Ćwiczenie 18

MODELOWANIE (SYMULACJA) PROCESÓW STOCHASTYCZNYCH

18.1. Wiadomości ogólne

Bardzo wiele zdarzeń, a także zjawisk fizycznych zachodzących w otaczającym nas świecie, nosi cechy procesów przypadkowych – stochastycznych, tzn. takich, dla których prawdopodobieństwo ich zajścia jest stałe w czasie (nie zależy od warunków zewnętrznych ani od historii danego obiektu). Należą do nich m.in. nieuporządkowany, chaotyczny ruch cząsteczek cieczy i gazów (ruchy Browna), naturalny rozpad promieniotwórczy, szumy urządzeń elektronicznych, rozpraszanie elektronów przechodzących przez ośrodki materialne, a nawet w pewnym sensie błądzenie człowieka zagubionego w lesie. Również proces pomiaru i występowanie niepewności przypadkowych należą do tej samej klasy procesów stochastycznych, których rozwój czasowy opisany jest przez wartości zmiennych losowych, a ich występowanie charakteryzuje się pewnym rozkładem statystycznym.

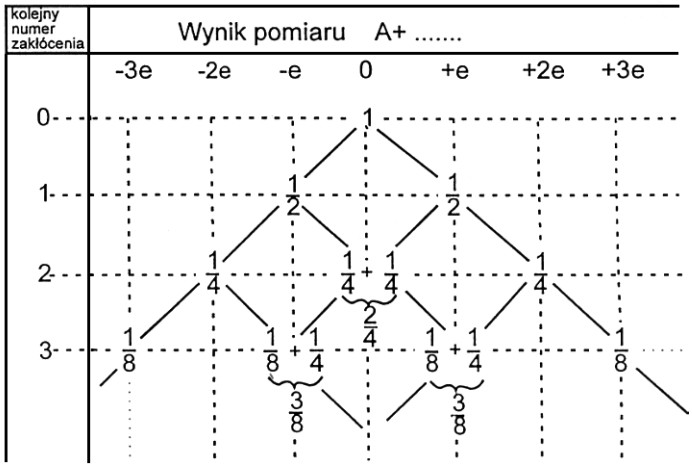
Od bardzo dawna uważano, że każda zmienna losowa ma rozkład normalny lub przynajmniej zbieżny do rozkładu normalnego. W bardzo wielu dziedzinach nauki i techniki znajdowano liczne przykłady zmiennych losowych o takim właśnie rozkładzie. Rozkład ten zaczęto uważać za uniwersalny rozkład statystyczny i dlatego właśnie nazwano rozkładem normalnym, dając tym wyraz, że jest to rozkład zazwyczaj występujący w praktyce. Dzisiaj wiemy, że rozkład normalny nie jest jedynym rozkładem teoretycznym, który daje dobre przybliżenie rozkładów eksperymentalnych, to jednak z całą pewnością można go zaliczyć do rozkładów, które najczęściej mogą być stosowane do aproksymacji danych doświadczalnych.

Najbardziej rozpowszechnionym poglądem, tłumaczącym przyczyny częstego występowania rozkładu normalnego, był pogląd wysunięty przez F. Bessela. Według niego wartości obserwacji statystycznych zależą od wielu drobnych przyczyn. Przyczyny te, to są swego rodzaju impulsy, które powodują, że wartości obserwacji statystycznych odchylają się od średniej arytmetycznej. Gdy impulsy mają jednakowy kierunek, wówczas zaobserwowane odchylenie wartości cechy od średniej arytmetycznej jest duże, jeżeli natomiast impulsy działają różnokierunkowo, to sumaryczne odchylenie wartości cechy od tej średniej jest małe.

Oczywiście prawdopodobieństwo tego, że znaczna ilość odchyleń będzie miała ten sam kierunek, jest małe, jeżeli przyjmie się nie budzące intuicyjnych sprzeciwów założenie, że odchylenia jednakowe co do wartości bezwzględnej, a różniące się jedynie kierunkiem (znakiem), mają jednakowe prawdopodobieństwo wystąpienia. Zakłada się również, że odchylenia są od siebie niezależne. Przy takich założeniach, na mocy twierdzenia o prawdopodobieństwie iloczynu zdarzeń niezależnych możemy uważać, że wystąpienie znacznej ilości odchyleń jednokierunkowych jest zdarzeniem mało prawdopodobnym. Prawdopodobieństwo tego zdarzenia równa się bowiem iloczynowi prawdopodobieństw zdarzeń elementarnych.

