מבני נתונים
פתרון תרגיל 3

גלעד אשרוב

1 במאי 2012

שאלה 1. הוכח את חסם האיחוד. כלומר, יהיו $A_1, \ldots, A_n$ מאורעות. אזי מתקיים:

$$\Pr\left[\bigvee_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \Pr\left[A_i\right]$$

פתרון: ראינו כי

$$\Pr\left[A \vee B\right] \leq \Pr\left[A\right] + \Pr\left[B\right]$$

הרעיון הוא פשוט להפעיל את אי השיוויון הנ"ל שוב ושוב. כלומר:

$$\begin{aligned} \Pr\left[\bigvee_{i=1}^{n} A_i\right] &= \Pr\left[A_1 \vee \bigvee_{i=2}^{n} A_i\right] \leq \Pr[A_1] + \Pr\left[\bigvee_{i=2}^{n} A_i\right] = \Pr\left[A_1\right] + \Pr\left[A_2 \vee \bigvee_{i=3}^{n} A_i\right] \\ &\leq \quad \Pr[A_1] + \Pr[A_2] + \Pr\left[\bigvee_{i=3}^{n} A_i\right] \leq \ldots \leq \Pr\left[A_1\right] + \Pr\left[A_2\right] + \ldots \Pr\left[A_n\right] = \sum_{i=1}^{n} \Pr\left[A_i\right] \end{aligned}$$

שאלה 2. נניח אלגוריתם אקראי A שעבור כל קלט x, מחזיר את התוצאה הנכונה בהסתברות $1/2$, ומחזיר "לא יודע" בהסתברות $1/2$. השאלה האם יחזיר את התשובה הנכונה או "לא יודע" תלויה באקראיות של האלגוריתם A בלבד, ולכן, ייתכן בהחלט שאם החזיר "לא יודע" בפעם הראשונה ונריץ אותו שוב, יחזיר את התוצאה הנכונה שוב בהסתברות $1/2$.

כעת, נניח אלגוריתם B העובד בצורה הבאה: B מריץ את A שוב ושוב, כל עוד A מחזיר "לא יודע". ברגע ש־A מחזיר תשובה אחרת, B עוצר, ומחזיר את התשובה ש־A החזיר.

1. מהי ההסתברות שהלאוגירתם B יצטרך להריץ את A יותר מ-100 פעמים בכדי לקבל תשובה נכונה?

   פתרון: האלגוריתם צריך להיכשל 100 פעמים, ולכן ההסתברות היא 2^{-100}.

2. מהי תוחלת מספר הפעמים שהאלגוריתם B מריץ את A?

   פתרון: החישוב הוא פשוט מה שראינו בכיתה לגבי תוחלת מספר הפעמים שמטילים מטבע עד שמקבלים "עץ". נקבל תוחלת - 2.

3. חזור על שני הסעיפים הקודמים תחת ההנחה שאלגוריתם A מחזיר תשובה נכונה בהסתברות p ו"לא יודע" בהתסברות $1-p$, עבור איזשהו $0 < p < 1$.

פתרון: ההסתברות שנרוץ 100 סיבובים היא: p^{100}.
ההסתברות שנעצור אחרי i סיבובים היא: $(1-p)^{i-1}p$. נסמן ב-X את המשתנה המקרי המייצג את מספר הסיבובים. נקבל:

$$\begin{aligned}
E\left[X\right] &= \sum_{i=1}^{\infty}(1-p)^{i-1}\cdot p\cdot i = p\cdot\sum_{i=1}^{\infty}(1-p)^{i-1}\cdot i \\
&= p\cdot\Big((1-p)^0+2\cdot(1-p)^1+2\cdot(1-p)^2+\ldots\Big) \\
&= p\cdot\Big(\Big((1-p)^0+(1-p)^1+(1-p)^2+\ldots\Big)+\Big((1-p)^1+(1-p)^2+\ldots\Big)+\ldots\Big) \\
&= p\cdot\left(\frac{1}{p}+\frac{1-p}{p}+\frac{(1-p)^2}{p}+\ldots\right) \\
&= 1+(1-p)+(1-p)^2+\ldots=\frac{1}{p}
\end{aligned}$$

שאלה 3. נניח אנו מטילים קוביה 10 פעמים.

1. מהי ההסתברות שבכל הפעמים נקבל 6?

   פתרון: $(\frac{1}{6})^{10}$.

2. מהי ההסתברות שבכל הפעמים נקבל את אותה התוצאה?

   פתרון: $(\frac{1}{6})^{9}$.

3. מהי ההסתברות שבדיוק ב־4 הטלות נקבל את התוצאה 5?

   פתרון: $\binom{10}{4}\left(\frac{1}{6}\right)^4\cdot\left(\frac{5}{6}\right)^6$.

4. תאר נוסחא כללית. מהי ההסתברות שב־k מהמקרים נקבל את התוצאה x? (עבור $x\in\{1,\ldots,6\}$

   פתרון: $\binom{10}{k}\left(\frac{1}{6}\right)^k\cdot\left(\frac{5}{6}\right)^{10-k}$.

