1 Przestrzeń zdarzeń elementarnych

Przestrzeń zdarzeń elementarnych jest pojęciem pierwotnym w teorii prawdopodobieństwa. W zastosowaniach tej teorii zdarzenia elementarne interpretuje się jako możliwe przypadki, wyniki doświadczenia, stany obiektów, wystąpienia zjawisk, itp., jednak zawsze w takiej sytuacji, kiedy istnieje niepewność, który z przypadków, wyników, stanów, itd. pojawił się bądź pojawi się w przyszłości.
Zbiory zdarzeń elementarnych, czyli podzbiory danej przestrzeni zdarzeń elementarnych nazywamy krótko *zdarzeniami.* Zbiór pusty nazywamy zdarzeniem niemożliwym, a całą przestrzeń — zdarzeniem pewnym. O zdarzeniach elementarnych, które należą do danego zdarzenia, mówimy, że mu sprzyjają.
Oznaczenia:

- Ω — przestrzeń zdarzeń elementarnych;
- ω — zdarzenie elementarne;
- $A \subset \Omega$ — zdarzenie;
- $A' = \Omega \setminus A$ — zdarzenie przeciwne.

Jeśli np. rozpatrujemy doświadczenie polegające na rzucie kostką, to zdarzeniami elementarnymi będą poszczególne wyniki: 1, 2, 3, 4, 5, 6. Przestrzeń zdarzeń elementarnych jest zbiorem $\Omega = \{1, 2, 3, 4, 5, 6\}$. Zdarzeniem jest każdy jego podzbiór, np. $A = \{1, 3, 5\}$ (wyrzucenie nieparzystej liczby oczek); zdarzeniem przeciwnym jest wtedy $A' = \{2, 4, 6\}$(wyrzucenie parzystej liczby oczek). (Wszystkich zdarzeń jest $2^6 = 64$ — dlaczego?)
Jak widać w konkretnym problemie zbiór Ω musi być zdefiniowany, i na ogół nie powinno to sprawiać problemu. Jednak trzeba sobie zdawać sprawę, że zdarzenia *nie są elementami zbioru Ω, lecz jego podzbiorami.* Zatem należy rozpatrywać jakąś klasę $\mathcal{F}$ podzbiorów zbioru Ω. Może to być klasa wszystkich podzbiorów, ale może też być mniejsza.
W wyborze tej klasy nie ma całkowitej dowolności: aby poprawnie rozwijać teorię klasa $\mathcal{F}$ musi mieć trzy cechy:

1. $\mathcal{F} \neq \emptyset$;
2. $A \in \mathcal{F} \Rightarrow A' \in \mathcal{F}$;
3. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Te własności zapewniają, że suma, przekrój (iloczyn), czy dopełnienie zdarzenia też jest zdarzeniem. Rodzinę zbiorów o takich własnościach nazywamy *ciałem zdarzeń.*

Definicja 1 *Niech* $\mathcal{F}$ *będzie ciałem zdarzeń. Funkcję liczbową* $P : \mathcal{F} \to [0, 1]$ *nazywamy* prawdopodobieństwem *jeśli:*

P1. $(A_1, A_2, \ldots \in \mathcal{F}$ oraz $A_i \cap A_j = \emptyset$ dla $i \neq j) \Rightarrow P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$;

P2. $P(\Omega) = 1$.

Trójkę $(\Omega, \mathcal{F}, P)$ nazywamy *przestrzenią probabilistyczną.*
Na danym ciele zdarzeń można określić różne prawdopodobieństwa. Umiejętne określenie zarówno zbioru Ω jak i prawdopodobieństwa ma podstawowe znaczenie dla skuteczności zastosowania teorii.
Definicja prawdopodobieństwa jest stosunkowo prosta, ale warunki P1 i P2 są na tyle "mocne", że można z nich wywnioskować następujące własności prawdopodobieństwa.
Własności prawdopodobieństwa

1. $P(\emptyset) = 0$;
2. $P(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k)$, o ile $A_i \cap A_j \neq \emptyset$;
3. Jeżeli $A, B \in \mathcal{F}$ oraz $A \subset B$, to $P(A) \leqslant P(B)$;
4. Jeżeli $A_1, A_2, \ldots \in \mathcal{F}$, to $P(\bigcup_{k=1}^{\infty} A_k) \leqslant \sum_{k=1}^{\infty} P(A_k)$;
5. Jeżeli $A, B \in \mathcal{F}$, to $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Przykładowo wykażemy własność 5. Mamy $B = (A \cap B) \cup (B \setminus A)$, zatem z warunku P1:

$$P(B) = P(A \cap B) + P(B \setminus A), \text{ czyli } P(B \setminus A) = P(B) - P(A \cap B).$$

Ale także $A \cup B = A \cup (B \setminus A)$, więc

$$P(A \cup B) = P(A) + P(B \setminus A),$$

i podstawiając $P(B \setminus A) = P(B) - P(A \cap B)$ otrzymujemy tezę.

2 Skończone przestrzenie probabilistyczne

Przestrzeń probabilistyczną nazywamy skończoną, gdy skończona jest przestrzeń zdarzeń elementarnych Ω.
Jeżeli $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, to ciałem zdarzeń $\mathcal{F}$ jest wtedy zbiór wszystkich podzbiorów zbioru Ω. $\mathcal{F}$ ma zatem 2^n elementów. Prawdopodobieństwo będzie określone, jeżeli każdemu zdarzeniu elementarnemu ω_i przyporządkujemy liczbę p_i taką, że $p_i \geqslant 0$ oraz $\sum_{i=1}^{n} p_i = 1$. Wtedy każdemu zdarzeniu $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ będzie przypisane prawdopodobieństwo

$$P(A) = P(\omega_{i_1}) + P(\omega_{i_2}) + \cdots + P(\omega_{i_k}).$$

W szczególności, jeśli przyjmiemy $p_i = \frac{1}{n}$, to otrzymamy klasyczną definicję prawdopodobieństwa (Laplace'a):
Prawdopodobieństwo zdarzenia A jest równe stosunkowi liczby zdarzeń sprzyjających zdarzeniu A do liczby wszystkich możliwych zdarzeń elementarnych.
Przykład Sześcian, którego wszystkie ściany są pomalowane rozpiłowano tworząc 1000 sześcianików jednakowej wielkości. Obliczyć prawdopodobieństwo, że losowo wybrany sześcianik będzie miał dwie ściany pomalowane.
Rozwiązanie. Tutaj $n = |\Omega| = 1000$. Ponieważ ściana ma 12 krawędzi, a na każdej z nich 8 sześcianików o dwóch ścianach pomalowanych, zatem $k = 12 \cdot 8 = 96$. Ostatecznie $p = \frac{k}{n} = 0,096$.

3 Nieskończone przestrzenie probabilistyczne

Jeżeli przestrzeń zdarzeń elementarnych jest nieskończona, to również przestrzeń probabilistyczną $(\Omega, \mathcal{F}, P)$ nazywamy nieskończoną. W tym przypadku klasyczna definicja prawdopodobieństwa nie może być zastosowana. Ale w wielu przypadkach można zastosować tzw. prawdopodobieństwo geometryczne. Mianowicie, jeśli całą przestrzeń zdarzeń elementarnych można zinterpretować jako obszar na płaszczyźnie (lub w przestrzeni), to można przyjąć, że prawdopodobieństwem jest zwykła miara geometryczna, a dokładniej

$$P(A) = \frac{|A|}{|\Omega|},$$

gdzie $|A|, |\Omega|$ oznaczają miary geometryczne obszarów (długość na prostej, pole na płaszczyźnie, objętość w przestrzeni).

