Теория на вероятностите I

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

Легенда. *Основен материал*: с (*) е видяно на лекции и/или се отнася до предварителни познания; с (**) – за по-любознателните. *Задачи*: с (*) – допълнителна подготовка; с (**) – трудни задачи.

1 Комбинаторика

Комбинаториката е анализ на крайните множества, напр. $M = \{a_1, \dots, a_n\}$ за $n \in \mathbb{N}$, като в частност служи за намиране на броя на техните елементи (в случая $|M| = n$) или броя на дадени комбинаторни конфигурации, използващи M като опорно множество, 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*

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

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 $\{a_1, \dots, a_{n-1}\}$ получаваме n на брой пермутации на $\{a_1, \dots, a_n\}$, като поставяме a_n съответно на първа, втора и т.н. до последна позиция; напр. от $(a_{i_1}, \dots, a_{i_{n-1}})$, за кои да е различни $i_1, \dots, i_{n-1} \in \{1, \dots, n-1\}$, получаваме

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

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

От друга страна, всяка пермутация на $\{a_1, \dots, a_n\}$ дефинира една единствена пермутация на $\{a_1, \dots, a_{n-1}\}$ след премахване на елемента a_n. Следователно $\mathbf{P}_n = B_1 \cup B_2 \cup \dots \cup B_{P_{n-1}}$, откъдето

$$P_n = \left|\bigcup_{l=1}^{P_{n-1}} B_l\right| = \sum_{l=1}^{P_{n-1}} |B_l| = n \cdot P_{n-1} = n!$$

□

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 (Комбинаторно доказателство). За да преброим елементите на дадено множество A, избираме друго подходящо множество B, за което знаем $|B|$, и доказваме съществуването на биективна функция от A в B, от което следва, че $|A| = |B|$.

*Забележка*** (Stirling's approximation). С нарастването на n числата $n!$ растат много бързо и непосредственото им пресмятане става практически невъзможно. Съществува обаче формула (на Стирлинг), 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 $k = 0, 1 \dots, n$, наричаме вариация на елементите на M от клас k без повторение, като множеството от вариации обозначаваме с*

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

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 $\mathbf{P}_n = B_1 \cup B_2 \cup \dots \cup B_{V_n^k}$, откъдето получаваме, че

$$P_n = \left| \bigcup_{l=1}^{V_n^k} B_l \right| = \sum_{l=1}^{V_n^k} |B_l| = P_{n-k} \cdot V_n^k \qquad \Longrightarrow \qquad V_n^k = \frac{P_n}{P_{n-k}}.$$

$\square$

Пример 1.1 (Урна с топки). 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. Ако допуснем, че извадените топки *не* се връщат в урната, а редът на записване на изтеглените номера *има* значение, тогава броят на всички различни извадки е V_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 от комбинации 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 $V_n^k = P_k \cdot C_n^k$. $\square$

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.

Пример 1.2 (Урна с топки - ctd.). Ако допуснем, че в Пример 1.1 извадените топки *не* се връщат и че редът на записване на изтеглените номера *няма* значение, тогава броят на различните извадки е C_n^k.

Твърдение 1.1 (Свойства).

(i) $\binom{n}{k} = \binom{n}{n-k}$;

(ii) $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$;

(iii) $k\binom{n}{k} = n\binom{n-1}{k-1}$;

(iv) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$;

(v) $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$.

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.

Твърдение (iii) следва от наблюдението, че a_n участва в $\frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{k}{n}$ от комбинациите на M от клас k, което ги превръща в елементи на $\mathbf{C}_{n-1}^{k-1}$ след премахване на a_n.

(iv) и (v) 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, (v) следва от принципа за включване-изключване, приложен върху множествата $A_i = \{a\}$, $i = 1, \ldots, n$. Относно (iv), $\mathbf{C}_n^0, \mathbf{C}_n^1, \ldots, \mathbf{C}_n^n$ са две по две непресичащи се множества, чието обединение $\mathcal{P}(M) := \mathbf{C}_n^0 \cup \mathbf{C}_n^1 \cup \cdots \cup \mathbf{C}_n^n$ 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 $(e_1, \ldots, e_n) \in \{0,1\}^n$ от нули и единици, такава че $a_i \in A \Longleftrightarrow e_i = 1$ за $i = 1, \ldots, n$, 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.

$$\begin{array}{c} C_0^0 \\ C_1^0 \quad C_1^1 \\ C_2^0 \quad C_2^1 \quad C_2^2 \\ \ldots\ldots\ldots\ldots\ldots \\ C_n^0 \quad C_n^1 \quad \ldots \quad C_n^{n-1} \quad C_n^n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array} \qquad \begin{array}{c} 1 \\ 1 \quad 1 \\ 1 \quad 2 \quad 1 \\ 1 \quad 3 \quad 3 \quad 1 \\ \underline{1} \quad \underline{\underline{4}} \quad \underline{6} \quad \underline{\underline{4}} \quad \underline{1} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array} \qquad \begin{array}{c} 2^0 \\ 2^1 \\ 2^2 \\ 2^3 \\ 2^4 \end{array}$$

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 $(k_1, \ldots, k_n)$, като множеството от пермутации обозначаваме с*

$$\boldsymbol{P}(k; k_1, \ldots, k_n) = \big\{(a_{i_1}, \ldots, a_{i_k}) : i_1, \ldots, i_k \in \{1, \ldots, n\}, \#\{j : a_{i_j} = a_1\} = k_1, \ldots, \#\{j : a_{i_j} = a_n\} = k_n\big\}.$$

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

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

Д-во.* Ако $(a_{i_1}, \ldots, a_{i_k})$ е такава k-орка, да означим с $X_1 = \{j \in \{1, \ldots, k\} : i_j = 1\}$ множеството от индекси j, за които $a_{i_j} = a_1$. Понеже $|X_1| = k_1$, то за X_1 има $C_k^{k_1}$ на брой възможни комбинации. Нека $(a_{i_1^*}, \ldots, a_{i_{k-k_1}^*})$ е редицата, получена след премахване на всичките k_1 елемента a_1 от $(a_{i_1}, \ldots, a_{i_k})$. Тогава $(a_{i_1^*}, \ldots, a_{i_{k-k_1}^*})$ представлява пермутация на $\{a_2, \ldots, a_n\}$ с повторение, дължина $k - k_1$ и честоти $(k_2, \ldots, k_n)$.

От друга страна, $(a_{i_1^*}, \ldots, a_{i_{k-k_1}^*})$ и X_1 определят еднозначно $(a_{i_1}, \ldots, a_{i_k})$. Следователно,

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

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

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 $\binom{k-k_1}{k_2}$ различни начина, по които можем да изберем k_2 от останалите $k - k_1$ места, на които да поставим елемента $a_2, \ldots$, 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.

Пример 1.3 (MISSISSIPPI). Броят на отделните анаграми на думата MISSISSIPPI е равен на броят на пермутациите на $\{M, I, S, P\}$ с повторение, дължина 11 и честоти $(1,4,4,2)$; т.е. $P(11;1,4,4,2)$.

*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 $M = \{1,\dots,n\}$ 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 $a_{i_j} = m \iff i_j \in X_m$, 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 $k \in \mathbb{N}$, наричаме вариация на елементите на M от клас k с повторение, като множеството от вариации 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 $V(n,k) = |M|^k$. $\square$

Пример 1.4 (Урна с топки - ctd.). Ако допуснем, че в Пример 1.1 извадените топки се връщат обратно и че редът на записване на изтеглените номера *има* значение, тогава броят на всички различни извадки е $V(n,k)$.

*Забележка*** (Разлагания на множество - ctd.). Нека E е k-елементно множество. Следвайки същите разсъждения като преди, 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 <u>мулти</u>-подмножество на M, за $k \in \mathbb{N}$, наричаме комбинация на елементите на M от клас k с повторение, като множеството от комбинации 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

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

Д-во (Техника).* Допълваме множеството M с $k-1$ нови елемента $a_{n+1}, a_{n+2},\dots,a_{n+k-1}$, при което получаваме ново множество $M^* = \{a_1,\dots,a_{n+k-1}\}$. Нека $\{a_{i_1},\dots,a_{i_k}\} \in \mathbf{C}(n,k)$ е k-елементно мулти-подмножество на M, където допускаме, че $i_1 \le i_2 \le \cdots \le i_k$ *без загуба на общност.* Полагаме $i_j^* := i_j + (j-1)$, за $j = 1,\dots,k$. Тогава $\{a_{i_1^*},\dots,a_{i_k^*}\}$ представлява комбинация от елементи на M^* *без повторение* от клас k, т.е. $\{a_{i_1^*},\dots,a_{i_k^*}\} \in \mathbf{C}_{n+k-1}^k$.

От друга страна лесно можем да възстановим $\{a_{i_1},\dots,a_{i_k}\}$ от $\{a_{i_1^*},\dots,a_{i_k^*}\}$, полагайки $i_j := i_j^* - (j-1)$. Следователно съответствието между $\mathbf{C}(n,k)$ и $\mathbf{C}_{n+k-1}^k$ е взаимно еднозначно, откъдето $C(n,k) = C_{n+k-1}^k$. $\square$

Пример 1.5 (Урна с топки - ctd.). Ако допуснем, че в Пример 1.1 извадените топки се връщат обратно и редът на записване на изтеглените номера *няма* значение, тогава броят на всички различни извадки е $C(n,k)$.

Твърдение 1.2** (Свойства). *$C(n,k)=\sum_{i=0}^{k}C(n-1,i)$, 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, където a_n се среща m пъти, за някое $m\in 0,1,\dots,k$. 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 към нея прибавим m пъти елемента 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 установената рекурентна 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. От друга страна, уравнението $C(n,k)=\sum_{i=0}^{k}C(n-1,i)$ се демонстрира от защрихованите елементи.

$$
\begin{array}{c}
\cancel{C(1,0)}\\
C(2,0)\quad \cancel{C(1,1)}\\
C(3,0)\quad C(2,1)\quad \cancel{C(1,2)}\\
C(4,0)\quad C(3,1)\quad \cancel{C(2,2)}\quad C(1,3)\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
C(m,0)\quad C(m-1,1)\quad \dots\quad C(2,m-2)\quad C(1,m-1)\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{array}
$$

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

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

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 $c_k:=b_{k-n}$, $k=n+1,\dots,n+m$. 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

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

*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 (виж Задача 2.10), ..., 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 удостоверяваме

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

т.е. елементът x е преброен точно по веднъж от двете страни на формулата (1). $\square$

Задачи

Задача 1.1. 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 (за $k \leq n$);

(б) частиците са различими/неразличими и клетките могат да съдържат произволен брой частици;

(в) частиците са различими/неразличими и няма празна клетка (за $k \geq n$);

(г) частиците са различими/неразличими, в клетка 1 има точно s частици, а в останалите клетки – най-много по една частица (за $k \geq s \geq k - n + 1$);

(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.

Задача 1.2. 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?

Задача 1.3. 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?

Задача 1.4. 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?

Задача 1.5. 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.

Задача 1.6. Нека $M = \{a_1, \ldots, a_m, b_1, \ldots, b_n\}$. Обозначаваме с A_i (респ. B_j) множеството от всички k-елементни подмножества на M, съдържащи точно i (респ. j) обекта от тип a (респ. b), за $i = 0, 1, \ldots, \min\{m, k\}$ (респ. $j = 0, 1, \ldots, \min\{n, k\}$).

(а) Да се пресметнат $|A_i|$ и $|B_j|$.

(б) Да се докаже комбинаторно, че $\sum_{i=\max\{0,k-n\}}^{\min\{m,k\}} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}$. *(формула на Вандермонд)*

(в) Колко са на брой k-елементните подмножества на M, които съдържат поне един обект от тип a и поне един обект от тип b?

(г) Колко са на брой подмножествата на M, които съдържат поне един обект от тип a и поне един обект от тип b?

Задача 1.7. Десет души се нареждат в редица. Колко са подрежданията, при които три фиксирани лица се намират едно до друго?

Задача 1.8. Колко е броят на думите с дължина n и съдържащи само символите a, b и c, които

(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 са a;

(г) съдържат съответно k_1, k_2 и k_3 пъти (за $k_1 + k_2 + k_3 = n$) символите a, b и c.

*Задача** 1.9. По колко различни начина $2n$ шахматиста могат да образуват $k \leq n$ 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?

*Задача*** 1.10. 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 $y = r$?

2 Вероятности

2.1 Дискретни вероятности

Нека множеството от възможните изходи (елементарни събития) на даден случаен експеримент е изброимо голямо, т.е. $\Omega = \{\omega_1, \omega_2, \ldots\}$. Към всяко елементарно събитие $\omega_n \in \Omega$ съотнасяме едно число $p(\omega_n)$, такова че $p(\omega_n) \geq 0$ и $\sum_n p(\omega_n) = 1$, а на всяко събитие $A \in \mathcal{P}(\Omega)$, да кажем $A = \{\omega_{i_1}, \omega_{i_2}, \ldots\}$, съпоставяме

$$\mathbb{P}(A) := \sum_m p(\omega_{i_m}) \equiv \sum_{n: w_n \in A} p(\omega_n). \tag{2}$$

Функцията $\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$ дефинира вероятностна мярка върху $(\Omega, \mathcal{P}(\Omega))$, наричана още дискретна, а тройката $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ образува т. нар. дискретно вероятностно пространство.

Обратното твърдение също е вярно. Ако $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ е вероятностно пространство, такова че Ω е изброимо голямо, то вероятностната мярка $\mathbb{P}$ има представяне (2), използвайки числата $p(\omega) := \mathbb{P}(\{\omega\})$, за $\omega \in \Omega$.

Равномерни вероятности. *Нека $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ е дискретно вероятностно пространство (в. п.), където $\Omega = \{\omega_1, \ldots, \omega_n\}$ е крайно множество. Ако елементарните събития са еднакво вероятни, т.е. $p(\omega_i) = p$ за всяко $i = 1, \ldots, n$, тогава от $\mathbb{P}(\Omega) = 1$ следва, че $p = \frac{1}{n}$, откъдето, за всяко $A = \{\omega_{i_1}, \ldots, \omega_{i_m}\} \in \mathcal{P}(\Omega)$,*

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

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, като при пресмятане на самите вероятности широко се използва комбинаторен анализ.

Пример 2.1 (Урна с топки). Урна съдържа M черни и $N - M$ бели топки. Правим случайна ненаредена извадка без връщане с обем $n \leq N$, т.е. $\Omega = \big\{\{a_1, \ldots, a_n\} : a_i \in \{\text{ч}_1, \ldots, \text{ч}_M, \text{б}_1, \ldots, \text{б}_{N-M}\}, a_i \neq a_j, i \neq j\big\}$. Ако допуснем, че елементарните събития са еднакво вероятни, то от Задача 1.6 следва, че вероятността извадката да съдържа точно $k \leq n$ черни топки, за някое $M - (N - n) \leq k \leq M$, е

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

От формулата на Вандермонд е видно, че $\sum_{k=0}^{n} \mathbb{P}(\{k \text{ черни от } n\}) = \sum_{k=0}^{n} \binom{M}{k}\binom{N-M}{n-k} / \binom{N}{n} = \binom{N}{n} / \binom{N}{n} = 1$.

Задачи

Задача 2.1. Да се пресметне вероятността при хвърлянето на два правилни зара да се падне поне една единица, ако приемете, че:
(а) заровете са различими и различните изходи са еднакво вероятни;
(б) заровете са неразличими и различните изходи са еднакво вероятни.

Описват ли вярно реалната действителност и двата модела?

Задача 2.2. Каква е вероятността случайно избрана плочка от домино да съдържа различни числа на двете си половини?

Задача 2.3. При игра на тото 6 от 49 да се пресметнат вероятностите за печалба на шестица, петица, четворка и тройка.

Задача 2.4. Във влак с три вагона по случаен начин се качват седем пътника. Каква е вероятността точно четирима души да се качат в първия вагон?

Задача 2.5. Ако номерата на колите са равномерно разпределени четирицифрени числа, 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?

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

Задача 2.7. Партида се състои от n доброкачествени и m бракувани изделия. За проверка по случаен начин са взети s изделия. При проверката се оказало, че първите k проверявани изделия са доброкачествени ($k < s$). Да се пресметне вероятността $(k+1)$-вото изделие да се окаже доброкачествено.

Задача 2.8. С цел намаляване броя на изиграните мачове, $2k$ отбора се разделят чрез жребий на две равни по брой групи. Да се определи вероятността двата най-силни отбора да са в различни групи.

Задача 2.9. Хвърлят се 10 зара. Каква е вероятността да се паднат равен брой единици и шестици?

Задача 2.10. От урна, съдържаща топки с номера $1, \ldots, n$, се вадят последователно k топки ($k \leq n$). Каква е вероятността номерата на извадените топки, записани по реда на изваждането им, да образуват растяща редица, ако (а) теглим без да връщаме; (б) изтеглените топки се връщат обратно в урната?

Задача 2.11. Около една маса сядат 10 мъже и 10 жени. Каква е вероятността лицата от един и същи пол да не седят един до друг?

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?

Задача 2.13. Група от n човека се нарежда в редица по случаен начин. Каква е вероятността между две фиксирани лица да има точно $r \leq n-2$ човека? А ако се нареждат в кръг и $r \neq \frac{n-2}{2}$?

*Задача** 2.14. Нека k *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 r частици ($r \leq k$);
- (б) точно m клетки да са празни ($m < n$);
- (в) във всяка клетка да има поне по две частици ($k \geq 2n$);
- (г) във всяка клетка да има най-много по четири частици ($k \leq 4n$).

*Задача** 2.15. Нека k *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 ($k < n$).

*Задача** 2.16. Нека k *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 $k_1 + \cdots + k_n = k$;
(б) при $n = k$ нито една клетка да не остане празна;
(в) при $n = k$ да остане празна точно една клетка;
(г) точно m клетки да са празни $(n - k \le m < n)$.

*Задача** 2.17. 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.

*Задача** 2.18. 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.

*Задача*** 2.19 (Bertrand's ballot problem). По време на избори за първия от двама кандидати са пуснати n бюлетини, а за втория кандидат – m бюлетини. Каква е вероятността при преброяването 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?

2.2 Условна вероятност и независимост на събития

Условна вероятност. *Нека $(\Omega, \mathcal{A}, \mathbb{P})$ е вероятностно пространство (в.п.), а $B \in \mathcal{A}$ е т.ч. $\mathbb{P}(B) > 0$. 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 $A_1, \ldots, A_n \in \mathcal{A}$ т.ч. $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, е в сила

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

*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 $\big(B, \mathcal{A} \cap B, \mathbb{P}(\cdot \mid B)\big)$, където $\mathbb{P}(\cdot \mid B) : \mathcal{A} \cap B \to [0,1]$ е добре дефинирана вероятностна мярка, за разлика от $\mathbb{P}(\cdot \cap B)$. В частност е вярно за $\mathbb{P}(\cdot \mid B)$, че $\mathbb{P}(B|B) = 1$,

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

и, за $A_1, A_2 \in \mathcal{A}$ т.ч. $A_1 \cap A_2 = \emptyset$,

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

Пример 2.2 (Урна с топки - ctd.). Нека в Пример 2.1 приемем, че топките се теглят една по една. Вероятността да изтеглим черна топка на първи ход е $\mathbb{P}(ч) = M/N$, а вероятността да изтеглим бяла топка на втори ход при положение, че сме изтеглили черна на първи ход, е $\mathbb{P}(ч,б|ч) = (N-M)/(N-1)$, и т.н. Тогава за всяка редица от n топки, k от които са черни (общо C_n^k на брой), напр. $\omega^{(k)} = (ч, б, ч, \ldots, б, ч)$, имаме

$$\mathbb{P}(ч, б, ч, \ldots, б, ч) = \mathbb{P}(ч)\mathbb{P}(ч, б|ч) \cdots \mathbb{P}(ч, б, ч, \ldots, б, ч|ч, б, ч, \ldots, б) = \frac{M}{N}\frac{N-M}{N-1} \cdots \frac{M-k+1}{N-n+1} = \frac{V_M^k V_{N-M}^{n-k}}{V_N^n},$$

което е инвариантно спрямо реда на изтеглените цветове. Следователно,

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

Сравнявайки горното решение с това в Пример 2.1, следва, че тегленето на всички n топки наведнъж или една по една е еквивалентно от вероятностна гледна точка.

Независимост. *Събитията $A_1, \ldots, A_n \in \mathcal{A}$ се наричат независими в съвкупност, ако за кои да е индекси $i_1, \ldots, i_k \in \{1, \ldots, n\} : i_j \neq i_l, j \neq l$, и $k = 2, \ldots, n$ имаме*

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

Две събития $A, B \in \mathcal{A}$ са независими, означено с $A \perp\!\!\!\perp B$, тогава и само тогава, когато $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Ако в допълнение $\mathbb{P}(B) > 0$, то

$$A \perp\!\!\!\perp B \qquad \Longleftrightarrow \qquad \mathbb{P}(A|B) = \mathbb{P}(A).$$

Забележка (Внимание). Две събития $A, B \in \mathcal{A}$ т.ч. $A \cap B = \emptyset$ не са непременно независими. Напротив, от настъпването на кое да е от тях следва, че другото събитие не може да се случи.

Пример 2.3. Хвърлят се два зара. Разглеждаме събитията $A_1 = \{$на първия зар се пада нечетно число$\}$, $A_2 = \{$на втория зар се пада нечетно число$\}$ и $A_3 = \{$сумата от точките на двата зара е нечетна$\}$. Тогава $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = 0.5$ и $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = 0.25$, но $\mathbb{P}(A_1 \cap A_2 \cap A_3) = 0$; т.е. събитията са две по две независими, но не са независими в съвкупност.

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. □

Формула за пълната вероятност. *Нека $H_1, H_2, \ldots \in \mathcal{A}$ образуват **пълна група от събития**, т.е. $\bigcup_n H_n = \Omega$, $H_n \cap H_m = \emptyset$, $n \neq m$, като допускаме, че $\mathbb{P}(H_n) > 0$. Тогава, за произволно събитие $A \in \mathcal{A}$,*

$$\mathbb{P}(A) = \mathbb{P}\Big(A \cap \Big(\bigcup_n H_n\Big)\Big) = \sum_n \mathbb{P}(A \cap H_n) = \sum_n \mathbb{P}(A|H_n)\mathbb{P}(H_n).$$

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

$$\mathbb{P}(H_n|A) = \frac{\mathbb{P}(A|H_n)\mathbb{P}(H_n)}{\sum_m \mathbb{P}(A|H_m)\mathbb{P}(H_m)}.$$

*Забележка** (ФПВ за УВ). Нека $A, B, C \in \mathcal{A}$. Ако $\mathbb{P}(B), \mathbb{P}(B \cap C), \mathbb{P}(B \cap C^c) > 0$, тогава е изпълнено

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

*Пример** 2.1. Две урни съдържат съответно $б_1$, $б_2$ бели топки и $ч_1$, $ч_2$ 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?

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

$$\mathbb{P}(A) = \sum_{i,j \in \{б,ч\}} \mathbb{P}(A|H_{i,j})\mathbb{P}(H_{i,j}) = \frac{2б_1б_2 + б_1ч_2 + ч_1б_2}{2(б_1 + ч_1)(б_2 + ч_2)}.$$

*Пример** 2.2. 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?

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

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

Задачи

Задача 2.20. Застрахователна компания води статистика за своите клиенти. Известно е, че: 1) всичи клиенти посещават лекар поне веднъж годишно; 2) 60% от тях посещават лекар повече от веднъж годишно; 3) 17% от тях 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?

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

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

Задача 2.23. Трима ловци стрелят едновременно по заек. Ако заекът е убит от точно един куршум, каква е вероятността той да е бил изстрелян от първия ловец, ако ловците уцелват съответно с вероятности 0.2, 0.4 и 0.6?

Задача 2.24 (Birthday Paradox). Какъв е най-малкият брой хора, които могат да бъдат избрани на случаен принцип, така че вероятността поне двама от тях да имат един и същ рожден ден да е по голяма от 50%?

Задача 2.25 (Boy or Girl Paradox). X има две деца. Ако по-голямото дете 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?

Задача 2.26. Кутия съдържа n билета, от които $m \leq n$ са печеливши. 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?

*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.

*Задача** 2.28. Монета с вероятност p за ези се хвърля n пъти. Нека E е събитието "пада се ези при първото хвърляне", а F_k е събитието "точно k пъти се пада тура". За кои двойки *цели* числа (n, k) събитията E и F_k са независими?

Задача 2.29. Разполагаме с тест за рядко заболяване, който е точен в 99% от случаите както при заразените (когато трябва да е положителен), така и при незаразените (когато трябва да е отрицателен). Ако знаете, че 0.5% от населението има това заболяване, каква е вероятността случайно избран човек с положителен тест да е болен?

Задача 2.30. 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 числа; в) последователни числа?

Задача 2.31. От урна, съдържаща топки с номера от 1 до n, се изваждат последователно две топки, като първата извадена топка се връща, ако номерът ѝ не е равен на 1. Да се пресметне вероятността топката с номер 2 да бъде извадена при второто теглене.

Задача 2.32. В кутия има 7 топки за тенис, 4 от които са нови. За първата игра по случаен начин се избират три от топките, като след играта те биват връщани обратно в кутията. За втората игра също случайно се избират три от топките. Каква е вероятността те да са съвсем нови?

Задача 2.33. На изпит се явяват 100 студенти, 55 момчета и 45 момичета. Момичетата взимат изпита с вероятност 0.7, а момчетата – с 0.4. След изпита се избират три резултата. Ако два от тях са били успешни, а един неуспешен, каква е вероятността и трите резултата да са на момичета?

Задача 2.34 (Bertrand's Box Paradox). Разполагате с 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 на масата. Ако горната страна на жетона е бяла, каква е вероятността другата му страна също да е бяла?

Задача 2.35 (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 която се оказва, че няма нищо. Сега трябва да решите – запазвате ли първоначалния си избор или го сменяте с другата неотворена врата?

Задача 2.36. 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 второ случайно избрано изделие от същата партида да се окаже бракувано.

Задача 2.37. Петнадесет изпитни билета съдържат по два въпроса и покриват целия конспект от 30 въпроса. Студент може да отговори на 25 от въпросите. Каква е вероятността той да си вземa изпита, ако за това е нужно той да отговори на двата въпроса от един билет или на един от двата въпроса, а след това и на посочен въпрос от друг билет?

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

*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?

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?

*Задача** 2.41. Раздаваме последователно картите от стандартно тесте. Ако за първи път видим червено асо на 6-та позиция, каква е вероятността след това да видим първо червено преди черно асо?

*Задача*** 2.42 (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?

3 Дискретни случайни величини

3.1 Основни понятия

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

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

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

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

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

X	x_1	x_2	$\cdots$	x_n	$\cdots$
$\mathbb{P}_X$	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

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

*Забележка*** (Интерпретация). Нека X е дискретна сл.в., приемаща стойности в $\mathbb{X} \subseteq \mathbb{R}$. Полагаме

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

където $\delta_x : 2^{\mathbb{X}} \to \{0,1\}$ е **Дирак масата** в точката $x \in \mathbb{X}$, зададена за всяко $A \subseteq \mathbb{X}$ 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 т. нар. дискретно в.п., чрез което свеждаме оригиналния случаен експеримент до негова трансформация, чиито елементарни изходи представляват възможните стойности на X. Освен това, за всяко $A \subseteq \mathbb{X}$, имаме $\{X \in A\} = \bigcup_{x\in A}\{X = x\} \in \mathcal{A}$, откъдето

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

В по-голямата част от случаите, ще работим директно върху $(\mathbb{X}, 2^{\mathbb{X}}, \mathbb{P}_X)$, приемайки съществуването на базовото в.п. $(\Omega, \mathcal{A}, \mathbb{P})$ имплицитно, като допускаме, че то е достатъчно богато, за да поддържа всяка сл.в., от която се нуждаем.

