

Rozdział III

Pojęcie i najprostsze własności prawdopodobieństwa

§ 1. Treść rozdziału

W niniejszym rozdziale przechodzimy do właściwego przedmiotu badań. W paragrafie 2 wprowadzamy pojęcie prawdopodobieństwa. Uważamy je za pojęcie pierwotne nie podlegające określeniu, charakteryzujemy je jednak za pomocą 6 aksjomatów. Nie jest to zatem określenie prawdopodobieństwa przez postulaty (analogiczne do określenia ciała Boole'a w rozdziale I), ale aksjomatyka, którą należy rozumieć mniej więcej tak, jak aksjomatykę geometrii lub aksjomatykę liczb naturalnych.

Dalej, podajemy najprostsze konsekwencje aksjomatów (§ 3 i § 4), omawiamy losową niezależność zdarzeń (§ 5) i przechodzimy w paragrafie 6 do pojęć zmiennej losowej i nadziei matematycznej. Pojęcia te wprowadzamy tu w tym tylko bardzo skromnym zakresie, w którym są niezbędne dla studiowania teorii gier losowych.

Krótki paragraf 7, omawiający ciała losowe nasycone, podajemy również ze względu na zastosowanie do gier losowych.

W paragrafie 8 omawiamy prawdopodobieństwo „klasyczne”, aby w ten sposób uwydatnić ścisłą łączność naszego przedstawienia rzeczy ze starą teorią prawdopodobieństwa, ujętą w traktacie Laplace'a.

§ 2. Aksjomatyka prawdopodobieństwa

Wprowadzimy przede wszystkim kilka pojęć pomocniczych.

OKREŚLENIE 1. Niech U oznacza pewne ciało Boole'a. Funkcję f określoną dla $u \in U$, której wartościami są liczby rzeczywiste lub $\pm\infty$, będziemy nazywali *funkcją rzeczywistą określoną w ciele U* . Funkcję rzeczywistą nie przyjmującą wartości $\pm\infty$ nazywamy *skończoną*.

Ponieważ zgodnie z powyższym określeniem funkcja rzeczywista określona w ciele Boole'a może przyjmować wartości niewłaściwe, przypominamy, że dla x rzeczywistego zachodzą wzory:

$$\begin{aligned} x + \infty &= +\infty, & x - \infty &= -\infty, \\ +\infty + \infty &= +\infty, & -\infty - \infty &= -\infty. \end{aligned}$$

Wartości $+\infty - \infty$ i $-\infty + \infty$ nie są określone. Dla $x_1, x_2, \dots$ rzeczywistych lub równych $\pm\infty$ mamy

$$\lim_{n \rightarrow \infty} x_n = +\infty \quad (\lim_{n \rightarrow \infty} x_n = -\infty)$$

wtedy i tylko wtedy, gdy dla każdego naturalnego k istnieje takie N , że dla $n > N$ mamy $x_n > k$ ($x_n < -k$). W szczególności, suma szeregu rozbieżnego liczb nieujemnych jest równa $+\infty$, suma szeregu rozbieżnego liczb niedodatnich jest równa $-\infty$.

Przypominamy również uwagi uczynione w paragrafie 9 (rozdział I, str. 46), dotyczące rozumienia nierówności między liczbami rzeczywistymi a wartościami niewłaściwymi $+\infty$ i $-\infty$.

OKREŚLENIE 2. Funkcja rzeczywista f określona w ciele U , rel I nazywa się *addytywną*, jeżeli dla wszelkich $u, v \in U$, rozłącznych rel I , mamy

$$(2.1) \quad f(u+v) = f(u) + f(v).$$

OKREŚLENIE 3. Funkcja rzeczywista f określona w podciele V , rel IV ciała przeliczalnie addytywnego U , rel I nazywa się *przeliczalnie addytywną*, jeżeli dla każdego ciągu $u_1, u_2, \dots$ elementów parami rozłącznych rel I ciała V , takiego, że $\sum_{i=1}^{\infty} u_i \in V$, mamy

$$(2.2) \quad f\left(\sum_{i=1}^{\infty} u_i\right) = \sum_{i=1}^{\infty} f(u_i).$$

OKREŚLENIE 4. Funkcja rzeczywista f określona w ciele U , rel I nazywa się *niemalejącą*, jeżeli warunek $u \rightarrow v$, rel I pociąga za sobą warunek $f(u) \leq f(v)$.

OKREŚLENIE 5. Funkcja f określona w ciele U nazywa się *rozszerzeniem funkcji g* określonej w podciele V ciała U , jeśli $f(u) = g(u)$ dla $u \in V$.

W związku z tymi określeniami zauważmy, że

(2.3) *Funkcja przeliczalnie addytywna w podciele V ciała U jest addytywna w ciele V .*

Podajemy teraz aksjomatykę prawdopodobieństwa:

(A1) *Prawdopodobieństwo jest funkcją p rzeczywistą i nieujemną, określoną w pewnym podciele zdarzeń zrelatywizowanym względem ideału właściwego.*

Ciała, w których jest określone prawdopodobieństwo, nazywamy losowymi. Przy formułowaniu następnych aksjomatów zakładamy, że p jest prawdopodobieństwem w ciele losowym $U, \text{rel } I$.

(A2) *p jest funkcją przeliczalnie addytywną,*

(A*2) *p jest funkcją addytywną,*

(A3) *$p(u) = 1$, dla $u \equiv 1, \text{rel } I$,*

to znaczy prawdopodobieństwo zdarzenia pewnego $\text{rel } I$ jest równe 1.

(A4) *Jeżeli $u_0 \in U$, $p(u_0) > 0$, to*

$$p_{u_0}(u) = \frac{p(u_0 u)}{p(u_0)}$$

jest prawdopodobieństwem w ciele losowym $U, \text{rel } (I \vee J(u'_0))$.

(A5) *Jeżeli $\{V_n \text{ rel } IV_n\}$ jest ciągiem podciał losowych ciała $U, \text{rel } I$ i jeżeli istnieje w ciele $U, \text{rel } I$ jedna i tylko jedna funkcja spełniająca aksjomaty (A1), (A2) i (A3) i będąca dla każdego u rozszerzeniem prawdopodobieństwa z V_n na U , wówczas jest ona prawdopodobieństwem w ciele $U, \text{rel } I$.*

(A6) *Jeżeli p jest prawdopodobieństwem w ciele losowym $U, \text{rel } I$, to jest ona również prawdopodobieństwem w każdym podciele $V, \text{rel } IV$ ciała $U, \text{rel } I$.*

W związku z podaną tu aksjomatyką nasuwają się następujące uwagi:

Aksjomaty (A1), (A2), (A3) i (A4) stanowią układ w zasadzie pokrywający się z najstarszym układem pewników teorii prawdopodobieństwa, podanym przez Bohlmana).

Aksjomat (A2) jest wzmocnieniem aksjomatu (A*2). Przeliczalna addytywność prawdopodobieństwa wydaje się nam nieodzowną do zbudowania matematycznej teorii prawdopodobieństwa, dlatego przyjmujemy, za przykładem Kołmogorowa, (A2) jako aksjomat właściwy. Formułujemy jednak osobno aksjomat (A*2) ze względów historycznych, a także, ponieważ wystarcza on na ogół

w elementarnej teorii prawdopodobieństwa. Nadmieniamy, że niektórzy matematycy podawali w wątpliwość aksjomat (A2) (np. Bruno de Finetti).

OKREŚLENIE 6. Występujące w (A4) prawdopodobieństwo $p_{u_0}(u)$ nazywamy *względny* lub też dokładniej: *prawdopodobieństwem zdarzenia u przy założeniu, że zaszło zdarzenie u_0 .*

§ 3. Najprostsze twierdzenia o prawdopodobieństwie

Zakładamy, że p jest prawdopodobieństwem w ciele losowym $U, \text{rel } I$.

TWIERDZENIE 1. *Jeżeli $u \equiv v, \text{rel } I$, to $p(u) = p(v)$, to znaczy zdarzenia równoważne mają równe prawdopodobieństwa.*

Z założenia, że $u \equiv v, \text{rel } I$, wynika, że $u' \equiv v', \text{rel } I$, a więc $u + u' \equiv v + u' \equiv 1, \text{rel } I$; $uu' \equiv vv' \equiv 0, \text{rel } I$; zatem na mocy (A*2) i (A3) $p(u + u') = p(u) + p(u') = 1 = p(v + u') = p(v) + p(u')$, a więc $p(u) = p(v)$, co należało okazać.

