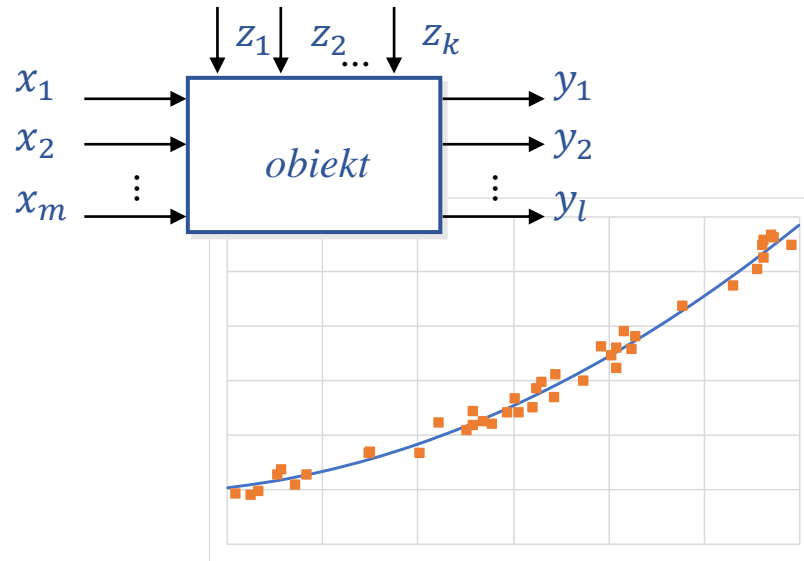


# Planowanie doświadczeń

## Wprowadzenie Podstawowe pojęcia statystyki matematycznej



Materiały

<http://pracownicy.uz.zgora.pl/ipajak/>

## ***Badania naukowe***

prace badawcze (teoretyczne i doświadczalne), zmierzające do osiągnięcia postępu wiedzy naukowej przez odkrywanie nowych praw i budowanie nowych teorii naukowych o rzeczywistości, ... *(Encyklopedia PWN)*

## ***Teoria eksperymentu***

dyscyplina naukowa zajmująca się podstawami teoretycznymi badań doświadczalnych.

## ***Eksperyment***

próba, doświadczenie naukowe; celowe wywoływanie określonego zjawiska (lub jego zmiany) w warunkach laboratoryjnych, w celu zbadania i wyjaśnienia jego przebiegu *(Encyklopedia PWN)*

eksperyment jest przeprowadzany na drodze badań doświadczalnych; może być przedstawiany jako ciąg doświadczeń.

## Doświadczenie

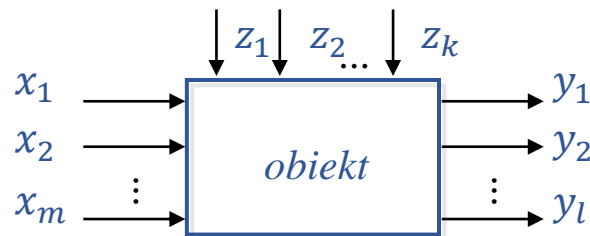
pojedyncza obserwacja wielkości objaśnianej (wyjściowej, zależnej) w zależności od ustalonych lub zaobserwowanych wielkości objaśniających (wejściowych, niezależnych).

## Eksperyment czynny

doświadczenia przeprowadzane według przygotowanego wcześniej planu.

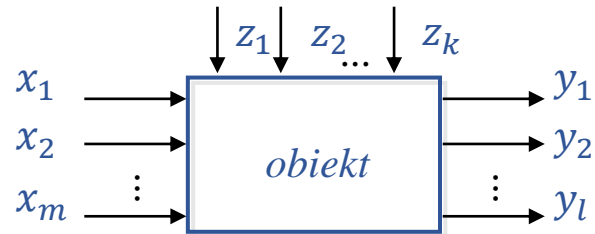
## Eksperyment bierny

doświadczenia dla wyłącznie obserwowanych wielkości objaśniających.



$x_1, x_2, \dots, x_m$  – wielkości wejściowe  
(objaśniające, niezależne),  
 $z_1, z_2, \dots, z_k$  – zakłócenia,  
 $y_1, y_2, \dots, y_l$  – wielkości wyjściowe  
(objaśniane, zależne)

Schemat typowego obiektu badań



## Plan eksperymentu

Zbiór wartości wielkości wejściowych, dla których należy przeprowadzić obserwacje wielkości wyjściowej. Matematycznie plan eksperymentu zapisywany jest w postaci *macierzy planu*.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

*Macierz planu*

$\mathbf{X}$  – macierz planu,

$n$  – liczba doświadczeń,

$x_{i1}, \dots, x_{im}$  – wielkości wejściowe  $i$ -tego doświadczenia.

## LITERATURA

- Dobosz M., *Wspomagana komputerowo statystyczna analiza wyników badań*, Akademicka Oficyna Wyd. EXIT, Warszawa, 2001,
- Górecka R., *Teoria i technika eksperymentu*, Wyd. Politechniki Krakowskiej, Kraków, 1998
- Korzyński M., *Metodyka eksperymentu. Planowanie, realizacja i statystyczne opracowanie wyników eksperymentów technologicznych*, Warszawa, PWN, 2006
- Kukielka L., *Podstawy badań inżynierskich*, PWN, Warszawa, 2002
- Mańczak K., *Technika planowania eksperymentu*, WNT, Warszawa, 1976
- Rumszyński L. Z., *Matematyczne opracowanie wyników eksperymentu*, WNT, Warszawa, 1973
- Wawrzynek J., *Planowanie eksperymentów zorientowane na doskonalenie jakości produktu*, Wyd. Uniw. Ekon. We Wrocławiu, Wrocław, 2009
- Brandt S., *Analiza Danych*, PWN, Warszawa 2002
- Montgomery D. C., *Design and Analysis of Experiments*, Wiley, 2012

## ***Podstawy teoretyczne badań doświadczalnych – teoria eksperymentu***

planowanie doświadczeń (plany zdeterminowane: kompletne i selekcyjne, dwupoziomowe, trzypoziomowe i wielopoziomowe, plany randomizowane: kompletne, blokowe, typu kwadrat łaciński i grecko-łaciński)

## ***Statystyczne metody analizy wyników pomiarów***

estymacja punktowa i przedziałowa, wyznaczanie minimalnej liczebności próby

sprawdzanie prawdziwości hipotez badawczych – weryfikacja hipotez statystycznych i analiza mocy testów

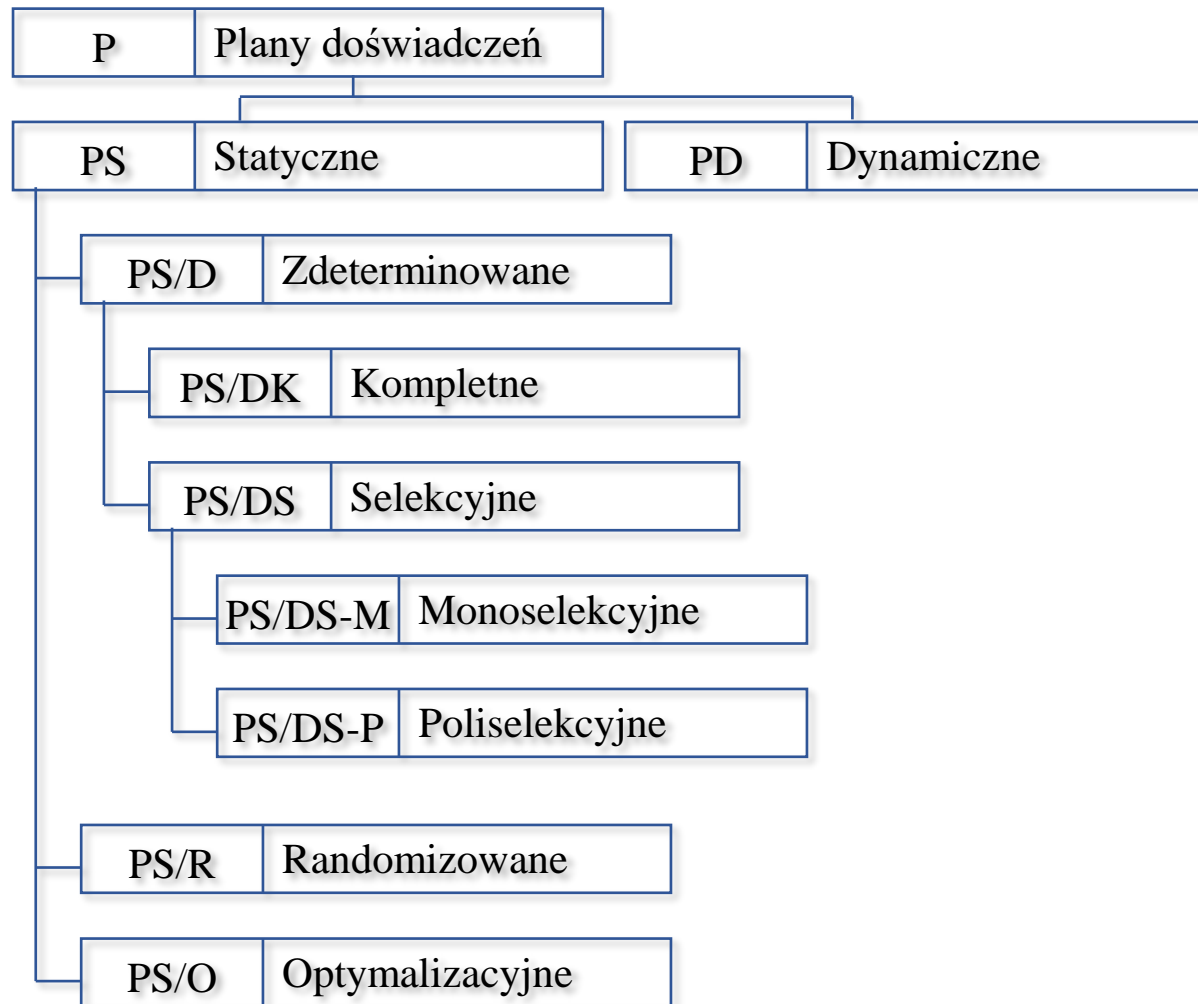
## ***badanie istotności wpływu czynników wejściowych na przebieg badanego procesu***

analiza wariancji

## ***aproksymacja funkcji obiektu badań***

metoda najmniejszych kwadratów, analiza regresji, statystyczna weryfikacja istotności funkcji obiektu oraz jej parametrów

# Klasyfikacja planów doświadczeń

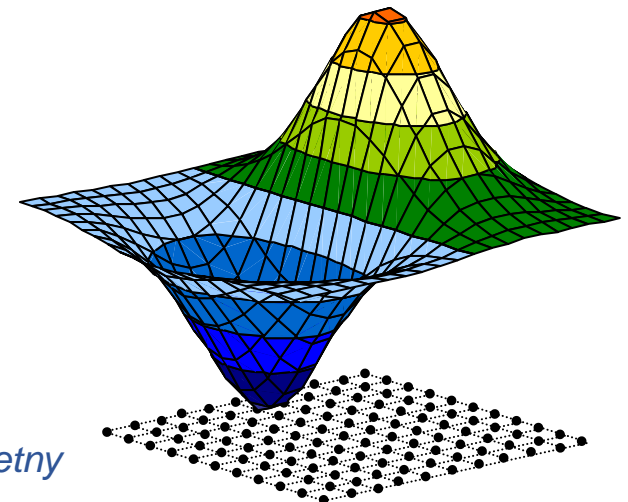


*Klasyfikacja planów doświadczeń*

- PS** – *plany statyczne*, stosowane do obiektów statycznych (tzn. opisanych zależnościami  $y = f(x)$ ).
- PD** – *plany dynamiczne*, stosowane do obiektów dynamicznych opisywanych równaniami różniczkowymi (różnicowymi) zwyczajnymi, cząstkowymi.
- PS/D** – *plany zdeterminowane*, plan doświadczenia jest ustalany na podstawie przyjętych założeń dotyczących np. postaci funkcji obiektu.
- PS/DK** – *plany kompletne*, plan doświadczenia tworzą wszystkie punkty z dziedziny funkcji przy ustalonym zakresie zmienności wielkości wejściowych i przyjętym sposobie ich dyskretyzacji; ogólna liczba pomiarów planu wynosi:

$$n = n_1 \cdot n_2 \cdot \dots \cdot n_m$$

( $m$  – liczba wielkości wejściowych,  $n_i$  – ilość wartości  $i$ -tej wielkości wejściowej, wynika z jej zakresu i dyskretyzacji).

