Wstęp do rachunku prawdopodobieństwa
Wykład 3: Prawdopodobieństwo całkowite. Wzór Bayesa.

Wzór łańcuchowy

Twierdzenie 1 (Wzór łańcuchowy). *Jeżeli zdarzenia* $A_1, A_2, \ldots, A_n$ *spełniają warunek* $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$*, to*

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

dowód. *Przeprowadzimy dowód indukcyjny ze względu na* n*. Dla* $n = 1$ *równanie oczywiście zachodzi, a dla* $n = 2$ *wynika z definicji prawdopodobieństwa warunkowego*

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

Załóżmy teraz, że twierdzenie zachodzi dla pewnego $n \geqslant 2$*. Niech zdarzenia* $A_1, A_2, \ldots, A_{n+1}$ *spełniają warunek* $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) > 0$*. W szczególności mamy* $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$*. Stosując najpierw definicję prawdopodobieństwa warunkowego do zdarzeń* $A_1 \cap A_2 \cap \ldots \cap A_n$ *i* A_{n+1}*, a następnie założenie indukcyjne do* $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n)$*, otrzymujemy:*

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n+1}) &= \mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) \cdot \mathbb{P}(A_{n+1}|A_1 \cap A_2 \cap \ldots \cap A_n) \\ &= \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \ldots \cdot \mathbb{P}(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \cdot \mathbb{P}(A_{n+1}|A_1 \cap A_2 \cap \ldots \cap A_n).\end{aligned}$$

□

Przykład 1. Z urny zawierającej 3 białe i 3 czarne kule losujemy trzykrotnie jedną kulę, przy czym jeśli wylosujemy kulę białą, to wrzucamy ją z powrotem do urny i dodajemy jeszcze jedną białą kulę, a jeśli wylosujemy kulę czarną, to odrzucamy ją na bok. Jakie jest prawdopodobieństwo wylosowania za każdym razem białej kuli?

B_i – w i-tym losowaniu wylosowaliśmy białą kulę, $i = 1, 2, 3$

$B_1 \cap B_2 \cap B_3$ – za każdym razem wylosowaliśmy białą kulę

Ze wzoru łańcuchowego otrzymujemy:

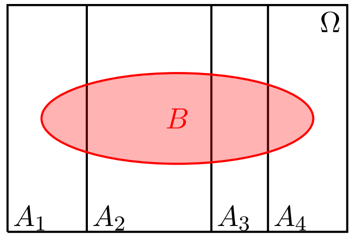
$$\mathbb{P}(B_1 \cap B_2 \cap B_3) = \mathbb{P}(B_1) \cdot \mathbb{P}(B_2|B_1) \cdot \mathbb{P}(B_3|B_1 \cap B_2) = \frac{3}{6} \cdot \frac{4}{7} \cdot \frac{5}{8} = \frac{5}{28}.$$

$$\boxed{\text{3B 3C}} \xrightarrow[3/6]{B_1} \boxed{\text{4B 3C}} \xrightarrow[4/7]{B_2} \boxed{\text{5B 3C}} \xrightarrow[5/8]{B_3} \boxed{\text{6B 3C}}$$

Uwaga 1. W przypadku, gdy zdarzenia $A_1, A_2, \ldots, A_n$ są niezależne mamy oczywiście $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \ldots \cdot \mathbb{P}(A_n)$.

Wzór na prawdopodobieństwo całkowite

Definicja 1. Rozbiciem przestrzeni Ω *nazywamy rodzinę zdarzeń* $\{A_i\}_{i\in I}$*, które wykluczają się parami, zaś ich suma jest równa* Ω*, tzn. dla każdych* $i \neq j$ *zachodzi* $A_i \cap A_j = \emptyset$ *oraz* $\bigcup_{i\in I} A_i = \Omega$*.*



Przykład rozbicia przestrzeni Ω za pomocą rodziny zdarzeń $\{A_1, A_2, A_3, A_4\}$.

