Wstęp do rachunku prawdopodobieństwa
Wykład 2: Aksjomatyczna definicja prawdopodobieństwa

Dla dowolnego zbioru X przez 2^X oznaczać będziemy rodzinę $\mathcal{F}_X$ wszystkich podzbiorów zbioru X. Przypomnijmy, że moc tej rodziny wynosi $2^{|X|}$.

Definicja 1. *Rodzinę $\mathcal{F} \subseteq 2^\Omega$ podzbiorów zbioru Ω nazywamy **σ-algebrą** jeśli spełnia poniższe warunki:*

(S1) $\Omega \in \mathcal{F}$;
(S2) *Jeśli $A \in \mathcal{F}$, to $A' \in \mathcal{F}$ ($\mathcal{F}$ jest zamknięta na dopełnienia);*
(S3) *Jeśli $A_1, A_2, \ldots \in \mathcal{F}$, to $\bigcup_{i=1}^\infty A_i \in \mathcal{F}$ ($\mathcal{F}$ jest zamknięta na przeliczalne sumy).*

Przykład 1. Dwie najłatwiejsze do zdefiniowana σ-algebry na zbiorze Ω to:

(1) $\mathcal{F} = \{\emptyset, \Omega\}$ – najmniejsze σ-algebra określona na Ω
(2) $\mathcal{F} = 2^\Omega$ – największa σ-algebra określona na Ω

Definicja 2. ***Przestrzenią probabilistyczną** nazywamy trójkę $(\Omega, \mathcal{F}, \mathbb{P})$, gdzie Ω jest **zbiorem zdarzeń elementarnych**, $\mathcal{F}$ jest pewną σ-algebrą podzbiorów zbioru Ω, której elementy nazywamy **zdarzeniami**, natomiast $\mathbb{P} : \mathcal{F} \to [0,1]$ jest **funkcją prawdopodobieństwa** spełniającą następujące aksjomaty:*

(A1) $0 \leqslant \mathbb{P}(A) \leqslant 1$ *dla każdego zdarzenia $A \in \mathcal{F}$;*
(A2) $\mathbb{P}(\Omega) = 1$;
(A3) *Dla dowolnego ciągu parami rozłącznych zdarzeń $A_1, A_2, \ldots$ zachodzi $\mathbb{P}(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mathbb{P}(A_i)$ (przeliczalna addytywność).*

Przykład 2. Rozważmy rzut dwiema kostkami. Jako przestrzeń probabilistyczną dla tego eksperymentu możemy przyjąć trójkę $(\Omega, \mathcal{F}, \mathbb{P})$, gdzie $\Omega = \{1, 2, \ldots, 6\} \times \{1, 2, \ldots, 6\}$, $\mathcal{F} = 2^\Omega$, $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$.

Przykład 3. Rozważmy rzut monetą do momentu wyrzucenia pierwszego orła. Tym razem jako przestrzeń probabilistyczną dla tego eksperymentu możemy przyjąć trójkę $(\Omega, \mathcal{F}, \mathbb{P})$, gdzie $\Omega = \{\omega_1, \omega_2, \ldots\}$, $\mathcal{F} = 2^\Omega$, natomiast zdarzenie elementarne ω_i oznacza wyrzucenia orła po raz pierwszy w i-tym rzucie, $i = 1, 2, \ldots$ Wówczas funkcję prawdopodobieństwa definiujemy najpierw dla zdarzeń elementarnych jako $\mathbb{P}(\{\omega_i\}) = (1/2)^i$. Następnie możemy rozszerzyć tę definicję na wszystkie zdarzenia $A \in \mathcal{F}$ w następujący sposób

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

Twierdzenie 1 (Własności funkcji prawdopodobieństwa)**.**

(W1) $\mathbb{P}(\emptyset) = 0$;
(W2) *Jeśli zdarzenia $A_1, A_2, \ldots, A_n$ są parami rozłączne, to $\mathbb{P}(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i)$ (skończona addytywność);*
(W3) $\mathbb{P}(A') = 1 - \mathbb{P}(A)$;
(W4) *Jeśli $A \subset B$, to $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$;*
(W5) *Jeśli $A \subset B$, to $\mathbb{P}(A) \leqslant \mathbb{P}(B)$ (monotoniczność);*
(W6) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
(W7) $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
(W8) $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$.

Przykład 4. Pokaż, że (W2) implikuje (W3) i (W4).
Zauważmy, że A i A' są zdarzeniami rozłącznymi, zatem na mocy (W2) mamy

$$\mathbb{P}(A) + \mathbb{P}(A') = \mathbb{P}(A \cup A') = \mathbb{P}(\Omega) = 1,$$

skąd otrzymujemy

$$\mathbb{P}(A') = 1 - \mathbb{P}(A).$$

Następnie załóżmy, że $A \subset B$. Zauważmy również, że zdarzenia $B \setminus A$ i A są rozłączne. Zatem na mocy (W2)

$$\mathbb{P}(B \setminus A) + \mathbb{P}(A) = \mathbb{P}((B \setminus A) \cup A) = \mathbb{P}(B),$$

skąd wnioskujemy, że

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A).$$

Przykład 5. Pokaż, że (W4) implikuje (W5).

