10.05.2023 ***ניתוח מידע אדפטיבי***

הרצאה 5: יציבות + אי שוויון אזומה־הופדינג

Source: Lecture notes by Aaron Roth and Adam Smith

מרצה: אורי שטמר

<u>תרגיל כיתה:</u> בהגדרה 3ג מהשיעור הקודם (יציבות של מכניזמים אינטראקטיביים) היריב A יכול להיות אקראי. הראו שכדי להוכיח שמכניזם כזה הוא DP-יציב מספיק להראות שתכונת היציבות מתקיימת לכל יריב A דטרמניסטי.

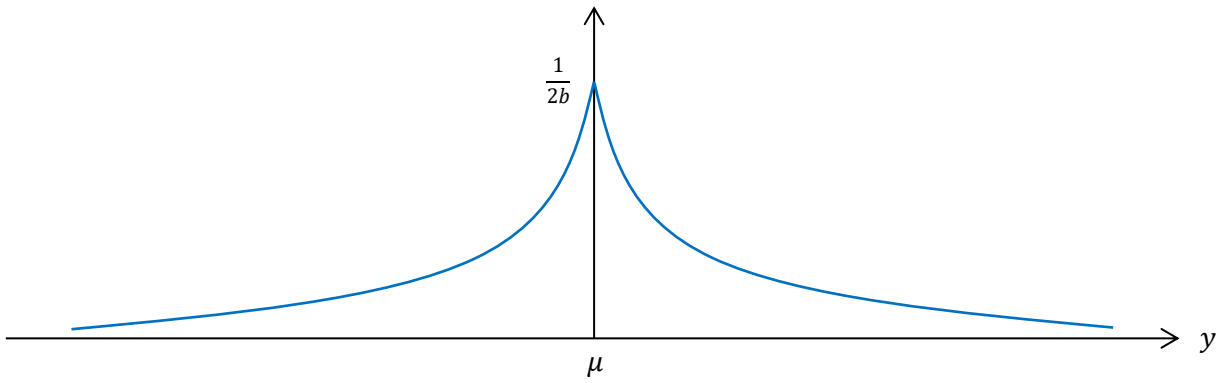
<u>שאלה:</u> האם מכניזם שעונה בעזרת הממוצע האמפירי המדויק הוא DP-יציב? מה לגבי מכניזם הקיטום?

עכשיו נתחיל לדבר על מכניזמים שמקיימים את תכונת היציבות.

<u>הגדרה 1 (התפלגות לפלס):</u> יהיו $b > 0$ ו- $\mu \in \mathbb{R}$ פרמטרים. למשתנה מקרי יש התפלגות $\text{Lap}(\mu, b)$ אם פונקציית הצפיפות שלו היא $h_{\mu,b}(x) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)$. עבור $\mu = 0$ נכתוב בקיצור $\text{Lap}(b)$.

(תזכורת: פונקציית צפיפות של משתנה מקרי היא מתארת את צפיפות המשתנה בכל נקודה במרחב המדגם. ההסתברות שמשתנה מקרי יימצא בקטע מסוים היא האינטגרל של הצפיפות בקטע ולכן המשתנה נוטה יותר לקבל ערכים שבהם הצפיפות גבוהה.)

