# Първо контролно по статистгика и емпирични методи спец. софтуерно инженерство

Вариант 1 ф.н. 62369

**Задача 1**. Три зара се хвърлят последователно 5 пъти. Каква е вероятността, броят на хвърлянията, при които се падат само нечетни точки да бъде четен? Да се намери средната стойност на този брой.

## Решение:

Дефинираме случайните величини  $A=\{$ брой паднали се нечетни точки при хвърлянето на 3 зара $\}$ , а  $A_i=\{$  на i  $^{-\text{тия}}$  зар се пада нечетен брой точки $\}$ , за i=1,2,3.

 $A_i$  е бернулиево разпределена случайна величина с вероятност за успех  $p=\frac{3}{6}=\frac{1}{2}.$ 

A е биномно разпределена случайна величина.

$$A\in Bin\left(n=3,\,p_A=rac{1}{2}
ight)$$
. Вероятността и на трите зара да се падне

нечетен брой точки е равна на 
$$\mathbb{P}(A=3) = \binom{3}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{3-3} = \frac{1}{8}.$$

Нека дефинираме още  $X=\{$  брой паднали се само нечетни цифри при n хвърляния на 3 зара $\}$ . В нашия експеримент имаме, че n=5.

Тъй като случайната величина X брой успехите на бернулиеви събития с вероятност за успех  $p_X = \frac{1}{8}$ , то  $X \in Bin\left(n = 5, p_X = \frac{1}{8}\right)$ .

В задачата се търси:

$$\mathbb{P}_{2k \le 5}(X = 2k) = \mathbb{P}(X = 0) + \mathbb{P}(X = 2) + \mathbb{P}(X = 4) = \\
= \binom{5}{0} p_X^0 (1 - p_X)^5 + \binom{5}{2} p_X^2 (1 - p_X)^3 + \binom{5}{4} p_X^4 (1 - p_X)^1 = \\
= \left(\frac{7}{8}\right)^5 + \frac{5!}{2!3!} \left(\frac{1}{8}\right)^2 \left(\frac{7}{8}\right)^3 + \frac{5!}{4!1!} \left(\frac{1}{8}\right)^4 \left(\frac{7}{8}\right)^1 = \\
= \left(\frac{7}{8}\right)^5 + 10 \left(\frac{1}{8}\right)^2 \left(\frac{7}{8}\right)^3 + 5 \left(\frac{1}{8}\right)^4 \frac{7}{8} = \\
= \frac{7^5 + 10 \times 7^3 + 5 \times 7}{8^5} = \frac{16807 + 3430 + 35}{32768} = \\
= \frac{20272}{32768} = 0.61865234375 \approx 0.618$$

$$\mathbb{E}[X]=\sum_{k=0}^n k inom{n}{k} p^k q^{n-k}$$
, където  $n=5,\, p:=p_X=rac{1}{8},\, q:=q_X=1-p_X=1-rac{1}{8}=rac{7}{8}$ 

Ще го смятаме в общия случай:

$$\mathbb{E}[X] = \sum_{k=0}^{n} k \binom{n}{k} p^{k} q^{n-k} = \sum_{k=1}^{n} k \binom{n}{k} p^{k} q^{n-k} = \sum_{k=1}^{n} k \times \frac{n}{k} \binom{n-1}{k-1} p^{k} q^{n-k} =$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \stackrel{m:=n-1}{\stackrel{j:=k-1}{=}} np \sum_{j=0}^{m} \binom{m}{j} p^{j} q^{m-j} =$$

$$= np(p+q)^{m} = np$$

$$\Rightarrow \mathbb{E}X = np = 5 \times 18 = \frac{5}{9} = 0.625.$$

**Задача 2**. Нека  $n\geq 2$  е естествено число. Дадени са  $\overline{n}$  независими събития  $A_1,\,A_2,\,\ldots,\,A_n$ . Да се докаже, че събитията  $\overline{A_1},\,\overline{A_2},\,\ldots,\,\overline{A_n}$  също са независими.

