Statystyka

Zmienne losowe

Zmienna losowa

Zmienna losowa jest funkcją, w której każdej wartości **x** $\in$ **R** odpowiada pewien podzbiór zbioru **Ω** będący zdarzeniem losowym.

Zmienna losowa powstaje poprzez przyporządkowanie każdemu zdarzeniu elementarnemu liczby rzeczywistej. Wtedy wartość zmiennej losowej ma swoje prawdopodobieństwo.

Czyli zmienna losowa ma wartości oraz prawdopodobieństwo

Typy zmiennej losowej

Ze względu na rodzaj zbioru zdarzeń elementarnych (skończony, nieskończony, przeliczalny, nie przeliczalny) wyróżniamy typy zmiennych losowych:

- zmienna losowa skokowa - gdy Ω skończony lub nieskończony ale przeliczalny
- zmienna losowa ciągła - gdy Ω nie przeliczalny i jest przedziałem lub sumą przedziałów

Funkcja rozkładu

Funkcję przyporządkowującą każdej wartości zmiennej losowej odpowiednie prawdopodobieństwo nazywamy funkcją prawdopodobieństwa zmiennej losowej

funkcją rozkładu prawdopodobieństwa dla zmiennej losowej **skokowej**

funkcją gęstości dla zmiennej losowej **ciągłej**

Zmienna losowa skokowa
funkcja rozkładu prawdopodobieństwa

$$f(x_i) = P(X = x_i) = p_i$$

$$0 \le f(x_i) \le 1$$

oraz

$$\sum_i f(x_i) = \sum_i p_i = 1$$

Zm. losowa ciągła   funkcja gęstości

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x}$$

$$f(x) \geq 0 \quad \text{dla} \quad x \in R$$

oraz

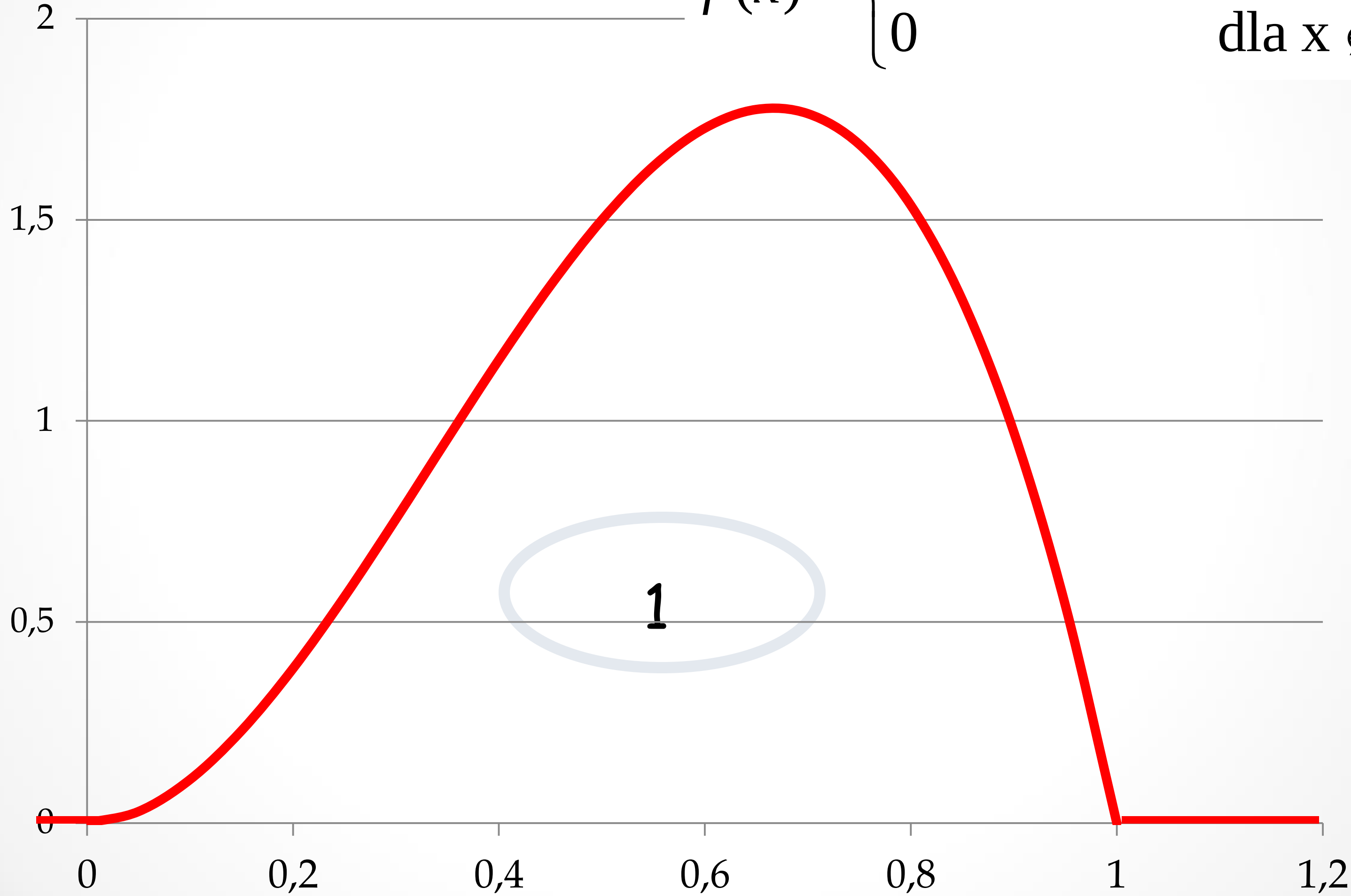
$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Zmienna losowa skokowa

x_i	$f(x_i)=p_i$
0	0,22
1	0,46
2	0,32
suma	1

Zmienna losowa ciągła

$$f(x) = \begin{cases} 12(x^2 - x^3) & \text{dla } x \in (0,1) \\ 0 & \text{dla } x \notin (0,1) \end{cases}$$



