

Statystyka

Magdalena Jakubek

kwiecień 2017

Nauka nie stara się wyjaśniać, a nawet niemal nie stara się interpretować, zajmuje się ona głównie budową modeli. Model rozumiany jest jako matematyczny twór, który opisuje badane zjawiska. Jedynym i właściwym uzasadnieniem takiego tworu matematycznego jest oczekiwanie, że sprawdzi się on w działaniu.

John von Neumann

1 Definicje podstawowych pojęć

Zbiór zdarzeń elementarnych – zbiór zdarzeń, które się wzajemnie wykluczają oraz wyczerpują wszystkie możliwości (tzn. w każdym możliwym przypadku przynajmniej jedno z nich musi zachodzić).

Zdarzeniem nazywamy dowolny podzbiór zdarzeń elementarnych ε .

Zdarzeniem losowym nazywamy zdarzenie spełniające warunki:

1. W zbiorze zdarzeń losowych znajduje się zdarzenie pewne oraz zdarzenie niemożliwe.
2. Jeżeli zdarzenia $A_1, A_2, \dots$ w ilości skończonej lub przeliczalnej są zdarzeniami losowymi to ich iloczyn i suma są również zdarzeniami losowymi.
3. Jeżeli A_1 i A_2 są zdarzeniami losowymi to ich różnica jest również zdarzeniem losowym.

Intuicyjne określenie: Zdarzenie losowe to takie, o którym nie możemy powiedzieć czy w danych warunkach zajdzie czy nie.

1.1 Definicja prawdopodobieństwa

Aksjomat 1: Każdemu zdarzeniu losowemu przyporządkowana jest jednoznacznie nieujemna liczba rzeczywista zwana prawdopodobieństwem zdarzenia: $0 \leq P(A) \leq 1$.

Aksjomat 2: Prawdopodobieństwo zdarzenia pewnego: $P(\Omega) = 1$.

Aksjomat 3: Jeżeli zdarzenie losowe A jest sumą skończonej lub przeliczalnej liczby rozłącznych zdarzeń losowych $A_1, A_2, \dots$ to prawdopodobieństwo zrealizowania zdarzenia A jest równe sumie prawdopodobieństw zdarzeń $A_1, A_2, \dots$.

Jeżeli dla dowolnych $i \neq j$ jest $A_i \cap A_j = \emptyset$,

$$P\left(\bigcup_i (A_i)\right) = \sum_i P(A_i)$$

Aksjomat 4: Prawdopodobieństwo warunkowe zdarzenia A pod warunkiem, że zachodzi zdarzenie B wyraża się wzorem:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Twierdzenie Bayesa

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

1.2 Zmienna losowa

Zmienną losową nazywamy jednoznacznie funkcję rzeczywistą określoną na zbiorze Ω zdarzeń elementarnych taką, że każdemu przedziałowi wartości funkcji oznaczanych jako X odpowiada zdarzenie losowe.

Zmienna losowa to inaczej wielkość, która może przyjmować wartości liczbowe z pewnego obszaru zmienności, z określonym prawdopodobieństwem.

- zmienną losową oznaczamy dużymi literami: $X, Y, \dots$
- zapisujemy bez argumentu: X zamiast $X(\omega)$
- argumentem funkcji *zmienna losowa* jest *zdarzenie elementarne*
- dla zdarzenia $\omega \in \Omega$ wartość $X(\omega)$ jest realizacją zmiennej losowej X
- realizacje zmiennej losowej oznaczamy małymi literami: $x, y, \dots$ ($x = 2$ dla $X(\omega) = 2$). $X = x$ oznacza, zdarzenie polegające na tym, że zmienna losowa X przyjęła wartość x .
- $X(\Omega)$ – wszystkie możliwe realizacje zmiennej losowej
- losowy jest argument zmiennej losowej

1.2.1 Rodzaje zmiennych losowych

Zmienna losowa:

- **typu skokowego (dyskretnego)** to taka, która przyjmuje tylko przeliczalny zbiór wartości (czyli zmienia się skokowo): $x_i, P(x_i)$.
- **typu ciągłego** może przyjmować dowolne wartości z przedziału zmienności (czyli zmienia się w sposób ciągły):

$$x, P(x_0 < X < x_0 + \Delta x), \quad P(x_i) = \frac{n(x_i)}{N} \quad \text{dla } N \longrightarrow \infty$$

$n(x_i)$ – liczba zdarzeń, którym przypisana jest zmienna x_i , N – liczba wszystkich zdarzeń.