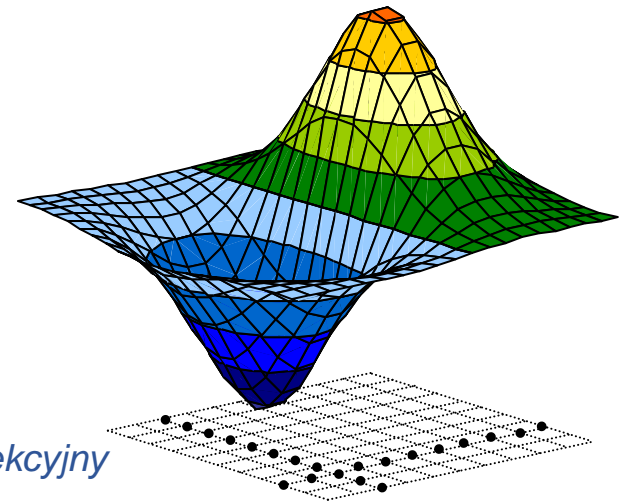


Plan kompletny



# Klasyfikacja planów doświadczeń

- PS/DS** – *plany selekcyjne*, plan doświadczenia jest podzbiorem planu kompletnego.
- PS/DK-M** – *plany monoselekcyjne*, każda z wielkości wejściowych badana jest w całym zakresie swojej zmienności przy ustalonych wartościach pozostałych wielkości wejściowych.
- PS/R** – *plany randomizowane*, plany wykorzystywane do oceny istotności wpływu wielkości wejściowych na wielkość wyjściową.
- PS/O** – *plany optymalizacyjne*, plany wykorzystywane do wyznaczania ekstremum funkcji obiektu (nie wymagają wyznaczania samej funkcji).



*Plan monoselekcyjny*

# Statystyka matematyczna – pojęcia podstawowe

**Populacja (populacja generalna)** – zbiór elementów (osób, rzeczy, zjawisk), podlegających badaniu ze względu na jedną lub więcej cech.

*Cechy statystyczne* mogą być:

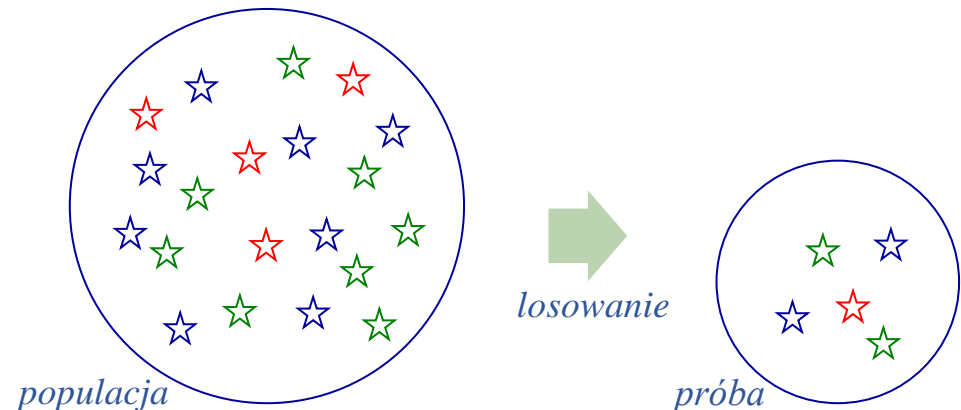
- *mieralne (ilościowe)*

przyjmują wartości ze zbioru liczbowego (np.: długość, waga)

- *niemierzalne (jakościowe)*

cechy których nie można wyrazić ilościowo, są opisywane słownie lub wyrażane przy pomocy wybranej skali (np.: płeć, kolor, funkcjonalność).

**Próba (populacja próbna)** – wybrany w określony sposób (np. przez losowanie) podzbiór populacji generalnej.



# Statystyka matematyczna – pojęcia podstawowe

Wartości prób mogą być prezentowane w formie tzw. **szeregów**.

## **Szereg prosty**

wartości porządkowane są rosnąco lub malejąco.

długość	2,9	3,0	3,2	3,3	3,4	3,5	3,5	3,6	4,0	4,1
---------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## **Szereg rozdzielczy**

wartości dzielone są na **klasy** (kategorie), dla każdej klasy podawana jest jej **liczebność** lub **częstość** (stosunek liczebności klasy do liczebności całej próby).

długość	[2,5 3,0)	[3,0 3,5)	[3,5 4,0)	[4,0 4,5]
liczebność	1	4	3	2
częstość	0,1	0,4	0,3	0,2

# Statystyka matematyczna – pojęcia podstawowe

**Zmienna** – to wielkość, która może przyjmować wartości z określonego zbioru.

**Zmienna losowa** – to zmienna, która w wyniku pewnego doświadczenia przyjmuje wartość z określonego zbioru z pewnym prawdopodobieństwem.

**Skokowa (dyskretna) zmienna losowa** – zmienna losowa która przyjmuje skończoną lub przeliczalną liczbę wartości.

**Ciągła zmienna losowa** – zmienna losowa której zbiór wartości jest nieskończony i nieprzeliczalny, może być np. przedstawiony w postaci przedziału liczbowego.

przykład	zmienna/wartości
zmienna losowa „Liczba wadliwych produktów” jako wynik kontroli jakości 6 wybranych produktów z linii produkcyjnej	skokowa / 0, 1, ..., 6
zmienna losowa „Liczba sprzedanych sztuk” jako dzienna ilość sprzedanych sztuk wybranego produktu	skokowa / 0, 1, ...
zmienna losowa „Długość detalu” jako wynik pomiaru długości wybranych detali z linii produkcyjnej	ciągła / 19,9, ..., 30,9

# Jednowymiarowe zmienne losowe

Jeżeli znany jest zbiór możliwych wartości zmiennej losowej oraz prawdopodobieństwa przyjęcia tych wartości przez zmienną losową (bądź też prawdopodobieństwa, że zmienna przyjmie wartość z określonego przedziału) to mówimy, że znany jest **rozkład tej zmiennej losowej**\*

**Rozkładem prawdopodobieństwa zmiennej losowej  $X$**  nazywana jest funkcja  $P(S)$  oznaczająca prawdopodobieństwo tego, że zmienna losowa przyjmie wartość z  $S$  (funkcja ta przedstawia związek między wartościami zmiennej losowej a prawdopodobieństwami, z jakimi te wartości występują). Sposób przedstawiania rozkładu prawdopodobieństwa zależy od typu zmiennej losowej:

- dla **zmiennej losowej skokowej** podaje się wartości tej zmiennej wraz z odpowiadającymi im prawdopodobieństwami,
- dla **zmiennej losowej ciągłej** rozkład zmiennej losowej podaje się za pomocą **funkcji gęstości prawdopodobieństwa**.

**Dystrybuanta zmiennej losowej  $X$ :  $F(X)$**  – to funkcja opisująca prawdopodobieństwo wystąpienia wartości zmiennej  $X$  mniejszych od  $x$ :

$$F(x) = P(X < x)$$

**Uwaga!**  $F(\infty) = 1$

\* Z.Pawłowski, *Wstęp do statystyki matematycznej*

# Jednowymiarowe zmienne losowe

Do opisania *rozkładu skokowej zmiennej losowej* wystarczy podać wszystkie prawdopodobieństwa:

$$p_i = P(X = x_i), \quad i = 1, 2, \dots$$

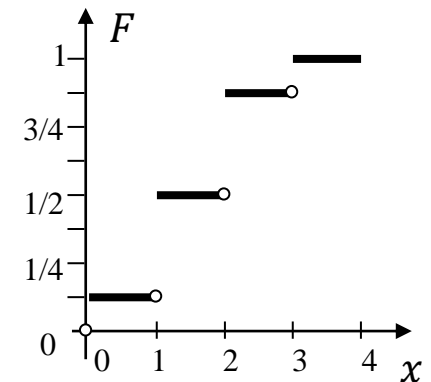
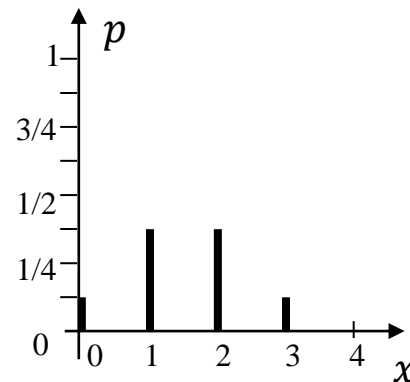
gdzie:  $X$  – zmienna losowa,  $x_i$  –  $i$ -ta wartość zmiennej losowej  $X$ ,  $P(X = x_i)$  – prawdopodobieństwo, że zmienna  $X$  przyjmie wartość  $x_i$ ,  $\sum P(X = x_i) = 1$ .

*Dystrybuantę dyskretnej zmiennej losowej* można zapisać wzorem:

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

Funkcja rozkładu prawdopodobieństwa i dystrybuanta dyskretnej zmiennej losowej przedstawiane są w formie tabelarycznej lub w postaci wykresu.

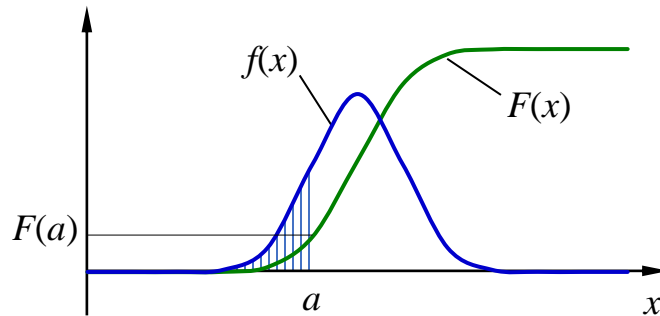
$x_i$	0	1	2	3	4
$p_i$	1/8	3/8	3/8	1/8	0
$F_i$	0	1/8	4/8	7/8	1



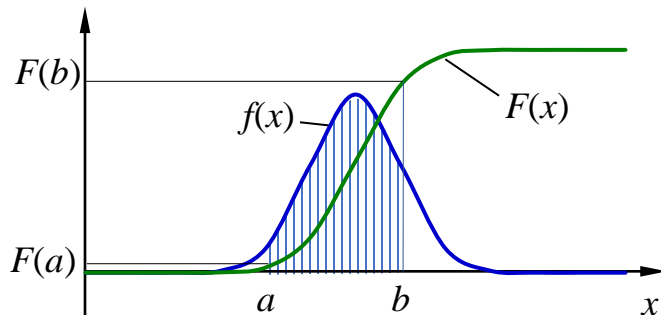
# Jednowymiarowe zmienne losowe

Do opisania *rozkładu ciągłej zmiennej losowej* wykorzystywana jest *funkcja gęstości prawdopodobieństwa*  $f$ , dla której spełniona jest zależność:

$$F(x) = P(X < x) = \int_{-\infty}^x f(t) dt$$



$$F(a) = P(X < a) = \int_{-\infty}^a f(x) dx$$



$$F(b) - F(a) = P(a \leq X < b) = \int_a^b f(x) dx$$

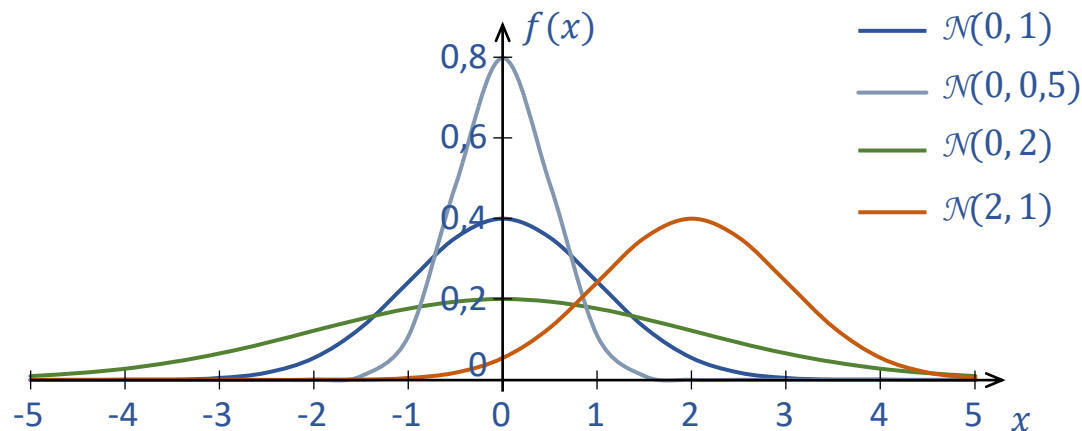
# Rozkład normalny

**Rozkład normalny (rozkład Gaussa)** jest jednym z częściej spotykanych rozkładów zmiennych losowych ciągłych (wiele zjawisk fizycznych ma rozkład normalny).