Podstawy rachunku całkowego

F(x) jest funkcją pierwotną dla f(x)

$$\bigwedge_{x\in R} F'(x) = f(x) \quad \Rightarrow \quad \int f(x)dx = F(x) + C$$

$$\int_a^b f(x)dx = \left[F(x)\right]_a^b = F(b) - F(a)$$

Podstawowe wzory

$$\int 0dx = C \qquad \int 1dx = x + C$$

$$\int x^a dx = \frac{x^{a+1}}{a+1} + C,\ a \neq -1$$

$$\int \frac{1}{x} dx = \ln|x| + C$$

$$\int a^x dx = \frac{a^x}{\ln a} + C,\ \ a > 0,\ \ a \neq 1$$

Przykład

$$f(x)=\begin{cases}12(x^2-x^3) & \text{dla x} \in (0,1)\\ 0 & \text{dla x} \notin (0,1)\end{cases}$$

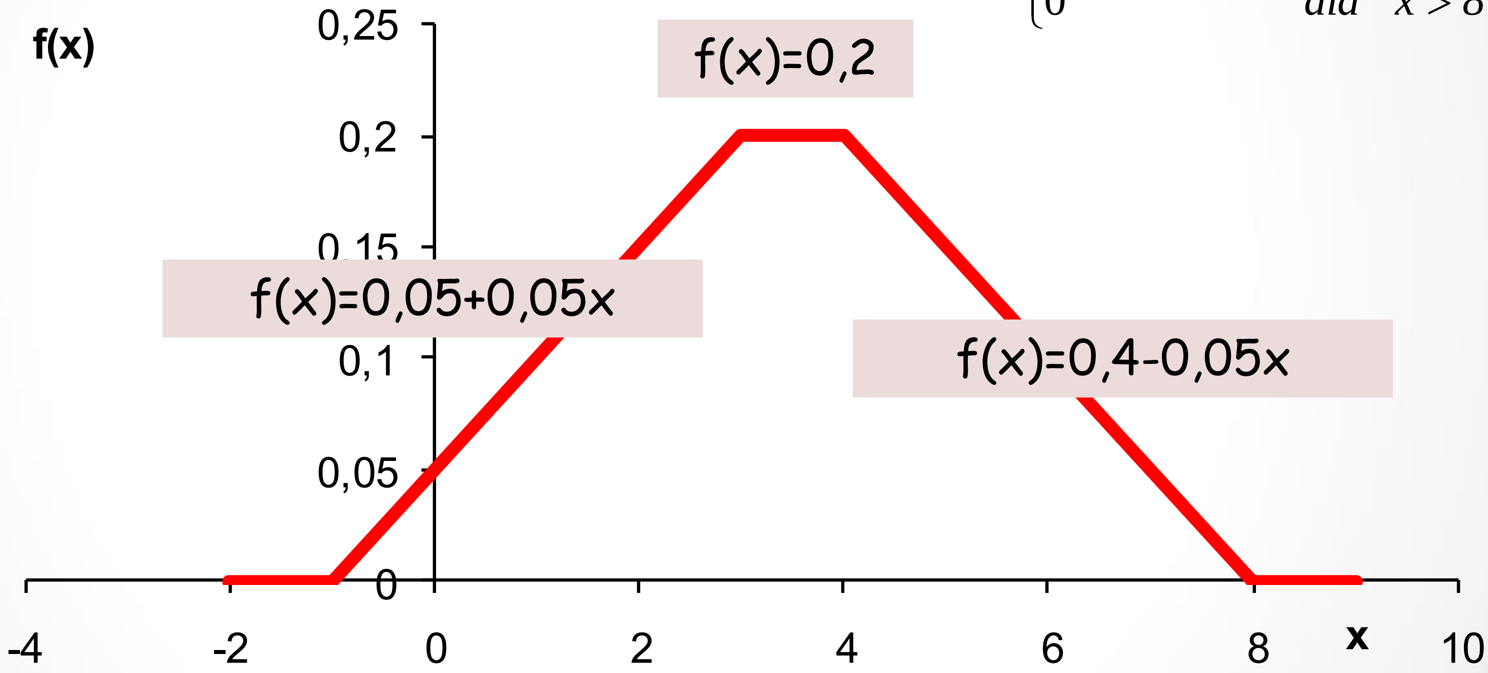
$$\int f(x)dx=\begin{cases}\int 0=0 & dla\ x<0\\ \int (12x^2-12x^3)dx & dla\ 0\le x\le 1\\ \int 0=0 & dla\ x>1\end{cases}$$

$$\int f(x)dx=\int_0^1 (12x^2-12x^3)dx=\left|12\frac{x^3}{3}\right|_0^1-\left|12\frac{x^4}{4}\right|_0^1=$$

$$=\left|4x^3\right|_0^1-\left|3x^4\right|_0^1=(4-0)-(3-0)=1$$

Zm. ciągła - przykład

$$f(x) = \begin{cases} 0 & dla \;\; x < -1 \\ 0{,}05 + 0{,}05x & dla \;\; -1 \le x \le 3 \\ 0{,}2 & dla \;\; 3 < x \le 4 \\ 0{,}4 - 0{,}05x & dla \;\; 4 < x \le 8 \\ 0 & dla \;\; x > 8 \end{cases}$$



Zm. ciągła - przykład

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-1}^{3}(0{,}05+0{,}05x)dx + \int_{3}^{4}(0{,}2)dx + \int_{4}^{8}(0{,}4-0{,}05x)dx =$$

$$[0{,}05x+0{,}05\frac{x^2}{2}\Big|_{-1}^{3} + \Big|0{,}2x\Big|_{3}^{4} + \Big|0{,}4x-0{,}05\frac{x^2}{2}\Big|_{4}^{8} =$$