Zauważmy, że przy okazji wyprowadziliśmy często stosowany wzór

$$(3.1) \quad p(u) + p(u') = 1.$$

W dalszym ciągu, mówiąc o funkcjach rzeczywistych f określonych w zrelatywizowanych ciałach Boole'a $U, \text{rel } I$, będziemy zawsze milcząco zakładali, że spełniają one „warunek równoważności”:

(R) *Jeżeli $u \equiv v, \text{rel } I$, to $f(u) = f(v)$.*

Twierdzenie 1 mówi nam właśnie, że prawdopodobieństwo spełnia warunek równoważności w każdym ciele losowym.

Udowodnimy teraz twierdzenia 2*5* dotyczące dowolnych funkcji g , addytywnych lub addytywnych nieujemnych, określonych w dowolnym ciele Boole'a $U, \text{rel } I$. Za pomocą aksjomatów (A1) i (A*2) otrzymujemy z tych twierdzeń jako ich szczególne przypadki twierdzenia 2-5 dotyczące prawdopodobieństwa.

TWIERDZENIE 2*. *Jeżeli g jest funkcją addytywną w ciele $U, \text{rel } I$, to $g(0) = 0$.*

Dowód. Mamy dla $u \in U$: $g(u) = g(u + 0) = g(u) + g(0)$, a więc $g(0) = 0$.

TWIERDZENIE 3*. *Funkcja g addytywna i nieujemna w ciele $U, \text{rel } I$ jest niemalejąca, to znaczy jeżeli $u \rightarrow v, \text{rel } I$, to $g(u) \leq g(v)$.*

Dowód. Jeżeli $u \rightarrow v, \text{rel } I$, to $v = u + u'v, \text{rel } I$, a więc na mocy warunku (R), addytywności i nieujemności funkcji g mamy $g(v) = g(u + u'v) = g(u) + g(u'v) \geq g(u)$.

Twierdzenie 4*. Jeżeli g jest funkcją addytywną w ciele $U, \text{rel } I$, to $g(u) + g(v) = g(u + v) - g(uv)$ dla wszystkich $u, v \in U$.

Dowód. Opierając się na (R) i wzorach algebry Boole'a mamy $g(u + v) = g(u + u'v) = g(u) + g(u'v)$ oraz $g(v) = g(uv + u'v) = g(uv) + g(u'v)$, a stąd eliminując $g(u'v)$ mamy $g(u + v) = g(u) + g(v) - g(uv)$, co należało okazać.

Twierdzenie 5*. Jeżeli g jest funkcją addytywną w ciele $U, \text{rel } I$, to dla parami rozłącznych $\text{rel } I$ elementów $u_1, u_2, \dots, u_n \in U$ mamy $g(\sum_{i=1}^n u_i) = \sum_{i=1}^n g(u_i)$. Jeżeli g jest ponadto nieujemna, to dla wszelkich $u_1, u_2, \dots, u_n \in U$ zachodzi wzór $g(\sum_{i=1}^n u_i) \leq \sum_{i=1}^n g(u_i)$.

Pierwszą tezę otrzymujemy przez łatwą indukcję, którą pozostawiamy czytelnikowi. Druga teza dla $n=2$ jest konsekwencją twierdzenia 4* i nierówności $g(uv) \geq 0$, dla dowolnych zaś u otrzymuje się ją znów przez indukcję.

Przechodząc teraz do prawdopodobieństwa, otrzymujemy jako konsekwencję twierdzeń 2*-5*

Twierdzenie 2. Jeżeli $u \equiv 0, \text{rel } I$, to $p(u) = 0$.

Twierdzenie 3. Jeżeli $u \rightarrow v, \text{rel } I$, to $p(u) \leq p(v)$, to znaczy prawdopodobieństwo jest funkcją niemalejącą.

Ponieważ $u \rightarrow 1, \text{rel } I$ dla każdego u , więc wynika stąd

$$(3.2) \quad p(u) \leq 1 \text{ dla każdego } u \in U,$$

a więc prawdopodobieństwo jest funkcją skończoną, a nawet ograniczoną.

Aksjomat (A*2), który możemy napisać w postaci

$$(3.3) \quad \text{Jeżeli } uv \equiv 0, \text{rel } I, \text{ to } p(u + v) = p(u) + p(v),$$

w klasycznej teorii prawdopodobieństwa nosi nazwę *twierdzenia o prawdopodobieństwie całkowitym*. (3.3) jest szczególnym przypadkiem twierdzenia następującego:

Twierdzenie 4. Dla wszelkich $u, v \in U$ mamy

$$(3.4) \quad p(u + v) = p(u) + p(v) - p(uv).$$

Wzory (3.3) oraz (3.4) można uogólnić na przypadek n składników.

Twierdzenie 5 (Ogólne twierdzenie o prawdopodobieństwie całkowitym). Jeżeli zdarzenia $u_1, \dots, u_n \in U$ wykluczają się parami (czyli $u_i \cdot u_j \equiv 0, \text{rel } I$ dla $i \neq j, i, j = 1, 2, \dots, n$), to

$$(3.5) \quad p\left(\sum_{i=1}^n u_i\right) = \sum_{i=1}^n p(u_i).$$

Uogólnienie wzoru (3.4) stanowi tak zwany wzór Poincarégo; aby wzór ten napisać w dogodnej formie, oznaczmy jak zwykle przez $(x_1, x_2, \dots, x_n)$ zbiór złożony z elementów $x_1, x_2, \dots, x_n$, przez $C_n^{(k)}$ zaś oznaczmy klasę kombinacji z n liczb $1, 2, \dots, n$ po k , innymi słowy klasę podzbiorów zbioru $(1, 2, \dots, n)$ mających dokładnie k elementów, gdzie $k \leq n$. Zachodzi niemal oczywiste twierdzenie:

$$(3.6) \quad \text{Klasa } C_{n+1}^{(k)} \text{ jest sumą klasy } C_n^{(k)} \text{ oraz klasy zbiorów } (j_1, j_2, \dots, j_{k-1}, j_k, \dots, j_{k-1}, n+1), \text{ gdzie } (j_1, j_2, \dots, j_{k-1}) \in C_n^{(k-1)} \text{ } (k \leq n).$$

Twierdzenie 6 (Wzór Poincarégo).

$$(3.7) \quad p(u_1 + u_2 + \dots + u_n) = \sum_{k=1}^n (-1)^{k-1} \sum_{(j_1, j_2, \dots, j_k) \in C_n^{(k)}} p(u_{j_1} u_{j_2} \dots u_{j_k}).$$

Dowód jest indukcyjny. Dla $n=2$ wzór (3.7) przechodzi we wzór (3.4). Załóżmy, że (3.7) zachodzi dla pewnego n ; stosując (3.4) do zdarzeń $u_1, u_2, \dots, u_n$ oraz u_{n+1} i uwzględniając (3.6) otrzymujemy

$$\begin{aligned} (3.8) \quad & p(u_1 + u_2 + \dots + u_n + u_{n+1}) = \\ & = p(u_1 + u_2 + \dots + u_n) + p(u_{n+1}) - p(u_1 u_{n+1} + u_2 u_{n+1} + \dots + u_n u_{n+1}) = \\ & = \sum_{j=1}^{n+1} p(u_j) + \sum_{k=2}^n (-1)^{k-1} \sum_{(j_1, \dots, j_k) \in C_n^{(k)}} p(u_{j_1} u_{j_2} \dots u_{j_k}) - \\ & \quad - \sum_{i=1}^n (-1)^{i-1} \sum_{(j_1, \dots, j_i) \in C_n^{(i)}} p(u_{j_1} u_{j_2} \dots u_{j_i} u_{n+1}) = \\ & = \sum_{j=1}^{n+1} p(u_j) + \sum_{k=2}^n (-1)^{k-1} \left\{ \sum_{(j_1, \dots, j_k) \in C_n^{(k)}} p(u_{j_1} u_{j_2} \dots u_{j_k}) + \right. \\ & \quad \left. + \sum_{(j_1, \dots, j_{k-1}) \in C_n^{(k-1)}} p(u_{j_1} u_{j_2} \dots u_{j_{k-1}} u_{n+1}) \right\} + (-1)^n p(u_1 u_2 \dots u_{n+1}) = \\ & = \sum_{k=1}^{n+1} (-1)^{k-1} \sum_{(j_1, \dots, j_k) \in C_{n+1}^{(k)}} p(u_{j_1} u_{j_2} \dots u_{j_k}), \end{aligned}$$

to znaczy wzór Poincarégo dla $n+1$.

Wzór Poincarégo używany jest rzadko, zazwyczaj dla małych wartości n (np. $n=3,4$). Dla $n=3$ mamy:

$$\begin{aligned} p(u_1 + u_2 + u_3) &= \\ &= p(u_1) + p(u_2) + p(u_3) - p(u_1 u_2) - p(u_2 u_3) - p(u_3 u_1) + p(u_1 u_2 u_3). \end{aligned}$$

Częściej używamy nierówności (konsekwencja twierdzenia 5*)

$$(3.9) \quad p\left(\sum_{i=1}^n u_i\right) \leq \sum_{i=1}^n p(u_i).$$

Podamy teraz dwa twierdzenia o funkcjach nieujemnych i przeliczalnie addytywnych w dowolnym ciele U , $\text{rel } I$, poprzedzone twierdzeniem pomocniczym.