Wynika stąd, że odchylenia duże co do bezwzględnej wartości (czyli odchylenia będące sumą małych odchyleń jednakowo skierowanych) mają małe prawdopodobieństwo wystąpienia, a odchylenia małe mają to prawdopodobieństwo duże. Bessel traktuje odchylenia od średniej arytmetycznej tak samo jak błędy obserwacji. Ponieważ błędy obserwacji, jak wykazali Gauss i Laplace, mają rozkład normalny, przeto z wywodów Bessela wypływa wniosek, że każda cecha statystyczna ma rozkład normalny.

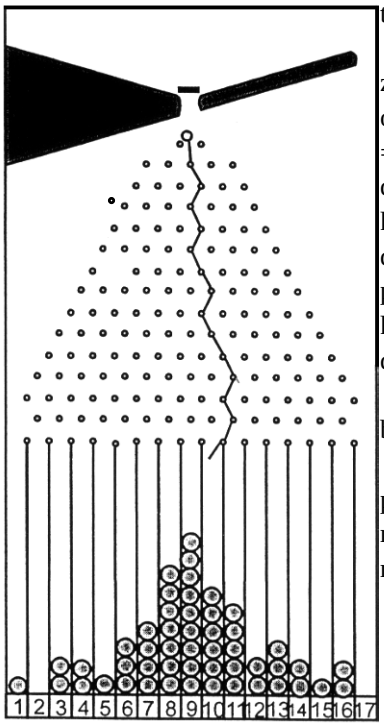
Wykorzystując zatem model Laplacea, można wykazać, jak niezależne zakłócenia (impulsy) występujące w pomiarze powodują narastanie niepewności pomiarowych, odchylając wynik pomiaru od rzeczywistej wartości (rys. 18.1).



Rys. 18.1

Załóżmy, że wartość rzeczywista mierzonej wielkości wynosi A, oraz że pomiar tej wielkości jest zakłócony przez N różnych, niezależnych czynników, z których każdy z prawdopodobieństwem p = 1/2 powoduje odchylenie wyniku pomiaru o wielkość ± e. Rysunek 18.1 pokazuje narastanie niepewności pomiarowych wraz ze wzrostem liczby czynników zakłócających pomiar. Prawdopodobieństwa otrzymania określonego wyniku podane są na przecięciu kolumn i wierszy tabeli, a ich suma w każdym wierszu równa jest pewności (P = 1). Taki proces narastania wypadkowego odchylenia można opisać rozkładem dwumianowym lub jego aproksymacją – rozkładem normalnym.

Prawo rozkładu zdarzeń przypadkowych zakłócanych wieloma różnymi, niezależnymi czynnikami, można symulować za pomocą deski Galtona (rys. 18.2). Jest to prostokątna deska, w której prostopadle do powierzchni wystaje kilkanaście rzędów kołków, których rozkład przypomina trójkąt Pascala. Odległość między kołkami jest taka, że może przejść między nimi tylko jedna kulka. Deska jest lekko pochylona. Z lejkowatego zbiornika, znajdującego się u góry deski, spadają kulki i zderzają się z kolejnymi warstwami kołków, doznając przy tym odchylenia w prawo lub w lewo z jednakowym prawdopodobieństwem p = q = 1/2. Wynika to stąd, że każdy kołek niższej warstwy znajduje się dokładnie pośrodku między dwoma kołkami wyższej warstwy, dlatego kulka trafia dokładnie swym środkiem na kołek stojący na jej drodze. Dziełem przypadku jest, czy po odbiciu od niego kulka przejdzie niżej z jego lewej, czy prawej strony. Gdy większa liczba kulek przejdzie takie przypadkowe drogi, ułożą się one w rowkach u dołu deski w kształcie krzywej symetrycznej. Krzywa wskaże maksymalne obsadzenie w okolicy swego środka, zaś ku brzegom obsadzenia będą zmierzały do zera (krzywa rozkładu Gaussa).



Symetria tej krzywej będzie zaburzona przez fluktuacje, tzn. niewielkie wahania w obsadzeniu poszczególnych rowków, które wynikają z równowagi chwiejnej kul. W każdym razie zarówno stałość krzywej rozkładu, jak i zmienność zachodzących fluktuacji świadczą o występujących w tym zjawisku prawidłowościach statystycznych.