Dzięki wprowadzeniu zmiennej losowej, prawdopodobieństwo zdarzenia można zastąpić prawdopodobieństwem przyjęcia przez zmienną losową określonej wartości x ze zbioru X . Zmienna losowa i rozkład prawdopodobieństwa pozwalają zapisać pewne zjawiska fizyczne w postaci modeli matematycznych.

1.2.2 Parametry rozkładu zmiennej losowej

Parametry rozkładu zmiennej losowej to wartości opisujące rozkład i pozwalające połączyć rachunek prawdopodobieństwa ze statystyką matematyczną.

- **Wartość oczekiwana** – $E(X)$ ($\bar{x}$, μ , m , m_1) inne nazwy: wartość przeciętna, wartość średnia, nadzieja matematyczna. Jest to jedna z miar określających "środek" zmiennej losowej.

- zmienna losowa dyskretna $E(X) = \sum_{i=1}^k x_i p_i$
- zmienna losowa ciągła $E(X) = \int_{-\infty}^{+\infty} x f(x) dx$

gdzie: X – zmienna losowa, x_i – argumenty zmiennej losowej, p_i – prawdopodobieństwo, k – liczba wartości, którą przyjmują argumenty zmiennej losowej.

Wybrane własności $E(X)$:

- $E(X) = c$
- $E(c \cdot X) = c \cdot E(X)$
- $E(X + Y) = E(X) + E(Y)$
- $E(X \cdot Y) = E(X) \cdot E(Y)$, X, Y **zmienne niezależne**
- **Odchylenie standardowe** ($\sigma, D(X)$) – najczęściej używana miara rozrzutu zmiennej losowej wokół wartości średniej. W celu określenia $D(X)$ definiujemy parametr – **wariancję** (dyspersję) oznaczaną: $\sigma^2(X), V(X), D^2(X)$

- dla dyskretnej zmiennej losowej $V(X) = \sum_{i=1}^k (x_i - E(X))^2 p_i$,
- dla ciągłej zmiennej losowej $V(X) = \sigma^2 = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx$

Odchylenie standardowe $\sigma = \sqrt{V(X)}$.

Własności wariancji:

- $V(c) = 0$
- $V(c \cdot X) = c^2 \cdot V(X)$
- $V(X \pm Y) = V(X) + V(Y)$, X, Y - zmienne niezależne

Wariancja : $D^2(X) = E(X^2) - [E(X)]^2$, inny zapis $D^2X = EX^2 - E^2X$

Wariancja zmiennej losowej jest równa wartości oczekiwanej kwadratu zmiennej losowej pomniejszonej o kwadrat wartości oczekiwanej.

$$D^2(X) = \left(\sum x_i^2 p_i \right) - \left(\sum x_i p_i \right)^2.$$

$$D^2(X) = \int_{-\infty}^{+\infty} x^2 f(x) dx - \left(\int_{-\infty}^{+\infty} x f(x) dx \right)^2.$$

2 Ilościowy opis zmiennych losowych

Ilościowy opis zmiennych losowych uzyskujemy stosując:

- **Dystrybuantę** $F(x)$ nazywaną funkcją rozkładu, która jest określona dla wszystkich x należących do $\mathbb{R}$, niezależnie od tego jakie wartości x przyjmuje zmienna losowa.
- **Rozkład prawdopodobieństwa** (inne nazwy: funkcja rozkładu prawdopodobieństwa, funkcja prawdopodobieństwa) – tylko w przypadku zmiennych dyskretnych. Funkcja, która realizacjom zmiennej losowej przyporządkowuje prawdopodobieństwo.
- **Funkcję gęstości prawdopodobieństwa** – tylko dla zmiennych ciągłych.

2.1 Rozkład zmiennej losowej typu skokowego

Jeżeli $x_1, x_2, \dots, x_k, \dots$ jest skończonym lub przeliczalnym zbiorem wartości zmiennej losowej X to funkcję

$$p_i = P(X = x_i) = P(\{\omega : X(\omega) = x_i\})$$

przyporządkowującą wartościom $x_1, x_2, \dots, x_k, \dots$ zmiennej losowej X odpowiednie prawdopodobieństwa $p_1, p_2, \dots, p_k, \dots$ nazywamy **rozkładem zmiennej losowej typu skokowego**.

