C Zliczanie i prawdopodobieństwo

Dodatek ten jest przeglądem podstawowych wiadomości z kombinatoryki i rachunku prawdopodobieństwa. Jeśli Czytelnik jest zaznajomiony z tymi zagadnieniami, to może pominąć początek dodatku i skoncentrować się na dalszej jego części. Do zrozumienia materiału zawartego w większości rozdziałów nie jest wymagana znajomość rachunku prawdopodobieństwa, lecz w niektórych miejscach jest ona bardzo ważna.

W dodatku C.1 podamy elementarne wyniki z teorii zliczania, w tym standardowe wzory na zliczanie permutacji i kombinacji. Aksjomaty teorii prawdopodobieństwa i podstawowe fakty dotyczące rozkładów prawdopodobieństwa zaprezentujemy w dodatku C.2, a zmienne losowe wraz z własnościami wartości oczekiwanej i wariancji wprowadzimy w dodatku C.3. Z kolei w dodatku C.4 zajmiemy się badaniem rozkładów geometrycznego i dwumianowego, pojawiających się przy badaniu prób Bernoulliego. Rozważania na temat rozkładu dwumianowego będziemy kontynuować w dodatku C.5, zajmując się „krańcami” tego rozkładu.

C.1 Zliczanie

Teoria zliczania stara się odpowiedzieć na pytanie „Ile?” bez faktycznego zliczania. Możemy na przykład zapytać „Ile jest różnych liczb n-bitowych?” lub „Ile jest sposobów uporządkowania n różnych elementów?”. W tym podrozdziale omawiamy elementy teorii zliczania. Zakładamy, że Czytelnik zna podstawowe zagadnienia teorii zbiorów – jeśli nie, to powinien uprzednio zapoznać się z materiałem zawartym w dodatku B.1.

Reguły sumy i iloczynu

Zbiór elementów, które chcemy policzyć, można czasami wyrazić jako sumę zbiorów rozłącznych lub jako iloczyn kartezjański zbiorów.

Reguła sumy mówi, że liczba możliwości, na jakie możemy wybrać element należący do jednego z dwóch zbiorów rozłącznych, jest równa sumie mocy tych zbiorów. To znaczy, że jeśli A i B są zbiorami skończonymi niemającymi wspólnych elementów, to $|A \cup B| = |A| + |B|$, co wynika z równania (B.3) na str. 1086. Na przykład jeśli każdy znak na tablicy rejestracyjnej samochodu jest cyfrą lub literą, to liczba możliwości dla każdej pozycji jest równa $26 + 10 = 36$, ponieważ jest 26 możliwości wyboru litery i 10 możliwości wyboru cyfry.

Reguła iloczynu mówi, że liczba możliwości, na jakie możemy wybrać uporządkowaną parę elementów, jest równa liczbie możliwości, na jakie możemy wybrać pierwszy element, pomnożonej przez liczbę możliwości, na jakie możemy wybrać drugi element. Czyli jeśli A i B są zbiorami skończonymi, to $|A \times B| = |A| \cdot |B|$, co jest po prostu równością (B.4) na str. 1086. Na przykład, jeśli cukiernia oferuje 28 smaków lodów i 4 rodzaje bakalii, to liczba możliwych zestawów złożonych z jednej gałki lodów i jednej porcji bakalii wynosi $28 \cdot 4 = 112$.

Słowa

Słowo nad zbiorem skończonym (alfabetem) S jest to ciąg elementów zbioru S. Na przykład istnieje 8 słów binarnych długości 3:

$$000, 001, 010, 011, 100, 101, 110, 111.$$

(Stosujemy tutaj skrótowy zapis, w którym pomijamy nawiasy kątowe w oznaczeniu ciągu). Czasami słowo długości k nazywamy ***k-słowem***. ***Podsłowo*** s' słowa s jest to uporządkowany ciąg następujących po sobie elementów słowa s. ***k-podsłowo*** danego słowa jest jego podsłowem długości k. Na przykład 010 jest 3-podsłowem słowa 01101001 (3-podsłowo zaczynające się na pozycji 4), ale 111 nie jest podsłowem słowa 01101001.

Każde k-słowo nad zbiorem S można traktować jako element iloczynu kartezjańskiego S^k; w takim razie istnieje $|S|^k$ słów długości k. Na przykład liczba binarnych k-słów wynosi 2^k. Intuicyjnie, aby skonstruować k-słowo nad n-zbiorem, mamy n sposobów na wybranie pierwszego elementu; dla każdego z tych wyborów mamy n możliwości na wybranie drugiego elementu itd. k razy. Konstrukcja ta prowadzi do k-krotnego iloczynu $n \cdot n \cdots n = n^k$, który jest liczbą k-słów.

Permutacje

Permutacją zbioru skończonego S nazywamy uporządkowany ciąg wszystkich elementów zbioru S, przy czym każdy element występuje dokładnie raz. Jeśli na przykład $S = \{a, b, c\}$, to mamy 6 permutacji zbioru S:

$$abc, acb, bac, bca, cab, cba.$$

(Tutaj również stosujemy skrótowy zapis, w którym pomijamy nawiasy kątowe w oznaczeniu ciągu). Liczba permutacji zbioru n-elementowego wynosi $n!$, ponieważ pierwszy element możemy wybrać na n sposobów, drugi na $n - 1$ sposobów, trzeci na $n - 2$ sposoby itd.

Uporządkowany ciąg k elementów zbioru S, w którym żaden element nie występuje więcej niż jeden raz, nazywamy ***k-permutacją*** zbioru S. (Zatem zwyczajna permutacja jest n-permutacją zbioru n-elementowego). Wszystkie 2-permutacje zbioru $\{a, b, c, d\}$ to:

$$ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc.$$

Liczba k-permutacji zbioru n-elementowego wynosi

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}, \tag{C.1}$$

jako że istnieje n sposobów wyboru pierwszego elementu, $n-1$ sposobów wyboru drugiego elementu itd., aż wybierzemy wszystkich k elementów, przy czym ostatni będzie wybierany spośród $n-k+1$ elementów. W powyższym przykładzie dla $n = 4$ i $k = 2$ wzór (C.1) daje wynik $4!/2! = 12$, co zgadza się z liczbą wypisanych 2-permutacji.

