Notatki do wykładu
„Rachunek prawdopodobieństwa dla informatyków”.

Marcin Milewski

Wrocław, 4 lutego 2009

Spis treści

Niektóre z wykorzystanych w tej pracy obrazków pochodzą z Wikipedii i są one dostępne jako public domain.

1 Prawdopodobieństwo

1.1 Oznaczenia i pojęcia

Ω - przestrzeń zdarzeń elementarnych.

Prawdopodobieństwo będzimy liczyć tylko dla niektórych podzbiorów Ω. Rodzinę tych podzbiorów oznaczamy przez $\mathcal{F}$. Zbiory z tej rodziny spełniają:

1. $A \in \mathcal{F} \Rightarrow A' \in \mathcal{F}$ (jeżeli zbiór należy do rodziny, to jego dopełnienie także)
2. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (jeżeli ileś elementów należy do rodziny to ich suma także)

Uwaga 1. Jeżeli Ω jest zbiorem przeliczalnym, to dobrym kandydatem na $\mathcal{F}$ jest 2^{Ω}.

Prawdopodobieństwo - funkcja $P\colon \mathcal{F} \to [0,1]$

$(\Omega, \mathcal{F}, P)$ - przestrzeń probabilistyczna.

Funkcja P spełnia:

1. $P(A) \in [0,1]$
2. $P(\Omega) = 1$
3. jeśli $A_1, A_2, \ldots \in \mathcal{F}$ i są parami rozłączne to $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Fakt 1. *Niech dana będzie przestrzeń probabilistyczna* $(\Omega, \mathcal{F}, P)$ *oraz* $A, B, A_1, A_2, \ldots, A_n \in \mathcal{F}$. *Wtedy:*

1. $P(\emptyset) = 0$
2. *jeśli* $A_i \cap A_j = \emptyset$ *dla* $i, j \in \{1, 2, \ldots, n\}, i \neq j$ *to* $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$
3. $P(A') = 1 - P(A)$
4. $A \subset B \Rightarrow P(B \setminus A) = P(B) - P(A)$
5. $A \subset B \Rightarrow P(A) \leq P(B)$
6. $P(A) \leq 1$
7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Dowód. ad 7.

$$\begin{aligned} A \cup B &= (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B) \\ P(A \cup B) &= P(A \setminus (A \cap B)) + P(B \setminus (A \cap B)) + P(A \cap B) \\ &= P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B) = P(A) + P(B) - P(A \cap B) \end{aligned}$$

□

1.2 Zasada włączeń i wyłączeń

Zasada włączeń i wyłączeń przydaje się, kiedy łatwiej niż sumę zdarzeń jest policzyć ich przekrój.

$$\begin{aligned} P(A_1 \cup A_2 \cup \ldots \cup A_n) = & \sum_{i=1}^{n} P(A_i) \\ & - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) \\ & + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) \\ & - \cdots \\ & + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \end{aligned} \tag{1}$$

1.3 Podstawe schematy kombinatoryczne

Przykład 2. Sieć komputerowa składa się z n różnych komputerów połączonych szeregowo. Na ile sposobów możemy ustawić komputery?

Rozwiązanie 3. $n(n-1)(n-2)\cdots 2 \cdot 1 = n!$ $\boxed{\text{PERMUTACJA}}$

Przykład 4. Załóżmy, że k z powyższych komputerów stoi w naszym pokoju, Na ile sposobów możemy skonfigurować sieć, gdy interesuje nas tylko nasz pokój?

Rozwiązanie 5. $n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$ $\boxed{\text{WARIACJA}}$

Przykład 6. Słowo maszynowe składa się z 32 bitów. Ile różnych słów można z tego uzyskać?

Rozwiązanie 7. $\underbrace{2 \cdot 2 \cdot \ \cdots \ \cdot 2}_{n \text{ dwójek}} = 2^{32}$ $\boxed{\text{WARIACJA Z POWTÓRZENIAMI}}$

Przykład 8. W jednym cyklu obliczeń program na wyjściu podaje zbiór k różnych liczb z zakresu $1, 2, \ldots, n$. Ile jest różnych wyników?

Rozwiązanie 9.

$$\text{kolejność liczb } \underline{\text{ma}} \text{ znaczenie} \longrightarrow \frac{n!}{(n-k)!}$$

$$\text{kolejność liczb } \underline{\text{nie ma}} \text{ znaczenia} \longrightarrow \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

$\boxed{\text{KOMBINACJA}}$

1.4 Prawdopodobieństwo geometryczne

Kiedy trzeba policzyć prawdopodobieństwo dla nieskończonego zbioru A lub Ω (gdzie oczywiście $A \subset \Omega$), wtedy za miarę uznajemy pole powierzchni, będące geometryczną reprezentacją zdarzeń A oraz Ω.

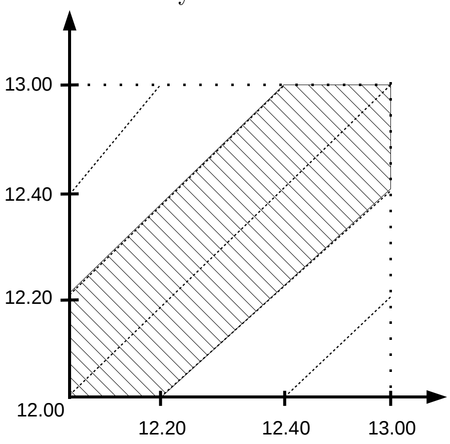
Przykład 10. Jaś i Małgosia umówili się pomiędzy 12:00 i 13:00. Osoba, która przyjdzie wcześniej czeka na drugą 20 minut. Jaka jest szansa spotkania się naszych bohaterów?

Rozwiązanie 11.

$$\Omega = [12, 13] \times [12, 13]$$

$$A - \text{spotkanie Jasia i Małgosi}$$

Rysunek 1



Zakreskowana część oznacza, że bohaterowie się spotkają.

$$P(A) = \frac{\text{pole zakreskowanej figury}}{\text{pole kwadratu o boku 1}} = \frac{2(\frac{1}{2}\cdot 1\cdot 1 - \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{2}{3})}{1\cdot 1} = \frac{5}{9}$$

Przykład 12. Rzucamy dwiema kośćiami do gry. Mamy kilka możliwości zdefiniowania przestrzeni zdarzeń (Ω).

1. jeśli interesuje nas co wypadło na której kostce to

$$\Omega = \{(i,j)\colon i,j \in \{1,2,\dots,6\}\}$$
$$|\Omega| = 6\cdot 6 = 36$$

2. jeżeli interesuje nas co wypadło na obu kościach (bez rozróżniania) to

$$\Omega = \{\{i,j\}\colon i,j \in \{1,2,\dots,6\}\}$$
$$|\Omega| = 6+5+4+3+2+1 = 21$$

3. jeżeli interesuje nas co wypadło na której kostce, ale chcemy mieć wąsy to musimy oznakować wartości na każdej z kostek

$$\Omega = \{\{i,j\}\colon i \in \{1',2',\dots,6'\}, j\in\{1'',2'',\dots,6''\}\}$$
$$|\Omega| = 6\cdot 6 = 36$$

Przyjmijmy Ω z punktu 1. Nazwijmy nasze kostki *zielona* i *czerwona* Jaka jest szansa, że na *zielonej* wypadła przysta liczba oczek a na *czerwonej* liczba oczek większa niż 4? (naturalne jest założenie o niezależności kostek).

$$A = \{(2,5),(2,6),(4,5),(4,6),(6,5),(6,6)\}$$
$$|A| = 6$$
$$A_1 - \text{na pierwszej kostce wypadła liczba parzysta}$$
$$A_2 - \text{na drugiej kostce wypadła liczba większa niż 4}$$
$$P(A_1) = \frac{|A_1|}{|\Omega|} = \frac{3}{6} \qquad P(A_2) = \frac{|A_2|}{|\Omega|} = \frac{2}{6}$$
$$P(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{3}{6}\cdot\frac{2}{6} = P(A_1)\cdot P(A_2)$$
$$A = A_1\cap A_2,\ P(A_1\cap A_2) = P(A_1)\cdot P(A_2)$$

Definicja 1. Zdarzenia $A_1, A_2, \dots, A_n$ nazywamy niezależnymi, gdy

$$P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i) \text{ dla każdego } I \subseteq \{1,2,\dots,5\} \tag{2}$$

Fakt 13. *Jeżeli zdarzenia A, B są niezależne oraz rozłączne, to*

$$P(A) = 0 \textit{ lub } P(B) = 0$$

Dowód.

$$P(A \cap B) = P(A) \cdot P(B) \text{ ale } A \cap B = \emptyset \text{ więc } P(A \cap B) = P(\emptyset) = 0$$

□

Uwaga 2. (nie wystarczy do niezależności analizować tylko pary). W urni są 4 kulki: czerwona, niebieska, żółta, czerwono-niebiesko-żółta.

$$\begin{aligned} A- &\text{ wyciągnięto kulę z kolorem czerwonym} \\ B- &\text{ wyciągnięto kulę z kolorem niebieskim} \\ C- &\text{ wyciągnięto kulę z kolorem zółtym} \\ P(A) = P(B) = P(C) &= \frac{2}{4} \\ P(A \cap B) &= \frac{1}{4} = P(A) \cdot P(B) = P(A \cap C) = P(B \cap C) \\ P(A \cap B \cap C) &= \frac{1}{4} \neq P(A) \cdot P(B) \cdot P(C) \end{aligned}$$

Fakt 14. *Jeśli zdarzenia $A_1, A_2, \ldots, A_n$ są niezależne to niezależne są też zdarzenia $B_1, B_2, \ldots, B_n$, gdzie $B_i = A_i$ lub $B_i = A_i'$.*

1.5 Prawdopodobieństwo warunkowe

Przykład 15. Sieć składa się z dwóch rozróżnialnych komputerów. W danej chwili komputer może wykonywać obliczenia/pracować albo czekać (z prawdopodobieństwem $\frac{1}{2}$). Jaka jest szansa tego, że komputer 1 jest w stanie uśpienia, gdy wiemy, ze sieć pracuje?

Rozwiązanie 16.

$$\begin{aligned} &\Omega = \{(0,0),(0,1),(1,0),(1,1)\} \\ &A- \text{ pierwszy komputer jest w stanie uśpienia} \\ &A = \{(0,1),(0,0)\} \end{aligned}$$

1. bez dodatkowej wiedzy: $P(A) = \frac{|A|}{|\Omega|} = \frac{1}{2}$

2. z dodatkową wiedzą

$$\begin{aligned} A \cap B &= \{(0,1)\} \\ B- &\text{ sieć pracuje} \qquad B = \{(0,1),(1,0),(1,1)\} \qquad |B| = 3 \\ P(A|B) &= \frac{1}{3} = \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A \cap B)}{P(B)} \end{aligned}$$

Definicja 2. Prawdopodobieństo warunkowe zajścia zdarzenia A pod warunkiem zajścia zdarzenia B zapisujemy

$$P(A|B) = \frac{P(A \cap B)}{P(B)},\ P(B) \neq 0 \tag{3}$$

Fakt 17. *Jeśli $P(A_1 \cap A_2 \cap \ldots \cap A_n) > 0$ to*
$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_n)$.

Dowód.

$$\begin{aligned} \text{prawa strona: } &P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdot \ldots \cdot \frac{P(A_1 \cap A_2 \cap \cdots \cap A_n)}{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1})} \\ &= P(A_1 \cap A_2 \cap \ldots \cap A_n) \end{aligned}$$

□

1.6 Prawdopodobieństwo całkowite

Przykład 18. W urnie mamy b kul białych i c kul czarnych. Wyciągamy 1 kulę i (nie podglądając) wyrzucamy ją. Jaka jest szansa, że w następnym ciągnięciu wylosujemy kulę białą?

