Elementy rachunku prawdopodobieństwa

Podstawowe pojęcia z kombinatoryki

Niech będzie dany zbiór *n* różnych elementów $\{a_1, a_2, a_3, ..., a_n\}$.

Permutacją bez powtórzeń nazywamy każde uporządkowanie wszystkich elementów tego zbioru, w którym każdy element występuje tylko raz. Liczba wszystkich permutacji tego zbioru jest równa:

$$P_n = n! = 1 \cdot 2 \cdot 3 \cdot ... \cdot n$$

Permutacją z powtórzeniami nazywamy zbiór składający się z n uporządkowanych elementów, wśród których pewne elementy powtarzają się n_1, n_2, ..., n_k razy. Liczbę permutacji z powtórzeniami wyznaczymy ze wzoru:

$$P_n^{n_1, n_2, \ldots, n_k} = \frac{n!}{(n_1)!(n_2)!\ldots(n_k)!}$$

Kombinacja bez powtórzeń to podzbiór *k* różnych elementów wybranych spośród *n* wszystkich elementów bez uwzględniania ich porządku. Liczba kombinacji bez powtórzeń jest równa:

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

przy czym 0!=1

Kombinacje te można traktować jako nieuporządkowane *k*-elementowe próbki z *n*-elementowej populacji, pobierane bez zwracania elementów, które różnią się wyłącznie składem elementów, a nie ich porządkiem.

Obliczyć prawdopodobieństwo wygranej w popularnej grze liczbowej toto-lotek. Gra ta polega na wylosowaniu 6 z 49 liczb. Zdarzeniem losowym jest wynik losowania 6 z 49 liczb. Elementami przestrzeni zdarzeń elementarnych są tu wszystkie sześcioelementowe podzbiory zbioru {1,2,3,…,49}. Liczba elementów przestrzeni jest równa:

$$\binom{49}{6} = \frac{49!}{6!(49-6)!} = 13983816$$

Na podstawie klasycznej definicji prawdopodobieństwa obliczymy prawdopodobieństwa następujących zdarzeń losowych:

A_1 – wylosowanie „trójki"

A_2 – wylosowanie „czwórki"

A_3 – wylosowanie „piątki"

A_4 – wylosowanie „szóstki"

$$P(A_1) = \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}} = 0{,}01765 \qquad P(A_2) = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}} = 0{,}0009686$$

$$P(A_3) = \frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}} = 0{,}00001845 \qquad P(A_4) = \frac{\binom{6}{6}\binom{43}{0}}{\binom{49}{6}} = 0{,}0000000715$$

Kombinacja z powtórzeniami to zbiór *k*-elementów, które się między sobą różnią (albo się nie różnią), wybranych spośród *n* elementów bez zachowania porządku. Liczba takich kombinacji wyrażona jest wzorem:

$$\overline{C}_n^k = \binom{n+k-1}{k}$$

Kombinacje z powtórzeniami można traktować jako k-elementowe próby pobierane z n-elementowej populacji ze zwracaniem, przy czym nieważna jest kolejność elementów.

Doświadczenie nazwiemy losowym, jeżeli pomimo znajomości warunków, w których ono się odbywa, nie jesteśmy w stanie przewidzieć jego wyniku.

W każdym doświadczeniu losowym można wyróżnić najprostsze zdarzenie, tzw. **zdarzenie elementarne**. Charakteryzuje się ono następującymi właściwościami:

- zdarzenie elementarne może zajść lub nie,
- jedno ze zdarzeń elementarnych na pewno zajdzie,
- w tym samym doświadczeniu zajście jednego zdarzenia wyklucza zajście innego zdarzenia.

Zbiór zdarzeń elementarnych dla danego doświadczenia losowego tworzy tzw. **przestrzeń zdarzeń elementarnych W.** Poszczególne elementy tej przestrzeni, czyli zdarzenia elementarne oznaczamy symbolem w. Przestrzeń zdarzeń elementarnych może składać się ze **skończonej liczby zdarzeń**, ale także może być **zbiorem nieskończonym przeliczalnym, jak też i nieprzeliczalnym**.

Wyniki jednokrotnego rzutu kostką do gry tworzą **skończoną przestrzeń** zdarzeń elementarnych, składającą się z 6 elementów (wypadnie "1", "2", "3", "4", "5", "6"). Wyniki oznaczeń wielkości skażeń chemicznych w warzywach są **nieskończoną i nieprzeliczalną** przestrzenią zdarzeń.