Kombinacje

***k*-kombinacją** zbioru n-elementowego S nazywamy po prostu k-elementowy podzbiór zbioru S. Istnieje sześć 2-kombinacji zbioru 4-elementowego $\{a, b, c, d\}$:

$$ab, ac, ad, bc, bd, cd.$$

(Używamy tutaj skrótu, pomijając nawiasy klamrowe w oznaczeniu każdego podzbioru). Możemy skonstruować k-kombinację zbioru n-elementowego, wybierając k różnych elementów z tego zbioru. Kolejność, w jakiej wybieramy te elementy, nie ma znaczenia.

Liczbę k-kombinacji zbioru n-elementowego można wyrazić przez liczbę k-permutacji tego zbioru. Dla każdej k-kombinacji istnieje dokładnie $k!$ permutacji jej elementów, przy czym każda z nich jest inną k-permutacją zbioru n-elementowego. Stąd liczba k-kombinacji zbioru n-elementowego jest równa liczbie k-permutacji podzielonej przez $k!$; w myśl równości (C.1) liczba ta wynosi

$$\frac{n!}{k!\,(n-k)!}. \tag{C.2}$$

Dla $k = 0$ ze wzoru tego wynika, że możliwości wyboru 0 elementów ze zbioru n-elementowego jest 1 (nie 0), bo $0! = 1$.

Współczynniki dwumianowe

Zapisu $\binom{n}{k}$ (czytaj „n nad k”) używamy do oznaczania liczby k-kombinacji zbioru n-elementowego. Z równości (C.2) mamy

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

Wzór ten jest symetryczny ze względu na k i $n-k$:

$$\binom{n}{k} = \binom{n}{n-k}. \tag{C.3}$$

Liczby te są także nazywane ***współczynnikami dwumianowymi*** z powodu ich występowania w ***twierdzeniu o dwumianie***:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}, \tag{C.4}$$

gdzie $n \in \mathbb{N}$ oraz $x, y \in \mathbb{R}$. Prawą stronę wzoru (C.4) nazywamy ***rozwinięciem dwumianu*** stanowiącego lewą stronę. Przypadek szczególny rozwinięcia dwumianu występuje, gdy $x = y = 1$:

$$2^n = \sum_{k=0}^{n} \binom{n}{k}.$$

Wzór ten odpowiada zliczaniu 2^n binarnych n-słów ze względu na liczbę zawartych w nich jedynek: jest $\binom{n}{k}$ binarnych n-słów zawierających dokładnie k jedynek, jako że istnieje $\binom{n}{k}$ sposobów wybrania k spośród n pozycji, na których można umieścić jedynki.

Istnieje wiele tożsamości związanych ze współczynnikami dwumianowymi. Zadania zamieszczone na końcu tego podrozdziału dadzą Czytelnikowi możliwość udowodnienia kilku z nich.

Szacowanie współczynników dwumianowych

Czasami potrzebujemy oszacować wartość współczynnika dwumianowego. Dla $1 \le k \le n$ mamy dolne ograniczenie

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \\ &= \left(\frac{n}{k}\right)\left(\frac{n-1}{k-1}\right)\cdots\left(\frac{n-k+1}{1}\right) \\ &\ge \left(\frac{n}{k}\right)^k .\end{aligned} \tag{C.5}$$

Korzystając z nierówności $k! \ge (k/e)^k$ wyprowadzonej ze wzoru Stirlinga (3.25) na str. 63, otrzymujemy górne ograniczenie

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} \\ &\le \frac{n^k}{k!} \\ &\le \left(\frac{en}{k}\right)^k .\end{aligned} \tag{C.6}$$

Dla wszystkich $0 \le k \le n$ możemy przez indukcję (patrz zad. C.1-12) udowodnić ograniczenie

$$\binom{n}{k} \le \frac{n^n}{k^k (n-k)^{n-k}}, \tag{C.7}$$

gdzie dla wygody przyjmujemy, że $0^0 = 1$. Dla $k = \lambda n$, gdzie $0 \le \lambda \le 1$, ograniczenie to możemy zapisać jako

$$\begin{aligned}\binom{n}{\lambda n} &\le \frac{n^n}{(\lambda n)^{\lambda n}((1-\lambda)n)^{(1-\lambda)n}} \\ &= \left(\left(\frac{1}{\lambda}\right)^{\lambda}\left(\frac{1}{1-\lambda}\right)^{1-\lambda}\right)^n \\ &= 2^{n\,H(\lambda)},\end{aligned}$$

gdzie

$$H(\lambda) = -\lambda \lg \lambda - (1-\lambda)\lg(1-\lambda) \tag{C.8}$$

jest ***(binarną) funkcją entropii*** i gdzie, dla wygody, przyjmujemy, że $0 \lg 0 = 0$, czyli że $H(0) = H(1) = 0$.

Zadania

C.1-1
Ile k-podsłów zawiera n-słowo? (Przyjmij identyczne k-podsłowa zaczynające się na różnych pozycjach za różne). Ile łącznie podsłów zawiera n-słowo?

C.1-2
Funkcja logiczna o n wejściach i m wyjściach to funkcja ze zbioru $\{\text{PRAWDA}, \text{FAŁSZ}\}^n$ w zbiór $\{\text{PRAWDA}, \text{FAŁSZ}\}^m$. Ile jest funkcji logicznych o n wejściach i 1 wyjściu? Ile jest funkcji logicznych o n wejściach i m wyjściach?

C.1-3
Na ile sposobów można posadzić n osób przy okrągłym stole? Uznajemy dwa ustawienia za takie same, jeśli jedno powstaje z drugiego przez obrót wokół stołu.

C.1-4
Na ile sposobów można wybrać trzy różne liczby ze zbioru $\{1, 2, \ldots, 99\}$, aby ich suma była parzysta?

C.1-5
Udowodnij tożsamość

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} \tag{C.9}$$

dla $0 < k \le n$.

C.1-6

Udowodnij tożsamość

$$\binom{n}{k} = \frac{n}{n-k}\binom{n-1}{k}$$

dla $0 \le k < n$.

C.1-7

Aby wybrać k przedmiotów z n, można jeden z nich wyróżnić i sprawdzić, czy ten wyróżniony przedmiot został wybrany. Użyj tej wskazówki do udowodnienia, że

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

C.1-8

Korzystając z wyniku zadania C.1-7, utwórz tabelę współczynników dwumianowych $\binom{n}{k}$ dla $n = 0, 1, \ldots, 6$ i $0 \le k \le n$ z $\binom{0}{0}$ na górze, $\binom{1}{0}$ i $\binom{1}{1}$ w następnym wierszu, dalej $\binom{2}{0}$, $\binom{2}{1}$ i $\binom{2}{2}$, itd. Tabela taka nosi nazwę ***trójkąta Pascala***.

