

# Statystyka

wydział medycyny i stomatologii

# Medycyna praktyczna - statystyka

<https://www.mp.pl/artykuly/naukipokrewne/statystyka.html>

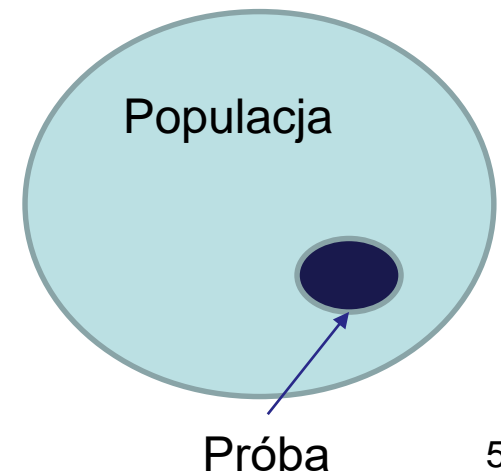
# Spis treści

- Dane .....4
- Dane ilościowe .....7
- Zmienne.....9
- Rozkłady prawdopodobieństw.....13
- Miary położenia .....21
- Miary rozproszenia .....33
- Miary symetrii rozkładu.....40
- Wielkość efektu.....41
- Przedział ufności .....44
- Hipotezy statystyczne.....54
- Testy parametryczne.....68
- Testy nieparametryczne.....85
- Analiza zależności.....102
- Ryzyko i szansa .....112
- Dane niemierzalne.....118
- Materiały uzupełniające.....135

Dane

# Dane

- Charakteryzują obserwacje jednej lub wielu zmiennych
- Są uzyskiwane z **próby**, reprezentującej **populację**
- Mają różne formy (postacie)



# Dane

**Populacja generalna** - zbiór elementów mający przynajmniej jedną właściwość wspólną dla wszystkich jego elementów kwalifikującą je do tego zbioru oraz przynajmniej jedną właściwość, ze względu na którą elementy tego zbioru mogą się różnić między sobą

# Dane ilościowe

**Dane ilościowe** – cecha będąca zmienną przyjmuje wartości liczbowe wyrażone w jednostkach na określonej skali

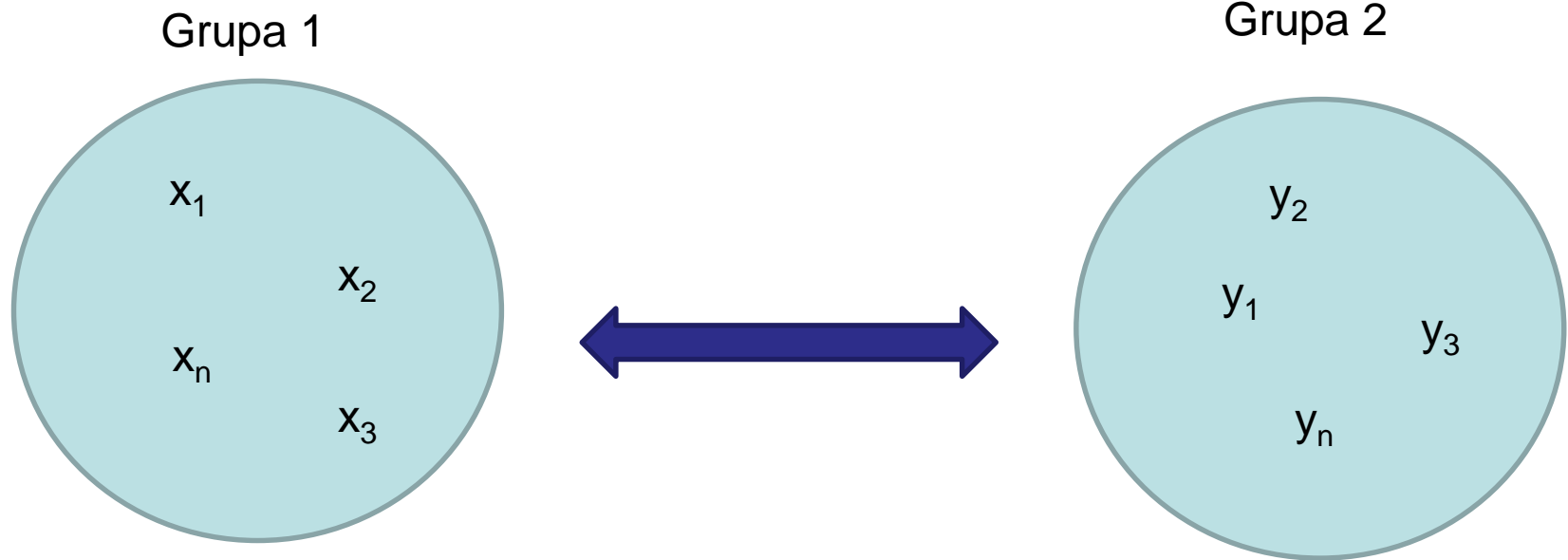
# Dane ilościowe

- **Dane ciągłe** – cecha zmienna przyjmuje dowolną wartość liczbową
- **Dane nieciągłe** – cecha zmienna przyjmuje tylko określone wartości np. całkowite
- **Proporcje (ilorazy)** – gdy taka forma danych jest czytelniejszym opisem cech badanego np.  $BMI = \text{masa} / \text{wzrost}^2$



# Zmienne

# Zmienne niepowiązane



*np. podajemy dwa różne leki obniżające ciśnienie tętnicze krwi dwóm grupom pacjentów i patrzymy na poziom obniżenia ciśnienia*

# Zmienne powiązane



*Najczęściej sytuacje „przed – po”, np. pomiar ciśnienia tętniczego przed i po podaniu leku, itp.*

# Wartości odstające

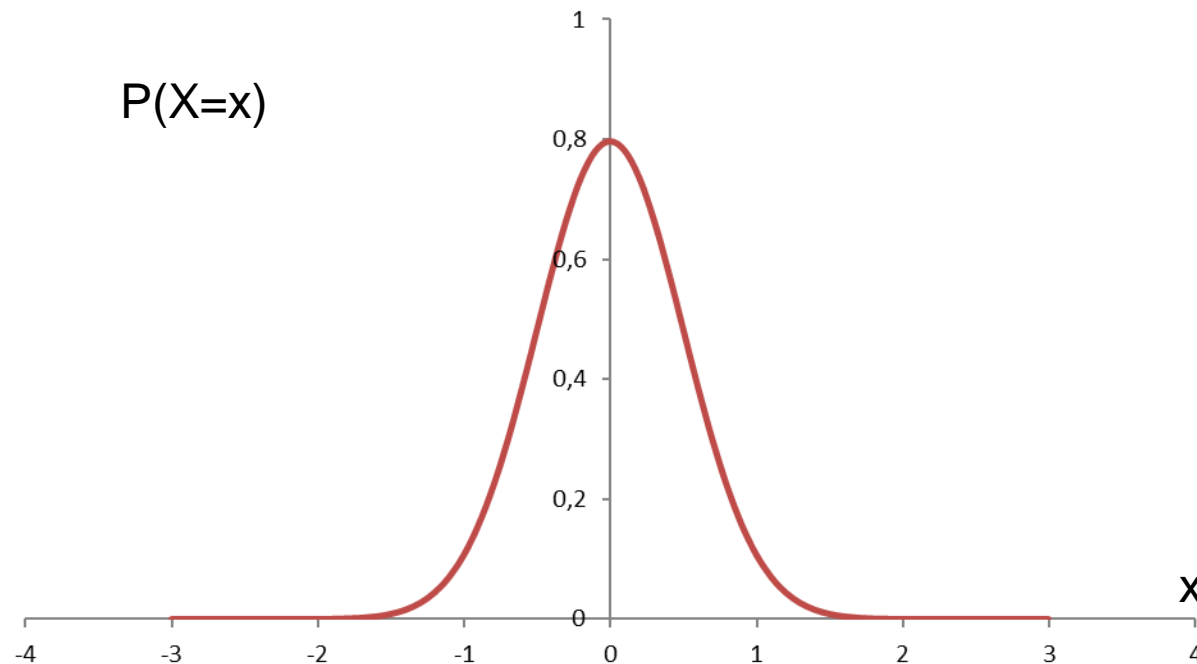
Znacznie różnią się od większości danych i są niezgodne z pozostałymi danymi.

Mogą być prawdziwe ale także mogą być wynikiem błędnego pomiaru np. kobieta o wzroście 204 cm. Przed analizą takich danych należy potwierdzić ich rzetelność sprawdzając dane źródłowe.

# Rozkłady prawdopodobieństw

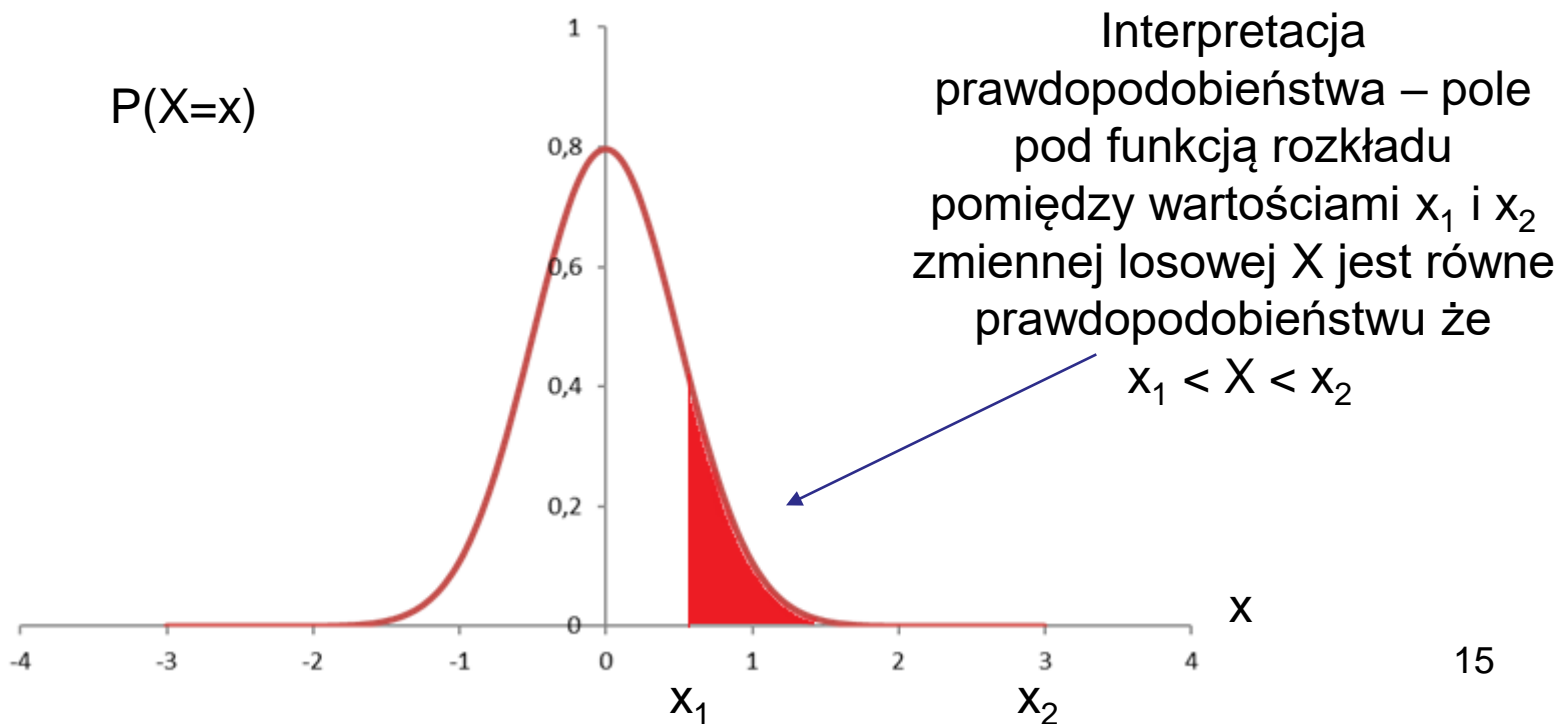
# Rozkłady prawdopodobieństw

Rozkład prawdopodobieństwa  $P(X=x)$  danej zmiennej losowej jest to funkcja przyporządkowująca wartościom zmiennej losowej  $X$  prawdopodobieństwa  $P$  przyjęcia danej wartości  $x$  przez tę zmienną ( $X=x$ ).



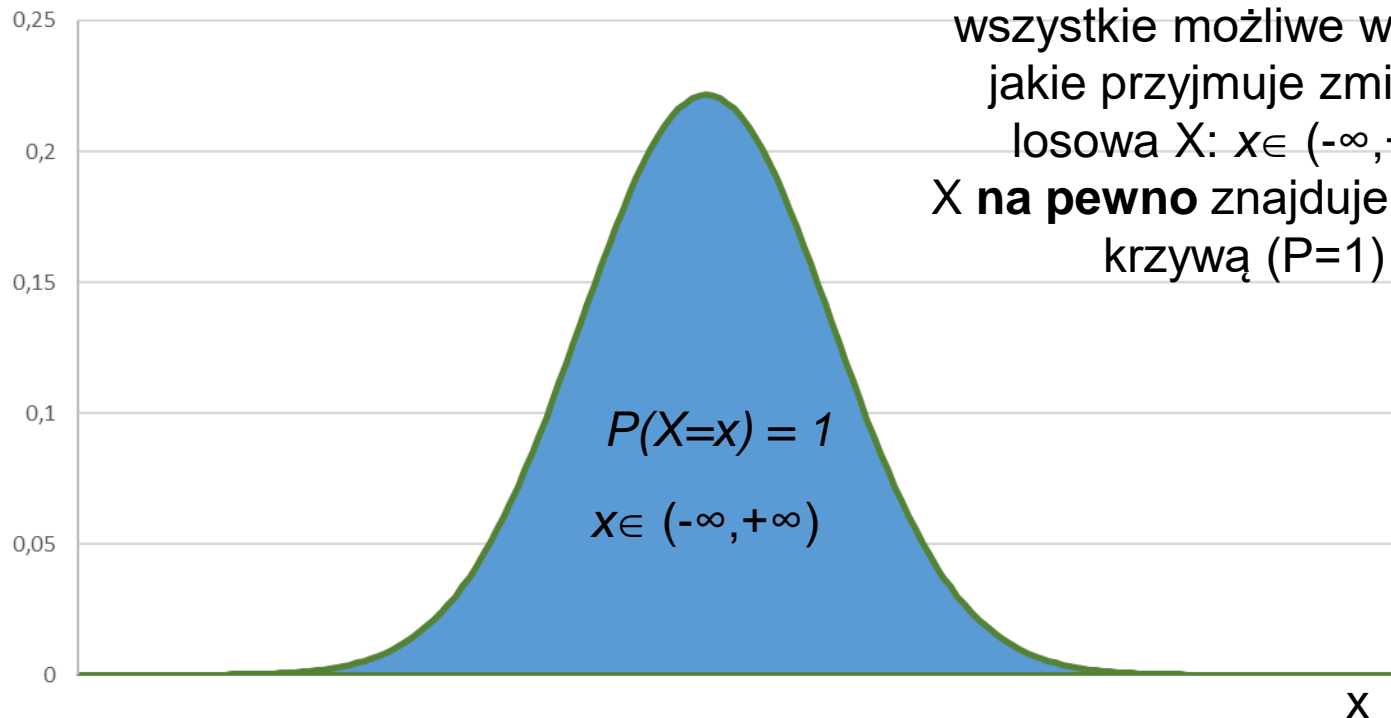
# Rozkłady prawdopodobieństw

Rozkład prawdopodobieństwa  $P(X=x)$  danej zmiennej losowej jest to funkcja przyporządkowująca wartościom zmiennej losowej  $X$  prawdopodobieństwa  $P$  przyjęcia danej wartości  $x$  przez tę zmienną ( $X=x$ ).