Twierdzenie 2 (**Wzór na prawdopodobieństwo całkowite**)**.** *Jeśli* $A_1, A_2, \ldots, A_n \in \mathcal{F}$ *jest rozbiciem przestrzeni* Ω *na zdarzenia o dodatnim prawdopodobieństwie, to dla każdego zdarzenia* $B \in \mathcal{F}$ *mamy:*

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\,\mathbb{P}(A_i)\,.$$

dowód. *Zauważmy, że* $B\cap A_1, B\cap A_2, \ldots, B\cap A_n$ *stanowi ciąg parami rozłącznych zdarzeń, które sumują się do zdarzenia* B*. Wówczas korzystając z własności funkcji prawdopodobieństwa oraz definicji prawdopodobieństwa warunkowego, otrzymujemy:*

$$\mathbb{P}(B) = \mathbb{P}\left(\bigcup_{i=1}^{n} B\cap A_i\right) = \sum_{i=1}^{n}\mathbb{P}(B\cap A_i) = \sum_{i=1}^{n}\mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

□

Uwaga 2. Analogiczny wzór działa również dla rozbicia przestrzeni Ω na przeliczalną rodzinę zdarzeń $\{A_i\}_{i\in I}$.

Przykład 2. Wróćmy do Przykładu 1. Stosując tę samą regułę losowania, zaczynając od urny z 3 białymi i 3 czarnymi kulami, oblicz prawdopodobieństwo wylosowania kuli białej w drugim losowaniu.

Ω – możliwe wyniki trzech kolejnych losowań

A_1 – w pierwszym losowaniu wylosowano białą kulę

A_2 – w pierwszym losowaniu wylosowano czarną kulę

$A_1 \cup A_2$ – rozbicie przestrzeni Ω

B – w drugim losowaniu wylosowano białą kulę

Stosując wzór na prawdopodobieństwo całkowite dostajemy:

$$\mathbb{P}(B) = \mathbb{P}(B|A_1)\cdot\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\cdot\mathbb{P}(A_2) = \frac{4}{7}\cdot\frac{1}{2} + \frac{3}{5}\cdot\frac{1}{2} = \frac{41}{70}.$$

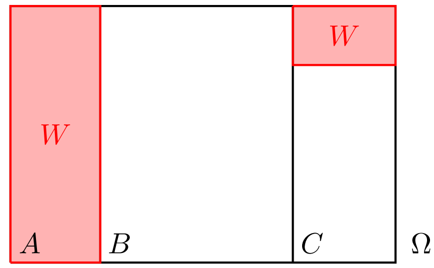
Przykład 3. W loterii fantowej szansa wygranej jest równa p, szansa przegranej wynosi q, przy czym $p+q<1$, natomiast z prawdopodobieństwem $r = 1-p-q$ można wyciągnąć los ”graj dalej”, który wrzucamy z powrotem do urny i dokonujemy ponownego losowania. Jaka jest szansa wygranej?

A – w pierwszym losowaniu wyciągnięto los wygrywający

B – w pierwszym losowaniu wyciągnięto los przegrywający

C – w pierwszym losowaniu wyciągnięto los ”graj dalej”

W – wygrana



Zawuażmy, że w przypadku wylosowania losu "graj dalej", a więc gdy zajdzie zdarzenie C, powtarzamy dokładnie ten sam eksperyment. Zatem $\mathbb{P}(W|C) = \mathbb{P}(W)$. Rodzina zdarzeń $\{A, B, C\}$ stanowi rozbicie przestrzeni Ω, zatem możemy zastosować wzór na prawdopodobieństwo całkowite:

$$\mathbb{P}(W) = \mathbb{P}(W|A) \cdot \mathbb{P}(A) + \mathbb{P}(W|B) \cdot \mathbb{P}(B) + \mathbb{P}(W|C) \cdot \mathbb{P}(C) = 1 \cdot p + 0 \cdot q + \mathbb{P}(W) \cdot (1 - p - q).$$

Po uproszczeniu otrzymujemy:

$$\mathbb{P}(W) = \frac{p}{p+q}.$$

Wzór Bayesa

Twierdzenie 3 (Wzór Bayesa). *Jeśli $A_1, A_2, \ldots, A_n \in \mathcal{F}$ jest rozbiciem przestrzeni Ω na zdarzenia o dodatnim prawdopodobieństwie oraz zdarzenie $B \in \mathcal{F}$ ma również dodatnie prawdopodobieństwo, to dla każdego $i = 1, 2, \ldots, n$ zachodzi:*

$$\mathbb{P}\left(A_i|B\right) = \frac{\mathbb{P}\left(B|A_i\right) \cdot \mathbb{P}\left(A_i\right)}{\mathbb{P}\left(B\right)} = \frac{\mathbb{P}\left(B|A_i\right) \cdot \mathbb{P}\left(A_i\right)}{\sum_{j=1}^{n} \mathbb{P}\left(B|A_j\right) \cdot \mathbb{P}\left(A_j\right)}.$$

Przykład 4. W urnie znajdują się 4 kule białe, 2 czarne i 1 czerwona. Losujemy jedną kulę i bez sprawdzania jej koloru odkładamy ją na bok. Następnie losujemy drugą kulę. Ile wynosi prawdopodobieństwo tego, że odłożona kula ma czerwony kolor, jeśli w drugim losowaniu wyciągnęliśmy białą kulę?

