Andrzej Sierociński

RACHUNEK PRAWDOPODOBIEŃSTWA

Elementy Rachunku Prawdopodobieństwa Wykład 2

4 Prawdopodobieństwo warunkowe

Prawdopodobieństwo dowolnego zdarzenia zależy od tego, co wiemy o badanym eksperymencie losowym. Jest intuicyjnie oczywiste, że informacja o zajściu jakiegoś zdarzenia losowego B może mieć istotny wpływ na naszą wiedzę o zajściu lub nie innego zdarzenia losowego A, a co za tym idzie na jego prawdopodobieństwo, jeżeli wiemy, że na pewno zaszło zdarzenie B.

Wprowadzimy jedno z najważniejszych pojęć w rachunku prawdopodobieństwa, mianowicie pojęcie **prawdopodobieństwa warunkowego**.

Pojęcie to jest ważne z dwóch powodów

1. Czasami zainteresowani jesteśmy wyznaczeniem prawdopodobieństwa pewnego zdarzenia w sytuacji, gdy dysponujemy częściową wiedzą na jego temat lub gdy chcemy je przeliczyć ponownie po uzyskaniu dodatkowych informacji.

2. Z kolei, w pewnych sytuacjach, łatwiej jest wyznaczyć prawdopodobieństwo interesującego nas zdarzenia pod warunkiem zajścia lub nie innego zdarzenia.

Przykład. *Wykonujemy dwa rzuty symetryczną kostką do gry. Załóżmy, że w wyniku pierwszego rzutu otrzymujemy* 1.

Pytanie: znając ten wynik, jakie jest prawdopodobieństwo uzyskania sumy oczek w dwóch rzutach równej 7*?*

Rozwiązanie.
Zbiór zdarzeń elementarnych Ω składa się z 36 jednakowo prawdopodobnych zdarzeń elementarnych $\Omega = \{(\omega_1, \omega_2) : \omega_i = 1, 2, 3, 4, 5, 6; \quad i = 1, 2\}$, gdzie ω_1 jest wynikiem pierwszego rzutu a ω_2 drugiego.
Oznaczmy przez A zdarzenie - suma oczek równa siedem. Wówczas $P(A) = 1/6$, ponieważ $A = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$.

Załóżmy, że w wyniku pierwszego rzutu otrzymaliśmy 1. Oznaczmy to zdarzenie przez B, wówczas $B = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\}$. Oczywiście $P(B) = 1/6$.

Zatem, jeżeli wiemy, że w pierwszym rzucie uzyskaliśmy 1, jakie jest prawdopodobieństwo, że suma oczek jest równa 7?

Ponieważ wynik pierwszego rzutu jest znany (1), mamy tylko sześć możliwych wyników naszego eksperymentu, mianowicie zdarzenia elementarne wchodzące w skład zdarzenia B.

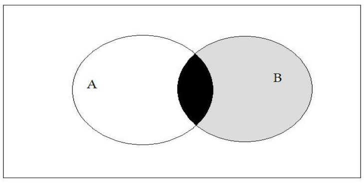
Pozostałe 30 wyników zawartych w $\overline{B}$ jest niemożliwe.

Zatem "warunkowe" prawdopodobieństwo każdego z wyników wchodzących w skład zdarzenia B pod warunkiem, że B zaszło, jest równe $1/6$.

Jeżeli oznaczymy to prawdopodobieństwo warunkowe zdarzenia A pod warunkiem, że zaszło zdarzenie B przez $P(A|B)$, to otrzymujemy $P(A|B) = 1/6$, ponieważ $(1,6)$ jest jedynym wynikiem spełniającym nasze wymagania.

Opisana w przykładzie sytuacja zilustrowana jest na poniższym diagramie Venna. Zauważmy, że $\{(1,6)\} = A \cap B$ oraz

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



Ciemny obszar to $A \cap B$.