# Rozkłady prawdopodobieństw

$P(X=x)$



Całkowite pole powierzchni pod krzywą rozkładu musi być równe 1, bo zakres wartości  $x$  pod krzywą obejmuje wszystkie możliwe wartości jakie przyjmuje zmienna losowa  $X$ :  $x \in (-\infty, +\infty)$ .  
 **$X$  na pewno znajduje się pod krzywą ( $P=1$ )**

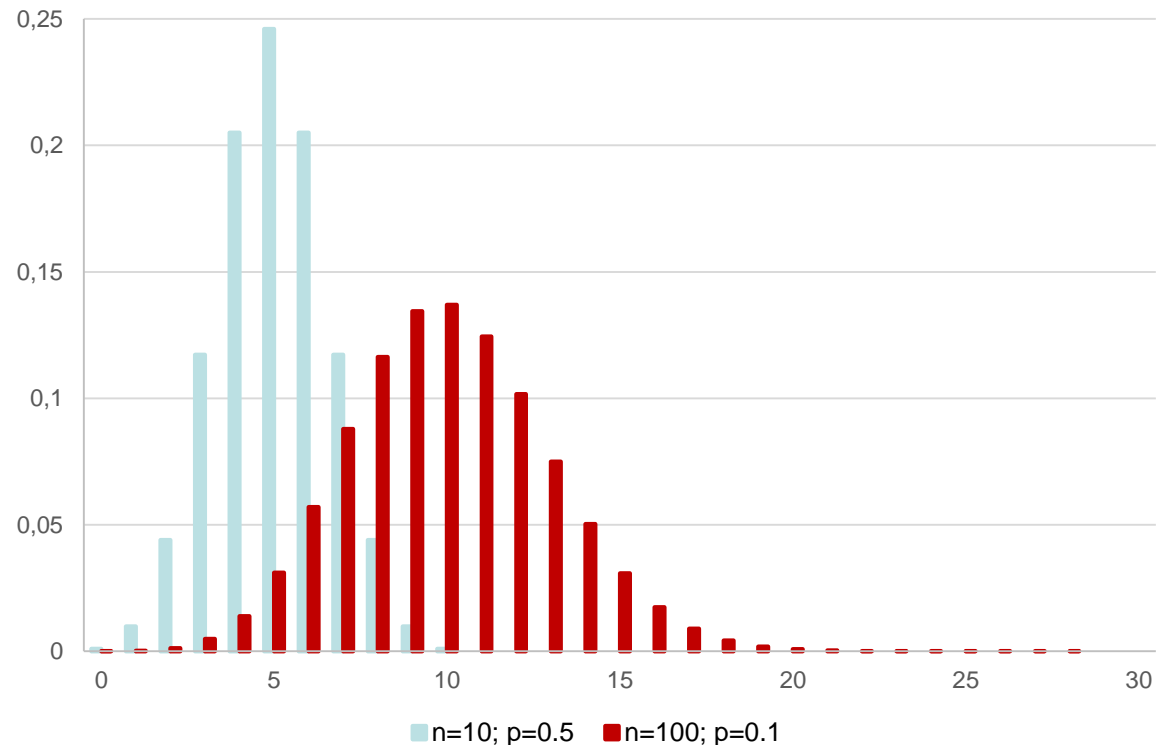


# Rozkłady prawdopodobieństw

**Dwumianowy** - dyskretny rozkład prawdopodobieństwa opisujący liczbę sukcesów  $k$  w ciągu  $N$  niezależnych prób, z których każda ma stałe prawdopodobieństwo sukcesu równe  $p$ .

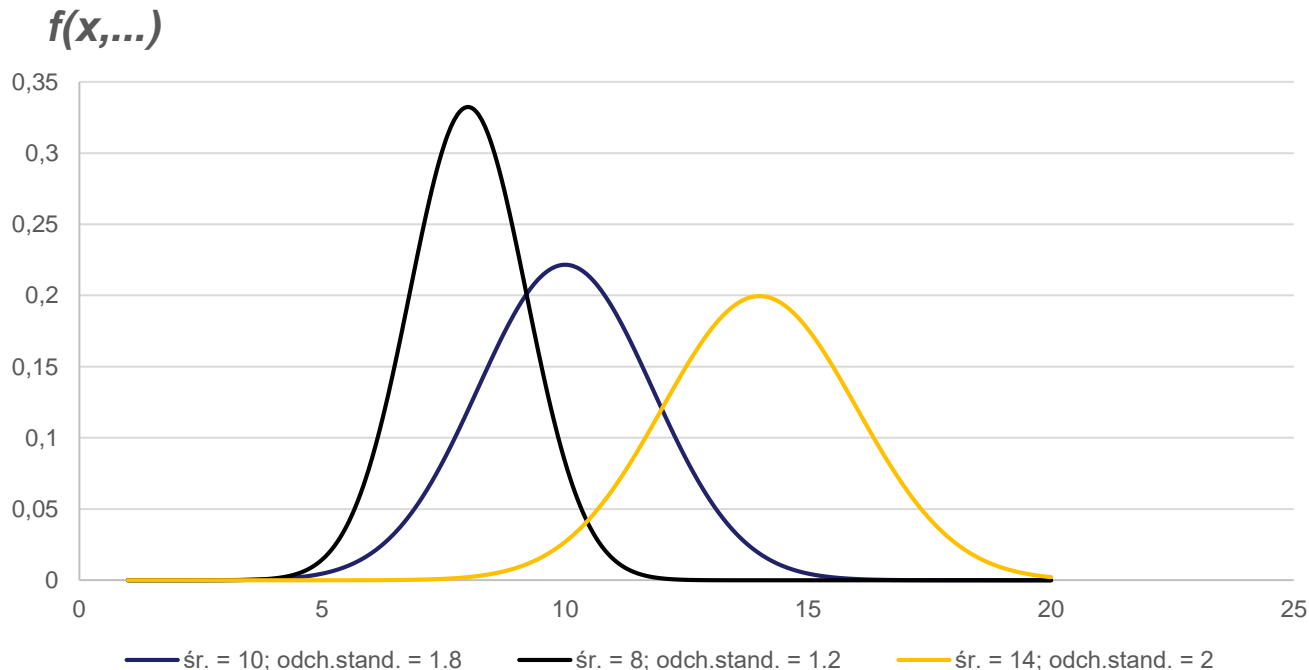


Rzut monetą  
jest bardzo  
dobrym  
przykładem

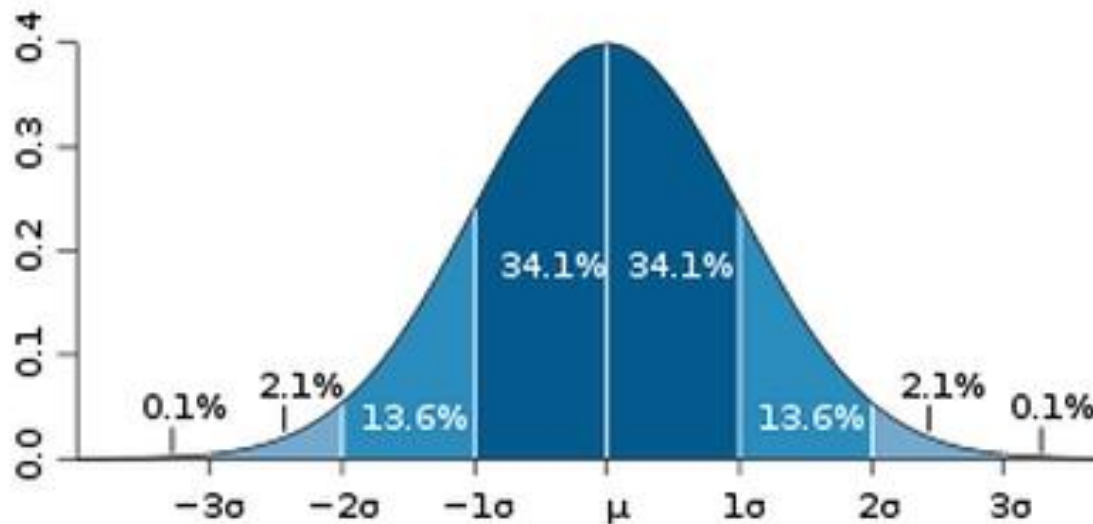


# Rozkłady prawdopodobieństw

**Rozkład normalny** - najbardziej intuicyjny symetryczny rozkład statystyczny. Opisuje on sytuacje w świecie, gdzie większość przypadków jest bliska średniemu wynikowi, a im dany wynik bardziej odchyła się od średniej tym jest mniej reprezentowany.



# Rozkłady prawdopodobieństw (zasada dwóch odchyłeń standardowych dla rozkładu normalnego)



Przedział	Procent obserwacji
$[\mu-\sigma, \mu+\sigma]$	68.2%
$[\mu-2\sigma, \mu+2\sigma]$	<b>95.4%</b> → ok. 5% wartości poza przedziałem
$[\mu-3\sigma, \mu+3\sigma]$	99.7%

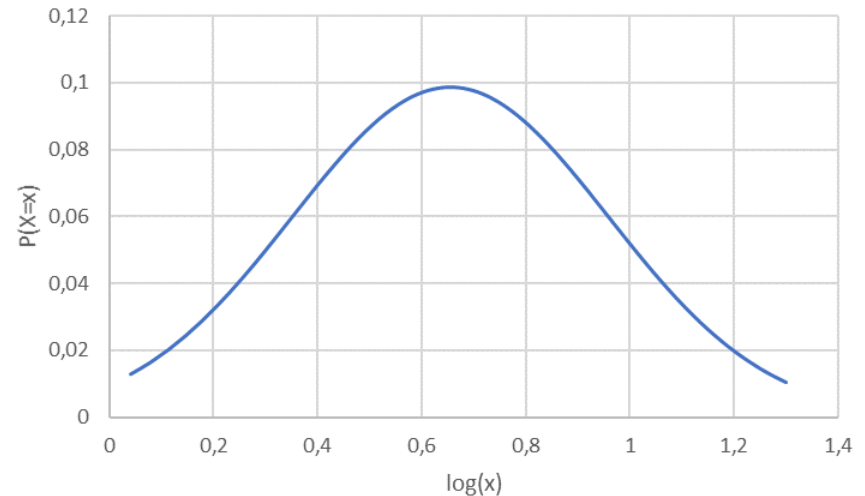
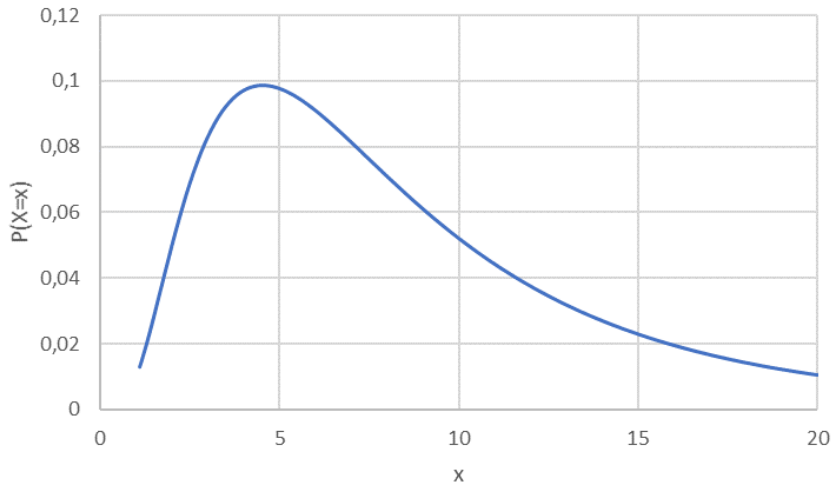
Przedział  $[\mu-2\sigma, \mu+2\sigma]$  najczęściej stosowany w naukach medycznych

# Rozkład log-normalny

ciągły rozkład prawdopodobieństwa dodatniej zmiennej losowej, której logarytm ma rozkład normalny

*(rozkład logarytmicznie normalny jest często lepszym od rozkładu normalnego przybliżeniem rozkładów cech, w których istotne są stosunki pomiędzy wartościami, a nie różnice pomiędzy nimi)*

*Stężenia różnorodnych związków chemicznych w płynach ustrojowych i tkankach mają często rozkład log-normalny w populacji. Optymalna jest wtedy ich prezentacja na skali logarytmicznej.*



# Miary położenia

# Miary położenia (tendencji centralnej)

Charakteryzuje „centralny” pomiar. Inaczej mówiąc jest **miarą położenia wartości reprezentatywnej**. W zależności od charakteru danych stosuje się różne definicje „centrum” i alternatywnie kilka miar położenia

# Miary położenia (tendencji centralnej)

## Średnia arytmetyczna $n$ pomiarów

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

inny zapis

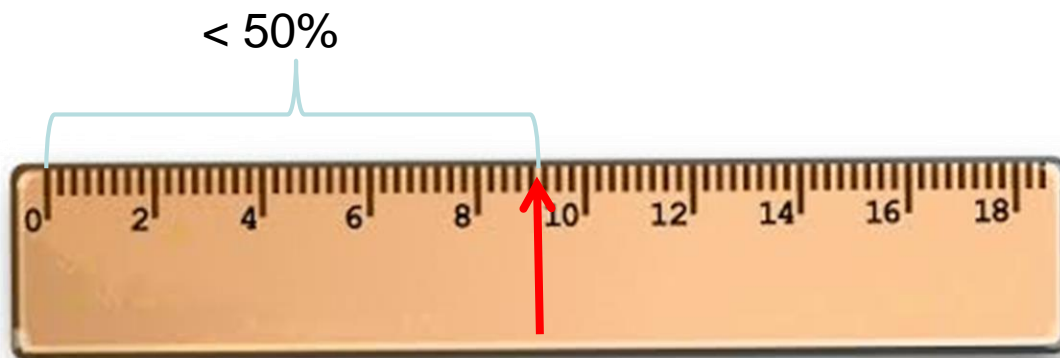
$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{lub} \quad \bar{x} = \frac{\sum x_i}{n} \quad \text{lub} \quad \bar{x} = \frac{\sum x}{n}$$

# Miary położenia (tendencji centralnej)

**Mediana** jest wartością znajdującą się w środku szeregu uporządkowanych danych, jeżeli liczba obserwacji jest nieparzysta

lub

jest średnią arytmetyczną z dwóch sąsiednich środkowych obserwacji, gdy liczba tych obserwacji jest parzysta





# Miary położenia (tendencji centralnej)

**Modalna** (moda, dominanta) jest wartością najczęściej występującą w zbiorze.

Jeżeli dane są ciągłe, to grupuje się je i wyznacza modalną dla grup.

Może istnieć kilka modalnych – gdy dwie lub więcej wartości występuje tyle samo razy a inne wartości występują mniej razy.

Może nie istnieć modalna – gdy każda z wartości występuje tylko raz

Jest mało odporna na przypadkowe odchylenia rozkładu próby od rozkładu populacji

# Miary położenia (tendencji centralnej)

Przykład 1:

Wyznaczyć modalną na podstawie liczby dzieci w 20 rodzinach.

