

## מבוא להסתברות

© ארזים

22 במאי 2016

ראינו בשיעור שעבר את משפט הגבול המרכזי. ניזכר בו. ראשית, נגדיר את התקנון של משתנה מקרי  $Y$  בעל שונות סופית להיות

$$\hat{Y} = \frac{Y - \mathbb{E}(Y)}{\sqrt{\text{Var}(Y)}}$$

נקבל כי תמיד

$$\begin{aligned}\mathbb{E}(\hat{Y}) &= 0 \\ \text{Var}(\hat{Y}) &= 1\end{aligned}$$

**משפט 0.1** (משפט הגבול המרכזי) יהיו  $X_1, X_2, \dots$  משתנים מקריים בלתי תלויים שווי התפלגות ובעלי שונות סופית. נגדיר  $S_n = X_1 + \dots + X_n$ . לכן

$$\hat{S}_n = \frac{S_n - n \cdot \mathbb{E}(X_1)}{\sqrt{n \cdot \text{Var}(X_1)}}$$

כעת, לכל  $-\infty \leq a < b \leq \infty$ , מתקיים

$$\mathbb{P}(a \leq \hat{S}_n \leq b) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

משמעות ההסתברות הזו היא ההסתברות שהמשתנה  $S_n$  סוטה מן התוחלת שלו בין  $a$  לבין  $b$  סטיות תקן.

**דוגמא** מטילים קוביה 12,000 פעם ומסמנים בתור  $S$  את כמות הפעמים שקיבלנו 6. מה ההסתברות, בקירוב, שנקבל כמות 6 בין 1800 לבין 2100?

**פתרון** נשים לב כי

$$\begin{aligned}S &\sim \text{Bin}\left(12000, \frac{1}{6}\right) \\ \mathbb{E}(S) &= 2000 \\ \text{Var}(S) &= 12000 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{10000}{6} \\ \text{std}(S) &= \sqrt{\text{Var}(S)} = 40.82 \dots\end{aligned}$$

לפי אי-שוויון צ'בישב:

$$\begin{aligned}\mathbb{P}(1800 \leq S \leq 2100) &= 1 - \mathbb{P}(S > 2100) - \mathbb{P}(S < 1800) \\ \mathbb{P}(S > 2100) &\leq \mathbb{P}(|S - 2000| > 100) \leq \frac{\text{Var}(S)}{100^2} = \frac{1}{6}\end{aligned}$$

לעומת זאת, לפי משפט הגבול המרכזי,

$$\mathbb{P}(1800 \leq S \leq 2100) = \mathbb{P}(-4.9 \leq \hat{S} \leq 2.45) \approx \frac{1}{\sqrt{2\pi}} \int_{-4.9}^{2.45} e^{-\frac{x^2}{2}} dx \approx 0.9929$$

### מקרה פרטי - משפט זה-מואבר-לפלס

זהו מקרה פרטי של משפט הגבול המרכזי, כאשר  $X_i \sim \text{Ber}(p)$ , ולכן  $S_n \sim \text{Bin}(n, p)$ .  
נדבר על ההוכחה:

$$\mathbb{P}(S_n = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

ובמקרה זה

$$\hat{S}_n = \frac{S_n - np}{\sqrt{np(1-p)}}$$

סימונים: עבור  $k$  שלם, נגדיר

$$\hat{k} = \frac{k - np}{\sqrt{np(1-p)}}$$

כך שמתקיים

$$\mathbb{P}(S_n = k) = \mathbb{P}(\hat{S}_n = \hat{k})$$

כמו כן, נגדיר

$$\begin{aligned}I_n &= \{k \in \mathbb{Z} \mid a \leq \hat{k} \leq b\} = \\ &= \{k \in \mathbb{Z} \mid np + a\sqrt{np(1-p)} \leq k \leq np + b\sqrt{np(1-p)}\}\end{aligned}$$

אנו רוצים להוכיח כי

$$\begin{aligned} \mathbb{P}(a \leq \hat{S}_n \leq b) &= \mathbb{P}(S_n \in I_n) = \sum_{k \in I_n} \mathbb{P}(S_n = k) = \\ &= \sum_{k \in I_n} \binom{n}{k} \cdot p^k (1-p)^{n-k} \xrightarrow{n \rightarrow \infty} \int_a^b \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} dx \end{aligned}$$

משפט עזר: עבור  $k \in I_n$  (במידה שווה)

$$\mathbb{P}(S_n = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \approx \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{k^2}{2}}$$

קירוב סטירלינג:

$$\frac{m!}{\sqrt{2\pi m} \left(\frac{m}{e}\right)^m} \xrightarrow{m \rightarrow \infty} 1$$

כעת,

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} \approx \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k \sqrt{2\pi(n-k)} \left(\frac{n-k}{e}\right)^{n-k}} = \sqrt{\frac{n}{2\pi k(n-k)}} \frac{n^n}{k^k (n-k)^{n-k}} \approx \\ &\approx \sqrt{\frac{1}{2\pi np(1-p)}} \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k} = \sqrt{\frac{1}{2\pi np(1-p)}} e^{k \log\left(\frac{n}{k}\right)} e^{(n-k) \log\left(\frac{n}{n-k}\right)} \end{aligned}$$

רעיון המשך ההוכחה: נשתמש בקירוב טיילור של  $\log$  כדי להעריך את מה שקיבלנו, ולבסוף נשתמש בקירוב רימן לאינטגרל כדי להראות את השאיפה שרצינו.