gdzie b oznacza wyciągnięcie kuli białej, a c kuli czarnej, a następnie prawdopodobieństwa

$$P(\{b\}) = 1, \qquad P(\{c\}) = 0$$

W zagadnieniach bardziej złożonych liczba możliwych do przyjęcia modeli bywa nieraz bardzo duża. Nie wszystkie modele są z punktu widzenia celu równoważne. Wielokrotnie przydatność modelu można sprawdzić jedynie drogą konfrontacji z rzeczywistością.

Dopasowywanie modeli do rzeczywistości nie należy do matematyki. Matematyczna teoria prawdopodobieństwa gwarantuje jedynie wewnętrzną poprawność modeli probabilistycznych. Ze sprawą dopasowywania modeli do rzeczywistości stykamy się jedynie w przykładach zastosowań teorii prawdopodobieństwa do konkretnych zagadnień.

§ 135. Przestrzeń probabilistyczna przeliczalna

Niech

$$X = \{e_1, e_2, \dots\}, \qquad \bigwedge_{\substack{j,k \in \mathfrak{N} \\ j \neq k}} e_j \neq e_k$$

i niech S będzie klasą wszystkich podzbiorów przestrzeni X. Klasa S jest nieprzeliczalna. Gdyby bowiem była co najwyżej przeliczalna, to tym bardziej byłaby co najwyżej przeliczalna klasa wszystkich nieskończonych podzbiorów przestrzeni X, czyli byłaby co najwyżej przeliczalna klasa wszystkich nieskończonych podciągów ciągu $(e_1, e_2, \dots)$. Ale wtedy wszystkie takie podciągi dałyby się ustawić w ciąg

$$\text{(i)} \qquad \begin{gathered} \left(e_{l_{1,1}},\ e_{l_{1,2}}, \dots\right) \\ \left(e_{l_{2,1}},\ e_{l_{2,2}}, \dots\right) \\ \dots\dots\dots\dots\dots \end{gathered}$$

gdzie

$$\bigwedge_{j,k \in \mathfrak{N}} l_{j,k+1} > l_{j,k}$$

Rozpatrzmy ciąg

$$\text{(ii)} \qquad \left(e_{l_{1,2}},\ e_{(l_{1,2} + l_{2,3})},\ e_{(l_{1,2} + l_{2,3} + l_{3,4})} \dots\right)$$

Ciąg ten jest podciągiem ciągu $(e_1, e_2, \ldots)$, a nie należy do ciągu (i), ponieważ dowolny k-ty z kolei podciąg z ciągu (i) różni się co najmniej swym k-tym wyrazem $e_{l_{k,k}}$ od k-tego wyrazu ciągu (ii), czyli od $e_{(l_{1,2}+\ldots+l_{k,k+1})}$. Przeczy to założeniu, że wszystkie nieskończone podciągi ciągu $(e_1, e_2, \ldots)$ dają się ustawić w ciąg. Oznacza to, że klasa wszystkich podciągów ciągu $(e_1, e_2, \ldots)$ jest nieprzeliczalna. Tym bardziej klasa S jest nieprzeliczalna.

Klasa S jest σ-ciałem, gdyż spełnia wszystkie warunki (σ1), (σ2), (σ3).

Możemy napisać

$$(60) \qquad X = \sum_{j=1}^{\infty} \{e_j\}$$

i wobec tego

$$(61) \qquad \bigwedge_{A \in S} A = A \cap X = \sum_{j=1}^{\infty} (A \cap \{e_j\})$$

Niech teraz będą dane liczby rzeczywiste $p_1, p_2, \ldots$ spełniające warunki

$$(62) \qquad p_1, p_2, \ldots \geq 0, \qquad \sum_{j=1}^{\infty} p_j = 1$$

Wprowadzamy funkcję rzeczywistą P o dziedzinie S będącej σ-ciałem, określoną wzorem

$$(63) \qquad \bigwedge_{A \in S} P(A) = \sum_{j=1}^{\infty} a_j p_j$$

gdzie

$$a_j = \begin{cases} 1 & \text{gdy } A \cap \{e_j\} \neq 0 \quad \text{czyli } e_j \in A \\ 0 & \text{gdy } A \cap \{e_j\} = 0 \quad \text{czyli } e_j \notin A \end{cases}$$

Wykażemy, że funkcja P jest miarą unormowaną. Istotnie, jest ona funkcją rzeczywistą zbioru o dziedzinie S będącej σ-ciałem przestrzeni X, a więc spełnia warunek (ν1) dla miary unormowanej. Na mocy (63) i (62) jest nieujemna, a więc spełnia warunek (ν2). Następnie na mocy (60), (63), (62) mamy

$$P(X) = \sum_{j=1}^{\infty} p_j = 1$$

czyli jest spełniony warunek $(v3)$. Niech teraz $A_1, A_2, \ldots$ będzie dowolnym ciągiem zbiorów należących do S i rozłącznych, tzn.

$$\bigwedge_{\substack{k,l \in \mathfrak{N} \\ k \neq l}} A_k \cap A_l = 0$$

Mamy wtedy na mocy (61)

$$\text{(iii)} \quad p\left(\sum_{k=1}^{\infty} A_k\right) = p\left(\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}(A_k \cap \{e_j\})\right) = p\left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(A_k \cap \{e_j\})\right) =$$

$$= p\left(\sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty} A_k\right) \cap \{e_j\}\right)$$

Niech teraz zgodnie z (63)

$$\text{(iv)} \qquad \bigwedge_{k \in \mathfrak{N}} p(A_k) = \sum_{j=1}^{\infty} a_{kj}\, p_j$$

gdzie

$$a_{kj} = \begin{cases} 1 & \text{gdy } e_j \in A_k \\ 0 & \text{gdy } e_j \notin A_k \end{cases}$$