Definicja.
Dla dowolnych dwóch zdarzeń losowych $A, B \in \mathcal{F}$, $P(B) > 0$, **prawdopodobieństwem zdarzenia A pod warunkiem, że zaszło zdarzenie B** jest

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

Przykład. *W fabryce są dwie linie produkcyjne A_1 i A_2. Linia produkcyjna A_1 wykorzystuje starsze oprzyrządowanie niż linia A_2, w związku z czym produkcja na niej jest wolniejsza i bardziej zawodna. Załóżmy, że pewnego dnia na linii A_1 wyprodukowano 80 wyrobów, z których 8 okazało się wadliwych (zdarzenie B) i 72 zgodnych z wymaganiami ($\overline{B}$), natomiast na linii A_2 był 1 wyrób wadliwy i 99 zgodnych z wymaganiami. Obliczyć prawdopodobieństwo tego, że jeżeli losowo wybrany z całej produkcji wyrób był wadliwy, to został wyprodukowany na linii A_1?*

		Stan produktu	
		B	$\overline{B}$
Linia	A_1	8	72
	A_2	1	99

Rozwiązanie. Wybieramy losowo 1 spośród 180 wyrobów, zatem prawdopodobieństwo zdarzenia, że wyrób został wyprodukowany na linii A_1 jest równe

$$P(A_1) = \frac{80}{180} = \frac{4}{9} = 0.444.$$

Jednakże wiadomo, że wyrób okazał się wadliwy (zaszło zdarzenie B) zatem prawdopodobieństwo warunkowe A_1 pod warunkiem B jest dwukrotnie większe

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{8}{9} = 0.889.$$

4.1 Wyznaczanie $P(A \cap B)$

Prostą konsekwencją definicji prawdopodobieństwa warunkowego jest następująca reguła służąca do wyznaczania prawdopodobieństwa iloczynu dwóch zdarzeń:

Jeżeli $B \in \mathcal{F}$ oraz $P(B) > 0$, to dla dowolnego zdarzenia $A \in \mathcal{F}$

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

Reguła wyznaczania prawdopodobieństwa przecięcia dwóch zdarzeń jest wyjątkowo użyteczna w przypadku, gdy eksperyment jest sekwencją kilu mniejszych eksperymentów.

Stwierdzenie. Jeżeli $A_1, A_2, \ldots, A_n \in \mathcal{F}$ oraz $P(A_1 \cap \ldots \cap A_n) > 0$, to

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \prod_{i=2}^{n} P(A_i|A_1 \cap \ldots \cap A_{i-1}).$$

Dowód. Jest to konsekwencja zastosowania reguły mnożenia prawdopodobieństw, zastosowanej krok po kroku do wszystkich czynników po prawej stronie równości. ■

Przykład. *Załóżmy, że mamy 100 żarówek, wśród których 5 jest uszkodzonych. Jakie jest prawdopodobieństwo, że wszystkie 3 losowo wybrane żarówki będą sprawne?*

Rozwiązanie. Zbiór zdarzeń elementarnych Ω składa się ze wszystkich kombinacji trójelementowych ze zbioru 100 żarówek. Oznaczmy przez A zdarzenie, że 3 losowo wybrane żarówki są sprawne. Zatem prawdopodobieństwo zdarzenia A jest równe

$$P(A) = \frac{\binom{95}{3}}{\binom{100}{3}} = \frac{\frac{95!}{3!92!}}{\frac{100!}{3!97!}} = \frac{95 \cdot 94 \cdot 93}{100 \cdot 99 \cdot 98}.$$

Rozwiązanie można uzyskać inną metodą, stosując pojęcie prawdopodobieństwa warunkowego. Oznaczmy przez A_i, $i = 1, 2, 3$ zdarzenie i-ta wybrana żarówka jest sprawna. Oczywiście $A = A_1 \cap A_2 \cap A_3$ oraz

$$\begin{aligned} P(A) &= P(A_1 \cap A_2 \cap A_3) = \\ &= P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) = \\ &= \frac{95}{100} \cdot \frac{94}{99} \cdot \frac{93}{98}, \end{aligned}$$

ponieważ $P(A_1) = 95/100$, $P(A_2|A_1) = 94/99$ oraz $P(A_3|A_1 \cap A_2) = 93/98$.