C.1-9

Udowodnij, że

$$\sum_{i=1}^{n} i = \binom{n+1}{2}.$$

C.1-10

Pokaż, że dla dowolnych liczb całkowitych $n \ge 0$ i $0 \le k \le n$ maksymalna wartość $\binom{n}{k}$ jest osiągana dla $k = \lfloor n/2 \rfloor$ lub $k = \lceil n/2 \rceil$.

★ ***C.1-11***

Uzasadnij, że dla dowolnych liczb całkowitych $n \ge 0$, $j \ge 0$, $k \ge 0$, takich że $j + k \le n$, zachodzi

$$\binom{n}{j+k} \le \binom{n}{j}\binom{n-j}{k}. \tag{C.10}$$

Podaj zarówno dowód algebraiczny, jak i argumentację opartą na metodzie wybierania $j + k$ przedmiotów spośród n. Podaj przykład, w którym nie zachodzi równość.

★ ***C.1-12***

Udowodnij nierówność (C.7) przez indukcję względem wszystkich całkowitych k, takich że $0 \le k \le n/2$, a następnie użyj równości (C.3), aby rozszerzyć ją na wszystkie całkowite k, takie że $0 \le k \le n$.

★ ***C.1-13***

Użyj wzoru Stirlinga, aby udowodnić, że

$$\binom{2n}{n} = \frac{2^{2n}}{\sqrt{\pi n}}(1 + O(1/n)). \tag{C.11}$$

★ ***C.1-14***

Różniczkując funkcję entropii $H(\lambda)$, pokaż, że osiąga ona maksimum dla $\lambda = 1/2$. Ile wynosi $H(1/2)$?

★ ***C.1-15***

Wykaż, że dla dowolnego całkowitego $n \geq 0$ zachodzi tożsamość

$$\sum_{k=0}^{n} \binom{n}{k} k = n2^{n-1}. \tag{C.12}$$

★ ***C.1-16***

Nierówność (C.5) stanowi dolne oszacowanie wartości współczynnika dwumianowego $\binom{n}{k}$. Dla małych k zachodzi silniejsze oszacowanie. Pokaż, że

$$\binom{n}{k} \geq \frac{n^k}{4k!}, \tag{C.13}$$

dla $k \leq \sqrt{n}$.

C.2 Prawdopodobieństwo

Teoria prawdopodobieństwa jest zasadniczym narzędziem służącym do tworzenia i analizy algorytmów randomizowanych (probabilistycznych). W tym dodatku przypominamy podstawy tej teorii.

Prawdopodobieństwo definiujemy w odniesieniu do ***przestrzeni zdarzeń*** S, będącej zbiorem, którego elementy nazywamy ***zdarzeniami elementarnymi***. Każde zdarzenie elementarne można traktować jako możliwy wynik eksperymentu. W eksperymencie polegającym na rzucie dwiema rozróżnialnymi monetami na przestrzeń zdarzeń możemy patrzeć jako na złożoną ze zbioru wszystkich możliwych 2-słów nad zbiorem $\{\text{O}, \text{R}\}$:

$$S = \{\text{OO}, \text{OR}, \text{RO}, \text{RR}\}.$$

Zdarzenie jest podzbiorem[1] przestrzeni zdarzeń S. Na przykład przy rzucie dwiema monetami zdarzenie polegające na uzyskaniu jednego orła i jednej reszki to $\{\text{OR}, \text{RO}\}$. Zdarzenie S jest nazywane ***zdarzeniem pewnym***, a zdarzenie $\emptyset$ jest nazywane ***zdarzeniem niemożliwym***. Mówimy, że dwa zdarzenia A i B są ***wzajemnie się wykluczające (rozłączne)***, jeśli $A \cap B = \emptyset$. Czasami zdarzenie elementarne $s \in S$ traktujemy jako zdarzenie $\{s\}$. Z definicji wszystkie zdarzenia elementarne wzajemnie się wykluczają.

Aksjomaty teorii prawdopodobieństwa

Rozkład prawdopodobieństwa $\Pr\{\}$ na przestrzeni zdarzeń S jest to przyporządkowanie zdarzeniom z S liczb rzeczywistych tak, że spełnione są następujące ***aksjomaty prawdopodobieństwa***:

1. $\Pr\{A\} \geq 0$ dla każdego zdarzenia A.
2. $\Pr\{S\} = 1$.
3. $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\}$ dla każdych dwóch wzajemnie wykluczających się zdarzeń A i B. Ogólniej, dla każdego (skończonego lub przeliczalnego) ciągu parami wykluczających się zdarzeń $A_1, A_2, \ldots$

$$\Pr\left\{\bigcup_i A_i\right\} = \sum_i \Pr\{A_i\}.$$

$\Pr\{A\}$ nazywamy ***prawdopodobieństwem*** zdarzenia A. Zwracamy tutaj uwagę, iż aksjomat 2 jest wymaganiem normalizacji: w rzeczywistości nie ma specjalnego powodu, aby wybrać 1 jako prawdopodobieństwo zdarzenia pewnego, poza tym, że jest to naturalne i wygodne.

Z tych aksjomatów i podstaw teorii zbiorów (patrz dodatek B.1) wynika natychmiast kilka wniosków. Prawdopodobieństwo zdarzenia niemożliwego $\emptyset$ wynosi $\Pr\{\emptyset\} = 0$. Jeśli $A \subseteq B$, to $\Pr\{A\} \leq \Pr\{B\}$. Stosując zapis $\overline{A}$ do oznaczenia zdarzenia $S - A$ (***dopełnienia*** zdarzenia A), mamy $\Pr\{\overline{A}\} = 1 - \Pr\{A\}$. Dla dowolnych dwóch zdarzeń A i B zachodzi

$$\begin{aligned} \Pr\{A \cup B\} &= \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\} && \text{(C.14)} \\ &\leq \Pr\{A\} + \Pr\{B\}. && \text{(C.15)} \end{aligned}$$

Przypuśćmy, że w naszym przykładzie z rzucaniem monetami każde z czterech zdarzeń elementarnych ma prawdopodobieństwo $1/4$. Wówczas prawdopodobieństwo otrzymania co

