



Podstawowe pojęcia i testy statystyczne

Piotr Knapik

Śląskie Centrum Chorób Serca w Zabrze

Zebranie w Domu Lekarza

- niehomogenna, wysoce wyselekcjonowana grupa słuchaczy,
- rozkład wiedzy na temat statystyki odbiega od dystrybucji normalnej,
- duże odchylenie standardowe z uwagi na liczne przypadki wartości skrajnych (wiem wszystko albo nie wiem nic),
- **dla niektórych słuchaczy informacje będą oczywiste, a nawet banalne.**

Fakty

- fachowcy którzy naprawdę rozumieją statystykę mają czasem trudności w przekazaniu tej wiedzy,
- tacy ludzie są na ogół uzdolnieni matematycznie – kontakt z laikiem, który takich uzdolnień nie ma, nie wie czego chce i nic nie rozumie może być bardzo irytujący!

Po co to wszystko?

Rozumienie podstaw wnioskowania statystycznego jest potrzebne do:

- zrozumienia literatury fachowej,
- zaplanowania metodyki badań własnych,
- zaplanowania tabel i wykresów,
- zrozumienia sensu uzyskanych wyników,
- konfrontowania swoich wyników z wynikami innych,
- wyciągnięcia prawidłowych wniosków z badań.

Problemy

- w literaturze fachowej jest mnóstwo (większość?) słabych prac, które nie posuwają do przodu wiedzy medycznej, tylko wprowadzają szum informacyjny,
- pracy nie da się ocenić na podstawie streszczenia,
- trzeba przeczytać dokładnie całą metodykę i przejrzeć wyniki,
- **znajomość podstaw statystyki jest potrzebna, aby ocenić wartość pracy.**

Po co są testy statystyczne?

- rozpoczynając eksperyment, mamy jakąś hipotezę (np. że lek A działa silniej niż lek B),
- nie sprawdzimy czy nasza hipoteza jest słuszna badając wszystkich ludzi na świecie,
- możemy jednak spróbować przetestować naszą hipotezę, przeprowadzając badania na pewnej ograniczonej populacji lub grupie (*sample*),
- **im większa grupa badana, tym większe prawdopodobieństwo że nie popełnimy błędu!**

Dane

- Jakościowe (*qualitative*)
- Ilościowe (*quantitative*)

Dane jakościowe

- **Binarne** (*binary data* - np. płeć, obecność lub brak danej choroby, dane typu 0/1)
- **Nominalne** (*nominal data* – np. rasa, diagnoza)
- **Porządkowe** lub rangowane (*ordinal data* – np. nasilenie objawów; łagodne, średnie, ciężkie)

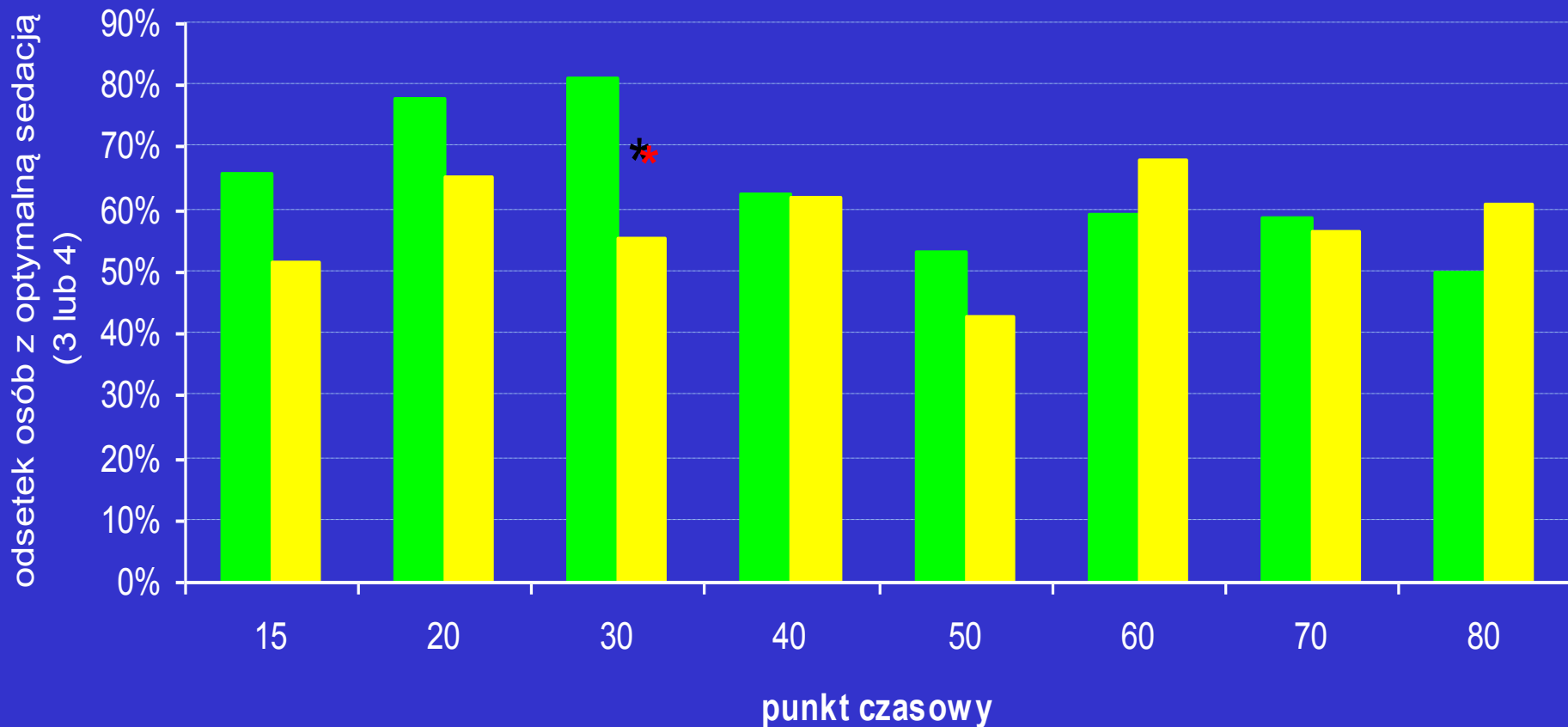
Dane jakościowe

- wynik podawany w procentach, rozsądne podejście do liczby miejsc po przecinku!
- przykład: $45/100 = 45\%$
- sposób wizualizacji: słupki (*bar charts*) lub wykresy kołowe (*pie charts*)

