

## Wykład 2

### 2.1. Gęstość prawdopodobieństwa funkcji zmiennej losowej

W zastosowaniach praktycznych często spotykamy się z funkcjami zmiennych losowych

$$z = f(x). \quad (2.1)$$

W tym przypadku zmienna  $z$  również staje się zmienna losową i problem sprowadza się do znalezienia gęstości prawdopodobieństwa  $p(z)$  zmiennej  $z$ .

Zgodnie z (1.6) średnia wartość statystyczna funkcji (2.1) wynosi

$$\bar{z} \equiv \overline{f(x)} = \int_{-\infty}^{\infty} f(x)p(x)dx = \int_{-\infty}^{\infty} zp(z)dz. \quad (2.2)$$

We wzorze (2.2)  $p(z)$  jest poszukiwaną gęstością prawdopodobieństwa zmiennej  $z$ .

Przypuśćmy, że istnieje przekształcenie odwrotne do przekształcenia (2.1)

$$x = f^{-1}(z). \quad (2.3)$$

Wtedy dla nieskończenie małych  $dx$  i  $dz$  możemy zapisać

$$dx = \left| \frac{dx}{dz} \right| dz. \quad (2.4)$$

Podstawiając (2.1) i (2.4) do pierwszej całki w (2.2) znajdujemy

$$\bar{z} = \int_{-\infty}^{\infty} zp[f^{-1}(z)] \left| \frac{dx}{dz} \right| dz = \int_{-\infty}^{\infty} zp(z)dz. \quad (2.5)$$

Skąd dla poszukiwanej gęstości prawdopodobieństwa zmienne  $z$  otrzymujemy

$$p(z) = p[f^{-1}(z)] \left| \frac{df^{-1}(z)}{dz} \right|. \quad (2.6)$$

Jako przykład zastosowania (2.6) rozważmy funkcję

$$z = f(x) = z_0 e^{x/a}, \quad (2.7)$$

gdzie  $z_0$  jest stałą, i przypuśćmy, że rozkład zmiennej losowej  $x$  jest rozkładem normalnym

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2}\right]. \quad (2.8)$$

Przykładem funkcji (2.8) w spektroskopii NMR jest czas korelacji  $\tau_c = \tau_0 \exp(U/RT)$ , który, w przypadku przypadkowego rozrzutu energii aktywacji  $U$ , staje się funkcją losową zmiennej  $U$ .

Ze wzoru (2.7) mamy

$$x = f^{-1}(z) = a \ln\left(\frac{z}{z_0}\right), \quad (2.9)$$

Różniczkując (2.9) względem  $z$  znajdujemy

$$\frac{dx}{dz} \equiv \frac{df^{-1}(z)}{dz} = \frac{a}{z}. \quad (2.10)$$

Po podstawieniu (2.9) do wzoru (2.8) mamy

$$p[f^{-1}(z)] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{a^2(\ln z - \overline{\ln z})^2}{2\sigma^2}\right]. \quad (2.11)$$

Ze wzoru (2.6) z uwzględnieniem (2.10) znajdujemy

$$\begin{aligned} p(z) &= p[f^{-1}(z)] \left| \frac{df^{-1}(z)}{dz} \right| = \\ &= \frac{1}{z\sqrt{2\pi}\sigma_z} \exp\left[-\frac{(\ln z - \overline{\ln z})^2}{2\sigma_z^2}\right]. \end{aligned} \quad (2.12)$$

Tu  $\sigma_z = \sigma/a$ .

Oznaczając

$$Z = \ln z - \overline{\ln z},$$

oraz biorąc pod uwagę, że

$$dZ = \frac{1}{z} dz,$$

ze wzoru (2.12) otrzymujemy

$$p(Z) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left[-\frac{Z^2}{2\sigma_z^2}\right]. \quad (2.13)$$

Rozkład prawdopodobieństwa (2.13) nosi nazwę logarytmicznie-normalnego (albo krótko - lognormalnego) rozkładu.