$$h_{\mu,b}(y) = \frac{1}{2b} \cdot \exp\left(-\frac{|y-\mu|}{b}\right)$$



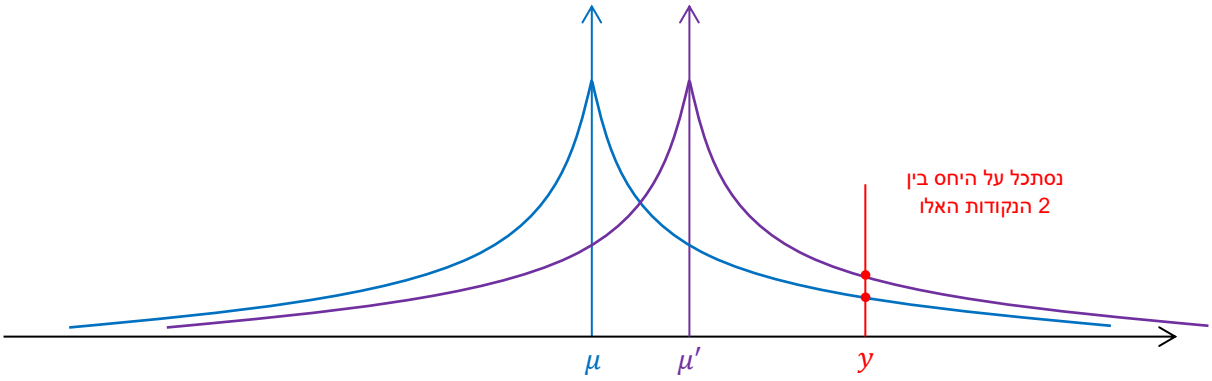
<u>אבחנה 2:</u> אם $Y \sim \text{Lap}(\mu, b)$ ונגדיר $X = Y + k$ עבור קבוע k כלשהו, אזי $X \sim \text{Lap}(\mu + k, b)$.

<u>הסבר:</u> זה נכון כי כאשר מוסיפים קבוע למשנה מקרי זה משנה את פונקציית הצפיפות על ידי "הזזה" של הגרף. ספציפית, אם f_Y היא פונקציית הצפיפות של מ"מ Y ואם $X = Y + k$ אזי פונקציית הצפיפות של X היא $f_X(x) = f_Y(x-k)$. במקרה שלנו, פונקציית הצפיפות של X היא

$$f_X(x) = h_{\mu,b}(x-k) = \frac{1}{2b}\exp\left(-\frac{|x-k-\mu|}{b}\right) = h_{\mu+k,b}(x)$$

<u>טענה 3 (תכונה של התפלגות לפלס):</u> לכל $\mu, \mu', b \in \mathbb{R}$ כך ש- $|\mu - \mu'| \leq \lambda$ ולכל $y \in \mathbb{R}$ מתקיים

$$\exp\left(-\frac{\lambda}{b}\right) \leq \frac{h_{\mu',b}(y)}{h_{\mu,b}(y)} \leq \exp\left(\frac{\lambda}{b}\right)$$



<u>הוכחה:</u>

$$\text{יחס} = \frac{h_{\mu',b}(y)}{h_{\mu,b}(y)} = \frac{\frac{1}{2b} \cdot \exp\left(-\frac{|y-\mu'|}{b}\right)}{\frac{1}{2b} \cdot \exp\left(-\frac{|y-\mu|}{b}\right)} = e^{\frac{1}{b} \cdot \underbrace{(|y-\mu|-|y-\mu'|)}_{\in[-\lambda,\lambda]}}$$

ולכן

$$e^{-\lambda/b} \leq \text{יחס} \leq e^{\lambda/b}$$

מ.ש.ל.

<u>הגדרה 4:</u> נגדיר את "מכניזם הלפלס" הבא, שיסומן כ- M^b_{Lap}, באופן הבא. המכניזם מקבל כקלט דטהבייס S ושאילתא q עם רגישות λ. המכניזם מחזיר $q(S) + Y$, עבור $Y \sim \text{Lap}(b\lambda)$.

<u>משפט 5:</u> יהי $\varepsilon > 0$ פרמטר. מכניזם הלפלס $M^{1/\varepsilon}_{\text{Lap}}$ הוא $(\varepsilon, 0)$-DP-יציב (עבור שאילתא אחת).

<u>הוכחה:</u> נקבע שני דטהבייסים שכנים $S, S' \in X^n$, ונקבע שאילתא q עם רגישות λ.

