ZMIENNA LOSOWA JEDNOWYMIAROWA

POJĘCIE ZMIENNEJ LOSOWEJ

Podstawowe pojęcia rachunku prawdopodobieństwa:

- zdarzenie losowe,
- zdarzenie elementarne,
- prawdopodobieństwo,
- zbiór zdarzeń elementarnych.

Def. Niech E będzie zbiorem zdarzeń elementarnych danego doświadczenia. Funkcję $X(e)$ przyporządkowującą każdemu zdarzeniu elementarnemu $e \in E$ jedną i tylko jedną liczbę $X(e)=x$ nazywamy *zmienną losową*.

Przykład

Rozpatrujemy doświadczenie polegające na rzucie symetryczną monetą. Wynikiem tego doświadczenia mogą być zdarzenia "pojawienie się orła" albo "pojawienie się reszki" tworzące zbiór zdarzeń elementarnych.

Na zbiorze zdarzeń elementarnych określamy zmienną losową X w sposób następujący:

X (orzeł) = 1; X (reszka) = 0

Zmienna losowa X przyjmuje wartość ze zbioru $\{0,1\}$. Ponieważ zdarzenia "pojawienie się orła" i "pojawienie się reszki" realizują się z prawdopodobieństwami równymi 1/2, można zapisać:

$P(X=1) = P\{\text{orzeł}\} = 1/2$,

$P(X=0) = P\{\text{reszka}\} = 1/2$.

TYPY ZMIENNYCH LOSOWYCH

Def. Zmienna losowa X jest *typu skokowego*, jeśli może przyjmować skończoną lub nieskończoną, ale przeliczalną liczbę wartości.

Wartości zmiennej losowej skokowej (określane często jako *punkty skokowe*) będziemy oznaczać przez $x_1, x_2, \ldots,$ natomiast prawdopodobieństwa, z jakimi są one realizowane (określane jako *skoki*), oznaczamy przez $p_1, p_2, \ldots$

Def. Zmienna losowa X jest *typu ciągłego*, jeśli jej możliwe wartości tworzą przedział ze zbioru liczb rzeczywistych.

Dla zmiennej losowej *typu ciągłego* możliwe jest określenie prawdopodobieństwa, że przyjmuje ona wartość należącą do dowolnego zbioru jej wartości. Sposób rozdysponowania całej "masy" prawdopodobieństwa (równej 1) pomiędzy wartości, jakie przyjmuje dana zmienna losowa, określamy mianem jej *rozkładu prawdopodobieństwa*.

ROZKŁAD ZMIENNEJ LOSOWEJ SKOKOWEJ

Założenia:

- zmienna losowa X typu skokowego przyjmuje wartość $x_1, x_2,...$ z prawdopodobieństwami, odpowiednio $p_1, p_2,...$,

- prawdopodobieństwa $p_1, p_2,...$ spełniają równość:

$$\sum_{i=1}^{n} p_i = 1, \qquad (1)$$

gdy zmienna losowa X przyjmuje skończoną liczbę n wielkości,

- prawdopodobieństwo $p_1, p_2,...$ spełniają równość:

$$\sum_{i=1}^{\infty} p_i = 1, \qquad (2)$$

gdy zmienna losowa X przyjmuje nieskończoną liczbę wartości.

Def. Zbiór prawdopodobieństw postaci:

$$P(X = x_i) = p_i \qquad (i = 1,2,...)$$

spełniających równość (1) lub (2) określamy mianem *funkcji prawdopodobieństwa* zmiennej losowej X typu skokowego.

Funkcję prawdopodobieństwa można przestawić tabelarycznie w poniższy sposób (przy założeniu, że zbiór wartości zmiennej losowej jest skończony):