5. מהי תוחלת סכום ההטלות?

   פתרון: 35.

שאלה 4. מגרילים גרף באקראי עם n קודקודים בצורה הבאה: לכל זוג קודקודים בגרף מטילים מטבע השווה 1 בהסתברות p (עבור $0<p<1$). אם המטבע יצא 1 - מגדירים קשת בין זוג הקודוקדים בגרף.

1. חשב את תוחלת מספר הקשתות בגרף?

פתרון: נסדר את הקשתות לפי סדר $1,\ldots,\binom{k}{2}$. יהי X_i משתנה מקרי המקבל 1 אם הקשת i הוגרלה, ו-0 אחרת. יהי X משתנה מקרי המסמל את מספר הקשתות בגרף. ברור כי $X=\sum_{i=1}^{\binom{k}{2}}X_i$. כמו־כן, ברור כי $E[X_i]=p$ לכל i. נקבל:

$$E[X]=E\left[\sum_{i=1}^{\binom{k}{2}}X_i\right]=\sum_{i=1}^{\binom{k}{2}}E[X_i]=\sum_{i=1}^{\binom{k}{2}}p=\binom{k}{2}p$$

2. קליק בגרף בגודל k היא תת קבוצה של קודקודים בגודל k כך שקיימת קשת בין כל זוג קודקודים בתת קבוצה זו. מצא חסם תחתון להסתברות שלא קיים בגרף קליק בגודל k?

פתרון: נקבע תת קבוצה של קודקודים בגודל k - A. מספר הקשתות בין k הקודקודים הנ"ל הוא - $\binom{k}{2}$. ההסתברות שיש קליק בתת קבוצה זו היא:

$$\Pr\left[A \text{ is clique}\right] = p^{\binom{k}{2}}$$

כעת, ישנן $\binom{n}{k}$ קבוצות בגודל k. לכן, נקבל כי:

$$\Pr\left[G \text{ has } k-\text{clique}\right] \leq \binom{n}{k} p^{\binom{k}{2}}$$

ולכן,

$$\Pr\left[G \text{ does not have } k-\text{clique}\right] = 1 - \Pr\left[G \text{ has } k-\text{clique}\right] \geq 1 - \binom{n}{k} p^{\binom{k}{2}}$$

שאלה. באוניברסיטה מסוימת במרכז הארץ, נלמד הקורס "מבני נתונים". הסטודנטים בקורס קיבלו תרגיל לשיעורי בית ובו רשימה של פונקציות שעליהם למיין לפי סדר גודל. בנוסף, הסטודנטים התבקשו לבחור מהרשימה 5 זוגות של פונקציות, ולהוכיח לגביהן טענה מסוימת.

1. בהינתן שברשימה היו סה"כ 16 פונקציות - כמה זוגות של פונקציות ניתן להרכיב בסה"כ?
2. תחת ההנחה שכל סטודנט בחר באקראי בהתפלגות אחידה ובצורה בלתי תלויה את חמשת הזוגות, מה ההסתברות ששני סטודנטים נתונים יבחרו **בדיוק** את אותם חמשת הזוגות?
3. אם ישנם 100 סטודנטים בקורס בסה"כ, מצא חסם עליון להסתברות שישנם שני סטודנטים כלשהם עם אותם חמישה זוגות בדיוק?
4. **בונוס:** בהינתן התוצאות שלעיל, מה מתרגלי הקורס עלולים להסיק אם נמצאו שני סטודנטים כלשהם עם אותם חמישה זוגות בדיוק? :-)

פתרון:

1. ניתן להרכיב $\binom{16}{2} = 16 \cdot 15/2 = 120$ זוגות בסה"כ.
2. אנחנו קובעים את הבחירה של הסטודנט הראשון, ושואלים מה ההסתברות שהסטודנט השני בוחר בדיוק אותן הפונקציות. כלומר - אנחנו שואלים מה ההסתברות שהסטודנט השני בחר חמש זוגות ספציפיים מתוך רשימה של 120 זוגות. נקבל:

$$\Pr\left[S_2 \text{ chooses the same functions as } S_1\right] = \frac{1}{\binom{120}{5}} = \frac{5!}{120 \cdot 119 \cdot 118 \cdot 117 \cdot 116} = 5.24 \cdot 10^{-9}$$

3. עבור שני סטודנטים נתונים, ההסתברות ששניהם בחרו בדיוק אותן שתי פונקציות היא $5.24 \cdot 10^{-9}$. ההסתברות שישנם שני סטודנטים כלשהם שבחרו את אותן הפונקציות חסומה ע"י:

$$\begin{aligned}&\Pr\left[\exists S_1, S_2 \text{ that chose the same 5 functions}\right] \\ \leq\ & \binom{100}{2} \Pr\left[S_2 \text{ chooses the same functions as } S_1\right] = 4950 \cdot 5.24 \cdot 10^{-9} = 2.59 \cdot 10^{-5}\end{aligned}$$