Вероятности и статистика

Факултет по математика и информатика, Софийски университет
h.sariev@math.bas.bg

Легенда. *Основен материал*: с (*) е видяно в клас, но няма пряко отношение към изпита; с (**) – за по-любознателните. *Задачи*: всичко без (*) се приема за решено в клас; с (*) – допълнителна подготовка; с (**) – трудни задачи.

1 Комбинаторика

Комбинаториката е анализ на крайните множества, напр. $M = \{a_1, \ldots, a_n\}$ за $n \in \mathbb{N}$, като в частност служи за намиране на броя на техните елементи (в случая $|M| = n$) без да се извършва непосредствено преброяване.

Пермутация. *Всяко нареждане на елементите на M в n-членна редица наричаме пермутация на елементите на M, като множеството от пермутации обозначаваме с*

$$\boldsymbol{P}_n = \big\{(a_{i_1}, \ldots, a_{i_n}) : i_1, \ldots, i_n \in \{1, \ldots, n\}, i_j \neq i_l, j \neq l\big\}.$$

Дефинирайки $P_n = |\boldsymbol{P}_n|$, имаме

$$P_n = n!$$

Д-во.* Равенството се доказва индуктивно по n. Очевидно, $\mathbf{P}_1 = \{(a_1)\}$ и $\mathbf{P}_2 = \{(a_1, a_2), (a_2, a_1)\}$, от което следват $P_1 = 1$ и $P_2 = 2$. Нека равенството важи за $n-1$. От всяка пермутация на $\{a_1, \ldots, a_{n-1}\}$ получаваме n на брой пермутации на $\{a_1, \ldots, a_n\}$, като поставяме a_n съответно на първа, втора и т.н. до последна позиция; напр. от $(a_{i_1}, \ldots, a_{i_{n-1}})$, за кои да е $i_1, \ldots, i_{n-1} \in \{1, \ldots, n-1\}$, получаваме

$$(a_n, a_{i_1}, a_{i_2}, \ldots, a_{i_{n-1}}), (a_{i_1}, a_n, a_{i_2}, \ldots, a_{i_{n-1}}), \ldots, (a_{i_1}, a_{i_2}, \ldots, a_{i_{n-1}}, a_n).$$

Нека обозначим с B_l множеството от така конструирани пермутации на $\{a_1, \ldots, a_n\}$, където индексът $l = 1, \ldots, P_{n-1}$ варира измежду всички пермутации на $\{a_1, \ldots, a_{n-1}\}$, като $|B_l| = n$ от по-горе. По замисъл, множествата $B_1, \ldots, B_{P_{n-1}}$ са две по две непресичащи се.

От друга страна, всяка пермутация на $\{a_1, \ldots, a_n\}$ дефинира една единствена пермутация на $\{a_1, \ldots, a_{n-1}\}$ след премахване на елемента a_n. Следователно $\mathbf{P}_n = B_1 \cup \cdots \cup B_{P_{n-1}}$, откъдето

$$P_n = |B_1 \cup \cdots \cup B_{P_{n-1}}| = \sum_{l=1}^{P_{n-1}} |B_l| = n \cdot P_{n-1} = n!$$

$\square$

Забележка (Интерпретация). Започвайки отляво надясно, първият елемент a_{i_1} можем да изберем по n начина, след което за a_{i_2} остават $(n-1)$ възможности, за a_{i_3} – $(n-2)$ и т.н.

*Забележка*** (Strling's Approximation). С нарастването на n числата $n!$ растат много бързо и непосредственото им пресмятане става практически невъзможно. Съществува обаче формула (на Стърлинг), която позволява сравнително лесно да се пресмята *приблизително*

$$n! \sim \sqrt{2n\pi}\left(\frac{n}{e}\right)^n \qquad (\text{при } n \to \infty).$$

Вариация без повторение. *Всяка наредена k-орка от елементи на M без повторение, за $k = 0, 1 \ldots, n$, наричаме вариация на елементите на M от клас k без повторение, като множеството обозначаваме с*

$$\boldsymbol{V}_n^k = \big\{(a_{i_1}, \ldots, a_{i_k}) : i_1, \ldots, i_k \in \{1, \ldots, n\}, i_j \neq i_l, j \neq l\big\}.$$

Дефинирайки $V_n^k = |\boldsymbol{V}_n^k|$, имаме

$$V_n^k \equiv (n)_k = \frac{n!}{(n-k)!}.$$

Д-во.* От всяка наредена k-орка $(a_{i_1}, \dots, a_{i_k})$ без повторение получаваме пермутация на M, като към нея добавим в произволен ред останалите елементи, $M \backslash \{a_{i_1}, \dots, a_{i_k}\}$; напр.,

$$(a_{i_1}, \dots, a_{i_k}, a_{j_1}, \dots, a_{j_{n-k}}),$$

където $(a_{j_1}, \dots, a_{j_{n-k}})$ е пермутация на елементите на $M \backslash \{a_{i_1}, \dots, a_{i_k}\}$. Нека обозначим с B_l множеството от пермутации на M, получени по описания начин от l-тата вариация, където индексът $l = 1, \dots, V_n^k$ преминава през всички възможни вариации на M от клас k. Тогава $|B_l| = P_{n-k}$. По замисъл, множествата $B_1, \dots, B_{V_n^k}$ са две по две непресичащи се.

От друга страна, всяка пермутация $(a_{i_1}, \dots, a_{i_k}, a_{i_{k+1}}, \dots, a_{i_n})$ на M дефинира една единствена вариация от клас k след премахване на последните $n-k$ елементи, т.е. $(a_{i_1}, \dots, a_{i_k})$. Следователно $\mathbf{P}_n = B_1 \cup \dots \cup B_{V_n^k}$, откъдето получаваме, че

$$P_n = |B_1 \cup \dots \cup B_{V_n^k}| = \sum_{l=1}^{V_n^k} |B_l| = P_{n-k} \cdot V_n^k \qquad \Longrightarrow \qquad V_n^k = \frac{P_n}{P_{n-k}}.$$

□

Комбинация без повторение. *Всяко k-елементно подмножество на M, за $k = 0, 1 \dots, n$, наричаме комбинация на елементите на M от клас k без повторение, като множеството обозначаваме с*

$$\boldsymbol{C}_n^k = \big\{\{a_{i_1}, \dots, a_{i_n}\} : i_1, \dots, i_k \in \{1, \dots, n\}, i_j \neq i_l, j \neq l\big\}.$$

Дефинирайки $C_n^k = |\boldsymbol{C}_n^k|$, имаме

$$C_n^k \equiv \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Д-во.* От всяко k-елементно подмножество на M можем да образуваме P_k на брой пермутации. Последните съвпадат с вариациите на M от клас k, откъдето следва, че

$$V_n^k = P_k \cdot C_n^k \qquad \Longrightarrow \qquad C_n^k = \frac{V_n^k}{P_k}.$$

□

Забележка (Интерпретация)*.* За разлика от вариациите от клас k без повторение, подредбата на елементите във всяка комбинация от клас k без повторение е без значение.

Твърдение 1.1. *(i)* $C_n^k = C_n^{n-k}$; *(ii)* $C_n^k = C_{n-1}^k + C_{n-1}^{k-1}$; *(iii)* $\sum_{k=0}^n C_n^k = 2^n$; *(iv)* $\sum_{k=0}^n (-1)^k C_n^k = 0$.

Д-во.* Относно (i), към всяко k-елементно подмножество $A \subseteq M$ съпоставяме неговото допълнение $M - A$, имащо $n-k$ елемента, и обратното, откъдето следва еднозначната обратима връзка между $\mathbf{C}_n^k$ и $\mathbf{C}_n^{n-k}$.

Относно (ii), множеството от комбинации на $\{a_1, \dots, a_n\}$ от клас k се разлага на две непресичащи се множества в зависимост от това дали a_n участва или не. От една страна, това са k-елементните подмножества на $\{a_1, \dots, a_{n-1}\}$, т.е. $\mathbf{C}_{n-1}^k$, а от друга – $(k-1)$-елементните подмножества на $\{a_1, \dots, a_{n-1}\}$, към всяко от които е добавен елементът a_n.

(iii) и (iv) следват от биномната формула, като $\sum_{k=0}^n \binom{n}{k} = (1+1)^n = 2^n$, а $\sum_{k=0}^n (-1)^k \binom{n}{k} = (1-1)^n = 0$.

Като алтернативен подход, (iv) следва от принципа за включване-изключване, приложен върху множествата $A_i = \{a\}$, $i = 1, \dots, n$. Относно (iii), $\mathbf{C}_n^0, \mathbf{C}_n^1, \dots, \mathbf{C}_n^n$ са две по две непресичащи се множества, чието обединение $\mathcal{P}(M) := \mathbf{C}_n^0 \cup \mathbf{C}_n^1 \cup \dots \cup \mathbf{C}_n^n$ е **множеството на всички подмножества**, образувани от елементите на M. От друга страна, на всяко подмножество $A \subseteq M$ можем да съпоставим една и само една n-орка $(e_1, \dots, e_n) \in \{0,1\}^n$ от нули и единици, такава че $a_i \in A \Longleftrightarrow e_i = 1$ за $i = 1, \dots, n$, откъдето

$$\sum_{k=0}^n C_n^k = \Big| \bigcup_{k=0}^n \mathbf{C}_n^k \Big| = |\mathcal{P}(M)| = |\{0,1\}|^n = 2^n.$$

□

*Забележка** (Триъгълника на Паскал). Установените свойства на числата C_n^k могат да се илюстрират по-нагледно, ако ги разположим в един безкраен, симетричен относно височината си триъгълник, при който по бедрата на триъгълника се разполагат единици, а всеки вътрешен елемент е сбор на двата си съседа от предния ред.

$$\begin{array}{ccccccccc} & & & & C_0^0 & & & & \\ & & & C_1^0 & & C_1^1 & & & \\ & & C_2^0 & & C_2^1 & & C_2^2 & & \\ & & & \cdots & \cdots & \cdots & & & \\ C_n^0 & & C_n^1 & & \dots & & C_n^{n-1} & & C_n^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{array} \qquad \begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ \underline{1} & & \underline{\underline{4}} & & \underline{6} & & \underline{\underline{4}} & & \underline{1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{array} \qquad \begin{array}{c} 2^0 \\ 2^1 \\ 2^2 \\ 2^3 \\ 2^4 \end{array}$$

В n-тия ред сумата на елементите е равна на 2^n, като сумата на елементите на четни позиции е равна на сумата на елементите на нечетни позиции.

Пермутация с повторение. *Всяка k-члена редица, при която елементът a_1 се повтаря k_1 пъти, a_2 – k_2 пъти и т.н., където $k_1 + \cdots + k_n = k$, се нарича пермутация на елементите на M с повторение, дължина k и честоти $(k_1, \dots, k_n)$, като множеството обозначаваме с*

$$\boldsymbol{P}(k; k_1, \dots, k_n) = \big\{(a_{i_1}, \dots, a_{i_k}) : i_1, \dots, i_k \in \{1, \dots, n\}, \#\{j : a_{i_j} = a_{i_1}\} = k_1, \dots, \#\{j : a_{i_j} = a_{i_n}\} = k_n\big\}.$$

Дефинирайки $P(k; k_1, \dots, k_n) = |\boldsymbol{P}(k; k_1, \dots, k_n)|$, имаме

$$P(k; k_1, k_2, \dots, k_n) \equiv \binom{k}{k_1, k_2, \dots, k_n} = \frac{k!}{k_1! k_2! \cdots k_n!}.$$

Д-во.* Ако $(a_{i_1}, \dots, a_{i_k})$ е такава k-орка, да означим с $X = \{j \in \{1, \dots, k\} : i_j = 1\}$ множеството от индекси j, за които $a_{i_j} = a_1$. Понеже $|X| = k_1$, то за X има $C_k^{k_1}$ на брой възможни комбинации. Нека $(a_{i_1^*}, \dots, a_{i_{k-k_1}^*})$ е редицата, получена след премахване на всичките k_1 елемента a_1 от $(a_{i_1}, \dots, a_{i_k})$. Тогава $(a_{i_1^*}, \dots, a_{i_{k-k_1}^*})$ представлява пермутация на $\{a_2, \dots, a_n\}$ с повторение, дължина $k - k_1$ и честоти $(k_2, \dots, k_n)$. Обратно, $(a_{i_1^*}, \dots, a_{i_{k-k_1}^*})$ и X определят еднозначно $(a_{i_1}, \dots, a_{i_k})$. Следователно,

$$P(k; k_1, \dots, k_n) = C_k^{k_1} \cdot P(k - k_1; k_2, \dots, k_n),$$

откъдето $P(k; k_1, k_2, \dots, k_n) = C_k^{k_1} C_{k-k_1}^{k_2} \cdots C_{k-k_1-\cdots-k_{n-1}}^{k_n}$, използвайки горната логика още $n - 1$ пъти. $\square$

Забележка (Интерпретация). Елемента a_1 можем да поставим по $\binom{k}{k_1}$ различни начина на k_1 от общо k места в редицата. На всеки от тези начини отговарят по $\binom{k-k_1}{k_2}$ различни начина, по които можем да изберем k_2 от останалите $k - k_1$ места, на които да поставим елемента a_2,..., елемента a_n поставяме по $\binom{k-k_1-\cdots-k_{n-1}}{k_n} = 1$ начин на последните $k - k_1 - \cdots - k_{n-1}$ места.

*Забележка*** (Разлагания на множество). Нека E е k-елементно множество. Всяка n-орка $(X_1, \dots, X_n)$ от подмножества $X_i \subseteq E$, такива че $X_i \cap X_j = \emptyset$, $i \neq j$ и $E = \bigcup_{i=1}^n X_i$, се нарича **n-разлагане на** E. В частност, полагайки $|X_i| = k_i$, редицата $(X_1, \dots, X_n)$ се нарича n-разлагане на E от тип $(k_1, \dots, k_n)$, като $k_1 + \cdots + k_n = k$. Тогава броят на n-разлаганията на E от тип $(k_1, \dots, k_n)$ е

$$\binom{k}{k_1}\binom{k-k_1}{k_2} \cdots \binom{k - k_1 - \cdots - k_{n-1}}{k_n} = \frac{k!}{k_1! \cdots k_n!}.$$

Нека сега $E = \{1, \dots, k\}$. Тогава за фиксирани $(k_1, \dots, k_n)$ съществува еднозначно обратимо съответствие между всички пермутации $(a_{i_1}, \dots, a_{i_k})$ на M с повторение, дължина k и честоти $(k_1, \dots, k_n)$ и всички n-разлагания $(X_1, \dots, X_n)$ на E от тип $(k_1, \dots, k_n)$, дефинирано чрез $a_{i_j} = a_m \iff i_j \in X_m$, за $m = 1, \dots, n$ и $j = 1, \dots, k$. С други думи, множеството X_m съдържа информация за местата в редицата $(a_{i_1}, \dots, a_{i_k})$, на които стои елементът a_m, като знанието на $(X_1, \dots, X_n)$ е достатъчно, за да възпроизведем $(a_{i_1}, \dots, a_{i_k})$; и обратното.

Вариация с повторение. *Всяка наредена k-орка от елементи на M, за k произволно, наричаме вариация на елементите на M от клас k с повторение, като множеството обозначаваме с*

$$\boldsymbol{V}(n,k) = \big\{(a_{i_1},\dots,a_{i_k}) : i_1,\dots,i_k \in \{1,\dots,n\}\big\}.$$

Дефинирайки $V(n,k) = |\,\boldsymbol{V}(n,k)|$, имаме

$$V(n,k) = n^k.$$

Д-во.* Не е трудно да се види, че $\mathbf{V}(n,k) = M \times M \times \cdots \times M$ е k-тата декартова степен на M, откъдето следва, че $V(n,k) = |M|^k$. $\square$

*Забележка*** (Разлагания на множество (Продължение))*.* Нека E е k-елементно множество. Следвайки предходните разсъждения, съществува еднозначно обратимо съответствие между вариациите на M с повторение от клас k и n-разлаганията на E, което означава, че броят на n-разлаганията на E е

$$\sum_{k_1,\dots,k_n \in \mathbb{N}_0 : k_1+\cdots+k_n = k} \frac{k!}{k_1!\cdots k_n!} = n^k.$$

Комбинация с повторение. *Всяко k-елементно мулти-подмножество на M, за k произволно, наричаме комбинация на елементите на M от клас k с повторение, като множеството обозначаваме с*

$$\boldsymbol{C}(n,k) = \big\{\{a_{i_1},\dots,a_{i_k}\} : i_1,\dots,i_k \in \{1,\dots,n\}\big\}.$$

Дефинирайки $C(n,k) = |\,\boldsymbol{C}(n,k)|$, имаме

$$C(n,k) = \binom{n+k-1}{k}.$$

Д-во.* Допълваме множеството M с $k-1$ нови елемента $a_{n+1}, a_{n+2}, \dots, a_{n+k-1}$. Получаваме множеството $M^* = \{a_1,\dots,a_{n+k-1}\}$. Нека $\{a_{i_1},\dots,a_{i_k}\} \in \mathbf{C}(n,k)$ е k-елементно мулти-подмножество на M, където $i_1 \le i_2 \le \cdots \le i_k$ без загуба на общност. Ако положим $i_j^* = i_j + (j-1)$, $j = 1,\dots,k$, получаваме комбинацията $\{a_{i_1^*},\dots,a_{i_k^*}\}$ от елементи на M^* без повторение от клас k, т.е. $\{a_{i_1^*},\dots,a_{i_k^*}\} \in \mathbf{C}^k_{n+k-1}$. Така дефинирано, съответствието между $\mathbf{C}(n,k)$ и $\mathbf{C}^k_{n+k-1}$ е взаимно еднозначно, от което следва, че $C(n,k) = C^k_{n+k-1}$. $\square$

Твърдение 1.2.** *$C(n,k) = \sum_{i=0}^k C(n-1,i)$, при което $C(n,k) = C(n,k-1) + C(n-1,k)$.*

*Д-во**.* Нека $\{a_{i_1},\dots,a_{i_{k-m}}, a_n, a_n, \dots, a_n\} \in \mathbf{C}(n,k)$ е k-елементно мулти-подмножество на M, при което a_n се среща m пъти, за някое $m = 0,1,\dots,k$. Тогава $\{a_{i_1},\dots,a_{i_{k-m}}\} \in \mathbf{C}(n-1,k-m)$ е комбинация на $\{a_1,\dots,a_{n-1}\}$ с повторение, чрез която можем да възстановим първоначалната комбинация, като добавим m пъти a_n. Т.е., съществува еднозначно обратимо съответствие между $\mathbf{C}(n,k)$ и $\bigcup_{i=0}^k \mathbf{C}(n-1,i)$. Понеже множествата $\mathbf{C}(n-1,k),\dots,\mathbf{C}(n-1,0)$ са две по две непресичащи се, имаме $|\mathbf{C}(n,k)| = \sum_{i=0}^k |\mathbf{C}(n-1,i)|$. $\square$

*Забележка*** (Триъгълника на Паскал)*.* По силата на рекурентната си връзка, числата $C(n,k)$ могат да бъдат разположени в един безкраен триъгълник, при който всеки вътрешен елемент е сбор на двата си съседа от предния ред, а по бедрата на триъгълника са разположени единици.

$$\begin{array}{c}
C(1,0) \\
C(2,0) \quad C(1,1) \\
C(3,0) \quad C(2,1) \quad C(1,2) \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
C(m,0) \quad C(m-1,1) \quad \dots \quad C(2,m-2) \quad C(1,m-1) \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{array}$$

Принцип за включване-изключване. *Нека $A_1, \dots, A_n$. са крайни множества. Тогава*

$$\Big|\bigcup_{i=1}^n A_i\Big| = \sum_{i=1}^n |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \dots + (-1)^{n-1}|A_1 \cap \dots \cap A_n|.$$

Д-во.* При $n = 2$, нека $A_1 = \{a_1, \dots, a_n\}$ и $A_2 = \{b_1, \dots, b_m\}$ удовлетворяват $A_1 \cap A_2 = \emptyset$, т.е. $a_i \neq b_j$. Да положим $c_k = a_k$, $k = 1, \dots, n$ и $c_k = b_{n-k}$, $k = n+1, \dots, n+m$. Тогава $A \cup B = \{c_1, \dots, c_{n+m}\}$, от което следва, че $|A_1 \cup A_2| = n + m = |A_1| + |A_2|$. Следователно, за произволни A_1, A_2, имаме

$$|A_1 \cup A_2| = \big|(A_1 \cap A_2^c) \cup (A_1^c \cap A_2) \cup (A_1 \cap A_2)\big| = |A_1 \cap A_2^c| + |A_1^c \cap A_2| + |A_1 \cap A_2|.$$

Но $|A_1| = \big|(A_1 \cap A_2) \cup (A_1 \cap A_2^c)\big| = |A_1 \cap A_2| + |A_1 \cap A_2^c|$ и $|A_2| = |A_2 \cap A_1| + |A_2 \cap A_1^c|$, откъдето

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|.$$

Нека равенството важи за $n - 1$. Тогава

$$\begin{aligned}\Big|\bigcup_{i=1}^n A_i\Big| &= \Big|\bigcup_{i=1}^{n-1} A_i\Big| + |A_n| - \Big|\bigcup_{i=1}^{n-1} (A_i \cap A_n)\Big| = \sum_{i=1}^{n-1} |A_i| - \sum_{i<j\le n-1} |A_i \cap A_j| + \dots + (-1)^{n-2}|A_1 \cap \dots \cap A_{n-1}| + |A_n| - \\ &\quad - \Big(\sum_{i=1}^{n-1} |A_i \cap A_n| - \sum_{i<j\le n-1} |A_i \cap A_j \cap A_n| + \dots + (-1)^{n-2}|A_1 \cap \dots \cap A_n|\Big) \\ &= \sum_{i=1}^n |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \dots + (-1)^{n-1}|A_1 \cap \dots \cap A_n|.\end{aligned}$$

*Комбинаторно д-во**.* Нека $x \in \bigcup_{i=1}^n A_i$, като x е общ елемент на $k \in \{1, \dots, n\}$ от множествата, да кажем $A_1, \dots, A_k$. Тогава в сумата $\sum_{i=1}^n |A_i|$ елементът x е преброен $\binom{k}{1}$ пъти, в $\sum_{i<j} |A_i \cap A_j|$ – е преброен $\binom{k}{2}$ пъти,..., в $\sum_{i_1<i_2<\dots<i_k} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}|$ – е преброен $\binom{k}{k} = 1$ път, откъдето

$$\binom{k}{1} - \binom{k}{2} + \dots + (-1)^{k-1}\binom{k}{k} = 1 - \sum_{i=0}^k \binom{k}{i}(-1)^i = 1 - (1-1)^k = 1.$$

□

1.1 Задачи

Задача 1.1. Урна съдържа топки, номерирани с числата $1, 2, \dots, n$. Изваждаме последователно k пъти по една топка, като всеки път записваме нейния номер. Полученият резултат наричаме извадка с обем k от n елемента. Да се пресметне броят на всички различни извадки, ако

- (а) не връщаме извадените топки и редът на записване на изтеглените номера има значение;
- (б) не връщаме извадените топки и редът на записване на изтеглените номера няма значение;
- (в) връщаме всяка извадена топка обратно в урната и редът на записване има значение;
- (г) връщаме всяка извадена топка обратно в урната и редът на записване няма значение.

Отговори. (а) V_n^k; (б) C_n^k; (в) $V(n,k)$; (г) $C(n,k)$. □

Задача 1.2. Намерете броят на отделните анаграми на думата MISSISSIPPI.

Отговор. $P(11; 1, 4, 4, 2)$ □

Задача 1.3. Намерете броя на възможните начини за разпределяне на k частици в n различими клетки, ако

- (а) частиците са различими/неразличими и всяка клетка може да съдържа най-много една частица;
- (б) частиците са различими/неразличими (и клетките могат да съдържат произволен брой частици);
- (в) частиците са различими/неразличими и няма празна клетка (за $k \ge n$);
- (г) частиците са различими/неразличими, в клетка 1 има точно s частици, а в останалите клетки – най-много по една частица (за $k \ge s \ge k - n + 1$);
- (д*) частиците са различими/неразличими, клетка 1 побира произволен брой частици, а останалите клетки – най-много по една частица;
- (е*) частиците са различими/неразличими, в клетка 1 има не повече от s частици, а в останалите клетки – най-много по една частица;
- (ж*) частиците са различими/неразличими, в клетка 1 има поне s частици, а в останалите клетки – най-много по една частица.

Решение (Техника). $M = \{1, \ldots, n\}$.

(а) За *различими*: V_n^k; за *неразличими*: C_n^k.

(б) За *различими*: $V(n, k)$; за *неразличими*: $C(n, k)$.

(в) За *различими*: нека $A_i = \{$клетка i е празна$\}$ за $i = 1, \ldots, n$. Тогава $\bigcup_{i=1}^n A_i = \{$поне една празна$\}$. От друга страна $|A_i| = V(n-1, k)$, $|A_i \cap A_j| = V(n-2, k)$ и т.н. Следователно общият брой разпределения без празна клетка е

$$V(n,k) - \Big|\bigcup_{i=1}^n A_i\Big| = n^k - \Big(n \cdot V(n-1,k) - \binom{n}{2}V(n-2,k) + \cdots + (-1)^{n-1} \cdot 0\Big) = \sum_{i=0}^n (-1)^i \binom{n}{i}(n-i)^k;$$

за *неразличими*: при условие че във всяка клетка има поне една топка, се търси комбинация от вида $\{1, 2, \ldots, n, i_1, \ldots, i_{k-n}\}$ за $i_1, \ldots, i_{k-n} \in M$, т.е. броят на разпределенията е $C(n, k-n) = C_{k-1}^{n-1}$.

(г) За *различими*: търси се редица от вида $(\ldots, i_1, \ldots, 1, \ldots, 1, \ldots, i_{k-s}, \ldots)$, имаща s "1"-ци на места $j_1, \ldots, j_s \in \{1, \ldots, k\}$ и $k - s$ елемента $i_1, \ldots, i_{k-s} \in \{2, \ldots, n\}$ без повторение. Тя се разлага еднозначно на редицата $(i_1, \ldots, i_{k-s})$ и комбинацията $\{j_1, \ldots, j_s\}$, т.е. общият брой е $\binom{k}{s}V_{n-1}^{k-s}$. За *неразличими*: търси се $\{1, \ldots, 1, i_1, \ldots, i_{k-s}\} : i_1, \ldots, i_{k-s} \in \{2, \ldots, n\}, i_j \neq i_l$, т.е. C_{n-1}^{k-s}.

(д) За *различими*: $\sum_{j=\max(0,k-n+1)}^k \binom{k}{j}V_{n-1}^{k-j}$; за *неразличими*: $\sum_{j=\max(0,k-n+1)}^k C_{n-1}^{k-j}$.

(е) За *различими*: $\sum_{j=\max(0,k-n+1)}^{\min(s,k)} \binom{k}{j}V_{n-1}^{k-j}$; за *неразличими*: $\sum_{j=\max(0,k-n+1)}^{\min(s,k)} C_{n-1}^{k-j}$.

(ж) За *различими*: $\sum_{j=s}^k \binom{k}{j}V_{n-1}^{k-j}$, ако $n-1 \geq k-s$, и $\sum_{j=k-n+1}^k \binom{k}{j}V_{n-1}^{k-j}$, ако $n-1 < k-s$; за *неразличими*: $\sum_{j=s}^k C_{n-1}^{k-j}$, ако $n-1 \geq k-s$, и $\sum_{j=k-n+1}^k C_{n-1}^{k-j}$, ако $n-1 < k-s$.

□

Задача 1.4. Колко решения има уравнението $x_1 + \cdots + x_k = n$, ако

(а) $x_1, \ldots, x_k$ са естествени числа (за $k \leq n$);

(б) $x_1, \ldots, x_k$ са неотрицателни цели числа.

**Колко на брой са решенията, ако редът на записването на x_i няма значение?

Решение (Техника).

(а) *Решение 1*: задачата е еквивалентна на разпределянето на n неразличими частици в k клетки, като не остава празна клетка (Задача 1.3в), т.е. броят на решенията е $C(k, n-k)$.
Решение 2 (stars & bars): всяко решение на задачата може да се визуализира като наредба от n звезди "$*$" и $k-1$ черти "$|$", напр. $**||**|*\cdots|***|$, където $x_1 = \{$брой $*$ преди първата $|\}$, $x_k = \{$брой $*$ след последната $|\}$ и $x_i = \{$брой $*$ между $(i-1)$-тата и i-тата $|\}$, $i = 2, \ldots, k-1$. Т.е., проблемът се свежда до избор на $k-1$ места, на които да поставим "$|$", избрани от $n+1$ възможни места в редицата от звезди $**\cdots**$. Понеже $x_i > 0$ в (а), не можем да имаме две последователни черти, нито да започваме или завършваме с "$|$". С други думи, търсим $\{j_1, \ldots, j_{k-1}\}$ за $j_1, \ldots, j_{k-1} \in \{1, \ldots, n-1\} : j_j \neq j_l$, откъдето намираме броя на решенията, C_{n-1}^{k-1}.

(б) *Решение 1*: задачата е еквивалентна на разпределянето на n неразличими частици в k клетки (Задача 1.3б), т.е. броят на решенията е $C(k, n)$.
Решение 2: съществува еднозначна обратимост между решенията на $x_1 + \cdots + x_k = n$ и тези на $(x_1+1) + \cdots + (x_k+1) = n+k$; т.е, C_{n+k-1}^{k-1} от (а).

□

Задача 1.5. Колко четирицифрени числа могат да се напишат с цифрите $1, 2, 3, 4$ и 5, ако

(а) не се допуска повторение на цифри;

(б) допуска се повторение на цифри;

(в) не се допускат повторения и числото е нечетно.

Отговори. (а) V_5^4; (б) $V(5, 4)$; (в) $3 \cdot V_4^3$. □

Задача 1.6. По колко начина може да се избере 4-членна делегация от 12 кандидати, ако

(а) няма ограничение за участие в нея;

(б) A и B не трябва да участват заедно;

(в) C и D могат да участват само заедно.

Отговори. (а) C_{12}^4; (б) $C_{10}^4 + 2 \cdot C_{10}^3$; (в) $C_{10}^2 + C_{10}^4$. □

Задача 1.7. Пет различими топки се разпределят в три различни кутии A, B и C. Да се намери броят на всички различни разпределения, за които:

(а) кутия A е празна;

(б) само кутия A е празна;
(в) точно една кутия е празна;
(г) поне една кутия е празна;
(д) няма празна кутия.

Отговори. (а) 2^5; (б) 2^5-2; (в) $3(2^5-2)$; (г) $3(2^5-2)+3$; (д) $3^5-(3(2^5-2)+3)$. □

Задача 1.8. Колко е броят на думите с дължина n и съдържащи само символите a, b и c, такива че
(а) започват с a;
(б) съдържат точно k пъти символа a;
(в) съдържат точно k пъти символа a, при което и първият и последният символ е a;
(г) съдържат съответно k_1, k_2 и k_3 пъти ($k_1+k_2+k_3=n$) от символите a, b и c.

Отговори. (а) 3^{n-1}; (б) $\binom{n}{k}2^{n-k}$; (в) $\binom{n-2}{k-2}2^{n-k}$; (г) $\binom{n}{k_1,k_2,k_3}$. □

Задача 1.9. Дадено е множеството $M=\{a_1,\dots,a_m,b_1,\dots,b_n\}$. С A_i (респ. B_i) обозначаваме множеството от комбинации на M без повторение от клас $k\le\min\{m,n\}$, които съдържат точно i обекта от типа a (респ. b), за $i=0,1,\dots,k$.
(а) Да се пресметнат $|A_i|$ и $|B_i|$.
(б) Да се докаже комбинатортно, че $\sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}=\binom{m+n}{k}$ (*формула на Вандермонд*).
(в) Колко са на брой комбинациите от клас k, които съдържат поне един обект от типа a и поне един обект от типа b?
(г) Колко са на брой подмножествата на M, които съдържат поне един обект от типа a и поне един обект от типа b?