**Funkcja gęstości rozkładu  $f$  i dystrybuanta  $F$**  rozkładu normalnego  $\mathcal{N}(\mu, \sigma)$  opisane są zależnościami ( $\mu, \sigma$  – parametry rozkładu).

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

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





# Rozkład normalny standaryzowany

Zmienna losowa  $U$  utworzona ze zmiennej losowej  $X$  o rozkładzie normalnym  $\mathcal{N}(\mu, \sigma)$  za pomocą przekształcenia:

$$Z = \frac{X - \mu}{\sigma}$$

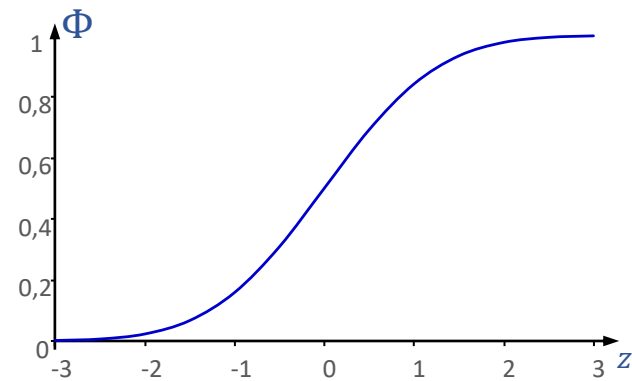
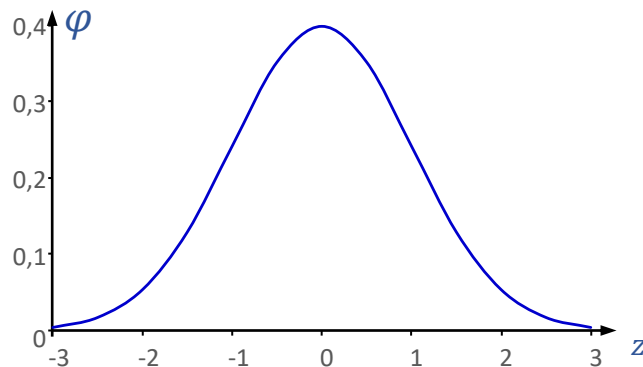
*Średnia populacji  $\mu$  jest odejmowana od każdej wartości cechy  $x$ , każda wyznaczona różnica dzielona jest przez odchylenie standardowe populacji  $\sigma$*

ma rozkład normalny  $\mathcal{N}(0, 1)$ .

Zmienna  $Z$  jest nazywana **zmienną losową normalną standaryzowaną**, a rozkład  $\mathcal{N}(0, 1)$  jest nazywany jest **rozkładem normalnym standaryzowanym**. Funkcja gęstości prawdopodobieństwa i dystrybuanta rozkładu opisane są zależnościami:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}},$$

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

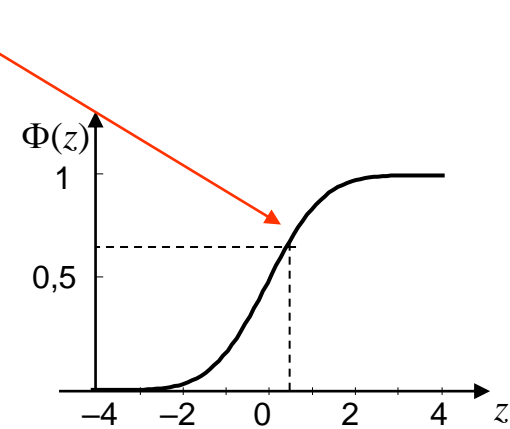
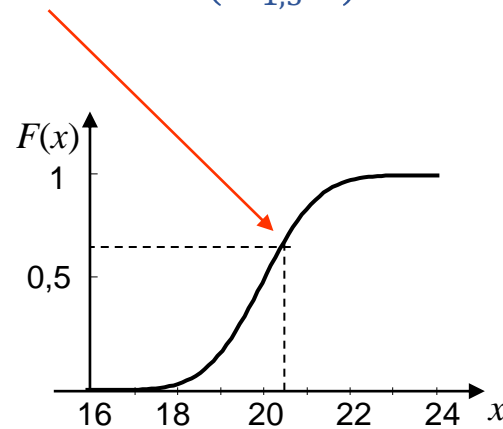
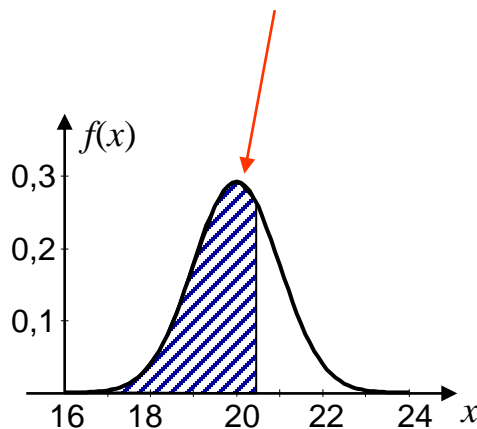


## Przykłady – rozkład normalny

Na podstawie pomiarów długości dużej partii detali wykonywanych na pewnym stanowisku stwierdzono, że rozkład długości jest rozkładem  $\mathcal{N}(20, 1,5)$ . Obliczyć prawdopodobieństwo, że długość losowo wybranego detalu:

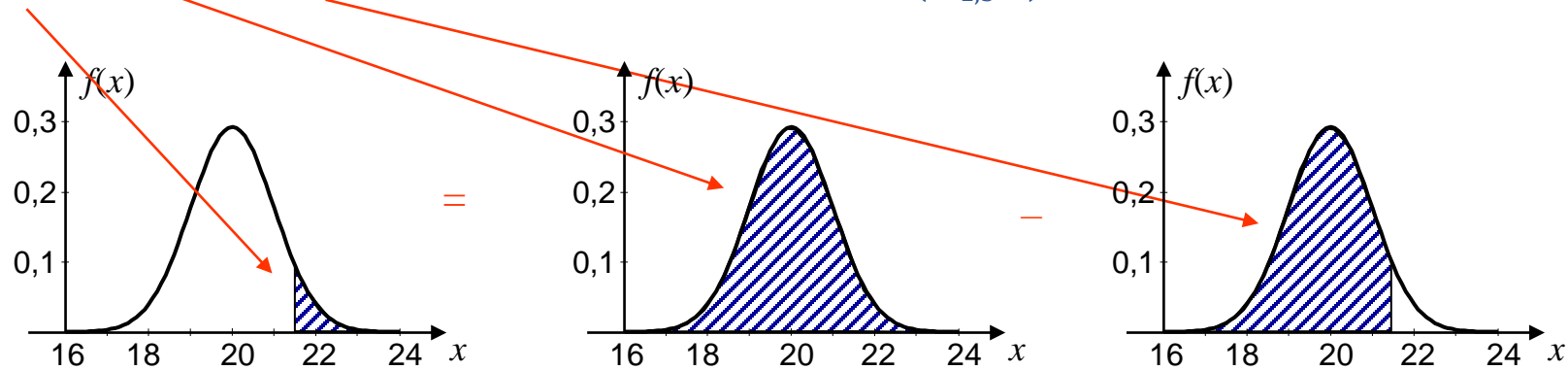
- jest mniejsza lub równa 20,5,
- jest większa od 21,5,
- mieści się w przedziale (20,5 21,5],
- co najmniej o 2 jednostki różni się od średniej,
- obliczyć odchylenie od średniej dla którego prawdopodobieństwo wystąpienia detali o długości przekraczającej wyznaczone odchylenie wyniesie 0,1.

a) 
$$P(x \leq 20,5) = F_{\mathcal{N}(20,1,5)}(20,5) = \Phi\left(\frac{20,5-20}{1,5}\right) = \Phi(0,3333) = 0,6306$$

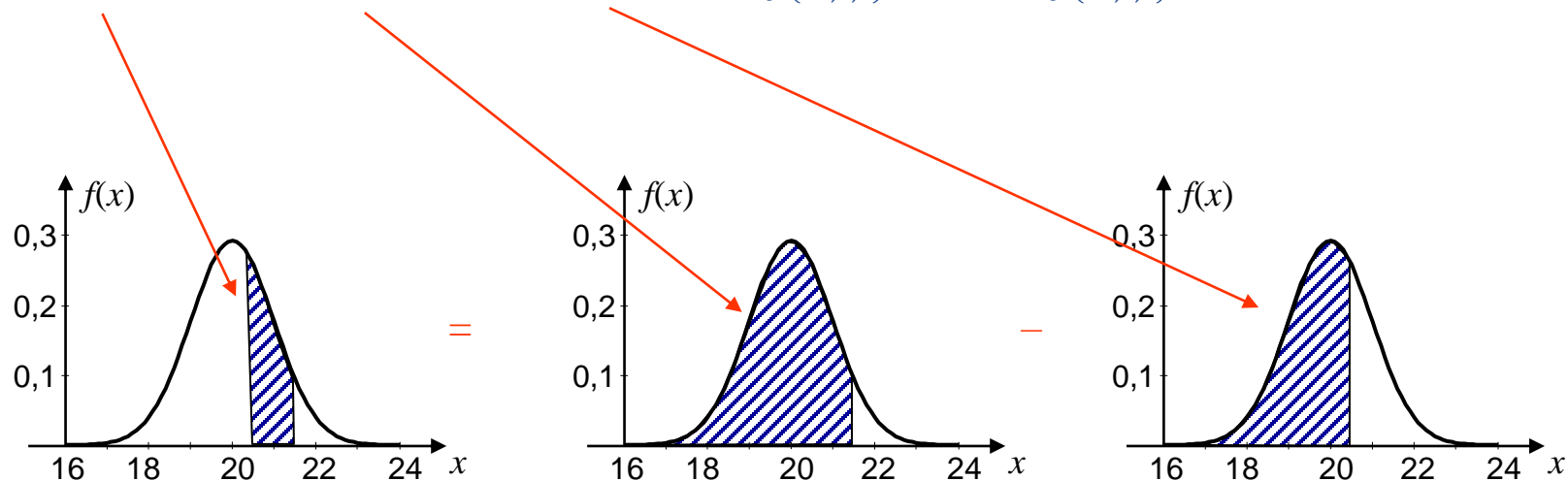


# Przykłady – rozkład normalny

$$b) P(x > 21,5) = 1 - P(x \leq 21,5) = 1 - F_{\mathcal{N}(20,1,5)}(21,5) = 1 - \Phi\left(\frac{21,5-20}{1,5}\right) = 1 - 0,8413 = 0,1587$$

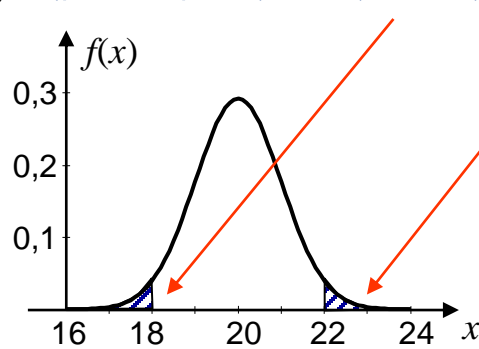


$$c) P(20,5 < x \leq 21,5) = P(x \leq 21,5) - P(x \leq 20,5) = F_{\mathcal{N}(20,1,5)}(21,5) - F_{\mathcal{N}(20,1,5)}(20,5) = 0,2108$$

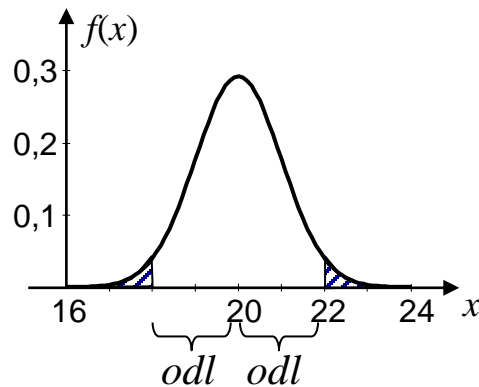


# Przykłady – rozkład normalny

$$d) P(|x - 20| \geq 2) = P(x \leq 18) + P(x \geq 22) = 2P(x \leq 18) = 2F_{\mathcal{N}(20,1,5)}(18) = 2\Phi\left(\frac{18-20}{1,5}\right) = 0,1824$$



$$e) P(|x - 20| \geq odl) = 0,1$$



$$P(|x - 20| \geq odl) = 0,1$$

$$P(|x - 20| \geq odl) = 2P(x \leq 20 - odl) = 2F_{\mathcal{N}(20,1,5)}(20 - odl)$$

$$\Rightarrow F_{\mathcal{N}(20,1,5)}(20 - odl) = 0,05 \Rightarrow$$

$$\Rightarrow 20 - odl = F_{\mathcal{N}(20,1,5)}^{-1}(0,05) \Rightarrow odl = 20 - F_{\mathcal{N}(20,1,5)}^{-1}(0,05) \Rightarrow odl = 2,4673$$