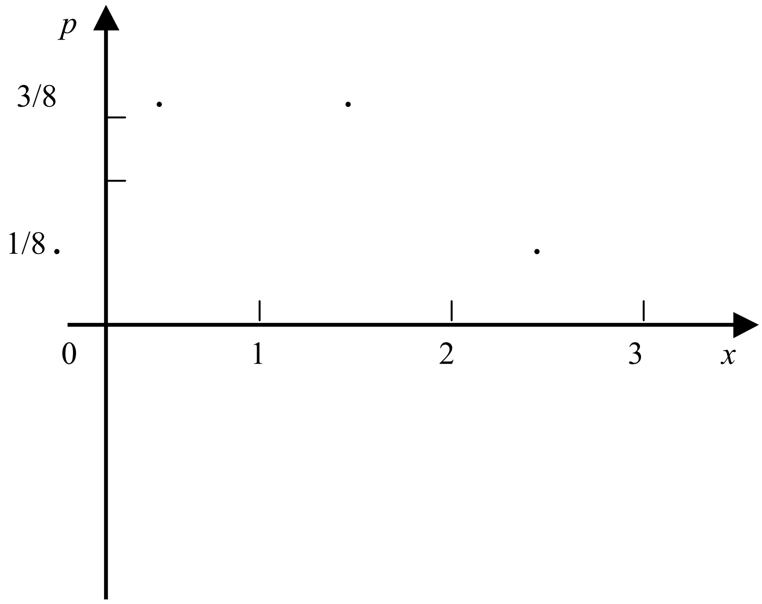
x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

Przykład

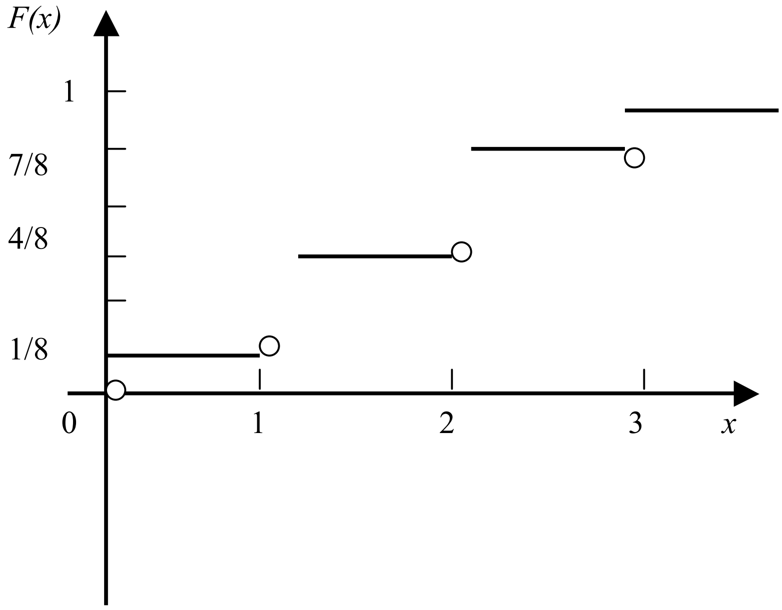
Funkcję prawdopodobieństwa zmiennej losowej przedstawia poniższa tabela:

x_i	*0*	*1*	*2*	*3*
p_i	*1/8*	*3/8*	*3/8*	*1/8*

Wykres funkcji prawdopodobieństwa x



Dystrybuanta zmiennej losowej x



Def. *Dystrybuantą* zmiennej losowej X nazywamy funkcję $F(x)$ określoną na zbiorze liczb rzeczywistych jako:

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

- Dla zmiennej losowej X skokowej, która przyjmuje wartości $x_1, x_2, \ldots$ z prawdopodobieństwami $p_1, p_2, \ldots$, dystrybuanta ma postać:

$$F(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_i \qquad (-\infty < x < \infty)$$

- Dystrybuantą *F(x)* zmiennej losowej *X* skokowej można zapisać też następująco (zakładamy, że zbiór wartości zmiennej losowej jest skończony oraz że został on uporządkowany według wzrastających wartości):

$$F(x)\begin{cases} 0 & dla \quad x < x_1 \\ p_1 & dla \quad x_1 \le x < x_2 \\ p_1 + p_2 & dla \\ . & . \\ . & . \\ . & . \\ p_1 + p_2 + \ldots + p_{n-1} & dla \quad x_{n-1} \le x < x_n \\ 1 & dla \quad x \ge x_n \end{cases}$$

Podstawowe własności dystrybuanty zmiennej losowej skokowej:

- $0 \le F(x) \le 1$,
- $\lim\limits_{x\to-\infty} F(x) = 0$ oraz $\lim\limits_{x\to+\infty} F(x) = 1$,
- $F(x)$ jest funkcją niemalejącą (dla $x_1<x_2$ zachodzi $F(x_1) \le F(x_2)$) i przedziałami stałą,
- $F(x)$ jest funkcją prawostronnie ciągłą.