Prawa statystyczne nie mogą dać odpowiedzi na pytania dotyczące określonego zachowania się pewnego obiektu, np. kulki w ruchu po desce Galtona. Tak więc nie potrafimy powiedzieć, do którego rowka wpadnie konkretna kulka, a także nie wiemy, w którą stronę się odchyli. Możemy tylko powiedzieć, że gdybyśmy kulkę

puszczali odpowiednio dużo razy, to średnio tyle a tyle razy wpadnie do określonego rowka, tzn. możemy określić, jakie jest prawdopodobieństwo tego zdarzenia.

Na rysunku 1.2 linią łamaną połączono kołki, na których spadająca kulka doznaje kolejnych odchyleń w pewnym przykładowo wybranym ruchu. Ruch kuleczek po desce Galtona jest przykładem ruchu zwanego błądzeniem przypadkowym i należy, podobnie jak przytoczone wcześniej przykłady, do obszernej klasy procesów stochastycznych, czyli ruchów, których rozwój czasowy opisywany jest przez wartości zmiennych losowych.

Kulka spadając w dół doznaje N odchyleń, kolejno na każdym rzędzie kołków w lewo albo w prawo, zawsze tylko o jeden rowek – to odchylenie jest zmienną losową. Jeżeli prawdopodobieństwo odchylenia kulki w prawo oznaczymy przez p, a w lewo przez q, to prawdopodobieństwo tego, że – w wyniku N odchyleń – n z nich wystąpi w prawo, można wyrazić rozkładem dwumianowym

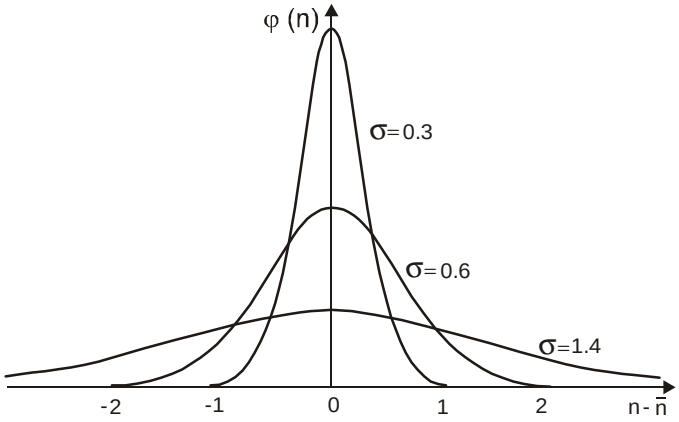
$$P_N(n) = \binom{N}{n} p^n q^{N-n} \ , \tag{18.1}$$

gdzie

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

jest liczbą n-elementowych kombinacji zbioru N-elementowego.

Rozkład dwumianowy stosujemy w przypadku, gdy dyskretna zmienna losowa przyjmuje niezbyt duże wartości n. Natomiast dla dużych wartości n, na mocy centralnego twierdzenia rozkład dwumianowy dąży do rozkładu normalnego, zwanego także rozkładem Gaussa. Rysunek 18.3 jest ilustracją rozkładu normalnego, w którym zmienna losowa przyjmuje ciągłe wartości.



Rys. 18.3

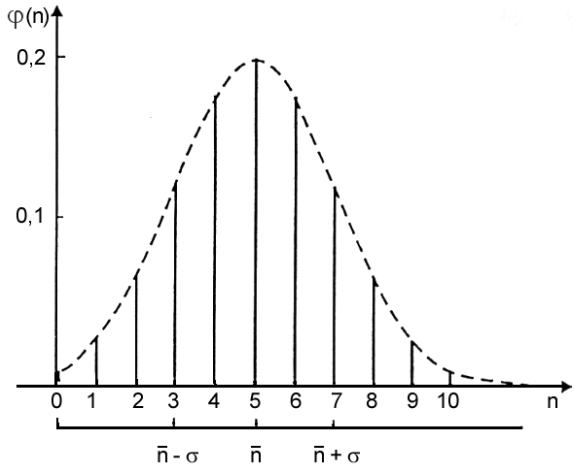
Przedstawione na rysunku 18.3 przebiegi można przybliżyć jedną wspólną krzywą gęstości prawdopodobieństwa rozkładu normalnego φ(n) opisaną wzorem

$$\varphi(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(n-\bar{n})^2}{2\sigma^2}\right], \tag{18.2}$$