Rozwiązanie 19.

$$\begin{aligned} B_1 - &\ \text{wyciągnięto białą kulę w pierwszym losowaniu} \\ B_2 - &\ \text{wyciągnięto białą kulę w drugim losowaniu} \\ P(B_2) &= P(B_2 \cap B_1) + P(B_2 \cap B_1^{'}) = P(B_2|B_1) \cdot P(B_1) + P(B_2|B_1') \cdot P(B_1') \\ &= \frac{b-1}{b+c-1} \cdot \frac{b}{b+c} + \frac{b}{b+c-1} \cdot \frac{c}{b+c} = \frac{b}{c+b} \end{aligned}$$

Definicja 3. Rozbiciem przestrzeni zdarzeń elementarnych nazywamy rodzinę zdarzeń $\{H_i\}_{i \in I}$, które są one wzajemnie rozłączne oraz $\bigcup\limits_{i \in I} H_i = \Omega$.

Twierdzenie 20. (WZÓR NA PRAWDOPODOBIEŃSTWO CAŁKOWITE) *Jeśli $\{B_1, B_2, \ldots, B_n\}$ jest rozbiciem Ω oraz $P(B_i) > 0$, to dla dowolnego zdarzenia A zachodzi*

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i) \tag{4}$$

Dowód.

$$P(A) = P\left(A \cap \left(\underbrace{\bigcup_{i=1}^{n} B_i}_{\Omega}\right)\right) = P\left(\bigcup_{i=1}^{n}(A \cap B_i)\right) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

□

Przykład 21. Algorytm daje odpowiedź poprawną z prawdopodobieństwem p, myli się z prawdopodobieństwem q oraz zawiesza się z prawdopodobieństwem r (wtedy ponownie uruchamiamy komputer). Jaka jest szansa, że ostatecznie uzyskamy poprawną odpowiedź?

Rozwiązanie 22.

$$\begin{aligned} W - &\ \text{uzyskano poprawną odpowiedź} \\ P(W) &= \underbrace{P(W|A) \cdot P(A)}_{1 \cdot p} + \underbrace{P(W|B) \cdot P(B)}_{0} + \underbrace{P(W|C) \cdot P(C)}_{P(W) \cdot r} \\ \text{wiemy jeszcze, że } &\ p + q + r = 1 \\ \text{zatem } P(W) &= \frac{p}{p+q} \end{aligned}$$

1.7 Prawdopodobieństwo przyczyny

Przykład 23. Zaobserwowane objawy awarii nowego komputera mogą być wynikiem uszkodzenia płyty głównej lub mogą być spowodowane innym czynnikiem (nie objętym gwarancją). Wiadomo, że awaria płyty głównej zdarza się raz na 1000 komputerów oraz powoduje zaobserwowane usterki z prawdopodobieństwem $\frac{4}{5}$. Z innych powodów występują one przeciętnie raz na 100 komputerów. Jaka jest szansa, że awarii uległa płyta główna?

Rozwiązanie 24.

$$\begin{aligned}
A -&\ \text{awarii uległa płyta główna}\\
B -&\ \text{zaobserwowano objawy awarii}\\
P(A) &= \frac{1}{1000} \Rightarrow P(A') = \frac{999}{1000}\\
P(B|A) &= \frac{4}{5}\\
P(B|A') &= \frac{1}{100}\\
P(A|B) &= ?\\
P(A|B) &= \frac{P(A\cap B)}{P(B)} = \frac{P(B|A)\cdot P(A)}{P(\underbrace{(B\cap A)\cup(B\cap A')}_{\text{składniki są rozłączne}})} = \frac{P(B|A)\cdot P(A)}{P(B\cap A)+P(B\cap A')}\\
&= \frac{P(B|A)\cdot P(A)}{P(B|A)\cdot P(A)+P(B|A')\cdot P(A')} \approx 0.074
\end{aligned}$$

Twierdzenie 25. (wzór Bayesa)[1] *Jeśli $\{H_i\}_{i\in I}$ jest przeliczalnym rozbiciem Ω takim, że $P(H_i)>0$, $i\in I$ oraz dane jest zdarzenie A takie, że $P(A)>0$, to*

$$P(H_j|A) = \frac{P(A|H_j)\cdot P(H_j)}{\sum\limits_{i\in I} P(A|H_i)\cdot P(H_i)} =^{z(4)} \frac{P(A|H_j)\cdot P(H_j)}{P(A)} \tag{5}$$

2 Zmienne losowe

Definicja 4. Jeśli Ω jest przeliczalny to dowolną funkcję określoną na Ω i o wartościach w $\mathbb{R}$ nazywamy **zmienną losową**.

Definicja 5. Niech dana będzie przestrzeń probabilistyczna $(\Omega,\mathcal{F},P)$. Funkcję $X\colon\Omega\to\mathbb{R}$ nazywamy **zmienną losową**, jeśli dla każdego $a\in\mathbb{R}$ $\{\omega\colon X(\omega)\le a\}\in\mathcal{F}$.

Definicja 6. Niech X będzie zmienną losową określoną na $(\Omega,\mathcal{F},P)$. Funkcję $F_X\colon\mathbb{R}\to[0,1]$, że $F_X(x) = P(\{\omega\colon X(\omega)\le x\})$ nazywamy dystrybuantą zmiennej losowej X.

Definicja 7. Zmienne losowe oznaczamy ostatnimi literami z alfabetu (X, Y, Z) oraz przyjmujemy zapis, że $P(\{\omega\colon X(\omega)\le x\}) = P(X\le x)$.

Własności dystrybuanty

1. jest funkcją niemalejącą
2. prawostronnie ciągła
3. $\lim\limits_{x\to-\infty} F_X(x)=0, \quad \lim\limits_{x\to+\infty} F_X(x)=1$

Przykład 26. Gramy w kości na pieniądze. Zasady są następujące:

- jeśli wyrzucimy 1 lub 2 to przegrywamy 1$.
- jeśli wyrzucimy 3, 4 lub 5 to nikt nikomu nie płaci.
- jeśli wyrzucimy 6 to wygrywamy 2$.

[1] Ciekawy przykład można znaleźć na Wikipedii

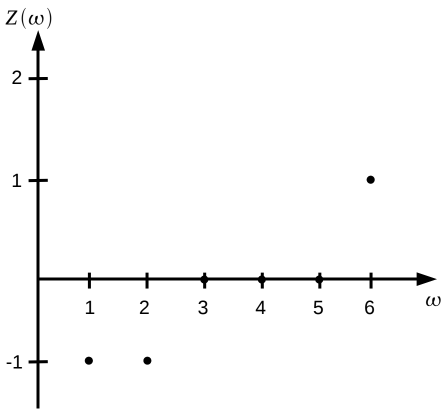
Zapiszmy to matematycznie:

$$\begin{aligned}
&\Omega = \{1, 2, \dots, 6\}\\
&Z - \text{ zysk z jednej kolejki}\\
&Z\colon \Omega \to \{-1, 0, 2\}
\end{aligned}$$

$$Z(\omega) = \begin{cases} -1 & \omega \in \{1, 2\} \\ 0 & \omega \in \{3, 4, 5\} \\ 2 & \omega = 6 \end{cases}$$

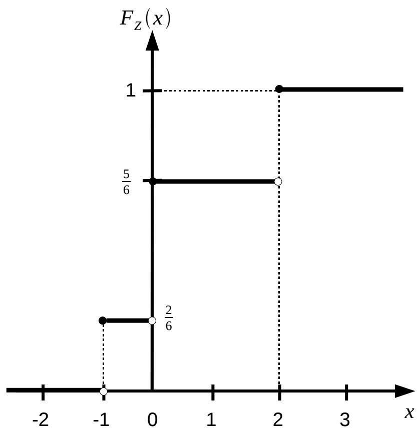
$$F_Z(x) = \begin{cases} 0 & x < -1 \\ \frac{2}{6} & x \in [-1, 0) \\ \frac{5}{6} & x \in [0, 2) \\ 1 & x \in [2, \infty) \end{cases}$$

Rysunek 2



Wykres zmiennej losowej Z

Rysunek 3



Wyres dystrybuanty zmiennej losowej Z

Uwaga 3. Jeśli znamy $F_X(x)$ to możemy liczyć także prawdopodobieństwo inne niż $P(X \leq x)$. Np.

$$\begin{aligned}
P(X \in (a, b]) &= P(X \leq b \wedge (X \leq a)') = P(X \leq b) - P(X \leq a) \\
&= P(X \leq b) + P(X > a) - 1 \text{ , bo } P(X > a) = 1 - P(X \leq a)
\end{aligned}$$

2.1 Niezależność zmiennych losowych

Przykład 27. Losujemy punkt z kwadratu $[0, 1]^2$.

$$\Omega = [0, 1]^2,\ P(A) = \frac{pole(A)}{pole(\Omega)} = pole(A)$$

W szczególności możemy myśleć o zmiennych losowych.

Niech

$$\begin{aligned}
&X - \text{zmienna losowa mówiąca jaka jest wartość pierwszej współrzędnej}\\
&Y - \text{zmienna losowa mówiąca jaka jest wartość drugiej współrzędnej}
\end{aligned}$$

wtedy

$$\begin{aligned}
F_X(x) &= P(\{\omega \in \Omega\colon X(\omega) \leq x\}) = P(\omega \in [0, x] \times [0, 1]) = x \text{ , gdy } x \in [0, 1] \\
F_Y(y) &= P(\{\omega \in \Omega\colon Y(\omega) \leq y\}) = P(\omega \in [0, 1] \times [0, y]) = y \text{ , gdy } y \in [0, 1]
\end{aligned}$$

Zauważmy, że

$$\begin{aligned}
P(X \leq x \wedge Y \leq y) &= P(\{\omega \in \Omega\colon \omega \in [0, x] \times [0, y]\}) = x \cdot y,\ x, y \in [0, 1] \\
&= P(X \leq x) \cdot P(Y \leq y)
\end{aligned}$$

Definicja 8. Niech $X_1, X_2, \ldots, X_n$ będą zmiennymi losowymi określonymi we wspólnej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$. Jeśli dla dowolnych $k \leq n$ oraz dowolnych $x_1, x_2, \ldots, x_n$ zachodzi

$$P(X_{j1} \leq x_1, X_{j2} \leq x_2, \ldots, X_{jk} \leq x_k) = \prod_{i=1}^{k} P(X_{ji} \leq x_i)$$

to zmienne te nazywamy **wzajemnie niezależnymi**.

2.2 Zmienne losowe dyskretne

Definicja 9. Zmienną losową $X \colon \Omega \to \{x_i := 1, 2, \ldots\}$ nazywamy dyskretną zmienną losową.

Definicja 10. Rozkładem zmiennej losowej dyskretnej $X \colon \Omega \to \{x_i := 1, 2, \ldots\}$ nazywamy ciąg par $(x_i, p_i)_{i=1,2,\ldots}$ gdzie $p_i = P(X = x_i)$.