## 2.2. Centralne twierdzenie graniczne

Centralne twierdzenie graniczne jest jednym z najważniejszych twierdzeń rachunku prawdopodobieństwa, uzasadniające częste występowanie w przyrodzie rozkładów prawdopodobieństw zmiennych losowych zbliżonych do rozkładu normalnego. Na podstawie centralnego twierdzenia granicznego można w wielu sytuacjach zakładać, że zmienna losowa, za pomocą której modelujemy dane zjawisko, ma rozkład bardzo zbliżony do rozkładu normalnego. Centralne twierdzenie graniczne to twierdzenie matematyczne, zgodnie z którym, jeżeli mamy zbiór niezależnych zmiennych losowych  $x_1, x_2, \dots, x_n$  o dowolnych (niekoniecznie takich samych) rozkładach prawdopodobieństwa i o średnich oraz wariancjach odpowiednio  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$  i  $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ , to funkcja gęstości prawdopodobieństwa  $p(z)$ , zmiennej losowej  $z$ , zdefiniowaną równością

$$z = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{x_i - \bar{x}_i}{\sigma_i}, \quad (2.14)$$

dla  $n$  dążącego do nieskończoności dąży do rozkładu normalnego

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}. \quad (2.15)$$

## Procesy stochastyczne

Przypomnijmy, że zmienną losową  $X$  nazywamy zbiór możliwych zdarzeń  $\{x\}$  oraz zbiór odpowiednich prawdopodobieństw  $\{P(x)\}$  tych zdarzeń. Pojęcie procesu stochastycznego jest uogólnieniem pojęcia zmiennej losowej. W przypadku procesu stochastycznego nieprzewidywalnymi są funkcje (zwykle czasu albo zmiennych przestrzennych), a nie liczby. W przypadku obserwacji zmiennej losowej w pewnym przedziale czasu (albo w określonym punkcie przestrzennym) będziemy rejestrowali konkretny przebieg zdarzeń losowych  $x_1, x_2, \dots, x_N$ , który nazywa się realizacją procesu stochastycznego. Powtarzając obserwacji zmiennej losowej wielokrotnie stwierdzimy, że za każdym razem będziemy mieli jaką inną kolejność zdarzeń, czyli inną realizację. Zbiór wszystkich realizacji nazywa się procesem stochastycznym. Proces stochastyczny jest więc funkcją rzeczywistą  $f(t, X(t))$  dwu zmiennych: zmiennej  $t$ , mającej sens czasu, oraz zmiennej losowej  $X(t)$ , mającej sens zdarzenia elementarnego. Proces stochastyczny nazywa

się skokowym lub ciągłym, gdy rozkład  $f(t, X(t))$  jest odpowiednio skokowy, lub ciągły dla każdego skończonego zbioru  $t_1, t_2, \dots, t_N$ .

Aby opisać proces stochastyczny, musimy określić tak zwane rozkłady prawdopodobieństwa.

### 2.3. Rozkłady prawdopodobieństwa. Średnie po zbiorze realizacji

Rozkłady prawdopodobieństwa są opisywane odpowiednio rozkładami rzędu pierwszego, drugiego, trzeciego itd.

$$P(x_1, t_1), \quad (2.16)$$

$$P(x_1, t_1; x_2, t_2), \quad (2.17)$$

.....

$$P(x_1, t_1; x_2, t_2; \dots; x_N, t_N). \quad (2.18)$$

Funkcja (2.16) określa prawdopodobieństwo tego, że w chwili  $t_1$  zmienna losowa  $X(t)$  przyjmuje wartość  $x_1$ . Rozkład (2.17) określa prawdopodobieństwo tego, że w chwili  $t_1$  zmienna losowa  $X(t)$  przyjmuje wartość  $x_1$ , a w chwili  $t_2$  ( $t_2 > t_1$ ) - wartość  $x_2$ . Funkcja (2.18) określa prawdopodobieństwo tego, że w chwili  $t_1$  zmienna losowa  $X(t)$  przyjmuje wartość  $x_1$ , w chwili  $t_2$  ( $t_2 > t_1$ ) - wartość  $x_2$ , ....., w chwili  $t_N$  ( $t_N > t_{N-1} > \dots > t_2 > t_1$ ) - wartość  $x_N$ .

