Statystyka matematyczna

Mieczysław Sobczyk

Wydawnictwo C.H. Beck

Statystyka
matematyczna

Statystyka matematyczna

Mieczysław Sobczyk

WYDAWNICTWO C.H. BECK
WARSZAWA 2010

Wydawca: Dorota Ostrowska-Furmanek
Redakcja merytoryczna: Danuta Kamińska-Hass
Recenzent: prof. dr hab. Józef Pociecha
Projekt okładki i stron tytułowych: Maryna Wiśniewska
Ilustracja na okładce: ©MarkEvans/iStockphoto.com

Seria: Metody ilościowe

Złożono programem TeX

Wydawnictwo C.H. Beck Sp. z o.o.
ul. Bonifraterska 17, 00-203 Warszawa

Skład i łamanie: Wydawnictwo C.H. Beck
Druk i oprawa: Poznańskie Zakłady Graficzne

ISBN 978-83-255-1606-2

Spis treści

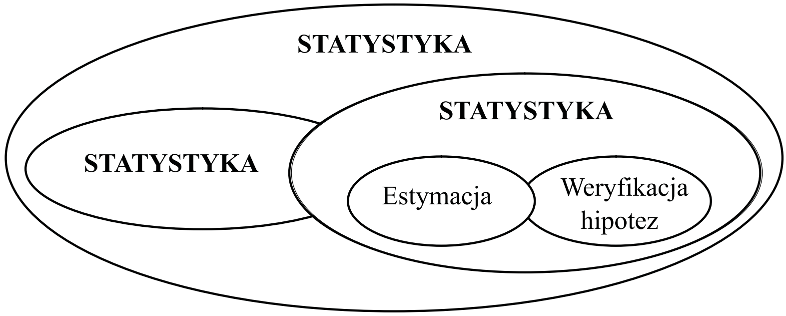
Wstęp

Do podejmowania prawidłowych decyzji i analizowania procesów gospodarczych, demograficznych, społecznych itp. niezbędna jest znajomość metod ilościowych. Z tego też względu w programach nauczania uczelni wyższych przewidziano zajęcia ze statystyki. Programy te obejmują zazwyczaj statystykę opisową (studia licencjackie) oraz statystykę matematyczną (studia magisterskie).

Przedmiotem zainteresowań statystyki opisowej są problemy programowania badań statystycznych, metody obserwacji statystycznej, sposoby opracowywania i prezentacji materiału statystycznego oraz syntetyzująca (sumaryczna) charakterystyka – za pomocą odpowiednich parametrów – właściwości zbioru danych. Metody statystyki opisowej są wykorzystywane wtedy, gdy obserwacją jest objęta cała zbiorowość (populacja generalna). Mówimy wówczas o badaniach pełnych, wyczerpujących.

Statystyka matematyczna (zwana też statystyką indukcyjną) zajmuje się metodami wnioskowania, tj. estymacji i weryfikacji hipotez o całej zbiorowości (populacji generalnej) na podstawie wybranej w sposób losowy pewnej jej części, określanej mianem próby losowej. Metody statystyki matematycznej znajdują szerokie zastosowanie w badaniach częściowych (niepełnych, niewyczerpujących). Fundamentalną sprawą w badaniu częściowym jest to, aby próba była reprezentatywna, tzn. jej struktura pod względem badanej właściwości (cechy) była zbliżona do struktury tej cechy w populacji generalnej. Zapewnienie wysokiej reprezentatywności próby jest zagadnieniem dość skomplikowanym. Wynika to z faktu, że populacja generalna jest nieznana, a dopiero badanie częściowe ma dostarczyć informacji o niej. Wykorzystanie metod wnioskowania statystycznego (statystyki matematycznej, statystyki indukcyjnej) wymaga znajomości rachunku prawdopodobieństwa. Z tego też względu w rozdziale 1 oraz w podrozdziale 5.1 (dotyczącym dwuwymiarowej zmiennej losowej). zamieszczono podstawowe wiadomości z tej dziedziny wiedzy. Zakres przedmiotowy statystyki przedstawia rysunek 1.