Решение. (а) $|A_i|=\binom{m}{i}\binom{n}{k-i}$ и $|B_i|=\binom{m}{k-i}\binom{n}{i}$; (б) множествата $A_0,\dots,A_k$ са две по две непресичащи се и изброяват всички възможни комбинации на M от клас k, т.е. $\bigcup_{i=0}^{k}A_i=\mathbf{C}_{m+n}^{k}$, от което следва, че $C_{m+n}^{k}=\sum_{i=0}^{k}|A_i|$; (в) $C_{m+n}^{k}-|A_k|-|B_k|$; (г)

$$|\mathcal{P}(M)|-|\mathcal{P}(\{a_1,\dots,a_m\})|-|\mathcal{P}(\{b_1,\dots,b_n\})|+1=2^{m+n}-2^m-2^n+1=(2^m-1)(2^n-1),$$

понеже празното множество е добавено веднъж и извадено два пъти. □

Задача 1.10. Десет души се нареждат на редица. Колко са подрежданията, при които три фиксирани лица се намират едно до друго?

Отговор. $8\cdot V_3^3V_7^7$ □

*Задача** 1.11. По колко различни начина от $2n$ шахматиста могат да се образуват $k\le n$ шахматни двойки, ако
(а) цветът на фигурите и номерът на дъските се взимат предвид;
(б) цветът на фигурите се взима предвид, но номерът на дъските няма значение;
(в) цветът на фигурите няма значение, но номерът на дъските се взима предвид;
(г) цветът на фигурите и номерът на дъските нямат значение.

Решение. Нека $M=\{1,\dots,2n\}$. (а) Броим $\{((i_1,i_2),\dots,(i_{2k-1},i_{2k})) : i_j\in M, i_j\ne i_l\}$, т.е. V_{2n}^{2k}. (б) Броим $\{\{(i_1,i_2),\dots,(i_{2k-1},i_{2k})\} : i_j\in M, i_j\ne i_l\}$, т.е. $\frac{1}{k!}V_{2n}^{2k}$. (в) Броим $\{(\{i_1,i_2\},\dots,\{i_{2k-1},i_{2k}\}) : i_j\in M, i_j\ne i_l\}$, т.е. $\prod_{i=0}^{k-1}C_{2n-2i}^2$. (г) Броим $\{\{\{i_1,i_2\},\dots,\{i_{2k-1},i_{2k}\}\} : i_j\in M, i_j\ne i_l\}$, т.е. $\frac{1}{k!}\prod_{i=0}^{k-1}C_{2n-2i}^2$. □

*Задача*** 1.12. Пресметнете броя на траекториите в правоъгълна координатна система, започващи в точка (x_1,y_1) и завършващи в точка (x_2,y_2), ако правим скокове с големина ±1. А колко на брой са траекториите от (x_1,y_1) до (x_2,y_2), които не докосват хоризонталата $y=r$?

Решение (Техника). Нека обозначим с u и d броя на скоковете, които правим, съответно с големина 1 и -1. За да свържем (x_1,y_1) и (x_2,y_2), трябва да извършим общо $u+d=x_2-x_1$ скока, като нетната промяна във височината ще бъде $u-d=y_2-y_1$. Тогава $u=\frac{(x_2-x_1)+(y_2-y_1)}{2}$, от което следва, че проблемът е добре дефиниран тогава и само тогава, когато u е цяло число. В този случай общият брой на траекториите, означен с $N_{x_2-x_1,y_2-y_1}$, е равен на броя на пермутациите на $\{1,-1\}$ с повторение, дължина $u+d$ и честоти (u,d),

$$N_{x_2-x_1,y_2-y_1}=\binom{u+d}{u}=\binom{x_2-x_1}{\frac{(x_2-x_1)+(y_2-y_1)}{2}}.$$

Що се отнася до втория въпрос, нека първо преброим траекториите, които докосват $y = r$. За всяка такава траектория можем да отразим частта до първия контакт с $y = r$ около $y = r$, и обратното, като отражението на (x_1, y_1) ще бъде в точка $(x_1, r + (r - y_1))$. По този начин определяме обратимо еднозначно съответствие между траекториите, които свързват (x_1, y_1) и (x_2, y_2) и докосват $y = r$, и тези от $(x_1, -y_1 + 2r)$ до (x_2, y_2) (т.нар. *reflection principle*). Общият брой на последните е $N_{x_2-x_1, y_2+y_1-2r}$, от което следва, че броят на траекториите, които не докосват $y = r$, е

$$N_{x_2-x_1,y_2-y_1} - N_{x_2-x_1,y_2+y_1-2r} = \binom{x_2 - x_1}{\frac{(x_2-x_1)+(y_2-y_1)}{2}} - \binom{x_2 - x_1}{\frac{(x_2-x_1)+(y_2+y_1)}{2} - r}.$$

□

2 Вероятности

2.1 Елементарни вероятности

Нека $\Omega = \{\omega_1, \omega_2, \dots\}$ е множеството от възможните изходи (елементарни събития) на случаен експеримент, имащ изброимо много на брой изходи. Към всяко елементарно събитие ω_n съпоставяме едно число $\mathbb{P}(\omega_n)$, такова че $\mathbb{P}(\omega_n) \geq 0$ и $\sum_{n=1}^{\infty} \mathbb{P}(\omega_n) = 1$. Тогава на всяко събитие $A \in \mathcal{P}(\Omega)$, да кажем $A = \{\omega_{i_1}, \omega_{i_2}, \dots\}$, съпоставяме неотрицателното число

$$\mathbb{P}(A) = \sum_{m=1}^{\infty} \mathbb{P}(\omega_{i_m}).$$

Горната формула дефинира вероятностна мярка върху $(\Omega, \mathcal{P}(\Omega))$, така че $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ е вероятностно пространство. В частност, за всички $A, B \in \mathcal{P}(\Omega) : A \cap B = \emptyset$, е изпълнено

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Равномерни вероятности. *Нека $\Omega = \{\omega_1, \dots, \omega_n\}$ е крайно множество, чиито елементарни събития са еднакво вероятни, $\mathbb{P}(\omega_i) = p$. Тогава за всяко $A = \{\omega_{i_1}, \dots, \omega_{i_m}\} \in \mathcal{P}(\Omega)$, е в сила*

$$\mathbb{P}(A) = \sum_{j=1}^{m} \mathbb{P}(\omega_{i_j}) = mp = \frac{m}{n} np = \frac{m}{n} = \frac{|A|}{|\Omega|};$$

т.е. $\mathbb{P}(A)$ е съотношението между броя на благоприятните изходи за събитието A и броя на възможните изходи на експеримента.

Забележка (Интуиция). Модели на вероятностни пространства, водещи до равномерни вероятности, се използват, когато елементарните събития се намират в едно и също отношение към условията, дефиниращи характера на случайния експеримент.

....

Задача 2.1. Куб, чиято повърхност е боядисана в червено, е разрязан на 1000 еднакви кубчета, които след това са добре размесени. Каква е вероятността случайно избрано кубче да има точно две червени стени?

Отговор. 0.096. □

Задача 2.2. Да се пресметне вероятността при хвърлянето на четири правилни зара да се падне поне една единица при предположение, че:

(а) заровете са различими и различимите изходи са еднакво вероятни;
(б) заровете са неразличими и различимите изходи са еднакво вероятни.

Описват ли вярно реалната действителност и двете предположения?

Решение. (а) $\Omega = \big\{(a_1, \dots, a_4) : a_i \in \{1, \dots, 6\}\big\}$. $\mathbb{P}(\{\text{поне една 1-ца}\}) = 1 - \frac{\text{изходи без единица}}{\text{всички изходи}} = 1 - \frac{5^4}{6^4} \approx 0.517$; (б) $\Omega = \big\{\{a_1, \dots, a_4\} : a_i \in \{1, \dots, 6\}\big\}$. $\mathbb{P}(\{\text{поне една 1-ца}\}) = 1 - \frac{C(5,4)}{C(6,4)} \approx 0.444$.

Ако и двете предположения, макар и математически издържани, описваха вярно реалността, щеше да се окаже, че вероятността да се падне поне една единица е съществено по-голяма, когато, например, заровете са оцветени различно. □

Забележка (Парадокс на Дьо Мере). Независимо дали заровете са субективно различни или не, те съществуват като различни реални обекти. Всъщност, дори да смятахме, че заровете са неразличими, не бихме приели елементарните събития в (б) за равновероятни, т.е. (б) не описва вярно експеримента във вероятностно отношение (напр., хвърлянето на $\{1,6,6,6\}$ очакваме да е (четири пъти) по-вероятно от $\{6,6,6,6\}$). Тогава е по-удобно да се приложи моделът с различими зарове, тъй като в този модел елементарните изходи са еднакво вероятни и е в сила формулата за равномерните вероятности.

Задача 2.3. Каква е вероятността случайно избрана плочка от домино да съдържа различни числа на двете си половинки?

Решение (Внимание). Нека $\Omega = \big\{\{a_1,a_2\} : a_1,a_2 \in \{0,\dots,6\}\big\}$. Тогава $|\Omega| = C(7,2)$, т.е. играта се играе с 28 плочки. В този случай, за разлика от задачата със заровете, така описаните елементарни събития са равновероятни, откъдето следва, че $\mathbb{P}(\{\{a_1,a_2\} : a_1 \neq a_2\}) = \frac{C_7^2}{C(7,2)} = \frac{3}{4}$. □

Задача 2.4. Урна съдържа M черни и $N-M$ бели топки. Правим случайна ненаредена извадка без връщане с обем $n \le N$. Да се пресметне вероятността извадката да съдържа точно $k \le n$ черни топки, където $k \le M$ и $n-k \le N-M$.

Решение 1. Нека $\Omega = \big\{\{a_1,\dots,a_n\} : a_i \in \{\text{ч}_1,\dots,\text{ч}_M,\text{б}_1,\dots,\text{б}_{N-M}\}\big\}$. Тогава

$$\mathbb{P}(\{k \text{ черни от } n\}) = \frac{\text{брой n-извадки с k черни}}{\text{брой n-извадки от N}} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

Решение 2 (Въведение на УВ).* Нека $\Omega = \big\{(a_1,\dots,a_m) : a_i \in \{\text{ч}_1,\dots,\text{ч}_M,\text{б}_1,\dots,\text{б}_{N-M}\}, m = 1,\dots,n\big\}$. Вероятността да изтеглим черна топка на първи ход е $\mathbb{P}(\text{ч}) = \frac{M}{N}$, а бяла – $\mathbb{P}(\text{б}) = \frac{N-M}{N}$. Вероятността да изтеглим черна топка на първи ход и бяла на втори е $\mathbb{P}(\text{ч},\text{б}) = \mathbb{P}(\text{ч}) \cdot \mathbb{P}(\text{ч},\text{б}|\text{ч}) = \frac{M}{N}\frac{N-M}{N-1}$, и т.н. Тогава за всяка редица от n топки, k от които са черни (общо C_n^k на брой), напр. $\omega^{(k)} = (\text{ч},\text{ч},\text{б},\dots,\text{ч})$, имаме

$$\mathbb{P}(\text{ч},\text{ч},\text{б},\dots,\text{ч}) = \frac{M}{N}\cdot\frac{M-1}{N-1}\cdot\frac{N-M}{N-2}\cdots\frac{M-k+1}{N-n+1} = \frac{M!(N-M)!(N-n)!}{(M-k)!(N-M-(n-k))!N!},$$

което е инвариантно спрямо реда на изтеглените цветове. Следователно,

$$\mathbb{P}(\{k \text{ черни от } n\}) = \mathbb{P}\big(\{\omega_1^{(k)},\dots,\omega_{C_n^k}^{(k)}\}\big) = \binom{n}{k}\frac{M!(N-M)!(N-n)!}{(M-k)!(N-M-(n-k))!N!} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

От двете решения следва, че тегленето на всички n топки наведнъж или една по една е еквивалентно от вероятностна гледна точка. □

Задача 2.5. Ако номерата на колите са равномерно разпределени, каква е вероятността номерът на случайна лека кола

(а) да не съдържа еднакви цифри;
(б) да има точно две еднакви цифри;
(в) да има точно три еднакви цифри;
(г) да има две двойки еднакви цифри;
(д) сумата на първите две цифри да съвпада със сумата на последните две?

Решение. Нека $\Omega = \big\{(a_1,\dots,a_4) : a_i \in \{0,1,\dots,9\}\big\}$. (а) $\frac{V_{10}^4}{10^4}$. (б)

Решение 1: можем да съставим $(a_1,\dots,a_4)$ като изберем множеството $\{a_1^*,a_2^*,a_3^*\}$, за $a_i^* \in \{0,1,\dots,9\}$, $a_1^* \neq a_2^* \neq a_3^*$, и пермутираме елементите му с повторение и честота една от $(1,1,2)/(1,2,1)/(2,1,1)$, в зависимост от това кое a_i^* повтаряме; т.е., търсената вероятност е $\frac{\binom{10}{3}\cdot 3\cdot\binom{4}{2,1,1}}{10^4}$.

Решение 2: избираме на кои две места да повторим някоя цифра, след което избираме вариация от клас 3; т.е., $\frac{\binom{4}{2}\cdot 10\cdot 9\cdot 8}{10^4}$.

(в) $\frac{\binom{10}{2}\cdot 2\cdot\binom{4}{3,1}}{10^4}$ или $\frac{\binom{4}{3}\cdot 10\cdot 9}{10^4}$; (г) $\frac{\binom{10}{2}\cdot 1\cdot\binom{4}{2,2}}{10^4}$ или $\frac{\binom{4}{2}\cdot\binom{10}{2}}{10^4} = \frac{1}{2}\frac{\binom{4}{2}\cdot 10\cdot 9}{10^4}$, като отново избираме на кои две места да повторим една от цифрите, но поради симетрията между двете двойки еднакви цифри броим всички вариации два пъти, което налага деление на две. (д) За (a,b,c,d), двойките $(a,b),(c,d)$ ще приемат различен брой стойности в зависимост от сумата $a+b=c+d$,

сума $a+b$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
варианти за (a,b)	1	2	3	4	5	6	7	8	9	10	9	8	7	6	5	4	3	2	1
варианти за (a,b,c,d)	1	4	9	16	25	36	49	64	81	100	81	64	49	36	25	16	9	4	1

Следователно, отговорът е $\frac{2\cdot\sum_{i=1}^{9} i^2+10^2}{10^4}$. $\square$

Задача 2.6. Картите от (случайно наредено) стандартно тесте се теглят последователно. Играч 1 печели, ако се обърне седмица спатия, а Играч 2, ако се обърнат общо два аса. Каква е вероятността Играч 1 да спечели?

Решение 1. $\Omega = \big\{(a_1,\dots,a_{52}) : a_i \in \{1,\dots,52\}\big\}$. Печеливши за играч 1 са пермутациите, в които имаме следните две подредби на асата и $7C$ измежду елементите на $a_1,\dots,a_{52}$: 7C-A_1-A_2-A_3-A_4 и A_1-7C-A_2-A_3-A_4. Имаме $\binom{52}{5}$ начина да изберем местата на асата и $7C$, на които можем да подредим четирите аса по $4!$ различни начина, след като вече сме фиксирали $7C$. Другите карти подреждаме по $47!$ начина на останалите 47 места, откъдето следва

$$\mathbb{P}(\{\text{Играч 1 печели}\}) = \frac{\text{пермутации с 7C-}A_1\text{-}A_2\text{-}A_3\text{-}A_4}{P_{52}} + \frac{\text{пермутации с }A_1\text{-7C-}A_2\text{-}A_3\text{-}A_4}{P_{52}} = 2\frac{\binom{52}{5}4!47!}{52!}.$$

Решение 2 (Внимание): Можем да се абстрахираме от останалите карти и точната подредбата на асата и да вземем $\Omega = \big\{(7,A,A,A,A),\dots(A,A,A,A,7)\big\}$, при което

$$\mathbb{P}(\{\text{Играч 1 печели}\}) = \frac{\big|\{(7,A,A,A,A),(A,7,A,A,A)\}\big|}{|\Omega|} = \frac{2}{5}.$$

$\square$

*Задача** 2.7. От партида изделия, от които n са доброкачествени и m – бракувани, за проверка по случаен начин са взети s изделия. При проверката се оказало, че първите k от проверяваните s изделия са доброкачествени ($k < s$). Да се пресметне вероятността $(k+1)$-вото изделие да се окаже доброкачествено.

Решение. Абстрахирайки се от първите k и от последните $s-k-1$ стъпки, $\mathbb{P}(\{\text{бракувано}\}) = \frac{m-k}{n+m-k}$. $\square$

Задача 2.8. Хвърлят се 10 различими зара. Каква е вероятността да се паднат равен брой единици и шестици?

Отговор. $\frac{\#\text{1ци}=\#\text{6ци}}{\text{всички вариации}} = \sum_{k=0}^{5}\frac{\#\text{1ци}=\#\text{6ци}=\text{k}}{\text{всички вариации}} = \sum_{k=0}^{5}\frac{\binom{10}{k}\binom{10-k}{k}4^{10-2k}}{6^{10}} \equiv \sum_{k=0}^{5}\frac{\binom{10}{2k}\binom{2k}{k,k}4^{10-2k}}{6^{10}}$. $\square$

Задача 2.9. От урна, съдържаща топки с номера $1,\dots,n$, се вадят последователно k топки ($k \le n$). Каква е вероятността номерата на извадените топки, записани по реда на изваждането, да образуват растяща редица, ако (а) теглим без да връщаме; (б) изтеглените топки се връщат обратно в урната?

Решение (Техника).

(а) Всяка редица $1 \le a_1 < a_2 < \dots < a_k \le n$ еднозначно определя комбинация на $\{1,\dots,n\}$ от k-ти ред без повторение, именно $\{a_1,\dots,a_k\}$, и обратното – елементите на всяка комбинация без повторение образуват една единствена растяща редица. От тук следва, че търсената вероятност е $\frac{C_n^k}{V_n^k} = \frac{1}{k!}$.

(б) Между редиците $a_1 \le a_2 \le \dots \le a_k$ и $a_1^* < a_2^* < \dots < a_k^*$, за $a_i^* = a_i + (i-1)$, $i = 1,\dots,n$, съществува еднозначна обратима връзка, като $a_i^* \in \{1,\dots,n+k-1\}$. От (а) търсената вероятност е $\frac{C_{n+k-1}^k}{V(n,k)}$.

$\square$

Задача 2.10. Около маса сядат 10 мъже и 10 жени. Каква е вероятността лица от еднакъв пол да не седят едно до друго?

Отговор. $\frac{2\cdot 10!\cdot 10!}{20!}$. $\square$

Задача 2.11. Група от n човека се нарежда в редица по случаен начин. Каква е вероятността между две фиксирани лица да има точно $r \le n-2$ човека? А ако се нареждат в кръг?

Решение. $\Omega = \big\{(a_1,\dots,a_n) : a_i \in \{1,\dots,n\}, a_i \ne a_j\big\}$. Нека фиксирани са $x, y \in \{1,\dots,n\} : x \ne y$.

Случай 1: за да получим крайната подредба $(a_1,\dots,a_n)$, можем първо да изберем хората между x и y, $(b_1,\dots,b_r)$ за $b_i \in \{1,\dots,n\}\backslash\{x,y\}$, след което заедно с x и y да ги "вмъкнем" в редицата от останали хора, $(c_1,\dots,c_{n-r-2})$ за $c_i \in \{1,\dots,n\}\backslash\{x,y,b_1,\dots,b_r\}$, като има $2!$ начина как да подредим x и y; т.е. те заедно ще заемат едно от $(n-r-1)$ възможни места, напр. $(c_1,\dots,c_k,y,b_1,\dots,b_r,x,c_{k+1},\dots,c_{n-r-2})$. От тук търсената вероятност е

$$\frac{2!\cdot V_{n-2}^r\cdot P_{n-r-2}\cdot(n-r-1)}{n!} = 2\cdot\frac{n-r-1}{n}\cdot\frac{1}{n-1}.$$

С други думи, първият от двамата ще седне вляво или вдясно от група от r души, като за това има $2(n-r-1)$ възможни места от n. Другият ще седне на точно 1 предварително определено място от общо $n-1$ останали места.

Случай 2: понеже места "n" и "1" са "долепени", в допълнение към пермутациите в случай 1 трябва да разгледаме случаите, когато между x и y има r човека, като вземем предвид хората в двата края на редицата, напр. $(y, \dots, a_{n-1}, a_n, a_1, a_2, \dots x)$. Тях можем да конструираме, като разделим редицата $(y, b_1, \dots, b_r, x)$ на две и след това построим $(b_{k+1}, \dots, b_r, x, c_1, \dots, c_{n-r-2}, y, b_1, \dots, b_k)$ за $k = 0, 1, \dots, r$; т.е. търсената вероятност е

$$\frac{2! \cdot V_{n-2}^r \cdot P_{n-r-2} \cdot (n-r-1)}{n!} + \frac{2! \cdot V_{n-2}^r \cdot P_{n-r-2} \cdot (r+1)}{n!} = \frac{2}{n-1}.$$

Условно, фиксираме позицията на единия човек, като за другия остават 2 възможни места от $n-1$, броейки $r+1$ места по часовниковата стрелка и $r+1$ места обратно на часовниковата стрелка.

□

Задача 2.12 (Occupancy Problem). Секретарка написала n писма, сложила ги в пликове и ги запечатала. Забравила кое писмо в кой плик е, но въпреки това написала отгоре n-те различни адреса и изпратила писмата. Каква е вероятността никой да не получи своето писмо?

Решение (Техника). Нека $A_i = \{$човек i е получил своето писмо$\}$, $i = 1, \dots, n$. Тогава имаме, че $\left(\bigcup_{i=1}^n A_i\right)^c = \{$никой не е получил своето писмо$\}$, $|A_i| = (n-1)!$, $|A_i \cap A_j| = (n-2)!$, $i \neq j$, и т.н., откъдето

$$\mathbb{P}(\{\text{никой не е получил своето писмо}\}) = 1 - \frac{|\bigcup_{i=1}^n A_i|}{n!} = \sum_{k=0}^n (-1)^k \binom{n}{k} \frac{(n-k)!}{n!} = \sum_{k=0}^n (-1)^k \frac{1}{k!}.$$

Проблемът е еквивалентен на този в Задача 1.3в, където n различими частици се разпределят в n клетки, така че нито една клетка не остава празна. □

*Забележка*** (Derangement). Броят на пермутациите на едно множество с n елементи, при които нито един елемент не се появява на първоначалното си място, се нарича n субфакториел и се означава с $!n$. От горната задача следва, че $!n = n! \sum_{k=0}^n (-1)^k \frac{1}{k!}$, откъдето $\lim_{n\to\infty} \frac{!n}{n!} = e^{-1}$.

*Задача** 2.13. Нека r неразличими частици се разпределят по случаен начин в n различими клетки. Всяка клетка може да побере произволен брой частици. Предполагаме, че всички различими разпределения са равновероятни. Да се определи вероятността:

(а) фиксирана клетка да съдържа точно k частици ($k \leq r$);
(б) точно m клетки да са празни ($m < n$);
(в) във всяка клетка да има поне по две частици ($r \geq 2n$);
(г) във всяка клетка да има най-много по четири частици ($r \leq 4n$).

Решение (Техника). (а) $\frac{C(n-1,r-k)}{C(n,r)}$. (б) $\binom{n}{m}\frac{C(n-m,r)}{C(n,r)}$. (в) $\frac{C(n,r-2n)}{C(n,r)}$. (г) Нека $A_i = \{$клетка i съдържа $\geq 5\}$, т.е. $A_i = \big\{\{i,i,i,i,i,a_1,\dots,a_{r-5}\} : a_j \in \{1,\dots,n\}\big\}$. Тогава $\left(\bigcup_{i=1}^n A_i\right)^c = \{$всяка клетка $\leq 4\}$, $|A_i| = C(n, r-5)$, $|A_i \cap A_j| = C(n, r-10)$, $i \neq j$, и т.н., откъдето

$$\mathbb{P}(\{\text{всяка клетка} \leq 4\}) = \frac{C(n,r) - |\bigcup_{i=1}^n A_i|}{C(n,r)} = \frac{\sum_{k=0}^{\lfloor r/5 \rfloor} (-1)^k \binom{n}{k} C(n, r-5k)}{C(n,r)}.$$

□

*Задача** 2.14. Нека r неразличими частици се разпределят по случаен начин в n различими клетки. Всяка клетка може да побере най-много една частица. Предполагаме, че всички различими разпределения са равновероятни. Да се определи вероятността фиксирана клетка да е празна ($r < n$).

Отговор. $\frac{n-r}{n}$. □

*Задача** 2.15. Нека r различими частици се разпределят по случаен начин в n различими клетки. Всяка клетка може да побере произволен брой частици. Предполагаме, че всички различими разпределения са равновероятни. Да се определи вероятността:

(а) първата клетка да съдържа k_1 частици, втората – k_2 частици и т.н, където $k_1 + \dots + k_n = r$;
(б) при $n = r$ нито една клетка да не остане празна;
(в) при $n = r$ да остане празна точно една клетка;
(г) точно m клетки да са празни ($m < n, r \geq n - m$).

Решение. (а) $\frac{\binom{r}{k_1,\ldots,k_n}}{n^r}$. (б) $\frac{n!}{n^n}$. (в) $\frac{n\binom{n}{2}(n-1)!}{n^n}$, като в една от останалите клетки ще има две топки. (г) В останалите $n-m$ клетки, избрани по $\binom{n}{n-m}$ начина, ще има съответно $k_1,\ldots,k_{n-m}$ топки, откъдето

$$\mathbb{P}(\{\text{точно m клетки са празни}\}) = \frac{\binom{n}{n-m}\sum_{k_1,\ldots,k_{n-m}\in\mathbb{N}:k_1+\cdots+k_{n-m}=r}\binom{r}{k_1,\ldots,k_{n-m}}}{n^n} = \frac{\binom{n}{n-m}\sum_{i=0}^{n-m}(-1)^i\binom{n-m}{i}(n-m-i)^r}{n^n},$$

като използваме, че $\sum_{k_1,\ldots,k_{n-m}\in\mathbb{N}_0:k_1+\cdots+k_{n-m}=r}\binom{r}{k_1,\ldots,k_{n-m}} = (n-m)^r$ от забележките след дефиницията на $V(n,k)$, и принципа за включване-изключване, за да извадим случаите, в които поне едно $k_i = 0$. □

*Задача** 2.16. Хвърлят се n зара. Да се пресметне вероятността сумата от падналите се точки да бъде равна на: а) n; б) $n+1$; в) дадено число s.

Решение. $\Omega = \{(a_1,\ldots,a_n) : a_i \in \{1,\ldots,6\}\}$, $|\Omega| = 6^n$. (а) $\frac{1}{6^n}$. (б) $\frac{n}{6^n}$. (в)

Решение 1: търсим вариациите с n_1 единици,..., n_6 шестици, такива че $n_1+2n_2+\cdots+6n_6 = s$, т.е.

$$\mathbb{P}(\{a_1+\cdots+a_n = s\}) = \frac{1}{6^n}\sum_{n_1,\ldots,n_6\in\mathbb{N}_0:n_1+2n_2+\cdots+6n_6=s}\binom{n}{n_1,\ldots,n_6}.$$

Решение 2: броят на търсените вариации е равен на броя на решенията на $x_1+\cdots+x_n = s$ за $x_i \in \mathbb{N} : x_i \le 6$. Нека $A_i = \{\text{решения с } x_i \ge 7\}$, така че $\left(\bigcup_{i=1}^n A_k\right)^c = \{\text{решения с } x_1 \le 6,\ldots,x_n \le 6\}$. Тогава $|A_i| = C(n, s-n-6)$, $|A_i\cap A_j| = C(n, s-n-12)$, $i \ne j$, и т.н., откъдето

$$\mathbb{P}(\{a_1+\cdots+a_n = s\}) = \frac{C(n,s-n) - |\bigcup_{i=1}^n A_i|}{6^n} = \frac{\sum_{k=0}^{\lfloor (s-n)/6\rfloor}(-1)^k\binom{n}{k}C(n,s-n-6k)}{6^n}.$$

□

*Задача** 2.17. От n чифта обувки случайно се избират $2r$ обувки ($2r < n$). Да се пресметне вероятността измежду избраните обувки:

(а) да няма нито един чифт;
(б) да има точно един чифт;
(в) да има точно два чифта.

Решение. $\Omega = \{(a_1,\ldots,a_{2r}) : a_i \in \{\text{л}_1,\ldots,\text{л}_n,\text{д}_1,\ldots,\text{д}_n\}, a_i \ne a_j\}$, $|\Omega| = V_{2n}^{2r}$.

(а) $\frac{2n(2n-2)(2n-4)\cdots(2n-4r+2)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = \frac{2^{2r}\cdot V_n^{2r-1}}{V_{2n}^{2r}} = \frac{2^{2r}\cdot\binom{n}{2r}}{\binom{2n}{2r}}$, т.е. без значение подредбата, избираме $2r$ чифта по $\binom{n}{2r}$ начина, след което решаваме да вземем дясна или лява обувка от всеки избран чифт.

(б) $\frac{\binom{2n}{2}\cdot 2\cdot n\cdot(2n-2)\cdot(2n-4)\cdots(2n-4r+4)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = \frac{n\cdot 2^{2r-2}\cdot\binom{n-1}{2r-2}}{\binom{2n}{2r}}$, където $\binom{2n}{2}$ са местата, на които по 2 начина поставяме лявата и дясната обувка от един и същи чифт, избран от всички n чифта обувки, т.е. $\binom{2n}{2}\cdot 2\cdot n$ възможни такива случаи.

(в) $\frac{\binom{2r}{2}\cdot 2\cdot n\cdot\binom{2r-2}{2}\cdot 2\cdot(n-1)\cdot(2n-4)\cdot(2n-6)\cdots(2n-4r+6)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = \frac{\binom{n}{2}\cdot 2^{2r-4}\cdot\binom{n-2}{2r-4}}{\binom{2n}{2r}}$.

□

*Задача*** 2.18 (Bertrand's Ballot Problem). По време на избори за първия от двама кандидати са пуснати n бюлетини, а за втория – m бюлетини. Каква е вероятността при преброяване на бюлетините броят на преброените гласове, подадени за първия кандидат, да бъде по-голям от броя на гласовете, подадени за втория кандидат, през цялото време?

Решение (Техника). За всяко $i = 1,\ldots,n+m$, полагаме $x_i = 1$, ако i-тата бюлетина е за кандидат 1, и $x_i = -1$, ако е за кандидат 2. Задачата е еквивалентна на това броят на нетните гласове $s_1 > 0, s_2 > 0\ldots, s_{n+m} = n-m > 0$ да бъде постоянно положителен, където $s_i = x_1+\cdots+x_i$. Геометрично погледнато, последното предполага, че начупената линията, свързваща точките $(0,0),(1,s_1),(2,s_2),\ldots,(n+m,n-m)$, не докосва абсцисната ос. Нещо повече, тъй като кандидат 1 е винаги начело, имаме $s_1 = x_1 = 1$ и траекториите, които удовлетворяват условието, са сред всички свързващи $(1,1)$ и $(n+m,n-m)$. Тогава от Задача 1.12 получаваме

$$\mathbb{P}(\{\text{кандидат 1 е винаги начело}\}) = \frac{\text{от } (1,1) \text{ до } (n+m,n-m) \text{ без да докосва}}{\text{всички от } (0,0) \text{ до } (n+m,n-m)} = \frac{\binom{n+m-1}{n-1}-\binom{n+m-1}{n}}{\binom{n+m}{n}} = \frac{n-m}{n+m}.$$

□

2.2 Условни вероятности и независимост на събития

Условна вероятност. *Нека* $(\Omega, \mathcal{A}, \mathbb{P})$ *е вероятностно пространство (в. п.) и* $B \in \mathcal{A} : \mathbb{P}(B) > 0$. *Условната вероятност на събитието* $A \in \mathcal{A}$ *при условие* B *дефинираме чрез количеството*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Ако в допълнение $\mathbb{P}(A) > 0$, *то* $\mathbb{P}(A|B) = \mathbb{P}(B|A)\frac{\mathbb{P}(A)}{\mathbb{P}(B)}$ ***(формула на Бейс)*** *или*

$$\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B) = \mathbb{P}(A)\mathbb{P}(B|A).$$

По индукция, за $A_1, \dots, A_n \in \mathcal{A} : \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$, *е в сила*

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}).$$

*Забележка** (Интерпретация). Ако условията на случайния експеримент са се променили така, че възможните изходи ω лежат изцяло в B, тогава има смисъл да ограничим оригиналното в. п. $(\Omega, \mathcal{A}, \mathbb{P})$ и да работим върху $\big(B, \mathcal{A} \cap B, \mathbb{P}(\cdot|B)\big)$, където $\mathbb{P}(\cdot|B) : \mathcal{A} \cap B \to [0, 1]$ е добре дефинирана вероятностна мярка, за разлика от $\mathbb{P}(\cdot \cap B)$. В частност е вярно за $\mathbb{P}(\cdot|B)$, че $\mathbb{P}(B|B) = 1$,

$$\mathbb{P}(A^c|B) = 1 - \mathbb{P}(A|B),$$

и, за $A_1, A_2 \in \mathcal{A} : A_1 \cap A_2 = \emptyset$,

$$\mathbb{P}(A_1 \cap A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B).$$

Независимост. *Събитията* $A_1, \dots, A_n \in \mathcal{A}$ *се наричат независими в съвкупност, ако* <u>*за кои да е*</u> $i_1, \dots, i_k \in \{1, \dots, n\}$ *при* $k = 2, \dots, n$, *имаме*

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_n}).$$

Ако $n = 2$ *и* $\mathbb{P}(A_2) > 0$, *следва, че* A_1 *и* A_2 *са независими, означено с* $A_1 \perp\!\!\!\perp A_2$, *тогава и само тогава, когато*

$$\mathbb{P}(A_1|A_2) = \mathbb{P}(A_1).$$

Твърдение 2.1. *Ако* $A \perp\!\!\!\perp B$, *то* $A \perp\!\!\!\perp B^c$, $A^c \perp\!\!\!\perp B$ *и* $A^c \perp\!\!\!\perp B^c$.