4.2 Twierdzenie o prawdopodobieństwie całkowitym

Definicja. Ciąg zdarzeń losowych $A_1, A_2, \ldots$ (ciąg $\{A_i \in \mathcal{F}\}$ jest skończony lub nie) nazywamy **układem zupełnym** zdarzeń jeżeli

1. $\bigcup_i A_i = \Omega$,
2. $\forall i \neq j \quad A_i \cap A_j = \emptyset$,
3. $\forall i \quad P(A_i) > 0$.

Twierdzenie o prawdopodobieństwie całkowitym (zupełnym).
Jeżeli zdarzenia $A_1, A_2, \ldots$ tworzą układ zupełny zdarzeń, to dla dowolnego zdarzenia losowego $B \in \mathcal{F}$, mamy

$$P(B) = \sum_i P(B|A_i) \cdot P(A_i).$$

Dowód. Ponieważ zdarzenia A_i są rozłączne i co najmniej jedno z nich musi zajść, to

$$B = B \cap \Omega = B \cap \left(\bigcup_i A_i \right) = \bigcup_i (B \cap A_i).$$

Oczywiście zdarzenia $(B \cap A_i)$ są także parami rozłączne. Zatem

$$P(B) = P\left(\bigcup_i (B \cap A_i) \right) = \sum_i P(B \cap A_i) = \sum_i P(A_i) \cdot P(B|A_i).$$

$\blacksquare$

4.3 Reguła Bayesa

Twierdzenie - Reguła Bayesa.
Niech zdarzenia $A_1, A_2, \ldots$ tworzą układ zupełny zdarzeń. Wówczas dla dowolnego zdarzenia losowego $B \in \mathcal{F}$ takiego, że $P(B) > 0$ oraz dowolnego j,

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_i P(B|A_i) \cdot P(A_i)}.$$

Dowód. Zauważmy, że

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j) \cdot P(A_j)}{\sum_i P(B|A_i) \cdot P(A_i)}.$$

Jest to prosta konsekwencja reguły mnożenia oraz twierdzenia o prawdopodobieństwie całkowitym. ■

Jeśli pomyślimy o zdarzeniach A_i jako wszystkich możliwych stanach rzeczywistości ("hipotezach") dotyczących naszego problemu, wówczas reguła Bayesa, nazwana od nazwiska angielskiego filozofa Thomasa Bayesa,

Thomas Bayes (1702-1761)

może być zinterpretowana jako zmiana opinii (prawdopodobieństw) o możliwych stanach rzeczywistości (hipotezach A_i) po zaobserwowaniu wyniku eksperymentu losowego (zdarzenie B), tzn.

- z prawdopodobieństw początkowych $P(A_j)$ - *a priori*
- na prawdopodobieństwa po uzyskaniu wyniku eksperymentu $P(A_j|B)$ - *a posteriori*.

Przykład. *Pewne detale pakowane są po 10 sztuk w paczce i wysyłane do odbiorcy. Załóżmy, że 50% paczek nie zawiera żadnych wadliwych detali, 30% zawiera jeden wadliwy detal, a pozostałe 20% zawiera 2 wadliwe detale. Odbiorca, z losowo wybranej paczki wybiera dwa przypadkowe detale i poddaje badaniu. Jakie są prawdopodobieństwa, że w wylosowanej paczce mamy 0, 1 lub 2 braki, jeżeli w wyniku eksperymentu zaobserwowaliśmy:*