$$=\{(0{,}05\cdot 3+0{,}025\cdot 3^2)-(0{,}05\cdot(-1)+0{,}025\cdot(-1)^2)\}+$$

$$+\{0{,}2\cdot 4-0{,}2\cdot 3\}+$$

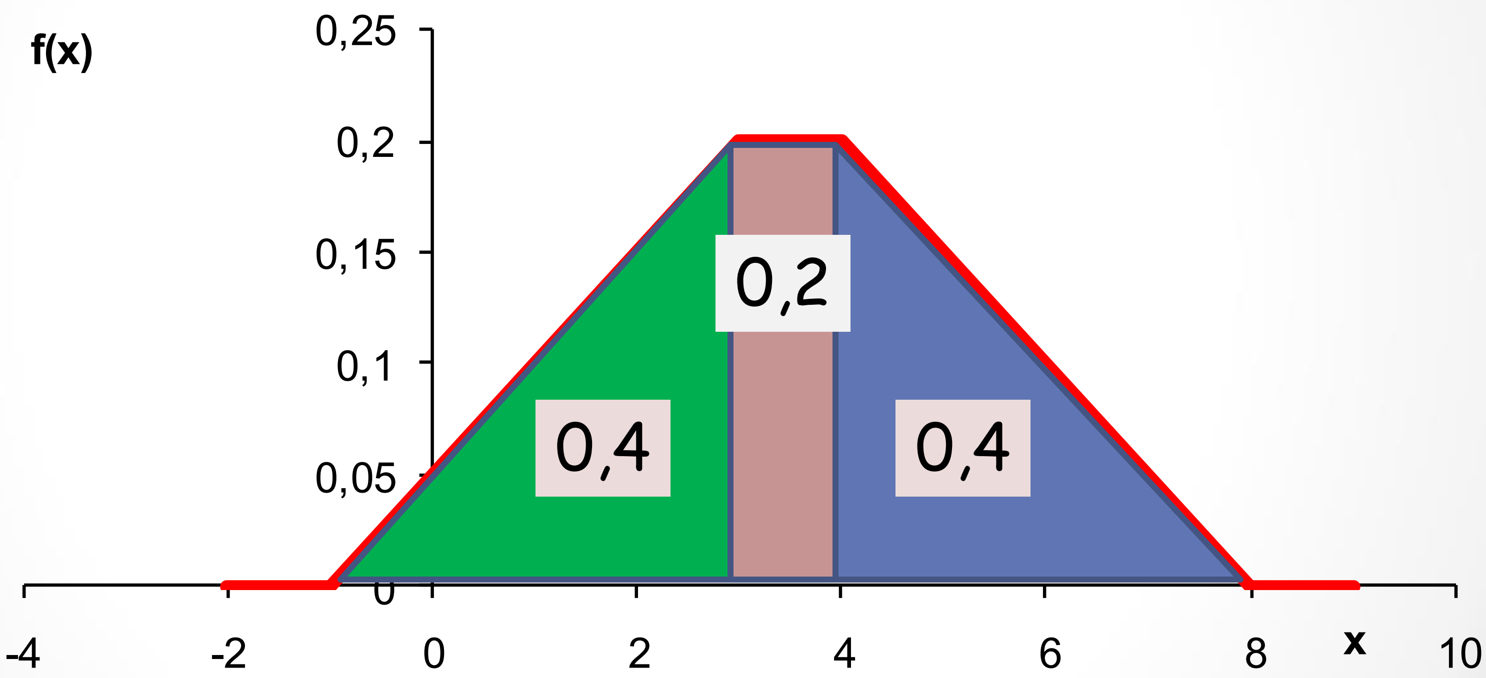
$$+\{(0{,}4\cdot 8-0{,}025\cdot 8^2)-(0{,}4\cdot 4-0{,}025\cdot 4^2)\}=$$

$$=\{(0{,}15+0{,}225)-(-0{,}05+0{,}025)\}+\{0{,}8-0{,}6\}+$$

$$+\{(3{,}2-1{,}6)-(1{,}6-0{,}4)\}=$$

$$=\{0{,}375+0{,}025\}+\{0{,}2\}+\{1{,}6-1{,}2\}=0{,}4+0{,}2+0{,}4=1$$

Zm. ciągła - przykład



Dystrybuanta zmiennej losowej - F(x)

Możemy utworzyć funkcję taka, która określi prawdopodobieństwo, że zmienna losowa nie przekroczy wartości tej funkcji:

$$F(x_0) = F(X = x_0) = P(X \leq x_0)$$

funkcję tę nazywamy dystrybuantą zmiennej losowej

$$F(x_o) = P(X \leq x_0) = \sum_{x_i \leq x_0} P(x_i) = \sum_{x_i \leq x_0} p_i$$

