מכניקה סטטיסטית: תרגול 1 - הסתברות

26.3.15

1 הגדרות

משתנה מקרי הוא משתנה שמקבל ערכים שונים על פי פונקצית הסתברות כלשהי.

משתנה מקרי בדיד מקבל ערכים מתוך קבוצה בדידה, למשל: $\{0,1\}$, $\{0,1,2,3...\}$

משתנה מקרי רציף מקבל ערכים מתוך רצף, למשל: $[0,1]$, $(0,\infty)$

	משתנה בדיד n	משתנה רציף x
ההתפלגות	$P(n)$= ההסתברות למדוד ערך n	$P(x)$= פונקצית צפיפות הסתברות $\mathcal{P}(a \le x \le b) = \int_a^b P(x)dx$ = ההסתברות למדוד ערך בקטע $[a,b]$
נרמול	$\sum_n P(n) = 1$	$\int_{-\infty}^{\infty} P(x)dx = 1$
תוחלת	$\langle n \rangle = \sum_n nP(n)$	$\langle x \rangle = \int_{-\infty}^{\infty} xP(x)dx$
תוחלת של פונקציה f	$\langle f(n) \rangle = \sum_n f(n)P(n)$	$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)P(x)dx$
שונות	$Var(n) = \langle n^2 \rangle - \langle n \rangle^2$	$Var(x) = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$
סטיית תקן	$\sigma_n = \sqrt{Var(n)}$	$\sigma_x = \sqrt{Var(x)}$

1.1 דוגמא: הטלת קוביה

מטילים קוביה הוגנת. n = תוצאת ההטלה. מהי ההתפלגות $P(n)$? מהן התוחלת, השונות וסטיית התקן של n?

ההתפלגות:
$$P(n) = \begin{cases} \frac{1}{6} & n = 1,2,...,6 \\ 0 & otherwise \end{cases}$$

התוחלת:

$$\langle n \rangle = \sum_{n=1}^{6} nP(n) = \frac{1}{6}\sum_{n=1}^{6} n = \frac{21}{6} = 3.5$$

השונות:

$$\langle n^2 \rangle = \sum_{n=1}^{6} n^2P(n) = \frac{1}{6}\sum_{n=1}^{6} n^2 = \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$$

$$Var(n) = \langle n^2 \rangle - \langle n \rangle^2 = \frac{91}{6} - 3.5^2 = \frac{105}{36} = 2.9167$$

סטיית התקן: $\sigma_n = \sqrt{Var(n)} \approx 1.7$

1.2 דוגמא: התפלגות נורמלית (גאוסית)

משתנה מקרי x המתפלג נורמלית הוא משתנה רציף המקבל ערכים בקטע $(-\infty, \infty)$ לפי פונקצית צפיפות ההסתברות:
$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

נראה שההתפלגות מנורמלת, ונחשב את התוחלת והשונות שלה.

נרמול:

$$\int_{-\infty}^{\infty} P(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1$$

(תזכורת: חישוב אינטגרל גאוסי - $\int_{-\infty}^{\infty} e^{-a(x+b)^2} dx = \sqrt{\frac{\pi}{a}}$)

תוחלת:

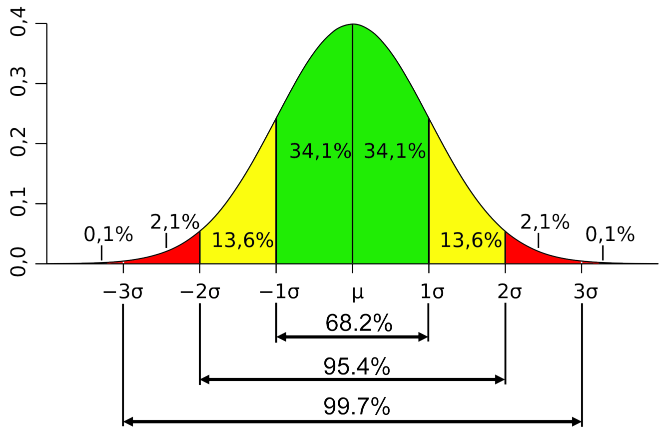
$$\begin{aligned}
\langle x \rangle &= \int_{-\infty}^{\infty} xP(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \stackrel{y=x-\mu}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y+\mu)e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \Big[\underbrace{\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2\sigma^2}} dy}_{=0} + \mu \underbrace{\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy}_{=\sqrt{2\pi\sigma^2}} \Big] = \mu
\end{aligned}$$

שונות:

$$\begin{aligned}
\langle x^2 \rangle &= \int_{-\infty}^{\infty} x^2P(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \stackrel{y=x-\mu}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y+\mu)^2e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \Big[\underbrace{\int_{-\infty}^{\infty} y^2e^{-\frac{y^2}{2\sigma^2}} dy}_{=\frac{1}{2}\sqrt{\pi}(2\sigma^2)^{3/2}} + 2\mu \underbrace{\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2\sigma^2}} dy}_{=0} + \mu^2 \underbrace{\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy}_{=\sqrt{2\pi\sigma^2}} \Big] = \sigma^2 + \mu^2
\end{aligned}$$

כאשר באינטגרל הראשון השתמשנו בנוסחה: $\int\limits_{-\infty}^{\infty} x^{2n}e^{-ax^2}dx = (-1)^n \frac{d^n}{dx^n}\sqrt{\frac{\pi}{a}}$

$$\Rightarrow Var(x) = \langle x^2\rangle - \langle x\rangle^2 = \sigma^2$$



2 משפט הגבול המרכזי

תהי $X_1, X_2, ..., X_N$ סדרת משתנים מקריים בלתי תלויים בעלי אותה התפלגות $P(X_i)$ עם תוחלת $\langle X_i\rangle = \mu$ ושונות $Var(X_i) = \sigma^2$.

