדול מ־b, צריך שיהיו לפחות $\frac{k}{2}$ קריאות שנתנו ערך גדול מ־b. בשני המקרים זה אומר שלפחות $\frac{k}{2}$ מהקריאות לא נתנו ערך בין a ל־b. כמו כן, בשאלה נתון שבכל קריאה בודדת הסיכוי לקבל ערך שאינו בין a ל־b חסום ע"י $\frac{1}{3}$.

נגדיר עתה מ"מ אינדיקטור X_i עבור המאורע "לא מתקיים $a \leq c_i \leq b$". לפי נתוני השאלה $X_1, \ldots, X_n$ הם מ"מ ב"ת שכל אחד מהם מקבל 1 בהסתברות שאינה עולה על $\frac{1}{3}$, ומקבל 0 אחרת. מהניתוח למעלה עולה שבפרט מתקיים $\sum_{i=1}^{k} X_i \geq \frac{k}{2}$ בכל מקרה שפלט האלגוריתם אינו בין a ל־b. לפי חסימת סטיות גדולות (אי־השוויון השני מהשיעור), ההסתברות עבור פלט שגוי חסומה ע"י $e^{-2(k/2-k/3)^2/k} = e^{-k/18}$. אם רוצים שחסם זה יהיה קטן מ־δ, אפשר לקחת למשל $k = 18\lceil \log(1/\delta) \rceil + 1 = O(\log(1/\delta))$.

פתרונות לתרגיל הרביעי

לרצות את כולם

השאלה מדברת על מרחבי הסתברות מרובים, אולם עיקר הפתרון הוא הזיהוי של מרחב ההסתברות שאנחנו מנתחים. קבוצת הבסיס שלנו תהיה קבוצת המאורעות האפשריים מעל S, ז"א $\mathcal{P}(S)$, והבחירה תהיה יוניפורמית מתוכה. כל משפחה $\mathcal{A}_i \subseteq \mathcal{P}(S)$ שמתאימה למאורע "$\mu_i(E) \geq \frac{1}{2}$" היא מונוטונית לא־יורדת.

משפט קלייטמן, במושגים הסתברותיים, אומר לנו שלכל שתי משפחות מונוטוניות לא־יורדות $\mathcal{A}$ ו־$\mathcal{B}$ מתקיים $\Pr[E \in \mathcal{A} \wedge E \in \mathcal{B}] \geq \Pr[E \in \mathcal{A}]\Pr[E \in \mathcal{B}]$. מכאן אפשר להוכיח באינדוקציה שמתקיים $\Pr[\bigwedge_{i=1}^{k} E \in \mathcal{A}_i] \geq \Pr[E \in \mathcal{A}_1]\Pr[\bigwedge_{i=2}^{k} E \in \mathcal{A}_i] \geq \cdots \geq \prod_{i=1}^{k} \Pr[E \in \mathcal{A}_i]$.

לבסוף, מכיוון שלכל קבוצה E תמיד מתקיים $\mu_i(E) \geq \frac{1}{2}$ או $\mu_i(S \setminus E) \geq \frac{1}{2}$ (או שניהם), מתקיים $\Pr[E \in \mathcal{A}_i] \geq \frac{1}{2}$ לכל $1 \leq i \leq k$. לכן מאי־השוויון למעלה מתקיים $\Pr[\bigwedge_{i=1}^{k} E \in \mathcal{A}_i] \geq 2^{-k}$ כנדרש.

בלי עצמות במרק

הפתרון כאן הוא יישום די מיידי של הפרדיגמה של פואסון שנלמדה בתרגול (החסם התחתון אפשרי להוכחה גם דרך קורלציות). אנחנו נשתמש בסימונים של משפט ינסון כפי שמופיעים בחוברת, ז"א שנתייחס לגרף תלויות D, שהצמתים שלו מתאימים למאורעות של הכלת קבוצות ב־U, והקשתות שלו מתאימות לזוגות של קבוצות שאינן זרות. אצלנו, קבוצת הצמתים של גרף התלויות D תהיה בדיוק קבוצת m הקשתות של G, ו"צמתים" e, e' של D יהוו קשת של D אם הקשתות המתאימות ב־G הן בעלות צומת משותף של G. לכל מאורע B_e מתקיים $\Pr[B_e] = p^2$ (כי כל צומת של G נבחר בהסתברות p באופן ב"ת). עבור e, e' שיש להן צומת משותף ב־G, מתקיים $\Pr[B_e \wedge B_{e'}] = p^3$. לפי הנתון על הדרגה המקסימלית של G, אפשר לחסום עתה $\Delta = p^3|\{e, e' \in E(G) : e \cap e' \neq \emptyset\}| \leq 2mdp^3$.