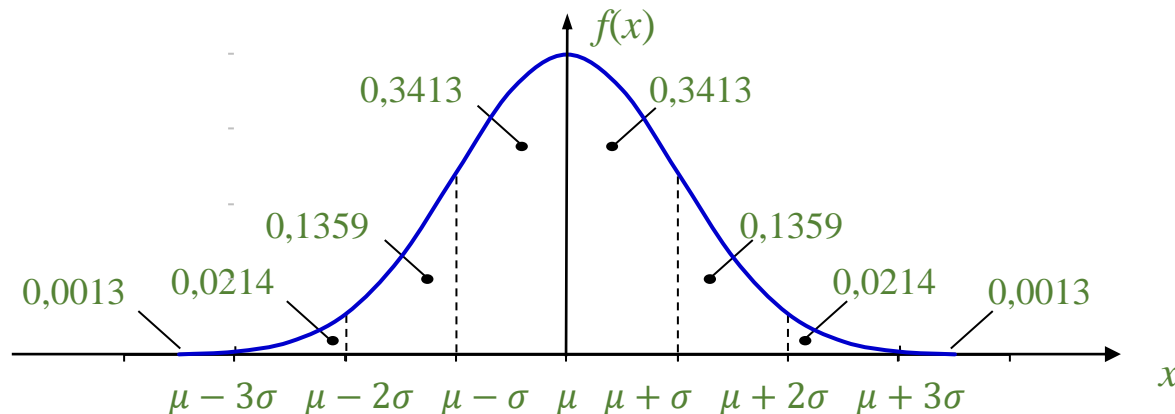
Dla rozkładu normalnego  $\mathcal{N}(\mu, \sigma)$  wyniki odległe od wartości średniej o

$\sigma$  – wypadają w ok. 68% przypadków,

$2\sigma$  – wypadają w ok. 95% przypadków,

$3\sigma$  – wypadają w ok. 99% przypadków (czyli praktycznie wszystkie).

**Prawo 3 sigma:** Prawie wszystkie wartości zmiennej losowej o rozkładzie normalnym mieszczą się w przedziale  $[\mu - 3\sigma, \mu + 3\sigma]$ .



$$P(\mu - \sigma \leq x \leq \mu + \sigma) = P(x \leq \mu + \sigma) - P(x \leq \mu - \sigma) = F(\mu + \sigma) - F(\mu - \sigma) = \Phi(1) - \Phi(-1) = 0,6827$$

$$P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = P(x \leq \mu + 2\sigma) - P(x \leq \mu - 2\sigma) = F(\mu + 2\sigma) - F(\mu - 2\sigma) = \Phi(2) - \Phi(-2) = 0,9545$$

$$P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) = P(x \leq \mu + 3\sigma) - P(x \leq \mu - 3\sigma) = F(\mu + 3\sigma) - F(\mu - 3\sigma) = \Phi(3) - \Phi(-3) = 0,9973$$

W rzeczywistości wiele wielkości losowych ma w przybliżeniu *rozkład normalny* – rozkład ten ma bardzo duże znaczenie w statystyce i w zastosowaniach praktycznych. Uzasadnieniem powszechności występowania rozkładów zbliżonych do normalnego jest *centralne twierdzenie graniczne*.

Jeżeli  $X_1, X_2, \dots, X_n$  są niezależnymi zmiennymi losowymi o jednakowym rozkładzie o wartości oczekiwanej  $\mu$  i wariancji  $\sigma^2$  to dla  $n \rightarrow \infty$  zmienna losowa:

$$\frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

ma w przybliżeniu rozkład  $\mathcal{N}(0, 1)$ .

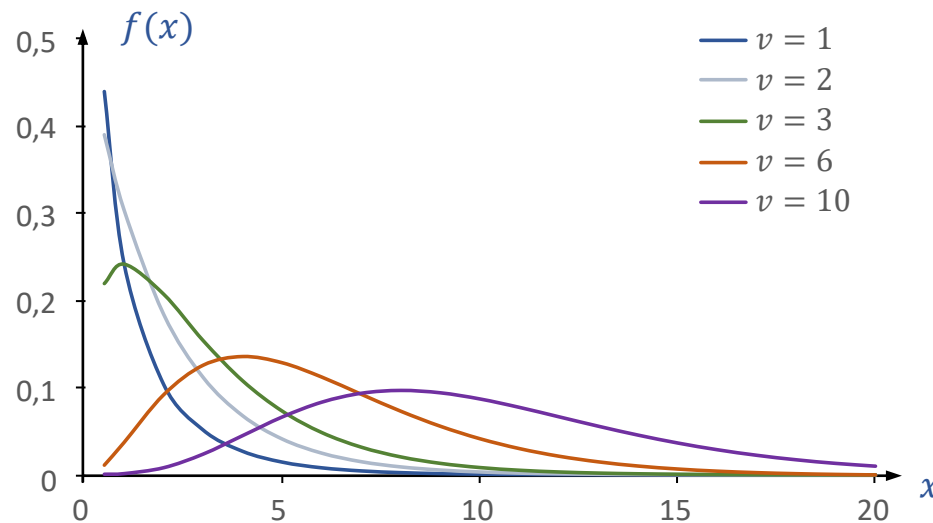
W konsekwencji, dla  $n \rightarrow \infty$  rozkłady poniższych zmiennych losowych są zbieżne do rozkładu normalnego:

- $X_1 + X_2 + \dots + X_n \rightarrow \mathcal{N}(n\mu, \sigma\sqrt{n})$
- $\frac{1}{n}(X_1 + X_2 + \dots + X_n) \rightarrow \mathcal{N}(\mu, \frac{\sigma}{\sqrt{n}})$

**Rozkład  $\chi^2$  (chi kwadrat).** Zmienną o rozkładzie  $\chi^2$  o  $n$  stopniach swobody nazywana jest zmienna zdefiniowana w postaci sumy kwadratów  $n$  niezależnych zmiennych o rozkładzie normalnym standaryzowanym:

$$\chi^2 = Z_1^2 + Z_2^2 + \dots + Z_n^2$$

gdzie:  $Z_1, Z_2, \dots, Z_n$  – zmienne o rozkładzie  $\mathcal{N}(0, 1)$ ,  $n$  – liczba zmiennych niezależnych  $Z_i$  w sumie, parametr rozkładu (jedyne) nazywany **liczbą stopni swobody**, **liczba stopni swobody** oznaczana jest także symbolem  $\nu$ .



Dla  $\nu \rightarrow \infty$  rozkład  $\chi^2$  o jest zbieżny do rozkładu normalnego.

# Zmienne losowe o rozkładzie $\chi^2$

Zakładając, że:

- $X_1, X_2, \dots, X_n$  są niezależnymi zmiennymi losowymi o rozkładach  $\mathcal{N}(\mu, \sigma)$ ,
- $\bar{X} = \frac{1}{n} \sum X_i$ ,

zmienna losowa:

$$\sum \left( \frac{X_i - \bar{X}}{\sigma} \right)^2$$

ma rozkład  $\chi^2$  o  $v = (n - 1)$  stopniach swobody.

Zmienna ta po przekształceniach, zapisywana jest także w postaci:

$$\sum \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 = \frac{1}{\sigma^2} \sum (X_i - \bar{X})^2 = \frac{ns^2}{\sigma^2}$$

1. niech  $n = 2$  i  $\bar{X} = \frac{X_1 + X_2}{2}$ ,

2.  $\left( \frac{X_1 - \bar{X}}{\sigma} \right)^2 + \left( \frac{X_2 - \bar{X}}{\sigma} \right)^2 =$

$$\left( \frac{X_1 - \frac{1}{2}(X_1 + X_2)}{\sigma} \right)^2 + \left( \frac{X_2 - \frac{1}{2}(X_1 + X_2)}{\sigma} \right)^2 =$$

$$\left( \frac{X_1 - X_2}{2\sigma} \right)^2 + \left( \frac{X_2 - X_1}{2\sigma} \right)^2 = \left( \frac{X_1 - X_2}{\sqrt{2}\sigma} \right)^2$$

3. zmienna  $X_1 - X_2 \sim \mathcal{N}(0, \sqrt{2}\sigma)$

więc:  $\frac{X_1 - X_2}{\sqrt{2}\sigma} \sim \mathcal{N}(0, 1)$

ostatecznie:

$$\left( \frac{X_1 - X_2}{\sqrt{2}\sigma} \right)^2 \sim \chi^2(1) \blacksquare$$

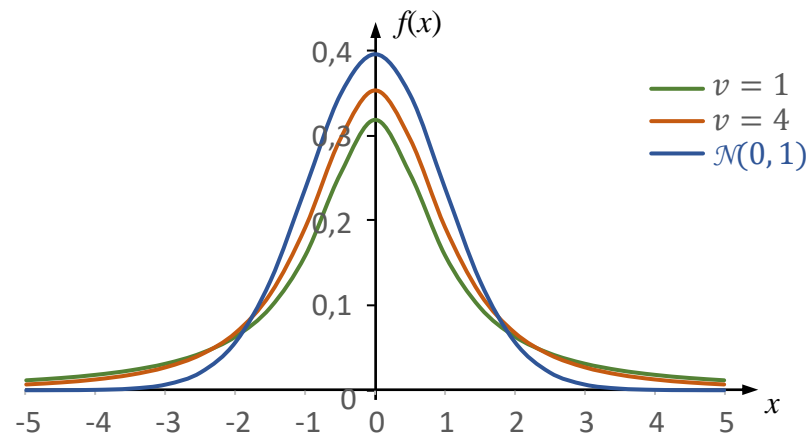


# Rozkład $t$ – Studenta

**Rozkład  $t$  – Studenta.** Zmienną o rozkładzie  $t$  – Studenta o  $n$  stopniach swobody nazywana jest zmienna zdefiniowana w postaci ilorazu zmiennej o rozkładzie normalnym standaryzowanym i zmiennej o rozkładzie  $\chi^2$  o  $n$  stopniach swobody:

$$t = \frac{Z\sqrt{n}}{\sqrt{\chi^2}}$$

gdzie:  $Z$  – zmienna o rozkładzie  $\mathcal{N}(0, 1)$ ,  $\chi^2$  – zmienna o rozkładzie  $\chi^2$  o  $n$  stopniach swobody,  $v = n$  – liczba stopni swobody.



Dla  $v > 30$  rozkład  $t$  – Studenta pokrywa się z rozkładem  $\mathcal{N}(0, 1)$ .

# Zmienne losowe o rozkładzie t – Studenta

Zakładając, że:

- $X_1, X_2, \dots, X_n$  są niezależnymi zmiennymi losowymi o rozkładach  $\mathcal{N}(\mu, \sigma)$ ,
- $\bar{X} = \frac{1}{n} \sum X_i, s^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$ ,

zmienna losowa:

$$\frac{\bar{X} - \mu}{s} \sqrt{n-1}$$

ma rozkład *t – Studenta* o

$v = (n - 1)$  stopniach swobody.