נסמן: ממוצע ה־X_i: $\bar{X}_N = \frac{X_1+X_2+...+X_N}{N}$

אז:

$$P(\bar{X}_N \le x) \underset{N\to\infty}{\longrightarrow} \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{(s-\mu)^2}{2\tilde{\sigma}^2}} ds \tag{1}$$

כאשר: $\tilde{\sigma} = \frac{\sigma}{\sqrt{N}}$.

ניתן לגזור לפי x ולקבל:

$$P(\bar{X}_N = x) \underset{N\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{(x-\mu)^2}{2\tilde{\sigma}^2}} \tag{2}$$

משמעות תוצאה זו היא שכאשר מספר המשתנים N גדול, ההתפלגות של הממוצע $\bar{X}_N$ שואפת להתפלגות נורמלית עם ממוצע μ וסטיית תקן $\tilde{\sigma} = \frac{\sigma}{\sqrt{N}}$. שימו לב שככל ש־$N$ יותר גדול, ההתפלגות יותר צרה.

אך כאן צריך להיות מעט זהירים, ולשים לב שכאשר המשתנים X_i, ולכן גם $\bar{X}_N$, בדידים, הביטוי בצד שמאל של (1) הוא סכום ולא אינטגרל ולכן התוצאה (2) אינה מדויקת.

כאשר $\bar{X}_N$ בדיד והמרווח בין הערכים שהוא יכול לקבל קבוע, נסמנו ב־Δx, אז מתקיים:

$$\frac{P(\bar{X}_N=x)}{\Delta x} \underset{N\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}}e^{-\frac{(x-\mu)^2}{2\tilde{\sigma}^2}}$$

3 חלקיקים במיכל

במיכל יש N חלקיקים. נסמן ב־Q את מספר החלקיקים שנמצאים בשליש המיכל השמאלי, ונגדיר גם את $q=\frac{Q}{N}$ ־ יחס החלקיקים בשליש המיכל השמאלי. מתקיים $N,Q\gg 1$.

בכיתה חישבתם במדויק את ההתפלגות $P(q)$ וראיתם שבגבול של N גדול ההתפלגות המתקבלת היא התפלגות נורמלית. עכשיו נגיע לתוצאה זו באמצעות שימוש במשפט הגבול המרכזי.

3.1 חישוב $P(q)$ באמצעות משפט הגבול המרכזי

נגדיר: $X_i=\begin{cases}1, & \text{particle i is in the left third of the box}\\ 0, & \text{otherwise}\end{cases}$

מתקיים: $q=\frac{1}{N}\sum\limits_{i=1}^{N}X_i$

נרצה להשתמש במשפט הגבול המרכזי עבור q. נשים לב ש: $P(X_i)=\begin{cases}\frac{1}{3}, & X_i=1\\ \frac{2}{3}, & X_i=0\end{cases}$

ולכן:

$$\mu=\langle X_i\rangle=\frac{1}{3}\cdot 1+\frac{2}{3}\cdot 0=\frac{1}{3}$$

$$\langle X_i^2\rangle=\frac{1}{3}\cdot 1^2+\frac{2}{3}\cdot 0^2=\frac{1}{3}$$

$$\sigma^2{=}\mathrm{Var}(X_i)=\langle X_i^2\rangle-\langle X_i\rangle^2=\frac{1}{3}-\frac{1}{9}=\frac{2}{9}$$

על פי משפט הגבול המרכזי, $\frac{1}{\sqrt{\frac{4}{9N}\pi}}e^{-\frac{9N(q-\frac{1}{3})^2}{4}} = \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}}e^{-\frac{(q-\mu)^2}{2\tilde{\sigma}^2}} \underset{N\to\infty}{\approx} P(q)$.

3.2 ההסתברות לסטיה מהממוצע

עבור מיכל של ליטר אוויר, מה ההסתברות למדוד בשליש מיכל כמות חלקיקים שרחוקה מהממוצע ביותר מ־1%?

$N_A \approx 6.022 \cdot 10^{23}$ מספר אבוגדרו = מספר האטומים של פחמן־12 הנמצאים ב־12 גרם של חומר זה. מספר זה מאפיין את כמות החלקיקים בסמ"ק של חומר עבור חומרים מוצקים רבים.

מספר החלקיקים בליטר של אוויר: $N \sim N_A \sim 10^{23}$.

יחס החלקיקים הממוצע בשליש מיכל הוא $\mu = \frac{1}{3}$. מדידת סטיה מהממוצע ביותר מ־1% משמעותה מדידת ערך של q שאינו בתחום $[0.33, 0.3367]$.

אחוז מהממוצע הוא $0.01 \cdot \frac{1}{3} = \frac{1}{300} \sim 2 \cdot 10^9 \tilde{\sigma}$

הסיכוי למדוד משתנה המתפלג נורמלית במרחק של יותר ממיליארד סטיות תקן מהממוצע שלו קטן מ־$10^{-10^{17}}$ (!).

$$P(|q - \frac{1}{3}| > 0.01 \cdot \frac{1}{3}) = P(|q - \mu| > 10^9 \sigma) < 10^{-10^{17}}$$