Вероятности и статистика

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

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

1 Комбинаторика

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

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

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

$$P_n = n!$$

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

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

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

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

$$P_n = |B_1 \cup B_2 \cup \cdots \cup B_{P_{n-1}}| = \sum_{l=1}^{P_{n-1}} |B_l| = n \cdot P_{n-1} = n!$$

□

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*

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

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

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

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

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

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

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

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

От друга страна, всяка пермутация $(a_{i_1}, \dots, a_{i_k}, a_{i_{k+1}}, \dots, a_{i_n})$ на M дефинира една единствена вариация от клас k след премахване на последните $n-k$ елементи, т.е. $(a_{i_1}, \dots, a_{i_k})$. Следователно $\mathbf{P}_n = B_1 \cup B_2 \cup \dots \cup B_{V_n^k}$, откъдето получаваме, че

$$P_n = |B_1 \cup B_2 \cup \dots \cup B_{V_n^k}| = \sum_{l=1}^{V_n^k} |B_l| = P_{n-k} \cdot V_n^k \qquad \Longrightarrow \qquad V_n^k = \frac{P_n}{P_{n-k}}.$$

$\square$

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 обозначаваме с*

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

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

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

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

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

$\square$

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. *(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 $\frac{k}{n}$ от комбинациите на M от клас k, което ги превръща в елементи на $\mathbf{C}_{n-1}^{k-1}$ след премахването 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 $A_i=\{a\}$, $i=1,\dots,n$. Относно (iv), $\mathbf{C}_n^0,\mathbf{C}_n^1,\dots,\mathbf{C}_n^n$ са две по две непресичащи се множества, чието обединение $\mathcal{P}(M):=\mathbf{C}_n^0\cup\mathbf{C}_n^1\cup\dots\cup\mathbf{C}_n^n$ е **множеството на всички подмножества**, образувани от елементите на M. От друга страна, на всяко подмножество $A\subseteq M$ можем да съпоставим една и само една n-орка $(e_1,\dots,e_n)\in\{0,1\}^n$ от нули и единици, такава че $a_i\in A\Longleftrightarrow e_i=1$ за $i=1,\dots,n$, откъдето

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

$\square$

*Забележка** (Триъгълника на Паскал). Установените свойства на числата C_n^k могат да се илюстрират понагледно, 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.

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

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

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

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

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

Д-во.* Ако $(a_{i_1},\dots,a_{i_k})$ е такава k-орка, да означим с $X_1=\{j\in\{1,\dots,k\}:i_j=1\}$ 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 $(a_{i_1^*},\dots,a_{i_{k-k_1}^*})$ е редицата, получена след премахване на всичките k_1 елемента a_1 от $(a_{i_1},\dots,a_{i_k})$. Тогава $(a_{i_1^*},\dots,a_{i_{k-k_1}^*})$ представлява пермутация на $\{a_2,\dots,a_n\}$ с повторение, дължина $k-k_1$ и честоти $(k_2,\dots,k_n)$.

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

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

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

Забележка (Интерпретация). Елемента a_1 можем да поставим по $\binom{k}{k_1}$ различни начина на k_1 от общо k места в редицата. На всеки от тези начини отговарят по $\binom{k-k_1}{k_2}$ различни начина, по които можем да изберем k_2 от останалите $k-k_1$ места, на които да поставим елемента a_2,..., елемента a_n поставяме по $\binom{k-k_1-\dots-k_{n-1}}{k_n}=1$ начин на последните $k-k_1-\dots-k_{n-1}$ 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 $(X_1, \ldots, X_n)$ 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 $(X_1, \ldots, X_n)$ се нарича n-разлагане на E от тип $(k_1, \ldots, k_n)$, като $k_1 + \cdots + k_n = k$. Тогава броят на n-разлаганията на E от тип $(k_1, \ldots, k_n)$ 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 $E = \{1, \ldots, k\}$. Тогава за фиксирани $(k_1, \ldots, k_n)$ съществува еднозначно обратимо съответствие между всички пермутации $(a_{i_1}, \ldots, a_{i_k})$ на $M = \{1, \ldots, n\}$ с повторение, дължина k и честоти $(k_1, \ldots, k_n)$ и всички n-разлагания $(X_1, \ldots, X_n)$ на E от тип $(k_1, \ldots, k_n)$, дефинирано чрез $a_{i_j} = m \iff i_j \in X_m$, за $m = 1, \ldots, n$ и $j = 1, \ldots, k$. С други думи, множеството X_m съдържа информация за местата в редицата $(a_{i_1}, \ldots, a_{i_k})$, на които стои елементът a_m, като знанието на $(X_1, \ldots, X_n)$ е достатъчно, за да възпроизведем $(a_{i_1}, \ldots, a_{i_k})$; 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 обозначаваме с*

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

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

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

Комбинация с повторение. *Всяко k-елементно мулти-подмножество на M, за $k \in \mathbb{N}$, наричаме комбинация на елементите на M от клас k с повторение, като множеството обозначаваме с*

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

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

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

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

От друга страна лесно можем да възстановим $\{a_{i_1}, \ldots, a_{i_k}\}$ от $\{a_{i_1^*}, \ldots, a_{i_k^*}\}$, 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.** *$C(n,k) = \sum_{i=0}^{k} C(n-1,i)$, при което $C(n,k) = C(n,k-1) + C(n-1,k)$.*

*Д-во**.* Нека $\{a_{i_1}, \ldots, a_{i_{k-m}}, a_n, a_n, \ldots, a_n\} \in \mathbf{C}(n,k)$ е k-елементно мулти-подмножество на M, където a_n се среща m пъти, за $m = 0, 1, \ldots, k$. Тогава $\{a_{i_1}, \ldots, a_{i_{k-m}}\} \in \mathbf{C}(n-1, k-m)$ е комбинация на $\{a_1, \ldots, a_{n-1}\}$ 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 $\mathbf{C}(n-1,k), \ldots, \mathbf{C}(n-1,0)$ 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.

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

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

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

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

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

Но $|A_1| = \left|(A_1 \cap A_2) \cup (A_1 \cap A_2^c)\right| = |A_1 \cap A_2| + |A_1 \cap A_2^c|$ 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

$$\begin{aligned}
\left|\bigcup_{i=1}^{n} A_i\right| = \left|\bigcup_{i=1}^{n-1} A_i\right| + |A_n| - \left|\bigcup_{i=1}^{n-1} (A_i \cap A_n)\right| = \sum_{i=1}^{n-1} |A_i| - \sum_{i<j\leq n-1} |A_i \cap A_j| + \cdots + (-1)^{n-2}|A_1 \cap \cdots \cap A_{n-1}| + |A_n| - \\
- \left(\sum_{i=1}^{n-1} |A_i \cap A_n| - \sum_{i<j\leq n-1} |A_i \cap A_j \cap A_n| + \cdots + (-1)^{n-2}|A_1 \cap \cdots \cap A_n|\right) \\
= \sum_{i=1}^{n} |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| + \cdots + (-1)^{n-1}|A_1 \cap \cdots \cap A_n|.
\end{aligned}$$

*Комбинаторно д-во**.* Нека $x \in \bigcup_{i=1}^{n} A_i$, като x е общ елемент на $k \in \{1, \ldots, n\}$ от множествата, да кажем $A_1, \ldots, A_k$. 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 2.7),..., в $\sum_{i_1<i_2<\cdots<i_k} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}|$ – е преброен $\binom{k}{k} = 1$ път, откъдето

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

$\square$

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;
- (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.