Przykład 28. (kontynuacja przykładu 26)

$$Z - \text{ zysk na rzucie 1 kością do gry.}$$
$$Z \colon \{1, 2, 3, 4, 5, 6\} \to \{-1, 0, 2\}$$

Mamy:

$$\text{dla } x_1 = -1 \longrightarrow P(Z = -1) = \frac{2}{6} = \frac{1}{3} =: p_1$$
$$\text{dla } x_2 = 0 \longrightarrow P(Z = 0) = \frac{3}{6} = \frac{1}{2} =: p_2$$
$$\text{dla } x_3 = 2 \longrightarrow P(Z = 1) = \frac{1}{6} =: p_3$$

Otrzymaliśmy następujący rozkład: $(-1, \frac{1}{3}); (0, \frac{1}{2}); (2; \frac{1}{6})$

Uwaga 4.

$$F_X(t) = P(X \leq t) = \sum_{x_i \leq t} P(X = x_i) = \sum_{x_i \leq t} p_i$$

2.2.1 Zmienna losowa 0-1 (Bernoulli'ego)

Definicja 11. Jeśli $P(X = 1) = 1 - P(X = 0) = p$, dla $p \in (0, 1)$ to zmienną losową nazywamy 0-1 lub Bernoulli'ego. W skrócie pisać będziemy

$$X \sim B(1, p) \qquad \text{czyt.: zmienna losowa } X \text{ ma rozkład Bernoulli'ego.}$$

Zastosowanie 1. Załóżmy, że interesuje nas zdarzenie $A \subset \Omega$ o prawdopodobieństwie $P(A) = p$. Chcemy wiedzieć czy A zaszło. Zdefiniujmy funkcję

$$X(\omega) := \mathbf{1}_A(\omega) \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Zauważmy, że

$$P(X = 1) = P(\{\omega \in \Omega : \omega \in A\}) = P(A) = p$$

2.2.2 Zmienna losowa o rozkładzie dwumianowym

Przykład 29. Interesuje nas ilość zajść zdarzenia A w n niezależnych doświadczeniach.

$$X_i(\omega) = \mathbf{1}_{A_i}(\omega) \qquad \text{gdzie } A_1, A_2, \ldots, A_n - \text{ zdarzenia niezależne, } P(A_1) = P(A_2) = \ldots = P(A_n) = p$$

$$S_n(\omega) = \sum_{i=1}^{n} X_i(\omega)$$

Zbadajmy rozkład S_n:

$$\begin{aligned} P(S_n = k) &= P\left(\sum X_i(\omega) = k\right) = \sum_{n_1,n_2,\ldots,n_k} P(X_{n_1} = 1, X_{n_2} = 1, \ldots, X_{n_k} = 1, X_{n_{k+1}}, \ldots, X_{n_n} = 0) \\ &= \sum_{n_1,n_2,\ldots,n_k} p^k(1-p) = \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

Definicja 12. Zmienna losowa S_n ma rozkład dwumianowy z parametrami n oraz p jeśli $P(S_n = k) = \binom{n}{k}p^k(1-p)^{n-k}$ dla $k = 0, 1, \ldots, n$. Prawdopodobieństwo to nazywamy też prawdopodobieństwem w n próbach Bernoulli'ego. Oznaczamy ten rozkład $S_n \sim B(n, p)$.

Uwaga 5. Musi zachodzić własność:

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

2.2.3 Zmienne losowe o rozkładzie geometrycznym

Przykład 30. Mamy do przesłania plik za pomocą sieci. Jak to w życiu bywa, sieć czasem ulega awarii. Jeśli nastąpi awaria, to musimy zacząć transfer od początku. Interesuje nas ile razy wysyłaliśmy plik aż do momentu, gdy transfer się udał.

Rozwiązanie 31. Załóżmy, że kolejne próby są niezależne od pozostałych oraz $P(X_i = 1) = p$ dla $i = 1, 2, \ldots$. Niech

$$X_i = \begin{cases} 1 & \text{transfer się powiódł} \\ 0 & \text{transfer się nie powiódł} \end{cases}$$

$$Y - \text{ ilość transferów}$$

Zauważmy, że Y może przyjmować wartości $1, 2, \ldots$. Ponadto

$$Y = \min\{n \in \mathbb{N} \colon X_n = 1\}$$

Rozkład Y: niech $k \in \mathbb{N}$,

$$\begin{aligned} P(Y = k) &= P(X_1 = 0, X_2 = 0, \ldots, X_{k-1} = 0, X_k = 1) \\ &= P(X_1 = 0) \cdot P(X_2 = 0) \cdot \ldots \cdot P(X_{k-1} = 0) \cdot P(X_k = 1) \\ &= (1-p)^{k-1} p = p_k \end{aligned}$$

$$\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} p(1-p)^{k-1} = 1$$

Definicja 13. Zmienna losowa X ma rozkład geometryczny z parametrem $p \in (0, 1)$, gdy $P(X = k) = p(1-p)^{k-1}$ dla $k = 1, 2, \ldots$.
Oznaczenie: $X \sim \mathcal{G}(p)$ - związane z czasem czekania na pierwszy sukces.

Twierdzenie 32. (Własność braku pamięci) *Załóżmy, że X ma rozkład geometryczny $\mathcal{G}(p)$. Wtedy $P(X \geq n_0 + k | X > n_0) = P(X \geq k)$*

Uwaga 6. Tę własność mają tylko dwa rodzaje rozkładów: geometryczny (2.2.3) i wykładniczy (2.3.2).

Dowód. Rozpisując lewą stronę otrzymujemy

$$\begin{aligned} L &= \frac{P(X \geq n_0 + k \wedge X > n_0)}{P(X > n_0)} = \frac{P(X \geq n_0 + k)}{P(X \geq n_0 + 1)} = \\ &\frac{\sum_{i=n_0+k}^{\infty} P(X = i)}{\sum_{i=n_0+1}^{\infty} P(X = i)} = \frac{\sum_{i=n_0+k}^{\infty} p(1-p)^{i-1}}{\sum_{i=n_0+1}^{\infty} p(1-p)^{i-1}} =^{*} \\ &= \frac{(1-p)^{n_0+k-1}}{(1-p)^{n_0+1-1}} = (1-p)^{k-1} \end{aligned}$$

gdzie w (*) skorzystaliśmy z tego, że

$$\sum_{i=l}^{\infty} p(1-p)^{i-1} = p(1-p)^{l-1} \sum_{i=0}^{\infty} (i-p)^i = (1-p)^{l-1}$$

Ale patrząc z prawej strony, mamy

$$P = P(X \geq k) = \sum_{i=k}^{\infty} P(X=i) = \sum_{i=k}^{\infty} p(1-p)^{i-1} = (1-p)^{k-1}$$

□

2.2.4 Zmienne losowe o rozkładzie Poissona

Definicja 14. Zmienna losowa X ma rozkład Poissona[2] z parametrem $\lambda > 0$ jeśli $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ dla $k = 0, 1, \ldots$.

Uwaga 7. Faktycznie uzyskujemy rozkład, bo:

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1$$

2.3 Ciągłe zmienne losowe

Będziemy zajmować się tylko pewną podklasą ciągłych zmiennych losowych, mianowicie zmiennymi losowymi z gęstością[3].

Przykład 33. Mamy odcinek $[0, 1]$, rzucamy na niego punkt w sposób całkowicie losowy. Niech X - współrzędna trafienia. Chcemy znać jej rozkład. Ponieważ punktów jest nieskończenie wiele, to dla każdego punktu prawdopodobieństwo trafienia w niego jest równe 0. Możemy jednak policzyć inne prawdopodobieństwo. Zbadajmy $F_X(t) = P(X \leq t)$ czyli dystrybuantę. Zastosujmy chwyt z prawdopodobieństwem geometrycznym. Wtedy

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & t < 0 \\ t & t \in [0, 1] \\ 1 & t > 1 \end{cases}$$

Zauważmy, że

$$F_X(t) = \int_{-\infty}^{t} \mathbf{1}_{[0,1]}(x) dx \tag{6}$$

W praktyce okazuje się, że wiele dystrybuant możemy zapisać za pomocą całki.

Definicja 15. Niech X będzie zmienną losową o dystrybuancie F_X. Jeśli istnieje nieujemna funkcja f_X taka, że $F_X(t) = \int_{-\infty}^{t} f_X(s)ds$ dla każdego $t \in \mathbb{R}$, to f_X nazywamy gęstością rozkładu zmiennej losowej X.

Uwaga 8. Od tej pory będziemy zakładać, że f_X istnieje.

Uwaga 9.

$$F'_X(t) = f_X(t) \qquad \text{(tak jest prawie zawsze)}$$

To, że zachodzi (6) nie oznacza, że F_X jest w każdym punkcie różniczkowalna. Prawdą jest jednak, że jest skończenie wiele punktów, w których nie jest różniczkowalna.

Gdybyśmy znali gęstość naszego rozkładu, to znalibyśmy prawdopodobieństwo zajścia wielu zdarzeń, które mogłyby nas interesować, np.

$$P(a < x \leq b) = P(X \leq b) - P(X \leq a) = \int_{-\infty}^{b} f_X(s)ds - \int_{-\infty}^{a} f_X(s)ds = \int_{a}^{b} f_X(s)ds$$

Ogólnie

$$P(X \in A) = \int_A f_X(s)ds$$

[2] czyt. płasona
[3] istnieje pochodna dystrybuanty

Własności

1.
$$1 = P\left(X \in (-\infty, \infty)\right) = \int_{-\infty}^{\infty} f_X(s)ds \tag{7}$$

2. $P(X = t) = 0$, czyli szansa trafienia w konkretny punkt wynosi 0. Probabiliści mówią, że nie mamy atomów[4].

2.3.1 Rozkład jednostajny $\mathcal{U}[a, b]$

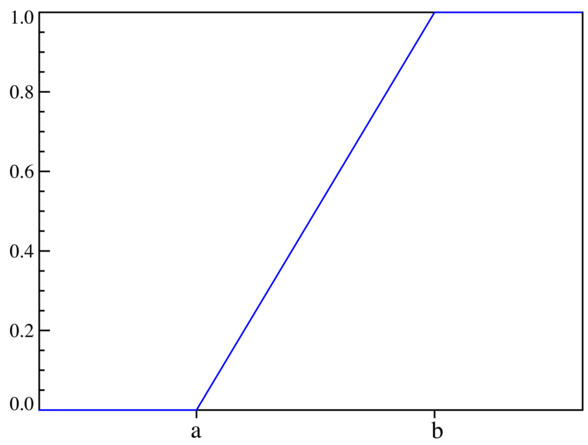
Definicja 16. Zmienna losowa X ma rozkład jednostajny[5] na odcinku $[a, b]$, gdy jej funkcja gęstości wynosi

$$f_X(t) = \begin{cases} 0 & t < a \\ 0 & t > b \\ \frac{1}{b-a} & t \in [a, b] \qquad \text{(wynika z (7) oraz poprzednich przypadków)} \end{cases} \tag{8}$$

$$P(X < t) = F_X(t) = \int_{-\infty}^{t} f_X(t)dt = \int_{a}^{t} f_X(t)dt = \frac{t-a}{b-a} \tag{9}$$

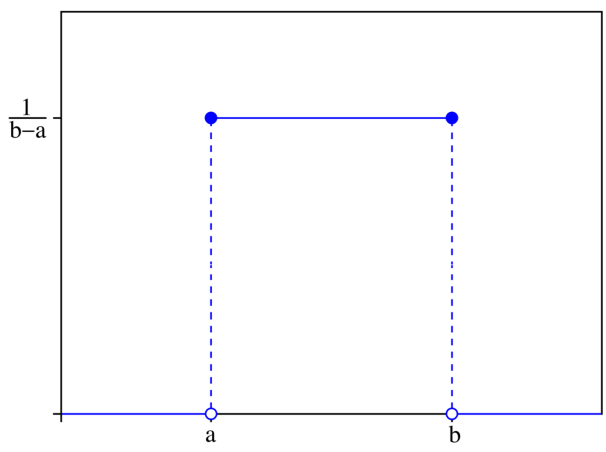
Uwaga 10. Ważny przypadek: $\mathcal{U}[0, 1]$.

Rysunek 4



Dystrybuanta, $F_X(t)$

Rysunek 5



Gęstość, $f_X(t)$

2.3.2 Rozkład wykładniczy $E(\lambda)$

Definicja 17. Zmienna losowa ma rozkład wykładniczy, gdy jej funkcja gęstości wynosi

$$f_X(t) = \begin{cases} 0 & t < 0 \\ \lambda e^{-\lambda t} & t \geq 0 \end{cases} \tag{10}$$

$$P(X < t) = F_X(t) = \int_{-\infty}^{t} f_X(t)dt = \int_{0}^{t} f_X(t)dt = 1 - e^{-\lambda t} \tag{11}$$

skąd

$$P(X > t) = \int_{t}^{\infty} \lambda e^{-\lambda s} ds = e^{-\lambda t} \tag{12}$$

[4] Ciekawe co na to chemicy….

[5] Rozkład jednostajny znany jest też jako jednorodny, równomierny, prostokątny oraz płaski.