dla zmiennej losowej skokowej

$$F(x_o) = \int_{-\infty}^{x_0} f(x)dx$$

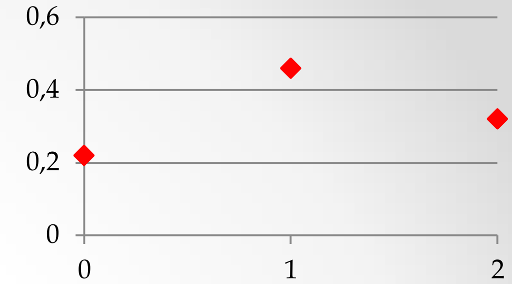
dla zmiennej losowej ciągłej

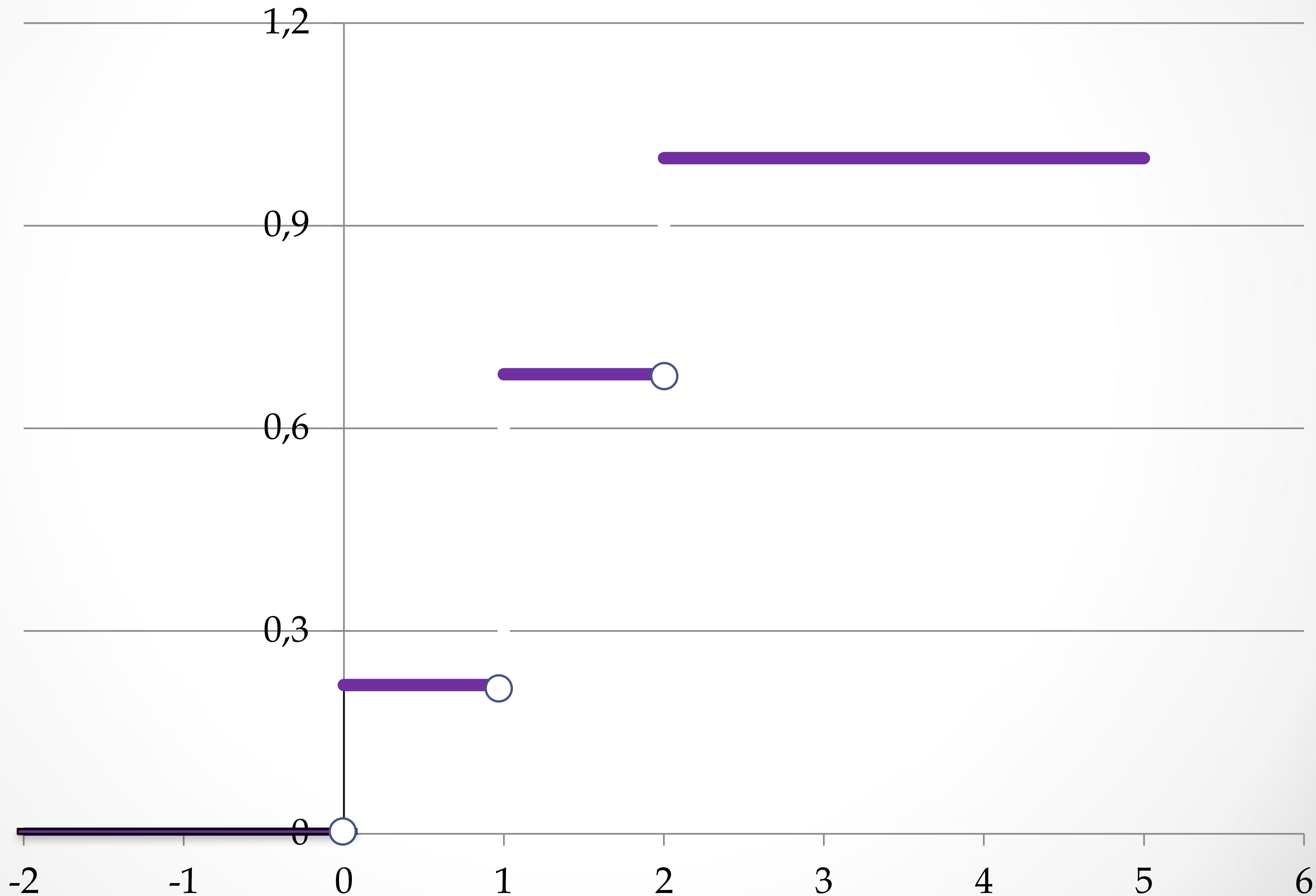
Dystrybuanta F(x)

x_i	p_i	$F(x_i)$
0	0,22	0,22
1	0,46	0,68
2	0,32	1
razem	1	

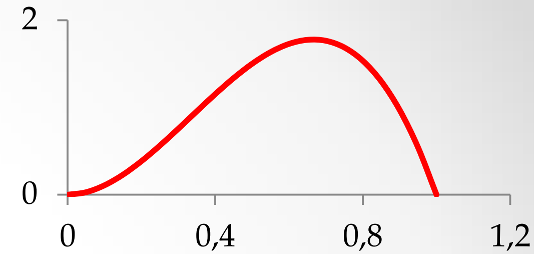
kumulacja prawdopodobieństwa

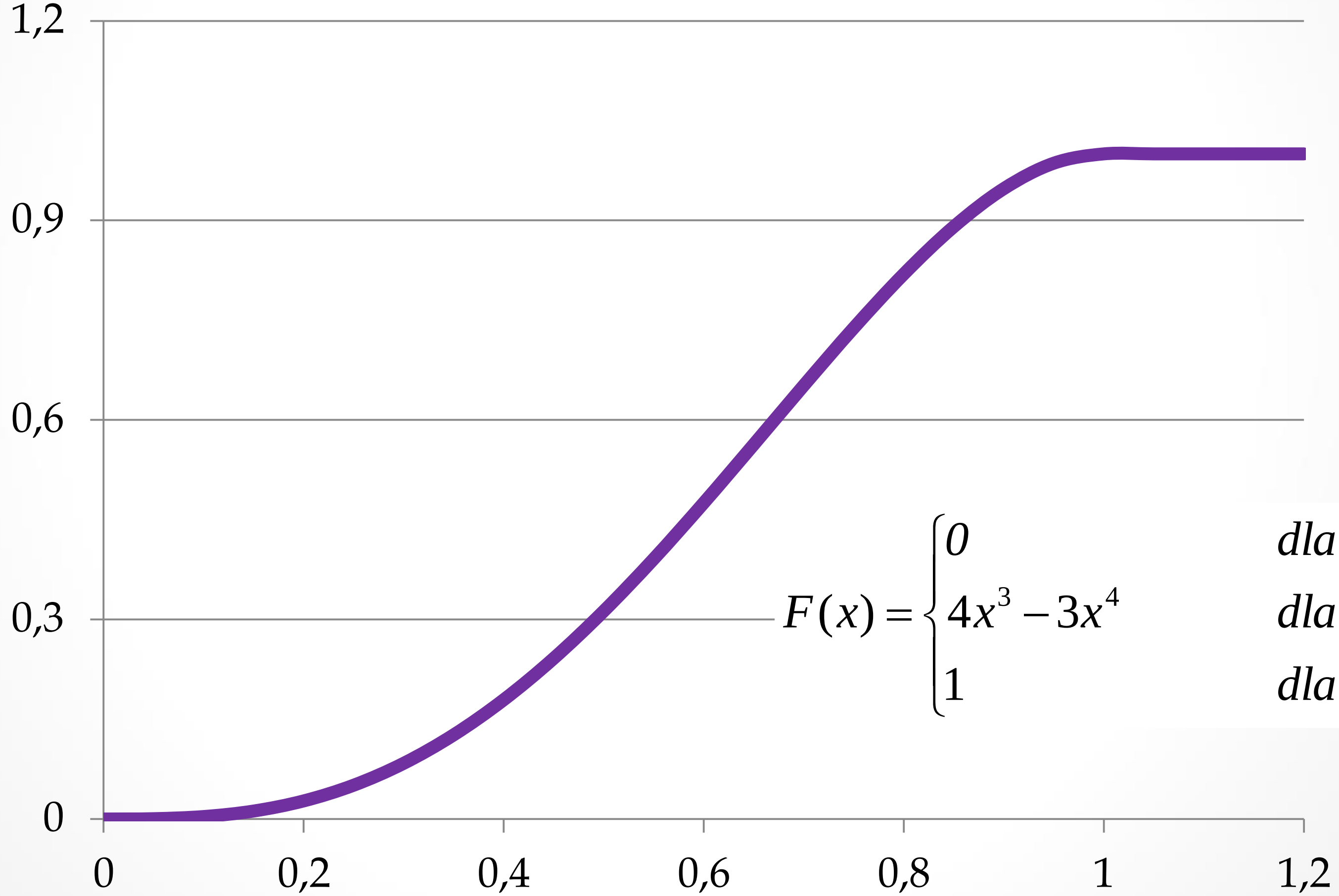
Dystrybuanta





Dystrybuanta





$$F(x) = \begin{cases} 0 & dla\ x < 0 \\ 4x^3 - 3x^4 & dla\ 0 \le x \le 1 \\ 1 & dla\ x > 1 \end{cases}$$

