1. Rachunek prawdopodobieństwa - przypomnienie wiadomości

Definicja 1.1. Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną, a B - dowolnie ustalonym zdarzeniem takim, że $P(B) > 0$.

Prawdopodobieństwo warunkowe zdarzenia A pod warunkiem zajścia zdarzenia B nazywamy liczbę

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

Wniosek 1.2. *Jeżeli $P(B) > 0$, to funkcja $P(\cdot|B)$ określona na $\mathcal{F}$ spełnia aksjomaty prawdopodobieństwa.*

Wniosek 1.3. *Niech A i B będą dowolnymi zdarzeniami takimi, że $P(A) > 0$ oraz $P(B) > 0$. Wówczas $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.*

Wniosek 1.4. *Posługując się zasadą indukcji matematycznej możemy udowodnić, że dla dowolnego n, przy założeniu, że $P(A_1 \cap \cdots \cap A_{n-1}) > 0$, prawdziwa jest równość*

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

Przykład 1.5. Rzucamy dwa razy symetryczną kostką.

1. Jakie jest prawdopodobieństwo wyrzucenia różnej liczby oczek?

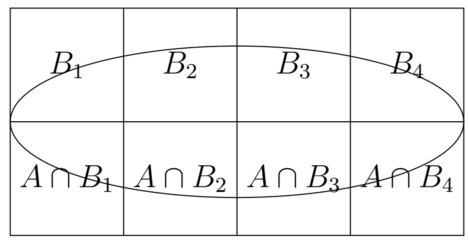
2. Jakie jest prawdopodobieństwo wyrzucenia różnej liczby oczek, jeżeli suma oczek, jeżeli suma oczek wynosi 11?

3. Jakie jest prawdopodobieństwo wyrzucenia różnej liczby oczek, jeżeli suma oczek, jeżeli suma oczek wynosi 10?

Przykład 1.6. Studenci Wydziału Elektroniki muszą zdać w I semestrze trzy egzaminy: z fizyki (A), analizy matematycznej (C) i z algebry (B). Z danych Dziekanatu wynika, że 70% studentów zalicza I semestr, a 90% - zdaje egzamin z fizyki. Jeżeli student zaliczy algebrę i fizykę, to prawdopodobieństwo, że zda analizę wynosi $\frac{4}{5}$. Jakie jest prawdopodobieństwo, że student, który zdał fizykę, zda algebrę?

Definicja 1.7. Mówimy, że zdarzenia $B_1, B_2, ..., B_n$ tworzą podział przestrzeni zdarzeń elementarnych Ω (Ω jest sumą rozłącznych zbiorów $B_i \in \mathcal{F}$ dla $i = 1, ..., n$), jeżeli

$$B_i \cap B_j = \emptyset, \text{ gdy } i \neq j \quad \text{oraz} \quad \bigcup_{i=1}^{n} B_i = \Omega$$



$$A = (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cup (A \cap B_4)$$
A - elipsa, B_1, B_2, B_3, B_4 - podział przestrzeni Ω

Figura: Diagram Venna ilustrujący **twierdzenie o prawdopodobieństwie całkowitym**

Twierdzenie 1.8 (Twierdzenie o prawdopodobieństwie całkowitym)**.** *Jeżeli Ω jest sumą rozłącznych zbiorów B_i, przy czym $P(B_i) > 0$ dla wszystkich $i = 1, ..., n$, to dla dowolnego zdarzenia A zachodzi równość*

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

Czasem zdarzenia B_i występujące we wzorze na prawdopodobieństwo całkowite nazywamy przyczynami a zdarzenie A - skutkiem. Rozważmy zagadnienie odwrotne: jakie jest prawdopodobieństwo przyczyny B_i, gdy znany jest skutek A?

Twierdzenie 1.9 (Wzór Bayesa)**.** *Jeżeli Ω jest sumą rozłącznych zbiorów B_i, przy czym $P(B_i) > 0$ dla wszystkich $i = 1, ..., n$, to dla dowolnego zdarzenia A, takiego że $P(A) > 0$, zachodzi równość*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}.$$

Przykład 1.10. Przeciętnie 3% wyprodukowanych elementów ma wadę. Do wykrywania wady stosuje się test, który z prawdopodobieństwem 0.9 wskazuje wadę (wynik testu pozytywny), jeżeli element ma wadę i z prawdopodobieństwem 0.95 nie wskazuje wady, jeżeli element jej nie ma.

a) Jakie jest prawdopodobieństwo, że element ma wadę, jeżeli wynik testu jest pozytywny?

b) Jakie jest powyższe prawdopodobieństwo, jeżeli wynik testu jest negatywny?

Przykład 1.11. Na stole leży 5 długopisów. Każde przypuszczenie (hipoteza) dotyczące liczby zepsutych długopisów jest jednakowo prawdopodobne. Wybrany losowo długopis okazał się zepsuty. Które przypuszczenie dotyczące liczby zepsutych długopisów jest teraz najbardziej prawdopodobne?

Definicja 1.12. Dwa zdarzenia A i B nazywamy niezależnymi, jeżeli

$$P(A \cap B) = P(A) \cdot P(B).$$

Jeżeli $P(B) > 0$, to z niezależności zdarzeń A i B wynika, że $P(A|B) = P(A)$. Dla dowolnego $A \in \mathcal{F}$ zdarzenia A i Ω są niezależne. Podobnie - zdarzenia A i $\emptyset$ są niezależne. Jeżeli zdarzenia A i B są rozłączne i mają niezerowe prawdopodobieństwa, to nie mogą być niezależne.