Twierdzenie 34. (WŁASNOŚĆ BRAKU PAMIĘCI)

$$P(X > t + s | X > s) = P(X > t) \tag{13}$$

Dowód. (Przez rozpisanie)

$$\frac{P(X > t + s \wedge X > s)}{P(X > s)} = \frac{P(X > t + s)}{P(X > s)} =^{(12)} \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t)$$

□

Uwaga 11. To, że czekamy na przystanku już s czasu nie oznacza, że autobus przyjedzie szybciej. Jest tak dlatego, że osoba, która dopiero co przyszła na przystanek dopiero zaczyna czekanie.

2.3.3 Rozkład gamma $\Gamma(n, \lambda)$

Definicja 18. Funkcja gęstości jest postaci

$$f_X(t) = \begin{cases} 0 & t < 0 \\ \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} & t \geq 0 \end{cases} \tag{14}$$

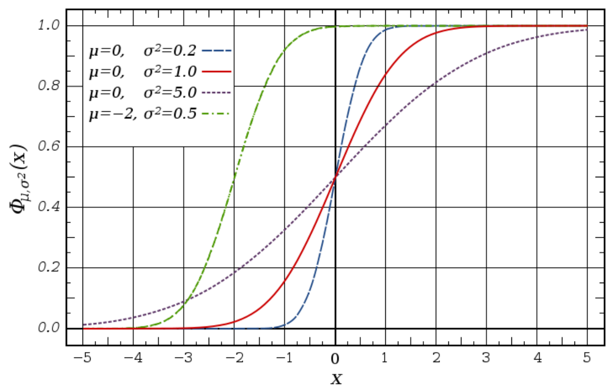
2.3.4 Rozkład normalny (Gaussa) $\mathcal{N}(m, \sigma^2)$

$$\mathcal{N}(m, \sigma^2), \quad m \in \mathbb{R}, \sigma^2 > 0 \tag{15}$$

Definicja 19. Zmienna losowa X ma rozkład $\mathcal{N}(m, \sigma^2)$ jeśli jej funkcja gęstości jest postaci

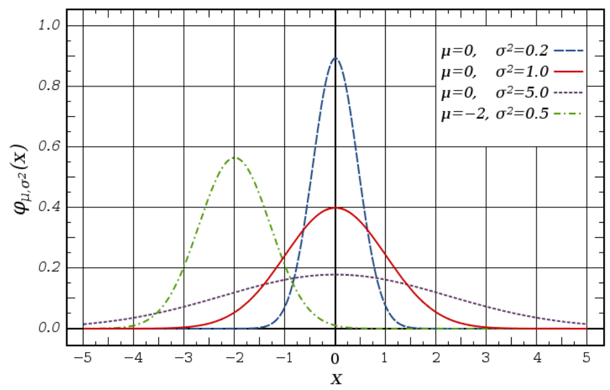
$$f_X(t) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right), \quad t \in \mathbb{R} \tag{16}$$

Rysunek 6



Dystrybuanta, $F_X(t)$

Rysunek 7



Gęstość, $f_X(t)$

Uwaga 12. Okazuje się, że nie ma wzoru na dystrybuantę! [6]. Nie można jej wyliczyć explicite.

$$F_X(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx$$

Ludzkość zdawszy sobie sprawę z powagi sytuacji doszła do wniosku, że rozkład Gaussa możemy zastąpić przez rozkład standardowy $\longrightarrow N(0,1) \leftarrow$ to jest stablicowane.

$$\Phi(t) = P(\mathcal{N} \leq t), \qquad \mathcal{N} \sim N(0,1)$$

Definicja 20. $\mathcal{N}$ rezerwujemy na oznaczenie zmiennej losowej o rozkładzie standardowym.

[6] Koniec świata jest bliski...

Fakt 35. (STANDARYZACJA) *Jeśli zmienna losowa $X \sim N(m, \sigma^2)$ (ma rozkład normalny z parametrami), to*

$$Y := \frac{X-m}{\sigma} \sim N(0,1) \tag{17}$$

Dowód.

$$f_Y(t) = F_Y'(t) = \frac{dP(Y \leq t)}{dt} = \frac{dP(\frac{X-m}{\sigma} \leq t)}{dt} = \frac{dP(X \leq \sigma t + r)}{dt} = F_X'(\sigma t + m) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp(-\frac{t^2}{2})$$

□

Przykład 36. Mamy dane, że $X \sim N(2,4)$. Chcemy zbadać ile wynosi $P(X \leq 3)$. Korzystamy z (17)

$$P(X \leq 3) = P(\frac{X-2}{2} \leq \frac{3-2}{2}) = P(\mathcal{N} \leq \frac{1}{2})$$

Reguła trzech sigm

W rozkładzie normalnym trzeba zwrócić uwagę na istotny szczegół. Mianowicie dla dowolnej liczby rzeczywistej gęstość jest niezerowa. Reguła trzech sigm mówi, że $\sim 99,7\%$ przypadków znajduje się w odległości mniejszej niż 3σ od wartości oczekiwanej.
Zobacz `http://pl.wikipedia.org/wiki/Reguła_trzech_sigm#Dla_rozk.C5.82adu_normalnego`

Lemat 2.1. *Załóżmy, że h jest funkcją ściśle rosnącą. Wtedy dla $Y = h(X)$*

$$F_Y(t) = F_X(h^{-1}(t)) \tag{18}$$

Dowód.

$$F_Y(t) = P(Y \leq t) = P(h(X) \leq t) = P(X \leq h^{-1}(t)) = F_X(h^{-1}(t))$$

□

2.4 Wektory losowe

Przykład 37. Rzucamy igłą w kwadrat $[0,1]^2$. Rzut jest losowy. Interesuje nas położenie trafionego punktu.

$$\begin{aligned} &X - \text{współrzędna } x\text{-owa} \\ &Y - \text{współrzędna } y\text{-owa} \\ &P(X \in A; Y \in B) = ? \end{aligned}$$

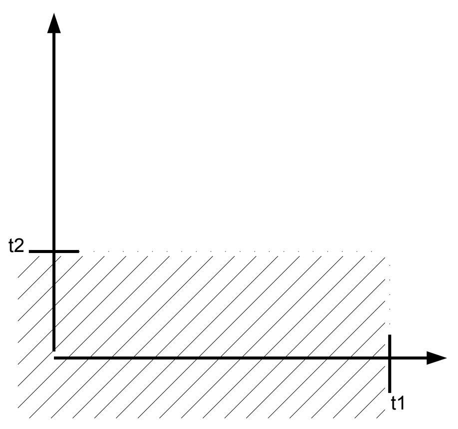
Uwaga 13. W przypadku 1-wymiarowym wystarczy znać $P(X \leq x)$. Jeżeli mamy gęstość to wystarczy ją scałkować po zbiorze, do którego chcemy, żeby należała igła.

W przypadku większej liczby wymiarów analogonem dystrybuanty 1-wymiarowej jest iloczyn dystrybuant. W szczególności $F_{(X,Y)}(t_1, t_2) := P(X \leq t_1; Y \leq t_2)$.

Wtedy na przykład dla $0 \leq a, b, c, d \leq 1$ mamy

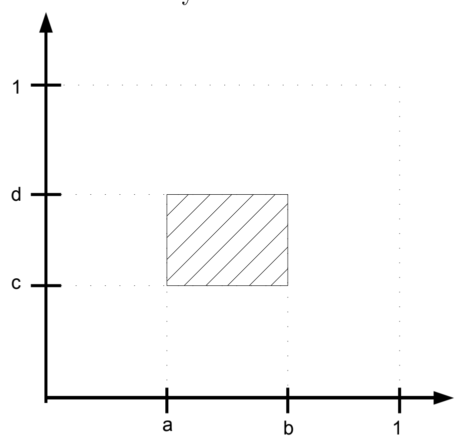
$$P(X \in (a,b], Y \in (c,d]) = F_{(X,Y)}(b,d) - F_{(X,Y)}(a,d) - F_{(X,Y)}(b,c) + F_{(X,Y)}(a,c)$$

Rysunek 8



$F_{(X,Y)}(t_1, t_2) := P(X \le t_1; Y \le t_2)$

Rysunek 9



$P\left(X \in (a, b], Y \in (c, d]\right)$

2.4.1 Rozkład. Przypadek dyskretny

Jeśli $p_{ij} = P(X = x_i; Y = y_j)$ dla każdego x_i, y_j to znamy rozkład.

2.4.2 Przypadek ciągły (z gęstością)

Uwaga 14. Dla jednego wymiaru było:

$$F_X(t) = \int_{-\infty}^{t} f_X(s) ds$$

Gęstość 2-wymiarowa to taka nieujemna funkcja $f_{(X,Y)}(s,t)$, że

$$F_{(X,Y)}(t_1, t_2) = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} f_{(X,Y)}(s,t)\ ds\ dt \tag{19}$$

Przykład 38. Załóżmy, że wektor losowy (X, Y) ma następujący rozkład:

X \ Y	1	2	
1	0.1	0.25	0.35
5	0.1	0.15	0.25
7	0.4	0	0.4
	0.6	0.4	1

Pytanie: Jaka jest szansa, że $Y = 1$? $Y = 2$?

$$\begin{aligned} P(Y = 1) &= P(Y = 1; X = 1) + P(Y = 1; X = 5) + P(Y = 1; X = 7) \\ &= 0.1 + 0.1 + 0.4 = 0.6 \qquad \text{zsumowaliśmy pierwszą kolumnę} \\ P(Y = 2) &= 0.4 \end{aligned}$$

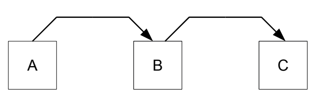
Definicja 21. Rozkłady X, Y liczone oddzielnie nazywamy rozkładami brzegowymi wektora (X, Y).

Uwaga 15. Rozkład (X, Y) wyznacza jednoznacznie rozkłady brzegowe, ale rozkłady brzegowe nie wyznaczają rozkładu łącznego.

2.4.3 Sumy niezależnych zmiennych losowych

Przykład 39.

Rysunek 10



Sposób transmisji pliku przez sieć.

Przesyłamy plik (rysunek) z prawdopodobieństwem sukcesu p w 1 próbie. Jeżeli udało się przesłać plik z A do B to przesyłamy dalej, czyli z B do C. Między fragmentami sieci oraz kolejnymi transmisjami pliku zachodzi niezależność. Interesuje nas łączna ilość prób aż do skutecznego dostarczenia pliku od C.

$$\begin{aligned}&T = X + Y\\&X - \text{ilość prób dokonanych na A} \curvearrowright \text{B}\\&Y - \text{ilość prób dokonanych na B} \curvearrowright \text{C}\end{aligned}$$

Zauważmy, że X jest niezależne od Y. Pytamy

$$\begin{aligned}P(T=k) &= \sum_{l=1}^{k-1} P(X=l; Y=k-l) = \sum_{l=1}^{k-1} P(X=l)P(Y=k-l)\\&= \sum_{l=1}^{k-1}(1-p)^{l-1}p(1-p)^{k-l-1}p = p^2\sum_{l=1}^{k-1}(1-p)^{k-2} = (k-1)p^2(1-p)^{k-2}\end{aligned}$$

Ponadto zauważmy, że

$$\begin{aligned}P(T=k) &= \sum_{l=1}^{k-1} P(X=l)\cdot P(Y=k-l)\\&= \sum_{l=1}^{\infty} P(X=l)\cdot P(Y=k-l) \qquad (\text{dla } l \geq k \ P(Y=k-l)=0)\end{aligned} \tag{20}$$

Wniosek 1. Jeśli X, Y - dyskretne zmienne losowe, niezależne oraz I_X-zbiór wartości X, to

$$P(X+Y=s) = \sum_{x\in I_X} P(X=x_i; Y=s-x_i) = \sum_{x\in I_X} P(X=x_i)\cdot P(Y=s-x_i)$$

Jeśli X, Y mają rozkład z gęstością odpowiednio f_X, f_Y, wtedy gęstość rozkładu $X+Y$ jest następująca

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(s)\cdot f_Y(t-s)\,ds$$

Wniosek 2. Jeśli $X \sim N(m_x, \sigma_x^2)$, $Y \sim N(m_Y, \sigma_Y^2)$ oraz X i Y są niezależne to

$$X+Y \sim N(m_X+m_Y, \sigma_X^2+\sigma_Y^2)$$

Ponadto dla danych a, b zachodzi $aX+b \sim N(am_X+b, a^2\sigma_X^2)$.