Przykładem **nieskończonej, ale przeliczalnej** przestrzeni zdarzeń elementarnych jest np. rzut monetą do pierwszego pojawienia się orła, w którym O oznacza otrzymanie orła w pierwszym rzucie, RO pojawienie się reszki w pierwszym rzucie i orła w drugim, RRO - otrzymanie w dwóch pierwszych rzutach reszki, a w trzecim po raz pierwszy orła itd.

Zdarzenie losowe jest podzbiorem przestrzeni zdarzeń elementarnych.

Z matematycznego punktu widzenia, zdarzenia są zbiorami. Można więc wykonywać na nich takie działania, jakie są wykonalne na zbiorach:

- **suma zdarzeń A i B** zapisywana jako $A \cup B$ oznacza zbiór zdarzeń elementarnych należących do A lub B. Niech **A** oznacza zdarzenie oznaczające pojawienie się ścianki (w kostce do gry) o parzystej liczbie oczek. Sprzyja temu zdarzeniu zdarzenie elementarne A={w_2, w_4, w_6}. Niech **B** oznacza zdarzenie, że pojawi się liczba oczek większa od trzech. Sprzyjają temu zdarzeniu zdarzenia elementarne B={w_4, w_5, w_6}. Sumą zdarzeń w tym przypadku są zdarzenia polegające na pojawieniu się 2, 4, 5, lub 6 oczek, czyli: $A \cup B$ ={w_2, w_4, w_6}∪{w_4, w_5, w_6}={w_2, w_4, w_5, w_6},

- **iloczyn zdarzeń A i B** (oznaczony symbolem $A \cap B$) to zdarzenie elementarne złożone z tych zdarzeń, które należą do zdarzenia A i zdarzenia B. Niech **A** oznacza zdarzenie polegające na tym, że z listy osób wylosujemy mężczyznę, zaś zdarzenie **B** polega na tym, że z listy osób wylosujemy osobę pracującą na stanowisku managera. Iloczyn zdarzeń oznacza wylosowanie z listy osób managera mężczyzny,

- **różnica zdarzeń A i B** (oznaczana symbolem A - B) to zbiór zdarzeń, które należą do zdarzenia A i nie należą do zdarzenia B. Niech **A** oznacza zdarzenie oznaczające pojawienie się ścianki (w kostce do gry) o nieparzystej liczbie oczek. Sprzyja temu zdarzeniu zdarzenie elementarne $A=\{w_1, w_3, w_5\}$. Niech **B** oznacza zdarzenie, że pojawi się liczba oczek większa od trzech. Sprzyja temu zdarzeniu zdarzenie elementarne $B=\{w_4, w_5, w_6\}$. Różnicą zdarzeń w tym przypadku jest zdarzenie polegające na pojawieniu się 1 lub 3, czyli:

  $A - B =\{w_1, w_3, w_5\} - \{w_4, w_5, w_6\}=\{w_1, w_3\}$.

Prawdopodobieństwo warunkowe

W rachunku prawdopodobieństwa bardzo ważną sprawą jest podział zdarzeń losowych na **zdarzenia niezależne i zależne**.

Zdarzenia losowe niezależne to takie, dla których zajście jednego ze zdarzeń nie ma wpływu na wystąpienie drugiego zdarzenia. Jeżeli natomiast zajście jednego zdarzenia warunkuje prawdopodobieństwo pojawienia się drugiego z nich, to zdarzenia takie nazywamy **zdarzeniami zależnymi** (warunkowymi, względnymi).

Prawdopodobieństwo zdarzenia A obliczone przy założeniu, że zaszło zdarzenie B, nazywamy **prawdopodobieństwem warunkowym zdarzenia A pod warunkiem B**. Prawdopodobieństwo warunkowe zdarzenia A pod warunkiem B oznacza się symbolem P(A/B) i oblicza się je ze wzoru:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

gdzie:

- **P(A/B)** - prawdopodobieństwo warunkowe zdarzenia A pod warunkiem zajścia zdarzenia B,
- $P(A \cap B)$ - prawdopodobieństwo iloczynu zdarzeń A i B,
- **P(B)** - prawdopodobieństwo zajścia zdarzenia B.

przy założeniu, że P(B)>0 czyli, że zdarzenie B nie jest zdarzeniem niemożliwym. Założenie to jest wystarczające, ponieważ jeśli zdarzenie B nie jest zdarzeniem niemożliwym, to prawdopodobieństwo jego zajścia jest większe od zera, ponieważ prawdopodobieństwo jest liczbą nieujemną (tzn. dodatnią lub równą zero).