Definicja 1.13. Zdarzenia $A_1, A_2, A_3, \ldots$ nazywamy rodziną zdarzeń niezależnych, jeżeli dla każdej skończonej ilości zdarzeń $A_{i_1}, A_{i_2}, \ldots, A_{i_n}$ z tej rodziny, gdzie $n \geqslant 2$, zachodzi równość

$$P(A_{i_1} \cap \cdots \cap A_{i_n}) = P(A_{i_1}) \ldots P(A_{i_n}).$$

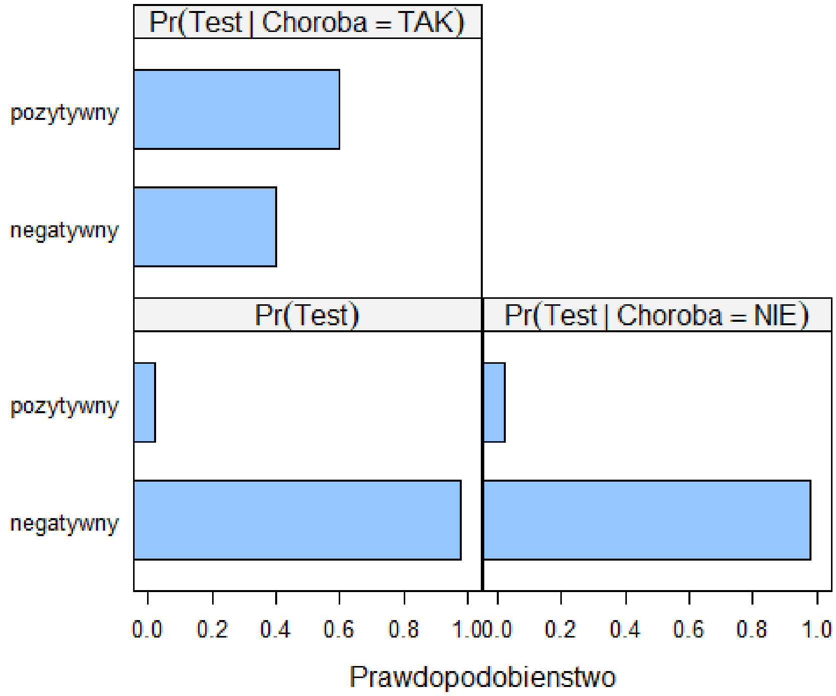
Wniosek 1.14. *Jeżeli zdarzenia A i B są niezależne, to także niezależne są pary zdarzeń A' i B, A i B' oraz A' i B'.*

Wniosek 1.15. *Jeżeli zdarzenia $A_1, A_2, ..., A_n$ są niezależne, to niezależne są także zdarzenia $B_1, B_2, ..., B_n$, gdzie $B_i = A_i$ lub $B_i = A_i'$ dla $i = 1, 2, ..., n$.*

Przykład 1.16. Trzech kontrolerów jakości pracuje niezależnie. Pierwszy wykrywa 90% wad, drugi - 80% a trzeci - 60%. Jaki procent wad wykrywają łącznie? Jaki procent wad wykrywa trzeci kontroler a nie wykrywa pierwszy ani drugi?

Przykład 1.17. Wybieramy jedną rodzinę spośród rodzin, mających n dzieci. Niech zdarzenie A polega na tym, że w losowo wybranej rodzinie jest co najwyżej jedna dziewczynka, B - w rodzinie są dziewczynki i chłopcy. Czy zdarzenia A i B są niezależne?

Przykład 1.18. Pacjent wykonuje test na obecność pewnej choroby. Wynik jest pozytywny. Prawdopodobieństwo otrzymania wyniku pozytywnego w przypadku choroby wynosi 0.6, natomiast prawdopodobieństwo, że wynik będzie negatywny jeśli pacjent nie jest chory 0.98. Prawdopodobieństwo występowania tej choroby w populacji wynosi 0.001.



Rysunek 1

(a) Jakie jest prawdopodobieństwo, że pacjent jest chory po wykonaniu jednego testu?

(b) Narysować rozkład prawdopodobieństw warunkowych. Obliczyć brakujące prawdopodobieństwa.

Przykład 1.19. W pewnym sklepie przeprowadzono następującą ankietę wśród klientów.

Pytanie 1. Jaka jest Twoja ogólna ocena zadowolenia z obsługi klienta w naszym sklepie w skali od 1 (bardzo niski poziom zadowolenia) do 10 (bardzo wysoki poziom zadowolenia)?

1	2	3	4	5	6	7	8	9	10

Pytanie 2. W jaki sposób dowiedziałeś się o naszym sklepie?

- internet
- znajomi
- telewizja

- radio
- prasa
- inne

Pytanie 3. Czy poleciłbyś nasz sklep znajomym?

- tak
- nie
- nie wiem

Wyniki poniższej ankiety zebrano w pliku `nowedane.csv`. Rozważmy zdarzenie polegające na wybraniu jednej ankietowanej osoby. Zakładamy, że prawdopodobieństwo wylosowania każdej osoby jest takie samo. Sprawdzić czy wśród zdarzeń:

- A_i - zdarzenie polegające na wybraniu osoby, której poziom zadowolenia z obsługi klienta jest określony na poziomie i, $i \in \{1, \ldots, 10\}$
- B_j - zdarzenie polegające na wybraniu osoby, która na pytanie, czy poleciłaby sklep odpowiedziała w sposób j-ty, $j \in \{1, 2, 3\}$
- C_k - zdarzenie polegające na wybraniu osoby, która dowiedziała się o sklepie w sposób k-ty, $k \in \{1, 2, \ldots, 6\}$

są zdarzenia niezależne lub warunkowo niezależne.