Załóżmy, że mamy dwa zdarzenia A i B takie, że $A \subset B$. Wówczas korzystając po kolei z własności (W4) oraz aksjomatu (A1) otrzymujemy

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geqslant \mathbb{P}(A).$$

Przykład 6. Pokaż, że jeśi $\mathbb{P}(A) + \mathbb{P}(B) > 1$, to zdarzenia A i B nie mogą się wykluczać.

Załóżmy, że zdarzenia A i B wykluczają się, czyli $A \cap B = \emptyset$. Z własności (W2) oraz aksjomatu (A1) wynika wówczas, że

$$\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cup B) \leqslant 1.$$

Sprzeczność!

Na poprzednim wykładzie zdefiniowaliśmy pojęcie niezależności zmiennych losowych. W dokładnie taki sam sposób definiujemy niezależność zmiennych losowych w dowolnych przestrzeniach probabilistycznych. Udowodnimy teraz twierdzenie z poprzedniego wykładu dotyczące zdarzeń niezależnych.

Twierdzenie 2 (Własności zdarzeń niezależnych)**.**

(1) *Jeśli A i B są zdarzeniami niezależnymi, to A i B' są niezależne, A' i B są niezależne oraz A' i B' są niezależne.*

(2) *Jeśli $A_1, A_2, \ldots, A_n$ są niezależne oraz przyjmiemy konwencję, że $A^0 = A$, $A^1 = A'$ dla dowolnego zdarzenia A, wówczas dla każdego ciągu $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, gdzie $\varepsilon_i \in \{0, 1\}$ dla $i = 1, 2, \ldots, n$, zdarzenia $A_1^{\varepsilon_1}, A_2^{\varepsilon_2}, \ldots, A_n^{\varepsilon_n}$ są niezależne.*

dowód. *(1) Załóżmy, że zdarzenia A i B są niezależne. Pokażemy, że wówczas zdarzenia A i B' są też niezależne.*

A zatem skoro A i B są niezależne, zachodzi $\mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(A \cap B)$. Wykorzystując po kolei własności (W3) oraz (W4) otrzymujemy

$$\mathbb{P}(A) \cdot \mathbb{P}(B') = \mathbb{P}(A) \cdot (1 - \mathbb{P}(B)) = \mathbb{P}(A) - \mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A \setminus A \cap B) = \mathbb{P}(A \cap B'),$$

skąd wnioskujemy, że zdarzenia A i B' są niezależne. Stosując to samo rozumowanie możemy pokazać, że również zdarzenia A' i B oraz A' i B' są niezależne.

W podobny sposób możemy pokazać, że (2) również zachodzi. Rozważmy rodzinę zdarzeń $A_1, A_2, \ldots, A_n$, które są niezależne. Pokażemy, że zastępując którekolwiek ze zdarzeń A_i, $i = 1, 2, \ldots, n$, zdarzeniem A_i', nasza rodzina nadal będzie niezależna. Bez straty ogólności możemy przyjąć, że zastępujemy zdarzenie A_n zdarzeniem A_n'. Wówczas korzystając z niezależności zdarzeń $A_1, A_2, \ldots, A_n$ oraz własności (W3) i (W4) mamy:

$$\begin{aligned}
\mathbb{P}(A_1) \cdot \ldots \cdot \mathbb{P}(A_{n-1}) \cdot \mathbb{P}(A_n') &= \mathbb{P}(A_1 \cap \ldots \cap A_{n-1})(1 - \mathbb{P}(A_n)) \\
&= \mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) - \mathbb{P}(A_1 \cap \ldots \cap A_{n-1})\mathbb{P}(A_n) \\
&= \mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) - \mathbb{P}(A_1 \cap \ldots \cap A_{n-1} \cap A_n) \\
&= \mathbb{P}((A_1 \cap \ldots \cap A_{n-1}) \setminus (A_1 \cap \ldots \cap A_{n-1} \cap A_n)) \\
&= \mathbb{P}((A_1 \cap \ldots \cap A_{n-1} \cap A_n'),
\end{aligned}$$

skąd otrzymujemy, że rodzina $A_1, A_2, \ldots, A_{n-1}, A_n'$ również stanowi rodzinę zdarzeń niezależnych. Sukcesywnie zamieniając inne zdarzenia A_i w tej rodzinie na A_i' i powtarzając powyższe rozumowanie jesteśmy w stanie pokazać, że dla dowolnego ciągu $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, gdzie $\varepsilon_i \in \{0, 1\}$ dla $i = 1, 2, \ldots, n$, $A_1^{\varepsilon_1}, A_2^{\varepsilon_2}, \ldots, A_n^{\varepsilon_n}$ jest rodziną zdarzeń niezależnych. □