Twierdzenie 7. *Jeżeli $u_1, u_2, \dots$ jest ciągiem elementów podciała V ciała przeliczalnie addytywnego U , to istnieje ciąg $w_1, w_2, \dots$ elementów parami rozłącznych ciała U taki, że*

$$(3.10) \quad \sum_{i=1}^{\infty} u_i = \sum_{j=1}^{\infty} w_j, \quad \sum_{i=1}^n u_i = \sum_{j=1}^n w_j \quad \text{dla } n=1, 2, \dots$$

Dowód. Możemy przyjąć

$$(3.11) \quad w_j = u_j \left(\sum_{i=1}^{j-1} u_i \right)' = u_j \prod_{i=1}^{j-1} u_i'.$$

Oczywiście, elementy $w_1, w_2, \dots$ są parami rozłączne, a ponadto $w_1 = u_1$, co daje drugą równość (3.10) w przypadku $n=1$. Przypuśćmy, że dla pewnego n mamy $\sum_{i=1}^n u_i = \sum_{j=1}^n w_j$ wtedy

$$\begin{aligned} \sum_{i=1}^{n+1} u_i &= u_{n+1} + \sum_{i=1}^n u_i = u_{n+1} \left(\sum_{i=1}^n u_i \right) + u_{n+1} \left(\sum_{i=1}^n u_i \right)' + \sum_{i=1}^n u_i = \\ &= u_{n+1} \left(\sum_{i=1}^n u_i \right)' + \sum_{i=1}^n u_i = w_{n+1} + \sum_{j=1}^n w_j = \sum_{j=1}^{n+1} w_j, \end{aligned}$$

co kończy dowód indukcyjny drugiej równości (3.10). Ponieważ

$$\sum_{i=1}^{\infty} u_i = \sum_{n=1}^{\infty} \sum_{i=1}^n u_i = \sum_{n=1}^{\infty} \sum_{j=1}^n w_j = \sum_{j=1}^{\infty} w_j,$$

więc zachodzi i pierwsza równość (3.10).

Twierdzenie 8*. *Jeżeli g jest nieujemną i przeliczalnie addytywną funkcją w ciele U , $\text{rel } I$ i jeżeli $u_1, u_2, \dots \in U$, $\sum_{i=1}^{\infty} u_i \in U$, to*

$$g\left(\sum_{i=1}^{\infty} u_i\right) = \lim_{n \rightarrow \infty} g\left(\sum_{i=1}^n u_i\right) \leq \sum_{i=1}^{\infty} g(u_i).$$

Dowód. Zachowując oznaczenia (3.11) mamy

$$\begin{aligned} g\left(\sum_{i=1}^{\infty} u_i\right) &= g\left(\sum_{j=1}^{\infty} w_j\right) = \sum_{j=1}^{\infty} g(w_j) = \lim_{n \rightarrow \infty} \sum_{j=1}^n g(w_j) = \\ &= \lim_{n \rightarrow \infty} g\left(\sum_{j=1}^n w_j\right) = \lim_{n \rightarrow \infty} g\left(\sum_{i=1}^n u_i\right). \end{aligned}$$

Ponieważ zaś g jest funkcją nieujemną, więc na mocy twierdzenia 5* mamy $g\left(\sum_{i=1}^n u_i\right) \leq \sum_{i=1}^n g(u_i) \leq \sum_{i=1}^{\infty} g(u_i)$ i wobec tego $\lim_{n \rightarrow \infty} g\left(\sum_{i=1}^n u_i\right) \leq \sum_{i=1}^{\infty} g(u_i)$, co należało okazać.

Twierdzenie 9*. *Jeżeli g jest nieujemną i przeliczalnie addytywną funkcją w ciele U , $\text{rel } I$, a ponadto $u_1, u_2, \dots \in U$, $\prod_{i=1}^{\infty} u_i \in U$ i dla pewnego n_0 jest $g\left(\prod_{i=1}^{n_0} u_i\right) < +\infty$, to*

$$g\left(\prod_{i=1}^{\infty} u_i\right) = \lim_{n \rightarrow \infty} g\left(\prod_{i=1}^n u_i\right).$$

Dowód. Przyjmijmy $u_0 = \prod_{i=1}^{n_0} u_i$. Mamy

$$\begin{aligned} g\left(\prod_{i=1}^{\infty} u_i\right) &= g(u_0) - g\left(u_0 \left(\prod_{i=1}^{\infty} u_i\right)'\right) = g(u_0) - g\left(u_0 \sum_{i=1}^{\infty} u_i'\right) = \\ &= g(u_0) - g\left(\sum_{i=1}^{\infty} u_0 u_i'\right) = g(u_0) - \lim_{n \rightarrow \infty} g\left(\sum_{i=1}^n u_0 u_i'\right) = \\ &= g(u_0) - \lim_{n \rightarrow \infty} g\left(u_0 \sum_{i=1}^n u_i'\right) = \lim_{n \rightarrow \infty} [g(u_0) - g\left(u_0 \sum_{i=1}^n u_i'\right)] = \\ &= \lim_{n \rightarrow \infty} g\left(u_0 - \sum_{i=1}^n u_i'\right) = \lim_{n \rightarrow \infty} g\left(u_0 \prod_{i=1}^n u_i\right) = \lim_{n \rightarrow \infty} g\left(\prod_{i=1}^n u_i\right). \end{aligned}$$

W przypadku, gdy funkcja g jest skończona (a więc np. w przypadku, gdy g jest prawdopodobieństwem), warunek $g(\prod_{i=1}^{n_0} u_i) < +\infty$ jest oczywiście spełniony już przy $n_0=1$.

Podajemy jeszcze pewien wniosek z twierdzeń 8* i 9*.

(3.12) *Jeżeli g jest przeliczalnie addytywną i nieujemną funkcją w ciele U , $u_i^{(j)} \in U$, $i, j=1, 2, \dots$, $\sum_{j=1}^{\infty} \prod_{i=1}^{\infty} u_i^{(j)} \in U$, $\prod_{i=1}^{\infty} u_i^{(j)} \in U$ i dla każdego m istnieje takie k , że $g(\sum_{j=1}^m \prod_{i=1}^k u_i^{(j)}) < +\infty$, to*

$$g(\sum_{j=1}^{\infty} \prod_{i=1}^{\infty} u_i^{(j)}) = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} g(\sum_{j=1}^m \prod_{i=1}^n u_i^{(j)}).$$

Jeżeli zaś

$$\prod_{j=1}^{\infty} \sum_{i=1}^{\infty} u_i^{(j)} \in U, \quad \sum_{i=1}^{\infty} u_i^{(j)} \in U$$

i dla pewnego n_0 jest $g(\prod_{j=1}^{n_0} \sum_{i=1}^{\infty} u_i^{(j)}) < \infty$, to

$$g(\prod_{j=1}^{\infty} \sum_{i=1}^{\infty} u_i^{(j)}) = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} g(\prod_{j=1}^m \sum_{i=1}^n u_i^{(j)}).$$

Udowodnimy pierwszą tezę (3.12), dowód drugiej, który jest podobny do dowodu pierwszej, pozostawiamy czytelnikowi.

Mamy oczywiście

$$\sum_{j=1}^m \prod_{i=1}^n u_i^{(j)} = \prod_{k=1}^n \sum_{j=1}^m \prod_{i=1}^k u_i^{(j)} \quad \text{oraz} \quad \sum_{j=1}^m \prod_{i=1}^{\infty} u_i^{(j)} = \prod_{k=1}^{\infty} \sum_{j=1}^m \prod_{i=1}^k u_i^{(j)}.$$

Korzystając z tych wzorów oraz z twierdzeń 8* i 9* otrzymujemy

$$\begin{aligned} g(\sum_{j=1}^{\infty} \prod_{i=1}^{\infty} u_i^{(j)}) &= \lim_{m \rightarrow \infty} g(\sum_{j=1}^m \prod_{i=1}^{\infty} u_i^{(j)}) = \lim_{m \rightarrow \infty} g(\prod_{k=1}^{\infty} \sum_{j=1}^m \prod_{i=1}^k u_i^{(j)}) = \\ &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} g(\prod_{k=1}^n \sum_{j=1}^m \prod_{i=1}^k u_i^{(j)}) = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} g(\sum_{j=1}^m \prod_{i=1}^n u_i^{(j)}). \end{aligned}$$

Przechodząc na mocy (A1) i (A2) do prawdopodobieństwa otrzymujemy z twierdzeń 8* i 9*:

Twierdzenie 8. *Jeżeli $u_1, u_2, \dots \in U$ i $\sum_{i=1}^{\infty} u_i \in U$, to*

$$(3.13) \quad p(\sum_{i=1}^{\infty} u_i) = \lim_{n \rightarrow \infty} p(\sum_{i=1}^n u_i).$$

Twierdzenie 9. *Jeżeli $u_1, u_2, \dots \in U$ i $\prod_{i=1}^{\infty} u_i \in U$, to*

$$(3.14) \quad p(\prod_{i=1}^{\infty} u_i) = \lim_{n \rightarrow \infty} p(\prod_{i=1}^n u_i).$$

Zauważmy wreszcie, że na mocy twierdzenia 8*:

$$(3.15) \quad p(\sum_{i=1}^{\infty} u_i) \leq \sum_{i=1}^{\infty} p(u_i).$$

Zastosujemy teraz udowodnione twierdzenia do ciał losowych atomowych, mających co najwyżej przeliczalny zbiór atomów reprezentacyjnych.