Na mocy (iii) jest

$$\text{(v)} \qquad p\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{kj}\right) p_j = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} a_{kj}\, p_j$$

Zauważmy, że z rozłączności zbiorów $A_1, A_2, \ldots$ wynika, iż

$$\bigwedge_{j \in \mathfrak{N}}\left(\bigvee_{k_0} e_j \in A_{k_0} \Rightarrow \bigwedge_{k \neq k_0} e_j \notin A_k\right)$$

Wobec tego

$$\sum_{k=1}^{\infty} a_{kj} = \begin{cases} 1 & \text{gdy } \bigvee_{k_0} e_j \in A_{k_0} \\ 0 & \text{gdy } \bigwedge_{k} e_j \notin A_k \end{cases}$$

skąd

$$\sum_{k=1}^{\infty} a_{kj}\, p_j = \begin{cases} p_j & \text{gdy } \bigvee_{k_0} e_j \in A_{k_0} \\ 0 & \text{gdy } \bigwedge_{k} e_j \notin A_k \end{cases}$$

Zatem na mocy (62)

$$\sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{kj}\, p_j\right) \leq \sum_{j=1}^{\infty} p_j = 1$$

Wynika stąd, że szereg

$$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} a_{kj}\, p_j = \sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{kj}\, p_j\right)$$

jako niemalejący i ograniczony od góry jest zawsze zbieżny.

Udowodnimy teraz następujący lemat:

Jeżeli istnieje suma

(vi) $$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\alpha_{kj} = s < \infty$$

gdzie

$$\bigwedge_{k,j} \quad \alpha_{kj} \in \mathfrak{R} \wedge \alpha_{kj} \geqslant 0$$

to istnieje również suma

(vii) $$\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\alpha_{kj}$$

i zachodzi równość

(viii) $$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\alpha_{kj} = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\alpha_{kj}$$

Dowód: Rozpatrzymy sumy

(ix) $$\sum_{k=1}^{m}\sum_{j=1}^{\infty}\alpha_{kj}$$

oraz

$$\sum_{k=1}^{m}\sum_{j=1}^{n_k}\alpha_{kj}$$

gdzie $m, n_1, \ldots, n_m$ są dowolnymi liczbami naturalnymi. Niech

$$N = \max\left\{ n_1, \ldots, n_m \right\}$$

Mamy wtedy

$$\sum_{k=1}^{m}\sum_{j=1}^{n_k}\alpha_{kj} \leq \sum_{k=1}^{m}\sum_{j=1}^{N}\alpha_{kj} = \sum_{j=1}^{N}\sum_{k=1}^{m}\alpha_{kj} \leq \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\alpha_{kj}$$

Na mocy (vi) możemy zatem napisać

(x) $$\bigwedge_{m\in\mathfrak{N}}\ \bigwedge_{n_1,\ldots,n_m\in\mathfrak{N}}\ \sum_{k=1}^{m}\sum_{j=1}^{n_k}\alpha_{kj} \leq s$$

Wynika stąd również, że

$$\bigwedge_{k \in \mathfrak{N}} \bigwedge_{n_k \in \mathfrak{N}} \sum_{j=1}^{n_k} \alpha_{kj} \leq s$$

Zatem wszystkie szeregi

$$\sum_{j=1}^{\infty} \alpha_{kj} \, , \qquad k = 1,2,\ldots$$

jako niemalejące i ograniczone od góry są zbieżne i dla każdego $m \in \mathfrak{N}$ istnieje suma (ix). Przechodząc w (x) kolejno do granicy dla

$$n_1 \to \infty, \ldots, n_m \to \infty$$

otrzymujemy nierówność

(xi) $$\bigwedge_{m \in \mathfrak{N}} \sum_{k=1}^{m} \sum_{j=1}^{\infty} \alpha_{kj} \leq s$$

Sumy (ix) tworzą zatem ciąg niemalejący i ograniczony od góry, a więc zbieżny. Istnieje zatem suma (vii), a z nierówności (xi) wynika, że

$$\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \alpha_{kj} \leq s$$

Z symetrii wynika, że gdybyśmy sumę (vii) oznaczyli symbolem s^*, to analogicznie można dowieść, że

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_{kj} \leq s^*$$

Mielibyśmy zatem nierówności

$$s^* \leq s \quad \text{i} \quad s \leq s^*$$

skąd wynika $s = s^*$ czyli równość (viii) stanowiąca tezę lematu.

Wykorzystując dowiedziony lemat możemy na mocy (v) napisać

$$p\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} a_{kj} \, p_j = \sum_{k=1}^{\infty} \left(\sum_{j=1}^{\infty} a_{kj} \, p_j\right)$$

skąd na mocy (iv)

$$p\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} p(A_k)$$

Funkcja P spełnia zatem także warunek (ν4) i jest miarą unormowaną. Wobec tego trójkę (X, S, P) możemy przyjąć jako przestrzeń probabilistyczną. Będziemy ją nazywać przestrzenią probabilistyczną przeliczalną. Przestrzenie probabilistyczne skończone i przestrzenie probabilistyczne przeliczalne obejmujemy wspólną nazwą przestrzeni probabilistycznych co najwyżej przeliczalnych.