Własności dystrybuanty

- Ograniczona (wynika z własności prawdopodobieństwa)
- Co najmniej prawostronnie ciągła
- Niemalejąca
- Określona dla liczb rzeczywistych

Przykład

Dla jakiej wartości parametru A dana funkcja może być funkcją gęstości zmiennej losowej X.

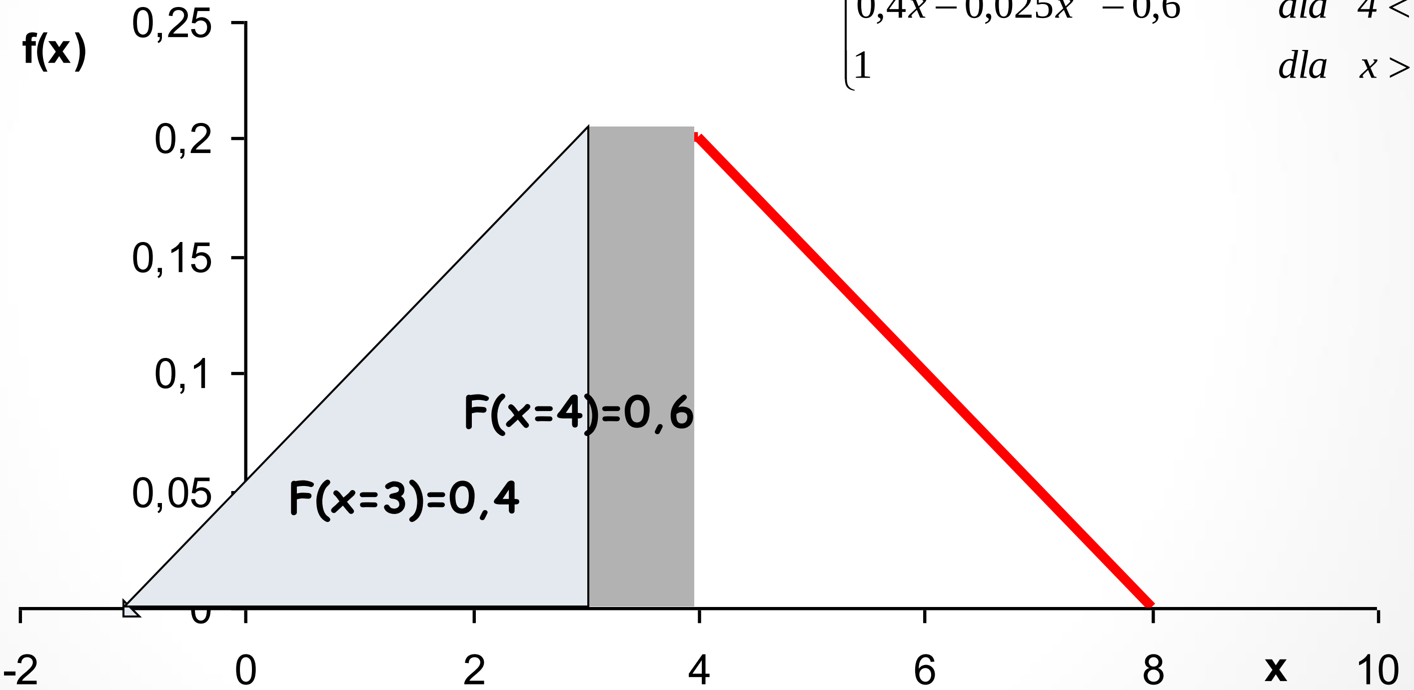
$$f(x) = \begin{cases} \frac{1}{\sqrt{x}} + A & \text{dla } 1 \le x \le 4 \\ 0 & \text{dla pozostalych} \end{cases}$$

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{1}^{4}\left(x^{-\frac{1}{2}} + A \right)dx = \left(2\sqrt{x} + Ax\right)\Big|_{1}^{4} = 2\sqrt{4} + 4A - 2\sqrt{1} - A =$$

$$= 3A + 2 = 1 \Rightarrow A = -\frac{1}{3}$$

Zastosowanie dystrybuanty

$$F(x)=\begin{cases}0 & dla\ \ x<-1\\ 0{,}05x+0{,}025x^2+0{,}025 & dla\ \ -1\le x\le 3\\ 0{,}2x-0{,}2 & dla\ \ 3<x\le 4\\ 0{,}4x-0{,}025x^2-0{,}6 & dla\ \ 4<x\le 8\\ 1 & dla\ \ x>8\end{cases}$$



$$P(3<X<4)=F(X=4)-F(X=3)=0{,}6-0{,}4=0{,}2$$

Przykład

x_i	$F(x_i)$
2	0,35
3	0,66
4	1

Które zdanie jest prawdą

$$P(3 < X < 4) = F(X = 4) - F(X = 3) = 1 - 0,66 = 0,34$$

$$P(3 \leq X < 4) = F(X = 4) - F(X = 3) = 1 - 0,66 = 0,34$$

$$P(3 < X \leq 4) = F(X = 4) - F(X = 3) = 1 - 0,66 = 0,34$$

$$P(3 \leq X \leq 4) = F(X = 4) - F(X = 3) = 1 - 0,66 = 0,34$$

Parametry zmiennej losowej

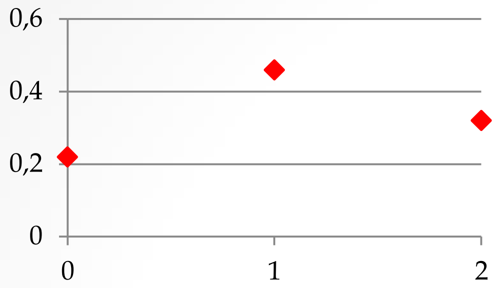
- EX – wartość oczekiwana
- D^2X – wariancja
- DX – odchylenie standardowe
- inne, np. kwantyle, mediana, wartość modalna itd.

