Rozkład gamma

Definicja 3.17. Mówimy, że zmienna losowa X ma *rozkład gamma*, jeśli jej funkcja gęstości jest określona wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x \le 0, \\ \dfrac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} \text{ dla } x > 0, \end{cases}$$

gdzie $b > 0$ i $p > 0$ oznaczają pewne stałe.

Przypomnijmy, że *funkcją gamma* nazywamy całkę Eulera drugiego rodzaju daną wzorem

$$\Gamma(p) = \int_0^\infty x^{p-1} \cdot e^{-x} dx, \quad p > 0.$$

Mamy przy tym

$$\Gamma(p+1) = p \cdot \Gamma(p)$$

oraz

$$\Gamma(n+1) = n!,$$

gdy wartość n jest liczbą naturalną.

Przykład 3.7. Zmienna losowa X podlega rozkładowi gamma dla $b = 1$ i $p = 2$. Obliczyć prawdopodobieństwo, że zmienna ta przyjmie wartość nie większą od 1 oraz wyznaczyć dystrybuantę.

Zgodnie z definicją 3.17 funkcja gęstości ma postać

$$f(x) = \begin{cases} 0 \text{ dla } x \le 0, \\ \dfrac{1}{\Gamma(2)} x \cdot e^{-x} \text{ dla } x > 0. \end{cases}$$

Ponieważ $\Gamma(2) = 1$, więc

$$P(X \le 1) = \int_0^1 x \cdot e^{-x} dx = -(x \cdot e^{-x} + e^{-x})\Big|_0^1 = \frac{e-2}{e} \approx 0{,}26$$

oraz

$$F(x) = \begin{cases} 0 \text{ dla } x \le 0, \\ 1-(x+1)e^{-x} \text{ dla } x > 0. \end{cases}$$

Szczególnym przypadkiem rozkładu gamma jest *rozkład wykładniczy*, który otrzymujemy dla $b = \lambda$ i $p = 1$. Wówczas

$$f(x) = \begin{cases} 0 \text{ dla } x \le 0, \\ \lambda \cdot e^{-\lambda x} \text{ dla } x > 0. \end{cases}$$

Rozkład beta

Definicja 3.18. Zmienna losowa X ma *rozkład beta*, jeżeli jej gęstość jest określona wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x < 0, \\ \dfrac{1}{B(p,q)} x^{p-1}(1-x)^{q-1} \text{ dla } 0 \le x \le 1, \\ 0 \text{ dla } x > 1, \end{cases}$$

gdzie $p > 0$ i $q > 0$ oznaczają pewne stałe.

Przypomnijmy, że *funkcją beta* nazywamy całkę Eulera pierwszego rodzaju, którą określa wzór

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx, \quad p > 0, \quad q > 0$$

oraz że zachodzą następujące zależności:

$$B(p,q) = \frac{q-1}{p+q-1} B(p, q-1),$$
$$B(p,q) = \frac{\Gamma(p)\cdot\Gamma(q)}{\Gamma(p+q)},$$
$$B(p,n) = \frac{(n-1)!}{p\cdot(p+1)\cdot\ldots\cdot(p+n-1)},$$
$$B(m,n) = \frac{(n-1)!(m-1)!}{(m+n-1)!}.$$

gdzie m i n oznaczają liczby naturalne.

Przykład 3.8. Zmienna losowa X podlega rozkładowi beta dla $p = 3$ i $q = 2$. Obliczyć prawdopodobieństwo, że przyjmie ona wartość mniejszą od 0,1.
Na podstawie definicji 3.18 funkcja gęstości dana jest wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x < 0 \text{ lub } x > 1, \\ \dfrac{1}{B(3,2)} x^2(1-x) \text{ dla } 0 \le x \le 1. \end{cases}$$

Ponieważ

$$\frac{1}{B(3,2)} = \frac{(3+2-1)!}{1!\cdot 2!} = \frac{4!}{2} = 12,$$

więc

$$P(X < 0{,}1) = 12\int_0^{0{,}1} x^2(1-x)dx = 12\left(\frac{1}{3}x^3 - \frac{1}{4}x^4\right)\Bigg|_0^{0{,}1} = 0{,}0037.$$

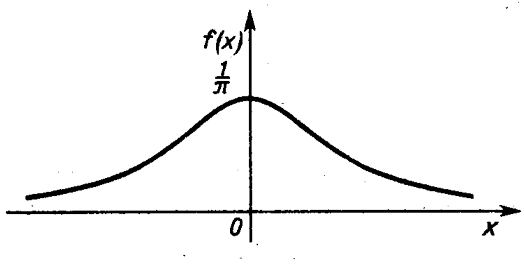
Rozkład Cauchy'ego

Definicja 3.19. Mówimy, że zmienna losowa X ma *rozkład Cauchy'ego*, gdy jej gęstość jest określona wzorem

$$f(x) = \frac{1}{\pi} \cdot \frac{\alpha}{\alpha^2 + (x-a)^2},$$

gdzie $\alpha > 0$, a a oznacza dowolną liczbę rzeczywistą, $-\infty < x < +\infty$..