Д-во. $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c)$. Останалите твърдение се доказват по аналогичен начин. □

Задача 2.19. Застрахователна компания води статистика за своите клиенти. Известно е, че 1) всичи клиенти посещават лекар поне веднъж годишно; 2) 60% посещават лекар повече от веднъж годишно; 3) 17% посещават хирург; 4) 15% от тези, които посещават лекар повече от веднъж годишно, посещават хирург. Каква е вероятността случайно избран клиент, който посещава лекар само веднъж годишно, да не е бил при хирург?

Решение. Нека $A = \{\text{посещава хирург}\}$ и $B = \{\text{на лекар } > 1 \text{ годишно}\}$. От условието, $\mathbb{P}(A) = 0.17$, $\mathbb{P}(B) = 0.6$ и $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = 0.09$, откъдето $\mathbb{P}(A^c|B^c) = \frac{\mathbb{P}((A \cup B)^c)}{\mathbb{P}(B^c)} = \frac{1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B)}{1 - \mathbb{P}(B)} = 0.80$. □

Задача 2.20. Хвърлят се два зара. Разглеждаме събитията $A_1 = \{\text{на първия зар се пада нечетно число}\}$, $A_2 = \{\text{на втория зар се пада нечетно число}\}$ и $A_3 = \{\text{сумата от падналите се точки е нечетна}\}$. Независими ли са тези събития две по две? Независими ли са в съвкупност?

Решение. $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \frac{1}{2}$, $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = \frac{1}{4}$, $\mathbb{P}(A_1 \cap A_2 \cap A_3) = 0$. □

Задача 2.21. При всеки опит едно събитие настъпва с вероятност p. Опитите се провеждат последователно до настъпване на събитието. Да се пресметне вероятността събитието да настъпи точно на $(k + 1)$-вия опит.

Отговор. $(1 - p)^k p$. □

Задача 2.22. Вероятността за излизане на строя на k-тия блок на една машина е равна на p_k, $k = 1, \dots, n$. Да се пресметне вероятността за излизане от строя поне на един от n-те блока на машината, ако работата на всички блокове е взаимно независима.

Отговор. $1 - \prod_{k=1}^{n}(1 - p_k)$. □

Задача 2.23. Трима ловци едновременно стрелят по заек. Заекът е убит от един куршум. Каква е вероятността той да е изстрелян от първия ловец, ако те уцелват с вероятност, съответно p_1, p_2 и p_3?

Решение. Нека $A = \{\text{убит от един куршум}\}$ и $H_i = \{\text{ловец i уцелва}\}$, $i = 1, 2, 3$. Тогава $A = (H_1 \cap H_2^c \cap H_3^c) \cup (H_1^c \cap H_2 \cap H_3^c) \cup (H_1^c \cap H_2^c \cap H_3)$, откъдето $\mathbb{P}(H_1|A) = \frac{\mathbb{P}(A \cap H_1)}{\mathbb{P}(A)} = \frac{p_1(1-p_2)(1-p_3)}{p_1(1-p_2)(1-p_3)+(1-p_1)p_2(1-p_3)+(1-p_1)(1-p_2)p_3}$. □

Задача 2.24 (Birthday Paradox). Какъв е най-малкият брой хора, които трябва да се изберат по случаен начин така, че вероятността поне двама от тях да имат един и същ рожден ден да е по голяма от 1/2?

Отговор. $\min\{n \in \mathbb{N} : 1 - \frac{V_{365}^n}{365^n} > \frac{1}{2}\} = 23$. □

Задача 2.25 (Boy or Girl Paradox). X има две деца. Ако по-старото е момиче, каква е вероятността и двете да са момичета? А ако знаете, че поне едно от тях е момиче, каква е вероятността и двете да са момичета?

Решение 1. Нека $A = \{\text{по-малкото е момиче}\}$ и $B = \{\text{по-старото е момиче}\}$. Предполага се, че $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$ и $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = \frac{1}{4}$. Тогава $\mathbb{P}(A|B) = \mathbb{P}(A) = \frac{1}{2}$ и

$$\mathbb{P}(A \cap B | A \cup B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cup B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Решение 2: Нека $\Omega = \big\{(a_1, a_2) : a_1, a_2 \in \{m, f\}\big\}$, където елементарните събития са равновероятни. Тогава $\mathbb{P}(\{\text{две момичета}\}) = \mathbb{P}((f, f)) = \frac{1}{4}$, $\mathbb{P}(\{\text{поне едно момиче}\}) = \mathbb{P}(\{(m, f), (f, m), (f, f)\}) = \frac{3}{4}$ и

$$\mathbb{P}\big(\{\text{две момичета}\}\big|\{\text{поне едно момиче}\}\big) = \frac{1}{3}.$$

□

*Задача** 2.26. Кутия съдържа n билета, от които $m \le n$ печелят, а останалите билети губят. Всеки от n играчи на свой ред избира по един билет. Какви са шансовете за печалба на всеки от играчите? Кога е по-изгодно да се изтегли билет?

Решение. Нека $A_k^0 = \{k\text{-тия играч тегли печеливш билет}\}$ и $A_k^1 = (A_k^0)^c$, $k = 1, \dots, n$. Тогава

$$\begin{aligned}
\mathbb{P}(A_k^0) &= \sum_{i_1,\dots,i_{k-1} \in \{0,1\}} \mathbb{P}(A_1^{i_1} \cap \dots \cap A_{k-1}^{i_{k-1}} \cap A_k^0) = \sum_{i_1,\dots,i_{k-1} \in \{0,1\}} \mathbb{P}(A_1^{i_1})\mathbb{P}(A_2^{i_2}|A_1^{i_1}) \cdots \mathbb{P}(A_k^0|A_1^{i_1} \cap \dots \cap A_{k-1}^{i_{k-1}}) \\
&= \sum_{i=\max\{0,k-(n-m)\}}^{\min\{m-1,k-1\}} \binom{k-1}{i} \frac{m \cdots (m-i+1) \cdot (m-i) \cdot (n-m) \cdots \big((n-m)-(k-i-1)+1\big)}{n \cdots (n-k+1)} \\
&= \frac{m}{n} \sum_{i=\max\{0,k-(n-m)\}}^{\min\{m-1,k-1\}} \frac{\binom{m-1}{i}\binom{(n-1)-(m-1)}{k-i-1}}{\binom{n-1}{k-1}} = \frac{m}{n},
\end{aligned}$$

където $\sum_{i=\max\{0,k-(n-m)\}}^{\min\{m-1,k-1\}} \binom{m-1}{i}\binom{(n-1)-(m-1)}{k-i-1} = \binom{n-1}{k-1}$ изброява начините, по които можем да изберем комбинация от $(k-1)$-ви клас на множество от $n-1$ елементи, състоящо се от $m-1$ печеливши билети и $(n-1)-(m-1)$ непечеливши такива (виж Задача 1.9). □

*Задача** 2.27. Двама души играят до победа, като за това е необходимо първият да спечели m партии, а вторият – n партии. Първият играч може да спечели всяка отделна партия с вероятност p, а вторият – с вероятност $1-p$. Да се пресметне вероятността първият играч да спечели цялата игра.

Решение.

$$\mathbb{P}(\{\text{I печели}\}) = \sum_{k=0}^{n-1} \mathbb{P}\big(\{\text{I печели от } m+k \text{ партии}\}\big) = \sum_{k=0}^{n-1} \binom{m+k-1}{k} p^m (1-p)^k,$$

където, за да спечели I от $m+k$ изиграни партии, означава, че I е спечелил последната партия и II е спечелил $k \le n-1$ партии от предходните $m+k-1$ партии. □

Формула за пълната вероятност. *Нека $\{H_n\}_{n=1}^{\infty} \subseteq \mathcal{A}$ образуват* ***пълна група от събития****, т.е. $\bigcup_{n=1}^{\infty} H_n = \Omega$, $H_n \cap H_m = \emptyset$, $n \neq m$, $n, m \in \mathbb{N}$. Допускаме, че $\mathbb{P}(H_n) > 0$, $n \in \mathbb{N}$. Тогава, за произволно събитие $A \in \mathcal{A}$,*

$$\mathbb{P}(A) = \mathbb{P}\Big(A \cap \bigcup_{n=1}^{\infty} H_n\Big) = \sum_{n=1}^{\infty} \mathbb{P}(A \cap H_n) = \sum_{n=1}^{\infty} \mathbb{P}(A|H_n)\mathbb{P}(H_n).$$

Ако в допълнение е изпълнено, че $\mathbb{P}(A) > 0$, тогава за всяко n,

$$\mathbb{P}(H_n|A) = \frac{\mathbb{P}(A|H_n)\mathbb{P}(H_n)}{\sum_{m=1}^{\infty} \mathbb{P}(A|H_m)\mathbb{P}(H_m)}.$$

*Забележка** (ФПВ за УВ). За $B, C \in \mathcal{A} : \mathbb{P}(B), \mathbb{P}(C) > 0$ е изпълнено

$$\mathbb{P}(A|B) = \mathbb{P}(A|B, C)\mathbb{P}(C|B) + \mathbb{P}(A|B, C^c)\mathbb{P}(C^c|B).$$

Задача 2.28. Две урни съдържат съответно $б_1$ и $б_2$ бели топки и $ч_1$ и $ч_2$ черни топки. От всяка урна случайно се изважда по една топка, а след това от тези две топки случайно се избира едната. Каква е вероятността тази топка да е бяла?

Решение. Нека $A = \{$бяла топка$\}$ и $H_{ij} = \{$цветът от първата урна е i, а от втората j$\}$, $i, j \in \{$б,ч$\}$. Тогава $\mathbb{P}(H_{ij}) = \frac{i_1 j_2}{(б_1+ч_1)(б_2+ч_2)}$, $\mathbb{P}(A|H_{бб}) = 1$, $\mathbb{P}(A|H_{бч}) = \mathbb{P}(A|H_{чб}) = \frac{1}{2}$, $\mathbb{P}(A|H_{чч}) = 0$, откъдето

$$\mathbb{P}(A) = \mathbb{P}(A|H_{бб})\mathbb{P}(H_{бб}) + \mathbb{P}(A|H_{чб})\mathbb{P}(H_{чб}) + \mathbb{P}(A|H_{бч})\mathbb{P}(H_{бч}) + \mathbb{P}(A|H_{чч})\mathbb{P}(H_{чч}) = \frac{2б_1б_2 + б_1ч_2 + ч_1б_2}{2(б_1 + ч_1)(б_2 + ч_2)}.$$

□

Задача 2.29. Разполагаме с три стандартни зара и един, чиито страни са само шетици. По случаен начин избираме три от заровете и ги хвърляме. Да се определи вероятността да се паднат а) три шестици; б) различни цифри; в) последователни цифри?

Отговори. а) $\mathbb{P}(6{,}6{,}6) = \frac{1}{4}\frac{1}{6^3} + \frac{3}{4}\frac{1}{6^2}$; б) $\mathbb{P}(\text{различни}) = \frac{1}{4}\frac{V_6^3}{6^3} + \frac{3}{4}\frac{V_5^2}{6^2} = \frac{V_5^2}{6^2}$; в) $\mathbb{P}(\text{последователни}) = \frac{1}{4}\frac{4 \cdot 3!}{6^3} + \frac{3}{4}\frac{2}{6^2}$. □

Задача 2.30. От урна, която съдържа топки с номера от 1 до n, последователно се изваждат две топки, като първата се връща, ако номерът ѝ не е равен на 1. Да се пресметне вероятността топката с номер 2 да бъде извадена при второто теглене.

Отговор. $\frac{1}{n-1} \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{n-1}{n}$. □

Задача 2.31. В кутия има 7 топки за тенис, от които 4 са нови. За първата игра по случаен начин се избират три топки, които след играта се връщат обратно в кутията. За втората игра също случайно се избират три топки. Каква е вероятността те да са нови?

Отговор. $\mathbb{P}(\{\text{втора игра с 3 нови}\}) = \frac{\binom{4}{3}}{\binom{7}{3}} \cdot \frac{\binom{3}{3}}{\binom{7}{3}} + \frac{\binom{3}{3}}{\binom{7}{3}} \cdot \frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}}$. □

Задача 2.32. Всяка от k_1 на брой урни съдържа по $б_1$ бели и $ч_1$ черни топки, а всяка от k_2 на брой урни – по $б_2$ бели и $ч_2$ черни топки. От случайно избрана урна е била изтеглена топка, която се оказва бяла. Каква е вероятността тази топка да е изтеглена от първата група урни?

Решение. Нека $A = \{$бяла топка$\}$ и $H_i = \{$топката е от i-тата група урни$\}$, $i \in \{1, 2\}$. Тогава $H_2 = H_1^c$, $\mathbb{P}(H_i) = \frac{k_i}{k_1+k_2}$ и $\mathbb{P}(A|H_i) = \frac{б_i}{б_i+ч_i}$, откъдето

$$\mathbb{P}(H_1|A) = \frac{P(A|H_1)\mathbb{P}(H_1)}{P(A|H_1)\mathbb{P}(H_1) + P(A|H_2)\mathbb{P}(H_2)} = \frac{\frac{k_1}{k_1+k_2}\frac{б_1}{б_1+ч_1}}{\frac{k_1}{k_1+k_2}\frac{б_1}{б_1+ч_1} + \frac{k_2}{k_1+k_2}\frac{б_2}{б_2+ч_2}} = \Big[1 + \frac{k_2б_2(б_1 + ч_1)}{k_1б_1(б_2 + ч_2)}\Big]^{-1}.$$

□

Задача 2.33. Разполагаме с тест за рядко заболяване, който е точен в 99% от случаите и при заразените (когато трябва да е положителен), и при незаразените (когато трябва да е отрицателен). Ако знаете че 0.5% от населението има това заболяване, какв е вероятността случайно избран човек с положителен тест да е болен?

Отговор. $\approx 34\%$. $\square$

Задача 2.34. На изпит се явяват 100 студенти, 55 момчета и 45 момичета. Момичетата взимат изпита с вероятност p_1, а момчетата – с p_2. След изпита се избират три резултата. Два от тях се оказали успешни, а един неуспешен. Каква е вероятността и трите резултата да са на момичета?

Решение. Нека $A = \{2 \text{ успешни и } 1 \text{ неуспешен}\}$ и $H_i = \{\text{избрани i резултата на момичета}\}$, $i = 0,1,2,3$. Тогава H_0, H_1, H_2, H_3 образуват пълна група и $\mathbb{P}(H_i) = \frac{\binom{45}{i}\binom{55}{3-i}}{\binom{100}{3}}$. Отделно $\mathbb{P}(A|H_0) = 3 \cdot p_2^2(1-p_2)$, $\mathbb{P}(A|H_1) = p_2^2(1-p_1) + 2 \cdot p_1 p_2(1-p_2)$, $\mathbb{P}(A|H_2) = p_1^2(1-p_2) + 2 \cdot p_1 p_2 (1-p_1)$, $\mathbb{P}(A|H_3) = 3 \cdot p_1^2(1-p_1)$ и решението следва от формулата $\mathbb{P}(H_3|A) = \frac{\mathbb{P}(A|H_3)\mathbb{P}(H_3)}{\sum_{j=0}^{3}\mathbb{P}(A|H_j)\mathbb{P}(H_j)}$. $\square$

Задача 2.35 (Bertrand's Box Paradox). Дадени са три жетона. Първият има две бели страни, вторият две черни, а третият една бяла и една черна страна. По случаен начин се избира жетон и се хвърля върху маса. Ако горната страна на жетона е бяла, каква е вероятността другата му страна, която не се вижда, също да е бяла?

Решение. Нека $A = \{\text{видяна е бяла страна}\}$ и $H_i = \{\text{монетата има i бели страни}\}$, $i = 0, 1, 2$. Тогава H_0, H_1, H_2 образуват пълна група, $\mathbb{P}(H_i) = \frac{1}{3}$ и $\mathbb{P}(A|H_i) = \frac{i}{2}$, откъдето

$$\mathbb{P}(H_2|A) = \frac{\mathbb{P}(A|H_2)\mathbb{P}(H_2)}{\mathbb{P}(A|H_0)\mathbb{P}(H_0) + \mathbb{P}(A|H_1)\mathbb{P}(H_1) + \mathbb{P}(A|H_2)\mathbb{P}(H_2)} = \frac{1 \cdot \frac{1}{3}}{0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{2}{3}.$$

$\square$

*Задача** 2.36 (Monty Hall Problem). Зад една от три затворени врати има чисто нова кола, а зад другите две няма нищо. Избирате врата, след което водещият отваря една от останалите две врати, зад които няма нищо. Сега трябва да решите – сменяте ли избраната врата или запазвате първоначалния си избор?

Решение. Нека сме избрали врата 1, при което водещият, избирайки между врати 2 и 3, отваря врата 3, зад която няма нищо. Дефинираме $A = \{\text{отваря врата } 3\}$ и $H_i = \{\text{врата i печели}\}$, $i = 1, 2, 3$. Знаем, че $\mathbb{P}(H_i) = \frac{1}{3}$. По условие $\mathbb{P}(A|H_2) = 1$ и $\mathbb{P}(A|H_3) = 0$. Нека обозначим с $\mathbb{P}(A|H_1) = p$ за $p \in [0, 1]$. Тогава

$$\mathbb{P}(H_1|A) = \frac{\mathbb{P}(A|H_1)\mathbb{P}(H_1)}{\mathbb{P}(A|H_1)\mathbb{P}(H_1) + \mathbb{P}(A|H_2)\mathbb{P}(H_2) + \mathbb{P}(A|H_3)\mathbb{P}(H_3)} = \frac{p\frac{1}{3}}{p\frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} = \frac{p}{p+1}.$$

От тук следва, че $\mathbb{P}(H_1|A) = 0$ за $p = 0$, $\mathbb{P}(H_1|A) = \frac{1}{3}$ за $p = \frac{1}{2}$, $\mathbb{P}(H_1|A) = \frac{1}{2}$ за $p = 1$. $\square$

*Задача** 2.37. Всички изделия в едната от две партиди са доброкачествени, а в другата $1/4$ от изделията са бракувани. Изделие, взето от случайно избрана партида, се оказало доброкачествено, след което е върнато обратно в своята партида. Да се пресметне вероятността второто случайно избрано изделие от същата партида да се окаже бракувано.

Решение. Нека $A = \{\text{2-ро бракувано}\}$, $B = \{\text{1-во добро}\}$ и $C = \{\text{от I-ва партида}\}$. Тогава

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B|C)\mathbb{P}(C) + \mathbb{P}(A \cap B|C^c)\mathbb{P}(C^c)}{\mathbb{P}(B|C)\mathbb{P}(C) + \mathbb{P}(B|C^c)\mathbb{P}(C^c)} = \frac{0 \cdot 1 \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + \frac{3}{4}\frac{1}{2}} = \frac{3}{28}.$$

$\square$

Задача 2.38. В урна има n бели, m зелени и l червени топки, които се изваждат по случаен начин една след друга: (а) без връщане; (б) с връщане. В двата случая да се пресметне вероятността бяла топка да бъде извадена преди зелена топка.

Решение (Техника). Нека $H_0 = \{\text{изтеглени само червени}\}$, $H_k^{\text{б}} = \{\text{бяла преди зелена на стъпка } k\}$ и $H_k^{\text{з}} = \{\text{зелена преди бяла на стъпка } k\}$, за $k = 1, 2, \ldots$, и $A = \{\text{бяла преди зелена}\}$, като в (б) играта може да продължи, докато в урната не остане нито една червена топка $+$ 1 последен ход.

(а) $H_0, H_1^{б}, H_1^{з}, \ldots$ образуват пълна група от събития върху $\Omega = \{б, з, чб, чз, ччб, ччз, \ldots, чччч\cdots\}$. Тогава $A \cap H_0 = \emptyset$ и $A \cap H_k^{з} = \emptyset$, за всяко $k = 1, 2, \ldots$, откъдето

$$\mathbb{P}(A) = \sum_{k=1}^{\infty} \mathbb{P}(A \cap H_k^{б}) = \sum_{k=1}^{\infty} \mathbb{P}(H_k^{б}) = \frac{n}{n+m+l} \sum_{k=0}^{\infty} \Big(\frac{l}{n+m+l}\Big)^k = \frac{n}{n+m+l} \cdot \frac{1}{1 - \frac{l}{n+m+l}} = \frac{n}{n+m}.$$

От друга страна, ако на първия ход няма победител, т.е. е настъпило събитието $G := (H_1^{б} \cup H_1^{з})^c$, то играта започва отначало от вероятностна гледна точка. Това може да бъде изведено от

$$\mathbb{P}(A) = \mathbb{P}(A \cap H_1^{б}) + \mathbb{P}(A \cap H_1^{з}) + \mathbb{P}(A \cap G) = \frac{n}{n+m+l} + \frac{l}{n+m+l}\mathbb{P}(A|G),$$

като $\mathbb{P}(A|G) = \sum_{k=1}^{\infty} \mathbb{P}(H_k^{б}|G)$. Но $\mathbb{P}(H_1^{б}|G) = 0$, $\mathbb{P}(H_2^{б}|G) = \mathbb{P}(H_1^{б})$, $\mathbb{P}(H_3^{б}|G) = \mathbb{P}(H_2^{б})$, т.н., откъдето $\mathbb{P}(A|G) = \sum_{k=1}^{\infty} \mathbb{P}(H_k^{б}) \equiv \mathbb{P}(A)$. Следователно $\mathbb{P}(A) = \frac{n}{n+m+l} + \frac{l}{n+m+l}\mathbb{P}(A)$ и $\mathbb{P}(A) = \frac{n}{n+m}$.

(б) В този случай $\Omega = \{б, з, чб, чз, ччб, ччз, \ldots, ччч\cdots чб, ччч\cdots чз\}$ и $H_1^{б}, H_1^{з}, \ldots, H_{l+1}^{б}, H_{l+1}^{з}$ образуват пълна група от събития, като $H_0 = \emptyset$ и $A^c = \{$зелена преди бяла$\}$. Тогава

$$\mathbb{P}(A) = \sum_{k=1}^{l+1} \mathbb{P}(A \cap H_k^{б}) = \sum_{k=1}^{l+1} \mathbb{P}(H_k^{б}) = \frac{n}{n+m+l}\Big[1 + \sum_{k=1}^{l} \frac{l\cdots(l-k+1)}{(n+m+l-1)\cdots(n+m+l-k)}\Big].$$

От друга страна $\mathbb{P}(A^c) = \frac{m}{n+m+l}\big[1 + \sum_{k=1}^{l} \frac{l\cdots(l-k+1)}{(n+m+l-1)\cdots(n+m+l-k)}\big]$, откъдето намираме, че $\frac{\mathbb{P}(A)}{\mathbb{P}(A^c)} = \frac{n}{m}$ и, съответно, $\mathbb{P}(A) = \frac{n}{n+m}$ и $\mathbb{P}(A^c) = \frac{m}{n+m}$.

□

Забележка.* Вероятността да няма победител в (а) е $\mathbb{P}(H_0) = \mathbb{P}(ччччч\cdots) = \lim_{n\to\infty}\big(\frac{l}{n+m+l}\big)^n = 0$. Вероятността да има победител е $\mathbb{P}(H_0^c) = \frac{n+m}{n+m+l}\sum_{n=0}^{\infty}(\frac{l}{n+m+l})^n = 1$.

Задача 2.39. Двама играчи последователно хвърлят монета. Играта печели този, който пръв хвърли ези. Каква е вероятността за спечелване на играта за всеки от играчите? А ако печели този, който хвърли същото като падналото се непосредствено преди това?

Решение (Техника). Нека $A = \{$I печели$\}$, $H_0 = \{$никой не печели$\}$, $H_{2n-1} = \{I$ печели на (2n-1)-ти опит$\}$ и $H_{2n} = \{II$ печели на (2n)-ти опит$\}$, за $n = 1, 2, \ldots$. Очевидно е, че $\mathbb{P}(H_0) = 0$, откъдето $\mathbb{P}(\{II \text{ печели}\}) = 1 - \mathbb{P}(A)$. От друга страна, $H_0, H_1, H_2, \ldots$ образуват пълна група върху $\Omega = \{E, TE, TTE, \ldots, TTTTTTT\cdots\}$. Тогава

$$\mathbb{P}(A) = \sum_{n=0}^{\infty} \mathbb{P}(A \cap H_n) = \sum_{n=1}^{\infty} \mathbb{P}(H_{2n-1}) = \mathbb{P}(E) + \mathbb{P}(TTE) + \mathbb{P}(TTTTE) + \cdots = \frac{1}{2}\sum_{n=0}^{\infty}\Big(\frac{1}{4}\Big)^n = \frac{1}{2}\frac{1}{1-\frac{1}{4}} = \frac{2}{3}.$$

Рекурсивно решение: Използвайки същата логика като в Задача 2.38, имаме

$$\mathbb{P}(A) = \mathbb{P}(A \cap H_1) + \mathbb{P}(A \cap H_2) + \mathbb{P}(A|(H_1 \cup H_2)^c)\mathbb{P}\big((H_1 \cup H_2)^c\big) = \frac{1}{2} + \frac{1}{4}\mathbb{P}(A) \quad \Longrightarrow \quad \mathbb{P}(A) = \frac{2}{3}.$$

Що се отнася до втория въпрос, ролите на играчите са един вид разменени, като

$$\mathbb{P}(A) = \mathbb{P}(ETT) + \mathbb{P}(TEE) + \mathbb{P}(ETETT) + \mathbb{P}(TETEE) + \cdots = 2 \cdot \frac{1}{2}\sum_{n=1}^{\infty}\Big(\frac{1}{4}\Big)^n = \frac{1}{1-\frac{1}{4}} - 1 = \frac{1}{3}.$$

□

Задача 2.40. Всяка от N урни съдържа по m бели и n черни топки. От първата урна случайно се избира една топка и се прехвърля във втората. След това от втората урна случайно се избира една топка и се прехвърля в третата и т.н. Каква е вероятността от последната урна да бъде извадена бяла топка?

Решение (Техника). Нека $A_k = \{$от k-тата урна вадим бяла топка$\}$, $k = 1, \ldots, N$. Тогава $\mathbb{P}(A_1) = \frac{m}{m+n}$. Допускаме, че $\mathbb{P}(A_k) = \frac{m}{m+n}$ е изпълнено за някое $k = 1, \ldots, N-1$. Тогава

$$\mathbb{P}(A_{k+1}) = \mathbb{P}(A_{k+1}|A_k)\mathbb{P}(A_k) + \mathbb{P}(A_{k+1}|A_k^c)\mathbb{P}(A_k^c) = \frac{m+1}{m+n+1}\frac{m}{m+n} + \frac{m}{m+n+1}\frac{n}{m+n} = \frac{m}{m+n},$$

откъдето по индукция следва, че $\mathbb{P}(A_N) = \frac{m}{m+n}$. □

*Задача*** 2.41 (Gambler's Ruin). Последователно се хвърля монета. Ако се падне ези, играчът печели 1 лв., а ако се падне тура – губи 1 лв. В началото на играта играчът има x лв. Играта завършва или когато играчът набере предварително определена сума от a лв., или когато проиграе всичките си пари. Каква е вероятността играчът да спечели? **А какъв е шансът играчът да спечели a лв. на n-ти ход?

Решение (Техника). Нека $A = \{\text{играчът печели}\}$ и $B = \{\text{играчът печели първото хвърляне}\}$. Дефинираме $p(x) = \mathbb{P}(\{\text{печели с } x \text{ лв. начален капитал}\})$, $x = 0, 1, \ldots, a$. Тогава $p(0) = 0$, $p(a) = 1$ и

$$p(x) = \mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c) = \frac{1}{2}\big(p(x+1) + p(x-1)\big),$$

откъдето $p(x+1)-p(x) = p(x)-p(x-1) = \cdots = p(1)-p(0) = p(1)$. Но $p(x+1)-p(1) = \sum_{i=1}^{x}[p(i+1)-p(i)] = xp(1)$, т.е. $p(x) = xp(1)$. Следователно $p(a) = ap(1)$, откъдето $p(x) = \frac{x}{a}$.

Що се отнася до втория въпрос, резултатът следва след преброяване на траекториите от $(0,0)$ до $(n-1, a-x-1)$ (валидно за n, такова че $n+a-x$ е четно), които не докосват хоризонталите $y = -x$ и $y = a-x$, като от условието е ясно, че играчът печели n-тото хвърляне.

За по-прегледно, нека положим $n^* = n-1$, $m^* = a-x-1$. От Задача 1.12, броят на траекториите от $(0,0)$ до (n^*, m^*), докосващи $y = -x$ (или $y = a-x$), е N_{n^*,m^*+2x} (или $N_{n^*,m^*-2(a-x)}$). Въпреки това $N_{n^*,m^*+2x} + N_{n^*,m^*-2(a-x)}$ брои по два пъти траекториите от $(0,0)$ до (n^*, m^*), които докосват $y = -x$ и после $y = a-x$, и тези, които докосват $y = a-x$ и после $y = -x$. Техният брой, прилагайки на два пъти принципа на отражението от Задача 1.12, е съответно $N_{n^*,m^*-2(a-x)-2x}$ и $N_{n^*,m^*+2(a-x)+2x}$.

Аналогично, в $N_{n^*,m^*+2x} + N_{n^*,m^*-2(a-x)} - N_{n^*,m^*-2(a-x)-2x} - N_{n^*,m^*+2(a-x)+2x}$ са извадени траекториите от $(0,0)$ до (n^*, m^*), които докосват последователно $y = -x$, $y = a-x$ и $y = -x$, и тези, които докосват последователно $y = a-x$, $y = -x$ и $y = a-x$.

Нека означим с $N(A_1) = N_{n^*,m^*+2x}$, $N(A_2) = N_{n^*,m^*-2(a-x)-2x}$, $N(A_3)$ броя на траекториите от $(0,0)$ до (n^*, m^*), които докосват последователно $y = -x$, $y = a-x$ и $y = -x$, и т.н., и с $N(B_1) = N_{n^*,m^*-2(a-x)}$, $N(B_2) = N_{n^*,m^*+2(a-x)+2x}$, $N(B_3)$ броя на траекториите от $(0,0)$ до (n^*, m^*), които докосват последователно $y = a-x$, $y = -x$ и $y = a-x$, и т.н. Лесно се вижда, за $j = 0, 1, 2, \ldots$, че

$$\begin{aligned} N(A_{2j+1}) &= N_{n^*,m^*+2j(a-x)+2(j+1)x}, & N(A_{2j}) &= N_{n^*,m^*-2j(a-x)-2jx}, \\ N(B_{2j+1}) &= N_{n^*,m^*-2(j+1)(a-x)-2jx}, & N(B_{2j}) &= N_{n^*,m^*+2j(a-x)+2jx}, \end{aligned}$$

откъдето

$$\mathbb{P}(\{\text{печели а лв. на n-ти ход}\}) = \Big(N_{n^*,m^*} + \sum_{j=1}^{\infty}(-1)^j\big[N(A_j) + N(B_j)\big]\Big)\frac{1}{2^n}.$$

□

Забележка (Край на играта). По същия начин се доказва, че $\mathbb{P}(\{\text{играчът фалира}\}) = \frac{a-x}{a}$, откъдето получаваме $\mathbb{P}(\{\text{играе се безкрайно дълго}\}) = 1 - \mathbb{P}(\{\text{играчът печели}\}) - \mathbb{P}(\{\text{играчът фалира}\}) = 0$.