[1] W ogólnym rozkładzie prawdopodobieństwa mogą znajdować się podzbiory przestrzeni zdarzeń S, które nie są uważane za zdarzenia. Sytuacja ta występuje zwykle, gdy przestrzeń zdarzeń jest nieprzeliczalna. Głównym wymogiem jest, aby zbiór zdarzeń był zamknięty ze względu na operacje brania dopełnienia zdarzenia, sumowania skończonej lub przeliczalnej liczby zdarzeń oraz brania części wspólnej skończonej lub przeliczalnej liczby zdarzeń. Większość rozkładów prawdopodobieństwa, którymi będziemy się zajmować, jest określona nad skończoną lub przeliczalną przestrzenią zdarzeń i będziemy ogólnie przyjmować, że wszystkie podzbiory przestrzeni zdarzeń są zdarzeniami. Ważnym wyjątkiem jest ciągły rozkład jednostajny prawdopodobieństwa, który zostanie przedstawiony wkrótce.

najmniej jednego orła wynosi

$$\begin{aligned}\Pr\{\text{OO}, \text{OR}, \text{RO}\} &= \Pr\{\text{OO}\} + \Pr\{\text{OR}\} + \Pr\{\text{RO}\} \\ &= 3/4 .\end{aligned}$$

Alternatywnie, skoro prawdopodobieństwo otrzymania mniej niż jednego orła wynosi $\Pr\{\text{RR}\} = 1/4$, to prawdopodobieństwo otrzymania co najmniej jednego orła wynosi $1 - 1/4 = 3/4$.

Dyskretne rozkłady prawdopodobieństwa

Rozkład prawdopodobieństwa jest ***dyskretny***, jeśli jest zdefiniowany nad skończoną lub przeliczalną przestrzenią zdarzeń. Niech S będzie taką przestrzenią zdarzeń. Wówczas dla każdego zdarzenia A zachodzi

$$\Pr\{A\} = \sum_{s \in A} \Pr\{s\} ,$$

jako że zdarzenia elementarne, a w szczególności te w A, wzajemnie się wykluczają. Jeśli przestrzeń S jest skończona i prawdopodobieństwo każdego zdarzenia elementarnego $s \in S$ wynosi

$$\Pr\{s\} = 1/|S| ,$$

to wtedy mamy ***jednostajny rozkład prawdopodobieństwa*** na S. W takich wypadkach eksperyment jest często opisywany jako „losowy wybór elementu z S”.

Jako przykład rozważmy rzut ***monetą symetryczną***, w którym prawdopodobieństwo uzyskania orła jest takie samo jak prawdopodobieństwo uzyskania reszki, czyli $1/2$. Jeśli rzucimy monetą n razy, to otrzymamy jednostajny rozkład prawdopodobieństwa określony na przestrzeni zdarzeń $S = \{\text{O}, \text{R}\}^n$ – zbiorze mocy 2^n. Każde zdarzenie elementarne w S można reprezentować jako słowo długości n nad zbiorem $\{\text{O}, \text{R}\}$ i każde występuje z prawdopodobieństwem $1/2^n$. Zdarzenie

$$A = \{\text{wypadło dokładnie } k \text{ orłów i dokładnie } n-k \text{ reszek}\}$$

jest podzbiorem S rozmiaru $|A| = \binom{n}{k}$, jako że istnieje $\binom{n}{k}$ słów długości n nad $\{\text{O}, \text{R}\}$, w których O występuje dokładnie k razy. Prawdopodobieństwo zdarzenia A wynosi więc $\Pr\{A\} = \binom{n}{k}/2^n$.

Ciągły rozkład jednostajny

Ciągły rozkład jednostajny jest przykładem rozkładu prawdopodobieństwa, w którym nie wszystkie podzbiory przestrzeni zdarzeń są uważane za zdarzenia. Ciągły rozkład jednostajny jest zdefiniowany nad przedziałem domkniętym $[a, b]$ liczb rzeczywistych, gdzie $a < b$. Intuicyjnie, chcielibyśmy, aby każdy punkt z przedziału $[a, b]$ był „jednakowo prawdopodobny”. Jest nieprzeliczalnie wiele punktów, więc jeśli przyporządkujemy każdemu punktowi to samo skończone, dodatnie prawdopodobieństwo, nie możemy jednocześnie spełnić aksjomatów 2 i 3. Z tego powodu chcielibyśmy przyporządkować prawdopodobieństwo jedynie *niektórym* podzbiorom S w taki sposób, aby spełnić aksjomaty dla tych zdarzeń.

Dla dowolnego przedziału domkniętego $[c, d]$, gdzie $a \leq c \leq d \leq b$, ***ciągły rozkład jednostajny prawdopodobieństwa*** definiuje prawdopodobieństwo zdarzenia $[c, d]$ wzorem

$$\Pr\{[c, d]\} = \frac{d-c}{b-a}.$$

Przyjmując $c = d$ widzimy, że prawdopodobieństwo dla pojedynczego punktu wynosi 0. Jeśli usuniemy końcowe punkty przedziału $[c, d]$, to otrzymamy przedział otwarty (c, d). Ponieważ $[c, d] = [c, c] \cup (c, d) \cup [d, d]$, na podstawie aksjomatu 3 dostajemy $\Pr\{[c, d]\} = \Pr\{(c, d)\}$. Ogólnie, zdarzeniem dla ciągłego rozkładu jednostajnego jest dowolny podzbiór przestrzeni zdarzeń $[a, b]$, który można otrzymać jako skończoną lub przeliczalną sumę przedziałów otwartych i domkniętych, jak również pewne bardziej skomplikowane zbiory.

Prawdopodobieństwo warunkowe i niezależność

Czasami dysponujemy już pewną wiedzą na temat wyniku danego doświadczenia. Przypuśćmy na przykład, że kolega rzucił dwiema monetami i powiedział nam, że co najmniej jedna z monet upadła orłem do góry. Jakie jest prawdopodobieństwo, że na obu monetach wypadł orzeł? Informacja, którą otrzymaliśmy, eliminuje możliwość pojawienia się dwóch reszek. Trzy pozostałe zdarzenia elementarne są jednakowo prawdopodobne, skąd wnioskujemy, że każde występuje z prawdopodobieństwem $1/3$. Ponieważ tylko w jednym ze zdarzeń elementarnych wypadają dwa orły, odpowiedź na nasze pytanie brzmi $1/3$.