Przykłady

1. Jakie jest prawdopodobieństwo, że suma dwóch na chybił trafił wybranych dodatnich liczb, z których każda jest nie większa od 1, jest nie większa od 1 a ich iloczyn jest nie większy od $\frac{2}{9}$?

   Rozwiązanie. Niech x, y będą wybranymi liczbami. Ponieważ $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$, więc obszarem Ω jest kwadrat o polu 1. Wartości sprzyjające spełniają nierówności

   $$x + y \leqslant 1, \quad xy \leqslant \frac{2}{9}.$$

   Pole obszaru sprzyjającego:

   $$|A| = \int_0^{\frac{1}{3}} (1-x)dx + \int_{\frac{1}{3}}^{\frac{2}{3}} \frac{1}{x}dx + \int_{\frac{2}{3}}^{1} (1-x)dx = \frac{1}{3} + \frac{2}{9}\ln 2 \approx 0,487$$

   zatem

   $$P(A) = 0,487.$$

2. Na płaszczyźnie poprowadzone są proste równoległe, odległości między nimi wynoszą na zmianę 1,5 cm i 8 cm. Obliczyć prawdopodobieństwo, że losowo rzucony na tę płaszczyznę okrąg o promieniu 2,5 cm nie przetnie ani jednej prostej.

   Odp. $\frac{3}{9{,}5} \approx 0,316$.

3. Dwie osoby mają jednakowe prawdopodobieństwo przybycia na dane miejsce w każdej chwili przedziału czasu o długości T. Obliczyć prawdopodobieństwo, że czas oczekiwania jednej osoby na drugą będzie nie dłuższy niż t.

   Rozwiązanie Przestrzeń Ω to kwadrat o boku T: $0 \leqslant x \leqslant T$, $0 \leqslant y \leqslant T$. Natomiast $(x, y) \in A \Leftrightarrow |y - x| \leqslant t$. Zatem

   $$|A| = T^2 - 2 \cdot \frac{1}{2}(T-t)^2 = 2Tt - t^2,$$

   więc $P(A) = 2\frac{T}{t} - \frac{t^2}{T^2}$.

Zadanie mniej typowe. W teleturnieju uczestnik staje przed trzema drzwiami i aby wygrać samochód powinien odgadnąć za którymi drzwiami on jest. Uczestnik wskazuje drzwi, po czym prowadzący (który wie, gdzie jest samochód) otwiera jedne z pozostałych drzwi i pokazuje, że tam nie ma samochodu. W tym momencie uczestnik ma prawo zmienić decyzję. Jak powinien postąpić: zmienić, czy nie?

Odpowiedź: zmienić. Prawdopodobieństwo, że samochód jest za drzwiami, które wskazał na początku, wynosi 1/3. Zatem za tymi drugimi musi być z prawdopodobieństwem 2/3. Zadanie jest mylące, bo wiele osób sądzi, że wprawdzie na początku szansa właściwego wyboru jest 1/3, ale skoro pozostało dwoje drzwi do wyboru, to szanse są równe. Nie jest to prawda, bo prowadzący nie otworzy drzwi za którymi jest samochód, a więc jego postępowanie nie jest całkiem losowe. Inaczej: gdyby na początku kazać uczestnikowi wybrać drzwi i zapytać, jaka jest szansa na to, że samochód jest za tymi drzwiami, a jaka jest szansa, że za pozostałymi dwoma, to bez wątpienia powiedziałby, że 1/3 i 2/3. Fakt otwarcia jednych drzwi nie zmienia tych prawdopodobieństw.

4 Prawdopodobieństwo warunkowe

Czasem dysponujemy informacjami, na podstawie których należy wyeliminować pewne zdarzenia elementarne, ponieważ nie będą mogły one wystąpić. W takiej sytuacji należy rozważać prawdopodobieństwo warunkowe.
Zakładamy, że określona jest przestrzeń probabilistyczna $(\Omega, \mathcal{F}, P)$ i mamy warunek ograniczający, tzn. pewne zdarzenie $B \in \mathcal{F}$, $P(B) > 0$. Przez wystąpienie zdarzenia A pod warunkiem zdarzenia B rozumiemy zdarzenie $A \cap B$. Prawdopodobieństwo warunkowe zdarzenia A pod warunkiem zdarzenia B oznaczamy $P(A|B)$.
Prawdopodobieństwo warunkowe można obliczyć tworząc przestrzeń probabilistyczną, w której przestrzenią zdarzeń elementarnych będzie B. Będziemy rozważać klasę

$$\mathcal{F}_B = \{A \cap B : \ A \in \mathcal{F}\},$$

która jest ciałem zdarzeń. Na tym ciele określamy nową funkcję prawdopodobieństwa:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad A, B \in \mathcal{F}, \ P(B) > 0.$$

Jeżeli $P(A|B) = P(A)$, to zdarzenia A i B nazywamy niezależnymi. Wtedy mamy

$$P(A \cap B) = P(A) \cdot P(B),$$

i zwykle ten warunek podaje się jako definicję zdarzeń niezależnych.
Przykłady

1. Niech Z_n, $n = 1, 2, \ldots$ oznacza zdarzenie polegające na tym, że nowo zainstalowane urządzenie znajdować się będzie w stanie zdatności przez okres n kolejnych miesięcy. Wiadomo, że $P(Z_6) = 0,89$ i

$P(Z_{12}) = 0,53$. Obliczyć prawdopodobieństwo, że urządzenie pracujące 6 miesięcy będzie zdatne do użytku przez dalsze 6 miesięcy.

$$P(Z_{12}|Z_6) = \frac{P(Z_{12} \cap Z_6)}{P(Z_6)} = \frac{P(Z_{12})}{P(Z_6)} = \frac{0,53}{0,89} = 0,6.$$

2. Prawdopodobieństwo występowania zakłóceń przy przesyłaniu sygnału impulsowego (zdarzenie A) jest równe $P(A) = 0,012$, natomiast prawdopodobieństwo stłumienia sygnału (zdarzenie B) przez zakłócenie wynosi $P(B|A) = 0,25$. Jakie jest prawdopodobieństwo, że wystąpi zakłócenie tłumiące sygnał?

$$P(A \cap B) = P(A) \cdot P(B|A) = 0,003.$$

Twierdzenie 1 (o prawdopodobieństwie całkowitym) *Jeżeli zdarzenia $H_1, H_2, \ldots, H_n$ tworzą podział przestrzeni zdarzeń elementarnych i $P(H_i) > 0$ dla $i = 1, 2, \ldots n$, to dla dowolnego zdarzenia A zachodzi równość:*

$$P(A) = \sum_{i=1}^{n} P(A|H_i)P(H_i).$$

D o w ó d. Ponieważ $A = \sum_{i=1}^{n}(A \cap H_i)$ oraz $A \cap H_i$ są parami rozłączne, więc

$$P(A) = \sum_{i=1}^{n} P(A \cap H_i) = \sum_{i=1}^{n} P(A|H_i)P(H_i).$$

Przykład Prawdopodobieństwo, że przedmiot wyprodukowany przez maszynę M będzie pierwszego gatunku wynosi 0,7. Dla maszyny N to prawdopodobieństwo wynosi 0,8. Na maszynie M zrobiono dwa przedmioty, a na N trzy. Obliczyć prawdopodobieństwo, że wszystkie przedmioty są pierwszego gatunku.
Rozwiązanie. Niech A_i oznacza zdarzenie, że i-ty przedmiot jest pierwszego gatunku, H_1, że został zrobiony na maszynie M, a H_2, że został zrobiony na maszynie N. Wtedy

$$P(A_i) = P(A_i|H_1)P(H_1) + P(A_i|H_2)P(H_2) = 0,7 \cdot 0,4 + 0,8 \cdot 0,6 = 0,76.$$