(3.16) *Jeżeli ciało losowe $U, \text{rel } I$ jest atomowe i ma przeliczalny zbiór atomów reprezentacyjnych, to $p(v)$ jest sumą prawdopodobieństw atomów reprezentacyjnych zawartych w v dla każdego $v \in U$.*

Oznaczamy przez $At^*(v)$ klasę atomów reprezentacyjnych zawartych w v i ustawiamy elementy tej klasy w ciąg $u_1, u_2, \dots, u_n, \dots$ skończony lub nieskończony. Jest wówczas (por. rozdział I, (12.7), str. 64) $v = \sum_{j=1}^{\infty} u_j, \text{rel } I$, a zarazem $u_i u_j = 0, \text{rel } I, i \neq j$.

Wobec tego na mocy twierdzeń 1, 5 i 8

$$(3.17) \quad p(v) = \sum_{j=1}^{\infty} p(u_j) = \sum_{u \in At^*(v)} p(u).$$

Wynika stąd w szczególności dla $v=1, \text{rel } I$ wzór

$$(3.18) \quad \sum_{u \in At(U, I)} p(u) = 1.$$

Jest to wzór, z którego często będziemy korzystali. Opierając się na (3.16) możemy wypowiedzieć twierdzenie:

(3.19) *Jeżeli $U, \text{rel } I$ jest ciałem atomowym co najwyżej przeliczalnym zbioru atomów, to wartości, jakie przybiera prawdopodobieństwo $p(u)$ w zbiorze $At(U, I)$, wyznaczają wartości funkcji $p(u)$ w całym ciele U .*

§ 4. Twierdzenia o prawdopodobieństwie złożonym

Z aksjomatu (A4) wynika, że przy założeniu $p(u) > 0$ mamy

$$(4.1) \quad p(uv) = p(u)p_u(v).$$

W tej postaci (A4) występuje w klasycznej teorii prawdopodobieństwa pod nazwą *twierdzenia o prawdopodobieństwie złożonym*.

LEMAT 1. Jeżeli $p(u) > 0$, to ideał $I \vee J(u')$ jest ideałem właściwym.

Z założenia mamy $p(u) > 0$, więc na podstawie twierdzenia 2 (§ 3) $u \not\equiv 0, \text{rel} I$, a zatem u nie należy do I , z czego wynika, że $I \vee J(u')$ jest ideałem właściwym.

Udowodnimy teraz twierdzenie nieco innej natury, które stanowi wyjaśnienie i w pewnej mierze usprawiedliwienie aksjomatu (A4).

TWIERDZENIE 10. Jeżeli funkcja g jest w ciele $U, \text{rel} I$ nieujemna, przeliczalnie addytywna i równa 1 dla elementów $u \equiv 1, \text{rel} I$, to dla każdego $u_0 \in U$ takiego, że $p(u_0) > 0$, funkcja $g_1(u) = \frac{g(u_0 u)}{g(u_0)}$ jest w ciele $U, \text{rel}(I \vee J(u'))$ również nieujemna, przeliczalnie addytywna i równa 1 dla elementów $u \equiv 1, \text{rel}(I \vee J(u'))$.

Nieujemność funkcji g_1 jest oczywista. Zauważmy dalej, że zgodnie z umową g spełnia warunek równoważności $\text{rel} I$, z czego wynika, że g_1 też spełnia ten warunek $\text{rel}(I \vee J(u'))$, z $u \equiv v, \text{rel}(I \vee J(u'))$ wynika bowiem $uu_0 \equiv vv_0, \text{rel} I$, więc $g(uu_0) = g(vu_0)$, stąd zaś otrzymujemy natychmiast, że $g_1(u) = g_1(v)$. Z udowodnionego spełniania przez g_1 warunku równoważności w ciele $U, \text{rel}(I \vee J(u'))$ wynika, że dla elementów $u \equiv 1, \text{rel}(I \vee J(u'))$ przyjmuje ona wartość 1.

Pozostaje do wykazania przeliczalna addytywność funkcji g_1 .

Niech $u_1, u_2, \dots \in U$, $\sum_{i=1}^{\infty} u_i \in U$, $u_i u_j \equiv 0, \text{rel}(I \vee J(u'))$ dla $i \neq j$, $i, j = 1, 2, \dots$; wtedy $u_i u_j u_0 \equiv 0, \text{rel} I$, więc z przeliczalnej addytywności funkcji g wynika, że

$$\sum_{i=1}^{\infty} g_1(u_i) = \sum_{i=1}^{\infty} \frac{g(u_0 u_i)}{g(u_0)} = \frac{\sum_{i=1}^{\infty} g(u_0 u_i)}{g(u_0)} = \frac{g(u_0 \sum_{i=1}^{\infty} u_i)}{g(u_0)} = g_1\left(\sum_{i=1}^{\infty} u_i\right),$$

co dowodzi naszego twierdzenia.

Twierdzenie 10 usprawiedliwia aksjomat (A4) jak też wyjaśnia fakt, że w warunkach aksjomatu (A5) nie ma wzmianki o aksjomacie (A4).

Funkcję $g_1(u) = \frac{g(uu_0)}{g(u_0)}$ rozpatrywaną w ciele $U, \text{rel}(I \vee J(u'))$ będziemy nazywali *relatywizacją funkcji g za pomocą elementu u_0* .

LEMAT 2. Jeżeli g_1 jest relatywizacją funkcji g za pomocą elementu u_1 , a g_2 jest relatywizacją funkcji g_1 za pomocą elementu u_2 , to g_2 jest relatywizacją funkcji g za pomocą elementu $u_1 u_2$.

Na podstawie założeń mamy

$$g_2(u) = \frac{g_1(u_2 u)}{g_1(u_2)} = \frac{\frac{g(u_1 u_2 u)}{g(u_1)}}{\frac{g(u_1 u_2)}{g(u_1)}} = \frac{g(u_1 u_2 u)}{g(u_1 u_2)}.$$

Z drugiej zaś strony $I \vee J(u'_1) \vee J(u'_2) = I \vee J(u'_1 + u'_2) = I \vee J((u_1 u_2)'),$ co dowodzi, że funkcja g_2 , jako relatywizacja funkcji g_1 względem elementu u_2 , jest rozpatrywana w tym samym ciele jako relatywizacja funkcji g względem elementu $u_1 u_2$. Przechodząc do prawdopodobieństwa uzyskujemy na podstawie lematu 2 następujący wynik (przy założeniu, że $p(u_1 u_2) > 0$):

$$(4.2) \quad p_{u_1 u_2}(v) = p_{(u_1 u_2)}(v).$$

TWIERDZENIE 11 (Ogólne twierdzenie o prawdopodobieństwie złożonym). Jeżeli $p(u_1 u_2 \dots u_n) > 0$, to

$$p(u_1 u_2 \dots u_n u_{n+1}) = p(u_1) p_{u_1}(u_2) \dots p_{u_1 u_2 \dots u_n}(u_{n+1}).$$

Dowód indukcyjny nie nastęrcza żadnych trudności, gdyż z założenia wynika, że $p(u_1) > 0$, $p(u_1 u_2) > 0, \dots, p(u_1 u_2 \dots u_n) > 0$.

Podamy teraz dwa twierdzenia o innym nieco charakterze.

