Andrzej Sierociński

RACHUNEK PRAWDOPODOBIEŃSTWA

Elementy Rachunku Prawdopodobieństwa
Wykład 1

1 Pojęcia wstępne

Początki rachunku prawdopodobieństwa sięgają połowy XVII wieku, kiedy to na zapotrzebowanie prominentnych graczy, dwaj francuscy matematycy: Błażej Pascal oraz Pierre de Fermat

Blaise Pascal (1623-1662)

Pierre de Fermat (1601-1665)

stworzyli pierwszy matematyczny model pozwalający na opis oraz wyznaczenie szans pojawienia się pewnych wyników w grach hazardowych, takich jak gra w kości.

Język prawdopodobieństwa jest powszechnie używany w mowie potocznej. Na przykład w następujących sformułowaniach:

1. "Jest wysokie prawdopodobieństwo wystąpienia gwałtownych burz wieczorem",
2. "Kandydat A ma 50-50% szans na ponowny wybór",
3. "Oczekuje się, że 90% klientów sieci sklepów B skorzysta z bieżącej promocji",
4. "W grupie osób palących istnieje wyższe ryzyko wystąpienia chorób nowotworowych niż w grupie osób niepalących".

Sformułowania te, choć nieprecyzyjne, coraz częściej pojawiają się w naszym życiu codziennym. Naszym celem będzie stworzenie teorii oraz metodologii, która pozwoli na matematyczne, precyzyjne przedstawienie stwierdzeń podobnych do zaprezentowanych powyżej.

1.1 Zbiór zdarzeń elementarnych

Przy budowaniu dowolnej teorii musimy w pierwszej kolejności zdefiniować pewne podstawowe obiekty i pojęcia.

W przypadku dowolnego eksperymentu (zarówno deterministycznego jak i niedeterministycznego) podstawową rolę odgrywa zbiór możliwych wyników tego eksperymentu. Nas będą interesowały jedynie eksperymenty niedeterministyczne, czyli losowe, tzn. takie, dla których przy spełnieniu identycznych warunków nie jesteśmy w stanie określić (zdeterminować) jego wyniku.

Do takich klasycznych eksperymentów losowych zaliczamy np. rzut symetryczną monetą czy rzut symetryczną kostką do gry, gdzie przy identycznych warunkach rzutu, w żaden sposób nie jesteśmy zdolni do 100% przewidywania jego wyniku. I chociaż nie jesteśmy zdolni do przewidywania wyniku, to zbiór możliwych wyników jest znany.

Definicja. Każdy możliwy wynik eksperymentu losowego nazywamy **zdarzeniem elementarnym** ω, a zbiór wszystkich możliwych wyników eksperymentu (wszystkich zdarzeń elementarnych) nazywamy **zbiorem zdarzeń elementarnych** i oznaczamy grecką literą $\Omega \quad (\omega \in \Omega)$.

Przykład. *Najprostszym eksperymentem losowym jest eksperyment z dwoma możliwymi wynikami,* $\Omega = \{\omega_1, \omega_2\}$.
Na przykład, przy rzucie monetą mamy dwa możliwe wyniki: orzeł lub reszka, wówczas $\Omega = \{O, R\}$.

Przykład. *Załóżmy, że sprawdzamy 3 żarówki. Niech* D *oznacza, że żarówka jest dobra, a* N*, że nie jest dobra. Wówczas wynikiem eksperymentu jest dowolny ciąg złożony z 3 symboli* N *lub* D*.*

$$\Omega = \{NNN,\ NND,\ NDN,\ DNN,\ NDD,\ DND,\ DDN,\ DDD\}.$$

Identycznie będzie wyglądał zbiór zdarzeń elementarnych w przypadku trzykrotnego rzutu monetą (jedynie zastępujemy D i N przez O i R.)

Przykład. *Załóżmy, że chcemy ustalić dawkę leku, którą należy podać pacjentowi , aby odniósł pozytywny skutek. Wiemy,że dawka leku, w żadnym przypadku, nie powinna przekroczyć pewnej ustalonej wielkości* M*.*

W tym przypadku zbiór zdarzeń elementarnych zawiera wszystkie dodatnie wielkości nie przekraczające M*, tzn.*

$$\Omega = \{\omega : 0 < \omega \leqslant M\} = (0, M],$$

gdzie ω *oznacza dawkę leku, na którą pacjent reaguje pozytywnie, ale nie ma pozytywnej reakcji na żadną dawkę mniejszą od* ω*.*

Jak widać z powyższych przykładów, zbiór zdarzeń elementarnych może być zbiorem skończonym lub nie, a nawet zbiorem nieprzeliczalnym.