W przypadku $a = 0$ i $\alpha = 1$ wykres gęstości jej następujący:



Dla $a = 0$ dystrybuanta tego rozkładu ma postać

$$F(x) = \frac{\alpha}{\pi} \int_{-\infty}^{x} \frac{dt}{t^2 + \alpha^2} = \frac{1}{\pi} \operatorname{arctg} \frac{t}{\alpha} \Big|_{-\infty}^{x} = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} \frac{x}{\alpha}.$$

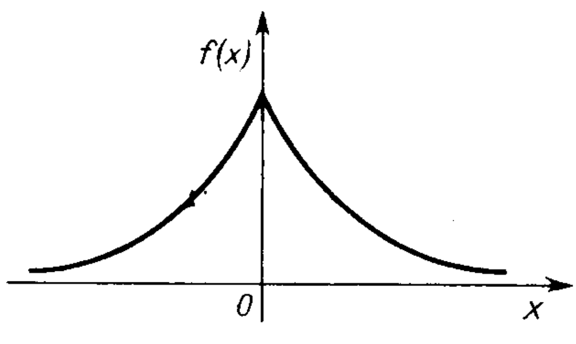
Rozkład Laplace'a

Definicja 3.20. Mówimy, e zmienna losowa X podlega *rozkładowi Laplace'a*, jeśli jej gęstość prawdopodobieństwa jest określona wzorem

$$f(x) = \frac{1}{2} \lambda \cdot \exp(-\lambda |x|),$$

gdzie $\lambda > 0$ i $-\infty < x < +\infty$.

Wykres tej funkcji jest następujący:



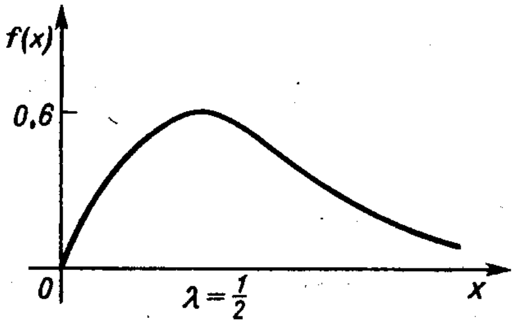
Rozkład Maxwella

Definicja 3.21. Mówimy, że zmienna losowa X ma *rozkład Maxwella*, gdy jej gęstość prawdopodobieństwa dana jest wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x \le 0, \\ \dfrac{4\lambda^{3/2}}{\sqrt{\pi}} x^2 \exp(-\lambda x^2) \text{ dla } x > 0, \end{cases}$$

przy czym $\lambda > 0$.

Wykres powyższej funkcji gęstości dla $\lambda = 1/2$ jest następujący:



Zadania

1. Zmienna losowa X ma rozkład według gęstości danej wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x < 0, \\ C \cdot \sin x \text{ dla } 0 \le x \le \dfrac{\pi}{3}, \\ 0 \text{ dla } x > \dfrac{\pi}{3}. \end{cases}$$

A. Obliczyć stałą C.
B. Podać dystrybuantę.
C. Obliczyć $P(\pi/6 \le X \le \pi/4)$.

2. Zmienna losowa X podlega rozkładowi według gęstości danej wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x < 1, \\ \ln x \text{ dla } 1 \le x \le a, \\ 0 \text{ dla } x > a. \end{cases}$$

A. Obliczyć stałą a.
B. Podać dystrybuantę.
C. Obliczyć $P(2 \le X \le e)$.

3. Strzałka minutowa zegara elektrycznego zmienia położenie w końcu każdej minuty. Jeżeli strzałka wskazuje a minut, to rzeczywisty czas t jest zmienną losową przyjmującą wartości z przedziału $[a, a+1]$. Znaleźć gęstość prawdopodobieństwa zmiennej t.

4. Zmienna losowa przyjmuje wartości z przedziału [1, 7], przy czym prawdopodobieństwo przyjęcia przez nią wartości z wycinka przedziału [3, 4] jest pięć razy większe od prawdopodobieństwa przyjęcia wartości z wycinka o tej samej długości z przedziału [1, 3), a także z przedziału (4, 7]. Podać gęstość, dystrybuantę i obliczyć $P(3{,}6 \le X \le 4{,}7)$.