Twierdzenie 3 (Nierówność Boole'a)**.** *Dla dowolnego ciągu zdarzeń $A_1, A_2, \ldots$ zachodzi*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

dowód. *Aby udowodnić nierówność Boole'a skorzystamy z aksjomatu (A3). W tym celu zdefiniujemy następujący ciąg zdarzeń losowych:*

$$\begin{aligned} B_1 &:= A_1 \\ B_2 &:= A_2 \setminus B_1 \\ B_3 &:= A_2 \setminus (B_1 \cup B_2) \\ &\vdots \\ B_k &:= A_k \setminus (B_1 \cup B_2 \cup \cdots \cup B_{k-1}) \\ &\vdots \end{aligned}$$

Wprost z definicji ciągu zdarzeń $(B_i)_{i \in \mathbb{N}_+}$ wynika, że

$$B_k := A_k \setminus (B_1 \cup B_2 \cup \cdots \cup B_{k-1}) = A_k \setminus (A_1 \cup A_2 \cup \cdots \cup A_{k-1}).$$

Ponadto każde ze zdarzeń B_i jest rozłączne z każdym zdarzeniem B_j, $j < i$. Zatem ciąg zdarzeń $(B_i)_{i \in \mathbb{N}_+}$ składa się z parami rozłącznych zdarzeń i możemy zastosować do niego aksjomat (A3). Co więcej, ponieważ dla każdego $i = 1, 2, \ldots$ zachodzi

$$B_i \subseteq A_i,$$

otrzymujemy

$$\mathbb{P}(B_i) \leqslant \mathbb{P}(A_i)$$

oraz

$$\bigcup_{i=1}^{\infty} B_i \subseteq \bigcup_{i=1}^{\infty} A_i.$$

Pokażemy, że powyższe zawieranie zachodzi również w drugą stroną, co oznacza, że oba zbiory są sobie równe. W tym celu weźmy dowolny element $x \in \bigcup_{i=1}^{\infty} A_i$. Niech k będzie najmniejszym indeksem takim, że $x \in A_k$ oraz $x \notin A_j$ dla $j < k$. Wówczas $x \in B_k = A_k \setminus (A_1 \cup A_2 \cup \cdots \cup A_{k-1})$, skąd wynika, że $x \in \bigcup_{i=1}^{\infty} B_i$. A zatem

$$\bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i=1}^{\infty} B_i.$$

Ostatecznie, stosując aksjomat (A3) do ciągu zdarzeń $(B_i)_{i \in \mathbb{N}_+}$, otrzymujemy

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) \leqslant \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

□

Twierdzenie 4 (o ciągłości)**.**

(1) *Niech $A_1, A_2, \ldots$ będzie wstępującym ciągiem zdarzeń, tzn. $A_1 \subseteq A_2 \subseteq \ldots$ Wówczas*

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \lim_{i \to \infty} \mathbb{P}(A_i).$$

(2) *Niech $A_1, A_2, \ldots$ będzie zstępującym ciągiem zdarzeń, tzn. $A_1 \supset A_2 \supset \ldots$ Wówczas*

$$\mathbb{P}(\bigcap_{i=1}^{\infty} A_i) = \lim_{i \to \infty} \mathbb{P}(A_i).$$

Definicja 3. *Niech* $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ *oraz* $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ *będą dwoma przestrzeniami probabilistycznymi.* **Iloczynem (produktem)** *tych przestrzeni nazywamy przestrzeń probabilistyczną* $(\Omega, \mathcal{F}, \mathbb{P})$*, gdzie* $\Omega = \Omega_1 \times \Omega_2$*,* $\mathcal{F}$ *zawiera wszystkie zbiory postaci* $A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2$*, oraz funkcja prawdopodobieństwa* $\mathbb{P}$ *spełnia warunek*

$$\mathbb{P}(A_1 \times A_2) = \mathbb{P}_1(A_1) \cdot \mathbb{P}_2(A_2).$$

Definicję tę można uogólnić na dowolną skończoną lub przeliczalną rodzinę przestrzeni probabilistycznych.

Uwaga 1. W powyższej definicji $\mathcal{F}$ jest σ-algebrą generowaną przez zbiory postaci $A_1 \times A_2$, czyli najmniejszą σ-algebrą, która zawiera wszystkie takie zbiory.