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

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

  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 на разпределенията, 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. Колко решения има уравнението $x_1 + \dots + x_k = n$, ако

- (а) $x_1, \dots, x_k$ са естествени числа (за $k \leq n$);
- (б) $x_1, \dots, x_k$ са неотрицателни цели числа.

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

Решение (Техника).

- (а) *Решение 1*: задачата е еквивалентна на разпределянето на n неразличими частици в k клетки 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 $x_1 = \{$брой $*$ преди първата $|\}$, $x_k = \{$брой $*$ след последната $|\}$ и $x_i = \{$брой $*$ между $(i-1)$-тата и i-тата $|\}$, $i = 2, \ldots, k-1$. Т.е., проблемът се свежда до избор на $k-1$ места, на които да поставим “ $|$ ”, избрани от $n+1$ възможни места в редицата от звезди $**\cdots**$. Понеже $x_i > 0$ в (а), не можем да имаме две последователни черти, нито да започваме или завършваме с “ $|$ ”. С други думи, търсим $\{j_1, \ldots, j_{k-1}\}$ за $j_1, \ldots, j_{k-1} \in \{1, \ldots, n-1\} : j_m \neq j_l$, 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 $x_1 + \cdots + x_k = n$ и тези на $(x_1 + 1) + \cdots + (x_k + 1) = n + k$; т.е. C_{n+k-1}^{k-1} от (а).

□

*Задача** 1.3. 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.

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

*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.

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

*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.

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

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

(а) Да се пресметнат $|A_i|$ и $|B_i|$.
(б) Да се докаже комбинаторно, че $\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}$ (*формула на Вандермонд*).
(в) Колко са на брой k-елементните подмножества на M, които съдържат поне един обект от типа a и поне един обект от типа b?
(г) Колко са на брой подмножествата на M, които съдържат поне един обект от типа a и поне един обект от типа b?

Решение. Нека $M^a = \{a_1, \ldots, a_m\}$ и $M^b = \{b_1, \ldots, b_n\}$. (а) $|A_i| = \binom{m}{i}\binom{n}{k-i}$ и $|B_i| = |A_{k-i}|$; (б) множествата $A_0, \ldots, A_k$ са две по две непресичащи се и изброяват всички възможни комбинации на M от клас k, т.е. $\bigcup_{i=0}^{k} A_i = \mathbf{C}_{m+n}^k$, откъдето следва, че $C_{m+n}^k = \sum_{i=0}^{k} |A_i|$; (в) $C_{m+n}^k - |A_0| - |B_0| = \sum_{i=1}^{k-1} |A_i|$; (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. □

*Задача** 1.7. По колко различни начина от $2n$ шахматиста могат да се 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.

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

*Задача*** 1.8. Пресметнете броя на траекториите в правоъгълна координатна система, започващи в точка (x_1, y_1) и завършващи в точка (x_2, y_2), ако правим скокове с големина ± 1. А колко на брой са траекториите от (x_1, y_1) до (x_2, y_2), които не докосват хоризонталата $y = r$?

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

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

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

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

$\square$

2 Вероятности

2.1 Дискретни вероятности

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

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

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.

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

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

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 комбинаториката.

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. $\Omega = \big\{\{a_1, \dots, a_n\} : a_i \in \{ч_1, \dots, ч_M, б_1, \dots, б_{N-M}\}, a_i \neq a_j, i \neq j\big\}$. 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 $M - (N - n) \leq k \leq M$, е

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

Забележете също така, 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 при предположение, 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 двете предположения?

Решение. (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. □

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

Решение. Нека $\Omega = \big\{\{a_1, a_2\} : a_1, a_2 \in \{0, \dots, 6\}\big\}$. Тогава $|\Omega| = C(7, 2)$, т.е. играта се играе с 28 плочки. В този случай, за разлика от задачата със заровете, така описаните елементарни събития са равновероятни, откъдето следва, че $\mathbb{P}(\{\{a_1, a_2\} : a_1 \neq a_2\}) = C_7^2/C(7, 2) = 0.75$. ☐

Задача 2.3. Ако номерата на колите са равномерно разпределени, каква е вероятността номерът на случайна лека кола

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

Решение. Нека $\Omega = \big\{(a_1, \dots, a_4) : a_i \in \{0, 1, \dots, 9\}\big\}$. (а) $V_{10}^4/10^4$. (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 2.4. 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?

Решение (Техника). 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 7C заедно, след това ги подредим по един от описаните начини и накрая запълним останалите места с останалите карти. Тогава

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

Всъщност ни интересува само позицията на 7C между четирите аса (независимо от цветовете им), така че за $\Omega^* = \{(7\text{C}, A, A, A, A), \dots (A, A, A, A, 7\text{C})\}$, елементарните изходи са равно вероятни, откъдето получаваме отново $\mathbb{P}^*(\{\text{Играч 1 печели}\}) = 2/5$ (сравнете със Задача 2.33). ☐

*Задача** 2.5. От партида изделия, от които n са доброкачествени и m – бракувани, за проверка по случаен начин са взети s изделия. При проверката се оказало, че първите k от проверяваните s изделия са доброкачествени $(k < s)$. Да се пресметне вероятността $(k + 1)$-вото изделие да се окаже доброкачествено.

Решение. 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 $\mathbb{P}(\{k + 1\text{-во добро}\}) = \frac{n-k}{n+m-k}$. ☐

*Задача** 2.6. Хвърлят се 10 зара. Каква е вероятността да се паднат равен брой единици и шестици?

Отговор. $\sum_{k=0}^{5} \frac{\binom{10}{k}\binom{10-k}{k}4^{10-2k}}{6^{10}} \equiv \sum_{k=0}^{5} \frac{\binom{10}{2k}\binom{2k}{k,k}4^{10-2k}}{6^{10}}$. ☐

Задача 2.7. От урна, съдържаща топки с номера $1, \dots, n$, се вадят последователно k топки $(k \leq n)$. Каква е вероятността номерата на извадените топки, записани по реда на изваждането, да образуват растяща редица, ако (а) теглим без да връщаме; (б) изтеглените топки се връщат обратно в урната?

Решение (Техника).

(а) $\Omega = \{(a_1, \dots, a_k) : 1 \le a_i \le n, a_i \ne a_j\}$. 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 2.8 (Occupancy Problem). Секретарка написала n писма, сложила ги в пликове и ги запечатала. Забравила кое писмо в кой плик е, но въпреки това написала отгоре n-те различни адреса и изпратила писмата. Каква е вероятността никой да не получи своето писмо?

Решение (Техника). $\Omega = \{(a_1, \dots, a_n) : 1 \le a_i \le n, a_i \ne a_j\}$. Нека $A = \{$никой не е получил своето писмо$\}$. Дефинираме $B_i := \{(a_1, \dots, a_{i-1}, i, a_{i+1}, \dots, a_n) : 1 \le a_j \le n, a_j \ne a_l\}$. Тогава $A = \left(\bigcup_{i=1}^n B_i\right)^c$, $|B_i| = (n-1)!$, $|B_i \cap B_j| = (n-2)!$, $i \ne j$, и т.н., откъдето

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

□

*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, константна вероятност горното събитие да се случи, при $n \to \infty$.