(4.3) Będziemy mówili, że zdarzenia $v_1, \dots, v_n$ są *jedynie możliwe i wyłączające się rel I względem zdarzenia u* , jeżeli

$$u \equiv \sum_{i=1}^n u v_i, \text{rel} I, \quad u v_i v_j \equiv 0, \text{rel} I \quad \text{dla} \quad i \neq j, \quad i, j = 1, 2, \dots, n.$$

TWIERDZENIE 12a (Wzór Laplace'a-Bayesa). Jeżeli zdarzenia $v_1, \dots, v_n$ są *jedynie możliwe i wyłączające się rel I względem u* i jeżeli

$p(u) > 0, p(v_i) > 0, i = 1, 2, \dots, n$, to

$$(4.4) \quad p_v(v_i) = \frac{p(v_i)p_{v_i}(u)}{\sum_{j=1}^n p(v_j)p_{v_j}(u)}.$$

Dowód. Opierając się na określeniu (4.3) i stosując twierdzenia o prawdopodobieństwie całkowitym i złożonym mamy

$$(4.5) \quad p(u) = \sum_{j=1}^n p(uv_j) = \sum_{j=1}^n p(v_j)p_{v_j}(u),$$

$$(4.6) \quad p_u(v_i) = \frac{p(uv_i)}{p(u)} = \frac{p(v_i)p_{v_i}(u)}{\sum_{j=1}^n p(v_j)p_{v_j}(u)}.$$

Twierdzenie 12b. Jeżeli zdarzenia $v_1, \dots, v_n$ są jedynie możliwe i wyłączające się relI względem u i jeżeli $p(uv_i) > 0, i = 1, 2, \dots, n$, to

$$(4.7) \quad p_u(w) = \frac{\sum_{i=1}^n p(v_i)p_{v_i}(u)p_{uv_i}(w)}{\sum_{i=1}^n p(v_i)p_{v_i}(u)}.$$

Dowód. Stosując kilkakrotnie twierdzenia o prawdopodobieństwie całkowitym i złożonym otrzymujemy

$$(4.8) \quad p(uw) = \sum_{i=1}^n p(uv_i w) = \sum_{i=1}^n p(uv_i)p_{uv_i}(w) = \sum_{i=1}^n p(v_i)p_{v_i}(u)p_{uv_i}(w),$$

skąd, stosując (4.5), otrzymujemy, że

$$(4.9) \quad p_u(w) = \frac{p(uw)}{p(u)} = \frac{\sum_{i=1}^n p(v_i)p_{v_i}(u)p_{uv_i}(w)}{\sum_{i=1}^n p(v_i)p_{v_i}(u)},$$

co należało okazać.

§ 5. Niezależność losowa zdarzeń

Określenie 7. Zdarzenia u i v ciała losowego nazywamy losowo niezależnymi, jeżeli spełniają warunek

$$(5.1) \quad p(uv) = p(u)p(v).$$

Twierdzenie 13. Jeżeli zdarzenia u i v są losowo niezależne, to zdarzenia u' i v ; u i v' ; u' i v' są również losowo niezależne.

Należy więc udowodnić, że warunek (5.1) pociąga za sobą warunki

$$(5.2) \quad p(u'v) = p(u')p(v), \quad p(uv') = p(u)p(v'), \quad p(u'v') = p(u')p(v').$$

Dowód. Na mocy (5.1) mamy $p(u'v) = p(v) - p(uv) = p(v) - p(u)p(v) = p(v)[1 - p(u)] = p(u')p(v)$, to znaczy pierwszy związek (5.2). Udowodniliśmy więc, że jeżeli zdarzenia u, v są losowo niezależne, to zdarzenia u', v też są losowo niezależne. Stosując to prawo do zdarzeń v, u , a następnie do zdarzeń v, u' , otrzymujemy dalsze związki (5.2).

Twierdzenie 14. Jeżeli $p(u) > 0$, to warunkiem koniecznym i dostatecznym losowej niezależności zdarzeń u, v jest, aby spełniony był wzór

$$(5.3) \quad p_u(v) = p(v).$$

Dowód. Istotnie, jeżeli $p(u) > 0$ i zachodzi (5.1), to $p_u(v) = \frac{p(uv)}{p(u)} = \frac{p(u)p(v)}{p(u)} = p(v)$. Jeżeli natomiast mamy (5.3), to mnożąc obie strony przez $p(u)$ otrzymujemy (5.1).

Określenie 8. Zdarzenia $u_1, \dots, u_n$ nazywamy losowo niezależnymi, jeżeli zachodzi wzór

$$(5.4) \quad p(u_1 u_2 \dots u_n) = p(u_1)p(u_2) \dots p(u_n)$$

oraz $2^n - 1$ związków powstających z (5.4) przez zastąpienie niektórych $u_1, \dots, u_n$ ich negacjami $u'_1, \dots, u'_n$; związki te nazwiemy pochodnymi związku (5.4).

Określenie 9. Zdarzenia $u_1, \dots, u_n$ są parami niezależne, jeżeli każde dwa spośród nich są losowo niezależne.

Twierdzenie 15. Jeżeli zdarzenia $u_1, \dots, u_n$ są losowo niezależne, to zdarzenia $u_1, \dots, u_{n-1}$ są również losowo niezależne.

Dowód. Istotnie, na mocy związku (5.4) oraz pochodnego związku $p(u_1 \dots u_{n-1} u'_n) = p(u_1)p(u_2) \dots p(u_{n-1})p(u'_n)$ otrzymujemy $p(u_1 u_2 \dots u_{n-1}) = p(u_1 \dots u_{n-1}(u_n + u'_n)) = p(u_1 \dots u_{n-1} u_n + u_1 \dots u_{n-1} u'_n) = p(u_1 \dots u_{n-1} u_n) + p(u_1 \dots u_{n-1} u'_n) = p(u_1) \dots p(u_{n-1})p(u_n) + p(u_1) \dots p(u_{n-1})p(u'_n) = p(u_1) \dots p(u_{n-1})(p(u_n) + p(u'_n)) = p(u_1) \dots p(u_{n-1})$.

W ten sposób można udowodnić, że zachodzą pochodne związku $p(u_1 \dots u_{n-2}) = p(u_1) \dots p(u_{n-2})$ itd. Stąd wynika natychmiast twierdzenie

(5.5) *Jeżeli $u_1, \dots, u_n$ są losowo niezależne, to są niezależne parami.*

Teza powyższa nie daje się odwrócić. Mianowicie, jeżeli $n \geq 3$, to z tego, że $u_1, \dots, u_n$ są niezależne parami, nie zawsze wynika, że są losowo niezależne. Wykazuje to następujący przykład:

Przykład 1. Weźmy pod uwagę ciało atomowe o 12 atomach reprezentacyjnych $a_1, \dots, a_{12}$. Załóżmy, że prawdopodobieństwo atomów zostało ustalone przez związek: $p(a_i) = \frac{1}{12}$, $i = 1, 2, \dots, 12$ (jest to tak zwane prawdopodobieństwo klasyczne; por. § 8). Weźmy pod uwagę zdarzenia:

$$u_1 = a_1 + a_2 + a_5 + a_6 + a_7 + a_8,$$

$$u_2 = a_1 + a_3 + a_7 + a_8 + a_9 + a_{10},$$

$$u_3 = a_1 + a_4 + a_5 + a_6 + a_9 + a_{10}.$$

Wówczas

$$u_1 u_2 = a_1 + a_7 + a_8, \quad u_2 u_3 = a_1 + a_9 + a_{10}, \quad u_3 u_1 = a_1 + a_5 + a_6, \quad u_1 u_2 u_3 = a_1.$$

Wobec tego $p(u_1) = p(u_2) = p(u_3) = \frac{1}{2}$, $p(u_1 u_2) = p(u_2 u_3) = p(u_3 u_1) = \frac{1}{4}$, zatem zdarzenia u_1, u_2, u_3 są parami niezależne.

Zdarzenia te jednak nie są losowo niezależne, gdyż $p(u_1 u_2 u_3) = \frac{1}{12}$ $p(u_1)p(u_2)p(u_3) = \frac{1}{8}$.

W myśl określenia 8, niezależność losowa n zdarzeń wymaga spełnienia warunku (5.4) oraz warunków pochodnych, ogółem więc spełnienia 2^n warunków. Niektóre z nich wynikają z pozostałych, a więc liczbę ich można zredukować. Zagadnieniem powyższej redukcji nie będziemy się jednak zajmowali; jest ono trudne, a nie ma szczególnego znaczenia.

Ograniczymy się do podania jeszcze jednego przykładu, w którym spełniony jest warunek (5.4), ale nie są spełnione warunki pochodne.