Prawdopodobieństwo warunkowe określa w formalny sposób pojęcie posiadania *a priori* częściowej wiedzy na temat wyniku eksperymentu. ***Prawdopodobieństwo warunkowe*** zdarzenia A w sytuacji, kiedy wiemy, że zachodzi inne zdarzenie B, jest zdefiniowane następująco:

$$\Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}}, \tag{C.16}$$

kiedy $\Pr\{B\} \neq 0$. (Zapis „$\Pr\{A \mid B\}$” czytamy jako „prawdopodobieństwo A pod warunkiem B”). Intuicyjnie, ponieważ wiemy, że B zachodzi, więc zdarzenie, że A również zachodzi, to $A \cap B$. Oznacza to, że $A \cap B$ jest zbiorem takich wyników, że zachodzi zarówno A, jak i B. Skoro wynik jest jednym ze zdarzeń elementarnych w B, normalizujemy prawdopodobieństwa wszystkich zdarzeń elementarnych w B, dzieląc je przez $\Pr\{B\}$, żeby ich suma wynosiła 1. Prawdopodobieństwo warunkowe zdarzenia A pod warunkiem B jest zatem stosunkiem prawdopodobieństwa zdarzenia $A \cap B$ do prawdopodobieństwa zdarzenia B. W powyższym przykładzie A jest zdarzeniem, w którym na obu monetach wypadł orzeł, a B jest zdarzeniem, w którym co najmniej na jednej monecie wypadł orzeł. Zatem $\Pr\{A \mid B\} = (1/4)/(3/4) = 1/3$.

Dwa zdarzenia są ***niezależne***, jeśli

$$\Pr\{A \cap B\} = \Pr\{A\}\Pr\{B\}, \tag{C.17}$$

co jest równoważne, o ile $\Pr\{B\} \neq 0$, warunkowi

$$\Pr\{A \mid B\} = \Pr\{A\}.$$

Przypuśćmy na przykład, że zostały rzucone dwie monety i wyniki rzutów są niezależne. Wówczas prawdopodobieństwo pojawienia się dwóch orłów wynosi $(1/2)(1/2) = 1/4$. Teraz przypuśćmy, że jedno zdarzenie polega na tym, że pierwsza moneta upadła orłem do góry, a drugie na tym, że monety upadły różnie. Każde z tych zdarzeń występuje z prawdopodobieństwem $1/2$, a prawdopodobieństwo, że wystąpią oba zdarzenia, wynosi $1/4$; zatem, na mocy definicji niezależności, zdarzenia te są niezależne – chociaż mogłoby się wydawać, że oba zdarzenia zależą od wyniku dla pierwszej monety. Przypuśćmy wreszcie, że monety są sklejone tak, że upadają obie orłem lub reszką do góry i że obydwie te możliwości są jednakowo prawdopodobne. Wówczas prawdopodobieństwo, że każda z monet upadnie orłem do góry wynosi $1/2$, ale prawdopodobieństwo, że obydwie upadną orłem do góry wynosi $1/2 \neq (1/2)(1/2)$. Zatem zdarzenie, że pierwsza moneta upadnie orłem do góry, i zdarzenie, że druga moneta upadnie orłem do góry, nie są niezależne.

Zdarzenia należące do rodziny zdarzeń $A_1, A_2, \ldots, A_n$ są ***parami niezależne***, jeśli

$$\Pr\{A_i \cap A_j\} = \Pr\{A_i\}\Pr\{A_j\}$$

dla wszystkich $1 \le i < j \le n$. Mówimy, że zdarzenia te są ***(wzajemnie) niezależne***, jeśli każdy k-podzbiór $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ tej rodziny, gdzie $2 \le k \le n$ i $1 \le i_1 < i_2 < \cdots < i_k \le n$, spełnia zależność

$$\Pr\{A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\} = \Pr\{A_{i_1}\}\Pr\{A_{i_2}\}\cdots\Pr\{A_{i_k}\}.$$

Przypuśćmy na przykład, że rzucamy dwiema monetami. Niech A_1 będzie zdarzeniem polegającym na tym, że na pierwszej monecie wypadł orzeł, niech A_2 będzie zdarzeniem polegającym na tym, że na drugiej monecie wypadł orzeł, i niech A_3 będzie zdarzeniem polegającym na tym, że na każdej z monet wypadło co innego. Mamy

$$\begin{aligned}
\Pr\{A_1\} &= 1/2,\\
\Pr\{A_2\} &= 1/2,\\
\Pr\{A_3\} &= 1/2,\\
\Pr\{A_1 \cap A_2\} &= 1/4,\\
\Pr\{A_1 \cap A_3\} &= 1/4,\\
\Pr\{A_2 \cap A_3\} &= 1/4,\\
\Pr\{A_1 \cap A_2 \cap A_3\} &= 0.
\end{aligned}$$

Ponieważ dla $1 \le i < j \le 3$ mamy $\Pr\{A_i \cap A_j\} = \Pr\{A_i\}\Pr\{A_j\} = 1/4$, zdarzenia A_1, A_2 i A_3 są parami niezależne. Zdarzenia te nie są jednak wzajemnie niezależne, gdyż $\Pr\{A_1 \cap A_2 \cap A_3\} = 0$ i $\Pr\{A_1\}\Pr\{A_2\}\Pr\{A_3\} = 1/8 \neq 0$.

Wzór Bayesa

Z definicji prawdopodobieństwa warunkowego (C.16) i z prawa przemienności $A \cap B = B \cap A$ wynika, iż dla dwóch zdarzeń A i B, każde o niezerowym prawdopodobieństwie, zachodzi

$$\begin{aligned}
\Pr\{A \cap B\} &= \Pr\{B\}\Pr\{A \mid B\} \qquad (C.18)\\
&= \Pr\{A\}\Pr\{B \mid A\}.
\end{aligned}$$

Rozwiązując to równanie ze względu na $\Pr\{A \mid B\}$, otrzymujemy tożsamość

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{B\}} \tag{C.19}$$

znaną jako ***wzór Bayesa***. Mianownik $\Pr\{B\}$ jest stałą normalizującą, którą inaczej można zapisać w następujący sposób. Ponieważ $B = (B \cap A) \cup (B \cap \overline{A})$, a $B \cap A$ i $B \cap \overline{A}$ są zdarzeniami wzajemnie się wykluczającymi, więc

$$\begin{aligned}\Pr\{B\} &= \Pr\{B \cap A\} + \Pr\left\{B \cap \overline{A}\right\} \\ &= \Pr\{A\}\Pr\{B \mid A\} + \Pr\left\{\overline{A}\right\}\Pr\left\{B \mid \overline{A}\right\}.\end{aligned}$$

Podstawiając to do równości (C.19), uzyskujemy równoważną postać wzoru Bayesa:

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{A\}\Pr\{B \mid A\} + \Pr\left\{\overline{A}\right\}\Pr\left\{B \mid \overline{A}\right\}}. \tag{C.20}$$

Wzór Bayesa może uprościć obliczanie prawdopodobieństw warunkowych. Przypuśćmy na przykład, że mamy monetę symetryczną i monetę fałszywą, która zawsze upada orłem do góry. Nasze doświadczenie składa się z trzech niezależnych zdarzeń: jedną z dwóch monet wybieramy losowo, rzucamy raz tą monetą, a następnie rzucamy ponownie. Przypuśćmy, że na wybranej monecie wypadł orzeł przy obydwu próbach. Jakie jest prawdopodobieństwo, że jest to ta fałszywa?