# Доказателство:

Да вземем n=2 и да проверим верността на Лемата от условието на задачата за две независими събития  $A_1$  и  $A_2$ .

От дефиницията за независимост имаме, че  $\mathbb{P}(A_1A_2)=\mathbb{P}(A_1)\times \mathbb{P}(A_2)$ , където  $A_1A_2:=A_1\cap A_2$  за краткост.

Сега,  $A_2$  и  $\overline{A}_2$  ни дават тривиалната пълна група от събития  $(A_2\cap \overline{A}_2= \mbox{\O}$  и  $A_2\cup \overline{A}_2=\Omega)$ . Следователно, от формулата за пълната вероятност имаме, че:  $\mathbb{P}(A_1)=\mathbb{P}(A_1A_2)+\mathbb{P}(A_1\overline{A}_2)$ .

Тогава,

$$\mathbb{P}(A_1\overline{A_2}) = \mathbb{P}(A_1) - \mathbb{P}(A_1A_2) = \mathbb{P}(A_1) - \mathbb{P}(A_1)\mathbb{P}(\overline{A_2}) = \mathbb{P}(A_1)\underbrace{\left(1 - \mathbb{P}(A_2)\right)}_{\mathbb{P}(\overline{A_2})} = \mathbb{P}(A_1)\mathbb{P}(\overline{A_2}),$$

което по дефиниция означава, че  $A_1$  и  $\overline{A_2}$  са независими събития.

Дуално, ако разбием  $\Omega$  на  $A_1$  и  $\overline{A}_1$  и отново приложим формулата за пълната вероятност, ще получим:  $\mathbb{P}(A_2) = \mathbb{P}(A_2A_1) + \mathbb{P}(A_2\overline{A}_1) \Rightarrow \mathbb{P}(\overline{A_1}A_2) = \mathbb{P}(A_2) - \mathbb{P}(A_1A_2) = \mathbb{P}(A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2) = \mathbb{P}(A_2)\left(1 - \mathbb{P}(A_1)\right) = \mathbb{P}(A_2)\mathbb{P}(\overline{A_1}),$  което отново по дефиниция показва, че и събитията  $\overline{A}_1$  и  $A_2$  също са независими.

Остана да докажем, че  $\overline{A}_1$  и  $\overline{A}_2$  също са независими събития.

За  $\mathbb{P}(\overline{A}_1\overline{A}_2)$  може да използваме същата техника както до сега.

$$\mathbb{P}(\overline{A_1}) = \mathbb{P}(\overline{A_1}A_2) + \mathbb{P}(\overline{A_1}\overline{A_2}) \Rightarrow \mathbb{P}(\overline{A_1}\overline{A_2}) = \mathbb{P}(\overline{A_1}) - \mathbb{P}(\overline{A_1}A_2)$$
, но ние вече сме доказали, че  $\overline{A_1}$  и  $A_2$  са независими събития, откъдето следва, че  $\mathbb{P}(\overline{A_1}\overline{A_2}) = \mathbb{P}(\overline{A_1}) - \mathbb{P}(\overline{A_1})\mathbb{P}(\overline{A_2}) = \mathbb{P}(\overline{A_1})\left(1 - \mathbb{P}(A_2)\right) = \mathbb{P}(\overline{A_1})\mathbb{P}(\overline{A_2})$ .

Този резултат ще ни послужи за база на пълната математическа индукция (пълна, защото използва всички резултати преди нея, а не

само предходния). Правим хипотезата, че Лемата е изпълнена за някое  $n=k\geq 2$ . Ще докажем, че е изпълнена и за n=k+1.