Dane - liczba dzieci w rodzinie:

1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5

Moda wynosi 3

# Miary położenia (tendencji centralnej)

## Średnia geometryczna

$$\bar{x}_g = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

dla rozkładu prawoskośnego jest zbliżona do mediany i mniejsza od średniej arytmetycznej. Dla rozkładu log-normalnego jest równa medianie oraz średniej arytmetycznej na skali przekształconej logarytmicznie ( $x \rightarrow \log x$ ), dlatego jest szczególnie użyteczna dla opisu tego rozkładu.

$$\log \bar{x}_g = \log \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n} = (\log x_1 + \log x_2 + \log x_3 + \dots + \log x_n) / n \quad 27$$

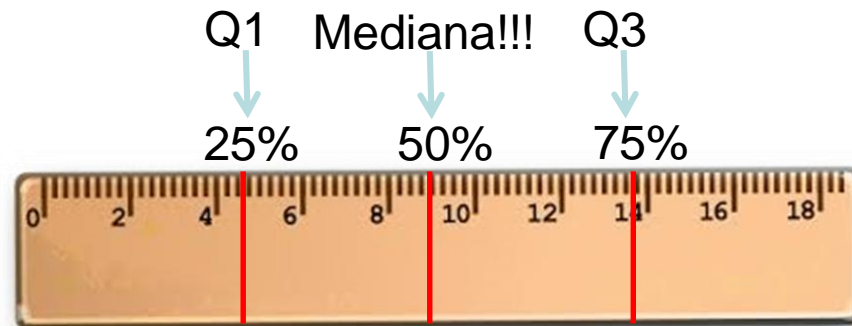
# Miary położenia (tendencji centralnej)

Średnia ważona - średnia elementów, którym przypisywane są różne wagi  $w_i$  w ten sposób, że elementy o większej wadze mają większy wpływ na średnią.

$$\bar{x} = \frac{w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + \dots + w_n \cdot x_n}{w_1 + w_2 + w_3 + \dots + w_n} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i}$$

# Miary położenia (tendencji centralnej)

**Kwartyl.** Wartości  $Q1(25\%)$ ,  $Q2(50\%)$ ,  $Q3(75\%)$ , zmiennej  $x$  które dzielą uporządkowany szereg na 4 równe pod względem liczebności części nazywa się kwartylami



# Miary położenia (tendencji centralnej)

Kwartyl pierwszy  $Q1$  dzieli obserwacje w taki sposób, że 25% obserwacji jest niższa bądź równa wartości tego kwartyla, a 75% obserwacji jest równa bądź większa niż wartość tego kwartyla

Analogicznie:

Kwartyl drugi – mediana

# Miary położenia (tendencji centralnej)

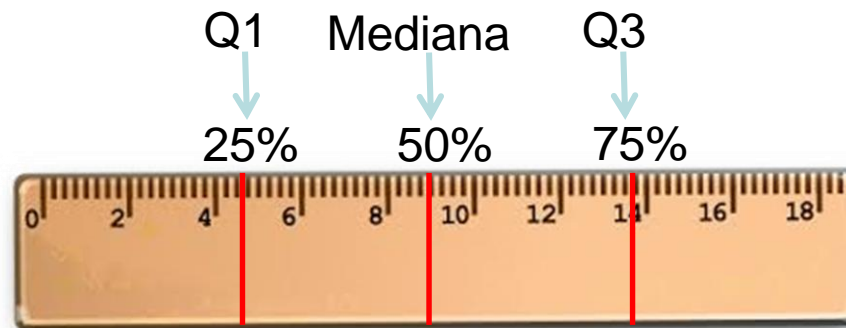
**Percentyl.** Porządkujemy szereg danych od najmniejszej do największej. Wartość zmiennej  $x$ , poniżej której w tym szeregu znajduje się 1% wartości nazywa się pierwszym percentylem. Analogicznie drugi percentyl, to taka wartość zmiennej  $x$ , poniżej której w tym uporządkowanym szeregu znajduje się 2% wartości

# Miary położenia (tendencji centralnej)

Czym jest dwudziesty piąty percentyl?

Czym jest pięćdziesiąty percentyl?

Czym jest siedemdziesiąty piąty percentyl?





# Miary rozproszenia

# Miary rozproszenia

**Miary rozproszenia** znane są także pod nazwą **dyspersji** lub **zmienności**

# Miary rozproszenia

**Rozstęp** to różnica pomiędzy największą (maksimum -  $x_{max}$ ) i najmniejszą (minimum -  $x_{min}$ ) wartością w zbiorze danych.

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

Może być mylący (zawyżony), gdy występują wartości odstające

# Miary rozproszenia

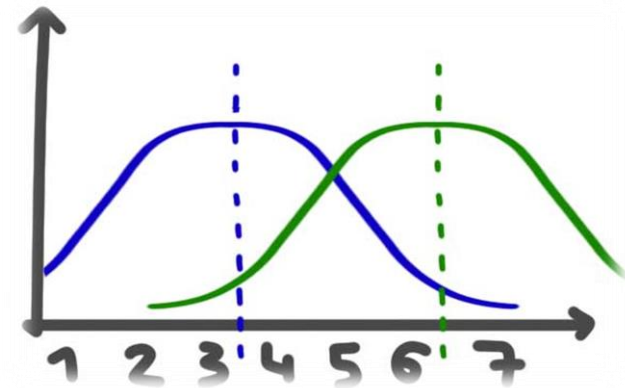
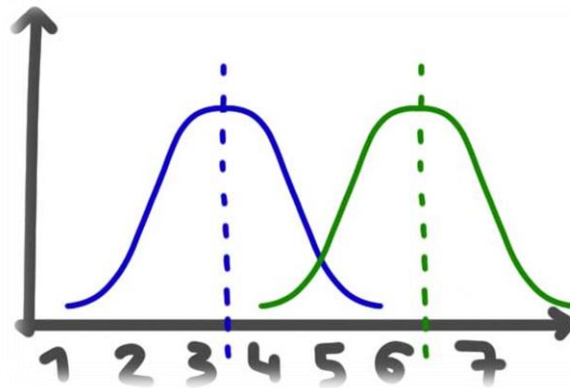
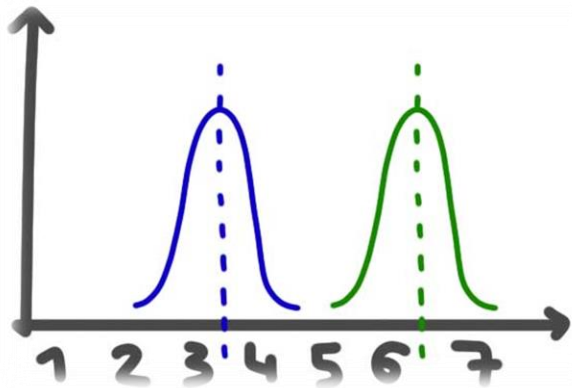
**Wariancja n pomiarów ( $n \geq 2$ )**

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

**Odchylenie standardowe**

$$s = \sqrt{s^2}$$

# Odchylenie standardowe



# Miary rozproszenia

## Współczynnik zmienności $w$

$$w = \frac{s}{\bar{x}} \cdot 100\%$$

*s* – odch. stand.  $\bar{x}$  – średnia

jest wartością względną (rozproszenie względem położenia)  
wyrażoną w procentach.

Stosowany gdy analizowana zmienna ma wartości tylko nieujemne  
(np. stężenie).

Wysokie wartości (>100%) świadczą o asymetrii rozkładu zmiennej.

# Miary rozproszenia

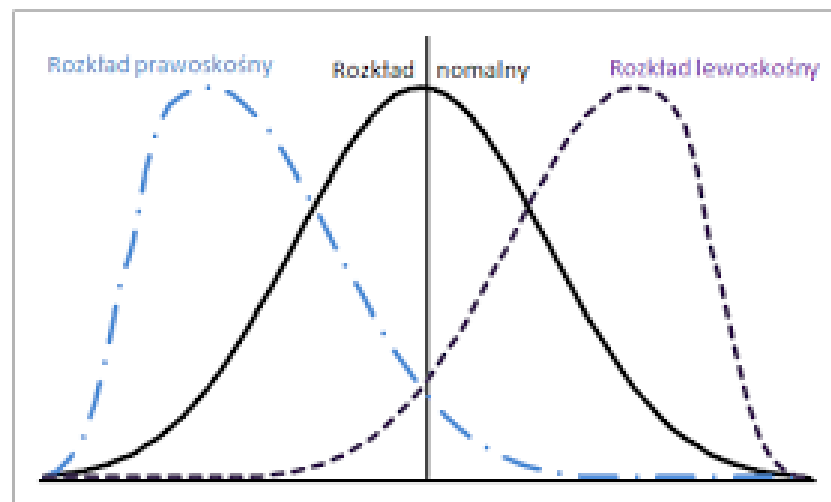
**Rozstęp międzykwartylowy (IQR –  
interquartile range)**

$$IQR = Q_3 - Q_1$$

# Miary symetrii rozkładu

Skośność:

- rozkład lewoskośny
- rozkład symetryczny
- rozkład prawoskośny – często spotykany w badaniach medycznych gdy pewna część populacji ma znacznie większe wartości zmiennej niż reszta np. parametry stanu zapalnego
  - Przykład: rozkład log-normalny





# Wielkość efektu

Podajemy dwa leki (A-właściwa substancja czynna i B-placebo) obniżające ciśnienie dwóm różnym grupom pacjentów. Czy działanie obu „leków” różni się istotnie pod względem obniżenia ciśnienia skurczowego pacjentów? Jak duża jest ta różnica odzwierciedlająca efekt działania substancji czynnej?

Wielkość efektu wyrażona różnicą średnich wynosi  
 $-3,7 - 1,3 = -5 \text{ mmHg}$

	A[mmHg]	B[mHg]
	-5	2
	-6	1
	-12	2
	-9	-1
	-8	1
	5	2
	-7	-2
	8	4
	1	3
średnia	-3,7	1,3
różnica	-5,0	

# Standaryzowana różnica średnich SMD

Często stosowanym wskaźnikiem wielkości efektu jest wskaźnik Cohena (SMD)

$$SMD = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Dzięki standaryzacji wielkość efektu nie zależy od jednostki w jakiej wyrażony jest mierzony parametr (np. mmHg czy kPa)

$\bar{x}_i$  – średnie;  $s$  – odchyl. stand.;  $n_i$  – liczebność prób;  $s_i^2$  – wariancje z prób  
 $i = 1, 2$

# Standaryzowana różnica średnich

## SMD

Wartości bezwzględne SMD >1 interpretuje się jako efekt znaczący,  
>2 jako duży,  
>3 jako bardzo duży

Znak SMD oznacza kierunek efektu: dodatni gdy w grupie z działającym czynnikiem średnia jest wyższa niż w grupie kontrolnej

$$\bar{x}_1 - \bar{x}_2 \text{ (różnica średnich)} = -5$$

$$s \text{ (uogólnione odch. stand.)} = 4,97$$

$$\text{SMD} = -5 / 4,97 = -1,01$$

**Interpretacja:**

***substancja czynna wywiera znaczący efekt obniżający ciśnienie tętnicze***

# Przedział ufności

# Przedział ufności

Jeżeli cecha  $X$  w **zbiorowości generalnej** ma rozkład normalny

$$X : N[\bar{X}, \sigma]$$

to średnia arytmetyczna z **próby z tej zbiorowości** ma rozkład normalny o parametrach:  $\bar{x}$

$$\bar{x} : N\left[\bar{X}, \frac{\sigma}{\sqrt{n}}\right] \quad \frac{\sigma}{\sqrt{n}} - \text{błąd standardowy}$$

# Przedział ufności

Wartość standaryzowana średniej arytmetycznej

$$u = \frac{\bar{x} - \bar{X}}{\sigma} \sqrt{n}$$

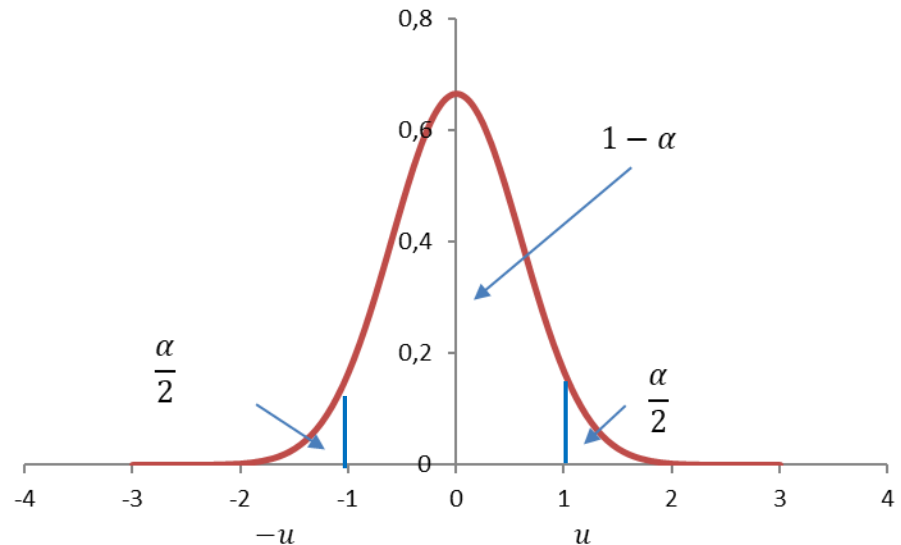
ma rozkład normalny o parametrach:

$$u : N[0,1]$$

# Przedział ufności

Prawdopodobieństwo, że wartość standaryzowana średniej arytmetycznej mieści się wewnątrz przedziału  $(-u_\alpha, u_\alpha)$  jest równe:

$$P\left(-u_\alpha < \frac{\bar{x} - \bar{X}}{\sigma} \sqrt{n} < u_\alpha\right) = 1 - \alpha$$



# Przedział ufności

Przedział ufności dla średniej arytmetycznej:

$$P\left(\bar{x} - u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \bar{x} + u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Poziom ufności:

$$1 - \alpha$$

*wyraża prawdopodobieństwo, że średnia w populacji estymowana na podstawie próby mieści się w granicach powyższego przedziału.*

*Wartość  $\alpha/2$  wyraża prawdopodobieństwo, że średnia w populacji znajduje się poniżej dolnej granicy tego przedziału. Ta sama wartość  $\alpha/2$  wyraża prawdopodobieństwo, że średnia w populacji znajduje się powyżej górnej granicy tego przedziału)*

Suma tych trzech prawdopodobieństw  $(1-\alpha)+\alpha/2+\alpha/2 = 1$  co oznacza że średnia w populacji musi się znajdować w przedziale ufności lub poniżej dolnej granicy tego przedziału lub powyżej górnej granicy tego przedziału



# Przedział ufności

Najczęściej stosowane w biostatystyce  
wartości:

$$1 - \alpha = 0,95, \quad \alpha = 0,05$$

$u_\alpha = 1,96$  (wynika z zasady 2SD rozkładu normalnego)

95% prawdopodobieństwo że parametr w populacji mieści się w określonym przedziale jest zwykle wystarczające z praktycznego punktu widzenia do podejmowania decyzji w medycynie i pokrewnych dziedzinach nauk

# Przedział ufności

## Przykład:

O populacji ludzi wiadomo, że czas reakcji na pewien bodziec ma rozkład normalny, o odchyleniu standardowym równym 12 minut. Wylosowano 36 zdrowych osób. Średni czas reakcji dla tej próby wynosi 23 minuty.

Obliczyć z prawdopodobieństwem równym 0,95 średni czas reakcji na ten bodziec u wszystkich zdrowych osób.

# Przedział ufności

$$P\left(\bar{x} - u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \bar{x} + u_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(23 - 1,96 \cdot \frac{12}{\sqrt{36}} < \bar{X} < 23 + 1,96 \cdot \frac{12}{\sqrt{36}}\right) = 0,95$$

$$P(19.1 < \bar{X} < 26.9) = 0.95$$

# Przedział ufności

Jeżeli cecha  $X$  w **zbiorowości generalnej** ma rozkład normalny

$$X : N[\bar{X}, \sigma]$$

oraz odchylenie standardowe populacji **nie jest znane**, przedziały ufności oblicza się w oparciu o tzw. rozkład t-Studenta, (zmienna losowa ma rozkład t-Studenta)

# Przedział ufności

Przedział ufności dla średniej arytmetycznej przy nieznanym SD populacji:

$$P \left( \bar{x} - t_{\alpha} \cdot \frac{s}{\sqrt{n}} < \bar{X} < \bar{x} + t_{\alpha} \cdot \frac{s}{\sqrt{n}} \right) = 1 - \alpha$$

*s* - odchylenie standardowe z próby

***W porównaniu z rozkładem o znanym SD zastępujemy SD populacji ( $\sigma$ ) przez SD próby (*s*) oraz standaryzowaną średnią  $u_{\alpha}$  o rozkładzie normalnym wartością  $t_{\alpha}$  o rozkładzie t-Studenta***

# Hipotezy statystyczne

# Hipotezy statystyczne

Hipoteza statystyczna to każde przypuszczenie o populacji generalnej, dotyczące jej cech statystycznych:

- rozkładu,
- miary tendencji centralnej,
- miary rozproszenia.

