Wykład 1

Elementy rachunku prawdopodobieństwa

1.1. Zdarzenie i zmienne losowe. Definicja prawdopodobieństwa

Zdarzeniem losowym lub przypadkowym będziemy nazywać takie zdarzenie B, które z góry nie da się przewidzieć. Na przykład jeżeli rozważamy wyniki jednoczesnego rzucenia N razy dwóch jednakowych monet, to wyniki całego takiego doświadczenia, jak i wyniki pojedynczych rzuceń będą zupełnie nieprzewidzialne. Wynikiem pojedynczego rzutu będzie jedno z możliwych zdarzeń: RR, RO, OR albo OO, gdzie R oznacza wyrzucenie reszki, a O - wyrzucenie orla. Każdemu możliwemu zdarzeniu elementarnemu B możemy przypisać pewną liczbę: $x_1 \to RR$; $x_2 \to RO$; $x_3 \to OR$; $x_4 \to OO$. W ciągu z N zdarzeń każde z tych elementarnych zdarzeń x_i zostaje rozmieszczone w sposób przypadkowy.

Niech w wyniku N - krotnego powtórzenia doświadczenia w jednakowych warunkach otrzymaliśmy w N_i przypadkach zdarzenie x_i. Prawdopodobieństwem $P(x_i)$ zdarzenia przypadkowego x_i nazywa się wielkość

$$P(x_i) = \lim_{N \to \infty} \frac{N_i}{N} \; . \tag{1.1}$$

Przypuśćmy, że zbiór wszystkich możliwych zdarzeń jest reprezentowany przez zbiór $\{x\}$, i każdej z możliwych wartości x_i tego zbioru wiadomo jest odpowiednie prawdopodobieństwo $P(x_i)$. Zbiór możliwych zdarzeń $\{x\}$ oraz zbiór odpowiednich prawdopodobieństw $\{P(x)\}$ określają wielkość, która nazywa się zmienną losową x. Jeżeli zbiór $\{x\}$ jest dyskretny, to zmienna losowa nazywa się dyskretną. Jeżeli możliwe wartości zmiennej losowej znajdują się w przedziale pewnego kontinuum $(a \le x \le b)$, zmienna losową nazywamy ciągłą.

Rachunek prawdopodobieństw opiera się na trzech aksjomatach:

1. Prawdopodobieństwo $P(x_i)$ zdarzenia x_i spełnia nierówność $P(x_i) \ge 0$.
2. Jeżeli wiemy, że zdarzenie x_i musi zajść, to $P(x_i) = 1$.
3. Dla zdarzeń x_i ($i = 1,2,\ldots,n$) wzajemnie wykluczających się, czyli dla zdarzeń dla których wiemy, że zachodzenie jednego ze zdarzeń jednoznacznie gwarantuje, że

pozostałe zdarzenia nie zachodzą, prawdopodobieństwo $P(R)$ tego, że chociażby jedno ze zdarzeń x_i zajdzie określa wzór

$$P(R) = \sum_{i=1}^{n} P(x_i).$$

Na podstawie tego aksjomatu, prawdopodobieństwo tego, że w wyniku monety chociażby na jednej z monet zostaje wyrzucona reszka wynosi

$$P(R) = P(x_1) + P(x_2) + P(x_3).$$

Tu $P(x_1)$ jest prawdopodobieństwem zdarzenia $x_1 = RR$, $P(x_2)$ jest prawdopodobieństwem zdarzenia $x_2 = RO$, a $P(x_3)$ jest prawdopodobieństwem zdarzenia $x_3 = OR$.