Индукционна стъпка (индукционен преход):

От индукционната хипотезата следва, че  $A_1^*, A_2^*, \ldots, A_k^*$  са независими събития и искаме да докежем, че и  $A_1^*, A_2^*, \ldots, A_k^*, A_{k+1}^*$  са независими, като знаем, че  $A_1, A_2, \ldots, A_k, A_{k+1}$  са независими.  $(A_j^*$  е някое от събитията:  $A_j, \overline{A_j}$ )

<u>Твърдение</u>: Двете събития  $A_1A_2\dots A_k = \bigcap_{i=1}^k A_i$  и  $A_{k+1}$  са независими т.с.т.к  $\Leftrightarrow A_1, A_2, \dots, A_k, A_{k+1}$  са независими.

Това е така, защото твърденията от двете страни са еквивалентни на  $\mathbb{P}(A_1A_2\dots A_k)\mathbb{P}(A_{k+1})$  (тъй като по дефиниция

$$\mathbb{P}(A_1)\mathbb{P}(A_2)\dots\mathbb{P}(A_k)\mathbb{P}(A_{k+1}) = \mathbb{P}\left(\left(\bigcap_{i=1}^k A_i\right)\cap A_{k+1}\right).$$

Отново прилагаме аналогична техника, както в индукционната база – разбиваме по новото събитие:

$$\mathbb{P}\left(\bigcap_{i=1}^k A_i\right) = \mathbb{P}\left(\bigcap_{\underline{i=1}}^k A_i \cap A_{k+1}\right) + \mathbb{P}\left(\bigcap_{\underline{i=1}}^k A_i \cap \overline{A}_{k+1}\right) \Rightarrow$$

$$\Rightarrow \mathbb{P}\left(\bigcap_{i=1}^{k} A_{i} \cap \overline{A_{k+1}}\right) = \mathbb{P}\left(\bigcap_{i=1}^{k} A_{i}\right) - \mathbb{P}\left(\bigcap_{i=1}^{k} A_{i} \cap A_{k+1}\right) =$$

$$= \mathbb{P}\left(\bigcap_{i=1}^{k} A_{i}\right) - \mathbb{P}\left(\bigcap_{i=1}^{k} A_{i}\right) \mathbb{P}(A_{k+1}) = \mathbb{P}\left(\bigcap_{i=1}^{k} A_{i}\right) \left(1 - \mathbb{P}(A_{k+1})\right) =$$

тъй като  $A_1, A_2, \ldots, A_k, A_{k+1}$ 

$$= \mathbb{P}\left(\bigcap_{i=1}^{k} A_i\right) \left(\overline{A_{k+1}}\right)$$

$$\Rightarrow \bigcap_{i=1}^k A_i$$
 и  $\overline{A_{k+1}}$  са независими  $\Rightarrow A_1,\,A_2,\,\ldots,\,A_k,\,\overline{A_{k+1}}$  са независими.

Аналогично, ако означим  $\bigcap_{i=1}^k A_i = B$  за улеснение, то ще получим, че

$$\mathbb{P}(A_{k+1}) = \mathbb{P}(BA_{k+1}) + \mathbb{P}(\overline{B}A_{k+1})$$
 или

$$\mathbb{P}(\overline{B}A_{k+1}) = \mathbb{P}(A_{k+1}) - \mathbb{P}(BA_{k+1}) = \mathbb{P}(A_{k+1}) \left(1 - \mathbb{P}(B)\right) = \mathbb{P}(A_{k+1})\mathbb{P}(\overline{B})$$

 $\Rightarrow$   $A_{k+1}$  и  $\overline{B}$  са независими.

Остана само да докажем, че и събитията  $\overline{B}$  и  $\overline{A_{k+1}}$  са независими:

$$\begin{split} \mathbb{P}(\overline{B}\,\overline{A_{k+1}}) &= \mathbb{P}(\overline{B}\,) - \mathbb{P}(\overline{B}A_{k+1}) = \mathbb{P}(\overline{B}\,) - \mathbb{P}(\overline{B}\,)\mathbb{P}(A_{k+1}) = \\ &= \mathbb{P}(\overline{B}\,) \Big(1 - \mathbb{P}(A_{k+1})\Big) = \mathbb{P}(\overline{B}\,)\mathbb{P}(\overline{A_{k+1}}), \text{ което искахме да докажем.} \end{split}$$

Очевидно резултата, който искахме да докажем в задача 2 е частен случай на доказаната Лема.