Przykład 2. Bierzemy pod uwagę ciało atomowe o atomach $a_1, \dots, a_{16}$; prawdopodobieństwo określamy wzorem $p(a_j) = \frac{1}{16}$, $j = 1, 2, \dots, 16$, i rozpatrujemy zdarzenia

$$u_1 = a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8,$$

$$u_2 = a_1 + a_2 + a_6 + a_7 + a_8 + a_9 + a_{10} + a_{11},$$

$$u_3 = a_1 + a_2 + a_3 + a_4 + a_5 + a_9 + a_{10} + a_{11}.$$

Mamy wówczas $p(u_j) = p(u'_j) = \frac{1}{2}$, $j = 1, 2, 3$, oraz $p(u_1 u_2 u_3) = \frac{1}{8}$, zatem warunek (5.4) jest spełniony. Natomiast $p(u_1 u_2 u'_3) = p(u_1 u'_2 u_3) = p(u'_1 u_2 u_3) = \frac{1}{16}$, zaś $p(u_1 u'_2 u'_3) = p(u'_1 u_2 u'_3) = p(u'_1 u'_2 u_3) = 0$, wreszcie $p(u'_1 u'_2 u'_3) = \frac{1}{16}$, a zatem żaden z warunków pochodnych nie jest spełniony.

Pojęcie niezależności jest jednym z najważniejszych pojęć teorii prawdopodobieństwa.

§ 6. Pojęcie zmiennej losowej i nadziei matematycznej

Pojęcia powyższe będą tu określone dla ciał dość specjalnych.

W toku rozważań niniejszego paragrafu będziemy mianowicie zakładali, że ciało losowe $U, \text{rel } I$ jest atomowe i że zbiór $At(U, I)$ jego atomów reprezentacyjnych jest co najwyżej przeliczalny.

OKREŚLENIE 10. *Zmienną losową ciała losowego $U, \text{rel } I$ nazywamy każdą funkcję $x(u)$ określoną dla $u \in At(U, I)$.*

Tak np. w 6-atomowym ciele jednorazowego rzutu kostką liczbą wyrzuconych oczek jest zmienną losową. Każda stała c może być uważana za zmienną losową $c(u) = c$ dla każdego $u \in At(U, I)$.

Poza tym z określenia 10 wynika, że jeżeli $x_i = x_i(u)$ są zmiennymi losowymi ciała $U, \text{rel } I$, $f(z_1, z_2, \dots, z_k)$ funkcją k zmiennych takich, że zmienna z_i przebiega zbiór wartości zmiennej losowej x_i , to $f(x_1, x_2, \dots, x_k) = f(x_1(u), x_2(u), \dots, x_k(u))$ jest też zmienną losową ciała $U, \text{rel } I$. W zasadzie, wartości zmiennej losowej $x(u)$ mogą należeć do dowolnego zbioru przedmiotów. Najważniejsze są jednak zmienne losowe *rzeczywiste* (to znaczy takie, dla których $x(u)$ jest liczbą rzeczywistą), które będziemy nazywali po prostu *zmiennymi losowymi*; ponadto ważne są zmienne losowe *zespólone* (gdy $x(u)$ jest liczbą zespoloną) i *wektorialne* (gdy $x(u)$ jest elementem przestrzeni wektorialnej, to znaczy takiej, w której określone jest dodawanie elementów oraz ich mnożenie przez liczbę rzeczywistą).

OKREŚLENIE 11. *Nadzieją matematyczną zmiennej losowej $x = x(u)$ nazywamy liczbę, którą oznaczamy bądź przez $\bar{x}$, bądź przez $E(x)$ i określamy przez wzór*

$$(6.1) \quad \bar{x} = E(x) = \sum_{u \in At(U, I)} x(u)p(u),$$

jeżeli (w przypadku gdy zbiór $At(U, I)$ jest przeliczalny) szereg po prawej stronie jest zbieżny bezwzględnie.

Twierdzenie 16. *Nadzieja matematyczna sumy dwu zmiennych losowych równa się sumie ich nadziei matematycznych*

$$(6.2) \quad \overline{x+y} = \bar{x} + \bar{y}.$$

Dowód. Istotnie, z określenia nadziei matematycznej mamy

$$\overline{x+y} = \sum_{u \in At(U, I)} [x(u) + y(u)] p(u) = \sum_{u \in At(U, I)} x(u) p(u) + \sum_{u \in At(U, I)} y(u) p(u) = \bar{x} + \bar{y}.$$

Twierdzenie powyższe można z łatwością uogólnić na sumę n zmiennych losowych, co daje wzór

$$(6.3) \quad \overline{\sum_{j=1}^n x_j} = \sum_{j=1}^n \bar{x}_j.$$

Podamy jeszcze twierdzenie, którego dowód jest natychmiastowy.

$$(6.4) \quad \text{Jeżeli } c \text{ jest dowolną stałą, to } \overline{c} = c, \quad \overline{cx} = c\bar{x}.$$

Określenie nadziei matematycznej można przenieść bez zmiany na zmienne zespolone i wektorialne; zachodzą wtedy również wzory (6.2), (6.3) i (6.4).

Określenie 12. Niech $U_1, \text{rel } I$ będzie podciałem ciała losowego $U, \text{rel } I$, $x(u)$ niech będzie zmienną losową tego podciała. *Przeniesieniem zmiennej losowej $x(u)$ z ciała $U_1, \text{rel } I$ na ciało $U, \text{rel } I$ będziemy nazywali zmienną losową $y(v), v \in At(U, I)$ określoną przez $y(v) = x(u)$, gdzie u oznacza atom reprezentacyjny podciała U_1 zawierający atom v .*

Udowodnimy twierdzenie:

(6.5) *Zmienna losowa i jej przeniesienie mają tę samą nadzieję matematyczną.*

Oznaczmy dla $u \in At(U_1, I)$ przez $At^*(u)$ zbiór tych $v \in At(U, I)$, które są zawarte w u . Tak określone zbiory są rozłączne i w sumie dają zbiór $At(U, I)$, a ponadto

$$\sum_{v \in At^*(u)} v = u, \text{rel } I.$$

Korzystając z tego otrzymujemy na podstawie określeń 11 i 12

$$\begin{aligned} \bar{y} &= \sum_{v \in At(U, I)} y(v) p(v) = \sum_{u \in At(U_1, I)} \sum_{v \in At^*(u)} y(v) p(v) = \\ &= \sum_{u \in At(U_1, I)} \left(x(u) \sum_{v \in At^*(u)} p(v) \right) = \sum_{u \in At(U_1, I)} x(u) p(u) = \bar{x}, \end{aligned}$$

co należało okazać.

Zmienną losową i jej przeniesienia będziemy oznaczali na ogół tą samą literą.

Niech $x_1 = x_1(u)$, $x_2 = x_2(u)$ będą dwiema zmiennymi losowymi ciała $U, \text{rel } I$, xRy stosunkiem określonym dla par x, y , gdzie x jest elementem zbioru wartości zmiennej losowej x_1 , y zaś elementem zbioru wartości zmiennej losowej x_2 . Np. w przypadku gdy x_1, x_2 są zmiennymi losowymi rzeczywistymi, xRy może być jednym ze stosunków $x < y$, $x = y$, $x^2 - y^2 = 0$ itd. Przy tych założeniach możemy mówić o zdarzeniu polegającym na tym, że x_1Rx_2 ; zgodnie z naszą intuicją będziemy uważali, że zdarzeniem tym, które oznaczamy przez $z(x_1Rx_2)$, jest suma tych atomów $u \in At(U, I)$, dla których $x_1(u)Rx_2(u)$. Mamy więc

$$(6.6) \quad z(x_1Rx_2) = \sum_{\substack{u \in At(U, I) \\ x_1(u)Rx_2(u)}} u.$$

Przyjmiamo, że

$$(6.7) \quad P(x_1Rx_2) = p(z(x_1Rx_2)).$$

Czytelnik udowodni łatwo, że w przypadku zmiennych losowych rzeczywistych są prawdziwe następujące wzory:

$$(6.8) \quad P(x_1 < x_2) \leq P(x_1 \leq x_2) = P(x_1 < x_2) + P(x_1 = x_2) = 1 - P(x_1 > x_2).$$

Zachodzą dwa następujące twierdzenia:

(6.9) *Jeżeli dla każdego $u \in At(U, I)$ z warunku $x_1(u)Rx_2(u)$ wynika warunek $x_1(u)Qx_2(u)$, to zachodzą wzory*

$$z(x_1Rx_2) \rightarrow z(x_1Qx_2) \quad \text{i} \quad P(x_1Rx_2) \leq P(x_1Qx_2).$$

Twierdzenie to wynika łatwo z określeń (6.6) i (6.7) oraz z własności prawdopodobieństwa.