מתקיים $|q(S) - q(S')| \leq \lambda$ ולכן לכל $H \subseteq \mathbb{R}$ מתקיים:

$$\Pr\left[M^{1/\varepsilon}_{\text{Lap}}(S) \in H\right] = \int_H h_{q(S),b\lambda}(y)\,dy \underbrace{\leq}_{\substack{\text{לפי התכונה} \\ \text{שראינו קודם}}} \int_H e^{\varepsilon} \cdot h_{q(S'),b\lambda}(y)\,dy = e^{\varepsilon} \cdot \Pr\left[M^{1/\varepsilon}_{\text{Lap}}(S') \in H\right]$$

מ.ש.ל.

אוקיי, אז מכניזם הלפלס $M^{1/\varepsilon}_{\text{Lap}}$ הוא $(\varepsilon, 0)$-DP-יציב. כמה התשובות שהוא מחזיר מדוייקות ביחס לממוצע האמפירי? נשים לב שעוצמת הרעש שאנחנו מוסיפים תלויה ברגישות של השאילתא. מה קורה עבור רגישות $\lambda = 1/n$?

טענה 6: יהי $Y \sim \text{Lap}\left(\frac{1}{\varepsilon n}\right)$ ויהי $\Delta > 0$ כלשהו. אזי $\Pr[|Y| > \Delta] = \exp(-\varepsilon n \Delta)$.

הוכחה:

$$\Pr[Y > \Delta] = \int_{\Delta}^{\infty} \frac{\varepsilon n}{2} \cdot \exp(-\varepsilon n \cdot y)\, dy = \frac{\varepsilon n}{2} \cdot \exp(-\varepsilon n \cdot y) \cdot \left(\frac{1}{-\varepsilon n}\right) \Bigg|_{\Delta}^{\infty} = 0 - \frac{\varepsilon n}{2} \cdot \exp(-\varepsilon n \Delta) \cdot \left(\frac{1}{-\varepsilon n}\right)$$

$$= \frac{1}{2} \cdot \exp(-\varepsilon n \Delta)$$

(והכיוון השני באופן דומה)

מ.ש.ל.

המסקנה היא שההסתברות שהשגיאה תהייה גדולה מ $\frac{1}{\varepsilon n}$ דועכת אקספוננציאלית.

לאן אנחנו הולכים? הראנו שמכניזם הלפלס הוא גם DP-יציב וגם מדוייק אמפירית. כמו שהראינו עבור דחיסת-טרנסקריפט, בהמשך אנחנו נראה שזה גורר שמכניזם הלפלס הוא גם מדוייק סטטיסטית. אבל זה רק עבור שאילתא אחת... מה קורה עבור הרבה שאילתות? בשביל זה נצטרך להראות משפט קומפוזיציה עבור אלגוריתמים DP-יציבים. פה יהיה הכח של DP-יציבות לעומת דחיסת-טרנסקריפט, כי משפט הקומפוזיציה שנקבל יהיה הרבה יותר חזק. אבל יקח לנו זמן להגיע לזה. נתחיל עם טיעוני קומפוזיציה פשוטים יחסית. בהמשך נחזק אותם מאוד (וזה יהיה הרבה יותר מסובך...)

Basic Composition

משפט 7: אם M_1 הוא $(\varepsilon_1, \delta_1)$-*DP*-יציב ואם M_2 הוא $(\varepsilon_2, \delta_2)$-*DP*-יציב אזי (M_1, M_2) הוא $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-*DP*-יציב.

* כאן (M_1, M_2) מוגדר באופן הבא:
קלט: דטהבייס S
(1) חשב $y_1 \leftarrow M_1(S)$
(2) חשב $y_2 \leftarrow M_2(S)$
(3) החזר (y_1, y_2)

בפרט, אם M הוא יציב ואני אריץ אותו על אותו דטהבייס הרבה פעמים אז לאט לאט היציבות תתדרדר.