2.2 Dystrybuanta zmiennej losowej typu skokowego

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

Dystrybuanta to funkcja sumująca wszystkie wcześniejsze prawdopodobieństwa. Rozkłady zmiennej losowej typu skokowego np: dwupunktowy, rozkład Bernoulliego, rozkład Poissona.

2.3 Dystrybuanta zmiennej losowej ciągłej

Dystrybuanta $F(x)$ to prawdopodobieństwo tego, że zmienna losowa X przyjmie wartość mniejszą od x . Dystrybuanta jest funkcją x .

$$F(x) \equiv P(X \leq x) = \int_{-\infty}^x f(u) du$$

Własności dystrybuanty:

- $0 \leq F(x) \leq 1$
- $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$
- $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$
- $F(x)$ jest funkcją niemalejącą
- $F(x)$ nie ma wymiaru
- $P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1), x_1, x_2 \in \mathbb{R}$

2.4 Funkcja rozkładu prawdopodobieństwa dla zmiennej losowej typu dyskretnego

Wartościom zmiennej losowej przypisujemy prawdopodobieństwo ich realizacji. Taka funkcja to [funkcja rozkładu prawdopodobieństwa](#) zmiennej losowej.

Wartość zmiennej losowej to argument funkcji.

Jeżeli x_i (dla $i = 1, 2, \dots$) są wartościami dyskretnej zmiennej losowej to rozkładem prawdopodobieństwa nazywamy zespół prawdopodobieństw:

$$p(x_i) = P(X = x_i) = p_i > 0$$
$$\sum_i p_i = 1 \quad \text{dla } i \in N$$

2.5 Funkcja gęstości prawdopodobieństwa dla zmiennej losowej typu ciągłego

Prawdopodobieństwo, że zmienna losowa trafi do przedziału x_1, x_2 , jest równe polu pod krzywą gęstości między punktami x_1, x_2 a także przyrostowi dystrybuanty na odcinku x_1, x_2 .

$$\int_{x_1}^{x_2} f(x) dx \equiv P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$$

Własności funkcji gęstości prawdopodobieństwa:

1. $f(x) \geq 0$,
2. $f(x)$ jest unormowana tj. $\int_{-\infty}^{+\infty} f(x) dx = 1$,
3. $f(x) = \frac{dF(x)}{dx}$,
4. wymiar $f(x) = \text{wymiar } (\frac{1}{x})$.

Jeżeli X jest zmienną losową typu ciągłego a $F(x)$ jest dystrybuantą tej zmiennej, to przy $\Delta x > 0$

$$\lim_{\Delta x \rightarrow 0^+} \frac{P(x \leq X < x + \Delta x)}{\Delta x} = \lim_{x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x).$$

Dystrybuanta F jest funkcją różniczkowalną a funkcję $f(x)$ nazywamy [gęstością prawdopodobieństwa](#) zmiennej losowej typu ciągłego

$$f(x) = F'(x) \geq 0.$$

Mówimy o zmiennej losowej ciągłej gdy w danym obszarze ma ona gęstość prawdopodobieństwa $f(x)$ i $f(x)$ jest funkcją ciągłą.

2.6 Dystrybuanta a gęstość prawdopodobieństwa

Zależności pomiędzy dystrybuantą F a gęstością f :

$$F'(x) = \frac{d}{dx}F(x) = f(x), \quad \text{gdy } F \text{ różniczkowalna}$$

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{x_0}^x f(t)dt + F(x_0),$$

$$P(a \leq X < b) = \int_a^b f(t)dt = F(b) - F(a),$$

$$\int_{-\infty}^{+\infty} f(t)dt = F(+\infty) - F(-\infty) = 1.$$

2.7 Rozkłady dyskretne

2.7.1 Rozkład dwupunktowy

Zmienna losowa X przyjmuje tylko dwie wartości:

- x_1 z prawdopodobieństwem $P(X = x_1) = p$
- x_2 z prawdopodobieństwem $P(X = x_2) = q$, $q = 1 - p$, $p \in (0, 1)$

Gdy $x_1 = 1$, $x_2 = 0$ zmienna X ma rozkład zero-jedynkowy z parametrem p . Parametr p nazywamy [prawdopodobieństwem sukcesu](#).

$$P(X = 1) = p, \quad P(X = 0) = q = 1 - p.$$

$$F(x) = \begin{cases} 0 & \text{dla } x \leq 0 \\ q = 1 - p & \text{dla } 0 < x \leq 1 \\ 1 & \text{dla } 1 < x \end{cases}$$

Gdy X ma rozkład zero-jedynkowy: $E(X) = p$, $D^2(X) = pq$.