Задача 2.9. Група от n човека се нарежда в редица по случаен начин. Каква е вероятността между две фиксирани лица да има точно $r \le n-2$ човека? А ако се нареждат в кръг и $r < n-2$?

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

Случай 1: за да получим крайната подредба $(a_1, \dots, a_n)$, можем първо да изберем хората между x и y, $(b_1, \dots, b_r)$, след което заедно с x и y да ги "вмъкнем" между останалите хора, $(c_1, \dots, c_{n-r-2})$, на една от $(n-r-1)$ възможни 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 местата "n" и "1" са "долепени", в допълнение към пермутациите в случай 1 трябва да разгледаме случаите, когато между x и y има r човека, като вземем предвид хората в двата края на редицата. 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 $n = r+2$, тогава двете места съвпадат и вероятността е $\frac{1}{n-1}$.

□

Забележка. Двете решения показват, че от вероятностна гледна точка е едно и също 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** 2.10. 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 ($k \leq 4n$).

Отговори. (а) $\frac{C(n-1,k-r)}{C(n,k)}$; (б) $\binom{n}{m}\frac{C(n-m,k)}{C(n,k)}$; (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** 2.11. 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 ($k < n$).

Отговор. $(n-k)/n$. □

*Задача** 2.12. 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 ($n - k \leq m < n$).

Отговори. (а) $\binom{k}{k_1,\ldots,k_n}/n^k$; (б) $n!/n^n$; (в) $n\binom{n}{2}(n-1)!/n^n$; (г) $\binom{n}{m}\sum_{i=0}^{n-m}(-1)^i\binom{n-m}{i}(n-m-i)^k/n^n$. □

*Задача** 2.13. 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.

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

*Задача** 2.14. От n чифта обувки случайно се избират $2r$ обувки ($2r < n$). Да се пресметне вероятността измежду избраните обувки:

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

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

(а) $\frac{2n(2n-2)(2n-4)\cdots(2n-4r+2)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = \frac{2^{2r}\cdot V_n^{2r-1}}{V_{2n}^{2r}} = \frac{2^{2r}\cdot\binom{n}{2r}}{\binom{2n}{2r}}$, т.е. без значение подредбата, избираме $2r$ чифта по $\binom{n}{2r}$ начина, след което решаваме да вземем дясна или лява обувка от всеки избран чифт.
(б) $\frac{\binom{2n}{2}\cdot 2\cdot n\cdot(2n-2)\cdot(2n-4)\cdots(2n-4r+4)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = \frac{n\cdot 2^{2r-2}\cdot\binom{n-1}{2r-2}}{\binom{2n}{2r}}$, където $\binom{2n}{2}$ са местата, на които по 2 начина поставяме лявата и дясната обувка от един и същи чифт, избран от всички n чифта обувки.
(в) $\frac{\binom{2r}{2}\cdot 2\cdot n\cdot\binom{2r-2}{2}\cdot 2\cdot(n-1)\cdot(2n-4)\cdot(2n-6)\cdots(2n-4r+6)}{2n(2n-1)(2n-2)\cdots(2n-2r+1)} = \frac{\binom{n}{2}\cdot 2^{2r-4}\cdot\binom{n-2}{2r-4}}{\binom{2n}{2r}}$.

□

*Задача*** 2.15 (Bertrand's Ballot Problem). По време на избори за първия от двама кандидати са пуснати n бюлетини, а за втория – m бюлетини. Каква е вероятността при преброяване на бюлетините броят на преброените гласове, подадени за първия кандидат, да бъде по-голям от броя на гласовете, подадени за втория кандидат, през цялото време?

Решение (Техника). За всяко $i = 1, \ldots, n+m$, полагаме $x_i = 1$, ако i-тата бюлетина е за кандидат 1, и $x_i = -1$, ако е за кандидат 2. Задачата е еквивалентна на това броят на нетните гласове $s_1 > 0, s_2 > 0 \ldots, s_{n+m} = n - m > 0$ да бъде постоянно положителен, където $s_i = x_1 + \cdots + x_i$. Геометрично погледнато, последното предполага, че начупената линията, свързваща точките $(0,0), (1, s_1), (2, s_2), \ldots, (n+m, n-m)$, не докосва абсцисната ос. Нещо повече, тъй като кандидат 1 е винаги начело, имаме $s_1 = x_1 = 1$ и траекториите, които удовлетворяват условието, са сред всички свързващи $(1,1)$ и $(n+m, n-m)$. Тогава от Задача 1.8 получаваме

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

□

2.2 Условни вероятности и независимост на събития

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

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

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

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

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

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

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

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

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

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

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 $\mathbb{P}(\text{ч}) = M/N$, а вероятността да изтеглим бяла топка на втори ход при положение, че сме изтеглили черна на първи ход, е $\mathbb{P}(\text{ч}, \text{б}|\text{ч}) = (N - M)/(N - 1)$, и т.н. Тогава за всяка редица от n топки, k от които са черни (общо C_n^k на брой), напр. $\omega^{(k)} = (\text{ч}, \text{б}, \text{ч}, \dots, \text{б}, \text{ч})$, имаме

$$\mathbb{P}(\text{ч}, \text{б}, \text{ч}, \dots, \text{б}, \text{ч}) = \mathbb{P}(\text{ч})\mathbb{P}(\text{ч}, \text{б}|\text{ч}) \cdots \mathbb{P}(\text{ч}, \text{б}, \text{ч}, \dots, \text{б}, \text{ч}|\text{ч}, \text{б}, \text{ч}, \dots, \text{б}) = \frac{M}{N} \cdot \frac{N - M}{N - 1} \cdot \frac{M - 1}{N - 2} \cdots \frac{M - k + 1}{N - n + 1} = \frac{V_M^k V_{N-m}^{n-k}}{V_N^n},$$

което е инвариантно спрямо реда на изтеглените цветове. Следователно,

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

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* $A_1, \dots, A_n \in \mathcal{A}$ *се наричат независими в съвкупност, ако* <u>*за кои да е*</u> $i_1, \dots, i_k \in \{1, \dots, n\} : i_j \neq i_l, j \neq l$, *и* $k = 2, \dots, n$, *имаме*

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

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

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

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 $A_1 = \{\text{на първия зар се пада нечетно число}\}$, $A_2 = \{\text{на втория зар се пада нечетно число}\}$ и $A_3 = \{\text{сумата от точките е нечетна}\}$. Тогава $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = \frac{1}{2}$ и $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = \frac{1}{4}$, 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. $\square$

Задачи

Задача 2.16. Застрахователна компания води статистика за своите клиенти. Известно е, че 1) всичи клиенти посещават лекар поне веднъж годишно; 2) 60% посещават лекар повече от веднъж годишно; 3) 17% посещават хирург; 4) 15% от тези, които посещават лекар повече от веднъж годишно, посещават хирург. Каква е вероятността случайно избран клиент, който посещава лекар само веднъж годишно, да не е бил при хирург?

Решение. Нека $A = \{$посещава хирург$\}$ и $B = \{$на лекар > 1 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 2.17. Трима ловци стрелят едновременно по заек. Заекът е убит от един куршум. Каква е вероятността той да е изстрелян от първия ловец, ако те уцелват с вероятност, съответно 0.2, 0.4 и 0.6?