Metody statystyczne dotyczące zarówno opisu, jak i wnioskowania mogą być użyteczne w liczbowym rozpoznaniu struktury, współzależności oraz dynamiki zjawisk masowych. O ile jednak opis statystyczny odpowiada na pytanie: „jak jest?”, o tyle wnioskowanie statystyczne jest związane z pytaniem: „jak prawdopo-



Rysunek 1. Zakres statystyki

Źródło: [Cieciura, Zacharski, 2007, s. 239].

dobnie może być?”. Tak więc, wnioskowanie statystyczne w obydwu zasadniczych procedurach, tj. estymacji i weryfikacji hipotez, odbywa się zawsze w warunkach niepewności. Mówimy, że jest ono wnioskowaniem indukcyjnym o charakterze stochastycznym (opartym na rachunku prawdopodobieństwa). Wzajemne relacje pomiędzy metodami analizy, opisem i wnioskowaniem statystycznym przedstawia tabela 1.

Tabela 1. Podstawowe metody analizy statystycznej

Metody analizy	**Opis statystyczny**	**Wnioskowanie statystyczne**	
Struktury zjawisk	Miary średnie Miary zmienności Miary asymetrii Miary koncentracji i spłaszczenia (kurtozy)	Estymacja parametrów (punktowa i przedziałowa)	Weryfikacja hipotez (parametrycznych i nieparametrycznych)
Współzależności zjawisk	Rachunek korelacji Rachunek regresji		
Dynamiki zjawisk	Szeregi czasowe bez sezonowości Szeregi czasowe z sezonowością Indeksy statystyczne		

Źródło: opracowanie własne na podstawie: [Luszniewicz, Słaby, 2008, s.14].

Całość problematyki statystyki jest ujęta w dwóch odrębnych, wzajemnie uzupełniających się pozycjach noszących tytuły: *Statystyka opisowa* i *Statystyka matematyczna*. Prezentowane w nich zagadnienia są podawane od podstaw i nie są nazbyt skomplikowane pod względem matematycznym. Cenną zaletą tych pozycji jest – zdaniem autora – zamieszczenie wielu rozwiązanych przykładów, które ilustrują istotę omawianych procedur statystycznych oraz pokazują możliwości interpretacyjne otrzymanych wyników. Ponadto zawierają one bogaty zestaw zadań do samodzielnego rozwiązania. Kontrolę stopnia opanowania wiedzy ułatwiają zamieszczone na końcu podręczników odpowiedzi do wszystkich zadań. Celowi temu służą również pytania kontrolne.

W pracach zamieszczono obszerną bibliografię, zawierającą publikacje książkowe z zakresu statystyki. Pozwoli ona na pogłębienie znajomości prezentowanych zagadnień, jak również na poszerzenie wiedzy w zakresie – pominiętych z konieczności – treści.

We współczesnym świecie nabycie umiejętności praktycznego stosowania metod statystycznych jest nieodzowne. Serwisy informacyjne, gazety codzienne, pisma popularne zawierają „gąszcz liczb". Należy umieć je interpretować. W przeciwnym przypadku nie jesteśmy w stanie trafnie ocenić tych doniesień, jak też wyciągać z nich odpowiednie wnioski. Autor daleki jest od popierania stwierdzeń typu: „statystyka jest prosta i oczywista" czy „statystyka jest intuicyjnie zrozumiała". Tak jak każda dziedzina wiedzy wymaga ona znacznego wysiłku i żmudnej pracy w celu zdobycia określonych umiejętności.

Autor żywi nadzieję, że zawarty w pracach zakres przedmiotowy i zrozumiały sposób ujęcia rozpatrywanych zagadnień, uczynią je przyjaznymi dla studentów wydziałów ekonomicznych i humanistycznych (socjologów, psychologów, pedagogów, historyków itp.), jak również dla szerokiego grona praktyków prowadzących analizy statystyczne. Oby przekonanie o nieuchronności związku „nudy" i „statystyki" stało się nieuzasadnione, a wątpliwości w zakresie możliwości zrozumienia świata opisanego za pomocą statystyki – błędne[1].