Zauważmy jeszcze, że wobec definicji (63) wystarczy dla określenia rozkładu prawdopodobieństwa w przestrzeni przeliczalnej określić liczby $p_1, p_2, \ldots$ z warunku

$$p_k = P(\{e_k\}) \quad \text{dla} \quad k = 1,2,\ldots \tag{64}$$

§ 136. Przykład

Powtarzamy wielokrotnie rzut monetą. Niech e_k $(k = 1,2,\ldots)$ oznacza zdarzenie polegające na tym, że orzeł pojawił się po raz pierwszy w k-tym z kolei rzucie, a e_0 zdarzenie, że orzeł w ogóle się nie pojawił. Wobec tego za przestrzeń zdarzeń elementarnych możemy przyjąć

$$X = \{e_0, e_1, e_2, \ldots\}$$

a za ciało zdarzeń S klasę wszystkich podzbiorów przestrzeni X. Aby określić przestrzeń probabilistyczną przeliczalną, wystarczy jeszcze określić rozkład prawdopodobieństwa za pośrednictwem wzoru (64). Zgodnie z naszą intuicją, możemy przede wszystkim, przyjąć, że

$$p_0 = P(\{e_0\}) = 0$$

Rozpatrzmy teraz przypadek, gdy orzeł pojawił się po raz pierwszy w k-tym rzucie. Wszystkich możliwych serii po k-rzutów każda jest 2^k. Z tego interesuje nas jedna, w której $k-1$ razy pojawiła się reszka, a dopiero za k-tym razem orzeł. Traktując wszystkie serie jako jednakowo prawdopodobne, możemy przyjąć, że

$$p_k = P(\{e_k\}) = \frac{1}{2^k} \qquad \text{dla} \quad k = 1,2,\ldots$$

Sprawdzamy, że

$$\bigwedge_{k \in \mathfrak{N}_0} p_k \geqslant 0, \qquad \sum_{k=0}^{\infty} p_k = 1$$

Określiliśmy w ten sposób rozkład prawdopodobieństwa.

Obliczmy dla przykładu prawdopodobieństwo, że orzeł pojawi się po raz pierwszy nie wcześniej jak za setnym rzutem. Oznaczmy to zdarzenie symbolem A.

Mamy

$$A = \{e_{100}\} + \{e_{101}\} + \dots$$

czyli

$$A = \{e_{100}, \; e_{101}, \dots\}$$

i na mocy (63)

$$P(A) = \sum_{k=100}^{\infty} p_k = \sum_{k=100}^{\infty} \frac{1}{2^k} = \frac{1}{2^{100}}\left(1 + \frac{1}{2} + \frac{1}{2^2} + \dots\right) = \frac{1}{2^{99}} \approx$$

$$\approx 1{,}6 \cdot 10^{-30}$$

§ 137. Przestrzeń probabilistyczna euklidesowa

W bardzo wielu zastosowaniach przyjmuje się za przestrzeń zdarzeń elementarnych X przestrzeń euklidesową $\mathfrak{R}^n$, a za ciało zdarzeń S ciało zbiorów borelowskich w $\mathfrak{R}^n$. Jest to klasa zbiorów wystarczająca dla praktyki, a jednocześnie pozwalająca na określenie na niej rozkładu prawdopodobieństwa P. Przestrzeń probabilistyczną (X, S, P) w przypadku, gdy X jest przestrzenią euklidesową $\mathfrak{R}^n$, a S ciałem zbiorów borelowskich w $\mathfrak{R}^n$, będziemy nazywać przestrzenią probabilistyczną euklidesową n-wymiarową albo po prostu przestrzenią probabilistyczną euklidesową.

Podobnie, jak w przypadku przestrzeni probabilistycznych co najwyżej przeliczalnych było wygodnie wprowadzać rozkład prawdopodobieństwa przez określenie prawdopodobieństw dla zdarzeń jednoelementowych (wzory (59) i (64)), tak rozkład prawdopodobieństwa P w przestrzeni probabilistycznej euklidesowej wprowadza się przez określenie prawdopodobieństw dla zdarzeń postaci przedziałów $\langle -\infty; x)$, gdzie $x \in \mathfrak{R}_0^n$, czyli przez określenie funkcji rzeczywistej

$$\bigwedge_{x \in \mathfrak{R}_0^n} F(x) \overset{\text{def}}{=} P[\langle -\infty \; ; \; x)] \tag{65}$$

Funkcję $F(x)$ nazywamy dystrybuantą rozkładu prawdopodobieństwa w przestrzeni probabilistycznej euklidesowej n-wymiarowej albo dystrybuantą n-wymiarowego rozkładu prawdopodobieństwa albo po prostu dystrybuantą rozkładu prawdopodobieństwa.

Ze wzoru (65) wynika, że dystrybuanta rozkładu prawdopodobieństwa jest dystrybuantą miary P i wobec tego na mocy twierdzenia § 114 jest dystrybuantą n-wymiarową, tzn. spełnia warunki (F1),...,(F5) z § 103 i ma własności (f1),...,(f5) z tegoż paragrafu.

Twierdzenie § 115 gwarantuje odpowiedniość wzajemnie jednoznaczną rozkładów prawdopodobieństwa w przestrzeni probabilistycznej euklidesowej n-wymiarowej i dystrybuant n-wymiarowych.

Ze wzoru (∗∗) podanego w dowodzie twierdzenia § 114 wynika, że dla dowolnego przedziału lewostronnie domkniętego $\langle a;b)$, gdzie $a,b\in\Re_0^n$,

$$P[\langle a;b)] = \Delta_b^n F(a) \tag{66}$$

Z łatwością wyprowadzimy wzory na prawdopodobieństwo dla innych przedziałów. Mamy mianowicie

$$(a;b) = \bigcup_{k=1}^{\infty} \langle c_k;b), \quad \text{gdzie} \quad c_k\in\Re_0^n,\ a<c_k\leq b,\ c_k\searrow a$$