Z drugiego aksjomatu wynika, że prawdopodobieństwa $P(x_i)$, spełniają równanie

$$\sum_{i=1}^{N} P(x_i) = 1. \tag{1.2}$$

Dla zmiennej losowej ciągłej postuluje się, że istnieje taka funkcja $p(x)$, że

$$P(a \leq x \leq b) = \int_a^b p(x)dx. \tag{1.3}$$

Tu $P(a \leq x \leq b)$ jest prawdopodobieństwem, iż wartości zmiennej losowej x znajdują się w przedziale (a,b). Wielkość $p(x)$ nazywa się gęstością prawdopodobieństwa. Na podstawie drugiego aksjomatu znajdujemy, że gęstość prawdopodobieństwa spełnia warunek

$$\int_{-\infty}^{\infty} p(x)dx = 1. \tag{1.4}$$

Wzory (1.3) i (1.4) będą również słuszne w przypadku zmiennej losowej dyskretnej, jeżeli, korzystając z właściwości funkcji δ - Diraca, wprowadźmy następującą funkcję gęstości prawdopodobieństwa

$$p(x) = \sum_i P(x_i)\delta(x - x_i). \tag{1.5}$$

1.2. Średnie statystyczne wartości. Funkcja charakterystyczna. Momenty

Średnie statystyczną wartością lub wartością oczekiwaną funkcji zmiennej losowej $f(x)$ nazywamy wielkość

$$\overline{f(x)} = \int_{-\infty}^{\infty} f(x)p(x)dx \quad . \tag{1.6}$$

Po podstawieniu do (1.6) wzoru (1.5), dla zmiennej losowej dyskretnej otrzymujemy

$$\overline{f(x)} = \sum_i P(x_i)f(x_i) \quad . \tag{1.7}$$

Średnie statystyczna wartość wielkości x^n

$$C_n \equiv \overline{x^n} = \int_{-\infty}^{\infty} x^n p(x)dx \quad , \tag{1.8a}$$

albo

$$C_n \equiv \overline{x^n} = \sum_i P(x_i)\cdot x_i^n \tag{1.8b}$$

nazywa się momentem n - tego rzędu zmiennej losowej.

Szczególnie ważną rolę w praktyce odgrywają momenty centralne

$$c_n \equiv \overline{\left(x - \bar{x}\right)^n} = \int_{-\infty}^{\infty} \left(x - \bar{x}\right)^n p(x)dx \quad , \tag{1.9}$$

które charakteryzują fluktuacje zmiennej losowej wokół jej wartości średniej $\bar{x}$.

Z określenia (1.9) łatwo otrzymać związki pomiędzy momentami zwykłymi C_n i momentami centralnymi c_n. Na przykład

$$c_1 = 0, \tag{1.10a}$$

$$c_2 = C_2 - C_1^2 \equiv \sigma^2(x), \tag{1.10b}$$

Drugi moment centralny $c_2 \equiv \sigma^2(x)$ nazywa się wariancją, a $\sigma(x)$ - pierwiastek kwadratowy z wariancji, - nazywa się odchyleniem standardowym.

Funkcja charakterystyczna $P(\omega)$ zmiennej losowej x jest określona jako wartość oczekiwana funkcji $\exp(i\omega\cdot x)$

$$P(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} p(x)dx \quad . \tag{1.11}$$

Funkcja charakterystyczna $P(\omega)$ jest więc transformatą Fouriera funkcji gęstości prawdopodobieństwa zmiennej losowej x. Ze wzoru (1.11) wynika, że

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega x} P(\omega) d\omega \quad , \tag{1.12}$$

a zatem funkcja charakterystyczna $P(\omega)$ zawiera taką samą informację o statystycznych własnościach zmiennej losowej x, co i funkcja gęstości prawdopodobieństwa $p(x)$.

Korzystając z możliwości rozwinięcia funkcji $\exp(i\omega x)$ w szereg potęgowy

$$\exp(i\omega x) = \sum_{n=0}^{\infty} \frac{(i\omega)^n}{n!} x^n , \tag{1.13}$$

ze wzoru (1.11) otrzymujemy

$$P(\omega) = \sum_{n=0}^{\infty} \frac{(i\omega)^n}{n!} \overline{x^n} . \tag{1.14}$$

Ze wzoru (1.14) wynika, że momenty zmiennej losowej możemy również obliczyć z funkcji charakterystycznej $P(\omega)$

$$\overline{x^n} = \frac{1}{i^n} \frac{d^n P(\omega)}{d\omega^n} \bigg|_{\omega = 0} . \tag{1.15}$$

Należy zauważyć, że nie zawsze funkcja charakterystyczna jest analityczna w punkcie $\omega = 0$. Na przykład dla tak zwanego rozkładu Cauchy'ego, który opisuje funkcja Lorentza

$$p(x) = \frac{1}{\pi \left(1 + x^2\right)} ,$$

funkcja charakterystyczna ma postać

$$P(\omega) = e^{-|\omega|} ,$$

a zatem nie jest ona funkcją analityczną w punkcie $\omega = 0$.