(6.10) *Jeżeli zmienna losowa x przybiera tylko wartości rzeczywiste $a_1, a_2, \dots, a_m$, to*

$$\bar{x} = \sum_{i=1}^m a_i P(x = a_i).$$

Dowód. Istotnie, mamy $P(x = a_i) = p\left(\sum_{\substack{u \in At(U, I) \\ x(u) = a_i}} u\right) = \sum_{\substack{u \in At(U, I) \\ x(u) = a_i}} p(u)$,

a więc

$$\sum_{i=1}^m a_i P(x = a_i) = \sum_{i=1}^m a_i \sum_{\substack{u \in At(U, I) \\ x(u) = a_i}} p(u) = \sum_{i=1}^m \sum_{\substack{u \in At(U, I) \\ x(u) = a_i}} a_i p(u) = \sum_{u \in At(U, I)} x(u) p(u) = \bar{x},$$

co należało okazać.

§ 7. Ciała losowe nasycone

Twierdzenie 17. *Nasylenie ciała losowego jest również ciałem losowym.*

Dowód. Niech $\bar{U}, \text{rel} I$ będzie nasyleniem ciała losowego $U, \text{rel} I$, p zaś niech będzie prawdopodobieństwem w tym ostatnim ciele. Określamy w ciele $\bar{U}$ funkcję $\bar{p}$ w następujący sposób: $\bar{p}(v) = p(u)$, jeżeli $v \in \bar{U}$, $v \equiv u, \text{rel} I$, $u \in U$.

Jest to rozszerzenie funkcji $p(u)$ z U na $\bar{U}$, przy czym, jak łatwo sprawdzić, spełnia ono aksjomaty (A1), (A2) i (A3). Niech $p_1(v)$ oznacza inne rozszerzenie $p(u)$ z U na $\bar{U}$, spełniające aksjomaty (A1), (A2) i (A3); ponieważ z aksjomatów tych na mocy twierdzenia 1 wynika, że $p_1(v_1) = p_1(v_2)$ dla $v_1 \equiv v_2, \text{rel} I$, więc dla $v \in \bar{U}$, $v \equiv u, \text{rel} I$, $u \in U$ mamy $p_1(v) = p_1(u) = p(u) = \bar{p}(v)$. Wobec tego $\bar{p}(v)$ jest jedynym rozszerzeniem p z U na $\bar{U}$, spełniającym aksjomaty (A1), (A2), (A3); na mocy aksjomatu (A5) p_1 jest prawdopodobieństwem, co dowodzi twierdzenia.

Z uwagi na powyższe twierdzenie można zawsze zakładać, że ciało losowe jest nasycone.

§ 8. Klasyczna definicja prawdopodobieństwa

Twierdzenie 18. *Jeżeli ciało skończone $U, \text{rel} I$ ma m atomów reprezentacyjnych, $m(u)$ zaś oznacza liczbę tych atomów reprezentacyjnych, które są zawarte w elemencie $u \in U$, to funkcja*

$$(8.1) \quad q(u) = \frac{m(u)}{m}$$

*spełnia aksjomaty (A1), (A*2) i (A3).*

Dowód. Jest to oczywiście dla aksjomatów (A1) i (A3). Jeżeli $uv \equiv 0, \text{rel} I$, $u, v \in U$, to nie ma atomów równocześnie zawartych w u i v , a więc $m(u+v) = m(u) + m(v)$, zatem $q(u+v) = q(u) + q(v)$. Spełniony jest więc aksjomat (A*2), a nawet aksjomat (A2), gdyż w przypadku, gdy ciało $U, \text{rel} I$ jest skończone, aksjomat (A2) orzeka to samo, co aksjomat (A*2).

Określenie 13. Jeżeli w ciele skończonym $U, \text{rel} I$ prawdopodobieństwo określone jest przez wzór $p(u) = q(u) = \frac{m(u)}{m}$, to nazywamy je *prawdopodobieństwem klasycznym*.

Innymi słowy prawdopodobieństwo klasyczne zdarzenia u jest stosunkiem liczby atomów reprezentacyjnych ciała zawartego w zdarzeniu u do liczby wszystkich atomów reprezentacyjnych.

Twierdzenie 19. *Jeżeli $p(u)$ jest prawdopodobieństwem klasycznym w ciele $U, \text{rel} I$, $u_0 \in U$ i $p(u_0) > 0$, to $p_{u_0}(u)$ jest prawdopodobieństwem klasycznym w ciele $U, \text{rel} (I \vee J(u'_0))$.*

Łatwy dowód, w którym należy się powołać na twierdzenie (6.10) z rozdziału II (str. 82), pozostawiamy czytelnikowi.

Szczególna nazwa „prawdopodobieństwo klasyczne” tłumaczy się względami historycznymi; nasze określenie prawdopodobieństwa klasycznego pokrywa się, co do istoty rzeczy, z określeniem prawdopodobieństwa podanym przez klasyka naszej teorii, Laplace'a, i zachowanym w ciągu całego XIX stulecia nawet przez tych autorów, którzy uważali je za niepoprawne. Terminologia nasza jest odmienna; to, co my nazywamy *atomami*, w teorii klasycznej nazywano *przypadkami możliwymi*, atomy zawarte w zdarzeniu nazywano *przypadkami możliwymi sprzyjającymi zdarzeniu*, przy czym dołączano do definicji warunek, aby przypadki możliwe były jednakowo możliwe. Pomijając na razie ten warunek, do którego powrócimy niżej, możemy określenie prawdopodobieństwa klasycznego sformułować w terminach klasycznych:

Prawdopodobieństwo klasyczne zdarzenia jest stosunkiem liczby przypadków możliwych sprzyjających zdarzeniu do liczby wszystkich przypadków możliwych.

Dla klasyków powyższe określenie było punktem wyjścia całej teorii (w toku rozwijania której samo pojęcie prawdopodobieństwa uległo zresztą pewnej ewolucji rozszerzającej). Dla nas określenie prawdopodobieństwa klasycznego jest jednym z możliwych sposobów liczbowego wyznaczenia prawdopodobieństwa, stosowanym przede wszystkim w teorii gier losowych.

Musimy jeszcze zastanowić się nad zagadnieniem następującym: W jakich przypadkach zastosowanie definicji klasycznej można uważać za uzasadnione, a tym samym przypisać prawdopodobieństwu klasycznemu, to znaczy obliczonemu na podstawie tej definicji, wartość obiektywną. Zagadnienie powyższe wiąże się ściśle z pojęciem jednakowej możliwości i z pewną zasadą nieco zagadkową, mianowicie tak zwaną *zasadą racji brakującej*.

Weźmy pod uwagę rzut zwykłą kostką i oznaczmy (por. § 9, rozdział II, str. 87) przez u_i , $i=1,2,\dots,6$ zdarzenie polegające na wyrzuceniu i oczek. Wówczas jedynie możliwymi przypadkami są zdarzenia $u_1, u_2, \dots, u_6$. Ponieważ nie mamy żadnej racji do przewidywania, że nastąpi raczej zdarzenie u_1 niż zdarzenie u_2 , przeto

wszystkie 6 możliwych przypadków uznajemy za „jednakowo możliwe” i przypisujemy im jednakowe prawdopodobieństwo, traktując to prawdopodobieństwo jako „miarę możliwości”. Zwracamy uwagę, że rozumując w ten sposób opieramy się na pewnych obiektywnych danych dotyczących kostki i rzutu, a mianowicie:

(a) rzetelna kostka (a o takiej tu mowa) została sporządzona z materiału możliwie jednorodnego jako prostopadłościan możliwie najbardziej zbliżony do umiarowego; słowem, przy jej sporządzaniu starano się wykluczyć wszelkiego rodzaju asymetrię struktury, która mogłaby sprzyjać któremuś z przypadków możliwych;

(b) rzut powinien być dokonany tak, aby rzucający nie miał możliwości świadomego oddziaływania na wynik rzutu (należy więc np. rzucać kostką z kubka, a nie bezpośrednio z ręki).

W rezultacie nasze rozumowanie i oparte na nim pojęcie prawdopodobieństwa ma wartość obiektywną, gdyż pozwala nam w pewnym stopniu przewidzieć rzeczywiste wyniki gier złożonych z wielu rzutów kostką.

Podobnie przedstawia się sytuacja przy ciągnięciach loteryjnych. Loteria składa się z losowań polegających na wyciąganiu z urny zawierającej zwitki numerowane jednego z tych zwitków. Dyspozycja tego ciągnięcia jest znowu obmyślona tak, aby wykluczyć wszelkie okoliczności, które mogłyby faworyzować którykolwiek numer na niekorzyść innego. A więc: wewnątrz urny jest niewidzialne dla ciążącego; zwitki czy gałki, które wyciągamy, mają tę samą wielkość i są starannie wykonane z jednorodnego materiału, tak aby jedne od drugich nie dawały się odróżnić dotykowo; przed ciągnięciem mieszamy je przez potrząsanie lub obracanie urny, wreszcie, co ma zapewne znaczenie raczej dekoracyjne, powierzymy zazwyczaj ciągnięcie dziecku z jakiegoś sierocińca jako istocie specjalnie naiwnej i niewinnej. Wobec tego *nie mamy żadnej racji do przewidywania*, że spośród znajdujących się w urnie numerów wyciągnięty będzie raczej numer k niż l ; znowu więc przypadki możliwe uznajemy za „jednakowo możliwe” i określamy ich prawdopodobieństwa jako równe. I w tym przypadku doświadczenie stwierdza obiektywną wartość tego określenia jako narzędzia przewidywań.