הוכחה עבור המקרה בו $\delta = 0$

נסמן:

$$M_1 : X^n \to R_1$$
$$M_2 : X^n \to R_2$$

* הנחה מפשטת: R_1, R_2 בנות מניה

יהיו S, S' שכנים ויהיו $r_1 \in R_1,\ r_2 \in R_2$.
נחשב:

$$\Pr[(M_1, M_2)(S) = (r_1, r_2)] = \Pr[M_1(S) = r_1] \cdot \Pr[M_2(S) = r_2]$$

$$\leq e^{\varepsilon_1} \cdot \Pr[M_1(S') = r_1] \cdot e^{\varepsilon_2} \cdot \Pr[M_2(S') = r_2]$$

$$\leq e^{\varepsilon_1 + \varepsilon_2} \cdot \Pr[(M_1, M_2)(S') = (r_1, r_2)]$$

מ.ש.ל.

שאלה: בהוכחה האחרונה הנחנו כי M_1, M_2 נקבעו מראש. האם אני יכול לבחור את המכניזם השני על סמך התוצאה של המכניזם הראשון?

משפט 8: יהי $M_1(\cdot)$ מכניזם המשמר $(\varepsilon_1, \delta_1)$-*DP*-יציבות ויהי $M_2(\cdot,\cdot)$ מכניזם כך שלכל פרמטר p מתקיים ש-$M_2(\cdot, p)$ משמר $(\varepsilon_2, \delta_2)$-*DP*-יציבות. אזי $M_3(S) = M_2\big(S, M_1(S)\big)$ משמר $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-*DP*-יציבות.

הוכחה עבור המקרה בו $\delta = 0$

$$\Pr\big[M_2\big(S, M_1(S)\big) = y\big] = \sum_p \Pr[M_2(S, p) = y] \cdot \Pr[M_1(S) = p]$$

$$\leq \sum_p e^{\varepsilon_2} \cdot \Pr[M_2(S', p) = y] \cdot e^{\varepsilon_1} \cdot \Pr[M_1(S') = p]$$

$$= e^{\varepsilon_1 + \varepsilon_2} \cdot \sum_p \Pr[M_2(S', p) = y] \cdot \Pr[M_1(S') = p]$$

$$= e^{\varepsilon_1 + \varepsilon_2} \cdot \Pr\big[M_2\big(S', M_1(S')\big) = y\big]$$

מ.ש.ל.

משפט הקומפוזיציה הנ"ל הם לא רעים. אבל הם לא יספיקו לנו כדי לקבל תוצאות טובות יותר מאשר מה שקיבלנו בעזרת דחיסה. לקראת הוכחת משפט קומפוזיציה חזק יותר עבור יציבות, נצטרך ללמוד הכללה של משפטי צ'רנוף-הופדינג שראינו בהרצאה הראשונה.

תזכורות מההרצאה הראשונה:

משפט 9 (חסם הופדינג):
יהיו $a, \mu \in \mathbb{R}$ ויהיו $X_1, X_2, \ldots, X_k$ משתנים מקריים בלתי תלויים כך שלכל $i \in [k]$ מתקיים $\Pr[|X_i| \leq a] = 1$ וגם $\mathbb{E}[X_i] \leq \mu$. אזי, לכל $z > 0$ מתקיים

$$\Pr\left[\sum_{i=1}^{n} X_i \geq k\mu + z \cdot \sqrt{k} \cdot a\right] \leq \exp\left(-\frac{z^2}{2}\right)$$

לקראת ההוכחה של משפט קומפוזיציה עבור DP-יציבות, נצטרך גרסה קצת יותר כללית של המשפט הנ"ל.

<u>משפט 10 (אי שוויון אזומה-הופדינג):</u>
יהיו $a,\mu \in \mathbb{R}$ ויהיו $X_1, X_2, \dots, X_k$ משתנים מקריים כך שלכל $i \in [k]$ מתקיים $\Pr[|X_i| \le a] = 1$ וגם לכל בחירה של $(x_1, \dots, x_{i-1}) \in \text{Support}(X_1, \dots, X_{i-1})$ מתקיים
$$\mathbb{E}[X_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}] \le \mu.$$
אזי, לכל $z > 0$ מתקיים
$$\Pr\left[\sum_{i=1}^{k} X_i \ge k\mu + z \cdot \sqrt{k} \cdot a\right] \le \exp\left(-\frac{z^2}{2}\right)$$