Забележка (Обобщение). Ако вероятността да се падне ези е p, тогава $p(x) = p \cdot p(x+1) + (1-p) \cdot p(x-1)$, откъдето $p(x+1)-p(x) = \frac{1-p}{p}[p(x)-p(x-1)] = \cdots = \left(\frac{1-p}{p}\right)^x p(1)$ и $p(x+1) = \sum_{i=0}^{x}\left(\frac{1-p}{p}\right)^i p(1) = \frac{1-(\frac{1-p}{p})^{x+1}}{1-\frac{1-p}{p}}p(1)$. Но $1 = p(a) = \frac{1-(\frac{1-p}{p})^a}{1-\frac{1-p}{p}}p(1)$, следователно

$$p(x) = \frac{1-\left(\frac{1-p}{p}\right)^x}{1-\left(\frac{1-p}{p}\right)^a}.$$

3 Дискретни случайни величини

3.1 Основни понятия

Дискретна случайна величина. *Нека $(\Omega, \mathcal{A}, \mathbb{P})$ е вероятностно пространство. Наричаме една функция $X : \Omega \to \mathbb{R}^m$ дискретна случайна величина (сл. в.), ако*

(i) нейният образ $\mathbb{X}$ в $\mathbb{R}^m$ е крайно или безкрайно, но изброимо множество;

(ii) за всяко $x \in \mathbb{X}$ е в сила

$$\{X = x\} \equiv X^{-1}(\{x\}) := \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{A};$$

Съвкупността от стойности $p(x) = \mathbb{P}(X = x)$ се нарича разпределение на X и удовлетворява $p(x) \geq 0$, $\sum_{x \in \mathbb{X}} p(x) = 1$, като често се записва в табличния вид

$$\begin{array}{c|ccccc} X & x_1 & x_2 & \cdots & x_n & \cdots \\ \hline \mathbb{P}_X & p(x_1) & p(x_2) & \cdots & p(x_n) & \cdots \end{array},$$

при определена подредба на $\mathbb{X} = \{x_1, x_2, \ldots, x_n, \ldots\}$.

Забележка (Интерпретация). Нека X е дискретна сл. в. и

$$\mathbb{P}_X(\cdot) := \sum_{x \in \mathbb{X}} p(x) \delta_x(\cdot),$$

където $\delta_x : 2^{\mathbb{X}} \to \{0, 1\}$ е *Дирак масата* в т. $x \in \mathbb{X}$, зададена за $A \subseteq \mathbb{X}$ чрез $\delta_x(A) = \begin{cases} 0, & x \notin A; \\ 1, & x \in A \end{cases}$.

Тогава $(\mathbb{X}, 2^{\mathbb{X}}, \mathbb{P}_X)$ дефинира вероятностно пространство, чрез което свеждаме оригиналния случаен експеримент до негов подексперимент, чиито елементарни изходи са възможните стойности на X. Освен това, за всяко $A \subseteq \mathbb{X}$, множеството $\{X \in A\} = X^{-1}(A) = \bigcup_{x \in A} \{X = x\} \in \mathcal{A}$ е измеримо, откъдето

$$\mathbb{P}(X \in A) = \mathbb{P}\big(\cup_{x \in A} \{X = x\}\big) = \sum_{x \in A} p(x) = \mathbb{P}_X(A).$$

Ще работим директно върху $(\mathbb{X}, 2^{\mathbb{X}}, \mathbb{P}_X)$, приемайки съществуването на базовото вероятностно пространство $(\Omega, \mathcal{A}, \mathbb{P})$ имплицитно, като допускаме, че то е достатъчно богато, за да поддържа сл. в., от които се нуждаем.

Забележка (Проста сл. в.). Нека $a_1, \ldots, a_n \in \mathbb{R}$ и $A_1, \ldots, A_n \in \mathcal{A}$ образува пълна група от събития върху Ω. Тогава

$$X(\omega) := \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(\omega), \qquad \omega \in \Omega,$$

дефинира сл. в., такава че $X(\omega) = a_i \iff \omega \in A_i$ за $i = 1, \ldots, n$, където $\mathbb{1}_A : \Omega \to \{0, 1\}$ е *индикаторната функция* на подмножеството $A \subseteq \Omega$, зададена за $\omega \in \Omega$ чрез $\mathbb{1}_A(\omega) = \delta_\omega(A)$.

*Забележка** (Функция на сл. в.). Нека X е дискретна сл. в. и $f : \mathbb{R}^m \to \mathbb{R}^n$ *произволна* функция. Тогава $Y = f(X)$ е дискретна сл. в. със стойности в $f(\mathbb{X}) := \{t \in \mathbb{R}^n : \exists x \in \mathbb{X}, f(x) = t\}$. Действително,

$$\{Y = y\} = \{X \in f^{-1}(\{y\})\} = \bigcup_{x \in \mathbb{X} : f(x) = y} \{X = x\} \in \mathcal{A}, \qquad y \in f(\mathbb{X}).$$

*Забележка** (Многомерна сл. в.). Нека дискретната сл. в. X приема стойности в $\mathbb{R}^m$. Дефинираме функцията $X_i(\omega) := (X(\omega))_i$, $\omega \in \Omega$ като i-тата компонента на X, която приема стойности в $\mathbb{X}_i \subseteq \mathbb{R}$, за $i = 1, \ldots, m$. Тогава, за всяко $x_i \in \mathbb{X}_i$, е в сила

$$\{X_i = x_i\} = \{X \in \mathbb{X}_1 \times \mathbb{X}_{i-1} \times \{x_i\} \times \mathbb{X}_{i+1} \times \mathbb{X}_m\} \in \mathcal{A};$$

т.е. $X_1, \ldots, X_m$ са реални дискретни сл. в. От това следва, че многомерната сл. в. $X = (X_1, \ldots, X_m)$ може да се разглежда еквивалентно като вектор от едномерни случайни величини. И обратното, ако $X_1, \ldots, X_m$ са реални дискретни сл. в., за всяко $x = (x_1, \ldots, x_m) \in \mathbb{X}_1 \times \cdots \times \mathbb{X}_m$ имаме

$$\{X = x\} = \{X_1 = x_1, \ldots, X_m = x_m\} = \bigcap_{i=1}^{m} \{X_i = x_i\} \in \mathcal{A}.$$

От последните две забележки следва, че сумата, произведението, частното, минимумът, максимумът, квадратът и т.н. на дискретни сл. в. сами по себе си са дискретни сл. в.

Задача 3.1. Човек хвърля монета, като прави крачка напред, ако се падне ези, и крачка назад, ако се падне тура. Каква е вероятността след 10 хвърляния да се намира (а) на мястото, откъдето е тръгнал; (б) на разстояние 2 крачки от началната си позиция; (в) на 5 крачки пред началната си позиция?

Решение. Нека сл. в. X е броят езита от 10 опита. Тогава $\mathbb{X} = \{0, 1, \ldots, 10\}$ и $\mathbb{P}(X = k) = \binom{10}{k}\frac{1}{2^{10}}$ за $k = 0, 1, \ldots, 10$, откъдето (а) $\mathbb{P}(X = 5) = 0.25$; (б) $\mathbb{P}(\{X = 4\} \cup \{X = 6\}) = 2 \cdot \mathbb{P}(X = 6) = 0.41$; (в) $\{X = 7.5\} = \emptyset$. □

Задача 3.2. Подводница стреля n пъти последователно по кораб. Всяко торпедо улучва с вероятност p. Корабът има m отсека и ако торпедото улучи кораба, вероятността да наводни кой да е от тях е една и съща. Каква е вероятността корабът да бъде потопен, ако за това е необходимо да се наводнят поне два отсека? А да бъдат наводнени точно $j \leq m$ отсека?

Отговор. Нека сл. в. X е броят наводнени отсека. Тогава $\mathbb{P}(X \geq 2) = 1 - (1-p)^n - m(\frac{p}{m} + (1-p))^n$ и $\mathbb{P}(X = j) = \binom{m}{j}\sum_{k=j}^{n}\binom{n}{k}p^k(1-p)^{n-k}\left(\frac{j}{m}\right)^k \frac{C(j,k-j)}{C(j,k)}$. □

. . . .

Моменти. *Нека X е реална дискретна сл. в. Ако редът $\sum_{x\in\mathbb{X}} xp(x)$ е абсолютно сходящ, числото*

$$\mathbb{E}[X] = \sum_{x\in\mathbb{X}} xp(x)$$

се нарича математическо очакване на X. За $k \in \mathbb{N}$, числата $\mathbb{E}[X^k]$ и $\mathbb{E}\big[(X - \mathbb{E}[X])^k\big]$ се наричат съответно k-ти начален и k-ти централен момент на X, като имат следното представяне

$$\mathbb{E}[X^k] = \sum_{x\in\mathbb{X}} x^k p(x) \qquad \text{и} \qquad \mathbb{E}\big[(X - \mathbb{E}[X])^k\big] = \sum_{x\in\mathbb{X}} \big(x - \mathbb{E}[X]\big)^k p(x),$$

при условие, че редовете са абсолютно сходящи. Вторият централен момент

$$\mathrm{Var}(X) \equiv \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2$$

се нарича дисперсия и се бележи още с σ^2. Числото $\sigma = \sqrt{\mathrm{Var}(X)}$ се нарича средно квадратично отклонение.

Забележка (Интуиция)*.* Ако съществуват, моментите на една сл. в. X описват числено определени характеристики на нейното вероятностното разпределение. Например математическото очакване като претеглена средна аритметична стойност измерва центъра на масата на $\mathbb{P}_X$. Нещо повече

$$\mathbb{E}[X] = \min_{a\in\mathbb{R}} \sum_{x\in\mathbb{X}} (x-a)^2 p(x);$$

т.е. $\mathbb{E}[X]$ минимизира очакваното квадратично отклонение, откъдето дисперсията $\mathrm{Var}(X)$ е точно минималното очакваното квадратично отклонение, асоциирано с $\mathbb{P}_X$.

*Забележка*** (Смяна на променливите)*.* Нека $Y = f(X)$ за произволна функция $f : \mathbb{R}^m \to \mathbb{R}$. Тогава

$$\mathbb{E}[f(X)] = \mathbb{E}[Y] = \sum_{y\in\mathbb{Y}} y\mathbb{P}(Y = y) = \sum_{y\in\mathbb{Y}} y\Big(\sum_{x:f(x)=y} \mathbb{P}(X = x)\Big) = \sum_{y\in\mathbb{Y}}\sum_{x:f(x)=y} f(x)\mathbb{P}(X = x) = \sum_{x\in\mathbb{X}} f(x)\mathbb{P}(X = x),$$

откъдето следват формулите за k-тия начален и k-тия централен момент на X.

Твърдение 3.1 (Свойства на очакването)**.**
позитивност: *ако $X \geq 0$, то $\mathbb{E}[X] \geq 0$, като $\mathbb{E}[X] = 0 \iff \mathbb{P}(X = 0) = 1$.*
линейност: *$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$, за $a, b \in \mathbb{R}$.*
базов случай: *$\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$, за $A \in \mathcal{A}$.*

Д-во. Ако $X \geq 0$, тогава $\mathbb{E}[X] = \sum_{x\in\mathbb{X}} xp(x) \geq 0$ тъй като $x \geq 0$. В случай, че $\mathbb{E}[X] = 0$, имаме $x = 0$ за всяко $x \in \mathbb{X}$, такова че $p(x) > 0$, откъдето $\mathbb{P}(X = 0) \geq \sum_{x\in\mathbb{X}:p(x)>0} p(x) = \sum_{x\in\mathbb{X}} p(x) = 1$, т.е. $\mathbb{P}(X = 0) = 1$. Относно второто свойство,

$$\mathbb{E}[aX + b] = \sum_{x\in\mathbb{X}} (ax + b)p(x) = a\sum_{x\in\mathbb{X}} xp(x) + b\sum_{x\in\mathbb{X}} p(x) = a\mathbb{E}[X] + b.$$

Последно $\mathbb{E}[\mathbb{1}_A] = 1 \cdot \mathbb{P}(\mathbb{1}_A = 1) + 0 \cdot \mathbb{P}(\mathbb{1}_A = 0) = 1 \cdot \mathbb{P}(A) + 0 \cdot \mathbb{P}(A^c) = \mathbb{P}(A)$. □

Забележка (Свойства на дисперсията). От горното твърдение следва, че $\mathrm{Var}(X) \geq 0$,

$$\mathrm{Var}(aX+b) = \mathbb{E}\big[\big((aX+b) - (\mathbb{E}[aX+b])\big)^2\big] = \mathbb{E}[a^2(X-\mathbb{E}[X])^2] = a^2\mathrm{Var}(X),$$

и $\mathrm{Var}(X) = 0 \iff \mathbb{P}(X = \mathbb{E}[X]) = 1 \iff X$ е (почти сигурно) константна сл. в.

Твърдение 3.2 (Неравенство на Марков)**.** *Нека X е дискретна сл. в., такава че $X \geq 0$. Тогава за всяко $\alpha > 0$ е в сила*

$$\mathbb{P}(X > \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}.$$

Д-во.* $\mathbb{E}[X] = \sum_{x\in\mathbb{X}} xp(x) \geq \sum_{x\in\mathbb{X}:x>\alpha} xp(x) \geq \sum_{x\in\mathbb{X}:x>\alpha} \alpha p(x) = \alpha\cdot\mathbb{P}(X>\alpha)$. □

Твърдение 3.3 (Неравенство на Чебишев)**.** *Нека X е дискретна сл. в. с очакване и дисперсия съответно $\mathbb{E}[X]$ и $\mathrm{Var}(X)$. Тогава за всяко $\alpha > 0$ е в сила*

$$\mathbb{P}\big(|X - \mathbb{E}[X]| > \alpha\big) \leq \frac{\mathrm{Var}(X)}{\alpha^2}.$$

Д-во.* Използвайки свойствата на математическото очакване,

$$\mathrm{Var}(X) \geq \mathbb{E}\big[(X-\mathbb{E}[X])^2 \cdot \mathbb{1}_{|X-\mathbb{E}[X]|>\alpha}\big] = \sum_{x\in\mathbb{X}:|x-\mathbb{E}[X]|>\alpha} (x-\mathbb{E}[X])^2 p(x) \geq \sum_{x\in\mathbb{X}:|x-\mathbb{E}[X]|>\alpha} \alpha^2 p(x) = \alpha^2\mathbb{P}(|X-\mathbb{E}[X]|>\alpha).$$

□

Забележка (Общо покритие). От неравенството на Чебишев, следва, че

$$\mathbb{P}\big(\mathbb{E}[X] - 2\sigma < X < \mathbb{E}[X] + 2\sigma\big) = 1 - \mathbb{P}\big(|X - \mathbb{E}[X]| > 2\sigma\big) \geq 0.75,$$

т.е. с вероятност поне 75% X приема стойност в интервала $(\mathbb{E}[X] - 2\sigma, \mathbb{E}[X] + 2\sigma)$.

Задача 3.3. Играч залага 5 лв. и има право да хвърли два зара. Ако хвърли две шестици печели 100 лв., а ако хвърли точно една шестица – 5 лв. Да се пресметне математическото очакване на печалбата на играча.

Отговор. $-5/6$. □

Задача 3.4 (St. Petersburg Paradox). Казино предлага следната игра: играч плаща A лв. След това хвърля монета, докато не се падне ези. Ако това се случи на n-тия ход, печели 2^n лв. Каква е очакваната му печалба? При какви стойности на A бихте участвали?

Решение. Нека сл. в. X е печалбата на играча. Тогава $\mathbb{X} = \{2-A, 2^2-A, \dots, 2^n-A, \dots\}$ и $\mathbb{P}(X = 2^n - A) = \frac{1}{2^n}$, $n \in \mathbb{N}$, откъдето $\mathbb{E}[X] = \sum_{n=1}^{\infty}(2^n - A)\frac{1}{2^n} = \infty$, т.е. очакваната печалба не е ограничена отгоре. Въпреки това сумата, която човек е готов да плати, за да участва, рядко е голяма. Това показва, че очакваната стойност, като една от многото характеристики, описващи разпределението на X, в някои случаи е недостатъчен критерий за вземане на решения. □

Задача 3.5 (Martingale Strategy). Казино предлага коефициент 2 при игра на ези/тура, т.е. при залог от A лв., бихме спечелили чисто A лв. Играч залага само на ези, докато спечели, като удвоява залога си всеки път, когато не спечели. Каква е очакваната му печалба? Бихте ли пробвали?

Решение. Нека сл. в. X е печалбата на играча. Ако играчът хвърли ези на n-тото залагане, печалбата му е

$$2^nA - (A + 2A + \dots + 2^{n-1}A) = A\big(2^n - (1 + 2 + \dots + 2^{n-1})\big) = A(2^n - (2^n - 1)) = A;$$

т.е. $X = A$ е *константна сл. в.*, откъдето $\mathbb{E}[X] = A$. Играта обаче предполага, че играчът разполага с безкраен времеви хоризонт и богатство, докато залозите нарастват експоненциално. □

Задача 3.6. Случайна величина X приема краен брой стойности $x_1, \dots, x_n$ с една и съща вероятност. Да се пресметнат $\mathbb{E}[X]$ и $\mathrm{Var}(X)$, ако (а) $x_i = \frac{i-1}{n-1}$ за $i = 1, \dots, n$; (б) $x_i = a + (b-a)\frac{i-1}{n-1}$, $b > a$, за $i = 1, \dots, n$.

Решение. Сл. в. X има *равномерно разпределение*, т.е. $\mathbb{P}(X = x_i) = \frac{1}{n}$, за $i = 1, \dots, n$. Тогава (а) $\mathbb{E}[X] = \sum_{i=1}^n \frac{i-1}{n-1}\frac{1}{n} = \frac{\sum_{i=1}^{n-1} i}{n(n-1)} = \frac{1}{2}$ и $\mathbb{E}[X^2] = \sum_{i=1}^n \big(\frac{i-1}{n-1}\big)^2\frac{1}{n} = \frac{\sum_{i=1}^{n-1} i^2}{n(n-1)^2} = \frac{2n-1}{6(n-1)}$, откъдето $\mathrm{Var}(X) = \frac{n+1}{12(n-1)}$; (б) $\mathbb{E}[X] = a + (b-a)\frac{1}{2} = \frac{a+b}{2}$ и $\mathrm{Var}(X) = (b-a)^2\frac{n+1}{12(n-1)}$. □

3.2 Колекция от случайни величини

Съвместно разпределение. *Нека X и Y са две дискретни сл. в., дефинирани в едно и също вероятностно пространство и приемащи стойности съответно в $\mathbb{X} = \{x_1, x_2, \dots\} \subseteq \mathbb{R}^m$ и $\mathbb{Y} = \{y_1, y_2, \dots\} \subseteq \mathbb{R}^n$. Съвместно разпределение на X и Y се нарича съвкупността на числата $p(x_i, y_j) = \mathbb{P}(X = x_i, Y = y_j)$ за различни $x_i \in \mathbb{X}$ и $y_j \in \mathbb{Y}$. Числата $p(x_i, y_j)$ удовлетворяват условията $p(x_i, y_j) > 0$, $\sum_{i,j} p(x_i, y_j) = 1$,*

$$\sum_j p(x_i, y_j) = \mathbb{P}(X = x_i) \qquad \text{и} \qquad \sum_i p(x_i, y_j) = \mathbb{P}(Y = y_j),$$

т.е. съвместното разпределение съдържа т.нар. маргинални разпределения на всяка от отделните сл. в., като често се записва в табличния вид

$Y \backslash X$	x_1	x_2	$\cdots$	x_i	$\cdots$	
y_1	$p(x_1, y_1)$	$p(x_2, y_1)$	$\cdots$	$p(x_i, y_1)$	$\cdots$	$\mathbb{P}(Y = y_1)$
y_2	$p(x_1, y_2)$	$p(x_2, y_2)$	$\cdots$	$p(x_i, y_2)$	$\cdots$	$\mathbb{P}(Y = y_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
y_j	$p(x_1, y_j)$	$p(x_2, y_j)$	$\cdots$	$p(x_i, y_j)$	$\cdots$	$\mathbb{P}(Y = y_j)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
	$\mathbb{P}(X = x_1)$	$\mathbb{P}(X = x_2)$	$\cdots$	$\mathbb{P}(X = x_i)$	$\cdots$	

Твърдение 3.4.
(i) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;
(ii) *ако* $X \geq Y$, *то* $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

Д-во.

$$\mathbb{E}[X + Y] = \sum_{(x,y) \in \mathbb{X} \times \mathbb{Y}} (x + y) p(x, y) = \sum_{x \in \mathbb{X}} x \sum_{y \in \mathbb{Y}} p(x, y) + \sum_{y \in \mathbb{Y}} y \sum_{x \in \mathbb{X}} p(x, y) = \sum_{x \in \mathbb{X}} x \mathbb{P}(X = x) + \sum_{y \in \mathbb{Y}} y \mathbb{P}(Y = y).$$

Относно второто свойство, от $X - Y \geq 0$ следва, че $\mathbb{E}[X] = \mathbb{E}[Y] + \mathbb{E}[X - Y] \geq \mathbb{E}[Y]$. □

Независимост. *Две дискретни сл. в. X и Y се наричат независими, ако за произволни техни стойности $x \in \mathbb{X}$ и $y \in \mathbb{Y}$ е изпълнено равенството*

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

В този случай ще обозначаваме $X \perp\!\!\!\perp Y$.

Група от дискретни случайни величини $X_1, \dots, X_n$, всяка от които приема стойности съответно в $\mathbb{X}_i \subseteq \mathbb{R}^{m_i}$, $i = 1, \dots, n$, се наричат независими в съвкупност, ако <u>за кои да е</u> $i_1, \dots, i_k \in \{1, \dots, n\}$ при $k = 2, \dots, n$ и стойности $x_{i_1} \in \mathbb{X}_{i_1}, \dots, x_{i_k} \in \mathbb{X}_{i_k}$, е изпълнено

$$\mathbb{P}(X_{i_1} = x_{i_1}, \dots, X_{i_1} = x_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(X_{i_j} = x_{i_j}).$$

Твърдение 3.5. *Ако $X \perp\!\!\!\perp Y$, то $f(X) \perp\!\!\!\perp g(Y)$ и $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ за $f : \mathbb{R}^m \to \mathbb{R}$ и $g : \mathbb{R}^n \to \mathbb{R}$.*

Д-во. Лесно се вижда, че $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ за $A \subseteq \mathbb{X}$ и $B \subseteq \mathbb{Y}$, ако $X \perp\!\!\!\perp Y$. Тогава, за всеки $u \in f(\mathbb{X})$ и $v \in g(\mathbb{Y})$, е валидно

$$\mathbb{P}(f(X) = u, g(Y) = v) = \mathbb{P}\big(X \in f^{-1}(\{u\}), Y \in g^{-1}(\{v\})\big) = \mathbb{P}(f(X) = u)\mathbb{P}(g(Y) = v).$$

Относно втория резултат,

$$\mathbb{E}[f(X)g(Y)] = \sum_{(x,y) \in \mathbb{X} \times \mathbb{Y}} f(x)g(y)\mathbb{P}(X = x, Y = y) = \sum_{x \in \mathbb{X}} f(x) \Big(\sum_{y \in \mathbb{Y}} g(y)\mathbb{P}(Y = y) \Big) \mathbb{P}(X = x) = \mathbb{E}[f(X)]\mathbb{E}[g(Y)].$$

□

Еднаквост по разпределение. *Две дискретни сл. в. X и Y, приемащи едни и същи възможни стойности в $\mathbb{X} \subseteq \mathbb{R}^m$, се наричат еднакво разпределени, ако за всяко $x \in \mathbb{X}$ е изпълнено*

$$\mathbb{P}(X = x) = \mathbb{P}(Y = x).$$

В този случай ще обозначаваме $X \stackrel{d}{=} Y$. Нещо повече, ако $X \stackrel{d}{=} Y$, тогава за $f : \mathbb{R}^m \to \mathbb{R}$,

$$\mathbb{E}[f(X)] = \sum_{x\in\mathbb{X}} f(x)\mathbb{P}(X = x) = \sum_{x\in\mathbb{X}} f(x)\mathbb{P}(Y = x) = \mathbb{E}[f(Y)].$$

Когато $X \perp\!\!\!\perp Y$ и $X \stackrel{d}{=} Y$, ще наричаме X и Y независимо и еднакво разпределени, накратко i.i.d.

Задача 3.7. А хвърля 3 монети, а Б – 2. Печели този, който хвърли повече езита, като взима всичките 5 монети. В случай на равен брой печели Б. Каква е вероятността А да спечели? Ако е спечелил А, каква е вероятността Б да хвърли точно едно ези? Каква е средната печалба на играчите?

Отговори. Ако сл. в. X и Y са броят на езитата, хвърлени съответно от А и Б, тогава $\mathbb{P}(X > Y) = \frac{1}{2}$ и $\mathbb{P}(Y = 1|X > Y) = \frac{1}{2}$. Печалбата на А е $f(X, Y) = (5 - 3) \cdot \mathbb{1}_{\{X>Y\}} + (0 - 3) \cdot \mathbb{1}_{\{X\le Y\}}$, откъдето

$$\mathbb{E}[f(X, Y)] = \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} \big(2 \cdot \mathbb{1}_{\{x>y\}}(x, y) - 3 \cdot \mathbb{1}_{\{x\le y\}}(x, y)\big)\mathbb{P}(X = x, Y = y) = 2 \cdot \mathbb{P}(X > Y) - 3 \cdot \mathbb{P}(X \le Y) = -\frac{1}{2}.$$

Понеже играта е от вида *"zero-sum game"*, очакваната печалба на Б е $\frac{1}{2}$. □

Задача 3.8. В урна има 3 бели и 2 черни топки, като теглим последователно топки без връщане. Нека сл. в. X е моментът на тегленето на първата бяла топка. След това продължаваме да теглим, докато се появи черна топка. Нека Y е моментът на тегленето на първата черна топка след първата бяла. Дефинираме $Y = 6$, ако няма такава. Да се определи (а) съвместното разпределение на X и Y; (б) $\mathbb{P}(Y > 2|X = 1)$ и $\mathbb{P}(Y = 3|X < 3)$.

Отговори. $\mathbb{P}(Y > 2|X = 1) = \frac{1}{2}$, $\mathbb{P}(Y = 3|X < 3) = \frac{1}{3}$,

$X\backslash Y$	2	3	4	5	6
1	3/10	2/10	1/10	0	0
2	0	1/10	1/10	1/10	0
3	0	0	0	0	1/10

□

Задача 3.9. От урна, съдържаща 5 бели и 3 черни топки се избират последователно, една по една топки, докато се появи бяла топка. Да се намери разпределението, очакването и дисперсията на случайната величина X = "брой на изтеглените черни топки" при извадка без връщане. Опитът се повтаря 1000 пъти. Да се оцени вероятността да са извадени повече от 900 черни топки.

Решение. От $\mathbb{P}(X = k) = \frac{V_3^k V_5^1}{V_8^{k+1}}$, $k = 0, 1, 2, 3$, следва, че $\mathbb{E}[X] = 0.5$ и $\mathrm{Var}(X) = 0.54$. Нека $X_1, \dots, X_{1000}$ са i.i.d. копия на X, и $Y = \sum_{i=1}^{1000} X_i$. Тогава $\mathbb{E}[Y] = 500$ и $\mathrm{Var}(Y) = 540$, откъдето, чрез неравенството на Чебишев,

$$\mathbb{P}(Y > 900) = \mathbb{P}\big(Y - \mathbb{E}[Y] > 400\big) \le \mathbb{P}\big(|Y - \mathbb{E}[Y]| > 400\big) \le \frac{\mathrm{Var}(Y)}{400^2} = 0.0034.$$

Разпределението на Y може да се изведе, напр. от нейната характеристична функция, но горната граница е достатъчна, за да заключим невъзможността на събитието да се случи. □

Задача 3.10. Хвърлят се два зара. Нека случайните величини X и Y са съответно сумата и произведението от падналите се точки. Да се намери разпределението, очакването и дисперсията на X и Y, ако (а) заровете са правилни; (б) вероятността да хвърлим 2, 3, 4 и 5 е 1/8, а 1 и 6 – 1/4. Ще бъде ли необичайно, ако при хвърлянето на 1000 зара сумата да е била повече от 3700?

Решение (Идея). Нека сл. в. X_1 и X_2 са точките, хвърлени съответно на първия и втория зар. Тогава $X_1 \perp\!\!\!\perp X_2$, $X_1 \stackrel{d}{=} X_2$, $X = X_1 + X_2$ и $Y = X_1X_2$, откъдето

$$\mathbb{E}[X] = 2 \cdot \mathbb{E}[X_1], \qquad \mathrm{Var}(X) = 2 \cdot \mathrm{Var}(X_1), \qquad \mathbb{E}[Y] = \big(\mathbb{E}[X_1]\big)^2, \qquad \mathrm{Var}(Y) = \big(\mathbb{E}[X_1^2]\big)^2 - \big(\mathbb{E}[X_1]\big)^4.$$

За втория въпрос приложете неравенството на Чебишев. □

Ковариация *на две реални дискретни сл. в. X и Y наричаме числото*

$$\operatorname{Cov}(X,Y) \equiv \mathbb{E}\big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\big] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y],$$

при условие, че $\mathbb{E}[X]$, $\mathbb{E}[Y]$ и $\mathbb{E}[XY]$ съществуват. В този случай числото

$$\rho_{X,Y} = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}}.$$

наричаме коефициент на корелация на X и Y.

Забележка (Интуиция)*.* Ковариацията измерва съвместното движение на две сл. в., което ни информира за линейната зависимост между тях. Тъй като ковариацията като число зависи от "мерните единици" на двете сл. в., въвеждаме нейната нормализирана стойност, а именно коефициента на корелация.

Твърдение 3.6 (Неравенство на Коши-Шварц)**.** *Нека X и Y са реални сл. в. Тогава*

$$|\rho_{X,Y}| \le 1,$$

като $|\rho_{X,Y}| = 1$ тогава и само тогава, когато $\mathbb{P}(Y = aX + b) = 1$ за някои $a, b \in \mathbb{R} : a \neq 0$.

*Д-во**.* Нека положим $\bar{X} = \frac{X-\mathbb{E}[X]}{\sqrt{\operatorname{Var}(X)}}$ и $\bar{Y} = \frac{Y-\mathbb{E}[Y]}{\sqrt{\operatorname{Var}(Y)}}$, така че $\mathbb{E}[\bar{X}], \mathbb{E}[\bar{Y}] = 0$, $\mathbb{E}[\bar{X}^2], \mathbb{E}[\bar{Y}^2] = 1$ и $\rho_{X,Y} = \mathbb{E}[\bar{X}\bar{Y}]$. Дефинираме $Z = \bar{X} - \rho_{X,Y}\cdot\bar{Y}$. Тогава $\mathbb{E}[Z\bar{Y}] = \mathbb{E}[\bar{X}\bar{Y}] - \rho_{X,Y}\mathbb{E}[\bar{Y}^2] = 0$, откъдето

$$1 = \mathbb{E}[\bar{X}^2] = \mathbb{E}\big[(Z + \rho_{X,Y}\cdot\bar{Y})^2\big] = \mathbb{E}[Z^2] + \rho_{X,Y}^2\cdot\mathbb{E}[\bar{Y}^2] \ge \rho_{X,Y}^2.$$

В случай, че $|\rho_{X,Y}| = 1$, тогава $\mathbb{E}[Z^2] = 0$, което предполага $\mathbb{P}(\bar{X} = \rho_{X,Y}\cdot\bar{Y}) = 1$. □

Твърдение 3.7. *Ако $X \perp\!\!\!\perp Y$, тогава $\operatorname{Cov}(X,Y) = 0$ и $\operatorname{Var}(aX \pm bY + c) = a^2\operatorname{Var}(X) + b^2\operatorname{Var}(Y)$ за $a, b, c \in \mathbb{R}$. Иначе*

$$\operatorname{Var}(aX \pm bY + c) = a^2\operatorname{Var}(X) + b^2\operatorname{Var}(Y) \pm 2ab\operatorname{Cov}(X,Y).$$

Д-во.

$$\begin{aligned}\operatorname{Var}(aX \pm bY + c) &= \mathbb{E}\Big[\big(a(X-\mathbb{E}[X]) \pm b(Y-\mathbb{E}[Y])\big)^2\Big] \\ &= a^2\mathbb{E}\big[(X-\mathbb{E}[X])^2\big] + b^2\mathbb{E}\big[(Y-\mathbb{E}[Y])^2\big] \pm 2ab\,\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] \\ &= a^2\operatorname{Var}(X) + b^2\operatorname{Var}(Y) \pm 2ab\operatorname{Cov}(X,Y).\end{aligned}$$

Ако $X \perp\!\!\!\perp Y$, то $\operatorname{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$. □

Забележка (Обобщение)*.* По индукция се извежда равенството за сума от n реални сл. в. $X_1, \dots, X_n$,

$$\operatorname{Var}\Big(\sum_{i=1}^n X_i\Big) = \sum_{i=1}^n \operatorname{Var}(X_i) + 2\sum_{i,j:i<j}\operatorname{Cov}(X_i, X_j) = \sum_{i,j}\operatorname{Cov}(X_i, X_j).$$

Задача 3.11. Независимите сл. в. X и Y имат следните разпределения:

X	-1	1
$\mathbb{P}_X$	0.5	0.5

;

Y	1	3	5
$\mathbb{P}_Y$	0.5	0.25	0.25

.