Z definicji rozkładów  $P(x_1, t_1; x_2, t_2; \dots; x_N, t_N)$  wynika, że

$$P(x_1, t_1; x_2, t_2; \dots; x_{n-1}, t_{n-1}) = \sum_{x_n} P(x_1, t_1; x_2, t_2; \dots; x_N, t_N). \quad (2.19)$$

Rozkłady  $P(x_1, t_1; x_2, t_2; \dots; x_N, t_N)$  zawierają całkowitą informację o procesie stochastycznym.

Zauważmy, że

$$\begin{aligned} \sum_{x_1} P(x_1, t_1) &= 1, \\ \sum_{x_1} \sum_{x_2} P(x_1, t_1; x_2, t_2) &= 1, \\ &\dots \\ \sum_{x_1} \sum_{x_2} \dots \sum_{x_n} P(x_1, t_1; x_2, t_2; \dots; x_N, t_N) &= 1. \end{aligned} \quad (2.20)$$

Jeżeli wartości zmiennych losowych  $x_1, x_2, \dots, x_N$ , które rejestrują się w chwilach  $t_1, t_2, \dots, t_N$ , są niezależne od siebie, to rozkład prawdopodobieństwa  $P(x_1, t_1; x_2, t_2; \dots; x_N, t_N)$  przyjmuje postać

$$P(x_1, t_1; x_2, t_2; \dots; x_N, t_N) = P(x_1, t_1)P(x_2, t_2) \cdots P(x_N, t_N), \quad (2.21)$$

gdzie  $P(x_k, t_k)$  - prawdopodobieństwo tego, że w chwili  $t_k$  losowa zmienna  $x$  przyjmuje wartość  $x_k$ .

Jeżeli proces stochastyczny opisuje rozkład prawdopodobieństwa (2.21), to mówimy, że zmienne losowe  $x_1, x_2, \dots, x_N$  są nieskorelowane.

Dla funkcji  $f(x_1, t_1; x_2, t_2; \dots; x_N, t_N)$  procesu stochastycznego wzór

$$\langle f(t = t_N) \rangle = \lim_{\substack{N \rightarrow \infty \\ t_i - t_{i-1} \rightarrow 0}} \sum_{x_1} \cdots \sum_{x_N} f(x_1, t_1; x_2, t_2; \dots; x_N, t_N) P(x_1, t_1; x_2, t_2; \dots; x_N, t_N = t) \quad (2.22)$$

określa średnią wartość funkcji  $f(x_1, t_1; x_2, t_2; \dots; x_N, t_N)$  w chwili  $t$  po zbiorze realizacji stochastycznego procesu.

Mówimy, że proces stochastyczny jest rzędu  $N$ , jeżeli on całkowicie jest określony przez swój  $N$ -ty rozkład prawdopodobieństwa  $P(x_1, t_1; x_2, t_2; \dots; x_N, t_N)$ , tj. przez rozkład rzędu  $N$ . Ze wzoru (2.21) wynika, że nieskorelowany stochastyczny proces jest procesem pierwszego rzędu.

## 2.4. Warunkowe rozkłady prawdopodobieństwa

Pojęcie prawdopodobieństwa warunkowego jest jednym z podstawowych narzędzi teorii procesów stochastycznych. Warunkowe rozkłady prawdopodobieństwa dla procesu stochastycznego definiuje wzór

$$\begin{aligned} P(x_1, t_1; x_2, t_2; \dots; x_m, t_m | x_{m+1}, t_{m+1}; \dots; x_N, t_N) = \\ = \frac{P(x_1, t_1; x_2, t_2; \dots; x_N, t_N)}{P(x_1, t_1; x_2, t_2; \dots; x_m, t_m)}. \end{aligned} \quad (2.23)$$

Rozkład warunkowego prawdopodobieństwa określa prawdopodobieństwo tego, że w chwili  $t_{m+1}$  zajdzie zdarzenie  $x_{m+1}, \dots$ , w chwili  $t_N$  zajdzie zdarzenie  $x_N$ , przy założeniu (spełnieniu warunku), że w chwili  $t_1$  zaszło zdarzenie  $x_1, \dots$ , w chwili  $t_m$  - zdarzenie  $x_m$ .