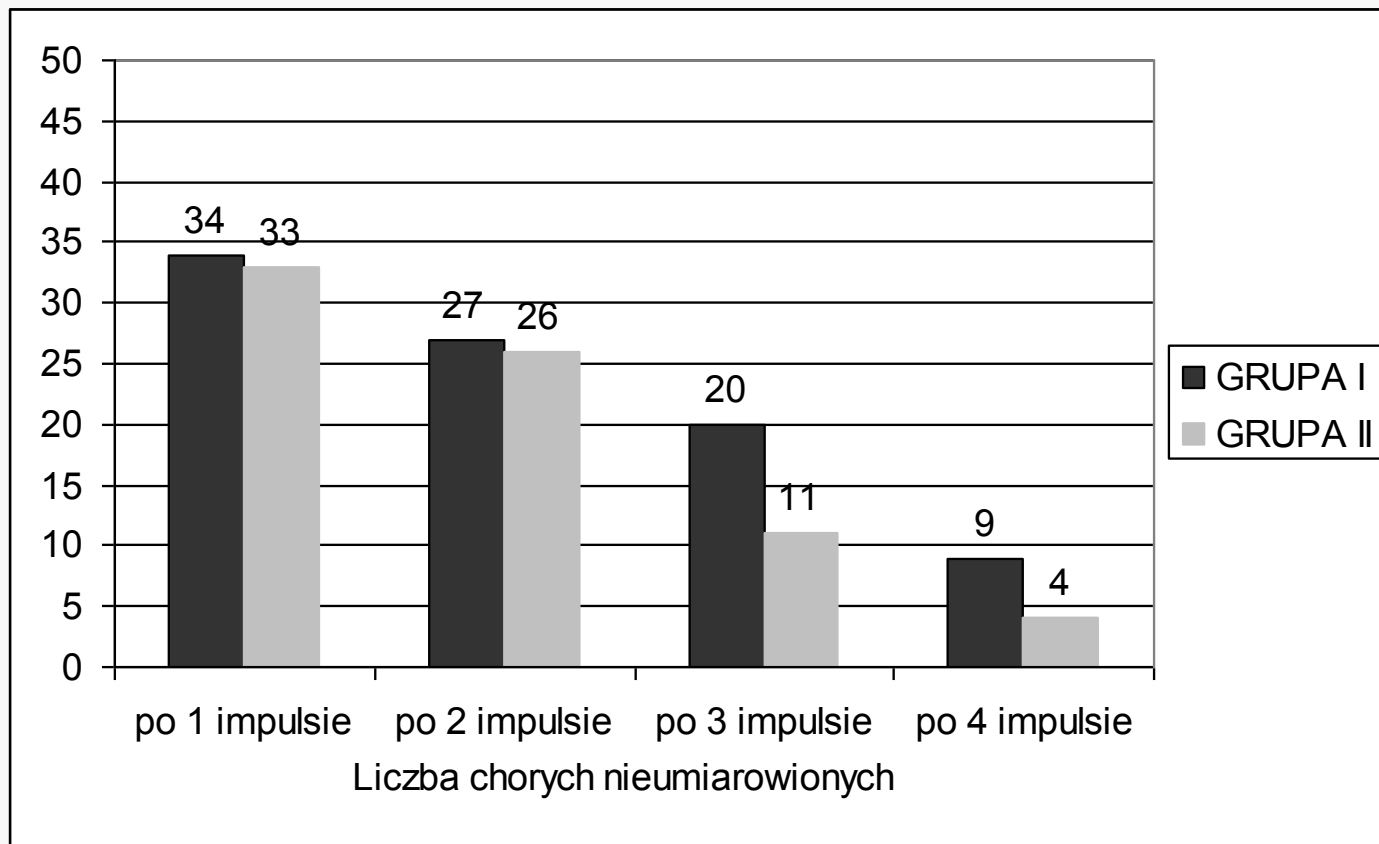
ODSETEK CHORYCH Z OPTYMALNĄ SEDACJĄ

■ Grupa I ■ Grupa II

* $p < 0,05$

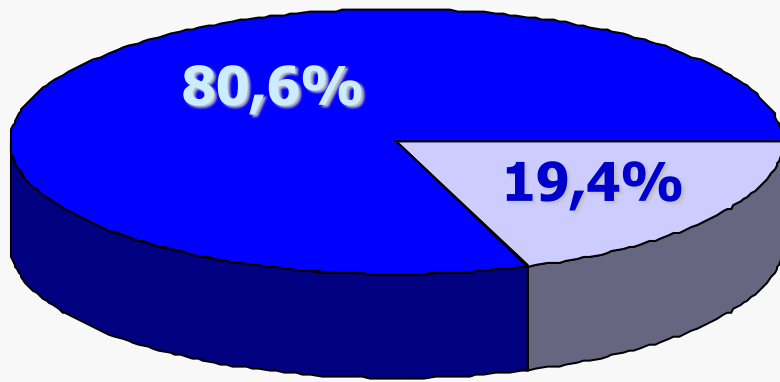


Porównanie nieskuteczności kardiowersji po zastosowaniu kolejnych dawek energii

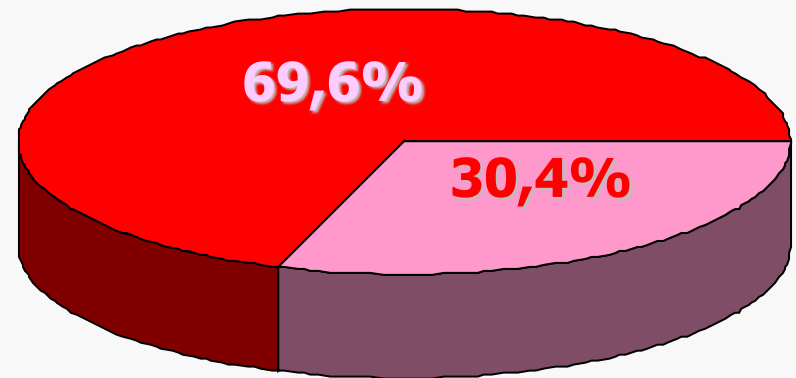