Да се намери разпределението, очакването и дисперсията на (а) $2X + Y + 1$; (б) XY.

Отговори. $\mathbb{E}[2X + Y + 1] = \frac{7}{2}$, $\operatorname{Var}(2X + Y + 1) = \frac{27}{4}$; $\mathbb{E}[XY] = \operatorname{Cov}(X,Y) = 0$, $\operatorname{Var}(XY) = \mathbb{E}[X^2]\mathbb{E}[Y^2] = 9$. □

Задача 3.12. Дискретни сл. в. X и Y имат следното съвместно разпределение:

$Y\backslash X$	-1	0	1
0	0.1	0.2	0
1	0.2	0.3	0.2

Да се намерят (а) разпределенята на X и Y; (б) $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\operatorname{Var}(X)$ и $\operatorname{Var}(Y)$; (в) $\operatorname{Cov}(X,Y)$ и $\rho_{X,Y}$; (г) разпределението на $Z = X^2 + 2Y$, $\mathbb{E}[Z]$ и $\operatorname{Var}(Z)$.

Отговори. $\mathbb{E}[X] = -0.1$, $\mathbb{E}[Y] = 0.7$, $\mathrm{Var}(X) = 0.49$, $\mathrm{Var}(Y) = 0.21$; $\mathrm{Cov}(X, Y) = 0.07$, $\rho_{X,Y} = 0.22$; $\mathbb{E}[Z] = 1.9$, $\mathrm{Var}(Z) = 1.29$, $\begin{array}{c|cccc} Z & 0 & 1 & 2 & 3 \\ \hline \mathbb{P}_Z & 0.2 & 0.1 & 0.3 & 0.4 \end{array}$. □

Задача 3.13. Хвърляме монета два пъти. Нека X и Y са съответно броят на хвърлените езита и тури. Да се намери $\rho_{X,Y}$.

Решение. Понеже $Y = 2 - X$, следва, че $\rho_{X,Y} = -1$. □

Задача 3.14. Правилен зар се хвърля n пъти. Нека X е сумата от първите $n-1$ хвърляния, а Y – сумата от последните $n-1$. Намерете $\rho_{X,Y}$.

Решение. Нека сл. в. X_i е хвърленото число на i-тия ход, $i = 1, \dots, n$. Тогава $X_1, \dots, X_n$ са i.i.d. Да положим $Z = \sum_{i=2}^{n-1} X_i$. По условие $X = X_1 + Z$ и $Y = Z + X_n$, откъдето следва, че

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathrm{Cov}(X_1 + Z, Z + X_n) = \mathbb{E}\Big[\big((X_1 - \mathbb{E}[X_1]) + (Z - \mathbb{E}[Z])\big)\big((Z - \mathbb{E}[Z]) + (X_n - \mathbb{E}[X_n])\big)\Big] \\ &= \mathrm{Cov}(Z, Z) + \mathrm{Cov}(X_1, Z) + \mathrm{Cov}(X_1, X_n) + \mathrm{Cov}(Z, X_n) \\ &= \mathrm{Var}(Z) = (n-2)\mathrm{Var}(X_1) = (n-2)\left[\frac{\sum_{i=1}^6 i^2}{6} - \left(\frac{\sum_{i=1}^6 i}{6}\right)^2\right] = \frac{35(n-2)}{12}.\end{aligned}$$

По същия начин $\mathrm{Var}(X) \equiv \mathrm{Var}(Y) = (n-1)\mathrm{Var}(X_1)$, откъдето $\rho_{X,Y} = \frac{n-2}{n-1}$. □

3.3 Условно очакване

Условно очакване. *Нека X и Y са две дискретни сл. в., дефинирани в едно и също вероятностно пространство и приемащи стойности съответно в $\mathbb{X} \subseteq \mathbb{R}$ и $\mathbb{Y} \subseteq \mathbb{R}^n$. Да допуснем, че $\mathbb{P}(Y = y) > 0$, $y \in \mathbb{Y}$. Ако $\mathbb{E}[X]$ съществува, условното математическо очакване на X относно Y се нарича случайната величина $\mathbb{E}[X|Y] = g(Y)$, където функцията $g : \mathbb{R}^n \to \mathbb{R}$ е зададена чрез*

$$g(y) \equiv \mathbb{E}[X|Y = y] = \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y), \qquad y \in \mathbb{Y}.$$

Тогава $\mathbb{E}[X|Y] = \sum_{y \in \mathbb{Y}} \mathbb{E}[X|Y = y]\mathbb{1}_{\{Y=y\}}$ приема възможни стойности $k \in \big\{\mathbb{E}[X|Y = y]\big\}_{y \in \mathbb{Y}}$ с вероятност

$$\mathbb{P}(\mathbb{E}[X|Y] = k) = \sum_{y \in \mathbb{Y}: \mathbb{E}[X|Y=y]=k} \mathbb{P}(Y = y).$$

Забележка (Интуиция). Математическото очакване на X при условие $Y = y$, т.е. $\mathbb{E}[X|Y = y]$, измерва среднопретеглената стойност на X при положение, че е известно, че $Y = y$.

Забележка (Условно разпределение). Условното разпределение на X относно Y е семейството от вероятностни разпределения $\big(\mathbb{P}_{X|Y=y}(\cdot)\big)_{y \in \mathbb{Y}}$, където за всяко $y \in \mathbb{Y}$ и $A \subseteq \mathbb{X}$,

$$P_{X|Y=y}(A) = \mathbb{P}(X \in A|Y = y) = \sum_{x \in \mathbb{X}} \mathbb{1}_A(x)\mathbb{P}(X = x|Y = y) = \mathbb{E}[\mathbb{1}_{\{X \in A\}}|Y = y].$$

Твърдение 3.8 (Свойства на условното очаване)**.**
- *(i)* $\mathbb{E}[aX + b|Y] = a\mathbb{E}[X|Y] + b$, *за* $a, b \in \mathbb{R}$;
- *(ii)* $\mathbb{E}\big[\mathbb{E}[X|Y]\big] = \mathbb{E}[X]$; *(Law of Total Expectation)*
- *(iii)* *ако* $X \perp\!\!\!\perp Y$, *то* $\mathbb{E}[X|Y] = \mathbb{E}[X]$;
- *(iv)* *ако* $X = f(Y)$ *за* $f : \mathbb{R}^n \to \mathbb{R}$, *то* $\mathbb{E}[X|Y] = X$.

Д-во.

$$\mathbb{E}[aX + b|Y] = \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} \mathbb{E}[aX + b|Y = y]\mathbb{1}_{\{Y=y\}} = a\sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y)\mathbb{1}_{\{Y=y\}} + b\sum_{y \in \mathbb{Y}} \mathbb{1}_{\{Y=y\}} = a\sum_{y \in \mathbb{Y}} \mathbb{E}[X|Y = y]\mathbb{1}_{\{Y=y\}} + b;$$

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X|Y]] &= \mathbb{E}\Big[\sum_{y \in \mathbb{Y}} \mathbb{E}[X|Y = y]\mathbb{1}_{\{Y=y\}}\Big] = \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y)\mathbb{E}[\mathbb{1}_{\{Y=y\}}] \\ &= \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y)\mathbb{P}(Y = y) = \sum_{x \in \mathbb{X}} x\Big(\sum_{y \in \mathbb{Y}} \mathbb{P}(X = x|Y = y)\mathbb{P}(Y = y)\Big) = \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x) = \mathbb{E}[X];\end{aligned}$$

$$\mathbb{E}[X|Y] = \sum_{y\in\mathbb{Y}}\sum_{x\in\mathbb{X}} x\mathbb{P}(X=x|Y=y)\mathbb{1}_{\{Y=y\}} = \sum_{y\in\mathbb{Y}}\sum_{x\in\mathbb{X}} x\mathbb{P}(X=x)\mathbb{1}_{\{Y=y\}} = \sum_{y\in\mathbb{Y}} \mathbb{E}[X]\mathbb{1}_{\{Y=y\}} = \mathbb{E}[X], \qquad \text{за } X \perp\!\!\!\perp Y;$$

$$\mathbb{E}[X|Y] = \sum_{y\in\mathbb{Y}}\sum_{x\in\mathbb{X}} x\mathbb{P}(X=x|Y=y)\mathbb{1}_{\{Y=y\}} = \sum_{y\in\mathbb{Y}} g(y)\Big(\sum_{x\in\mathbb{X}} \mathbb{P}(X=x|Y=y)\Big)\mathbb{1}_{\{Y=y\}} = \sum_{y\in\mathbb{Y}} g(y)\mathbb{1}_{\{Y=y\}} = g(Y), \qquad \text{за } X = g(Y).$$

□

Задача 3.15. Четири пъти последователно се хвърля монета. Нека X е броят езита, паднали се при първите три хвърляния, а Y – броят езита от последните две. Да се намери (а) съвместното разпределение на X и Y; (б) условното разпределение на X относно Y и на Y относно X; (в) $\mathbb{P}(X = Y)$, $\mathbb{P}(X > 1|Y = 1)$ и $\mathbb{P}(X + Y > 2|X = 2)$; (г) разпределенията на $\mathbb{E}[X|Y]$ и $\mathbb{E}[Y|X]$.

Отговори. $\mathbb{E}[X] = 1.5$, $\mathbb{E}[Y] = 1$, $\mathbb{P}(X = Y) = \frac{3}{8}$, $\mathbb{P}(X > 1|Y = 1) = \frac{7}{8}$, $\mathbb{P}(X + Y > 2|X = 2) = \frac{5}{6}$,

X	0	1	2	3
$\mathbb{P}_{X\|Y=0}$	1/4	1/2	1/4	0
$\mathbb{P}_{X\|Y=1}$	1/8	3/8	3/8	1/8
$\mathbb{P}_{X\|Y=2}$	0	1/4	1/2	1/4

Y	0	1	2
$\mathbb{P}_{Y\|X=0}$	1/2	1/2	0
$\mathbb{P}_{Y\|X=1}$	1/3	1/2	1/6
$\mathbb{P}_{Y\|X=2}$	1/6	1/2	1/3
$\mathbb{P}_{Y\|X=3}$	0	1/2	1/2

$\mathbb{E}[X\|Y]$	1	1.5	2
$\mathbb{P}$	1/4	1/2	1/4

$\mathbb{E}[Y\|X]$	0.5	0.8334	1.1667	1.5
$\mathbb{P}$	1/8	3/8	3/8	1/8

□

Задача 3.16. На всяка от страните на зар се записва случайно число от 1 до 6, след което зарът се хвърля n пъти. Нека X е сумата от първите $n-1$ хвърляния, а Y – сумата от последните $n-1$. Намерете $\rho_{X,Y}$.

Решение. Нека сл. в. X_i е хвърленото число на i-тия ход, $i = 1, \ldots, n$, а $N = (N_1, \ldots, N_6)$ – векторът, представляващ разгъвката на зара. Да положим $Z = \sum_{i=2}^{n-1} X_i$. По условие $X = X_1 + Z$, $Y = Z + X_n$ и $\mathbb{P}(N = (i_1, \ldots, i_6)) = \frac{1}{6^6}$, за $i_1, \ldots, i_6 \in \{1, \ldots, 6\}$, от което следва, че $N_1, \ldots, N_6$ са i.i.d. и $\mathbb{P}(N_1 = i) = \frac{1}{6}$ за $i = 1, \ldots, 6$. От друга страна, за разлика от Задача 3.14, $X_1, \ldots, X_n$ са еднакво разпределени, но не са независими. Въпреки това те са **условно независими** относно N, откъдето

$$\mathrm{Var}(X_i) = \mathbb{E}\big[\mathbb{E}[X_i^2|N]\big] - \big(\mathbb{E}\big[\mathbb{E}[X_i|N]\big]\big)^2 = \frac{1}{6}\mathbb{E}\Big[\sum_{k=1}^{6} N_k^2\Big] - \frac{1}{36}\Big(\mathbb{E}\Big[\sum_{k=1}^{6} N_k\Big]\Big)^2 = \mathbb{E}[N_1^2] - (\mathbb{E}[N_1])^2 = \frac{1}{6}\sum_{i=1}^{6} i^2 - \frac{1}{36}\Big(\sum_{i=1}^{6} i\Big)^2 = 2.917,$$

като всъщност X_1 има маргинално същото разпределение като на правилен зар, и

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}\big[\mathbb{E}[X_i X_j|N]\big] - \mathbb{E}\big[\mathbb{E}[X_i|N]\big]\mathbb{E}\big[\mathbb{E}[X_j|N]\big] = \frac{1}{36}\mathbb{E}\Big[\sum_{k,l} N_k N_l\Big] - \frac{1}{36}\Big(\mathbb{E}\Big[\sum_{k=1}^{6} N_k\Big]\Big)^2 = \frac{1}{36}\mathbb{E}\Big[\sum_{j=1}^{6} N_j^2 + \sum_{k\neq l} N_k N_l\Big] - (\mathbb{E}[N_1])^2$$

$$= \frac{1}{6}\mathbb{E}[N_1^2] + \frac{1}{36}\sum_{k\neq l}\mathbb{E}[N_k]\mathbb{E}[N_l] - (\mathbb{E}[N_1])^2 = \frac{1}{36}\sum_{i=1}^{6} i^2 + \frac{30}{36*36}\Big(\sum_{i=1}^{6} i\Big)^2 - \frac{1}{36}\Big(\sum_{i=1}^{6} i\Big)^2 = \frac{1}{6}\mathrm{Var}(X_1) = 0.486 > 0!$$

От тук следва, че

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(Z, Z) + \mathrm{Cov}(X_1, Z) + \mathrm{Cov}(Z, X_n) + \mathrm{Cov}(X_1, X_n) = (n-2)\mathrm{Var}(X_1) + (n^2 - 3n + 3)\frac{1}{6}\mathrm{Var}(X_1).$$

По същия начин $\mathrm{Var}(X) \equiv \mathrm{Var}(Y) = (n-1)\mathrm{Var}(X_1) + (n^2 - 3n + 2)\frac{1}{6}\mathrm{Var}(X_1)$, откъдето $\rho_{X,Y} = \frac{n^2+3n-9}{n^2+3n-4}$. □

*Задача** 3.17 (Two Envelopes Problem). Пред вас има два еднакви плика с пари, единият от които съдържа два пъти повече пари от другия. Избирате плик, като можете да задържите парите в него. Преди да сте отворили плика обаче, ви се дава възможност да го смените. Ще го направите ли?

Решение. Ако сумите в пликовете са x и $2x$ лв., то очакваната печалба преди започване на играта е $1.5x$ лв. Нека сл. в. Y отбелязва дали сме избрали плика с по-голямата сума на първа стъпка, като $\mathbb{P}(Y = 1) = \mathbb{P}(Y = 0) = 0.5$, а сл. в. X – отчита печалбата при смяна на писмата. Тогава

$$\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X|Y]\big] = \mathbb{E}[X|Y=1]\mathbb{P}(Y=1) + \mathbb{E}[X|Y=0]\mathbb{P}(X=0) = 0.5\cdot x + 0.5\cdot 2x = 1.5x;$$

т.е. няма промяна в очакваната печалба. □

3.4 Характеристики на случайни величини

"A generating function is a device somewhat similar to a bag. Instead of carrying many little objects detachedly, which could be embarrassing, we put them all in a bag, and then we have only one object to carry, the bag."

– George Pólya , *Mathematics and Plausible Reasoning* (1954)

В зависимост от конкретната задача използването на някоя от следните функции може значително да улесни пресмятането на вероятности или други количества.

Функция на разпределение *на сл. в.* X *наричаме функцията*

$$F_X(x) = \mathbb{P}(X < x), \qquad x \in \mathbb{R}.$$

Функция на разпределение на многомерна сл. в. $X = (X_1, \ldots, X_m)$ *наричаме функцията*

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_m) = \mathbb{P}(X_1 < x_1, \ldots, X_m < x_m) = \mathbb{P}\big(X \in (-\infty, x_1) \times \cdots \times (-\infty, x_m)\big), \qquad x_1, \ldots, x_m \in \mathbb{R}.$$

Забележка (Следствия за дискретни сл. в.). Понеже $F_X(x) = \sum_{y<x} \mathbb{P}(X = y)$, следва, че $\mathbb{P}(X = x) = F_X(x+1) - F_X(x)$, т.е. F_X определя напълно и еднозначно вероятностното разпределение на X. Следователно за две дискретни сл. в. X и Y, такива че $F_X = F_Y$, имаме $X \stackrel{d}{=} Y$. По подобен начин разпределението на X се определя напълно и еднозначно от всяко едно от $\{\mathbb{P}(X > x)\}_{x\in\mathbb{R}}$, $\{\mathbb{P}(X \le x)\}_{x\in\mathbb{R}}$, $\{\mathbb{P}(X \ge x)\}_{x\in\mathbb{R}}$.

Забележка (Свойства). Функцията F_X удовлетворява следните условия: (*i*) F_X е ненамаляваща, т.е. ако $x_1 < x_2$, то $F_X(x_1) \le F_X(x_2)$; (*ii*) F_X е непрекъсната отляво, т.е. $\lim_{y\uparrow x} F_X(u) = F_X(x)$ за всяко $x \in \mathbb{R}$; (*iii*) $\lim_{x\to\infty} F_X(x) = 1$ и $\lim_{x\to-\infty} F_X(x) = 0$.

Пораждаща функция *на целочислена дискретна сл. в.* X *наричаме функцията*

$$G_X(s) = \mathbb{E}[s^X] = \sum_{n=0}^{\infty} s^n \mathbb{P}(X = n), \qquad s \in \mathbb{R} : |s| \le 1.$$

Пораждаща функция на многомерна сл. в. $X = (X_1, \ldots, X_m)$*, приемаща стойности на решетката* $\mathbb{N}_0^m$*, наричаме функцията*

$$G_{X_1,\ldots,X_m}(s_1, \ldots, s_m) = \sum_{k_1,\ldots,k_n} s_1^{k_1} s_2^{k_2} \cdots s_m^{k_m} \mathbb{P}(X_1 = k_1, \ldots, X_m = k_m), \qquad s_1, \ldots, s_m \in \mathbb{R} : |s_i| \le 1.$$

*Забележка** (Следствия). От $G_{X_1,\ldots,X_m}(s_1, \ldots, s_m)$ лесно се извежда пораждащата функция на $X_{i_1}, \ldots, X_{i_k}$ за кои да е $i_1, \ldots, i_k \in \{1, \ldots, m\}$, $i_j \ne i_l$, при $k = 1, \ldots, m$. Например

$$G_{X_1,X_2}(s_1, s_2) = \sum_{k_1,k_2} s_1^{k_1} s_2^{k_2} \mathbb{P}(X_1 = k_1, X_2 = k_2) = \sum_{k_1,\ldots,k_m} s_1^{k_1} s_2^{k_2} \mathbb{P}(X_1 = k_1, \ldots, X_m = k_m) = G_{X_1,\ldots,X_m}(s_1, s_2, 1, \ldots, 1).$$

Твърдение 3.9. *Нека* X*,* X_1 *и* X_2 *са целочислени дискретни сл. в. Тогава*

$$\mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!}, \qquad \mathbb{E}[X] = G'_X(1), \qquad \mathrm{Var}(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2,$$

при положение, че съответните количества съществуват. Ако $X_1 \perp\!\!\!\perp X_2$*, то*

$$G_{X_1+X_2}(s) = G_{X_1}(s) G_{X_2}(s).$$

Забележка (Еднозначност). От горното следва, че пораждащата функция на X определя напълно и еднозначно вероятностното разпределение $\mathbb{P}_X$. Следователно за сл. в. X и Y, такива че $G_X = G_Y$, имаме $X \stackrel{d}{=} Y$.

Твърдение 3.10.** *Нека* X *е целочислена сл. в. с пораждаща функция* G_X *и функция на разпределение* F_X*. Тогава*

$$\sum_{n=0}^{\infty} s^n F_X(n) = \frac{s}{1-s} G_X(s).$$

Д-во. Използвайки равенството $\sum_{n=0}^{\infty}\sum_{m=n}^{\infty} a_{mn} = \sum_{n\le m} a_{mn} = \sum_{m=0}^{\infty}\sum_{n=0}^{m} a_{mn}$ за абсолютно сходящи редове, имаме

$$\sum_{n=0}^{\infty} s^n \mathbb{P}(X \ge n) = \sum_{n=0}^{\infty}\sum_{m=n}^{\infty} s^n \mathbb{P}(X=m) = \sum_{m=0}^{\infty}\left(\sum_{n=0}^{m} s^n\right)\mathbb{P}(X=m) = \sum_{m=0}^{\infty} \frac{1-s^{m+1}}{1-s}\mathbb{P}(X=m)$$
$$= \frac{1}{1-s}\sum_{m=0}^{\infty}\mathbb{P}(X=m) - \frac{s}{1-s}\sum_{m=0}^{\infty} s^m \mathbb{P}(X=m) = \frac{1}{1-s} - \frac{s}{1-s}G_X(s) = \frac{1-sG_X(s)}{1-s},$$

откъдето $\sum_{n=0}^{\infty} s^n F_X(n) = \sum_{n=0}^{\infty} s^n (1-\mathbb{P}(X\ge n)) = \frac{1}{1-s} - \frac{1-sG_X(s)}{1-s} = \frac{sG_X(s)}{1-s}$. □

Твърдение 3.11. *Ако X е целочислена сл. в., тогава*

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} \mathbb{P}(X > n).$$

Д-во.

$$\mathbb{E}[X] = \sum_{m=1}^{\infty} m\mathbb{P}(X=m) = \sum_{m=1}^{\infty}\sum_{n=1}^{m}\mathbb{P}(X=m) = \sum_{n=1}^{\infty}\sum_{m=n}^{\infty} m\mathbb{P}(X=m) = \sum_{n=1}^{\infty}\mathbb{P}(X\ge n).$$

От друга страна $1-(1-s)\sum_{n=0}^{\infty} s^n\mathbb{P}(X>n) = G_X(s)$, откъдето директно $\mathbb{E}[X] = G'_X(1) = \sum_{n=0}^{\infty}\mathbb{P}(X>n)$. □

3.5 Основни дискретни разпределения

Разпределение на Бернули *с параметър $p \in [0,1]$ е дискретно разпределение на булева сл. в. X, приемаща стойностите 0 и 1 с вероятности*

$$\mathbb{P}(X=1) = p, \qquad \textit{и} \qquad \mathbb{P}(X=0) = 1-p.$$

В този случай ще обозначаваме $X \sim \mathrm{Ber}(p)$.

Забележка (Интерпретация). Тази сл. в. може да се разглежда като резултата от експеримент, който може да се окачестви единствено като "успех" с вероятност p и "неуспех" с вероятност $(1-p)$.

Твърдение 3.12. *Ако $X \sim \mathrm{Ber}(p)$, тогава $\mathbb{E}[X] = p$, $\mathrm{Var}(X) = p(1-p)$ и $G_X(s) = ps + (1-p)$.*

Д-во. Резултатът следва директно от дефиницията на съответното количество. □

. . . .

Биномно разпределение *с параметри $n \in \mathbb{N}$ и $p \in [0,1]$ е дискретно разпределение на сл. в. X, приемаща стойностите $0,1,\dots,n$ съответно с вероятности*

$$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0,1,\dots,n.$$

В този случай ще обозначаваме $X \sim \mathrm{Bin}(n,p)$.

Забележка (Интерпретация). Тази сл. в. може да се интерпретира като броя на успехите в поредица от n независими Бернулиеви опити.

Забележка (Пример). Нечестна монета с вероятност за ези $p \in [0,1]$ се хвърля n пъти. Да се намери разпределението на броя на падналите се езита.

*Забележка*** (Най-вероятна стойност). Лесно се вижда, че

$$\frac{\mathbb{P}(X=k)}{\mathbb{P}(X=k-1)} = \frac{n-k+1}{k}\left(\frac{1}{1-p} - 1\right) = 1 + \frac{1}{k(1-p)}\left[p(n+1) - k\right];$$

т.е. $\mathbb{P}(X=k) \ge \mathbb{P}(X=k-1)$ за $k < p(n+1)$ и $\mathbb{P}(X=k) \le \mathbb{P}(X=k-1)$ за $k \ge p(n+1)$, като достига своя максимум в $k^* = \lfloor p(n+1) \rfloor$. Ако числото $p(n+1)$ е цяло, то $\mathbb{P}(X=k^*) = \mathbb{P}(X=k^*-1)$ от по-горе. С други думи, най-вероятните биномни вероятности са групирани около математическото очакване на X.

Твърдение 3.13. *Ако $X \sim \mathrm{Bin}(n,p)$, тогава $\mathbb{E}[X] = np$, $\mathrm{Var}(X) = np(1-p)$, $G_X(s) = (ps + (1-p))^n$.*

Д-во.* Първо $\sum_{k=0}^n \mathbb{P}(X=k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1$, използвайки биномната теорема; т.е. биномното разпределение е добре дефинирано. По същия начин

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^n k\binom{n}{k} p^k(1-p)^{n-k} = \sum_{k=1}^n n\binom{n-1}{k-1}p^k(1-p)^{n-k} = np\sum_{j=0}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-j} = np\big(p+(1-p)\big)^{n-1} = np,\\
\mathbb{E}[X^2] &= \sum_{k=1}^n k^2\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=1}^n kn\binom{n-1}{k-1}p^k(1-p)^{n-k} = \sum_{k=1}^n (k-1)n\binom{n-1}{k-1}p^k(1-p)^{n-k} + \sum_{k=1}^n n\binom{n-1}{k-1}p^k(1-p)^{n-k}\\
&= \sum_{k=2}^n n(n-1)\binom{n-2}{k-2}p^k(1-p)^{n-k} + np = n(n-1)p^2 + np,\\
\mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = n(n-1)p^2 + np - (np)^2 = np(1-p),\\
G_X(s) &= \sum_{k=0}^n s^k\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k}(ps)^k(1-p)^{n-k} = \big(ps + (1-p)\big)^n.
\end{aligned}$$

Алт. д-во. Нека приемем, че съществуват $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathrm{Ber}(p)$. Тогава, за всяко $k = 0, 1, \ldots, n$,

$$\begin{aligned}
\mathbb{P}(X_1 + \cdots + X_n = k) &= \sum_{j_1,\ldots,j_n \in \{0,1\}: j_1+\cdots+j_n = k} \mathbb{P}(X_1 = j_1, \ldots, X_n = j_n)\\
&= \binom{n}{k}\mathbb{P}(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0) = \binom{n}{k}p^k(1-p)^{n-k};
\end{aligned}$$

т.е. $X \overset{d}{=} \sum_{i=1}^n X_i$, откъдето $\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^n X_i] = np$, $\mathrm{Var}(X) = np(1-p)$ и $G_X(s) = G_{\sum_{i=1}^n X_i}(s) = (G_{X_1}(s))^n$. □

Задача 3.18. Нека $X_1 \sim \mathrm{Bin}(n_1, p)$ и $X_2 \sim \mathrm{Bin}(n_2, p)$ са две независими случайни величини. Да се докаже, че $X = X_1 + X_2 \sim \mathrm{Bin}(n_1 + n_2, p)$.

Решение. За всяко $k = 0, 1, \ldots, n_1 + n_2$,

$$\begin{aligned}
\mathbb{P}(X = k) &= \sum_{i=\max\{0,k-n_2\}}^{\min\{n_1,k\}} \mathbb{P}(X = k, X_1 = i) = \sum_{i=\max\{0,k-n_2\}}^{\min\{n_1,k\}} \mathbb{P}(X_1 = i)\mathbb{P}(X_2 = k-i)\\
&= \sum_i \binom{n_1}{i}p^i(1-p)^{n_1-i}\binom{n_2}{k-i}p^{k-i}(1-p)^{n_2-k+i} = p^k(1-p)^{n_1+n_2-k}\sum_i\binom{n_1}{i}\binom{n_2}{k-i} = \binom{n_1+n_2}{k}p^k(1-p)^{n_1+n_2-k}.
\end{aligned}$$

От друга страна, директно, $G_X(s) = G_{X_1}(s)G_{X_2}(s) = (ps - (1-p))^{n_1}(ps - (1-p))^{n_2} = (ps - (1-p))^{n_1+n_2}$. □

Задача 3.19. Нека $N \sim \mathrm{Bin}(M, q)$ и $X|N = n \sim \mathrm{Bin}(n, p)$, $n = 1, \ldots, M$, за $p, q \in [0,1]$ и $M \in \mathbb{N}$. Да се докаже, че $X \sim \mathrm{Bin}(M, pq)$.

Решение. От формулата за пълната вероятност получаваме, за всяко $k = 0, 1, \ldots, M$,

$$\begin{aligned}
\mathbb{P}(X = k) &= \sum_{n=k}^M \mathbb{P}(X = k|N = n)\mathbb{P}(N = n) = \sum_{n=k}^M \binom{n}{k}p^k(1-p)^{n-k}\binom{M}{n}q^n(1-q)^{M-n}\\
&= \binom{M}{k}p^kq^k\sum_{n=k}^M\binom{M-k}{n-k}(q(1-p))^{n-k}(1-q)^{M-n} = \binom{M}{k}(pq)^k\big(q(1-p) + (1-q)\big)^{M-k} = \binom{M}{k}(pq)^k\big(1-pq\big)^{M-k}.
\end{aligned}$$

□

Задача 3.20. Нека $X|P = p \sim \mathrm{Bin}(200, p)$, където параметърът P е сл. в., такава че $\mathbb{P}(P = \frac{1}{2}) = \mathbb{P}(P = \frac{2}{3}) = \frac{1}{2}$. Коя стойност на параметъра P има по-голяма *апостериорна* вероятност при условие, че $X = 120$.

Решение. От свойствата на условните вероятности,

$$\frac{\mathbb{P}(P = \frac{1}{2}|X = 120)}{\mathbb{P}(P = \frac{2}{3}|X = 120)} = \frac{\mathbb{P}(X = 120|P = \frac{1}{2})\mathbb{P}(P = \frac{1}{2})}{\mathbb{P}(X = 120)} \cdot \frac{\mathbb{P}(X = 120)}{\mathbb{P}(X = 120|P = \frac{2}{3})\mathbb{P}(P = \frac{2}{3})} = \frac{\binom{200}{120}\left(\frac{1}{2}\right)^{200}\frac{1}{2}}{\binom{200}{120}\left(\frac{2}{3}\right)^{120}\left(\frac{1}{3}\right)^{80}\frac{1}{2}} = \frac{3^{200}}{2^{320}} = 0.124.$$

□

Задача 3.21 (Нееднородно биномно разпределение). Вероятностите за изгаряне на първа, втора и трета лампа са равни съответно на 0.1, 0.2 и 0.3. Вероятностите за излизане на прибора от строя при изгарянето на една, две и три лампи са равни съответно на 0.25, 0.6 и 0.9. Да се пресметне вероятността приборът да излезе от строя.

Решение (Техника). Нека $A = \{\text{приборът излиза от строя}\}$ и $X_1 \sim \mathrm{Ber}(0.1)$, $X_2 \sim \mathrm{Ber}(0.2)$, $X_3 \sim \mathrm{Ber}(0.3)$ са независими сл. в. Тогава вероятността да изгорят точно $k = 1, 2, 3$ лампи, $\mathbb{P}(X_1 + X_2 + X_3 = k)$, може да се изведе по стандартния начин. От друга страна, можем да използваме пораждащата функция на $X_1+X_2+X_3$, като $\mathbb{P}(X_1 + X_2 + X_3 = k)$ ще бъде равно на коефициента пред s^k в развитието

$$\mathbb{E}[s^{X_1+X_2+X_3}] = \mathbb{E}[s^{X_1}]\mathbb{E}[s^{X_2}]\mathbb{E}[s^{X_3}] = (0.1s + 0.9)(0.2s + 0.8)(0.3s + 0.7) = (0.006s^3 + 0.092s^2 + 0.398s + 0.504),$$

откъдето $\mathbb{P}(A) = \sum_{k=1}^{3} \mathbb{P}(A|X_1 + X_2 + X_3 = k)\mathbb{P}(X_1 + X_2 + X_3 = k) = 0.16.$ □

*Задача*** 3.22.* Нека $X \sim \mathrm{Bin}(n, p)$. Каква е вероятността X да бъде четно число?