Wartość oczekiwana

$$EX = \begin{cases} \sum_i x_i \cdot p_i & dla\ zm.losowej\ skokowej \\ \int_{-\infty}^{+\infty} x \cdot f(x)dx & dla\ zm.losowej\ ciaglej \end{cases}$$

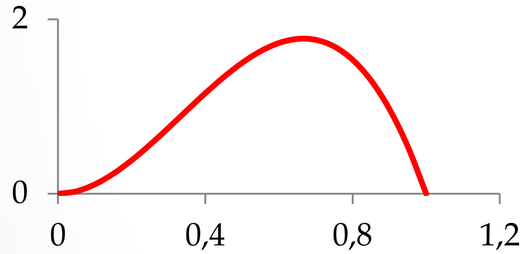
Wartość oczekiwana EX wyznacza położenie najbardziej prawdopodobnej wartości zmiennej losowej

Wartość oczekiwana



$$EX = \sum_i x_i \cdot p_i = 0 \cdot 0{,}22 + 1 \cdot 0{,}46 + 2 \cdot 0{,}32 =$$

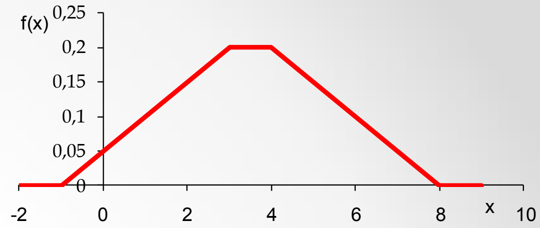
$$= 0 + 0{,}46 + 0{,}64 = 1{,}1$$



$$EX = \int x \cdot f(x)dx = \int_0^1 (x \cdot 12x^2 - x \cdot 12x^3)dx = \left| 12\frac{x^4}{4} \right|_0^1 - \left| 12\frac{x^5}{5} \right|_0^1 =$$

$$= \left|3x^4\right|_0^1 - \left|2{,}4x^5\right|_0^1 = (3-0) - (2{,}4-0) = 0{,}6$$

Wartość oczekiwana



$$EX=\int_{-\infty}^{+\infty} x\cdot f(x)dx=\int_{-1}^{3}(0{,}05x+0{,}05x^2)dx+\int_{3}^{4}(0{,}2x)dx+\int_{4}^{8}(0{,}4x-0{,}05x^2)dx=$$

$$[0{,}05\frac{x^2}{2}+0{,}05\frac{x^3}{3}\Bigg|_{-1}^{3}+\Bigg|0{,}2\frac{x^2}{2}\Bigg|_{3}^{4}+\Bigg|0{,}4\frac{x^2}{2}-0{,}05\frac{x^3}{3}\Bigg|_{4}^{8}=$$

$$=\{(0{,}025\cdot 3^2+0{,}05\cdot {}^{3^3}\!/_{3})-(0{,}025\cdot(-1)^2+0{,}05\cdot {}^{(-1)^3}\!/_{3})\}+$$

$$+\{0{,}1\cdot 4^2-0{,}1\cdot 3^2\}+$$

$$+\{(0{,}2\cdot 8^2-0{,}05\cdot {}^{8^3}\!/_{3})-(0{,}2\cdot 4^2-0{,}05\cdot {}^{4^3}\!/_{3})\}=$$

$$=\{(0{,}225+0{,}45)-(0{,}025-{}^{0{,}05}\!/_{3})\}+\{1{,}6-0{,}9\}+$$

$$+\{(12{,}8-{}^{25{,}6}\!/_{3})-(3{,}2-{}^{3{,}2}\!/_{3})\}=$$

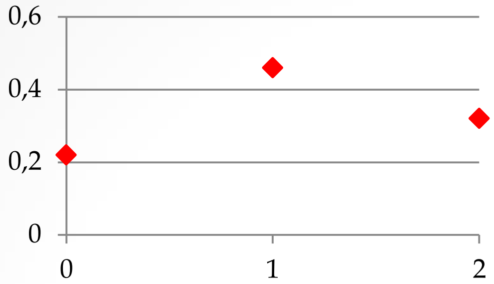
$$=\{0{,}65+{}^{0{,}05}\!/_{3}\}+\{0{,}7\}+\{9{,}6-{}^{22{,}4}\!/_{3}\}=10{,}95-{}^{22{,}35}\!/_{3}=10{,}95-7{,}45=3{,}5$$

Wariancja

$$D^2X = \begin{cases} \sum_i (x_i - EX)^2 \cdot p_i = \sum_i x_i^2 p_i - EX^2 & dla\ skokowej \\ \int_{-\infty}^{+\infty} (x - EX)^2 \cdot f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx - EX^2 & dla\ ciaglej \end{cases}$$

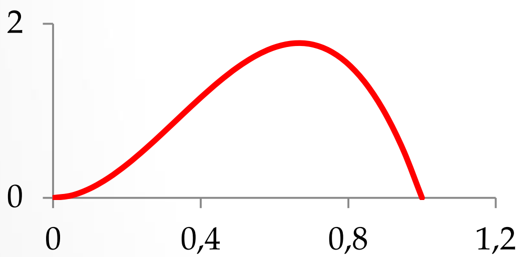
Wariancja jest miarą rozrzutu wartości zmiennej losowej wokół wartości oczekiwanej

Wariancja



$$D^2X = \sum_i x_i^2 \cdot p_i - EX^2 = 0^2 \cdot 0{,}22 + 1^2 \cdot 0{,}46 + 2^2 \cdot 0{,}32 - 1{,}1^2 =$$