*Забележка** (Представяне на дискретна сл.в.). Нека $A \subseteq \Omega$. Под **индикаторна функция** на A ще разбираме функцията $\mathbb{1}_A : \Omega \to \{0,1\}$, зададена за всяко $\omega \in \Omega$ чрез

$$\mathbb{1}_A(\omega) = \begin{cases} 0, & \omega \notin A; \\ 1, & \omega \in A. \end{cases}$$

От равенството $\{\mathbb{1}_A = 1\} = A$ следва, че $Z := \mathbb{1}_A$ е (дискретна) сл.в. тогава и само тогава, когато $A \in \mathcal{A}$. Полагайки $p := \mathbb{P}(A) = \mathbb{P}(Z = 1)$, виждаме, че Z е бинарна сл.в., показваща с вероятност p дали събитието A е настъпило.

Нека X е дискретна сл.в., приемаща стойности в $\mathbb{X} \subseteq \mathbb{R}$. Полагаме $H_x := \{X = x\}$, за всяко $x \in \mathbb{X}$. Тогава колекцията от събития $\{H_x\}_{x\in\mathbb{X}}$ образува пълна група върху Ω, т.е. $H_{x'} \cap H_{x''} = \emptyset$, $x' \neq x''$, и $\Omega = \bigcup_{x\in\mathbb{X}} H_x$.Освен това $X(\omega) = x$ тогава и само тогава, когато $\omega \in H_x$, откъдето следва компактното представяне на X,

$$X(\omega) = \sum_{x\in\mathbb{X}} x \cdot \mathbb{1}_{H_x}(\omega), \qquad \text{за } \omega \in \Omega.$$

В случаите, когато $|\mathbb{X}| < \infty$, ще наричаме X **проста** сл.в.

*Забележка*** (Трансформация на сл.в.). Нека X е дискретна сл.в., приемаща стойности в $\mathbb{X} \subseteq \mathbb{R}^m$, и $f : \mathbb{R}^m \to \mathbb{R}^n$ е *произволна* функция. Тогава $Y := f(X)$ е дискретна сл.в. с възможни стойности $f(\mathbb{X}) := \{t \in \mathbb{R}^n : f(x) = t, \text{ за някое } x \in \mathbb{X}\}$. Това следва от факта, че за всяко $y \in f(\mathbb{X})$ имаме

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

Следователно

$$aX + b, \qquad |X|, \qquad X^k,$$

са сл.в., за всички $a, b \in \mathbb{R}$ и $k \in \mathbb{N}$.

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

$$\mathbb{E}[X] := \sum_{x\in\mathbb{X}} x\, p(x)$$

*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 (*виж втората забележка*)*

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

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

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

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

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, ако X има краен втори момент,

$$\mathbb{E}[X] = \operatorname{argmin}_{a\in\mathbb{R}} \mathbb{E}\big[(X-a)^2\big];$$

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

Забележка (Смяна на променливите/Law of the unconscious statistician). Нека $Y = f(X)$ приема стойности в $\mathbb{Y} \subseteq \mathbb{R}$, за произволна функция $f : \mathbb{R}^m \to \mathbb{R}$. Тогава

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

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 всички* $a, b \in \mathbb{R}$.

базов случай: $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$, *за всяко* $A \in \mathcal{A}$.

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

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

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

Твърдение 3.2 (Свойства на дисперсията). *Нека X е дискретна сл.в., такава че* $\mathrm{Var}(X) < \infty$. *Тогава*

(а) $\mathrm{Var}(X) \geq 0$, *откъдето* $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$;

(б) $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$, *за всички* $a, b \in \mathbb{R}$.

Д-во. Свойство *(а)* следва от свойствата на очакването. От друга страна, свойство *(б)* следва от фактът, че

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

□

Забележка (Константна сл.в.). От позитивността на очакването следва, че

$\mathrm{Var}(X) = 0 \iff \mathbb{P}\big((X-\mathbb{E}[X])^2 = 0\big) = 1 \iff \mathbb{P}\big(X = \mathbb{E}[X]\big) = 1 \iff X$ е (п.с.) константна сл.в.

Твърдение 3.3 (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. $\mathbb{E}[X] = \sum_{x\in\mathbb{X}} x\,p(x) \geq \sum_{x\in\mathbb{X}:x>\alpha} x\,p(x) \geq \sum_{x\in\mathbb{X}:x>\alpha} \alpha\,p(x) = \alpha\cdot\mathbb{P}(X>\alpha)$. □

Твърдение 3.4 (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.* От неравенството на Марков, $\mathbb{P}(|X-\mathbb{E}[X]| > \alpha) = \mathbb{P}\big((X-\mathbb{E}[X])^2 > \alpha^2\big) \leq \frac{\mathbb{E}[(X-\mathbb{E}[X])^2]}{\alpha^2} = \frac{\mathrm{Var}(X)}{\alpha^2}$. □

Задачи

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?

Задача 3.2. От числата от 1 до 10 по случаен начин се избират три числа без връщане. Нека X е сл.в. "средното по големина число от избраните три". Да се намери разпределението на X и да се пресметнат вероятностите на събитията $\{X \geq 7\}$, $\{3 \leq X < 7\}$ и $\{|X-2| < 2\}$. Да се намери разпределението на $Y = (X-5)^2$.

*Задача** 3.3. Подводница стреля последователно n пъти 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. Нека X е сл.в. "брой наводнени отсека". Каква е вероятността корабът да бъде потопен, ако за това е необходимо да бъдат наводнени поне два отсека? А да бъдат наводнени точно $k \leq m$ отсека?

Задача 3.4. Играч залага 5 лв. и има правото да хвърли два зара. Ако хвърли две шестици получава 100 лв., а ако хвърли точно една шестица получава 5 лв. Да се пресметне математическото очакване на печалбата на играча.

Задача 3.5. Нека сл.в. X приема краен брой стойности $x_1, \ldots, x_n$ с една и съща вероятност. Пресметнете $\mathbb{E}[X]$ и $\mathrm{Var}(X)$, ако (а) $x_i = \frac{i-1}{n-1}$, за $i = 1, \ldots, n$; (б) $x_i = a + (b-a)\frac{i-1}{n-1}$, при $b > a$, за $i = 1, \ldots, n$.

Задача 3.6 (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?

Задача 3.7 (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? А ако разполагахме с общо 127 лв. за залагания и $A = 1$ лв.?

*Задача** 3.8. A и B играят последователно партии, като A печели една партия с вероятност $2/3$. Равни партии не са възможни. Играта продължава докато някой спечели две последователни партии. Нека X е сл.в. "брой изиграни партии". Да се определи разпределението, очакването и дисперсията на X.

3.2 Колекция от случайни величини

Съвместно разпределение. *Нека X и Y са две дискретни сл.в., дефинирани върху едно и също в.п. и приемащи стойности съответно в $\mathbb{X} = \{x_1, x_2, \ldots\} \subseteq \mathbb{R}^m$ и $\mathbb{Y} = \{y_1, y_2, \ldots\} \subseteq \mathbb{R}^n$. Съвместното разпределение на X и Y се нарича съвкупността от стойности $p(x,y) := \mathbb{P}(X = x, Y = y)$, $x \in \mathbb{X}$, $y \in \mathbb{Y}$, които имат свойствата $p(x,y) \geq 0$, $\sum_{x \in \mathbb{X}, y \in \mathbb{Y}} p(x,y) = 1$,*

$$\sum_{y \in \mathbb{Y}} p(x,y) = \mathbb{P}(X = x) \qquad \text{и} \qquad \sum_{x \in \mathbb{X}} p(x,y) = \mathbb{P}(Y = y),$$

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

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

$$\{X_i = x_i\} = \{X \in \mathbb{X}_1 \times \cdots \times \mathbb{X}_{i-1} \times \{x_i\} \times \mathbb{X}_{i+1} \times \cdots \times \mathbb{X}_m\} = \bigcup_{x \in \mathbb{X} : (x)_i = x_i} \{X = x\} \in \mathcal{A};$$

т.е. $X_1, \dots, X_m$ са реални дискретни сл.в. От това следва, че многомерната сл.в. $X = (X_1, \dots, X_m)$ може да се разглежда еквивалентно като вектор от едномерни случайни величини.

Обратното твърдение също е вярно. Нека $X_1, \dots, X_m$ са дискретни сл.в., приемащи стойности съответно в $\mathbb{X}_1, \dots, \mathbb{X}_m \subseteq \mathbb{R}$. Дефинираме $X(\omega) := (X_1(\omega), \dots, X_m(\omega))$, за $\omega \in \Omega$. Тогава за всяко $x = (x_1, \dots, x_m) \in \mathbb{X}_1 \times \cdots \times \mathbb{X}_m$ е в сила

$$\{X = x\} = \big\{(X_1, \dots, X_m) = (x_1, \dots, x_m)\big\} = \{X_1 = x_1, \dots, X_m = x_m\} = \bigcap_{i=1}^{m} \{X_i = x_i\} \in \mathcal{A};$$

т.е. X е многомерна дискретна сл. в., примаща стойности в $\mathbb{X} := \mathbb{X}_1 \times \cdots \times \mathbb{X}_m$.

Нека сега X и Y са две дискретни реални сл.в., и $f : \mathbb{R}^2 \to \mathbb{R}$. От забележката след дефиницията на дискретна сл.в. имаме, че $f(X, Y)$ е дискретна сл.в. Следователно

$$X + Y, \qquad XY, \qquad X/Y, \text{ за } Y \neq 0, \qquad \min\{X, Y\}, \qquad \max\{X, Y\},$$

са дискретни сл.в.

Твърдение 3.5.

(а) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;

(б) *ако* $X \geq Y$, *то* $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

*Д-во**.* Относно *(а)*,

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

Относно *(б)*, от $X - Y \geq 0$ и *(а)* следва, че $\mathbb{E}[X] = \mathbb{E}[Y] + \mathbb{E}[X - Y] \geq \mathbb{E}[Y]$. □

Допълнителни понятия

Независимост. *Две дискретни сл.в.* X *и* Y *се наричат независими, ако за всички* $x \in \mathbb{X}$ *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 сл.в. $X_1, \dots, X_n$, *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 е изпълнено*

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i), \qquad \text{за всички } x_1 \in \mathbb{X}_1, \dots, x_n \in \mathbb{X}_n.$$

Сл.в. в една безкрайна редица $X_1, X_2, \dots$ *са независими в съвкупност, ако* $X_{i_1}, \dots, X_{i_n}$ *са независими в съвкупност, за кои да е* $n, i_1, \dots, i_n \in \mathbb{N}$, $i_j \neq i_k$, $j \neq k$.

Твърдение 3.6. *Нека* $f : \mathbb{R}^m \to \mathbb{R}$ *и* $g : \mathbb{R}^n \to \mathbb{R}$. *Ако* $X \perp\!\!\!\perp Y$, *то* $f(X) \perp\!\!\!\perp g(Y)$ *и*

$$\mathbb{E}[f(X) g(Y)] = \mathbb{E}[f(X)]\, \mathbb{E}[g(Y)].$$

*Д-во**.* Нека $A \subseteq \mathbb{X}$ и $B \subseteq \mathbb{Y}$, да кажем $A = \{x_{i_1}, x_{i_2}, \dots\}$ и $B = \{y_{j_1}, y_{j_2}, \dots\}$. От $X \perp\!\!\!\perp Y$ получаваме

$$\mathbb{P}(X \in A, Y \in B) = \sum_n \sum_m \mathbb{P}(X = x_{i_n}, Y = y_{j_m}) = \sum_n \sum_m \mathbb{P}(X = x_{i_n}) \mathbb{P}(Y = y_{j_m}) = \mathbb{P}(X \in A)\, \mathbb{P}(Y \in B).$$

Тогава за всички $u \in f(\mathbb{X})$ и $v \in g(\mathbb{Y})$ е валидно

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

т.е. $f(X) \perp\!\!\!\perp g(Y)$. Относно втория резултат,

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

□

Твърдение 3.7. *Ако $X \perp\!\!\!\perp Y$, то* $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Д-во.

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathbb{E}[(X+Y)^2] - \big(\mathbb{E}[X+Y]\big)^2 \\ &= \mathbb{E}[X^2] + 2\,\mathbb{E}[X]\,\mathbb{E}[Y] + \mathbb{E}[Y^2] - \big(\mathbb{E}[X]\big)^2 - 2\,\mathbb{E}[X]\,\mathbb{E}[Y] - \big(\mathbb{E}[Y]\big)^2 = \mathrm{Var}(X) + \mathrm{Var}(Y)\end{aligned}$$

□

Равенство 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 $X \overset{d}{=} Y$. Нещо повече, ако $X \overset{d}{=} Y$, 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 $X \overset{d}{=} Y$, ще наричаме X и Y независимо и еднакво разпределени, накратко i.i.d.

*Забележка** (Мотивация). Нищо не изисква сл.в. X и Y да бъдат дефинирани в едно и също в.п., за да бъдат еднакви по разпределение. Напротив, ако X идва от $(\Omega, \mathcal{A}, \mathbb{P})$, а Y от $(\Omega', \mathcal{A}', \mathbb{P}')$, тогава фактът, че $\mathbb{P}(X = x) = \mathbb{P}'(Y = x)$, за всяко $x \in \mathbb{X}$, означава, че двата модела, които потенциално произтичат от случайни експерименти с различна сложност, могат да бъдат третирани по един и същи начин от вероятностна гледна точка.

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

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

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

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

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 3.8 (Неравенство на Коши-Шварц)**.** *Нека X и Y са реални дискретни 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 работим с центрираните и нормирани сл.в. $\bar{X} := \frac{X-\mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}}$ и $\bar{Y} := \frac{Y-\mathbb{E}[Y]}{\sqrt{\mathrm{Var}(Y)}}$, така че $\mathbb{E}[\bar{X}], \mathbb{E}[\bar{Y}] = 0$, $\mathbb{E}[\bar{X}^2], \mathbb{E}[\bar{Y}^2] = 1$ и $\rho_{X,Y} = \mathbb{E}[\bar{X}\bar{Y}]$. Тогава

$$0 \le \mathbb{E}\big[(\bar{X} \pm \bar{Y})^2\big] = \mathbb{E}[\bar{X}^2] \pm 2\mathbb{E}[\bar{X}\bar{Y}] + \mathbb{E}[\bar{Y}^2] = 2 \pm 2\,\rho_{X,Y},$$

откъдето $|\rho_{X,Y}| \le 1$. Ако $\rho_{X,Y} = 1$ (или $\rho_{X,Y} = -1$), от горното неравенство следва, че $\mathbb{E}[(\bar{X} - \bar{Y})^2] = 0$ (респ. $\mathbb{E}[(\bar{X} + \bar{Y})^2] = 0$), като $\mathbb{P}(\bar{X} = \bar{Y}) = 1$ от забележката след Твърдение 3.2 (респ. $\mathbb{P}(\bar{X} = -\bar{Y}) = 1$). □

Твърдение 3.9. *Нека $a, b, c \in \mathbb{R}$. Ако съответните количества съществуват,*

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

Ако $X \perp\!\!\!\perp Y$, то $\mathrm{Cov}(X,Y) = 0$.

Д-во.

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

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

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

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

Задачи

Задача 3.9. 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 и теглим от празна урна. Да се определи (а) съвместното разпределение на X и Y; (б) $\mathbb{P}(Y > 2|X = 1)$ и $\mathbb{P}(Y = 3|X < 3)$.

Задача 3.10. А хвърля 3 монети, а Б хвърля 2 монети. Печели този, който хвърли повече езита, като взима всичките 5 монети. В случай на равенство, печели Б. Каква е вероятността А да спечели? Ако А е спечелил, каква е вероятността Б да хвърли точно едно ези? Каква е средната печалба на двамата играчи?

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.

Задача 3.12. Монета с вероятност p за ези се хвърля n пъти. Нека E е събитието "пада се ези при първото хвърляне", а F_k е събитието "точно k пъти се пада тура". За кои двойки цели числа (n, k) събитията E и F_k са независими?

Задача 3.13. 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 сл.в. X = "брой на изтеглените черни топки" при извадка без връщане. Опитът се повтаря 1000 пъти. Да се оцени вероятността да бъдат извадени повече от 900 черни топки.

Задача 3.14. Хвърлят се два зара. Да се намери очакването и дисперсията на сумата и произведението от падналите се точки, 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?

Задача 3.15. Известно е, че 75% от произвежданите в даден цех еднотипни изделия са първо качество. Да се оцени вероятността броят на първокачествените изделия от 100 000 произведени в цеха изделия да се различава от математическото очакване на този брой с по-малко от 1000.

Задача 3.16. Дискретните сл.в. X и Y имат следното съвместно разпределение:

$Y\backslash X$	-1	0	1
0	1/10	2/10	0
1	2/10	3/10	2/10

Да се намерят

(а) маргиналните разпределеня на X и Y;

(б) $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\mathrm{Var}(X)$ и $\mathrm{Var}(Y)$;

(в) $\mathrm{Cov}(X,Y)$ и $\rho_{X,Y}$;

(г) $\mathbb{E}[Z]$ и $\mathrm{Var}(Z)$ за $Z = 3X - 2Y$.

Задача 3.17. Нека X е случайно число измежду $\{0, 1, 2\}$ и Y е независима от него сл.в., такава че $\mathbb{P}(Y = 1) = 2/3 = 1 - \mathbb{P}(Y = 0)$. Намерете съвместното разпределение и корелацията на X и $Z := (X + Y)(\text{mod } 3)$.

Задача 3.18. Хвърляме монета два пъти. Нека X и Y са съответно броят на хвърлените езита и тури. Да се намери $\rho_{X,Y}$.

Задача 3.19. Нека X, Y, Z са сл.в. Докажете, че $\mathrm{Cov}(X + Y, Z) = Cov(X, Z) + \mathrm{Cov}(Y, Z)$.

Задача 3.20. Правилен зар се хвърля n пъти. Нека X е сумата от първите $n-1$ хвърляния, а Y – сумата от последните $n - 1$. Намерете $\rho_{X,Y}$.

3.3 Условно очакване

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

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

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

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

*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 $Y = y$. Тогава $\mathbb{E}[X|Y]$ е съвкупността от тези очаквания при различните хипотези за Y. Нещо повече, ако съответният ред е сходящ,

$$\mathbb{E}[X|Y] = \mathrm{argmin}_{g(Y)} \mathbb{E}\big[(X - g(Y))^2\big];$$

т.е. $\mathbb{E}[X|Y]$ е оптималната в средно квадратичен смисъл оценка на X, използвайки само и единствено информацията от Y.

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

$$P_{X|Y=y}(A) := \mathbb{P}(X \in A|Y = y).$$

Пример 3.1. Нека $X = \mathbb{1}_B$ и $Y = \mathbb{1}_A$, за $A, B \in \mathcal{A}$. Ако $0 < \mathbb{P}(A) < 1$, то по дефиниция

$$\mathbb{E}[X|Y] = \mathbb{E}[X|Y=1] \cdot \mathbb{1}_{\{Y=1\}} + \mathbb{E}[X|Y=0] \cdot \mathbb{1}_{\{Y=0\}} = \mathbb{P}(B|A) \cdot \mathbb{1}_A + \mathbb{P}(B|A^c) \cdot \mathbb{1}_{A^c}.$$

От друга страна всяка функция g от Y има вида $g(Y) = a \cdot \mathbb{1}_{\{Y=1\}} + b \cdot \mathbb{1}_{\{Y=0\}}$ за някои $a, b \in \mathbb{R}$. Тогава

$$\mathbb{E}\big[(X - g(Y))^2\big] = \mathbb{E}\big[(\mathbb{1}_B - a \cdot \mathbb{1}_A - b \cdot \mathbb{1}_{A^c})^2\big] = \mathbb{P}(B) + a^2\mathbb{P}(A) + b^2\mathbb{P}(A^c) - 2a\,\mathbb{P}(A \cap B) - 2b\,\mathbb{P}(A^c \cap B),$$

като постига своя минимум за $a = \mathbb{P}(B|A)$ и $b = \mathbb{P}(B|A^c)$.

Пример 3.2. Разполагаме със стандартно зарче (Зарче 1) и едно, запълнено само с шестици (Зарче 2). Избираме случайно едно от зарчетата и го хвърляме, като сл.в. X отчита падналите се точки. Нека индикаторната сл.в. Y показва кой зар сме хвърлили, като $\mathbb{P}(Y = 1) = 0.8 = 1 - \mathbb{P}(Y = 2)$. Тогава $\mathbb{E}[X|Y = 1] = 3.5$ и $\mathbb{E}[X|Y = 2] = 6$, откъдето

$$\begin{array}{c|cc} \mathbb{E}[X|Y] & 3.5 & 6 \\ \hline & 0.8 & 0.2 \end{array}.$$

Твърдение 3.10 (Свойства на условното очаване)**.**

(i) $\mathbb{E}[aX + bZ|Y] = a\mathbb{E}[X|Y] + b\mathbb{E}[Z|Y]$ *за всички* $a, b \in \mathbb{R}$*;*

(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* $\mathbb{E}[X|Y] = X$*;*

(v) *за* $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$*, имаме* $\mathbb{E}[g(X,Y)|Y = y] = \mathbb{E}[g(X,y)|Y = y]$.

*Д-во**. Лесно се вижда, че

$$\mathbb{E}[X|Y] = \sum_{y \in \mathbb{Y}} \mathbb{E}[X|Y = y] \cdot \mathbb{1}_{\{Y=y\}} = \sum_{y \in \mathbb{Y}} \left(\sum_{x \in \mathbb{X}} (x \cdot 1) \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \right) \mathbb{1}_{\{Y=y\}} = \sum_{y \in \mathbb{Y}} \frac{\mathbb{E}[X \cdot \mathbb{1}_{\{Y=y\}}]}{\mathbb{P}(Y = y)} \mathbb{1}_{\{Y=y\}},$$

откъдето

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

$$\mathbb{E}\big[\mathbb{E}[X|Y]\big] = \mathbb{E}\Big[\sum_{y \in \mathbb{Y}} \mathbb{E}[X|Y = y] \cdot \mathbb{1}_{\{Y=y\}}\Big] = \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y)\mathbb{E}[\mathbb{1}_{\{Y=y\}}]$$

$$= \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y)\mathbb{P}(Y = y) = \sum_{x \in \mathbb{X}} x\Big(\sum_{y \in \mathbb{Y}} \mathbb{P}(X = x|Y = y)\mathbb{P}(Y = y)\Big) = \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x) = \mathbb{E}[X];$$

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

$$\mathbb{E}[X|Y] = \sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} x\mathbb{P}(X = x|Y = y) \cdot \mathbb{1}_{\{Y=y\}} = \sum_{y \in \mathbb{Y}} f(y)\Big(\sum_{x \in \mathbb{X}} \mathbb{P}(X = x|Y = y)\Big) \cdot \mathbb{1}_{\{Y=y\}} = \sum_{y \in \mathbb{Y}} f(y)\mathbb{1}_{\{Y=y\}} = f(Y), \quad \text{за } X = f(Y);$$

$$\mathbb{E}[g(X,Y)|Y = y] = \frac{\mathbb{E}[g(X,Y) \cdot \mathbb{1}_{\{Y=y\}}]}{\mathbb{P}(Y = y)} = \frac{\mathbb{E}[g(X,y) \cdot \mathbb{1}_{\{Y=y\}}]}{\mathbb{P}(Y = y)} = \mathbb{E}[g(X,y)|Y = y].$$

□

Задачи

Задача 3.21. 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 намерят

(а) съвместното разпределение на X и Y;

(б) условните разпределения 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 $\mathbb{E}[Y|X]$.

Задача 3.22. Дискретни сл.в. X и Y имат следното съвместно разпределение:

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

Да се определят (а) условните разпределения на X относно Y и на Y относно X; (б) разпределенията на $\mathbb{E}[X|Y]$ и $\mathbb{E}[Y|X]$.

*Задача** 3.23 (Равенство на Валд). Нека $X_1, X_2, \ldots$ са еднакво разпределени сл.в. с крайно очакване, а T е независима от тях сл.в., приемаща стойности в $\{1, 2, \ldots\}$. Покажете, че

$$\mathbb{E}[X_1 + \cdots + X_T] = \mathbb{E}[T] \cdot \mathbb{E}[X_1].$$

Задача 3.24. Разполагаме с две кутии с топки. В първата има 4 бели и 6 черни, а във втората – 7 бели и 3 черни. От всяка урна случайно се изважда по 1 топка. Към извадените две топки се прибавя една бяла и се поставят в трета кутия. Да допуснем, че теглим от третата кутия по 1 топка с връщане, докато изтеглим бяла. Нека X е броят изтеглени топки от третата кутия. Какво е очакването на X?

Задача 3.25. На всяка от страните на зар се записва случайно число от 1 до 6. Каква е вероятността да се падне 6-ца при хвърляне на зара? Колко е очакваният брой хвърляния докато това се случи?

*Задача*** 3.26 (Продължение на Задача 3.20). 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 $\rho_{X,Y}$.

*Задача** 3.27. Трима приятели играят следната игра:

- Всеки играч заплаща по 1 лев такса участие.
- Хвърля се честна монета до първия път, когато се падне тура. За всяко ези казиното изплаща по 1 лев.
- При всяка печалба приятелите избират на случаен принцип кой да вземе този 1 лев.

Какво е разпределението, очакването и дисперсията на печалбата на всеки един от приятелите? (Използвайте формулите $\sum_{n=0}^{\infty} nx^n = \frac{x}{(1-x)^2}$, $\sum_{n=0}^{\infty} n^2 x^n = \frac{x(1+x)}{(1-x)^3}$, $\sum_{n=0}^{\infty} \binom{n+k}{k} x^n = \frac{1}{(1-x)^{k+1}}$, за $|x| < 1$) **Каква е вероятността и тримата да не са на загуба?

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. По-специално, пораждащата функция може да се използва за извеждането на някои свойства на сл.в. X чрез аналитични методи.

Функция на разпределение *на (реална) сл.в. X наричаме функцията*

$$F_X(x) := \mathbb{P}(X \leq x), \qquad \text{за } x \in \mathbb{R}.$$

**Функция на разпределение на вектор от (реални) сл.в. $(X_1, \ldots, X_m)$ наричаме функцията*

$$F_{X_1,\ldots,X_m}(x_1, \ldots, x_m) := \mathbb{P}(X_1 \leq x_1, \ldots, X_m \leq x_m), \qquad \text{за } x_1, \ldots, x_m \in \mathbb{R}.$$

Забележка (Следствия). Нека X е дискретна сл.в., приемаща стойности в $\mathbb{X} = \{x_1, x_2, \ldots\} \subseteq \mathbb{R}$, където $x_1 < x_2 < \cdots$, без загуба на общност. Тогава $\mathbb{P}(X = x_n) = F_X(x_n) - F_X(x_{n-1})$, $n \geq 1$, т.е. F_X определя напълно и еднозначно вероятностното разпределение на X. Следователно за две дискретни сл.в. X и Y,

$$F_X = F_Y \qquad \Longleftrightarrow \qquad X \stackrel{d}{=} Y.$$

По подобен начин $\mathbb{P}_X$ се определя напълно и еднозначно от $\{\mathbb{P}(X > x)\}_{x\in\mathbb{R}}$, $\{\mathbb{P}(X < x)\}_{x\in\mathbb{R}}$, $\{\mathbb{P}(X \geq x)\}_{x\in\mathbb{R}}$.

Забележка (Свойства). Функцията $F_X : \mathbb{R} \to [0,1]$ има следните свойства:

(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 отдясно, т.е. $\lim_{y \downarrow x} F_X(y) = F_X(x)$, за всяко $x \in \mathbb{R}$;

(iii) $\lim_{x\to\infty} F_X(x) = 1$ и $\lim_{x\to-\infty} F_X(x) = 0$.

**Всъщност всяка такава функция $F : \mathbb{R} \to [0,1]$, удовлетворяваща (i)-(iii), еднозначно определя една вероятностна мярка P върху $\mathbb{R}$, такава че $F(x) = P((-\infty, x])$, за всяко $x \in \mathbb{R}$.