2.7.2 Rozkład Bernoulliego (dwumianowy) – B(n,p)

Zmienna losowa X przyjmuje wartości $k \in \{0, 1, 2, \dots, n-1, n\}$ odpowiednio z prawdopodobieństwem

$$p_k = P_n(k) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n,$$

gdzie: $p \in (0, 1)$ jest ustalone, $q = 1 - p$, $n \in \mathbb{N}$.

(prawdopodobieństwo uzyskania k sukcesów w n próbach).

Gdy X ma rozkład dwumianowy: $E(X) = np$, $D^2(X) = npq$.

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

$$F(x) = \sum_{k < x} P_n(k) = \sum_{k < x}^n \binom{n}{k} p^k q^{n-k}$$

2.7.3 Doświadczenie Bernoulliego

W przypadku zmiennej o rozkładzie dwumianowym zakłada się, że mamy do czynienia z eksperymentem losowym polegającym na wykonaniu ciągu doświadczeń Bernoulliego.

Doświadczenie Bernoulliego to ciąg n identycznych doświadczeń losowych, spełniających warunki:

1. są dwa możliwe wyniki każdego doświadczenia (sukces i porażka)
2. prawdopodobieństwo sukcesu p jest w każdym doświadczeniu stałe
3. doświadczenia są niezależne, co oznacza, że wynik jednego doświadczenia nie ma wpływu na wyniki pozostałych doświadczeń

2.7.4 Rozkład Poissona – $P(\lambda)$

Zmienna losowa X przyjmuje odpowiednio wartości będące liczbami naturalnymi $k \in \{0, 1, 2, \dots\}$ z prawdopodobieństwami

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda},$$

gdzie λ jest ustalonym dodatnim parametrem.

Rozkład Poissona jest granicznym przypadkiem rozkładu dwumianowego przy $n \rightarrow \infty$ i zmiennym $p = \lambda/n$, $p \rightarrow 0$.

Gdy X ma rozkład Poissona: $E(X) = \lambda$, $D^2(X) = \lambda$

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

Zmienna losowa przyjmuje przeliczalnie wiele wartości.

2.8 Najważniejsze rozkłady zmiennej losowej typu ciągłego

- rozkład jednostajny
- rozkład Gaussa
- rozkład log-normalny (wytrzymałość zmęczeniowa, trwałość). Jeżeli zmienna losowa X jest dodatnia to rozkład normalny nazywamy logarytmiczno normalnym.
- rozkład Gumbela – maksymalne lub minimalne pomiary
- rozkład Weibulla – wytrzymałość materiałów na rozciąganie

2.8.1 Rozkład jednostajny

Zmienna losowa ciągła X ma rozkład jednostajny w przedziale $[a, b]$, jeżeli jej funkcja gęstości wyraża się wzorem:

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

Dla takiej zmiennej losowej: $E(X) = \frac{b+a}{2}$, $D^2(X) = \frac{1}{12}(b-a)^2$.

2.8.2 Rozkład normalny

Gęstość rozkładu normalnego (rozkładu Gaussa):

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

gdzie $\mu \in \mathbb{R}$, $\sigma > 0$ są ustalonymi parametrami, które określają odpowiednio położenie środka (przesunięcie krzywej) i szerokość (smukłość) krzywej Gaussa. Wartość funkcji gęstości jest niezerowa dla dowolnego x .

Rozkład Gaussa oznaczamy $N(\mu, \sigma)$. Dla zmiennej X o rozkładzie normalnym $E(X) = \mu$, $D^2(X) = \sigma^2$.

