למידת מכונה - מטלה 1

מגישים: איתי רפיעי, אלמוג יעקב מעטוף

1. עבור כל צד בקובייה i אנו צריכים שיתקיים $|X_i - p_i| \le \epsilon$
כלומר, $P(|X_1 - p_1| > \epsilon)$ ***OR*** $P(|X_2 - p_2| > \epsilon)$ ***OR*** $P(|X_3 - p_2| > \epsilon) \dots < \delta$
כאשר X_i הוא משתנה מקרי ברנולי הסופר את מספר המופעים של הספרה i בקובייה בהתאמה

$$X_i = \begin{cases} 0, & \text{If the digit } \boldsymbol{i} \text{ did not appear} \\ 1, & \text{If the digit } \boldsymbol{i} \text{ appeared} \end{cases}$$

ע"פ נתוני השאלה $\epsilon = 0.01$ שזהו ה"חלון" (המרחק [בערך מוחלט] מהממוצע).
כמו כן, $\delta = 0.01$ (חסם על הטעות) כיוון שאנו רוצים הצלחה בלפחות 0.99
(כלומר, אנו רוצים שתמיד יתקיים $\delta \le 0.01$ ולכן נבחר $\delta = 0.01$).
מכאן, צריך להתקיים:

$$\sum_{i=1}^{6} P(|X_i - p_i| > 0.01) \le 0.01$$

המשתנים המקריים שהגדרנו הינם משתנים ברנולים לכן נוכל להפעיל את הופדינג על המשתנים. מכאן, ע"פ שווין הופדינג:

$$\sum_{i=1}^{6} P(|X_i - p_i| > 0.01) < 6 \cdot 2e^{-2n(0.01)^2} \le 0.01$$

כלומר, נדרוש:

$$6 \cdot 2e^{-2n(0.01)^2} \le 0.01$$

נבודד את n כפי שעשינו בהרצאה ונקבל:

$$n \ge 35451 \ge \frac{ln\left(\frac{12}{0.01}\right)}{2(0.01)^2} = 35450.384$$

2.

א. ע"פ הנתון, הטעות האמפירית עבור החוק הנ"ל הינה *0.15*. כלומר, $\bar{e}(h) = 0.15$.
נשים לב כי H מוגדר להיות סט החוקים המכיל את כל התוכניות בגודל 50 ביטים בדיוק.
לכן, מספר החוקים הקיימים לנו הוא 2^{50}.
נשתמש בחסם ההכללה שלמדנו בהרצאה על מנת לקבל חסם עליון על הטעות האמיתית:

$$e(h) \le \bar{e}(h) + \sqrt{\frac{ln\, 2|H| + ln\frac{1}{\delta}}{2n}}$$

כאשר δ היא הטעות. וכאשר n הוא מספר הנקודות.
נקבל:

$$e(h) \le 0.15 + \sqrt{\frac{ln(2 \cdot 2^{50}) + ln\frac{1}{\delta}}{2n}}$$

אבל אנו רוצים שהטעות האמיתית $e(h)$ תהיה חסומה מלמעלה ע"י $0.15 + \epsilon$. לכן נדרוש:

$$0.15 + \sqrt{\frac{ln(2 \cdot 2^{50}) + ln\frac{1}{\delta}}{2n}} \le 0.15 + \epsilon$$

נבודד את δ:

$$\sqrt{\frac{ln(2 \cdot 2^{50}) + ln\frac{1}{\delta}}{2n}} \le \epsilon$$

$$\frac{ln(2\cdot 2^{50})+ln\frac{1}{\delta}}{2n}\le\epsilon^2$$
$$ln\left(\frac{2\cdot 2^{50}}{\delta}\right)\le 2n\epsilon^2$$
$$\frac{2\cdot 2^{50}}{\delta}\le e^{2n\epsilon^2}$$
$$\boldsymbol{\frac{2\cdot 2^{50}}{e^{2n\epsilon^2}}\le\delta}$$

ב. נשתמש בסעיף א' על מנת לקבל את הערך המינימלי עבור ϵ.
על פי הנתון $\delta=0.05$, $n=1000$. לכן, נציב את הנתונים במשוואה שקיבלנו בסעיף א'

$$\frac{2\cdot 2^{50}}{e^{2\cdot 1000\cdot\epsilon^2}}\le 0.05$$

נבודד את ϵ:

$$\frac{2\cdot 2^{50}}{0.05}\le e^{2\cdot 1000\cdot\epsilon^2}$$
$$ln\left(\frac{2\cdot 2^{50}}{0.05}\right)\le 2\cdot 1000\cdot\epsilon^2$$
$$\frac{ln\left(\frac{2\cdot 2^{50}}{0.05}\right)}{2000}\le\epsilon^2$$
$$0.13846..=\sqrt{\frac{ln\left(\frac{2\cdot 2^{50}}{0.05}\right)}{2000}}\le\epsilon$$

כלומר, הערך המינימלי הינו $\epsilon=0.13846..$

ג. נשים לב כי מדובר על התפלגות אחרת, עבורה אין לנו שום מידע.
אם חוק עקבי על נקודות של התפלגות מסוימת אינו מחייב שהחוק יהיה עקבי על התפלגות אחרת.
לכן, נוכל לקבל טעות אמיתית שונה מזו שקיבלנו בסעיף ב' ואולי טעות אמיתית זהה.
מכאן, לא נוכל להסיק דבר מכך.