5. Zmienna losowa X podlega rozkładowi według trójkąta utworzonego przez oś Ox oraz proste $y = ax + a$ i $y = -x + 4$. Dobrać odpowiednio stałą a i podać gęstość prawdopodobieństwa tej zmiennej.

6. Zmienna losowa podlega rozkładowi według trójkąta utworzonego przez oś Ox oraz proste $y = ax$ $(a > 0)$ i $y = -2a^2x + 4\sqrt{5}$. Dobrać odpowiednio stałą a i podać gęstość prawdopodobieństwa tej zmiennej.

7. Zmienna losowa X podlega rozkładowi według gęstości danej wzorem

$$f(x) = \begin{cases} 0 \text{ dla } x < 1, \\ \dfrac{x}{4} \text{ dla } 1 \le x \le 3, \\ 0 \text{ dla } x > 3. \end{cases}$$

Wyznaczyć dystrybuantę oraz obliczyć $P(1{,}4 \le X \le 2)$. Wykonać wykresy gęstości i dystrybuanty. Zaznaczyć wartość obliczonego prawdopodobieństwa na wykonanych wykresach.

8. Zmienna losowa podlega rozkładowi według trapezu równoramiennego o podstawie x i kącie nachylenia $\alpha = \pi/6$, przy czym $a \le x \le b$. Napisać równanie gęstości.

9. Zmienna losowa podlega rozkładowi według trójkąta równoramiennego o podstawie $a \le x \le 3a$. Napisać równanie gęstości.

10. Zmienna losowa X przyjmuje wartości $x_1 = 1$, $x_2 = 2$ i $x_3 = 4$ z prawdopodobieństwami odpowiednio 1/3, 1/4 i 5/12. Wyznaczyć wartości dystrybuanty $F(1)$, $F(2{,}5)$, $F(4)$ i $F(6)$.

11. Pewna gra polega na rzucie trzema monetami i otrzymaniu wygranej 10 zł w przypadku wyrzucenia trzech orłów i przegraniu 6 zł w pozostałych przypadkach. Traktując wygraną jako zmienną losową podać jej funkcję prawdopodobieństwa i dystrybuantę.

3.4. Funkcje zmiennych losowych

Ograniczymy się do podania twierdzenia.

Twierdzenie 3.3. *Jeżeli zmienna X jest zmienną losową, a funkcja g(x) jest funkcją B-mierzalną, to zmienna Y=g(X) jest też zmienną losową.*

Dowód pomijamy. ■

Funkcję $g(x)$ nazywa się funkcją B-mierzalną, gdy określony przez nierówność $g(x) < y$ zbiór argumentów x jest zbiorem borelowskim dla każdej wartości y. W szczególności każda funkcja ciągła jest B-mierzalna.

3.5. Określenia momentów

Momenty są charakterystykami liczbowymi rozkładów zmiennych losowych, które umożliwiają szybkie porównanie rozkładów między sobą.

Definicja 3.22. *Momentem rzędu l* (l = 1, 2, ...) względem liczby c zmiennej losowej X nazywamy dla zmiennej losowej skokowej sumę

$$\mu_l' = \sum_k (x_k - c)^l p_k,$$

a dla zmiennej losowej ciągłej całkę

$$\mu_l' = \int_{-\infty}^{+\infty} (x-c)^l f(x) dx,$$

jeśli odpowiednio szereg lub całka są bezwzględnie zbieżne, gdzie p_k oznacza prawdopodobieństwo przyjęcia przez zmienną losową X wartości x_k, a $f(x)$ – gęstość prawdopodobieństwa.

Uwagi
1. Jeśli zmienna losowa X jest skokowa o skończonej liczbie punktów skokowych x_k lub ciągła o gęstości ograniczonej i większej od zera na przedziale skończonym $[a, b]$, to warunek bezwzględnej zbieżności odpada, bo w pierwszym przypadku mamy do czynienia z sumą skończoną, a w drugim – z całką oznaczoną.
2. Z definicji wynika, że moment zależy tylko od rozkładu. Jeśli zatem X_1 i X_2 oznaczają dwie zmienne losowe o tym samym rozkładzie, to momenty tych zmiennych są równe. Może jednak zdarzyć się, że rozkład nie ma momentów.