1.3. Łączny rozkład dwóch zmiennych losowych

Rozważmy eksperyment z rzucaniem dwóch monet. Jedną monetę będziemy rzucali prawą ręką, a drugą – lewą. Wyniki doświadczenia dla dwóch monet tworzą dwa zbiory $\{x\}$ i $\{y\}$ z odpowiednimi zbiorami prawdopodobieństw $\{P(x)\}$ i $\{P(y)\}$. Z dwóch zbiorów zdarzeń $\{x\}$ i $\{y\}$ możemy utworzyć nowy zbiór zdarzeń łącznych $\{x \times y\}$, utworzony z par zdarzeń po jednemu z każdego zbioru. Niech w wyniku N- krotnego powtórzenia doświadczenia, czyli po rzucaniu dwóch monet N razy w jednakowych warunkach, otrzymaliśmy w N_i przypadkach zdarzenie x_i, w N_k przypadkach – zdarzenie y_k, a w N_{ik} przypadkach zaszło jednocześnie dwa zdarzenia x_i i y_k. Prawdopodobieństwem $P(x_i, y_k)$ zdarzenia łącznego x_i i y_k nazywa się wielkość

$$P(x_i, y_k) = \lim_{N \to \infty} \frac{N_{ik}}{N} \quad . \tag{1.16}$$

Prawdopodobieństwo zdarzenia łącznego $P(x_i, y_k)$, zgodnie z pierwszym aksjomatem rachunku prawdopodobieństwa, musi spełniać warunek

$$\sum_i \sum_k P(x_i, y_k) = 1 \, . \tag{1.17}$$

Jeżeli rozważmy wszystkie łączne zdarzenia (x_i, y_k) dla których zdarzenie x_i jest takie same, ale zdarzenia y_k są różne, to liczba takich zdarzeń wynosi N_i. Stąd otrzymujemy, że prawdopodobieństwo zdarzenia łącznego $P(x_i, y_k)$ musi również spełniać równania

$$P(x_i) = \sum_k P(x_i, y_k) \, , \tag{1.18a}$$

$$P(y_k) = \sum_i P(x_i, y_k) \, . \tag{1.18b}$$

Prawdopodobieństwem warunkowym $P(x_i \mid y_k)$ zdarzenia x_i nazywa się wielkość

$$P(x_i \mid y_k) = \lim_{N \to \infty} \frac{N_{ik}}{N_k} \quad . \tag{1.19a}$$

Prawdopodobieństwo warunkowo $P(x_i \mid y_k)$ określa prawdopodobieństwo tego, że w łącznym zdarzeniu x_i i y_k zajdzie zdarzenie x_i, przy spełnieniu warunku, że zdarzenie ze zbioru $\{y\}$ jest zderzeniem y_k.

W podobny sposób prawdopodobieństwem warunkowym $P(y_k \mid x_i)$ zdarzenia y_k nazywa się wielkość

$$P(y_k \mid x_i) = \lim_{N \to \infty} \frac{N_{ik}}{N_i} \quad . \tag{1.19b}$$

Prawdopodobieństwo warunkowo $P(y_k \mid x_i)$ określa prawdopodobieństwo tego, że w łącznym zdarzeniu x_i i y_k zajdzie zdarzenie y_k, przy spełnieniu warunku, że zdarzenie ze zbioru $\{x\}$ jest zderzeniem x_i.