נשים לב שמשפט 9 הוא מקרה פרטי של משפט 10. למה? כי אם $X_1, X_2, \dots, X_k$ הם בלתי תלויים וגם $\mathbb{E}[X_i] \le \mu$ אזי
$$\mathbb{E}[X_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}] = \mathbb{E}[X_i] \le \mu$$

כדי להוכיח את המשפט הזה, נצטרך להיעזר בעובדה הבאה:

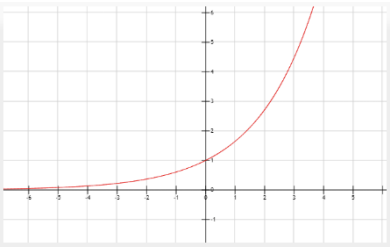
<u>משפט 11 (למת הופדינג):</u>
יהי X משתנה מקרי ממשי המקיים $\Pr[a \le X \le b] = 1$ וגם $\mathbb{E}[X] \le \mu$. אזי, לכל $\lambda \in \mathbb{R}$ מתקיים
$$\mathbb{E}\left[e^{\lambda X}\right] \le \exp\left(\lambda\mu + \frac{\lambda^2 (b-a)^2}{8}\right)$$

<u>רעיון ההוכחה:</u>
ראשית נשים לב כי מספיק להוכיח את המשפט עבור המקרה בו $\mathbb{E}[X] = 0$.

מדוע? נניח כי הוכחנו את המשפט עבור המקרה בו $\mathbb{E}[X] = 0$. כעת יהי X משתנה מקרי עם תוחלת שונה מאפס. נגדיר משתנה מקרי חדש $\tilde{X} = X - \mathbb{E}[X]$ ונשים לב כי $\mathbb{E}[\tilde{X}] = 0$. לכן:
$$\mathbb{E}\left[e^{\lambda X}\right] = \mathbb{E}\left[e^{\lambda(\tilde{X} + \mathbb{E}[X])}\right] = e^{\lambda \cdot \mathbb{E}[X]} \cdot \mathbb{E}\left[e^{\lambda \tilde{X}}\right] \le e^{\lambda \cdot \mu} \cdot \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

אז נותר להוכיח את המשפט בהנחה שמתקיים $\mathbb{E}[X] = 0$.

נתבונן בפונקציה $f(y) = e^{\lambda y}$, ונשים לב שזוהי פונקציה קמורה.



תזכורת: לפי הגדרה, פונקציה $f: \mathbb{R} \to \mathbb{R}$ היא קמורה אם לכל $y_1, y_2 \in \mathbb{R}$ ולכל $t \in [0,1]$ מתקיים
$$f\big((1-t) \cdot y_1 + t \cdot y_2\big) \le (1-t) \cdot f(y_1) + t \cdot f(y_2)$$

לכן, במקרה שלנו, עבור $y_1 = a,\ y_2 = b$ ועבור $t = \frac{y-a}{b-a}$ כאשר $y \in [a,b]$ מתקיים
$$e^{\lambda y} = f(y) = f\big((1-t)a + tb\big) \le (1-t) \cdot f(a) + t \cdot f(b) = \frac{b-y}{b-a} \cdot e^{\lambda a} + \frac{y-a}{b-a} \cdot e^{\lambda b}$$
וזה נכון נקודתית לכל $y \in [a,b]$.

כעת נביט במשתנה המקרי שלנו $X \in [a,b]$. מתקיים:

$$\mathbb{E}\left[e^{\lambda X}\right] \le \mathbb{E}\left[\frac{b-X}{b-a}\cdot e^{\lambda a} + \frac{X-a}{b-a}\cdot e^{\lambda b}\right] =$$

$$= \frac{b-\mathbb{E}[X]}{b-a}\cdot e^{\lambda a} + \frac{\mathbb{E}[X]-a}{b-a}\cdot e^{\lambda b} \underbrace{=}_{\substack{\text{נחנו כי} \\ \mathbb{E}[X]=0}} \frac{b}{b-a}\cdot e^{\lambda a} + \frac{-a}{b-a}\cdot e^{\lambda b}$$

אז חסמנו את $\mathbb{E}\left[e^{\lambda X}\right]$ בעזרת איזשהו ביטוי שלא תלוי ב- X (תלוי רק ב- a,b,λ).