2.8.3 Charakterystyki rozkładu ciągłego

Wartość średnią rozkładu ciągłego o gęstości $f(x)$ określa się wzorem

$$\mu = \int_{-\infty}^{+\infty} x f(x) dx.$$

A wariancję wzorem:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$

Liczbę λ_p , gdzie $0 < p < 1$, spełniającą warunek

$$F(\lambda_p) = p,$$

nazywamy kwantylem rzędu p . Kwantyl rzędu $1/2$ nazywamy medianą - m_e , kwantyl rzędu $1/4$ - kwantylem dolnym, a kwantyl rzędu $3/4$ - kwantylem górnym

2.8.4 Rozkład zmiennej losowej typu ciągłego

Prawdopodobieństwo w rozkładzie normalnym wyznaczone jest dla wartości zmiennej losowej z określonego przedziału $P(a < X \leq b)$

$$P(a < X \leq b) = F(b) - F(a).$$

Dystrybuanta $F(x)$ ma postać

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

2.8.5 Rozkład Gaussa-Laplace'a – normalny standaryzowany

Dowolną zmienną o rozkładzie normalnym można standaryzować tworząc wielkość Z o [rozkładzie standardowym normalnym](#) $N(0, 1)$:

$$\text{Standaryzowana zmienna} \quad Z = \frac{X - \mu}{\sigma}$$

$\phi(x)$ i $\Phi(x)$ to funkcje Laplace'a.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} = \phi(x), \quad \text{dla } x \in \mathbb{R}$$

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

Własności funkcji rozkładu normalnego

- dla $\mu = 0$ jest funkcją parzystą,
- posiada maximum w punkcie $E(X) = \frac{1}{\sqrt{2\pi}\sigma}$,
- pole ograniczone tą funkcją i osią x jest równe jedności,
- wraz ze wzrostem parametru σ rozkład normalny ulega spłaszczeniu,
- w punktach $x = \mu \pm \sigma$ funkcja ma punkty przegięcia.
- funkcja gęstości przyjmuje wartości nieujemne

2.9 Rozkład Weibulla

[Rozkład Weibulla](#) zmienna losowa X ma rozkład Weibulla o parametrach $k, \lambda > 0$ jeżeli jej gęstość jest określona wzorem

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right) \quad \text{dla } x > 0$$

$$\text{średnia} \quad \lambda \Gamma\left(1 + \frac{1}{k}\right)$$

$$\text{mediana} \quad \lambda (\ln(2))^{1/k}$$

$$\text{moda} \quad \begin{cases} \lambda \left(\frac{k-1}{k}\right)^{1/k} & k > 1 \\ 0 & k = 1 \end{cases}$$

2.10 Rozkład Gumbela

Rozkład zmiennej losowej służący do wyznaczania ekstremalnych wartości zmiennej losowej. Większość losowych zjawisk naturalnych (takich jak temperatura otoczenia, prędkość wiatru) daje się dobrze opisywać rozkładem Gumbela.

$$F(x, \mu, \beta) = \exp^{-\exp^{-(x-\mu)/\beta}}$$

Parametry tego rozkładu μ, β są w następujący sposób związane ze standardowymi cechami rozkładów prawdopodobieństwa:

$$\text{średnia} \quad \mu + \beta\gamma, \quad \gamma = 0.5772 \quad \text{stała Eulera}$$

$$\text{mediana} \quad \mu - \beta \ln(\ln 2), \quad \text{moda} \quad \mu$$

$$\text{wariancja} \quad \frac{\pi^2}{6}\beta^2.$$

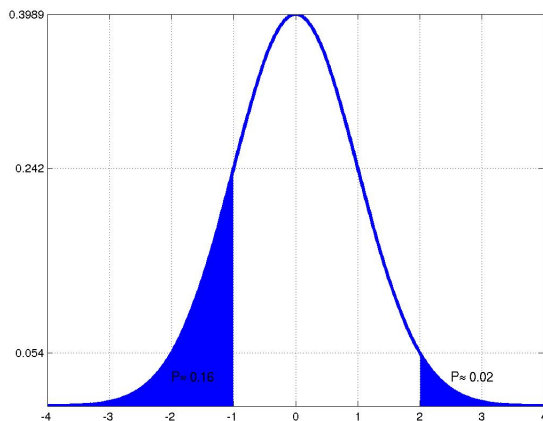
2.11 Parametry rozkładu zmiennej losowej