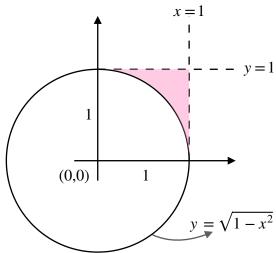
**Задача 3**. По случаен начин се избират две числа в интервала [0,1]. Каква е вероятността сумата от квадратите им да бъде по-голяма от 1?

## Решение:

Нека с x и y бележим двете случайно избрани числа. В задачата се търси:

$$\mathbb{P}(\{x^2 + y^2 > 1\} \mid \{0 \le x, y, \le 1\}) = \frac{\mathbb{P}(\{x^2 + y^2 > 1\} \cap \{0 \le x, y, \le 1\})}{\mathbb{P}(\{0 \le x, y \le 1\})} = \frac{\mu(\{x^2 + y^2 > 1\} \cap \{0 \le x, y \le 1\})}{\mu(\{0 \le x, y \le 1\})} = \mu(\{x^2 + y^2 > 1\} \cap \{0 \le x, y \le 1\}) = M.$$

 $x^2 + y^2 = 1$  е уравнение на окръжност с радиус 1 и център точката (0,0). А пък другите две ограничения отсичат от  $O\overrightarrow{xy}$  квадрат със страна 1.



Вероятността, която търсим се свежда до геометрична вероятност и е равна на лицето (мярката) на розовата част от чертежа по-горе.

Лицето на тази част може да пресметнем по два начина:

**I<sup>-BИ</sup> н/н**: (Чрез съображения и наблюдения по чертежа).

Съобразяваме, че лицето на розовата част е равно на лицето на квадрата заключен между положителните посоки на абсцисата и

ординатата минус  $\frac{1}{4}$  от лицето на окръжността. Следователно  $M=1-\frac{\pi r^2}{4}=1-\frac{\pi}{4}\approx 0.2146018366.$ 

**II<sup>-ри</sup> н/н**: Чрез груба сила и директно интегриране без мислене (универсален начин):

$$\begin{split} M &= \int_0^1 \left(1 - \sqrt{1 - x^2}\right) \mathrm{d} \, x = \\ &= x \Big|_0^1 - \int_0^1 \sqrt{1 - x^2} \, \mathrm{d} \, x \stackrel{\text{полагания}}{=} \left\{ \begin{aligned} x &= \sin(u) \\ \mathrm{d} \, x &= \cos(u) \mathrm{d} \, u \\ x &\in [0, 1] \to u \in \left[0, \frac{\pi}{2}\right] \end{aligned} \right. \\ &= 1 - \int_0^{\frac{\pi}{2}} \cos(u) \sqrt{1 - \sin^2(u)} \, \mathrm{d} \, u = 1 - \int_0^{\frac{\pi}{2}} \frac{\cos(2u) + 1}{2} \, \mathrm{d} \, u = \\ &= 1 - \frac{1}{4} \int_0^{\frac{\pi}{2}} \cos(2u) \mathrm{d} \, 2u - \frac{1}{2} \int_0^{\frac{\pi}{2}} \mathrm{d} \, u = \\ &= 1 - \frac{1}{4} \sin(2u) \Big|_0^{\frac{\pi}{2}} - \frac{1}{2} \times \frac{\pi}{2} = 1 - \frac{\pi}{4} \approx 0.2146018366. \end{split}$$

Задача 4. На състезание участват 25 отбора: 8 отбора в категория джипове, 10 при камиони и 7 при мотоциклети. Джиповете завършват състезанието с вероятност 0.9, камионите с 0.7, а моторите с вероятност 0.6. След състезанието на случаен принцип се избират три отбора , за провеждане на технически контрол. Известно е че един от избраните три отбора е завършил състезанието, а другите два не. Каква е вероятността избраните три отбора да са от различни категории?

