מבוא להסתברות

© ארזים

29 במאי 2016

1 משפט הגבול המרכזי

משפט 1.1 יהיו $X_1, \ldots, X_n$ משתנים מקריים בלתי תלויים שווי התפלגות עם תוחלת סופית μ ושונות סופית σ^2. אזי

$$\mathbb{P}\left(a \leq \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} x_i - n\mu\right) \leq b\right) \xrightarrow[n\to\infty]{} \phi\left(b\right) - \phi\left(a\right)$$

כאשר

$$\phi\left(t\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} e^{-\frac{r^2}{2}}\,\mathrm{d}r$$

אם לוקחים משתנה מקרי Y, מחסרים ממנו את התוחלת ומחלקים את התוצאה בסטיית התקן, מקבלים משתנה מקרי $\tilde{Y}$ עם תוחלת 0 ושונות 1.

כלל אצבע

עבור $n \geq 20$ לערך,

$$\mathbb{P}\left(a \leq \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} x_i - n\mu\right) \leq b\right) \cong \phi\left(b\right) - \phi\left(a\right)$$

את הערכים של ϕ מוציאים מטבלת ההתפלגות הנורמלית.
ניתן לחלק את הביטוי בתוך פונקציית ההסתברות בערך n ולכפול בו, ולקבל

$$\mathbb{P}\left(a \leq \frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}\sum_{i=1}^{n} x_i - \mu\right) \leq b\right) \cong \phi\left(b\right) - \phi\left(a\right)$$

1.1 תכונות של ϕ

.1

$$\lim_{t\to\infty}\phi\left(t\right)=1$$

.2

$$\lim_{t\to-\infty}\phi\left(t\right)=0$$

.3

$$\phi\left(-t\right)=1-\phi\left(t\right)$$

תרגיל יוסי ודנה מטילים כל אחד מטבע הוגן n פעמים באון בלתי תלוי. S - מספר העצים של יוסף. T - מספר העצים של דנה. חשבו בקירוב את $\mathbb{P}\left(S-T\right)\geq\sqrt{\frac{n}{2}}$.

פתרון נגדיר משתנים מקריי חדשים. לכל $1\leq i\leq n$, נגדיר S_i להיות האינדיקטור של עץ בהטלה i של יוסי, T_i באותו אופן עבור דנה. כעת

$$S=\sum S_i,T=\sum T_i,S-T=\sum_{i=1}^{n}\left(S_i-T_i\right)$$

כמו כן,

$$\begin{aligned}
\mathbb{E}\left(S-T\right) &= \sum_{i=1}^{n}\mathbb{E}\left(S_i-T_i\right)=0\\
\mathrm{Var}\left(S-T\right) &= \sum_{i=1}^{n}\mathrm{Var}\left(S_i-T_i\right)=\sum_{i=1}^{n}\mathrm{Var}\left(S_i\right)+\mathrm{Var}\left(T_i\right)=2n\cdot\mathrm{Var}\left(S_1\right)=\frac{n}{2}
\end{aligned}$$

נשתמש בכל אלה לחישוב ההסתברות (Z הוא סימון למשתנה שמתפלג נורמלית):

$$\begin{aligned}
\mathbb{P}\left(S-T\geq\sqrt{\frac{n}{2}}\right) &= \mathbb{P}\left(\frac{\sum\left(S_i-T_i\right)-n\cdot0}{\sqrt{\frac{n}{2}}}\geq\frac{\sqrt{\frac{n}{2}}-n\cdot0}{\sqrt{\frac{n}{2}}}\right)=\mathbb{P}\left(\frac{\sum\left(S_i-T_i\right)-n\cdot0}{\sqrt{\frac{n}{2}}}\geq1\right)\cong\\
&\cong \mathbb{P}\left(Z\geq1\right)=1-\mathbb{P}\left(Z<1\right)=1-\phi\left(1\right)=\phi\left(-1\right)\cong0.8413
\end{aligned}$$

תרגיל מסדרים 10 כדורים בשורה באקראי - 3 לבנים, 7 שחורים. חוזרים על הניסוי 100 פעמים. נסמן X_i - כמות הכדורים הלבנים במקומות הזוגיים בפעם i. $\overline{X}=\frac{1}{100}\sum X_i$. העריכו בקירוב את $p=\mathbb{P}\left(1.25\leq\overline{X}\leq1.51\right)$.