Ze wzoru (2.23) wynika, że

$$P(x_1, t_1; \dots; x_N, t_N) = P(x_1, t_1; \dots; x_{N-1}, t_{N-1})P(x_1, t_1; \dots; x_{N-1}, t_m | x_N, t_N). \quad (2.24)$$

## 2.5. Stacjonarne procesy stochastyczne. Właściwość ergodyczna

Procesy stochastyczne dzielą się na stacjonarne i niestacjonarne. Dla wielu fizycznych układów mechanizm, który powoduje chaotyczne fluktuacje zmiennej  $x$  jest stałym, tj. nie zależnym od początku liczenia czasu. Takie procesy stochastyczne nazywają stacjonarnymi. Dla stacjonarnych procesów stochastycznych każda z jego dystrybuant pozostaje bez zmian, jeżeli przesuniemy wszystkie chwile  $t_1, t_2, \dots, t_N$  o  $t_0$  dowolnej wartości, czyli dla procesów stochastycznych stacjonarnych

$$P(x_1, t_1; \dots; x_N, t_N) = P(x_1, t_1 + t_0; \dots; x_{N-1}, t_{N-1} + t_0) . \quad (2.25)$$

Istnieją dwa sposoby badania procesów stochastycznych. Pierwszy sposób polega na obserwacji realizacji procesu w różnych chwilach. Drugi sposób polega na obserwacji procesu stochastycznego w jednej chwili ale na dużym zbiorze podobnych realizacji procesu stochastycznego. Badanie doświadczalne dużego zbioru realizacji procesu stochastycznego zwykle jest bardzo trudne do realizacji. Natomiast łatwiej można przeprowadzić pomiary jednej realizacji w wielu różnych chwilach. Znaczna część stacjonarnych procesów ma tak zwaną właściwość ergodyczną. Twierdzenie ergodyczne dla procesów stacjonarnych stwierdza, że dowolna statystyczna charakterystyka procesu, otrzymana ze zbioru realizacji w dowolnej chwili  $t$  jest równa (z prawdopodobieństwem zbliżonym do jedynki) podobnej charakterystyce otrzymanej z jednej realizacji procesu przy obliczeniu jej jako średniej względem dostatecznie długiego czasu. Z twierdzenia ergodycznego wynika, że obydwa sposoby badania procesów stochastycznych są prawie równoważne, a zatem uśrednienie losowej funkcji względem różnych realizacji procesu można zastąpić uśrednieniem jednej realizacji względem czasu.

## 2.6. Stochastyczne procesy Markowa

Ważnym stochastycznym procesem, znajdującym szerokie zastosowanie w praktyce, jest proces Markowa. Proces Markowa – ciąg zdarzeń, w którym prawdopodobieństwo każdego zdarzenia zależy jedynie od wyniku poprzedniego. Dla tego procesu żeby opisać możliwe zdarzenia w przyszłości wystarczy wiedzieć tylko zdarzenia które zaszły teraz. Proces stochastyczny ciągły lub skokowy nazywają procesem Markowa, gdy dla każdego zbioru skończonego  $t_1 < t_2 < \dots < t_{n-1} < t_n$  ( $n > 1$ ) i dowolnych  $x_1, x_2, \dots, x_{n-1}, x_n$  jest odpowiednio

$$\begin{aligned}
P(x_1, t_1; x_2, t_2; \dots; x_{n-1}, t_{n-1} | x_n, t_n) &= P(x_{n-1}, t_{n-1} | x_n, t_n) = \\
&= \frac{P(x_{n-1}, t_{n-1}; x_n, t_n)}{P(x_n, t_n)} .
\end{aligned} \tag{2.26}$$

Więc dla procesu Markowa prawdopodobieństwo tego, że w chwili  $t_n$  losowa zmienna  $X(t)$  przyjmuje wartość  $X(t_n) = x_n$  zależy tylko od znajomości  $X(t)$  w chwili poprzedniej  $t_{n-1}$ . Znajomość  $X(t_{n-2}), X(t_{n-3}), \dots, X(t_1)$  nie dodaje żadnej nowej informacji o rozkładzie zmiennej  $X(t)$ .