Przykład

Dystrybuanta zmiennej losowej ma postać:

$$F(x) = \begin{cases} 0 & dla \quad x < 0, \\ 1/8 & dla \quad 0 \le x < 1, \\ 4/8 & dla \quad 1 \le x < 2, \\ 7/8 & dla \quad 2 \le x < 3, \\ 1 & dla \quad x \ge 3, \end{cases}$$

1 ROZKŁAD ZMIENNEJ LOSOWEJ CIĄGŁEJ

Def. *Funkcją gęstości* prawdopodobieństwa zmiennej losowej typu ciągłego nazywamy funkcję *f(x)*, określoną na zbiorze liczb rzeczywistych o następujących własnościach:

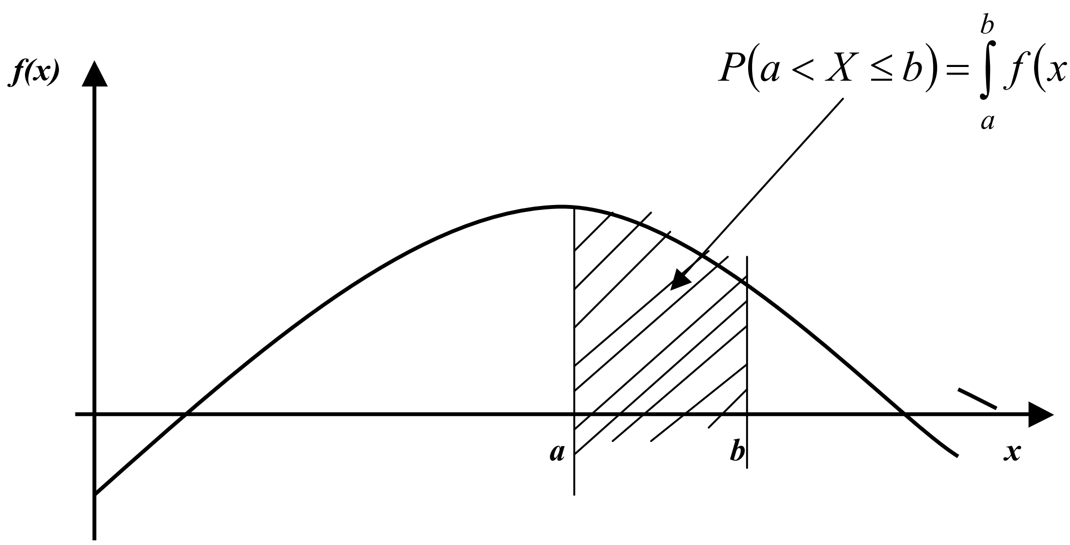
- $f(x) \geq 0,$
- $\int_a^b f(x)dx = P(a < X \leq b)$ dla dowolnych $a<b$.
- $\int_{-\infty}^{+\infty} f(x)dx = P(-\infty < X \leq +\infty) = 1$

Def. *Funkcją gęstości* zmiennej losowej *X* typu ciągłego nazywamy funkcję *f(x)* określoną następująco:

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

$$P(x < X \leq x + \Delta x) \cong f(x)\Delta x$$

Przykładowy wykres funkcji gęstości prawdopodobieństwa i graficzna interpretacja $P(a < X \leq b)$