Твърдение 3.11. *Ако X е сл.в., примаща стойности в $\{0,1,\ldots\}$, тогава*

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

Д-во.

$$\mathbb{E}[X] = \sum_{m=1}^{\infty} m\,\mathbb{P}(X = m) = \sum_{m=1}^{\infty}\sum_{n=1}^{m} \mathbb{P}(X = m) = \sum_{n=1}^{\infty}\sum_{m=n}^{\infty} \mathbb{P}(X = m) = \sum_{n=1}^{\infty} \mathbb{P}(X \ge n) = \sum_{n=0}^{\infty} \mathbb{P}(X > n).$$

Алтернативно, от Твърдение 3.13, $1-(1-s)\sum_{n=0}^{\infty} s^n \mathbb{P}(X > n) = g_X(s)$, откъдето $g_X'(1) = \sum_{n=0}^{\infty} \mathbb{P}(X > n)$. $\square$

Пораждаща функция *на дискретна сл.в. X, приемаща стойности в $\{0,1,\ldots\}$, наричаме функцията*

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

**Пораждаща функция на вектор от сл.в. $(X_1,\ldots,X_m)$, приемащи стойности в $\{0,1,\ldots\}$, наричаме функцията*

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

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

$$\begin{aligned} g_{X_1,X_2}(s_1,s_2) &= \sum_{k_1,k_2} s_1^{k_1} s_2^{k_2} \mathbb{P}(X_1 = k_1, X_2 = k_2) \\ &= \sum_{k_1,\ldots,k_m} s_1^{k_1} s_2^{k_2} \mathbb{P}(X_1 = k_1,\ldots,X_m = k_m) = g_{X_1,\ldots,X_m}(s_1,s_2,1,\ldots,1). \end{aligned}$$

Твърдение 3.12. *Нека X, X_1 и X_2 са дискретни сл.в., примащи стойности в $\{0,1,\ldots\}$. Тогава*

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

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 твърдение следва, че пораждащата функция на X определя напълно и еднозначно вероятностното разпределение на X. Следователно за две целочислени сл.в. X и Y,

$$g_X = g_Y \qquad \Longleftrightarrow \qquad X \overset{d}{=} Y.$$

Твърдение 3.13.** *Ако X е сл.в., примаща стойности в $\{0,1,\ldots\}$, тогава*

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

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

$$\begin{aligned}\sum_{n=0}^{\infty} s^n \mathbb{P}(X \ge n) &= \sum_{n=0}^{\infty}\sum_{m=n}^{\infty} s^n \mathbb{P}(X = m) = \sum_{m=0}^{\infty}\left(\sum_{n=0}^{m} s^n\right)\mathbb{P}(X=m) = \sum_{m=0}^{\infty}\frac{1-s^{m+1}}{1-s}\mathbb{P}(X=m)\\ &= \frac{1}{1-s}\sum_{m=0}^{\infty}\mathbb{P}(X=m) - \frac{s}{1-s}\sum_{m=0}^{\infty} s^m\mathbb{P}(X=m) = \frac{1}{1-s} - \frac{s}{1-s}g_X(s) = \frac{1-s\,g_X(s)}{1-s},\end{aligned}$$

откъдето $\sum_{n=0}^{\infty} s^n F_X(n) = \sum_{n=0}^{\infty} s^n\big(1-\mathbb{P}(X\ge n)+\mathbb{P}(X=n)\big) = \frac{1}{1-s} - \frac{1-s\,g_X(s)}{1-s} + g_X(s) = \frac{g_X(s)}{1-s}$. ☐

Задачи

Задача 3.28. Нека $X_1, \dots, X_n$ са i.i.d. сл.в. Дефинираме $X_{\max} := \max\{X_1, \dots, X_n\}$ и $X_{\min} := \min\{X_1, \dots, X_n\}$. Изразете функциите на разпределение на $X_{\max}$, $X_{\min}$ и $(X_{\max}, X_1)$ през тази на X_1.

Задача 3.29. Нека X е сл.в. с пораждаща ф-ия $g_X(s) = \frac{s}{4-3s}$. Намерете $\mathbb{E}[X]$, $\mathrm{Var}(X)$, $\mathbb{P}(X=1)$, $\mathbb{P}(X=2)$.

Задача 3.30. Върху инструмент са монтирани три лампи. Вероятностите за изгаряне на първа, втора и трета лампа в даден ден са равни съответно на 0.1, 0.2 и 0.3. Вероятностите за излизане на инструмента от строя при изгарянето на една, две и три лампи са равни съответно на 0.25, 0.6 и 0.9. Намерете очаквания брой изгорели лампи. Пресметнете вероятността инструментът да излезе от строя.

3.5 Основни дискретни разпределения

Разпределение на Бернули

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 т. нар. Бернулиев опит, който може да се окачестви единствено като "успех" с вероятност p и "неуспех" с вероятност $(1-p)$.

Твърдение 3.14. *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

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

Пример 3.3. Нечестна монета с вероятност за ези $p \in [0,1]$ се хвърля n пъти. Тогава броят на падналите се езита има биномно разпределение.

Твърдение 3.15. *Ако* $X \sim \mathrm{Bin}(n,p)$*, тогава* $\mathbb{E}[X] = np$, $\mathrm{Var}(X) = np(1-p)$ *и* $g_X(s) = (ps + (1-p))^n$.

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 $X_1, \dots, X_n \overset{i.i.d.}{\sim} \mathrm{Ber}(p)$. Тогава, за всяко $k = 0, 1, \dots, n$,

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

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

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

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

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 $\mathbb{P}_X$ достига своя максимум в $k^* = \lfloor p(n+1) \rfloor$. Ако числото $p(n+1)$ е цяло, то $\mathbb{P}(X=k^*) = \mathbb{P}(X=k^*-1)$. С други думи, най-вероятните биномни стойности са групирани около математическото очакване на X.

Геометрично разпределение

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

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

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). За всяко $k = 0, 1, 2, \dots$ имаме

$$\mathbb{P}(X \ge k) = \sum_{i=k}^\infty (1-p)^i p = \sum_{j=0}^\infty (1-p)^{j+k} p = (1-p)^k \frac{p}{1-(1-p)} = (1-p)^k,$$

откъдето, 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; т.е. от вероятностна гледна точка експериментът започва отначало.

Всъщност геометричното разпределение е единственото дискретно разпределение, което има това свойство. Нека $f(k) := \mathbb{P}(X \geq k)$, за $k \in \mathbb{N}$. Лесно се вижда, че функционалното уравнение $f(m+k) = f(m)f(k)$ има решение $f(k) = f(1)^k$, откъдето $\mathbb{P}(X = k) = f(k) - f(k+1) = (1 - f(1))f(1)^k$. т.е. $X \sim \mathrm{Geo}(f(1))$

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

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

т.е. $\min\{n \in \mathbb{N} : \sum_{i=1}^{n} X_i = 1\} - 1 \sim \mathrm{Geo}(p)$.

Твърдение 3.16. *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.

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k(1-p)^k p = \sum_{k=1}^{\infty}(k-1)(1-p)^k p + \sum_{k=1}^{\infty}(1-p)^k p = (1-p)(\mathbb{E}[X]+1) \quad \Longrightarrow \quad \mathbb{E}[X] = \frac{1-p}{p},\\
\mathbb{E}[X] &= \sum_{k=0}^{\infty}\mathbb{P}(X > k) = \sum_{k=0}^{\infty}(1-p)^{k+1} = (1-p)\sum_{k=0}^{\infty}(1-p)^k = \frac{1-p}{1-(1-p)} = \frac{1-p}{p},\\
\mathbb{E}[X^2] &= \sum_{k=1}^{\infty} k^2(1-p)^k p = \sum_{k=1}^{\infty}(k-1+1)^2(1-p)^k p = \sum_{k=1}^{\infty}(k-1)^2(1-p)^k p + 2\sum_{k=1}^{\infty}(k-1)(1-p)^k p + \sum_{k=1}^{\infty}(1-p)^k p\\
&= (1-p)\big(\mathbb{E}[X^2] + 2\mathbb{E}[X] + 1\big) \quad \Longrightarrow \quad \mathbb{E}[X^2] = \frac{(1-p)(2-p)}{p^2}.
\end{aligned}$$

□

Отрицателно биномно разпределение

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* $0, 1, 2, \ldots$ *с вероятности*

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

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*** (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 3.17. *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 $\sum_{k=0}^{\infty}\binom{n+k-1}{k}p^n(1-p)^k = p^n\sum_{k=0}^{\infty}\binom{n+k-1}{k}(1-p)^k = p^n(1-(1-p))^{-n} = p^np^{-n} = 1$, 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, използвайки, че $k\binom{n+k-1}{k} = n\binom{n+k-1}{k-1}$,

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

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

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

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

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

Алт. д-во. Човек може да заключи от формата на $g_X(s) = \left(\frac{p}{1-s(1-p)}\right)^n$, че $X \overset{d}{=} \sum_{i=1}^n X_i$, където $X_1,\dots,X_n \overset{i.i.d.}{\sim} \mathrm{Geo}(p)$ понеже $g_{\sum_{i=1}^n X_i}(s) = (g_{X_1}(s))^n = \left(\frac{p}{1-s(1-p)}\right)^n$. Алтернативно, за всяко $k = 0,1,2,\dots$,

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

$$= \sum_{j_1,\dots,j_n\in\{0,\dots,k\}:j_1+\dots+j_n=k}(1-p)^{j_1}p\cdots(1-p)^{j_n}p = \binom{n+k-1}{k}p^n(1-p)^k.$$

Тогава $\mathbb{E}[X] = n\frac{1-p}{p}$ и $\mathrm{Var}(X) = n\frac{1-p}{p^2}$. □

*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 $\mathbb{P}_X$ достига своя максимум в $k^* = \lfloor\frac{(n-1)(1-p)}{p}\rfloor$. Ако числото $\frac{(n-1)(1-p)}{p}$ е цяло, то $\mathbb{P}(X=k^*) = \mathbb{P}(X=k^*-1)$.

Хипергеометрично разпределение

Хипергеометрично разпределение *с параметри $M,N,n\in\mathbb{N}$, такива че $M<N$ и $n<N$, е дискретно разпределение на сл. в. X, приемаща стойности в $\{0,1,\dots,n\}$ с вероятности*

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

и нула иначе. 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.

Пример 3.4 (Пример 2.1)*.* 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 и без връщане n топки ($n<N$). Тогава броят на извадените черни топки има хипергеометрично разпределение.

Твърдение 3.18. *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 $\sum_{k=\max\{0,M-N+n\}}^{\min(M,n)} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{\sum_{k=\max\{0,M-N+n\}}^{\min(M,n)}\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1$, използвайки резултатите от Задача 1.6; т.е. хипергеометричното разпределение е добре дефинирано. По същия начин

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_k k\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \sum_k M\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\frac{N}{n}\binom{N-1}{n-1}} = \frac{nM}{N}\sum_{j=\max\{0,M-N+n-1\}}^{\min(M-1,n-1)}\frac{\binom{M-1}{j}\binom{N-M}{(n-1)-j}}{\binom{N-1}{n-1}} = \frac{nM}{N}\frac{\binom{N-1}{n-1}}{\binom{N-1}{n-1}} = \frac{nM}{N},\\
\mathbb{E}[X^2] &= \sum_k k^2\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{nM}{N}\sum_k k\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N-1}{n-1}} = \frac{nM}{N}\sum_k (k-1)\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N-1}{n-1}} + \frac{nM}{N}\sum_k\frac{\binom{M-1}{k-1}\binom{N-M}{n-k}}{\binom{N-1}{n-1}}\\
&= \frac{nM}{N}\sum_k (M-1)\frac{\binom{M-2}{k-2}\binom{N-M}{n-k}}{\frac{N-1}{n-1}\binom{N-2}{n-2}} + \frac{nM}{N} = \frac{nM(M-1)(n-1)}{N(N-1)} + \frac{nM}{N} = \frac{nM(Mn+N-M-n)}{N(N-1)},\\
\mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{nM(Mn+N-M-n)}{N(N-1)} - \frac{n^2M^2}{N^2} = \frac{nM(N-M)(N-n)}{N^2(N-1)}.
\end{aligned}
$$

□

*Забележка*** (Връзка с биномното разпределение). За разлика от хипергеометричното разпределение, биномното разпределение описва 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 коефициент $\frac{N-n}{N-1}$. 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 $\lim_{N\to\infty}\frac{M}{N} = p > 0$, то

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

Разпределение на Поасон

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

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

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

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.

Пример 3.5. Център за обслужване на клиенти получава средно 180 обаждания на час, 24 часа в денонощието (т.е. без спадове и пикове в натоварването). Обажданията са независими; получаването на едно от тях не променя вероятността за това кога ще пристигне следващото. Тогава броят на обажданията, получени през която и да е минута, има Поасоново разпределение с параметър 3.

Твърдение 3.19. *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 $x_0 = 0$; т.е. разпределението на Поасон е добре дефинирано. По същия начин

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

□

*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 $\mathbb{P}(X = k) \le \mathbb{P}(X = k-1)$, за $k \ge \lambda$, като $\mathbb{P}_X$ достига своя максимум в $k^* = \lfloor \lambda \rfloor$. Ако параметърът λ е цяло число, то $\mathbb{P}(X = k^*) = \mathbb{P}(X = k^* - 1)$.

*Забележка*** (Теорема на Поасон). Нека $X \sim \mathrm{Po}(\lambda)$ и $X_n \sim \mathrm{Bin}(n, p_n)$, като $np_n \to \lambda > 0$ при $n \to \infty$, т.е. с увеличаване на броя на опитите вероятността за успех намалява пропорционално. Тогава, за $k = 0, 1, 2, \ldots$,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(X_n = k) &= \lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k} \\ &= \lim_{n\to\infty} \frac{(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})}{k!} n^k p_n^k (1-p_n)^{n-k} = \frac{e^{-\lambda}\lambda^k}{k!} = \mathbb{P}(X = k),\end{aligned}$$

където използваме, че $\lim_{n\to\infty}(1-p_n)^{n-k} = \lim_{n\to\infty}(1-\frac{\lambda}{n})^{n-k} = e^{-\lambda}$.

Горното приближение позволява пресмятането на биномни вероятности, което в противен случай би могло да доведе до значителни технически трудности. На практика, ако $Y \sim \mathrm{Bin}(n, p)$, където $np \le 20$ и $n \ge 100$, тогава $\mathbb{P}(Y = k) \approx e^{-np}(np)^k/k!$ осигурява достатъчно добро приближение (виж Задача 3.47).

Полиномно разпределение

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* $(k_1, \ldots, k_r) \in \mathbb{N}_0^r$, *такива че* $k_1 + \cdots + k_r = n$, *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 $X^{(i)} \sim \mathrm{Multi}(1; p_1, \ldots, p_r)$, за $i = 1, \ldots, n$, като $X^{(1)}, \ldots, X^{(n)}$ са независими в съвкупност. Тогава

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

където $e_1, \ldots, e_n$ са вектори в $\mathbb{R}^r$, чиито компоненти са 0, с изключение на един, който е равен на 1; т.е. $X \stackrel{d}{=} X^{(1)} + \cdots + X^{(n)}$.

Пример 3.6. Урна съдържа l бели, m зелени и n червени топки. Ако вадим топките една по една с връщане, каква е вероятността да извадим l_1 бели, m_1 зелени и n_1 червени топки от $l_1 + m_1 + n_1$ тегления?

Твърдение 3.20. *Ако* $X \sim \mathrm{Multi}(n; p_1, \ldots, p_r)$, *тогава* $g_X(s_1, \ldots, s_r) = (p_1 s_1 + \cdots + p_r s_r)^n$.

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

Задачи

Задача 3.31. Два зара се хвърлят последователно пет пъти. Каква е вероятността броят на хвърлянията, при които сумата от точките е шест, да бъде точно 2? Да се намери средната стойност на този брой.

Задача 3.32. 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 намери разпределението на $X = X_1 + X_2$.

Задача 3.33. 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 $X = 120$?

Задача 3.34. Нека $N \sim \mathrm{Bin}(M, q)$ и $X|N = n \sim \mathrm{Bin}(n, p)$, $n = 0, 1, \ldots, M$, за $p, q \in (0, 1)$ и $M \in \mathbb{N}$. Да се намери разпределението на X.

*Задача*** 3.35. Нека $X \sim \mathrm{Bin}(n, p)$. Каква е вероятността X да бъде четно число?

Задача 3.36. 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) $\max\{X, Y\}$.

Задача 3.37. 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 очакваният брой изстрели, необходими за уцелване на мишената? *А колко средно изстрела биха направили двамата, ако имаха право на максимум n опита?

Задача 3.38. Хвърлят се 5 стандартни зара едновременно. Ако на някой/някои от заровете се падне шестица, той/те се отстранява, а останалите се хвърлят отново, докато не останат зарове. Какъв е очакваният брой хвърляния?

Задача 3.39 (Coupon Collector's Problem). Всеки пакет зрънчо съдържа една от n различни вида играчки. Намерете очаквания брой пакети, докато бъдат събрани всички n играчки. *А колко е очакваният брой пакети, докато ни се падне повтаряща се играчка?

Задача 3.40 (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 $k \leq n$ клечки?

Задача 3.41. В кутия има 7 лампи, от които 3 са дефектни. По случаен начин се избират за проверка 4 лампи. Да се намери разпределението на сл. в. $X =$ "брой на изпробваните дефектни лампи" и да се пресметне нейното очакване и дисперсия.

Задача 3.42. 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?

Задача 3.43. От урна, съдържаща M черни и $N-M$ бели топки, последователно теглим n топки без връщане. Нека индикаторната сл. в. X_i отчита дали сме изтеглили черна топка на i-ти ход, за $i = 1, \ldots, n$. Намерете $\mathrm{Corr}(X_i, X_j)$, разпределението, очакването и дисперсията на $X_1 + \cdots + X_n$.

Задача 3.44. Нека $X_1 \sim \mathrm{Po}(\lambda_1)$ и $X_2 \sim \mathrm{Po}(\lambda_2)$ са независими случайни величини. Да се намери разпределението на $X = X_1 + X_2$.

Задача 3.45. В Патагония на месец се регистрират средно две слаби земетресения. Каква е вероятността за три месеца да има по-малко от четири слаби земетресения?

Задача 3.46 (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.

Задача 3.47. 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?

Задача 3.48. 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.

Задача 3.49. Урна съдържа три еднакви топки: бяла, зелена и червена. Изваждаме с връщане пет пъти по една топка. Каква е вероятността всяка от бялата и червената топка да бъде извадена поне два пъти?

*Задача*** 3.50. Нека $(X_1, \ldots, X_r) \sim \mathrm{Multi}(n; p_1, \ldots, p_r)$. Да се покаже, че $\rho_{X_i, X_j} = -\sqrt{\frac{p_i p_j}{(1-p_i)(1-p_j)}}$, $i \neq j$.

*Задача** 3.51. На колко е равна вероятността при хвърлянето на n зара сумата от падналите точки да бъде равна на дадено число k $(n \leq k \leq 6n)$?

*Задача** 3.52. Урна съдържа 2 бели, 2 червени и 6 зелени топки. Изваждат се една по една с връщане 20 топки. Каква е вероятността да изтеглим с 4 повече бели топки отколкото червени?

*Задача** 3.53. 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?

4 Непрекъснати случайни величини

4.1 Основни понятия

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

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

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

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

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

като множеството

$$D_X := \{x \in \mathbb{R} : f_X(x) > 0\}$$

се нарича носител на f_X. В този случай функцията на разпределение на X има вида

$$F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(t)dt, \qquad \text{за } x \in \mathbb{R}.$$

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

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

откъдето $\mathbb{P}(X = a) = 0$ и

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

За произволни $B \in \mathcal{B}(\mathbb{R})$ вероятността $\mathbb{P}(X \in B) = \int_B f_X(x)dx$ е добре дефинирана, където интегрирането е в смисъла на Лебег.

Условие *(ii)* пък влече

(а) $f_X(x) \geq 0$ за почти всички (п. вс.) x (спрямо мярката на Лебег);

(б) $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Всъщност (а) и (б) са необходими и достатъчни условия за задаването на непрекъсната сл. в. Действително, ако f е (интегруема) функция, удовлетворяваща (а) и (б), то $F(x) := \int_{-\infty}^x f(t)dt$, $x \in \mathbb{R}$, е (i) ненамаляваща; (ii) непрекъсната отдясно; (iii) $\lim_{x\to-\infty} F(x) = 0$ и $\lim_{x\to\infty} F(x) = 1$; т.е. F 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, в смисъл, че

$$X \overset{d}{=} Y \qquad \Longleftrightarrow \qquad f_X = f_Y \quad \text{(почти навсякъде).}$$

*Забележка*** (Анализ). Ако f_X е непрекъсната функция, то F_X е диференцируема и $f_X = F_X'$ по Нютон-Лайбниц. Обратно, ако ни е дадена функция на разпределение F_X, която има непрекъсната производна, то X е непрекъсната сл. в. с плътност F_X'.

В общия случай X е непрекъсната сл. в. тогава и само тогава, когато функцията ѝ на разпределение F_X е **абсолютно непрекъсната:** за всички $a, b \in \mathbb{R} : a < b$ и $\epsilon > 0$, съществува $\delta > 0$, такова че за всички непресичащи се интервали $(a_1, b_1), \ldots, (a_n, b_n) \subseteq [a, b]$,

$$\sum_{k=1}^n (b_k - a_k) < \delta \qquad \Longrightarrow \qquad \sum_{k=1}^n \big(F_X(b_k) - F_X(a_k)\big) < \epsilon.$$

Тогава $F_X'(x)$ съществува за п. вс. $x \in \mathbb{R}$ и $F_X(x) = \int_{-\infty}^x F_X'(t)dt$, за всяко $x \in \mathbb{X}$ (*Фундаментална теорема за Лебеговия интеграл*), т.е. X е непрекъсната сл. в. с плътност

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

За да бъде F_X абсолютно непрекъсната, е *необходимо* (но не непременно достатъчно) F_X да е непрекъсната и диференцируема извън изброимо на брой точки.

*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 $a, b \in \mathbb{R} : a < b$,

$$\{a < g(X) \leq b\} = \left\{X \in g^{-1}((a, b])\right\} \in \mathcal{A},$$

което е изпълнено, ако $g^{-1}((a, b]) \in \mathcal{B}(\mathbb{R})$. Такива функции g се наричат **Борелови** или най-общо измерими. В частност, всички непрекъснати и/или монотонни функции са измерими, откъдето

$$aX + b, \qquad |X|, \qquad X^k,$$

са сл. в., за всички $a, b \in \mathbb{R}$ и $k \in \mathbb{N}$.

Обърнете внимание, че ако X е освен това непрекъсната сл. в., то не е задължително $g(X)$ да бъде също непрекъсната сл. в.

Контрапример. Нека X е непрекъсната сл. в. Да дефинираме $g(x) := 1$ при $x \geq 0$, и $g(x) := 0$ при $x < 0$. Тогава g е измерима функция и $g(X) \sim \mathrm{Ber}(\mathbb{P}(X \geq 0))$, т.е. $g(X)$ е дискретна сл. в.

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

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

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

$$\begin{aligned}\mathbb{P}(a < g(X) \leq b) = \mathbb{P}\big(h(a) < X \leq h(b)\big) &= \int_{h(a)}^{h(b)} f_X(x)dx \\ = \int_a^b f_X(h(y)) \cdot \mathbb{1}_{D_X}(h(y))dh(y) &= \int_a^b f_X(h(y))h'(y) \cdot \mathbb{1}_{g(D_X)}(y)dy.\end{aligned}$$

Ако g е строго намаляваща, $\mathbb{P}(a < g(X) \leq b) = \mathbb{P}(h(b) < X \leq h(a)) = \cdots = -\int_a^b f_X(h(y))h'(y)\mathbb{1}_{g(D_X)}(y)dy$. Понеже в този случай h е също строго намаляваща, то $h'(y) < 0$, откъдето следва общото твърдение. $\square$

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

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

се нарича дисперсия на X.

Забележка (Смяна на променливите). Нека $Y = g(X)$, където $g : \mathbb{R} \to \mathbb{R}$ е измерима функция. Ако интегралът е абсолютно сходящ, тогава, независимо дали Y е непрекъсната сл. в. или не, е изпълнено

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

В случай, че g е строго растяща диференцируема функция, то Y е непрекъсната сл. в. от Твърдение 4.1 и има плътност f_Y, откъдето

$$\begin{aligned}\mathbb{E}[g(X)] = \mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy &= \int_{-\infty}^{\infty} y f_X(g^{-1}(y))(g^{-1}(y))'dy \\ = \int_{-\infty}^{\infty} g(u) f_X\big(g^{-1}(g(u))\big)(g^{-1}(g(u)))'g'(u)du &= \int_{-\infty}^{\infty} g(u) f_X(u)du,\end{aligned}$$

като използваме смяната $y = g(u)$ и фактът, че $(g^{-1}(g(u)))' = \frac{1}{g'(g^{-1}(g(u)))} = \frac{1}{g'(u)}$.

Твърдение 4.2.

позитивност: *ако* $X \geq 0$, *то* $\mathbb{E}[X] \geq 0$, *откъдето* $\mathrm{Var}(X) \geq 0$ *за произволни* X;

линейност: $\mathbb{E}[aX + b] = a, \mathbb{E}[X] + b$ *и* $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$, *за всички* $a, b \in \mathbb{R}$.

Д-во. От свойствата на интегралите. □

Задачи

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

Задача 4.2. Нека X е непрекъсната сл. в. с четна плътност, т.е. $f_X(-x) = f_X(x)$, $x \in \mathbb{R}$. Да се намери $F_X(0)$ и да се представят $F_X(-x)$ и $\mathbb{P}(|X| \leq x)$ чрез $F_X(x)$, за всяко $x \in \mathbb{R}$.

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

*Задача** 4.4. Нека X е непрекъсната сл. в. с плътност f_X. Да се намери разпределението на сл. в. X^2.

Задача 4.5. 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) $\mathbb{P}(X \leq \mathbb{E}[X])$;

(г) $\mathbb{E}[X^2 + 3X]$.

4.2 Колекция от случайни величини

Непрекъснат вектор от случайни величини. *Нека X и Y са две сл. в., дефинирани в едно и също вероятностно пространство. Тогава (X, Y) е непрекъснат вектор от сл. в., ако съществува 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 $y_1, y_2 \in \mathbb{R} : y_1 < y_2$,*

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

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

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

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

$$F_{X,Y}(x, y) := \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t, s) ds dt, \qquad \text{за } x, y \in \mathbb{R},$$

се нарича функция на разпределение на вектора (X, Y). Горните дефиниции се обобщават за вектор от повече от две сл. в. по очевиден начин.

Забележка (Свойства). Подобно на едномерния случай, функцията $f_{X,Y}$ удовлетворява характеризиращите я свойства:

(а) $f_{X,Y}(x, y) \geq 0$ за почти всички (п. вс.) $x, y \in \mathbb{R}$ (спрямо мярката на Лебег върху $\mathbb{R}^2$);

(б) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy dx = 1$.

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

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

използвайки многомерна смяна на променливите. В частност,

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

**От друга страна, ако двойният интеграл съществува, то теоремата на Фубини позволява редът на интегриране да бъде сменен, като

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

Твърдение 4.3.

(а) $\mathbb{E}[aX+bY+c] = a\mathbb{E}[X]+b\mathbb{E}[Y]+c$, *за всички* $a,b,c\in\mathbb{R}$;

(б) *ако* $X \geq Y$, *то* $\mathbb{E}[X]\geq\mathbb{E}[Y]$;

(в) $\mathrm{Var}(aX\pm bY+c) = a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y)\pm 2ab\,\mathrm{Cov}(X,Y)$, *за всички* $a,b,c\in\mathbb{R}$.

Д-во.* От свойствата на интегралите. □

Допълнителни понятия

Независимост. *Две непрекъснати сл. в.* X *и* Y *се наричат независими, обозначено* $X \perp\!\!\!\perp Y$, *ако за всички* $x,y\in\mathbb{R}$ *е изпълнено равенството*

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

Може да се покаже, че $X \perp\!\!\!\perp Y$ *тогава и само тогава, когато* $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ *за п. вс.* $x,y\in\mathbb{R}$.