Dla stacjonarnych procesów Markowa, zgodnie z (2.25)

$$P(x_{n-1}, t_{n-1} | x_n, t_n) = P(x_{n-1}, t_{n-1} + t_0 | x_n, t_n + t_0) , \tag{2.27}$$

gdzie  $t_0$  może przyjmować dowolną wartość.

Jeżeli wybierzemy  $t_0 = -t_{n-1}$  i wprowadzimy zmienną  $\tau = t_n - t_{n-1}$ , to wzór (2.27) możemy zapisać w postaci

$$P(x_{n-1}, t_{n-1} | x_n, t_n) = P(x_{n-1} | x_n, \tau) , \tag{2.28}$$

gdzie  $x_{n-1}$  - wartość  $X(t)$  w chwili  $t = 0$ .

Biorąc pod uwagę (2.24) i (2.28) dla rozkładu prawdopodobieństwa rzędu  $N$  stochastycznego procesu Markowa możemy zapisać

$$P(x_1, t_1; \dots; x_N, t_N) = P(x_1, 0) \prod_{i=2}^N P(x_{i-1} | x_i, \tau_i) , \tag{2.29}$$

gdzie  $\tau_i = t_i - t_{i-1}$ .

## 2.7. Równanie Chapmana-Kolmogorowa-Smoluchowskiego

Z definicji (2.26) procesu Markowa wynika podstawowe równanie procesów Markowa - równanie Chapmana-Kolmogorowa-Smoluchowskiego

$$\begin{aligned}
P(x_1, t_1 | x_2, t_2) &= \sum_m P(x_1, t_1 | x_m, t_3) P(x_m, t_3 | x_2, t_2), \\
t_1 \leq t_3 \leq t_2.
\end{aligned} \tag{2.30}$$

Dla stacjonarnego procesu Markowa równanie Chapmana-Kolmogorowa-Smoluchowskiego przyjmuje postać

$$P(x_1 | x_2, t) = \sum_m P(x_1 | x_m, t - \tau) P(x_m | x_2, \tau) , \tag{2.31}$$

gdzie  $t = t_2 - t_1$ ,  $\tau = t_2 - t_3$  i  $t - \tau = t_3 - t_1$ .

## 2.8. Równanie Smoluchowskiego

Równanie Chapmana – Kolmogorowa - Smoluchowskiego jest słusznym przy dowolnej wartości  $\tau$  :  $t \geq \tau \geq 0$ . Jeżeli  $\tau$  jest bardzo małym, to  $P(x_m | x_j, \tau)$  możemy przedstawić w postaci potęgowego względem  $\tau$  szeregu Taylora

$$P(x_m | x_j, \tau) = P(x_m | x_j, 0) + \left. \frac{\partial P(x_m | x_j, \tau)}{\partial \tau} \right|_{\tau=0} \tau + \dots \quad (2.32)$$

Oznaczając

$$W_{mj} = \left. \frac{\partial P(x_m | x_j, \tau)}{\partial \tau} \right|_{\tau=0} \tau \quad (2.33)$$

i porzucając we wzorze (2.32) człony zawierające  $\tau^n$  przy  $n > 1$  otrzymujemy

$$P(x_m | x_j, \tau) = P(x_m | x_j, 0) + W_{mj} \tau \quad (2.34)$$

Ponieważ w chwili  $t = 0$ , zmienna  $X(t)$  przyjmuje określoną wartość  $x_m$ , to możemy zapisać

$$P(x_m | x_j, 0) = \delta_{mj} \quad (2.35)$$

gdzie  $\delta_{mj}$  - symbol Kronekiera.