Rozwiążemy ten problem, posługując się wzorem Bayesa. Niech A będzie zdarzeniem polegającym na tym, że została wybrana moneta fałszywa, i niech B będzie zdarzeniem polegającym na tym, że moneta upadnie dwukrotnie orłem do góry. Chcemy znaleźć $\Pr\{A \mid B\}$. Mamy $\Pr\{A\} = 1/2$, $\Pr\{B \mid A\} = 1$, $\Pr\left\{\overline{A}\right\} = 1/2$ i $\Pr\left\{B \mid \overline{A}\right\} = 1/4$; tak więc

$$\begin{aligned}\Pr\{A \mid B\} &= \frac{(1/2)\cdot 1}{(1/2)\cdot 1 + (1/2)\cdot(1/4)} \\ &= 4/5.\end{aligned}$$

Zadania

C.2-1

Profesor Rosencrantz rzuca monetą (symetryczną) dwa razy. Profesor Guildenstern rzuca monetą (symetryczną) jeden raz. Jakie jest prawdopodobieństwo, że profesor Rosencrantz wyrzuci więcej orłów niż profesor Guildenstern?

C.2-2

Udowodnij ***nierówność Boole'a***: dla każdej skończonej lub przeliczalnej rodziny zdarzeń $A_1, A_2, \ldots$ zachodzi

$$\Pr\{A_1 \cup A_2 \cup \cdots\} \le \Pr\{A_1\} + \Pr\{A_2\} + \cdots . \tag{C.21}$$

C.2-3

Mamy 10 kart, każda z innym numerem od 1 do 10. Dokładnie tasujemy karty, po czym wyciągamy trzy (po kolei). Jakie jest prawdopodobieństwo, że te trzy karty będą miały coraz większe numery?

C.2-4

Udowodnij, że

$$\Pr\{A \mid B\} + \Pr\left\{\overline{A} \mid B\right\} = 1.$$

C.2-5

Udowodnij, że dla dowolnej rodziny zdarzeń $A_1, A_2, \ldots, A_n$ zachodzi

$$\begin{aligned}\Pr\{A_1 \cap A_2 \cap \cdots \cap A_n\} &= \Pr\{A_1\} \cdot \Pr\{A_2 \mid A_1\} \cdot \Pr\{A_3 \mid A_1 \cap A_2\} \cdots \\ &\quad \cdots \Pr\{A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}\}.\end{aligned} \tag{C.22}$$

★ ***C.2-6***

Pokaż, jak skonstruować zbiór n zdarzeń, które są parami niezależne, ale żaden podzbiór $k > 2$ spośród nich nie jest wzajemnie niezależny.

★ ***C.2-7***

Dwa zdarzenia A i B są ***warunkowo niezależne*** od C, jeśli

$$\Pr\{A \cap B \mid C\} = \Pr\{A \mid C\} \cdot \Pr\{B \mid C\}.$$

Podaj prosty, ale nietrywialny przykład dwóch zdarzeń, które nie są niezależne, ale są warunkowo niezależne od trzeciego zdarzenia.

★ ***C.2-8***

Profesor Gore prowadzi zajęcia z rytmiki w szkole muzycznej, a trzem jego uczniom – Jeffowi, Timowi i Carmine – grozi niezaliczenie. Profesor poinformował całą trójkę, że jeden z nich zaliczy kurs, a dwóch pozostałych go obleje. Carmine pyta profesora na osobności, kto z pary Jeff i Tim obleje kurs, argumentując, że skoro i tak wie, że co najmniej jeden z nich obleje, profesor nie ujawni żadnej informacji o jego własnym wyniku. Profesor powiedział mu, że Jeff obleje. Carmine czuje się teraz spokojniejszy, bo wie, że albo on, albo Tim zaliczy kurs, co oznacza, że prawdopodobieństwo jego pozytywnego wyniku wynosi teraz $1/2$. Czy ma rację, czy też jego szanse nadal wynoszą $1/3$? Odpowiedź uzasadnij.

C.3 Dyskretne zmienne losowe

(Dyskretna) zmienna losowa X jest funkcją ze skończonej lub przeliczalnej przestrzeni zdarzeń S w zbiór liczb rzeczywistych. Przyporządkowuje ona liczbę rzeczywistą każdemu możliwemu wynikowi doświadczenia, co pozwala nam operować indukowanym rozkładem prawdopodobieństwa na zbiorze liczb. Zmienne losowe można również określać na nieprzeliczalnych przestrzeniach zdarzeń, ale wiąże się to z kwestiami technicznymi, których nie chcemy tutaj poruszać. Będziemy zatem zakładać, że zmienne losowe są dyskretne.

Dla zmiennej losowej X i liczby rzeczywistej x definiujemy zdarzenie $X = x$ jako $\{s \in S : X(s) = x\}$; zatem

$$\Pr\{X = x\} = \sum_{s \in S :\, X(s) = x} \Pr\{s\}.$$

Funkcja

$$f(x) = \Pr\{X = x\}$$

jest ***funkcją gęstości prawdopodobieństwa*** zmiennej losowej X. Z aksjomatów prawdopodobieństwa mamy $\Pr\{X = x\} \geq 0$ i $\sum_x \Pr\{X = x\} = 1$.

Jako przykład rozważmy rzut dwiema sześciennymi kostkami do gry. W przestrzeni zdarzeń jest 36 możliwych zdarzeń elementarnych. Przyjmujemy, że rozkład prawdopodobieństwa jest jednostajny, więc każde zdarzenie elementarne $s \in S$ jest jednakowo prawdopodobne: $\Pr\{s\} = 1/36$. Zdefiniujmy zmienną losową X jako *większą* z dwóch wartości widocznych na kostkach. Mamy $\Pr\{X = 3\} = 5/36$, bo X przypisuje wartość 3 pięciu z trzydziestu sześciu możliwych zdarzeń elementarnych, a mianowicie $(1,3)$, $(2,3)$, $(3,3)$, $(3,2)$ i $(3,1)$.