#### Решение:

Означаваме следните събития:

 $\mathscr{D} = \{$ случайно избран отбор е в категория "джипове" $\}$   $\mathscr{K} = \{$ случайно избран отбор е в категория "камиони" $\}$   $\mathscr{M} = \{$ случайно избран отбор е в категория "мотоциклети" $\}$ 

Тъй като  $\mathcal{D}\cap\mathcal{K}=\mathcal{K}\cap\mathcal{M}=\mathcal{M}\cap\mathcal{D}=\emptyset$  и  $\mathcal{D}\cup\mathcal{K}\cup\mathcal{M}=\Omega$ , то  $\mathcal{D},$   $\mathcal{K}$  и  $\mathcal{M}$  са разбиване на  $\Omega$  (пълна група от събития). Имаме, че  $\mathbb{P}(\mathcal{D})=\frac{8}{25};\,\mathbb{P}(\mathcal{K})=\frac{10}{25}$  и  $\mathbb{P}(\mathcal{M})=\frac{7}{25}.$ 

сумират се до 1, тъй като са разбиване

Нека  $\Phi$  е събитието {отбор финишира}. От условието ще паднат следните условни вероятности:

$$\mathbb{P}(\Phi \,|\, \mathcal{D}) = 0.9$$
,  $\mathbb{P}(\Phi \,|\, \mathcal{K}) = 0.7$  и  $\mathbb{P}(\Phi \,|\, \mathcal{M}) = 0.6$ .

Нека A е събитието {**точно 1** от отборите, на които се провежда технически преглед е финиширал}

От формулата за пълната вероятност ще имаме, че  $\mathbb{P}(A) = \sum_{k \in I} \mathbb{P}(A \,|\, H_k) \mathbb{P}(H_k)$ , където  $I = \{$  множеството от всички

възможни сценарии за това с какво са участвали 3 отбора}

Питаме се кое е това множество I? Ще го изброим с груба сила, но все пак ще очакваме, то да е от кардиналност |I|=10.

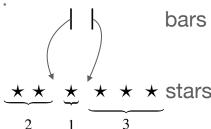
Ще очакваме тази кардиналност за множеството I, тъй като дефинираме  $I=\left\{H_{ijk}$ , където  $H_{ijk}=\left\{i\times \mathscr{D}+j\times K+k\times \mathscr{M}\right.$  за i+j+k=3 и  $0\leq u,j,k\leq 3\}\right\}$ .

Как пресмятаме кардиналността без грубо изброяване? Ами много лесно:

|I| е равно на броя на решенията на уравнението  $x_1+x_2+x_3=3$  в цели неотрицателни числа. Последното е равно на броя на решенията на уравнението  $y_1+y_2+y_3=6$  в цели положителни числа (положихме  $x_i=y_i+1$  за да премахнем възможните нули от събираемите). Сега прилагайки трика – "stars and bars", получаваме, че

$$|I| = {6-1 \choose 3-1} = {5 \choose 2} = 10.$$

Пояснение на "stars and bars":



Така поставените "bars" между "stars", дават решението:  $2+1+3=y_1+y_2+y_3=6$ . Може да поставим 2 "bars" измежду 6 "stars" като изберем двете позиции измежду пет (без най-отпред и найотзад) възможни.

В крайна сметка, с така въведената нотация, търсим:

$$\mathbb{P}(H_{111}|A) \longrightarrow \text{точно един отбор е завршил състезанието}$$

Това е:

$$\mathbb{P}(H_{111} | A) = \frac{\mathbb{P}(H_{111} \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \, | \, H_{111}) \mathbb{P}(H_{111})}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \, | \, H_{111}) \mathbb{P}(H_{111})}{\sum_{k \in I} \mathbb{P}(A \, | \, H_k) \mathbb{P}(H_k)}$$
 имаме го разбито

Всъщност, ние просто доказахме формулата на Бейс, чрез формулата за усовна и пълна вероятности.