1.2 Zdarzenia losowe

Zdarzenia elementarne są najprostszymi z możliwych wynikami eksperymentu losowego. W praktyce zwykle jesteśmy zainteresowani zdarzenaimi bardziej złożonymi.

Np. dla przykładu z żarówkami możemy zapytać o zdarzenie pojawienia się tylko jednej żarówki wadliwej, wówczas takie zdarzenie nie będzie zdarzeniem elementarnym i będzie się składało z 3 zdarzeń elementarnych, mianowicie

$$\begin{aligned} A &= \{\text{wśród 3 badanych żarówek jest jedna wadliwa}\} = \\ &= \{NDD,\ DND,\ DDN\}. \end{aligned}$$

Oczywiście $A \subset \Omega$ *jest podzbiorem zbioru zdarzeń elementarnych.*

Powstaje pytanie, czy dowolny podzbiór zbioru zdarzeń elementarnych jest zdarzeniem losowym?

W przykładzie, ponieważ zbiór Ω jest skończony, tak jest.

Rodzinę wszystkich zdarzeń losowych w dalszym ciągu będziemy oznaczali przez $\mathcal{F}$.

Dla przykładu z żarówkami $\mathcal{F} = 2^{\Omega}$, (2^{Ω} oznacza zbiór wszystkich podzbiorów).

Jednakże nie zawsze można postępować w ten sposób. Jeżeli zbiór zdarzeń elementarnych jest zbiorem nieprzeliczalnym, to z pewnych względów musimy zawęzić rodzinę zdarzeń losowych do rodziny podzbiorów Ω posiadających strukturę mnogościową tzw. σ-ciała. Wiąże się to z niemożliwością poprawnego zdefiniowania prawdopodobieństwa, dla pewnych, szczególnych, podzbiorów zbioru Ω, nie mających dla nas praktycznego znaczenia.

Definicja. Rodzinę $\mathcal{F}$ podzbiorów zbioru Ω nazywamy σ-ciałem zdarzeń losowych jeżeli

1. $\Omega \in \mathcal{F}$,
2. jeżeli $A \in \mathcal{F}$ to $\bar{A} \in \mathcal{F}$;
3. jeżeli $\forall n \in \mathcal{N}\ A_n \in \mathcal{F}$ to $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$.

Najprostszym σ-ciałem jest tzw. **σ-ciało trywialne** $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Najprostszym nietrywialnym σ-ciałem jest σ-ciało zawierające cztery zbiory $\mathcal{F}_1 = \{\emptyset, A, \bar{A}, \Omega\}$, gdzie $A \subset \Omega$ jest dowolnym, istotnym i niepustym podzbiorem zbioru zdarzeń elementarnych.

Jeżeli zbiór zdarzeń elementarnych jest zbiorem przeliczalnym, to przyjmujemy $\mathcal{F} = 2^{\Omega}$.

Aby zdefiniować szanse pojawienia się poszczególnych zdarzeń losowych posłużymy się doświadczeniem. Załóżmy, że chcemy określić szansę (prawdopodobieństwo) pewnego zdarzenia losowego $A \subset \Omega$. Zdarzenie to może zajść lub nie w wyniku przeprowadzonego eksperymentu losowego. Załóżmy, że eksperyment ten można przeprowadzić w niezmienionych warunkach dowolną liczbę razy. Oznaczmy przez k_n liczbę zajść zdarzenia A w n eksperymentach.

Wówczas można zdefiniować empiryczną (względną) częstość pojawienia się zdarzenia A jako iloraz

$$\frac{k_n}{n} \to p_A \quad przy \quad n \to \infty.$$

W wyniku takiego eksperymentu można zaobserwować tzw. prawo stabilności empirycznej częstości wokół pewnej wielkości $p_A \in [0, 1]$ zależnej jedynie od zdarzenia A.

Prawo to jest znane jako **empiryczne prawo wielkich liczb**. Stałą, którą w ten sposób uzyskujemy, jesteśmy skłoni nazwać prawdopodobieństwem zdarzenia A.

1.3 Aksjomaty prawdopodobieństwa

Aby poprawnie zdefiniować prawdopodobieństwo zdarzeń losowych niezbędne jest przyjęcie pewnych postulatów jako fundamentu teorii matematycznej.