Ponieważ zdarzenia A_i są niezależne, więc

$$P(A_1 \cap \cdots \cap A_5) = P(A_i)^5 = 0,254.$$

Twierdzenie 2 (wzór Bayesa) *Jeżeli hipotezy H_i, $i = 1, 2, \ldots, n$ stanowią pełny układ wyłączających się zdarzeń, to prawdopodobieństwo hipotezy H_k pod warunkiem, że zaszło zdarzenie A wynosi*

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{P(A)} = \frac{P(H_k)P(A|H_k)}{\sum_{i=1}^{n} P(H_i)P(A|H_i)}.$$

Przykłady

1. W partii 1000 żarówek może występować $0, 1, \ldots, 5$ wadliwych żarówek. Możliwości te są jednakowo prawdopodobne. Wzięto losowo 100 żarówek, i wszystkie okazały się dobre. Jakie jest prawdopodobieństwo, że wszystkie są dobre?

   Rozwiązanie. Niech H_i będzie zdarzeniem, że i żarówek jest wadliwych, A — że losowo wybranych 100 jest dobrych. Wtedy mamy

   $$P(H_i) = \frac{1}{6}, \quad P(A|H_i) = \frac{\binom{1000-i}{100}}{\binom{1000}{100}}$$

   dla $i = 0, 1, \ldots, 5$, więc

   $$P(H_0|A) = \frac{P(H_0)P(A|H_0)}{\sum_{i=0}^{5} P(H_i)P(A|H_i)} = \frac{1 \cdot \frac{1}{6}}{1 \cdot \frac{1}{6} + \frac{9}{10} \cdot \frac{1}{6} + \frac{899}{999} \cdot \frac{9}{10} \cdot \frac{1}{6} \cdots} = 0,214.$$

2. Wiadomo, że 90% produkcji pewnego wyrobu spełnia wymagania normy. Stosując uproszczoną kontrolę wyrobu uznaje się wybraną sztukę za dobrą z prawdopodobieństwem 0,98 gdy jest ona rzeczywiście dobra, i z prawdopodobieństwem 0,05, gdy jest zła. Jakie jest prawdopodobieństwo, że sztuka uznana za dobrą jest rzeczywiście dobra?

   Rozwiązanie. Niech H_1: sztuka dobra; $H_2 = H_1'$: sztuka zła; A_1: uznanie za dobrą; $A_2 = A_1'$: uznanie za złą. Po rachunkach:

   $$P(H_1|A_1) \approx 0,998.$$

   Można też obliczyć, że $P(H_1|A_2) = 0,159 = 15,9\%$. Zatem jest dość duże prawdopodobieństwo, że sztuka uznana za złą jest dobra.

3. W czasie egzaminu 50% studentów odpisywało. Spośród studentów uczciwych zdało 60%, a z grupy oszukujących 40%. Jakie jest prawdopodobieństwo, że student, który zdał egzamin, jest uczciwy?

Rozwiązanie. Niech A jest zdarzeniem, że zdał; U, że jest uczciwy.

$$P(U|A) = \frac{P(U)P(A|U)}{P(U)P(A|U) + P(U')P(A|U')} = \frac{\frac{1}{2} \cdot \frac{3}{5}}{\frac{1}{2} \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{2}{5}} = \frac{3}{5} = 60\%.$$

Przy innych danych: 60% odpisuje, zdaje 70% uczciwych i 40% nieuczciwych mamy wynik

$$P(U|A) = \frac{28}{52} \approx 54\%.$$

4. Pewna choroba występuje u $0,2\%$ ogółu ludności. Test do wykrycia choroby daje wynik pozytywny (tj. osoba jest chora) u 97% chorych i 1% zdrowych. U pewnej osoby test dał wynik pozytywny. Jakie jest prawdopodobieństwo, że jest ona chora?

   Rozwiązanie. A: chora; T: test pozytywny.

   $$P(A|T) = \frac{P(A)P(T|A)}{P(A)P(T|A) + P(A')P(T|A')} = 0,1628 = 16,28\%.$$

   Bardziej "namacalna" ilustracja tego problemu jest taka. Rozważmy 1 000 000 osób. Z nich $0,2\%$, tj. 2 000 osób jest chorych, a test da wynik dodatni dla 1940 osób. 998 000 jest zdrowych , a test da wynik dodatni dla 9980 z nich. Łącznie wynik dodatni ma 11 920 osób, z czego tylko 1940 jest chorych. Prawdopodobieństwo jest równe 1940/11920=16,28%.

5. Jakie jest prawdopodobieństwo, że kobieta w wieku 40-50 lat, która ma dodatni wynik mammografii, rzeczywiście jest chora na raka piersi? Wiadomo, że:
   – mammografia wskazuje na raka u 7% kobiet, które go nie mają
   – mammografia nie wykazuje raka u 10% kobiet, które go mają
   – choroba występuje u 0,8% kobiet

   Rozwiązanie. Niech A będzie zdarzeniem, że kobieta ma raka, a M zdarzeniem, że wynik mammografii jest pozytywny.

   $$\begin{aligned} P(A|M) &= \frac{P(A)P(M|A)}{P(A)P(M|A) + P(A')P(M|A')} = \\ &= \frac{0,9 \cdot 0,008}{0,9 \cdot 0,008 + 0,07 \cdot 0,992} = 9,39\%. \end{aligned}$$

Manipulacja i niezrozumienie. Nieznajomość prawdopodobieństwa powoduje, że jak wykazały badania statystyczne, lekarze znacznie przeceniają wyniki badań. Np. na pytanie postawione w przykładzie 5 większość lekarzy amerykańskich odpowiadała, że prawdopodobieństwo raka to ok. 75%. Poważne skutki mogą też być (i są) na sali sądowej. Np. w głośnym w 1994 r. procesie O.J. Simpsona oskarżonego o zabicie żony, prokuratura prezentowała jako argument historię wcześniejszego znęcania się nad żoną. Obrona przedstawiła następujące rozumowanie: co roku 4 miliony kobiet w USA są bite przez swoich partnerów, ale w 1992 roku ofiarami morderstw popełnionych przez partnerów padły 1432 takie kobiety, czyli 1 na 2500. Dowodzi to, że niewielu domowych tyranów posuwa się do morderstwa.

Ta argumentacja jest manipulacją, bo wprawdzie wyliczone prawdopodobieństwo jest prawdziwe, ale nie o to chodzi. Istotne jest prawdopodobieństwo zdarzenia: kobieta, która była bita i zamordowana, została zamordowana przez partnera. Dane z 1992 roku wskazywały, że to prawdopodobieństwo wynosi ok. 90%. Oskarżeniu jednak zabrakło wiedzy, aby odeprzeć manipulację obrony.