- *0 braków*
- *1 brak.*

Rozwiązanie. Oznaczmy przez A_i, $i = 1, 2, 3$ zdarzenia odpowiadające odpowiednio 0, 1 lub 2 brakom w paczce. Zdarzenia A_i tworzą układ zupełny zdarzeń oraz $P(A_1) = 0.5$, $P(A_2) = 0.3$, $P(A_3) = 0.2$. Niech B oznacza zdarzenie, że żaden z testowanych detali nie jest wadliwy a C, że dokładnie jeden detal jest wadliwy. Zauważmy, że

$$P(B|A_1) = \frac{\binom{10}{2}}{\binom{10}{2}} = 1, \quad P(B|A_2) = \frac{\binom{9}{2}}{\binom{10}{2}} = 0.8, \quad P(B|A_3) = \frac{\binom{8}{2}}{\binom{10}{2}} = 0.622.$$

Stosując twierdzenie o prawdopodobieństwie całkowitym, otrzymujemy

$$P(B) = \sum_{i=1}^{3} P(B|A_i) \cdot P(A_i) = 1 \cdot 0.5 + 0.8 \cdot 0.3 + 0.622 \cdot 0.2 = 0.864.$$

Prawdopodobieństwa zdarzeń, że mamy $i - 1$ braków w paczce pod warunkiem, że żaden z testowanych detali nie okazał się brakiem są równe $P(A_i|B)$, $i = 1, 2, 3$. Zatem na podstawie reguły Bayesa otrzymujemy

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B)}, \quad i = 1, 2, 3.$$

Ostatecznie mamy

$$\begin{aligned} P(A_1|B) &= \frac{1 \cdot 0.5}{0.864} = 0.578, \\ P(A_2|B) &= \frac{0.8 \cdot 0.3}{0.864} = 0.278, \\ P(A_3|B) &= \frac{0.622 \cdot 0.2}{0.864} = 0.144. \end{aligned}$$

Podobnie postępujemy w drugim przypadku, gdy jeden spośród testowanych detali jest brakiem.

$$P(C|A_1) = \frac{\binom{10}{1} \cdot \binom{0}{1}}{\binom{10}{2}} = 0,$$

$$P(C|A_2) = \frac{\binom{9}{1} \cdot \binom{1}{1}}{\binom{10}{2}} = 0.2,$$

$$P(C|A_3) = \frac{\binom{8}{1} \cdot \binom{2}{1}}{\binom{10}{2}} = 0.356,$$

oraz

$$P(C) = \sum_{i=1}^{3} P(C|A_i) \cdot P(A_i) = 0 \cdot 0.5 + 0.2 \cdot 0.3 + 0.356 \cdot 0.2 = 0.131.$$

Zatem

$$\begin{aligned} P(A_1|C) &= \frac{0 \cdot 0.5}{0.131} = 0, \\ P(A_2|C) &= \frac{0.2 \cdot 0.3}{0.131} = 0.458, \\ P(A_3|C) &= \frac{0.356 \cdot 0.2}{0.131} = 0.542. \end{aligned}$$

Przykład. *Tylko jedna na tysiąc dorosłych osób zapada na bardzo rzadką chorobę. Opracowano pewien test medyczny pozwalający zdiagnozować tę chorobę. W przypadku, gdy dana osoba jest chora, test ten daje wynik pozytywny (stwierdzający wystąpienie choroby) w 99% przypadków, w przypadku osób zdrowych wynik pozytywny (w tym przypadku błędny) występuje tylko w 2% przypadków.*

Na przypadkowo wybranej osobie przeprowadzono ten test i dał on wynik pozytywny. Jakie jest prawdopodobieństwo, że osoba ta w istocie ma tę chorobę.

Rozwiązanie. Oznaczmy przez A_1 zdarzenie, że wybrana osoba ma tę chorobę a przez $A_2 = \overline{A}_1$, że jej nie ma, natomiast przez B, że wynik testu jest pozytywny.

Wówczas $P(A_1) = 0.001$, $P(A_2) = 0.999$, $P(B|A_1) = 0.99$ oraz $P(B|A_2) = 0.02$.