gdzie $\bar{n}$ jest wartością średnią zmiennej losowej n, określającą położenie maksimum rozkładu, a σ – odchylenie standardowe. Kształt krzywej Gaussa silnie zależy od wartości odchylenia standardowego – i jak widać z rysunku – im większa jest jego wartość, tym szersza i bardziej spłaszczona jest krzywa rozkładu. Można wykazać, że wartość odchylenia standardowego σ jest równa $|n - \bar{n}|$, dla takich wartości n, dla których krzywa rozkładu ma punkty przegięcia. Położenia tych punktów są symetryczne względem położenia maksimum krzywej rozkładu, a ich wartości wynoszą odpowiednio: $\bar{n} - \sigma$ i $\bar{n} + \sigma$.

Wykonując natomiast całkowanie funkcji rozkładu Gaussa w granicach od $\bar{n} - \sigma$ do $\bar{n} + \sigma$, możemy wykazać, że odchylenie standardowe określa szerokość przedziału $(< \bar{n} - \sigma, \ \bar{n} + \sigma >)$, w którym z

prawdopodobieństwem P = 0,683 znajduje się dowolna wartość zmiennej losowej. Przykład rozkładu normalnego o parametrach $\bar{n} = 5$ oraz $\sigma = 2$ pokazano na rys. 18.4.



Rys. 18.4

Między parametrami rozkładu normalnego a dwumianowego występują ścisłe relacje określone wyrażeniami:

$$\bar{n} = pN, \quad \sigma = \sqrt{Npq}\,. \tag{18.3}$$

Dla omawianego tu przypadku deski Galtona, dla której p = q = 1/2, otrzymujemy: $\bar{n} = N/2$ oraz $\sigma = (N/4)^{1/2}$, przy czym N jest liczbą rowków deski Galtona (N = 17 na rys. 18.2).

Możliwość wykorzystania rozkładu Gaussa do obliczenia wartości prawdopodobieństwa P(n) w konkretnym zastosowaniu wymaga znajomości wartości przybliżonych, czyli estymat zmiennej losowej $\bar{n}$ i odchylenia standardowego σ – w praktyce wartości te obliczamy z próby. Wykonując pomiary ruchu serii ν kulek na desce Galtona, możemy obliczyć estymaty z następujących wyrażeń:

$$\overline{n_\nu} = \sum_{n=1}^{N} \frac{n\nu_n}{\nu}\,, \quad \sigma_\nu = \left(\sum_{n=1}^{N} \frac{\left(n - \overline{n_\nu}\right)^2 \nu_n}{\nu - 1} \right)^{\frac{1}{2}} \approx \left(\sum_{n=1}^{N} \frac{\left(n - \overline{n_\nu}\right)^2 \nu_n}{\nu - 1} \right)^{\frac{1}{2}} \tag{18.4}$$

Znaczenie symboli w wyrażeniach (18.4) zostanie wyjaśnione w punkcie 1.3, gdzie opisano zasadę i przebieg pomiaru.

18.2. Zadania

18.2.1. Wyznaczyć eksperymentalne wartości prawdopodobieństwa wpadnięcia kulki do n-tego rowka dla serii ν = 100, 200, 300 kulek.

18.2.2. Wyznaczyć prawdopodobieństwo P(n) wpadnięcia kulki do n-tego rowka na podstawie rozkładu normalnego (wz. (18.4)) dla ν = 300 kulek.

18.2.3. Wykreślić otrzymane w doświadczeniu rozkłady dla całkowitej liczby ν = 100, 200, 300 kulek. Na ten sam wykres nanieść rozkład Gaussa dla ν = 300 kulek, wyliczony w pkt. 18.2.2.

18.2.4. Obliczyć względną liczbę kulek, które znalazły się w rowkach deski zawartych w przedziale $\bar{n} \pm \sigma$ dla całkowitej liczby 300 kulek.

18.3. Zasada i przebieg pomiaru

18.3.1. Przed przystąpieniem do pomiarów należy przechylić deskę i przesypać wszystkie kulki do lejka zsypowego. Następnie spuszczamy pierwszą serię 100 kulek, kolejno jedna po drugiej. Zliczamy kulki w poszczególnych rowkach, a wyniki zapisujemy w tabeli. Powtarzamy cykl pomiarów dla następnych serii po 100 kulek każda, zwiększając w ten sposób całkowitą liczbę kulek do 200 w drugim i 300 w trzecim pomiarze. Wyniki zapisujemy w odpowiednich wierszach tabeli 18.1.