Ponieważ

$$\frac{N_{ik}}{N} = \frac{N_{ik}}{N_i}\frac{N_i}{N} = \frac{N_{ik}}{N_k}\frac{N_k}{N},$$

ze wzorów (1.16) i (1.19) otrzymujemy wzór Bayesa

$$P(x_i, y_k) = P(x_i)P(y_k \mid x_i) = P(y_k)P(x_i \mid y_k) \ . \tag{1.20}$$

Zgodnie z (1.17) i (1.2) prawdopodobieństwa warunkowe muszą spełniać równość

$$\sum_k P(y_k \mid x_i) = \sum_i P(x_i \mid y_k) = 1 \ . \tag{1.21}$$

Dwie zmienne losowe x i y nazywamy statystycznie niezależnymi, jeżeli

$$N_{ik} = \frac{N_i N_k}{N} \ . \tag{1.22}$$

Wtedy ze wzorów (1.19a) i (1.19b) znajdujemy

$$P(x_i \mid y_k) = \lim_{N\to\infty}\frac{N_{ik}}{N_k} = \lim_{N\to\infty}\frac{N_i}{N} = P(x_i) \ . \tag{1.23a}$$

$$P(y_k \mid x_i) = \lim_{N\to\infty}\frac{N_{ik}}{N_i} = \lim_{N\to\infty}\frac{N_k}{N} = P(y_k) \ . \tag{1.23b}$$

Ze wzorów (1.23) wynika, że dla zmiennych statystycznie niezależnych znajomość wartości jednej ze zmiennych nie wpływa na prawdopodobieństwo drugiej zmiennej.

Dla zmiennych niezależnych statystycznie ze wzoru Bayesa (1.20) mamy

$$P(x_i, y_k) = P(x_i)P(y_k) \ . \tag{1.24}$$

W przypadku ciągłych zmiennych losowych x i y łączny rozkład prawdopodobieństwa charakteryzuje łączna funkcja gęstości prawdopodobieństwa $p(x, y)$. Zakłada się, że

$$\int_{-\infty}^{\infty} p(x, y)dxdy = 1 \ . \tag{1.25}$$

W tym przypadku, korzystając ze wzoru Bayesa (1.20), warunkowe funkcje gęstości prawdopodobieństwa definiujemy jako

$$p(y \mid x) = \frac{p(x,y)}{p(x)}, \tag{1.26a}$$

$$p(x \mid y) = \frac{p(x,y)}{p(y)}. \tag{1.26b}$$

Dla łącznego rozkładu dwóch zmiennych losowych momenty mieszane zmiennych x i y określamy wzorem

$$\overline{x^n y^m} = \int_{-\infty}^{\infty} x^n y^m p(x,y) dxdy \ . \tag{1.27}$$

Szczególnie ważnymi momentami dwóch zmiennych losowych są korelacja, kowariancja i współczynnik korelacji zmiennych x i y.

Korelację (albo moment korelacyjny) określa wzór

$$K(x,y) = \overline{xy} = \int_{-\infty}^{\infty} (xy) p(x,y) dxdy \ . \tag{1.28}$$

Kowariancja jest określona wzorem

$$c(x,y) = \overline{(x - \overline{x})(y - \overline{y})} = K(x,y) - \overline{x} \cdot \overline{y} \ . \tag{1.29}$$

Natomiast współczynnikiem korelacji jest wielkość

$$\rho = \frac{c(x,y)}{\sigma(x)\sigma(y)} \ . \tag{1.30}$$

Tu $\sigma(x)$ jest odchyleniem standardowym zmiennej x, a $\sigma(y)$ jest odchyleniem standardowym zmiennej y.

Zgodnie z (1.24) dla zmiennych niezależnych statystycznie

$$K(x,y) = \overline{x} \cdot \overline{y}, \tag{1.31}$$

a zatem współczynnik korelacji $\rho = 0$.