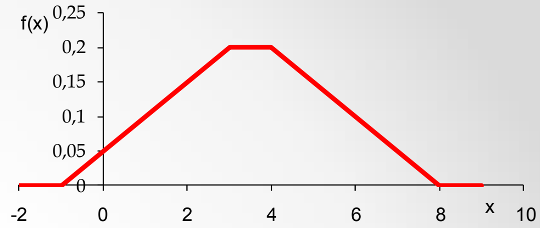
$$= 0 + 0{,}46 + 1{,}28 - 1{,}21 = 0{,}53$$



$$EX = \int x^2 \cdot f(x)dx - EX^2 = \int_0^1 (12x^4 - 12x^5)dx - 0{,}6^2 = \left| 12\frac{x^5}{5} \right|_0^1 - \left| 12\frac{x^6}{6} \right|_0^1 - 0{,}36 =$$

$$= \left| 2{,}4x^5 \right|_0^1 - \left| 2x^6 \right|_0^1 - 0{,}36 = (2{,}4 - 0) - (2 - 0) - 0{,}36 = 0{,}04$$

Wariancja



$$D^2X=\int_{-\infty}^{+\infty} x^2 f(x)dx - EX^2 = \int_{-1}^{3}(0{,}05x^2+0{,}05x^3)dx+\int_{3}^{4}(0{,}2x^2)dx+\int_{4}^{8}(0{,}4x^2-0{,}05x^3)dx-3{,}5^2=$$

$$=[0{,}05\frac{x^3}{3}+0{,}05\frac{x^4}{4}\Big|_{-1}^{3}+\Big|0{,}2\frac{x^3}{3}\Big|_{3}^{4}+\Big|0{,}4\frac{x^3}{3}-0{,}05\frac{x^4}{4}\Big|_{4}^{8}-3{,}5^2=$$

$$=\{(0{,}05\cdot 3^3/3+0{,}05\cdot 3^4/4)-(0{,}05\cdot(-1)^3/3+0{,}05\cdot(-1)^4/4)\}+$$

$$+\{0{,}2\cdot 4^3/3-0{,}2\cdot 3^3/3\}+$$

$$+\{(0{,}4\cdot 8^3/3-0{,}05\cdot 8^4/4)-(0{,}4\cdot 4^3/3-0{,}05\cdot 4^4/4)\}-3{,}5^2=$$

$$=\{(0{,}45+0{,}0125\cdot 81)-(-0{,}05/3+0{,}0125)\}+\{12{,}8/3-1{,}8\}+$$

$$+\{(204{,}8/3-51{,}2)-(25{,}6/3-3{,}2)\}-3{,}5^2=$$

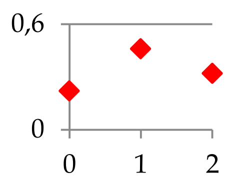
$$=\{1{,}45+0{,}05/3\}+\{12.8/3-1{,}8\}+\{179{,}2/3-48\}-12{,}25=192{,}05/3-60{,}6=64{,}0167-60{,}6=3{,}4167$$

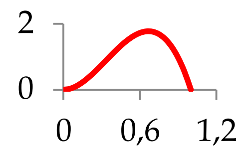
Odchylenie standardowe

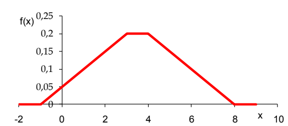
$$DX = \sqrt{D^2 X}$$

Wskaźnik zmienności

$$VX = \frac{DX}{EX}$$







EX	1,1	0,6	3,5
D^2X	0,53	0,04	3,4167
DX	0,7280	0,2	1,8484
VX	0,6618	0,3333	0,5281

Kwantyle

Kwantylem rzędu q zmiennej losowej X jest taka liczba x_q, że

$$P(X \le x_q) = q \text{ czyli } F(x_q) = q$$

Kwantyl rzędu $^1/_2$ nazywamy medianą (Me), lub inaczej wartością środkową (ED_{50} - *dosis effectiva* - mediana dawek efektywnych, LD_{50} - *dosis lethal* - mediana dawek śmiertelnych - w medycynie czy przy truciznach).

- Kwantyle $^k/_4$ dla k=1,2,3 to kwartyle
- Kwantyle $^k/_{10}$ dla k=1,2...9 to decyle
- Kwantyle $^k/_{100}$ dla k=1,2...99 to centyle