Trzecim przykładem są gry losowe w karty. Elementami gwarantującymi jednakową możliwość rozdań a zarazem niemożliwość rozpoznania zakrytej karty są: jednakowa wielkość kart, jednorodność papieru, z którego są wykonane, jednakowy wygląd odwrotu, wreszcie mieszanie przed grą (tasowanie i zbieranie), które

zresztą (według opinii Culbertsona) jest na ogół niewystarczające. Do tego dochodzą: częsta zmiana talii (gdyż karty odkształcają się nierównomiernie w toku gry) oraz selekcja graczy.

Podkreślone przy omawianiu pierwszych dwu przykładów zdania: „nie mamy żadnej racji itd.” stanowią cząstkowe sformułowanie tego, cośmy nazwali *zasadą brakującej racji*. Podanie jednak ogólnego sformułowania tej zasady jest niezmiernie trudne. Sądzę wszakże, że można powiedzieć, co następuje:

Możemy konstruować pewne gry w taki sposób, aby na podstawie możliwie dokładnej znajomości narzędzi gry i metod jej dyspozycji nie było żadnej racji do przewidywania, że z dwóch dowolnych przypadków możliwych nastąpi raczej jeden niż drugi.

Uznajemy wówczas przypadki możliwe za jednakowo możliwe i przypisujemy im prawdopodobieństwa według definicji klasycznej. Tak określone prawdopodobieństwa mają wartość obiektywną jako narzędzie przewidywania przebiegu gry.

§ 9. Przykłady i zadania

(9.1) Niech $A_1, A_2, \dots, A_k$ będą zbiorami przeliczalnymi, parami rozłącznymi. Oznaczmy przez $V(A_i)$ klasę podzbiorów skończonych zbioru A_i i ich dopełnień; każdy więc zbiór $X_i \in W(A_i)$ bądź jest skończony, bądź też jego uzupełnienie $A_i - X_i$ jest skończone. Klasa zbiorów kształtu $X_1 + X_2 + \dots + X_k$, gdzie $X_i \in V(A_i)$, jest ciałem zbiorów, w szczególności podciałem ciała $W\left(\sum_{i=1}^k A_i\right)$. W podciele tym określamy funkcję g wzorem $g(X_1 + X_2 + \dots + X_k) = l/k$, gdzie l jest liczbą tych zbiorów X_i ($i=1, \dots, k$), które są nieskończone. Udowodnić, że funkcja g jest w tym ciele addytywna, ale nie jest przeliczalnie addytywna.

(9.2) Udowodnić, że w ciele losowym $U, \text{rel } I$ zbiór I_0 tych elementów $u \in U$, dla których $p(u) = 0$, jest ideałem przeliczalnie addytywnym, zawierającym I .

(9.3) Mówimy, że prawdopodobieństwo $p(u)$ w ciele losowym $U, \text{rel } I$ jest *ściśle rosnące*, jeżeli warunek $p(u) = 0$ jest spełniony tylko dla elementów niemożliwych, $\text{rel } I$ (a więc $p(u) = 0$ wtedy i tylko wtedy, gdy $u \equiv 0, \text{rel } I$).

Udowodnić, że jeżeli p jest prawdopodobieństwem w ciele $U, \text{rel } I$, to p jest prawdopodobieństwem ściśle rosnącym w ciele $U, \text{rel } I_0$, gdzie I_0 ma to samo znaczenie co w zadaniu (9.2).

(9.4) W ciele ilorazowym U/I ciała losowego U , rel I określamy funkcję $P([u])$ wzorem $P([u]) = p(u)$ (możliwość przyjęcia takiej definicji wynika z twierdzenia 1).

Udowodnić, że $P([u])$ jest funkcją przeliczalnie addytywną w ciele U/I .

(9.5) Udowodnić, że jeżeli U , rel I jest ciałem losowym, prawdopodobieństwo p jest ściśle rosnące w U , rel I , a $P(X)$ jest funkcją w U/I zdefiniowaną w zadaniu (9.4), to U/I jest przestrzenią metryczną względem metryki zdefiniowanej dla $[u], [v] \in U/I$ przez wzór

$$\rho([u], [v]) = P([u] \div [v]).$$

(9.6) Udowodnić, że jeżeli zdarzenia $u_1, u_2, \dots, u_n$ są losowo niezależne w ciele U , rel I , to zbiór $(u_1, u_2, \dots, u_n)$ jest zbiorem swobodnym.

W szczególności, w takim przypadku elementy zbioru $(u_1, u_2, \dots, u_n)$ są niezależne w sensie działań algebry Boole'a w ciele U , rel I (por. (9.2), rozdział II, str. 89).

(9.7) Udowodnić, że jeżeli A_1 i A_2 są dwoma podzbiorami skończonymi ciała losowego U , rel I , przy czym elementy zbiorów A_i ($i=1,2$) wykluczają się parami, każda zaś para elementów u_1, u_2 , gdzie $u_1 \in A_1$, $u_2 \in A_2$, jest losowo niezależna, to elementy $\sum_{u \in A_1} u$ i $\sum_{u \in A_2} u$ są również losowo niezależne.

(9.8) Niech $V(A)$, rel $I(A)$ będzie losowym, kanonicznym ciałem zbioru A , składającego się z m elementów, $z = z(u)$, a $u \in A$ nieujemną zmienną losową ciała $V(A)$, rel $I(A)$, którą będziemy nazywali *zmienną losową zapłaty*; *stawką* będziemy nazywali każdą nieujemną zmienną losową $s = s(u)$ ciała $V(A)$, rel $I(A)$, a *szansą gracza przy stawce s* będziemy nazywali zmienną losową $x_s = x_s(u) = z(u)s(u)$. Zmienną zapłaty będziemy nazywali *sprawiedliwą*, jeżeli nadzieja matematyczna szansy gracza przy każdej stawce jest równa $\sum_{u \in A} s(u)$.

Udowodnić, że zapłata jest sprawiedliwa tylko wtedy, gdy $z(u) = \frac{1}{p(u)}$ dla każdego $u \in A$.

Wyprowadzić stąd wniosek, że jedyną sprawiedliwą zapłatą przy grze „rzetelnej” (to znaczy o prawdopodobieństwie klasycznym) sześcienną kostką do gry jest zapłata sześciokrotnej stawki przy wyrzuceniu obstawionego numeru.

Rozdział IV

Schemat Bernoulliego

§ 1. Treść rozdziału

Przedmiotem niniejszego rozdziału są pewne specjalne zespoły regularne będące ciałami losowymi, a mianowicie tak zwane zespoły prób niezależnych nad pewnym zjawiskiem C lub pewnymi zjawiskami $C^{(1)}, C^{(2)}, \dots, C^{(m)}$.

Nie zakładamy tu, że prawdopodobieństwo jest dane klasyczną definicją, zakładamy natomiast stałość występujących tu prawdopodobieństw, od próby do próby, i wzajemną niezależność prób, co prowadzi do tak zwanego *schematu Bernoulliego*.

Rozpoczynamy od określenia *zjawiska*, *próby* i *zespołu prób* (§ 2), po czym w paragrafie 3 określamy *niezależność prób* i schemat Bernoulliego, a dalej omawiamy *częstość* zjawiska i podajemy zasadniczy wzór na prawdopodobieństwo, że dane zjawisko w n pierwszych próbach pojawi się dokładnie k razy. W paragrafie 4 obliczamy nadzieję matematyczną częstości, paragraf 5 zaś zawiera wzór asymptotyczny na P_{nk} oraz ustalenie wartości k , dla której prawdopodobieństwo to ma maksimum.

W paragrafie 6 podajemy najprostsze prawo wielkich liczb Jakuba Bernoulliego.

W paragrafie 7 wyprowadzimy wzór graniczny de Moivre'a-Laplace'a, przy czym natrafimy na zasadniczą w teorii prawdopod

podobieństwa całkę $\frac{1}{\sqrt{2\pi}} \int_{\lambda_1}^{\lambda_2} e^{-\frac{1}{2}t^2} dt$.

W paragrafie 8 podajemy wzór graniczny Poissona na prawdopodobieństwo P_{nk} dla zdarzeń rzadkich. Paragraf 9 zawiera pewne wiadomości z kombinatoryki, potrzebne do paragrafów 10-13,