W teorii prawdopodobieństwa te postulaty nazywamy **aksjomatami (pewnikami) prawdopodobieństwa**.

Aksjomaty prawdopodobieństwa są zdaniami wyodrębnionymi spośród wszystkich twierdzeń teorii prawdopodobieństwa, wybranymi tak, aby wynikały z nich wszystkie pozostałe twierdzenia tej teorii.

Taki układ aksjomatów teorii prawdopodobieństwa nazywany jest **aksjomatyką prawdopodobieństwa**. Ich twórcą jest rosyjski matematyk **Andriej Kołmogorow (1903 - 1987)**.

Definicja. *Dla danego zbioru zdarzeń elementarnych Ω oraz σ-ciała zdarzeń losowych $\mathcal{F}$,* ***prawdopodobieństwem (funkcją prawdopodobieństwa)*** *nazywamy funkcję zbioru P, $P : \mathcal{F} \to \mathcal{R}$ spełniającą następujące aksjomaty prawdopodobieństwa:*

Aksjomat 1. Dla dowolnego zdarzenia losowego $A \in \mathcal{F}$,

$$P(A) \geqslant 0.$$

Aksjomat 2. $P(\Omega) = 1$.

Aksjomat 3. Dla dowolnego, nieskończonego ciągu zdarzeń losowych $A_1, A_2, \ldots$, $\forall n \in \mathcal{N}\ A_n \in \mathcal{F}$, parami rozłącznych $\forall i \neq j\ A_i \cap A_j = \emptyset$, mamy

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Do stworzenia opisu matematycznego dowolnego eksperymentu losowego niezbędne jest zdefiniowanie trzech obiektów:

1. zbioru zdarzeń elementarnych Ω,
2. rodziny (σ-ciała) zdarzeń losowych $\mathcal{F}$,
3. funkcji prawdopodobieństwa P.

Definicja. ***Przestrzenią probabilistyczną*** *nazywamy trójkę*

$$(\Omega, \mathcal{F}, P).$$

- Jeżeli zbiór zdarzeń elementarnych Ω jest zbiorem przeliczalnym, to jako σ-ciało zdarzeń losowych przyjmujemy $\mathcal{F} = 2^{\Omega}$.
- Jeżeli $\Omega \subset \mathcal{R}^d$ jest podzbiorem przestrzeni euklidesowej d-wymiarowej, to jako σ-ciało zdarzeń losowych przyjmujemy $\mathcal{F} = \mathcal{B}(\Omega)$, gdzie $\mathcal{B}(\Omega)$ jest tzw. σ-ciałem zbiorów borelowskich (najmniejszym σ-ciałem zawierającym wszystkie podzbiory otwarte i domknięte zbioru Ω). Dowód istnienia takiego σ-ciała pomijamy, ponieważ leży on poza zakresem tego wykładu.

2 Podstawowe własności prawdopodobieństwa

Korzystając z aksjomatyki prawdopodobieństwa udowodnimy kilka ważnych własności prawdopodobieństwa.

Własność 1.

$$P(\emptyset) = 0.$$

Dowód. Korzystając z aksjomatu 3, kładąc $A_n = \emptyset$, $n = 1, 2, \ldots$, otrzymujemy

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P(\emptyset) = p = \sum_{n=1}^{\infty} P(A_n) = p + \sum_{n=2}^{\infty} P(A_n).$$

Stąd

$$\sum_{n=2}^{\infty} P(A_n) = 0.$$

Ostatecznie, ponieważ $P(A_n) = P(\emptyset) = p \geqslant 0$ to z aksjomatu 1 mamy $P(\emptyset) = p = 0$. ■

Uwaga. Jeżeli zdarzenia A i B są wzajemnie rozłączne, to $P(A \cap B) = 0$.