Решение. Нека $A = \{$убит от един куршум$\}$ и $H_i = \{$ловец i уцелва$\}$, $i = 1, 2, 3$. Тогава $A = (H_1 \cap H_2^c \cap H_3^c) \cup (H_1^c \cap H_2 \cap H_3^c) \cup (H_1^c \cap H_2^c \cap H_3)$, откъдето $\mathbb{P}(H_1|A) = \frac{\mathbb{P}(A\cap H_1)}{\mathbb{P}(A)} \approx 10\%$. □

Задача 2.18 (Birthday Paradox). Какъв е най-малкият брой хора, които трябва да се изберат по случаен начин така, че вероятността поне двама от тях да имат един и същ рожден ден да е по голяма от 1/2?

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

Задача 2.19 (Boy or Girl Paradox). X има две деца. Ако по-старото е момиче, каква е вероятността и двете да са момичета? А ако знаете, че поне едно от тях е момиче, каква е вероятността и двете да са момичета?

Решение. Нека $\Omega = \big\{(a_1, a_2) : a_1, a_2 \in \{m, f\}\big\}$, $A = \{$по-малкото е момиче$\} \equiv \{(m, f), (f, f)\}$ и $B = \{$по-старото е момиче$\} \equiv \{(f, m), (f, f)\}$, като $\mathbb{P}(\{\omega\}) = 1/4$, за $\omega \in \Omega$. Тогава

$$\mathbb{P}(A \cap B|B) = \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{(f, f)\})}{\mathbb{P}(\{(f, m), (f, f)\})} = \frac{1}{2},$$
$$\mathbb{P}(A \cap B|A \cup B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cup B)} = \frac{\mathbb{P}(\{(f, f)\})}{\mathbb{P}(\{(f, m), (m, f), (f, f)\})} = \frac{1}{3}.$$

□

*Задача** 2.20. Кутия съдържа n билета, от които $m \le n$ печелят, а останалите билети губят. Всеки от n играчи на свой ред избира по един билет. Какви са шансовете за печалба на всеки от играчите? Кога е по-изгодно да се изтегли билет?

Решение 1. Нека си представим, че играчите теглят билети по реда на техните номера. Поставяме $A_k^0 = \{k$-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. $\mathbb{P}(A_k^0) = \binom{n-1}{m-1}\big/\binom{n}{m} = m/n$. □

*Задача** 2.21. 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.

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

Формула за пълната вероятност. *Нека* $\{H_n\}_n \subseteq \mathcal{A}$ *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* $A \in \mathcal{A}$,

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

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

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

*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 $A = \{\text{бяла топка}\}$ и $H_i = \{\text{топката е от i-тата група урни}\}$, $i \in \{1, 2\}$. 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 2.22. 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?

Отговор. $\approx 34\%$. □

Задача 2.23. Разполагаме с три стандартни зара и един, чиито страни са само шетици. По случаен начин избираме три от заровете и ги хвърляме. Да се определи вероятността да се паднат а) три шестици; б) различни цифри; в) последователни цифри?

Отговори. а) $\frac{1}{4}\frac{1}{6^3} + \frac{3}{4}\frac{1}{6^2}$; б) $\frac{1}{4}\frac{V_6^3}{6^3} + \frac{3}{4}\frac{V_5^2}{6^2}$; в) $\frac{1}{4}\frac{4\cdot3!}{6^3} + \frac{3}{4}\frac{2}{6^2}$. □

*Задача** 2.24. В кутия има 7 топки за тенис, от които 4 са нови. За първата игра по случаен начин се избират три топки, които след играта се връщат обратно в кутията. За втората игра също случайно се избират три топки. Каква е вероятността те да са нови?

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

Задача 2.25. 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 се оказали успешни, а един неуспешен, каква е вероятността и трите резултата да са на момичета?

Решение. Нека $A = \{2 \text{ усп. и } 1 \text{ неусп.}\}$ и $H_i = \{\text{избрани i момичета}\}$, $i = 0, 1, 2, 3$. Тогава $H_0, \ldots, H_3$ образуват пълна група и $\mathbb{P}(H_i) = \frac{\binom{45}{i}\binom{55}{3-i}}{\binom{100}{3}}$. Отделно $\mathbb{P}(A|H_0) = 3 \cdot 0.4^2 \cdot 0.6$, $\mathbb{P}(A|H_1) = 0.4^2 \cdot 0.3 + 2 \cdot 0.4 \cdot 0.7 \cdot 0.6$, $\mathbb{P}(A|H_2) = 0.7^2 \cdot 0.6 + 2 \cdot 0.7 \cdot 0.4 \cdot 0.3$, $\mathbb{P}(A|H_3) = 3 \cdot 0.7^2 \cdot 0.3$, 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 2.26 (Bertrand's Box Paradox). Дадени са три жетона. Първият има две бели страни, вторият две черни, а третият една бяла и една черна страна. По случаен начин се избира жетон и се хвърля върху маса. Ако горната страна на жетона е бяла, каква е вероятността другата му страна също да е бяла?

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

Забележете, че $\mathbb{P}(H_2|H_1 \cup H_2) = 1/2$, но $A \subsetneq H_1 \cup H_2$. □

*Задача** 2.27 (Monty Hall Problem). Зад една от три затворени врати има чисто нова кола, а зад другите две няма нищо. Избирате врата, след което водещият отваря една от останалите две врати, зад които няма нищо. Сега трябва да решите – сменяте ли избраната врата или запазвате първоначалния си избор?

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

Задача 2.28. Всички изделия в едната от две партиди са доброкачествени, а в другата $1/4$ от изделията са бракувани. Изделие, взето от случайно избрана партида, се оказало доброкачествено, след което е върнато обратно в своята партида. Да се пресметне вероятността второто случайно избрано изделие от същата партида да се окаже бракувано.

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

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

Алтернативно, от факта, че $A \subseteq C$, получаваме $\mathbb{P}(A|B) = \mathbb{P}(B|A,C)\mathbb{P}(A)/\mathbb{P}(B) = 3/28$. □

*Задача** 2.29. Петднадесет изпитни билета съдържат по два въпроса и покриват целия конспект от 30 въпроса. Студент може да отговори на 25 въпроса. Каква е вероятността той да вземе изпита, ако за това е нужно той да отговори на двата въпроса в един билет или на един от двата въпроса, а след това и на посочен въпрос от друг билет?

Решение. Нека $A = \{\text{взема изпита}\}$, $H_i = \{\text{отговаря на i въпроси}\}$, $F = \{\text{отговаря на допълнителен въпрос}\}$, за $i = 0, 1, 2$. Тогава

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

□

Задача 2.30. 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.

Решение (Техника). 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

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

като $\mathbb{P}(A|\text{ч}) = \mathbb{P}(\text{чб}|\text{ч}) + \mathbb{P}(\text{ччб}|\text{ч}) + \cdots$. 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 $\mathbb{P}(A|\text{ч}) = \mathbb{P}(A)$. Следователно $\mathbb{P}(A) = \frac{n}{n+m+l} + \frac{l}{n+m+l}\mathbb{P}(A)$ и $\mathbb{P}(A) = \frac{n}{n+m}$.