Kładąc

$$A_k = \langle c_k;b) \quad \text{dla} \quad k=1,2,\ldots$$

mamy

$$A_1\subset A_2\subset\ldots$$

i na mocy (47)

$$P[(a;b)] = \lim_{k\to\infty} \Delta_b^n F(c_k) = \Delta_b^n F(a+) \tag{67}$$

Analogicznie, przyjmując, że

$$b=(b_1,\ldots,b_n), \qquad b_1,\ldots,b_n\in\Re_0$$

mamy

$$\langle a;b\rangle = \bigcap_{k=1}^{\infty} \langle a;c_k)$$

gdzie

$$c_k = (c_{k1},\ldots,c_{kn})\in\Re_0^n, \qquad c_{k1},\ldots,c_{kn}\in\Re_0,$$

$$c_{kj}\searrow b_j \quad \text{gdy} \quad b_j<\infty,$$

$$c_{kj}=\infty \quad \text{gdy} \quad b_j=\infty, \quad j=1,\ldots,n; \quad k=1,2,\ldots$$

Otrzymujemy wtedy na mocy (48)

$$P\left[\langle a\,;\,b\rangle\right] = \lim_{k\to\infty} \Delta^n_{c_k} F(a)$$

co będziemy pisać w postaci

(68) $$P\left[\langle a\,;b\rangle\right] = \Delta^n_{b+} F(a)$$

W ten sam sposób otrzymujemy ze wzoru

$$(a\,;b\rangle = \bigcap_{k=1}^{\infty} (a;\, c_k)$$

wzór

(69) $$P\left[(a;\,b\rangle\right] = \Delta^n_{b+} F(a+)$$

Ze wzoru (65) wynika, że

(70) $$P\left[\langle -\infty\,;b)\right] = P\left[(-\infty\,;b)\right] = F(b)$$

i na mocy (68)

(71) $$P\left[\langle -\infty\,;\,b\rangle\right] = P\left[(-\infty\,;\,b\rangle\right] = F(b+)$$

Ze wzoru (66) wynika, że

(72) $$P\left[\langle a\,;\,\infty)\right] = P\left[\langle a\,;\infty\rangle\right] = \Delta^n_{\infty} F(a)$$

a ze wzoru (67)

(73) $$P\left[(a;\,\infty)\right] = P\left[(a;\,\infty\rangle\right] = \Delta^n_{\infty} F(a+)$$

Wreszcie ze wzoru (65) na mocy (39) otrzymujemy

(74) $$P\left[(-\infty\,;\,\infty)\right] = P\left[\langle -\infty\,;\,\infty)\right] = P\left[(-\infty;\,\infty\rangle\right] = P\left[\langle -\infty\,;\,\infty\rangle\right] = 1$$

Ponieważ zbiór jednopunktowy $\{a\}$, gdzie $a \in \mathfrak{R}^n$ można przedstawić w postaci

$$\{a\} = \langle a\,;a\rangle$$

więc na mocy (68) mamy

(75) $$P\left[\{a\}\right] = \Delta^n_{a+} F(a)$$

Wynika stąd, że w przypadku, gdy dystrybuanta F jest funkcją ciągłą

$$\bigwedge_{a\,\in\,\mathfrak{R}^n} P\left[\{a\}\right] = 0$$

§ 138. Przestrzeń probabilistyczna prosta

Przestrzenią probabilistyczną prostą będziemy nazywać przestrzeń probabilistyczną euklidesową 1-wymiarową. Przestrzeń probabilistyczna prosta jest najprostszym a równocześnie najczęściej spotykanym przypadkiem przestrzeni probabilistycznych euklidesowych.

W przypadku przestrzeni probabilistycznej prostej, czyli w przypadku $n = 1$, wzory (66)-(75) ulegają uproszczeniu:

(76) $P[\langle a;b)] = F(b) - F(a)$

(77) $P[(a;b)] = F(b) - F(a+)$

(78) $P[\langle a;b\rangle] = F(b+) - F(a)$

(79) $P[(a;b\rangle] = F(b+) - F(a+)$

(80) $P[(-\infty;b)] = F(b)$

(81) $P[(-\infty;b\rangle] = F(b+)$

(82) $P[(a;\infty)] = 1 - F(a+)$

(83) $P[\langle a;\infty)] = 1 - F(a)$

(84) $P[(-\infty;\infty)] = 1$

(85) $P[\{a\}] = F(a+) - F(a)$

§ 139. Przykłady

I. Sprawdzić, że funkcja

$$(*) \qquad F(x_1, x_2) = \frac{1}{4}\left(\frac{x_1}{1+|x_1|} + 1\right)\left(\frac{x_2}{1+|x_2|} + 1\right)$$

gdzie $x_1, x_2 \in \Re_0$, jest dystrybuantą dwuwymiarową.

Istotnie, funkcja $(*)$ jest funkcją rzeczywistą, dla której dziedziną jest cała przestrzeń $\Re_0^2$, a zatem spełniającą warunek (F1) dla dystrybuanty. Następnie mamy na mocy twierdzenia § 78 dla $x = (x_1, x_2)$, $y = (y_1, y_2) \in \Re_0^2$, $x \leq y$

$$\Delta_y^2 F(x) = F(y_1, y_2) - F(x_1, y_2) - F(y_1, x_2) + F(x_1, x_2) =$$

$$= \frac{1}{4}\left[\left(\frac{y_1}{1+|y_1|}+1\right)\left(\frac{y_2}{1+|y_2|}+1\right)-\right.$$

$$-\left(\frac{x_1}{1+|x_1|}+1\right)\left(\frac{y_2}{1+|y_2|}+1\right)-\left(\frac{y_1}{1+|y_1|}+1\right)\left(\frac{x_2}{1+|x_2|}+1\right)+$$

$$\left.+\left(\frac{x_1}{1+|x_1|}+1\right)\left(\frac{x_2}{1+|x_2|}+1\right)\right] =$$

$$= \frac{1}{4}\left[\left(\frac{y_1}{1+|y_1|}-\frac{x_1}{1+|x_1|}\right)\left(\frac{y_2}{1+|y_2|}+1\right)-\right.$$

$$\left.-\left(\frac{y_1}{1+|y_1|}-\frac{x_1}{1+|x_1|}\right)\left(\frac{x_2}{1+|x_2|}+1\right)\right] =$$

$$= \frac{1}{4}\left(\frac{y_1}{1+|y_1|}-\frac{x_1}{1+|x_1|}\right)\left(\frac{y_2}{1+|y_2|}-\frac{x_2}{1+|x_2|}\right)$$