# Hipotezy statystyczne

Hipotezy statystyczne można podzielić na:

- **parametryczne** - hipoteza dotyczy wartości parametrów rozkładu (średnia, odch. stand.)
- **nieparametryczne** - hipoteza dotyczy rang (kolejności uporządkowania wartości) odzwierciedlających prawdopodobieństwo że wartość jednej zmiennej losowej jest większa lub mniejsza niż wartość innej zmiennej losowej, niezależnie od parametrów ich rozkładów



# Hipotezy statystyczne

Hipotezę statystyczną podlegającą weryfikacji nazywa się **zerową**  $H_0$ .

Jest przeciwieństwem hipotezy **alternatywnej**  $H_1$ .

# Hipotezy statystyczne

Tradycyjnie hipotezę **zerową**  $H_0$  formułuje się jako hipotezę o braku związku (np.: średnie dwóch populacji nie różnią się),

a hipotezę **alternatywną**  $H_1$  jako hipotezę o istnieniu związku (np.: średnie dwóch populacji różnią się).

Brak związku zakładany przez  $H_0$  można interpretować jako zerową wartość wielkości efektu definiowanego np. jako różnica średnich rozkładów parametru porównywanych populacji: jeśli średnie rozkładów parametru tych populacji nie różnią się od siebie, to różnica średnich wynosi zero.

# Hipotezy statystyczne

Weryfikacja hipotezy statystycznej dotyczącej jednej lub wielu populacji odbywa się na podstawie danych z próby wylosowanej z tej (tych) populacji.

Zatem, wnioski można formułować z pewnym prawdopodobieństwem a nie z pewnością.

# Hipotezy statystyczne

Można popełnić błędy dwojakiemu rodzaju:

- **błąd I rodzaju z prawdopodobieństwem  $\alpha$**  - odrzucić hipotezę  $H_0$  mimo, że jest prawdziwa

Stwierdzenie że populacje się różnią gdy w rzeczywistości się nie różnią

- Im mniejsze ryzyko błędu I rodzaju tym bardziej test statystyczny jest **konserwatywny**

- **błąd II rodzaju z prawdopodobieństwem  $\beta$**  - przyjąć hipotezę  $H_0$  mimo, że jest fałszywa

Stwierdzenie że populacje się nie różnią gdy w rzeczywistości się różnią

- Im mniejsze ryzyko błędu II rodzaju tym większą **moc** ma test statystyczny

# Hipotezy statystyczne

Prawdopodobieństwo popełnienia błędu I rodzajuu nosi nazwę **poziomu istotności**  $\alpha$ .

Poziom istotności ustalany jest a priori. W naukach biologiczno-medycznych wynosi zwykle  $\alpha = 0,05$ , rzadziej  $\alpha = 0,01$ .

# Hipotezy statystyczne

W praktyce hipoteza zerowa  $H_0$  o braku związku (średnie dwóch populacji nie różnią się) jest nieprawdopodobna, gdyż prawdopodobieństwo, że określony ciągły parametr (np. średni wiek) wyznaczony z dokładnością do dowolnie dużej liczby miejsc po przecinku jest identyczny w dwóch populacjach, jest nieskończenie małe.

# Hipotezy statystyczne a przedział ufności

Weryfikację nieprawdopodobnej hipotezy zerowej  $H_0$  o braku związku należy traktować tylko jako teoretyczne narzędzie przydatne do wnioskowania statystycznego o wartości parametrów rozkładów w populacjach i wielkości efektów na podstawie parametrów rozkładów w próbach z tych populacji.

**Poziom istotności  $\alpha$**  (prawdopodobieństwo odrzucenia hipotezy  $H_0$  gdyby była prawdziwa) jest bowiem ściśle powiązany z poziomem ufności  $1-\alpha$  dla **przedziału ufności** odzwierciedlającym prawdopodobieństwo, że wartość parametru w populacji estymowana na podstawie próby mieści się w granicach powyższego przedziału.

# Interpretacja wartości $p$ w oparciu o przedział ufności

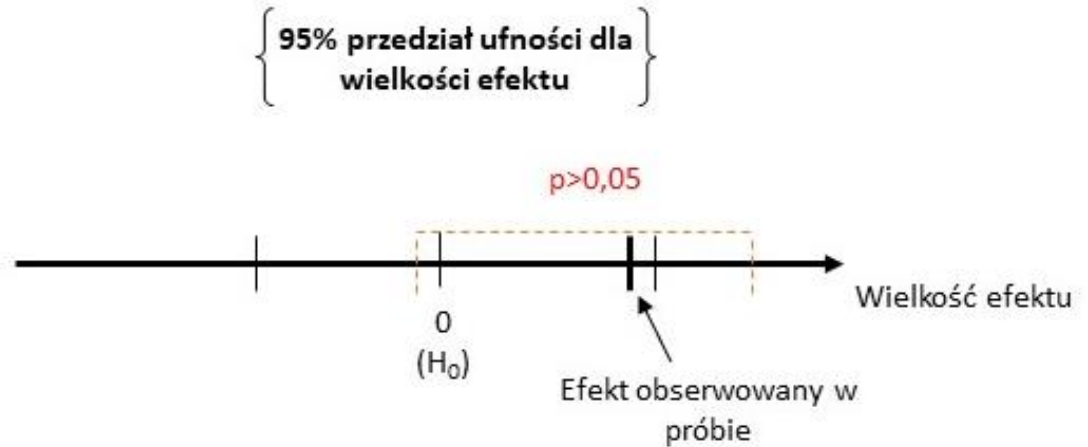
Wartość  $p < 0.05$  dla weryfikacji hipotezy  $H_0$  o braku efektu oznacza, że **brak efektu** znajduje się **poza przedziałem 95% ufności** dla wielkości efektu, a więc **możemy z dużym prawdopodobieństwem wykluczyć w populacji brak efektu lub efekt przeciwny do obserwowanego w próbie.**

Wartość  $p > 0.05$  dla weryfikacji hipotezy  $H_0$  o braku efektu oznacza, że **brak efektu** znajduje się **wewnątrz przedziału 95% ufności** dla wielkości efektu, a więc **nie możemy z dużym prawdopodobieństwem wykluczyć w populacji braku efektu lub efektu przeciwnego do obserwowanego w próbie.**

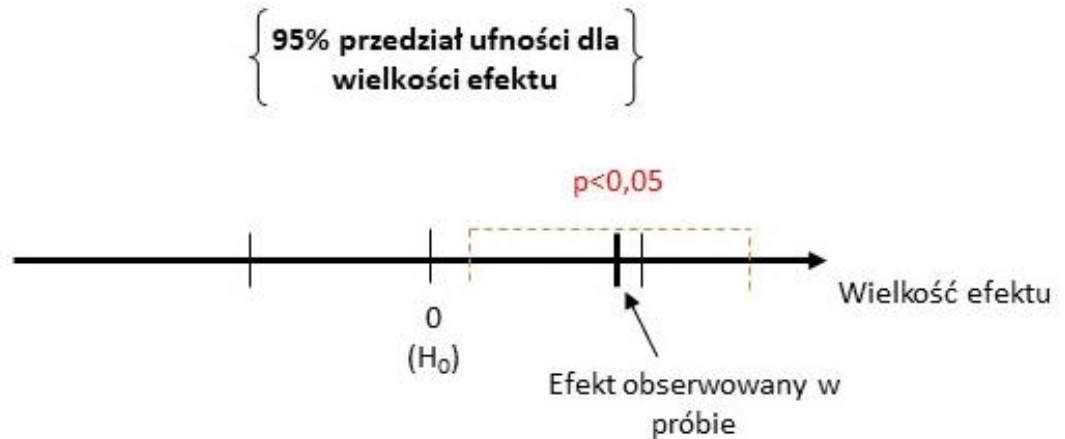


Wielkość efektu = 0 – brak efektu

brak efektu znajduje się wewnątrz przedziału 95% ufności dla wielkości efektu



brak efektu znajduje się poza przedziałem 95% ufności dla wielkości efektu



# Wybór typu testu

O wyborze decyduje liczebność porównywanych grup, rodzaj zmiennych (powiązane, niepowiązane) oraz kształt rozkładu prawdopodobieństwa:

- rozkład zgodny z rozkładem normalnym – test parametryczny
- rozkład niezgodny z rozkładem normalnym – test nieparametryczny

# Wybór typu testu

Badanie kształtu rozkładu na podstawie próby przeprowadzić można przy pomocy np.

testu **Shapiro-Wilka**

*Im niższa wartość statystyki **W** tym rozkład bardziej odbiega od normalnego*

*Jeśli odpowiadająca statystyce **W** wartość  $p < 0.05$ , dowodzi to istotnego odchylenia badanego rozkładu od rozkładu normalnego*

# Testy parametryczne

# Testy parametryczne

- Badana cecha ma rozkład normalny
- Różnica wariancji w badanych populacjach jest nieistotna statystycznie (wariancje są jednorodne)

# Testy parametryczne

Test t-Studenta dla dwóch niewiązanych prób losowych (dla różnicy średnich) gdy wariancje w populacjach są nieznane ale równe sobie:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{K \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$K = \sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}}$$

$s_i^2$  – wariancje z próby ( $i = 1, 2$ )

Przykład: Podajemy dwa leki (A-10 osobom i B-12 osobom) obniżające ciśnienie dwóm różnym grupom pacjentów. Czy działanie leków różni się istotnie?

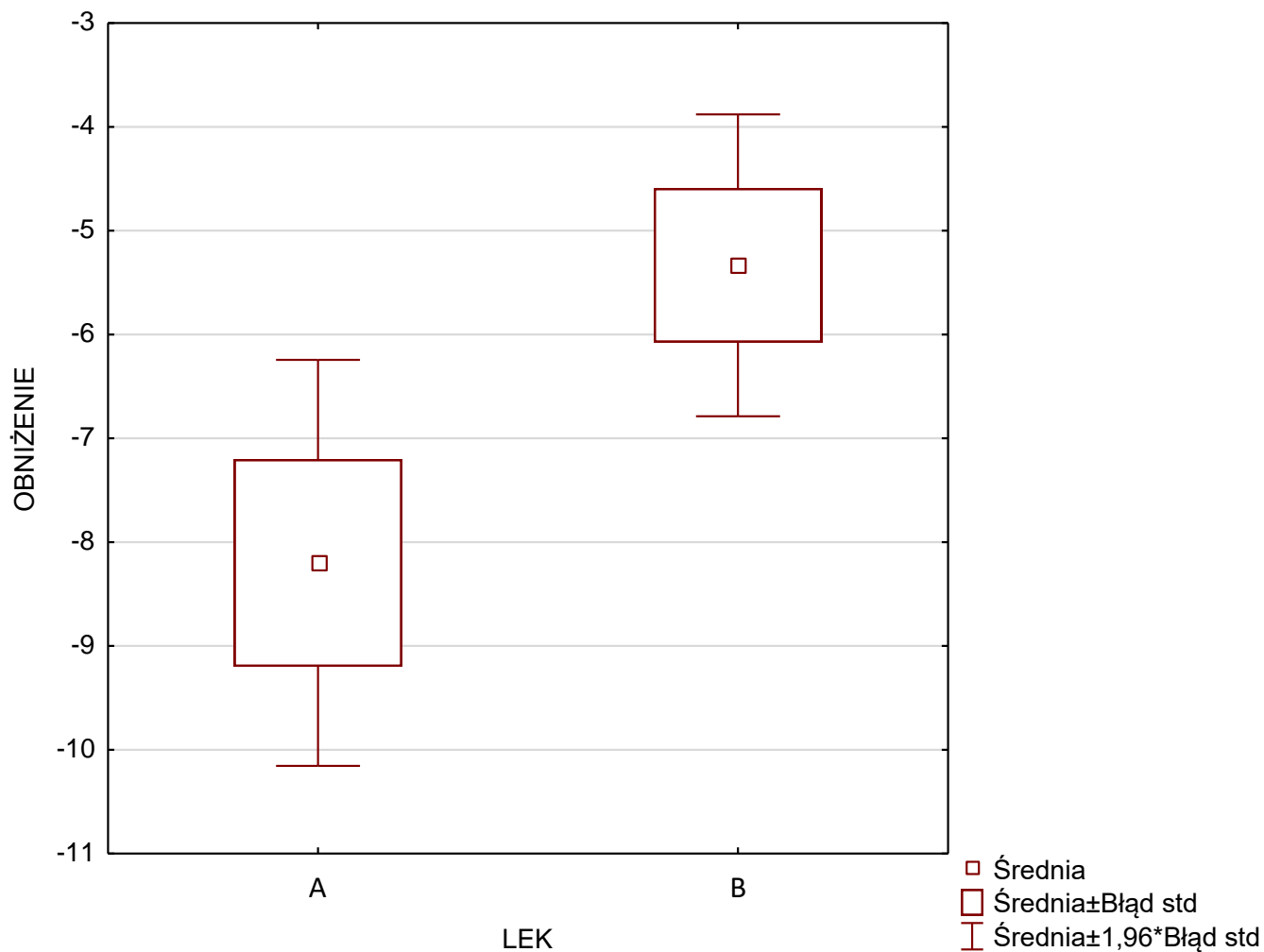
OBNIŻENIE	LEK
-5	Lek A
-6	Lek A
-12	Lek A
-9	Lek A
-8	Lek A
-5	Lek A
-7	Lek A
-8	Lek A
-15	Lek A
-7	Lek A
-6	Lek B
-5	Lek B
-11	Lek B
-5	Lek B
-3	Lek B
-4	Lek B
-6	Lek B
-6	Lek B
-4	Lek B
-9	Lek B
-3	Lek B
-2	Lek B

Zmienna	Średnia A	Średnia B	t	df	p	Nważnych A	Nważnych B
OBNIŻENIE	-8,20	-5,33	-2,35	20	0,03	10	12

$p < 0,05$



Istotne statystycznie różnice  
w działaniu leku



Obserwujemy, iż Lek A „wykazał” się większą (w sensie istotnym matematycznie, ale niekoniecznie klinicznie) skutecznością w obniżaniu ciśnienia krwi



# Testy parametryczne

**Jeżeli wariacje różnią się w sposób istotny statystycznie należy skorzystać z testów nieparametrycznych**

# Testy parametryczne

Test t – Studenta dla dwóch powiązanych prób losowych (dla średniej różnic  $\bar{d}$ ), gdy wariancja tych różnic w populacji nie jest znana

$$t = \frac{\bar{d}}{s} \cdot \sqrt{n}$$

$$d_i = x_{1i} - x_{2i} \quad \bar{d} = \frac{\sum_{i=1}^n d_i}{n} \quad s = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

Przykład: w pewnym eksperymencie porównywano miano przeciwciał przed i po szczepieniu. Czy miana przeciwciał przed i po szczepieniu statystycznie istotnie się różnią?