Група от непрекъснати сл. в. $X_1,\dots,X_n$ *се наричат независими в съвкупност, ако за произволни* $x_1,\dots,x_n\in\mathbb{R}$ *е изпълнено*

$$F_{X_1,\dots,X_n}(x_1,\dots,x_n) = \prod_{i=1}^n F_{X_i}(x_i).$$

Тогава $(X_1,\dots,X_n)$ *образува непрекъснат вектор от сл. в. с плътност*

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n).$$

Обратното твърдение също е вярно. Именно, ако $f_{X_1,\dots,X_n}(x_1,\dots,x_n) = \prod_{i=1}^n f_{X_i}(x_i)$, *за п. вс.* $x_1,\dots x_n\in\mathbb{R}$, *то* $X_1,\dots,X_n$ *са независими в съвкупност.*

Твърдение 4.4. *Нека* X *и* Y *са непрекъснати сл. в., такива че* $X \perp\!\!\!\perp Y$. *Тогава*

(i) $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$;

(ii) $\mathrm{Var}(X+Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)$ *и* $\mathrm{Cov}(X,Y)=0$;

(iii) $f(X) \perp\!\!\!\perp g(Y)$ *за измерими функции* $f,g:\mathbb{R}\to\mathbb{R}$.

Д-во.* Относно *(i)*,

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dydx = \int_{-\infty}^{\infty} x\Big(\int_{-\infty}^{\infty} yf_Y(y)dy\Big)f_X(x)dx = \mathbb{E}[X]\mathbb{E}[Y].$$

Твърдение *(ii)* следва от $\mathrm{Var}(X+Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)+2\,\mathrm{Cov}(X,Y)$ и фактът, че $\mathrm{Cov}(X,Y) = \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y] = 0$.

Относно *(iii)*, от $X \perp\!\!\!\perp Y$ имаме за всички $A,B\in\mathcal{B}(\mathbb{R})$, че

$$\mathbb{P}(X\in A, Y\in B) = \int_A\int_B f_{X,Y}(x,y)dxdy = \Big(\int_A f_X(x)dx\Big)\Big(\int_B f_Y(y)dy\Big) = \mathbb{P}(X\in A)\,\mathbb{P}(Y\in B).$$

Тогава за всички $C, D \in \mathcal{B}(\mathbb{R})$ получаваме

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

$\square$

Равенство по разпределение. *Две непрекъснати сл. в. X и Y се наричат еднакво разпределени, обозначено $X \stackrel{d}{=} Y$, ако $F_X(x) = F_Y(x)$ за всички $x \in \mathbb{R}$, което е вярно тогава и само тогава, когато за п. вс. $x \in \mathbb{R}$ е изпълнено*

$$f_X(x) = f_Y(x).$$

В този случай, за измерими $g : \mathbb{R} \to \mathbb{R}$, ако съответният интеграл е съществува,

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

Твърдение 4.5 (Многомерна смяна на променливи). *Нека (X, Y) е непрекъснат вектор от сл. в. с плътност $f_{X,Y}$, и функцията $g : D_{X,Y} \to \mathbb{R}^2$ е биективна с обратна функция $h = g^{-1}$. Ако h и g са непрекъснати, h е диференцируема, а якобианът $J(u, v) := \det\big[\frac{\partial}{\partial \boldsymbol{z}} h(\boldsymbol{z})\big|_{\boldsymbol{z}=(u,v)}\big] \neq 0$ за всички $(u, v) \in g(D_{X,Y})$, то $g(X, Y)$ е непрекъснат вектор от сл. в. с плътност*

$$f_{g(X,Y)}(u, v) = f_{X,Y}\big(h(u, v)\big) \cdot |J(u, v)| \cdot \mathbb{1}_{g(D_{X,Y})}(u, v), \qquad \text{за } (u, v) \in \mathbb{R}^2.$$

Д-во.* За всяко $A \in \mathcal{B}(\mathbb{R}^2) \cap g(D_{X,Y})$, използвайки смяната $(x, y) = h(u, v)$, имаме

$$\mathbb{P}(g(X, Y) \in A) = \mathbb{P}\big((X, Y) \in h(A)\big) = \int_{h(A)} f_{X,Y}(x, y) dx dy = \int_A f_{X,Y}(h(u, v)) |J(u, v)| du dv.$$

$\square$

Твърдение 4.6 (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*

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_Y(z - x) f_X(x) dx, \qquad \text{за } z \in \mathbb{R}.$$

Д-во. Поставяме $U = X + Y$ и $V = X$. Функцията $(x, y) \mapsto (x + y, x)$ е дифеоморфизъм с обратна функция $(u, v) \mapsto (v, u - v)$. Освен това $J(u, v) = \det\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = -1 \neq 0$, за всички $(u, v) \in \mathbb{R}^2$. Следователно от Твърдение 4.5 получаваме, че (U, V) е вектор от непрекъснати сл. в. с плътност

$$f_{U,V}(u, v) = f_{X,Y}(v, u - v) \cdot |J(u, v)| = f_X(v) \cdot f_Y(u - v).$$

Тогава

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u, v) dv = \int_{-\infty}^{\infty} f_X(v) f_Y(u - v) dv.$$

Алтернативно, за всички $z \in \mathbb{R}$, получаваме, полагайки $C = \{(x, y) \in \mathbb{R}^2 : x + y \leq z\}$,

$$\begin{aligned} F_{X+Y}(z) &= \mathbb{P}(X + Y \leq z) = \mathbb{P}\big((X, Y) \in C\big) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_X(x) f_Y(y) dy dx = \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{z-x} f_Y(y) dy\Big) f_X(x) dx = \int_{-\infty}^{\infty} F_Y(z - x) f_X(x) dx. \end{aligned}$$

Тогава F_{X+Y} е абсолютно непрекъсната функция и следователно е диференцируема почти навсякъде. Като използваме правилото на Лайбниц за диференциране на интеграли, получаваме, че $X + Y$ е непрекъсната сл. в. с плътност, определена за (п.вс.) $z \in \mathbb{R}$ чрез

$$f_{X+Y}(z) = \frac{d}{dz} F_{X+Y}(z) = \int_{-\infty}^{\infty} \Big(\frac{d}{dz} F_Y(z - x)\Big) f_X(x) dx = \int_{-\infty}^{\infty} f_Y(z - x) f_X(x) dx.$$

$\square$

Задачи

Задача 4.6. Електронно устройство за предпазване от крадци автоматично променя осветлението в дома. То е настроено така, че през фиксиран час, в случаен момент X ще запали лампите, а в момент Y ще ги угаси. Нека съвместната плътност на случайните величини X и Y е $f_{X,Y}(x,y) = cxy$, за $0 < x < y < 1$, и 0 иначе. Да се намери:

(а) константата c;
(б) маргиналните плътности и математическите очаквания на X и Y;
(в) вероятността лампите да бъдат запалени преди 45-тата минута и да светят по-малко от 10 минути;
(г) вероятността лампите да светят по-малко от 20 минути.

Задача 4.7. Нека (X,Y) имат съвместна плътност $f_{X,Y}(x,y) = ax^2 + bxy$, за $x \in (0,1), y \in (0,2)$, и 0 иначе, където a, b са положителни константи. Намерете a, b, ако знаете, че $\mathbb{E}[X] = \frac{13}{18}$, и пресметнете $\mathbb{P}(X + Y \geq 1)$.

*Задача** 4.8. Нека X и Y са сл. в. със съвместна функция на разпределение $F_{X,Y}$. Изразете функцията на разпределение на $(\max\{X,Y\}, \min\{X,Y\})$ чрез $F_{X,Y}$.

Задача 4.9. Нека X и Y са независими и еднакво разпределени непрекъснати сл. в. с плътност $f(x) = \mathbb{1}_{(0,1)}(x)$. Пресметнете дисперсията на $Z := |X - Y|$.

Задача 4.10. Нека (X,Y) имат съвместна плътност $f_{X,Y}(x,y) = 4xy$, за $x \in (0,1), y \in (0,1)$, и 0 иначе. Намерете $\mathrm{Cov}(X,Y)$ и съвместната плътност на (X^2, Y^2).

4.3 Условно очакване

Условно очакване. *Нека (X,Y) е непрекъснат вектор от сл.в. със съвместна плътност $f_{X,Y}$. Функцията*

$$f_{Y|X}(y|x) := \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)} & \text{ако } f_X(x) > 0; \\ 0 & \text{иначе,} \end{cases} \qquad \text{за } x, y \in \mathbb{R},$$

наричаме условна плътност на Y относно X, а

$$\mathbb{P}(a < Y \leq b | X = x) := \int_a^b f_{Y|X}(y|x) dy, \qquad \text{за } a, b \in \mathbb{R} : a < b,$$

условната вероятност на $\{a < Y \leq b\}$ при условие $X = x$.

*Забележка*** (Мотивация). За разлика от дискретния случай, тук имаме, че $\mathbb{P}(X = x) = 0$, т.е. не можем да използваме стандартната дефиниция на условна вероятност. От друга страна, в дискретния случай използваме условното разпределение на Y относно X, за да пресмятаме

$$\mathbb{P}(X \in A, Y \in B) = \sum_{x \in A} \mathbb{P}(Y \in B | X = x) \mathbb{P}(X = x),$$

за всички измерими събития A и B. Следователно в непрекъснатия случай очакваме условната вероятност на $\{Y \in B\}$ относно X да бъде някаква функция $h : \mathbb{R} \to \mathbb{R}$, такава че

$$\mathbb{P}(X \in A, Y \in B) = \int_A h(x) f_X(x) dx.$$

Последното е изпълнено, ако вземем $h(x) = \mathbb{P}(Y \in B | X = x) = \int_B f_{Y|X}(y|x) dy$. Обърнете внимание, че $\mathbb{P}(X \in D_X) = 1$, т.е. $f_X(x) > 0$ за п.вс. x. (спрямо вероятностната мярка $\mathbb{P}_X$), така че дефиницията на $f_{Y|X}(y|x)$ не е ограничаваща.

Освен това, за фиксирани $x \in \mathbb{R}$, изображението $\mathbb{P}(Y \in \cdot \mid X = x)$ е добре дефинирана вероятностна мярка върху $\mathbb{R}$. Тогава, ако $g : \mathbb{R}^2 \to \mathbb{R}$ е измерима функция, а интегралът абсолютно сходящ, то

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

Твърдение 4.7. *Нека (X,Y) е непрекъснат вектор от сл.в. със съвместна плътност $f_{X,Y}$ и $g:\mathbb{R}^2\to\mathbb{R}$ е измерима функция. Тогава*

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

Ако в допълнение $X \perp\!\!\!\perp Y$, то $\mathbb{E}[g(X,Y)|X=x] = \mathbb{E}[g(x,Y)]$ за п.вс. x.

Д-во.* Използвайки, че $\mathbb{P}(X\in D_X) = \int_{D_X} f_X(x)dx = \int_{-\infty}^{\infty} f_X(x)dx = 1$, т.е. $f_X(x)>0$ за п.вс. x (спрямо $\mathbb{P}_X$),

$$\begin{aligned}\int_{-\infty}^{\infty} \mathbb{E}[g(x,Y)|X=x] f_X(x)dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{Y|X}(y|x) f_X(x)dydx \\ &= \int_{D_X}\int_{-\infty}^{\infty} g(x,y)\frac{f_{X,Y}(x,y)}{f_X(x)} f_X(x)dydx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)dydx = \mathbb{E}[g(X,Y)].\end{aligned}$$

Относно втория резултат, за всички x, такива че $f_X(x)>0$, имаме

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

□

Задачи

Задача 4.11. Нека (X,Y) е непрекъснат вектор от сл.в. със съвместна плътност

$$f_{X,Y}(x,y) = \begin{cases} cx+y & \text{ако } x,y\geq 0, x+2y\leq 1; \\ 0 & \text{иначе,}\end{cases}$$

за всички $x,y\in\mathbb{R}$ и някоя константа $c\in\mathbb{R}$. Намерете c, $\mathbb{E}[Y|X=1/2]$ и плътността на $Z=X+2Y$.

Задача 4.12. Точка (X,Y) попада по случаен начин в триъгълник с върхове точките с координати $(0,0)$, $(0,2)$ и $(3,0)$. Да се намери съвместната плътност, функцията на разпределение и корелацията на X и Y.

Задача 4.13. Във вътрешността на кръг с радиус R случайно се избират точките A и B. Да се намери вероятността окръжността с център A и радиус $|AB|$ да лежи във вътрешността на кръга.

4.4 Основни непрекъснати разпределения

Равномерно разпределение

Равномерно разпределение *върху интервала (a,b), за $a,b\in\mathbb{R}: a<b$, е разпределение на непрекъснатата 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 (Интуиция). От горната дефиниция следва, че всички интервали с една и съща дължина в (a,b) са еднакво вероятни. По този начин равномерното разпределение изразява еднаквата предразположеност на изследователя към възможните стойности в (a,b) при все, че $\mathbb{P}(X=x)=0$ за всички $x\in(a,b)$.

Забележка (Стандартно равномерно разпределение). Нека $Y := \frac{X-a}{b-a}$. От Твърдение 4.1 следва, че Y е непрекъсната сл. в. с плътност

$$f_Y(y) = f_X\big(a+(b-a)y\big)\cdot|b-a| = 1\cdot\mathbb{1}_{(0,1)}(y),$$

т.е. $Y\sim\mathrm{Unif}(0,1)$.

Твърдение 4.8. *Ако* $X \sim \mathrm{Unif}(a,b)$*, тогава* $\mathbb{E}[X] = \frac{a+b}{2}$, $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$, $F_X(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & x \in (a,b) \\ 1 & x \geq b \end{cases}$, $\mathbb{E}[e^{tX}] = \frac{e^{tb}-e^{ta}}{t(b-a)}$*, за* $t \neq 0$*, и* $\mathbb{E}[e^{tX}] = 1$*, за* $t = 0$.

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 $F_X(x) = \int_a^x f_X(y)dy = \frac{y}{b-a}\Big|_a^x = \frac{x-a}{b-a}$, за $x \in (a,b)$,

$$\mathbb{E}[X] = \int_a^b \frac{x}{b-a}dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{a+b}{2},$$
$$\mathbb{E}[X^2] = \int_a^b \frac{x^2}{b-a}dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^2+ab+a^2}{3} \qquad \Longrightarrow \qquad \mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{(b-a)^2}{12},$$
$$\mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx = \frac{1}{b-a}\int_a^b e^{tx}dx = \frac{1}{t(b-a)}\int_{ta}^{tb} e^u du = \frac{e^{tb}-e^{ta}}{t(b-a)}.$$

□

Експоненциално разпределение

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 (Безпаметност). Това разпределение може да се интерпретира като непрекъснатия аналог на геометричното. Всъщност $F_X(x) = \int_0^x \lambda e^{-\lambda t}dt = e^t|_{-\lambda x}^0 = 1 - e^{-\lambda x}$, за $x > 0$, откъдето

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

т.е. експоненциалното разпределение притежава свойството безпаметност, като може да се покаже, че то е единственото непрекъснато такова.

Твърдение 4.9. *Ако* $X \sim \mathrm{Exp}(\lambda)$*, тогава* $\mathbb{E}[X] = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$ *и* $\mathbb{E}[e^{tX}] = \frac{\lambda}{\lambda - t}$*, за* $t < \lambda$.

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. Използвайки съответно интегриране по части и метода на Файнман, получаваме

$$\mathbb{E}[X] = \int_0^\infty x\lambda e^{-\lambda x}dx = -\int_0^\infty x de^{-\lambda x} = -(xe^{-\lambda x})|_0^\infty + \int_0^\infty e^{-\lambda x}dx = -\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda},$$
$$\frac{d}{d\lambda}\frac{1}{\lambda} = \frac{d}{d\lambda}\mathbb{E}[X] \quad \Longleftrightarrow \quad -\frac{1}{\lambda^2} = \int_0^\infty xe^{-\lambda x}dx - \int_0^\infty x^2\lambda e^{-\lambda x}dx = \frac{1}{\lambda^2} - \mathbb{E}[X^2] \qquad \Longrightarrow \qquad \mathbb{E}[X^2] = \frac{2}{\lambda^2},$$
$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}, \qquad \mathbb{E}[e^{tX}] = \int_{-\infty}^\infty e^{tx}f_X(x)dx = \int_0^\infty \lambda e^{-(\lambda - t)x}dx = -\frac{\lambda}{\lambda - t}\int_0^{-\infty} e^u du = \frac{\lambda}{\lambda - t}.$$

□

Нормално разпределение

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

$$f_Z(z) = f_X(\sigma z + \mu) \cdot |\sigma| = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot \mathbb{1}_{(-\infty,\infty)}(z);$$

т.е. $Z \sim \mathrm{N}(0,1)$.

Забележка (Гаусов интеграл). $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

Твърдение 4.10. *Ако* $X \sim \mathrm{N}(\mu, \sigma^2)$, *тогава* $\mathbb{E}[X] = \mu$, $\mathrm{Var}(X) = \sigma^2$ *и* $\mathbb{E}[e^{tX}] = e^{t\mu + \frac{\sigma^2 t^2}{2}}$, *за* $|t| < 1$.

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

$$\mathbb{E}[Z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{z^2}{2}} dx = 0, \qquad \text{понеже } z \mapsto z e^{-\frac{z^2}{2}} \text{ е нечетна интегруема функция} \qquad \Longrightarrow \qquad \mathbb{E}[X] = \mu,$$

$$\mathrm{Var}(Z) = \mathbb{E}[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz = \frac{-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z d e^{-\frac{z^2}{2}} = \frac{-1}{\sqrt{2\pi}} z e^{-\frac{z^2}{2}} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = 1 \quad \Longrightarrow \quad \mathrm{Var}(X) = \sigma^2,$$

$$\mathbb{E}[e^{tZ}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t)^2}{2} + \frac{t^2}{2}} dz = e^{\frac{t^2}{2}} \qquad \Longrightarrow \qquad \mathbb{E}[e^{tX}] = \mathbb{E}[e^{t(\mu + \sigma Z)}] = e^{t\mu} \mathbb{E}[e^{(t\sigma)Z}] = e^{t\mu + \frac{\sigma^2 t^2}{2}}.$$

По същия начин се доказва, че $\mathbb{E}[Z^{2n+1}] = 0$ за всяко $n \in \mathbb{N}_0$. □

Теорема 4.11 (Моавър-Лаплас). *Нека* $Z \sim \mathrm{N}(0,1)$ *и* $X_n \sim \mathrm{Bin}(n,p)$, *за* $p \in (0,1)$. *Тогава*

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

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

Забележка (Полза). Нека $Z \sim \mathrm{N}(0,1)$. Тогава функцията на разпределение на Z, която често се обозначава с $\Phi(z) \equiv F_Z(z)$, $z \in \mathbb{R}$, няма явен вид, но съществуват таблици с приблизителни нейни стойности. Тяхната полза идва от приближението, което Централната гранична теорема (и в частност Теорема 4.11) дава за функцията на разпределение на средно аритметично от група независими еднакво-разпределени сл. в. (с краен втори момент).

Забележка (Покритие). От неравенството на Чебишев следва, че за всяка сл. в. X с краен втори момент и дисперсия $\sigma^2 = \mathrm{Var}(X)$ имаме долна граница

$$\mathbb{P}(\mathbb{E}[X] - 2\sigma < X < \mathbb{E}[X] + 2\sigma) = \mathbb{P}(|X - \mathbb{E}[X]| < 2\sigma) = 1 - \mathbb{P}(|X - \mathbb{E}[X]| \geq 2\sigma) \geq 1 - \frac{\mathrm{Var}(X)}{(2\sigma)^2} = 0.75.$$

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

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

където Φ е функцията на разпределение на стандартно нормално разпределена сл. в.

Гама разпределение

Гама разпределение *с параметри* $\alpha > 0$ *и* $\beta > 0$ *е непрекъснато разпределение на неотрицателна сл. в.* X *с плътност*

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

В този случай ще обозначаваме $X \sim \Gamma(\alpha, \beta)$.

*Забележка** (Гама функция). Ойлеровата Гама функция е дефинирана като

$$\Gamma(z) := \int_0^\infty t^{z-1} e^{-t} dt, \qquad \text{за } z \in \mathbb{C} : \mathcal{R}(z) > 0.$$

Понеже $\Gamma(z+1) = z\Gamma(z)$, то $\Gamma(n) = (n-1)!$ за $n \in \mathbb{N}$, откъдето Гама функцията може да бъде раглеждана като обобщение на функцията факториел.

Твърдение 4.12. *Ако $X \sim \Gamma(\alpha, \beta)$, тогава $\mathbb{E}[X] = \frac{\alpha}{\beta}$, $\mathrm{Var}(X) = \frac{\alpha}{\beta^2}$ и $\mathbb{E}[e^{tX}] = \left(\frac{\beta}{\beta - t}\right)^\alpha$, за $t < \beta$.*

Задачи

Задача 4.14. Върху окръжност с център O и радиус r е фиксирана точката A, а точка B попада по случаен начин върху окръжността. Да се намери математическото очакване на лицето на $\triangle AOB$.

Задача 4.15. Нека $X \sim \mathrm{Unif}(0, 7)$ е времето за 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?

Задача 4.16. 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.

Задача 4.17. Нека X е непрекъсната сл. в. с плътност f_X. Да допуснем, че функцията ѝ на разпределение F_X е строго нарастваща. Намерете разпределенията на $F_X(X)$ и $F_X^{-1}(U)$, където $U \sim \mathrm{Unif}(0, 1)$.

Задача 4.18. Нека $X, Y \sim \mathrm{Unif}(0, 1)$ са независими. Да се намери разпределението на $X + Y$.

Задача 4.19. Нека $X \sim \mathrm{Unif}(0, 10)$. Пресметнете вероятността $\mathbb{P}(2 \le X \le 8)$, след което я оценете с неравенството на Чебишев и сравнете получените резултати.

Задача 4.20. Върху отсечката OA с дължина 1 избираме случайна точка X. След това избираме случайна точка Y върху отсечката XA. Каква е вероятността да можем да съставим триъгълник от отсечките OX, XY и YA?

Задача 4.21. Нека $U_1, U_2, U_3 \sim \mathrm{Unif}(0, 1)$ са независими сл.в. Означаваме с $U_{(1)} = \min\{U_1, U_2, U_3\}, \dots, U_{(3)} = \max\{U_1, U_2, U_3\}$ наредените в нарастващ ред U_1, U_2, U_3.

(а) Намерете разпределенията на $U_{(1)}$ и $U_{(3)}$.
(б) Вярно ли е, че $U_{(2)} \sim \mathrm{Unif}(0, 1)$? Ако да, докажете го, а ако не – намерете разпределението му.
(в) Намерете дисперсията на $-\ln U_1$.

Задача 4.22. 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?

Задача 4.23. 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.

Задача 4.24. Застрахователна компания "Инс 1" моделира размера на исковете, които изплаща чрез независими експоненциални сл. в. със средно 100 лв. "Инс 1" сключва презастраховка на цена от $x > 0$ лв с "Инс 2", която гласи, че ако постъпи иск над 300 лв към "Инс 1", "Инс 2" ще покрие 200 лв от тях. Каква е вероятността "Инс 1" да трябва да плати от своя бюджет по-малко от 200 лв за един иск?

Задача 4.25. Нека $X_1, X_2, \dots$ са независими $\mathrm{Exp}(1)$ сл. в.

(а) Намерете разпределението $\max\{X_1, \dots, X_n\}$ и $\min\{X_1, \dots, X_n\}$.
(б) Покажете, че $Y_n = \max\{X_1, \dots, X_n\} - \log(n)$ се схожда по разпределение.

Задача 4.26. Нека $X_1, X_2 \sim \mathrm{Exp}(\lambda)$ са независими. Да се намери разпределението на $Y = \frac{X_1}{X_1+X_2}$.

Задача 4.27. Нека $X \sim \mathrm{Exp}(2)$ и $Y \sim \mathrm{Unif}(0,3)$ са независими. Да се намери разпределението на $Z = X/Y$.

Задача 4.28. 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?

*Задача** 4.29. 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 $a, b \in \mathbb{R}$. Докажете, че $aX_1 + bX_2 \sim \mathrm{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$.

Задача 4.30. На спирките на градски транспорт се инсталират информационни табла с размери 10×100 диода. Доставени са качествени материали, като можем да моделираме времето на изправност на един диод чрез експоненциална сл. в. със средно 10 години. Опитът показва, че ако работят по-малко от 75% от диодите, информацията често е неразбираема и таблото трябва да се ремонтира. Каква е вероятността да трябва да бъде извършен ремонт след 3 години експлоатация?

Задача 4.31. Два инструмента се използват за измерването на прахови частици във въздуха. Да допуснем, че реалното количество е $x\ g/m^3$. В такъв случай, първият уред дава показание $X \sim \mathrm{N}(x, 0.0025x^2)$, а вторият $Y \sim \mathrm{N}(x, 0.01x^2)$. Кой апарат бихте използвали? Колко е вероятността за всеки от апаратите да допусне грешка, която е повече от $0.1x$? Човек решава да използва средно аритметичното от двата апарата. Ако измерванията им са независими, каква е вероятността за грешка над $0.1x$ при тази процедура?

Задача 4.32. Да предположим, че можем да моделираме възвращаемостта на три актива A, B, C като независими нормално разпределени случайни величини $\mathrm{N}(3,2), \mathrm{N}(3,3), \mathrm{N}(1,10)$ и че разполагате с 5 единици за инвестиции.

(а) Как бихте разпределили парите си, за да максимизирате очакваната печалба?

(б) Между всички възможности от (а) един начин за избор е да предпочетем разпределението с най-малка дисперсия. Кое е то?

(в) Рисков инвеститор залага 5-те си единици в независим актив $D \sim \mathrm{N}(-2, 20)$. Каква е вероятността неговата инвестиция да е по-успешна от тази в (б).

Задача 4.33. Размерът на пъпешите е $\mathrm{N}(25, 36)$. Пъпешите по-малки от 20 см са трето качество, а останалите се разделят на две равни по брой групи, като по-големите са първо качество, а по-малките второ. Каква част от пъпешите са трето качество? Колко голям трябва да е един пъпеш, за да е първо качество?

*Задача** 4.34. Според компания за производство на чипове, само 1 на всеки 1000 чипа е неизправен. Как бихте оценили вероятността от 100 чипа да има поне 1 неизправен?

Неравенството на Бери-Есен гласи, че ако $X_1, X_2, \ldots$ са i.i.d. сл. в. и

$$\mu := \mathbb{E}[X_1], \qquad \sigma := \sqrt{\mathrm{Var}(X_1)}, \qquad \rho := \mathbb{E}|X_1 - \mu|^3 < \infty,$$

то, полагайки $Z_n := (\sum_{i=1}^n X_i - n\mu)/(\sigma\sqrt{n})$, имаме

$$\sup_{x \in \mathbb{R}} \big|\mathbb{P}(Z_n \le x) - \Phi(x)\big| \le \frac{\rho}{2\sigma^3\sqrt{n}}.$$

При какви n грешката при приближение чрез ЦГТ би била под 0.001, ако приемете, че информацията, дадена от компанията, е вярна?

*Задача** 4.35. Нека $X_1, X_2 \sim \Gamma(\alpha, \beta)$ са независими. Намерете плътността на $Y := X_1/(X_1 + X_2)$ и $\mathrm{Corr}(X_1 + X_2, Y)$. Независими ли са $X_1 + X_2$ и Y?

*Задача** 4.36. Нека $X_1, X_2, \ldots$ са i.i.d. $\mathrm{Exp}(1)$ сл. в. Дефинираме $S_n := \sum_{i=1}^n X_i$ и $N_t := \max\{n \ge 1 : S_n \le t\}$, за $n \ge 1$ и $t \in \mathbb{R}_+$. Намерете плътността на S_n и разпределението на N_t.

5 Решения и отговори

5.1 Комбинаторика

Задача 1.1. $M = \{1, \dots, n\}$.

(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 $A := \{\text{няма празна клетка}\}$. Тогава $A^c = \{\text{поне една празна клетка}\}$. Всъщност, ако $A_i := \{\text{клетка } i \text{ не е празна}\}$, за $i = 1, \dots, n$, то A_i^c е множеството от всички k-орки на M, които не съдържат елемента i, при което $A = \bigcap_{i=1}^n A_i$ и $A^c = \bigcup_{i=1}^n A_i^c$.

