Подготовка за контролно 2 по СЕМ

Задача 1. Правилен зар се хвърля 6 милиона пъти. Каква е вероятността да се паднат повече от 1 млн. шестици?

Решение: (Със сив цвят са маркирани подробните разсъждения)

Нека $X_i = \{$пада се 6-ца на i-тото хвърляне на зара$\}$, $i = 1, 2, \ldots, n$.

$X_i \in \mathrm{Ber}\left(p = \dfrac{1}{6}\right)$ и $\underbrace{\perp\!\!\!\perp_{i=1}^{n} X_i}_{\text{независими}}$.

$\mathbb{E}X_1 = 1 \times p + 0 \times (1-p) = p;\qquad \mathbb{E}X_1^2 = 1^2 \times p + 0^2 \times (1-p) = p;$

$DX = \mathbb{E}X_1^2 - (\mathbb{E}X_1^2) = p - p^2 = p(1-p).$

Нека $S_n = X_1 + X_2 + \ldots + X_n \Rightarrow S_n \in \mathrm{Bin}\left(n, p = \dfrac{1}{6}\right)$.

$$\mathbb{E}S_n = \mathbb{E}\sum_{i-1}^{n} X_i \overset{\substack{\mathbb{E}\text{ е линеен}\\ \text{функционал}}}{=} \sum_{i=1}^{n} \mathbb{E}X_i = \sum_{i=1}^{n} \mathbb{E}X_1 = n \times \mathbb{E}X_1 = np = \mu;$$

$$DS_n = D\sum_{i=1}^{n} \overset{\text{независимост}}{=} \sum_{i=1}^{n} DX_1 = n \times DX_1 = np(1-p) = \sigma^2.$$

S_n е случайна величина, която е сума от еднакво разпределени случайни величини (тук не се интересуваме от това какви точно са разпределенията, които се сумират за да образуват S_n, а само от това, че са еднакви и имат добре дефинирано средно)

$\overset{\text{ЗГЧ}}{\Rightarrow} \lim\limits_{n\to\infty} \dfrac{S_n}{n} = \mu.$ От друга страна, ЦГТ ни дава информацията относно това как точно схожда тази редица $R_n = \dfrac{S_n}{n}$ към средното (с какъв порядък/ колко бързо/ каква е грешката (теорема на Берн-Есеен)).

$\overset{\text{ЦГТ}}{\Rightarrow} \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1)$. Тоест за достатъчно големи n е напълно резонно да направим приближението $S_n \sim \sqrt{n\sigma^2} \times \mathcal{N}(0,1) + n\mu = \mathcal{N}(n\mu, n\sigma^2)$.

Следователно за $\underbrace{n \gg 30}_{\text{Rule of thumb}}$: $\mathbb{P}(l \leq S_n \leq r) \approx \Phi\left(\frac{l - n\mu}{\sqrt{n\sigma^2}}\right) - \Phi\left(\frac{r - n\mu}{\sqrt{n\sigma^2}}\right)$.

Тук използвахме лиинеините свойства на нормалното разпределение.

За биномно разпределената случайна величина $S_n = X_1 + X_2 + \ldots + X_n$ с $X_i \sim \mathrm{Ber}(n, p)$ знаем, че за $n = 6$ млн. $\gg 30$ и $p = \frac{1}{6}$ имаме $\mu = np = 1$ млн. и $\sigma^2 = np(1-p) = 1$ млн. $\times \frac{5}{6}$.

Следователно,

$$\mathbb{P}\left(S_{6 \text{ млн.}} > 1 \text{ млн.}\right) \approx \mathbb{P}\left(\mathcal{N}(\mu, \sigma^2) > 1 \text{ млн.}\right) = \mathbb{P}\left(\mathcal{N}(0{,}1) > \frac{1 \text{ млн.} - n\mu}{\sigma}\right) =$$

$$= \mathbb{P}\left(\mathcal{N}(0{,}1) > 0\right) = 1 - \mathbb{P}(\mathcal{N}(0{,}1) \leq 0) = 1 - \Phi(0) = \frac{1}{2}.$$

Или директно от горната изведена формула:

$$\mathbb{P}\left(S_n > 1 \text{ млн.}\right) = 1 - \mathbb{P}\left(S_n \leq 1 \text{ млн.}\right) \approx 1 - \Phi\left(\frac{1 \text{ млн.} - 6 \text{ млн.} \times \frac{1}{6}}{\sqrt{6 \text{ млн.} \times 1 \text{ млн.} \times \frac{5}{6}}}\right) = 1 - \Phi(0) = \frac{1}{2}$$