Wzór na **prawdopodobieństwo iloczynu zdarzeń** (w oparciu o wzór przedstawiony powyżej) przedstawia się następująco:

$$P(A \cap B) = P(A)P(B / A), gdy\, P(A) > 0$$

$$P(A \cap B) = P(B)P(A / B), gdy\, P(B) > 0$$

gdzie:

- $P(A \cap B)$ - prawdopodobieństwo iloczynu zdarzeń A i B,
- **P(A)** - prawdopodobieństwo zajścia zdarzenia A,
- **P(B)** - prawdopodobieństwo zajścia zdarzenia B,
- **P(B/A)** - prawdopodobieństwo warunkowe zdarzenia B pod warunkiem zajścia zdarzenia A,
- **P(A/B)** - prawdopodobieństwo warunkowe zdarzenia A pod warunkiem zajścia zdarzenia B.

W urnie znajduje się 10 kul: 7 białych i 3 czerwone. **Należy obliczyć prawdopodobieństwo, że w dwóch kolejnych losowaniach obie kule będą białe.** Losowanie odbywa się bez zwracania kul do urny, tak więc po każdym losowaniu ubywa z urny jedna kula i zmienia się prawdopodobieństwo zajścia następnego zdarzenia.

Korzystamy ze wzoru na prawdopodobieństwo iloczynu zdarzeń. **Zdarzenie A** oznacza, że pierwsza z wylosowanych kul jest biała, a **zdarzenie B**, że druga jest biała.

$$P(A \cap B) = P(A)P(B/A) = \frac{7}{10} \cdot \frac{6}{9} = \frac{42}{90} = \frac{7}{15}$$

- **P(A)** oznacza prawdopodobieństwo wylosowania pierwszej białej kuli. Wynosi ono **7/10**, ponieważ w urnie znajduje się 7 białych kul, a wszystkich kul jest 10.
- **P(B/A)** oznacza prawdopodobieństwo wylosowania drugiej białej kuli pod warunkiem, że została wylosowana pierwsza biała kula. P(B/A) wynosi **6/9**, ponieważ po wylosowaniu pierwszej kuli w urnie pozostało już tylko 9 kul, w tym 6 białych.

Teraz chcemy dowiedzieć się, jakie jest prawdopodobieństwo wylosowania jako pierwszej kuli białej i jako drugiej kuli czerwonej. Tym razem **zdarzenie A** oznacza, że pierwsza z wylosowanych kul jest biała, a **zdarzenie B**, że druga jest czerwona.

$$P(A \cap B) = P(A)P(B/A) = \frac{7}{10} \cdot \frac{3}{9} = \frac{21}{90} = \frac{7}{30}$$

- **P(A)** oznacza prawdopodobieństwo wylosowania pierwszej białej kuli. Wynosi ono **7/10**, ponieważ w urnie znajduje się 7 białych kul, a wszystkich kul jest 10.
- **P(B/A)** oznacza prawdopodobieństwo wylosowania drugiej kuli, tym razem czerwonej pod warunkiem, że pierwsza wylosowana kula jest biała. P(B/A) wynosi **3/9**, ponieważ po wylosowaniu pierwszej kuli w urnie pozostało już tylko 9 kul, z których 3 są czerwone.

Prawdopodobieństwo całkowite i twierdzenie Bayesa

Przyjmijmy założenie, że zajście zdarzenia B jest uwarunkowane zajściem jednego spośród *n* wzajemnie wykluczających się zdarzeń A_i (gdzie *i* = 1, 2, ..., *n*), które zajmują całą przestrzeń zdarzeń elementarnych (przy czym $P(A_i)>0$). W tym przypadku prawdopodobieństwo zajścia zdarzenia B określamy jako **prawdopodobieństwo całkowite** i obliczamy z następującego wzoru:

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + \ldots + P(A_n)P(B/A_n) =$$

$$= \sum_{i=1}^{n} P(A_i)P(B/A_i)$$

gdzie:

- P(B) - prawdopodobieństwo całkowite zajścia zdarzenia B,
- $P(A_i)$ - prawdopodobieństwo zajścia zdarzenia A_i,
- $P(B/A_i)$ - prawdopodobieństwo zajścia zdarzenia B uwarunkowane zajściem jednego spośród wzajemnie wykluczających się zdarzeń A_i.

Prawdopodobieństwa $P(A_i)$ nazywane są prawdopodobieństwami **a priori** i odgrywają ważną rolę w zagadnieniach dotyczących podejmowania decyzji w warunkach niepewności. Przykładowo pewna firma otrzymuje dostawy od kilku dostawców i część każdej z dostaw zawiera wadliwe lub nie spełniające wymogów produkty. Wzory na prawdopodobieństwo warunkowe i całkowite pozwalają wyznaczyć dostawcę od którego pochodzi najwięcej wadliwych produktów i dzięki temu firma może podjąć decyzję z którym z dostawców kontynuować współpracę, a z którym tą współpracę przerwać.