Z twierdzenia o prawdopodobieństwie całkowitym mamy

$$\begin{aligned} P(B) &= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) = \\ &= 0.99 \cdot 0.001 + 0.02 \cdot 0.999 = 0.02097. \end{aligned}$$

Następnie, stosując regułę Bayesa

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B)} = \frac{0.99 \cdot 0.001}{0.02097} = 0.0472.$$

Otrzymany wynik wydaje się mocno nieintuicyjny. Można sądzić, że osoba z pozytywnym wynikiem testu z dużym prawdopodobieństwem jest chora, gdy w istocie prawdopodobieństwa wynosi zaledwie 0.047 (co prawda zwiększyło się 47 razy - z prawdopodobieństwa a priori 0.001). Jednakże, ponieważ choroba jest niezwykle rzadka, to większość pozytywnych wyników testu wynika z błędów niż z pozytywnego rozpoznania choroby.

Jeśli choroba nie byłaby tak rzadka np. $P(A_1) = 0.25$, to $P(A_1|B) = 0.943$ test dawałby bardzo dobrą diagnozę.

Uwaga.
Dla dowolnego $B \in \mathcal{F}$, $P(B) > 0$ funkcja

$$P_B(A) = P(A|B), \quad \text{dla każdego } A \in \mathcal{F},$$

jest funkcją $P_B(\cdot)$ prawdopodobieństwa (spełnia aksjomaty 1-3).

5 Niezależność

Na ogół $P(A|B) \neq P(A)$ i definicja prawdopodobieństwa warunkowego pozwala zweryfikować naszą wiedzę o prawdopodobieństwie interesującego nas zdarzenia A jeśli posiadamy informację o zajściu pewnego innego zdarzenia B.

Jednakże w przypadku, gdy $P(A|B) = P(A)$ możemy wnioskować, że zajście zdarzenia B nie ma wpływu na szansę zajścia zdarzenia A. Naturalnym wydaje się stwierdzenie, że wówczas zdarzenie A nie zależy od B.

Ponieważ przy założeniu, że $P(A) > 0$ oraz $P(B) > 0$ warunki $P(A|B) = P(A)$ i $P(B|A) = P(B)$ są równoważne, to także prawdziwe jest stwierdzenie, że B nie zależy od A. Założenie o niezerowaniu się prawdopodobieństw zdarzeń A i B nie jest istotne i można je pominąć, zapisując warunek niezależności w nieco inny (równoważny) sposób.

5.1 Niezależność dwóch zdarzeń

Definicja. Zdarzenia $A \in \mathcal{F}$ i $B \in \mathcal{F}$ są ***(wzajemnie)*** **niezależne** wtedy i tylko wtedy, gdy

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

W przeciwnym przypadku mówimy, ze zdarzenia te są zależne.

Przykład. *Rozważmy jednokrotny rzut symetryczną kostką do gry oraz zdefiniujmy następujące zdarzenia:* $A = \{2, 4, 6\}$, $B = \{1, 2, 3\}$, $C = \{1, 2, 3, 4\}$ *i* $D = \{5, 6\}$. *Pokazać, że*

1. A *i* B *oraz* B *i* D *są zależne, a*

2. A *i* C *oraz* A *i* D *są niezależne.*

Rozwiązanie. Zbiór zdarzeń elementarnych Ω składa się z 6 jednakowo prawdopodobnych zdarzeń elementarnych, zatem $P(A) = 0.5$, $P(B) = 0.5$, $P(C) = 2/3$ oraz $P(D) = 1/3$. Ponieważ $A \cap B = \{2\}$, $A \cap C = \{2, 4\}$, $A \cap D = \{6\}$ oraz $B \cap D = \emptyset$, to $P(A \cap B) = P(A \cap D) = 1/6$, $P(A \cap C) = 1/3$ oraz $P(B \cap D) = 0$. Stąd otrzymujemy, że

$$P(A \cap B) = \frac{1}{6} \neq \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B),$$

tym samym A i B są zależne.
Podobnie zależne są B i D, ponieważ

$$P(B \cap D) = 0 \neq \frac{1}{2} \cdot \frac{1}{3} = P(B) \cdot P(D).$$

W podobny sposób sprawdzamy, że

$$P(A \cap C) = \frac{1}{3} = \frac{1}{2} \cdot \frac{2}{3} = P(A) \cdot P(C),$$

co oznacza, że A i C są niezależne.
Również niezależne są A i D, gdyż

$$P(A \cap D) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A) \cdot P(D).$$

Uwaga.
Jeżeli dwa zdarzenia A i B, $P(A) > 0$, $P(B) > 0$ są rozłączne, to są zależne.