5 Schemat Bernoullego

Rozpatrzmy serię n kolejnych powtórzeń tego samego doświadczenia. W wyniku każdego z tych powtórzeń może zostać zrealizowane pewne zdarzenie A. Jeżeli prawdopodobieństwo wystąpienia zdarzenia A jest w każdym powtórzeniu takie samo, to model probabilistyczny takiej serii doświadczeń nazywamy schematem Bernoullego.
Wystąpienie zdarzenia A nazywamy sukcesem, a zdarzenie A' porażką. Załóżmy, że dla pojedynczej próby $P(A) = p$. Jakie jest prawdopodobieństwo $P(k)$ wystąpienia k sukcesów w n próbach?
Ponieważ kolejne próby są niezależne, więc wystąpienie sekwencji $sppssp \dots p$ sukcesów i porażek z liczbą sukcesów równą k wynosi $p^k(1-p)^{n-k}$. Ale takich sekwencji jest $\binom{n}{k}$, stąd

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Przykład Rzucamy kostką. Jaka musi być liczba rzutów, aby szóstka pojawiła się z prawdopodobieństwem 0,99?
Rozwiązanie. Zdarzenie przeciwne: prawdopodobieństwo, że szóstka się nie pojawi wynosi

$$P(0) = \binom{n}{0} (\frac{1}{6})^0 (\frac{5}{6})^n = (1\frac{5}{6}^{\,n}.$$

Ma być

$$(\frac{5}{6})^n < 0,001,$$

stąd

$$n > \frac{\log 0,01}{\log(\frac{5}{6})} = \frac{-2}{\log(\frac{5}{6})} \approx 22{,}26,$$

a więc $n = 26$.

Przykład Stacja radiolokacyjna wykrywa obiekt będący w polu obserwacji z prawdopodobieństwem 0,93 przy każdym obrocie anteny. Jaka musi być liczba obrotów, aby obiekt został wykryty z prawdopodobieństwem 0,999?

Rozwiązanie. Zdarzenie przeciwne: prawdopodobieństwo, że obiekt nie będzie wykryty wynosi

$$P(0) = \binom{n}{0} p^0 (1-p)^n = (1-p)^n.$$

Ma być

$$(1-p)^n < 0,001,$$

stąd

$$n > \frac{\log 0,001}{\log(1-p)} \approx 2,6,$$

a więc $n = 3$.

6 Zmienne losowe

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną.

Definicja 2 *Zmienną losową nazywamy funkcję*

$$X : \Omega \to \mathbb{R},$$

taką, że dla dowolnego $x \in \mathbb{R}$

$$\{\omega \in \Omega : \ X(\omega) < x\} \in \mathcal{F}.$$

Warunek podany w definicji oznacza, że zbiór zdarzeń elementarnych dla których wartości X są mniejsze od x jest zdarzeniem. W typowych zagadnieniach praktycznych ten warunek jest spełniony, i dlatego najważniejsze jest zapamiętanie, że zmienna losowa przyporządkowuje zdarzeniom liczby. Tradycyjnie zmienne losowe oznaczamy dużymi literami łacińskimi: $X, Y, Z, \ldots$.

Przykład Rozpatrzmy doświadczenie polegające na rzucie dwiema kostkami. Niech ω_{ij} oznacza zdarzenie elementarne polegające na tym, że na pierwszej kostce wypadnie i a na drugiej j. Funkcja określona wzorem

$$X(\omega_{ij}) = i + j$$

jest zmienną losową, bo każdy ze zbiorów $\{X < x\}$, tj.

$$\{\omega_{ij} \in \Omega : \; X(\omega_{ij}) < x\}$$

jest zdarzeniem. Zauważmy np., że dla $x \leqslant 2$ zbiór $\{X < x\}$ jest pusty, bo zmienna X nie może mieć wartości mniejszych niż 2. Jeśli $2 < x \leqslant 3$, to $\{X < x\} = \{\omega_{11}\}$; jeśli $3 < x \leqslant 4$, to $\{X < x\} = \{\omega_{11}, \omega_{12}, \omega_{21}\}$, itd. Jeśli $x > 12$, to zdarzenie $\{X < 12\}$ jest pewne.
Każdemu z powyższych zdarzeń odpowiada jego prawdopodobieństwo. Rozkład prawdopodobieństwa można zapisać w tabeli:

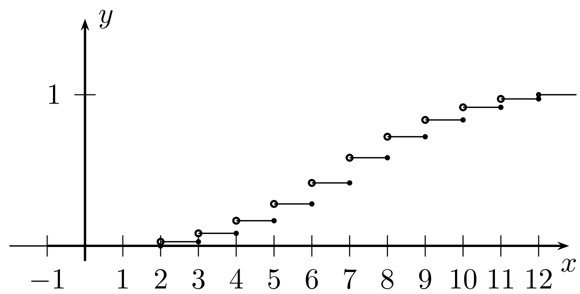
$x \in$	$(-\infty, 2]$	$(2,3]$	$(3,4]$	$(4,5]$	$(5,6]$	$(6,7]$	$(7,8]$	$(8,9]$	$(9,10]$	$(10,11]$	$(11,12]$	$(12,\infty)$
$P(X < x)$	0	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	1

Definicja 3 *Niech X będzie zmienną losową. Funkcję*

$$F : \mathbb{R} \to [0,1], \quad F(x) = P(X < x)$$

nazywamy dystrybuantą zmiennej losowej.

Przykład Powyższa tabela jest w istocie tabelą dystrybuanty. Wykres:



Rysunek 1: Wykres dystrybuanty

Twierdzenie 3 *Jeżeli $F(x)$ jest dystrybuantą zmiennej losowej X, $x_1, x_2 \in \mathbb{R}$ oraz $x_1 < x_2$, to*

$$P(x_1 \leqslant X < x_2) = F(x_2) - F(x_1).$$

D o w ó d.

$$\begin{aligned} P(x_1 \leqslant X < x_2) &= P(X \geqslant x_1 \cap X < x_2) = \\ &= -P(X \geqslant x_1 \cup X < x_2) + P(X \geqslant x_1) + P(X < x_2) = \\ &= -1 + (1 - P(X < x_1)) + P(X < x_2) = F(x_2) - F(x_1). \end{aligned}$$

Twierdzenie 4 *Dystrybuanta $F(x)$ zmiennej losowej jest funkcją niemalejącą, lewostronnie ciągłą. Ponadto*

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1.$$

Ze względu na charakter dystrybuanty wyróżniamy dwa typy zmiennych losowych.

Definicja 4 *Zmienną losową nazywamy* typu dyskretnego *(lub dyskretną, skokową), gdy jej dystrybuanta jest funkcją przedziałami stałą i posiada przeliczalną ilość punktów nieciągłości (skoków).*

Obrazem (przeciwdziedziną) zmiennej typu dyskretnego jest przeliczalny podzbiór zbioru $\mathbb{R}$ składający się z tych wartości argumentu, dla których dystrybuanta nie jest ciągła.
Oznaczmy $P(x_i) = p_i$. Wtedy zbiór par $\{(x_i, p_i)\}$ przekazuje pełną informację o zmiennej losowej skokowej.
Przykład Dla rzutu dwiema kostkami i zmiennej $X(\omega_{ij}) = i + j$ mamy tabelę:

x_i	2	3	4	5	6	7	8	9	10	11	12
p_i	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Funkcję $P(x) \stackrel{\text{def}}{=} P(X = x)$ nazywamy funkcją prawdopodobieństwa zmiennej typu dyskretnego.
Przykład Z partii 100 części, wśród których jest 10 braków, wybrano losowo 5 części. Niech X będzie zmienną losową oznaczającą ilość braków w tej próbie. Wyznaczyć funkcję prawdopodobieństwa.
Zmienna X przyjmuje wartości x_i, $x_i = 0, 1, \ldots, 5$ oraz

$$P(x_i) = \frac{\binom{90}{5-x_i}\binom{10}{x_i}}{\binom{100}{5}}.$$

Tabela (przybliżone wartości do 3 miejsc po przecinku)

x_i	0	1	2	3	4	5
p_i	$0,583$	$0,340$	$0,070$	$0,007$	$0,000$	$0,000$

Rozkład Bernoullego

Ze schematem Bernoullego związana jest zmienna losowa określona jako liczba sukcesów w n próbach. Dokładniej: mówimy, że zmienna losowa ma rozkład Bernoullego, jeżeli przyjmuje wartości $k = 0, 1, 2, \ldots, n$ z prawdopodobieństwem odpowiednio

$$P(p, n, k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Zmienna ta ma dwa parametry: p i n.