***

Podręcznik *Statystyka matematyczna* obejmuje problematykę wnioskowania statystycznego (statystyki indukcyjnej). Pod względem formalnym zawiera on pięć rozdziałów poświęconych kolejno: probabilistycznym podstawom statystyki matematycznej, próbie losowej i rozkładom z próby, estymacji i weryfikacji hipotez w analizie struktury, współzależności i dynamiki zjawisk masowych. Integralną częścią podręcznika jest aneks, w którym zamieszczono – niezbędne przy rozwiązywaniu problemów dotyczących badań częściowych – tablice statystyczne.

W tej pracy, podobnie jak w *Statystyce opisowej* zamieszczono dużo rozwiązanych przykładów oraz zadań do samodzielnego „zmagania się" ze statystycznymi problemami. Poprawność otrzymanych wyników można sprawdzić z podanymi odpowiedziami do zadań.

Zarówno sposób prezentacji materiału statystycznego, jak i przyjęta symbolika wskazują na to, że obydwie części: *Statystyka opisowa* i *Statystyka matematyczna*, stanowią integralną całość. W zależności od liczby godzin przeznaczonych w programach studiów różnych kierunków na zgłębianie statystyki, niektóre treści mogą być pominięte.

[1] Jak mawiał jeden ze studentów: „jeśli pozostałby mi tylko jeden dzień życia, spędziłbym go na statystyce – dzięki temu wydałby mi się on znacznie dłuższy", por. [Szwed, 2009, s. 11].

Jakkolwiek w podręcznikach występuje duża liczba wzorów, to nie powinny one wzbudzać niepokoju. Ich zrozumienie i wykorzystanie nie wymaga bowiem pogłębionej znajomości wiedzy z matematyki. Świadomie pominięto wyprowadzanie wzorów, podając je – w gotowej do praktycznych zastosowań – postaci. Pogłębieniu metod statystycznych służy zamieszczona w obydwu podręcznikach bogata literatura przedmiotu.

Autor składa serdeczne podziękowania na ręce Pana Profesora Józefa Pociechy z Uniwersytetu Ekonomicznego w Krakowie za niezwykle trafne i cenne uwagi, które miały istotny wpływ na ostateczny kształt tej pracy.

Dziękuję również córce Ewelinie oraz Monice Sobolewskiej za ich trud włożony w techniczne redagowanie tekstu.

Autor

Rozdział 1. Probabilistyczne podstawy statystyki matematycznej

1.1. Prawdopodobieństwo i jego własności

W różnych dziedzinach ludzkiej działalności mamy do czynienia ze **zjawiskami losowymi** (przypadkowymi), tj. takimi, których wyniku nie można z góry przewidzieć. Tego rodzaju zjawiska są bowiem zależne od przyczyn, których nie jesteśmy w stanie wyeliminować. Przykładowo, nie jesteśmy w stanie przewidzieć liczby wypadków drogowych, wad produkowanych wyrobów, ceny akcji na giełdzie, dokładnego czasu żarzenia się żarówki itp. Badaniem prawidłowości rządzących zjawiskami losowymi zajmuje się **rachunek prawdopodobieństwa**.

W każdym doświadczeniu losowym można wyróżnić najprostsze, nierozkładalne zdarzenia (wyniki doświadczenia). Zdarzenia te nazywamy **elementarnymi**. Zbiór wszystkich zdarzeń elementarnych związanych z danym doświadczeniem losowym nazywamy **przestrzenią zdarzeń elementarnych**. Elementy przestrzeni zdarzeń elementarnych będziemy oznaczać symbolem e_i, natomiast przestrzeń zdarzeń elementarnych – symbolem E. Każdy podzbiór przestrzeni zdarzeń elementarnych nazywamy **zdarzeniem losowym**. Zdarzenia losowe oznaczamy dużymi literami alfabetu A, B, C, ...