Definicja 3.23. Jeżeli $c = 0$, to moment nazywamy *zwykłym* i oznaczamy przez m_l, tj.

$$m_l = \sum_k x_k^l p_x \quad \text{dla zmiennej losowej skokowej,}$$

$$m_l = \int_{-\infty}^{+\infty} x^l f(x) dx \quad \text{dla zmiennej losowej ciągłej.}$$

Definicja 3.24. Moment zwykły rzędu pierwszego nazywamy *wartością przeciętną* (*wartością oczekiwaną*, *nadzieją matematyczną*) i oznaczamy symbolem $E(X)$, tj.

$$E(X) = \sum_k x_k p_k \quad \text{dla zmiennej losowej skokowej,}$$

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx \text{ dla zmiennej losowej ciągłej.}$$

Przykład 3.9. Zmienna losowa X podlega rozkładowi $P(X = x_k) = p_k$, $k = 1, 2, \ldots$, gdzie

$$p_k = \frac{2}{3^k}, \quad x_k = \frac{(-1)^k 3^k}{k}.$$

Zbadać istnienie momentu rzędu pierwszego.
Na podstawie definicji 3.24 należy obliczyć sumę iloczynów wartości zmiennej losowej przez odpowiadające im prawdopodobieństwa. Mamy

$$\sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} \frac{2}{3^k} \cdot \frac{(-1)^k 3^k}{k} = 2 \sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -2 \cdot \ln 2.$$

Poprzestanie na tych obliczeniach prowadzi do błędnego wniosku, że wartość oczekiwana istnieje. Mamy jednak

$$\sum_{k=1}^{\infty} |x_k| \cdot p_k = 2 \sum_{k=1}^{\infty} \frac{1}{k} = \infty,$$

co oznacza, że szereg nie jest bezwzględnie zbieżny (jest to szereg harmoniczny rzędu pierwszego).

Przykład 3.10. Zmienna losowa X ma rozkład Cauchy'ego postaci

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}, \quad -\infty < x < +\infty.$$

Zbadać istnienie wartości oczekiwanej w tym rozkładzie.
Mamy

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x}{1+x^2} dx = \frac{2}{\pi} \int_{0}^{+\infty} \frac{x}{1+x^2} dx = \frac{1}{\pi} \cdot \ln(1+x^2)\Big|_0^{\infty} = \infty,$$

co oznacza, że rozkład Cauchy'ego nie ma momentu rzędu pierwszego. Wynika z tego, że rozkład ten nie ma żadnego momentu.

Definicja 3.25. Jeśli w definicji 3.22 przyjmiemy $c = m_1 = E(X)$, to taki moment nazywamy *momentem centralnym rzędu* l i oznaczamy przez μ_l, czyli

$$\mu_l = \sum_k (x_k - m_1)^l p_k \text{ dla zmiennej losowej skokowej,}$$

$$\mu_l = \int_{-\infty}^{+\infty} (x - m_1)^l f(x)dx \text{ dla zmiennej losowej ciągłej.}$$

Definicja 3.26. Moment centralny rzędu drugiego μ_2 nazywamy *wariancją*, a pierwiastek z niego – *odchyleniem standardowym*. Wariancję oznaczamy przez σ^2 lub $D^2(X)$, a odchylenie standardowe przez σ lub $D(X)$, tj.

$$D^2(X) = \sum_k [x_k - E(X)]^2 p_k \text{ dla zmiennej losowej skokowej,}$$

$$D^2(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)dx \text{ dla zmiennej losowej ciągłej.}$$

Moment centralny można przedstawić za pomocą wartości oczekiwanej:

$$\mu_l = E((X - E(X))^l).$$

W przypadku zmiennej losowej skokowej mamy bowiem

$$\mu_l = \sum_k (x_k - m_1)^l p_k = \sum_k \xi_k p_k = E(\xi) = E((X - m_1)^l) = E((X - E(X))^l),$$

gdzie przyjęto oznaczenie $(x_k - m_1)^l = \xi_k$. W przypadku zmiennej losowej ciągłej uzasadnienie jest podobne.

Przykład 3.10. Wyrazić trzy pierwsze momenty centralne poprzez momenty zwykłe. Rozważymy przypadek zmiennej losowej skokowej (dla zmiennej losowej ciągłej rozumowanie jest podobne). Mamy:

$$\begin{aligned}
\mu_1 &= E(X - E(X)) = \sum_k (x_k - m_1)p_k = \sum_k x_k p_k - m_1 \sum_k p_k = m_1 - m_1 = 0,\\
\mu_2 &= E((X - E(X))^2) = \sum_k (x_k - m_1)^2 p_k\\
&= \sum_k x_k^2 p_k - 2m_1 \sum_k x_k p_k + m_1^2 \sum_k p_k = m_2 - 2m_1^2 + m_1^2 = m_2 - m_1^2,\\
\mu_3 &= E((X - E(X))^3) = \sum_k (x_k - m_1)^3 p_k\\
&= \sum_k x_k^3 p_k - 3m_1 \sum_k x_k^2 p_k + 3m_1^2 \sum_k x_k p_k - m_1^3 \sum_k p_k\\
&= m_3 - 3m_1 m_2 + 3m_1^3 - m_1^3 = m_3 - 3m_1 m_2 + 2m_1^3.
\end{aligned}$$

Momenty $m_1 = E(X)$ (wartość oczekiwana) i $\mu_2 = D^2(X)$ (wariancja) odgrywają dużą rolę w rachunku prawdopodobieństwa i statystyce matematycznej i dlatego ich własności rozważymy w kolejnych punktach.