Dowód. Trzeba skorzystać z lematu 2.1. □

3 Wartość oczekiwana

Przykład 40. (przytoczona wcześniej gra)

$$\begin{aligned}\{1,2\} &\longrightarrow -1\\\{3,4,5\} &\longrightarrow 0\\\{6\} &\longrightarrow 2\end{aligned}$$

W $\leftarrow$ wygrana w 1 rundzie. Czy ta gra jest sprawiedliwa, czyli czy przy długim graniu każdy z graczy średnio nic nie traci i nie zarabia? Załóżmy, że mamy $6n$ rozgrywek. Spodziewamy się, że -1 trafi się $2n$ razy, 0 – $3n$ a 2 – n razy. Średni zysk na rozgrywkę wynosi $\frac{(-1)\cdot 2n+0\cdot 3n+2\cdot n}{6n} = 0$ (zatem gra jest sprawiedliwa) $= (-1)\cdot\frac{2n}{6n} + 0\cdot\frac{3n}{6n} + 2\cdot\frac{n}{6n} = (-1)\cdot P(W=-1) + 0\cdot P(W=0) + 2P\cdot(W=2) = E(W) \rightarrow$ wartość oczekiwana zmiennej losowej W.

Definicja 22. Niech X będzie dyskretną zmienną losową o rozkładzie $\{(x_i, p_i) : i = 1, 2, \ldots\}$. Jeśli $\sum_i |x_i| \cdot p_i < \infty$ to istnieje wartość oczekiwana zmiennej losowej X i wynosi $E(X) = \sum_i x_i \cdot P(X = x_i) = \sum_i x_i \cdot p_i$. W parktyce rzadko będziemy obserwować, że to założenie nie jest spełnione.

Definicja 23. Niech X będzie zmienną losową z gęstością f_X[7]. Jeśli $\int_{-\infty}^{\infty} |x| \cdot f_X(x)\, dx < \infty$, to istnieje wartość oczekiwana $E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$.

Uwaga 16. Wartość oczekiwana = wartość średnia = wartość przeciętna = pierwszy moment zmiennej losowej.

Fakt 41. *Załóżmy, że istnieją EX, EY. Wtedy*

1. $E(aX + b) = aE(X) + b \longleftarrow$ *liniowość*
2. $E(X + Y) = EX + EY$
3. *jeśli X, Y - niezależne, to $E(XY) = EX \cdot EY$*

Dowód.
(ad 1)
$$E(ax + b) = \sum_k (ax_k + b)P(X = x_k) = a\sum_k x_k P(X = x_k) + b\sum_k P(X = x_k)$$
(ad 3)
$$\begin{aligned} E(X \cdot Y) &=^{def} \sum_i \sum_j x_i y_j P(X = x_i; Y = y_j) \\ &= \sum_i \sum_j x_i y_j P(X = x_i)P(Y = y_j) \\ &= \sum_i x_i P(X = x_i) \sum_j y_j P(Y = y_j) \\ &= \sum x_i P(X = x_i)E(Y) = E(X)E(Y) \end{aligned}$$

□

Przykład 42.
T - długość pliku
$$P(T > t) \sim C \cdot t^{-\alpha}, \qquad \alpha \approx 1.7$$

Załóżmy, że $P(T < t) = 1 - \frac{1}{t^2}$, dla $t \geq 1$. Jaka jest średnia długość pliku, który przesyłamy, $E(T) =?$. Widać, że mamy model ciągły, więc skorzystamy ze wzoru $EX = \int_{\mathbb{R}} x f_X(x) dx$. Gęstość to nic innego jak pochodna, więc ją wyliczmy i będzie fajnie ;)

$$f_T(x) = F'_T(x) = \begin{cases} \dfrac{2}{x^3} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

Zatem
$$ET = \int_{\mathbb{R}} x f_T(x) dx = \int_1^{\infty} x \frac{2}{x^3} dx = -\frac{2}{x}\Big|_1^{\infty} = 2$$

Przykłady, dla których nie istnieje wartość oczekiwana

Podamy teraz kilka przykładów, dla których nie istnieje wartość oczekiwana.

1. przypadek ciągły: zmienna losowa o rozkładzie Cauchy'ego. Jej gęstość wyraża się jako $f_X(t) = \frac{1}{\pi(1+t^2)}$, $t \in \mathbb{R}$.

[7] gęstość wyznacza nam też rozkład

2. przypadek dyskretny: niech X ma rozkład $P(X=n^2)=\frac{c}{n^2}$, $n=1,2,\ldots$. Wtedy $EX=\sum_n n^2\frac{c}{n^2}=\infty$.

Co zatem należy robić, żeby wyliczyć wartość oczekiwaną $g(X)$, g-skomplikowana funkcja? Formalnie:

1. przypadek dyskretny: $Eg(X)=^{def}\sum_k g(x_k)\cdot P(X=x_k)$.
2. przypadek ciągły: $Eg(X)=\int_{\mathbb{R}} g(x)f_X(x)dx$.

Twierdzenie 43. (Nierówność Jensena) *Niech f będzie funkcją wypukłą dwukrotnie różniczkowalną. Wtedy $E(f(X))\geq f(EX)$.*

Dowód. Niech $m=E(X)$. Rozwijamy funkcję f w szereg Taylora w punkcie m:

$$f(x)=f(m)+(x-m)f'(m)+\frac{(x-m)^2}{2}f''(\Theta)\geq f(m)+(x-m)f'(m)$$

Zatem

$$\begin{aligned}E(f(X))\geq E(f(m)+(X-m)f'(m)) &= E(f(m))+E[(X-m)f'(m)]\\ &= f(E(X))+f'(m)[E(X)-m]=f(E(X))\end{aligned}$$

□

Uwaga 17. Jeśli X-nieujemna zmienna losowa o wartościach w $\mathbb{N}$, to

$$E(X)=\sum_{i=1}^{\infty}P(X\geq i)$$

$$\begin{array}{lllll}P(X\geq 1)= & P(X=1)+ & P(X=2)+ & P(X=3)+\ldots & \\ P(X\geq 2)= & & P(X=2)+ & P(X=3)+P(X=4)+\ldots & \\ P(X\geq 3)= & & & P(X=3)+P(X=4)+P(X=5)+\ldots & \\ \hline \end{array}$$

$$=P(X=1)+2P(X=2)+3P(X=3)+4P(X=4)+\ldots=E(X)$$

4 Ryzyko i jego pomiar

Przykład 44.

I) Powraca nieśmiertelny przykład 26. $E(X)=0$, więc gra jest sprawiedliwa. $X\in\{-1,0,2\}$

II) Zmodyfikujmy grę tak, że płacimy nie jednego \$, ale 10^4\$ gdy wypadnie 1 lub 2; nic się nie dzieje, gdy wypadnie 3,4 lub 5; dostajemy $2\cdot 10^4$\$, gdy wypadnie 6 oczek.

Y- wygrywamy w 1 rozgrywce

$E(Y)=-10^4\cdot\frac{1}{3}+0\cdot\frac{1}{2}+2\cdot 10^4\cdot\frac{1}{6}=0.$

Aby mierzyć ryzyko odchylenia wartość średniej możemy postąpić następująco:

$$Var(X):=E(X-E(X))^2 \qquad \leftarrow \text{wariancja zmiennej losowej } X \tag{21}$$

Ryzyko będziemy mierzyć licząc

$$\sigma_X:=\sqrt{Var(X)} \qquad \leftarrow \text{odchylenie standardowe.} \tag{22}$$

Uwaga 18. Odchylenie standardowe nazywane jest także dyspersją. Oznacza się, że

$$\mathcal{D}^2(X):=Var(x)=E(X^2)-(E(X))^2 \tag{23}$$

$\mathcal{D}^2(x)=E(X-E(X))^2=E(X^2-2X\cdot E(X)+(E(X))^2)=E(X^2)-2(EX)^2+E(X)^2=E(X^2)-(E(X))^2$

Przykład 45. (kontynuacja przykładu 44.)

I) $E(X^2) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 4 \cdot \frac{1}{6} = 1$.
$Var(X) = 1 - 0^2 = 1$, $\sigma_X = \sqrt{1} = 1$

II) $E(Y^2) = 10^4 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + (2 \cdot 10^4)^2 \cdot \frac{1}{6} = 10^8$.
$Var(Y) = 10^8 - 0^2 = 10^8$, $\sigma_Y = \sqrt{10^8} = 10^4$

Uwaga 19.

- $E(X^2)$ – drugi moment zmiennej losowej
- $Var(X) \equiv \sigma_X^2$

Własności wariancji

1. $Var(aX) = a^2 \cdot Var(X)$
2. $Var(X + b) = Var(X)$
3. jeśli X, Y są niezależne, to $Var(X \pm Y) = Var(X) + Var(Y)$
4. $Var(X \pm Y) = Var(X) + Var(Y) \pm 2 \cdot Cov(X,Y)$ (przypadek ogólny)

Uwaga 20. $Var(X) \geq 0$!!!

Dowód.
(1) $Var(a \cdot X) =^{def} E(aX - E(aX))^2 = Ea^2(X - E(X))^2 = a^2 \cdot E(X - E(X))^2 =^{(21)} a^2 \cdot Var(X)$
(2) $Var(X + b) =^{def} E(X + b - E(X + b))^2 = E(X + b - EX - b)^2 =^{def} Var(X)$
(3) $Var(X \pm Y) =^{(23)} E(X \pm Y)^2 - (E(X \pm Y))^2 = EX^2 \pm E(2XY) + EY^2 - \big((EX)^2 \pm 2(EX)(EY) + (EY)^2\big) = EX^2 - (EX)^2 + EY^2 - (EY)^2 =^{(23)} Var(X) + Var(Y)$ □

4.1 Kowariancja, $Cov(X,Y)$

Definicja 24. Do mierzenia niezależnośći zmiennych losowych X, Y używamy kowariancji.

$$Cov(X,Y) = E(X \cdot Y) - EX \cdot EY \tag{24}$$

- jeżeli $Cov(X,Y) \neq 0$ to zmienne są zależne
- jeżeli $Cov(X,Y) = 0$ to nie ma pewności, że są niezależne!

Fakt 46.

$$Cov(aX, bY) = abCov(X,Y) \qquad a, b \in \mathbb{R}$$

4.2 Współczynnik korelacji

Definicja 25. Współczynnik korelacji ρ zmiennych losowych X, Y definiujemy jako

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{25}$$

Fakt 47. $\rho(aX, bY) = \rho(X,Y) \quad a, b > 0$

Dowód.

$$\rho(aX, bY) = \frac{ab \cdot Cov(X,Y)}{a\sqrt{Var(X)} \cdot b\sqrt{Var(Y)}} = \rho(X,Y)$$

□

Własności korelacji

1. jeśli zmienne losowe X, Y są niezależne, to $\rho(X,Y) = 0$
2. odwrotna implikacja nie zachodzi (chyba, że zmienne X, Y mają rozkłady normalne)
3. jeśli $\rho(X,Y) \neq 0$ to X i Y są wzajemnie zależne

Przykład 48.

X\ Y	-1	1	
-1	0.4	0.1	0.5
1	0.1	0.4	0.5
	0.5	0.5	1

$Cov(X,Y) = E(XY) - EX \cdot EY,$

$EX = 0 = EY,$

$E(XY) = (-1)^2 \cdot 0.4 + (-1) \cdot 1 \cdot 0.1 + 1 \cdot (-1) \cdot 0.1 + 1^2 \cdot 0.4 = 0.6$

$Var(X) = 1 = Var(Y) \leftarrow$ trzeba sprawdzić!

$\rho(X,Y) = 0.6$

Przykład 49.

X\ Y	-1	1	
-1	0.1	0.4	0.5
1	0.4	0.1	0.5
	0.5	0.5	1

$Var(X) = Var(Y) = 1 \leftarrow$ trzeba sprawdzić!