Rozkład Poissona

Załóżmy, że w rozkładzie Bernoullego parametry p i n związane są zależnością

$$pn = \lambda = \mathrm{const}$$

i obliczmy granicę, gdy $n \to \infty$ (czyli $p \to 0$):

$$\begin{aligned}
\lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} &= \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} = \\
&= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \lim_{n\to\infty} \frac{n!}{(n-k)!(n-\lambda)^k} = \frac{\lambda^k e^{-\lambda}}{k!}.
\end{aligned}$$

Otrzymany rozkład graniczny, którego prawdopodobieństwo wynosi

$$P(\lambda, k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

nazywamy rozkładem Poissona.
Rozkład Bernoullego zastępujemy rozkładem Poissona, gdyż w tym drugim łatwiej jest wykonywać rachunki. Jest tylko jeden parametr $\lambda = pn$. Praktycznie, gdy $n > 20$ i $p < 0,2$, to niedokładność wynikająca z zastąpienia rozkładu Bernoullego rozkładem Poissona jest niewielka.
Przykłady

1. Partia wyrobów zawiera 3% braków. Z partii losujemy próbę liczącą $n = 100$ sztuk. Obliczyć prawdopodobieństwo, że w próbie znajdzie się $k = 0, 1, \ldots, 100$ braków.

   Rozwiązanie. Stosując rozkład Poissona należy przyjąć $\lambda = 3$. Wtedy

   $$P(3,k) = \frac{3^k e^{-3}}{k!},$$

   więc dla $k = 0, 1, \ldots, 100$ mamy 0,050, 0,149, 0,224, itd.

k	0	1	2	3	4	5	6
$P(3,k)$	0,0498	0,1494	0,2240	0,2240	0,1680	0,1008	0,0504

   ...

...	7	8	9	10	11	12
...	0,0216	0,0081	0,0027	0,0008	0,0002	0,0001

2. Przy transmisji n bitów dodajemy jeszcze jeden bit tak, aby liczba wszystkich jedynek była parzysta. Błąd wykryjemy jeśli podczas przesyłu wystąpi nieparzysta liczba błędów. Obliczyć prawdopodobieństwo niewykrycia błędu, gdy prawdopodobieństwo przekłamania wynosi 10^{-6} a $n = 10^5$. Zakładamy, że przekłamania są niezależne.

   Rozwiązanie. Prawdopodobieństwo niewykrycia jest w przybliżeniu równe

   $$P(X = 2) = e^{-0{,}1}\frac{0,1^2}{2} \approx 0,0045.$$

   Prawdopodobieństwa $P(X = 4)$, $P(X = 6)$, itd. można pominąć, bo są bardzo małe. Np. $P(X = 4) = 0,0000038$.

3. Elementy wadliwe stanowią średnio 1% wszystkich elementów. Znaleźć prawdopodobieństwo tego, że wśród 200 elementów będzie więcej niż 3 wadliwe.

   Rozwiązanie. Stosujemy rozkład Poissona z $\lambda = 200 \cdot 0,01 = 2$.

   $$\begin{aligned} p &= 1 - P(2,0) - P(2,1) - P(2,2) - P(2,3) = \\ &= 1 - e^{-2}\left(\frac{1}{1} + \frac{2}{1} + \frac{2^2}{2!} + \frac{2^3}{3!}\right) = 0,1429. \end{aligned}$$

7 Wartość oczekiwana zmiennej losowej

Definicja 5 *Jeżeli zmienna losowa X może przyjmować wartości $x_1, x_2, \ldots$ z prawdopodobieństwami $p_1, p_2, \ldots$ odpowiednio, i jeżeli szereg $\sum_{i=1}^{\infty} p_i x_i$ jest* *bezwzględnie zbieżny, to jego sumę nazywamy* wartością oczekiwaną (przeciętną) *zmiennej losowej* X *i oznaczamy* $\mathrm{E}(X)$ *lub* $\mathrm{E}X$.

Przykład Student potrafi odpowiedzieć na a) 5 z 10 pytań egzaminacyjnych; b) 10 z 20 pytań egzaminacyjnych. Na egzaminie losuje 3 pytania. Jaka jest wartość oczekiwana liczby pytań na które odpowie?
Rozwiązanie. a)

x_i	0	1	2	3
p_i	$\frac{1}{12}$	$\frac{5}{12}$	$\frac{5}{12}$	$\frac{1}{12}$

Zatem

$$\mathrm{E}X = \frac{5 + 2 \cdot 5 + 3 \cdot 1}{12} = \frac{18}{12} = 1{,}5.$$

b)

x_i	0	1	2	3
p_i	$\frac{12}{114}$	$\frac{45}{114}$	$\frac{45}{114}$	$\frac{12}{114}$

Zatem

$$\mathrm{E}X = \frac{45 + 2 \cdot 45 + 3 \cdot 12}{114} = \frac{171}{114} = 1{,}5.$$

Własności wartości oczekiwanej.

1. Jeżeli zmienna losowa $X = c =$const, to $\mathrm{E}X = c$.
2. $\mathrm{E}(cX) = c\mathrm{E}X$.
3. $\mathrm{E}(X + Y) = \mathrm{E}X + \mathrm{E}Y$.
4. Jeżeli zmienne losowe X, Y są niezależne, to $\mathrm{E}(XY) = \mathrm{E}X \cdot \mathrm{E}Y$.

Uwaga. Zmienne losowe X, Y nazywamy niezależnymi, gdy

$$P(X < x \wedge Y < y) = P(X < x) \cdot P(Y < y).$$

Przykład Trzy przyrządy niezależnie od siebie podlegają kontroli. Prawdopodobieństwo, że poszczególne przyrządy nie będą działały są równe odpowiednio p_1, p_2, p_3. Wykazać, że wartość oczekiwana liczby niedziałających przyrządów jest równa $p_1 + p_2 + p_3$.
Rozwiązanie. Określamy zmienne X_1, X_2, X_3 wzorem

$$X_i = \begin{cases} 1 & i - \text{ty przyrząd nie działa} \\ 0 & i - \text{ty przyrząd działa} \end{cases}$$

Wtedy $\mathrm{E}X_i = p_i$ dla $i = 1, 2, 3$. Niech Z będzie liczbą niedziałających przyrządów: $Z = X_1 + X_2 + X_3$. Wtedy

$$\mathrm{E}Z = \mathrm{E}X_1 + \mathrm{E}X_2 + \mathrm{E}X_3 = p_1 + p_2 + p_3.$$

Korzystając z definicji obliczymy wartości oczekiwane niektórych zmiennych losowych.

Wartość oczekiwana rozkładu Bernoullego

Wiemy, że $P(p, n, k) = \binom{n}{k} p^k (1-p)^{n-k}$ dla $k = 0, 1, \ldots, n$. Zatem

$$\mathrm{E}X = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}.$$

Aby obliczyć tę sumę korzystamy z dwumianu Newtona.