Ocena znieczulenia przez chore

pozytywna



pozytywna



negatywna

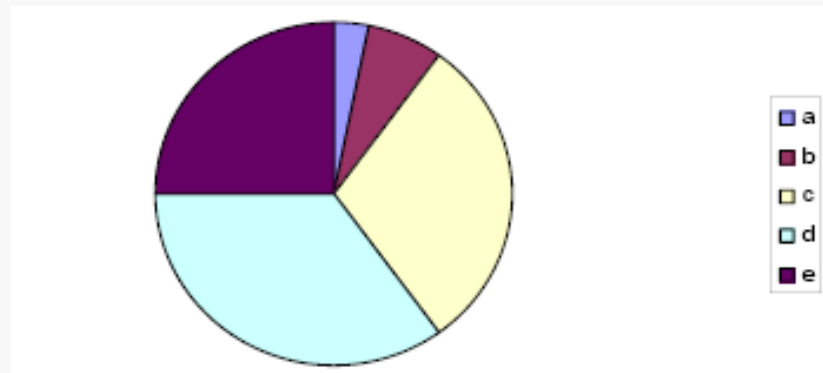
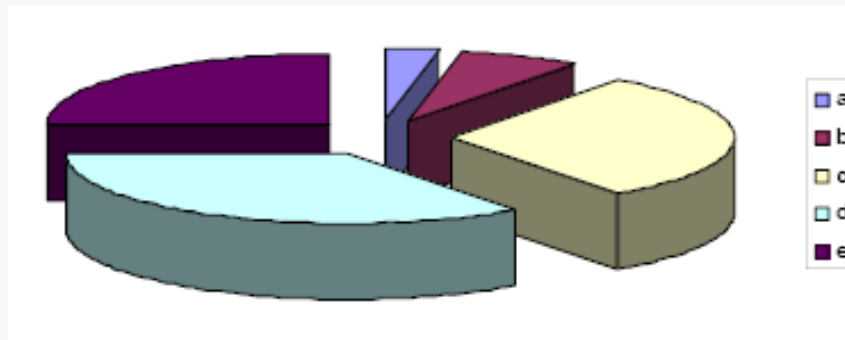
negatywna

NS

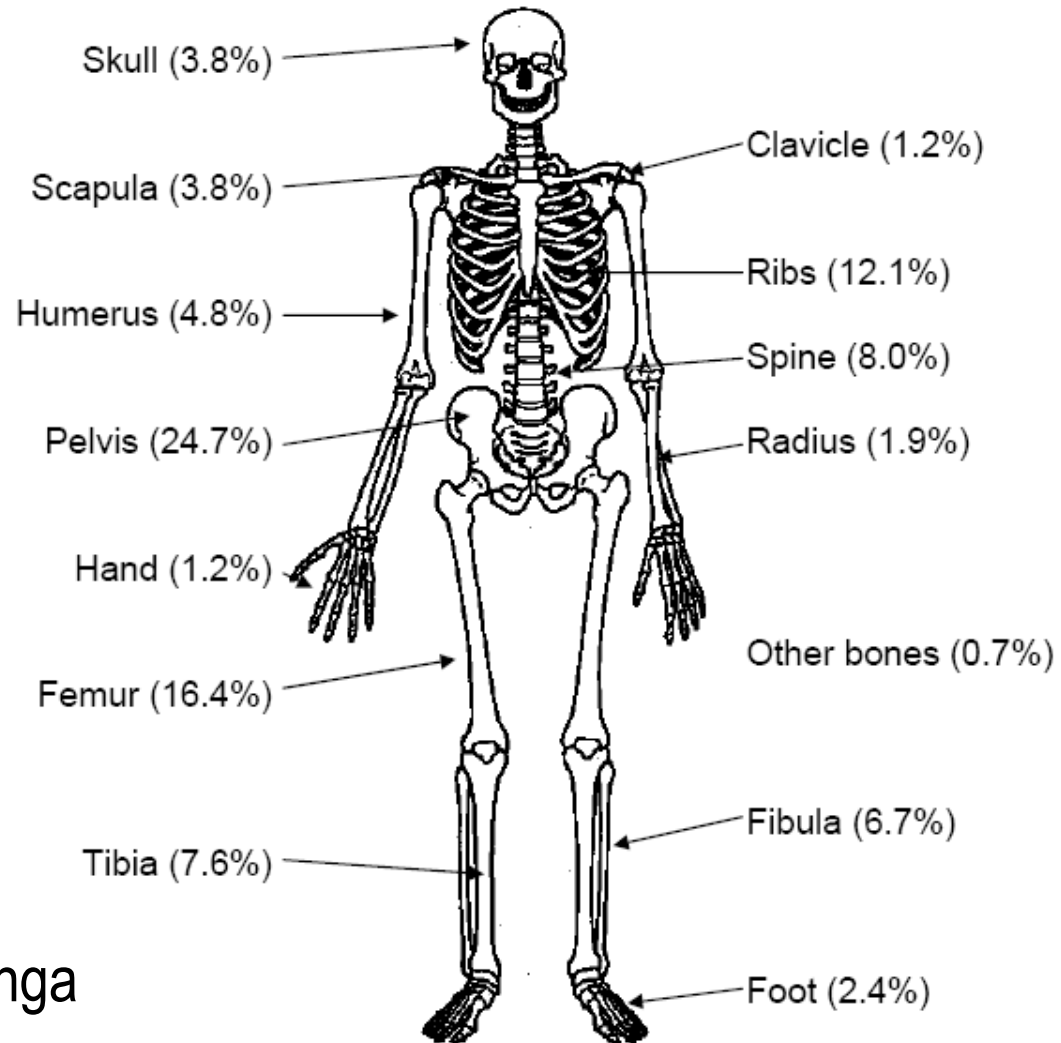
Błędy wizualizacji – na czym polegają?