Dla rozkładu łącznego dwóch zmiennych losowych x i y funkcję charakterystyczną $P(\omega_1, \omega_2)$ określamy wzorem

$$P(\omega_1,\omega_2) = \int\limits_{-\infty}^{\infty} \exp[i(\omega_1 x + \omega_2 y)] \cdot p(x,y)dxdy. \tag{1.32}$$

Zgodnie z (1.24), dla zmiennych niezależnych statystycznie

$$P(\omega_1,\omega_2) = P(\omega_1) \cdot P(\omega_2). \tag{1.33}$$

Momenty mieszane zmiennych losowych x i możemy obliczyć przez funkcję charakterystyczną jako

$$\overline{x^n y^m} = \frac{1}{(i)^{n+m}} \frac{\partial^{n+m}}{\partial \omega_1^n \partial \omega_2^m} P(\omega_1,\omega_2)\Bigg|_{\omega_1=\omega_2=0}. \tag{1.34}$$

2. Przykłady rozkładów prawdopodobieństwa

2.1. Rozkład dwumianowy. Błądzenie przypadkowe

Rozważmy cząstkę, która może się poruszać krokami o długości a wzdłuż linii prostej. Niech p jest prawdopodobieństwem wykonania kroku w prawo, a q jest prawdopodobieństwem wykonania kroku w lewo. Zakładając, że każdy krok jest niezależny od pozostałych, znajdziemy prawdopodobieństwo tego, że spośród wykonanych N kroków n kroków będą wykonane w prawą stronę, a $N-n$ kroków – w lewą stronę. Wypadkowe przemieszczenie w prawą stronę będzie wtedy wynosiło $(n-(N-n))a = (2n-N)a$.

Oznaczmy przez P krok w prawo, a przez L - krok w lewo. Ponieważ każdy krok jest niezależny od pozostałych, prawdopodobieństwo tego, że w danej sekwencji z N kroków

$$\underbrace{PLLP\cdots LPLL}_{N}, \tag{1.35}$$

będzie n kroków w prawo i ($N-n$) kroków w lewo wynosi

$$pqqp\cdots qpqq = p^n q^{N-n}. \tag{1.36}$$

Prawdopodobieństwo (1.36) nie zależy od tego w jakiej kolejności pojawiły się w sekwencji (1.35) n kroków w prawo. Liczba sekwencji typu (1.35), zawierających n kroków w prawo i ($N-n$) kroków w lewo, jest równa liczbie kombinacji n elementów spośród N

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!}. \tag{1.37}$$

Korzystając z reguły sumowania niezależnych prawdopodobieństw (z trzeciego aksjomatu rachunku prawdopodobieństwa) otrzymujemy

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \ . \tag{1.38}$$

Rozkład prawdopodobieństwa (1.38) nazywa się dwumianowym rozkładem.

2.2. Rozkład Poissona

Rozkład Poissona otrzymuje się z rozkładu dwumianowego jako przypadek graniczny przy $N \to \infty$ i $p \to 0$, tak że $Np = const = a$.

Podstawiając do wzoru (1.38) zamiast p wielkość a/N otrzymujemy

$$P_N(n) = \frac{N(N-1)\cdots[N-(n-1)]}{n!}\left(\frac{a}{N}\right)^n\left(1-\frac{a}{N}\right)^{N-n} =$$

$$= \frac{a^n}{n!}\cdot 1\cdot\left(1-\frac{1}{N}\right)\cdot\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)\cdot\left(1-\frac{a}{N}\right)^{N-n} \equiv \frac{a^n}{n!}e^{-a} \ . \tag{1.39}$$

Ostatni człon w (1.39) otrzymaliśmy, korzystając ze wzoru

$$\lim_{N\to\infty}\left(1-\frac{a}{N}\right)^N \cong e^{-a} \ . \tag{1.40}$$

Biorąc pod uwagę wzór (1.39), dla funkcji charakterystycznej rozkładu Poissona znajdujemy

$$P_N(\omega) = \sum_{n=0}^{\infty} e^{in\omega} P_N(n) = e^{-a}\sum_{n=0}^{\infty}\frac{\left(ae^{i\omega}\right)^n}{n!} = \exp\left\{a\left(e^{i\omega}-1\right)\right\} \ . \tag{1.41}$$