Przykład

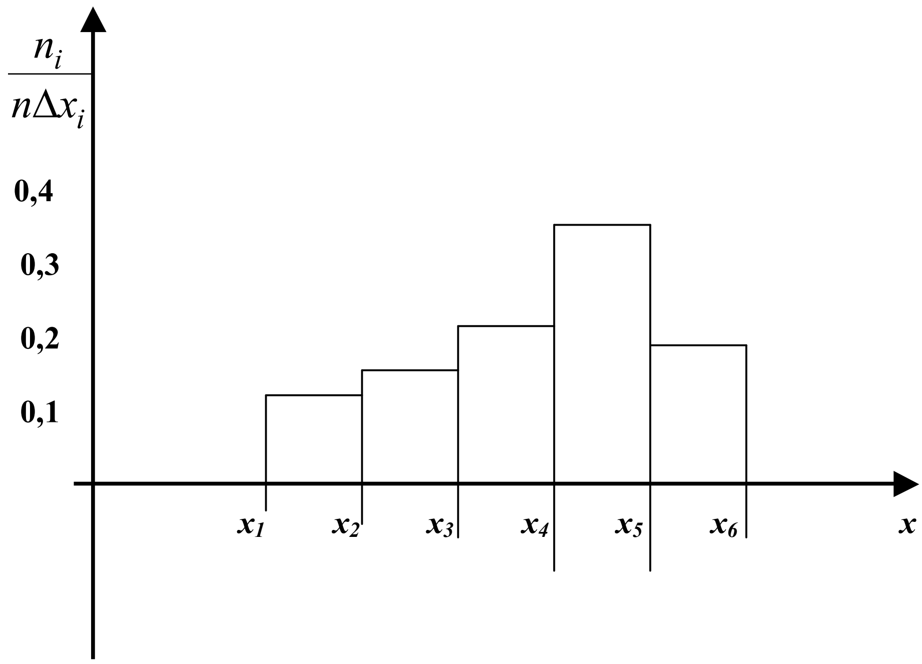
- przeprowadzamy pomiar wagi pewnego typu odkuwek tłoczonych przez prasę hydrauliczną,
- waga pojedynczych odkuwek odchyla się w sposób przypadkowy od wagi nominalnej, tym samym wyniki pomiarów wagi odkuwek można traktować jako *realizacje zmiennej losowej ciągłej*,
- dokonujemy *n* pomiarów, grupując uzyskane wyniki w *l* rozłącznych przedziałach $\left(x_i , x_{i+1} \right\langle$, gdzie $i=1,...,l$,
- długość przedziału $\left(x_i , x_{i+1} \right\langle$ oznaczamy przez $\Delta x_i = x_{i+1} - x_i$, natomiast liczbę pomiarów, które trafiły do tego przedziału, przez n_i,
- przedstawiamy uzyskane dane w postaci *histogramu*, konstruując go przy następujących założeniach:

  - podstawy poszczególnych prostokątów reprezentują wyróżnione przedziały $\left(x_i , x_{i+1} \right\langle$ wartości pomiarów wagi podkuwek,

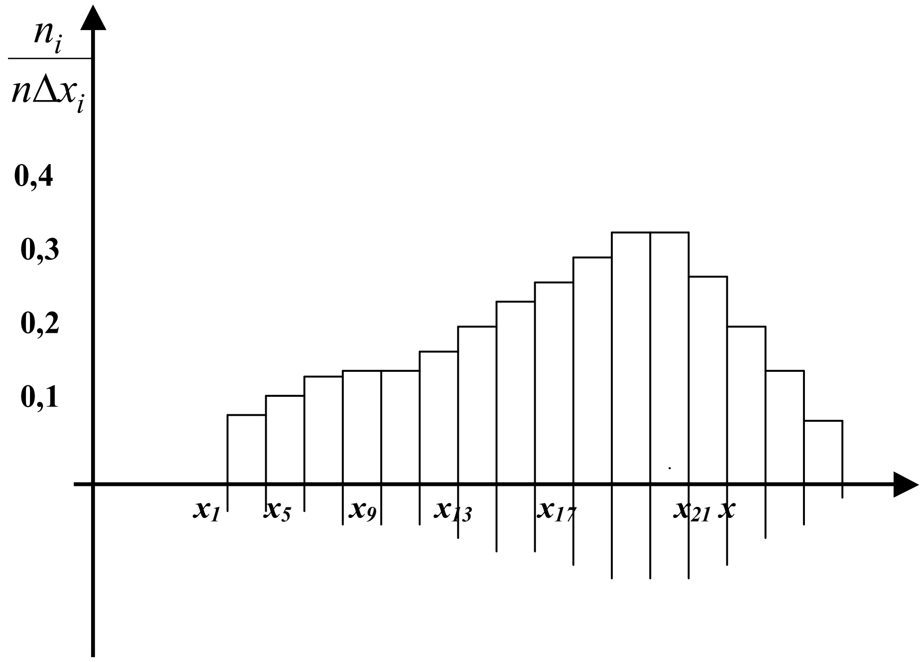
  - wysokości poszczególnych prostokątów $\frac{n_i}{n} \cdot \frac{1}{\Delta x_i}$ są ustalone w taki sposób, aby pola prostokątów były równe częstościom $\frac{n_i}{n}$:

$$\Delta x_i \cdot \frac{n_i}{n} \cdot \frac{1}{\Delta x_i} = \frac{n_i}{n}$$

Histogram wyników pomiarów wagi odkuwek $\left(l = 5, \Delta x_i = 1 \right)$



Histogram wyników pomiarów wagi odkuwek $\left(l = 20, \Delta x_i = 0{,}25\right)$



Def. Dystrybuantę zmiennej losowej *X* typu ciągłego można określić następująco:

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

gdzie *f(t)* jest funkcją gęstości zmiennej losowej *X*

Własności dystrybuanty F(x) zmiennej losowej X typu ciągłego:

- $0 \le F(x) \le 1$ dla $-\infty < x < +\infty$,
- $\lim\limits_{x \to -\infty} F(x) = 0$ oraz $\lim\limits_{x \to +\infty} F(x) = 1$,
- *F(x)* jest funkcją niemalejącą i ciągłą.