Ponieważ dla $y_j \geqslant x_j$ $(j = 1,2)$ **mamy**

w przypadku $x_j \leqslant y_j \leqslant 0$:

$$\frac{y_j}{1+|y_j|}-\frac{x_j}{1+|x_j|} = \frac{y_j}{1-y_j}-\frac{x_j}{1-x_j} = \frac{y_j - x_j}{(1-x_j)(1-y_j)} \geqslant 0$$

w przypadku $x_j \leqslant 0 < y_j$:

$$\frac{y_j}{1+|y_j|}-\frac{x_j}{1+|x_j|} = \frac{y_j}{1+y_j}-\frac{x_j}{1-x_j} = \frac{y_j - x_j - 2x_j y_j}{(1-x_j)(1+y_j)} \geqslant 0$$

w przypadku $0 < x_j \leqslant y_j$:

$$\frac{y_j}{1+|y_j|}-\frac{x_j}{1+|x_j|} = \frac{y_j}{1+y_j}-\frac{x_j}{1+x_j} = \frac{y_j - x_j}{(1+x_j)(1+y_j)} \geqslant 0$$

zatem

$$\bigwedge_{x \leq y} \Delta_y^2 F(x) \geqslant 0$$

czyli funkcja F **jest dwuwymiarowo niemalejąca i tym samym spełnia warunek (F2) dla dystrybuanty.**

Funkcje

$$(**) \qquad f_j(x_j) = \frac{x_j}{1+|x_j|} + 1, \qquad j = 1,2,\dots$$

są ciągłe, ponieważ

$$\bigwedge_{x_j, y_j \in \mathcal{R}} f_j(y_j) - f_j(x_j) = \frac{y_j}{1+|y_j|} - \frac{x_j}{1+|x_j|} =$$

$$= \frac{y_j - x_j + y_j\,|x_j| - x_j\,|y_j|}{(1+|x_j|)(1+|y_j|)}$$

i dla $x_j y_j \geqslant 0$

$$\left| f_j(y_j) - f_j(x_j) \right| = \frac{|y_j - x_j|}{(1+|x_j|)(1+|y_j|)} \leqslant \left| y_j - x_j \right|$$

a

$$\bigwedge_{x_j} \bigvee_{\delta > 0} \bigwedge_{0<\varepsilon<\delta} \bigwedge_{y_j} (|y_j - x_j| < \varepsilon \Rightarrow x_j y_j \geqslant 0)$$

gdyż dla $x_j \neq 0$ wystarczy przyjąć $\delta = |x_j|$ a dla $x_j = 0$ wartość δ może być dowolna.

Ponadto

(***) $$f_j(-\infty) = \lim_{x_j \to -\infty} f_j(x_j) = 0,$$

$$f_j(\infty) = \lim_{x_j \to \infty} f_j(x_j) = 2$$

czyli funkcje (**) są ciągłe w całej przestrzeni $\mathcal{R}_0$. Funkcję (*) można przedstawić w postaci

(****) $$F(x_1, x_2) = \frac{1}{4} f_1(x_1) f_2(x_2)$$

Z ciągłości funkcji f_1 i f_2 wynika zatem ciągłość funkcji F, a tym bardziej lewostronna ciągłość funkcji F. Funkcja (*) spełnia zatem warunek (F3) dla dystrybuanty. Spełnienie warunków (F4) i (F5) wynika ze wzorów (***) i (****). Wykazaliśmy w ten sposób, że funkcja (*) jest dystrybuantą 2-wymiarową.

II. Niech dystrybuanta (*) określona w poprzednim przykładzie będzie dystrybuantą rozkładu prawdopodobieństwa na płaszczyźnie. Obliczyć prawdopodobieństwo dla przedziału $\langle a ; b \rangle$, gdzie $a = (-1,-1)$, $b = (1,1)$.

Rozwiązanie:

Ze względu na ciągłość dystrybuanty mamy

$$\bigwedge_{x \in \mathcal{R}^2} F(x+) = F(x)$$

i wobec tego wzór (68) przyjmie postać

$$P(\langle a; b\rangle) = \Delta_b^2 F(a) = F(1,1) - F(-1,1) - F(1,-1) +$$

$$+ F(-1,-1) = \frac{9}{16} - \frac{3}{16} - \frac{3}{16} + \frac{1}{16} = \frac{1}{4}$$

III. Określić dystrybuantę rozkładu prawdopodobieństwa na prostej, jeżeli wiadomo, że prawdopodobieństwa dla punktów

$$a_1 = -1, \quad a_2 = 0, \quad a_3 = 1$$

są odpowiednio równe

$$p_1 = \frac{1}{6}, \quad p_2 = \frac{1}{3}, \quad p_3 = \frac{1}{2}$$

a prawdopodobieństwo każdego przedziału nie zawierającego żadnego z tych punktów jest zerem.