Przestrzeń zdarzeń elementarnych może być zbiorem **skończonym** lub **nieskończonym**. Ze skończoną przestrzenią zdarzeń elementarnych mamy do czynienia np. w doświadczeniu polegającym na jednokrotnym rzucie kostką do gry. Zdarzeniami losowymi mogą być tutaj [Oczkoś, 2005, s. 3]:

A – wypadła parzysta liczba oczek,

B – wypadła liczba oczek większa od 3,

C – wypadła liczba oczek będąca liczbą pierwszą.

W tym doświadczeniu zbiór zdarzeń elementarnych jest sześcioelementowy $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, gdzie e_i $(i = 1, 2, \ldots, 6)$ oznaczają zdarzenie elementarne: wypadło i oczek. Zdarzeniem A określamy podzbiór $\{e_2, e_4, e_6\}$. Zachodzi ono wtedy i tylko wtedy, gdy wypada 2 lub 4, lub 6 oczek. Zdarzenie B to podzbiór $\{e_4, e_5, e_6\}$. Zachodzi ono wtedy i tylko wtedy, gdy wypada 4 lub 5, lub 6 oczek. Wreszcie zdarzenie C określamy jako podzbiór $\{e_2, e_3, e_5\}$. Zachodzi ono wtedy i tylko wtedy, gdy wypada 2 lub 3, lub 5 oczek.

Jeżeli jakieś zdarzenie elementarne należy do zdarzenia losowego (zbioru) A, to mówimy, że **sprzyja** ono zdarzeniu A. W przeciwnym przypadku mówimy o zdarzeniu **przeciwnym** (niesprzyjającym). Spośród wszystkich podzbiorów zbioru E, na szczególną uwagę zasługują dwa z nich: podzbiór identyczny z całym zbiorem E i podzbiór Ø niezawierający żadnego elementu. W pierwszym przypadku mówimy o **zdarzeniu pewnym**, a w drugim – o **zdarzeniu niemożliwym**. Inaczej mówiąc, zdarzenie A jest **zdarzeniem pewnym**, jeśli każde zdarzenie elementarne ze zbioru E sprzyja zdarzeniu A. Podobnie, zdarzenie A jest **zdarzeniem niemożliwym**, jeśli żadne zdarzenie elementarne ze zbioru E nie sprzyja zajściu zdarzenia A.

Jeżeli przestrzeń zdarzeń elementarnych jest zbiorem n-elementowym, to liczba jego podzbiorów jest równa 2^n. Ze zbioru $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ można więc utworzyć $2^6 = 64$ podzbiory (zdarzenia), a mianowicie:

1) 2 podzbiory niewłaściwe: Ø, E;
2) 6 podzbiorów jednoelementowych: $\{e_1\}, \ldots, \{e_6\}$;
3) 15 podzbiorów dwuelementowych: $\{e_1, e_2\}, \{e_1, e_3\}, \ldots, \{e_1, e_6\}$;
4) 20 podzbiorów trzyelementowych: $\{e_1, e_2, e_3\}, \ldots, \{e_4, e_5, e_6\}$;
5) 15 podzbiorów czteroelementowych: $\{e_1, e_2, e_3, e_4\}, \ldots, \{e_3, e_4, e_5, e_6\}$;
6) 6 podzbiorów pięcioelementowych: $\{e_1, e_2, e_3, e_4, e_5\}, \ldots, \{e_2, e_3, e_4, e_5, e_6\}$.

Z **nieskończoną** przestrzenią zdarzeń elementarnych mamy np. do czynienia w doświadczeniu polegającym na obserwacji wzrostu przypadkowo wybranej osoby. Przestrzeń tę stanowi pewien przedział liczbowy możliwych wartości wzrostu.