1.  $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1),$

2.  $\frac{ns^2}{\sigma^2} \sim \chi^2(n-1)$

3. zdefiniujemy zmienną o rozkładzie  $t(n-1)$ :

$$t = \frac{Z\sqrt{n-1}}{\sqrt{\chi^2}},$$

$$t = \frac{\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sqrt{n-1}}{\sqrt{\frac{ns^2}{\sigma^2}}} = \frac{\bar{X} - \mu \sqrt{n} \sqrt{n-1} \sigma}{\sqrt{ns}},$$

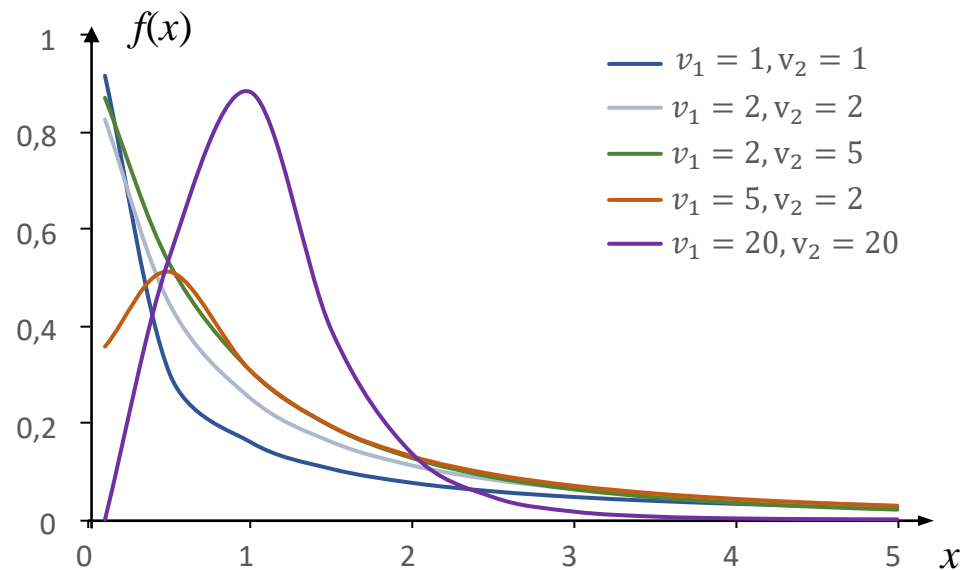
$$t = \frac{(\bar{X} - \mu) \sqrt{n-1}}{s} \blacksquare$$

# Rozkład F Snedecora (Fishera)

**Rozkład F Snedecora (Fishera).** Zmienną o rozkładzie  $F$  i stopniach swobody  $v_1$  i  $v_2$  nazywana jest zmienna zdefiniowana w postaci ilorazu zmiennych o rozkładzie  $\chi^2$ :

$$F = \frac{\chi_1^2}{v_1} / \frac{\chi_2^2}{v_2}$$

gdzie:  $\chi_1^2, \chi_2^2$  – zmienne o rozkładzie  $\chi^2$  z odpowiednio  $v_1$  i  $v_2$  stopniami swobody.



# Statystyka matematyczna – dziedziny zastosowań

---

*Statystyka matematyczna* zajmuje się *wnioskowaniem statystycznym*, tzn. wnioskowaniem o *populacji generalnej* na podstawie znajomości *próby*. Podstawowymi działami statystyki są:

## *teoria estymacji*

zajmuje się wnioskowaniem o własnościach rozkładu prawdopodobieństwa populacji generalnej na podstawie próby,

*estymacja parametryczna* zajmuje się wyznaczaniem (szacowaniem) wartości nieznanymi parametrów rozkładu, *estymacja nieparametryczna* – poszukuje postaci funkcyjnej rozkładu,

szacowanie wartości parametru rozkładu populacji na podstawie próby nazywane jest *estymacją punktową*, *estymacją przedziałową* wyznacza pewien przedział, do którego z określonym prawdopodobieństwem należy szacowana wartość parametru rozkładu,

## *teoria weryfikacji hipotez statystycznych*

zajmuje się tworzeniem reguł umożliwiającymi rozstrzygnięcie o słuszności sądów (hipotez statystycznych),

*testy parametryczne* służą do weryfikacji hipotez o nieznanymi parametrach rozkładu ale znanym samym rozkładzie, *testy nieparametryczne* weryfikują hipotezy w których nie ma założeń o postaci rozkładu.

# Estymacja punktowa – miary położenia

W praktyce rozkład prawdopodobieństwa badanej zmiennej losowej może nie być znany – mogą być mierzone natomiast pewne wielkości wyznaczające przybliżony opis rozkładu. **Miary położenia** stosowane są do oceny miejsca skupienia wyników.

## Średnia arytmetyczna

$$\bar{x} = \frac{1}{n} \sum x_i$$

## Średnia ważona

$$\bar{x} = \frac{\sum w_i x_i}{\sum w_i}$$

## Średnia geometryczna

$$G = \sqrt[n]{\prod x_i}$$

## Średnia harmoniczna

$$H = \frac{n}{\sum \frac{1}{x_i}}$$

gdzie:  $x_i$ ,  $w_i$  –  $i$ -ta wartość badanej cechy i jej waga,  $n$  – liczebność próby.

Wykonano 10 niezależnych pomiarów długości losowo wybranego detalu.

18	21	22,4	23	21,3	21,9	17,6	21	17,8	19,4
----	----	------	----	------	------	------	----	------	------

Średnie wynoszą odpowiednio:

$$\bar{x} = \frac{18 + 21 + \dots + 19,4}{10} = \frac{203,4}{10} = 20,34$$

zakładając, że waga pierwszego i ostatniego pomiaru są równe 2 a pozostałe 1:

$$\bar{x} = \frac{2 \cdot 18 + \dots + 2 \cdot 19,4}{12} = \frac{240,8}{12} \approx 20,07$$

$$G = \sqrt[10]{18 \cdot \dots \cdot 19,4} \approx 20,25$$

$$H = \frac{10}{\frac{1}{18} + \dots + \frac{1}{19,4}} \approx \frac{10}{0,496} \approx 20,16$$

# Estymacja punktowa – miary położenia

**Moda**  $M_0$  (wartość modalna, wartość najczęstsza) wartość najczęściej występująca w próbie.

**Kwantyl rzędu  $p$**  ( $0 < p < 1$ ) to wartość cechy  $x_p$ , która dzieli szereg na 2 części:

- w części 1. znajduje się  $100p$  [%] elementów próbki (wartości tych elementów są mniejsze lub równe *kwantylowi*  $x_p$ ),
- w części 2. znajduje się  $100(1 - p)$  [%] elementów (wartości tych elementów są większe bądź równe *kwantylowi*  $x_p$ ).

**Kwantyle rzędu  $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$**  są nazywane **kwartylami**.

kwantyl	rzęd kwantyla
kwantyl dolny (pierwszy) $Q_1$	$p = \frac{1}{4}$
mediana $Q_2, Me$	$p = \frac{1}{2}$
kwantyl górny (trzeci) $Q_3$	$p = \frac{3}{4}$

Na podstawie wyników pomiarów z poprzedniego przykładu

18	21	22,4	23	21,3	21,9	17,6	21	17,8	19,4
----	----	------	----	------	------	------	----	------	------

można określić *modę*:

$$M_0 = 21.$$

*Kwartyle* można wyznaczyć po uporządkowaniu danych, tzn. po utworzeniu z wyników szeregu prostego:

17,6	17,8	18	19,4	21	21	21,3	21,9	22,4	23
------	------	----	------	----	----	------	------	------	----

$$Q_1 = 18$$

$$Q_2 = M_e = (21 + 21)/2 = 21$$

$$Q_3 = 21,9$$

Jeśli liczebności poszczególnych ćwiartek szeregu prostego są:

- liczbami całkowitymi, to *kwartyle* wyznaczone są jako średnia arytmetyczna
- w przeciwnym przypadku rozdzielają kolejne ćwiartki szeregu

# Estymacja punktowa – miary rozproszenia

Miary rozproszenia (rozrzutu) stosowane są do oceny stopnia rozproszenia wartości badanej cechy.

## Odchylenie standardowe $s$

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}^*, \quad s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

## Wariancja $s^2$

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}, \quad s^2 = \frac{\sum(x_i - \bar{x})^2}{n}$$

**Rozstęp  $r$**  to różnica pomiędzy wartością największą i najmniejszą:

$$r = x_{max} - x_{min}$$

## Rozstęp międzykwartyłowy **IQR**

$$IQR = Q_3 - Q_1$$

gdzie:  $n$  – liczebność próby,  $x_i$  –  $i$ -ta wartość badanej cechy,  $\bar{x}$  – średnia arytmetyczna,  $Q_3$ ,  $Q_1$  – kwartył górny i dolny, \* małe próby.

Miary rozproszenia dla wyników pomiarów:

18	21	22,4	23	21,3	21,9	17,6	21	17,8	19,4
----	----	------	----	------	------	------	----	------	------

wynoszą:

$$s = \sqrt{((18 - 20,34)^2 + \dots)/9} \approx 2,0$$

$$s = \sqrt{((18 - 20,34)^2 + \dots)/10} \approx 1,9$$

$$s^2 = ((18 - 20,34)^2 + \dots)/9 \approx 4,0$$

$$s^2 = ((18 - 20,34)^2 + \dots)/10 \approx 3,6$$

$$r = 23 - 17,6 = 5,4$$

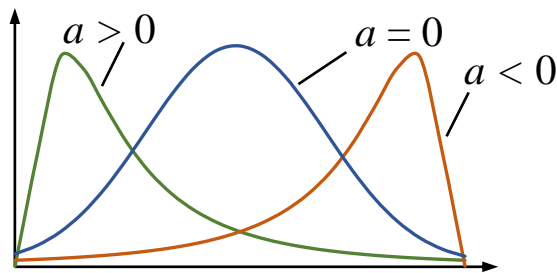
$$IQR = 21,9 - 18 = 3,9$$

# Estymacja punktowa – miary zniekształcenia

Miary zniekształcenia stosowane są do oceny asymetrii i stopnia spłaszczenia rozkładu w stosunku do rozkładu normalnego.

## Współczynnik skośności

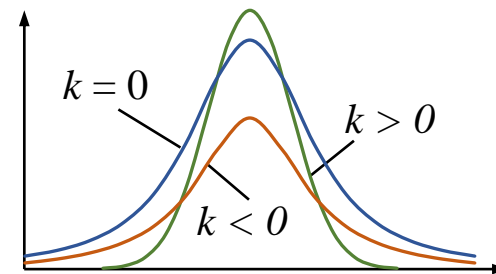
$$a = \frac{M_3}{s^3} = \frac{\frac{1}{n} \sum (x_i - \bar{x})^3}{s^3}$$



$a$	wygląd rozkładu
$= 0$	symetryczny
$> 0$	wydłużona prawa strona
$< 0$	wydłużona lewa strona

## Współczynnik spłaszczenia (kurtoza)

$$k = \frac{M_4}{s^4} = \frac{\frac{1}{n} \sum (x_i - \bar{x})^4}{s^4} - 3$$



$k$	wygląd rozkładu
$= 0$	podobny do r. normalnego
$> 0$	bardziej stromy niż r. normalny
$< 0$	bardziej spłaszczony niż r. normalny

gdzie:  $n$  – liczebność próby,  $x_i$  –  $i$ -ta wartość badanej cechy,  $\bar{x}$  – średnia arytmetyczna,  $s$  – odchylenie standardowe,  $M_3$ ,  $M_4$  – momenty centralne rzędu 3 i 4.

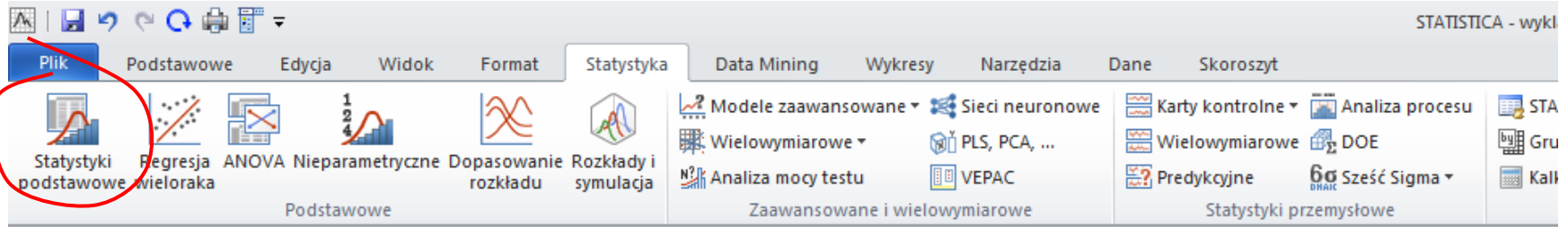


# Estymacja punktowa – miary rozkładu

Opis rozkładu zmiennej uzyskany w wyniku próby jest tylko przybliżonym opisem dla całej populacji.

