הסתברות 1 - תרגול 9
חסמיי צ'רנוף והופדינג

12 בדצמבר 2018

בשבועות האחרונים למדתם איך להעריך הסתברויות מסוימות באמצעות א"ש מרקוב וצ'בישב. אך גם ראינו שהא"ש הללו במקרים רבים אינם מספקים. בתרגול זה נעסוק בהערכות נוספות שלמדתם בשבוע האחרון.

א"ש צ'רנוף

הרעיון הוא מאוד פשוט. כפי שהצבנו בהוכחת א"ש צ'בישב פונקציה ריבועית בא"ש מרקוב נציב כעת פונקציה אקספוננציאלית. כלומר, כל מ"מ X ו $t > 0$

$$P(X \geq a) = P(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}}$$

נגדיר אם כן את **הפונקציה יוצרת מומנטים**

$$M_X(t) := E(e^{tX})$$

לכל t עבורו תוחלת זו מוגדרת היטב. קיבלנו אם כן, את א"ש צרנוף:

טענה 0.1 יהי X מ"מ בעל מומנט מעריכי ($M_X(t)$ מוגדר), $a \in \mathbb{R}$ ו t עבורו $M_X(t)$ מוגדר. אזי

$$P(X \geq a) \leq M_X(t)e^{-ta}$$

הערה: מתקיים (וזה שימושי מאוד ברמה החישובית) כי אם X_1, X_2 מ"מ ב"ת בעלי מומנט מעריכי אזי

$$M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t)$$

דוגמא:

טענה 0.2 יהי $X = \sum_{i=1}^{n} X_i$ כאשר $X_i \sim Ber(p_i)$ ב"ת. נסמן $\mu = E(X) = \sum p_i$.
אזי

1.

$$P(X \geq (1+\delta)\mu) \leq \exp(-\mu\frac{\delta^2}{2+\delta}), \quad \forall \delta > 0$$

2.

$$P(X \geq (1-\delta)\mu) \leq \exp(-\mu\frac{\delta^2}{2}), \quad \forall 0 < \delta < 1$$

3.

$$P(|X-\mu| \geq \delta\mu) \leq 2\exp(-\mu\frac{\delta^2}{3})$$

הוכחה (סעיף 1)

ראשית נעריך את $M_{X_i}(t)$:

$$M_{X_i}(t) = E(e^{tX_i}) = p_i e^t + (1-p_i)e^0 = 1 + p(e^t - 1)$$

כעת, מאחר ולכל $y \in \mathbb{R}$, $1 + y \leq e^y$ נקבל ע"י הצבה של $y = p(e^t - 1)$ ש

$$M_{X_i}(t) \leq e^{p(e^t - 1)}$$

כעת,

$$M_X(t) = \Pi_{i=1}^{n} M_{X_i}(t) \leq \Pi_{i=1}^{n} e^{p_i(e^t-1)} = \exp\left((e^t - 1)\sum_{i=1}^{n} p_i\right) = e^{(e^t-1)\mu}$$

כעת נפעיל את א"ש צ'רנוף עבור $a = (1+\delta)\mu$ ונקבל

$$\begin{aligned} P(X \geq (1+\delta)\mu) &\leq M_X(t)e^{-t(1+\delta)\mu} \leq e^{-t(1+\delta)\mu}e^{(e^t-1)\mu} \\ &= \exp(-t(1+\delta)\mu + (e^t - 1)\mu) \end{aligned}$$

כעת נחפש t שימזער את החסם שמצאנו. לשם כך נזכור ש $\exp$ מונ' עולה ולכן מספיק למזער את הארגומנט

$$\phi(t) = -t(1+\delta)\mu + (e^t - 1)\mu$$

נגזור

$$\phi'(t) = -(1+\delta)\mu + e^t\mu$$

ולכן

$$\phi'(t) = 0 \iff e^t = (1+\delta) \iff t = \ln(1+\delta)$$

בדיקה פשוטה תראה שזו אכן נקודת מינימום. נציב ונקבל

$$P(X \geq (1+\delta)\mu) \leq \exp(-\ln(1+\delta)(1+\delta)\mu + \delta\mu)$$

כעת ניעזר בכך שלכל $x > 0$ (תרגיל)

$$\ln(1+x) \geq \frac{x}{1+x/2}$$

ונקבל ש

$$-\ln(1+\delta)(1+\delta)\mu + \delta\mu \leq \mu\delta - \frac{\delta}{1+\delta/2}(1+\delta)\mu = \mu\frac{-\delta^2/2}{1+\delta/2} = -\frac{\delta^2}{2+\delta}\mu$$

ולכן

$$P(X \geq (1+\delta)\mu) \leq \exp(-\frac{\delta^2}{2+\delta}\mu)$$

כנדרש!
הוכחת החלק השני דומה ומושארת כתרגיל בית. סעיף (3) הוא מסקנה ישירה מהסעיפים הקודמים.