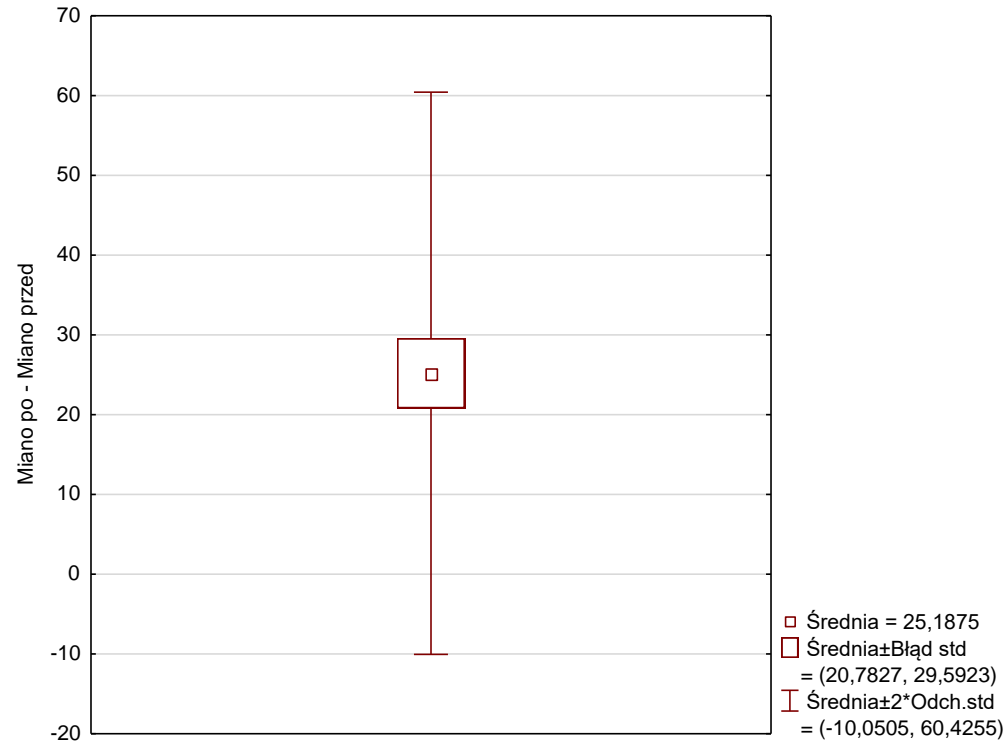
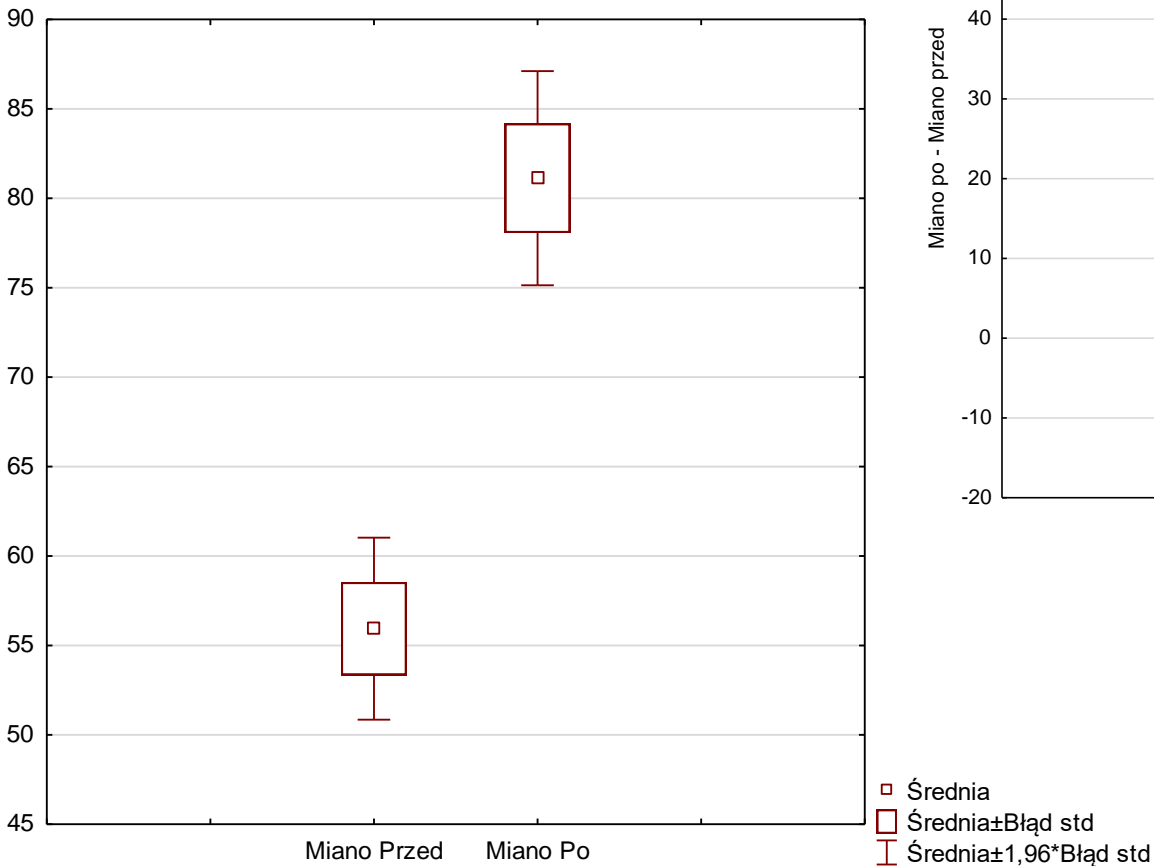
Zmienna	Srednia	Odch.st.	Ważnych	Różnica	Odch.st. Różnica	t	p
Miano Po	81,13	12,22					
Miano Przed	55,94	10,38	16	25,19	17,62	5,72	<0,01

Miano Przed	Miano Po	Różnica
74	63	-11
49	90	41
55	64	9
75	96	21
49	66	17
41	97	56
75	72	-3
58	86	28
52	88	36
47	91	44
55	68	13
50	92	42
57	91	34
54	80	26
46	68	22
58	86	28

$p < 0,05$



Istotne statystycznie różnice:  
miano przeciwciał przed i po  
szczepieniu różni się istotnie



Obserwujemy, iż miano przeciwciał po szczepieniu wzrosło istotnie (w sensie matematycznym, niekoniecznie klinicznym)

# ANOVA

(dla więcej niż dwóch grup)

Analiza wariancji - jednoczynnikowa:

- każda populacja musi mieć rozkład normalny,
- próby pobrane z każdej populacji muszą być niezależnymi próbami losowymi,
- wariancje w populacjach są zbliżone.

ROZSZERZENIE TESTU T-STUDENTA DLA PRÓB  
NIEPOWIĄZANYCH

# ANOVA

(dla więcej niż dwóch zmiennych powiązanych)

ANOVA z powtarzanymi pomiarami:

- każda zmienna musi mieć rozkład normalny,
- wariancje zmiennych są zbliżone.

ROZSZERZENIE DLA TESTU T-STUDENTA DLA  
ZMIENNYCH POWIĄZANYCH

# ANOVA

jeśli odrzucamy  $H_0$  (są istotne statystycznie różnice między średnimi, ale nie wiemy między którymi) to kolejnym krokiem analizy jest porównanie wielokrotne *post hoc* dla wszystkich par średnich

- **procedura Bonferroniego** – test *post hoc*: dzielimy poziom istotności przez liczbę wykonanych porównań (w tym wypadku wartość  $p$  testu t-Studenta dla danej pary średnich przez liczbę wszystkich analizowanych par średnich)
- w naukach medycznych często jest stosowany **test post hoc Tukeya**, który ma większą moc statystyczną niż procedura Bonferroniego

Porównania wielokrotne wiążą się ze zwiększeniem **błędu I rodzaju** (odrzućenie hipotezy zerowej wówczas, gdy jest uzasadniona) – poprawka np. Bonferroniego zmniejsza to ryzyko, ale kosztem spadku mocy testu.

**Wniosek:** jeśli będą wykonywane porównania wielokrotne należy zwiększyć liczebność próby z populacji.

# ANOVA

## Przykład 1

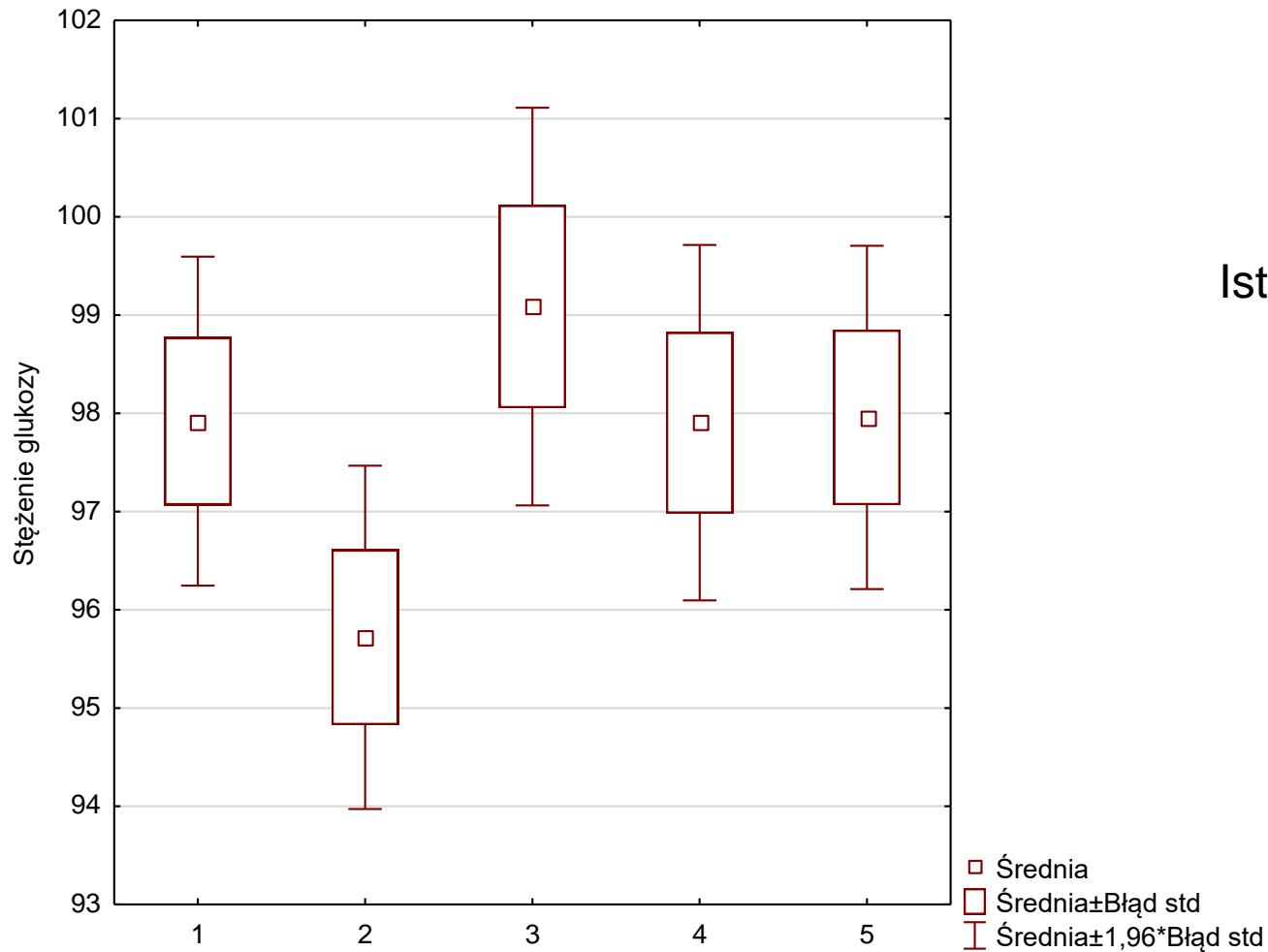
Tabela zawiera wyniki pomiaru stężenia glukozy we krwi u osób stosujących różne diety. Czy poziom glikemii zależy od wybranej diety?



1	2	3	4	5
101	93	102	92	90
105	100	104	104	97
98	93	93	102	97
96	94	103	101	97
90	98	103	100	104
100	102	96	98	93
101	90	105	95	97
102	100	99	105	99
96	103	91	100	99
91	101	101	97	97
91	92	92	97	103
100	92	104	101	103
97	96	101	92	104
94	104	93	100	92
101	93	103	91	100
102	90	96	99	100
95	93	101	104	93
99	95	103	97	98
92	95	94	94	95
99	91	95	93	103
96	92	105	94	101
104	95	91		104
94	90	104		90
102	99			95
102	102			

**Weryfikacja normalności  
rozkładów oraz  
jednorodności wariancji  
wykazała, że można  
zastosować ANOVA**

# ANOVA



$$p_{ANOVA} < 0,05$$



Istotne statystycznie  
różnice

# ANOVA

Pomiędzy którymi grupami?



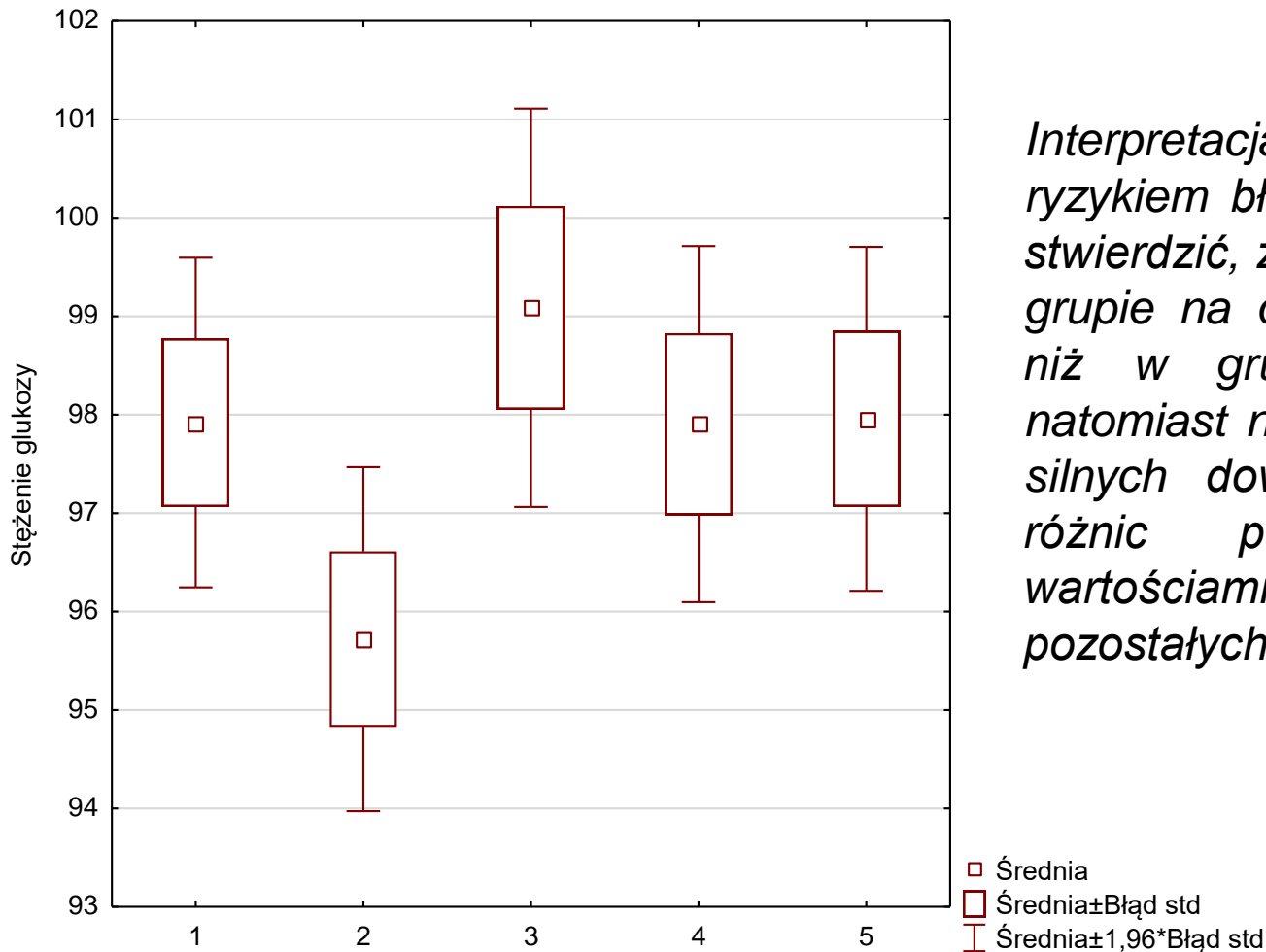
Test post-hoc Tukeya

Istotna statystycznie jest tylko różnica między grupami **2 i 3** ( $p=0.02$ )

Test RIR Tukeya; zmienna: Zmn1 (Zeszyt1)  
Zaznaczone różnice są istotne z  $p < ,05000$

Zmn2	{1}	{2}	{3}	{4}	{5}
	M=97,920	M=95,720	M=99,087	M=97,905	M=97,958
1 {1}		0,41	0,89	1,00	1,00
2 {2}	0,41		0,02	0,47	0,41
3 {3}	0,89	0,02		0,90	0,91
4 {4}	1,00	0,47	0,90		1,00
5 {5}	1,00	0,41	0,91	1,00	

# ANOVA+test post-hoc Tukeya



*Interpretacja: Z niewielkim ryzykiem błędu ( $p < 0.05$ ) można stwierdzić, że średnia glikemia w grupie na diecie 3 jest wyższa niż w grupie na diecie 2, natomiast nie ma wystarczająco silnych dowodów na istnienie różnic pomiędzy średnimi wartościami glikemii dla pozostałych par diet.*

# Testy nieparametryczne

# Testy nieparametryczne

- Stosowane gdy zmienne losowe znacznie odbiegają od rozkładu normalnego lub wyrażone są na skali porządkowej (rangowej)
- nie wymagają założeń odnośnie rozkładu populacji, z której losowana jest próba
- są mniej wrażliwe na obserwacje odstające
- formalnie wymagają zgodności rozkładów w porównywanych populacjach, ale są mało wrażliwe na różnice między tymi rozkładami

# Testy nieparametryczne

Dwie próby niepowiązane?