המשך ההוכחה מתבצעת על ידי חקירת הביטוי הזה שקיבלנו. ניתן להראות כי ביטוי זה חסום על ידי

$$\exp\left(\frac{\lambda^2(b-a)^2}{8}\right)$$

מ.ש.ל. (משפט 11)

בנוסף, נצטרך להיזכר במשפט התוחלת השלמה, לפיו התוחלת של התוחלת המותנית של משתנה מקרי כלשהו שווה לתוחלת של אותו משתנה מקרי.

<u>משפט 12 (משפט התוחלת השלמה):</u>

יהיו X,Y משתנים מקריים כך ש- $\mathbb{E}[|X|] < \infty$. אזי מתקיים $\mathbb{E}[X] = \mathbb{E}\left[\ \mathbb{E}[X|Y]\ \right]$

לפני שניזכר בהוכחה של משפט 12, ננסה להבין מה כתוב כאן בכלל, כלומר מה זה $\mathbb{E}[\ \mathbb{E}[X|Y]\ ]$.

קודם נזכר מה זה תוחלת מותנית: עבור $y \in \text{Support}(Y)$ מתקיים (במקרה הדיסקרטי)

$$\mathbb{E}[X|Y=y] = \sum_{x\in\text{Supp}(X)} x\cdot \Pr[X=x|Y=y] = \sum_{x\in\text{Supp}(X)} x\cdot\frac{\Pr[X=x,\ \ Y=y]}{\Pr[Y=y]}$$

כעת, התוחלת המותנית של X ביחס ל Y היא פונקציה של הערך של Y, המוגדרת באופן הבא:

$$\mathbb{E}[X|Y](y) = \mathbb{E}[X|Y=y]$$

כלומר $\mathbb{E}[X|Y]$ הוא משתנה מקרי (פונקציה של Y).

אז משפט 12 אומר שהתוחלת של $\mathbb{E}[X|Y]$ שווה לתוחלת של X.

אולי דרך ברורה יותר לנסח את המשפט הזה הייתה $\mathbb{E}[X] = \mathbb{E}_{y\leftarrow Y}\left[\ \mathbb{E}[X|Y=y]\ \right]$ אבל כמו שזה כתוב למעלה זאת הצורה הסטנדרטית...

<u>הוכחת משפט 12 (עבור המקרה הדיסקרטי והסופי):</u>

$$\mathbb{E}\Big[\ \mathbb{E}[X|Y]\ \Big] = \sum_y \mathbb{E}[X|Y=y]\cdot \Pr[Y=y] = \sum_y \left[\sum_x x\cdot \Pr[X=x|Y=y]\right]\cdot \Pr[Y=y]$$

$$= \sum_y\sum_x x\cdot \Pr[X=x,\ \ Y=y] = \sum_x\sum_y x\cdot \Pr[X=x,\ \ Y=y]$$

$$= \sum_x x\sum_y \Pr[X=x,\ \ Y=y] = \sum_x x\cdot \Pr[X=x] = \mathbb{E}[X]$$

מ.ש.ל. (משפט 12)

<u>תרגיל כיתה:</u> נתון שק אשר בהתחלה מכיל R כדורים אדומים ו- B כדורים כחולים.
בכל שלב, נוציא כדור מהשק באקראי (בהתפ' אחידה ובאופן ב"ת) ונחזיר לשק 2 כדורים בצבע שהוצאנו.
כלומר אחרי k שלבים השק מכיל $R+B+k$ כדורים.