Przykład 7 (Schemat Bernoulliego). Niech $p \in (0,1)$ oraz $n \in \mathbb{N}_+$. Przypuśćmy, że powtarzamy n razy eksperyment losowy, w którym prawdopodobieństwo sukcesu wynosi p (a porażki $1-p$) oraz wyniki poszczególnych prób nie mają na siebie nawzajem wpływu. Jeśli τ_k oznacza prawdopodobieństwo uzyskania dokładnie k sukcesów w n próbach, to

$$\tau_k = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{dla} \quad k = 0, 1, \ldots, n.$$

Przykład 8 (Rozkład geometryczny). Powtarzamy eksperyment losowy, w którym prawdopodobieństwo sukcesu wynosi $p \in (0,1)$, do momentu osiągnięcia pierwszego sukcesu. Jeśli przez σ_k oznaczymy prawdopodobieństwo, że potrzebujemy na to k prób, to

$$\sigma_k = p(1-p)^{k-1} \quad \text{dla} \quad k = 1, 2, \ldots$$

Przykład 9. Ile wynosi prawdopodobieństwo, że w dziesięciu rzutach monetą wypadną co najmniej trzy orły? Wykorzystamy schemat Bernoulliego, gdzie przez sukces rozumiemy wyrzucenie orła. Wówczas przyjmujemy $p = 1/2$ oraz $n = 10$. Szukane prawdopodobieństwo jest równe

$$1 - \tau_0 - \tau_1 - \tau_2 = 1 - \binom{10}{0}\left(\frac{1}{2}\right)^{10} - \binom{10}{1}\left(\frac{1}{2}\right)^{10} - \binom{10}{2}\left(\frac{1}{2}\right)^{10}.$$

Definicja 4. *Jeśli* $(\Omega, \mathcal{F}, \mathbb{P})$ *jest przestrzenią probabilistyczną, a* $B \in \mathcal{F}$ *jest zdarzeniem,* **dla którego** $\mathbb{P}(B) > 0$*, wtedy możemy skonstruować* **przestrzeń warunkową** $(B, \mathcal{F}_B, \mathbb{P}_B)$ *przyjmując*

$$\mathcal{F}_B = \{A \cap B : A \in \mathcal{F}\},$$

a dla każdego $A \in \mathcal{F}$

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

Prawdopodobieństwo $\mathbb{P}_B(A)$ *zwykle zapisujemy jako* $\mathbb{P}(A|B)$ *i czytamy jako* **prawdopodobieństwo (warunkowe) zdarzenia** A **pod warunkiem (zajścia) zdarzenia** B.

Uwaga 2. Jeśli zdarzenia A i B są niezależne, to

$$\mathbb{P}(A|B) = \mathbb{P}(A) \quad \text{i} \quad \mathbb{P}(B|A) = \mathbb{P}(B)\,,$$

zakładając, że $\mathbb{P}(A), \mathbb{P}(B) > 0$ (bez tego założenia prawdopodobieństwa warunkowe nie są dobrze zdefiniowane).

dowód. *Z definicji prawdopodobieństwa warunkowego oraz niezależności zdarzeń* A *i* B *mamy*

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

□

Obserwacja 1. Przy założeniu, że $\mathbb{P}(A), \mathbb{P}(B) > 0$ mamy

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

A zatem

$$\mathbb{P}(B|A) > \mathbb{P}(B) \iff \mathbb{P}(A|B) > \mathbb{P}(A),$$

co można interpretować jako: zdarzenie A "sprzyja" zdarzeniu B wtedy i tylko wtedy, gdy zdarzenie B "sprzyja" zdarzeniu A.

Przykład 10. W trzykrotnym rzucie monetą wypadła nieparzysta liczba orłów. Ile wynosi prawdopodobieństwo, że wypadły dokładnie trzy orły?

A – nieparzysta liczba orłów, $A = \{OOO, ORR, ROR, RRO\}$

B – dokładnie trzy orły, $B = \{OOO\}$

$\mathbb{P}(B|A)$ – prawdopodobieństwo, że wypadły dokładnie trzy orły, jeśli wiemy, że wypadła nieparzysta liczba orłów

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\frac{1}{8}}{\frac{4}{8}} = \frac{1}{4}$$

Przykład 11. Jaś strzela do tarczy tak długo aż trafi dziesiątkę po raz pierwszy. Prawdopodobieństwo trafienia dziesiątki w pojedynczym strzale wynosi $1/10$. Z jakim prawdopodobieństwem Jaś odda dokładnie sześć strzałów, jeśli trzy pierwsze były pudłem?

A – pierwsze trzy strzały spudłowane

B – Jaś odda dokładnie sześć strzałów

Możemy zastosować tutaj rozkład geometryczny, gdzie przez sukces rozumiemy strzał w dziesiątkę, zatem przyjmujemy $p = 1/10$.

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\frac{1}{10}\left(1 - \frac{1}{10}\right)^5}{\left(1 - \frac{1}{10}\right)^3} = \frac{1}{10} \cdot \left(\frac{9}{10}\right)^2 = 0,081$$