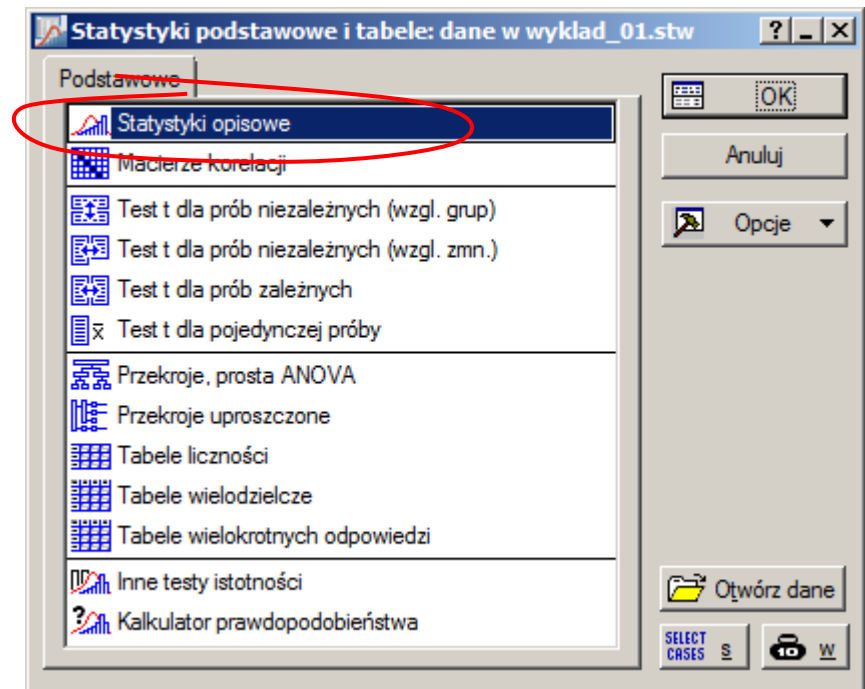
<i>Parametry opisujące własności rozkładu zmiennej</i>	
<i>populacja generalna</i>	<i>próba (estymatory parametrów populacji)</i>
<i>wartość oczekiwana <math>\mu</math></i>	<i>średnia arytmetyczna</i> $\bar{x} = \frac{1}{n} \sum x_i$
<i>odchylenie standardowe <math>\sigma</math></i>	<i>odchylenie standardowe <math>s</math></i> $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$ dla małych prób ( $n < 30$ ) $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

# STATISTICA – estymacja punktowa



The screenshot shows a data table with 10 rows and 2 columns. The first column is labeled '1' and the second column is labeled 'X'. The data values are as follows:

1	X
1	18
2	21
3	22,4
4	23
5	21,3
6	21,9
7	17,6
8	21
9	17,8
10	19,4



# STATISTICA – estymacja punktowa

Statystyki opisowe: dane w wyklad\_01.stw

Zmienne: X Podsumowanie

Podstawowe Więcej Odporne Normalność W. prawd. i rozrzutu W. skategoryzowane Opcje

Statystyki W1 W2 Oblicz statystyki:

**Położenia, N**

- N ważnych
- % ważnych
- Średnia
- Suma
- Mediana
- Moda
- Śr. geometr.
- Śr. harmon.

**Zmienności, momenty**

- Odchylenie standardowe
- PU dla odch. std.
- Przedział: 95,00 %
- Wsp. zmienności
- Wariancja
- Błąd stand. średniej
- Przedz. ufn. średniej
- Przedział: 95,00 %
- Skośność
- Błąd st. skośności
- Kurtozja
- Błąd st. kurtozy

**Percentyle, zakresy**

- Minimum i maksimum
- Dolny i górny kwartyl
- Zakres percentyli
  - Pierwszy: 10,00 % Drugi: 90,00 %
  - Granice
  - Przedziały
- Rozstęp
- Rozstęp kwartyłowy

Wszystkie Od nowa

Zapisz ustawienia jako domyślne

Anuluj

Opcje

Grupami...

SELECT CASES

Momenty ważne

DF =

W-1  N-1

Usuwanie BD

- Przypadkami
- Parami

Dane: Statystyki opisowe (dane w wyklad\_01.stw)

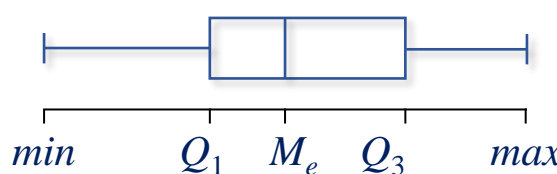
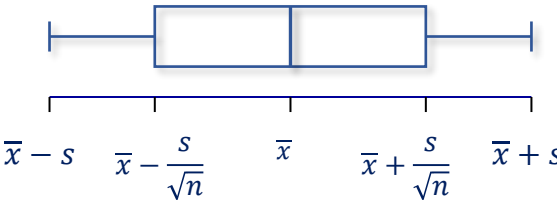
Zmienna	Statystyki opisowe (dane w wyklad_01.stw)										
	Nważnych	Średnia	Geometr. Średnia	Harmon. Średnia	Mediana	Moda	Liczność Mody	Dolny Kwartył.	Górny Kwartył.	Rozstęp	Kwartył. Rozstęp
X	10	20,34000	20,24998	20,15846	21,00000	21,00000	2	18,00000	21,90000	5,400000	3,900000

Wariancja	Odch.std	Skośność	Kurtozja
3,984889	1,996219	-0,310479	-1,53576

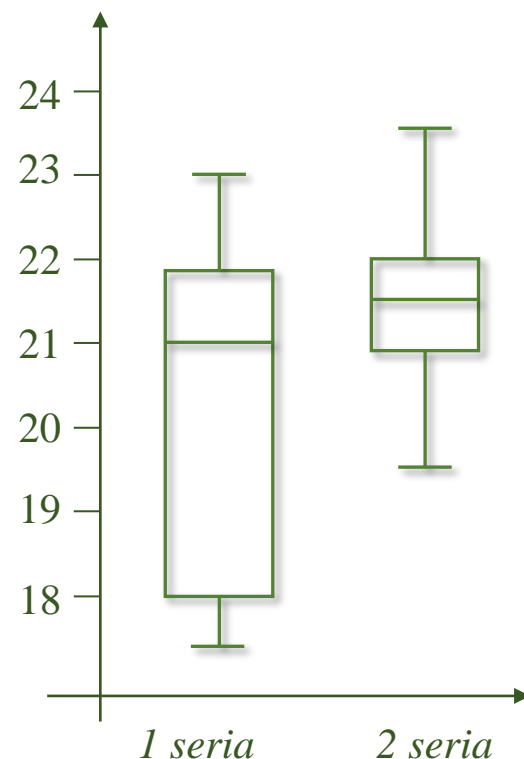
# Graficzna prezentacja danych

**Wykres pudełkowy** – pozwala na ilustrację miar rozkładu jednowymiarowej zmiennej losowej. Jest rysowany w kilku odmianach.

<i>Ilustrowane miary</i>	<i>Wygląd</i>
$min, max$ $Q_1, Me, Q_3$	
$\bar{x}, s$	

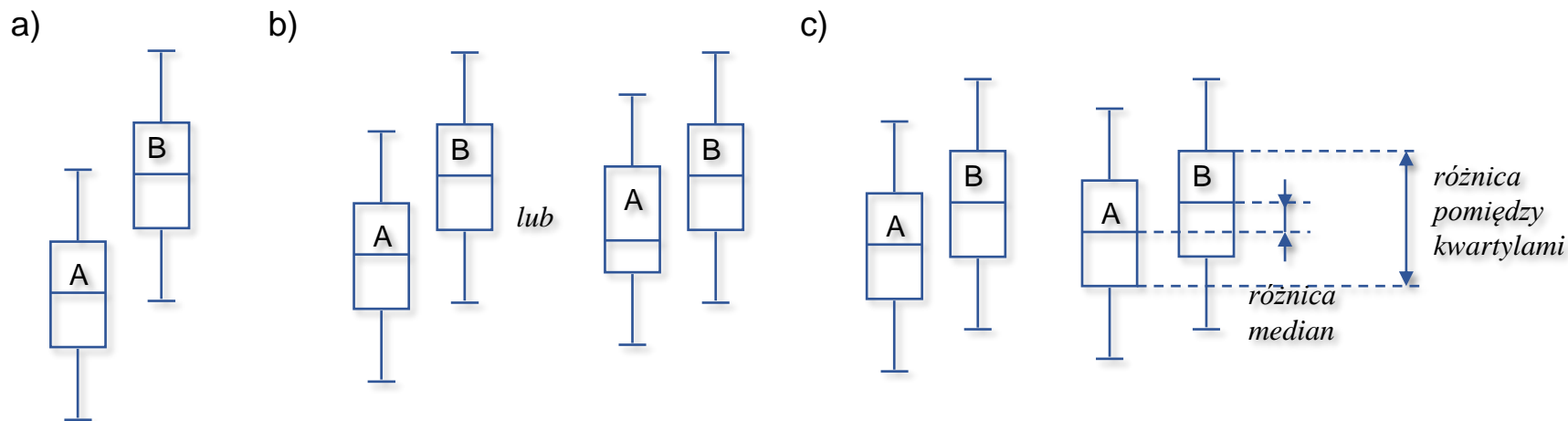
**Wykres pudełkowy** może być rysowany dla wielu grup danych w celu ich porównania.

Na wykresie przedstawiony zostały **wykres pudełkowy** przedstawiający dwie serie pomiarów długości pewnego detalu mierzone dwoma różnymi przyrządami.



# Graficzna prezentacja danych

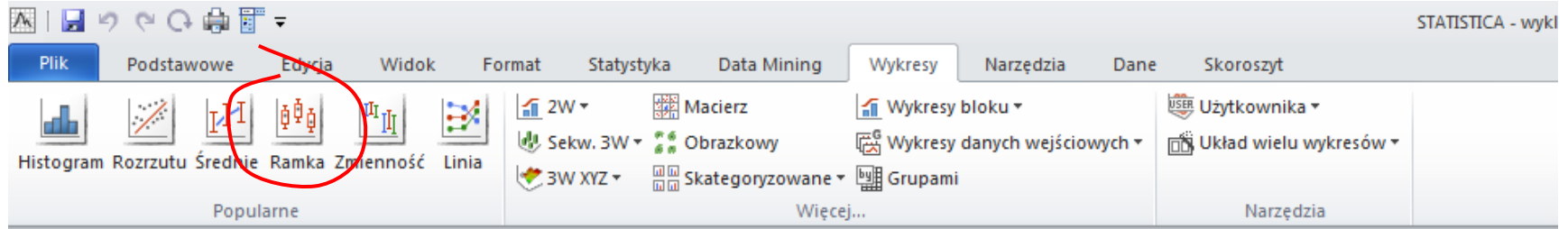
Poniżej rozważone zostały typowe układy dwóch wykresów pudełkowych.



- a) prostokąty wykresów nie nachodzą na siebie – w tej sytuacji uznaje się że **B jest większe od A**,
- b) prostokąty nachodzą na siebie ale jedna z median jest poza prostokątem drugiego wykresu – w tej sytuacji wyciągany jest wniosek, że **B jest prawdopodobnie większe od A**,
- c) prostokąty nachodzą na siebie, obydwie mediany są wewnątrz prostokątów sąsiedniego wykresu – w tej sytuacji **nie można stwierdzić, że istnieje różnica pomiędzy A i B**.

W przypadku c) można wyznaczyć iloraz różnicy median do różnicy pomiędzy dalej położonym dolnym i górnym kwartylem. Jeśli wielkość tak wyznaczonego wskaźnika przekracza pewną wartość graniczną uznaje się że jest **tendencja do różnicy pomiędzy A i B**, za wartość graniczną uznaje się wielkość: **0,33** dla próbek o rozmiarze **30**, **0,2** dla próbek o rozmiarze **100**, **0,1** dla próbek o rozmiarze **1000**.

# STATISTICA – graficzna prezentacja danych



wykład\_01.stw - dane2

wykład\_01.stw<

- dane1
- dane2**
- dane3

	1	2
	nr	X
1	1	18
2	1	21
3	1	22,4
4	1	23
5	1	21,3
6	1	21,9
7	1	17,6
8	1	21
9	1	17,8
10	1	19,4
11	2	19,5
12	2	20
13	2	20,9
14	2	21,2
15	2	21,6
16	2	21,6
17	2	22
18	2	22
19	2	22,6
20	2	23,6

dane1 dane2 dane3

wykład\_01.stw\* - dane3

wykład\_01.stw<

- dane1
- dane2
- dane3**

	1	2
	X1	X2
1	18	19,5
2	21	20
3	22,4	20,9
4	23	21,2
5	21,3	21,6
6	21,9	21,6
7	17,6	22
8	21	22
9	17,8	22,6
10	19,4	23,6

dane1 dane2 dane3

# STATISTICA – graficzna prezentacja danych

Wykresy ramka-wąsy 2W

Podstawowe Więcej Wygląd Skategoryzowane Opcje 1 Opcje 2

Rodzaj wykresu:

Ramka-wąsy Zwykły Zmienne:

Maks-min-zamkn. Wielokrotny

Grupująca: nr

Zależna: X

Przedziały

Tryb całkowity  Automatyczny

Unikalne wartości

Niesort.  Ros.  Mal.