נגדיר משתנים מקריים $X_i = 1$ אם בשלב ה i הוצאנו כדור אדום. אחרת $X_i = 0$.
הוכיחו כי

$$\mathbb{E}\left[\sum_{i=1}^{k} X_i\right] = k\cdot\frac{R}{R+B}$$

פתרון:

נוכיח באינדוקציה כי לכל i מתקיים $\mathbb{E}[X_i] = \frac{R}{R+B}$ ואז הטענה נובעת מלינאריות התוחלת.
מקרה הבסיס עבור $i=1$ טריוויאלי. נניח שזה נכון לכל $\ell \le i$. נסמן $S_\ell = \sum_{i=1}^{\ell} X_i$. אזי

$$\mathbb{E}[X_{\ell+1}] = \mathbb{E}\big[\mathbb{E}[X_{\ell+1}|S_\ell]\big] = \mathbb{E}\left[\frac{R+S_\ell}{R+B+\ell}\right] = \frac{R+\mathbb{E}[S_\ell]}{R+B+\ell} \underbrace{=}_{\substack{\text{הנחת}\\ \text{האינדוקציה}}} \frac{R+\ell\cdot\frac{R}{R+B}}{R+B+\ell} = \frac{R}{R+B}$$

עכשיו אנחנו יכולים להוכיח את משפט 10 (אי שוויון אזומה-הופדינג). כמו שנראה, ההוכחה תהייה דומה להוכחת חסם צ'רנוף שראינו בהרצאה הראשונה.

<u>הוכחה של משפט 10:</u>

יהיו $a,\mu \in \mathbb{R}$ ויהיו $X_1, X_2, \dots, X_k$ משתנים מקריים כך שלכל $i \in [k]$ מתקיים $\Pr[|X_i| \le a] = 1$ וגם לכל בחירה של $(x_1,\dots,x_{i-1}) \in \text{Support}(X_1,\dots,X_{i-1})$ מתקיים

$$\mathbb{E}[X_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}] \le \mu.$$

יהי $z > 0$. נסמן $t = k\mu + z\cdot\sqrt{k}\cdot a$. עלינו להראות שמתקיים

$$\Pr\left[\sum_{i=1}^{k} X_i \ge t\right] \le \exp\left(-\frac{z^2}{2}\right)$$

נסמן $c = \frac{z}{a\sqrt{k}}$ ונחשב:

$$\Pr\left[\sum_{i=1}^{k} X_i \geq t\right] = \Pr\left[c \cdot \sum_{i=1}^{k} X_i \geq c \cdot t\right] = \Pr\left[e^{c\cdot\sum_{i=1}^{k} X_i} \geq e^{c\cdot t}\right] \underbrace{\leq}_{\text{אי שוויון מרקוב}} e^{-ct} \cdot \mathbb{E}\left[e^{c\cdot\sum_{i=1}^{k} X_i}\right] = ((1))$$

בהוכחת חסם צ'רנוף השתמשנו בהנחה שהמשתנים המקריים הם בלתי תלויים כדי לפצל את התוחלת הנ"ל למכפלת התוחלות, ואז ניתחנו כל תוחלת בודדת. כאן לא נוכל לפצל למכפלת תוחלות כי אין לנו אי-תלות. במה כן נוכל להשתמש? משפט התוחלת השלמה:

$$((1)) = e^{-ct} \cdot \mathbb{E}\left[\mathbb{E}\left[e^{c\cdot\sum_{i=1}^{k} X_i}\middle| X_1, \dots, X_{k-1}\right]\right] = ((2))$$

כאן התוחלת החיצונית היא על פני הגרלת $X_1, \dots, X_{k-1}$ והתוחלת הפנימית היא על פני הגרלת X_k (מהמרחב המותנה המתאים). נשים לב שבתוך התוחלת הפנימית $X_1, \dots, X_{k-1}$ הם קבועים (כי אנחנו מתנים עליהם) ולכן הם יוצאים החוצה מהתוחלת הפנימית. נקבל

$$((2)) = e^{-ct} \cdot \mathbb{E}\left[e^{c\cdot\sum_{i=1}^{k-1} X_i} \cdot \mathbb{E}[e^{c\cdot X_k}|X_1, \dots, X_{k-1}]\right] = ((3))$$