לבסוף נחשב את $M = (1-p^2)^m$. ממשפט ינסון נקבל עתה

$$(1-p^2)^m \leq \Pr[\bigwedge_{e \in E(G)} \neg B_e] \leq (1-p^2)^m e^{mdp^3/(1-p^2)} = (1-p^2)^{m(1+dp^3/(1-p^2)\ln(1-p^2))}$$

נותר רק לשים לב שלפי הנתון של $p = \frac{1}{\log(m)}$ מתקיים $p = o(1)$, ולכן גם $\frac{-dp^3}{(1-p^2)\ln(1-p^2)} = O(dp) = o(1)$, מה שמביא לנו את האיבר הימני לצורה הנדרשת $(1-p^2)^{(1-o(1))m}$.

לא ממריא

פתרון ראשון: נסמן ב־a את התוחלת של X, ונניח $a > 0$ (אחרת $\mathrm{H}[X] = 0$ וסיימנו). זה נוח (אם כי אפשר גם בלי) להגדיר את המשתנה המקרי הנוסף הבא: $Y = \max\{0, \lfloor \log(X/a) \rfloor\}$. במילים אחרות, Y הוא המספר הטבעי המינימלי k שעבורו מתקיים $X \leq a2^Y$. מכיוון ש־Y הוא פונקציה של X, מתקיים $\mathrm{H}[X] = \mathrm{H}[X,Y] = \mathrm{H}[Y] + \mathrm{H}[X|Y]$. עתה נחסום את שני המחוברים מימין.

העיקר הוא לשים לב שלפי אי־שוויון מרקוב מתקיים $\Pr[Y = k] \leq \Pr[X > a2^{k-1}] \leq 2^{1-k}$. מכאן חוסמים את המחובר הראשון לפי $\mathrm{H}[Y] = \sum_{k=0}^{\infty} \Pr[Y=k] \log \frac{1}{\Pr[Y=k]} \leq 3 + \sum_{k=3}^{\infty} k2^{1-k} \leq 6$ (עבור $0 < x < \frac{1}{e}$ הפונקציה $x \log \frac{1}{x} = (x \ln \frac{1}{x})/\ln 2$ עולה לפי גזירה, ומפתחים $\sum_{k=1}^{\infty} k2^{-k} = \sum_{i=1}^{\infty}(\sum_{j=i}^{\infty} 2^{-j}) = 2$).

עבור חסימת המחובר השני, ניזכר בהגדרה $\mathrm{H}[X|Y] = \sum_{k=0}^{\infty} \Pr[Y=k]\mathrm{H}[X|Y=k]$. ניתן לחסום כל מחובר לחוד ע"י $\Pr[Y=k]\mathrm{H}[X|Y=k] \leq 2^{1-k}\log(a2^k+1) \leq 2^{1-k}(\log(a)+1+k)$ (השתמשנו בכך ש־X יכול לקבל לא יותר מ־$a2^k+1$ ערכים שונים כאשר מתנים על $Y=k$). לכן יש לנו את החסם

$\mathrm{H}[Y] \leq 4\log(a)+14$ וסה"כ $\mathrm{H}[X|Y] \leq \sum_{k=0}^{\infty} 2^{1-k}(\log(a)+1+k) \leq 4\log(a)+8$, **(כמובן שממש לא ניסינו לתת כאן את הביטוי הכי טוב האפשרי).**

פתרון שני: לשם הנוחות לכל $i \in \mathbb{N}$ נסמן $p_i = \Pr[X=i]$. נחלק את קבוצת המספרים הטבעיים לפי ההסתברות שלהם ל־$A = \{i \in \mathbb{N} : p_i \leq 2^{-i}\}$ ו־$B = \{i \in \mathbb{N} : p_i > 2^{-i}\}$. מתקיים השוויון $\mathrm{H}[X] = \sum_{i\in A} p_i \log \frac{1}{p_i} + \sum_{i\in B} p_i \log \frac{1}{p_i}$, ונחסום כל מחובר בנפרד.

עבור המחובר הראשון, מתקיים $\sum_{i\in A} p_i \log \frac{1}{p_i} \leq 3 + \sum_{i=2}^{\infty} i \cdot 2^{-i} \leq 5$ (גם כאן השתמשנו במונוטוניות של $x \log \frac{1}{x}$ עבור $0 < x < \frac{1}{e}$). עבור המחובר השני, נשתמש ב־$\sum_{i\in B} p_i \log \frac{1}{p_i} \leq \sum_{i\in B} i \cdot p_i \leq \mathrm{E}[X]$.

סודות ושקרים

כפי שנאמר בהדרכה, התשובה הסופית של האלגוריתם היא פונקציה של סדרת התשובות שניתנו ב־q השלבים. נסמן את התשובה בשלב ה־i ב־$a_i \in \{0,1\}$ (0 בשביל "לא" ו־1 בשביל "כן"), ואת סידרת כל התשובות ב־$X = (a_1,\ldots,a_q) \in \{0,1\}^q$. נניח ש־$k$ נבחר באופן יוניפורמי מתוך $\{1,\ldots,n\}$, ונסמן אותו כמשתנה מקרי $Y = k$. לבסוף נסמן את התשובה של האלגוריתם עצמו גם כמשתנה מקרי, $Z = k'$. שימו לב בפרט ש־Z היא פונקציה של X. נניח שהאלגוריתם מקיים $Y = Z$ בהסתברות לפחות $1-\alpha$ (את הערך הספציפי של α נבחר לקראת סוף ההוכחה), ונראה חסם תחתון על האנטרופיה $\mathrm{H}[X]$. מכיוון שתמיד מתקיים $\mathrm{H}[X] \leq \log|\{0,1\}^q| = q$, זה ייתן לנו חסם תחתון על q.