Решение (Техника). Нека p_k е вероятността за четен брой успехи при $k \le n$ опити с вероятност за успех p. Тогава

$$p_k = \mathbb{P}(\text{четен от } k-1 \text{ опита, неуспех } k\text{-ти опит}) + \mathbb{P}(\text{нечетен от } k-1 \text{ опита, успех } k\text{-ти опит}) = p_{k-1}(1-p) + (1-p_{k-1})p,$$

откъдето извеждаме $p_k - \frac{1}{2} = (1-2p)(p_{k-1} - \frac{1}{2})$. Чрез заместване, $p_n - \frac{1}{2} = (1-2p)^n(p_0 - \frac{1}{2})$, като от началното условие, $p_0 = 1$, имаме $\mathbb{P}(X = 0, 2, 4, \ldots, 2\lfloor \frac{n}{2} \rfloor) = p_n = \frac{1}{2}\big[1 + (1-2p)^n\big]$. □

....

Геометрично разпределение *с параметър $p \in [0,1]$ е дискретно разпределение на сл. в. X, приемаща безброй много стойности $0, 1, 2, \ldots$ съответно с вероятности*

$$\mathbb{P}(X = k) = (1-p)^k p, \quad k = 0, 1, 2, \ldots.$$

В този случай ще обозначаваме $X \sim \mathrm{Geo}(p)$.

Забележка (Интерпретация). Тази сл. в. може да се интерпретира като броя на неуспехите до настъпването на първия успех в поредица от независими Бернулиеви опити.

Забележка (Безпаметност). Геометричното разпределение е единственото дискретно разпределение, което има свойството, за $k, m \in \mathbb{N}_0$,

$$\mathbb{P}(X \ge m + k | X \ge k) = \frac{\mathbb{P}(X \ge m + k)}{\mathbb{P}(X \ge k)} = \frac{(1-p)^{m+k}}{(1-p)^k} = (1-p)^m = \mathbb{P}(X \ge m).$$

С други думи, ако първият успех все още не е настъпил, условното вероятностно разпределение на броя на допълнителните опити не зависи от това колко неуспеха са били вече наблюдавани.

*Забележка*** (Схема на Бернули). Нека приемем, че съществуват $X_1, X_2, \ldots \overset{i.i.d.}{\sim} \mathrm{Ber}(p)$. Тогава, за всяко $k = 0, 1, 2, \ldots$,

$$\mathbb{P}\Big(\min_n\Big\{\sum_{i=1}^{n} X_i = 1\Big\} - 1 = k\Big) = \mathbb{P}(X_1 = 0, \ldots, X_k = 0, X_{k+1} = 1) = (1-p)^k p;$$

т.е. $X \overset{d}{=} \min_n\{\sum_{i=1}^{n} X_i = 1\} - 1$.

Твърдение 3.14. *Ако $X \sim \mathrm{Geo}(p)$, тогава $\mathbb{E}[X] = \frac{1-p}{p}$, $\mathrm{Var}(X) = \frac{1-p}{p^2}$ и $G_X(s) = \frac{p}{1-s(1-p)}$.*

Д-во.* Първо $\sum_{k=0}^{\infty}(1-p)^k p = p\sum_{k=0}^{\infty}(1-p)^k = \frac{p}{1-(1-p)} = 1$, т.е. геометричното разпределение е добре дефинирано. Използвайки развитието в ред на $(1-x)^{-n} = \sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j$ в точката $x_0 = 0$,

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k(1-p)^k p = p\sum_{j=0}^{\infty}(j+1)(1-p)^{j+1} = p(1-p)\sum_{j=0}^{\infty}\binom{j+1}{j}(1-p)^j = \frac{p(1-p)}{(1-(1-p))^2} = \frac{1-p}{p}, \\
\mathbb{E}[X^2] &= \sum_{k=1}^{\infty} k^2(1-p)^k p = \sum_{k=1}^{\infty} k\binom{k}{k-1}(1-p)^k p = \sum_{k=1}^{\infty}(k-1)\binom{k}{k-1}(1-p)^k p + \sum_{k=1}^{\infty} k(1-p)^k p \\
&= 2p(1-p)^2\sum_{k=2}^{\infty}\binom{k}{k-2}(1-p)^{k-2} + \frac{(1-p)}{p} = \frac{2p(1-p)^2}{p^3} + \frac{(1-p)}{p} = \frac{(1-p)(2-p)}{p^2}, \\
\mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{(1-p)(2-p)}{p^2} - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2}, \\
G_X(s) &= \sum_{k=0}^{\infty} s^k(1-p)^k p = p\sum_{k=0}^{\infty}(s(1-p))^k = \frac{p}{1-s(1-p)}.
\end{aligned}$$

Алт. д-во.

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k(1-p)^k p = \sum_{k=1}^{\infty}(k-1)(1-p)^k p + \sum_{k=1}^{\infty}(1-p)^k p = (1-p)(\mathbb{E}[X]+1) \quad \Longrightarrow \quad \mathbb{E}[X] = \frac{1-p}{p}, \\
\mathbb{E}[X^2] &= \sum_{k=1}^{\infty} k^2(1-p)^k p = \sum_{k=1}^{\infty}(k-1+1)^2(1-p)^k p = \sum_{k=1}^{\infty}(k-1)^2(1-p)^k p + 2\sum_{k=1}^{\infty}(k-1)(1-p)^k p + \sum_{k=1}^{\infty}(1-p)^k p \\
&= (1-p)\big(\mathbb{E}[X^2] + 2\mathbb{E}[X] + 1\big) \quad \Longrightarrow \quad \mathbb{E}[X^2] = \frac{(1-p)(2-p)}{p^2}.
\end{aligned}$$

□

Задача 3.23. A и B стрелят едновременно по мишена, която A улучва с вероятност p_1, а B – с p_2. Ако никой не улучи мишената, двамата стрелят отново. Какъв е средният брой изстрели, необходими за уцелване на мишената? А колко средно изстрела биха направили двамата, ако имаха право на максимум n опита?

Решение. Нека сл. в. $X \sim \mathrm{Geo}(p)$ брои изстрелите преди уцелването на мишената, където $p = p_1 + p_2 - p_1p_2$. Тогава $\mathbb{E}[X+1] = \frac{1}{p}$. По отношение на втория въпрос, нека сл. в. Y отчита изстрелите, направени от n опита,

Y	1	2	$\cdots$	$n-1$	n
$\mathbb{P}_Y$	p	$(1-p)p$	$\cdots$	$(1-p)^{n-2}p$	$(1-p)^{n-1}$

като $\sum_{i=1}^{n}\mathbb{P}(Y=i) = p\sum_{i=1}^{n-1}(1-p)^{i-1} + (1-p)^{n-1} = p\frac{1-(1-p)^{n-1}}{1-(1-p)} + (1-p)^{n-1} = 1$, използвайки равенството $\sum_{k=0}^{n-1} x^k = \frac{1-x^n}{1-x}$ за $x<1$. По същия начин, от $\sum_{k=0}^{n-1}(k+1)x^k = \frac{1-x^n}{(1-x)^2} - \frac{nx^n}{1-x}$ за $x<1$ следва

$$\begin{aligned}
\mathbb{E}[Y] &= p\sum_{i=1}^{n-1} i(1-p)^{i-1} + n(1-p)^{n-1} \\
&= p\left(\frac{1-(1-p)^{n-1}}{p^2} - \frac{(n-1)(1-p)^{n-1}}{p}\right) + n(1-p)^{n-1} = \frac{1-(1-p)^{n-1}}{p} + (1-p)^{n-1} = \frac{1-(1-p)^n}{p}.
\end{aligned}$$

□

Задача 3.24. Нека $X \sim \mathrm{Geo}(p)$ и $Y \sim \mathrm{Geo}(p)$ са две независими случайни величини. Да се намери разпределението на (а) $\min\{X,Y\}$; (б) $\max\{X,Y\}$.

Решение. (а) За всяко $k = 0,1,2,\ldots$,

$$\begin{aligned}
\mathbb{P}(\min\{X,Y\}=k) &= \mathbb{P}(X=k, Y>k) + \mathbb{P}(Y=k, X>k) + \mathbb{P}(X=k, Y=k) = 2(1-p)^k p(1-p)^{k+1} + \big((1-p)^k p\big)^2 \\
&= \big((1-p)^2\big)^k\big(2p(1-p)+p^2\big) = \big(1-(2p-p^2)\big)^k(2p-p^2),
\end{aligned}$$

т.е. $\min\{X,Y\} \sim \mathrm{Geo}(2p-p^2)$. Алтернативно, $\mathbb{P}(\min\{X,Y\} \geq k) = \mathbb{P}(X \geq k, Y \geq k) = (1-p)^k(1-p)^k = (1-(2p-p^2))^k$. (б) $\mathbb{P}(\max\{X,Y\} < k) = \mathbb{P}(X<k, Y<k) = (1-(1-p)^k)^2$. □

Задача 3.25 (Coupon Collector's Problem). Всеки пакет зърнчо съдържа една от n различни вида играчки. Намерете очаквания брой пакети, докато бъдат събрани всички n играчки.

Решение. Нека сл. в. X_i отчита броя на опаковките до намирането на i-тата играчка, след като вече имаме $i-1$ играчки, $i = 1, \dots, n$. Тогава $X = \sum_{i=1}^n X_i + n$ е броят на пакетите, необходими за събирането на всички играчки. Понеже $X_i \sim \mathrm{Geo}(\frac{n-i+1}{n})$, имаме

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] + n = \sum_{i=1}^n \frac{1 - \frac{n-i+1}{n}}{\frac{n-i+1}{n}} + n = \sum_{i=1}^n \left(\frac{i-1}{n-i+1} + 1 \right) = \sum_{i=1}^n \frac{n}{n-i+1} = n \sum_{k=1}^n \frac{1}{k}.$$

От друга страна разпределението на X може да се изведе подобно на Задача 1.3в, като $\{X = k\}$ е събитието "k различими частици в n клетки, където нито една клетка не остава празна и с последната частица запълваме последната празна клетка". □

. . . .

Отрицателно биномно разпределение *с параметри $n \in \mathbb{N}$ и $p \in [0,1]$ е дискретно разпределение на сл. в. X, приемаща безброй много стойности $0, 1, 2, \dots$ съответно с вероятности*

$$\mathbb{P}(X = k) = \binom{n+k-1}{k} p^n (1-p)^k, \quad k = 0, 1, 2, \dots.$$

В този случай ще обозначаваме $X \sim \mathrm{NB}(n, p)$.

Забележка (Интерпретация). Тази сл. в. може да се интерпретира като броя на неуспехите до настъпването на n-тия успех в поредица от независими Бернулиеви опити.

*Забележка*** (Най-вероятна стойност). Лесно се вижда, че

$$\frac{\mathbb{P}(X = k)}{\mathbb{P}(X = k-1)} = \frac{(n+k-1)(1-p)}{k} = 1 + \frac{p}{k}\Big[(n-1)\frac{1-p}{p} - k\Big];$$

т.е. $\mathbb{P}(X = k) \geq \mathbb{P}(X = k-1)$ за $k < \frac{(n-1)(1-p)}{p}$ и $\mathbb{P}(X = k) \leq \mathbb{P}(X = k-1)$ за $k \geq \frac{(n-1)(1-p)}{p}$, като достига своя максимум в $k^* = \lfloor \frac{(n-1)(1-p)}{p} \rfloor$. Ако числото $\frac{(n-1)(1-p)}{p}$ е цяло, то $\mathbb{P}(X = k^*) = \mathbb{P}(X = k^* - 1)$ от по-горе.

*Забележка*** (Отрицателни биномни коефициенти). Нека $\binom{-n}{k} = \frac{-n \cdot (-n-1) \cdots (-n-k+1)}{k!}$. Тогава

$$\binom{n+k-1}{k} = \frac{n \cdot (n+1) \cdots (n+k-1)}{k!} = (-1)^k \frac{-n \cdot (-n-1) \cdots (-n-k+1)}{k!} = (-1)^k \binom{-n}{k}.$$

Твърдение 3.15. *Ако $X \sim \mathrm{NB}(n,p)$, тогава $\mathbb{E}[X] = n\frac{1-p}{p}$, $\mathrm{Var}(X) = n\frac{1-p}{p^2}$, $G_X(s) = \left(\frac{p}{1-s(1-p)}\right)^n$.*

Д-во.* Първо $\sum_{k=0}^\infty \binom{n+k-1}{k} p^n (1-p)^k = p^n \sum_{k=0}^\infty \binom{n+k-1}{k} (1-p)^k = p^n (1-(1-p))^{-n} = p^n p^{-n} = 1$, използвайки развитието в ред на $(1-x)^{-n} = \sum_{k=0}^\infty \binom{n+k-1}{k} x^k$ в точката $x_0 = 0$; т.е. отрицателното биномно разпределение е добре дефинирано. По същия начин

$$\mathbb{E}[X] = \sum_{k=1}^\infty k \binom{n+k-1}{k} p^n (1-p)^k = p^n \sum_{k=1}^\infty n \binom{n+k-1}{k-1} (1-p)^k = np^n(1-p) \sum_{j=0}^\infty \binom{n+j}{j} (1-p)^j = \frac{np^n(1-p)}{(1-(1-p))^{n+1}} = \frac{n(1-p)}{p},$$

$$\mathbb{E}[X^2] = \sum_{k=1}^\infty k^2 \binom{n+k-1}{k} p^n (1-p)^k = np^n \sum_{k=1}^\infty k \binom{n+k-1}{k-1} (1-p)^k = np^n \sum_{k=1}^\infty (k-1) \binom{n+k-1}{k-1} (1-p)^k + \frac{n(1-p)}{p}$$

$$= n(n+1)p^n \sum_{k=2}^\infty \binom{n+k-1}{k-2} (1-p)^k + \frac{n(1-p)}{p} = \frac{n(n+1)p^n(1-p)^2}{(1-(1-p))^{n+2}} + \frac{n(1-p)}{p} = \frac{n(1-p)(n-np+1)}{p^2},$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{n(1-p)(n-np+1)}{p^2} - \frac{n^2(1-p)^2}{p^2} = \frac{n(1-p)}{p^2},$$

$$G_X(s) = \sum_{k=0}^\infty s^k \binom{n+k-1}{k} p^n (1-p)^k = p^n \sum_{k=0}^\infty \binom{n+k-1}{k} (s(1-p))^k = p^n (1 - s(1-p))^{-n}.$$

Алт. д-во. Нека приемем, че съществуват $X_1, \dots, X_n \overset{i.i.d.}{\sim} \mathrm{Geo}(p)$. Тогава, за всяко $k = 0, 1, 2, \dots$,

$$\mathbb{P}(X_1 + \cdots + X_n = k) = \sum_{j_1, \dots, j_n \in \{0, 1, \dots, k\} : j_1 + \cdots + j_n = k} \mathbb{P}(X_1 = j_1, \dots, X_n = j_n)$$

$$= \binom{n+k-1}{k}\mathbb{P}(X_1 = k, X_2 = 0, \dots, X_n = 0) = \binom{n+k-1}{k} p^n(1-p)^k;$$

т.е. $X \stackrel{d}{=} \sum_{i=1}^n X_i$, откъдето $\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^n X_i] = n\frac{1-p}{n}$, $\mathrm{Var}(X) = n\frac{1-p}{p^2}$ и $G_X(s) = G_{\sum_{i=1}^n X_i}(s) = (G_{X_1}(s))^n$. $\square$

Задача 3.26 (Banach's Matchbox Problem). Пушач носи в джоба си две кутии кибрит с по n клечки. Всеки път, когато иска да запали, той избира произволна кутия и вади една клечка. След известно време той забелязва, че едната кутия е празна. Каква е вероятността в този момент в другата да са останали точно $k \le n$ клечки?

Решение. Нека сл. в. $X_1 \sim \mathrm{NB}(n+1, \frac{1}{2})$ и $X_2 \sim \mathrm{NB}(n+1, \frac{1}{2})$ проследяват броя на тегленията съответно от кутия 2 и 1, докато пушачът види, че другата кутия е празна. Тогава $\{X_1 = n-k\} \cup \{X_2 = n-k\}$ е събитието "пушачът за пръв път забелязва, че едната кутия е празна, докато в другата има k останали клечки", откъдето

$$\mathbb{P}(\{X_1 = n-k\} \cup \{X_2 = n-k\}) = \mathbb{P}(X_1 = n-k) + \mathbb{P}(X_2 = n-k) = 2\binom{2n-k}{n-k}\frac{1}{2^{2n-k+1}} = \binom{2n-k}{n-k}\frac{1}{2^{2n-k}},$$

използвайки, че $\{X_1 = n-k\} \cap \{X_2 = n-k\} = \emptyset$. $\square$

. . . .

Хипергеометрично разпределение *с параметри* $M, N, n \in \mathbb{N}$, *такива че* $M < N, n < N, n < N - M$, *е дискретно разпределение на сл. в.* X, *приемаща стойностите* $0, 1, \dots, \min(M, n)$ *съответно с вероятности*

$$\mathbb{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad k = 0, 1, \dots, \min(M, n).$$

В този случай ще обозначаваме $X \sim \mathrm{HG}(M, N, n)$.

Забележка (Интерпретация). Тази сл. в. може да се интерпретира като броя на успехите (т.е. изтегленият обект да има определена характеристика) в n тегления без връщане от крайна популация с размер N, която съдържа точно M обекта с тази характеристика.

Забележка (Пример/Задача 2.4). В урна има N топки, от които M са черни, а останалите $N - M$ са бели ($M < N$). По случаен начин от урната се изваждат последователно без връщане n топки ($n < N, n < N - m$). Да се намери вероятността точно $k \le \min(M, n)$ от извадените топки да са черни.

Твърдение 3.16. *Ако* $X \sim \mathrm{HG}(M, N, n)$, *тогава* $\mathbb{E}[X] = n\frac{M}{N}$ *и* $\mathrm{Var}(X) = n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$.

Д-во.* Първо $\sum_{k=0}^{\min(M,n)} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{\sum_{k=0}^{\min(M,n)}\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1$, използвайки резултатите от Задача 1.9; т.е. хипергеометричното разпределение е добре дефинирано. По същия начин

$$\mathbb{E}[X] = \sum_{k=1}^{\min(M,n)} k\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \sum_{k=1}^{\min(M,n)} M\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\frac{N}{n}\binom{N-1}{n-1}} = \frac{nM}{N}\sum_{j=0}^{\min(M-1,n-1)}\frac{\binom{M-1}{j}\binom{N-M}{(n-1)-j}}{\binom{N-1}{n-1}} = \frac{nM}{N}\frac{\binom{N-1}{n-1}}{\binom{N-1}{n-1}} = \frac{nM}{N},$$

$$\mathbb{E}[X^2] = \sum_{k=1}^{\min(M,n)} k^2\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{nM}{N}\sum_{k=1}^{\min(M,n)} k\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N-1}{n-1}} = \frac{nM}{N}\sum_{k=1}^{\min(M,n)}(k-1)\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N-1}{n-1}} + \frac{nM}{N}\sum_{k=1}^{\min(M,n)}\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N-1}{n-1}}$$

$$= \frac{nM}{N}\sum_{k=2}^{\min(M,n)}(M-1)\frac{\binom{M-2}{k-2}\binom{N-M}{n-k}}{\frac{N-1}{n-1}\binom{N-2}{n-2}} + \frac{nM}{N} = \frac{nM(M-1)(n-1)}{N(N-1)} + \frac{nM}{N} = \frac{nM(Mn+N-M-n)}{N(N-1)},$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{nM(Mn+N-M-n)}{N(N-1)} - \frac{n^2M^2}{N^2} = \frac{nM(N-M)(N-n)}{N^2(N-1)}.$$

$\square$

*Забележка*** (Връзка с биномното разпределение). За разлика от хипергеометричното разпределение биномното разпределение описва вероятността за броя на успехите в n тегления *с връщане*. Ако положим $p = \frac{M}{N}$, виждаме, че математическото очакване на биномното и хипергеометричното разпределение съвпадат, а дисперсиите им се различават с множителя $\frac{N-n}{N-1}$. Тогава при достатъчно голяма популация, спазвайки съотношението $\frac{M}{N} = p$, и малък размер на извадката n двете процедури за вземане на проби са приблизително еднакви. Действително, ако $\lim_{N\to\infty} \frac{M}{N} = p$, то

$$\lim_{N\to\infty} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \lim_{N\to\infty} \left\{ \frac{n!}{k!(n-k)!} \cdot \frac{M}{N} \cdots \frac{M-k+1}{N-k+1} \cdot \frac{N-M}{N-k} \cdots \frac{N-M-(n-k)+1}{N-n+1} \right\} = \binom{n}{k} p^k (1-p)^{n-k}.$$

*Забележка*** (Зависима схема на Бернули). Нека X е като в Задача 2.4, а сл. в. X_i отчита "успеха" на i-тия ход в n тегления без връщане. Тогава $X_1 + \cdots + X_n \sim \text{HG}(M, N, n)$. От Задача 2.26 следва, че $X_i \sim \text{Ber}(\frac{M}{N})$ за всяко $i = 1, \ldots, n$ откъдето $\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = n\frac{M}{N}$. Въпреки това сл. в. X_i не са независими. И все пак

$$\mathbb{E}[X_i X_j] = \mathbb{P}(X_i = 1, X_j = 1) = \mathbb{P}(X_i = 1)\mathbb{P}(X_j = 1 | X_i = 1) = \frac{M}{N}\frac{M-1}{N-1},$$

откъдето $\text{Cov}(X_i, X_j) = \frac{M}{N}\frac{M-1}{N-1} - \frac{M}{N}\frac{M}{N} = -\frac{M}{N}\frac{N-M}{N(N-1)} < 0$, $i \neq j$. Следователно

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) + 2\sum_{i<j} \text{Cov}(X_i, X_j) = n\frac{M}{N}\frac{N-M}{N} - 2\frac{n(n-1)}{2}\frac{M}{N}\frac{N-M}{N(N-1)} = n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}.$$

Задача 3.27. В кутия има 7 лампи, от които 3 са дефектни. По случаен начин се избират за проверка 4 лампи. Да се намери разпределението на случайната величина $X =$ "брой на изпробваните дефектни лампи$''$ и да се пресметне нейното очакване и дисперсия.

Отговор. $\mathbb{E}[X] = \frac{12}{7}$, $\text{Var}(X) = \frac{24}{49}$. □

Задача 3.28. От колода, съдържаща 52 карти, случайно се теглят 13 карти. Каква е вероятността две от избраните карти да бъдат червени, ако (а) картите се теглят с връщане; (б) картите се теглят без връщане?

Отговор. Нека $X \sim \text{Bin}(13, \frac{26}{52})$ и $Y \sim \text{HG}(26, 52, 13)$. (а) $\mathbb{P}(X = 2) = 0.01$; (б) $\mathbb{P}(Y = 2) = 0.004$. □

. . . .

Разпределение на Поасон *с параметър* $\lambda \in (0, \infty)$ *е дискретно разпределение на сл. в.* X*, приемаща безброй много стойности* $0, 1, 2, \ldots$ *с вероятности*

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

В този случай ще обозначаваме $X \sim \text{Po}(\lambda)$.

Забележка (Интерпретация). Тази сл. в. изразява вероятността даден брой събития да се случат в определен интервал от време или пространство, ако събитията се случват с постоянна средна честота и независимо от времето, изминало от последното събитие.

Забележка (Пример). Център за обслужване на клиенти получава средно 180 обаждания на час, 24 часа в денонощието. Обажданията са независими; получаването на едно от тях не променя вероятността за това кога ще пристигне следващото. Броят на обажданията, получени през която и да е минута, има Поасоново разпределение с параметър 3.

*Забележка*** (Най-вероятна стойност). Лесно се вижда, че $\frac{\mathbb{P}(X=k)}{\mathbb{P}(X=k-1)} = \frac{\lambda}{k}$; т.е. $\mathbb{P}(X = k) > \mathbb{P}(X = k-1)$ за $k < \lambda$ и $\mathbb{P}(X = k) \leq \mathbb{P}(X = k-1)$ за $k \geq \lambda$, като достига своя максимум в $k^* = \lfloor \lambda \rfloor$. Ако параметърът λ е цяло число, то $\mathbb{P}(X = k^*) = \mathbb{P}(X = k^* - 1)$ от по-горе.

*Забележка*** (Връзка с биномното разпределение). Нека $p = p_n \to 0$ по такъв начин, че $np \to \lambda \geq 0$ при $n \to \infty$, т.е. с увеличаване на броя на опитите вероятността за успех намалява пропорционално. Тогава

$$\begin{aligned}\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} &= \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n} + o\Big(\frac{1}{n}\Big)\right)^k \left(1 - \frac{\lambda}{n} + o\Big(\frac{1}{n}\Big)\right)^{n-k} \\ &= \lim_{n\to\infty} \frac{\lambda^k}{k!}\Big(1 - \frac{1}{n}\Big)\cdots\Big(1 - \frac{k-1}{n}\Big)\Big(1 - \frac{\lambda}{n}\Big)^{n-k}(1 + o(1)) = \frac{e^{-\lambda}\lambda^k}{k!}.\end{aligned}$$

Горното приближение позволява пресмятането на биномните вероятности, което в противен случай би предизвикало значителни технически трудности.

Твърдение 3.17. *Ако $X \sim \mathrm{Po}(\lambda)$, тогава $\mathbb{E}[X] = \lambda$, $\mathrm{Var}(X) = \lambda$ и $G_X(s) = e^{\lambda(s-1)}$.*

Д-во. Първо $\sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$, използвайки развитието в ред на $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ в точката $x{=}0$; т.е. разпределението на Поасон е добре дефинирано. По същия начин

$$\begin{aligned}
&\mathbb{E}[X] = \sum_{k=1}^{\infty} k\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda,\\
&\mathbb{E}[X^2] = \sum_{k=1}^{\infty} k^2\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^k}{(k-1)!} = e^{-\lambda}\sum_{k=1}^{\infty}(k-1)\frac{\lambda^k}{(k-1)!} + \lambda = \lambda^2 e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} + \lambda = \lambda^2+\lambda,\\
&\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \lambda^2+\lambda-\lambda^2 = \lambda,\\
&G_X(s) = \sum_{k=0}^{\infty} s^k\frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(c\lambda)^k}{k!} = e^{-\lambda}e^{s\lambda} = e^{\lambda(s-1)}.
\end{aligned}$$

□

Задача 3.29. Нека $X_1 \sim \mathrm{Po}(\lambda_1)$ и $X_2 \sim \mathrm{Po}(\lambda_2)$ са независими сл. в. Да се докаже, че $X = X_1+X_2 \sim \mathrm{Po}(\lambda_1+\lambda_2)$.

Решение. За всяко $k = 0, 1, 2, \ldots$,

$$\mathbb{P}(X=k) = \sum_{i=0}^{k}\mathbb{P}(X_1=i, X_2=k-i) = \sum_{i=0}^{k}\frac{e^{-\lambda_1}\lambda_1^i}{i!}\frac{e^{-\lambda_2}\lambda_2^{k-i}}{(k-i)!} = \frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k}\binom{k}{i}\lambda_1^i\lambda_2^{k-i} = \frac{e^{-(\lambda_1+\lambda_2)}}{k!}(\lambda_1+\lambda_2)^k.$$

От друга страна, директно, $G_X(s) = G_{X_1}(s)G_{X_2}(s) = e^{\lambda_1(s-1)}e^{\lambda_2(s-1)} = e^{(\lambda_1+\lambda_2)(s-1)}$. □

Задача 3.30. В Патагония на месец се регистрират средно две слаби земетресения. Каква е вероятността за три месеца да има по-малко от четири слаби земетресения?

Решение. Нека $X = X_1 + X_2 + X_3$, където $X_i \sim \mathrm{Po}(\lambda)$ е броят на земетресенията през i-тия месец, $i = 1, 2, 3$. По условие $\mathbb{E}[X_i] = 2$, откъдето $\lambda = 2$. Ако приемем, че X_1, X_2 и X_3 са независими в съвкупност, тогава $X \sim \mathrm{Po}(6)$ и $\mathbb{P}(X < 4) = 0.15$. □

*Задача** 3.31 (Poisson Thinning). Броят на проведените опити има Поасоново разпределение с параметър λ. Всеки опит може да бъде успешен с вероятност p и неуспешен с вероятност $1-p$. Да се построи разпределението на броя на успешните опити X.

Решение. За да има $k = 0, 1, 2, \ldots$ успешни опита, общият брой на опитите трябва да е поне k, откъдето

$$\mathbb{P}(X=k) = \sum_{n=k}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k} = \frac{e^{-\lambda}\lambda^k p^k}{k!}\sum_{n=k}^{\infty}\frac{\lambda^{n-k}(1-p)^{n-k}}{(n-k)!} = \frac{e^{-\lambda}(\lambda p)^k}{k!}\sum_{m=0}^{\infty}\frac{(\lambda(1-p))^m}{m!} = \frac{e^{-\lambda p}(\lambda p)^k}{k!},$$

т.е. $X \sim \mathrm{Po}(\lambda p)$. □

*Задача** 3.32. Вероятността за улучване на цел при един изстрел е 0.001. За поразяване на целта са необходими поне две попадения. Каква е вероятността за поразяване на целта, ако са направени 5000 изстрела?

Отговор. $X \sim \mathrm{Bin}(5000, 0.001)$, $\mathbb{P}(X \geq 2) = 1-(1-\frac{1}{1000})^{5000} - 5000(1-\frac{1}{1000})^{4999}\frac{1}{1000} \approx 1-e^{-5}-5e^{-5} \approx 0.959$. □

. . . .

Полиномно разпределение *с параметри $n, r \in \mathbb{N}$ и $p_1, \ldots, p_r \in [0,1]$, такива че $p_1 + \cdots + p_r = 1$, е дискретно разпределение на многомерна сл. в. $X = (X_1, \ldots, X_r)$, приемаща стойности $(k_1, \ldots, k_r) \in \mathbb{N}_0^r$: $k_1 + \cdots + k_r = n$ с вероятности*

$$\mathbb{P}(X = (k_1, \ldots, k_r)) = \frac{n!}{k_1!k_2!\ldots k_r!}p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r}$$

В този случай ще обозначаваме $X \sim \mathrm{Multi}(n; p_1, \ldots, p_r)$.

Забележка (Интерпретация). Тази сл. в. моделира резултата от n независими опита, при всеки от които настъпва точно едно от r несъвместими събития. Нека $X^{(i)} \sim \text{Multi}(1; p_1, \dots, p_r)$, $i = 1, \dots, n$, като $X^{(1)}, \dots, X^{(n)}$ са независими в съвкупност. Тогава

$$\mathbb{P}\big(X^{(1)} + \dots + X^{(n)} = (k_1, \dots, k_r)\big) = \sum_{e_1,\dots,e_n : e_1+\dots+e_n=(k_1,\dots,k_r)} \mathbb{P}(X^{(1)} = e_1, \dots, X^{(n)} = e_n) = \frac{n!}{k_1!k_2!\dots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r},$$

където $e_1, \dots, e_n$ са вектори в $\mathbb{R}^r$, чиито компоненти са 0, с изключение на един, който е равен на 1; т.е. $(X_1, \dots, X_r) \equiv X \overset{d}{=} X^{(1)} + \dots + X^{(n)} \equiv (\sum_{i=1}^n X_1^{(i)}, \dots, \sum_{i=1}^n X_r^{(i)})$.

Забележка (Пример). Урна съдържа l бели, m зелени и n червени топки. Изваждат се една по една с връщане $l_1 + m_1 + n_1$ топки. Каква е вероятността да бъдат извадени l_1 бели, m_1 зелени и n_1 червени топки в произволен ред?

Твърдение 3.18. *Ако* $X \sim \text{Multi}(n; p_1, \dots, p_r)$*, тогава* $G_X(s_1, \dots, s_r) = (p_1 s_1 + \dots + p_r s_r)^n$.