Kategorie: 10

Kody: brak

Maks. nieodstających

75%

Mediana

25%

Min. nieodstających

Punkt środkowy

Wartość: Mediana

Styl: Punkt

Wariancja wspólna

OK

Anuluj

Opcje

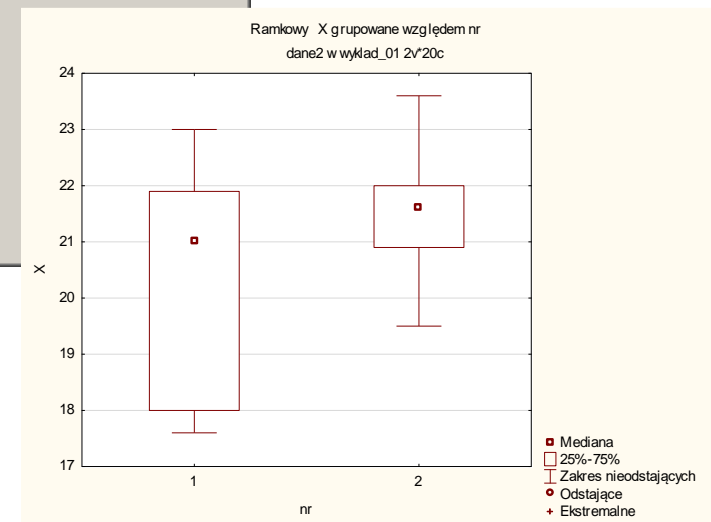
Grupami

SELECT CASES Warunki selekcji

Wagi przypadków

Galeria wykresów

Aktualizuj: Auto



# STATISTICA – graficzna prezentacja danych

Wykresy ramka-wąsy 2W

Podstawowe | Więcej | Wygląd | Skategoryzowane | Opcje 1 | Opcje 2

Rodzaj wykresu:

Ramka-wąsy | Zwykły | Zmienne: | Maks-min-zamkn. | Wielokrotny | Grupująca: brak | Zależna: WSZYSTKIE

Przedziały

Tryb całkowity  Automatyczny

Unikalne wartości

Niesort.  Ros.  Mal.

Kategorie: 10

Kody: brak

Maks. nieodstających 75%

Mediana

25%

Min. nieodstających

Punkt środkowy

Wartość: Mediana

Styl: Linia

Wariancja wspólna

OK

Anuluj

Opcje

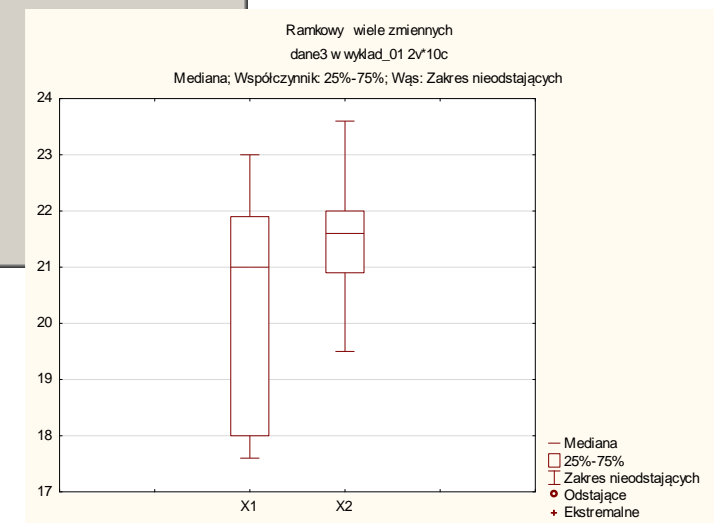
Grupami

Warunki selekcji

Wagi przypadków

Galeria wykresów

Aktualizuj: Auto





# Graficzna prezentacja danych

**Histogram** – wykres słupkowy ilustrujący rozkład prawdopodobieństwa określonej cechy.

W przypadku małych prób wykres jest budowany w oparciu o kolejne wartości odpowiedniego szeregu prostego, dla większych prób konstruowany na podstawie danych szeregu rozdzielczego.

Podstawę słupków stanowią wartości lub przedziały klasowe szeregu. Wysokość słupków jest ustalana dla określonej wartości lub przedziału na podstawie:

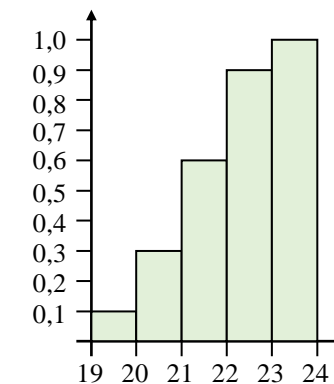
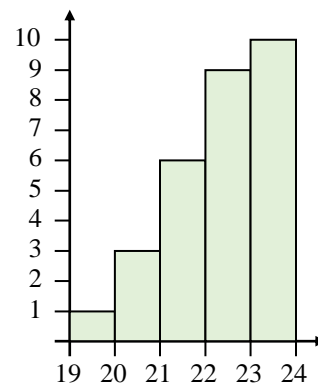
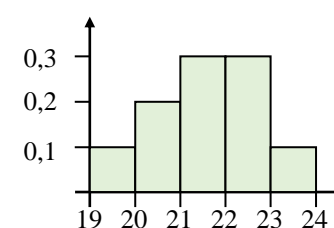
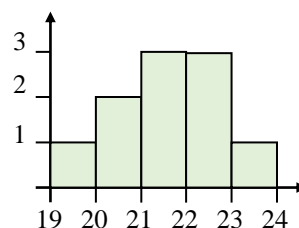
- **liczebności** (histogram liczebności)
- **częstości** (histogram częstości)
- **skumulowanych liczebności** (histogram liczebności skumulowanych)
- **skumulowanych częstości** (histogram częstości skumulowanych).

Wyniki pomiarów długości pewnego detalu:

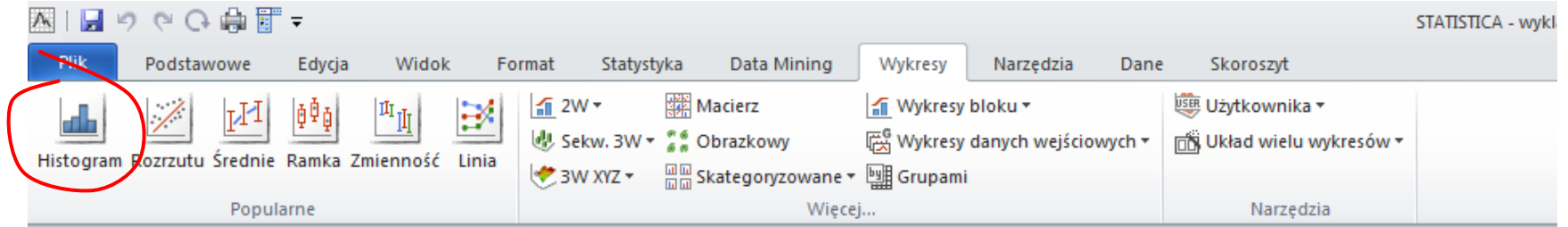
19,5	20	20,9	21,2	21,6	21,6	22	22	22,6	24
------	----	------	------	------	------	----	----	------	----

Szereg rozdzielczy:

	[19,20)	[20,21)	[21,22)	[22,23)	[23,24)
<i>liczebności</i>	1	2	3	3	1
<i>częstości</i>	0,1	0,2	0,3	0,3	0,1
<i>l. skumul.</i>	1	3	6	9	10
<i>c. skumul.</i>	0,1	0,3	0,6	0,9	1,0



# STATISTICA – graficzna prezentacja danych



wykład\_01.stw - dane4

	1
	X
1	19,5
2	20
3	20,9
4	21,2
5	21,6
6	21,6
7	22
8	22
9	22,6
10	23,6

Histogramy 2W

Podstawowe Więcej Wygląd Skategoryzowane Opcje 1 Opcje 2

Zmienne Zmienna: X

Rodzaj wykresu: Zwykły

Dopasuj: Wylaczone, Normalny, Beta, Wykładniczy, Ekstremalny, Gamma, Geometryczny

Sposób wyświetlania: Standardowe

Przedziały: Zmienna: X, Tryb całkowity, Automatyczny, Unikalne wartości, Niesort., Ros., Mal., Kategorie: 10, Granice: 19/1/24, Kody: brak, Podzbiory wielowarunkowe, Określ granice

Statystyki: Test Shapiro Wilka, Całkowita licznosc, Statystyki opisowe, Test Kolmogorowa-Smimowa

Wykres Pareto, Odstępy między kolumnami, Pokaż procenty, Pokaż liczby, Oś Y: N

OK, Anuluj, Opcje, Grupami, Warunki selekcji, Wagi przypadków, Galeria wykresów, Aktualizuj: Auto

# STATISTICA – graficzna prezentacja danych

**Określ granice dla X** [?] [X]

Określ granice:

Minimum:

Krok:

Maksimum:

Wprowadź granice

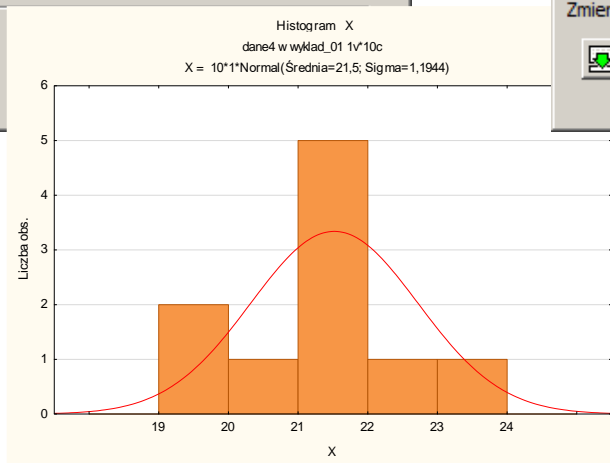
Wpisz wartości stanowiące granice klas na wykresie.

Np. jeżeli wpiszesz trzy poniższe liczby:  
2 2.5 5  
to otrzymasz cztery przedziały:  
(1)  $x < 2$ , (2)  $2 < x < 2.5$ , (3)  $2.5 < x < 5$  i (4)  $5 < x$

Podajemy dolne albo górne granice, w zależności od wyboru poniżej.

Wprowadzane granice:

Zmienne



**Określ granice dla X** [?] [X]

Określ granice:

Minimum:

Krok:

Maksimum:

Wprowadź granice

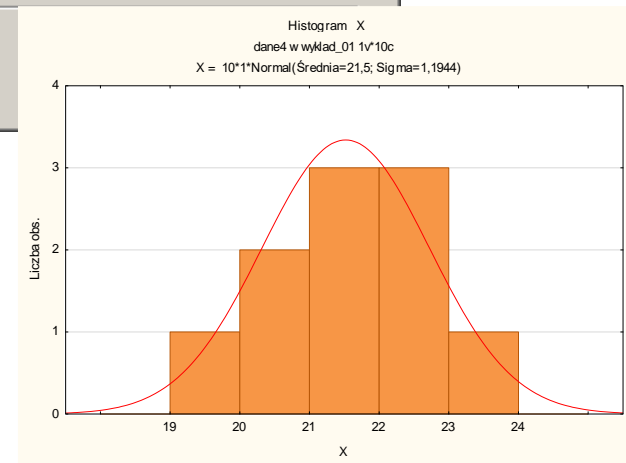
Wpisz wartości stanowiące granice klas na wykresie.

Np. jeżeli wpiszesz trzy poniższe liczby:  
2 2.5 5  
to otrzymasz cztery przedziały:  
(1)  $x < 2$  (2)  $2 < x < 2,5$  (3)  $2,5 < x < 5$  i (4)  $5 < x$

Podajemy dolne albo górne granice, w zależności od wyboru poniżej.