От друга страна, $|A_i^c| = V(n-1,k)$, $|A_i^c \cap A_j^c| = V(n-2,k)$ и т.н. Следователно общият брой разпределения, при които не остава празна клетка, е равен на общият брой разпределения минус тези с поне една празна клетка, откъдето получаваме, използвайки принципа за включване-изключване,

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

За *неразличими*: при условие, че във всяка клетка има поне една топка, се търси комбинация от вида $\{1, 2, \dots, n, i_1, \dots, i_{k-n}\}$ за $i_1, \dots, i_{k-n} \in M$; т.е. общият брой е $C(n, k-n)$.

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

За *неразличими*: търси се комбинация от вида $\{1, \dots, 1, i_1, \dots, i_{k-s}\}$, такава че $i_j \neq i_l$ и $i_1, \dots, i_{k-s} \in \{2, \dots, n\}$; т.е. общият брой 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 $n-1 < k-s$.

$\square$

*Забележка** (Stirling numbers of the second kind). Числото на Стирлинг от втори род брои начините за разделяне на множество от k различими обекта (частици) на n *неразличими* непразни подмножества (клетки) и се обозначава с $S(k,n)$. Тогава от подточка (в) следва, че

$$S(k,n) = \frac{1}{n!} \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)^k.$$

Задача 1.2.

(а) *Решение 1*: задачата е еквивалентна на разпределянето на n неразличими частици в k клетки в случая, когато не остава празна клетка (Задача 1.1в); 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 $x_1 = \{\text{брой } * \text{ преди първата } |\}$, $x_k = \{\text{брой } * \text{ след последната } |\}$ и $x_i = \{\text{брой } * \text{ между } (i-1)\text{-тата и } i\text{-тата } |\}$, $i = 2, \dots, k-1$. 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 $\{j_1, \dots, j_{k-1}\}$ за $j_1, \dots, j_{k-1} \in \{1, \dots, n-1\} : j_m \neq j_l$, 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 1.1б); т.е. броят на решенията е $C(k,n)$.

Решение 2: съществува еднозначна обратимост между решенията на $x_1 + \dots + x_k = n$ и тези на $(x_1+1) + \dots + (x_k+1) = n+k$; т.е. C_{n+k-1}^{k-1} от (а).

$\square$

Задача 1.3. (а) V_5^4; (б) $V(5,4)$; (в) $3 \cdot V_4^3$ или $\frac{3}{5}V_5^4$. $\square$

Задача 1.4. (а) C_{12}^4; (б) $C_{10}^4 + 2 \cdot C_{10}^3$; (в) $C_{10}^2 + C_{10}^4$. $\square$

Задача 1.5. (а) 2^5; (б) $2^5 - 2$; (в) $3(2^5 - 2)$; (г) $3(2^5 - 2) + 3$; (д) $3^5 - (3(2^5 - 2) + 3)$. $\square$

Задача 1.6. Нека $M^a = \{a_1, \dots, a_m\}$ и $M^b = \{b_1, \dots, b_n\}$. (а) $|A_i| = \binom{m}{i}\binom{n}{k-i}$ и $|B_j| = |A_{k-j}|$; (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. $\bigcup_{i=\max\{0,k-n\}}^{\min\{m,k\}} A_i = \mathbf{C}_{m+n}^k$, откъдето $C_{m+n}^k = \sum_i |A_i|$; (в) $C_{m+n}^k - |A_0| - |B_0|$; (г) това са всички подмножества на M, с изключение на тези, състоящи се изцяло от елементи от един и същи тип,

$$|\mathcal{P}(M)| - |\mathcal{P}(M^a)| - |\mathcal{P}(M^b)| + 1 = (2^m - 1)(2^n - 1) = \big(|\mathcal{P}(M^a)| - 1\big)\big(|\mathcal{P}(M^b)| - 1\big),$$

където добавяме една единица, понеже празното множество е добавено веднъж и извадено два пъти. $\square$

Задача 1.7. $8 \cdot P_3 P_7$ $\square$

Задача 1.8. (а) 3^{n-1}; (б) $\binom{n}{k}2^{n-k}$; (в) $\binom{n-2}{k-2}2^{n-k}$; (г) $\binom{n}{k_1,k_2,k_3}$. $\square$

Задача 1.9. (а) V_{2n}^{2k}; (б) $\frac{1}{k!}V_{2n}^{2k}$; (в) $\prod_{i=0}^{k-1} C_{2n-2i}^2$; (г) $\frac{1}{k!}\prod_{i=0}^{k-1} C_{2n-2i}^2$. $\square$

Задача 1.10. 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

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

$\square$

Забележка (Числа на Каталан). n-тото число на Каталан се задава от

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n+1},$$

и отговаря на броя на траекториите, започващи в $(0, 0)$ и завършващи в $(2n, 0)$, които докосват, но не пресичат абсцисната ос отгоре.

5.2 Вероятности

Дискретни вероятности

Задача 2.1.

(а) $\Omega = \big\{(a_1, a_2) : a_1, a_2 \in \{1, \dots, 6\}\big\}$. $\mathbb{P}(\{\text{поне една 1-ца}\}) = 1 - \frac{\text{изходи без единица}}{\text{всички изходи}} = 1 - \frac{5^2}{6^2} \approx 0.306$;

(б) $\Omega = \big\{\{a_1, a_2\} : a_1, a_2 \in \{1, \dots, 6\}\big\}$. $\mathbb{P}(\{\text{поне една 1-ца}\}) = 1 - \frac{C(5,2)}{C(6,2)} \approx 0.286$.

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. $\square$

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 $\{1,6\}$ очакваме да е (два пъти) по-вероятно от $\{6,6\}$). 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 2.2. $\Omega = \big\{\{a_1, a_2\} : a_1, a_2 \in \{0,\dots,6\}\big\}$, като $|\Omega| = C(7,2)$, понеже играта се играе с 28 плочки. В този случай, за разлика от задачата със заровете, така описаните елементарни събития са равновероятни, откъдето следва, че $\mathbb{P}(\{\{a_1,a_2\} : a_1 \neq a_2\}) = C_7^2/C(7,2) = 0.75$. □

Задача 2.3. $\binom{6}{i}\binom{43}{6-i}/\binom{49}{6}$, за $i = 3,\dots,6$. □

Задача 2.4. $\binom{7}{4}2^3/3^7$. □

Задача 2.5. $\Omega = \big\{(a_1,\dots,a_4) : a_i \in \{0,1,\dots,9\}\big\}$. (а) $V_{10}^4/10^4$. (б)

Решение 1: ще съставим такова $(a_1,\dots,a_4)$ като първо изберем уникалните му цифри $\{a_1^*, a_2^*, a_3^*\}$ и след това пермутираме елементите на $\{a_1^*, a_2^*, a_3^*\}$ с повторение и честота $(1,1,2)$ или $(1,2,1)$ или $(2,1,1)$, в зависимост от това кое a_i^* повтаряме; т.е. търсената вероятност е $\binom{10}{3}\big\{\binom{4}{2,1,1} + \binom{4}{1,2,1} + \binom{4}{1,1,2}\big\}/10^4$.

Решение 2: избираме коя цифра да повторим, на кои две места да я повторим, след което попълваме останалите 2 места; т.е. $10\binom{4}{2}V_9^2/10^4$.

(в) $\binom{10}{2}\big\{\binom{4}{3,1} + \binom{4}{1,3}\big\}/10^4$ или $10\binom{4}{3}V_9^1/10^4$; (г) $\binom{10}{2}\binom{4}{2,2}/10^4$ или $\binom{4}{2}\binom{10}{2}/10^4$; (д) за двойките (a_1,a_2) и (a_3,a_4) има различен брой възможности в зависимост от сумата $a_1 + a_2 = a_3 + a_4$, като

$a_1 + a_2 =$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
варианти за (a_1,a_2)	1	2	3	4	5	6	7	8	9	10	9	8	7	6	5	4	3	2	1
варианти за (a_1,a_2,a_3,a_4)	1	4	9	16	25	36	49	64	81	100	81	64	49	36	25	16	9	4	1

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

Задача 2.6. Можем да си представим, че изплащането на наградите се случва след като се изтеглят всички карти, т.е. $\Omega = \big\{(a_1,\dots,a_{52}) : a_i \in \{1,\dots,52\}, a_i \neq a_j\big\}$.

Печеливши за Играч 1 са пермутациите, при които наблюдаваме $(\dots, 7\text{C}, \dots, A_1, \dots, A_2, \dots, A_3, \dots, A_4, \dots)$ или $(\dots, A_1, \dots 7\text{C}, \dots, A_2, \dots, A_3, \dots, A_4, \dots)$. Можем да конструираме еднозначно всяка една от тях, като първо изберем позициите на асата и 7С заедно, след това ги подредим по един от описаните начини и накрая запълним останалите места с останалите карти. Тогава

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

Всъщност от цялата информация ни интересува единствено позицията на 7С между четирите аса (независимо от цветовете им), т.е. можем да работим с $\Omega^* = \{(7\text{C}, A, A, A, A), \dots (A, A, A, A, 7\text{C})\}$. В този случай елементарните изходи остават равно вероятни, откъдето получаваме отново $\mathbb{P}^*(\{\text{Играч 1 печели}\}) = 2/5$ (сравнете със Задача 2.41). □

Задача 2.7. По условие $\Omega = \big\{(a_1,\dots,a_s) : a_i \in \{\text{д}_1,\dots,\text{д}_n, \text{б}_1,\dots,\text{б}_m\}, a_i \neq a_j, i \neq j, a_1,\dots,a_k = \text{“добри”}\big\}$. Тогава

$$\mathbb{P}(\{k+1\text{-во добро}\}) = \frac{V_n^k\cdot(n-k)\cdot(n+m-k-1)!}{V_n^k\cdot(n+m-k)!} = \frac{n-k}{n-m-k}.$$

Всъщност можем да си представим, че случайният експеримент се провежда след първата проверка, когато партидата вече се състои от $n-k$ доброкачествени и m бракувани изделия. Абстрахирайки се допълнително от проверката на последните $s-k-1$ изделия, получаваме отново $\mathbb{P}(\{k+1\text{-во добро}\}) = \frac{n-k}{n+m-k}$. □

Задача 2.8. $2\binom{2k-2}{k-1}/\binom{2k}{k}$. □

Задача 2.9. $\sum_{k=0}^5 \binom{10}{k}\binom{10-k}{k}4^{10-2k}/6^{10}$. □

Задача 2.10.

(а) $\Omega = \{(a_1, \dots, a_k) : 1 \le a_i \le n, a_i \neq a_j\}$. Нека $A = \{(a_1, \dots, a_k) : 1 \le a_1 < a_2 < \dots < a_k \le n\}$. Всяка редица $a_1 < \dots < a_k$ определя еднозначно комбинация на $\{1, \dots, n\}$ от k-ти ред без повторение, именно $\{a_1, \dots, a_k\}$, и обратно – една единствена подредба на $\{a_1, \dots, a_k\}$ образува растяща редица. От тук следва съществуването на биекция от A в $\mathbf{C}_n^k$, откъдето $\mathbb{P}(A) = C_n^k / V_n^k = \frac{1}{k!}$.

(б) $\Omega = \{(a_1, \dots, a_k) : 1 \le a_i \le n\}$. Нека $A = \{(a_1, \dots, a_k) : 1 \le a_1 \le a_2 \le \dots \le a_k \le n\}$. От всяка редица $a_1 \le \dots \le a_k$ можем да образуваме строго растяща такава $a_1^* < \dots < a_k^*$, използвайки трансформацията $a_i^* = a_i + (i-1)$, като тогава $a_i^* \in \{1, \dots, n+k-1\}$. Понеже тази трансформация е обратима, от (а) следва, че $\mathbb{P}(A) = C(n,k)/V(n,k)$.

□

Задача 2.11. $2 \cdot 10!10!/20!$. □

Задача 2.12. $\Omega = \{(a_1, \dots, a_n) : 1 \le a_i \le n, a_i \neq a_j\}$. Нека $A = \{$никой не е получил своето писмо$\}$. Дефинираме $A_i := \{a_i \neq i\}$, за $i = 1, \dots, n$. Тогава $A = \bigcap_{i=1}^n A_i$. От друга страна $A_i^c = \{a_i = i\}$, откъдето $|A_i^c| = (n-1)!$, $|A_i^c \cap A_j^c| = (n-2)!$, $i \neq j$, и т.н. Следователно, използвайки принципа за включване-изключване,

$$\mathbb{P}(A) = \frac{|\Omega| - |\bigcup_{i=1}^n A_i^c|}{|\Omega|} = 1 - \frac{|\bigcup_{i=1}^n A_i^c|}{n!} = \sum_{k=0}^n (-1)^k \binom{n}{k} \frac{(n-k)!}{n!} = \sum_{k=0}^n (-1)^k \frac{1}{k!}.$$

□

*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 $\lim_{n\to\infty} \frac{!n}{n!} = e^{-1}$, т.е. има ненулева вероятност горното събитие да се случи, независимо от това колко голям е броят на получателите.

Задача 2.13. $\Omega = \big\{(a_1, \dots, a_n) : a_i \in \{1, \dots, n\}, a_i \neq a_j\big\}$. Нека фиксирани са $x, y \in \{1, \dots, n\} : x \neq y$.

Случай 1: За да получим крайната подредба $(a_1, \dots, a_n)$, можем първо да изберем хората между x и y, $(b_1, \dots, b_r)$, след което заедно с x и y да ги "вмъкнем" между останалите хора, $(c_1, \dots, c_{n-r-2})$, на една от възможните $(n - r - 2 + 1)$ позиции, напр. $(c_1, \dots, c_k, y, b_1, \dots, b_r, x, c_{k+1}, \dots, c_{n-r-2})$. Понеже има 2 начина да се подредят x и y, търсената вероятност е

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

Еквивалентно, първият човек ще седне отляво или отдясно на група от $r+1$ души, оставяйки свободно място за другия човек, като има за тази цел $2(n-r-1)$ възможни места от n. Вторият ще седне на точно 1 определено място от $n-1$ оставащи места, за да се изпълни условието.

Случай 2: Нека приемем, че местата са номерирани и че номерацията е от значение при подредбата, като тогава ще можем да използваме същата логика като в случай 1. Тъй като обаче места "n" и "1" са "съседни", 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. 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

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

Еквивалентно, първият човек може да седне на всяко едно място, а вторият – на 2 възможни места от $n-1$, броейки $r+1$ места по часовниковата стрелка и $r+1$ места обратно на часовниковата стрелка. Когато $r = \frac{n-2}{2}$ е цяло число, тогава двете места съвпадат и вероятността е $\frac{1}{n-1}$.

□

Забележка. Двете решения показват, че от вероятностна гледна точка е все едно дали хората са настанени заедно, или един по един. При втория подход лице 2 ще седне "условно" на това къде е седнало лице 1.

Задача 2.14. (а) $\frac{C(n-1,k-r)}{C(n,k)}$; (б) $\binom{n}{m}\frac{C(n-m,k-n+m)}{C(n,k)}$; (в) $\frac{C(n,k-2n)}{C(n,k)}$; (г) $\frac{\sum_{i=0}^{\lfloor k/5 \rfloor}(-1)^i \binom{n}{i} C(n,k-5i)}{C(n,k)}$. □

Задача 2.15. $\frac{n-k}{n}$. □

Задача 2.16. (а) $\frac{\binom{k}{k_1,\dots,k_n}}{n^k}$; (б) $\frac{n!}{n^n}$; (в) $\frac{n\binom{n}{2}(n-1)!}{n^n} = \frac{n\sum_{i=0}^{n-1}(-1)^i\binom{n-1}{i}(n-1-i)^n}{n^n}$ от свойствата на числата на Стирлинг от втори ред; (г) $\binom{n}{m}\frac{\sum_{i=0}^{n-m}(-1)^i\binom{n-m}{i}(n-m-i)^k}{n^n}$. □

Задача 2.17. (а) $\frac{1}{6^n}$; (б) $\frac{n}{6^n}$; (в) $\frac{\sum_{k=0}^{\lfloor (s-n)/6\rfloor}(-1)^k\binom{n}{k}C(n,s-n-6k)}{6^n}$. □

Задача 2.18. $\Omega = \big\{(a_1,\dots,a_{2r}) : a_i \in \{\text{л}_1,\dots,\text{л}_n,\text{д}_1,\dots,\text{д}_n\}, a_i \neq a_j\big\}$, $|\Omega| = V_{2n}^{2r}$.

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

(б) $\frac{\binom{2n}{2}\cdot 2\cdot n\cdot(2n-2)\cdot(2n-4)\cdots(2n-4r+4)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = n2^{2r-2}\binom{n-1}{2r-2}/\binom{2n}{2r}$, 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.

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

□

Задача 2.19. За всяко $i = 1,\dots,n+m$, 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 $s_1,\dots,s_{n+m}$ да бъде положителен във всеки един момент от преброяването, където $s_i = x_1 + \cdots + x_i$. Геометрично погледнато, последното предполага, че начупената линията, свързваща точките $(0,0),(1,s_1),(2,s_2),\dots,(n+m,n-m)$, 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 1.10 получаваме

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

□

Условна вероятност и независимост на събития

Задача 2.20. Нека $A = \{\text{посещава хирург}\}$ и $B = \{\text{на лекар } > 1 \text{ годишно}\}$. По условие $\mathbb{P}(A) = 0.17$, $\mathbb{P}(B) = 0.6$ и $\mathbb{P}(A|B) = 0.15$, откъдето

$$\mathbb{P}(A^c|B^c) = \frac{\mathbb{P}((A\cup B)^c)}{\mathbb{P}(B^c)} = \frac{1-\mathbb{P}(A)-\mathbb{P}(B)+\mathbb{P}(A|B)\mathbb{P}(B)}{1-\mathbb{P}(B)} = 0.80.$$

□

Задача 2.21. $(1-p)^k p$. □

Задача 2.22. $1-\prod_{k=1}^n(1-p_k)$. □

Задача 2.23. Нека $A = \{\text{убит от един куршум}\}$ и $H_i = \{\text{ловец } i \text{ уцелва}\}$, $i = 1,2,3$. 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 $\mathbb{P}(H_1|A) = \mathbb{P}(A\cap H_1)/\mathbb{P}(A) \approx 10\%$. □

Задача 2.24. $\min\{n\in\mathbb{N} : 1 - V_{365}^n/365^n > \frac{1}{2}\} = 23$. □

Задача 2.25. Нека $\Omega = \big\{(a_1,a_2) : a_1,a_2 \in \{m,f\}\big\}$, $A = \{\text{по-малкото е момиче}\} \equiv \{(m,f),(f,f)\}$ и $B = \{\text{по-старото е момиче}\} \equiv \{(f,m),(f,f)\}$, като $\mathbb{P}(\{\omega\}) = 1/4$, за $\omega\in\Omega$. Тогава вероятността за две момичета при условие, че по-голямото е момиче, е

$$\mathbb{P}(A\cap B|B) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{(f,f)\})}{\mathbb{P}(\{(f,m),(f,f)\})} = \frac{1}{2},$$

а вероятността за две момичета при условие, че едно е момиче, е

$$\mathbb{P}(A\cap B|A\cup B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A\cup B)} = \frac{\mathbb{P}(\{(f,f)\})}{\mathbb{P}(\{(f,m),(m,f),(f,f)\})} = \frac{1}{3}.$$

□

Задача 2.26. Решение 1. Нека си представим, че играчите теглят билети по реда на техните номера. Поставяме $A_k^0 = \{k\text{-тия играч тегли печеливш билет}\}$ и $A_k^1 = (A_k^0)^c$, за $k = 1, \dots, n$. Тогава

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

където $\sum_{i=\max\{0,k-(n-m)\}}^{\min\{m-1,k-1\}}\binom{m-1}{i}\binom{(n-1)-(m-1)}{k-i-1} = \binom{n-1}{k-1}$ от формулата на Вандермонд (виж Задача 1.6).

Решение 2. При едновременно теглене, начините за разпределяне на m печеливши билети между n души са $\binom{n}{m}$, от които $\binom{n-1}{m-1}$ са тези, при които печели k-тият играч; т.е. $\mathbb{P}(A_k^0) = \binom{n-1}{m-1}/\binom{n}{m} = m/n$. □

Задача 2.27. $\sum_{k=0}^{n-1}\binom{m+k-1}{k}p^m(1-p)^k$. □

Задача 2.28. $(n,k)\in\{(lr, l(m-r)) : l\in\mathbb{N}\}$, при условие, че $p = \frac{r}{m}$ е рационално число, за някои $r, m\in\mathbb{N}$. □

Задача 2.29. По условие $\mathbb{P}(+|\text{болен}) = 0.99$, $\mathbb{P}(-|\text{здрав}) = 0.99$ и $\mathbb{P}(\text{болен}) = 0.005$, откъдето

$$\mathbb{P}(\text{болен}|+) = \mathbb{P}(+|\text{болен})\frac{\mathbb{P}(\text{болен})}{\mathbb{P}(+|\text{болен})\mathbb{P}(\text{болен}) + (1-\mathbb{P}(-|\text{здрав}))\mathbb{P}(\text{здрав})} \approx 33\%.$$

□

Задача 2.30. Нека $A = \{\text{търсеното събитие}\}$ и $H = \{\text{избрани три стандартни зара}\}$. Тогава $\mathbb{P}(H) = 1/4$ и

а) $\mathbb{P}(A) = \mathbb{P}(A|H)\mathbb{P}(H) + \mathbb{P}(A|H^c)\mathbb{P}(H^c) = \frac{1}{4}\frac{1}{6^3} + \frac{3}{4}\frac{1}{6^2}$;

б) $\mathbb{P}(A) = \mathbb{P}(A|H)\mathbb{P}(H) + \mathbb{P}(A|H^c)\mathbb{P}(H^c) = \frac{1}{4}\frac{V_6^3}{6^3} + \frac{3}{4}\frac{V_5^2}{6^2}$;

в) $\mathbb{P}(A) = \mathbb{P}(A|H)\mathbb{P}(H) + \mathbb{P}(A|H^c)\mathbb{P}(H^c) = \frac{1}{4}\frac{4\cdot 3!}{6^3} + \frac{3}{4}\frac{2}{6^2}$.

□

Задача 2.31. $\frac{1}{n-1}\cdot\frac{1}{n} + \frac{1}{n}\cdot\frac{n-1}{n}$. □

Задача 2.32. $\frac{\binom{4}{3}}{\binom{7}{3}}\cdot\frac{\binom{3}{3}}{\binom{7}{3}} + \frac{\binom{3}{3}}{\binom{7}{3}}\cdot\frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}}$. □

Задача 2.33. Нека $A = \{2\text{ усп. и }1\text{ неусп.}\}$ и $H_i = \{\text{избрани } i \text{ момичета}\}$, $i = 0,1,2,3$. Тогава $H_0,\dots,H_3$ образуват пълна група и $\mathbb{P}(H_i) = \binom{45}{i}\binom{55}{3-i}/\binom{100}{3}$. Отделно $\mathbb{P}(A|H_0) = 3\cdot 0.4^2\cdot 0.6$, $\mathbb{P}(A|H_1) = 0.4^2\cdot 0.3 + 2\cdot 0.4\cdot 0.7\cdot 0.6$, $\mathbb{P}(A|H_2) = 0.7^2\cdot 0.6 + 2\cdot 0.7\cdot 0.4\cdot 0.3$ и $\mathbb{P}(A|H_3) = 3\cdot 0.7^2\cdot 0.3$, откъдето

$$\mathbb{P}(H_3|A) = \frac{\mathbb{P}(A|H_3)\mathbb{P}(H_3)}{\sum_{i=0}^3\mathbb{P}(A|H_i)\mathbb{P}(H_i)} \approx 10\%.$$

□

Задача 2.34. Нека $A = \{\text{горна бяла страна}\}$ и $H_i = \{\text{избрана е монета с } i \text{ бели страни}\}$, $i = 0,1,2$. Тогава H_0, H_1, H_2 образуват пълна група, $\mathbb{P}(H_i) = 1/3$ и $\mathbb{P}(A|H_i) = i/2$, откъдето

$$\mathbb{P}(H_2|A) = \frac{\mathbb{P}(A|H_2)\mathbb{P}(H_2)}{\sum_{i=0}^2\mathbb{P}(A|H_i)\mathbb{P}(H_i)} = \frac{2}{3}.$$

Забележете, че $\mathbb{P}(H_2|H_1\cup H_2) = 1/2$, но $A \subsetneq H_1\cup H_2$, понеже събитието $H_1\cup H_2$ позволява да сме видяли отгоре черната страна на смесената монета. □

Задача 2.35. Нека сме избрали врата 1, при което водещият, избирайки между врати 2 и 3, отваря врата 3, зад която няма нищо. Дефинираме $A = \{\text{отваря врата 3 при избрана врата 1}\}$ и $H_i = \{\text{врата } i \text{ печели}\}$, $i = 1, 2, 3$. Знаем, че $\mathbb{P}(H_i) = 1/3$. По условие $\mathbb{P}(A|H_1) = 1/2$, $\mathbb{P}(A|H_2) = 1$ и $\mathbb{P}(A|H_3) = 0$. Тогава

$$\frac{\mathbb{P}(H_1|A)}{\mathbb{P}(H_2|A)} = \frac{\mathbb{P}(A|H_1)\mathbb{P}(H_1)}{\mathbb{P}(A|H_2)\mathbb{P}(H_2)} = \frac{1}{2},$$

т.е. има двойно по-голям шанс колата да се намира зад врата 2. Забележете, че $\mathbb{P}(H_1|H_1 \cup H_2) = 1/2$, но $A \subsetneq H_1 \cup H_2$.

Всъщност, ако променим условието на играта така: играчът <u>сам</u> избира врата 3 да бъде отворена (игра тип "сделка или не") и зад нея няма нищо, тогава вероятността колата да бъде зад врата 1 действително е 50%, т.е. в действието на водещия има допълнителна информация, която следва да се вземе предвид. Това е така, защото водещият винаги ще отвори врата, зад която няма нищо, докато избирайки сам, играчът може да отвори вратата, зад която е колата. □

Задача 2.36. Нека $A = \{\text{2-ро бракувано}\}$, $B = \{\text{1-во добро}\}$ и $C = \{\text{от I-ва партида}\}$. Тогава

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

□

Задача 2.37. Нека $A = \{\text{взема изпита}\}$, $H_i = \{\text{отговаря на } i \text{ въпроса}\}$, $F = \{\text{отговаря на екстра въпрос}\}$, за $i = 0, 1, 2$. Тогава

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A|H_2)\mathbb{P}(H_2) + \mathbb{P}(A|H_1)\mathbb{P}(H_1)\\ &= \mathbb{P}(A|H_2)\mathbb{P}(H_2) + \mathbb{P}(A|H_1,F)\mathbb{P}(F|H_1)\mathbb{P}(H_1) = 1\cdot\frac{25\cdot 24}{30\cdot 29} + 1\cdot\frac{24}{28}\cdot\frac{2\cdot 25\cdot 5}{30\cdot 29} \approx 94\%.\end{aligned}$$

□

Задача 2.38. Нека $A = \{\text{бяла преди зелена}\}$.

(а) $\mathbb{P}(A) = \mathbb{P}(\text{б}) + \mathbb{P}(\text{чб}) + \mathbb{P}(\text{ччб}) + \cdots = \frac{n}{n+m+l}\sum_{k=0}^{\infty}\left(\frac{l}{n+m+l}\right)^k = \frac{n}{n+m+l}\cdot\frac{1}{1-\frac{l}{n+m+l}} = \frac{n}{n+m}$.