Д-во. Резултатът следва от мултиномната теорема, $(x_1 + \dots + x_r)^n = \sum_{k_1+\dots+k_r=n} \frac{n!}{k_1!\cdots k_r!} x_1^{k_1} \cdots x_r^{k_r}$. □

*Задача** 3.33. Нека $(X_1, \dots, X_r) \sim \text{Multi}(n; p_1, \dots, p_r)$. Да се покаже, че $\rho_{X_i,X_j} = -\sqrt{\frac{p_i p_j}{(1-p_i)(1-p_j)}}$, $i \neq j$.

Решение. Нека разгледаме съвместното разпределение на X_1 и X_2. За $k_1, k_2 \in \{0, 1, \dots, n\} : k_1 + k_2 \leq n$,

$$\begin{aligned}
\mathbb{P}(X_1 = k_1, X_2 = k_2) &= \sum_{k_3+\dots+k_r=n-k_1-k_2} \mathbb{P}(X_1 = k_1, X_2 = k_2, \dots, X_r = k_r) = \sum_{k_3+\dots+k_r=n-k_1-k_2} \frac{n!}{k_1!k_2!k_3!\cdots k_r!} p_1^{k_1} p_2^{k_2} p_3^{k_3} \cdots p_r^{k_r} \\
&= \frac{n! p_1^{k_1} p_2^{k_2} (1-p_1-p_2)^{n-k_1-k_2}}{k_1!k_2!(n-k_1-k_2)!} \sum_{k_3+\dots+k_r=n-k_1-k_2} \frac{(n-k_1-k_2)!}{k_3!\cdots k_r!} \Big(\frac{p_3}{1-p_1-p_2}\Big)^{k_3} \cdots \Big(\frac{p_r}{1-p_1-p_2}\Big)^{k_r} \\
&= \binom{n}{k_1}\binom{n-k_1}{k_2} p_1^{k_1} p_2^{k_2} (1-p_1-p_2)^{n-k_1-k_2},
\end{aligned}$$

където $\sum_{k_3+\dots+k_r=n-k_1-k_2} \frac{(n-k_1-k_2)!}{k_3!\cdots k_r!} \big(\frac{p_3}{1-p_1-p_2}\big)^{k_3} \cdots \big(\frac{p_r}{1-p_1-p_2}\big)^{k_r} = \big(\frac{p_3+\dots+p_r}{1-p_1-p_2}\big)^{n-k_1-k_2} = 1$. Тогава

$$\begin{aligned}
\mathbb{E}[X_1 X_2] &= \sum_{i=0}^{n} \sum_{j=0}^{n-i} ij \binom{n}{i}\binom{n-i}{j} p_1^i p_2^j (1-p_1-p_2)^{n-i-j} = np_1p_2 \sum_{i=1}^{n} (n-i) \sum_{j=1}^{n-i} \binom{n-1}{i-1}\binom{n-i-1}{j-1} p_1^{i-1} p_2^{j-1} (1-p_1-p_2)^{n-i-j} \\
&= np_1p_2 \sum_{i=0}^{n-1} (n-i-1) \binom{n-1}{i} p_1^i \Big\{ \sum_{j=0}^{n-i-2} \binom{n-i-2}{j} p_2^j (1-p_1-p_2)^{n-i-j-2} \Big\} \\
&= n(n-1)p_1p_2 \sum_{i=0}^{n-1} \binom{n-1}{i} p_1^i (1-p_1)^{n-i-2} - np_1p_2 \sum_{i=0}^{n-1} i \binom{n-1}{i} p_1^i (1-p_1)^{n-i-2} \\
&= \frac{n(n-1)p_1p_2}{1-p_1} \sum_{i=0}^{n-1} \binom{n-1}{i} p_1^i (1-p_1)^{n-i-1} - \frac{n(n-1)p_1^2 p_2}{1-p_1} \sum_{i=0}^{n-2} \binom{n-2}{i} p_1^i (1-p_1)^{n-i-2} \\
&= \frac{n(n-1)p_1p_2}{1-p_1} - \frac{n(n-1)p_1^2p_2}{1-p_1} = n(n-1)p_1p_2.
\end{aligned}$$

По същия начин $X_1 \sim \text{Bin}(n, p_1)$ и $X_2 \sim \text{Bin}(n, p_2)$, откъдето

$$\rho_{X_1,X_2} = \frac{\mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]}{\sqrt{\text{Var}(X_1)}\sqrt{\text{Var}(X_2)}} = \frac{n(n-1)p_1p_2 - n^2p_1p_2}{\sqrt{n^2p_1(1-p_1)p_2(1-p_2)}} = -\sqrt{\frac{p_1p_2}{(1-p_1)(1-p_2)}}.$$

Резултатът е аналогичен за всяка двойка (X_i, X_j), $i \neq j$. □

Задача 3.34. Всяка от девет топки може с една и съща вероятност да попадне в едната от три празни клетки. Да се пресметне вероятността: (а) във всяка клетка да попаднат по три топки; (б) в една клетка да паднат четири топки, в друга – три и в останалата – две топки.

Отговори. (а) $\frac{9!}{(3!)^3}\frac{1}{3^9}$; (б) $3!\frac{9!}{4!3!2!}\frac{1}{3^9}$. □

Задача 3.35. Урна съдържа три еднакви топки: бяла, зелена и черна. Изваждаме с връщане пет пъти по една топка. Каква е вероятността всяка от бялата и червената топка да бъде извадена поне два пъти?

Отговор. $\frac{5!}{2!2!}\frac{1}{3^5}+2\frac{5!}{3!2!}\frac{1}{3^5}$. □

Задача 3.36. Урна съдържа 2 бели, 2 черни и 6 зелени топки. Изваждат се една по една с връщане 20 топки. Каква е вероятността да изтеглим с 4 повече бели топки отколкото черни?

Решение. Нека $X=(X_1,X_2,X_3)\sim \mathrm{Multi}(20;\frac{1}{5},\frac{1}{5},\frac{3}{5})$, където сл. в. X_1, X_2 и X_3 са съответно броят на изтеглените бели, черни и зелени топки. От преди,

$$\mathbb{P}(X_1-X_2=4)=\sum_{i=4}^{20}\mathbb{P}(X_1=i,X_2=i-4)=\sum_{i=4}^{20}\binom{n}{i}\binom{20-i}{i-4}\Big(\frac{1}{5}\Big)^{2i-4}\Big(\frac{3}{5}\Big)^{n-2i+4}.$$

Алт. решение (Техника).* Търсената вероятност е равна на коефициента пред s^4 в развитието на

$$\begin{aligned}\mathbb{E}[s^{X_1-X_2}]&=\sum_{k_1,k_2,k_3\in\mathbb{N}_0:k_1+k_2+k_3=20}s^{k_1}s^{-k_2}\mathbb{P}(X_1=k_1,X_2=k_2,X_3=k_3)\\&=G_X\left(s,s^{-1},1\right)=\Big(\frac{1}{5}s+\frac{1}{5s}+\frac{3}{5}\Big)^{20}=\frac{1}{5^{20}}\sum_{n=0}^{20}\sum_{m=0}^{20-n}\frac{20!}{m!n!(20-m-n)!}s^{n-m}3^{20-n-m},\end{aligned}$$

откъдето $\mathbb{P}(X_1-X_2=4)=\frac{1}{5^{20}}\sum_{k=0}^{8}\frac{20!}{k!(k+4)!(20-2k-4)!}3^{20-2k-4}\approx 0.052$. □

Задача 3.37. На колко е равна вероятността при хвърлянето на n зара сумата от падналите точки да бъде равна на дадено число k $(n\le k\le 6n)$?

Решение. Нека $X=(X_1,\dots,X_6)\sim\mathrm{Multi}(n;\frac{1}{6},\dots,\frac{1}{6})$, където сл. в. X_i отчита броя на хвърлянията, в които на зара са се паднали точно i точки, $i=1,\dots,6$. Тогава

$$\mathbb{P}(X_1+2X_2+\cdots+6X_6=k)=\sum_{k_1+2k_2+\cdots+6k_6=k}\frac{n!}{k_1!k_2!\cdots k_6!}\frac{1}{6^n}.$$

Алт. решение (Техника).* Търсената вероятност е равна на коефициента пред s^k в развитието на

$$\begin{aligned}\mathbb{E}[s^{X_1+2X_2\cdots+6X_6}]&=\sum_{k_1+k_2+\cdots+k_6=n}s^{k_1}s^{2k_2}\cdots s^{6k_6}\mathbb{P}(X_1=k_1,X_2=k_2,\dots,X_6=k_6)=G_X\left(s,\dots,s^6\right)=\Big(\frac{s+\cdots+s^6}{6}\Big)^n\\&=\frac{s^n}{6^n}\frac{(1-s^6)^n}{(1-s)^n}=\frac{s^n}{6^n}\sum_{i=0}^{n}\binom{n}{i}(-s)^{6i}\sum_{j=0}^{\infty}\binom{n+j-1}{j}s^j=\frac{1}{6^n}\sum_{i=0}^{n}(-1)^i\binom{n}{i}\sum_{j=0}^{\infty}\binom{n+j-1}{j}s^{n+6i+j},\end{aligned}$$

откъдето $\mathbb{P}(X_1+2X_2+\cdots+6X_6=k)=\frac{1}{6^n}\sum_{i=0}^{\lfloor k-n\rfloor/6}(-1)^i\binom{n}{i}\binom{k-6i}{k-n-6i}$ (сравнете със Задача 2.16). □

*Задача*** 3.38. Каква е вероятността в случайно избрано число от съвкупността на числата от 000000 до 999999 сумата от първите три цифри да е равна на сумата от последните три цифри?

Решение. Нека $X=(X_0,X_1,\dots,X_9)\sim\mathrm{Multi}(3;\frac{1}{10},\dots,\frac{1}{10})$ и $Y=(Y_0,Y_1,\dots,Y_9)\sim\mathrm{Multi}(3;\frac{1}{10},\dots,\frac{1}{10})$, такива че $X\perp\!\!\!\perp Y$, където сл. в. X и Y отчитат съответно цифрите в първите три и последните три цифри на числото. Търсената вероятност е равна на коефициента пред свободния член в развитието на

$$\begin{aligned}\mathbb{E}[s^{(X_1+\cdots+9X_9)-(Y_1+\cdots+9Y_9)}]&=G_X(1,s,s^2,\dots,s^9)\cdot G_Y(1,s^{-1},s^{-2},\dots,s^{-9})=\frac{(1+s+s^2\cdots+s^9)^3}{10^3}\frac{(1+\frac{1}{s}+\frac{1}{s^2}+\cdots+\frac{1}{s^9})^3}{10^3}\\&=\frac{1}{10^6}\Big(\frac{1-s^{10}}{1-s}\Big)^3\Big(\frac{1-s^{-10}}{1-s^{-1}}\Big)^3=\frac{s^{-27}}{10^6}\frac{(1-s^{10})^6}{(1-s)^6}=\frac{1}{10^6}\sum_{i=0}^{6}(-1)^i\binom{6}{i}\sum_{j=0}^{\infty}\binom{6+j-1}{j}s^{10i+j-27},\end{aligned}$$

откъдето $\mathbb{P}(X_1+\cdots+9X_9=Y_1+\cdots+9Y_9)=\frac{1}{10^6}\sum_{i=0}^{3}(-1)^i\binom{6}{i}\binom{32-10i}{27-10i}$ (сравнете със Задача 2.5). □

4 Непрекъснати случайни величини

4.1 Основни понятия

Непрекъсната случайна величина. *Нека $(\Omega, \mathcal{A}, \mathbb{P})$ е вероятностно пространство. Наричаме една реална функция $X : \Omega \to \mathbb{R}$ непрекъсната сл. в., ако*

(i) *за всички $a, b \in \mathbb{R} : a < b$ е в сила*

$$\{a < X < b\} \equiv X^{-1}\big((a,b)\big) = \{\omega \in \Omega : a < X(\omega) < b\} \in \mathcal{A};$$

(ii) *съществува неотрицателна функция $f_X : \mathbb{R} \to \mathbb{R}_+$, наречена функция на плътност на X, такава че, за всички $a, b \in \mathbb{R} : a < b$,*

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x)dx.$$

Тогава функцията на разпределение на X има следния вид

$$F_X(x) = \int_{-\infty}^x f_X(t)dt, \qquad x \in \mathbb{R}.$$

Забележка (Следствия). От условие (i) следва, че $\{X \in A\} \in \mathcal{A}$ за всяко Борелово множество $A \in \mathcal{B}(\mathbb{R})$ и в частност за $A = (a,b], [a,b), [a,b], (-\infty, b], (a, \infty), \{a\}$. Тогава

$$\mathbb{P}(X = a) \leq \mathbb{P}(a - n^{-1} < X < a + n^{-1}) = \int_{a-\frac{1}{n}}^{a+\frac{1}{n}} f_X(x)dx \longrightarrow 0,$$

откъдето $\mathbb{P}(X = a) = 0$ и

$$\mathbb{P}(a < X < b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a \leq X \leq b).$$

Освен това имаме, че $\mathbb{P}(X \in A) = \int_A f_X(x)dx$, където интегрирането е в смисъла на Лебег.

Условие (ii) пък предполага, че (а) $f_X \geq 0$ и (б) $\int_{-\infty}^{\infty} f_X(x)dx = 1$, но не и $f_X \leq 1$. Всъщност (а) и (б) са достатъчни условия за задаването на непрекъсната сл. в. Действително, ако f_X е функция, удовлетворяваща (а) и (б), то функцията $F_X(x) = \int_{-\infty}^x f_X(t)dt$, $tx \in \mathbb{R}$ е (i) ненамаляваща; (ii) ляво непрекъсната; (iii) $\lim_{x\to-\infty} F_X(x) = 0$ и $\lim_{x\to\infty} F_X(x) = 1$; т.е. F_X е функция на разпределение и определя вероятностното разпределение на някоя сл. в. X. Следователно задаването на непрекъсната сл. в. е еквивалентно на задаването на нейната плътност.

Забележка (Анализ). Ако f_X е непрекъсната, то $f_X(x) = \frac{d}{dx}F_X(x)$ по Нютон-Лайбниц. В общия случай производната $\frac{d}{dx}\int_{-\infty}^x f_X(t)dt$ съществува за почти всяко $x \in \mathbb{R}$, като тогава можем да работим реално с

$$f_X(x) = \begin{cases} \frac{d}{dx}F_X(x), & \text{ако съществува} \\ 0, & \text{иначе} \end{cases}.$$

Обратното, ако ни е дадена функция на разпределение F_X, която има непрекъсната производна, то X е непрекъсната сл. в. с плътност $f_X(x) = \frac{d}{dx}F_X(x)$, $x \in \mathbb{R}$. Последното може да се обобщи за *абсолютно непрекъснати* F_X (виж Фундаментална теорема за Лебеговия интеграл).

*Забележка** (Функция на сл. в.). Нека X е сл. в. и $g : \mathbb{R} \to \mathbb{R}$ е произволна функция. За да бъде $g(X)$ сл. в. е достатъчно, за всички $a, b \in \mathbb{R} : a < b$,

$$\{a < g(X) < b\} = \big\{X \in g^{-1}(a,b)\big\} \in \mathcal{A},$$

което е изпълнено, ако $g^{-1}(a,b) \equiv \{x \in \mathbb{R} : a < g(x) < b\} \in \mathcal{B}(\mathbb{R})$. Такива функции g се наричат Борелови или най-общо измерими. В частност, всички непрекъснати функции са измерими.

Задача 4.1. Във вътрешността на кръг с център O и радиус R случайно се избира точка A. Нека $X = |OA|$. Да се намери разпределението на X.

Решение. Нека $x \in (0,R)$. Тогава $F_X(x) = \mathbb{P}(X < x) = \frac{x^2}{R^2}$ е вероятността A да попадне в кръга с център O и радиус x. В допълнение $F_X(x) = 0$, за $x \leq 0$, и $F_X(x) = 1$ за $x \geq R$; т.е. F_X е непрекъснато диференцируема функция, откъдето следва, че X е непрекъсната сл. в. с плътност $f_X(x) = \frac{2x}{R^2} \cdot \mathbb{1}_{(0,R)}(x)$. $\square$

Задача 4.2. Нека е X непрекъсната сл. в. със симетрична плътност, т.е. $f_X(-x) = f_X(x)$ за всяко $x \in \mathbb{R}$. Да се докаже, че за произволно $x \in \mathbb{R}$ е вярно: (а) $F_X(0) = 0.5$; (б) $F_X(-x) = 1 - F_X(x)$; (в) $\mathbb{P}(|X| < x) = 2F_X(x) - 1$.

Решение. (а) $F_X(0) = 1 - \int_0^\infty f_X(x)dx = 1 + \int_0^{-\infty} f_X(-y)dy = 1 - F_X(0)$, откъдето $F_X(0) = 0.5$; (б) $F_X(-x) = \int_{-\infty}^{-x} f_X(t)dt = \int_x^\infty f_X(u)du = 1 - F_X(x)$; (в) $\mathbb{P}(|X| < x) = \mathbb{P}(-x < X < x) = F_X(x) - F_X(-x) = 2F_X(x) - 1$. ☐

. . . .

Твърдение 4.1 (Смяна на променливи). *Нека X е непрекъсната сл. в. с плътност f_X и $g : \mathbb{R} \to \mathbb{R}$ е диференцируема строго монотонна функция. Тогава $g(X)$ е непрекъсната сл. в. с плътност*

$$f_{g(X)}(y) = f_X(g^{-1}(y))\big|(g^{-1}(y))'\big| \cdot \mathbb{1}_{(\inf_x g(x), \sup_x g(x))}(y), \qquad \text{за } y \in \mathbb{R}.$$

Д-во. Нека g е строго растяща. Полагаме $h(y) = g^{-1}(y)$, $y \in \mathbb{R}$. Тогава за $a, b \in \mathbb{R} : a < b$, използвайки смяната $y = g(x)$,

$$\mathbb{P}(a < g(X) < b) = \mathbb{P}\big(h(a) < X < h(b)\big) = \int_{h(a)}^{h(b)} f_X(x)dx = \int_a^b f_X(h(y))dh(y) = \int_a^b f_X(h(y))h'(y)dy.$$

Ако g е строго намаляваща, $\mathbb{P}(a < g(X) < b) = \mathbb{P}(h(b) < X < h(a)) = \cdots = -\int_a^b f_X(h(y))h'(y)dy$. Понеже h е също строго намаляваща, то $h'(y) < 0$, откъдето следва общото твърдение. ☐

Забележка (Контрапример). Не за всички измерими функции g трансформацията $g(X)$ е непрекъсната сл. в. Например, за $g(x) = 1$, при $x \geq 0$, и $g(x) = 0$, при $x < 0$, получаваме $g(X) \sim \mathrm{Ber}(\mathbb{P}(X \geq 0))$.

Задача 4.3. Нека X е непрекъсната сл. в. с плътност $f_X(x) = \lambda e^{-\lambda x} \cdot \mathbb{1}_{(0,\infty)}(x)$. Да се намерят разпределенията на случайните величини (а) $Y = -X$; (б) $Y = 2X - 1$; (в) $Y = \sqrt{X}$; (г) $Y = X^\alpha$ за $\alpha > 0$.

Решение. От Твърдение 4.1 сл. в. са непрекъснати с плътности (а) $f_{-X}(y) = f_X(-y) \cdot \mathbb{1}_{(-\infty,0)}(y) = \lambda e^{\lambda y} \cdot \mathbb{1}_{(-\infty,0)}(y)$; (б) $f_{2X-1}(y) = f_X(\frac{y+1}{2})\frac{1}{2} \cdot \mathbb{1}_{(-1,\infty)}(y) = \frac{1}{2}\lambda e^{-\lambda\frac{y+1}{2}} \cdot \mathbb{1}_{(-1,\infty)}(y)$; (в) $f_{\sqrt{X}}(x) = f_X(y^2)|2y| \cdot \mathbb{1}_{(0,\infty)}(y) = 2\lambda y e^{-\lambda y^2} \cdot \mathbb{1}_{(0,\infty)}(y)$; (г) $f_{X^\alpha}(y) = f_X(y^{\frac{1}{\alpha}})|\frac{1}{\alpha}y^{\frac{1-\alpha}{\alpha}}| \cdot \mathbb{1}_{(0,\infty)}(y) = \frac{\lambda}{\alpha}y^{\frac{1-\alpha}{\alpha}}e^{-\lambda y^{\frac{1}{\alpha}}} \cdot \mathbb{1}_{(0,\infty)}(y)$. ☐

. . . .

Моменти. *Нека X е непрекъсната сл. в. с плътност f_X. Ако $\int_{-\infty}^\infty |x|f_X(x)dx < \infty$, числото*

$$\mathbb{E}[X] = \int_{-\infty}^\infty x f_X(x)dx$$

се нарича математическо очакване на X. Ако $\int_{-\infty}^\infty x^2 f_X(x)dx < \infty$, числото

$$\mathrm{Var}(X) \equiv \mathbb{E}\big[(X - \mathbb{E}[X])^2\big] = \int_{-\infty}^\infty (x - \mathbb{E}[X])^2 f(x)dx$$

се нарича дисперсия на X.

Забележка (Смяна на променливите). Нека $Y = g(X)$ за *измерима* функция $g : \mathbb{R} \to \mathbb{R}$. Ако интегралът е абсолютно сходящ, то е изпълнено

$$\mathbb{E}[Y] = \int_{-\infty}^\infty g(x)f_X(x)dx.$$

В случай, че g е диференцируема строго растяща функция, директно,

$$\mathbb{E}[Y] = \int_{-\infty}^\infty y f_Y(y)dy = \int_{-\infty}^\infty y f_X(g^{-1}(y))(g^{-1}(y))'dy = \int_{-\infty}^\infty g(x)f_X(x)dx.$$

Твърдение 4.2.

позитивност: *ако $X \geq 0$, то $\mathbb{E}[X] \geq 0$.*

линейност: $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$, *за $a, b \in \mathbb{R}$.*

Д-во. От свойствата на интегралите. $\square$

Забележка (Свойства на дисперсията)*.* От горното следва, че $\mathrm{Var}(X) \geq 0$ и $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$.

Задача 4.4. Дадена е сл. в. X с плътност $f_X(x) = c(x^2+2x)\cdot \mathbb{1}_{[0,1]}(x)$. Намерете (а) константата c; (б) $\mathbb{E}[X]$ и $\mathrm{Var}(X)$; (в) $\mathbb{P}\big(X \leq \mathbb{E}[X]\big)$; (г) $\mathbb{E}[X^2+3X]$.

Отговори. (а) $\int_0^1 c(x^2+2x)dx = 1 \iff c = \frac{3}{4}$; (б) $\mathbb{E}[X] = \frac{11}{16}$, $\mathbb{E}[X^2] = \frac{21}{40}$, $\mathrm{Var}(X) \approx 0.052$; (в) $\mathbb{P}\big(X \leq \mathbb{E}[X]\big) \approx 0.436$; (г) $\mathbb{E}[X^2+3X] \approx 2.59$. $\square$

4.2 Основни непрекъснати разпределения

Равномерно разпределение *върху интервала* (a,b)*, за* $a,b \in \mathbb{R} : a < b$*, е непрекъснато разпределение на сл. в.* X *с плътност*

$$f_X(x) = \frac{1}{b-a}\cdot \mathbb{1}_{(a,b)}(x).$$

В този случай ще обозначаваме $X \sim \mathrm{Unif}(a,b)$*.*

Забележка (Стандартно равномерно разпределение)*.* Нека $Y = \frac{X-a}{b-a}$. От Твърдение 4.1 следва, че Y е непрекъсната сл. в. с плътност $f_Y(y) = f_X\big(a+(b-a)y\big)|b-a|\cdot \mathbb{1}_{(0,1)}(y) = \mathbb{1}_{(0,1)}(y)$; т.е. $Y \sim \mathrm{Unif}(0,1)$.

Твърдение 4.3. *Ако* $X \sim \mathrm{Unif}(a,b)$*, тогава* $\mathbb{E}[X] = \frac{a+b}{2}$, $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$, $F_X(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & x \in (a,b) \\ 1, & x \geq b \end{cases}$ *.*

Д-во. Първо $\int_a^b f_X(x)dx = \int_a^b \frac{1}{b-a}dx = \frac{x}{b-a}\Big|_a^b = 1$; т.е. равномерното разпределение е добре дефинирано. По същия начин $F_X(x) = \int_a^x f_X(y)dy = \frac{y}{b-a}\Big|_a^x = 1 = \frac{x-a}{b-a}$ за $x \in (a,b)$,

$$\mathbb{E}[X] = \int_a^b \frac{x}{b-a}dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{a+b}{2}, \qquad \mathbb{E}[X^2] = \int_a^b \frac{x^2}{b-a}dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^2+ab+a^2}{3}, \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{2}.$$

$\square$

Задача 4.5. Върху окръжност с център O и радиус r е фиксирана точка A, а точка B попада по случаен начин върху окръжността. Да се намери математическото очакване на лицето на $\triangle AOB$.

Решение. Нека т. A има полярни координати $(r,0)$, а т. B – (r,Φ), където $\Phi \sim \mathrm{Unif}(0,2\pi)$. Да дефинираме $g(x) = \frac{r^2|\sin x|}{2}$, $x \in [0,2\pi]$. Тогава очакваното лице е

$$\mathbb{E}[g(\Phi)] = \int_0^{2\pi} g(x)\frac{1}{2\pi}dx = \frac{r^2}{4\pi}\int_0^{\pi}\sin x dx - \frac{r^2}{4\pi}\int_{\pi}^{2\pi}\sin x dx = -\frac{r^2}{4\pi}\cos x\Big|_0^{\pi} + \frac{r^2}{4\pi}\cos x\Big|_{\pi}^{2\pi} = \frac{r^2}{\pi}.$$

$\square$

Задача 4.6. Нека $X \sim \mathrm{Unif}(0,7)$ е времето на безотказна работа в години на даден апарат. Съгласно гаранцията на апарата, той ще бъде заменен с нов на петата годна или преди това в случай на дефект. Нека Y е времето до смяната на апарата. Да се пресметнат $\mathbb{P}(Y < 4)$, $\mathbb{E}[Y]$ и $\mathrm{Var}(Y)$. Ако са продадени 1000 апарата, колко средно ще трябва да се подменят преди петата година?

Отговори. $Y = X\cdot \mathbb{1}_{\{X\leq 5\}} + 5\cdot \mathbb{1}_{\{X>5\}}$, откъдето $\mathbb{P}(Y<4) = \mathbb{P}(X<4) = \frac{4}{7}$, $\mathbb{E}[Y] = \frac{45}{14}$, $\mathbb{E}[Y^2] = \frac{275}{21}$ и $\mathrm{Var}(Y) \approx 2.764$. Относно втория въпрос, нека $Z \sim \mathrm{Bin}(1000, \mathbb{P}(X<5))$. Тогава $\mathbb{E}[Z] \approx 714.29$. $\square$

Задача 4.7. Лъч светлина минава от т. $(0,2)$ към т. $(0,1)$ и се пречупва случайно, сключвайки ъгъл $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ с Oy. Нека X е точката, в която пречупеният лъч пресича Ox. Да се намери плътността на X.

Решение. $\Theta \sim \mathrm{Unif}(-\frac{\pi}{2}, \frac{\pi}{2})$ и $X = \tan\Theta$, откъдето $f_X(x) = f_\Theta(\arctan x)(\arctan x)' = \frac{1}{\pi(1+x^2)}\cdot \mathbb{1}_{(-\infty,\infty)}(x)$; т.е. X има **разпределение на Коши** с параметри $(0,1)$. $\square$

*Забележка** (Тежки опашки). Отгоре $\mathbb{P}(X > t) = \int_t^\infty \frac{1}{\pi(1+x^2)}dx \approx \int_t^\infty \frac{1}{\pi x^2}dx = \frac{1}{\pi t}$ за големи t, т.е. "опашката" на X намалява полиномиално. Нещо повече

$$\int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)}dx \int_0^\infty \frac{x}{\pi(1+x^2)}dx = \frac{1}{2\pi}\int_0^\infty \frac{1}{u}du = \frac{1}{2\pi}\ln(1+x^2)\big|_0^\infty = \infty,$$

откъдето следва, че $\mathbb{E}[X]$ (и следователно всеки друг момент на X) не съществува. Това е характерно свойство за случайни величини, чиито вероятностните разпределения имат тежки опашки, т.е. X, такова че

$$\mathbb{P}(X > t) \sim t^{-\alpha}, \qquad \text{при } t \to \infty, \text{ за } \alpha > 0.$$

Може да се покаже, че $\mathbb{E}[X^k]$ не съществува за $k > \alpha - 1$.

*Задача** 4.8. Нека X е сл. в. с непрекъсната функция на разпределение F_X. Намерете разпределението на $F_X(X)$.

Д-во. Нека положим $G(u) := \sup\{x \in \mathbb{R} : F_X(x) < u\}$ за $u \in [0,1]$. Фиксираме $u \in [0,1]$. Ще докажем, че $F_X(G(u)) = u$. От горната дефиниция имаме $F_X(G(u)) \leq u$. Да допуснем, че $F_X(G(u)) < u$. Понеже F_X е непрекъсната в $G(u)$, то $\exists y > G(u)$, за което $F_X(G(u)) < F_X(y) < F_X(G(u)) + \frac{u - F_X(G(u))}{2} < u$. Тогава $G(u) \geq y$, абсурдно; т.е. $F_X(G(u)) = u$.

Нека $x \in \mathbb{R} : G(u) > x$. От свойствата на супремума съществува $y \in \mathbb{R}$ т. ч. $G(u) - \frac{G(u)-x}{2} < y < G(u)$, за което $F_X(y) < u$; т.е. имаме $F_X(x) \leq F_X(y) < u$.

Нека $x \in \mathbb{R} : F_X(x) < u$. Тогава $F_X(x) < F_X(G(u)$ и $G(u) > x$, понеже F_X е ненамаляваща функция. От горните резултати следва, че $\{x \in \mathbb{R} : F_X(x) < u\} = (-\infty, G(u))$, откъдето

$$\mathbb{P}\big(F_X(X) < u\big) = \mathbb{P}\big(X < G(u)\big) = F_X(G(u)) = u;$$

т.е. $F_X(X) \sim \mathrm{Unif}(0,1)$. □

Забележка (Обобщение). За произволна функция на разпределение F_X, ако $F_X(G(u)) = u$, то $\{x \in \mathbb{R} : F_X(x) < u\} = (-\infty, G(u))$ отпреди, където F_X не е непременно непрекъсната в $G(u)$. Ако $F_X(G(u)) < u$, тогава $\{x \in \mathbb{R} : F_X(x) < u\} = (-\infty, G(u)]$, като F_X има скок в $G(u)$ и граница отдясно $F_X(G(u)^+) := \lim_{x \downarrow G(u)} F_X(x)$. Обединявайки двата резултата,

$$\mathbb{P}\big(F_X(X) < u\big) = \begin{cases} \mathbb{P}\big(X < G(u)\big) = u, & \text{за } u \in [0,1] : F_X(G(u)) = u; \\ \mathbb{P}\big(X \leq G(u)\big) = F_X(G(u)^+), & \text{за } u \in [0,1] : F_X(G(u)) < u. \end{cases}$$

В случай, че F_X е стриктно нарастваща, то тя е обратима, при което $G(u) = F_X^{-1}(u)$.

Забележка (Квантилна функция). Функцията G се нарича **квантилната функция**, асоциирана с разпределението на X. От свойствата на супремума (i) G е ненамаляваща; (ii) G е дясно непрекъсната.

. . . .

Експоненциално разпределение *с параметър $\lambda > 0$ е непрекъснато разпределение на неотрицателна сл. в. X с плътност*

$$f_X(x) = \lambda e^{-\lambda x} \cdot \mathbb{1}_{(0,\infty)}(x).$$

В този случай ще обозначаваме $X \sim \mathrm{Exp}(\lambda)$.