ראשית נחסום את $\mathrm{H}[X,Y] - \mathrm{H}[X] = \mathrm{H}[Y|X] = \mathrm{H}[Y|X,Z] \leq \mathrm{H}[Y|Z] = \mathrm{H}[Y,Z] - \mathrm{H}[Z]$ (השתמשנו בשוויון השני בכך ש־Z הוא פונקציה של X): מכיוון ש־Z שווה ל־Y בהסתברות לפחות $1-\alpha$, לכל k מתקיים $\Pr[Y=Z=k] \geq (1-\alpha)/n$. סה"כ ההסתברות לכל המאורעות מהצורה $Y = k \neq k' = Z$ חסום ע"י α. נשים לב שתחת האילוצים האלו, הביטוי

$$\mathrm{H}[Y,Z] = \sum_{k=1}^{n} \Pr[Y=Z=k] \log \frac{1}{\Pr[Y=Z=k]} + \sum_{k\neq k'} \Pr[Y=k \wedge Z=k'] \log \frac{1}{\Pr[Y=k \wedge Z=k']}$$

יקבל את המקסימום כאשר ההתפלגות על ערכי Y,Z היא יוניפורמית בהתניה על המקרה $Y \neq Z$ (בדומה לכך שהאנטרופיה של מ"מ כל שהוא מקבלת את המקסימום כאשר הוא מתפלג יוניפורמית). זה נותן לנו

$$\mathrm{H}[Y,Z] \leq \log(n) + n(n-1) \cdot \frac{\alpha}{n(n-1)} \log \frac{n(n-1)}{\alpha} \leq \log(n) + \alpha(\log(1/\alpha) + 2\log(n))$$

נניח ש־n גדול מספיק (ביחס ל־α שאח"כ נבחר) ע"מ לקיים $\log(1/\alpha) < \log(n)$. לכל k מתקיים $\Pr[Z=k] \geq \Pr[Y=Z=k] \geq (1-\alpha)/n$, ואם נניח גם ש־$n \geq 3$ נובע מזה (ומהמונוטוניות של $x \log \frac{1}{x}$ בתחום $0 < x < \frac{1}{e}$) שמתקיים $\Pr[Z=k] \log \frac{1}{\Pr[Z=k]} \geq \frac{1-\alpha}{n} \log \frac{n}{1-\alpha}$. כשסוכמים על $1 \leq k \leq n$ מקבלים $\mathrm{H}[Z] \geq (1-\alpha)\log \frac{n}{1-\alpha} > (1-\alpha)\log(n)$, וסה"כ $\mathrm{H}[X,Y] - \mathrm{H}[X] \leq \mathrm{H}[Y,Z] - \mathrm{H}[Z] \leq 4\alpha\log(n)$.

עתה נחסום מהצד השני את $\mathrm{H}[X,Y] = \mathrm{H}[X|Y] + \mathrm{H}[Y]$. לפי ההגדרות מתקיים $\mathrm{H}[Y] = \log(n)$. כמו כן, לפי הנתון שההסתברות ל"שקר" כל פעם שהאלגוריתם שואל שאלה היא בדיוק $\frac{1}{10}$ (באופן ב"ת בשאלות קודמות), אפילו בהתניה על ערך ספציפי של Y מתקיים $\mathrm{H}[X|Y=k] = qH(\frac{1}{10})$ (בצד ימין יש את "פונקצית האנטרופיה" המספרית, ובפרט יש שם מקדם קבוע גדול מ־0), ולכן $\mathrm{H}[X|Y] = qH(\frac{1}{10})$.

סה"כ אנחנו מקבלים $\mathrm{H}[X,Y] = qH(\frac{1}{10}) + \log(n)$. יחד עם חסם ההפרש בין $\mathrm{H}[X]$ ו־$\mathrm{H}[X,Y]$ נקבל $q \geq \mathrm{H}[X] \geq \mathrm{H}[X,Y] - 4\alpha\log(n) = qH(\frac{1}{10}) + (1-4\alpha)\log(n)$, ובהעברת אגפים $q \geq \frac{1-4\alpha}{1-H(1/10)}\log(n)$. עבור $\alpha = H(\frac{1}{10})/8$ ו־$\beta = H(\frac{1}{10})/16$ למשל נקבל את הנדרש $q \geq (1+\beta)\log(n)$.

הערה: חישוב יותר מדוייק של מספר השאלות (חסם עליון ותחתון), כאשר כל שקר הוא בהסתברות p, הושג ע"י Rényi בשנת **1961**, ועומד על $(1 \pm o(1))\log(n)/(1-H(p))$ עבור שגיאה חסומה ע"י קבוע.