Często na tej samej przestrzeni zdarzeń definiuje się wiele zmiennych losowych. Jeśli X i Y są zmiennymi losowymi, to funkcja

$$f(x, y) = \Pr\{X = x \text{ i } Y = y\}$$

jest ***funkcją gęstości łącznego prawdopodobieństwa*** zmiennych X i Y. Dla ustalonej wartości y

$$\Pr\{Y = y\} = \sum_x \Pr\{X = x \text{ i } Y = y\};$$

podobnie, dla ustalonej wartości x

$$\Pr\{X = x\} = \sum_y \Pr\{X = x \text{ i } Y = y\}.$$

Korzystając z definicji prawdopodobieństwa warunkowego (C.16) na str. 1115, mamy

$$\Pr\{X = x \mid Y = y\} = \frac{\Pr\{X = x \text{ i } Y = y\}}{\Pr\{Y = y\}}.$$

Definiujemy dwie zmienne losowe X i Y jako ***niezależne***, jeśli dla wszystkich x i y zdarzenia $X = x$ i $Y = y$ są niezależne lub, równoważnie, jeśli dla wszystkich x i y mamy $\Pr\{X = x \text{ i } Y = y\} = \Pr\{X = x\}\Pr\{Y = y\}$.

Mając dany zbiór zmiennych losowych zdefiniowanych nad tą samą przestrzenią zdarzeń, możemy definiować nowe zmienne losowe jako sumy, iloczyny lub inne funkcje zmiennych początkowych.

Wartość oczekiwana zmiennej losowej

Najprostszą i najbardziej użyteczną charakterystyką rozkładu zmiennej losowej jest „średnia" z wartości, jakie ona przyjmuje. ***Wartość oczekiwana (średnia)*** dyskretnej zmiennej losowej X jest określona wzorem

$$\mathrm{E}[X] = \sum_x x \cdot \Pr\{X = x\}; \tag{C.23}$$

wielkość ta jest dobrze określona, jeśli suma jest skończona lub zbieżna bezwględnie. Czasami wartość oczekiwaną X zapisuje się jako μ_X lub – kiedy zmienna losowa, o którą chodzi, wynika z kontekstu – po prostu jako μ.

Wyobraźmy sobie grę, w której rzucamy dwiema monetami. Zarabiamy 3 zł za każdego orła, ale tracimy 2 zł za każdą reszkę. Wartość oczekiwana zmiennej losowej X reprezentującej nasz zarobek wynosi

$$\begin{aligned}
\mathrm{E}[X] &= 6 \cdot \Pr\{2\ \textsc{o}\} + 1 \cdot \Pr\{1\ \textsc{o}, 1\ \textsc{r}\} - 4 \cdot \Pr\{2\ \textsc{r}\} \\
&= 6(1/4) + 1(1/2) - 4(1/4) \\
&= 1.
\end{aligned}$$

Wartość oczekiwana sumy dwóch zmiennych losowych jest równa sumie ich wartości oczekiwanych, czyli

$$\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y], \tag{C.24}$$

kiedy tylko są określone $\mathrm{E}[X]$ i $\mathrm{E}[Y]$. Własność tę nazywamy ***liniowością wartości oczekiwanej*** i zachodzi ona nawet wtedy, kiedy X i Y nie są niezależne. Rozszerza się ona również na skończone i zbieżne bezwzględnie sumy wartości oczekiwanych. Liniowość wartości oczekiwanej jest kluczową własnością umożliwiającą nam przeprowadzanie analiz probabilistycznych z zastosowaniem wskaźnikowych zmiennych losowych (patrz podrozdz. 5.2).

Jeśli X jest dowolną zmienną losową, to dowolna funkcja $g(x)$ definiuje nową zmienną losową $g(X)$. Jeśli wartość oczekiwana $g(X)$ jest określona, to

$$\mathrm{E}[g(X)] = \sum_x g(x) \cdot \Pr\{X = x\}.$$

Przyjmując $g(x) = ax$, dla dowolnej stałej a mamy

$$\mathrm{E}[aX] = a\mathrm{E}[X]. \tag{C.25}$$

Operacja brania wartości oczekiwanej jest zatem liniowa: dla dowolnych dwóch zmiennych losowych X i Y oraz dowolnej stałej a zachodzi

$$\mathrm{E}[aX + Y] = a\mathrm{E}[X] + \mathrm{E}[Y]. \tag{C.26}$$

Jeśli dwie zmienne losowe X i Y są niezależne i każda z nich ma określoną wartość oczekiwaną, to

$$\begin{aligned}
\mathrm{E}[XY] &= \sum_x \sum_y xy \cdot \Pr\{X = x \text{ i } Y = y\} \\
&= \sum_x \sum_y xy \cdot \Pr\{X = x\} \Pr\{Y = y\} && \text{(z niezależności } X \text{ i } Y) \\
&= \left(\sum_x x \cdot \Pr\{X = x\} \right) \left(\sum_y y \cdot \Pr\{Y = y\} \right) \\
&= \mathrm{E}[X]\,\mathrm{E}[Y] && \text{(ze wzoru (C.23)).}
\end{aligned}$$

Ogólnie, jeśli n zmiennych losowych $X_1, X_2, \ldots, X_n$ jest wzajemnie niezależnych, to

$$\mathrm{E}[X_1 X_2 \cdots X_n] = \mathrm{E}[X_1]\,\mathrm{E}[X_2] \cdots \mathrm{E}[X_n]. \tag{C.27}$$

Kiedy zmienna losowa X przyjmuje wartości ze zbioru liczb naturalnych $\mathbb{N} = \{0, 1, 2, \ldots\}$, istnieje zgrabny wzór na jej wartość oczekiwaną:

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr\{X = i\} \\
&= \sum_{i=0}^{\infty} i(\Pr\{X \geq i\} - \Pr\{X \geq i + 1\}) \\
&= \sum_{i=1}^{\infty} \Pr\{X \geq i\}, && \text{(C.28)}
\end{aligned}$$

ponieważ każdy składnik $\Pr\{X \geq i\}$ jest dodany i razy i odjęty $i - 1$ razy (oprócz składnika $\Pr\{X \geq 0\}$, który jest dodany 0 razy i nie odjęty w ogóle).