Грубо изброяване и пресмятане:

$$\left( \begin{array}{c}
\mathbb{P}(H_{ijk}) \\
i, j, k \in \mathbb{N}_0 \\
i+j+k=3
\end{array} \right) = \frac{\binom{8}{i} \binom{10}{j} \binom{7}{k}}{\binom{25}{3}} = \frac{\binom{25}{3}}{\binom{25}{3}}$$

$$\begin{split} \mathbb{P}(H_{012}) &= \frac{\binom{8}{0}\binom{10}{1}\binom{7}{2}}{\binom{25}{3}} = \frac{350}{2300}; \, \mathbb{P}(H_{003}) = \frac{\binom{8}{0}\binom{10}{0}\binom{7}{3}}{\binom{25}{3}} = \frac{35}{2300}; \\ \mathbb{P}(H_{030}) &= \frac{\binom{8}{0}\binom{10}{3}\binom{7}{0}}{\binom{25}{3}} = \frac{120}{2300}; \, \mathbb{P}(H_{021}) = \frac{\binom{8}{0}\binom{10}{0}\binom{7}{1}}{\binom{25}{3}} = \frac{315}{2300}; \\ \mathbb{P}(H_{120}) &= \frac{\binom{8}{1}\binom{10}{2}\binom{7}{0}}{\binom{25}{3}} = \frac{320}{2300}; \, \mathbb{P}(H_{102}) = \frac{\binom{8}{1}\binom{10}{0}\binom{7}{1}}{\binom{25}{3}} = \frac{56}{2300}; \\ \mathbb{P}(H_{111}) &= \frac{\binom{8}{1}\binom{10}{1}\binom{7}{1}}{\binom{25}{3}} = \frac{560}{2300}; \\ \mathbb{P}(H_{201}) &= \frac{\binom{8}{2}\binom{10}{0}\binom{7}{1}}{\binom{25}{3}} = \frac{56}{2300}; \\ \mathbb{P}(H_{300}) &= \frac{\binom{8}{1}\binom{10}{0}\binom{7}{0}}{\binom{25}{3}} = \frac{56}{2300}. \end{split}$$

Припомняме означенията:

$$\begin{array}{l} \begin{subarray}{l} \begin{subarray}{l$$

Сега вече имаме всичко необходимо пресметнато. Следователно,

$$\mathbb{P}(H_{111}|A) = \frac{\mathbb{P}(A|H_{111})\mathbb{P}(H_{111})}{\sum_{k \in I} \mathbb{P}(A|H_k)\mathbb{P}(H_k)} =$$

$$= \frac{\frac{0.154 \times \frac{560}{2300}}{\frac{1}{2300} \times \left(0.256 \times 350 + 0.288 \times 35 + 0.189 \times 120 + 0.222 \times 315 + 0.123 \times 320 + 0.192 \times 56 + 0.154 \times 560 + 0.078 \times 56 + 0.061 \times 280 + 0.027 \times 56\right)}{1.54 \times 56} = \frac{\frac{1.54 \times 56}{2.56 \times 35 + 0.288 \times 35 + 1.89 \times 12 + 0.222 \times 315 + 1.23 \times 32 + 0.192 \times 56 + 1.54 \times 56 + 0.078 \times 56 + 0.61 \times 28 + 0.027 \times 56}}{86.24} = \frac{\frac{86.24}{192.29 + 136.352 + 21.448 + 1.512}}{192.29 + 136.352 + 21.448 + 1.512} = \frac{\frac{86.24}{351.602}}{351.602} \approx 0.24527733061 \approx 0.245$$