.3

א. נשים לב כי ישנם 7 סכומים שונים היכולים להתקבל משליפת שני קלפים (*2-8*)
נחשב את ההסתברות לקבלת כל סכום: (כאשר מס' הצירופים האפשריים הוא *4*4=16*)

הסכום	הצירופים	ההסתברות
2	(1,1)	1/16
3	(1,2), (2,1)	2/16
4	(1,3), (2,2), (3,1)	3/16
5	(1,4), (2,3), (3,2), (4,1)	4/16
6	(2,4), (3,3), (4,2)	3/16
7	(3,4), (4,3)	2/16
8	(4,4)	1/16

ולכן ההסתברות לתוצאת "תיקו" היא ההסתברות של כל סכום כפול ההסתברות לקבלת אותו סכום שנית. לכן ההסתברות לתוצאות תיקו הינה:

$$\frac{1}{16}\cdot\frac{1}{16}+\frac{2}{16}\cdot\frac{2}{16}+\frac{3}{16}\cdot\frac{3}{16}+\frac{4}{16}\cdot\frac{4}{16}+\frac{3}{16}\cdot\frac{3}{16}+\frac{2}{16}\cdot\frac{2}{16}+\frac{1}{16}\cdot\frac{1}{16}=\frac{11}{64}$$

נשים לב כי הסיכוי שאחד מהשחקנים ינצח הוא המשלים של ההסתברות לקבלת תיקו.
כלומר, $1-\frac{11}{64}=\frac{53}{64}$.
אך ע"פ הנתון הסיכוי של כל שחקן לנצח שווה ולכן בפרט ההסתברות של שחקן A לנצח
הינה: $\frac{\frac{53}{64}}{2}=\frac{53}{128}$. ובאופן דומה ההסתברות של שחקן B לנצח הינה: $\frac{\frac{53}{64}}{2}=\frac{53}{128}$.

ב. אם ישנם 100 סיבובים ממוצע הניצחונות של שחקן A היא כפל ההסתברות של שחקן A לנצח בסיבוב בודד כפול מספר הסיבובים ולכן נקבל: $\frac{53}{128}\cdot 100=\frac{1325}{32}=41.40625$.
באופן דומה נחשב את ממוצע הסיבובים בהם נקבל תיקו: $\frac{11}{64}\cdot 100=\frac{275}{16}=17.1875$

ג. עבור 100 סיבובים ניתן חסם עליון לסיכוי ששחקן A ינצח לכל הפחות ב50 סיבובים.
נשים לב כי נוכל להשתמש בממוצע הניצחונות שקיבלנו בסעיף ב' על מנת להציב בנוסחאות המתאימות.
נסמן: X משתנה מקרי הסופר את מספר הניצחונות של שחקן A.

<u>חישוב החסם העליון בעזרת מרקוב:</u>

$$P(X\geq 50)\leq\frac{E(x)}{50}=\frac{41.40625}{50}$$

<u>חישוב החסם העליון בעזרת צ'רנוף:</u>

נחשב את ערך δ המתאים (על מנת לקבל את המאורעות בהם $X\geq 50$):

$$(1+\delta)41.40625=50$$
$$\delta\cdot 41.40625=8.59375$$
$$\delta=\frac{11}{53}$$

לכן:

$$P(X\geq 50)=P\left(X\geq\left(1+\frac{11}{53}\right)41.40625\right)$$
$$=P\big(X\geq(1+\delta)E(X)\big)\leq e^{-(\delta)^2(E(X))/3}=e^{-\left(\frac{11}{53}\right)^2(41.40625)/3}$$

נשים לב כי אכן ערך δ מקיים את התנאי לשימוש במשוואת צ'רנוף.

<u>חישוב החסם העליון בעזרת הופדינג:</u>

$$P(X\geq 50)=P(X>49)$$
$$=P\left(\frac{X}{100}-\frac{53}{128}>\frac{49}{100}-\frac{53}{128}\right)\leq P\left(\left|\frac{X}{100}-\frac{53}{128}\right|>\frac{49}{100}-\frac{53}{128}\right)$$
$$=P\left(|X_i-E(X_i)|>\frac{243}{3200}\right)<2e^{-2(n)(\epsilon^2)}=2e^{-2(100)(\frac{243}{3200})^2}$$

ד.