3.6. Własności wartości oczekiwanej

Definicja 3.27. *Wartością oczekiwaną funkcji g zmiennej losowej X* (funkcji $g(X)$, gdzie X oznacza zmienną losową) nazywamy wyrażenie

$$E(g(x)) = \sum_k g(x_k) p_k,$$

gdy zmienna losowa X jest skokowa o punktach skokowych x_k i skokach p_k oraz wyrażenie

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

gdy zmienna losowa X jest ciągła i ma gęstość $f(x)$.

Uwaga: W powyższej definicji szereg i całka powinny być bezwzględnie zbieżne.

Podstawowe własności wartości oczekiwanej podamy w kilku twierdzeniach.

Twierdzenie 3.4. *Jeśli $g_1(X)$ i $g_2(X)$ oznaczają dwie jednoznaczne funkcje zmiennej losowej X oraz jeśli istnieją wartości oczekiwane $E(g_1(X))$ i $E(g_2(X))$, to*

$$E(g_1(X) + g_2(X)) = E(g_1(X)) + E(g_2(X)).$$

Dowód. Rozpatrzymy przypadek zmiennej losowej skokowej (dla zmiennej losowej ciągłej dowód jest podobny). Zauważmy przede wszystkim, że wartość oczekiwana sumy

$$E(g_1(X) + g_2(X)) = \sum_k (g_1(x_k) + g_2(x_k))p_k$$

istnieje, bo szereg z prawej strony jest bezwzględnie zbieżny, co wynika z nierówności

$$\sum_k |g_1(x_k) + g_2(x_k)|p_k \le \sum_k |g_1(x_k)|p_k + \sum_k |g_2(x_k)|p_k$$

i faktu, że z założenia oba szeregi po prawej stronie tej nierówności są bezwzględnie zbieżne. Z bezwzględnej zbieżności szeregów wynika ich zbieżność, a więc

$$\sum_k (g_1(x_k) + g_2(x_k)) = \sum_k g_1(x_k) + \sum_k g_2(x_k) = E(g_1(X)) + E(g_2(X)).$$ ■

Twierdzenie 3.5. *Wartość oczekiwana stałej a równa się tej stałej, tj.*

$$E(a) = a.$$

Dowód jest oczywisty. ■

Twierdzenie 3.6. *Zachodzi równość*

$$E((aX)^n) = a^n E(X^n),$$

gdzie a oznacza pewną stałą, a n – liczbę naturalną.

Dowód (dla zmiennej losowej skokowej). Mamy

$$E((aX)^n) = \sum_k (ax_k)^n p_k^n = a^n \sum_k (x_k p_k)^n = a^n E(X^n).$$ ■

Bezpośrednią konsekwencją twierdzeń 3.4 – 3.6 jest

Twierdzenie 3.7. *Wartość oczekiwana przekształcenia liniowego zmiennej losowej X jest równa przekształceniu liniowemu wartości oczekiwanej tej zmiennej, tj.*

$$E(aX+b)=aE(X)+b,$$

gdzie a i b oznaczają pewne stałe.

Twierdzenie 3.8. *Jeżeli zmienna losowa Y ma postać* $Y = X - E(X)$, *to* $E(Y) = 0$.

Dowód. Wartość oczekiwana jest stałą, a więc wartość oczekiwana z tej wartości jest niej równa (na podstawie twierdzenia 3.5), czyli $E(E(X)) = E(X)$. Uwzględniając twierdzenie 3.7 mamy

$$E(Y)=E(X-E(X))=E(X)-E(E(X))=E(X)-E(X)=0. \qquad \blacksquare$$

Twierdzenie 3.9 (nierówność Schwarza). *Jeśli X i Y oznaczają zmienne losowe o rozkładach, dla których istnieją momenty zwykłe rzędu pierwszego i drugiego, to*

$$|E(X\cdot Y)|\le \sqrt{E(X^2)\cdot E(Y^2)}.$$

Dowód. Niech $Z = (X - aY)^2$ oznacza zmienną losową z parametrem a, który jest liczbą rzeczywistą. Na podstawie twierdzeń 3.4 i 3.6 wartość oczekiwana tej zmiennej losowej jest równa

$$\begin{aligned}E(Z)&=E((X-aY)^2)=E(X^2)+E(-2a\cdot X\cdot Y)+E(a^2Y^2)\\&=E(X^2)-2aE(X\cdot Y)+a^2E(Y^2).\end{aligned}$$