Na zdarzeniach losowych, podobnie jak na zbiorach, możemy wykonywać działania, a mianowicie:

1) Zdarzenia A i B nazywamy **identycznymi** ($A = B$), jeżeli zdarzenie A zachodzi wtedy i tylko wtedy, gdy zachodzi zdarzenie B. Zdarzeniom tym sprzyjają te same zdarzenia elementarne.
2) Zdarzenie A **pociąga** za sobą (**implikuje**) zdarzenie B ($A \subset B$), jeśli zdarzenia elementarne sprzyjające zdarzeniu A sprzyjają również zdarzeniu B (jeśli zaszło zdarzenie A, to zaszło również zdarzenie B).
3) **Sumą (alternatywą) zdarzeń** A i B ($A \cup B$) nazywamy takie zdarzenie C, które zachodzi wtedy i tylko wtedy, gdy zachodzi zdarzenie A lub zdarzenie B.

   Niech A oznacza zdarzenie polegające na tym, że w rzucie kostką wypadła nieparzysta liczba oczek, a B – zdarzenie polegające na otrzymaniu liczby oczek mniejszej od 4. Mamy wówczas:

$$C = A \cup B = \{e_1, e_3, e_5\} \cup \{e_1, e_2, e_3\} = \{e_1, e_2, e_3, e_5\}.$$

4) **Iloczynem (koniunkcją, częścią wspólną)** zdarzeń A i B nazywamy takie zdarzenie C, które zachodzi wtedy i tylko wtedy, gdy zachodzi zdarzenie A i zdarzenie B, tj.

$$C = A \cap B.$$

Niech A oznacza zdarzenie polegające na wylosowaniu z listy mężczyzny, a B – na wylosowaniu osoby palącej papierosy. Wtedy $A \cap B$ oznacza zdarzenie polegające na wylosowaniu mężczyzny palącego papierosy.

Jeżeli iloczyn zdarzeń A i B tworzy zbiór pusty ($A \cap B = \emptyset$), to zdarzenia A i B nazywamy **wykluczającymi się (wyłączającymi się)**. Przykładowo w rzucie kostką do gry zdarzeniami wykluczającymi się są: zdarzenie A polegające na wyrzuceniu liczby oczek mniejszej od 2 oraz zdarzenie B polegające na otrzymaniu liczby oczek większej od 4.

5) **Różnicą zdarzeń** A i B ($A - B$) nazywamy takie zdarzenie C, które zachodzi wtedy i tylko wtedy, gdy zachodzi zdarzenie A i nie zachodzi zdarzenie B.

Niech A oznacza zdarzenie polegające na wyrzuceniu w rzucie kostką parzystej liczby oczek, a B oznacza wyrzucenie liczby oczek większej od 3. Wówczas:

$$C = A - B = \{e_2, e_4, e_6\} - \{e_4, e_5, e_6\} = \{e_2\}.$$

Zdarzeniem przeciwnym do zdarzenia A ($\overline{A}$) nazywamy zdarzenie, do którego należą zdarzenia elementarne nienależące do A, tzn. $\overline{A} = E - A$.

Niech A oznacza zdarzenie polegające na wypadnięciu nieparzystej liczby oczek w rzucie kostką do gry. Zdarzenie $\overline{A}$ jest następującym zbiorem zdarzeń elementarnych:

$$\overline{A} = E - \{e_1, e_3, e_5\} = \{e_2, e_4, e_6\}.$$

Zdarzenie $\overline{A}$ oznacza: wypadła parzysta liczba oczek.

Graficzną ilustrację działań na zdarzeniach są **wykresy Eulera**, zwane też **diagramami Venna** (rys. 1.1).

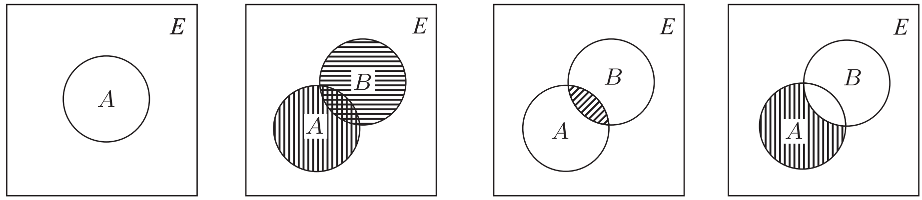
Na rysunku 1.1 przestrzeń zdarzeń elementarnych symbolizuje kwadrat, a zdarzenia A lub B – koła w tym kwadracie. Zakreskowany obszar ilustruje rozważane zdarzenie.