Własność 2. Jeśli $A_1, A_2, \ldots, A_n$ są parami rozłącznymi zdarzeniami losowymi $A_i \in \mathcal{F}$, $i = 1, 2, \ldots, n$, $\forall i \neq j,\ i, j \leqslant n \quad A_i \cap A_j = \emptyset$, to

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

Dowód. Połóżmy $A_i = \emptyset$, $i = n+1, n+2, \ldots$, wówczas

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$$

i z własności 1 mamy

$$\sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{n} P(A_i).$$

Korzystając z aksjomatu 3 otrzymujemy

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

■

Własność 3. Dla dowolnego zdarzenia losowego $A \in \mathcal{F}$

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

Dowód. Ponieważ zdarzenia losowe A i $\bar{A}$ są wzajemnie rozłączne oraz $A \cup \bar{A} = \Omega$, to korzystając z własności 2 oraz aksjomatu 2 otrzymujemy, że

$$P(A) + P(\bar{A}) = P(A \cup \bar{A}) = P(\Omega) = 1.$$

■

Własność 4. Dla dowolnych zdarzeń losowych $A, B \in \mathcal{F}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Dowód. Zauważmy, że

$$(A \cup B) = B \cup (A - B) = B \cup (A \cap \bar{B}),$$

oraz

$$A = (A \cap B) \cup (A \cap \bar{B}).$$

Ponieważ B i $(A \cap \bar{B})$ są wzajemnie rozłączne podobnie jak $(A \cap B)$ i $(A \cap \bar{B})$, to z własności 2 mamy

$$P(A \cup B) = P(B) + P(A \cap \bar{B}),$$

oraz

$$P(A) = P(A \cap B) + P(A \cap \bar{B}).$$

Łącząc te dwa wyniki otrzymujemy tezę

$$\begin{aligned} P(A \cup B) &= P(B) + P(A \cap \bar{B}) = P(B) + [P(A) - P(A \cap B)] = \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

■

Przykład. *Pokazać, że dla dowolnych zdarzeń* $A, B, C \in \mathcal{F}$

$$\begin{aligned} P(A \cup B \cup C) = \quad & P(A) + P(B) + P(C) - P(A \cap B) + \\ - \quad & P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

Dowód.

$$\begin{aligned}
P(A\cup B\cup C) &= P((A\cup B)\cup C)= \\
&= P(A\cup B)+P(C)-P((A\cup B)\cap C)= \\
&= P(A\cup B)+P(C)-P((A\cap C)\cup(B\cap C))= \\
&= P(A\cup B)+P(C)+ \\
&\quad -[P((A\cap C))+P(B\cap C)-P((A\cap C)\cap(B\cap C))]= \\
&= P(A)+P(B)+P(C)-P(A\cap B)+ \\
&\quad -P(A\cap C)-P(B\cap C)+P(A\cap B\cap C).
\end{aligned}$$

Własność 5. Dla dowolnych zdarzeń losowych $A, B\in\Omega$ takich, że $A\subset B$, mamy

$$P(A)\leqslant P(B).$$

Dowód. Jeżeli $A\subset B$, to $B=A\cup(B-A)$. Stąd

$$P(B)=P(A)+P(B-A)\geqslant P(A).$$

■

Uwaga. Jeżeli $A\subset B$, to $P(B-A)=P(B)-P(A)$.

Uwaga. Dla dowolnego zdarzenia losowego $A\in\mathcal{F}$, $P(A)\leqslant 1$.

Własność 6. (ciągłość prawdopodobieństwa)
Jeżeli zdarzenia losowe $A_n\in\mathcal{F}$, $n\in\mathcal{N}$ tworzą ciąg wstępujący zdarzeń, tzn. $\forall n\in\mathcal{N}\ A_n\subset A_{n+1}$, to

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}P(A_n).$$

Dowód. Niech $A_0=\emptyset$ oraz $\forall n\in\mathcal{N}\ B_n=A_n-A_{n-1}$.
Wówczas zdarzenia losowe B_n są wzajemnie rozłączne ($\forall i\neq j\ B_i\cap B_j=\emptyset$) oraz $\forall n\in\mathcal{N}$

$$A_n=\bigcup_{i=1}^{n}A_i=\bigcup_{i=1}^{n}B_i.$$

Z własności 2 otrzymujemy, że

$$P(A_n)=\sum_{i=1}^{n}P(B_i).$$

Zatem

$$\begin{aligned}\lim_{n\to\infty} P(A_n) &= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \sum_{i=1}^{\infty} P(B_i) = \\ &= P\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} A_n\right).\end{aligned}$$

■

Uwaga. Korzystając z prawa de Morgana, własność tę można sformułować dualnie dla ciągu zstępującego zdarzeń ($\forall n \in \mathcal{N}\ A_n \supset A_{n+1}$). Wówczas

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n).$$

3 Przykłady definiowania funkcji prawdopodobieństwa

3.1 Skończony zbiór zdarzeń elementarnych jednakowo prawdopodobnych

W przypadku, gdy zbiór zdarzeń elementarnych $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ składa się z N różnych jednakowo prawdopodobnych wyników, to zadanie obliczenia prawdopodobieństwa dowolnego zdarzenia losowego ogranicza się do wyznaczenia liczności tego zdarzenia.