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \qquad \Big| \frac{d}{dp}$$

$$n(p+q)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} p^{k-1} q^{n-k} \qquad | \cdot p$$

$$pn(p+q)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k}.$$

Podstawiając $q = 1 - p$ otrzymujemy

$$np = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k},$$

czyli

$$\mathrm{E}X = np.$$

Wartość oczekiwana rozkładu Poissona

Ponieważ $P(\lambda, k) = \frac{\lambda^k e^{-\lambda}}{k!}$ dla $k = 0, 1, 2, \ldots$, więc

$$\mathrm{E}X = \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

8 Wariancja zmiennej losowej

Rozważmy dwa zakłady o wynik rzutu monetą:

1. jeżeli wypadnie orzeł wygrywamy 1 zł, jeżeli reszka tracimy 1 zł;
2. jeżeli wypadnie orzeł wygrywamy 100 zł, jeżeli reszka tracimy 100 zł.

Wartość oczekiwana odpowiednich zmiennych losowych jest w obu przypadkach taka sama (równa 0), ale w pierwszym przypadku wartości zmiennej nie odbiegają zbytnio od wartości oczekiwanej, a w drugim — tak. Druga zmienna jest bardziej "rozproszona".

Definicja 6 Wariancją *zmiennej losowej nazywamy liczbę*

$$\mathrm{Var}X = \mathrm{E}(X - \mathrm{E}X)^2.$$

Spotyka się także oznaczenie D^2X.
Zauważmy, że $\mathrm{E}(X - \mathrm{E}X)^2 = \mathrm{E}(X^2 - 2X\mathrm{E}X + (\mathrm{E}X)^2) = \mathrm{E}(X^2) - (\mathrm{E}X)^2$, zatem

$$\mathrm{Var}X = \mathrm{E}(X^2) - (\mathrm{E}X)^2.$$

Dla zmiennych wspomnianych wyżej mamy $\mathrm{Var}(X) = \mathrm{E}(X^2)$ (bo $\mathrm{E}X = 0$). W pierwszym przypadku jest to 1, w drugim $100^2 = 10000$.
Wariancja ma wymiar kwadratu zmiennej losowej. Czasem wygodniej jest posługiwać się pierwiastkiem z wariancji. Nazywa się go *odchyleniem standardowym* i oznacza σ_X:

$$\sigma_X = \sqrt{\mathrm{Var}X}.$$

Twierdzenie 5 *Niech X ma rozkład Bernoullego. Wtedy*

$$\mathrm{Var}X = npq, \quad \text{gdzie } q = 1 - p.$$

D o w ó d.

$$\begin{aligned}
\mathrm{Var}X &= \mathrm{E}(X^2) - (\mathrm{E}X)^2 = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} - (np)^2 = \\
&= \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k} - (np)^2 = \\
&= np \sum_{k=0}^{n-1} (k+1) \frac{(n-1)!}{k!(n-k-1)!} p^k q^{n-k-1} - (np)^2 = \\
&= np \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} p^k q^{n-k-1} - (np)^2 = \\
&= np\left((n-1)p + (p+q)^{n-1}\right) - (np)^2 = np(np-p+1) - (np)^2 = \\
&= np(np+q) - (np)^2 = n^2p^2 + npq - (np)^2 = npq.
\end{aligned}$$

Twierdzenie 6 *Niech X ma rozkład Poissona z parametrem λ. Wtedy*

$$\mathrm{Var}X = \lambda.$$

D o w ó d.

$$\begin{aligned}
\mathrm{Var}X &= \mathrm{E}(X^2) - (\mathrm{E}X)^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k e^{-\lambda}}{k!} - \lambda^2 = \\
&= e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} - \lambda^2 = e^{-\lambda} \lambda \sum_{k=0}^{\infty} (k+1) \frac{\lambda^k}{k!} - \lambda^2 = \\
&= \lambda^2 + \lambda - \lambda^2 = \lambda
\end{aligned}$$

9 Zmienne losowe typu ciągłego

Niech X będzie zmienną losową. Jeżeli dystrybuantę $F(x)$ można przedstawić w postaci

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

to zmienną losową X nazywamy *zmienną losową typu ciągłego*. Każdą funkcję $f(t)$ dla której spełniona jest powyższa równość nazywamy *gęstością* zmiennej losowej.
Jeżeli $F(x)$ jest funkcją różniczkowalną, to $F'(x) = f(x)$.

Twierdzenie 7 *Funkcja $f(x)$ jest gęstością zmiennej losowej wtedy, i tylko wtedy, gdy:*

1. $f(x) \geqslant 0$*;*

2. $\int_{-\infty}^{\infty} f(x)dx = 1$*.*

Przykłady

1 . Dobrać stałą a tak, aby funkcja

$$f(x) = \begin{cases} a \sin x & , \quad x \in [0, \pi] \\ 0 & , \quad x \in \mathbb{R} \setminus [0, \pi] \end{cases}$$

była funkcją gęstości.

2. Sprawdzić, że funkcja

$$F(x) = \begin{cases} 1 - e^{-x^2} & , \quad x \geqslant 0 \\ 0 & , \quad x < 0 \end{cases}$$

jest dystrybuantą zmiennej losowej X. Wyznaczyć $P(X > 1)$, $P(1 < X \leqslant 2)$, $P(X = 3)$.

Definicja 7 Wartość oczekiwaną *zmiennej losowej typu ciągłego określamy następująco*

$$\mathrm{E}X = \int_{-\infty}^{\infty} x f(x) dx,$$

o ile całka jest zbieżna.

Jeżeli nowa zmienna losowa Y jest funkcją zmiennej X, np. $Y = g(X)$, to można wykazać, że

$$\mathrm{E}Y = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Wariancję określamy jak poprzednio:

$$\mathrm{Var}X = \mathrm{D}^2 X = \mathrm{E}(X - \mathrm{E}X)^2 = \mathrm{E}X^2 - (\mathrm{E}X)^2.$$

Omówimy teraz niektóre rozkłady typu ciągłego.

Definicja 8 Rozkładem jednostajnym *na przedziale* $[a, b]$ *nazywamy rozkład określony funkcją gęstości*

$$f(x) = \begin{cases} \frac{1}{b-a} & , \quad x \in [a, b] \\ 0 & , \quad x \notin [a, b] \end{cases}$$

Dystrybuanta wynosi więc

$$F(x) = \int_a^x \frac{dt}{b-a} = \frac{x-a}{b-a} \quad \text{dla } x \in [a,b]$$

Dla $x < a$ jest $F(x) = 0$, dla $x > b$ jest $F(x) = 1$.
Możemy obliczyć wartość oczekiwaną:

$$\mathrm{E}X = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2},$$

i wariancję:

$$\mathrm{Var}X = \int_a^b x^2 \frac{1}{b-a} dx - (\mathrm{E}X)^2 = \frac{b^3-a^3}{3(b-a)} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Przykłady

1. Załóżmy, że napięcie $U = U_{\max} \sin\varphi$ prądu zmiennego ma losową fazę φ o rozkładzie jednostajnym na przedziale $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Znaleźć dystrybuantę i gęstość napięcia U.