(б) $\mathbb{P}(A) = \mathbb{P}(\text{б}) + \mathbb{P}(\text{чб}) + \cdots + \mathbb{P}(\text{ччч}\cdots\text{чб}) = \frac{n}{n+m+l}\Big[1 + \sum_{k=1}^{l} \frac{l\cdots(l-k+1)}{(n+m+l-1)\cdots(n+m+l-k)}\Big]$.
Но $A^c = \{\text{зелена преди бяла}\}$ 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 2.4 можем да (i) 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** 2.31. Двама играчи последователно хвърлят монета. Играта печели този, който пръв хвърли ези. Каква е вероятността за спечелване на играта за всеки от играчите? А ако печели този, който хвърли същото като падналото се непосредствено преди това?

Решение (Техника). Нека $A = \{\text{I печели}\}$. 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 2.30, 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 2.32. Всяка от N урни съдържа по m бели и n черни топки. От първата урна случайно се избира една топка и се прехвърля във втората. След това от втората урна случайно се избира една топка и се прехвърля в третата и т.н. Каква е вероятността от последната урна да бъде извадена бяла топка?

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

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

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

Алтернативно, нека $A_k^0 = A_k$ и $A_k^1 = A_k^c$, за $k = 1,\dots,n$. Понеже вероятността на $A_k^{i_k}$ зависи само от $A_{k-1}^{i_{k-1}}$,

$$\mathbb{P}(A_N) = \sum_{i_1,\dots,i_{N-1}\in\{0,1\}} \mathbb{P}(A_1^{i_1})\mathbb{P}(A_2^{i_2}|A_1^{i_1})\cdots\mathbb{P}(A_N|A_1^{i_1},\dots,A_{N-1}^{i_{N-1}}) = \sum_{i_1,\dots,i_{N-1}\in\{0,1\}} \mathbb{P}(A_1^{i_1})\mathbb{P}(A_N|A_{N-1}^{i_{N-1}})\prod_{k=2}^{N-1}\mathbb{P}(A_k^{i_k}|A_{k-1}^{i_{k-1}})$$

$$= \sum_{i_1,\dots,i_{N-1}\in\{0,1\}} \Big(i_1\frac{m}{m+n} + (1-i_1)\frac{n}{m+n}\Big)\frac{m+i_{N-1}}{m+n+1}\prod_{k=2}^{N-1}\Big(i_k\frac{m+i_{k-1}}{m+n+1} + (1-i_k)\frac{n+(1-i_{k-1})}{m+n+1}\Big) = \frac{m}{m+n}.$$

□

*Задача** 2.33. 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?

Решение. Можем да подходим към проблема подобно на Задача 2.4. Абстрахирайки се от другите 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 едно от събитията

$$\Omega = \Big\{(b,b,\boxed{\text{r}},r),(b,\boxed{\text{r}},b,r),(b,\boxed{\text{r}},r,b),(\boxed{\text{r}},b,b,r),(\boxed{\text{r}},b,r,b),(\boxed{\text{r}},r,b,b)\Big\},$$

като в 50% от случаите ще видим първо червено преди черно асо след 6-та позиция. Вероятността обаче на това събитие, A, не е 50% понеже елементарните събития в Ω не са равно вероятни. Нека H_i е събитието "има i черни аса преди 6-та позиция", за $i = 0, 1, 2$. Тогава $\mathbb{P}(H_i)$ зависи от броя на възможните комбинации при раздаването на цялото тесте, като съотношението $\mathbb{P}(H_2) : \mathbb{P}(H_1) : \mathbb{P}(H_0)$ е равно на $\binom{5}{2} : \binom{5}{1}\binom{45}{1} : \binom{45}{2}$. От друга страна, не е трудно да се види, че $\mathbb{P}(A|H_i) = 1/(1+i)$, откъдето

$$\mathbb{P}(A) = \sum_{i=0}^{2} \mathbb{P}(A|H_i)\mathbb{P}(H_i) = \frac{1}{\binom{5}{2} + \binom{5}{1}\binom{45}{1} + \binom{45}{2}} \left(1 \cdot \binom{5}{2} + \frac{1}{2} \cdot \binom{5}{1}\binom{45}{1} + \frac{1}{3} \cdot \binom{45}{2}\right) \approx 0.3694.$$

□

*Задача*** 2.34 (Gambler's Ruin). Последователно се хвърля монета. Ако се падне ези, играчът печели 1 лв., а ако се падне тура – губи 1 лв. В началото на играта играчът има x лв. Играта завършва или когато играчът набере предварително определена сума от a лв., или когато проиграе всичките си пари. Каква е вероятността играчът да спечели? **А какъв е шансът играчът да спечели a лв. на n-ти ход?

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

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

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

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

За по-прегледно, нека положим $n^* = n-1$, $m^* = a-x-1$. От Задача 1.8, 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 1.8, 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

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

откъдето

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

□

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

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

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

3 Дискретни случайни величини

3.1 Основни понятия

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

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

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

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

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

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

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

Забележка (Интерпретация)*.* Нека X е дискретна сл. в. 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 дискретно вероятностно пространство, чрез което свеждаме оригиналния случаен експеримент до негова част, чиито елементарни изходи представляват възможните стойности на X. Освен това, за всяко $A \subseteq \mathbb{X}$, множеството $\{X \in A\} = \bigcup_{x \in A} \{X = x\} \in \mathcal{A}$ е измеримо, откъдето

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

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

*Забележка** (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

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

Тогава $Z := \mathbb{1}_A$ е (дискретна) сл. в. тогава и само тогава, когато $\{Z = 1\} = A \in \mathcal{A}$ е измеримо събитие, като имаме $\mathbb{P}(Z = 1) = \mathbb{P}(A) =: p$. В този случай Z е бинарна сл. в., показваща с вероятност p дали събитието A е настъпило.

Нека X е дискретна сл. в., приемаща стойности в $\mathbb{X} \subseteq \mathbb{R}$. Тогава за всяко $x \in \mathbb{X}$ имаме $X(\omega) = x \iff \omega \in H_x := \{X = x\}$, откъдето следва компактното представяне на X,

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

Нещо повече, 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 $\Omega = \bigcup_{x \in \mathbb{X}} H_x$. В случаите, когато $|X| < \infty$, ще наричаме X **проста** сл. в.

Забележка (Трансформация на сл. в.)*.* Нека X е дискретна сл. в. и $f : \mathbb{R}^m \to \mathbb{R}^n$ е *произволна* функция. Тогава $Y = f(X)$ е дискретна сл. в. със стойности в $f(\mathbb{X}) := \{t \in \mathbb{R}^n : \exists x \in \mathbb{X}, f(x) = t\}$. Последното следва от факта, че за всяко $y \in f(\mathbb{X})$,

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

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

Решение. Нека $\Omega = \{(a_1, \ldots, a_{10}) : a_i \in \{E, T\}\}$ и $\mathbb{P}(\{\omega\}) = 2^{-10}$, $\omega \in \Omega$. Тогава $X(\omega) := \sum_{i=1}^{n} \mathbb{1}_{\{E\}}((\omega)_i)$, $\omega \in \Omega$, е дискретна сл. в., описваща броят езита от 10-те опита, като приема стойности в $\mathbb{X} = \{0, 1, \ldots, 10\}$. Освен това $\mathbb{P}(X = x) = \binom{10}{x} 2^{-10}$ за $x \in \mathbb{X}$, откъдето (а) $\mathbb{P}(X = 5) = 0.25$; (б) $\mathbb{P}(\{X = 4\} \cup \{X = 6\}) = 2 \cdot \mathbb{P}(X = 6) = 0.41$; (в) $\{X = 7.5\} = \emptyset$. □

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

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

....