Uwzględniając (2.35), ze wzoru (2.34) otrzymujemy dla  $m \neq j$

$$W_{mj} = \frac{P(x_m | x_j, \tau)}{\tau} \quad (2.36)$$

A więc  $W_{mj}$  jest prawdopodobieństwem przejścia układu fizycznego za czas  $\tau$  od stanu  $x_m$  do stanu  $x_j$ .

Podstawiając (2.34) do równania Chapmana-Kolmogorowa-Smoluchowskiego otrzymujemy

$$\begin{aligned} P(x_i | x_j, t) &= \sum_m P(x_i | x_m, t - \tau) P(x_m | x_j, \tau) = \\ &= P(x_i | x_j, t - \tau) + \tau \sum_m P(x_i | x_m, t - \tau) W_{mj} \end{aligned} \quad (2.37)$$

Skąd znajdujemy następujące równanie różnicowe



$$\frac{P(x_i | x_j, t) - P(x_i | x_j, t - \tau)}{\tau} = \sum_m P(x_i | x_m, t - \tau) W_{mj}. \quad (2.38)$$

Przy  $\tau \rightarrow 0$  równanie to sprowadza się do układu równań różniczkowych

$$\frac{\partial P(x_i | x_j, t)}{\partial t} = \sum_m P(x_i | x_m, t) W_{mj}. \quad (2.39)$$

Równanie (2.39) czasami nazywa się równaniem Smoluchowskiego.

Z definicji funkcji warunkowego prawdopodobieństwa (2.26) i wzorów (2.19) i (2.20) wynika, że

$$\sum_m P(x_i | x_m, t) = \frac{1}{P(x_i, 0)} \sum_m P(x_i, 0; x_m, t) = 1. \quad (2.40)$$

Podstawiając (2.34) do wzoru (2.40) znajdujemy

$$\sum_m P(x_i | x_m, t) = 1 + \tau \sum_m W_{mj} = 1. \quad (2.41)$$

Skąd mamy

$$W_{ii} = - \sum_{m \neq i} W_{mi}. \quad (2.42)$$

Przypomnijmy, że  $|W_{ii}|$  jest prawdopodobieństwem tego, że układ fizyczny za czas  $\tau$  pozostaje w stanie  $x_i$ .

Układ równań (2.39) w jawnej postaci wygląda następująco

$$\begin{aligned} \frac{\partial P(x_i | x_1, t)}{\partial t} &= W_{11}P(x_i | x_1, t) + W_{21}P(x_i | x_2, t) + \dots + W_{n1}P(x_i | x_n, t), \\ &\dots \\ \frac{\partial P(x_i | x_k, t)}{\partial t} &= W_{1k}P(x_i | x_1, t) + W_{2k}P(x_i | x_2, t) + \dots + W_{nk}P(x_i | x_n, t), \\ &\dots \\ \frac{\partial P(x_i | x_n, t)}{\partial t} &= W_{1n}P(x_i | x_1, t) + W_{2n}P(x_i | x_2, t) + \dots + W_{nn}P(x_i | x_n, t), \end{aligned} \quad (2.43)$$

gdzie  $n$  - liczba możliwych wartości losowej zmiennej  $X(t)$ .

Wprowadzając macierz  $\vec{W}$

$$\vec{W} = \begin{bmatrix} W_{11} & W_{21} & \cdots & W_{n1} \\ W_{12} & W_{22} & \cdots & W_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ W_{1n} & W_{2n} & \cdots & W_{nn} \end{bmatrix}, \quad (2.44)$$

oraz wektor  $P(x_i | \vec{X}(t))$  ze składowymi

$$P(x_i | x_1, t), P(x_i | x_2, t), \dots, P(x_i | x_n, t), \quad (2.45)$$

układ równań (2.43) możemy zapisać w następującej postaci

$$\frac{\partial P(x_i | \vec{X}(t))}{\partial t} = \vec{W} \cdot P(x_i | \vec{X}(t)). \quad (2.46)$$

Zauważmy, że we wzorze (2.46) „dynamiczna” macierz  $\vec{W}$  nie zależy od czasu  $t$ , a zatem ogólne rozwiązanie układu różniczkowych równań będzie sumą funkcji exponentialnych. Pokażemy to w następującym wykładzie.