Wartość oczekiwana $E(Z)$ jest nieujemna, bo zmienna losowa Z jest nieujemna. Rozwiążmy nierówność

$$E(X^2)-2aE(X\cdot Y)+a^2E(Y^2)\ge 0.$$

Aby ta nierówność była prawdziwa, wyróżnik Δ tego trójmianu kwadratowego (z uwagi na wartość a) musi być niedodatni, czyli

$$\frac{1}{4}\Delta=(E(X\cdot Y))^2-E(X^2)\cdot E(Y^2)\le 0.$$

Z nierówności tej wynika nierówność podana w twierdzeniu. $\blacksquare$

Przykład 3.11. Zmienna losowa X przyjmuje wartości $x_1, x_2, \ldots, x_n$, każdą z jednakowym prawdopodobieństwem p. Obliczyć wartość oczekiwaną tej zmiennej losowej.

Jeżeli zmienna losowa X przyjmuje każdą z n wartości z jednakowym prawdopodobieństwem p, to $n \cdot p = 1$, skąd $p = 1/n$. Wówczas

$$E(X)=\sum_{k=1}^{n}x_kp=p\sum_{k=1}^{n}x_k=\frac{1}{n}\sum_{k=1}^{n}x_k,$$

czyli wartość oczekiwana podanej zmiennej losowej jest średnią arytmetyczną.

Przykład 3.12. Zmienna losowa X przyjmuje wartości $x_1, x_2, \ldots, x_i$ z prawdopodobieństwami odpowiednio $n_1p, n_2p, \ldots, n_ip$. Obliczyć wartość oczekiwaną tej zmiennej.

Suma wszystkich prawdopodobieństw musi być równa 1, tj.

$$n_1 p + n_2 p + \dots + n_i p = 1,$$

skąd

$$p = \frac{1}{n_1 + n_2 + \dots + n_i}.$$

Jeśli oznaczymy $n = n_1 + n_2 + \dots + n_i$, to otrzymamy $p = 1/n$. Poszczególne prawdopodobieństwa będą zatem równe

$$\frac{n_1}{n}, \frac{n_2}{n}, \dots, \frac{n_i}{n},$$

które można interpretować jako częstości względne sukcesów w n doświadczeniach. Dla wartości oczekiwanej otrzymujemy

$$E(X) = \sum_{k=1}^{i} x_k \cdot \frac{n_k}{n} = \frac{1}{n} \sum_{k=1}^{i} x_k n_k.$$

Okazuje się, że w tym przypadku wartość oczekiwana jest tzw. średnią arytmetyczną ważoną.

Zadania

1. Rzucamy sześcienną kostką do gry. Obliczyć wartość oczekiwaną liczby wyrzuconych oczek.

2. Rzucamy dwiema sześciennymi kostkami do gry. Jaka jest wartość oczekiwana sumy otrzymanej liczby oczek na obu kostkach?

3. Zorganizowano dwa turnieje: I i II. Gracz może wybrać tylko jeden z nich. Orientuje się, że w turnieju I ma szansę wygrania pierwszej nagrody w wysokości 1200 zł, drugiej – 800 zł i trzeciej – 200 zł z prawdopodobieństwami odpowiednio 0,2, 0,6 i 0,1, a w turnieju II nagrody o tej samej wysokości ma szansę wygrać z prawdopodobieństwami odpowiednio 0,3, 0,2 i 0,4. W którym turnieju jest korzystniej uczestniczyć graczowi?

4. Zorganizowano następującą loterię: rzucamy trzema kostkami i w przypadku trzech szóstek wygrywamy 100 zł plus stawkę, a w przypadku dwóch szóstek wygrywamy 10 zł plus stawkę. Ile powinna wynosić stawka, aby gra była sprawiedliwa?
   Uwaga: Gra jest sprawiedliwa, gdy wartość oczekiwana zmiennej losowej opisującej grę jest równa zero.

5. W urnie mamy 6 kul białych i 4 czarne. Ciągniemy z urny kule ze zwrotem, a ż do otrzymania kuli białej, lecz co najwyżej trzy razy. Obliczyć w tych warunkach wartość oczekiwaną liczby ciągnięć.