Istotnie,

$$P(A \cap B) = 0 \neq P(A) \cdot P(B) > 0.$$

Innymi słowy, kiedy A *i* B *są rozłączne, to informacja, że zaszło zdarzenie* A *mówi coś o zdarzeniu* B *i tym samym wykluczona jest niezależność tych zdarzeń.*

Twierdzenie. Jeżeli zdarzenia losowe A i B są niezależne, to niezależne są również zdarzenia

- A i $\overline{B}$ $\quad (P(A \cap \overline{B}) = P(A) \cdot P(\overline{B}))$;
- $\overline{A}$ i B $\quad (P(\overline{A} \cap B) = P(\overline{A}) \cdot P(B))$;
- $\overline{A}$ i $\overline{B}$ $\quad (P(\overline{A} \cap \overline{B}) = P(\overline{A}) \cdot P(\overline{B}))$.

Dowód. Udowodnimy jedynie pierwszą własność, pozostałe dowodzi się w podobny sposób.

Załóżmy, że A i B są niezależne, wówczas $P(A \cap B) = P(A) \cdot P(B)$. Zatem

$$\begin{aligned} P(A \cap \overline{B}) &= P(A - (A \cap B)) = P(A) - P(A \cap B) = \\ &= P(A) - P(A) \cdot P(B) = P(A) \cdot (1 - P(B)) = \\ &= P(A) \cdot P(\overline{B}). \end{aligned}$$

■

5.2 Niezależność większej liczby zdarzeń

Definicja. Zdarzenia losowe $A_1, A_2, \ldots, A_n$ są **wzajemnie niezależne** jeżeli dla dowolnego k ($k = 2, 3, \ldots, n$) i dowolnego podzbioru indeksów $i_1, i_2, \ldots, i_k$,

$$P(\bigcap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}).$$

Przykład.
Załóżmy, że zbiór zdarzeń elementarnych jest postaci $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$ i wszystkie zdarzenia elementarne są jednakowo prawdopodobne. Rozważmy następujące zdarzenia losowe:

$$A = \{1, 2, 3, 4\}, \quad B = \{1, 2, 5, 6\}, \quad C = \{2, 3, 6, 7\} \quad \textit{oraz} \quad D = \{1, 2, 7, 8\}.$$

Pokazać, że zdarzenia A, B i C są wzajemnie niezależne a zdarzenia A, B i D są parami niezależne ale nie są wzajemnie niezależne.

Rozwiązanie. Oczywiste jest, że $P(A) = P(B) = P(C) = P(D) = 0.5$ oraz $A \cap B = A \cap D = B \cap D = A \cap B \cap D = \{1, 2\}$, $A \cap C = \{2, 3\}$, $B \cap C = \{2, 6\}$ oraz $A \cap B \cap C = \{2\}$.

Stąd, mamy

$$P(A) \cdot P(B) = 0.5 \cdot 0.5 = 0.25 = P(A \cap B), \; A \text{ i } B \text{ są niezależne},$$

$$P(A) \cdot P(C) = 0.5 \cdot 0.5 = 0.25 = P(A \cap C), \; A \text{ i } C \text{ są niezależne},$$
$$P(A) \cdot P(D) = 0.5 \cdot 0.5 = 0.25 = P(A \cap D), \; A \text{ i } D \text{ są niezależne},$$
$$P(B) \cdot P(C) = 0.5 \cdot 0.5 = 0.25 = P(B \cap C), \; B \text{ i } C \text{ są niezależne},$$
$$P(B) \cdot P(D) = 0.5 \cdot 0.5 = 0.25 = P(B \cap D), \; B \text{ i } D \text{ są niezależne},$$