Ω – możliwe wyniki dwóch losowań

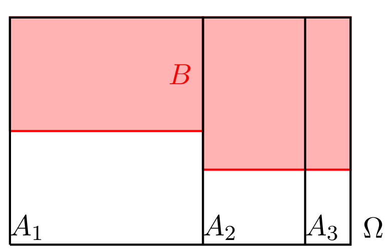
A_1 – w pierwszym losowaniu wyciągnęliśmy białą kulę

A_2 – w pierwszym losowaniu wyciągnęliśmy czarną kulę

A_3 – w pierwszym losowaniu wyciągnęliśmy czerwoną kulę

$A_1 \cup A_2 \cup A_3$ – rozbicie przestrzeni Ω

B – w drugim losowaniu wyciągnęliśmy białą kulę



Interesuje nas prawdopodobieństwo zdarzenia A_3 pod warunkiem zajścia zdarzenia B. Korzystając ze wzoru Bayesa otrzymujemy:

$$\mathbb{P}(A_3|B) = \frac{\mathbb{P}(B|A_3) \cdot \mathbb{P}(A_3)}{\sum_{i=1}^{3} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)} = \frac{\frac{4}{6} \cdot \frac{1}{7}}{\frac{3}{6} \cdot \frac{4}{7} + \frac{4}{6} \cdot \frac{2}{7} + \frac{4}{6} \cdot \frac{1}{7}} = \frac{4}{24} = \frac{1}{6}.$$

Przykład 5. W zbiorze 33 monet jedna ma po obu stronach orły, pozostałe są standardowe, czyli mają po jednej stronie orła, a po drugiej reszkę. W wyniku 10 rzutów losowo wybraną monetą otrzymaliśmy 10 orłów. Oblicz prawdopodobieństwo, że rzucaliśmy monetą z dwoma orłami.

Ω – możliwe przebiegi eksperymentu (wybór monety, a następnie oddanie 10 rzutów wylosowaną monetą)

A_1 – wylosowaliśmy monetę z dwoma orłami

$A_2 = A_1'$ – wylosowaliśmy standardową monetę

$A_1 \cup A_2$ – rozbicie przestrzeni Ω

B – w 10 rzutach wylosowaną monetą wypadło 10 orłów

Interesuje nas prawdopodobieństwo zdarzenia A_1 poz warunkiem zajścia zdarzenia B. Ze wzoru Bayesa otrzymujemy:

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B|A_1) \cdot \mathbb{P}(A_1)}{\mathbb{P}(B|A_1) \cdot \mathbb{P}(A_1) + \mathbb{P}(B|A_2) \cdot \mathbb{P}(A_2)} = \frac{1 \cdot \frac{1}{33}}{1 \cdot \frac{1}{33} + \left(\frac{1}{2}\right)^{10} \cdot \frac{32}{33}} = \frac{1}{1 + \frac{1}{32}} = \frac{32}{33}\,.$$

Przykład 6. Test na rzadką chorobę, którą dotknięta jest średnio jedna osoba na tysiąc, daje tzw. pozytywną odpowiedź u 5% zdrowych osób (u chorego zawsze daje odpowiedź pozytywną). Jaka jest szansa na to, że osoba, u której test dał odpowiedź pozytywną jest faktycznie chora?

Ω – cała populacja

A_1 – losowa osoba jest chora

A_2 – losowa osoba jest zdrowa

$A_1 \cup A_2$ – rozbicie przestrzeni Ω

B – test wykonany na losowej osobie dał odpowiedź pozytywną

Chcemy wyznaczyć prawdopodobieństwo zdarzenia A_1 pod warunkiem zajścia zdarzenia B. Stosując wzór Bayesa otrzymujemy:

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B|A_1) \cdot \mathbb{P}(A_1)}{\mathbb{P}(B|A_1) \cdot \mathbb{P}(A_1) + \mathbb{P}(B|A_2) \cdot \mathbb{P}(A_2)} = \frac{1 \cdot 0,001}{1 \cdot 0,001 + 0,05 \cdot 0,999} = \frac{20}{1019} \approx 0,01963\,.$$