- wszystko jest przedstawione, tylko nic z tego nie wiadomo
- nie da się tego zdefiniować, chodzi o zdrowy rozsądek

Nie



Tak



Lokalizacja guza Ewinga

Dane ilościowe

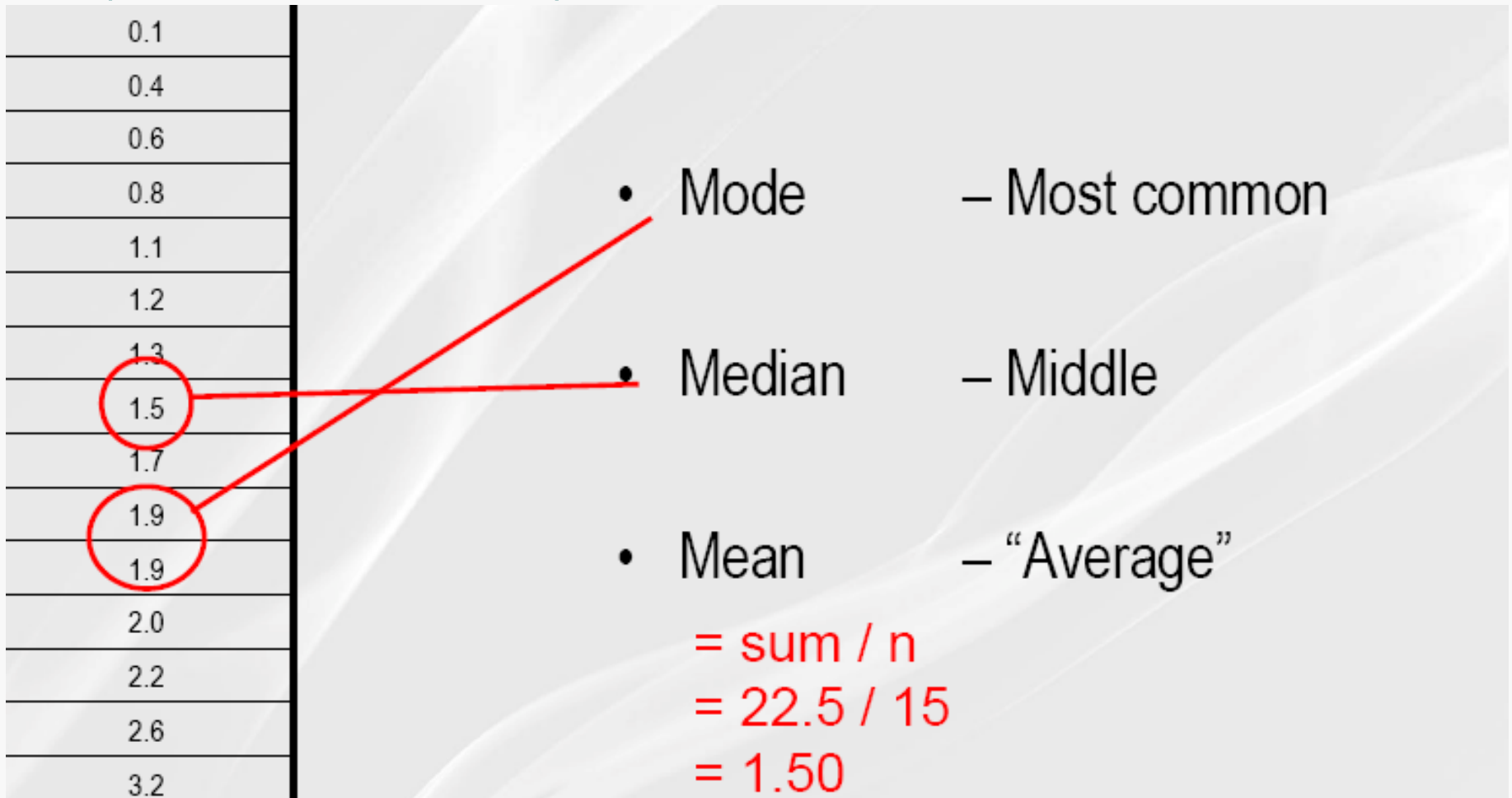
- **ciągłe** (*continuous data* - np. poziom cholesterolu)
- **nieciągłe** (*discrete data* - np. liczba przebytych zaostrzeń choroby)

Dane ilościowe - charakterystyka

- miary położenia – średnia, mediana, najczęstsza wartość (*mean, median, mode*)
- miary rozproszenia – zakres (*range*), odchylenie standardowe (*standard deviation*), odstęp pomiędzy dolnym i górnym kwartylem (*interquartile range*)

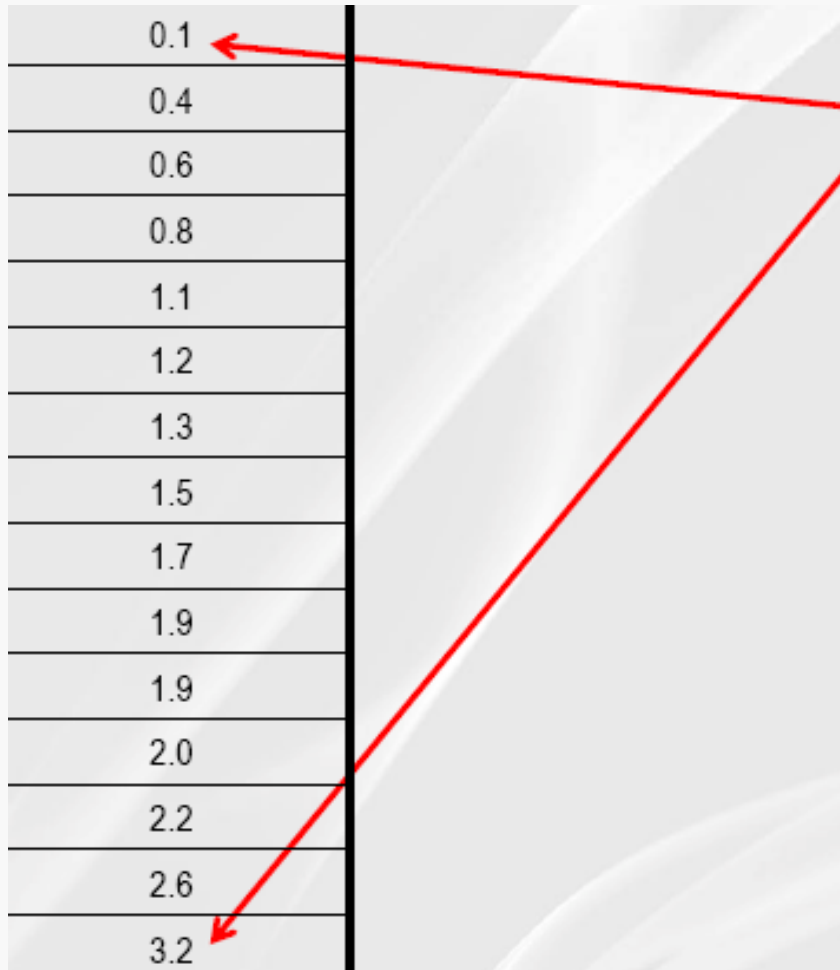
Miary położenia

(measure of location)