   Mamy

$$F(u) = \begin{cases} 0 & \text{dla} \quad u \leqslant -U_{\max} \\ P\Big(\varphi < \arcsin \frac{u}{U_{\max}}\Big) & \text{dla} \quad -U_{\max} < u < U_{\max} \\ 1 & \text{dla} \quad u \geqslant U_{\max} \end{cases},$$

   bo

$$P\Big(U_{\max} \sin\varphi < u\Big) = P\Big(\varphi < \arcsin \frac{u}{U_{\max}}\Big).$$

   Ale

$$P(\varphi < \varphi_0) = \frac{\varphi_0 + \pi/2}{\pi} = \frac{1}{\pi}\varphi_0 + \frac{1}{2}$$

   dla $|\varphi_0| < \pi/2$ (najprostszy sposób obliczenia: zauważyć, że to prawdopodobieństwo można interpretować jako pole prostokąta o podstawie $\varphi_0 + \pi/2$ i wysokości $1/\pi$). Zatem

$$P\Big(\varphi < \arcsin \frac{u}{U_{\max}}\Big) = \frac{1}{\pi} \arcsin \frac{u}{U_{\max}} + \frac{1}{2},$$

   więc

$$f(u) = \frac{1}{\pi} \frac{1}{\sqrt{1 - \frac{u^2}{U_{\max}^2}}} \frac{1}{U_{\max}} = \frac{1}{\pi} \frac{1}{\sqrt{U_{\max}^2 - u^2}} \quad \text{dla } |u| < U_{\max}$$

   oraz

$$f(u) = 0 \quad \text{dla } |u| \geqslant U_{\max}$$

2. X ma rozkład jednostajny na $[-1,2]$. Znaleźć dystrybuantę i gęstość zmiennej losowej $Y = X^2$.

$$F(y) = \begin{cases} 0 & \text{dla} \quad y \leqslant 0 \\ P(X^2 < y) & \text{dla} \quad 0 < y \leqslant 4 \\ 1 & \text{dla} \quad y > 4 \end{cases}$$

ale

$$\begin{aligned} P(X^2 < y) &= P(|X| < \sqrt{y}) = \\ &= \begin{cases} P(-\sqrt{y} < X < \sqrt{y}), & 0 < y \leqslant 1 \\ P(X < \sqrt{y}), & 1 < y \leqslant 4 \end{cases} = \\ &= \begin{cases} \frac{2}{3}\sqrt{y} & , \quad 0 < y \leqslant 1 \\ \frac{1+\sqrt{y}}{3} & , \quad 1 < y \leqslant 4 \end{cases}. \end{aligned}$$

Zatem gęstość wynosi

$$f(y) = \begin{cases} 0 & , \quad y \leqslant 0 \\ \frac{1}{3\sqrt{y}} & , \quad 0 < y \leqslant 1 \\ \frac{1}{6\sqrt{y}} & , \quad 1 < y \leqslant 4 \\ 0 & , \quad y > 4 \end{cases}.$$

Definicja 9 Rozkładem wykładniczym *nazywamy rozkład określony funkcją gęstości*

$$f(x) = \begin{cases} 0 & \text{dla} \quad x < 0 \\ \lambda e^{-\lambda x} & \text{dla} \quad x \geqslant 0 \end{cases}$$

gdzie λ *jest parametrem rozkładu.*

Zatem dystrybuanta wynosi

$$F(x) = \begin{cases} 0 & \text{dla} \quad x < 0 \\ 1 - e^{-\lambda x} & \text{dla} \quad x \geqslant 0 \end{cases}$$

Można sprawdzić, że

$$\mathrm{E}X = \frac{1}{\lambda}, \quad \mathrm{Var}X = \frac{1}{\lambda^2}.$$

Rozkład wykładniczy jest stosowany w teorii niezawodności, bo dobrze opisuje czas pracy elementów niestarzejących się. Jeżeli T jest czasem pracy takiego elementu, to $\tau = \frac{1}{\lambda}$ jest *średnim czasem pracy*, a parametr $\lambda = \frac{1}{\tau}$ jest nazywany *intensywnością uszkodzeń*. Niestarzenie się elementu oznacza, że prawdopodobieństwo awarii w danym przedziale czasu nie zależy od wieku elementu. Określamy tę własność jako *brak pamięci* rozkładu wykładniczego. Mamy więc

$$P(T > t_0 + t | T > t_0) = P(T > t),$$

co łatwo sprawdzić bezpośrednio, bo

$$P(T > t) = 1 - P(T \leqslant t) = 1 - (1 - e^{-\lambda t}) = e^{-t/\tau}$$

więc

$$P(T > t_0 + t | T > t_0) = \frac{P(T > t_0 + t)}{P(T > t_0)} = \frac{e^{-\frac{t+t_0}{\tau}}}{e^{-\frac{t_0}{\tau}}} = e^{-t/\tau}.$$

Przykład Zmienna losowa X ma rozkład wykładniczy o wartości oczekiwanej 1. Znaleźć gęstość zmiennej losowej $Y = \ln X$ oraz obliczyć $P(Y > 0)$.
Rozwiązanie. Obliczamy dystrybuantę:

$$G(y) = P(Y < y) = P(\ln X < y) = P(X < e^y) = F(e^y) = 1 - e^{-e^y}.$$

Zatem

$$g(y) = \left(1 - e^{-e^y}\right)' = e^{-e^y} e^y = e^{y - e^y},$$

oraz

$$P(Y > 0) = 1 - P(Y \leqslant 0) = 1 - \left(1 - e^{-e^0}\right) = e^{-1} = 0,3679.$$

Definicja 10 Rozkładem normalnym *nazywamy rozkład określony funkcją gęstości*

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Mówimy wtedy, że zmienna losowa ma rozkład $N(0, 1)$.

Uwaga: można wykazać, że:

$$\int_{-\infty}^{\infty} e^{-x^2/2} = \sqrt{2\pi}.$$

Jeżeli zmienna losowa Y ma rozkład $N(0, 1)$, to zmienna

$$X = \sigma Y + m$$

ma rozkład $N(m, \sigma)$. Piszemy wtedy $X \sim N(m, \sigma)$ i mówimy, że zmienna ma rozkład normalny z parametrami m, σ.

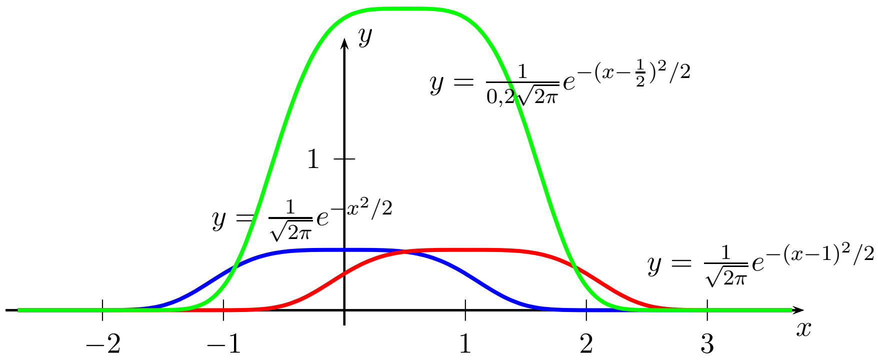
Odwrotnie, jeżeli $X \sim N(m, \sigma)$, to zmienna

$$\tilde{X} = \frac{X - m}{\sigma} \sim N(0, 1).$$

To przekształcenie nazywa się *standaryzacją* zmiennej losowej X.
Można wykazać, że gęstość zmiennej losowej o rozkładzie $N(m, \sigma)$ wynosi

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}.$$

Wykres tej funkcji nazywamy *krzywą Gaussa*. Osią symetrii wykresu jest prosta $x = m$. Dla $x = m$ funkcja osiąga maksimum.