W praktyce jesteśmy zainteresowani nie tyle pojedynczymi zdarzeniami elementarnymi, ale ich zbiorami (czyli podzbiorami E). Rodzinę podzbiorów nazywamy **ciałem zdarzeń losowych** (przestrzenią zdarzeń losowych) i oznaczamy symbolem $\mathcal{S}$. Jeżeli E jest zbiorem przeliczalnym, to $\mathcal{S}$ składa się ze wszystkich podzbiorów E. Jeśli natomiast E jest zbiorem nieprzeliczalnym, to $\mathcal{S}$ jest pewną rodziną podzbiorów E zwaną σ-**ciałem zdarzeń** lub borelowskim ciałem zdarzeń, spełniającą następujące warunki [Grzegorzewski, Bobecka, Dembińska, Pusz, 2001, s. 10]:

- $E \in \mathcal{S}$,
- jeżeli $A \in \mathcal{S}$, to $\overline{A} \in \mathcal{S}$,
- jeżeli $A_1, A_2, \ldots \in \mathcal{S}$, to $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{S}$.

Nazwa „zdarzenie losowe” odnosi się wyłącznie do tych zbiorów zdarzeń elementarnych, które zostały zaliczone do ciała zdarzeń.

Przestrzeń zdarzeń elementarnych wraz z ciałem zdarzeń, tj. parę $(E, \mathcal{S})$, nazywamy **przestrzenią mierzalną**, a zdarzenie $A \in \mathcal{S}$ – S-mierzalnym. Na

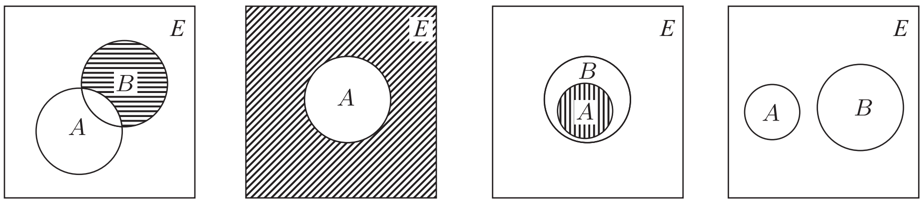


a) zdarzenie A w przestrzeni E

b) suma zdarzeń $A \cup B$

c) iloczyn zdarzeń $A \cap B$

d) różnica zdarzeń $A - B$



e) różnica zdarzeń $B - A$

f) zdarzenie przeciwne do A lub dopełnienie zdarzenia A

g) zdarzenie A zawiera się w zdarzeniu B $(A \subset B)$

h) zdarzenia A i B wykluczają się $(A \cap B = \varnothing)$

Rysunek 1.1. Graficzna ilustracja działań na zdarzeniach losowych

Źródło: [Sobczyk, 2000, s. 86].

przestrzeni mierzalnej można określać różne funkcje (miary). Niech P będzie funkcją rzeczywistą określoną na zdarzeniach S-mierzalnych, spełniającą następujące warunki:

1) dla każdego zdarzenia $A \in \mathcal{S}$, $P(A) \geqslant 0$,
2) $P(E) = 1$,
3) jeżeli $A_1, A_2, \ldots \in \mathcal{S}$ jest dowolnym ciągiem zdarzeń parami rozłącznych, tzn. $A_i \cap A_j = \varnothing$ dla $i \neq j$, to

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

Funkcję P, która spełnia powyższe warunki, określa się mianem miary probabilistycznej lub krótko – **prawdopodobieństwem**. Wartość liczbową $P(A)$ nazywamy **prawdopodobieństwem zdarzenia** A. Trójkę $(E, \mathcal{S}, P)$ nazywamy **przestrzenią probabilistyczną** związaną z danym doświadczeniem losowym.