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

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

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

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

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

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

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

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

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

Д-во. Ако $X \geq 0$, тогава $\mathbb{E}[X] = \sum_{x \in \mathbb{X}} x p(x) \geq 0$ 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* $a, b \in \mathbb{R}$*.*

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

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

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

Отговор. $-5/6$. □

Задача 3.4. Случайна величина X приема краен брой стойности $x_1, \dots, x_n$ с една и съща вероятност. Да се пресметнат $\mathbb{E}[X]$ и $\mathrm{Var}(X)$, ако (а) $x_i = \frac{i-1}{n-1}$ за $i = 1, \dots, n$; (б) $x_i = a + (b-a)\frac{i-1}{n-1}$, $b > a$, за $i = 1, \dots, n$.

Решение. Сл. в. X има *равномерно разпределение*, т.е. $\mathbb{P}(X = x_i) = \frac{1}{n}$, за $i = 1, \dots, n$. Тогава (а) $\mathbb{E}[X] = \sum_{i=1}^n \frac{i-1}{n-1}\frac{1}{n} = \frac{\sum_{j=1}^{n-1} j}{n(n-1)} = \frac{1}{2}$ 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 3.5 (St. Petersburg Paradox)*.* Казино предлага следната игра: играч плаща A лв. След това хвърля монета, докато не се падне ези. Ако това се случи на n-тия ход, печели 2^n лв. Каква е очакваната му печалба? При какви стойности на A бихте участвали?

Решение. Нека сл. в. X е моментът на печалба. Тогава $\mathbb{P}(X = n) = 2^{-n}$, $n \in \mathbb{N}$, 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 3.6 (Martingale Strategy)*.* Казино предлага коефициент 2 при игра на ези/тура, т.е. при залог от A лв. бихме спечелили чисто A лв. Играч залага само на ези, докато спечели, като удвоява залога си всеки път, когато не спечели. Каква е очакваната му печалба? Бихте ли пробвали?

Решение. Нека сл. в. X е моментът на печалба. Тогава печалбата му 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. □

3.2 Колекция от случайни величини

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

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

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

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

Забележка (Многомерна сл. в.). Нека X е дискретна сл. в., приемаща стойности в $\mathbb{R}^m$. Дефинираме функцията $X_i(\omega) := (X(\omega))_i$, $\omega \in \Omega$. С други думи, X_i е i-тата компонента на X, като 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

$$\{X_i = x_i\} = \{X \in \mathbb{X}_1 \times \mathbb{X}_2 \times \cdots \times \mathbb{X}_{i-1} \times \{x_i\} \times \mathbb{X}_{i+1} \times \cdots \times \mathbb{X}_m\} \in \mathcal{A};$$

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. И обратното, ако $X_1, \dots, X_m$ са реални дискретни сл. в., за всяко $x = (x_1, \dots, x_m) \in \mathbb{X}_1 \times \cdots \times \mathbb{X}_m$ имаме

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

Нека сега X и Y са две дискретни реални сл. в., и $c \in \mathbb{R}$. От последните две забележки следва, че cX, $X + Y$, XY, X/Y за $Y \neq 0$, $\min\{X, Y\}$, $\max\{X, Y\}$, X^2 са дискретни сл. в.

Твърдение 3.3.
(i) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;
(ii) *ако* $X \geq Y$, *то* $\mathbb{E}[X] \geq \mathbb{E}[Y]$.

Д-во. Използвайки смяна на променливите,

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

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

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

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

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*

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

Случайните величини 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* $i_1, \dots, i_n \in \mathbb{N}$, $n \geq 1$.

Твърдение 3.4. *Ако* $X \perp\!\!\!\perp Y$, *то* $f(X) \perp\!\!\!\perp g(Y)$ *и* $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ *за* $f : \mathbb{R}^m \to \mathbb{R}$ *и* $g : \mathbb{R}^n \to \mathbb{R}$.

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

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

Относно втория резултат,

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

□

Твърдение 3.5. *Ако* $X \perp\!\!\!\perp Y$, *то* $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Д-во.

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

□

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

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

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

Забележка (Интерпретация). Нищо не изисква случайните величини X и Y да бъдат дефинирани в едно и също вероятностно пространство, 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.

Задача 3.7. В урна има 3 бели и 2 черни топки, като теглим последователно топки без връщане. Нека сл. в. X е моментът на тегленето на първата бяла топка. След това продължаваме да теглим, докато се появи черна топка. Нека Y е моментът на тегленето на първата черна топка след първата бяла. Дефинираме $Y = 6$, ако няма такава. Да се определи (а) съвместното разпределение на X и Y; (б) $\mathbb{P}(Y > 2|X = 1)$ и $\mathbb{P}(Y = 3|X < 3)$.

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

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

$\square$

Задача 3.8. А хвърля 3 монети, а Б – 2. Печели този, който хвърли повече езита, като взима всичките 5 монети. В случай на равен брой печели Б. Каква е вероятността А да спечели? Ако е спечелил А, каква е вероятността Б да хвърли точно едно ези? Каква е средната печалба на играчите?

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

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

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

. . . .

Твърдение 3.6 (Неравенство на Марков). *Нека X е дискретна сл. в., такава че $X \geq 0$. Тогава за всяко $\alpha > 0$ е в сила*

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

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

Твърдение 3.7 (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.* $\mathrm{Var}(X) \geq \mathbb{E}\big[(X - \mathbb{E}[X])^2 \cdot \mathbb{1}_{|X-\mathbb{E}[X]|>\alpha}\big] \geq \sum_{x\in\mathbb{X}:|x-\mathbb{E}[X]|>\alpha} \alpha^2 p(x) = \alpha^2 \cdot \mathbb{P}(|X - \mathbb{E}[X]| > \alpha)$. $\square$

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

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

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

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

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

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

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

Задача 3.10. Хвърлят се два зара. 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?

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

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

За втория въпрос приложете неравенството на Чебишев. □

....

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

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

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 сл. в. Тогава*

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

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 $\bar{X} = \frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}}$ 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

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

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

Твърдение 3.9. *За всяко $a, b, c \in \mathbb{R}$, 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 $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$. □

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

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

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

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

;

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

.

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

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

$2X+Y+1$	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

.