zatem A, B, C i D są zdarzeniami parami niezależnymi. Ponieważ

$$P(A) \cdot P(B) \cdot P(C) = 0.5 \cdot 0.5 \cdot 0.5 = 0.125 = P(A \cap B \cap C),$$

to A, B and C są wzajemnie niezależne. Z drugiej strony

$$P(A) \cdot P(B) \cdot P(D) = 0.5 \cdot 0.5 \cdot 0.5 = 0.125 \neq 0,25 = P(A \cap B \cap D),$$

zatem A, B i D są (wzajemnie) zależne.
Zdefiniujmy dla dowolnego zdarzenia losowego A

$$A^\epsilon = \begin{cases} A & dla \quad \epsilon = 0 \\ \overline{A} & dla \quad \epsilon = 1. \end{cases}$$

Twierdzenie.
Dla dowolnych n niezależnych zdarzeń losowych $A_1, A_2, \ldots, A_n$ oraz dowolnego ciągu $(\epsilon_1, \ldots, \epsilon_n)$, gdzie $\epsilon_i = 0$ lub 1, zdarzenia losowe $A_1^{\epsilon_1}, A_2^{\epsilon_2}, \ldots, A_n^{\epsilon_n}$ są również wzajemnie niezależne.

Obliczanie prawdopodobieństwa sumy dużej liczby zdarzeń losowych, które nie są wzajemnie rozłączne, jest zadaniem niezwykle żmudnym i pracochłonnym. Jednakże w przypadku wzajemnej niezależności znakomicie się upraszcza.

Twierdzenie. Niech zdarzenia losowe $A_1, A_2, \ldots, A_n$ będą zdarzeniami wzajemnie niezależnymi. Wówczas

$$P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \prod_{i=1}^{n}(1 - P(A_i)).$$

Dowód. Stosując prawo De Morgana otrzymujemy

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &= 1 - P\left(\overline{\bigcup_{i=1}^{n} A_i}\right) = \\
&= 1 - P\left(\bigcap_{i=1}^{n} \overline{A}_i\right) = \\
&= 1 - \prod_{i=1}^{n} P(\overline{A}_i) = \\
&= 1 - \prod_{i=1}^{n} (1 - P(A_i)).
\end{aligned}$$

■

Przykład. *Rozważmy doświadczenie polegające na pięciokrotnym rzucie symetryczną kostką do gry. Niech A będzie zdarzeniem losowym, że wypadnie co najmniej jedna szóstka. Obliczyć $P(A)$.*

Rozwiązanie.
Oznaczmy przez A_i zdarzenie: w i-tym rzucie szóstka, $i = 1, \ldots, 5$. Ponieważ $P(A_i) = 1/6$ to,

$$\begin{aligned}
P(A) &= P\left(\bigcup_{i=1}^{5} A_i\right) = 1 - \prod_{i=1}^{5} (1 - P(A_i)) = \\
&= 1 - \left(\frac{5}{6}\right)^5 = 0.598.
\end{aligned}$$

5.3 Doświadczenia Bernoulliego

Rozważmy doświadczenie losowe E, które kończy się jednym z dwóch możliwych wyników, powiedzmy "sukces" (1) oraz "porażka" (0), z prawdopodobieństwami równymi odpowiednio p i $q = 1 - p$.

Oznaczmy przez E_n doświadczenie składające się z n niezależnych powtórzeń zdarzenia E. Wynik doświadczenia E_n można zapisać jako n wyrazowy ciąg o elementach 1 i 0.

Na przykład, zbiór możliwych wyników doświadczenia E_2 to $\Omega_2 = \{11, 10, 01, 00\}$, z prawdopodobieństwami równymi odpowiednio p^2, pq, qp i q^2.

Doświadczenie kończące się jednym z dwóch wyników, umownie zwanych "sukcesem" i "porażką", nazywamy **doświadczeniem Bernoulliego**, a doświadczenie będące ciągiem niezależnych powtórzeń tego samego doświadczenia Bernoulliego, nazywamy **ciągiem doświadczeń Bernoulliego**, od nazwiska szwajcarskiego matematyka Jakuba Bernoulliego.