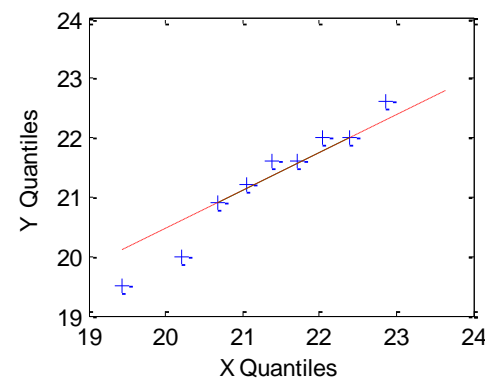
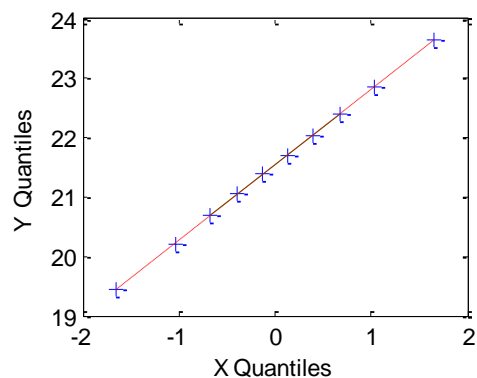
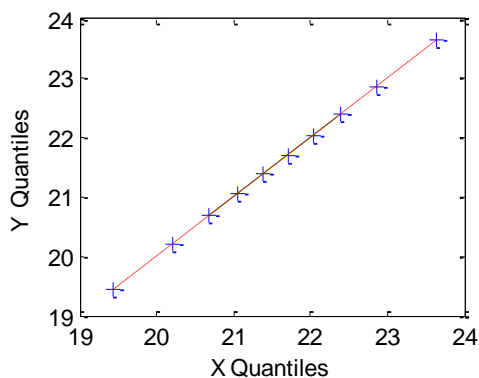
Wprowadzane granice:

Zmienne



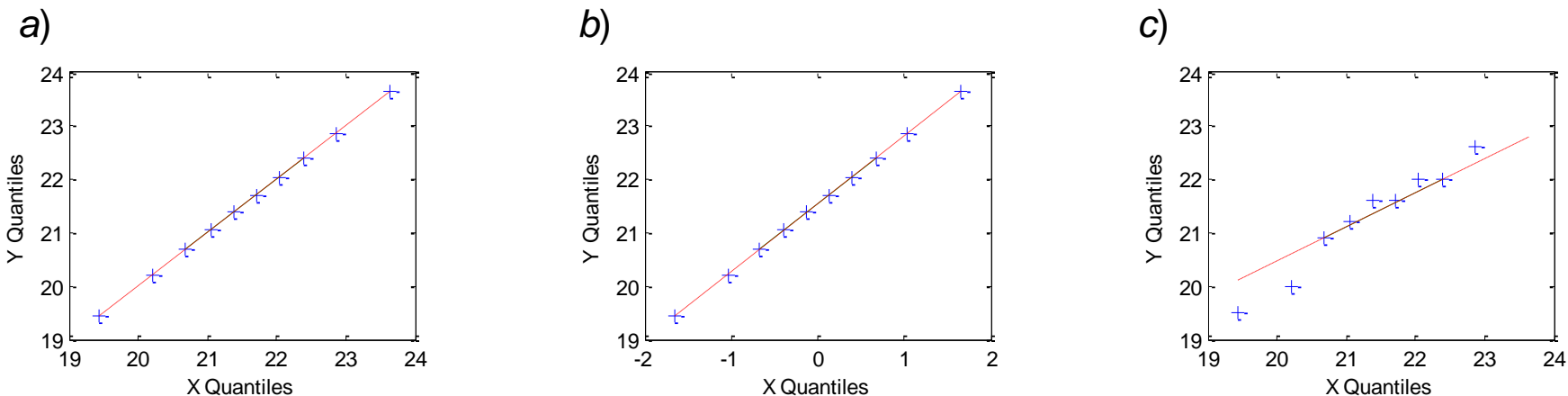
## Wykres kwanty–kwantyl (wykres K-K lub Q-Q)

Wykres  $Q - Q$  umożliwia porównanie kwantyli dwóch zmiennych, jest często wykorzystywany do porównania rozkładu zmiennej empirycznej z teoretycznym rozkładem tej zmiennej. Na osi poziomej odkładane są kwantyle wyznaczone z rozkładu pierwszej zmiennej a na osi pionowej drugiej zmiennej (lub kwantyle wynikające z teoretycznego rozkładu badanej zmiennej).



- porównanie rozkładu kwantyli teoretycznych tego samego rozkładu normalnego,*
- porównanie rozkładu kwantyli teoretycznych dwóch różnych rozkładów normalnych,*
- porównanie rozkładu kwantyli z próby z kwantylami teoretycznymi rozkładu o średniej i odchyleniu standardowym wyznaczonym na podstawie próby.*

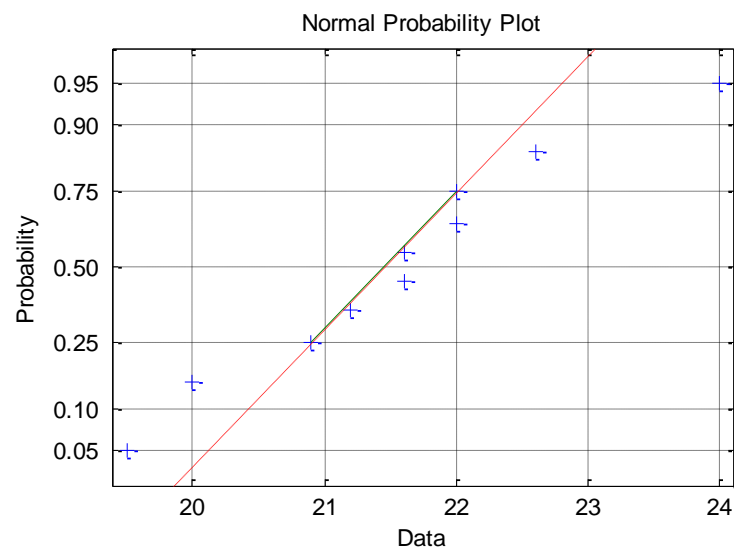
## Wykres kwanty–kwantyl (wykres K-K lub Q-Q)



- Jeżeli kwantyle obydwu rozkładów układają się na wykresie tworząc **linię prostą** to obydwa rozkłady są **idealnie dopasowane** (idealne dopasowanie można uzyskać np. porównując ze sobą kwantyle tego samego rozkładu teoretycznego).
- W przypadku porównywania kwantyli z próby (empirycznych) z kwantylami teoretycznymi punkty na wykresie **tylko bardziej dopasowane są do linii prostej im lepiej rozkład empiryczny jest dopasowany do rozkładu teoretycznego**.
- Wykres  $Q - Q$  jest bardzo często uzupełniany linią prostą pokazującą liniowy trend danych.

## Wykres prawdopodobieństwo-prawdopodobieństwo (wykres $P-P$ )

- Wykres  $P - P$  umożliwia porównanie skumulowanego prawdopodobieństwa rozkładu dwóch zmiennych, jest często wykorzystywany, podobnie jak wykres  $Q - Q$  do porównania rozkładu zmiennej empirycznej z teoretycznym rozkładem tej zmiennej.
- Na osi poziomej odkładane są skumulowane prawdopodobieństwa wyznaczone z próby, na osi pionowej odkładane są skumulowane prawdopodobieństwa rozkładu teoretycznego.
- Podobnie jak w przypadku wykresu  $Q - Q$  punkty na wykresie  $P - P$  są tym bardziej dopasowane są do linii prostej im bardziej rozkłady są zgodne, wykres  $P - P$  jest często uzupełniany liniowym trendem ułatwiającym sprawdzenie zgodności rozkładów.

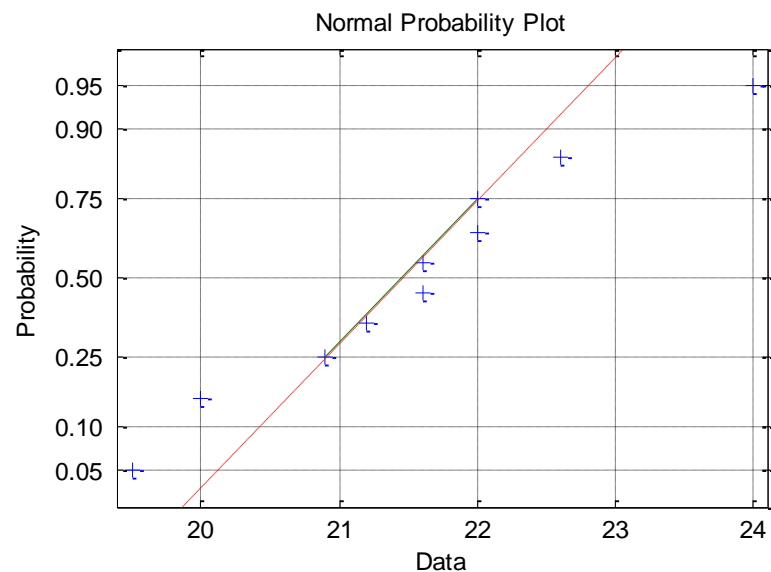


## Wykresy normalności

Wykres porównujący rozkład danych empirycznych z teoretycznym rozkładem normalnym nazywane są **wykresami normalności**.

Jeżeli na wykresach  $Q - Q$  i  $P - P$  porównywany jest rozkład z próby z rozkładem normalnym wykresy te nazywane są **wykresami normalności  $Q - Q$**  i **wykresami normalności  $P - P$** .

Na wykresach normalności  $Q - Q$  oś pionowa, na której odkładane są kwantyle wynikające z teoretycznego rozkładu normalnego badanej zmiennej, może być również opisywana rzędami kwantyli.



# STATISTICA – graficzna prezentacja danych

The screenshot displays the STATISTICA software interface. The 'Wykresy' (Charts) menu is open, showing various chart options. The 'Wykresy kwantyl-kwantyl...' option is highlighted with a red circle. In the background, a data table is visible with the following data:

	1	X
1	19,5	
2	20	
3	20,9	
4	21,2	
5	21,6	
6	21,6	
7	22	
8	22	
9	22,6	
10	23,6	



# STATISTICA – graficzna prezentacja danych

Wykresy kwantyl-quantyl

Podstawowe Więcej Wygląd Skategoryzowane Opcje 1 Opcje 2

Zmienne

Rozkład:

- Normalny
- Beta
- Wykładniczy
- Ekstremalny
- Gamma
- Lognormalny
- Rayleigha
- Weibulla

Dla tego rozkładu nie wymagane są żadne parametry.

Linia dopasowania

Pokaż liniowe

Użytkownika:

Skala: 1

Próg: 0

Poprawki

Rangi: .375

N: .25

Skala prawdopodob. .01 .05 .1 .25 .5 .75 .9 .95 .99

Aby wykreślić kwantyle jednej zmiennej względem drugiej użyj wykresów rozrzutu 2W.

Układ wykresu

Wiele wykresów na jednym rysunku

OK

Anuluj

Opcje

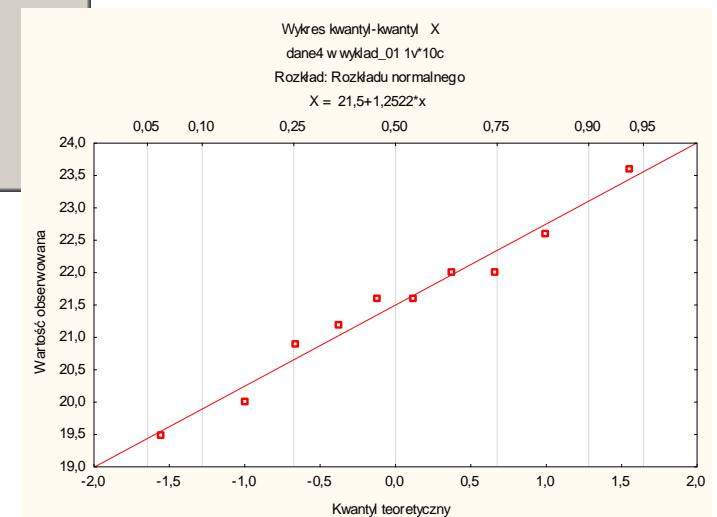
Grupami

Warunki selekcji

Wagi przypadków

Galeria wykresów

Aktualizuj: Auto



# STATISTICA – graficzna prezentacja danych

The screenshot displays the STATISTICA software interface. The main menu bar includes 'Plik', 'Podstawowe', 'Edycja', 'Widok', 'Format', 'Statystyka', 'Data Mining', 'Wykresy', 'Narzędzia', 'Dane', and 'Skoroszyt'. The 'Wykresy' menu is open, showing a list of chart types. The option 'Wykresy prawdopodob.-prawdopodob...' is highlighted with a red circle. The 'Populärne' section of the ribbon shows icons for Histogram, Rozrzut, Średnie, Ramka, Zmienność, and Linia. A data window titled 'wyklad\_01.stw - dane4' is visible, containing a table with 10 rows of data. The 'Narzędzia' section of the ribbon includes 'Użytkownika' and 'Układ wielu wykresów'.

	1
	X
1	19,5
2	20
3	20,9
4	21,2
5	21,6
6	21,6
7	22
8	22
9	22,6
10	23,6

# STATISTICA – graficzna prezentacja danych