6. Obliczyć wartość oczekiwaną dla rozkładu jednostajnego.

7. Zmienna losowa X podlega rozkładowi Bernoulliego. Obliczyć wartość oczekiwaną tej zmiennej.

8. Zmienna losowa X podlega rozkładowi Poissona. Obliczyć wartość oczekiwaną tej zmiennej.

3.7. Wskaźniki rozrzutu

Znajomość tzw. *wskaźników rozrzutu* jest konieczna, gdy przy rozpatrywaniu dwu lub więcej rozkładów zmiennej losowej stwierdza się, że wartości oczekiwane w tych rozkładach są takie same, a trzeba rozstrzygnąć, który rozkład jest w danych warunkach lepszy. Na przykład otrzymanie tych samych średnich trafień przy strzelaniu do tarczy nie rozstrzyga, który ze strzelców jest lepszy. Decyzję można podjąć po zbadaniu rozrzutów, czyli skupień trafień wokół celu.

Jednym ze wskaźników rozrzutu jest odchylenie standardowe, czyli pierwiastek z wariancji. W kolejnych twierdzeniach podano podstawowe własności wariancji i odchylenia standardowego.

Twierdzenie 3.10. *Dla każdej wartości $c \neq m_1$ zachodzi nierówność*

$$D^2(X) < E((X-c)^2).$$

Dowód. Mamy

$$E((X-c)^2) = E((X-m_1+m_1-c)^2) = E((X-m_1)^2) + 2(m_1-c)E(X-m_1) + E((m_1-c)^2).$$

Ale

$$E((X-m_1)^2) = E((X-E(X))^2) = D^2(X),$$

na podstawie twierdzenia 3.8

$$E(X-m_1) = E(X-E(X)) = 0,$$

a na podstawie twierdzenia 3.5

$$E((m_1-c)^2) = (m_1-c)^2.$$

Zatem

$$E((X-c)^2) = D^2(X) + (m_1-c)^2 > D^2(X)$$

dla $c \neq m_1$. ■

Twierdzenie 3.11. *Dodanie stałej do wartości zmiennej losowej nie zmienia jej wariancji, tj.*

$$D^2(X+b) = D^2(X).$$

Dowód. Mamy

$$\begin{aligned} D^2(X+b) &= E((X+b-E(X+b))^2) = E((X+b-E(X)-E(b))^2) \\ &= E((X+b-E(X)-b)^2) = E((X-E(X))^2) = D^2(X). \end{aligned}$$ ■

Twierdzenie 3.12. *Jeżeli $Y = aX$, to*

$$D^2(Y) = a^2 D^2(X),$$

czyli

$$\sigma = D(Y) = |\, a \,| \cdot D(X).$$

Dowód. Mamy

$$\begin{aligned} D^2(Y) &= D^2(aX) = E((aX - E(aX))^2) = E((aX)^2 - 2aXE(aX) + (E(aX))^2) \\ &= E((aX)^2) - 2E(aX \cdot E(aX)) + (E(aX))^2 \\ &= E((aX)^2) - 2E(aX) \cdot E(aX) + (E(aX))^2 = E((aX)^2) - (E(aX))^2 \\ &= a^2 E(X^2) - a^2 (E(X))^2 = a^2 (E(X^2) - E^2(X)) = a^2 D^2(X), \end{aligned}$$

co kończy dowód. ■

Twierdzenie 3.13. *Wariancja dowolnej stałej b jest równa zeru, tj.*

$$D^2(b) = 0.$$

Dowód wynika bezpośrednio z twierdzenia 3.5 i definicji wariancji. ■

Przykład 3.13. Sklep ma do wyboru zakup z hurtowni jednej z dwóch partii konserw A i B po 100 sztuk, każda o jednakowej cenie i jakości. Dane są ciężary (w kg) puszek partii A i B (odpowiednio A_k i B_k) oraz liczby puszek n_k o wadze A_k i m_k o wadze B_k:

A_k	0,7	0,8	0,9	1,0	1,1	1,2	1,3
n_k	3	2	36	21	28	9	1

B_k	0,7	0,8	0,9	1,0	1,1	1,2	1,3
n_k	10	10	21	18	15	22	4

Którą partię należy polecić do wyboru?
Zmienna losowa dla kupujących z partii A przyjmuje wartości A_k z prawdopodobieństwami $n_k/100$, a z partii B – wartości B_k z prawdopodobieństwami $m_k/100$ (ciężary obu partii są jednakowe i wynoszą 100).
Mamy $E(A) = E(B) = 1$ (w kg), tzn należałoby sprzedać puszkę jako ważącą 1 kg. Mogłoby wydawać się, że partie A i B są równoważne. Lepsza jest jednak ta paria, której ciężary puszek są bardziej zbliżone do nominalnych (1 kg). Lepsza jest zatem paria, która przy jednakowych wartościach oczekiwanych ma mniejszy rozrzut. Obliczamy zatem wariancje:

$$\begin{aligned} D^2(A) = \frac{1}{100}[&(0{,}7-1)^2 \cdot 3 + (0{,}8-1)^2 \cdot 2 + (0{,}9-1)^2 \cdot 36 \\ &+ (1{,}0-1)^2 \cdot 21 + (1{,}1-1)^2 \cdot 28 + (1{,}2-1)^2 \cdot 9 \\ &+ (1{,}3-1)^2 \cdot 1] = 0{,}0144 \end{aligned}$$

oraz

$$D^2(B) = \frac{1}{100}[(0{,}7-1)^2 \cdot 10 + (0{,}8-1)^2 \cdot 10 + (0{,}9-1)^2 \cdot 21 + (1{,}0-1)^2 \cdot 18 + (1{,}1-1)^2 \cdot 15 + (1{,}2-1)^2 \cdot 22 + (1{,}3-1)^2 \cdot 4] = 0{,}0290.$$

Ponieważ $D^2(A) < D^2(B)$, więc należy wybrać partię A.

Przykład 3.14. Obliczyć wariancję rozkładu jednostajnego.
Ponieważ

$$D^2(X) = E(X^2) - (E(X))^2 = m_2 - m_1^2,$$

a na podstawie zadania 6 po p. 3.6 wiemy, że

$$m_1 = E(X) = \frac{a+b}{2},$$

więc wystarczy obliczyć m_2. Mamy

$$m_2 = \int_{-\infty}^{+\infty} x^2 f(x) = \frac{1}{b-a}\int_a^b x^2 dx = \frac{1}{3}\cdot\frac{1}{b-a}\cdot x^2\Big|_a^b = \frac{1}{3}(a^2+ab+b^2).$$

Zatem

$$D^2(X) = \frac{1}{3}(a^2+ab+b^2) - \frac{1}{4}(a^2+2ab+b^2) = \frac{(a-b)^2}{12}.$$

Zadania

1. Rzucamy sześcienną kostką do gry. Obliczyć wariancję liczby wyrzuconych oczek na kostce.

2. Rzucamy dwiema sześciennymi kostkami do gry. Obliczyć wariancję sumy liczb oczek otrzymanych na kostkach.

3. Obliczyć wariancję rozkładu Bernoulliego.

4. Obliczyć wariancję rozkładu Poissona.

5. Rzucamy dwie kości do gry. Oznaczmy przez X_1 zmienną losową przyjmującą wartości równe liczbie oczek na pierwszej kostce, a przez X_2 zmienną losową przyjmującą wartość 1, o ile na pierwszej kostce jest piątka i na drugiej kostce jest piątka, natomiast wartość 0 w pozostałych przypadkach.
   A. Znaleźć rozkład zmiennej losowej $X = X_1 + X_2$.
   B. Wyznaczyć dystrybuantę zmiennej losowej X i wykonać jej wykres.
   C. Obliczyć $P(4 \le X < 6)$ oraz zaznaczyć obliczone prawdopodobieństwo na wykresie dystrybuanty.
   D. Obliczyć wartość oczekiwaną i odchylenie standardowe zmiennej losowej X.

6. Myśliwy ma trzy naboje i strzela do chwili trafienia do celu lub do chwili wystrzelenia wszystkich naboi. Liczba naboi jest zmienną losową.
   A. Podać rozkład tej zmiennej wiedząc, że prawdopodobieństwo trafienia celu przy każdym strzale jest równe 0,8.
   B. Obliczyć wartość oczekiwaną i odchylenie standardowe.

7. A. Dla jakiej wartości C funkcja

$$f(x) = \begin{cases} Cx^2 \text{ dla } 0 \le x \le 3, \\ 0 \text{ dla } x < 0 \text{ i } x > 3 \end{cases}$$

   jest gęstością prawdopodobieństwa?
   B. Znaleźć dystrybuantę wyznaczonego rozkładu i wykonać jej wykres.
   C. Obliczyć $P(1 \le X < 4)$ i zaznaczyć wartość obliczonego prawdopodobieństwa na wykresie dystrybuanty.
   D. Obliczyć wartość oczekiwaną i odchylenie standardowe.

8. Zmienna losowa X przyjmuje wartości -2, 2 i 3 z prawdopodobieństwami odpowiednio 0,25, 0,5 i 0,25.
   A. Znaleźć rozkład zmiennej losowej $Y = X^2$.
   B. Obliczyć $E(Y)$ i $D(Y)$.