Miary rozproszenia

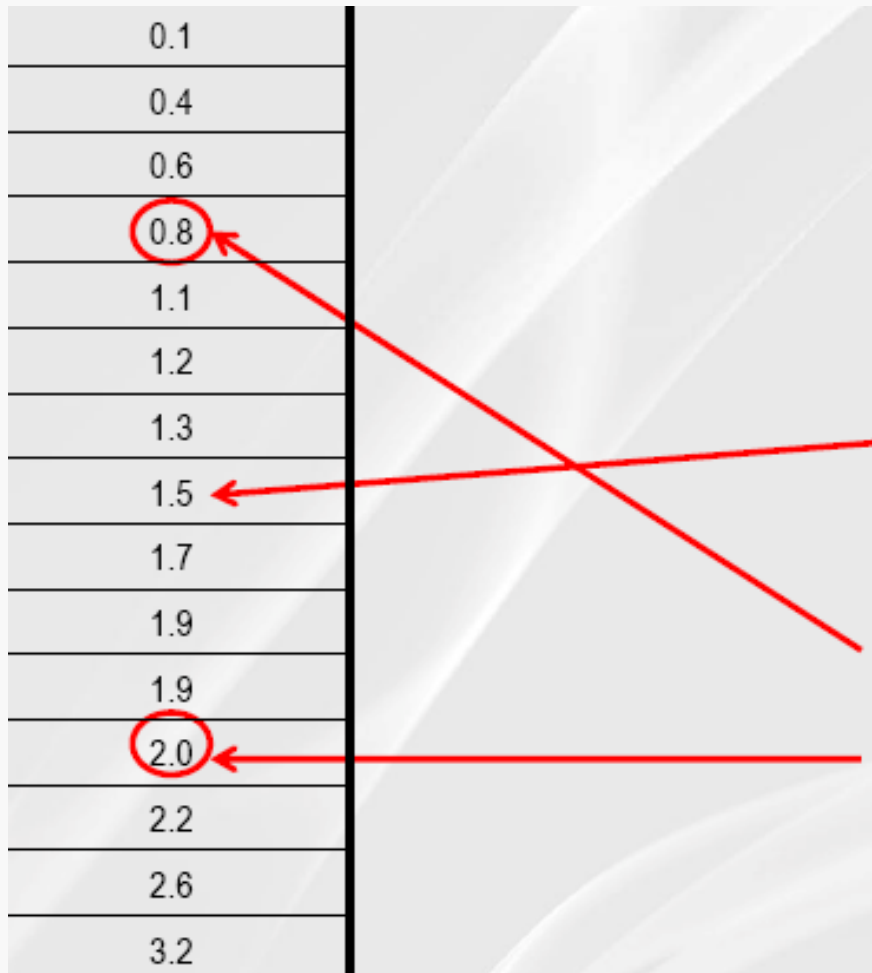
(measure of spread)



Zakres (*range*)
„0.1-3.2”

Miary rozproszenia

(measure of spread)



Mediana (50%)

Górny kwartył (25%)

Dolny kwartył (25%)

Odstęp pomiędzy dolnym i górnym kwartyłem

(*interquartile range*)

0.1
0.4
0.6
0.8
1.1
1.2
1.3
1.5
1.7
1.9
1.9
2.0
2.2
2.6
3.2

Interquartile range:
„0.8 – 2.0” lub „1.2”

W tym przedziale mieści się połowa wyników!

Odchylenie standardowe

(*standard deviation - SD*)

0.1	-1.4
0.4	-1.1
0.6	-0.9
0.8	-0.7
1.1	-0.4
1.2	-0.3
1.3	-0.2
1.5	0.0
1.7	0.2
1.9	0.4
1.9	0.4
2.0	0.5
2.2	0.7
2.6	1.1
3.2	1.7

Średnia (mean) wynosi 1.5

Średnie odchylenie od średniej wynosi 0

Zamiast tego oblicza się ile wynoszą odchylenia podniesione do kwadratu

W tym celu trzeba zsumować wszystkie odchylenia podniesione do kwadratu

Odchylenie standardowe

(*standard deviation - SD*)

0.1	-1.4	1.96
0.4	-1.1	1.21
0.6	-0.9	0.81
0.8	-0.7	0.49
1.1	-0.4	0.16
1.2	-0.3	0.09
1.3	-0.2	0.04
1.5	0.0	0
1.7	0.2	0.04
1.9	0.4	0.16
1.9	0.4	0.16
2.0	0.5	0.25
2.2	0.7	0.49
2.6	1.1	1.21
3.2	1.7	2.89

Suma odchyłeń podniesionych do kwadratu wynosi 9.96

Suma odchyłeń podzielona przez liczbę obserwacji (pomniejszoną o jeden) daje tzw. wariancję
 $9,96 / (n-1) = 0.71$

Pierwiastek kwadratowy z wariancji to odchylenie standardowe (czyli 0.84)

1.5 ± 0.84 (średnia ± SD)