Wartość środkowa - mediana

Mediana (Me) to kwantyl rzędu 0,5

$$P(X \leq x_{0,5}) = 0{,}5$$

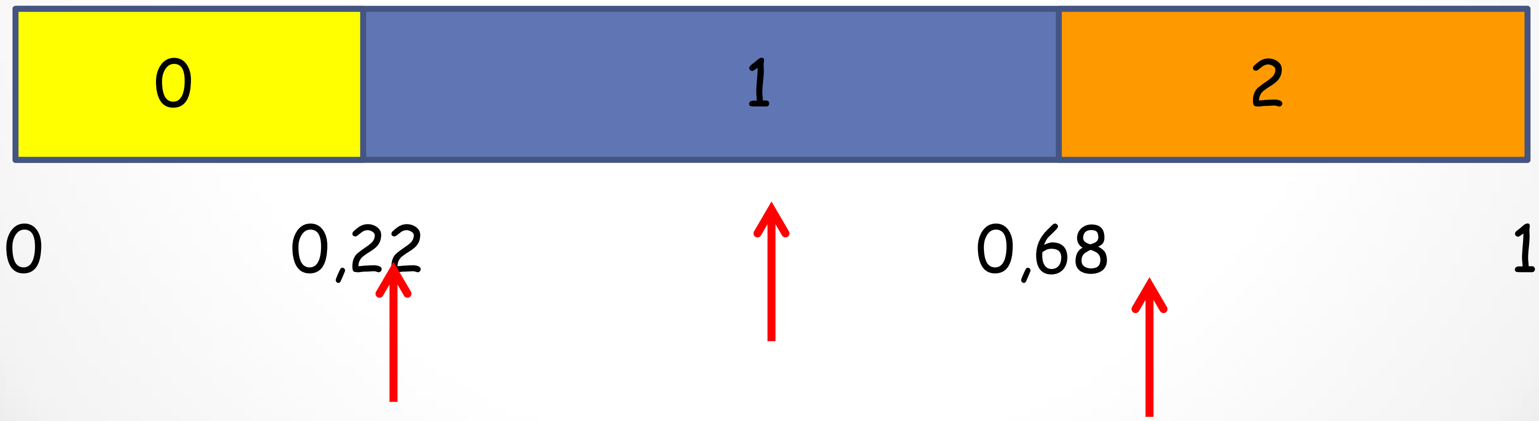
czyli

$$F(x_{0,5}) = 0{,}5$$

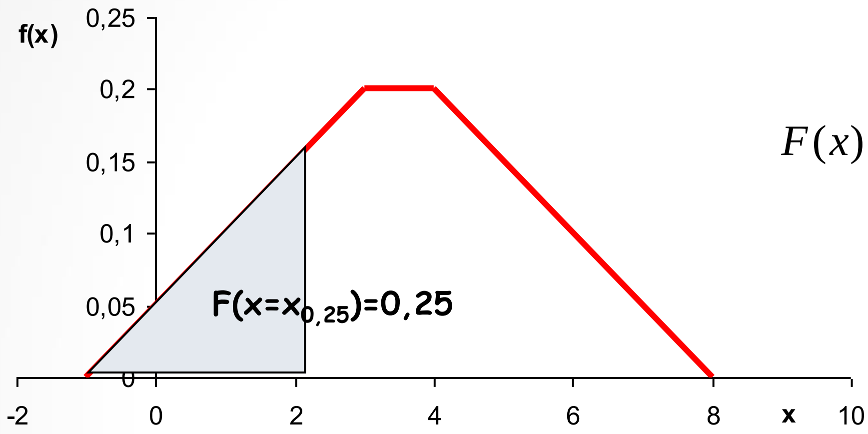
Jeżeli EX = Me to rozkład zmiennej losowej jest symetryczny

Kwantyle

x_i	p_i
0	0,22
1	0,46
2	0,32
suma	1



Kwantyle



$$F(x)=\begin{cases}0 & \textit{dla}\ \ x<-1\\ 0{,}05x+0{,}025x^2+0{,}025 & \textit{dla}\ \ -1\le x\le 3\\ 0{,}2x-0{,}2 & \textit{dla}\ \ 3<x\le 4\\ 0{,}4x-0{,}025x^2-0{,}6 & \textit{dla}\ \ 4<x\le 8\\ 1 & \textit{dla}\ \ x>8\end{cases}$$

$$0{,}05x_{0,25}+0{,}025x_{0,25}^2+0{,}025=0{,}25$$

$$0{,}025x_{0,25}^2+0{,}05x_{0,25}-0{,}225=0$$

$$\Delta=0{,}05^2-4\cdot 0{,}025\cdot(-0{,}225)=0{,}0025+0{,}0225=0{,}025$$

$$x_{0,25}=(-0{,}05+\sqrt{0{,}025})/(2\cdot 0{,}025)=2{,}1623$$

Wartość modalna

Wartość modalna (Mo) to taka wartość zmiennej losowej X dla której funkcja gęstości przyjmuje maksimum lokalne.

Dla zmiennej losowej skokowej wartość o najwyższym prawdopodobieństwie.

Rozkład może być jedno i dwumodalny

Współczynnik asymetrii

Współczynnik asymetrii (skośności) rozkładu zmiennej losowej X :

$$\gamma_1 = \begin{cases} \sum_i (x_i - EX)^3 \cdot p_i / D^3 X & \text{dla skokowej} \\ \int_{-\infty}^{+\infty} (x - EX)^3 \cdot f(x) dx / D^3 X & \text{dla ciaglej} \end{cases}$$

Dla rozkładów jednomodalnych:

$$\gamma_1 = \frac{EX - Mo}{DX}$$

Gdy $\gamma_1 = 0$ to rozkład symetryczny

Przykład

W klatce znajdują się cztery białe myszy i dwie szare. Myszy przechodzą tunelem do innej klatki, przy czym zakładamy, że wchodzą do tunelu niezależnie. Wartością zmiennej losowej jest numer pierwszej szarej myszy przechodzącej tunelem. Wyznaczyć rozkład zmiennej losowej.
Obliczyć EX, DX zmiennej losowej

x_i	p_i
1	5/15
2	4/15
3	3/15
4	2/15
5	1/15

$$EX = 1\cdot\frac{5}{15} + 2\cdot\frac{4}{15} + 3\cdot\frac{3}{15} + 4\cdot\frac{2}{15} + 5\cdot\frac{1}{15} =$$

$$= \frac{5+8+9+8+5}{15} = \frac{35}{15} = \frac{7}{3}$$

$$D^2X = 1^2\cdot\frac{5}{15} + 2^2\cdot\frac{4}{15} + 3^2\cdot\frac{3}{15} + 4^2\cdot\frac{2}{15} + 5^2\cdot\frac{1}{15} - \left(\frac{7}{3}\right)^2 =$$

$$= \frac{5+16+27+32+25}{15} - \frac{49}{9} = \frac{105}{15} - \frac{49}{9} = \frac{63-49}{9} = \frac{14}{9}$$

Me = 2 Mo =1

Przykład

Dla jakiej wartości parametru A dana funkcja może być funkcją gęstości zmiennej losowej X. Obliczyć EX, DX

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{x}}+A & \text{dla } 1\le x\le 4\\ 0 & \text{dla pozostalych}\end{cases}$$

$$A=-\frac{1}{3}$$

$$EX=\int_1^4 xf(x)dx=\int_1^4\left(\frac{x}{\sqrt{x}}-\frac{1}{3}x\right)dx=\int_1^4\left(x^{\frac{1}{2}}-\frac{1}{3}x\right)dx=\left(\frac{2}{3}x^{\frac{3}{2}}-\frac{1}{6}x^2\right)\Bigg|_1^4=\frac{16}{3}-\frac{16}{6}-\frac{2}{3}+\frac{1}{6}=\frac{13}{6}$$

$$D^2X=\int_1^4 x^2 f(x)dx-E^2X=\int_1^4\left(\frac{x^2}{\sqrt{x}}-\frac{1}{3}x^2\right)dx-\frac{169}{36}=$$

$$\int_1^4\left(x^{\frac{3}{2}}-\frac{1}{3}x^2\right)dx-\frac{169}{36}=\left(\frac{2}{5}x^{\frac{5}{2}}-\frac{1}{9}x^3\right)\Bigg|_1^4-\frac{169}{36}=\frac{64}{5}-\frac{64}{9}-\frac{2}{5}+\frac{1}{9}-\frac{169}{36}=\frac{62}{5}-\frac{421}{36}=\frac{127}{180}$$


www.wildanimalsonline.com
Ovis musimon