א"ש הופדינג

הבעיה עם א"ש צ'רנוף היא שהוא דורש לחשב כל פעם מחדש את הפונקציה יוצרת המומנטים של המ"מ של הבעיה. א"ש הופדינג חוסך את החישוב הזה אך בתמורה, מציב כמה תנאים נוספים.

טענה (אי שוויון הופדינג): יהיו $X_1, ..., X_n$ מ"מ ב"ת, בעלי תוחלת אפס המקיימים $|X_i| \leq 1$. אזי לכל $a \geq 0$ מתקיים

$$P(\sum_{i=1}^{n} X_i \geq a) \leq \exp(\frac{-a^2}{2n})$$

ע"פ רוב, המ"מ שנעסוק בהם לא יתאימו בדיוק לתנאי המשפט. כלומר ייתכן בהחלט ש $E(X_i) \neq 0$ או $|X_i| > 1$. גרוע מזה, לעיתים X_i לא יהיו ב"ת. כדי להשתמש בא"ש זה נצטרך במקרה הראשון לעשות חילוף משתנה כפי שתראו בדוגמא הבאה. לגבי המקרה השני נדבר בהמשך.

דוגמא (הילוך שיכור):

שיכור עומד על ציר המספרים השלמים. הוא מטיל מטבע שנופל על ראש בהסתברות $p > \frac{1}{2}$. אם יצא לו ראש הוא הולך צעד ימינה (ואז נוחת על המספר 1), ואם יצא לו פלי הוא הולך צעד אחד שמאלה (ואז נוחת על המספר -1). הוא מבצע n הטלות מטבע כנ"ל באופן ב"ת. יהא X_n מיקומו אחרי n הצעדים הנ"ל. ניעזר במשפט הופדינג על מנת לחסום את $P(X_n = 0)$.

נכתוב, $X_n = \sum_{i=1}^{n} Y_i$ כאשר $P(Y_i = 1) = p$ וכן $P(Y_i = -1) = 1 - p = q$. אז

$$\mathbb{E}(Y_i) = p + (-1)(1-p) = p - q.$$

נגדיר כעת מ"מ חדשים, לכל $1 \leq i \leq n$, ע"י

$$Z_i = \frac{(p-q) - Y_i}{1+p-q}$$

וכעת

$$\mathbb{E}(Z_i) = 0$$

מלינאריות התוחלת, נוסף, מאחר ו $Y_i \in \{-1, 1\}$ נסיק כי $|Z_i| \leq 1$. אכן,

$$-1 \leq \frac{(p-q)-1}{1+p-q} \leq \frac{(p-q)-Y_i}{1+p-q} \leq \frac{(p-q)+1}{1+p-q} \leq 1$$

נגדיר $S_n = \sum_{i=1}^{n} Z_i$. המ"מ S_n עומד בתנאי א"ש הופדינג. נמצא את הקשר בין S_n ל X_n.

$$S_n = \sum_{i=1}^{n} Z_i = \frac{1}{1+p-q} \sum_{i=1}^{n} ((p-q) - Y_i) = \frac{1}{1+p-q}(n(p-q) - X_n)$$

ולכן

$$X_n \leq 0 \iff S_n \geq \frac{n(p-q)}{1+p-q} > 0$$

נקבל אם כן,

$$P(X_n = 0) \leq P(X_n \leq 0) = P(S_n \geq \frac{n(p-q)}{1+p-q}) \leq e^{\frac{-(\frac{n(p-q)}{1+p-q})^2}{2n}} = e^{-n\frac{(2p-1)^2}{8p^2}}$$

כנדרש.

הערה על תרגיל בית הקרוב:
טרם עסקנו בשאלה כיצד להפעיל את א"ש הופדינג אם המ"מ X_i אינם ב"ת. במקרים רבים למרות שאין אי־תלות מלאה. ניתן לפצל את המ"מ X_i למספר קבוצות כך שבכל קבוצה המ"מ הם ב"ת.

למשלפן בתרגיל הבית תתבקשו לחסום את ההסתברות לקבלת מספר רצפים של HH בסדרה של n הטלות מטבע. את המ"מ הנ"ל ניתן להציג כפי שראינו כסכום של מ"מ ברנולי המתאימים לרצף המתחיל במקום ה i כלומר $X = \sum_{i=1}^{n-1} X_i$ (שימו לב שגם כאן צריך לעשות חילוף משתנה) אבל X_i, X_{i+1} כמובן אינם ב"ת. הפתרון הוא להסתכל בנפרד על הרצפים המתחילים במקום זוגי או אי־זוגי. שימו לב שזה לא מסיים את השאלה. עדיין נותר לקשור בין הסתברויות הקשורות לסכומים הזוגיים והאי זוגיים להסתברות אותה אנו רוצים להעריך.

$$S_n = \boxed{\frac{1}{1+p-q}(n(p-q) - X_n)}$$

אם $X_n < 0$ אז

$$\boxed{\frac{n(p-q)}{1+p-q}} < \boxed{\frac{1}{1+p-q}(n(p-q) - X_n)}$$