□

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

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

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

Отговори. $\mathbb{E}[X] = -0.1$, $\mathbb{E}[Y] = 0.7$, $\mathrm{Var}(X) = 0.49$, $\mathrm{Var}(Y) = 0.21$; $\mathrm{Cov}(X,Y) = 0.07$, $\rho_{X,Y} = 1/\sqrt{21}$; $\mathbb{E}[Z] = 1.9$, $\mathrm{Var}(Z) = 1.29$,

Z	0	1	2	3
$\mathbb{P}_Z$	0.2	0.1	0.3	0.4

. □

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

Отговор. Понеже $X + Y = 2$, следва, че $\mathrm{Cov}(X,Y) = \mathrm{Cov}(X, 2 - X) = \mathrm{Cov}(X, -X) = -\mathrm{Var}(X)$ и $\mathrm{Var}(Y) = \mathrm{Var}(X)$, откъдето $\rho_{X,Y} = -1$. □

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

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

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

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

3.3 Условно очакване

Условно очакване. *Нека X и Y са две дискретни сл. в., дефинирани в едно и също вероятностно пространство и приемащи стойности съответно в $\mathbb{X} \subseteq \mathbb{R}$ и $\mathbb{Y} \subseteq \mathbb{R}^n$, 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 $g : \mathbb{R}^n \to \mathbb{R}$ е зададена чрез*

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

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

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

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,

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

т.е. $\mathbb{E}[X|Y]$ е оптималната в средно квадратичен смисъл оценка на X, използвайки 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 $A, B \subseteq \Omega$. Тогава $\mathbb{E}[X|Y] = \mathbb{P}(B|A) \cdot \mathbb{1}_A + \mathbb{P}(B|A^c) \cdot \mathbb{1}_{A^c}$.

Твърдение 3.10 (Свойства на условното очаване)**.**
- *(i)* $\mathbb{E}[aX + bZ|Y] = a\mathbb{E}[X|Y] + b\mathbb{E}[Z|Y]$, *за* $a, b \in \mathbb{R}$;
- *(ii)* $\mathbb{E}\big[\mathbb{E}[X|Y]\big] = \mathbb{E}[X]$; *(Law of Total Expectation)*
- *(iii)* *ако* $X \perp\!\!\!\perp Y$, *то* $\mathbb{E}[X|Y] = \mathbb{E}[X]$;
- *(iv)* *ако* $X = f(Y)$ *за* $f : \mathbb{R}^n \to \mathbb{R}$, *то* $\mathbb{E}[X|Y] = X$.

Д-во.* Лесно се вижда, че

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

откъдето

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

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

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

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

□

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

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

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

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

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

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

□

*Задача** 3.16 (Продължение на Задача 3.14). На всяка от страните на зар се записва случайно число от 1 до 6, след което зарът се хвърля n пъти. Нека X е сумата от първите $n-1$ хвърляния, а Y – сумата от последните $n-1$. Намерете $\rho_{X,Y}$.

Решение. Нека $U_1, \dots, U_n$ 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 (в Задача 3.14 имаме просто $X_i = U_i$). Тогава, за $k = 1, \dots, 6$,

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

т.е. $X_1, \dots, X_n$ са еднакво разпределени като при правилен зар, откъдето $\mathbb{E}[X_1] = 3.5$ и $\operatorname{Var}(X_1) = 2.917$. Въпреки това, за разлика от Задача 3.14, сл. в. $X_1, \dots, X_n$ не са независими. Всъщност

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

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

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

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

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

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

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

т.е. няма промяна в очакваната печалба. □

3.4 Характеристики на случайни величини

> *"A generating function is a device somewhat similar to a bag. Instead of carrying many little objects detachedly, which could be embarrassing, we put them all in a bag, and then we have only one object to carry, the bag."*
>
> – George Pólya , *Mathematics and Plausible Reasoning* (1954)

В зависимост от конкретната задача използването на някоя от следните функции може значително да улесни пресмятането на вероятности или други количества. По-специално, пораждащата функция може да се използва за извеждането на някои свойства на X чрез аналитични методи.

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

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

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

$$F_{X_1,\dots,X_n}(x_1, \dots, x_m) := \mathbb{P}(X_1 \le x_1, \dots, X_m \le x_m), \qquad x_1, \dots, x_m \in \mathbb{R}.$$

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

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

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

Забележка (Свойства). Функцията F_X удовлетворява следните условия:

(i) F_X е ненамаляваща, т.е. ако $x_1 < x_2$, то $F_X(x_1) \le F_X(x_2)$;
(ii) F_X е непрекъсната 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 $\lim_{x \to -\infty} F_X(x) = 0$.

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

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

Пораждаща функция на вектор от целочислени сл. в. $(X_1, \dots, X_m)$ наричаме функцията

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

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

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

Твърдение 3.11. *Нека X, X_1 и X_2 са целочислени дискретни сл. в. 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 $\mathbb{P}_X$. Следователно за две целочислени сл. в. X и Y,

$$G_X = G_Y \qquad \Longleftrightarrow \qquad X \overset{d}{=} Y$$

Твърдение 3.12.** *Ако X е целочислена сл. в., 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

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

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

Твърдение 3.13. *Ако X е целочислена сл. в., тогава*

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

Д-во.

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

Алтернативно, от Твърдение 3.12, $1-(1-s)\sum_{n=0}^{\infty} s^n\mathbb{P}(X>n) = G_X(s)$, откъдето $G'_X(1) = \sum_{n=0}^{\infty}\mathbb{P}(X>n)$. □

3.5 Основни дискретни разпределения

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

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

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

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 $k=0,1,\dots,n$,

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

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.

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

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

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

използвайки формулата на Вандермонд. От друга страна, директно,

$$G_X(s) = G_{X_1}(s)G_{X_2}(s) = (ps-(1-p))^{n_1}(ps-(1-p))^{n_2} = (ps-(1-p))^{n_1+n_2}.$$

□

Задача 3.19. Нека $X|P=p \sim \mathrm{Bin}(200,p)$, където параметърът P е сл. в., такава че $\mathbb{P}(P=\frac{1}{2}) = \mathbb{P}(P=\frac{2}{3}) = \frac{1}{2}$. Коя стойност на параметъра P има по-голяма *апостериорна* вероятност при условие, че $X=120$?

Решение. От свойствата на условните вероятности,

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

□

Задача 3.20. Нека $N \sim \mathrm{Bin}(M,q)$ и $X|N=n \sim \mathrm{Bin}(n,p)$, $n=1,\dots,M$, за $p,q \in (0,1)$ и $M \in \mathbb{N}$. Да се докаже, че $X \sim \mathrm{Bin}(M,pq)$.

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

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

□

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

Решение (Техника). Нека $A = \{\text{приборът излиза от строя}\}$ и $X_1 \sim \mathrm{Ber}(0.1)$, $X_2 \sim \mathrm{Ber}(0.2)$, $X_3 \sim \mathrm{Ber}(0.3)$ са независими сл. в. Тогава вероятността да изгорят точно $k = 0, 1, 2, 3$ лампи, $\mathbb{P}(X_1 + X_2 + X_3 = k)$, може да се изведе по стандартния начин, откъдето

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

От друга страна, $\mathbb{P}(X_1 + X_2 + X_3 = k)$ е равно на коефициента пред s^k в развитието

$$G_{X_1+X_2+X_3}(s) = \prod_{i=1}^{3} G_{X_i}(s) = (0.1s + 0.9)(0.2s + 0.8)(0.3s + 0.7) = (0.006s^3 + 0.092s^2 + 0.398s + 0.504).$$

□

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

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

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

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 $\mathbb{P}(X = 0, 2, 4, \ldots, 2\lfloor \frac{n}{2} \rfloor) = p_n = \frac{1}{2}(1 + (1-2p)^n)$. □

. . . .

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

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

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

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

Забележка (Безпаметност). За всяко $k = 0, 1, 2, \ldots$ имаме

$$\mathbb{P}(X \geq k) = 1 - \mathbb{P}(X < k) = 1 - \sum_{j=0}^{k-1} p(1-p)^j = 1 - p\frac{1-(1-p)^k}{1-(1-p)} = (1-p)^k.$$

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

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

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.