## **Test U-Manna-Whitneya**

- badane są dwie populacje
- nieparametryczna alternatywa dla testu t-Studenta dla prób niepowiązanych
- porównuje rangi zmiennych zamiast ich wartości

Przykład: W dwóch grupach chorych na pewną chorobę neurologiczną przeprowadzono badania stężenia adrenaliny w surowicy krwi. Czy stężenie adrenaliny w obu grupach jest jednakowe?

ADRENALINA	GRUPA
14,34	a
20,33	a
18,79	a
8,22	a
31,5	a
12,08	a
22	a
9,22	a
19,5	a
78,89	a
30,48	a
45,86	a
5,33	b
22,5	b
11,74	b
7,39	b
12,34	b
13,22	b
8,53	b
22,8	b
12,7	b
7,78	b
9,63	b
8,9	b

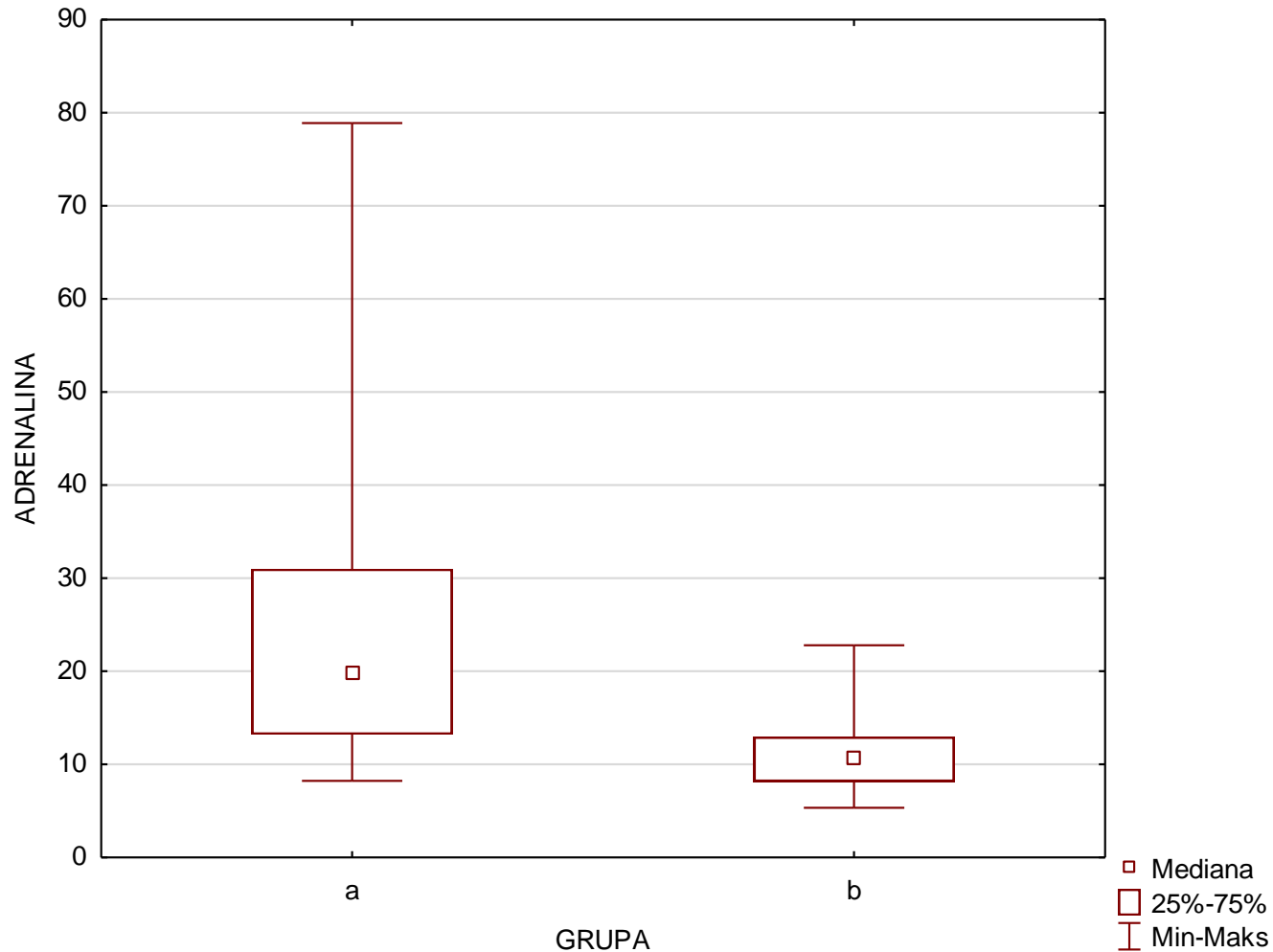
Sum.rang a	Sum.rang b	U	Z	p
191,00	109,00	31,00	2,34	0,02

$p < 0,05$



Istotne statystycznie różnice:  
stężenie adrenaliny istotnie  
się różni w obu badanych  
grupach





Stężenie adrenaliny jest istotnie wyższe w grupie a, w której wszystkie pokazane na rycinie miary położenia (mediana, kwartyle, minimum i maksimum) są wyższe niż w grupie b

# Testy nieparametryczne

Więcej niż dwie próby niepowiązane?



## **Test Kruskala-Wallisa**

- nieparametryczna alternatywa dla ANOVA
- porównuje rangi zmiennych zamiast ich wartości

# Testy nieparametryczne

## Test Kruskala-Wallisa:

- zmienna losowa jest zmienną ilościową lecz nie ma rozkładu normalnego lub zmienna losowa wyrażona jest w skali porządkowej (rangowej)
- formalnie wymaga zgodności rozkładów w porównywanych populacjach, ale jest mało wrażliwy na różnice między tymi rozkładami

CZAS_K	TEST
9,10	I
8,90	I
8,40	I
10,00	I
8,70	I
9,20	I
7,60	I
8,60	I
8,90	I
7,90	I
10,00	II
10,20	II
9,80	II
11,60	II
9,50	II
9,20	II
8,60	II
10,30	II
9,40	II
8,50	II
10,00	III
9,90	III
9,80	III
12,90	III
11,20	III
9,90	III
8,50	III
9,80	III
9,20	III
8,20	III
10,90	IV
11,10	IV
12,20	IV
14,40	IV
9,80	IV
12,00	IV
8,50	IV
10,90	IV
10,40	IV
10,00	IV

Badano czterema różnymi metodami czas krzepnięcia osocza krwi u dziesięciu losowo wybranych pacjentów. Czy czas krzepnięcia krwi zmierzony wybranymi metodami różni się istotnie?

Test Kruskala-Wallisa:  $H = 14,03$   $p = 0,001$

$p < 0,05$



Istotne statystycznie różnice:  
 czas krzepnięcia krwi  
 określony przez wybrane  
 metody pomiaru jest istotnie  
 różny

*Ale pomiędzy, którymi grupami?*

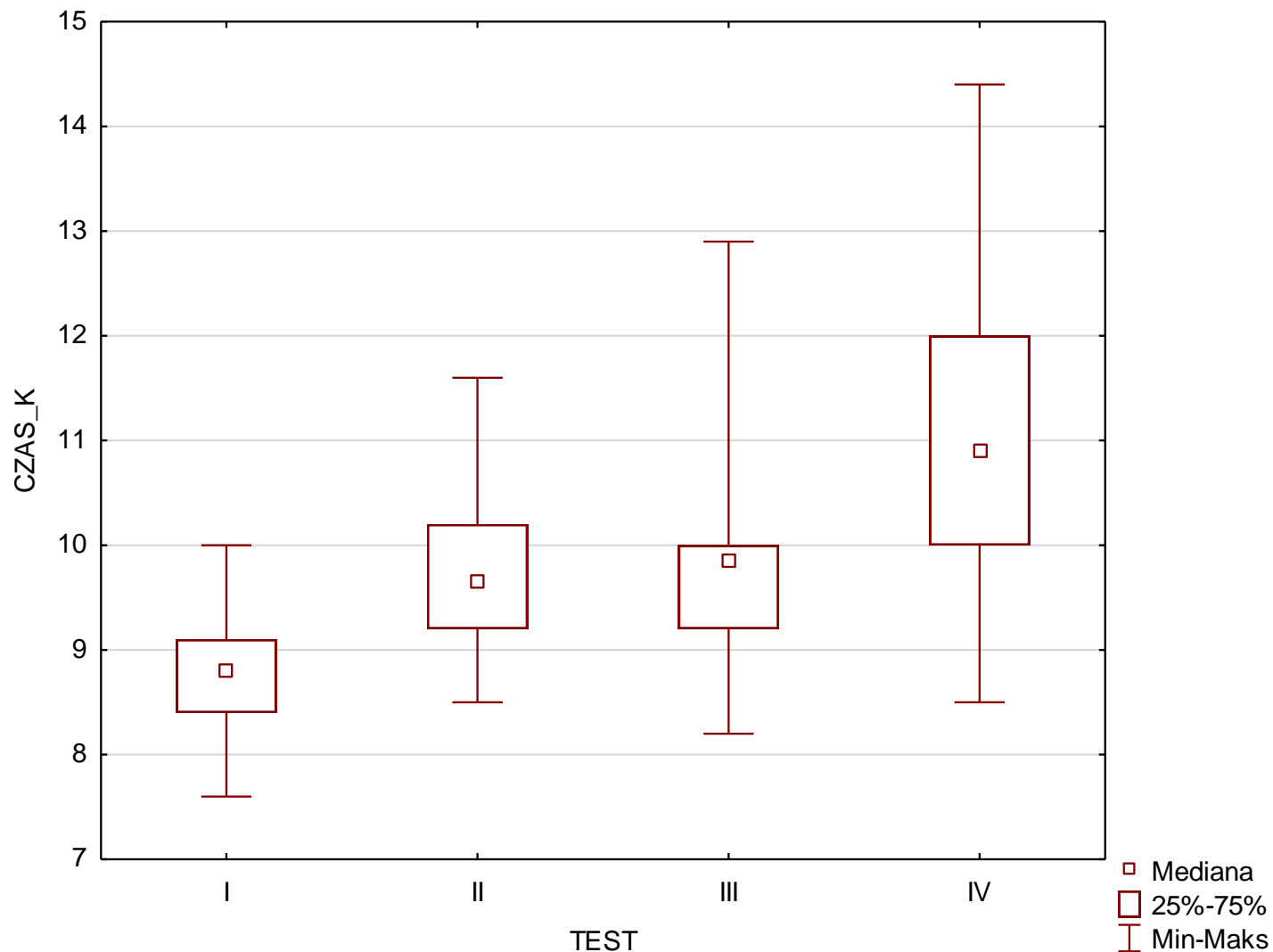
# Test post - hoc

Zależna: CZAS K	I R:10,300	II R:20,650	III R:21,250	IV R:29,800
I		0,29	0,22	0,00
II	0,29		1,00	0,48
III	0,22	1,00		0,61
IV	0,00	0,48	0,61	

$p < 0,05$



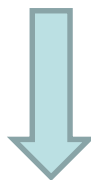
Istotne statystycznie różnice:  
czas krzepnięcia krwi różni  
się istotnie pomiędzy metodą  
I i IV



*Interpretacja: Z niewielkim ryzykiem błędu ( $p < 0.05$ ) można stwierdzić, że czas krzepnięcia zmierzony metodą IV jest dłuższy niż zmierzony metodą I, natomiast nie ma wystarczająco silnych dowodów na istnienie różnic pomiędzy czasami zmierzonymi pozostałymi parami metod.*

# Testy nieparametryczne

Dwie próby powiązane?



## **Test kolejności par Wilcoxona**

Przykład 1: Tabela zawiera wyniki badania masy ciała 15 osób przed kuracją odchudzającą i po kuracji.

Czy kuracja działa?

# Testy nieparametryczne

## Przykład 1

<b>pacjent</b>	<b>przed kuracją [kg]</b>	<b>po kuracji [kg]</b>	<b>różnica [kg]</b>	<b>mediana różnicy</b>
1	96	90	-6	-5
2	87	82	-5	
3	90	83	-7	
4	101	97	-4	
5	97	95	-2	
6	97	92	-5	
7	85	88	3	
8	92	86	-6	
9	103	97	-6	
10	88	89	1	
11	97	93	-4	
12	101	95	-6	
13	95	91	-4	
14	91	90	-1	
15	88	82	-6	



# Testy nieparametryczne

Przykład 1 (na podstawie D. Miller, St. Orzeszyna *Elementy statystyki medycznej*):

Przyjęty poziom istotności  $\alpha = 0,05$ .