Rozwiązanie:

Z definicji mamy wzór na dystrybuantę

$$F(x) = P(\langle -\infty; x))$$

Gdy $x \leq a_1$, przedział $\langle -\infty; x)$ nie zawiera żadnego z punktów a_1, a_2, a_3 i wobec tego $F(x) = 0$. Gdy $a_1 < x \leq a_2$, możemy napisać

$$\langle -\infty; x) = \langle -\infty; a_1) + \{a_1\} + (a_1; x)$$

skąd

$$P(\langle -\infty; x)) = P(\langle -\infty; a_1)) + P(\{a_1\}) + P\Big((a_1; x)\Big)$$

Ponieważ przedziały $\langle -\infty; a_1)$ i $(a_1; x)$ nie zawierają żadnego z punktów a_1, a_2, a_3, otrzymujemy

$$P(\langle -\infty; x)) = P(\{a_1\}) = p_1 = \frac{1}{6}$$

Postępując dalej analogicznie otrzymujemy

$$F(x) = \begin{cases} 0 & \text{dla} \quad x \leq a_1 \\ p_1 & \text{dla} \quad a_1 < x \leq a_2 \\ p_1 + p_2 & \text{dla} \quad a_2 < x \leq a_3 \\ 1 & \text{dla} \quad x > a_3 \end{cases}$$

czyli

$$F(x) = \begin{cases} 0 & \text{dla } x \leq -1 \\ \frac{1}{6} & \text{dla } -1 < x \leq 0 \\ \frac{1}{2} & \text{dla } 0 < x \leq 1 \\ 1 & \text{dla } x > 1 \end{cases}$$

§ 140. Prawdopodobieństwo warunkowe

Niech będzie dana przestrzeń probabilistyczna (X, S, P) i dowolne zdarzenie $B \in S$ spełniające warunek

$$(86) \qquad P(B) \neq 0$$

W danej przestrzeni probabilistycznej obok miary P używamy również miar względem zbiorów B. Miary te będziemy nazywać prawdopodobieństwami warunkowymi. Zgodnie z definicją podaną w § 122 mamy

$$(87) \qquad \bigwedge_{A \in S} P_B(A) \overset{\text{def}}{=} \frac{P(A \cap B)}{P(B)}$$

Prawdopodobieństwo warunkowe $P_B(A)$ nazywamy również prawdopodobieństwem warunkowym zdarzenia A pod warunkiem, że zaszło zdarzenie B. Jak wynika z rozważań § 122, prawdopodobieństwo warunkowe jest prawdopodobieństwem i wobec tego ma wszystkie własności prawdopodobieństwa.

Z rozważań § 122 wynika, że trójki (X, S, P_B) oraz $(B, S|B, P_B)$ są również przestrzeniami probabilistycznymi. Przestrzeń $(B, S|B, P_B)$ będziemy nazywali przestrzenią probabilistyczną zredukowaną do zbioru B.

Zauważmy, że w przestrzeni probabilistycznej (X, S, P_B) miara P_B jest rozkładem prawdopodobieństwa, którego nie będziemy nazywać warunkowym. Pojęcie prawdopodobieństwa warunkowego jest bowiem związane z wprowadzeniem w danej przestrzeni probabilistycznej (X, S, P) obok podstawowego rozkładu prawdopodobieństwa P innych miar, definiowanych za pośrednictwem miary P wzorem (87).

§ 141. Przykład

Na polowaniu oddano n strzałów, w tym strzelec B oddał m strzałów $(m>0)$ i trafił k razy $(k \leq m \leq n)$. Niech A oznacza zdarzenie, polegające na tym, że strzał był celny. Przyjmiemy za punkt wyjścia przestrzeń probabilistyczną skończoną o równych prawdopodobieństwach trafienia dla wszystkich strzałów. Mamy wtedy

$$P_B(A) = \frac{k}{m} = \frac{\frac{k}{n}}{\frac{m}{n}} = \frac{P(A \cap B)}{P(B)}$$

Istotnie, $P_B(A)$ oznacza prawdopodobieństwo celnego strzału pod warunkiem, że strzela strzelec B, czyli prawdopodobieństwo trafienia dla strzelca B, a zatem prawdopodobieństwo równe $\frac{k}{m}$. Następnie $P(A \cap B)$ oznacza prawdopodobieństwo, że strzał był celny i że strzelał strzelec B, a zatem prawdopodobieństwo równe $\frac{k}{n}$. Wreszcie $P(B)$ oznacza prawdopodobieństwo zdarzenia, polegającego na tym, że strzelał strzelec B, czyli prawdopodobieństwo równe $\frac{m}{n}$.

§ 142. Prawdopodobieństwo iloczynu zdarzeń

Ze wzoru (87) wynika, że

$$(88) \qquad \bigwedge_{\substack{A, B \in S \\ P(A) \neq 0 \\ P(B) \neq 0}} P(A \cap B) = P(A) \cdot P_A(B) = P(B) \cdot P_B(A)$$

Stosując ten wzór rekurencyjnie otrzymujemy

$$(89) \qquad \bigwedge_{\substack{A_1, \ldots, A_n \in S \\ P(A_1) \neq 0 \\ P(A_1 \cap A_2) \neq 0 \\ \vdots \\ P(A_1 \cap \ldots \cap A_{n-1}) \neq 0}} P(A_1 \cap \ldots \cap A_n) =$$

$$= P(A_1) \cdot P_{A_1}(A_2) \cdot P_{A_1 \cap A_2}(A_3), \ldots, P_{A_1 \cap \ldots \cap A_{n-1}}(A_n)$$

§ 143. Przykłady

I. W urnie jest 19 kul, w tym 8 białych i 11 czarnych. Jakie jest prawdopodobieństwo, że - nie znając położenia wzajemnego kul - wyciągniemy kolejno 3 kule białe:

a) w przypadku, gdy kul raz wyciągniętych nie zwracamy już do urny,

b) w przypadku, gdy każdą kulę po wyciągnięciu i obejrzeniu zwracamy do urny i przed wyciągnięciem następnej kuli zawartość urny dokładnie mieszamy.