1. $p < \frac{1}{4}$
2. $\frac{1}{4} \leq p < \frac{1}{2}$
3. $\frac{1}{2} \leq p < \frac{3}{4}$
4. $\frac{3}{4} \leq p$

פתרון נרצה להשתמש במשפט הגבול המרכזי לחישוב p. לשם כך צריך לחשב את $\mathbb{E}\left(X_i\right), \mathrm{Var}\left(X_i\right)$.
נשים לב כי

$$X_i \sim \mathrm{HG}\left(10, 5, 3\right)$$

ולכן

$$\begin{aligned} \mathbb{E}\left(X_i\right) &= \frac{Dn}{N} = \frac{5 \cdot 3}{10} = \frac{3}{2} \\ \mathrm{Var}\left(X_i\right) &= \frac{\frac{Dn}{N}\left(\frac{N-D}{N}\right)\left(N-n\right)}{N-1} = \frac{7}{12} \end{aligned}$$

כעת

$$\begin{aligned} \mathbb{P}\left(1.45 \leq \frac{1}{100}\sum_{i=1}^{n} X_i \leq 1.51\right) &= \mathbb{P}\left(\frac{1.45-1.5}{\frac{\sqrt{\frac{7}{12}}}{\sqrt{100}}} \leq \frac{\frac{1}{100}\sum X_i - 1.5}{\frac{\sqrt{\frac{7}{12}}}{\sqrt{100}}} \leq \frac{1.51-1.5}{\frac{\sqrt{\frac{7}{12}}}{\sqrt{100}}}\right) = \\ &\cong \mathbb{P}\left(\frac{\sqrt{3}}{5\sqrt{7}} \geq Z \geq -\sqrt{\frac{3}{7}}\right) = \phi\left(0.13\right) - \left(1 - \phi\left(0.65\right)\right) \cong 0.29 \end{aligned}$$

2 שרשראות מרקוב

S היא קבוצת מצבים סופית. P היא מטריצת מעברים, כאשר התא x, y, לכל $x, y \in S$, מסמל את ההסתברות לעבור ממצב x למצב y. μ_0 היא התפלגות התחלתית על S. התנאי על P - כל הערכים בה אי שליליים, וכן

$$\sum_{y \in S} P\left(x, y\right) = 1$$

נקודה מרכזית: בהינתן מצב כלשהו, הסיכוי לעבור למצב אחר הוא בלתי תלוי במצבים הקודמים.

2.1 קידום התפלגויות

נסמן X_t את המצב בו נהיה בזמן t. אזי

$$\mathbb{P}\left(X_t = s\right) = \left(\mu \cdot P^t\right)\left(s\right)$$

אם נסמן μ_t להיות ההתפלגות בזמן t נקבל

$$\mu_t = \mu \cdot P^t$$

2.2 התפלגות סטציונרית

זו התפלגות πשהיא ווקטור עצמי של P כך שמתקיים $\pi \cdot P = \pi$.

תרגיל אינסוף סטודנטים הנבחנים מנסים להעביר ביניהם תשובה לשאלת כן ולא. האדם הראשון הוא היחיד שיודע את התשובה. הוא מעביר תשובה כלשהי ליושב לידו, שמעביר אותה הלאה וכן הלאה. אם התשובה היא כן, אז היא עוברת בהסתברות 1 לבא בתור. אם התשובה היא לא אז בהסתברות $\frac{1}{2}$ מועברת כל אחת מהתשובות. מצאו כיצד תראה ההתפלגות הגבולית של התשובות.

פתרון $S = \{Y, N\}$ היא קבוצת המצבים. מטריצת המעברים תהיה

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

נראה באינדוקציה כי

$$P^n = \begin{pmatrix} 1 & 0 \\ 1 - \frac{1}{2^n} & \frac{1}{2^n} \end{pmatrix}$$

הצעד הוא:

$$P^{n+1} = P^n \cdot P = \begin{pmatrix} 1 & 0 \\ 1 - \frac{1}{2^n} & \frac{1}{2^n} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 - \frac{1}{2^{n+1}} & \frac{1}{2^{n+1}} \end{pmatrix}$$

לכן בגבול

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

ולכן ההתפלגות הגבולית היא $(1, 0)$.

תרגיל יש משחק מחשב עם 5 שלבים. בכל נקודת זמן ניתן לעלות שלב בהסתברות α או לרדת בהסתברות $1 - \alpha$ באופן בלתי תלוי. משלב 1 יורדים לשלב 5 ולהיפך. תארו את הבעיה כשרשרת מרקוב ומצאו התפלגות סטציונרית.

פתרון $S = \{1, 2, 3, 4, 5\}$ הם המצבים. מטריצת המעבר היא

$$P = \begin{pmatrix} 0 & \alpha & 0 & 0 & 1-\alpha \\ 1-\alpha & 0 & \alpha & 0 & 0 \\ 0 & 1-\alpha & 0 & \alpha & 0 \\ 0 & 0 & 1-\alpha & 0 & \alpha \\ \alpha & 0 & 0 & 1-\alpha & 0 \end{pmatrix}$$

כעת יש למצוא וקטור $\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5)$ כך שמתקיים $\pi \cdot P = \pi$. החישוב בחוברת, ומוצאים כי

$$\pi = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$$