Rysunek 2: Wykresy krzywych Gaussa

Dla rozkładu $N(0, 1)$ mamy $\mathrm{E}X = 0$, bo funkcja podcałkowa $xe^{-x^2/2}$ jest funkcją nieparzystą. Natomiast

$$\begin{aligned} \mathrm{E}X^2 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x \cdot xe^{-x^2/2} dx = \\ &= \frac{2}{\sqrt{2\pi}} \left[-xe^{-x^2/2} |_0^{\infty} + \int_0^{\infty} e^{-x^2/2} dx \right] = \frac{2}{\sqrt{2\pi}} \cdot \frac{\sqrt{2\pi}}{2} = 1. \end{aligned}$$

Stąd otrzymujemy

$$\mathrm{Var}X = \mathrm{E}X^2 - (\mathrm{E}X)^2 = 1.$$

Zatem dla zmiennej $X \sim N(m, \sigma)$ mamy:

$$\mathrm{E}X = \mathrm{E}(\sigma\tilde{X} + m) = \sigma \cdot 0 + m = m,$$

$$\mathrm{Var}X = \mathrm{Var}(\sigma\tilde{X} + m) = \sigma^2 \mathrm{Var}\tilde{X} = \sigma^2.$$

Wynika z tego, że rozkład normalny jest całkowicie określony przez dwa parametry $m = \mathrm{E}X$ i $\sigma = \sqrt{\mathrm{Var}X}$.
Dystrybuantę zmiennej losowej $X \sim N(0,1)$ oznacza się $\Phi(x)$. Jest to funkcja nieelementarna. Jej wartości dostępne są w tablicach statystycznych lub programach komputerowych. Tablice zawierają na ogół wartości

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \quad \text{dla } x > 0.$$

Aby obliczyć wartość dla ujemnego x korzystamy ze wzoru

$$\Phi(-x) = 1 - \Phi(x), \quad \text{dla } x > 0.$$

Wynika on z symetrii funkcji gęstości, bo $\int_{-\infty}^{-x} e^{-t^2/2} dt = \int_{x}^{\infty} e^{-t^2/2} dt$.
Następująca własność znana jest jako reguła 3-sigmowa.

Twierdzenie 8 *Jeżeli zmienna losowa X ma rozkład normalny $N(m, \sigma)$, to*

$$P(|X - m| > 3\sigma) < 0{,}01.$$

D o w ó d. Niech $\tilde{X} = \frac{X-m}{\sigma}$ będzie zmienną standaryzowaną, $\tilde{X} \sim N(0,1)$. Wtedy

$$\begin{aligned} P(|X - m| > 3\sigma) &= P\Big(\Big|\frac{X-m}{\sigma}\Big| > 3\Big) = P(|\tilde{X}| > 3) = \\ &= 1 - P(|\tilde{X}| \leqslant 3) = 1 - (\Phi(3) - \Phi(-3)) = \\ &= 1 - \Phi(3) + 1 - \Phi(3) = 2(1 - \Phi(3)) = 0,0027 < 0{,}01. \end{aligned}$$

Przykłady

1. Pomiar odległości obiektu obarczony jest błędem systematycznym i losowym. Błąd systematyczny wynosi 50 m w stronę zaniżenia odległości. Błędy losowe mają rozkład normalny z odchyleniem standardowym $\sigma = 100$ m. Znaleźć:
   1) prawdopodobieństwo pomiaru odległości z błędem niewiększym niż 150 m;
   2) prawdopodobieństwo, że zmierzona odległość nie przekroczy prawdziwej odległości.

   Rozwiązanie.

   1) Niech X będzie sumarycznym błędem pomiaru. Błąd systematyczny wynosi -50 m i jest to wartość oczekiwana zmiennej losowej X. Zatem

$X \sim N(-50, 100)$. Obliczamy

$$\begin{aligned} P(|X| < 150) &= P(-150 < X < 150) = P(-100 < X + 50 < 200) = \\ &= P\Big(-1 < \frac{X+50}{100} < 2\Big) = \Phi(2) - \Phi(-1) = \\ &= \Phi(2) - 1 + \Phi(1) = 0,9772 - 1 + 0{,}8413 = 0{,}8185. \end{aligned}$$

2) $P(X < 0) < P(X + 50 < 50) = P(\frac{X+50}{100} < \frac{1}{2}) = \Phi(0{,}5) = 0{,}6915.$

2. Przedmiot jest traktowany jak produkt wysokiej jakości, jeżeli wartość bezwzględna odchylenia jego rozmiarów od rozmiarów nominalnych nie przekracza 2,5 mm. Losowe odchylenia rozmiarów przedmiotu od rozmiarów nominalnych mają rozkład $N(0; 2{,}1)$. Określić średnią ilość produktów wysokiej jakości, jeżeli wytworzono 100 przedmiotów.

   Rozwiązanie. Niech X będzie odchyleniem rozmiaru od rozmiaru nominalnego. Potraktujemy próbę 100 przedmiotów jako próbę Bernoullego i wyliczymy prawdopodobieństwo sukcesu p:

$$\begin{aligned} p &= P(|X| < 2{,}5) = P(2{,}5 < X < 2{,}5) = P\Big(-\frac{25}{21} < \frac{X}{2{,}1} < \frac{25}{21}\Big) = \\ &= \Phi\Big(\frac{25}{21}\Big) - \Phi\Big(-\frac{25}{21}\Big) = 2\Phi(1{,}19) - 1 = 0{,}766. \end{aligned}$$

   Zatem wartość oczekiwana liczby sukcesów w 100 próbach wynosi $np = 100 \cdot 0{,}766 = 76{,}6$.

3. Odchylenie losowe wymiaru detalu wyprodukowanego na danej maszynie od wymiaru nominalnego ma zerową wartość oczekiwaną i odchylenie standardowe równe 5μm. Ile trzeba wyprodukować detali aby z prawdopodobieństwem niemniejszym niż 0,9 był wśród nich choćby 1 detal dobry, jeżeli dla detalu dobrego dopuszczalne jest odchylenie wymiaru od wymiaru nominalnego niewiększe niż 2μm?

   Rozwiązanie. Obliczymy najpierw prawdopodobieństwo wyprodukowania detalu dopuszczalnego.

$$\begin{aligned} P(|X| < 2) &= P(-2 < X < 2) = P\Big(-\frac{2}{5} < \frac{X}{5} < \frac{2}{5}\Big) = \\ &= \Phi(\frac{2}{5}) - \Phi(-\frac{2}{5}) = 2\Phi(\frac{2}{5}) - 1 = 0{,}3108. \end{aligned}$$

   Stosujemy teraz schemat Bernoullego. Prawdopodobieństwo zera sukcesów w n próbach wynosi $(0{,}6892)^n$, więc n musi spełniać nierówność

$1 - (0{,}6892)^n \geqslant 0{,}9$. Zatem

$$n \geqslant \frac{\log 0{,}1}{\log 0{,}6892} = \frac{-1}{\log 0{,}6892} = 6{,}19.$$

Należy wyprodukować 7 detali.

Rozkład χ^2.

Niech X_i, $i = 1, 2, \dots, n$ będą niezależnymi zmiennymi losowymi o rozkładzie $N(0, 1)$. Określamy

$$\chi^2 = \chi^2_n = \sum_{i=1}^{n} X_i^2$$

i mówimy, że zmienna losowa χ^2 ma *rozkład chi-kwadrat* o n stopniach swobody.

Zmienna χ^2 ma dla małych n rozkład asymetryczny, ale wraz ze wzrostem n jej rozkład zbliża się do rozkładu normalnego. Przyjmuje się, że dla $n > 30$ dystrybuanta $\Phi(x)$ rozkładu normalnego przybliża dystrybuantę rozkładu chi-kwadrat wystarczająco dobrze.