Rozwiązanie:

Załóżmy, że kule są ponumerowane od 1 do 19. Niech kule białe mają numery od 1 do 8 a czarne od 9 do 19. Jako przestrzeń zdarzeń elementarnych X przyjmujemy zbiór wszystkich ciągów 3-elementowych złożonych z numerów kul możliwych do wyciągnięcia. W przypadku a) przestrzeń X składa się z $19 \cdot 18 \cdot 17 = 5814$ ciągów, w których numery nie mogą się powtarzać. Mamy bowiem 19 możliwości na obsadzenie pierwszego wyrazu ciągu, dla każdej z nich już tylko 18 możliwości na obsadzenie drugiego wyrazu ciągu, a dla każdej pary pierwszych dwu wyrazów ciągu już tylko 17 możliwości na obsadzenie trzeciego wyrazu ciągu. Możemy napisać

$$X = \left\{ (1,2,3),\ (1,2,4),\ \ldots,\ (19,18,17) \right\}$$

W przypadku b) przestrzeń X składa się z $19 \cdot 19 \cdot 19 = 6859$ ciągów, w których numery mogą się powtarzać, i możemy napisać

$$X = \left\{ (1,1,1),\ (1,1,2),\ \ldots,\ (19,19,19) \right\}$$

Niech A_j oznacza wyciągnięcie kuli białej za j-tym razem ($j = 1,2,3$). Zatem zdarzenie A_j składa się z tych ciągów, których j-ty wyraz jest numerem od 1 do 8. Interesuje nas prawdopodobieństwo zdarzenia

$$A = A_1 \cap A_2 \cap A_3 .$$

Na mocy (89) mamy

$$P(A) = P(A_1) \cdot P_{A_1}(A_2) \cdot P_{A_1 \cap A_2}(A_3).$$

Zakładając, że w danej skończonej przestrzeni probabilistycznej (X, S, P), gdzie S jest ciałem wszystkich podzbiorów przestrzeni X, prawdopodobieństwa dla wszystkich zdarzeń jednoelementowych są równe, otrzymujemy

w przypadku a): $P(A_1) = \dfrac{8 \cdot 18 \cdot 17}{19 \cdot 18 \cdot 17} = \dfrac{8}{19}$

w przypadku b): $P(A_1) = \dfrac{8 \cdot 19 \cdot 19}{19 \cdot 19 \cdot 19} = \dfrac{8}{19}$

Zakładając, że zdarzenie A_1 zaszło, tzn. za pierwszym razem wyciągnięto kulę białą, otrzymujemy dalej

w przypadku a): $P_{A_1}(A_2) = \dfrac{8 \cdot 7 \cdot 17}{8 \cdot 18 \cdot 17} = \dfrac{7}{18}$

w przypadku b): $P_{A_1}(A_2) = \frac{8 \cdot 8 \cdot 19}{8 \cdot 19 \cdot 19} = \frac{8}{19}$

Zakładając, że zaszły zdarzenia A_1 i A_2, tzn. że za 2 pierwszymi ciągnięciami wyciągnięto 2 białe kule, otrzymujemy

w przypadku a): $P_{A_1 \cap A_2}(A_3) = \frac{8 \cdot 7 \cdot 6}{8 \cdot 7 \cdot 17} = \frac{6}{17}$

w przypadku b): $P_{A_1 \cap A_2}(A_3) = \frac{8 \cdot 8 \cdot 8}{8 \cdot 8 \cdot 19} = \frac{8}{19}$

Ostatecznie otrzymujemy

w przypadku a): $P(A) = \frac{8}{19} \cdot \frac{7}{18} \cdot \frac{6}{17} = \frac{336}{5814} \approx 0,0578$

w przypadku b): $P(A) = \frac{8}{19} \cdot \frac{8}{19} \cdot \frac{8}{19} = \frac{512}{6859} \approx 0,0746$

Uwaga: Analizując sposób rozwiązania powyższego przykładu, zauważamy, że prawdopodobieństwa $P(A_1)$, $P_{A_1}(A_2)$, $P_{A_1 \cap A_2}(A_3)$ otrzymuje się również drogą następującego skróconego rozumowania:

W przypadku a):

Dla zdarzenia A_1 mamy 8 możliwości korzystnych na 19 możliwości ogółem, skąd

$$P(A_1) = \frac{8}{19}$$

Zakładając, że zdarzenie A_1 zaszło, pozostaje nam 18 kul, w tym 7 białych, skąd

$$P_{A_1}(A_2) = \frac{7}{18}$$

Wreszcie zakładając, że zaszły zdarzenia A_1 i A_2, pozostaje 17 kul, w tym 6 białych, skąd

$$P_{A_1 \cap A_2}(A_2) = \frac{6}{17}$$

W przypadku b) postępujemy analogicznie.

II. W urnie jest n kul $(n > 1)$, w tym b białych i $(n-b)$ czarnych. Ktoś wyjął z urny 1 kulę, ale nie wiemy jaką. Jakie jest teraz prawdopodobieństwo wyciągnięcia kuli białej?

Rozwiązanie:

Powołując się na analogię z przykładem poprzednim, oznaczamy przez:

A zdarzenie, polegające na tym, że zabrano kulę białą,
B zdarzenie, polegające na tym, że zabrano kulę czarną,
C zdarzenie, polegające na tym, że z pozostałych kul wyciągnie się kulę białą.

Mamy kolejno:

$$A + B = X$$

$$C = C \cap X = C \cap (A + B) = A \cap C + B \cap C$$

$$P(C) = P(A) \cdot P_A(C) + P(B) \cdot P_B(C)$$

Ponieważ nie wiemy, w jaki sposób została zabrana 1 kula, przyjmujemy, że z jednakowym prawdopodobieństwem mogła być zabrana każda z kul.