W szczególności, jeżeli N jest licznością Ω a $N(A) = \bar{\bar{A}}$ jest licznością zdarzenia A, to

$$P(A) = \frac{N(A)}{N} - tzw. \text{ **klasyczna definicja prawdopodobieństwa**.}$$

W tym przypadku ograniczamy się jedynie do wykorzystywania technik kombinatorycznych wyznaczania liczności interesujących zbiorów bez konieczności wypisywania wszystkich zdarzeń elementarnych.

Klasyczna definicja prawdopodobieństwa po raz pierwszy pojawiła się w pracach Pierre-Simona Laplace'a.

Pierre-Simon Laplace (1749–1827)

3.2 Zbiór zdarzeń elementarnych przeliczalny

Bez większego problemu możemy zdefiniować prawdopodobieństwo dla zdarzeń losowych będących podzbiorami co najwyżej przeliczalnego zbioru zdarzeń elementarnych $\Omega = \{\omega_1, \omega_2, \ldots\}$. Wystarczy wówczas zdefiniować funkcję prawdopodobieństwa dla poszczególnych zdarzeń elementarnych. Przyjmijmy zatem, że

$$P(\omega_i) = p_i, \quad \forall i,$$

gdzie

1. $p_i \geqslant 0 \quad \forall i;$

2. $\sum\limits_i p_i = 1.$

Oba warunki są warunkami koniecznymi z uwagi na to, że funkcja P musi spełniać aksjomaty prawdopodobieństwa.

Nietrudno zauważyć, że są one również warunkami wystarczającymi, gdyż dla dowolnego zdarzenia losowego $A \subset \Omega$ mamy $A = \{\omega_{i_1}, \omega_{i_2}, \ldots\}$ dla pewnego podciągu skończonego lub nie indeksów $\{i_k\}$. Z uwagi na rozłączność zdarzeń elementarnych mamy

$$P(A) = \sum_k p_{i_k}.$$

Klasyczna definicja prawdopodobieństwa jest szczególnym przypadkiem sposobu definiowania prawdopodobieństwa dla przestrzeni probabilistycznej ze skończonym (a więc i przeliczalnym) zbiorem zdarzeń elementarnych.

3.3 Prawdopodobieństwo geometryczne

Jeżeli przestrzeń probabilistyczna jest konstruowana na zbiorze zdarzeń elementarnych nieprzeliczalnym, to nie jest możliwe bezpośrednie przeniesienie definicji prawdopodobieństwa stosowanej dla zbioru zdarzeń elementarnych, który jest przeliczalny i zdefiniowania prawdopodobieństwa jedynie dla zdarzeń elementarnych. Musimy to zrobić całkowicie w inny sposób.

Podamy teraz jeden z możliwych sposobów definiowania prawdopodobieństwa, w przypadku, gdy zbiór zdarzeń elementarnych jest podzbiorem przestrzeni euklidesowej $\mathcal{R}^d$ dla pewnego $d \geqslant 1$ a poszczególne punkty zbioru zdarzeń elementarnych mają "jednakowe szanse", w tym sensie, że dwa zbiory o tej samej mierze Jordana w tej przestrzeni euklidesowej mają te same prawdopodobieństwa.

Niech $\Omega \subset \mathcal{R}^d$, $d \geqslant 1$ taki, że $\mu_d(\Omega) < \infty$, gdzie μ_d oznacza miarę Jordana na przestrzeni euklidesowej $\mathcal{R}_d$ (dla R^1 - odległość, dla $\mathcal{R}^2$ - pole, dla $\mathcal{R}^3$ - objętość, i.t.d.).

Prawdopodobieństwo geometryczne Dla dowolnego $A \subset \Omega$ dla którego zdefiniowana jest miara Jordana $\mu_d(A)$ prawdopodobieństwo tego zdarzenia jest zdefiniowane jako

$$P(A) = \frac{\mu_d(A)}{\mu_d(\Omega)}.$$

Nietrudno sprawdzić, że tak zdefiniowana funkcja prawdopodobieństwa spełnia aksjomaty prawdopodobieństwa.