Ze wzoru (1.15) dla pierwszych dwóch momentów rozkładu Poissona otrzymujemy

$$\bar{n} = \frac{1}{i}\frac{dP_N(\omega)}{d\omega}\bigg|_{\omega=0} = a = Np \ , \tag{1.42}$$

$$\overline{n^2} = -\frac{dP^2{}_N(\omega)}{d\omega^2}\bigg|_{\omega=0} = a^2 + a = \left(\bar{n}\right)^2 + \bar{n} \ . \tag{1.43}$$

Stąd odchylenie standardowe dla rozkładu Poissona wynosi

$$\sigma_N = \sqrt{\overline{n^2} - \bar{n}^2} = \sqrt{\bar{n}} = \sqrt{Np} \ . \tag{1.44}$$

Z uwzględnieniem (1.42) rozkład Poissona możemy zapisać w postaci

$$P_N(n) = \frac{\left(\bar{n}\right)^n}{n!} e^{-\bar{n}} \quad . \tag{1.45}$$

2.3. Rozkład Gaussa

Rozkład Gaussa otrzymuje się jako inny przypadek graniczny rozkładu dwumianowego albo rozkładu Poissona gdy $N \to \infty$ lecz $p = const$, a więc $\bar{n} = Np \to \infty$. Udowodnimy to, korzystając ze wzoru Stirllinga

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\left(1 + \frac{1}{12n} + \cdots\right) \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad . \tag{1.46}$$

Biorąc pod uwagę ten wzór, ze wzoru (1.45) otrzymujemy

$$P_N(n) = \frac{\left(\bar{n}\right)^n}{n!} e^{-\bar{n}} \approx \frac{1}{\sqrt{2\pi n}}\left(\frac{\bar{n}}{n}\right)^n e^{-\bar{n}+n} = \frac{1}{\sqrt{2\pi n}} \exp\left\{n - \bar{n} - n \ln\left(\frac{n}{\bar{n}}\right)\right\} \quad . \tag{1.47}$$

Oznaczając

$$x + 1 = \frac{n}{\bar{n}}, \tag{1.48}$$

zapiszmy wzór (1.47) w postaci

$$P_N(n) = \frac{1}{\sqrt{2\pi \bar{n}(1+x)}} \exp\left\{\bar{n}\left[x - (1+x)\ln(1+x)\right]\right\} \quad . \tag{1.49}$$

Korzystając z rozwinięcia funkcji $\ln(1+x)$

$$\ln(1+x) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{x^m}{m} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots, \tag{1.50}$$

otrzymujemy

$$P_N(n) = \frac{1}{\sqrt{2\pi \bar{n}(1+x)}} \exp\left\{\bar{n}\left[x - (1+x)\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \right)\right]\right\} =$$

$$= \frac{1}{\sqrt{2\pi \bar{n}(1+x)}} \exp\left\{\bar{n}\left[x - x + \frac{x^2}{2} - \frac{x^3}{3} + \frac{x^4}{4} - \cdots - x^2 + \frac{x^3}{2} - \frac{x^4}{3} + \cdots\right]\right\} =$$

$$= \frac{1}{\sqrt{2\pi \bar{n}(1+x)}} \exp\left\{\bar{n}\left[-\frac{x^2}{1\cdot 2} + \frac{x^3}{2\cdot 3} - \frac{x^4}{3\cdot 4} - \cdots\right]\right\} =$$

$$= \frac{1}{\sqrt{2\pi \bar{n}(1+x)}} \exp\left\{-\bar{n}\frac{x^2}{2}\left[1-\frac{x}{3}+\frac{x^2}{2\cdot 3}-\cdots\right]\right\}. \tag{1.51}$$

Gdy $n \to \bar{n}$, a zatem $x \to 0$, ze wzoru (1.51) mamy

$$P(n) = \frac{1}{\sqrt{2\pi \bar{n}}} \exp\left(-\bar{n}\frac{x^2}{2}\right) \equiv \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(n-\bar{n})^2}{2\sigma^2}\right]. \tag{1.52}$$

Tu $\sigma^2 = \bar{n}$ jest wariancją.

Rozkład prawdopodobieństwa (1.52) nazywa się rozkładem Gaussa lub rozkładem normalnym.