Рекурсивно решение: Ако на първия ход няма победител, т.е. е настъпило събитието $\{\text{ч}\}$, то играта започва отначало от вероятностна гледна точка. Това може да бъде изведено от

$$\mathbb{P}(A) = \mathbb{P}(A|\text{б})\mathbb{P}(\text{б}) + \mathbb{P}(A|\text{з})\mathbb{P}(\text{з}) + \mathbb{P}(A|\text{ч})\mathbb{P}(\text{ч}) = \frac{n}{n+m+l} + \frac{l}{n+m+l}\mathbb{P}(A|\text{ч}).$$

Но $\mathbb{P}(\text{чб}|\text{ч}) = \mathbb{P}(\text{б})$, $\mathbb{P}(\text{ччб}|\text{ч}) = \mathbb{P}(\text{чб})$, и т.н., откъдето $\mathbb{P}(A|\text{ч}) = \mathbb{P}(\text{чб}|\text{ч}) + \mathbb{P}(\text{ччб}|\text{ч}) + \cdots = \mathbb{P}(A)$. Следователно $\mathbb{P}(A) = \frac{n}{n+m+l} + \frac{l}{n+m+l}\mathbb{P}(A)$, чието решение е $\mathbb{P}(A) = \frac{n}{n+m}$.

(б) В този случай играта има победител след краен брой ходове, като

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

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

Алтернативно решение: Подобно на Задача 2.6 можем (i) да се абстрахираме от червените топки, и (ii) да наредим всички бели и зелени топки, след което да видим цвета на първата топка в редицата. Тогава $\mathbb{P}(A) = \mathbb{P}(\text{бяла на първо място}) = \frac{n\cdot(n+m-1)!}{(n+m)!} = \frac{n}{n+m}$.

□

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

Задача 2.39. Нека $A = \{\text{I печели}\}$. Тогава

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

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

$$\mathbb{P}(A) = \mathbb{P}(A|E)\mathbb{P}(E) + \mathbb{P}(A|TE)\mathbb{P}(TE) + \mathbb{P}(A|TT)\mathbb{P}(TT) = \frac{1}{2} + 0 + \frac{1}{4}\mathbb{P}(A) \quad \Longrightarrow \quad \mathbb{P}(A) = \frac{2}{3}.$$

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. Нека $A_k = \{\text{от } k\text{-тата урна вадим бяла топка}\}$, $k = 1, \ldots, N$. 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

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

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

Алтернативно решение: Нека $A_k^0 = A_k$ и $A_k^1 = A_k^c$, за $k = 1, \ldots, n$. Понеже вероятността на $A_k^{i_k}$ зависи само от $A_{k-1}^{i_{k-1}}$ (т.нар. **Марковско свойство**),

$$\begin{aligned}
\mathbb{P}(A_N) &= \sum_{i_1,\ldots,i_{N-1}\in\{0,1\}} \mathbb{P}(A_1^{i_1})\mathbb{P}(A_2^{i_2}|A_1^{i_1})\cdots\mathbb{P}(A_N|A_1^{i_1},\ldots,A_{N-1}^{i_{N-1}})\\
&= \sum_{i_1,\ldots,i_{N-1}\in\{0,1\}} \mathbb{P}(A_1^{i_1})\mathbb{P}(A_N|A_{N-1}^{i_{N-1}})\prod_{k=2}^{N-1}\mathbb{P}(A_k^{i_k}|A_{k-1}^{i_{k-1}})\\
&= \sum_{i_1,\ldots,i_{N-1}\in\{0,1\}} \Big(i_1\frac{m}{m+n} + (1-i_1)\frac{n}{m+n}\Big)\frac{m+i_{N-1}}{m+n+1}\prod_{k=2}^{N-1}\Big(i_k\frac{m+i_{k-1}}{m+n+1} + (1-i_k)\frac{n+(1-i_{k-1})}{m+n+1}\Big)\\
&= \frac{m}{m+n}.
\end{aligned}$$

□

Задача 2.41. Нека $A = \{\text{червено преди черно след 6-та}\}$, $B = \{\text{червено на 6-та}\}$ и $C = \{\text{червено след 6-та}\}$. Можем да подходим към проблема подобно на Задача 2.6, като

$$\mathbb{P}(A|B) = \frac{2!\binom{46}{3}2!48!}{2!51!} + \frac{2!\binom{5}{1}\binom{46}{2}2!48!}{2!51!} + \frac{2!\binom{5}{2}\binom{46}{1}2!48!}{2!51!} \approx 0.33.$$

Алтернативно решение: Абстрахирайки се от останалите карти и от конкретната боя на асата, и концентрирайки се само върху случаите, когато има червено асо на 6-та позиция **и** второ червено асо след 6-та позиция, ще наблюдаваме, условно на B, едно от събитията

$$\big\{(b,b,\boxed{\text{r}},r),(b,\boxed{\text{r}},b,r),(b,\boxed{\text{r}},r,b),(\boxed{\text{r}},b,b,r),(\boxed{\text{r}},b,r,b),(\boxed{\text{r}},r,b,b)\big\},$$

като в 50% от случаите ще видим първо червено преди черно асо след 6-та позиция. Вероятността на A обаче не е 50% понеже елементарните събития не са равно вероятни. Нека H_i е събитието "има i черни аса преди 6-та позиция", за $i = 0,1,2$. Тогава $\mathbb{P}(H_i|B) = 2\binom{5}{i}\binom{46}{2-i}/(51\cdot 50)$ и $\mathbb{P}(C|B,H_i) = (44+i)/49$, откъдето $\mathbb{P}(b,b,\boxed{\text{r}},r|B) = \mathbb{P}(H_2\cap C|B)$, $\mathbb{P}(b,\boxed{\text{r}},r,b|B) = \frac{1}{2}\mathbb{P}(H_1\cap C|B)$ и $\mathbb{P}(\boxed{\text{r}},r,b,b|B) = \frac{1}{3}\mathbb{P}(H_1\cap C|B)$. Следователно

$$\mathbb{P}(A|B) = \mathbb{P}(b,b,\boxed{\text{r}},r|B) + \mathbb{P}(b,\boxed{\text{r}},r,b|B) + \mathbb{P}(\boxed{\text{r}},r,b,b|B) \approx 0.33.$$

□

Задача 2.42. 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 $p(x) = x/a$.

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 1.10, 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 1.10, 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

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

$\square$

Забележка (Край на играта). По същия начин се доказва, че $\mathbb{P}(\{\text{играчът фалира}\}) = (a-x)/a$, 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

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

5.3 Дискретни случайни величини

Основни понятия

Задача 3.1. Нека $\Omega = \{(a_1, \ldots, a_{10}) : a_i \in \{E, T\}\}$ и $\mathbb{P}(\{\omega\}) = 2^{-10}$, за $\omega \in \Omega$. Дефинираме

$$X(\omega) := \sum_{i=1}^{n} \mathbb{1}_{\{a_i = E\}}, \qquad \text{за } \omega = (a_1, \ldots, a_{10}) \in \Omega.$$

Тогава X е дискретна сл.в., описваща броя на езитата от 10 опита, като приема стойности в $\mathbb{X} = \{0, 1, \ldots, 10\}$ с вероятности

$$\mathbb{P}(X = x) = \binom{10}{x}\left(\frac{1}{2}\right)^{10}, \qquad \text{за всяко } x \in \mathbb{X}.$$

От тук следва, че (а) $\mathbb{P}(X = 5) \approx 0.25$; (б) $\mathbb{P}(\{X = 4\} \cup \{X = 6\}) = 2 \cdot \mathbb{P}(X = 6) \approx 0.41$; (в) $\{X = 7.5\} = \emptyset$. $\square$

Задача 3.2.

X	2	3	4	5	6	7	8	9
$\mathbb{P}_X$	$\frac{8}{120}$	$\frac{14}{120}$	$\frac{18}{120}$	$\frac{20}{120}$	$\frac{20}{120}$	$\frac{18}{120}$	$\frac{14}{120}$	$\frac{8}{120}$

,

Y	0	1	4	9	16
$\mathbb{P}_Y$	$\frac{20}{120}$	$\frac{38}{120}$	$\frac{32}{120}$	$\frac{22}{120}$	$\frac{8}{120}$

,

$\mathbb{P}(X \geq 7) = 1/3$, $\mathbb{P}(3 \leq X < 7) = 9/15$, $\mathbb{P}(|X-2| < 2) = 11/60$. □

Задача 3.3. Нека сл.в. X описва броя наводнени отсека. Тогава

$$\mathbb{P}(X \geq 2) = 1 - (1-p)^n - m\sum_{i=1}^{n}\binom{n}{i}p^i(1-p)^{n-i}\frac{1}{m^i}.$$

Относно втория въпрос, $\mathbb{P}(X = k) = \binom{m}{k}\sum_{i=k}^{n}\binom{n}{i}p^i(1-p)^{n-i}\frac{\sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^i}{m^i}$. □

Задача 3.4. $-5/6$. □

Задача 3.5. Сл.в. X има *равномерно разпределение*, т.е.

$$\mathbb{P}(X = x_i) = \frac{1}{n}, \qquad \text{за всяко } i = 1, \ldots, n.$$

Тогава (а) $\mathbb{E}[X] = \sum_{i=1}^{n}\frac{i-1}{n-1}\frac{1}{n} = \frac{\sum_{j=1}^{n-1}j}{n(n-1)} = \frac{1}{2}$ и $\mathbb{E}[X^2] = \sum_{i=1}^{n}\left(\frac{i-1}{n-1}\right)^2\frac{1}{n} = \frac{\sum_{j=1}^{n-1}j^2}{n(n-1)^2} = \frac{2n-1}{6(n-1)}$, 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 $\mathrm{Var}(X) = (b-a)^2\frac{n+1}{12(n-1)}$. □

Задача 3.6. Нека сл.в. X описва момента на печалба на играча. Тогава $\mathbb{P}(X = n) = 2^{-n}$, за $n = 1, 2, \ldots$, откъдето $\mathbb{E}[2^X - A] = \sum_{n=1}^{\infty}(2^n - A)2^{-n} = \infty$, 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 3.7. Нека сл.в. X описва момента на печалба. Тогава печалбата на играча е $Y = 2^X A - (A + 2A + \cdots + 2^{X-1}A) = 2^X A - (2^X - 1)A = A$ *константна сл.в.*, откъдето $\mathbb{E}[Y] = A$. Играта обаче предполага, че играчът разполага с безкраен времеви хоризонт и богатство, докато залозите нарастват експоненциално.

Ако разполагаме с 127 лв. и $A = 1$ лв., тогава имаме пари за най-много 6 залагания, като

X	1	2	3	4	5	6	7	8
$\mathbb{P}_X$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{128}$

,

където $\{X = 8\} = \{\text{фалит}\}$. Както по-горе, печалбата винаги е $Y = 1$, ако $X \leq 7$, но губим $Y = -127$, ако $X = 8$, така че $\mathbb{E}[Y] = 1 \cdot \frac{127}{128} - 127 \cdot \frac{1}{64} = 0$. □

Задача 3.8. Лесно се съобразява, че $\mathbb{P}(X = 2n) = 5\frac{2^{n-1}}{3^{2n}}$ и $\mathbb{P}(X = 2n+1) = 6\frac{2^{n-1}}{3^{2n+1}}$, за $n = 1, 2, \ldots$. Използвайки формулите $\sum_{n=1}^{\infty}nx^n = \frac{x}{(1-x)^2}$ и $\sum_{n=1}^{\infty}n^2x^n = \frac{x(x+1)}{(1-x)^3}$, за $|x| < 1$, получаваме

$$\mathbb{E}[X] = \sum_{n=1}^{\infty}2n\,\mathbb{P}(X = 2n) + \sum_{n=1}^{\infty}(2n+1)\,\mathbb{P}(X = 2n+1) = 7\sum_{n=1}^{\infty}n\left(\frac{2}{9}\right)^n + \sum_{n=1}^{\infty}\left(\frac{2}{9}\right)^n = \frac{20}{7},$$

$$\mathbb{E}[X^2] = \sum_{n=1}^{\infty}(2n)^2\,\mathbb{P}(X = 2n) + \sum_{n=1}^{\infty}(2n+1)^2\,\mathbb{P}(X = 2n+1) = 14\sum_{n=1}^{\infty}n^2\left(\frac{2}{9}\right)^n + 4\sum_{n=1}^{\infty}n\left(\frac{2}{9}\right)^n + \sum_{n=1}^{\infty}\left(\frac{2}{9}\right)^n = \frac{482}{49},$$

откъдето $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 82/49$. □

Задача 3.9. $\mathbb{P}(Y > 2|X = 1) = 1/2$, $\mathbb{P}(Y = 3|X < 3) = 1/3$,

$X\backslash Y$	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

Забележете, че въведохме $Y = 6$, понеже събитието (Ч, Ч, Б, Б, Б) има ненулева вероятност и не бе отчетено никъде първоначално, т.е. $\bigcup_{i=1,2,3,\, j=2,3,4,5}\{X = i, Y = j\} \subsetneq \Omega$. ☐

Задача 3.10. Ако сл.в. X и Y описват броя на езитата, хвърлени съответно от А и Б, тогава $\mathbb{P}(X > Y) = 1/2$ и $\mathbb{P}(Y = 1|X > Y) = 1/2$. Печалбата на А е $f(X,Y) = (5-3)\cdot \mathbb{1}_{\{X>Y\}} + (0-3)\cdot \mathbb{1}_{\{X\le Y\}}$, откъдето

$$\mathbb{E}[f(X,Y)] = 2\cdot \mathbb{P}(X > Y) - 3\cdot \mathbb{P}(X \le Y) = -0.5.$$

Понеже играта е от вида *"zero-sum game"*, очакваната печалба на Б е 0.5. ☐

Задача 3.11.

$2X+Y+1$	0	2	4	6	8
$\mathbb{P}_{2X+Y+1}$	0.25	0.125	0.375	0.125	0.125

XY	-5	-3	-1	1	3	5
$\mathbb{P}_{XY}$	0.125	0.125	0.25	0.25	0.125	0.125

$\mathbb{E}[2X + Y + 1] = 7/2$, $\mathrm{Var}(2X + Y + 1) = 27/4$; $\mathbb{E}[XY] = 0$, $\mathrm{Var}(XY) = 9$. ☐

Задача 3.12. Търсим $(k, n) \in \mathbb{N}^2$, за които е изпълнено $k \le n$ и

$$\binom{n-1}{k-1}p^{n-k}(1-p)^k = \mathbb{P}(E \cap F_k) = \mathbb{P}(E)\mathbb{P}(F_k) = p\binom{n}{k}p^{n-k}(1-p)^k,$$

т.е., след съкращаване, $p = k/n$. От $k, n \in \mathbb{N}$ следва, че $p = l/r$ е рационално число за някои $l, r \in \mathbb{N} : l \le r$. Тогава възможните отговори са $(k, n) \in \{(ml, mr), m \in \mathbb{N}\}$. ☐

Задача 3.13. Лесно се вижда, че $\mathbb{P}(X = k) = V_3^k V_5^1 / V_8^{k+1}$, за $k = 0, 1, 2, 3$. Тогава $\mathbb{E}[X] = 0.5$ и $\mathrm{Var}(X) = 0.54$.

Относно последния въпрос, нека $X_1, \dots, X_{1000}$ са i.i.d. копия на X. Полагаме $Y := \sum_{i=1}^{1000} X_i$. 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.

Забележете, че неравенството на Марков ни дава в случая $\mathbb{P}(Y > 900) \le \frac{\mathbb{E}[Y]}{900} = \frac{5}{9}$, което обаче е крайно незадоволително като оценка. Въпреки това неравенството на Чебишев не е задължително да дава по-точна горна оценка от неравенството на Марков в общия случай. ☐

Задача 3.14. Нека сл.в. X_1 и X_2 са точките, хвърлени съответно на първия и втория зар, като $X_1 \perp\!\!\!\perp X_2$ и $X_1 \overset{d}{=} X_2$. Полагаме $X := X_1 + X_2$ и $Y := X_1 X_2$. Тогава

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

За втория въпрос приложете неравенството на Чебишев. ☐

Задача 3.15. $\mathbb{P}(|X - \mathbb{E}[X]| \le 1000) \ge 0.98$. ☐

Задача 3.16. $\mathbb{E}[X] = -1/10$, $\mathbb{E}[Y] = 7/10$, $\mathrm{Var}(X) = 49/100$, $\mathrm{Var}(Y) = 21/100$; $\mathrm{Cov}(X,Y) = 7/100$, $\rho_{X,Y} = \sqrt{21}/21$; $\mathbb{E}[Z] = -17/10$, $\mathrm{Var}(Z) = 441/100$. ☐

Задача 3.17. $\mathrm{Corr}(X,Z)=0$,

$Z\backslash X$	0	1	2	
0	1/9	0	2/9	1/3
1	2/9	1/9	0	1/3
2	0	2/9	1/9	1/3
	1/3	1/3	1/3	

.

□

Задача 3.18. В следствие на симетрията в задачата, $\mathrm{Var}(Y)=\mathrm{Var}(X)$. От друга страна $X+Y=2$, откъдето получаваме

$$\mathrm{Cov}(X,Y)=\mathrm{Cov}(X,2-X)=\mathrm{Cov}(X,-X)=-\mathrm{Var}(X).$$

Следователно $\rho_{X,Y}=-1$. □

Задача 3.19.

$$\begin{aligned}\mathrm{Cov}(X+Y,Z)&=\mathbb{E}[(X+Y)Z]-\mathbb{E}[X+Y]\mathbb{E}[Z]\\&=\mathbb{E}[XZ]+\mathbb{E}[YZ]-\mathbb{E}[X]\mathbb{E}[Z]-\mathbb{E}[Y]\mathbb{E}[Z]=\mathrm{Cov}(X,Z)+\mathrm{Cov}(Y,Z).\end{aligned}$$

□

Задача 3.20. Нека сл.в. X_i е хвърленото число на i-тия ход, за $i=1,\dots,n$. 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

$$\begin{aligned}\mathrm{Cov}(X,Y)&=\mathrm{Cov}(X_1+Z,Z+X_n)\\&=\mathrm{Cov}(Z,Z)+\mathrm{Cov}(X_1,Z)+\mathrm{Cov}(X_1,X_n)+\mathrm{Cov}(Z,X_n)=\mathrm{Var}(Z)=(n-2)\mathrm{Var}(X_1).\end{aligned}$$

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

Условно очакване

Задача 3.21. $\mathbb{P}(X=Y)=\frac{3}{8}$, $\mathbb{P}(X>1|Y=1)=\frac{1}{2}$, $\mathbb{P}(X+Y>2|X=2)=\frac{5}{6}$,

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 3.22.

X	-1	0	4
$\mathbb{P}_{X\|Y=2}$	1/2	1/4	1/4
$\mathbb{P}_{X\|Y=5}$	1/6	1/2	1/3

Y	2	5
$\mathbb{P}_{Y\|X=-1}$	2/3	1/3
$\mathbb{P}_{Y\|X=0}$	1/4	3/4
$\mathbb{P}_{Y\|X=4}$	1/3	2/3

$\mathbb{E}[X\|Y]$	1/2	7/6
$\mathbb{P}$	2/5	3/5

$\mathbb{E}[Y\|X]$	3	4	17/4
$\mathbb{P}$	3/10	3/10	2/5

□

Задача 3.23. Използвайки свойствата на условното очакване,

$$\mathbb{E}\big[\textstyle\sum_{i=1}^{T}X_i\big]=\mathbb{E}\big[\mathbb{E}\big[\sum_{i=1}^{T}X_i|T\big]\big]=\sum_{k=1}^{\infty}\sum_{i=1}^{k}\mathbb{E}[X_i|T=k]\mathbb{P}(T=k)=\sum_{k=1}^{\infty}k\,\mathbb{E}[X_1]\mathbb{P}(T=k)=\mathbb{E}[T]\cdot\mathbb{E}[X_1].$$

□

Задача 3.24. Нека сл.в. Y е броят на белите топки в третата урна. Тогава

$$\mathbb{E}[X]=\mathbb{E}\big[\mathbb{E}[X|Y]\big]=\sum_{i=1}^{3}\mathbb{E}[X|Y=i]\mathbb{P}(Y=i)=3\cdot\frac{6}{10}\cdot\frac{3}{10}+\frac{3}{2}\cdot\Big(\frac{4}{10}\cdot\frac{3}{10}+\frac{6}{10}\cdot\frac{7}{10}\Big)+1\cdot\frac{4}{10}\cdot\frac{7}{10}=\frac{163}{100}.$$

□

Задача 3.25. Нека сл.в. X е хвърленото число, а Y – броят шестици на зарчето. Тогава от биномната формула

$$\mathbb{P}(X=6)=\sum_{i=0}^{6}\mathbb{P}(X=6|Y=i)\mathbb{P}(Y=i)=\sum_{i=0}^{6}\frac{i}{6}\binom{6}{i}\Big(\frac{1}{6}\Big)^i\Big(\frac{5}{6}\Big)^{6-i}=\sum_{j=0}^{6-1}\binom{6-1}{j}\Big(\frac{1}{6}\Big)^i\Big(\frac{5}{6}\Big)^{6-i}=\frac{1}{6}.$$

По втория въпрос нека сл.в. T е броят хвърляния до 6-ца. Тогава $T|Y=i\sim\mathrm{Geo}(i/6)$, откъдето

$$\mathbb{E}[T]=\mathbb{E}\big[\mathbb{E}[T|Y]\big]=\sum_{i=0}^{n}\mathbb{E}[T|Y=i]\mathbb{P}(Y=i)\geq\mathbb{E}[T|Y=0]\mathbb{P}(Y=0)=\infty.$$

□

Задача 3.26. Алтернативно на Задача 3.25, нека $U_1,\dots,U_n$ и $N_1,\dots,N_6$ са две независими редици от независими сл.в., такива че $\mathbb{P}(U_i=k)=\mathbb{P}(N_j=k)=1/6$ за $i=1,\dots,n$ и $j,k=1,\dots,6$. Чрез $X_i=N_{U_i}$ ще моделираме хвърленото число на i-тия ход, като векторът $(N_1,\dots,N_6)$ представлява разгъвката на зара, а U_i показва страната на зара, която се е паднала на i-тото хвърляне (в Задача 3.20 имаме просто $X_i=U_i$). Тогава за $k=1,\dots,6$ имаме отново, че

$$\mathbb{P}(X_i=k)=\mathbb{P}(N_{U_i}=k)=\sum_{j=1}^{6}\mathbb{P}(N_j=k,U_1=j)=\frac{1}{6}\sum_{j=1}^{6}\mathbb{P}(N_j=k)=\mathbb{P}(N_1=k)=\frac{1}{6};$$

т.е. $X_1,\dots,X_n$ са еднакво разпределени като при правилен зар, откъдето $\mathbb{E}[X_1]=3.5$ и $\mathrm{Var}(X_1)=2.917$. Въпреки това, за разлика от Задача 3.20, сл.в. $X_1,\dots,X_n$ не са независими. Всъщност

$$\begin{aligned}\mathrm{Cov}(X_i,X_j)&=\mathbb{E}\big[N_{U_i}N_{U_j}\big]-\mathbb{E}\big[N_{U_i}\big]\mathbb{E}\big[N_{U_j}\big]=\frac{1}{36}\mathbb{E}\Big[\sum_{k=1}^{6}\sum_{l=1}^{6}N_kN_l\Big]-\frac{1}{36}\Big(\mathbb{E}\Big[\sum_{k=1}^{6}N_k\Big]\Big)^2\\&=\frac{1}{36}\mathbb{E}\Big[\sum_{k=1}^{6}N_k^2+\sum_{k\neq l}N_kN_l\Big]-\big(\mathbb{E}[N_1]\big)^2=\frac{1}{6}\mathbb{E}[N_1^2]+\frac{1}{36}\sum_{k\neq l}\mathbb{E}[N_k]\mathbb{E}[N_l]-\big(\mathbb{E}[N_1]\big)^2\\&=\frac{1}{36}\sum_{i=1}^{6}i^2+\frac{30}{36*36}\Big(\sum_{i=1}^{6}i\Big)^2-\frac{1}{36}\Big(\sum_{i=1}^{6}i\Big)^2=\frac{1}{6}\mathrm{Var}(X_1)=0.486>0!\end{aligned}$$

Да положим $Z=\sum_{i=2}^{n-1}X_i$. По условие $X=X_1+Z$ и $Y=Z+X_n$. От тук следва, че

$$\begin{aligned}\mathrm{Cov}(X,Y)&=\mathrm{Cov}(Z,Z)+\mathrm{Cov}(X_1,Z)+\mathrm{Cov}(Z,X_n)+\mathrm{Cov}(X_1,X_n)\\&=\sum_{i=2}^{n-1}\mathrm{Var}(X_i)+2\sum_{2\leq i<j\leq n-1}\mathrm{Cov}(X_i,X_j)+\sum_{j=2}^{n-1}\mathrm{Cov}(X_1,X_j)+\sum_{j=2}^{n-1}\mathrm{Cov}(X_j,X_n)+\mathrm{Cov}(X_1,X_n)\\&=(n-2)\mathrm{Var}(X_1)+(n^2-3n+3)\frac{1}{6}\mathrm{Var}(X_1).\end{aligned}$$

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

Задача 3.27. Нека X_1,X_2,X_3 са печалбите на играчите, а сл.в. T брои хвърлените езита. Тогава $\mathbb{P}(T=m)=\frac{1}{2^{m+1}}$, за $m=0,1,2,\dots$, откъдето, за $k=0,1,2,\dots$,

$$\mathbb{P}(X_1=k)=\sum_{m=k}^{\infty}\mathbb{P}(X_1=k|T=m)\frac{1}{2^{m+1}}=\sum_{m=k}^{\infty}\binom{m}{k}\Big(\frac{1}{3}\Big)^k\Big(\frac{2}{3}\Big)^{m-k}\frac{1}{2^{m+1}}=\frac{1}{2^{k+1}}\sum_{n=0}^{\infty}\binom{n+k}{k}\frac{1}{3^{n+k}}=\frac{3}{4}\Big(\frac{1}{4}\Big)^k,$$

използвайки, че $\sum_{n=0}^{\infty}\binom{n+k}{k}x^n=\frac{1}{(1-x)^{k+1}}$, за $|x|<1$. Следователно $\mathbb{E}[X_1-1]=-\frac{2}{3}$ и $\mathrm{Var}(X_1-1)=\frac{4}{9}$.