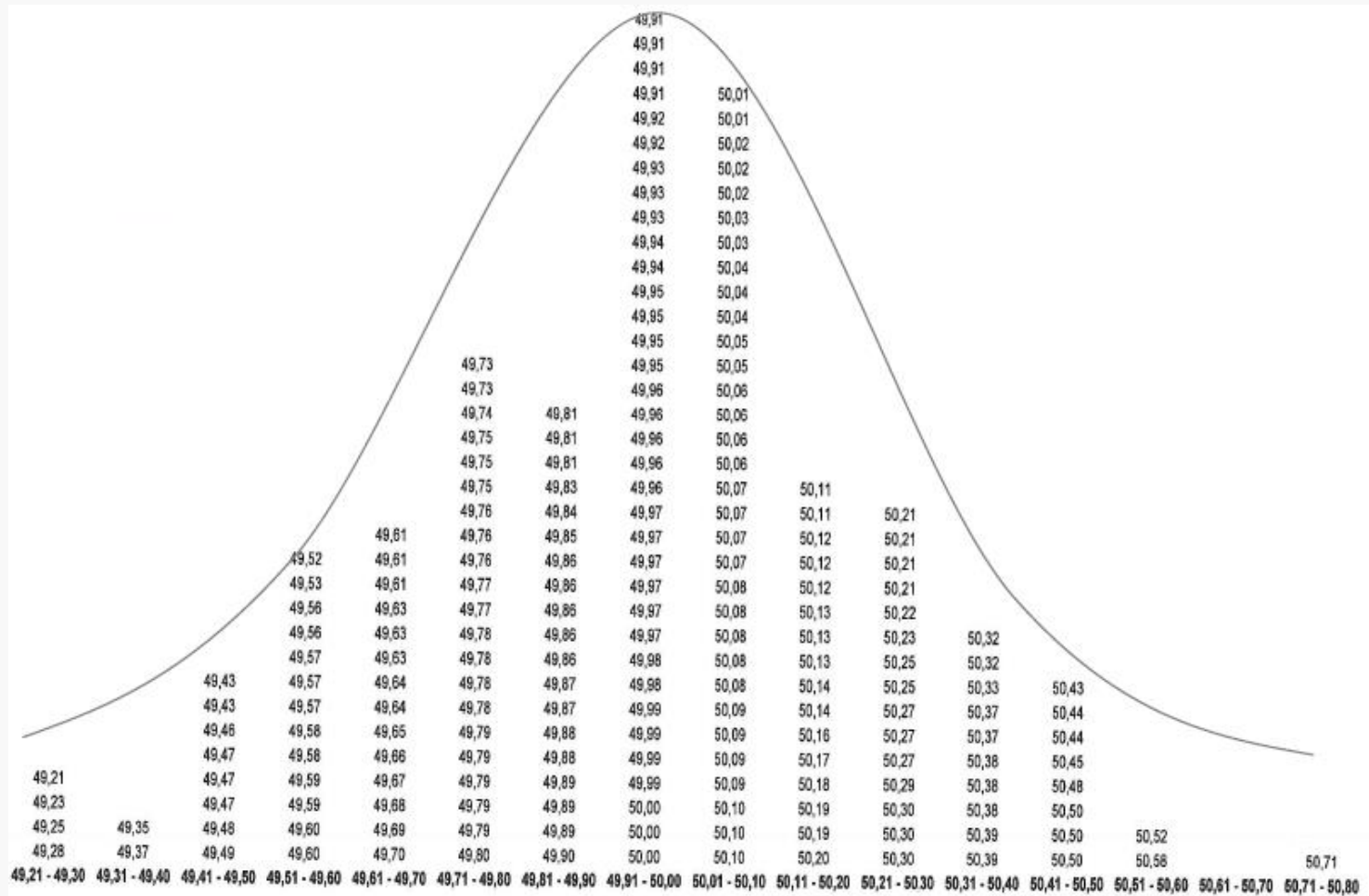
Odchylenie standardowe

(standard deviation - SD)

- do zastosowania **raczej** jeżeli zmienne mają rozkład normalny,
- testy parametryczne można stosować **tylko** jeżeli jest zmienne mają rozkład normalny,
- w przeciwnym przypadku stosuje się testy nieparametryczne,
- miarą położenia i rozproszenia dla testów nieparametrycznych jest mediana i zakres (a nie średnia i odchylenie standardowe!)

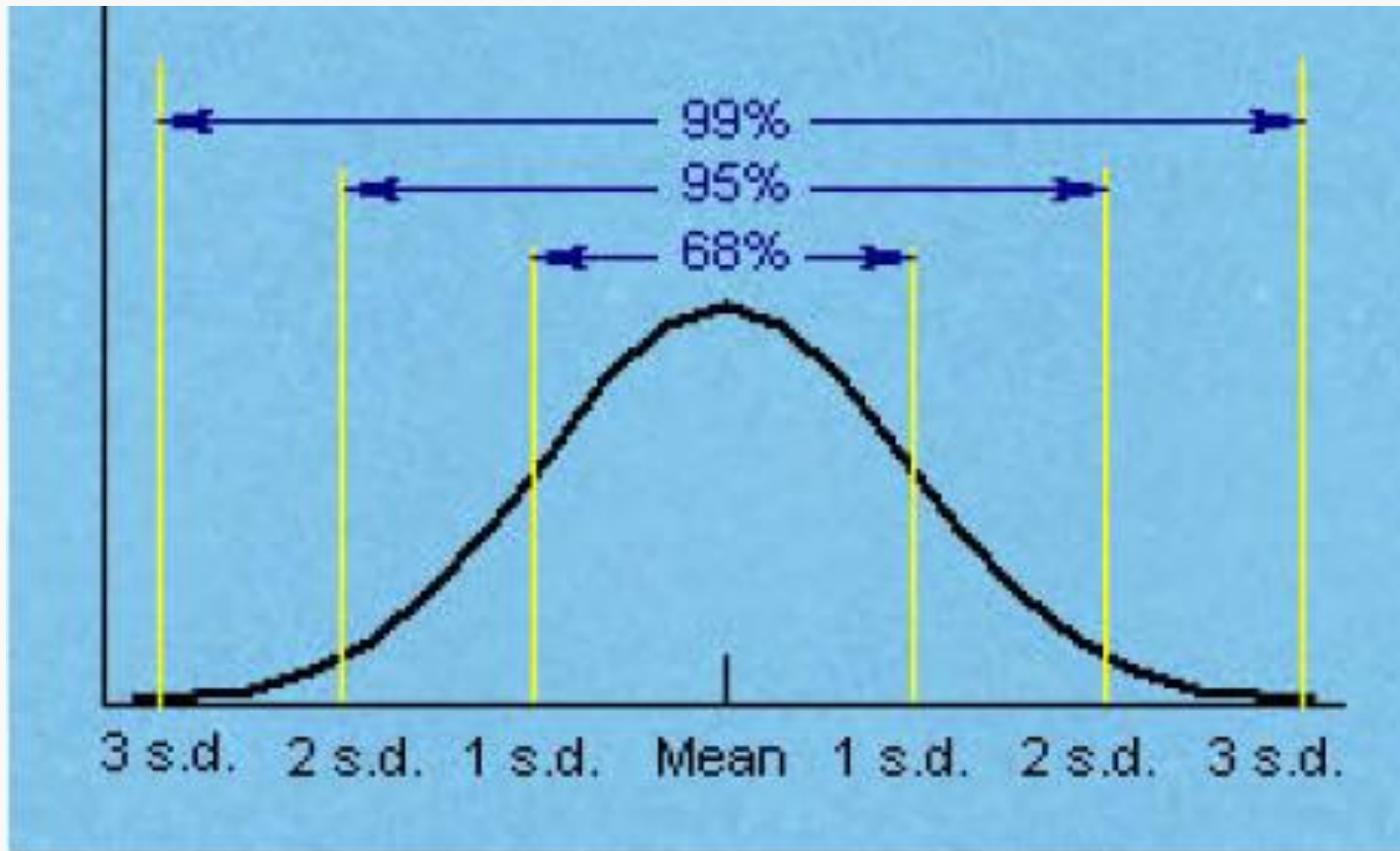
Rozkład normalny

(normal distribution)