$EX = 0 = EY,$

$Cov(X,Y) = E(XY) - 0 \cdot 0 = -0.6$

$\rho(X,Y) = -0.6$

Uwaga 21.

- jeśli $\rho(X,Y) > 0$, to duża wartość jednej zmiennej implikuje dużą wartość drugiej zmiennej
- jeśli $\rho(X,Y) < 0$, to duża wartość jednej zmiennej implikuje małą wartość drugiej zmiennej

Fakt 50.

$$-1 \leq \rho(X,Y) \leq 1$$

- *im bliżej jedynki tym zmienne są bardziej skorelowane/zależne - zachowują się podobnie.*
- *dla* $\rho(X,Y) = 1$ *jest* $Y = aX + b$ $a > 0$
- *dla* $\rho(X,Y) = -1$ *jest* $Y = aX + b$ $a < 0$
- *dla* $\rho(X,Y) \to 0$ *nie możemy stwierdzić czy zmienne są zależne*

Zastosowanie 2.

$$\begin{aligned} Var(X,Y) &= E\left[X + Y - E(X+Y)\right]^2 = E\left[(X - EX) + (Y - EY)\right]^2 \\ &= E\left[(X-EX)^2 + (Y+EY)^2 + 2 \cdot (X - EX)(Y - EY)\right] \\ &= Var(X) + Var(Y) + 2E\left[(X-EX)(Y-EY)\right] \\ &=^{(*)} Var(X) + Var(Y) + 2Cov(X,Y) \end{aligned}$$

gdzie

$$\begin{aligned} (*) =& E(X-EX)(Y-EY) = E\left[XY - XEY - YEX + EX \cdot EY\right] \\ =& E(XY) - 2EX \cdot EY + EX \cdot EY = Cov(X,Y) \end{aligned}$$

5 Warunkowa wartość oczekiwana

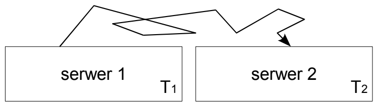
Będziemy zajmowali się wyłącznie dyskretnymi zmiennymi losowymi.

Przykład 51. Załóżmy, że mamy program, który zawiera 1 wywołanie prosedury S. Każde wywołanie S wywołuje nowe kopie procedury S z rozkładem $B(n,p)$ - maksymalnie n wywołań, każde z prawdopodobieństwem p. Zakładamy, że zmienne losowe są niezależne dla każdego wywołania S. Jaka jest spodziewana (średnia) liczba wywołań procedury S?

Na to pytanie jeszcze nie umiemy odpowiedzieć, rozwiązanie będzie niedługo….

Przykład 52. Przesyłamy plik przez sieć.

Rysunek 11



Sposób transmisji pliku przez sieć.

Interesuje nas całkowity czas przesyłu pliku, gdy na pierwszym serwerze plik może być przetwarzany w losowym czasie równym $T_1 \in \{1,2,3,4,6\}$ z jednakowym prawdopodobieństwem. Na drugim serwerze tak samo.

$$T = T_1 + T_2$$
$$E(T) = E(T_1 + T_2) = 2 \cdot 3.5 = 7$$

Pytamy: jak długo średnio będzie trwał transfer, gdy wiemy, że na pierwszym serwerze trwał on 2?

Korzystamy ze wzory na wartość oczekiwaną (jednocześnie definiujemy warunkową wartość oczekiwaną)

$$E(T|T_1 = 2) = \sum_{x=T_1+1}^{T_1+6} x \cdot P(T = x \big| T_1 = 2) = \sum_{x=3}^{8} x \cdot \frac{1}{6} = \frac{33}{6} = 5.5$$

Definicja 26. Niech X, Y będą dyskretnymi zmiennymi losowymi (określonymi na przestrzeni probablistycznej). Wtedy dla każdego y takiego, że $P(Y = y) > 0$ mamy

$$E(X \big| Y = y) = \sum_x x \cdot P(X = x \big| Y = y) \tag{26}$$

Własności warunkowej wartości oczekiwanej

Lemat 5.1. $E(X) = \sum_y P(Y = y) \cdot E(X \big| Y = y)$

Dowód. (przez rozpisanie)

$$\text{prawa strona} = \sum_y P(Y = y) \cdot \sum_x P(X = x \big| Y = y) = \sum_x \sum_y P(Y = y) \cdot x \cdot P(X = x \big| Y = y)$$
$$= \sum_x \sum_y P(X = x \wedge Y = y) = \sum_x x \cdot P(X = x) = EX = \text{lewa strona}$$

□

Fakt 53. $E(X_1 + X_2 \big| Y = y) = E(X_1 \big| Y = y) + E(X_2 \big| Y = y)$

Dowód. dowód analogicznie jak dla zwykłej wartości oczekiwanej □

$$T_1 = 1 \to E(T|T_1 = 1) = 4.5, \qquad T_1 = 2 \to E(T|T_1 = 2) = 5.5,$$
$$T_1 = 3 \to E(T|T_1 = 3) = 6.5, \qquad T_1 = 4 \to E(T|T_1 = 4) = 7.5,$$
$$T_1 = 5 \to E(T|T_1 = 5) = 8.5, \qquad T_1 = 6 \to E(T|T_1 = 6) = 9.5$$

Możemy myśleć o <u>zmiennej losowej</u> $E(T|T_1)$ o rozkładzie

$$P\big(E(T|T_1) = 4.5\big) = \frac{1}{6}, \qquad P\big(E(T|T_1) = 5.5\big) = \frac{1}{6}, \qquad P\big(E(T|T_1) = 6.5\big) = \frac{1}{6},$$
$$P\big(E(T|T_1) = 7.5\big) = \frac{1}{6}, \qquad P\big(E(T|T_1) = 8.5\big) = \frac{1}{6}, \qquad P\big(E(T|T_1) = 9.5\big) = \frac{1}{6}$$

Definicja 27. $E(X|Y)$ jest zmienną losową, która przyjmuje wartości $E(X|Y = y)$ z prawdopodobieństwem $P(Y = y)$.

Policzmy $E(T|T_1)$

$$E(T|T_1) = \sum_{x=T_1+1}^{T_1+6} x \cdot \frac{1}{6} = T_1 + \frac{7}{2}$$

$$E\left[E(T|T_1)\right] = E(T_1 + \frac{7}{2}) = 7 = ET$$

To nie przypadek! Tak wychodzi zawsze.

Twierdzenie 54.

$$E\left(E(X|Y)\right) = E(X) \tag{27}$$

<u>*Rozwiązanie*</u> 55. (do przykładu 51)

Proces należy do generacji i jeśli został wywołany przez proces z generacji z $i-1$.

$$\begin{aligned} &Y_i \text{ - ilość procesów } S \text{ w } i\text{-tej generacji} \\ &Y_0 = 1 \\ E(&Y_1) = n \cdot p \end{aligned}$$

Na moment założymy, że wiemy ile było procesów w $(i-1)$-szej generacji (oznaczmy ją $Y_{i-1} = y_{i-1}$) i policzymy ile było w i-tej. Niech Z_k oznacza ilość kopii procesu S, które zostały wywołane przez k-ty proces $(i-1)$-szej generacji $(k = 1, 2, \ldots, y_{i-1})$.

Pytamy jaki rozkład ma Z_k. Otóż ma ono zawsze taki sam rozkład, $Z_k \sim B(n, p)$.

Wyliczmy warunkową wartość oczekiwaną przy znajomości y_{i-1}

$$E\left[Y_i|Y_{i-1} = y_{i-1}\right] = E\left[\sum_{k=1}^{y_{i-1}} Z_k|Y_{i-1} = y_{i-1}\right]$$

Ponieważ warunek nie daje nam żadnej informacji, więc możemy się go pozbyć. Otrzymujemy

$$E\left[\sum_{k=1}^{y_{i-1}} Z_k|Y_{i-1} = y_{i-1}\right] = E\left[\sum_{k=1}^{y_{i-1}} Z_k\right] = \sum_{k=1}^{y_{i-1}} EZ_k = y_{i-1} \cdot n \cdot p$$

Stąd wynika, że $E[Y_i|Y_{i-1}] = Y_{i-1} \cdot n \cdot p$. Jest to klucz do rozwiązania naszego zadania. Korzystając ze wzoru (27) mamy:

$$E(Y_i) = E\left[E[Y_i|Y_{i-1}]\right] = E(Y_{i-1} \cdot n \cdot p) = n \cdot p \cdot E(Y_{i-1})$$

Zatem

$$E(Y_i) = (n \cdot p)^i$$

Jeśli Y=całkowita ilość wywołań procesu S, to

$$E(Y) = E\left(\sum_{i=0}^{\infty} Y_i\right) = \sum_{i=1}^{\infty} (n \cdot p)^i =^{szereggeom}= \begin{cases} \frac{1}{1-np} & np < 1 \\ \infty & np \geq 1 \end{cases}$$

6 Nierówności

6.1 Nierówność Markowa

Twierdzenie 56. (Markow) *Załóżmy, że X jest nieujemną zmienną losową oraz EX istnieje. Wtedy prawdziwa jest następująca nierówność*

$$P(X > a) \leq \frac{EX}{a} \tag{28}$$

Uwaga 22. Zauważmy, że wzór (28) ma sens tylko, gdy $\frac{EX}{a} < 1$. Inaczej uzyskana informacja nic nam nie mówi.

Uwaga 23. Ponieważ założenia nie są bardzo mocne, więc oszacowanie też nie jest bardzo dobre. Choć w rzeczywistośći można podać dużo lepsze oszacowania, o czym przekonamy się już niedługo, to nierówność Markowa jest przydatna chociażby dlatego, że można ją szybko wyliczyć.

$$\text{indykator: } \mathbf{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

$$\begin{aligned} &X - \text{zmienna losowa} \\ &E(\mathbf{1}_A(X)) = 0 \cdot P(X \notin A) + 1 \cdot P(X \in A) = P(X \in A) \end{aligned}$$

Dowód. Niech dany będzie $\mathbf{1}_{(a,\infty)}(X)$. Zauważmy, że $\mathbf{1}_{(a,\infty)}(X) \leq \frac{X}{a}$.

$$E\left(\mathbf{1}_{(a,\infty)}(X)\right) \leq E\left(\frac{X}{a}\right)$$

$$P(X > a) \leq \frac{EX}{a}$$

□

6.2 Nierówność Czebyszewa

Twierdzenie 57. *Jeśli istnieje $Var(X)$, to dla dowolnego $a > 0$*

$$P\left(|X - EX| > a\right) < \frac{Var(X)}{a^2} \tag{29}$$

co jest równoważne

$$P\left(|X - EX| \leq a\right) \geq 1 - \frac{Var(X)}{a^2} \tag{30}$$

Uwaga 24. Ponieważ zakładamy więcej niż w przypadku nierówności Markowa, dostajemy lepsze oszacowanie.

Dowód. Zauważmy, że

$$P\left(|X - EX| > a\right) = P\left(|X - EX|^2 > a^2\right)$$

Wiemy, że $Var(X)$ istnieje. Przypomnijmy, że $Var(X) = E(X - EX)^2$. Zatem z twierdzenia 28 mamy

$$P\left(|X - EX| > a\right) \leq \frac{E(X - EX)^2}{a^2} = \frac{Var(X)}{a^2}$$

□

Uwaga 25. Jeśli istnieje $E\left|X - EX\right|^k$, to

$$P\left(|X - EX| > a\right) \leq \frac{E\left(|X - EX|\right)^k}{a^k}$$

Wnioski

- $P\left(|X-EX| \geq t \cdot \sqrt{Var(X)}\right) \leq \frac{1}{t^2}$. W szczególności dla $t=3$ jest $P\left(|X-EX| \geq 3\sqrt{Var(X)}\right) \leq \frac{1}{9}$.