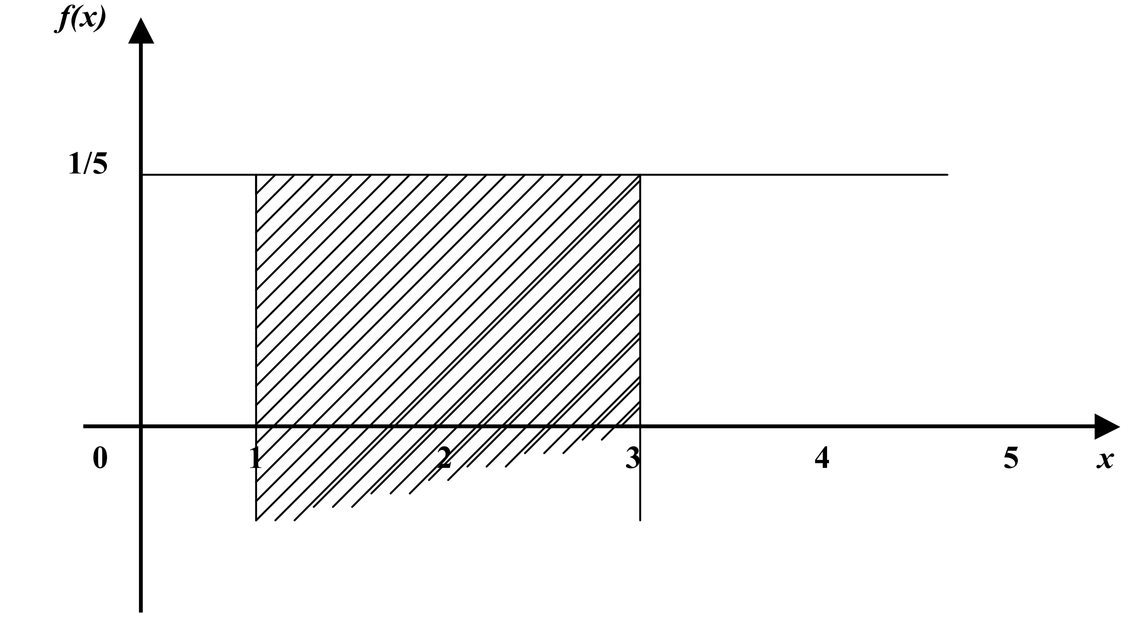
Przykład

Funkcja gęstości zmiennej losowej czasu oczekiwania na autobus ma postać:

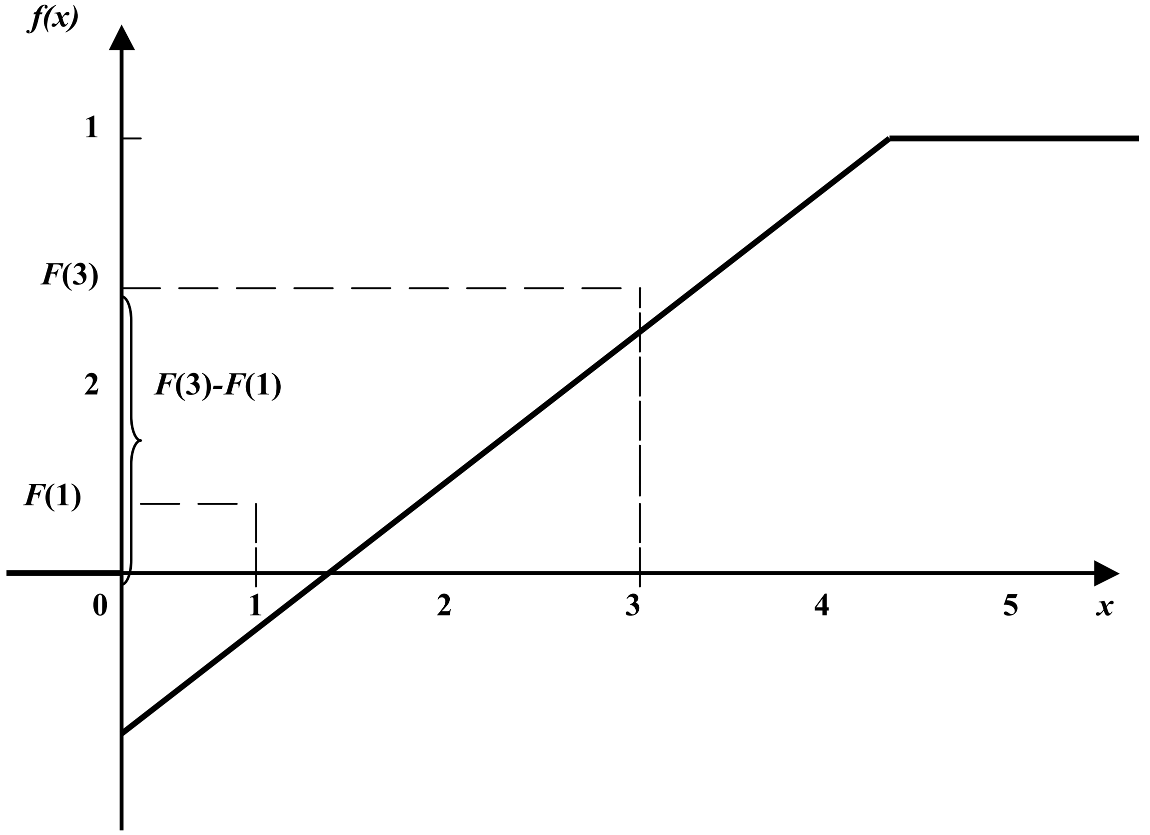
$$f(x) = \begin{cases} 0, dla x < 0, \\ c, dla 0 \le x \le 5, \\ 0, dla x > 5, \end{cases}$$

gdzie *c* jest pewną stałą.

Wykres funkcji gęstości czasu oczekiwania na autobus oraz ilustracja graficzna $P(1 < X \leq 3)$



Dystrybuanta czasu oczekiwania na autobus oraz ilustracja graficzna $P(1 < X \leq 3)$



WARTOŚĆ OCZEKIWANA I WARIANCJA ZMIENNEJ LOSOWEJ

Def. *Wartością oczekiwanej zmiennej losowej X* nazywamy wyrażenie:

$$E(X)=\begin{cases}\sum\limits_{i} x_i p_i & \text{dla zmiennej losowej skokowej}\\ \int\limits_{-\infty}^{\infty} xf(x)dx & \text{dla zmiennej losowej ciągłej}\end{cases}$$

gdzie p_i oznacza funkcję prawdopodobieństwa zmiennej losowej X przyjmującej wartości $x_i (i=1,2,...)$, natomiast $f(x)$ jest funkcją gęstości.

Def. Wariancją zmiennej losowej X nazywamy wyrażenie:

$$D^2(X)=E[X-E(X)]^2=\begin{cases}\sum\limits_{i}[x_i-E(X)]^2 p_i\\ \int\limits_{-\infty}^{\infty}[x-E(X)]^2 f(x)dx\end{cases}$$

MOMENTY

Def. *Momentem zwykłym* (lub po prostu momentem) *rzędu* $k(k=1,2,...)$ zmiennej losowej X nazywamy wartość oczekiwaną k-tej potęgi tej zmiennej, tzn.:

$$m_k=E\left(X^k\right)=\begin{cases}\sum\limits_{i} x_i^k p_i & \text{dla zmiennej losowej skokowej}\\ \int\limits_{-\infty}^{\infty} x^k f(x)dx & \text{dla zmiennej losowej ciągłej}\end{cases}$$

Def. *Momentem centralnym rzędu* $k(k=1,2,...)$ zmiennej losowej X nazywamy wartość oczekiwaną funkcji $g(X)=[X-E(X)]^k$ tej zmiennej, tzn.:

$$u_k=E[X-E(X)]^k=\begin{cases}\sum\limits_{i}[x_i-E(x)]^k_{pi} & \text{dla zmiennej losowej skokowej}\\ \int\limits_{-\infty}^{\infty}[x-E(x)]^k f(x)dx & \text{dla zmiennej losowej ciągłej}\end{cases}$$

ROZKŁAD ZERO - JEDYNKOWY

Założenia:

- przeprowadzamy doświadczenie, którego rezultatem mogą być dwa wzajemnie wykluczające się zdarzenia losowe A oraz $\overline{A}$,
- prawdopodobieństwo realizacji zdarzenia A wynosi p, przy czym $0<p<1$,
- prawdopodobieństwo zdarzenia $\overline{A}$ wynosi $q=1-p$,
- przyporządkowujemy zdarzeniu A liczbę 1 oraz zdarzeniu $\overline{A}$ liczbę 0, otrzymując zmienną losową X, której funkcja prawdopodobieństwa ma postać:

$$P(X=1)=p,$$

$$P(X=0)=1-p \quad ; \qquad (0<p<1)$$

Def. Zmienna losowa X ma *rozkład zero-jedynkowy*, jeśli przyjmuje wartość 1 z prawdopodobieństwem $0<p<1$ oraz wartość 0 z prawdopodobieństwem $q=1-p$.

Funkcja prawdopodobieństwa rozkładu zero-jedynkowego:

x_i	0	1
p_i	$1-p$	p

Dystrybuanta zmiennej losowej zero-jedynkowej:

$$F(x)=\begin{cases} 0, & dla \quad x<0 \\ 1-p, & dla \quad 0\le x<1 \\ 1, & dla \quad x\ge 1 \end{cases}$$

Wartość oczekiwana i wariancja zmiennej losowej zero-jedynkowej:

$$E(X)=0\cdot(1-p)+1\cdot p=p,$$

$$D^2(X)=(0-p)^2(1-p)+(1-p)^2p=p(1-p).$$

Schemat Bernoulliego

- wykonujemy doświadczenie, którego rezultatem może być zdarzenie A (sukces) z prawdopodobieństwem p lub zdarzenie przeciwne $\overline{A}$ (porażka) z prawdopodobieństwem $q=1-p$,

- doświadczenie powtarzamy n-krotnie w sposób niezależny co oznacza, że prawdopodobieństwo sukcesu pozostaje w pojedynczych próbach stałe i równe p,

- liczba sukcesów jaką zaobserwujemy w wyniku n-krotnego powtórzenia doświadczenia, może być równa $k=0,1,2,...,n$.

Założenia

- zmienna losowa X jest liczbą sukcesów zaobserwowanych w eksperymencie przeprowadzonym zgodnie ze schematem Bernoulliego,

- wyznaczamy dla niej funkcję prawdopodobieństwa, czyli $P(X=k)$, dla $k=0,1,...,n$:

  - zdarzenie $X=k$ zachodzi wtedy, gdy w wyniku n-krotnego powtórzenia pojedynczego doświadczenia zaobserwujemy dowolny ciąg zdarzeń, w którym pojawiło się k razy zdarzenie A, zaś $n-k$ razy zdarzenie $\overline{A}$:

$$\underbrace{A, A, ..., A}_{k} \qquad \underbrace{\overline{A}, \overline{A}, ..., \overline{A}}_{n-k}$$

  - prawdopodobieństwo otrzymania w eksperymencie takiego ciągu zdarzeń jest równe, ze względu na niezależność pojedynczych doświadczeń, iloczynowi prawdopodobieństw poszczególnych n zdarzeń, czyli $p^k(1-p)^{n-k}$,

  - prawdopodobieństwo otrzymania każdego innego ciągu zdarzeń, w którym A występuje k razy, zaś zdarzenie $\overline{A}$ $n-k$ razy, będzie takie samo,

  - liczba różnych możliwych ciągów n-elementowych, w których zdarzenie A występuje k razy, jest równa liczbie kombinacji z n elementów po k, czyli $\binom{n}{k}$,

  - wszystkie te kombinacje składają się na zdarzenie $X=k$, zatem jego prawdopodobieństwo jest sumą prawdopodobieństw dla poszczególnych kombinacji:

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k} \qquad (3)$$

Def. Zmienna losowa X ma *rozkład dwumianowy*, jeśli przyjmuje wartości $k=0,1,2,...,n$ z prawdopodobieństwami określonymi wzorem (3). Liczbę doświadczeń n oraz prawdopodobieństwo sukcesu p nazywamy parametrami tego rozkładu.