Twierdzenie Bayesa

Po zrealizowaniu zdarzenia B interesuje nas jakie będzie prawdopodobieństwo każdego ze zdarzeń A_i. Odpowiedź na to pytanie uzyskamy stosując do obliczeń **wzór Bayesa**, który ma następującą postać:

$$P(A_i / B) = \frac{P(A_i)P(B / A_i)}{P(B)}$$

gdzie:

- P(B) - prawdopodobieństwo całkowite zajścia zdarzenia B,
- $P(A_i)$ - prawdopodobieństwo zajścia zdarzenia A_i,
- $P(B/A_i)$ - prawdopodobieństwo zajścia zdarzenia B uwarunkowane zajściem jednego spośród *n* wzajemnie wykluczających się zdarzeń $A_{i,}$
- $P(A_i/B)$ - prawdopodobieństwo zajścia zdarzenia A_i po zrealizowaniu zdarzenia B.

Firma produkująca samochody osobowe zamawia łożyska u trzech dostawców. Od **dostawcy I** firma otrzymuje **połowę** wszystkich zamawianych łożysk. Od **dostawcy II** - **30%**, natomiast od **dostawcy III - 20%**. Wśród łożysk dostarczanych przez **dostawcę I** znajduje się **3%** łożysk z ukrytymi wadami, **dostawca II** przekazuje **7%** wyrobów z wadami, natomiast **dostawca III** - **5%**. **Należy obliczyć prawdopodobieństwo dostarczenia łożyska posiadającego ukrytą wadę i prawdopodobieństwo, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy II.**

Przez B oznaczamy zdarzenie, że dostarczone łożysko posiada ukrytą wadę. Przez $\mathbf{A_1}$ oznaczmy zdarzenie, że łożysko pochodzi od dostawcy I, przez $\mathbf{A_2}$, że łożysko pochodzi od dostawcy II, a przez $\mathbf{A_3}$ - od dostawcy III.

$P(B/A_1)$ oznacza prawdopodobieństwo, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy I,
$P(B/A_2)$ oznacza prawdopodobieństwo, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy II, natomiast
$P(B/A_3)$ oznacza prawdopodobieństwo, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy III.

Łożysko z ukrytą wadą może pochodzić od I dostawcy, który realizuje **50%** zamówień (z których **3%** łożysk jest wadliwych), może też pochodzić od drugiego dostawcy, który realizuje **30%** zamówień, z których **7%** jest wadliwych lub też może pochodzić od dostawcy III, który dostarcza **20%** ogółu wyrobów, wśród których jest **5%** wyrobów wadliwych.

Prawdopodobieństwo dostarczenia łożyska z ukrytą wadą obliczymy zatem ze wzoru na prawdopodobieństwo całkowite:

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + P(A_3)P(B/A_3) =$$
$$= 0{,}5 \cdot 0{,}03 + 0{,}3 \cdot 0{,}07 + 0{,}2 \cdot 0{,}05 = 0{,}046$$

W celu obliczenia prawdopodobieństwa, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy II należy skorzystać ze **wzoru Bayesa**. Dostawca II realizuje **30%** ogółu zamówień, z których **7%** jest wadliwych, zaś prawdopodobieństwo dostarczenia łożyska z ukrytą wadą jest równe **0,046**.

Prawdopodobieństwo, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy II obliczymy więc:

$$P(A_2 / B) = \frac{P(A_2)P(B / A_2)}{P(B)} = \frac{0,3 \cdot 0,07}{0,046} = 0,457$$

Gdybyśmy chcieli znać prawdopodobieństwo zdarzenia takiego, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy III, liczylibyśmy w następujący sposób:

$$P(A_3 / B) = \frac{P(A_3)P(B / A_3)}{P(B)} = \frac{0{,}2 \cdot 0{,}05}{0{,}046} = 0{,}217$$

Natomiast prawdopodobieństwo zdarzenia takiego, że dostarczone łożysko z ukrytą wadą pochodzi od dostawcy I, jest równe:

$$P(A_1 / B) = \frac{P(A_1)P(B / A_1)}{P(B)} = \frac{0,5 \cdot 0,03}{0,046} = 0,326$$

Podsumowując, prawdopodobieństwo dostarczenia łożyska z ukrytą wadą przez dostawcę II jest największe i wynosi 0,457. Ten dostawca dostarcza również relatywnie najwięcej wyrobów wadliwych.