<u>חישוב בעזרת צ'רנוף:</u>

נגדיר משתנה מקרי Y הסופר את מספר הסיבובים בהם שחקן A **לא ניצח**.
התוחלת עבור משתנה המקרי Y הינה $n\cdot\frac{75}{128}$. (המשלים להסתברות שA מנצח)
נחשב את ערך δ המתאים (על מנת לקבל את המאורעות בהם $Y>n-35$):

$$(1+\delta)(n\cdot\frac{75}{128})=n-34$$
$$(1+\delta)\cdot(75n)=128n-4352$$
$$1+\delta=\frac{128n-4352}{75n}=1.706-\frac{4352}{75n}$$
$$\delta=0.706-\frac{4352}{75n}$$

נחשב את ההסתברות לטעות:

$$P(Y > n - 35) = P(Y \geq n - 34) = P\left(Y \geq (1+\delta)(n \cdot \frac{75}{128})\right)$$

$$\leq e^{-(\delta)^2(E(Y))/3} = e^{-\left(0.706 - \frac{4352}{75n}\right)^2 (n \cdot \frac{75}{128})/3}$$

אבל אנו רוצים שהטעות תהיה קטנה ממש מ*0.01*.

לכן, נדרוש:

$$e^{-\left(0.706 - \frac{4352}{75n}\right)^2 (n \cdot \frac{75}{128}) \cdot (\frac{1}{3})} < 0.01$$

$$-((0.706 - 58.026/n)^2) \cdot \left(\frac{25n}{128}\right) < \ln(0.01)$$

$$-(0.195n) * ((0.706 - 58.026/n)^2) < \ln(0.01)$$

$$-\ln(0.01) < (0.195n) * ((0.706 - 58.026/n)^2)$$

$$\frac{-\ln(0.01)}{0.195} < \left(0.498436 + \frac{3367.094}{n^2} - \frac{81.9337}{n}\right) \cdot (n)$$

$$\frac{-\ln(0.01)}{0.195} < \left(0.498436n + \frac{3367.094}{n} - 81.9337\right)$$

$$\frac{-\ln(0.01)}{0.195} + 81.9337 < \left(0.498436n + \frac{3367.094}{n}\right)$$

$$0 < 0.498436n^2 - n(\frac{-\ln(0.01)}{0.195} + 81.9337) + 3367.094$$

קיבלנו שני פתרונות: $n_1 = 172.6,\ n_2 = 39.1$

הפרבולה מחייכת ולכן רק עבור $n > 172.6$ נקבל חסם תחתון על קיום המשוואה.

כלומר, שחקן *A* יצטרך לשחק לפחות 173 סיבובים על מנת לנצח בלפחות 35 סיבובים בהסתברות של לפחות 99%.

<u>חישוב בעזרת הופדינג:</u>

$$P(Y > n - 35) = P\left(\frac{Y}{n} - \frac{75}{128} > 1 - \frac{35}{n} - \frac{75}{128}\right)$$

$$= P\left(Y_i - E(Y_i) > \frac{53}{128} - \frac{35}{n}\right) \leq P\left(|Y_i - E(Y_i)| > \frac{53}{128} - \frac{35}{n}\right) < 2e^{-2(n)(\epsilon)^2}$$

אבל אנו רוצים שהטעות תהיה קטנה מ*0.01*.

לכן, נדרוש:

$$2e^{-2(n)(\frac{53}{128} - \frac{35}{n})^2} \leq 0.01$$

$$e^{-2(n)(\frac{53}{128} - \frac{35}{n})^2} \leq 0.005$$

$$-2(n)\left(\frac{53}{128} - \frac{35}{n}\right)^2 - \ln 0.005 \leq 0$$

$$-2(n)\left(\frac{1225}{n^2} - \frac{1855}{64n} + \frac{2809}{16384}\right) - \ln 0.005 \leq 0$$

$$63.2671 - \frac{2809n}{8192} - \frac{2450}{n} \leq 0$$

הפתרונות למשוואה הן: $n \geq 129.211$ וגם $0 < n \leq 55.2975$

הפתרון $0 < n \leq 55.2975$ נפסל מכיוון שעבור כל הצבה של n בטווח זה מביאה לכדי אפסילון שלילי.

מכאן, הפתרון האפשרי הינו $n \geq 129.211$. כלומר $n \geq 130$.

ואכן עבור $n \geq 130$ נקבל אפסילון חיובי.

נסיק כי שחקן *A* יצטרך לשחק לפחות 130 סיבובים על מנת לנצח בלפחות 35 סיבובים בהסתברות של לפחות 99%.