Относно втория въпрос, $\mathbb{P}(X_1\geq1,X_2\geq1,X_3\geq1)=1-\mathbb{P}(\bigcup_{i=1}^{3}\{X_i=0\})$, като от принципа за включване-изключване имаме

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{i=1}^{3}\{X_i=0\}\Big)&=\sum_{i=1}^{3}\mathbb{P}(X_i=0)-\sum_{i<j}\mathbb{P}(X_i=0,X_j=0)+\mathbb{P}(X_1=0,X_2=0,X_3=0)\\&=3\cdot\frac{3}{4}-3\sum_{m=0}^{\infty}\mathbb{P}(X_1=0,X_2=0|T=m)\mathbb{P}(T=m)+\mathbb{P}(T=0)=\frac{19}{20}.\end{aligned}$$

□

Характеристики на случайни величини

Задача 3.28. За всяко $x \in \mathbb{R}$,

$$F_{X_{\max}}(x) = \mathbb{P}(X_1 \le x, \ldots, X_n \le x) = \prod_{i=1}^{n} \mathbb{P}(X_i \le x) = \big(F_{X_1}(x)\big)^n,$$

$$F_{X_{\min}}(x) = 1 - \mathbb{P}(X_{\min} > x) = 1 - \mathbb{P}(X_1 > x, \ldots, X_n > x) = 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > x) = 1 - \big(1 - F_{X_1}(x)\big)^n.$$

От друга страна $F_{(X_{\max}, X_1)}(x, y) = \begin{cases} F_{X_1}(y)\big(F_{X_1}(x)\big)^{n-1} & \text{за } y < x, \\ \big(F_{X_1}(x)\big)^n & \text{за } y < x. \end{cases}$ □

Задача 3.29. $\mathbb{E}[X] = 4$, $\mathrm{Var}(X) = 12$, $\mathbb{P}(X = 1) = 1/4$, $\mathbb{P}(X = 2) = 3/16$. □

Задача 3.30. Нека $A = \{$приборът излиза от строя$\}$ и X_1, X_2, X_3 са независими сл.в. със стойности в $\{0, 1\}$, такива че $\mathbb{P}(X_1 = 1) = 0.1$, $\mathbb{P}(X_2 = 1) = 0.2$ и $\mathbb{P}(X_3 = 1) = 0.3$. Тогава $X = X_1 + X_2 + X_3$ е броят изгорели лампи, откъдето $\mathbb{E}[X] = 0.6$. От друга страна $\mathbb{P}(X = k)$, за $k = 0, 1, 2, 3$, е равно на коефициента пред s^k в развитието на

$$g_X(s) = \prod_{i=1}^{3} g_{X_i}(s) = (0.1s + 0.9)(0.2s + 0.8)(0.3s + 0.7) = (0.006s^3 + 0.092s^2 + 0.398s + 0.504),$$

откъдето намираме, че $\mathbb{P}(A) = \sum_{k=1}^{3} \mathbb{P}(A | X = k)\mathbb{P}(X = k) = 0.16.$ □

Основни дискретни разпределения

Задача 3.31. Нека сл. в. X е броят хвърляния на заровете, при които сумата от точките е равна на шест. Тогава $X \sim \mathrm{Bin}(5, \frac{7}{36})$, откъдето $\mathbb{P}(X = 2) = \binom{5}{2}\left(\frac{7}{36}\right)^2\left(\frac{19}{36}\right)^3$ и $\mathbb{E}[X] = \frac{35}{36}$. □

Задача 3.32. За всяко $k = 0, 1, \ldots, n_1 + n_2$,

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

използвайки формулата на Вандермонд; т.е. $X \sim \mathrm{Bin}(n_1 + n_2, p)$. 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 3.33. От свойствата на условните вероятности,

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

□

Задача 3.34. От формулата за пълната вероятност получаваме, за всяко $k = 0, 1, \ldots, M$,

$$\mathbb{P}(X = k) = \sum_{n=k}^{M} \mathbb{P}(X = k | N = n)\mathbb{P}(N = n)$$

$$= \sum_{n=k}^{M} \binom{n}{k} p^k (1-p)^{n-k} \binom{M}{n} q^n (1-q)^{M-n} = \binom{M}{k} p^k q^k \sum_{n=k}^{M} \binom{M-k}{n-k} (q(1-p))^{n-k} (1-q)^{M-n}$$
$$= \binom{M}{k} (pq)^k \big(q(1-p) + (1-q)\big)^{M-k} = \binom{M}{k} (pq)^k \big(1-pq\big)^{M-k}.$$

т.е. $X \sim \mathrm{Bin}(M, pq)$. От друга страна, директно,

$$g_X(s) = \mathbb{E}\big[\mathbb{E}[s^X|N]\big] = \sum_{n=0}^{M} \mathbb{E}[s^X|N=n]\mathbb{P}(N=n) = \sum_{n=0}^{M} \big(sp + (1-p)\big)^n \binom{M}{n} q^n (1-q)^{M-n} = \big(spq + (1-pq)\big)^M.$$

□

Задача 3.35. Нека $X_1, \dots, X_n \overset{i.i.d.}{\sim} \mathrm{Ber}(p)$. Полагаме $p_k = \mathbb{P}(\sum_{i=1}^k X_i \text{ четно})$, за $k = 1, \dots, n$. Тогава

$$p_k = \mathbb{P}\Big(\sum_{i=1}^{k-1} X_i \text{ нечетно}, X_k = 1\Big) + \mathbb{P}\Big(\sum_{i=1}^{k-1} X_i \text{ четно}, X_k = 0\Big) = (1-p_{k-1})p + p_{k-1}(1-p),$$

откъдето извеждаме $p_k - \frac{1}{2} = (1-2p)(p_{k-1} - \frac{1}{2})$ и, чрез заместване, $p_n - \frac{1}{2} = (1-2p)^{n-1}(p_1 - \frac{1}{2})$. Но $p_1 = p$ и $X \overset{d}{=} \sum_{i=1}^n X_i$, откъдето $\mathbb{P}(X = 0, 2, 4, \dots, 2\lfloor \frac{n}{2} \rfloor) = p_n = \frac{1}{2}(1 + (1-2p)^n)$. □

Задача 3.36. (а) За всяко $k = 0, 1, 2, \dots$,

$$\mathbb{P}(\min\{X, Y\} \geq k) = \mathbb{P}(X \geq k, Y \geq k) = (1-p)^k (1-p)^k = (1 - (2p - p^2))^k,$$

т.е. $\min\{X, Y\} \sim \mathrm{Geo}(2p - p^2)$. От друга страна

$$\mathbb{P}(\min\{X, Y\} = k) = \mathbb{P}(X = k, Y > k) + \mathbb{P}(X > k, Y = k) + \mathbb{P}(X = k, Y = k) = \cdots.$$

Относно (б), $\mathbb{P}(\max\{X, Y\} < k) = \mathbb{P}(X < k, Y < k) = (1 - (1-p)^k)^2$. □

Задача 3.37. Нека сл. в. $X \sim \mathrm{Geo}(p)$ брои изстрелите преди уцелването на мишената, където $p := 1 - (1 - p_1)(1 - p_2)$. Тогава $\mathbb{E}[X + 1] = \frac{1}{p}$.

От друга страна, ако $X_1 \sim \mathrm{Geo}(p_1)$ и $X_2 \sim \mathrm{Geo}(p_2)$, то ние се интересуваме от $\mathbb{E}[\min\{X_1, X_2\} + 1]$, където $\min\{X_1, X_2\} \sim \mathrm{Geo}(1 - (1-p_1)(1-p_2))$, подобно на Задача 3.37.

По отношение на втория въпрос, нека сл. в. Y отчита изстрелите, направени от n опита, като

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

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

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

Забележете, че $\mathbb{E}[Y] \approx \mathbb{E}[X]$ за големи n. □

Задача 3.38. Нека X_i е очакваният брой неуспехи, докато на зарче i се падне 6-ца, за $i = 1, 2$. Тогава $X_1, X_2 \overset{i.i.d.}{\sim} \mathrm{Geo}(\frac{1}{6})$. Полагаме $X := \max\{X_1, X_2\}$. От Задача 3.36 и Твърдение 3.11 следва, че

$$\mathbb{E}[X + 1] = \sum_{k=0}^{\infty} \big(1 - \mathbb{P}(X < k)\big) + 1 = \sum_{k=0}^{\infty} \left(1 - \left(1 - \left(1 - \frac{1}{6}\right)^k\right)^2\right) + 1 = \frac{107}{11} \approx 9.73.$$

□

Задача 3.39. Нека сл. в. X_i отчита броя на опаковките до намирането на i-тата уникална играчка, след като вече имаме $i-1$ уникални играчки, за $i = 1, \dots, n$. 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, ако $Y_j \sim \mathrm{Geo}(\frac{1}{10})$ е броят на опаковките до намирането на първата играчка от вид j, за $j = 1, \ldots, n$, тогава $X = \max\{Y_1, \ldots, Y_n\}$. За разлика от Задача 3.36 обаче, тук случайните величини са зависими една от друга.

Относно втория въпрос, нека сл. в. Y брои пакетите до първата повтаряща се играчка. Тогава $\mathbb{P}(Y = k) = \frac{(k-1)V_n^{k-1}}{n^k}$, $k = 2, \ldots, n+1$, откъдето $\mathbb{E}[Y] = \sum_{k=2}^{n+1} k \frac{(k-1)V_n^{k-1}}{n^k}$. □

Задача 3.40. 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 джоб 2 и 1, докато пушачът види, че другия джоб е празен. 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". Понеже $\{X_1 = n-k\} \cap \{X_2 = n-k\} = \emptyset$, имаме

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

□

Задача 3.41. $X \sim \mathrm{HG}(3, 7, 4)$, $\mathbb{E}[X] = \frac{12}{7}$, $\mathrm{Var}(X) = \frac{24}{49}$. □

Задача 3.42. Нека $X \sim \mathrm{Bin}(13, \frac{26}{52})$ и $Y \sim \mathrm{HG}(26, 52, 13)$. (а) $\mathbb{P}(X = 2) = 0.01$; (б) $\mathbb{P}(Y = 2) = 0.004$. □

Задача 3.43. От Задача 2.26 следва, че $X_i \sim \mathrm{Ber}(\frac{M}{N})$, за всяко $i = 1, \ldots, n$. От друга страна, за $i \neq j$,

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

откъдето $\mathrm{Cov}(X_i, X_j) = \frac{M}{N}\frac{M-1}{N-1} - \frac{M}{N}\frac{M}{N} = -\frac{M}{N}\frac{N-M}{N(N-1)}$ и $\mathrm{Corr}(X_i, X_j) = -\frac{1}{N-1}$.

Нека $X := X_1 + \cdots X_n$. Тогава $X \sim \mathrm{HG}(M, N, n)$, $\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = n\frac{M}{N}$ и

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

□

Задача 3.44. За всяко $k = 0, 1, 2, \ldots$,

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

т.е. $X \sim \mathrm{Po}(\lambda_1 + \lambda_2)$. 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 3.45. 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 3.46. Нека $N \sim \mathrm{Po}(\lambda)$, като $X|N = n \sim \mathrm{Bin}(n, p)$. За да има $k = 0, 1, 2, \ldots$ успешни опита, общият брой на опитите трябва да е поне k, откъдето

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

т.е. $X \sim \mathrm{Po}(\lambda p)$. От друга страна, директно,

$$g_X(s) = \mathbb{E}\big[\mathbb{E}[s^X|N]\big] = \sum_{n=0}^\infty \mathbb{E}[s^X|N = n]\mathbb{P}(N = n) = \sum_{n=0}^M \big(sp + (1-p)\big)^n \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda+\lambda(sp-(1-p))} = e^{\lambda p(s-1)}.$$

□

Задача 3.47. $X \sim \mathrm{Bin}(5000, 0.001)$, $\mathbb{P}(X \geq 2) = 1 - (1 - \frac{1}{1000})^{5000} - 5000(1 - \frac{1}{1000})^{4999}\frac{1}{1000} \approx 1 - e^{-5} - 5e^{-5} \approx 0.959$. □

Задача 3.48. (а) $\frac{9!}{(3!)^3}\frac{1}{3^9}$; (б) $3!\frac{9!}{4!3!2!}\frac{1}{3^9}$. □

Задача 3.49. $\frac{5!}{2!2!}\frac{1}{3^5} + 2\frac{5!}{3!2!}\frac{1}{3^5}$. □

Задача 3.50. Нека разгледаме съвместното разпределение на X_1 и X_2. За $k_1, k_2 \in \{0, 1, \dots, n\} : k_1 + k_2 \le n$,

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

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

$$\begin{aligned}\mathbb{E}[X_1X_2] &= \sum_{i=0}^{n}\sum_{j=0}^{n-i} ij\binom{n}{i}\binom{n-i}{j} p_1^i p_2^j (1-p_1-p_2)^{n-i-j}\\ &= np_1p_2 \sum_{i=1}^{n}(n-i)\sum_{j=1}^{n-i}\binom{n-1}{i-1}\binom{n-i-1}{j-1} p_1^{i-1} p_2^{j-1}(1-p_1-p_2)^{n-i-j}\\ &= np_1p_2 \sum_{i=0}^{n-1}(n-i-1)\binom{n-1}{i} p_1^i \Big\{\sum_{j=0}^{n-i-2}\binom{n-i-2}{j} p_2^j (1-p_1-p_2)^{n-i-j-2}\Big\}\\ &= n(n-1)p_1p_2 \sum_{i=0}^{n-1}\binom{n-1}{i} p_1^i (1-p_1)^{n-i-2} - np_1p_2\sum_{i=0}^{n-1} i\binom{n-1}{i} p_1^i (1-p_1)^{n-i-2}\\ &= \frac{n(n-1)p_1p_2}{1-p_1}\sum_{i=0}^{n-1}\binom{n-1}{i} p_1^i(1-p_1)^{n-i-1} - \frac{n(n-1)p_1^2p_2}{1-p_1}\sum_{i=0}^{n-2}\binom{n-2}{i} p_1^i (1-p_1)^{n-i-2}\\ &= \frac{n(n-1)p_1p_2}{1-p_1} - \frac{n(n-1)p_1^2p_2}{1-p_1} = n(n-1)p_1p_2.\end{aligned}$$

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

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

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

Задача 3.51. Нека $(X_1, \dots, X_6) \sim \mathrm{Multi}(n; \frac{1}{6}, \dots, \frac{1}{6})$, където сл. в. X_i отчита броя на хвърлянията, в които на зара са се паднали точно i точки, $i = 1, \dots, 6$. Тогава

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

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

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

откъдето $\mathbb{P}(X_1 + 2X_2 + \dots + 6X_6 = k) = \frac{1}{6^n}\sum_{i=0}^{\lfloor k-n\rfloor/6}(-1)^i\binom{n}{i}\binom{k-6i}{k-n-6i}$ (сравнете със Задача 2.17). □

Задача 3.52. Нека $(X_1, X_2, X_3) \sim \mathrm{Multi}(20; \frac{1}{5}, \frac{1}{5}, \frac{3}{5})$, където сл. в. X_1, X_2 и X_3 са съответно броят на изтеглените бели, червени и зелени топки. От Задача 3.50,

$$\mathbb{P}(X_1 - X_2 = 4) = \sum_{i=4}^{12} \mathbb{P}(X_1 = i, X_2 = i-4) = \sum_{i=4}^{12} \binom{20}{i}\binom{20-i}{i-4}\left(\frac{1}{5}\right)^{2i-4}\left(\frac{3}{5}\right)^{20-2i+4}.$$

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

$$\mathbb{E}[s^{X_1-X_2}] = g_X\left(s, s^{-1}, 1\right) = \left(\frac{1}{5}s + \frac{1}{5s} + \frac{3}{5}\right)^{20} = \frac{1}{5^{20}}\sum_{n=0}^{20}\sum_{m=0}^{20-n}\frac{20!}{m!n!(20-m-n)!}s^{n-m}3^{20-n-m},$$

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 3.53. Нека $X = (X_0, X_1, \ldots, X_9) \sim \mathrm{Multi}(3; \frac{1}{10}, \ldots, \frac{1}{10})$ и $Y = (Y_0, Y_1, \ldots, Y_9) \sim \mathrm{Multi}(3; \frac{1}{10}, \ldots, \frac{1}{10})$, 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

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

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). □

5.4 Непрекъснати случайни величини

Основни понятия

Задача 4.1. Нека $x \in (0, R)$. От геометрични съображения $F_X(x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}$, което съответства на вероятността A да попадне в кръга с център O и радиус x. Освен това $F_X(x) = 0$ за $x \le 0$, и $F_X(x) = 1$ за $x \ge R$. Тогава F_X е абсолютно непрекъсната функция, а в часност – непрекъснато диференцируема (освен в т. $x = R$), откъдето следва, че X е непрекъсната сл. в. с плътност

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{2x}{R^2}\cdot \mathbb{1}_{(0,R)}(x).$$

□

Задача 4.2.

$$F_X(0) = 1 - \int_0^{\infty} f_X(x)dx = 1 + \int_0^{-\infty} f_X(-y)dy = 1 - F_X(0) \qquad \Longrightarrow \qquad F_X(0) = 0.5,$$

$$F_X(-x) = \int_{-\infty}^{-x} f_X(t)dt = \int_x^{\infty} f_X(u)du = 1 - F_X(x),$$

$$\mathbb{P}(|X| \le x) = \mathbb{P}(-x \le X \le x) = \mathbb{P}(-x < X \le x) = F_X(x) - F_X(-x) = 2\,F_X(x) - 1.$$

□

Задача 4.3. От Твърдение 4.1 следва, че съответните сл. в. са непрекъснати и имат плътности

(а) $f_{-X}(y) = f_X(-y)\cdot|-1|\cdot\mathbb{1}_{(-\infty,0)}(y) = \lambda e^{\lambda y}\cdot\mathbb{1}_{(-\infty,0)}(y)$;

(б) $f_{2X-1}(y) = f_X(\frac{y+1}{2})\cdot|\frac{1}{2}|\cdot\mathbb{1}_{(-1,\infty)}(y) = \frac{1}{2}\lambda e^{-\lambda\frac{y+1}{2}}\cdot\mathbb{1}_{(-1,\infty)}(y)$;

(в) $f_{\sqrt{X}}(x) = f_X(y^2)\cdot|2y|\cdot\mathbb{1}_{(0,\infty)}(y) = 2\lambda y e^{-\lambda y^2}\cdot\mathbb{1}_{(0,\infty)}(y)$.

□

Задача 4.4. За всяко $x > 0$ имаме

$$\mathbb{P}(X^2 \le x) = \int_{-\sqrt{x}}^{0} f_X(t)dt + \int_{0}^{\sqrt{x}} f_X(t)dt = \int_0^x \frac{f_X(\sqrt{u}) + f_X(-\sqrt{u})}{2\sqrt{u}} du,$$

където използваме смяната $u = t^2$, така че $t = -\sqrt{u}$ при $t < 0$, и $t = \sqrt{u}$ при $t > 0$. Следователно X^2 е непрекъсната сл. в. с плътност $f_{X^2}(x) = \frac{f_X(\sqrt{x})+f_X(-\sqrt{x})}{2\sqrt{x}} \cdot \mathbb{1}_{(0,\infty)}(x)$. □

Задача 4.5. (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) $\mathbb{P}\big(X \le \mathbb{E}[X]\big) \approx 0.436$; (г) $\mathbb{E}[X^2 + 3X] = \frac{207}{80}$. □

Колекция от случайни величини

Задача 4.6. 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, че $\mathbb{1}_{D_{X,Y}}(x,y) = \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(x,1)}(y) \equiv \mathbb{1}_{(0,y)}(x) \cdot \mathbb{1}_{(0,1)}(y)$. От тук следва, че

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = 8x\Big(\int_x^1 ydy\Big) \cdot \mathbb{1}_{(0,1)}(x) = 4x(1-x^2) \cdot \mathbb{1}_{(0,1)}(x),$$

откъдето $\mathbb{E}[X] = 4\int_0^1 x^2(1-x^2)dx = \frac{8}{15}$, и $f_Y(y) = 4y^3 \cdot \mathbb{1}_{(0,1)}(y)$ 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

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

като използваме наблюдението, че $\mathbb{1}_{D_{X,Y}}(x,y) \cdot \mathbb{1}_B(x,y) = \mathbb{1}_{(0,\frac{3}{4})}(x) \cdot \mathbb{1}_{(x,x+\frac{1}{6})}(y)$.

Относно (г), нека $C = \{(x,y) \in \mathbb{R}^2 : y < x + \frac{1}{3}\}$. Тогава

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

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

Задача 4.7. От една страна имаме

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = \int_0^1\int_0^2 (ax^2 + bxy)dydx = 2\int_0^1 (ax^2 + bx)dx = \frac{2}{3}a + b.$$

От друга страна $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^2 (ax^2 + bxy)dy \cdot \mathbb{1}_{(0,1)}(x) = 2(ax^2 + bx) \cdot \mathbb{1}_{(0,1)}(x)$, откъдето

$$\frac{13}{18} = \mathbb{E}[X] = \int_{-\infty}^{\infty} xf_X(x)dx = 2\int_0^1 x(ax^2 + bx)dx = \frac{1}{2}a + \frac{2}{3}b.$$

Следователно $a = 1$ и $b = \frac{1}{3}$.

Що се отнася до втория въпрос, нека $C = \{(x,y) \in \mathbb{R}^2 : x + y \ge 1\}$. Тогава $\mathbb{1}_C(x,y) = \mathbb{1}_{(1-x,\infty)}(y)$ и

$$\mathbb{P}(X + Y \ge 1) = \mathbb{P}\big((X,Y) \in C\big) = \int_0^1\int_{1-x}^2 \Big(x^2 + \frac{1}{3}xy\Big)dydx \approx 0.45.$$

□

Задача 4.8. За всички $s, t \in \mathbb{R} : s < t$, имаме

$$\begin{aligned}
F_{\max\{X,Y\},\min\{X,Y\}}(t,s) &= \mathbb{P}(\max\{X,Y\} \leq t, \min\{X,Y\} \leq s) \\
&= \mathbb{P}(\max\{X,Y\} \leq t) - \mathbb{P}(\max\{X,Y\} \leq t, \min\{X,Y\} > s) \\
&= \mathbb{P}(X \leq t, Y \leq t) - \mathbb{P}(X \leq t, Y \leq t, X > s, Y > s) \\
&= \mathbb{P}(X \leq t, Y \leq t) - \mathbb{P}(s < X \leq t, s < Y \leq t) \\
&= F_{X,Y}(t,t) - \big(F_{X,Y}(t,t) - F_{X,Y}(t,s) - F_{X,Y}(s,t) + F_{X,Y}(s,s)\big) \\
&= F_{X,Y}(t,s) + F_{X,Y}(s,t) - F_{X,Y}(s,s).
\end{aligned}$$

От друга страна $F_{\max\{X,Y\},\min\{X,Y\}}(t,s) = \mathbb{P}(\max\{X,Y\} \leq t) = F_{X,Y}(t,t)$, за $s, t \in \mathbb{R} : s \geq t$. □

Задача 4.9. По условие $f_{X,Y}(x,y) = \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(0,1)}(y)$, откъдето

$$\begin{aligned}
\mathbb{E}|X - Y| &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - y| f_{X,Y}(x,y) dy dx \\
&= 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - y) \cdot \mathbb{1}_{(0,x)}(y) f_{X,Y}(x,y) dy dx = 2 \int_0^1 \int_0^x (x - y) dy dx = 2 \int_0^1 \Big(x^2 - \frac{x^2}{2}\Big) dx = \frac{1}{3}.
\end{aligned}$$

От друга страна

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[X^2 - 2XY + Y^2] = \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2] = \mathrm{Var}(X) + \mathrm{Var}(Y) + (\mathbb{E}[X] - \mathbb{E}[Y])^2.$$

Но $\mathbb{E}[X] = \mathbb{E}[Y] = \frac{1}{2}$ и $\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{1}{12}$, откъдето $\mathrm{Var}(Z) = \mathbb{E}[(X - Y)^2] - (\mathbb{E}|X - Y|)^2 = \frac{1}{18}$. □

Задача 4.10. По условие $f_{X,Y}(x,y) = 4xy \cdot \mathbb{1}_{D_{X,Y}}(x,y)$, за $D_{X,Y} = (0,1)^2$. Забележете, че $f_{X,Y}$ се разлага на

$$f_{X,Y}(x,y) = \big(2x \cdot \mathbb{1}_{0,1)}(x)\big) \cdot \big(2y \cdot \mathbb{1}_{(0,1)}(y)\big) = f_X(x) \cdot f_Y(y),$$

т.е. $X \perp\!\!\!\perp Y$, откъдето $\mathrm{Cov}(X,Y) = 0$.

Относно втория въпрос, функцията $g(x,y) = (x^2, y^2)$ е дифеоморфизъм върху $D_{X,Y}$ с обратна функция $(u,v) \mapsto (\sqrt{u}, \sqrt{v})$. Освен това $J(u,v) = \det\begin{bmatrix} \frac{1}{2\sqrt{u}} & 0 \\ 0 & \frac{1}{2\sqrt{v}} \end{bmatrix} = \frac{1}{4\sqrt{uv}} \neq 0$, за всички $(u,v) \in g(D_{X,Y}) = (0,1)^2$. Тогава от Твърдение 4.5 следва, че (X^2, Y^2) е вектор от непрекъснати сл. в. с плътност

$$f_{X^2,Y^2}(u,v) = f_{X,Y}(\sqrt{u}, \sqrt{v}) \cdot |J(u,v)| \cdot \mathbb{1}_{(0,1)^2}(u,v) = 4\sqrt{uv} \cdot \left|\frac{1}{4\sqrt{uv}}\right| \cdot \mathbb{1}_{(0,1)^2}(u,v) = 1 \cdot \mathbb{1}_{(0,1)}(u)\mathbb{1}_{(0,1)}(v).$$

*Всъщност $X^2, Y^2 \overset{i.i.d.}{\sim} \mathrm{Unif}(0,1)$, което следва директно от Задача 4.17, след като забележим, че $X^2 = F_X(X)$ и $Y^2 = F_Y(Y)$, където $F_X(t) = F_Y(t) = t^2$ за $t \in (0,1)$. □

Условно очакване

Задача 4.11. $c = \frac{23}{2}$ подобно на Задача 4.6. По условие $f_{X,Y}(x,y) = (\frac{23}{2}x + y) \cdot \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(0,\frac{1-x}{2})}(y)$, откъдето

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy = \int_0^{\frac{1-x}{2}} \Big(\frac{23}{2}x + y\Big) dy \cdot \mathbb{1}_{(0,1)}(x) = \frac{(1-x)(45x+1)}{8} \cdot \mathbb{1}_{(0,1)}(x)$$

и

$$f_{Y|X}(y|x) = \frac{8(\frac{23}{2}x + y)}{(1-x)(45x+1)} \cdot \mathbb{1}_{(0,1)}(x) \cdot \mathbb{1}_{(0,\frac{1-x}{2})}(y).$$

Тогава

$$\mathbb{E}[Y|X = 1/2] = \int_{-\infty}^{\infty} y f_{Y|X}(y|1/2) dy = \int_0^{\frac{1-1/2}{2}} y \frac{8(\frac{23}{4} + y)}{(1 - \frac{1}{2})(\frac{45}{2} + 1)} dy \approx 0.126.$$

Относно последния въпрос можем първо да намерим съвместната плътност на $(X + 2Y, X)$ от Твърдение 4.5 и след това маргиналната плътност на $X + 2Y$,

$$f_{X+2Y}(z) = \int_{-\infty}^{\infty} f_{X+2Y,X}(z,u) du = 3z^2 \cdot \mathbb{1}_{(0,1)}(z).$$

□

Задача 4.12. Нека положим $D_{X,Y} = \{(x,y) \in \mathbb{R}^2 : 0 < x < 3, 0 < y < 2, 2x + 3y \le 6\}$. Използвайки метода за пресмятане на т. нар. "геометрични" вероятности,

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

Понеже $F_{X,Y}$ е абсолютно напрекъсната функция, то (X,Y) е непрекъснат вектор от сл. в. с плътност $f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \frac{1}{3} \cdot \mathbb{1}_{D_{X,Y}}(x,y)$ (в следствие на Фундаменталната теорема за Лебеговия интеграл). Тогава

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y) dx dy = \frac{1}{3} \int_0^3 \int_0^{2-\frac{2}{3}x} xy \, dy dx = \frac{1}{6} \int_0^3 x \Big(2 - \frac{2}{3}x\Big)^2 dx = \frac{1}{2},$$

където използваме, че $\mathbb{1}_{D_{X,Y}}(x,y) = \mathbb{1}_{(0,3)}(x) \cdot \mathbb{1}_{(0,2-\frac{2}{3}x)}(y)$. 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. Разглеждайки само първата координата, X, получваме за функцията ѝ на разпределение, че

$$F_X(x) = \frac{\text{лице трапец с т. } (0,0), (x,0), (x, 2 - \frac{2}{3}x), (0,2)}{\text{лице триъгълник с т. } (0,0), (3,0), (0,2)} = \frac{2}{3}x - \frac{1}{9}x^2 \qquad \text{за } x \in (0,3),$$

откъдето $f_X(x) = (\frac{2}{3} - \frac{2}{9}x) \cdot \mathbb{1}_{(0,3)}(x)$. От друга страна, при фиксирана първа координата x, стойността на втората координата, Y, по условие, може да варира случайно в интервала $(0, 2 - \frac{2}{3}x)$, т.е. $Y \mid X = x \sim \mathrm{Unif}(0, 2 - \frac{2}{3}x)$ или

$$f_{Y|X}(y|x) = \frac{1}{2 - \frac{2}{3}x} \cdot \mathbb{1}_{(0,2-\frac{2}{3}x)}(y).$$

Тогава

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = \frac{1}{3} \cdot \mathbb{1}_{(0,3)}(x) \mathbb{1}_{(0,2-\frac{2}{3}x)}(y) = \frac{1}{3} \cdot \mathbb{1}_{D_{X,Y}}(x,y).$$

□

Забележка (Друг поглед върху геометричните вероятности). Забележете, че положителните стойности на $f_{X,Y}$ лежат на разстояние $\frac{1}{3}$ от равнината $(x,y,0)$ в $\mathbb{R}^3$, като обема $1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy dx$ е равен на обема на триъгълна призма с височина $\frac{1}{3}$ и основа триъгълника $D_{X,Y}$. Аналогично на едномерния случай, "плоската" форма на $f_{X,Y}$ показва, че точките (X,Y) са разпределени равномерно в $D_{X,Y}$. С други думи, пресмятането на "геометрични" вероятности се свежда до пресмятане на вероятности на равномерно разпределен непрекъснат вектор от сл. в.