Prawdopodobieństwo  $p$  ( $p$ -value) związane z wynikiem testu  $p = 0,002$

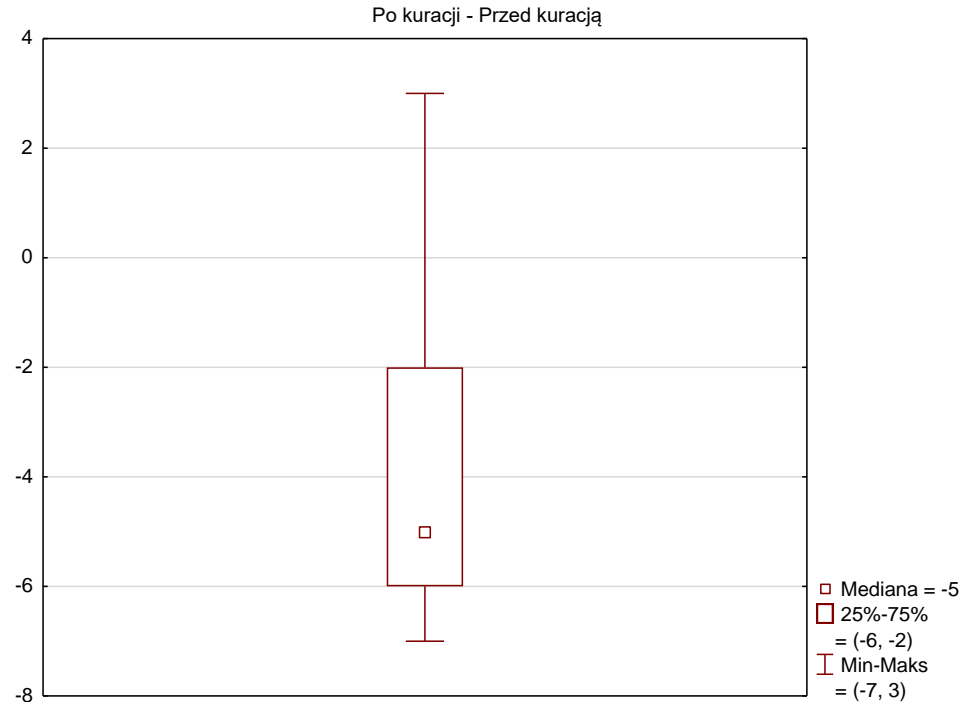
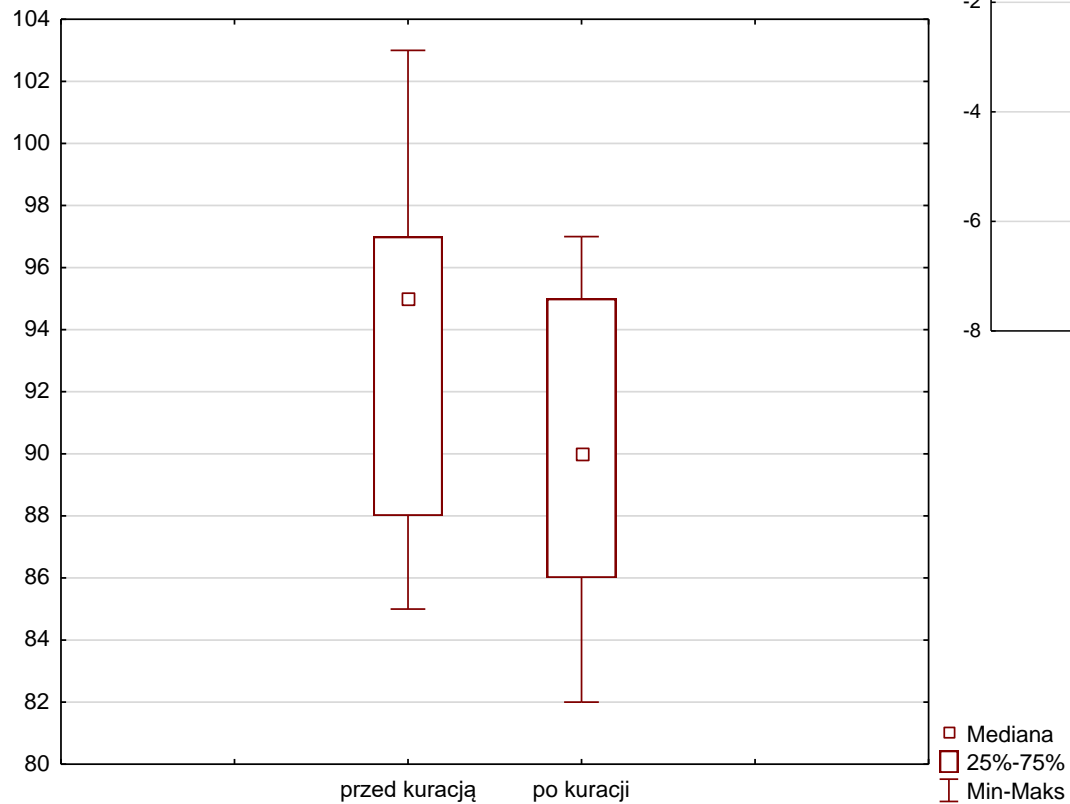
**Wynik testu istotny statystycznie** ( $p < \alpha$ ) – odnotowano **istotne statystycznie** różnice w masach ciał przed i po kuracji

**Kierunek efektu:** interpretując np. wartość mediany różnicy można stwierdzić **spadek** masy ciała po zastosowanej kuracji

**Wielkość efektu** wyrażona wartością mediany różnicy wynosi **-5 kg**

Podsumowanie: kuracja spowodowała **istotny** statystycznie **spadek** masy ciała, którego **mediana** wyniosła **-5 kg**

# Testy nieparametryczne



# Testy nieparametryczne

Więcej niż dwie próby powiązane?



**Test Friedmana**

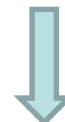
„rozwinięcie” testu Wilcoxon na więcej niż dwie próby

W pewnej klinice oceniano lek przeciwko niedokrwistości złośliwej. W trakcie leczenia oznaczono pięciokrotnie stężenie witaminy B12 w surowicy chorego. Czy stężenie witaminy B12 zmieniało się istotnie pod wpływem leczenia?

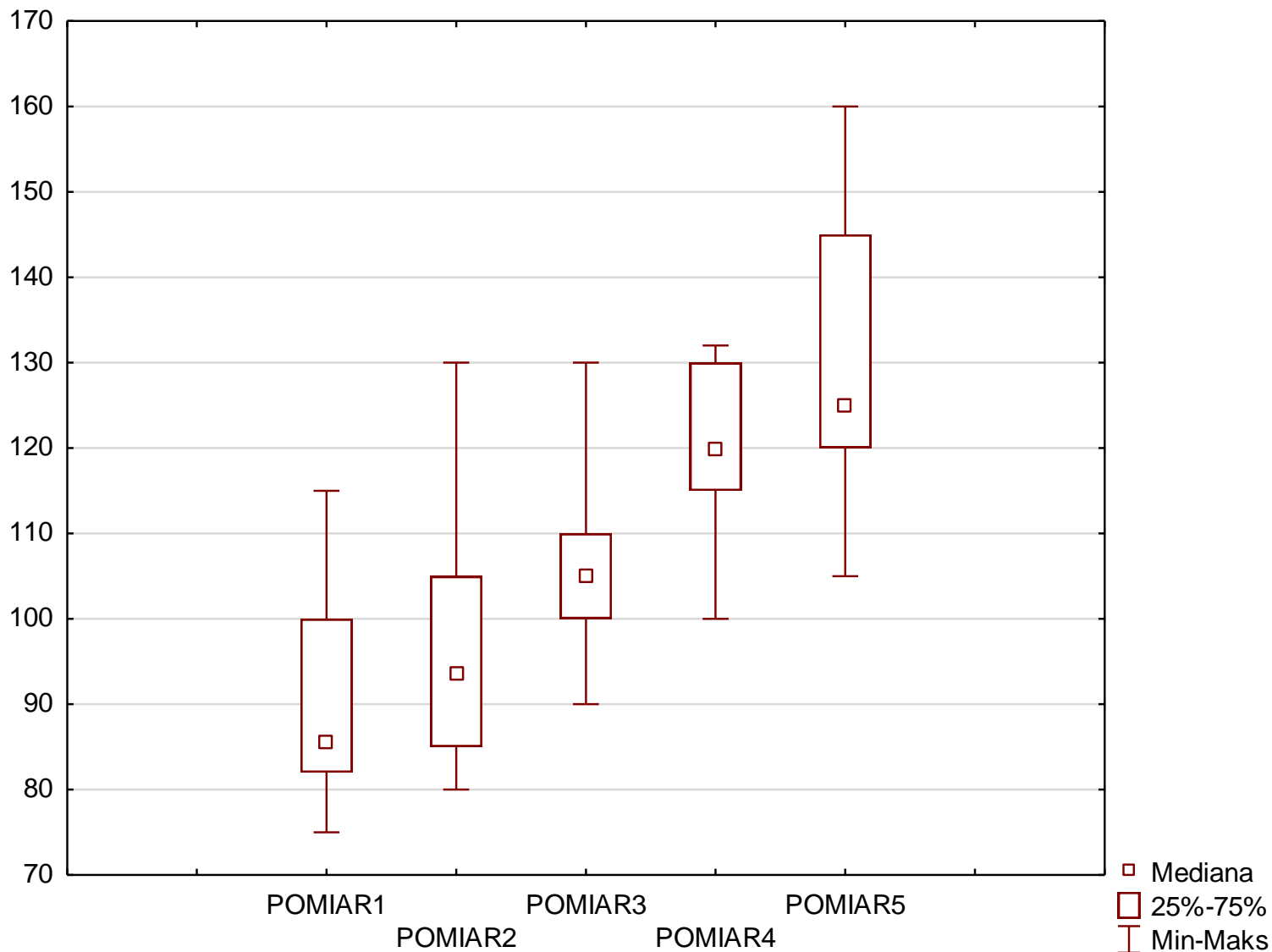
Test Friedmana= 36,22  $p < 0,001$

POMIAR1	POMIAR2	POMIAR3	POMIAR4	POMIAR5
82	90	105	115	125
86	80	100	130	160
100	105	115	130	150
75	80	90	105	115
82	92	110	120	125
90	110	110	115	130
85	85	100	120	125
115	130	130	132	145
105	100	105	125	120
85	95	90	100	105

$p < 0,05$



Istotne statystycznie różnice:  
stężenie witaminy B12 w  
surowicy istotnie zmienia się  
w czasie



Obserwujemy wysoce istotny (w sensie matematycznym:  $p < 0.001$ , niekoniecznie klinicznym) wzrost stężenia wit. B12 w czasie leczenia

# Analiza zależności

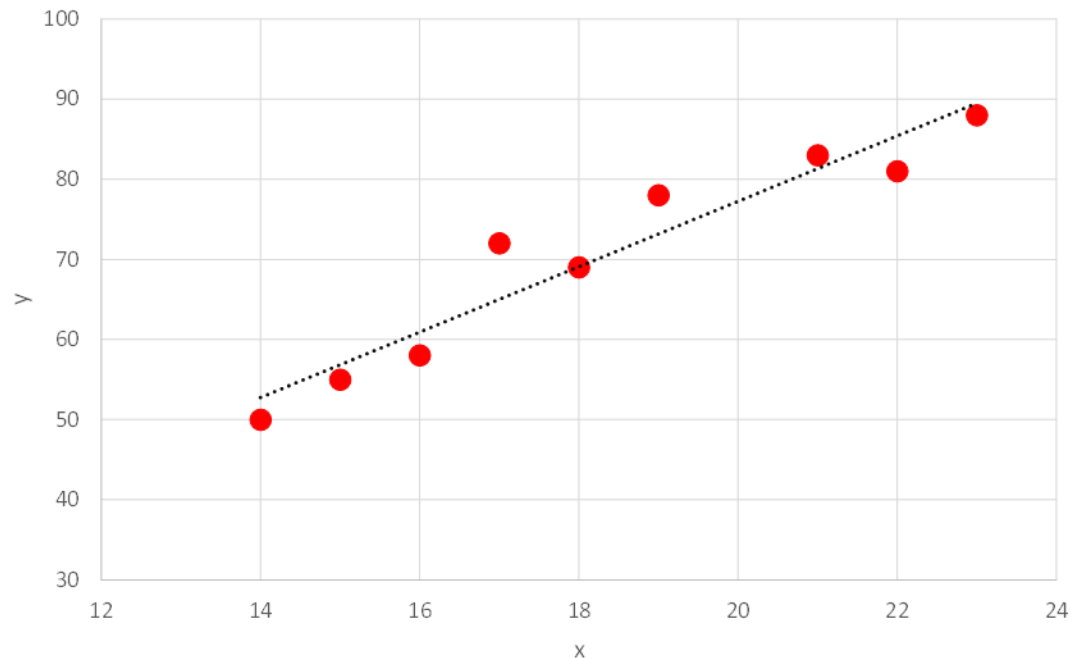
# Analiza zależności między dwiema cechami mierzalnymi

Powiązanie funkcyjne zmiennej zależnej od zmiennej niezależnej:

- zależność liniowa
- zależność nieliniowa

# Analiza zależności między dwiema cechami mierzalnymi

Zależność liniowa jest opisana liniowym równaniem regresji, inaczej równaniem regresji stopnia pierwszego:



$$y = a \cdot x + b$$



# Analiza zależności między dwiema cechami mierzalnymi

Równanie regresji stopnia pierwszego pozwala na predykcję wartości wewnątrz przedziału obserwacji zmiennej niezależnej!

Predykcja poza tym przedziałem jest ekstrapolacją i grozi znacznymi błędami gdy np. zależność staje się nieliniowa

# Analiza zależności między dwiema cechami mierzalnymi

Miara korelacji prostoliniowej – współczynnik korelacji Pearsona:

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2}}$$

Warunek stosowania: rozkłady obu cech są normalne  
a zależność między nimi jest liniowa

$r$  zmienia się od -1 (-100%) do +1 (+100%)

# Analiza zależności między dwiema cechami mierzalnymi

- $r > 0$  korelacja dodatnia - wzrost zmiennej niezależnej jest związany ze wzrostem zmiennej zależnej
- $r < 0$  korelacja ujemna - wzrost zmiennej niezależnej jest związany ze spadkiem zmiennej zależnej

# Analiza zależności między dwiema cechami mierzalnymi

- wartość bezwzględna równa 1 – pełny związek funkcyjny
- wartość równa 0 – całkowity brak związku funkcyjnego
- wartości pomiędzy 0 a 1 – słabsza lub silniejsza zależność liniowa

# Skala zależności

$$r_{xy} = 0$$

brak zależności

$$0 < r_{xy} < 0,1$$

nikła zależność

$$0,1 \leq r_{xy} < 0,3$$

słaba zależność

$$0,3 \leq r_{xy} < 0,5$$

przeciętna zależność

$$0,5 \leq r_{xy} < 0,7$$

silna zależność

$$0,7 \leq r_{xy} < 0,9$$

bardzo silna zależność

$$0,9 \leq r_{xy} < 1$$

prawie pełna zależność

$$r_{xy} = 1$$

pełna zależność

# Analiza zależności między dwiema cechami rangowanymi

Rozkłady nie są zbliżone do normalnego?

Zależność nie jest liniowa?



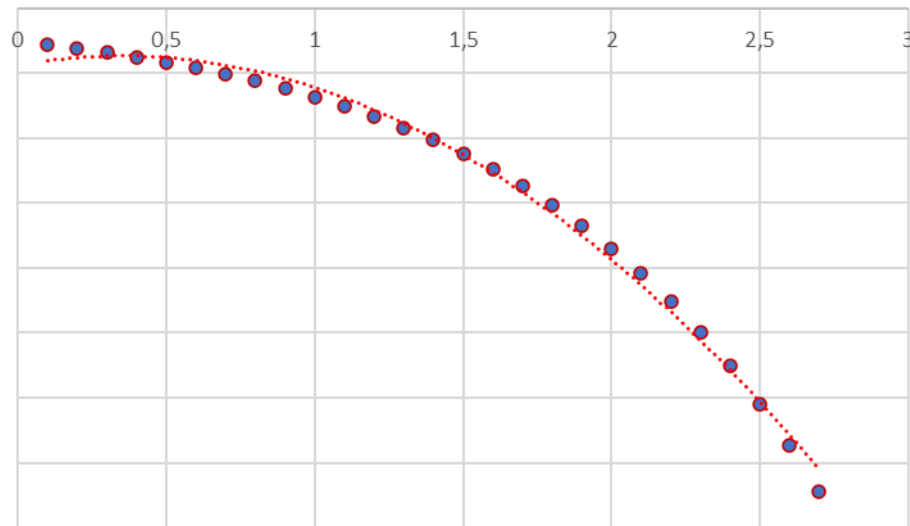
Współczynnik korelacji rang Spearmana ( $\rho$ )

Zależność może być nieliniowa, ale musi być monotoniczna (ze wzrostem zmiennej niezależnej zmienna zależna albo zawsze rośnie albo zawsze maleje)

# Analiza zależności między dwiema cechami rangowanymi

$$np. \rho = -0,95$$

Związek jest silny ale nieliniowy  
– współczynnik korelacji liniowej  
byłby nieadekwatny



Ujemny współczynnik korelacji rang wskazuje, iż wyższa dawka podawanego leku (oś OX) związana jest z obniżeniem poziomu ciśnienia tętniczego (oś OY). Wartość bezwzględna bliska 1 wskazuje na bardzo silną zależność pomiędzy stosowaną dawką leku a poziomem ciśnienia tętniczego.