Rozkład normalny

(normal distribution)



Błąd standardowy średniej

(standard error of the mean - SEM)

- SEM = SD : (pierwiastek z liczebności)
- SEM uwzględnia liczebność próbki, SD nie
- **95% przedział ufności = 1,96 x SEM**
- 95% przedział ufności (*95% confidence interval - 95%CI*)
to przedział który na 95% zawiera średnią z całej populacji

Błąd w statystyce

- **Typ I** (*type I error*) – stwierdzenie różnicy kiedy de facto jej nie ma
- **Typ II** (*type II error*) – stwierdzenie braku różnicy, gdy de facto różnica istnieje

Hipoteza zerowa

(null hypothesis)

- porównując próbki zakłada się (hipotetycznie) że nie ma pomiędzy nimi różnicy
- za pomocą testów statystycznych można odrzucić (lub nie) hipotezę zerową
- współczynnik istotności p mówi nam jaka jest szansa że hipoteza zerowa została błędnie odrzucona (czyli że popełniliśmy błąd typu I
- **$p < 0.05 = < 5\%$ szansy że wykazaliśmy różnicę, a tak naprawdę jej nie ma. Czyli mamy $> 95\%$ szansy na to że się nie pomyliliśmy!**

Testy parametryczne

- test t dla zmiennych niepowiązanych (dla porównania niezależnych próbek)
- test t dla zmiennych powiązanych (dla porównania powtarzanych obserwacji)
- test Pearsona (dla korelacji)

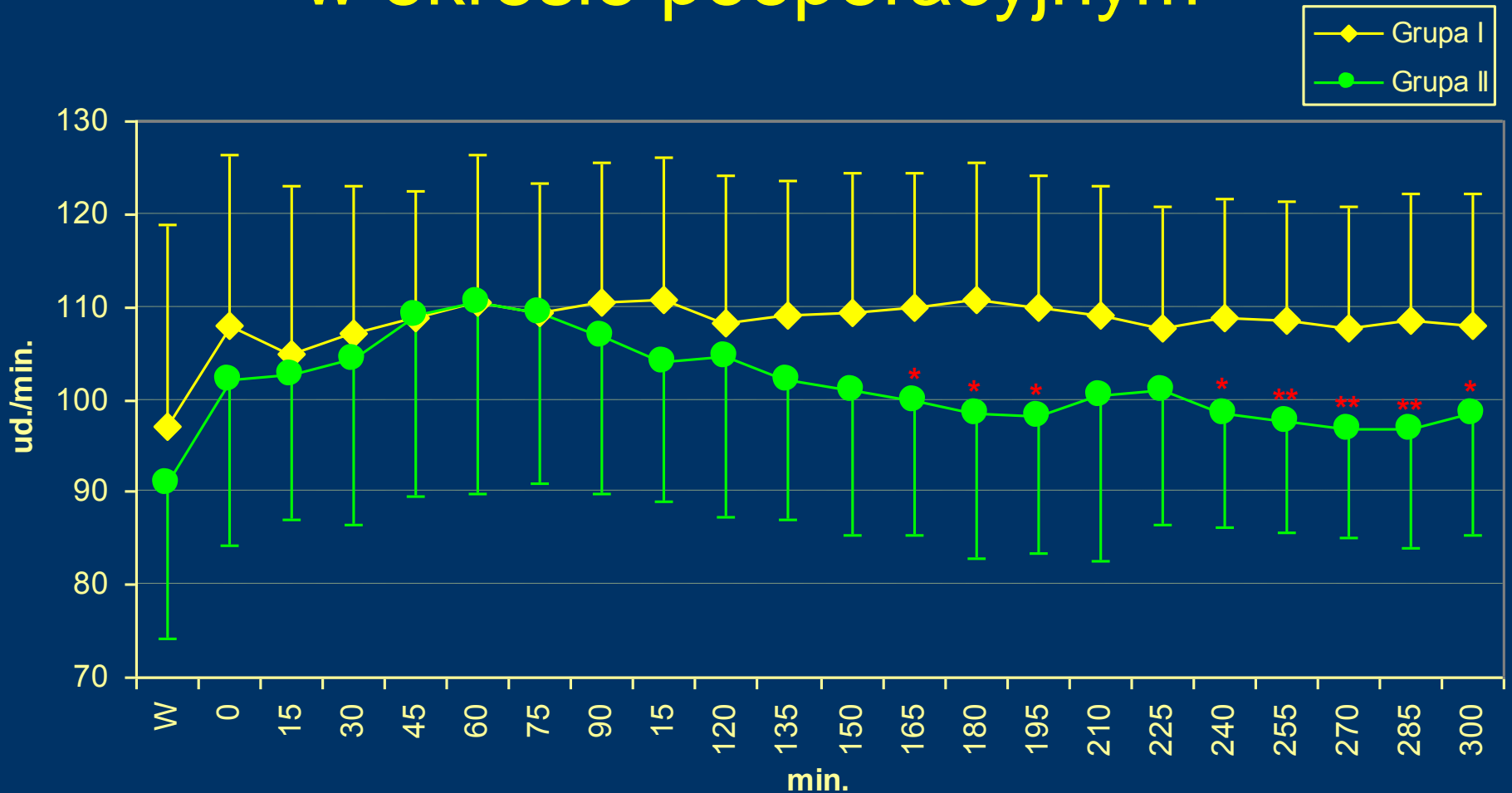
Testy nieparametryczne

- test Mann-Whitney'a dla zmiennych niepowiązanych (dla porównania niezależnych próbek)
- test Wilcoxona dla zmiennych powiązanych (dla porównania powtarzanych obserwacji)
- test Spearmana (dla korelacji)

Punktowa ocena stanu chorych przy przyjęciu do OIT

	Grupa I (n=82)	Grupa II (n=32)	P
APACHE II	25,7 ±4,4	25,5 ±9,5	NS
GCS	4,43 ±2,83	7,31 ±4,96	<0,01
TISS-28	35,8 ±5,1	40,6 ±6,9	<0,001