Przykład 58. (nieśmiertelny przykład raz jeszcze) Przeprowadzamy 1000 rozgrywek. Pytamy jaka jest szansa, że po nich będziemy co najmniej 100 do przodu. Niech

$$\begin{aligned} X_i &- \text{wynik } i\text{-tej rundy} \\ W &- \text{wygrana} \\ W &= \sum_{i=1}^{1000} X_i \\ P(W > 100) &= ? \end{aligned}$$

$$\begin{aligned} EW &= 1000 \cdot EX_1 = 0 \qquad \text{bo gra jest sprawiedliwa} \\ Var(W) &= 1000 \cdot Var(X_1) = 1000 \cdot 1 \end{aligned}$$

$$\begin{aligned} P(W > 100) &= P(W - EW > 100 - EW) = P(W - EW > 100) \leq P(|W - EW| > 100) \\ &\leq^{(30)} \frac{Var(W)}{100^2} = \frac{1000}{10000} = 10\% \end{aligned}$$

6.3 Funkcje tworzące momenty

Definicja 28. Funkcję $M_X(t) = E\left[e^{tX}\right]$ nazywamy funkcją tworzącą momenty zmiennej losowej X.

Przykład 59. Niech $X \sim \mathcal{G}(p)$. Wtedy

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \sum_{k=1}^{\infty} e^{tk} P(X=k) = \sum_{k=1}^{\infty} e^{tk} p(1-p)^{k-1} = \frac{p}{1-p} \sum_{k=1}^{\infty} e^{tk}(1-p)^k = \frac{p}{1-p} \sum_{k=1}^{\infty} \left[e^t(1-p)\right]^k \\ &= \frac{p}{1-p}\left[\frac{1}{1-e^t(1-p)} - 1\right] = \frac{pe^t}{1-(1-p)e^t} \qquad \text{dla } t < -\log(1-p) \end{aligned}$$

Twierdzenie 60. *Niech X będzie zmienną losową o funkcji tworzącej momenty $M_X(t)$. Jeśli istnieje M_X w otoczeniu 0, to*

$$E(X^n) = M_X^{(n)}(0) \qquad (n\text{-te pochodne})$$

Dowód.

$$M_X^{(n)}(t) = (Ee^{tX})^{(n)} =^{*} E([e^{tX}]^{(n)}) = E(X^n \cdot e^{tX})$$

Zachodzi przemienność (wynika z definicji pochodnej). Zatem

$$M_X^{(n)} = E(X^n \cdot 1) = E(X^n)$$

Zastosujmy to dla rozkładu $\mathcal{G}(p)$. Mamy

$$\begin{aligned} M'_X(t) &= p(1-(1-p)e^t)^{-2}e^t \\ E(X) = M'_X(0) &= \frac{1}{p} \\ M''_X(t) &= 2p(1-p)(1-(1-p)e^t)^{-3}e^{2t} + p(1-(1-p)e^t)^{-2}e^t \\ E(X^2) = M''_X(0) &= \frac{2-p}{p^2} \end{aligned}$$

□

Twierdzenie 61. *Niech X, Y - zmienne losowe o funkcjach tworzących momenty odpowiednio $M_X(t)$, $M_Y(t)$. Jeśli $M_X(t) \equiv M_Y(t)$ dla $t \in (-\theta, \theta)$, to X i Y mają takie same rozkłady.*

Twierdzenie 62. *Niech X, Y będą niezależnymi zmiennymi losowymi o funkcjach tworzących momenty $M_X(t)$ i $M_Y(t)$. Wtedy $X+Y$ ma funkcję tworzącą momenty $M_{X+Y}(t)$ daną w postaci*

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) \tag{31}$$

Dowód.

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX} \cdot e^{tY}] = E[e^{tX}] \cdot E[e^{tY}] = M_X(t) \cdot M_Y(t)$$

□

Zbadamy teraz odpowiedniość funkcji tworzących i rozkładów. Istnieją na to wzory[8]. Ogólnie:

1. liczymy momenty tworzące X,Y
2. mnożymy te momenty; dostajemy M_{X+Y}
3. odgadujemy jaki stoi za tym rozkład

Tabela funkcji tworzących momenty

X	$M_X(t)$
Geometryczny $\mathcal{G}(p)$	$\frac{pe^t}{1-(1-p)e^t}$ dla $t < -\log(1-p)$
Poisson(λ)	$e^{\lambda(e^t-1)}$ dla $t \in \mathbb{R}$
Wykładniczy $E(\lambda)$	$\frac{\lambda}{\lambda - t}$ dla $t < \lambda$
Normalny $N(m, \sigma^2)$	$e^{mt}e^{\frac{t^2\sigma^2}{2}}$ dla $t \in \mathbb{R}$

Przykład 63.

$$\begin{aligned} X &\sim N(m, \sigma_X^2) \\ Y &\sim N(m, \sigma_Y^2) \\ X, Y &- \text{niezależne} \\ X+Y &\sim \quad ? \end{aligned}$$

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) = e^{m_X t}e^{\frac{\sigma_X^2 t^2}{2}} e^{m_Y t}e^{\frac{\sigma_Y^2 t^2}{2}} = e^{(m_X+m_y)t}e^{\frac{(\sigma_X^2+\sigma_Y^2)t^2}{2}}$$

Przykład 64.

$$X \sim N(m_X, \sigma_X^2)$$

Pytamy o rozkład zmiennej losowej Y

$$Y = \frac{X - EX}{\sqrt{Var(X)}} = \frac{X - m_X}{\sigma_X}$$

Liczymy

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{\frac{t}{\sigma_X}(X-m_X)}\right) = E(e^{\frac{Xt}{\sigma_X}} \cdot e^{-\frac{m_X t}{\sigma_X}}) = e^{-\frac{m_X t}{\sigma_X}} \cdot Ee^{\frac{Xt}{\sigma_X}} = e^{-m_X \frac{t}{\sigma_X}} \cdot e^{m_X \frac{t}{\sigma_X}} \cdot e^{\frac{t^2}{2}} = e^{\frac{t^2}{2}}$$

6.4 Nierówność Chernoffa

Twierdzenie 65. *Załóżmy, że istnieje $M_X(t)$ dla zmiennej losowej X i jest ona dobrze określona dla $t \in (-\delta, \delta)$ dla pewnego $\delta > 0$.*
Wtedy dla $t \in (0, \delta)$ mamy

$$P(X > a) = P(tX > ta) = P(e^{tX} > e^{ta}) =^{28} \leq \frac{Ee^{tX}}{e^{ta}} = Ee^{tX} \cdot e^{-at}$$

Ponieważ t było dowolne, to

$$P(X > a) \leq \min_{t \in (0,\delta)} M_X(t) \cdot e^{-at}$$

[8] Wzory na odwrócenie, odpowiednie całkowanie pozwala odzyskać to co trzeba

Przykład 66. X-ilość orłów w n rzutach idealną monetą. Korzystając z nierówności Chernoffa można pokazać, że

$$P\left(\left|X - \frac{n}{2}\right| \geq \frac{1}{2}\sqrt{n \log n}\right) \leq \frac{2}{n}$$

7 Prawo wielkich liczb

Znane w XVII wieku, kiedy ludzie prawdopodobieństwo postrzegali wyłącznie intuicyjnie. Matematyk Bernoulli zaczął zastanawiać się jak to z tym wszystkim jest i doszedł do prawa wielkich liczb.

Twierdzenie 67. (Prawo wielkich liczb Bernoulliego) *Jeśli S_n jest liczbą sukcesów w n próbach Bernoulli'ego z prawdopodobieństwem sukcesu p, to*

$$\forall_{\epsilon>0}\ P\left(\left|\frac{S_n}{n} - p\right| > \epsilon\right) \longrightarrow^{n\to\infty} 0$$

Dowód. Zauważmy, że

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n}E(S_n) = \frac{1}{n}\cdot n \cdot p = p$$

$$Var(S_n) = np(1-p) \qquad \longleftarrow \text{z listy zadań. Stąd}$$
$$Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}$$

Korzystamy z nierówności Czebyszewa dla zmiennej losowej S_n

$$P\left(\left|\frac{S_n}{n} - p\right| > \epsilon\right) = P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| > \epsilon\right) \leq \frac{Var\left(\frac{S_n}{n}\right)}{\epsilon^2} = \frac{p(1-p)}{\epsilon^2}\cdot\frac{1}{n}$$

□

Podamy teraz ogólniejsze twierdzenie

Twierdzenie 68. (Prawo wielkich liczb) *Jeśli $X_1, X_2, \ldots$ są niezależnymi zmiennymi losowymi o takim samym rozkładzie oraz istnieje wartość oczekiwana ($E|X_1| < \infty$), to*

$$P\left(\left|\frac{\sum_{i=1}^n X_i}{n} - E(X_1)\right| > \epsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

Co się dzieje, kiedy nie możemy dobrze określić funkcji tworzącej momenty? Oto przykład takiego rozkładu

Przykład 69.

$$P(X > t) = \frac{1}{1+t} \qquad t \geq 0$$

Można sprawdzić, że wartość oczekiwana jest nieskończona (całka bardzo szybko rośnie do ∞)

$$Ee^{tX} = \int_0^\infty e^{tX} f_X(t)dt$$

Na takie niesforne przypadki mamy funkcje charakterystyczne

7.1 Funkcja charakterystyczna

Definicja 29. Funkcję $\phi : \mathbb{R} \to \mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ nazywamy funkcją charakterystyczną zmiennej losowej X jeśli

$$\phi_X(t) = E\left(e^{itX}\right) = E\left(\cos(tX) + i\sin(tX)\right) \tag{32}$$

Uwaga 26.

1. Funkcja charakterystyczna ma wszystkie własności jak funkcja tworzaca momenty.
2. Jest dobrze określona dla wszystkich zmiennych losowych, dla każdego $t \in \mathbb{R}$, bo $\left|e^{itX}\right| \equiv 1$.

Fakt 70. *Jeśli istnieje $E(X^2)$, to*

$$EX = \frac{1}{i}\phi'_X(0)$$
$$EX^2 = \frac{1}{i^2}\phi''_X(0) = -\phi''_X(0)$$

Fakt 71. *Jeśli X, Y-niezależne, to $\phi_{X+Y}(t) = \phi_X(t) \cdot \phi_Y(t)$.*

7.2 Rozkłady a ich funkcje charakterystyczne

Rozkład X	Funkcja charakterystyczna $\phi_X(t)$
geometryczny	$\frac{p}{1-qe^{it}}$
Bernoulli'ego $B(n,p)$	$\left(q+pe^{it}\right)^n$
Poissona $P(\lambda)$	$e^{\lambda(e^{it}-1)}$
normalny $N(0,1)$	$e^{-\frac{t^2}{2}}$
jednostajny $U[0,1]$	$(e^{iat}-1)/iat$
wykładniczy $E(\lambda)$	$\frac{\lambda}{\lambda - it}$

Fakt 72. *Dla danej funkcji charakterystycznej istnieje dokładnie jeden rozkład probabilistyczny, który ma taką funckję charakterystyczną.*

Fakt 73. *Jeśli mamy ciąg zmiennych losowych $X_1, X_2, \ldots$ o dystrybuantach $F_1(t), F_2(t), \ldots$ oraz odpowiadających im funkcjach charakterystycznych $\phi_{X_1}(t)\phi_{X_2}(t), \ldots$, które spełniają*

$$\forall_t \phi_X(t) \overset{i\to\infty}{\longrightarrow} \phi(t) \leftarrow \textit{f.charakterystyczna pewnej zmiennej losowej } X$$

to wtedy

$$F_i(t) \overset{i\to\infty}{\longrightarrow} F(t) \leftarrow \textit{dystrybuanta zmiennej losowej } X$$

w punktach ciągłości $F(t)$.

8 Twierdzenia graniczne

Kiedy wszystko inne zawiedzie...

8.1 Centralne twierdzenie graniczne

Chcemy przesłać info z Wrocławia do Nowego Jorku. Korzystamy z sieci, który ma strukturę tandemu (idziemy od oczka do oczka). Zakładamy, że mamy na drodze 100 węzłów. Zakładamy, że na każdym węźle może zdarzyć się pewne opóźnienie, które na każdym węźle jest losowe oraz niezależne pomiędzy węzłami oraz jednakowo rozłożone (takie samy rozkłady). Wiadomo, że opóźnienie X_i na i-tym węźle jest takie, że $E(X_i) = 10$, $Var(X_i) = 4$. Jaka jest szansa, że nasza informacja dotrze do Nowego Jorku przed upływem określonego czasu ($= 1050$)?