כעת נזכור כי לפי ההנחות שלנו, לכל קביעה של $X_1, \dots, X_{k-1}$ מתקיים ש- $\mathbb{E}[X_i|X_1 = x_1, \dots, X_{i-1} = x_{i-1}] \leq \mu$.
לכן נוכל להפעיל את למת הופדינג (משפט 11) על התוחלת הפנימית ולקבל

$$((3)) \leq e^{-ct} \cdot \mathbb{E}\left[e^{c\cdot\sum_{i=1}^{k-1} X_i} \cdot \exp\left(c \cdot \mu + \frac{c^2a^2}{2}\right)\right] = e^{-ct} \cdot e^{c\mu} \cdot e^{c^2a^2/2} \cdot \mathbb{E}\left[e^{c\cdot\sum_{i=1}^{k-1} X_i}\right] = ((4))$$

אז מה קיבלנו? התחלנו מ- $e^{-ct} \cdot \mathbb{E}\left[e^{c\cdot\sum_{i=1}^{k} X_i}\right]$ ונפטרנו מאחד המשתנים המקריים (נפטרנו מ- X_k) ובתמורה צברנו פקטור כפלי $e^{c\mu} \cdot e^{c^2a^2/2}$.

באינדוקציה,

$$((4)) \leq e^{-ct} \cdot e^{kc\mu} \cdot e^{kc^2a^2/2} = ((5))$$

נציב את הבחירה שלנו של $t = k\mu + z \cdot \sqrt{k} \cdot a$ ונקבל

$$((5)) = e^{-ck\mu - cz\sqrt{k}a} \cdot e^{kc\mu} \cdot e^{kc^2a^2/2} = e^{-cz\sqrt{k}a} \cdot e^{kc^2a^2/2} = ((6))$$

נציב את הבחירה שלנו של $c = \frac{z}{a\sqrt{k}}$ ונקבל

$$((6)) = e^{-z^2} \cdot e^{z^2/2} = e^{-z^2/2}$$

מ.ש.ל. (משפט 10)

<u>דוגמה פשוטה לשימוש באי שוויון אזומה-הופדינג (לא מדויק)</u>

נחזור לתרגיל הכיתה עם הכדורים.

הנחה מפשטת: מספר הכדורים ההתחלתי $N = R + B$ הוא הרבה יותר גדול ממספר השלבים בתהליך k (מספר הפעמים שאנחנו מוציאים כדור אחד ומחזירים שניים). בנוסף נניח $R = B = \frac{N}{2}$.

נשים לב שהמשתנים המקריים שהגדרנו בתרגיל $X_1, \dots, X_k$ הם <u>לא בלתי תלויים</u>.

עדיין נוכל להפעיל את אי-שוויון אזומה:

- המשתנים שלנו חסומים בקטע *[0,1]*
- לכל קביעה של $X_1, \dots, X_i$ מתקיים

$$\mathbb{E}[X_{i+1} | X_1 = x_1, \dots, X_i = x_i] \leq \frac{R+i}{N+i} \leq \frac{R}{N} + \frac{k}{N} \approx \frac{1}{2}$$

נקבל

$$\Pr\left[\sum_{i=1}^{k} X_i \gtrsim \frac{k}{2} + z \cdot \sqrt{k}\right] \leq \exp\left(-\frac{z^2}{2}\right)$$

כלומר בהסתברות גבוהה $\sum_{i=1}^{k} X_i$ לא יהיה הרבה יותר גדול מהתוחלת שלו, שהיא $\frac{k}{2}$ לפי תרגיל הכיתה (ולפי הבחירה $R = B = \frac{N}{2}$).

הערה: זה לא הדוק עבור הדוגמה הזאת. אפשר לקבל תוצאה טובה יותר בעזרת גרסה קצת אחרת של אזומה.