# Ryzyko i szansa



# Ryzyko i szansa

**Ryzyko** – ogólnie: wskaźnik stanu lub zdarzenia, które może prowadzić do strat lub zysków. Jest ono proporcjonalne do prawdopodobieństwa wystąpienia tego zdarzenia i do wielkości strat lub zysków, które może spowodować.

*W biostatystyce: prawdopodobieństwo wystąpienia określonego stanu klinicznego*

**Szansa** – ogólnie: prawdopodobieństwo, możliwość powodzenia w jakiejś sprawie, osiągnięcia sukcesu lub zaistnienia jakichś pożądaných okoliczności  
*W biostatystyce: stosunek prawdopodobieństwa wystąpienia określonego stanu klinicznego do prawdopodobieństwa jego niewystąpienia w danej grupie.*

# Ryzyko względne (RR)

*„Ryzyko względne (relative risk) - stosunek prawdopodobieństwa zajścia jakiegoś zdarzenia (np. zachorowanie, śmierć) w jednej badanej grupie względem drugiej grupy. Za pomocą ryzyka względnego ocenia się o ile jest większe lub mniejsze prawdopodobieństwo wystąpienia danego zdarzenia w grupie A w porównaniu do grupy B.”*

$$RR_{AB} = \frac{P(A)}{P(B)}$$

*P(A) - prawdopodobieństwo wystąpienia zdarzenia w grupie A*

*P(B) - prawdopodobieństwo wystąpienia zdarzenia w grupie B*

# Ryzyko względne (RR)

*„Ryzyko względne (relative risk) - stosunek prawdopodobieństwa zajścia jakiegoś zdarzenia (np. zachorowanie, śmierć) w jednej badanej grupie względem drugiej grupy. Za pomocą ryzyka względnego ocenia się o ile jest większe lub mniejsze prawdopodobieństwo wystąpienia danego zdarzenia w grupie A w porównaniu do grupy B.”*

RR przyjmuje wartości od 0 do  $+\infty$

- **RR $\approx$ 1**: określa sytuację, w której prawdopodobieństwo wystąpienia zdarzenia w grupie A jest zbliżone do prawdopodobieństwa wystąpienia zdarzenia w grupie B
- **RR $>$ 1**: prawdopodobieństwo wystąpienia zdarzenia w grupie A jest większe niż w B
- **RR = 2**: wystąpienie zdarzenia w grupie A jest 2-krotnie (o 100%) bardziej prawdopodobne niż w grupie B (RR = 3 oznacza, że 3-krotnie)
- **RR = 1,5**: o 50% większe w A niż w B
- **RR $<$ 1**: sytuacja odwrotna - w grupie A jest mniejsze niż w B
- **RR=0,75**: wystąpienie zdarzenia w grupie B jest o 25% mniej prawdopodobne w grupie A niż B, ale o 33% bardziej prawdopodobne w grupie B niż A, bo  $1/0.75=1.33$

# Iloraz szans (OR)

*„Iloraz szans (odds ratio) - stosunek szansy wystąpienia danego zdarzenia w danej grupie do szansy wystąpienia tego samego zdarzenia w innej porównywanej grupie. Za pomocą wskaźnika OR określamy zatem o ile większa bądź mniejsza jest szansa wystąpienia zdarzenia, np.: choroba, śmierć w jednej grupie w porównaniu do innej grupy.”*

$$OR_{AB} = \frac{S(A)}{S(B)}$$

$$S(A) = \frac{P(A)}{1-P(A)} - \text{szansa wystąpienia zjawiska w grupie A}$$

$$S(B) = \frac{P(B)}{1-P(B)} - \text{szansa wystąpienia zjawiska w grupie B}$$

# Iloraz szans (OR)

*„Iloraz szans (odds ratio) - stosunek szansy wystąpienia danego zdarzenia w danej grupie do szansy wystąpienia tego samego zdarzenia w innej porównywanej grupie. Za pomocą wskaźnika OR określamy zatem o ile większa bądź mniejsza jest szansa wystąpienia zdarzenia, np.: choroba, śmierć w jednej grupie w porównaniu do innej grupy.”*

- **OR $\approx$ 1**: szansa wystąpienia zdarzenia w jednej i drugiej grupie jest zbliżona
- **OR $<$ 1**: w badanej grupie (w porównaniu do grupy odniesienia) istnieje mniejsza szansa wystąpienia zdarzenia
- **OR $>$ 1**: w badanej grupie (w porównaniu do grupy odniesienia) istnieje większa szansa wystąpienia danego zdarzenia.

# Dane niemierzalne

# Analiza zależności między dwiema cechami niemierzalnymi

Współczynnik Spearmana ( $\rho$ ), tau Kendalla,  
V Cramera

dane zebrane są w tzw. tablicy czteropolowej

# Analiza zależności między dwiema cechami niemierzalnymi

	odmiany cechy B		
odmiany cechy A	I	II	razem
I	a	b	a+b
II	c	d	c+d
razem	a+c	b+d	a+b+c+d



# Analiza zależności między dwiema cechami niemierzalnymi

*Interpretacja współczynnika  $\rho$  Spearmana dla tablicy czteropolowej*

- **$\rho = 0$**  brak zależności
- **$\rho > 0$**  I odmiana cechy A współwystępuje z I odmianą cechy B, a II odmiana cechy A współwystępuje z II odmianą cechy B
- **$\rho < 0$**  I odmiana cechy A współwystępuje z II odmianą cechy B, a II odmiana cechy A współwystępuje z I odmianą cechy B

# Analiza zależności między dwiema cechami niemierzalnymi

Przykład: Czy istnieje zależność pomiędzy szczepieniem a odpornością populacji na chorobę?

Dane zawiera poniższa tabela:

	zachorowali	nie zachorowali	razem
szczepieni	13 (19,1%)	55 (80,9%)	68 (100%)
nieszczepieni	28 (51,9%)	26 (48,1%)	54 (100%)
razem	41 (33,6%)	81 (66,4%)	122 (100%)

# Analiza zależności między dwiema cechami niemierzalnymi

współczynnik Spearmana:

$$\rho = 0,54; p < 0,001$$

$\rho > 0$  – „umiarkowana” pozytywna zależność pomiędzy brakiem zachorowania a szczepieniem;  
współczynnik korelacji istotny statystycznie ( $p < 0,05$ )

*Współczynnik  $\rho$  Spearmana może być stosowany do oceny siły związku między cechami mierzalnymi, rangowymi i dychotomicznymi*

# Analiza zależności między dwiema cechami niemierzalnymi

	zachorowali	nie zachorowali
szczepieni	13 (19,1%)	55 (80,9%)
nieszczepieni	28 (51,9%)	26 (48,1%)

Nieszczepieni

Szczepieni

Ryzyko zachorowania =  $28/(28+26) = 0,5185$

Ryzyko zachorowania =  $13/(13+55) = 0,1912$

Szansa zachorowania =  $28/26 = 1,0769$

Szansa zachorowania =  $13/55 = 0,2364$

Ryzyko względne (RR) zachorowania szczepieni vs nieszczepieni  
=  $0,1912/0,5185 = 0,369$

Iloraz szans (OR) zachorowania szczepieni vs nieszczepieni  
=  $0,2364/1,0769 = 0,220$

# Dane nominalne

Test chi-kwadrat:

$$\chi^2 = \sum_{i=1}^k \frac{(E_i - T_i)^2}{T_i}$$

$E_i$  – liczebność eksperymentalna w kategorii  $i$

$T_i$  – liczebność teoretyczna w kategorii  $i$

# Dane nominalne

## Test chi-kwadrat:

- liczebność w wartościach absolutnych,
- minimalna, dopuszczalna liczebność dowolnej kategorii wynosi 1,
- maksymalnie  $1/5$  kategorii może mieć wartość mniejszą od 5

# Dane nominalne

Przykład 1 (na podstawie D. Schwartz, P.Lazar *Elementy statystyki medycznej i biologicznej*):

Tabela zawiera wyniki badania 298 osób ze stwierdzonym nowotworem żołądka.

Czy istnieje związek pomiędzy umiejscowieniem nowotworu i płcią?

	odźwiernik	trzon żołądka	razem
mężczyźni	53 (44,5%)	66 (55,5%)	194 (100%)
kobiety	48 (59,3%)	33 (40,7%)	104 (100%)
razem	101 (50,5%)	99 (49,5%)	298 (100%)

# Dane nominalne

Alternatywa dla testu chi  
– kwadrat  
(charakteryzuje się  
większą  
mocą/wiarygodnością)

$$\chi^2=4,18$$

$$p=0,041$$

Dokładny test  
Fishera (tabela 2x2)

$$p=0,043$$



Wynik istotny statystycznie

(odnotowano istotny statystycznie związek pomiędzy umiejscowieniem nowotworu i płcią ( $p < 0,05$ ): u kobiet częściej niż u mężczyzn nowotwór jest zlokalizowany w odźwierniku)



# Dane nominalne

	odźwiernik	trzon żołądka
mężczyźni	53 (44,5%)	66 (55,5%)
kobiety	48 (59,3%)	33 (40,7%)

Mężczyźni

Kobiety

$$\text{Ryzyko (odźwiernik)} = 53/(53+66) = 0,4453$$

$$\text{Ryzyko (odźwiernik)} = 48/(48+33) = 0,5926$$

$$\text{Ryzyko względne (RR) lokalizacji w odźwierniku vs w trzonie dla mężczyzn vs kobiet} \\ = 0,4453/0,5926 = \mathbf{0,752}$$

$$\text{Szansa (odźwiernik)} = 53/66 = 0,8030$$

$$\text{Szansa (odźwiernik)} = 48/33 = 1,4545$$

$$\text{Iloraz szans (OR) lokalizacji w odźwierniku vs w trzonie dla mężczyzn vs kobiet} \\ = 0,8030/1,4545 = \mathbf{0,552}$$

**Przedział 95% ufności (95%CI)** dla ilorazu szans umiejscowienia nowotworu żołądka w odźwierniku w porównaniu z trzonem u mężczyzn w porównaniu z kobietami : **0,31–0,98**  
**nie zawiera wartości OR=1** oznaczającej brak efektu.

Wniosek: u **mężczyzn rzadziej** niż u kobiet nowotwór jest zlokalizowany w **odźwierniku**.  
Kierunek i wielkość efektu jaki wywiera płeć (męska) na lokalizację nowotworu (odźwiernik) wyraża wartość **OR (95%CI)=0,552 (0,31–0,98)**, a przedział ufności nie zawierający wartości 1 jest pośrednim dowodem **istotności statystycznej** ( $p < 0,05$ ) obserwowanego efektu.

# Dane nominalne

Przykład 1 (na podstawie D. Schwartz, P.Lazar *Elementy statystyki medycznej i biologicznej*):

Grupa 348 dzieci została podzielona losowo na dwie podgrupy. Jedna została zaszczepiona szczepionką BCG firmy A a druga szczepionką BCG firmy B. Wyniki obserwacji reakcji na szczepionki zawiera tabela.

Czy szczepionki działają tak samo?

# Dane nominalne

Przykład 2:

Tabela liczności (dane obserwowane)

szczepionka	Brak reakcji lub reakcja lekka	reakcja średnia	reakcja ciężka	razem
A	156 (88,1%)	12 (6,8%)	9 (5,1%)	177 (100%)
B	135 (78,9%)	29 (17,0%)	7 (4,1%)	171 (100%)

# Dane nominalne

$$\chi^2=58,08$$

$$p=0,008$$



Wynik istotny statystycznie  
(odnotowano istotną  
statystycznie różnicę w  
działaniu szczepionek  
( $p<0,05$ ))

# Dane nominalne

szczepionka	Brak reakcji lub reakcja lekka	reakcja ciężka
A	156 (88,1%)	9 (5,1%)
B	135 (78,9%)	7 (4,1%)

A

B

Ryzyko reakcji ciężkiej =  $9/(9+156) = 0,0545$

Ryzyko reakcji ciężkiej =  $7/(7+135) = 0,0493$

Ryzyko względne (RR) reakcji ciężkiej: A vs B  
 $= 0,0545/0,0493 = 1,1055$

Szansa reakcji ciężkiej =  $9/156 = 0,0577$

Szansa reakcji ciężkiej =  $7/135 = 0,0516$

Iloraz szans (OR) reakcji ciężkiej: A vs B  
 $= 0,0577/0,0516 = 1,127$

Przedział 95% ufności dla ilorazu szans wystąpienia reakcji ciężkiej po szczepionce A w porównaniu z B wynoszący 0,63 – 1,97 zawiera wartość OR=1 oznaczającej brak efektu.

# Dane nominalne

szczepionka	Brak reakcji lub reakcja lekka	reakcja średnia
A	156 (88,1%)	12 (6,8%)
B	135 (78,9%)	29 (17,0%)

A

B

Ryzyko reakcji średniej =  $12/(12+156)=0,0714$

Ryzyko reakcji średniej =  $29/(29+135) = 0,1768$

Ryzyko względne (RR) reakcji średniej: A vs B  
=  $0,0714 / 0,1768 = 0,4038$

Szansa reakcji średniej =  $12/156 = 0,0769$

Szansa reakcji średniej =  $29/135 = 0,2148$

Iloraz szans (OR) reakcji średniej: A vs B  
=  $0,0769/0,2148 = 0,3581$

Przedział 95% ufności dla ilorazu szans wystąpienia reakcji średniej po szczepionce A w porównaniu z B wynoszący 0,20 – 0,63 **nie zawiera wartości OR=1** oznaczającej brak efektu.

# Materiały uzupełniające

# Nieściłości w nomenklaturze statystycznej

zmienna niepowiązana (niesparowana)  
≠  
zmienna niezależna\*

\*w oprogramowaniu Statistica pojęcie zmiennej niezależnej jest niesłusznie stosowane do zmiennych niepowiązanych (dotyczy większości testów statystycznych)

Prawidłowa definicja:

**Zmienna niezależna** – zmienna, której wartości zmieniamy

$$y = x$$



# Nieścisłości w nomenklaturze statystycznej

zmienna powiązana (sparowana)

≠

zmienna zależna\*

\*w oprogramowaniu Statistica pojęcie zmiennej zależnej jest niesłusznie stosowane do zmiennych powiązanych (dotyczy większości testów statystycznych)

Prawidłowa definicja:

**Zmienna zależna** – zmienna, którą mierzymy

$$y \neq x$$

# Nieścisłości w nomenklaturze statystycznej

## Wariancja dla próby a wariancja dla populacji

Obliczając wariancję lub odchylenie standardowe w programie Excel należy wybierać odpowiednio funkcję WARIANCJA() lub ODCH.STANDARDOWE(), które wyliczają te miary rozproszenia „na podstawie **próby** z populacji”.

Nie należy w tym celu korzystać z funkcji WARIANCJA.POPUL() ani ODCH.STANDARD.POPUL() obliczających te miary „na podstawie **całej** populacji”, gdyż otrzymana wartość jest niższa od rzeczywistego estymatora wariancji w populacji o współczynnik  $(n-1)/n$  w odniesieniu do wartości „dla **próby** z populacji”, gdzie  $n$  oznacza liczebność badanej próby. Dla odchylenia standardowego otrzymana wartość jest zaniżona o współczynnik równy pierwiastkowi z  $(n-1)/n$ .