Z podanych powyżej warunków 1)−3), zwanych aksjomatami Kołmogorowa (znanych od 1933 r.), oraz z własności działań na zdarzeniach, wynikają następujące **własności prawdopodobieństwa** [Major, 2007, s. 23−24]:

1) prawdopodobieństwo każdego zdarzenia A jest liczbą z przedziału [0; 1]

$$0 \leqslant P(A) \leqslant 1, \tag{1.1}$$

2) prawdopodobieństwo zdarzenia niemożliwego jest równe 0

$$P(\emptyset) = 0, \tag{1.2}$$

3) prawdopodobieństwo zdarzenia $\overline{A}$ (przeciwnego do zdarzenia A) jest równe

$$P(\overline{A}) = 1 - P(A), \tag{1.3}$$

4) jeżeli zdarzenie losowe A zawiera się w zdarzeniu losowym B ($A \subset B$), to

$$P(B) \geqslant P(A), \tag{1.4}$$

5) prawdopodobieństwo sumy dowolnych dwóch zdarzeń losowych A i B jest równe sumie prawdopodobieństw tych zdarzeń pomniejszonej o prawdopodobieństwo ich iloczynu

$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \tag{1.5}$$

6) jeżeli $A \cap B = \emptyset$, to

$$P(A \cup B) = P(A) + P(B), \tag{1.6}$$

7) jeżeli A i B są zdarzeniami losowymi należącymi do pewnej przestrzeni E oraz jeśli $P(B) \neq 0$, to prawdopodobieństwo warunkowe (względne) zdarzenia A pod warunkiem B jest określone wzorem

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \ \ P(B) > 0, \tag{1.7}$$

stąd

$$P(A \cap B) = P(B) \cdot P(A/B); \tag{1.8}$$

podobnie

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \ \ P(A) > 0, \tag{1.9}$$

$$P(A \cap B) = P(A) \cdot P(B/A), \tag{1.10}$$

8) jeżeli A i B są zdarzeniami niezależnymi, to

$$P(A/B) = P(A), \tag{1.11}$$

$$P(B/A) = P(B) \tag{1.12}$$

oraz

$$P(A \cap B) = P(A) \cdot P(B). \tag{1.13}$$

Równość (1.13) definiuje niezależność zdarzeń. Dwa zdarzenia A i B nazywamy **niezależnymi**, jeśli zajście jednego z nich nie ma wpływu na prawdopodobieństwo zajścia drugiego.

Przy obliczaniu prawdopodobieństwa sumy zdarzeń, należy najpierw określić czy zdarzenia są wykluczające się (wyłączające się, rozłączne), czy też nie. Dwa zdarzenia A i B nazywamy **wykluczającymi się**, jeśli zajście jednego z nich wyklucza zajście drugiego (w tym samym doświadczeniu). Sposób obliczania prawdopodobieństwa iloczynu zdarzeń jest uwarunkowany tym, czy zdarzenia te są **zależne**, czy **niezależne.**

Jakkolwiek aksjomatyczna definicja prawdopodobieństwa podana w 1933 r. przez A. Kołmogorowa jest poprawna z teoretycznego punktu widzenia, to jest ona niewygodna w praktycznych zastosowaniach. Nie podaje ona bowiem sposobu obliczania prawdopodobieństwa $P(A)$ dla konkretnych zdarzeń losowych. W związku z tym stosuje się inne definicje prawdopodobieństwa, np. **klasyczną definicję Laplace'a** czy **definicję geometryczną**. Klasyczna definicja prawdopodobieństwa, sformułowana w 1812 r. przez P.S. Laplace'a, brzmi następująco: jeżeli przestrzeń zdarzeń elementarnych E składa się z n jednakowo prawdopodobnych zdarzeń, z których k realizuje (sprzyja) zdarzeniu A, to