Częstość akcji serca w okresie pooperacyjnym



Zmienne jakościowe

- Test χ^2 (dla większych liczebności) lub test dokładny Fischera (dla mniejszych liczebności) - dla zmiennych niepowiązanych (dla porównania niezależnych próbek).
To jest popularna „czteropolówka”
- Test χ^2 McNemary dla zmiennych powiązanych (dla porównania powtarzanych obserwacji)

DZIAŁANIA NIEPOŻĄDANE

Rodzaj	Grupa I n=32	Grupa II n=29	p
Bradykardia < 60/min.	6 (18,8%)	1 (3,4%)	NS
Tachykardia > 90/min.	1 (3,1%)	4 (13,8%)	NS
Skurcze dodatkowe	3 (9,4%)	4 (13,8%)	NS
SBP < 90 mmHg	3 (9,4%)	1 (3,4%)	NS
SBP > 200 mmHg	2 (6,3%)	4 (13,8%)	NS
SpO ₂ < 93%	4 (12,5%)	11 (37,9%)	<0,05
Ból głowy	2 (6,9%)	0	NS
Suchość w ustach	2 (6,3%)	1 (3,4%)	NS
Nudności	3 (9,4%)	1 (3,4%)	NS
Zbyt głęboka sedacja (R6 ⁰)	6 (18,8%)	11 (37,9%)	NS

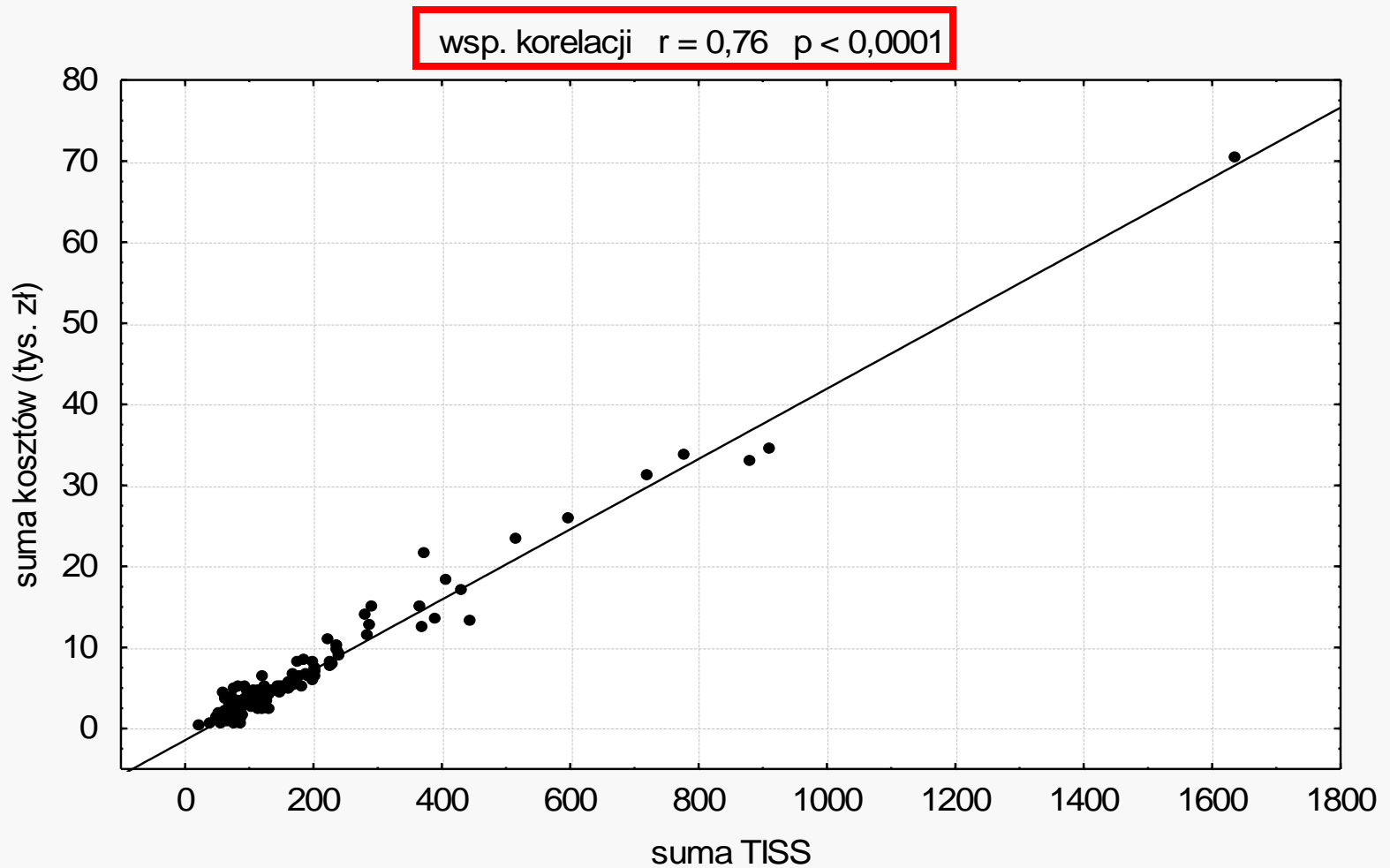
Uwaga - to była wersja skrócona!

Rodzaj	Grupa I n=32	Grupa II n=29	p
Bradykardia < 60/min.	6 (18,8%)	1 (3,4%)	NS
Bradykardia < 60/min. - TAK	6 (18,8%)	1 (3,4%)	
Bradykardia < 60/min. - NIE	26 (81,2%)	28 (96,8%)	NS

Korelacja

- określa związek pomiędzy zmiennymi
- ważny jest zarówno współczynnik korelacji r ($>0,5$), jak i współczynnik istotności p ($<0,05$)
- częsta nadinterpretacja wyników (p istotne bo duża próbka, ale co z tego skoro są niskie wartości r)
- **pokusa aby skorelować wszystko ze wszystkim – może coś istotnego się znajdzie? (*fishing expedition*)**

Korelacja łącznej punktacji TISS-28 i łącznych kosztów leczenia



Pytanie badawcze trzeba postawić na samym początku



Nie należy poszukiwać gorączkowo coraz to nowych wyników!

Wiadomo, że coś zawsze się znajdzie...



Ciąg dalszy nastąpi...