Przyglądając się z dystansu, zauważamy, że to co nas interesuje to zjawisko, które opisać możemy jako sumę dużej liczby podzjawisk (każde z nich jest łatwiejsze w analizie niż cała sieć jednocześnie), które mają jednorodną naturę i są wzajemnie niezależne.

Uwaga 27. Jeśli $X_1, X_2, \ldots$ są wzajemnie niezależne oraz mają takie same rozkłady, to piszemy, że $X_1, X_2, \ldots$ są i.i.d. (independent identically distributed).

Niech X_i - opóźnienia. Interesuje nas ile wynosi

$$P\left(\sum_{i=1}^{100} X_i \leq 1050\right) \approx ?$$

Twierdzenie 74. (CENTRALNE TWIERDZENIE GRANICZNE) *Niech $X_1, X_2, \ldots$ będą i.i.d. oraz $E(X_i) = m$, $Var(X_i) = \sigma^2 > 0$. Wtedy dla każdego $x \in \mathbb{R}$ mamy*

$$P\left(\frac{\sum_{i=1}^n X_i - n \cdot m}{\sqrt{n \cdot \sigma^2}} \leq x\right) \xrightarrow{n\to\infty} P(\mathcal{N} \leq x), \qquad gdzie\ \mathcal{N} \sim N(0,1) \tag{33}$$

Dowód. Niech

$$X_k^* = \frac{X - m}{\sigma}$$

Wtedy

$$\frac{\sum_{i=1}^n X_i - n \cdot m}{\sqrt{n \cdot \sigma^2}} = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i^* =: R$$

Niech $\phi_R(t)$ będzie funkcją charakterystyczną R. Zauważmy, że

$$\phi_n(t) = Ee^{i(t\frac{1}{\sqrt{n}})\sum X_i^*} = \left(Ee^{i\frac{t}{\sqrt{n}}X_1^*}\right)^n = \phi_{X_1^*}\left(\frac{t}{\sqrt{n}}\right)^n \tag{34}$$

Korzystając z rozwinięcia Taylora dla funkcji e^t w punkcie $t = 0$ mamy

$$\phi_{X_1^*}\left(\frac{t}{\sqrt{n}}\right) = Ee^{i\frac{t}{\sqrt{n}}X_1^*} = E\left(1 + i\frac{t}{\sqrt{n}}X_1^* + \frac{\left(i\frac{t}{\sqrt{n}}X_1^*\right)}{2}\right) - (-1 + o(1)) = \left(1 + i\frac{t}{\sqrt{n}}EX_1^* - \frac{t^2}{2n}E\left(X_1^*\right)^2\right)(1 + o(1))$$

Wstawiając momenty

$$E(X_1^*) = 0$$
$$E(X_1^*)^2 = 1$$

otrzymujemy

$$1 - \frac{t^2}{2n}(1 + o(1))$$

zatem nasze $\phi_n(t)$ z (34) dąży do $\left(1 - \frac{t^2}{2n}\right)^n \to e^{-\frac{t^2}{2}}$, czyli funkcja charakterystyczna ma rozkład $N(0,1)$. □

Uwaga 28. Jeśli $P(X_i = 1) = 1 - P(X_i = 0) = p$, to wtedy

$$P\left(\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq x\right) \xrightarrow{n\to\infty} P\left(\mathcal{N} \leq x\right) \qquad \text{gdzie } \mathcal{N} \sim N(0,1)$$

Uwaga 29. Oznaczenie

$$\Psi(x) = P(\mathcal{N} > x), \qquad \mathcal{N} \sim N(0,1)$$
$$\Phi(x) = P(\mathcal{N} \leq x)$$

Wróćmy teraz do problemu podanego na początku rozdziału. Mamy już odpowiednie narzędzia, aby je rozwiązać. Zauważmy, że $m = EX_1 = 10$, $\sigma^2 = 4$, $n = 100$.

$$\begin{aligned} P\left(\sum_{i=1}^n X_i \leq 1050\right) &= P\left(\sum X_i - 100 \cdot 10 \leq 1050 - 100 \cdot 10\right) \\ &= P\left(\frac{\sum X_i - 1000}{\sqrt{100 \cdot 4}} \leq \frac{1050 - 1000}{\sqrt{100 \cdot 4}}\right) \approx P\left(\mathcal{N} \leq \frac{50}{20}\right) \\ &= P(\mathcal{N} \leq 2,5) \approx 0.9938 \end{aligned}$$

8.2 Twierdzenie Poissona

Prawdopodobieństwo awarii danego komputera wynosi $p = \frac{1}{100}$. Komputery psują się niezależnie. Zakładamy, że w sieci jest 100 komputerów.

1. jaka jest szansa, że wszystkie komputery pracują?
2. jaka jest szansa, że co najmniej 2 komputery są uszkodzone?
3. jak duża powinna być sieć, aby z prawdopodobieństwem 95% w sieci znalazł się co najmniej 1 uszkodzony

odpowiedzi

1. $\left(\frac{99}{100}\right)^{100}$
2. $1 - \left(\frac{99}{100}\right)^{100} - \binom{100}{1}\frac{1}{100}\left(\frac{99}{100}\right)^{99}$
3. $1 - \left(\frac{99}{100}\right)^{n} \geq \frac{95}{100}$. $n \geq ?$

Twierdzenie 75. (Poisson) *Niech $B_n \sim B(n, p_n)$. Jeśli $n \cdot p_n \overset{n\to\infty}{\longrightarrow} \lambda > 0$, to*

$$P(B_n = k) \overset{n\to\infty}{\longrightarrow} e^{-\lambda}\frac{\lambda^k}{k!} \tag{35}$$

Dowód. Niech $\phi_n(t)$ będzie funkcją charakterystyczną rozkładu $B(n, p_n)$. Korzystając z tabelki na stronie 27

$$\phi_n(t) = (p_n e^{it} + (1-p_n))^n = (1 - p_n(1-e^{it}))^n = \left(1 - \frac{p_n n(1-e^{it})}{n}\right)^n \overset{n\to\infty}{\longrightarrow} e^{\lambda(1-e^{it})}$$

co odpowiada rozkładowi Poissona. □

A wracając do przykładu...

1. $P(B_{100} = 0) \sim e^{-\lambda} = e^{-1} \approx 0.368$
2. $1 - P(B_{100} = 0) - P(B_{100} = 1) \approx 1 - e^{-1} - e^{-1} \approx 0.264$
3. $1 - P(B_n = 0) \approx 1 - e^{-n\frac{1}{100}} \geq 0.95 \Longrightarrow n \geq 300$.

9 Zagadnienia estymacji

Czyli skąd biorą się EX oraz $VarX$. Będziemy od tej pory zakładać model, w którym

1. możemy wielokrotnie obserwować dane zjawisko
2. wyniki obserwacji są niezależne
3. warunki, w których przeprowadzamy eksperymenty są niezmienne

Formalnie obserwujemy $X_1, X_2, \ldots, X_n$ - i.i.d. (są wzajemnie niezależne oraz mają takie same rozkłady). Po dokonaniu obserwacji zjawiska mamy:

$$X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$$

9.1 Średnia

Przybliżenia prawdziwej EX_1 będziemy liczyć na podstawie tzw. *średniej próbkowej*

$$\bar{X} := \frac{\sum_{i=1}^{n} X_i}{n}$$

Z prawa wielkich liczb wynika, że dla odpowiednio dużych n jest $\bar{X} \to EX_1$.

9.2 Wariancja

Przypomnijmy, że

$$Var(X) = E(X^2) - (EX)^2 = E(X - EX)^2$$

Z wielkością $E(X^2)$ postępujemy analogicznie jak z estymacją EX. Jeśli mamy daną próbę obserwacyjną $X_1, X_2, \ldots, X_n$, to

$$EX^2 \approx \frac{\sum_{i=1}^n X_i^2}{n}$$

Ostatecznie estymatorem wariancji jest *wariancja próbkowa*

$$S^2 = \frac{\sum_{i=1}^n X_i^2}{n} - \left(\frac{\sum_{i=1}^n X_i}{n}\right)^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2, \qquad \text{gdzie } \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \tag{36}$$

9.3 Przykłady

Przykład 76. W kolejnych 10 dniach obserwowany był czas potrzebny na połączenie z internetem poprzez modem (od momentu włączenia komputera). Otrzymane wyniki: $1, 5, 4, 4, 2, 1, 3, 2, 1, 2$ $[s]$. Wtedy średnia próbkowa $\bar{X} = \frac{1+5+\cdots+2}{10} = 2.5$, wariancja próbkowa $S^2 = \frac{1+25+16+\cdots+4}{10} - 2.5^2 = 1.85$.

Pytanie: jak duży popełniamy błąd biorąc za wartość oczekiwaną EX średnią próbkową $\bar{X}$?

Oznaczenia:

m - prawdziwa wartość oczekiwana, EX
$\bar{X}$ - średnia próbkowa (jest losowa) z próby $(X_1, X_2, \ldots, X_n)$

Cel: oszacować $P\left(\left|\bar{X}_n - m\right| > a\right) = ?$. (na chwilę założymy, że znamy prawdziwe odchylenie standardowe σ zmiennej losowej X).

Rozwiązanie:

$$\begin{aligned} P\left(\left|\bar{X}_n - m\right| > a\right) &= P\left(\left|\frac{\sum_{i=1}^n X_i - nm}{n}\right| > a\right) \\ &= P\left(\left|\frac{\sum_{i=1}^n X_i - nm}{\sigma\sqrt{n}}\right| > \frac{\sqrt{n}a}{\sigma}\right) \sim P\left(|\mathcal{N}| > \frac{\sqrt{n}a}{\sigma}\right) \\ &= 2 - 2P\left(\mathcal{N} < \frac{a\sqrt{n}}{\sigma}\right) \end{aligned}$$

Np. jeśli $n = 50$, $a = 0.02$, $\sigma = 1$, to $P\left(\left|\bar{X}_n - m\right| > a\right) \approx 0.888$

Przykład 77. Znajdźmy liczbę doświadczeń, które należy wykonać aby błąd oszacowania nieznanej wartości oczekiwanej m przez wartość średniej próbkowej $\bar{X}_n$ był nie większy niż 10% odchylenia standardowego σ z prawdopodobieństwem 99%.

Rozwiązanie:

$$P\left(\left|\bar{X}_n - m\right| \leq 0.1\sigma\right) \geq 0.99$$

Ile wynosi n?

$$P\left(\left|\bar{X}_n - m\right| \leq 0.1\sigma\right) = P\left(\frac{\sum_{i=1}^n X_i - nm}{\sigma\sqrt{n}} \leq 0.1\sqrt{n}\right) \approx P\left(|\mathcal{N}| \leq 0.1\sqrt{n}\right)$$

dla jakiego n zachodzi

$$P\left(|\mathcal{N}| \leq 0.1\sqrt{n}\right) = 0.99 \quad ?$$

Policzymy korzystając ze zdarzenia przeciwnego

$$\begin{aligned} P\left(|\mathcal{N}| > 0.1\sqrt{n}\right) &= 0.01 \\ 2 - 2P\left(\mathcal{N} < 0.1\sqrt{n}\right) &= 0.01 \\ P\left(\mathcal{N} < 0.1\sqrt{n}\right) &= 0.995 \end{aligned}$$

Z tablic rozkładu normalnego

$$P(\mathcal{N} < 2.58) \approx 0.995$$

Zatem

$$0.1\sqrt{n} \approx 2.58 \Longleftrightarrow n \approx 665.6$$

10 Tablica dystrybuanty rozkładu normalnego

Tablica dla $\Phi(u) := P(\mathcal{N} < u)$, gdzie $\mathcal{N} \sim N(0,1)$. Warto pamiętać, że $\Phi(u) + \Phi(-u) = 1$.

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
4.1	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09

źródło: `http://www.mimuw.edu.pl/~aweber/zadania/geo/normalny.html`