Jacob Bernoulli (1654-1705)

Ciąg n doświadczeń Bernoulliego z prawdopodobieństwem sukcesu p oznaczamy przez $b(n, p)$.

Zbiór zdarzeń elementarnych dla doświadczenia $b(n, p)$ składa się z 2^n ciągów o długości n.

Prawdopodobieństwo uzyskania w doświadczeniu $b(n, p)$ ciągu składającego się z k sukcesów i $n - k$ porażek, przy założeniu niezależności doświadczeń, wynosi

$$p^k(1-p)^{n-k}.$$

Ponadto, liczba takich zdarzeń elementarnych, w których mamy dokładnie k "sukcesów", jest równa liczbie kombinacji k elementowych ze zbioru n elementowego, czyli

$$\binom{n}{k}.$$

Twierdzenie.

Prawdopodobieństwo zdarzenia, że uzyskamy dokładnie k sukcesów w n niezależnych doświadczeniach Bernoulliego z prawdopodobieństwem sukcesu $p \in [0, 1]$, wynosi

$$b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.$$

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

Zatem ciąg $\{b(k; n, p)\},\ k = 0, 1, \ldots, n$ jest rozkładem prawdopodobieństwa. Rozkład ten nazywamy **rozkładem dwumianowym** $(b(n, p))$.

Lemat Poissona.
Jeżeli $\{p_n\}$, $n = 1, 2, \ldots$, jest ciągiem liczbowym takim, że $p_n \in (0, 1)$, dla wszystkich n, oraz istnieje granica $\lim_{n\to\infty} n \cdot p_n = \lambda > 0$, to
$$\lim_{n\to\infty} \binom{n}{k} {p_n}^k (1-p_n)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \text{ for all } k = 0, 1, 2, \ldots.$$

Uwaga.
$$\sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} P(k; \lambda) = 1.$$

Zatem ciąg prawdopodobieństw $\{P(k; \lambda)\},\ k = 0, 1, \ldots$ jest rozkładem prawdopodobieństw zwanym **rozkładem Poissona** $(P(\lambda))$.

Używając lematu Poissona (nazwa pochodzi od sławnego francuskiego matematyka) możemy dla dużych n przybliżać wartości prawdopodobieństw w rozkładzie dwumianowym.

Siméon Denis Poisson (1781-1840)

Przybliżanie rozkładu dwumianowego rozkładem Poissona

Dla $n \geqslant 25$ i $n \cdot p \leqslant 10$ oraz dowolnego $k = 0, 1, 2, \ldots$ mamy
$$b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k} \approx e^{-n\cdot p} \cdot \frac{(n \cdot p)^k}{k!}.$$

Przykład. *Prawdopodobieństwo tego, że pewien wirus przeżyje po zaaplikowaniu szczepionki wynosi 0.02. Jakie jest prawdopodobieństwo, że spośród 500 wirusów po zaaplikowaniu szczepionki przeżyje co najmniej 15?*
Rozwiązanie. Nietrudno zauważyć, że mamy $n = 500$ doświadczeń Bernoulliego z prawdopodobieństwem sukcesu $p = 0.02$. Stąd szukane prawdopodobieństwo wynosi

$$\begin{aligned}\sum_{k=15}^{500} b(k; 500, 0.02) &= 1 - \sum_{k=0}^{14} b(k; 500, 0.02) \approx \\ &\approx 1 - F(14; 500 \cdot 0.02) = \\ &= 1 - 0.916542 = 0.083458,\end{aligned}$$

gdzie

$$F(k; \lambda) = \sum_{i=0}^{k} e^{-\lambda} \cdot \frac{\lambda^i}{i!}.$$

Wynik ten uzyskaliśmy stosując przybliżenie Poissona. Funkcja F jest stablicowana.

Dokładna wartość prawdopodobieństwa wynosi $0,0814$.