Stąd

$$P(A) = \frac{b}{n}, \qquad P(B) = \frac{n-b}{n}$$

a następnie

$$P_A(C) = \frac{b-1}{n-1}, \qquad P_B(C) = \frac{b}{n-1}$$

Zatem

$$P(C) = \frac{b}{n} \cdot \frac{b-1}{n-1} + \frac{n-b}{n} \cdot \frac{b}{n-1} = \frac{b}{n(n-1)}(b-1+n-b) = \frac{b}{n}$$

Widzimy, że informacja o zabraniu 1 kuli, ale nie wiadomo której, nie wpłynęła na zmianę prawdopodobieństwa wyciągnięcia kuli białej.

§ 144. Wzór na prawdopodobieństwo całkowite

Niech $A_1, \ldots, A_k$ (lub $A_1, A_2, \ldots$) tworzą pełny układ zdarzeń, skończony lub przeliczalny w przestrzeni probabilistycznej (X, S, P). Mamy zatem

$$A_1 + \ldots + A_k = X \qquad (\text{lub } A_1 + A_2 + \ldots = X)$$

Dla dowolnego zdarzenia $A \in S$ otrzymujemy

$$A = A \cap X = A \cap (A_1 + \ldots + A_k) = A_1 \cap A + \ldots + A_k \cap A$$

$$(\text{lub } A = A \cap X = A \cap (A_1 + A_2 + \ldots) = A_1 \cap A + A_2 \cap A + \ldots)$$

a stąd w przypadku skończonego pełnego układu zdarzeń $A_1, \ldots, A_k$

$$(90) \qquad P(A) = P(A_1)\cdot P_{A_1}(A) + \ldots + P(A_k)\cdot P_{A_k}(A)$$

W przypadku przeliczalnego pełnego układu zdarzeń $A_1, A_2, \ldots$ wzór (90) przyjmuje postać

$$(91) \qquad P(A) = P(A_1)\cdot P_{A_1}(A) + P(A_2)\cdot P_{A_2}(A) + \ldots$$

Wzory (90) i (91) noszą nazwę wzorów na prawdopodobieństwo całkowite.

§ 145. Wzór Bayesa

Obserwujemy zdarzenie B, dla którego $P(B) \neq 0$, w danej przestrzeni probabilistycznej (X, S, P). Przyczyną zdarzenia B mogły być zdarzenia $A_1, \ldots, A_n$ stanowiące pełny układ zdarzeń. Interesują nas prawdopodobieństwa, że przyczyną zdarzenia B było zdarzenie A_j $(j = 1, \ldots, n)$, tzn. prawdopodobieństwa warunkowe

$$P_B(A_j) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j)\cdot P_{A_j}(B)}{P(B)} \quad (j = 1, \ldots, n)$$

Wykorzystując wzór na prawdopodobieństwo całkowite (90) otrzymujemy

$$(92) \qquad P_B(A_j) = \frac{P(A_j)\cdot P_{A_j}(B)}{P(A_1)\cdot P_{A_1}(B) + \ldots + P(A_n)\, P_{A_n}(B)} \quad (j = 1, \ldots, n)$$

Wzór (92) nosi nazwę wzoru Bayesa. We wzorze tym prawdopodobieństwa $P(A_j)$ nazywamy prawdopodobieństwami a priori, a prawdopodobieństwa $P_B(A_j)$ prawdopodobieństwami a posteriori.

W przypadku, gdy nie mamy żadnych informacji o prawdopodobieństwach a priori, przyjmujemy

$$P(A_1) = \ldots = P(A_n) = \frac{1}{n}$$

Regułę tę nazywamy regułą Bayesa.

Wzór Bayesa ma liczne zastosowania, głównie w zagadnieniach diagnostycznych.

§ 146. Przykład

W magazynie jest 2500 żarówek, w tym 2000 z fabryki A_1, a 500 z fabryki A_2. Żarówki są tak wymieszane, że już nie wiadomo, które żarówki z której fabryki pochodzą. Fabryka A_1 wypuszcza średnio 2 żarówki wadliwe na 1000, a fabryka A_2 5 żarówek wadliwych na 1000. Wzięto z magazynu żarówkę i okazało się, że jest zła. Jakie jest prawdopodobieństwo, że żarówka pochodziła z fabryki A_1, a jakie, że pochodziła z fabryki A_2?

Rozwiązanie:

Niech B oznacza zdarzenie, że żarówka okazała się zła. Mamy

$$P(A_1) = \frac{2000}{2500} = 0,8 \qquad P(A_2) = \frac{500}{2500} = 0,2$$

$$P_{A_1}(B) = \frac{2}{1000} = 0,002 \qquad P_{A_2}(B) = \frac{5}{1000} = 0,005$$

Stosując wzór Bayesa otrzymujemy

$$P_B(A_1) = \frac{0,8 \cdot 0,002}{0,8 \cdot 0,002 + 0,2 \cdot 0,005} = \frac{8}{13}$$

$$P_B(A_2) = \frac{0,2 \cdot 0,005}{0,8 \cdot 0,002 + 0,2 \cdot 0,005} = \frac{5}{13}$$

W danym przypadku można było znaleźć prawdopodobieństwo $P_B(A_2)$ ze wzoru

$$P_B(A_2) = 1 - P_B(A_1)$$

§ 147. Zdarzenia niezależne

Niech będzie dana przestrzeń probabilistyczna (X, S, P) i dwa zdarzenia $A, B \in S$ takie, że $P(A) \neq 0$ i $P(B) \neq 0$.

Intuicyjnie wyczuwamy, że zdarzenie A nazwalibyśmy niezależnym od zdarzenia B, gdyby

$$P(A) = P_B(A)$$

Ale wtedy mielibyśmy

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

czyli