Funkcja $f(x)$ jest ***wypukła***, jeśli dla każdych x i y oraz dla wszystkich $0 \leq \lambda \leq 1$ zachodzi

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y). \tag{C.29}$$

Stosując funkcję wypukłą $f(x)$ do zmiennej losowej X, z ***nierówności Jensena*** otrzymujemy

$$\mathrm{E}[f(x)] \geq f(\mathrm{E}[X]), \tag{C.30}$$

pod warunkiem że obydwie wartości oczekiwane istnieją i są skończone.

Wariancja i odchylenie standardowe

Wartość oczekiwana zmiennej losowej nie mówi nam nic o „rozrzucie" wartości przyjmowanych przez tę zmienną. Jeśli na przykład mamy zmienne losowe X i Y, dla których $\Pr\{X = 1/4\} = \Pr\{X = 3/4\} = 1/2$ oraz $\Pr\{Y = 0\} = \Pr\{Y = 1\} = 1/2$, to zarówno $\mathrm{E}[X]$, jak i $\mathrm{E}[Y]$ jest równe $1/2$, ale faktyczne wartości przyjmowane przez Y są dalej od wartości oczekiwanej niż wartości przyjmowane przez X.

Pojęcie wariancji wyraża matematycznie to, jak daleko od średniej mogą być wartości przyjmowane przez zmienną losową. ***Wariancja*** zmiennej losowej X o wartości oczekiwanej $\mathrm{E}[X]$ jest równa

$$\begin{aligned} \mathrm{Var}[X] &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\ &= \mathrm{E}\left[X^2 - 2X\mathrm{E}[X] + \mathrm{E}^2[X]\right] \\ &= \mathrm{E}\left[X^2\right] - 2\mathrm{E}[X\mathrm{E}[X]] + \mathrm{E}^2[X] \\ &= \mathrm{E}\left[X^2\right] - 2\mathrm{E}^2[X] + \mathrm{E}^2[X] \\ &= \mathrm{E}\left[X^2\right] - \mathrm{E}^2[X]. \end{aligned} \tag{C.31}$$

Uzasadnieniem równości $\mathrm{E}[\mathrm{E}^2[X]] = \mathrm{E}^2[X]$ jest to, że $\mathrm{E}[X]$ nie jest zmienną losową, a po prostu liczbą rzeczywistą, zatem $\mathrm{E}^2[X]$ jest także liczbą rzeczywistą. Równość $\mathrm{E}[X\mathrm{E}[X]] = \mathrm{E}^2[X]$ wynika z zależności (C.25) z $a = \mathrm{E}[X]$. Równość (C.31) można przekształcić, aby uzyskać wyrażenie określające wartość oczekiwaną kwadratu zmiennej losowej:

$$\mathrm{E}\left[X^2\right] = \mathrm{Var}[X] + \mathrm{E}^2[X]. \tag{C.32}$$

Wariancja zmiennej losowej X i wariancja aX są związane następującą zależnością (patrz zad. C.3-10):

$$\mathrm{Var}[aX] = a^2\mathrm{Var}[X].$$

Jeśli X i Y są niezależnymi zmiennymi losowymi, to

$$\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y].$$

W ogólności, jeśli n zmiennych losowych $X_1, X_2, \ldots, X_n$ jest parami niezależnych, to

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i]. \tag{C.33}$$

Odchylenie standardowe zmiennej losowej X jest to nieujemny pierwiastek kwadratowy z wariancji X. Odchylenie standardowe zmiennej losowej X czasami zapisuje się jako σ_X lub po prostu σ, jeśli zmienna X wynika z kontekstu. Używając tej notacji, wariancję X można zapisać jako σ^2.

Zadania

C.3-1

Rzucamy dwiema sześciennymi kostkami do gry. Jaka jest wartość oczekiwana sumy liczb oczek widocznych na dwóch kostkach? Jaka jest wartość oczekiwana większej z liczb oczek widocznych na tych kostkach?

C.3-2
Tablica $A[1:n]$ zawiera n różnych liczb uporządkowanych losowo, tzn. każda permutacja tych n liczb jest jednakowo prawdopodobna. Jaka jest wartość oczekiwana indeksu największego z elementów tablicy? Jaka jest wartość oczekiwana indeksu najmniejszego z elementów tablicy?

C.3-3
Gra polega na rzucaniu trzema kostkami. Gracz może postawić 1 zł na jedną z liczb od 1 do 6. Zasady wypłacania wygranych po wykonaniu rzutu są następujące: jeśli liczba obstawiona przez gracza nie pojawiła się na żadnej z kostek, to traci on swoją złotówkę; jeśli liczba ta pojawiła się dokładnie na k kostkach, dla $k = 1, 2, 3$, to zachowuje on swoją złotówkę i dostaje dodatkowo k złotówek. Jaki jest oczekiwany zysk w jednej grze?

C.3-4
Uzasadnij, że jeśli X i Y są nieujemnymi zmiennymi losowymi, to

$$\mathrm{E}[\max\{X, Y\}] \le \mathrm{E}[X] + \mathrm{E}[Y].$$

★ ***C.3-5***
Niech X i Y będą niezależnymi zmiennymi losowymi. Udowodnij, że $f(X)$ i $g(Y)$ są niezależne dla dowolnego wyboru funkcji f i g.

★ ***C.3-6***
Niech X będzie nieujemną zmienną losową i załóżmy, że $\mathrm{E}[X]$ jest dobrze określona. Udowodnij ***nierówność Markowa***:

$$\Pr\{X \ge t\} \le \mathrm{E}[X]/t \tag{C.34}$$

dla każdego $t > 0$.

★ ***C.3-7***
Niech S będzie przestrzenią zdarzeń i niech X i X' będą zmiennymi losowymi takimi, że $X(s) \ge X'(s)$ dla każdego $s \in S$. Udowodnij, że dla dowolnej stałej rzeczywistej t zachodzi nierówność

$$\Pr\{X \ge t\} \ge \Pr\{X' \ge t\}.$$

C.3-8
Co jest większe: wartość oczekiwana kwadratu zmiennej losowej czy też kwadrat jej wartości oczekiwanej?

C.3-9
Pokaż, że dla dowolnej zmiennej losowej X, która przyjmuje wyłącznie wartości 0 i 1, mamy $\mathrm{Var}[X] = \mathrm{E}[X]\,\mathrm{E}[1 - X]$.

C.3-10
Udowodnij, że $\mathrm{Var}[aX] = a^2\mathrm{Var}[x]$ z definicji wariancji (C.31).