Tabela 18.1

Wyniki pomiarów

ν	n	1	2	3	4	5	6	...	...	N = 23
100	ν_n									
	$P_{exp}(n)$									
200	ν_n									
	$P_{exp}(n)$									
300	ν_n									
	$P_{exp}(n)$									
	P(n)									

W tabeli 18.1 zastosowano następujące oznaczenia: ν – całkowita liczba kulek w kolejnym pomiarze (100, 200, 300), ν_n – liczba kulek w n-tym rowku (n = 1, 2, 3, ...N), gdzie N = 23 jest całkowitą liczbą rowków deski Galtona. Korzystając z zależności: $P_{exp}(n) = \nu_n / \nu$, dla trzech serii pomiarów obliczamy doświadczalne prawdopodobieństwa wpadnięcia kulki do n-tego rowka. Wyniki zapisujemy w odpowiednim wierszu tabeli.

18.3.2. Bazując na danych zawartych w tabeli dla całkowitej liczby ν = 300 kulek, obliczamy estymaty wartości średniej n i odchylenia standardowego σ ze wzorów (18.4), przy czym ze względu na dużą wartość ν, możemy pominąć 1 w wyrażeniu na estymatę odchylenia standardowego. Następnie obliczamy prawdopodobieństwo P(n) wpadnięcia kulki do n-tego rowka ze wzoru na rozkład normalny (wz. (18.2)), przystosowany dla dyskretnej zmiennej losowej n

$$P(n) = \int_n \varphi(n)dn \equiv \varphi(n) \cdot \Delta n, \text{ gdzie } \Delta n = 1 . \qquad (18.5)$$

Wyniki zapisujemy w ostatnim wierszu tabeli 18.1.

18.3.3. Sporządzić wykres słupkowy (taki jak na rys. 18.4) rozkładu normalnego dla trzech serii pomiarów (ν = 100, 200, 300), wykorzystując wartości $P_{exp}(n)$ zawarte w tabeli. Na ten sam wykres nanieść punkty wyliczone ze wzoru na rozkład Gaussa – odpowiednie wartości P(n) zawarte są w ostatnim wierszu tabeli. Obliczyć względną liczbę kulek, które znalazły się w rowkach deski Galtona, w przedziale $\bar{n} \pm \sigma$ i porównać wynik z teoretyczną wartością prawdopodobieństwa znalezienia się kulek w rowkach z tego samego przedziału, wynikającą z rozkładu Gaussa (P = 0,683, wz. (18), Wstęp).

18.4. Ocena niepewności pomiarów

Deska Galtona jest swego rodzaju analogiem przyrządu pomiarowego, którego najmniejsza działka jest szerokością rowka, równą dokładnie odległości między środkami kołków w każdym rzędzie. Ruch kulki jest symulacją procesu pomiaru, a wynik pomiaru to wartość zmiennej losowej n (numer rowka, do którego ostatecznie trafi kulka). Zatem odchylenie standardowe σ jest miarą niepewności pojedynczego pomiaru (kulka doznaje zawsze elementarnego poziomego przemieszczenia po zderzeniu z kołkiem – w pomiarze zawsze występują niepewności przypadkowe). Odchylenie standardowe σ w teorii błędu nazywa się błędem średnim kwadratowym pojedynczego pomiaru (szersza dyskusja tego zagadnienia znajduje się w pkt. 4 – Wstęp).

Należy pamiętać, że zmienna losowa n przyjmuje wartości dyskretne, zatem estymaty zmiennej losowej n i odchylenia standardowego σ są liczbami całkowitymi!

Wyniki obliczeń wartości prawdopodobieństw $P_{exp}(n)$ i P(n) podać z dokładnością do trzech cyfr znaczących.

Literatura

[1] Wróblewski A. K., Zakrzewski A. J.: Wstęp do fizyki. Warszawa: PWN 1976.
[2] Szydłowski H.: Teoria pomiarów. Warszawa: PWN 1979.
[3] Bartnicki S., Borys W., Kostrzyński T.: Fizyka ogólna. Ćwiczenia laboratoryjne cz. I. Warszawa: Skrypt WAT 1991.
[4] Hrabowska J., Tykarski R.: Laboratorium podstaw fizyki. Opis ćwiczeń. Zeszyt 2. Warszawa: WPW 1986.