Prawdopodobieństwo znalezienia wartości zmiennej losowej. Wartość całki z funkcji Gaussa.

$$P(E(X) - \sigma(X) \leq X \leq E(X) + \sigma(X)) = 0.6827$$

$$P(E(X) - 2\sigma(X) \leq X \leq E(X) + 2\sigma(X)) = 0.9545$$

$$P(E(X) - 3\sigma(X) \leq X \leq E(X) + 3\sigma(X)) = 0.9973$$



2.12 Charakterystyki zmiennej losowej

W praktycznych zastosowaniach często nie trzeba znać pełnej informacji o rozkładzie, wystarczy poznanie wartości pewnych wielkości, które opisują rozkład prawdopodobieństwa.

Liczby takie to [charakterystyki zmiennej losowej](#).

	zmienna losowa dyskretna	zmienna losowa ciągła
Momenty zwykłe rzędu r	$\alpha_r = \sum_i x_i^r \cdot p_i$	$\alpha_r = \int_{-\infty}^{+\infty} x^r \cdot f(x)dx$
Pierwszy moment zwykły wartość oczekiwana	$EX = \sum_i x_i \cdot p_i$	$EX = \int_{-\infty}^{+\infty} x \cdot f(x)dx$
Momenty centralne rzędu r	$\mu_r = \sum_i (x_i - EX)^r \cdot p_i$	$\mu_r = \int_{-\infty}^{+\infty} (x - EX)^r \cdot f(x)dx$
Drugi moment centralny D^2X czyli wariancja	$D^2X = \sum_i (x_i - EX)^2 \cdot p_i$	$D^2X = \int_{-\infty}^{+\infty} (x - EX)^2 \cdot f(x)dx$

Parametry, które liczymy bez względu na typ rozkładu:

- Współczynnik zmienności ν

$$\nu = \frac{\sigma}{EX}$$

- Skośność inaczej współczynnik asymetrii γ (3 moment centralny)

$$\gamma = \frac{\mu_3}{\sigma^3}$$

- Kurtoza inaczej współczynnik skupienia K

$$K = \frac{\mu_4}{\sigma^4}$$

Fraktyl x_p (inaczej kwantyl) – wartość zmiennej losowej taka, że prawdopodobieństwo znalezienia mniejszych od niej wartości wynosi p : $P(X < x_p) \equiv f(x_p) = p$.

dyskretna	ciągła
$\sum_{x_i < x_p} p_i \leq p \leq \sum_{x_i \leq x_p} p_i$	$F(x_p) = \int_{-\infty}^{+\infty} f(x)dx = p$

Najważniejsze fraktyle to: dolny fraktyl: $x_{0.25}$, górny fraktyl: $x_{0.75}$ oraz mediana: $x_{0.5}$ (czyli wartość zmiennej losowej, która dzieli obszar jej zmienności na dwa obszary o równym prawdopodobieństwie).

- **Moda** jest to taka wartość zmiennej losowej, dla której rozkład prawdopodobieństwa (lub funkcja gęstości prawdopodobieństwa) ma maksimum. Dla zmiennej losowej typu dyskretnego, moda występuje z największym prawdopodobieństwem.
- Rozkłady prawdopodobieństwa posiadające jedną modę zwane są jednomodalnymi.

- **Mediana** to taka liczba $x_{0.5}$, że połowa wszystkich przyjmowanych przez zmienną losową wartości leży poniżej jej wartości:

zmienna losowa dyskretna	zmienna losowa ciągła
$\sum_{x_i < x_{0.5}} p_i \leq 0.5 \leq \sum_{x_i \leq x_{0.5}} p_i$	$F(x_{0.5}) = \int_{-\infty}^{+\infty} f(x)dx = 0.5.$

2.13 Ilościowy opis zmiennych losowych

Rozkład częstości – spis danych wartości wraz z ich częstościami czy zliczeniami.

dolna/górna granica klasy – min/max liczba należąca do klasy

punkty środkowe klas – czyli punkty w środku klasy

szerokości klas – różnica pomiędzy dwoma kolejnymi granicami klas.