Задача 4.13. Нека $X = |OA|$ и $C = \{\text{В лежи в окръжността с център А и радиус } R - X\}$. От Задача 4.1 знаем, че $X \sim f_X(x) = \frac{2x}{R^2} \cdot \mathbb{1}_{(0,R)}(x)$. Тогава търсената вероятност е

$$\mathbb{P}(C) = \int_{-\infty}^{\infty} \mathbb{P}(C|X = x) f_X(x) dx = \int_0^R \frac{\pi(R-x)^2}{\pi R^2} \frac{2x}{R^2} dx = \frac{1}{6}.$$

□

Основни непрекъснати разпределения

Задача 4.14. 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

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

□

Задача 4.15. По условие
$$Y = \min\{X, 5\} = X \cdot \mathbb{1}_{\{X \le 5\}} + 5 \cdot \mathbb{1}_{\{X > 5\}},$$
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 $\text{Var}(Y) \approx 2.764$.

Относно втория въпрос, нека $Z \sim \text{Bin}(1000, \mathbb{P}(X < 5))$. Тогава $\mathbb{E}[Z] \approx 714.29$. □

Задача 4.16. По условие $X = \tan\Theta$ за $\Theta \sim \text{Unif}(-\frac{\pi}{2}, \frac{\pi}{2})$, така че от Твърдение 4.1
$$f_X(x) = f_\Theta(\arctan x)(\arctan x)' = \frac{1}{\pi(1+x^2)} \cdot \mathbb{1}_{(-\infty,\infty)}(x);$$
т.е. X има **разпределение на Коши** с параметри $(0, 1)$. □

*Забележка** (Тежки опашки). От горното следва, че $\mathbb{P}(X > t) = \int_t^\infty \frac{1}{\pi(1+x^2)}dx \approx \int_t^\infty \frac{1}{\pi x^2}dx = \frac{1}{\pi t}$ за големи t, т.е. "опашката" на разпределението на X намалява полиномиално. Нещо повече,
$$\mathbb{E}|X| = \int_{-\infty}^\infty \frac{|x|}{\pi(1+x^2)}dx = 2\int_0^\infty \frac{x}{\pi(1+x^2)}dx = \frac{1}{\pi}\int_0^\infty \frac{1}{u}du = \frac{1}{\pi}\ln u\Big|_0^\infty = \infty,$$
откъдето получаваме, че $\mathbb{E}[X]$ (и следователно всеки друг момент на X) не съществува. Това е характерно свойство на сл. в., чиито вероятностни разпределения имат *тежки опашки*, т.е. X, такова че
$$\mathbb{P}(X > t) \sim t^{-\alpha}, \qquad \text{при } t \to \infty, \text{ за } \alpha > 0.$$
В този случай може да се покаже, че $\mathbb{E}[X^k]$ не съществува за всички $k > \alpha - 1$.

Задача 4.17. Понеже F_X е строго нарастваща и непрекъсната функция, последното от факта, че $\mathbb{P}(X = x) = 0$ за всяко $x \in \mathbb{R}$, то F_X е обратима върху $\mathbb{R}$, при което
$$\mathbb{P}(F_X(X) \le u) = \mathbb{P}\big(X \le F_X^{-1}(u)\big) = F_X\big(F_X^{-1}(u)\big) = u, \qquad \text{за } u \in (0,1),$$
т.е. $F_X(X) \stackrel{d}{=} U$; и
$$\mathbb{P}(F_X^{-1}(U) \le x) = \mathbb{P}(U \le F_X(x)) = F_X(x), \qquad \text{за } x \in \mathbb{R},$$
т.е. $F_X^{-1}(U) \stackrel{d}{=} X$. □

Задача 4.18. От Твърдение 4.6 следва, че $X + Y$ е непрекъсната сл. в. с плътност
$$\begin{aligned} f_{X+Y}(z) &= \int_{-\infty}^\infty f_Y(z-x)f_X(x)dx = \int_{-\infty}^\infty \mathbb{1}_{(0,1)}(x)\cdot\mathbb{1}_{(0,1)}(z-x)dx = \int_{-\infty}^\infty \mathbb{1}_{(0,1)}(x)\cdot\mathbb{1}_{(z-1,z)}(x)dx \\ &= \int_{-\infty}^\infty \mathbb{1}_{(\max\{0,z-1\},\min\{1,z\})}(x)dx = \int_{-\infty}^\infty \mathbb{1}_{(0,z)}(x)dx\cdot\mathbb{1}_{(0,1)}(z) + \int_{-\infty}^\infty \mathbb{1}_{(z-1,1)}(x)dx\cdot\mathbb{1}_{[1,2)}(z) \\ &= \int_0^z dx\cdot\mathbb{1}_{(0,1)}(z) + \int_{z-1}^1 dx\cdot\mathbb{1}_{[1,2)}(z) = z\cdot\mathbb{1}_{(0,1)}(z) + (2-z)\cdot\mathbb{1}_{[1,2)}(z). \end{aligned}$$
□

Задача 4.19. $\mathbb{P}(2 \le X \le 8) = \frac{3}{5}$, $\mathbb{E}[X] = 5$ и $\text{Var}(X) = \frac{10^2}{12}$, докато от неравенството на Чебишев получаваме
$$\mathbb{P}(2 \le X \le 8) = \mathbb{P}(-3 \le X - 5 \le 3) = \mathbb{P}\big(|X - \mathbb{E}[X]| \le 3\big) = 1 - \mathbb{P}\big(|X - \mathbb{E}[X]| > 3\big) \ge 1 - \frac{\text{Var}(X)}{9} = \frac{2}{27}.$$
□

Задача 4.20. Нека $X \sim \text{Unif}(0,1)$ е дължината на първата отсечка, а $Y|X = x \sim \text{Unif}(0, 1-x)$ е дължината на втората отсечка върху остатъка от OA, който е с дължина $1 - X$. Дефинираме
$$C = \big\{(x,y) \in \mathbb{R}^2 : x + y > 1 - x - y, x + 1 - x - y > y, y + 1 - x - y > x\big\}.$$
Тогава $\mathbb{1}_C(x,y) = \mathbb{1}_{(0,\frac{1}{2})}(x)\cdot\mathbb{1}_{(\frac{1}{2}-x,\frac{1}{2})}(y)$, откъдето търсената вероятност е
$$\mathbb{P}\big((X,Y) \in C\big) = \int_{-\infty}^\infty\int_{-\infty}^\infty \mathbb{1}_C(x,y)f_{Y|X}(y|x)f_X(x)dydx = \int_0^{\frac{1}{2}}\int_{\frac{1}{2}-x}^{\frac{1}{2}} \frac{1}{1-x}dydx = \ln 2 - \frac{1}{2}.$$
□

Задача 4.21. Относно (а), за всяко $u \in (0,1)$,

$$F_{U_{(3)}}(u) = \mathbb{P}(\max\{U_1, U_2, U_3\} \leq u) = \mathbb{P}(U_1 \leq u, U_2 \leq u, U_3 \leq u) = \prod_{i=1}^{3} \mathbb{P}(U_i \leq u) = u^3,$$

$F_{U_{(3)}}(u) = 0$ за $u \leq 0$, и $F_{U_{(3)}}(u) = 1$ за $u \geq 1$. Понеже $F_{U_{(3)}}$ е абсолютно непрекъсната функция, а в частност – непрекъснато диференцируема (освен в т. $u = 1$), то $U_{(3)}$ е непрекъсната сл. в. с плътност

$$f_{U_{(3)}}(u) = \frac{d}{du} F_{U_{(3)}}(u) \cdot \mathbb{1}_{(0,1)}(u) = 3u^2 \cdot \mathbb{1}_{(0,1)}(u).$$

По същия начин, за всяко $u \in (0,1)$,

$$F_{U_{(1)}}(u) = 1 - \mathbb{P}(\min\{U_1, U_2, U_3\} > u) = 1 - \mathbb{P}(U_1 > u, U_2 > u, U_3 > u) = 1 - \prod_{i=1}^{3} \mathbb{P}(U_i > u) = 1 - (1-u)^3,$$

откъдето $U_{(1)}$ е непрекъсната сл. в. с плътност $f_{U_{(1)}}(u) = \frac{d}{du} F_{U_{(1)}}(u) \cdot \mathbb{1}_{(0,1)}(u) = 3(1-u)^2 \cdot \mathbb{1}_{(0,1)}(u)$.

Относно (б), нека $u \in (0,1)$. Тогава

$$F_{U_{(2)}}(u) = \mathbb{P}(U_{(2)} \leq u) = 6 \cdot \mathbb{P}(U_2 \leq u, U_1 < U_2 < U_3) = 6 \int_0^u \Big[\int_0^{u_2} \Big(\int_{u_2}^1 du_3 \Big) du_1 \Big] du_2 = 6 \int_0^u u_2(1-u_2) du,$$

откъдето $U_{(2)}$ е непрекъсната сл. в. с плътност $f_{U_{(2)}}(u) = 6u(1-u) \cdot \mathbb{1}_{(0,1)}(u)$.

По отношение на (в) можем да подходим стандартно, използайки Твърдение 4.1. Алтернативно, нека $X \sim \text{Exp}(1)$, така че $F_X(x) = 1 - e^{-\lambda x}$, за $x \in (0, \infty)$. Тогава $F_X^{-1}(u) = -\ln(1-u)$, за $u \in (0,1)$. От Задача 4.17 следва, че

$$-\ln(1 - U_1) = F_X^{-1}(U) \sim \text{Exp}(1).$$

Но $U_1 \stackrel{d}{=} 1 - U_1$ по симетрия, откъдето $-\ln(U_1) \stackrel{d}{=} -\ln(1-U_1)$. □

Задача 4.22. Нека $A = \{$чака на първа опашка$\}$, $X_1 \sim \text{Exp}(\frac{1}{8})$ и $X_2 \sim \text{Exp}(\frac{1}{5})$. Времето, прекарано от клиента на касите, е $X := X_1 \cdot \mathbb{1}_A + X_2 \cdot \mathbb{1}_{A^c}$, откъдето

$$\mathbb{P}(A | X \leq 4) = \frac{\mathbb{P}(X \leq 4 | A)\mathbb{P}(A)}{\mathbb{P}(X \leq 4 | A)\mathbb{P}(A) + \mathbb{P}(X \leq 4 | A^c)\mathbb{P}(A^c)} = \frac{\mathbb{P}(X_1 \leq 4)}{\mathbb{P}(X_1 \leq 4) + \mathbb{P}(X_2 \leq 4)} \approx 0.42.$$

□

Задача 4.23. Нека $X_1 \sim \text{Exp}(\frac{1}{30})$ и $X_2 \sim \text{Exp}(\frac{1}{30})$, като $X_1 \perp\!\!\!\perp X_2$. Времето, прекарано от втория пациент в поликлиниката, е $Y := X_2 + (X_1 - 30) \cdot \mathbb{1}_{\{X_1 > 30\}}$, откъдето

$$\mathbb{E}[Y] = 30 + \int_{30}^{\infty} \frac{1}{30} x e^{-\frac{1}{30}x} dx - 30 \cdot \mathbb{P}(X_1 > 30) = 30 + x e^{-\frac{1}{30}x} \Big|_{30}^{\infty} + \int_{30}^{\infty} e^{-\frac{1}{30}x} dx - 30e^{-1} = 30(1 + e^{-1}).$$

□

Задача 4.24. Нека $X \sim \text{Exp}(\frac{1}{100})$. Сумата, която "Инс 1" ще изплати за един иск, е $Y := X + x - 200 \cdot \mathbb{1}_{\{X > 300\}}$, откъдето

$$\begin{aligned}
\mathbb{P}(Y \leq 200) &= \mathbb{P}(X + x \leq 200, X \leq 300) + \mathbb{P}(X + x - 200 \leq 200, X > 300) \\
&= \mathbb{P}(X \leq 200 - x) + \mathbb{P}(300 < X \leq 400 - x) \cdot \mathbb{1}_{\{x < 100\}}.
\end{aligned}$$

□

Задача 4.25. Относно (а), за всяко $z \in (0, \infty)$,

$$F_{\max\{X_1, \ldots, X_n\}}(z) = \mathbb{P}(X_1 \leq z, \ldots, X_n \leq z) = \mathbb{P}(X_1 \leq z) \cdots \mathbb{P}(X_n \leq z) = (1 - e^{-z})^n,$$

и $F_{\max\{X_1, \ldots, X_n\}}(z) = 0$, за $z \leq 0$. От друга страна, за всяко $z \in (0, \infty)$,

$$F_{\min\{X_1, \ldots, X_n\}}(z) = 1 - \mathbb{P}(\min\{X_1, \ldots, X_n\} > z) = 1 - \mathbb{P}(X_1 > z, \ldots, X_n > z) = 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > z) = 1 - e^{-nz},$$

и $F_{\min\{X_1,\ldots,X_n\}}(z) = 0$, за $z \le 0$; т.е. $\min\{X_1,\ldots,X_n\} \sim \mathrm{Exp}(n)$.

Относно (б), нека $x \in \mathbb{R}$. Тогава

$$F_{Y_n}(x) = \mathbb{P}\big(\max\{X_1,\ldots,X_n\} \le x + \log n\big) = \Big(1 - e^{-(x+\log n)}\Big)^n \longrightarrow e^{-e^{-x}},$$

От друга страна, функцията $F(x) := e^{-e^{-x}}$, $x \in \mathbb{R}$, е ненамаляваща, дясно непрекъсната и удовлетворяваща $\lim_{x\to-\infty} F(x) = 0$ и $\lim_{x\to\infty} F(x) = 1$, от което следва, че F е функция на разпределение; т.е. Y_n се схожда по разпределение към сл. в. с функция на разпределение F. □

Задача 4.26. По условие $f_{X_1,X_2}(x_1,x_2) = \lambda^2 e^{-\lambda(x_1+x_2)} \cdot \mathbb{1}_{D_{X_1,X_2}}(x_1,x_2)$, където $D_{X_1,X_2} = (0,\infty)^2$. Полагаме $Z := X_1$. Функцията $g(x_1,x_2) = (\frac{x_1}{x_1+x_2}, x_1)$ е дифеоморфизъм с обратна функция $(y,z) \mapsto (z, \frac{z}{y} - z)$ и

$$J(y,z) = \det\begin{bmatrix} 0 & 1 \\ -\frac{z}{y^2} & \frac{1}{y} - 1 \end{bmatrix} = \frac{z}{y^2} \neq 0,$$

за всички $(y,z) \in g(D_{X_1,X_2}) = (0,1)\times(0,\infty)$. Тогава от Твърдение 4.5 следва, че (Y,Z) е вектор от непрекъснати сл. в. с плътност

$$f_{Y,Z}(y,z) = f_{X_1,X_2}\Big(z, \frac{z}{y} - z\Big) \cdot |J(y,z)| \cdot \mathbb{1}_{g(D_{X_1,X_2})}(y,z) = \lambda^2 e^{-\lambda\frac{z}{y}}\frac{z}{y^2} \cdot \mathbb{1}_{(0,1)}(y) \cdot \mathbb{1}_{(0,\infty)}(z).$$

Тогава, използвайки смяната $u = \frac{\lambda z}{y}$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y,Z}(y,z)dz = \int_0^\infty \frac{\lambda^2 z}{y^2} e^{-\lambda\frac{z}{y}} dz \cdot \mathbb{1}_{(0,1)}(y) = \int_0^\infty u e^{-u} du \cdot \mathbb{1}_{(0,1)}(y) = \mathbb{1}_{(0,1)}(y);$$

т.е. $Y \sim \mathrm{Unif}(0,1)$.

Алт. решение. Нека $y \in (0,1)$. От Твърдения 4.4 и 4.7 получаваме

$$\begin{aligned} F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}\Big(\frac{1-y}{y}X_1 \le X_2\Big) &= \int_{-\infty}^{\infty} \mathbb{P}\Big(\frac{1-y}{y}X_1 \le X_2 \big| X_1 = x_1\Big) f_{X_1}(x_1)dx_1 \\ &= \int_{-\infty}^{\infty} \mathbb{P}\Big(\frac{1-y}{y}x_1 \le X_2\Big) f_{X_1}(x_1)dx_1 = \int_{-\infty}^{\infty} e^{-\lambda\frac{1-y}{y}x_1}\lambda e^{-\lambda x_1}dx_1 = \int_0^\infty \lambda e^{-\frac{\lambda}{y}x_1}dx_1 = y. \end{aligned}$$

В допълнение, $F_Y(y) = 0$ за $y \le 0$, и $F_Y(y) = 1$ за $y \ge 1$. Тогава F_Y е абсолютно непрекъсната функция (в частност непрекъснато диференцируема освен в т. $y = 0, 1$), откъдето следва, че Y е непрекъсната сл. в. с плътност

$$f_Y(y) = \frac{d}{dy}F_Y(y) \cdot \mathbb{1}_{(0,1)}(y) = 1 \cdot \mathbb{1}_{(0,1)}(y).$$

□

Задача 4.27. Нека $z \in (0,\infty)$. От Твърдения 4.4 и 4.7 получаваме

$$\begin{aligned} F_Z(z) = \mathbb{P}(X \le zY) &= \int_{-\infty}^{\infty} \mathbb{P}(x \le zY | X = x) f_X(x)dx \\ &= \int_{-\infty}^{\infty} \mathbb{P}\Big(\frac{x}{z} \le Y\Big) f_X(x)dx = \int_{-\infty}^{\infty} \mathbb{P}\Big(\frac{x}{z} \le Y\Big) \cdot \mathbb{1}_{(0,3)}\Big(\frac{x}{z}\Big) \cdot f_X(x)dx \\ &= \int_{-\infty}^{\infty} \frac{3 - \frac{x}{z}}{3} \cdot \mathbb{1}_{(0,3z)}(x) \cdot 2e^{-2x} \cdot \mathbb{1}_{(0,\infty)}(x)dx = \int_0^{3z} 2\frac{3 - \frac{x}{z}}{3} e^{-2x}dx = 1 - \frac{1}{6z} + \frac{e^{-6z}}{6z}, \end{aligned}$$

където сме използвали, че $\mathbb{P}(\frac{x}{z} \le Y) = 0$, ако $\frac{x}{z} \ge 3$. В допълнение, $F_Z(z) = 0$ за $z \le 0$. Тогава F_Z е абсолютно непрекъсната функция (в частност непрекъснато диференцируема освен в т. $z = 0$), откъдето следва, че Z е непрекъсната сл. в. с плътност

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \frac{1 - (1+6z)e^{-6z}}{6z^2} \cdot \mathbb{1}_{(0,\infty)}(z).$$

□

Задача 4.28. Нека $X \sim \mathrm{Bin}(2000, \frac{3}{4})$. От Теорема 4.11 (или по-общо от ЦГТ) следва, че

$$\mathbb{P}(1475 \le X \le 1535) = \mathbb{P}\left(\frac{-25}{\sqrt{375}} \le \frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}} \le \frac{35}{\sqrt{375}}\right) \approx \mathbb{P}(-1.29 \le Z \le 1.81) = \Phi(1.81) - \Phi(-1.29) = 0.867,$$

където $Z \sim \mathrm{N}(0,1)$. □

Задача 4.29. От свойствата на пораждащата функция на моментите, за всяко $|t| < 1$,

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

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

Алтернативно, от Твърдение 4.6 получаваме, например за $Z_1, Z_2 \sim \mathrm{N}(0,1)$, че

$$f_{Z_1+Z_2}(t) = \int_{-\infty}^{\infty} f_{Z_2}(t-x) f_{Z_1}(x) dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{(t-x)^2}{2} - \frac{x^2}{2}} dx = \frac{1}{2\pi} e^{-\frac{t^2}{4}} \int_{-\infty}^{\infty} e^{-(x-\frac{t}{2})^2} dx = \frac{1}{\sqrt{4\pi}} e^{-\frac{t^2}{4}},$$

т.е. $Z_1 + Z_2 \sim \mathrm{N}(0,2)$. □

Задача 4.30. Нека $X \sim \mathrm{Exp}(\frac{1}{10})$ и $Y_1, \dots, Y_{1000} \overset{i.i.d.}{\sim} \mathrm{Ber}(p)$, където $p = \mathbb{P}(X > 3) = e^{-\frac{3}{10}}$. Тогава

$$\mathbb{P}\left(\sum_{i=1}^{1000} Y_i \le 750\right) = \mathbb{P}\left(\frac{\sum_{i=1}^{1000} Y_i - 1000p}{\sqrt{1000p(1-p)}} \le 0.66\right) \approx \Phi(0.66) \approx 74.54\%.$$

□

Задача 4.31. Нека $Z \sim \mathrm{N}(0,1)$. Тогава $\frac{X-x}{0.05x} \overset{d}{=} Z$ и $\frac{Y-x}{0.1x} \overset{d}{=} Z$, откъдето

$$\mathbb{P}\big(|X-x| > 0.1x\big) = \mathbb{P}(|Z| > 2) = 2(1-\Phi(2)) \approx 5\% \quad \text{и} \quad \mathbb{P}\big(|Y-x| > 0.1x\big) = \mathbb{P}(|Z| > 1) = 2(1-\Phi(1)) \approx 32\%.$$

Относно втория въпрос $\mathrm{Var}(\frac{X+Y}{2}) = \frac{1}{4}\big(\mathrm{Var}(X) + \mathrm{Var}(Y)\big)$ и $\frac{X+Y}{2} \sim \mathrm{N}(x, 0.003125x^2)$ от Задача 4.29, откъдето

$$\mathbb{P}\left(\left|\frac{X+Y}{2} - x\right| > 0.1x\right) = \mathbb{P}(|Z| > 1.789) = 2(1-\Phi(1.789)) \approx 7.3\%.$$

□

Задача 4.32. Нека $\alpha, \beta, \gamma \in \mathbb{R}_+ : \alpha + \beta + \gamma = 5$ е позицията съответно в актив A, B, C. Тогава

$$\mathbb{E}[\alpha A + \beta B + \gamma C] = \alpha\mathbb{E}[A] + \beta\mathbb{E}[B] + \gamma\mathbb{E}[C] = 3\alpha + 3\beta + \gamma,$$

което е максимално върху $\{(\alpha, 5-\alpha, 0) : \alpha \in [0,5]\}$. В този случай

$$\mathrm{Var}(\alpha A + (5-\alpha)B) = \alpha^2\mathrm{Var}(A) + (5-\alpha)^2\mathrm{Var}(B) = 2\alpha^2 + 3(5-\alpha)^2$$

е максимално за $\alpha = 3$, въпреки че $\mathrm{Var}(A) < \mathrm{Var}(B)$ (диверсификация). По отношение на последния въпрос,

$$\mathbb{P}(3A + 2B \le 5D) = \mathbb{P}(3A + 2B - 5D \le 0) = \mathbb{P}\left(\frac{3A + 2B - 5D - 25}{\sqrt{530}} \le -\frac{25}{\sqrt{530}}\right) = \Phi(-1.09) \approx 14\%,$$

където използваме, че $3A + 2B - 5D \sim \mathrm{N}(25, 530)$ от Задача 4.29. □

Задача 4.33. $\mathbb{P}(X \le 20) \approx 20.3\%$ и $\mathbb{P}(X > x) = \frac{1-\mathbb{P}(X \le 20)}{2}$ за $x \approx 29.38$. □

Задача 4.34. Нека $X_1, X_2, \dots$ са i.i.d. $\mathrm{Ber}(\frac{1}{1000})$. От Теорема 4.11,

$$\mathbb{P}\left(\sum_{i=1}^{100} X_i \ge 1\right) = \mathbb{P}\left(\frac{\sum_{i=1}^{100} X_i - n\mu}{\sqrt{n\sigma^2}} \ge \frac{1 - \frac{1}{10}}{\sqrt{\frac{999}{10000}}}\right) \approx 1 - \Phi(2.85) \approx 0.2\%.$$

Относно втория въпрос, $\rho = \mathbb{E}|X_1 - \mu|^3 \approx 0.001$, откъдето $\frac{\rho}{2\sigma^3\sqrt{n}} = \frac{0.001}{0.032\sqrt{n}} \le 0.001$ за $n \ge 1000$. □

Задача 4.35. По условие $f_{X_1,X_2}(x_1,x_2) = \frac{\beta^{2\alpha}}{(\Gamma(\alpha))^2} x_1^{\alpha-1} e^{-\beta x_1} x_2^{\alpha-1} e^{-\beta x_2} \cdot \mathbb{1}_{D_{X_1,X_2}}(x_1,x_2)$, където $D_{X_1,X_2} = (0,\infty)^2$. Полагаме $Z = X_1 + X_2$. Функцията $g(x_1,x_2) = (\frac{x_1}{x_1+x_2}, x_1 + x_2)$ е дифеоморфизъм с обратна функция $(y,z) \mapsto (yz, z - yz)$. Освен това

$$J(y,z) = \det\begin{bmatrix} z & y \\ -z & 1-y \end{bmatrix} = z \neq 0,$$

за всички $(y,z) \in g(D_{X_1,X_2}) = (0,1) \times (0,\infty)$. Тогава от Твърдение 4.5 следва, че (Y,Z) е вектор от непрекъснати сл. в. с плътност

$$\begin{aligned} f_{Y,Z}(y,z) &= f_{X_1,X_2}(yz, z-yz) \cdot |J(y,z)| \cdot \mathbb{1}_{g(D_{X_1,X_2})}(y,z) \\ &= \frac{\beta^{2\alpha}}{(\Gamma(\alpha))^2} (yz)^{\alpha-1}(z-yz)^{\alpha-1} e^{-\beta yz} e^{-\beta(z-yz)} z \cdot \mathbb{1}_{(0,1)}(y) \cdot \mathbb{1}_{(0,\infty)}(z) \\ &= \frac{\Gamma(2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)} \frac{\beta^{2\alpha}}{\Gamma(2\alpha)} \big(y^{\alpha-1}(1-y)^{\alpha-1} \cdot \mathbb{1}_{(0,1)}(y)\big) \cdot \big(z^{2\alpha-1} e^{-\beta z} \cdot \mathbb{1}_{(0,\infty)}(z)\big). \end{aligned}$$

Понеже съвместната плътност може да бъде факторизирана като $f_{Y,Z}(y,z) = c \cdot h(y) \cdot g(z)$, то $Y \perp\!\!\!\perp Z$. Нещо повече $Y \sim \mathrm{Beta}(\alpha,\alpha)$ и $Z \sim \Gamma(2\alpha,\beta)$. □

Задача 4.36. Използвайки част от разсъжденията в решението на Задача 4.35, се вижда, че $S_n \sim \Gamma(n,1)$. От друга страна $\{N_t \geq n\} = \{S_n \leq t\}$ и $\{N_t = n\} = \{N_t \geq n\} \setminus \{N_t \geq n+1\}$, откъдето

$$\begin{aligned} \mathbb{P}(N_t = n) &= \mathbb{P}(S_n \leq t) - \mathbb{P}(S_{n+1} \leq t) \\ &= \int_0^t \frac{1}{(n-1)!} x^{n-1} e^{-x} dx - \int_0^t \frac{1}{n!} x^{n-1} e^{-x} dx = \frac{1}{n!}\int_0^t e^{-x}(nx^{n-1} - x^n) dx = \frac{1}{n!}\int_0^t d(e^{-x}x^n) = \frac{e^{-t}t^n}{n!}, \end{aligned}$$

т.е. $N_t \sim \mathrm{Po}(t)$. □

6 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.

Billingsley, P. (1995). *Probability and Measure*, 3rd Edition, John Wiley & Sons, Inc., New York.

Grimmett, G. and Stirzaker, D. (2001). *Probability and Random Processes*, 3rd Edition, Oxford University Press, UK.