$$P(A) = \frac{k}{n}. \tag{1.14}$$

Niech A oznacza zdarzenie polegające na otrzymaniu parzystej liczby oczek w rzucie kostką do gry, a zdarzenie B – zdarzenie polegające na otrzymaniu liczby oczek większej od 4. Przestrzeń zdarzeń elementarnych E składa się tu z 6 elementów ($n = 6$): $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$. Zdarzeniu A sprzyjają trzy wyniki ($k = 3$): $\{e_2, e_4, e_6\}$, a zdarzeniu B – 2 wyniki ($k = 2$): $\{e_5, e_6\}$. Korzystając z klasycznej definicji prawdopodobieństwa określonej wzorem (1.14), otrzymujemy

$$P(A) = \frac{3}{6} = 0{,}5$$

oraz

$$P(B) = \frac{2}{6} = 0{,}33(3).$$

Podczas obliczania prawdopodobieństw zdarzeń za pomocą klasycznej definicji prawdopodobieństwa często wykorzystuje się pojęcie **kombinacji**. Z kombinacjami mamy do czynienia wtedy, gdy tworzymy podzbiory z elementów określonego zbioru. W kombinacjach tego typu nie jest ważna kolejność występujących elementów. Stąd też jest używany termin **kombinacje bez powtórzeń**. Liczba k-elementowych kombinacji bez powtórzeń ze zbioru n-elementowego jest określona wzorem:

$$C_k^n = \frac{n!}{k!(n-k)!}, \tag{1.15}$$

gdzie symbol $n!$ (n silnia) jest iloczynem kolejnych liczb naturalnych od 1 do n, tzn. $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n$. Warto zapamiętać, że $0! = 1$.

Rozważmy popularną grę w toto-lotka. Doświadczenie losowe polega tu na wykonaniu losowania 6 spośród 49 liczb. Zdarzeniem losowym jest wynik losowania. Przestrzeń zdarzeń elementarnych E składa się ze wszystkich sześcioelementowych podzbiorów, jakie można utworzyć z 49 liczb. Jest ich tyle, ile jest kombinacji bez powtórzeń z 49 po 6 elementów, czyli

$$\binom{49}{6} = \frac{49!}{6!(49-6)!} = 13\,983\,816.$$

Zakładamy, że wszystkie zdarzenia elementarne są jednakowo możliwe (jednakowo prawdopodobne). Obliczmy prawdopodobieństwa zajścia następujących zdarzeń:

1) A_1 – wylosowanie (trafienie) trójki,
2) A_2 – wylosowanie czwórki,
3) A_3 – wylosowanie piątki,
4) A_4 – wylosowanie szóstki (największa wygrana).

Z sześciu wylosowanych liczb można utworzyć $\binom{6}{3}$ różnych trójek. Do każdej z tych trójek można dołączyć dowolną trójkę liczb spośród pozostałych 43 liczb. Takich trójek jest $\binom{43}{3}$, a więc wszystkich możliwych kombinacji sprzyjających zdarzeniu A_1 jest $\binom{6}{3}\binom{43}{3}$. Prawdopodobieństwo wygrania trójki w toto-lotka jest zatem równe

$$P(A_1) = \frac{\binom{6}{3}\binom{43}{3}}{\binom{49}{6}} = 0{,}01765040.$$

Prawdopodobieństwa dla pozostałych zdarzeń są równe:

$$P(A_2) = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}} = 0{,}0009686,$$

$$P(A_3) = \frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}} = 0{,}00001845,$$

$$P(A_4) = \frac{\binom{6}{6}\binom{43}{0}}{\binom{49}{6}} = 0{,}0000000715.$$

Chociaż klasyczna definicja prawdopodobieństwa jest intuicyjnie zrozumiała, to pod jej adresem formułowany jest zarzut **tautologii** (błąd idem per idem) polegający na użyciu w definicji słowa definiowanego. Dotyczy on słowa: zdarzenie (mówiąc o zdarzeniach jednakowo możliwych, mamy na myśli zdarzenia jednakowo prawdopodobne). Ponadto klasyczna definicja prawdopodobieństwa wymaga istnienia skończonej liczby elementów przestrzeni zdarzeń elementarnych i zbioru zdarzeń sprzyjających, co nie zawsze jest możliwe.