Procedura postępowania

1. ilość klas (5 -20)
2. wyliczenie szerokości klas
3. wyznaczenie przynależności danych do poszczególnych klas

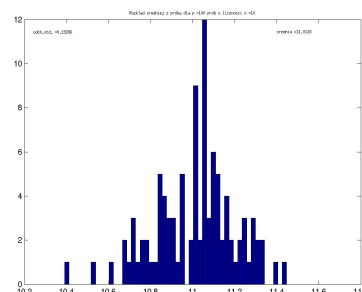
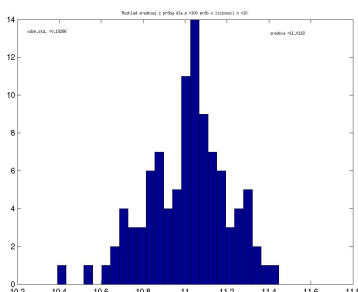
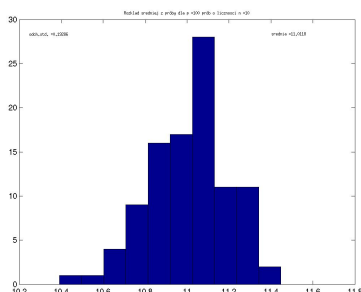
2.14 Graficzne przedstawienie wyników pomiarów

Z pomiarów otrzymujemy ciąg liczb, który nazywamy próbą losową. W zależności od liczebności przyjmujemy nazwy prób:

$n \leq 10$	próba bardzo mała
$10 < n \leq 30$	próba mała
$30 < n \leq 100$	próba duża
$n \geq 100$	próba bardzo duża

Dla $n > 30$ tworzymy szereg rozdzielczy. Budujemy przedziały lewostronnie domknięte $[x_{01}, x_{1j})$ i nazywamy je klasami. Ustalamy ile elementów należy do poszczególnych klas.

Taki szereg pozwala utworzyć **histogram empiryczny**. Histogram składa się ze słupków o szerokości pojedynczej klasy i wysokości równej liczebności w danym słupku.



3 Przykładowe zadania

1. Z przedziału $< -1, 1 >$ wybieramy losowo dwie liczby x i y . Oblicz prawdopodobieństwo, że spełnią one warunek: $2x^2 - 1 \leq y \leq -x$.
2. Obliczyć prawdopodobieństwo tego, że zmienna losowa określona funkcją gęstości

$$f(x) = \begin{cases} \exp(-x) & \text{dla } x \in < 0, \infty), \\ 0 & \text{dla } x < 0 \end{cases}$$

przyjmie wartości z przedziału $< 1, 2 >$.

3. Test egzaminacyjny składa się z 10 pytań. Każde pytanie ma 4 możliwe odpowiedzi ale tylko jedna jest prawidłowa. Student odpowiada losowo. Oblicz prawdopodobieństwo, że odpowie prawidłowo na co najmniej 9 pytań.
4. Prawdopodobieństwo trafienia w pojedynczym rzucie do kosza jest stałe i wynosi $p = 0.3$. Rzucamy 5 razy. Obliczyć rozkład prawdopodobieństwa, wartość oczekiwaną i wariancję.
5. Określić dla jakiej wartości stałej rzeczywistej A funkcja

$$f(x) = \frac{A}{1+x^2}$$

jest gęstością prawdopodobieństwa rozkładu zmiennej losowej X .

6. Mając daną dystrybuantę

$$F(x) = \begin{cases} 0 & x \leq 0 \\ x^2 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

wyznaczyć funkcję gęstości rozkładu zmiennej losowej X .

7. Dana jest funkcja

$$f(x) = \begin{cases} 0 & \text{gdy } x \leq 0 \\ A \cos(x) & \text{gdy } 0 < x \leq \pi/2 \\ 0 & \text{gdy } x > \pi/2 \end{cases}$$

Wyznaczyć:

- stałą A tak by $f(x)$ opisywała gęstość prawdopodobieństwa zmiennej losowej X typu ciągłego
 - dystrybuantę tego rozkładu
 - $P(|X| \leq \frac{\pi}{4})$
8. Wyznacz moment zwykły rzędu k oraz moment centralny rzędu k zmiennej losowej X o gęstości prawdopodobieństwa danej wzorem

$$f(x) = \begin{cases} 0 & x \notin < 0, 2 > \\ 0.5 & x \in < 0, 2 > . \end{cases}$$