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

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

т.е. $\min_n\{\sum_{i=1}^n X_i = 1\} - 1 \sim \mathrm{Geo}(p)$.

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 $x_0 = 0$,

$$\mathbb{E}[X] = \sum_{k=1}^\infty k(1-p)^k p = p\sum_{j=0}^\infty (j+1)(1-p)^{j+1} = p(1-p)\sum_{j=0}^\infty \binom{j+1}{j}(1-p)^j = \frac{p(1-p)}{(1-(1-p))^2} = \frac{1-p}{p},$$

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

$$= 2p(1-p)^2\sum_{k=2}^\infty \binom{k}{k-2}(1-p)^{k-2} + \frac{(1-p)}{p} = \frac{2p(1-p)^2}{p^3} + \frac{(1-p)}{p} = \frac{(1-p)(2-p)}{p^2},$$

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

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

Алт. д-во.

$$\mathbb{E}[X] = \sum_{k=1}^\infty k(1-p)^k p = \sum_{k=1}^\infty (k-1)(1-p)^k p + \sum_{k=1}^\infty (1-p)^k p = (1-p)(\mathbb{E}[X]+1) \quad\Longrightarrow\quad \mathbb{E}[X] = \frac{1-p}{p},$$

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

$$= (1-p)\big(\mathbb{E}[X^2] + 2\mathbb{E}[X] + 1\big) \quad\Longrightarrow\quad \mathbb{E}[X^2] = \frac{(1-p)(2-p)}{p^2}.$$

□

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

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

$$\begin{array}{c|ccccc} Y & 1 & 2 & \cdots & n-1 & n \\ \hline \mathbb{P}_Y & p & (1-p)p & \cdots & (1-p)^{n-2}p & (1-p)^{n-1} \end{array}.$$

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 $\sum_{k=0}^{n-1}(k+1)x^k = \frac{1-x^n}{(1-x)^2} - \frac{nx^n}{1-x}$ следва

$$\mathbb{E}[Y] = p\sum_{i=1}^{n-1} i(1-p)^{i-1} + n(1-p)^{n-1}$$

$$= p\Big(\frac{1-(1-p)^{n-1}}{p^2} - \frac{(n-1)(1-p)^{n-1}}{p}\Big) + n(1-p)^{n-1} = \frac{1-(1-p)^{n-1}}{p} + (1-p)^{n-1} = \frac{1-(1-p)^n}{p}.$$

□

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

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

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

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

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

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

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

От друга страна разпределението на X може да се изведе подобно на Задача 1.1в, като $\{X = k\}$ е събитието "k различими частици в n клетки, където нито една клетка не остава празна и с последната частица запълваме последната празна клетка". $\square$

. . . .

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 $\sum_{k=0}^\infty \binom{n+k-1}{k} p^n (1-p)^k = p^n \sum_{k=0}^\infty \binom{n+k-1}{k}(1-p)^k = p^n(1-(1-p))^{-n} = p^n p^{-n} = 1$, използвайки развитието в ред на $(1-x)^{-n} = \sum_{k=0}^\infty \binom{n+k-1}{k} x^k$ в точката $x_0 = 0$; т.е. отрицателното биномно разпределение е добре дефинирано. По същия начин, 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. Нека $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathrm{Geo}(p)$. Тогава, за всяко $k = 0, 1, 2, \ldots$,

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

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

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

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

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 $\mathbb{P}(X=k^*)=\mathbb{P}(X=k^*-1)$.

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

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

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

□

. . . .

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

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

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

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

Пример 3.3 (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 без връщане n топки ($n<N,n<N-m$). Да се намери вероятността точно $k\leq\min(M,n)$ от извадените топки да са черни.

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

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

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

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

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

$\square$

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

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

*Забележка*** (Зависима схема на Бернули). Нека X е като в Пример 2.1, а сл. в. X_i отчита "успеха" на i-тия ход в n тегления без връщане. Тогава $X_1 + \cdots + X_n \sim \mathrm{HG}(M, N, n)$.

От Задача 2.20 следва, че $X_i \sim \mathrm{Ber}(\frac{M}{N})$ за всяко $i = 1, \ldots, N$ и в частност за $i \leq n$, откъдето получаваме $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\frac{M}{N}$. Въпреки това сл. в. X_i не са независими. И все пак

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

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

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

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

Отговор. $X \sim \mathrm{HG}(3, 7, 4)$, $\mathbb{E}[X] = \frac{12}{7}$, $\mathrm{Var}(X) = \frac{24}{49}$. $\square$

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

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

. . . .

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

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

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

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 3.4. 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.

*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 $k = 0, 1, 2, \ldots$,

$$\lim_{n\to\infty} \mathbb{P}(X_n = k) = \lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n} + o\Big(\frac{1}{n}\Big)\right)^k \left(1 - \frac{\lambda}{n} + o\Big(\frac{1}{n}\Big)\right)^{n-k}$$

$$= \lim_{n\to\infty} \frac{\lambda^k}{k!}\Big(1-\frac{1}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big)\Big(1-\frac{\lambda}{n}\Big)^{n-k}(1+o(1)) = \frac{e^{-\lambda}\lambda^k}{k!} = \mathbb{P}(X=k).$$

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 3.32).

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 $\mathbb{P}(X=k^*) = \mathbb{P}(X=k^*-1)$.

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

Решение. За всяко $k=0,1,2,\ldots$,

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

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

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

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

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

Решение. 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

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

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

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

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

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

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

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

Забележка (Интерпретация). Тази сл. в. моделира резултата от n *независими* опита, при всеки от които настъпва точно едно от r *несъвместими* събития. Нека $X^{(i)} \sim \mathrm{Multi}(1; p_1, \ldots, p_r)$, $i = 1, \ldots, n$, 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 3.5. 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, $(x_1 + \cdots + x_r)^n = \sum_{k_1 + \cdots + k_r = n} \frac{n!}{k_1! \cdots k_r!} x_1^{k_1} \cdots x_r^{k_r}$. □

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

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

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

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

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

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

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

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

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

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

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

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

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

Решение. Нека $(X_1, X_2, X_3) \sim \text{Multi}(20; \frac{1}{5}, \frac{1}{5}, \frac{3}{5})$, където сл. в. X_1, X_2 и X_3 са съответно броят на изтеглените бели, черни и зелени топки. От Задача 3.33,

$$\mathbb{P}(X_1 - X_2 = 4) = \sum_{i=4}^{12} \mathbb{P}(X_1 = i, X_2 = i-4) = \sum_{i=4}^{12} \binom{20}{i}\binom{20-i}{i-4}\left(\frac{1}{5}\right)^{2i-4}\left(\frac{3}{5}\right)^{20-2i+4}.$$

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

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

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

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

Решение. Нека $(X_1, \dots, X_6) \sim \text{Multi}(n; \frac{1}{6}, \dots, \frac{1}{6})$, 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

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

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 2.13). □

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

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

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

откъдето $\mathbb{P}(X_1 + \dots + 9X_9 = Y_1 + \dots + 9Y_9) = \frac{1}{10^6}\sum_{i=0}^{3}(-1)^i\binom{6}{i}\binom{32-10i}{27-10i}$ (сравнете със Задача 2.3). □