Dystrybuanta zmiennej losowej X o rozkładzie dwumianowym:

$$F(x) = P(X \le x) = \sum_{k \le x} \binom{n}{k} p^k (1-p)^{n-k}$$

Wartość oczekiwana i wariancja

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = np,$$
$$D^2(X) = D^2\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} D^2(X_i) = np(1-p).$$

ROZKŁAD POISSONA

Def. Rozkład zmiennej losowej X przyjmującej wartość $k=0,1,2,...$ nazywamy *rozkładem Poissona* o parametrze λ, jeżeli jej funkcja prawdopodobieństwa opisana jest wzorem:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad \text{dla} \qquad k=0,1,2,...$$

gdzie: λ jest dodatnią stałą ($\lambda>0$)

Def. *Dystrybuantą* zmiennej losowej X mającej *rozkład Poissona* jest funkcja $F(x)$ o postaci:

$$F(x) = \sum_{k \le x} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Paramerty:

$E(X)=\lambda$
$D^2(X)=\lambda$

Wykorzystanie rozkładu Poissona do aproksymacji prawdopodobieństw w rozkładzie dwumianowym

Niech X_n oznacza zmienną losową o rozkładzie dwumianowym, z parametrami n oraz p, której rozkład opisany jest wzorem:

$$P(X_k = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad (k = 0,1,2,\ldots n)$$

Jeżeli dla $n \to \infty$ spełniona jest równość $np=\lambda$, gdzie λ jest wielkością stałą, to spełniona jest zależność:

$$\lim_{n \to \infty} P(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

ROZKŁAD NORMALNY

Def. Zmienna losowa *X* ma *rozkład normalny* o parametrach *m* oraz σ, co w skrócie zapisuje się jako $X : N(m, \sigma)$, jeżeli jej funkcja gęstości wyraża się wzorem:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \qquad -\infty < x < +\infty$$

przy czym $\sigma > 0$.

Def. *Dystrybuantą* zmiennej losowej *X* mającej *rozkład normalny* jest funkcją *F(x)* określona na zbiorze liczb rzeczywistych o postaci:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

Def. Rozkład normalny ze średnią *m=0* oraz odchyleniem standardowym *σ=1 nazywamy standardowym rozkładem normalnym* i oznaczamy *N(0,1)*

WŁASNOŚCI KRZYWEJ GĘSTOŚCI ROZKŁADU NORMALNEGO

a) jest symetryczna względem prostej $x = m$,

b) osiąga maksimum równe $\frac{1}{\sigma\sqrt{2\pi}}\ dla\ x = m$,

c) jej ramiona mają punkty przegięcia dla $x = m - \sigma\ oraz\ x = m + \sigma$.

Zestaw VII

1. Udowodnić, że jeżeli zmienna losowa X ma rozkład normalny, to Y = aX + b ma też rozkład normalny.
2. Błąd pomiaru temperatury w skali Fahrenheita jest zmienną losową o rozkładzie N(0, 1). Wyniki pomiaru przenosimy na skalę Celsjusza. Niech Y będzie zmienną losową wyrażającą błąd w skali Celsjusza (Y=. Znaleźć gęstość g(y) zmiennej Y.
3. Znaleźć gęstość i dystrybuantę zmiennej Y wyrażającej objętość sześcianu, jeżeli długość krawędzi X jest zmienna losową o rozkładzie jednostajnym na przedziale (0, a) a następnie obliczyć P (a < Y <2a) oraz E(Y).
4. Zmienna losowa X ma rozkład o gęstości f(x). Znaleźć gęstość zmiennej Y = sX.
5. Wykazać, że jeżeli zmienna losowa X ma rozkład jednostajny na przedziale (0,1), to Y = -lnX ma rozkład wykładniczy.
6. Zmienna losowa X ma rozkład o funkcji gęstości f(x). Znaleźć gęstość g(y) zmiennej Y =h(X) . Rozważyć przypadek szczególny, gdy X ma rozkład wykładniczy.
7. Zmienna losowa X ma rozkład N(0, 1). Znaleźć rozkład zmiennej Y=h(X), jeżeli $h(X)=X^2$.
8. Promień koła jest zmienną losową R o gęstości f(r). Znaleźć gęstość zmiennej $S = \pi R^2$.
9. Znaleźć gęstość prawdopodobieństwa zmiennej a) Y = min(X, 1); b) Y = max(X, 1), jeżeli X ma gęstość f(x).
10. Wyrazić dystrybuantę zmiennej losowej Y za pomocą dystrybuanty zmiennej losowej X, jeżeli Y=h(X) oraz
    a) h(x) = aX + b,
    b) $h(x) = x^2$
    c) h(x) = |x|;
    d) $h(x) = 2x^3$
    e) h(x) = sinx
    f) h(x) = atanx
11. Zmienna losowa ma rozkład N(0, 1). Znaleźć rozkład zmiennej Y, jeżeli:
    a) Y = 2X+3,
    b) Y = |X|