Забележка (Интерпретация/Безпаметност). Това разпределение може да се интерпретира като непрекъснатия аналог на геометричното. Всъщност

$$\mathbb{P}(X > x + y | X > x) = \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} = \mathbb{P}(X > y), \qquad \text{за } x, y \geq 0.$$

Твърдение 4.4. *Ако $X \sim \mathrm{Exp}(\lambda)$, тогава $\mathbb{E}[X] = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$, $F_X(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\lambda x}, & x > 0 \end{cases}$.*

*Д-во**. Първо $\int_0^\infty f_X(x)dx = \int_0^\infty \lambda e^{-\lambda x}dx = -\int_0^{-\infty} e^y dy = e^y|_{-\infty}^0 = 1$, използвайки смяната $y = -\lambda x$; т.е. експоненциалното разпределение е добре дефинирано. По същия начин $F_X(x) = 1 - e^{-\lambda x}$, $x \in (0, \infty)$. Използвайки съответно интегриране по части и метода на Файнман,

$$\mathbb{E}[X] = \int_0^\infty x\lambda e^{-\lambda x}dx = -\int_0^\infty x d e^{-\lambda x} = -(xe^{-\lambda x})\big|_0^\infty + \int_0^\infty e^{-\lambda x}dx = -\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda},$$
$$\frac{d}{d\lambda}\frac{1}{\lambda} = \frac{d}{d\lambda}\mathbb{E}[X] \iff -\frac{1}{\lambda^2} = \int_0^\infty xe^{-\lambda x}dx - \int_0^\infty x^2\lambda e^{-\lambda x}dx = \frac{1}{\lambda^2} + \mathbb{E}[X^2] \implies \mathbb{E}[X^2] = \frac{2}{\lambda^2},$$
$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

□

Задача 4.9. В магазин работят две касиерки. Предполагаме, че времето необходимо за обслужване на клиент на всяка от двете опашки е експоненциално разпределена случайна величина с математическо очакване 8 (мин) за първата опашка и 5 (мин) за втората. Клиент, избрал по случаен начин опашка, е чакал по-малко от 4 минути. Каква е вероятността той да е бил на първата опашка?

Решение. Нека $A = \{\text{чака на първа опашка}\}$, $X_1 \sim \mathrm{Exp}(\frac{1}{8})$ и $X_2 \sim \mathrm{Exp}(\frac{1}{5})$. Да положим $X = X_1 \cdot \mathbb{1}_A + X_2 \cdot \mathbb{1}_{A^c}$. Тогава $\mathbb{P}(A|X<4) = \frac{\mathbb{P}(X<4|A)\mathbb{P}(A)}{\mathbb{P}(X<4|A)\mathbb{P}(A)+\mathbb{P}(X<4|A^c)\mathbb{P}(A^c)} = \frac{\mathbb{P}(X_1<4)}{\mathbb{P}(X_1<4)+\mathbb{P}(X_2<4)} \approx 0.42$. □

Задача 4.10. Времето за преглед на пациент е експоненциално разпределена случайна величина с очакване 30 (мин). За преглед има записани двама пациенти – първият за 11:00, а вторият за 11:30, като и двамата пристигат в точно определения час. Ако прегледът на първия не е завършил, вторият изчаква. Да се пресметне средно колко време ще прекара вторият пациент в поликлиниката.

Решение. Нека $X_1 \sim \mathrm{Exp}(\frac{1}{30})$ и $X_2 \sim \mathrm{Exp}(\frac{1}{30})$, като $X_1 \perp X_2$. Да положим $Y = X_2 + (X_1 - 30)\cdot \mathbb{1}_{\{X_1>30\}}$. Тогава $\mathbb{E}[Y] = 30 + \int_{30}^\infty \frac{1}{30}xe^{-\frac{1}{30}x}dx - 30\cdot\mathbb{P}(X_1>30) = 30\mathbb{P}(X_1 \le 30) - xe^{-\frac{1}{30}x}\big|_{30}^\infty + \int_{30}^\infty e^{-\frac{1}{30}x}dx = 30(1+e^{-1})$. □

. . . .

Нормално разпределение *с параметри $\mu \in \mathbb{R}$ и $\sigma > 0$ е непрекъснато разпределение на сл. в. X с плътност*
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \mathbb{1}_{(-\infty,\infty)}(x).$$
В този случай ще обозначаваме $X \sim \mathrm{N}(\mu, \sigma^2)$.

Забележка (Стандартно нормално разпределение). Нека $Z = \frac{X-\mu}{\sigma}$. От Твърдение 4.1 следва, че Z е непрекъсната сл. в. с плътност $f_Z(z) = \sigma f_X(\sigma z + \mu)\cdot\mathbb{1}_{(-\infty,\infty)}(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\cdot\mathbb{1}_{(-\infty,\infty)}(z)$; т.е. $Z \sim \mathrm{N}(0,1)$.

Забележка (Полза). Нека $Z \sim \mathrm{N}(0,1)$. Тогава функцията на разпределение $\Phi(z) := F_Z(z)$, $z \in \mathbb{R}$, няма явен вид, но съществуват таблици с приблизителни нейни стойности. Тяхната полза идва от приближението, което Централната гранична теорема (и в частност Теорема 4.6) дава за функцията на разпределение на средно аритметично от група независими еднакво-разпределени сл. в. (с краен втори момент).

Твърдение 4.5. *Ако $X \sim \mathrm{N}(\mu,\sigma^2)$, тогава $\mathbb{E}[X] = \mu$, $\mathrm{Var}(X) = \sigma^2$ и $\mathbb{E}[e^{tX}] = e^{t\mu + \frac{\sigma^2 t^2}{2}}$, $|t| < 1$.*

Д-во.* Нека $Z = \frac{X-\mu}{\sigma}$. Тогава $Z \sim \mathrm{N}(0,1)$ и

$$\mathbb{E}[Z] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty ze^{-\frac{z^2}{2}}dx = 0, \qquad \text{понеже } z \mapsto ze^{-\frac{z^2}{2}} \text{ е нечетна функция} \qquad \implies \qquad \mathbb{E}[X] = \mu,$$
$$\mathrm{Var}(Z) = \mathbb{E}[Z^2] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty z^2e^{-\frac{z^2}{2}}dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^0 z^2e^{-\frac{z^2}{2}}dz + \frac{1}{\sqrt{2\pi}}\int_0^\infty z^2e^{-\frac{z^2}{2}}dz$$
$$= -\frac{1}{\sqrt{2\pi}}ze^{-\frac{z^2}{2}}\big|_{-\infty}^0 - -\frac{1}{\sqrt{2\pi}}ze^{-\frac{z^2}{2}}\big|_0^\infty + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{z^2}{2}}dz = 1 \qquad \implies \qquad \mathrm{Var}(X) = \sigma^2,$$
$$\mathbb{E}[e^{tX}] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^\infty e^{tx}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{t(\mu+\sigma z)}e^{-\frac{z^2}{2}}dz = e^{t\mu+\frac{\sigma^2t^2}{2}}\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{(z-t\sigma)^2}{2}}dz = e^{t\mu+\frac{\sigma^2t^2}{2}}.$$

По същия начин се доказва, че $\mathbb{E}[Z^{2n+1}] = 0$ за всяко $n \in \mathbb{N}$. □

Забележка (Покритие). От неравенството на Чебишев видяхме, че за всяка сл. в. X с краен втори момент

$$\mathbb{P}(\mathbb{E}[X] - 2\sigma < X < \mathbb{E}[X] + 2\sigma) \geq 0.75.$$

Нека $X \sim \mathrm{N}(\mu, \sigma^2)$. Тогава

$$\mathbb{P}(\mathbb{E}[X] - 2\sigma < X < \mathbb{E}[X] + 2\sigma) = \mathbb{P}\Big(-2 < \frac{X-\mu}{\sigma} < 2\Big) = \mathbb{P}\Big(\Big|\frac{X-\mu}{\sigma}\Big| < 2\Big) = 2\Phi(2) - 1 \approx 95.45\%$$

Теорема 4.6 (Теорема на Моавър-Лаплас)**.** *Нека $Z \sim \mathrm{N}(0,1)$ и $X_n \sim \mathrm{Bin}(n,p)$, за $p \in (0,1)$. Тогава*

$$\mathbb{P}(X_n = k) \sim \frac{1}{\sqrt{np(1-p)}} f_Z\Big(\frac{k-np}{\sqrt{np(1-p)}}\Big), \qquad \text{за } k = 0,1,2,\dots,$$

и

$$\mathbb{P}\Big(\frac{X_n - np}{\sqrt{np(1-p)}} \leq x\Big) \longrightarrow \mathbb{P}(Z \leq x), \qquad \text{за } x \in \mathbb{R}.$$

Задача 4.11. Монета, за която вероятността да се падне ези е 3/4, се хвърля 2000 пъти. Каква е вероятността броят на падналите се езита да е между 1475 и 1535?

Отговор. Нека $X \sim \mathrm{Bin}(2000, \frac{3}{4})$. От Теорема 4.6 (или по-общо от Централната гранична теорема) следва, че $\mathbb{P}(1475 \leq X \leq 1535) = \mathbb{P}(\frac{-25}{\sqrt{375}} \leq \frac{X - \frac{6000}{4}}{\sqrt{\frac{6000}{16}}} \leq \frac{35}{\sqrt{375}}) \approx \int_{-1.29}^{1.81} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \Phi(1.81) - \Phi(1.29) = 0.867.$ □

*Задача** 4.12. Нека $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ и $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ са независими сл. в и $a, b \in \mathbb{R}$. Тогава $aX_1 + bX_2 \sim \mathrm{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$.

Решение. Използвайки свойствата на пораждащата функция на моментите,

$$\mathbb{E}[e^{t(aX_1+bX_2)}] = \mathbb{E}[e^{taX_1}] \cdot \mathbb{E}[e^{tbX_2}] = e^{ta\mu_1 + \frac{a^2\sigma_1^2 t^2}{2}} \cdot e^{tb\mu_2 + \frac{b^2\sigma_2^2 t^2}{2}} = e^{t(a\mu_1+b\mu_2) + \frac{(a^2\sigma_1^2 + b^2\sigma_2^2)t^2}{2}},$$

откъдето следва, че $aX_1 + bX_2 \sim \mathrm{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$. □

4.3 Колекция от случайни величини

Вектор от непрекъснати случайни величини. *Нека X и Y са две сл. в., дефинирани в едно и също вероятностно пространство. Тогава (X,Y) е двумерен вектор от непрекъснати сл. в., ако съществува неотрицателна функция $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$, такава че, за всички $x_1, x_2 \in \mathbb{R} : x_1 < x_2$ и $y_1, y_2 \in \mathbb{R} : y_1 < y_2$,*

$$\mathbb{P}(x_1 < X < x_2, y_1 < Y < y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{X,Y}(x,y) dy dx,$$

като множеството $D_{X,Y} = \{(x,y) \in \mathbb{R}^2 : f_{X,Y}(x,y) > 0\}$ се нарича носител на (X,Y). От тук следва, че

$$\mathbb{P}(x_1 < X < x_2) = \int_{x_1}^{x_2} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy dx = \int_{x_1}^{x_2} f_X(x) dx,$$

където $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$; т.е. X е непрекъсната сл. в. с (маргинална) плътност f_X. Функцията

$$\mathbb{P}(X < x, Y < y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y) dy dx, \qquad \text{за } x, y \in \mathbb{R}$$

се нарича функция на разпределение на (X,Y). По аналогия с дискретния случай, наричаме условна плътност на Y относно X функцията

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)},$$

която е дефинирана за почти всяко $x \in \mathbb{R}$. Горните дефиниции се обобщават за вектор от повече от две сл. в. по очевиден начин.

Забележка (Очакване). Нека $g : \mathbb{R}^2 \to \mathbb{R}$ е измерима функция. Ако интегралът е абсолютно сходящ,

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dydx.$$

В частност, ако съществува,

$$\mathrm{Cov}(X,Y) \equiv \mathbb{E}\big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\big] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mathbb{E}[X])(y-\mathbb{E}[Y]) f_{X,Y}(x,y)dydx.$$

Забележка (Условно очакване). Нека $g : \mathbb{R}^2 \to \mathbb{R}$ е измерима функция. Ако интегралът е абсолютно сходящ, за почти всяко $x \in \mathbb{R}$,

$$\mathbb{E}[g(X,Y)|X=x] = \int_{-\infty}^{\infty} g(x,y) f_{Y|X}(y|x)dy.$$

Твърдение 4.7.
(i) $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;
(ii) *ако* $X \geq Y$, *то* $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

Д-во. От свойствата на интегралите. □

Задача 4.13. Нека съвместната плътност на сл. в. X и Y е $f_{X,Y} = cxy$, $0 < x < y < 1$. Да се намерят (а) константата c; (б) f_X, f_Y, $\mathbb{E}[X]$, $\mathbb{E}[Y]$; (в) $\mathbb{P}(X < \frac{3}{4}, Y - X < \frac{1}{6})$; (г) $\mathbb{P}(Y - X < \frac{1}{3})$; $\mathbb{E}[Y - X | X = \frac{1}{4}]$.

Решение. По условие $f_{X,Y}(x,y) = cxy \cdot \mathbb{1}_{D_{X,Y}}(x,y)$, където $D_{X,Y} = \{(x,y) \in \mathbb{R}^2 : 0 < x < y < 1\}$. Тогава

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = c\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot \mathbb{1}_{D_{X,Y}}(x,y)dydx = c\int_0^1 x\Big(\int_x^1 ydy\Big)dx = \frac{c}{2}\int_0^1 x(1-x^2)dx = \frac{c}{8},$$

т.е. $c = 8$, като използваме $\mathbb{1}_{D_{X,Y}}(x,y) = \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(x,1)}(y)$. От тук $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = 8x(\int_x^1 ydy) \cdot \mathbb{1}_{(0,1)}(x) = 4x(1-x^2) \cdot \mathbb{1}_{(0,1)}(x)$ и $f_Y(y) = 4y^3 \cdot \mathbb{1}_{(0,1)}(y)$, откъдето $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)dx = 4\int_0^1 x^2(1-x^2)dx = \frac{8}{15}$ и $\mathbb{E}[Y] = \frac{4}{5}$. Относно (в), нека $B = \{(x,y) \in \mathbb{R}^2 : x < \frac{3}{4}, y < x + \frac{1}{6}\}$. Тогава

$$\mathbb{P}(X < 3/4, Y - X < 1/6) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{1}_B(x,y) f_{X,Y}(x,y)dydx = 8\int_0^{\frac{3}{4}} x\Big(\int_x^{x+\frac{1}{6}} ydy\Big)dx \approx 0.22,$$

като използваме $\mathbb{1}_{D_{X,Y}}(x,y) \cdot \mathbb{1}_B(x,y) = \mathbb{1}_{(0,\frac{3}{4})}(x) \cdot \mathbb{1}_{(x,x+\frac{1}{6})}(y)$. Относно (г), нека $C = \{(x,y) \in \mathbb{R}^2 : y < x + \frac{1}{3}\}$. Тогава

$$\mathbb{P}(Y - X < 1/3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{1}_C(x,y) f_{X,Y}(x,y)dydx = \int_0^{\frac{2}{3}}\int_x^{x+\frac{1}{3}} 8xydydx + \int_{\frac{2}{3}}^1\int_x^1 8xydydx \approx 0.67,$$

като използваме $\mathbb{1}_{D_{X,Y}}(x,y) \cdot \mathbb{1}_C(x,y) = \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(x,\min\{x+\frac{1}{3},1\})}(y) = \mathbb{1}_{(0,\frac{2}{3})}(x) \cdot \mathbb{1}_{(x,x+\frac{1}{3})}(y) + \mathbb{1}_{(\frac{2}{3},1)}(x) \cdot \mathbb{1}_{(x,1)}(y)$. Относно (д)

$$\mathbb{E}[Y - X | X = 1/4] = \int_{-\infty}^{\infty}\Big(y - \frac{1}{4}\Big) f_{Y|X}(y|1/4)dy = \int_{-\infty}^{\infty}\Big(y - \frac{1}{4}\Big)\frac{f_{X,Y}(y,\frac{1}{4})}{f_X(\frac{1}{4})}dy = 8\int_{\frac{1}{4}}^1\Big(y - \frac{1}{4}\Big)\frac{2y}{1 - \frac{1}{16}}dy = 0.45.$$

□

Задача 4.14. Точка (X,Y) попада по случаен начин в триъгълник с върхове в точките с координати $(0,0)$, $(0,2)$ и $(3,0)$. Да се намери съвместната плътност, функцията на разпределение и корелацията на X и Y.

Решение. Използвайки метода за пресмятане на т. нар. "геометрични" вероятности,

$$F_{X,Y}(x,y) = \begin{cases} 0, & x < 0 \text{ или } y < 0 \\ \frac{xy}{3}, & x \in (0,3), y \in (0,2), 2x + 3y \leq 6 \\ 1 - \frac{(3-x)^2}{9} - \frac{(2-y)^2}{4}, & x \in (0,3), y \in (0,2), 2x + 3y > 6 \\ \frac{(4-y)y}{4}, & x > 3, y \in (0,2) \\ \frac{(6-x)x}{9}, & x \in (0,3), y > 2 \\ 1, & x > 3, y > 2 \end{cases}$$

откъдето следва, че $f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \frac{1}{3} \cdot \mathbb{1}_{D_{X,Y}}(x,y)$. Всъщност, ако (X,Y) има плътност $f_{X,Y}$, то $f_{X,Y}$ е дефинирана само върху $D_{X,Y} = \{(x,y) \in [0,3] \times [0,2] : 3y + 2x \leq 6\}$, откъдето и полученият резултат. Тогава

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dxdy = \frac{1}{3}\int_0^3\int_0^{2-\frac{2}{3}x} xydydx = \frac{1}{3}\int_0^3 x\Big(2-\frac{2}{3}x\Big)dx = \frac{1}{2}.$$

По същия начин $\mathbb{E}[X] = 1$ и $\mathbb{E}[Y] = \frac{2}{3}$, откъдето $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = -\frac{1}{6}$.

Алт. решение. Маргинално, $F_X(x) = \frac{\text{лице трапец с т. } (0,0),(x,0),(x,2-\frac{2}{3}x),(0,2)}{\text{лице триъгълник с т. } (0,0),(3,0),(0,2)}$ за $x \in (0,3)$, т.е. $F_X(x) = \frac{2}{3}x - \frac{1}{9}x^2$, откъдето $f_X(x) = (\frac{2}{3} - \frac{2}{9}x)\cdot \mathbb{1}_{(0,3)}(x)$. От друга страна, $Y|X = x \sim \mathrm{Unif}(0, 2-\frac{2}{3}x)$ за $x \in (0,3)$. Следователно $f_{Y|X}(y|x) = \frac{1}{2-\frac{2}{3}x}\cdot \mathbb{1}_{(0,2-\frac{2}{3}x)}(y)$, откъдето

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) = \frac{1}{3}\cdot \mathbb{1}_{D_{X,Y}}(x,y).$$

Аналогично на едномерния случай, формата на $f_{X,Y}$ показва, че точките (X,Y) са разпределени равномерно. □

*Задача** 4.15. Нека X и Y са сл. в. с функция на разпределение $F_{X,Y}$. Изразете функцията на разпределение на $(\max\{X,Y\}, \min\{X,Y\})$ чрез $F_{X,Y}$.

Решение. За всички $s,t \in \mathbb{R} : s < t$, имаме

$$\begin{aligned} F_{\max\{X,Y\},\min\{X,Y\}}(t,s) &= \mathbb{P}(\max\{X,Y\} < t, \min\{X,Y\} < s) = \mathbb{P}(\max\{X,Y\} < t) - \mathbb{P}(\max\{X,Y\} < t, \min\{X,Y\} \geq s) \\ &= \mathbb{P}(X < t, Y < t) - \mathbb{P}(X < t, Y < t, X \geq s, Y \geq s) = \mathbb{P}(X < t, Y < t) - \mathbb{P}(s \leq X < t, s \leq Y < t) = \\ &= F_{X,Y}(t,t) - \big(F_{X,Y}(t,t) - F_{X,Y}(t,s) - F_{X,Y}(s,t) + F_{X,Y}(s,s)\big) = F_{X,Y}(t,s) + F_{X,Y}(s,t) - F_{X,Y}(s,s). \end{aligned}$$

От друга страна $F_{\max\{X,Y\},\min\{X,Y\}}(t,s) = \mathbb{P}(\max\{X,Y\} < t) = F_{X,Y}(t,t)$, за $s,t \in \mathbb{R} : s \geq t$. □

. . . .

Независимост. *Група от непрекъснати сл. в.* $X_1,\dots,X_n$ *се наричат независими, ако за произволни* $x_1,\dots,x_n \in \mathbb{R}$ *е изпълнено*

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n).$$

Забележка (Линейна зависимост). Ако X,Y са непрекъсната сл. в. и $X \perp\!\!\!\perp Y$, тогава

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mathbb{E}[X])(y-\mathbb{E}[Y])f_X(x)f_Y(y)dydx = \int_{-\infty}^{\infty}(x-\mathbb{E}[X])\Big(\int_{-\infty}^{\infty} yf_Y(y)dy - \mathbb{E}[Y]\Big)f_X(x)dx = 0.$$

Обратното, ако $\mathrm{Cov}(X,Y) = 0$, това не значи непременно, че $X \perp\!\!\!\perp Y$. Ковариацията между две сл. в. ни информира единствено за тяхната линейна зависимост.

Твърдение 4.8. *Ако* $X \perp\!\!\!\perp Y$, *то* $f(X) \perp\!\!\!\perp g(Y)$ *и* $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ *за измерими функции* $f : \mathbb{R} \to \mathbb{R}$ *и* $g : \mathbb{R} \to \mathbb{R}$.

Д-во. Лесно се вижда, че $\mathbb{P}(X \in A, Y \in B) = \int_A\int_B f_{X,Y}(x,y)dxdy = \int_A f_X(x)dx \cdot \int_B f_Y(y)dy = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ за $A \subseteq \mathbb{X}$ и $B \subseteq \mathbb{Y}$, ако $X \perp\!\!\!\perp Y$. Тогава, за $C, D \in \mathcal{B}(\mathbb{R})$, е валидно

$$\mathbb{P}(f(X) \in C, g(Y) \in D) = \mathbb{P}\big(X \in f^{-1}(C), Y \in g^{-1}(D)\big) = \mathbb{P}(f(X) \in C)\mathbb{P}(g(Y) \in D).$$

Относно втория резултат,

$$\mathbb{E}[f(X)g(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)g(y)f_{X,Y}(x,y)dxdy = \int_{-\infty}^{\infty} f(x)\Big(\int_{-\infty}^{\infty} g(y)f_Y(y)dy\Big)f_X(x)dx = \mathbb{E}[f(X)]\mathbb{E}[g(Y)].$$

□

Задача 4.16. Нека $X \sim \mathrm{Exp}(2)$ и $Y \sim \mathrm{Unif}(0,3)$ са независими. Да се намери разпределението на X/Y.

Решение. По условие $f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y) = \frac{2}{3}e^{-2x}\cdot\mathbb{1}_{D_{X,Y}}(x,y)$, където $D_{X,Y} = (0,\infty)\times(0,3)$. Означаваме $Z = \frac{X}{Y}$. Нека $z \in (0,\infty)$. Да положим $C = \{(x,y)\in\mathbb{R}^2 : x < zy\}$. Тогава

$$F_Z(z) = \mathbb{P}(Z<z) = \mathbb{P}(X<zY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathbb{1}_C(x,y)f_{X,Y}(x,y)dydx = \frac{2}{3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-2x}\cdot\mathbb{1}_{C\cap D_{X,Y}}(x,y)(x,y)dydx$$
$$= \frac{2}{3}\int_0^{3z}\int_{\frac{x}{z}}^{3}e^{-2x}dydx = \frac{2}{3}\int_0^{3z}\Big(3-\frac{x}{z}\Big)e^{-2x}dx = \int_{-6z}^{0}e^u du + \frac{1}{6z}\int_{-6z}^{0}ue^u du = 1 - \frac{1}{6z} + \frac{e^{-6z}}{6z},$$

като използваме, че $\mathbb{1}_{C\cap D_{X,Y}}(x,y) = \mathbb{1}_{(0,3z)}(x)\cdot\mathbb{1}_{(\frac{x}{z},3)}(y)$. В допълнение, $F_Z(z) = 0$, за $z \le 0$; т.е. F_Z е непрекъснато диференцируема функция, откъдето Z има плътност $f_Z(z) = \frac{d}{dy}F_Z(z) = \frac{1-(1+6z)e^{-6z}}{6z^2}\cdot\mathbb{1}_{(0,\infty)}(z)$. $\square$

Задача 4.17. Нека $X_1 \sim \mathrm{Exp}(\lambda_1)$ и $X_2 \sim \mathrm{Exp}(\lambda_2)$ са независими сл. в. Да се намери разпределението на (а) $\max\{X_1,X_2\}$; (б) $\min\{X_1,X_2\}$.

Решение. Относно (а), за всяко $z\in(0,\infty)$,

$$F_{\max\{X_1,X_2\}}(z) = \mathbb{P}(X_1<z, X_2<z) = \mathbb{P}(X_1<z)\mathbb{P}(X_2<z) = (1-e^{-\lambda_1 z})(1-e^{-\lambda_2 z}),$$

$F_{\max\{X_1,X_2\}}(z) = 0$ за $z\le 0$; т.е. $F_{\max\{X_1,X_2\}}$ е непрекъснато диференцируема функция, откъдето $\max\{X_1,X_2\}$ е непрекъсната сл. в. с плътност $f_{\max\{X_1,X_2\}}(z) = \frac{d}{dz}F_{\max\{X_1,X_2\}}(z)\cdot\mathbb{1}_{(0,\infty)}(z)$. Относно (б), за всяко $z\in(0,\infty)$,

$$F_{\min\{X_1,X_2\}}(z) = 1-\mathbb{P}(\min\{X_1,X_2\}\ge z) = 1-\mathbb{P}(X_1\ge z, X_2\ge z) = 1-\mathbb{P}(X_1\ge z)\mathbb{P}(X_2\ge z) = 1-e^{-(\lambda_1+\lambda_2)z},$$

$F_{\min\{X_1,X_2\}}(z) = 0$ за $z\le 0$; т.е. $F_{\min\{X_1,X_2\}}$ е непрекъснато диференцируема функция, откъдето $\min\{X_1,X_2\}$ е непрекъсната сл. в. с плътност $f_{\min\{X_1,X_2\}}(z) = (\lambda_1+\lambda_2)e^{-(\lambda_1+\lambda_2)z}\cdot\mathbb{1}_{(0,\infty)}(z)$. Следователно $\min\{X_1,X_2\}\sim \mathrm{Exp}(\lambda_1+\lambda_2)$. $\square$

....

Твърдение 4.9 (Конволюция)**.** *Нека X, Y са непрекъснати и независими сл. в., съответно с плътности f_X и f_Y. Тогава $X+Y$ е непрекъсната сл. в. с плътност*

$$f_{X+Y}(z) = \int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx, \qquad \text{за } z\in\mathbb{R}.$$

Д-во. За всяко $z\in\mathbb{R}$ имаме

$$F_{X+Y}(z) = \mathbb{P}(X+Y<z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x}f_{XY}(x,y)dydx = \int_{-\infty}^{\infty}f_X(x) = \int_{-\infty}^{\infty}f_X(x)F_Y(z-x)dx,$$

откъдето $f_{X+Y}(z) = \frac{d}{dz}F_{X+Y}(z) = \int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx$. $\square$

Задача 4.18. Нека $X, Y\sim \mathrm{Unif}(0,1)$ са независими. Да се намери разпределението на $X+Y$.

Решение. От Твърдение 4.9 следва, че $X+Y$ е непрекъсната сл. в. с плътност

$$f_{X+Y}(z) = \int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx = \int_{-\infty}^{\infty}\mathbb{1}_{(0,1)}(x)\cdot\mathbb{1}_{(0,1)}(z-x)dx = \int_{-\infty}^{\infty}\mathbb{1}_{(0,1)}(x)\cdot\mathbb{1}_{(z-1,z)}(x)dx$$
$$= \int_{-\infty}^{\infty}\mathbb{1}_{(\max\{0,z-1\},\min\{1,z\})}(x)dx = \int_{-\infty}^{\infty}\mathbb{1}_{(0,z)}(x)dx\cdot\mathbb{1}_{(0,1)}(z) + \int_{-\infty}^{\infty}\mathbb{1}_{(z-1,1)}(x)dx\cdot\mathbb{1}_{[1,2)}(z) =$$
$$= \int_0^z dx\cdot\mathbb{1}_{(0,1)}(z) + \int_{z-1}^{1}dx\cdot\mathbb{1}_{[1,2)}(z) = z\cdot\mathbb{1}_{(0,1)}(z) + (2-z)\cdot\mathbb{1}_{[1,2)}(z).$$

$\square$

Твърдение 4.10 (Смяна на променливи)**.** *Нека (X,Y) е вектор от непрекъснати сл. в. с плътност $f_{X,Y}$ и $g:\mathbb{R}^2\to\mathbb{R}^2$ е биективна функция върху $D_{X,Y}$ с обратна функция $h = g^{-1}$. Ако h и g са непрекъснати, а h е диференцируема, то $g(X,Y)$ е вектор от непрекъснати сл. в. с плътност*

$$f_{g(X,Y)}(u,v) = f_{X,Y}\big(h(u,v)\big)|J(u,v)|\cdot\mathbb{1}_{g(D_{X,Y})}(u,v),$$

където $J(u,v) := \det\big[\frac{\partial}{\partial \boldsymbol{z}}h(\boldsymbol{z})\big|_{\boldsymbol{z}=(u,v)}\big]$ детерминантата на матрицата на Якоби в т. (u,v).

Задача 4.19. Нека $X_1, X_2 \sim \mathrm{Exp}(\lambda)$ са независими. Да се намери разпределението на $Y = \frac{X_1}{X_1+X_2}$.

Решение. По условие $f_{X_1,X_2}(x_1,x_2) = \lambda^2 e^{-\lambda(x_1+x_2)} \cdot \mathbb{1}_{(0,\infty)^2}(x_1,x_2)$. Обозначаваме $Z = X_1$. Функцията $(x_1,x_2) \mapsto (\frac{x_1}{x_1+x_2}, x_1)$ е дифеоморфизъм с обратна функция $(y,z) \mapsto (z, \frac{z}{y} - z)$. От Твърдение 4.10 следва, че (Y,Z) е вектор от непрекъснати сл. в. с плътност

$$f_{Y,Z}(y,z) = f_{X_1,X_2}(z, z/y - z)|J(y,z)| \cdot \mathbb{1}_{(0,\infty)^2}(z, z/y - z) = \lambda^2 e^{-\lambda\frac{z}{y}} \frac{z}{y^2} \cdot \mathbb{1}_{(0,1)}(y) \cdot \mathbb{1}_{(0,\infty)}(z),$$

където $J(y,z) = \det\begin{bmatrix} 0 & 1 \\ -\frac{z}{y^2} & \frac{1}{y} - 1 \end{bmatrix} = \frac{z}{y^2}$. Тогава

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y,Z}(y,z)dz = \int_0^{\infty} \frac{\lambda^2 z}{y^2} e^{-\lambda\frac{z}{y}} dy \cdot \mathbb{1}_{(0,1)}(y) = \int_0^{\infty} u e^{-u} du \cdot \mathbb{1}_{(0,1)}(y) = \mathbb{1}_{(0,1)}(y);$$

т.е. $Y \sim \mathrm{Unif}(0,1)$.

Алт. решение. Нека $y \in (0,1)$. Да положим $C = \{(x_1,x_2) \in \mathbb{R}^2 : \frac{1-y}{y} x_1 < x_2\}$. Тогава

$$\begin{aligned} F_Y(y) = \mathbb{P}(Y < y) = \mathbb{P}\Big(\frac{1-y}{y} X_1 < X_2\Big) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{1}_C(x_1,x_2) f_{X_1,X_2}(x_1,x_2) dx_2 dx_1 \\ &= \int_0^{\infty}\int_{\frac{1-y}{y}x_1}^{\infty} \lambda^2 e^{-\lambda(x_1+x_2)} dx_2 dx_1 = \int_0^{\infty} \lambda e^{-\lambda x_1} \int_{-\infty}^{-\lambda\frac{1-y}{y}x_1} e^u du dx_1 = \int_0^{\infty} \lambda e^{-\frac{\lambda}{y}x_1} dx_1 = y \int_{-\infty}^{0} e^v dv = y. \end{aligned}$$

В допълнение, $F_Y(y) = 0$, за $y \le 0$, и $F_Y(y) = 1$, за $y \ge 1$; т.е. F_Y е непрекъснато диференцируема функция, откъдето следва, че Y е непрекъсната сл. в. с плътност $f_Y(y) = \mathbb{1}_{(0,1)}(y)$. $\square$

5 Източници

ДИМИТРОВ, Б. и ЯНЕВ, Н. (2007). *Вероятности и статистика*, Софтех, София.

СТОЯНОВ, Й., МИРАЗЧИЙСКИ, И., ИГНАТОВ, Ц. и ТАНУШЕВ, М. (1976). *Ръководство за упражнения по теория на вероятностите*, Наука и изкуство, София.

ЧУКАНОВ, В. (1977). *Комбинаторика*, Народна просвета, София.

GRIMMETT, G. and STIRZAKER, D. (2001). *Probability and Random Processes*, 3rd Edition, Oxford University Press, UK.