





# https://kiero.net edu@kiero.net


Przejdź do treści ZALOGUJ

**Kiero.net** | *lokujemy wyłącznie na Twoją obsługę 4 pliki cookie* | *stosujemy dyrektywy RODO*

[START](#) [CZAT](#)

AMiSNS 2022/2023		Tematów	Postów	Ostatni post
	Informatyka i Biostatystyka 2-semestr Lekarski <b>Moderowane przez:</b> Administrator	3	3	01-03-2023 11:30 przez kiero
	Statystyka medyczna Pielęgniarstwo	4	4	02-03-2023 20:58 przez kiero
	Informatyka i Biostatystyka Ratownictwo	3	3	02-03-2023 20:58 przez kiero

 - Nowe treści od Twojej ostatniej wizyty.

 - Brak nowych treści od Twojej ostatniej wizyty.

# Wstęp do statystyki

## Typologia zbiorów danych, etapy ich opracowania i porównywania.

Zmienne objaśniające i wynikowe.

Tablice statystyczne i skale pomiarowe stosowane do charakterystyki analizowanych danych.

Nomogram Altmana do obliczeń wielkości próby.

# Typologia zbiorów danych

**Zbiór danych** statystycznych jest to rejestr zebranych danych i zestawionych przeważnie w formie tabelarycznej.

W takim zbiorze danych kolumny zazwyczaj odpowiadają obserwowanym cechom statystycznym, wiersze zaś opisują dane przypadki z próby.

Dane zestawione w kolumnach i wierszach nazywamy macierzą, a sama macierz składa się z komórek.

W komórkach przechowujemy pojedyncze dane/zmienne zebrane podczas procesu pomiaru/obserwacji.

# Typologia zbiorów danych

Nr	Rzut piłką lek.	Szybkość	Moc	Zwinność	Wytrzymałość	Bieg 20m	skok w dal z m.	Pch. kulą	Bieg 300m	Ciężar c.	Wysokość c.
1	9,50	8,60	42	16,50	30	2,5	2,10	9,28	50,30	81,80	186,10
2	6,20	8,10	49	15,00	34	2,2	2,40	8,20	45,60	53,00	170,80
3	7,00	9,00	45	16,00	32	2,3	2,20	8,25	53,00	60,70	169,80
4	9,90	8,00	56	13,70	30	2,2	2,35	9,50	47,40	68,00	166,00
5	11,10	8,20	45	15,40	31	2,9	2,40	9,40	46,80	81,70	183,50
6	9,90	8,00	50	16,00	30	2,3	2,80	9,20	46,70	71,40	175,10
7	12,00	7,50	52	14,50	28	2,3	2,85	10,40	43,30	72,60	183,00
8	7,90	8,00	45	14,50	32	2,3	2,40	8,10	46,50	60,80	167,20
9	9,90	7,70	40	15,80	31	2,3	2,50	8,50	44,30	63,90	173,60
10	7,70	8,20	40	13,20	31	2,6	2,55	8,80	44,20	67,30	176,10
11	10,40	7,60	50	14,50	34	2,2	2,84	10,00	45,00	70,60	178,10
12	7,20	8,20	48	16,00	30	2,4	2,30	8,40	48,60	54,90	170,00
13	7,50	8,40	39	16,00	30	2,3	2,40	9,00	46,20	70,60	173,70
14	10,60	7,00	49	16,00	31	2,3	2,65	11,00	43,60	83,20	181,10
15	8,90	7,80	48	14,00	32	2,2	2,40	9,00	44,50	73,00	174,60
16	9,40	8,00	47	15,00	30	2,3	2,44	10,50	43,80	67,40	176,40
17	9,90	8,20	49	15,00	32	2,3	2,15	9,10	44,50	61,00	171,40
18	10,10	7,40	53	14,80	30	2,2	2,85	12,30	48,90	66,00	181,60
19	9,80	8,40	54	15,00	31	2,5	2,65	8,90	41,30	59,00	167,70
20	9,50	7,40	41	15,00	30	2,1	2,55	11,80	45,80	72,20	175,20

# Typologia zbiorów danych

A1										
	A	B	C	D	E	F	G	H	I	J
1	Badany nr	rzut piłką lek.	Szybkość	Moc	Zwinność	Wytrzymałość	Bieg 20m	ok w dal z 10m	Pch. kulą	Bieg 300m
2	1	9,5	8,6	42	16,5	30	2,5	2,1	9,28	50,3
3	2	6,2	8,1	49	15	34	2,2	2,4	8,2	45,6
4	3	7	9	45	16	32	2,3	2,2	8,25	53
5	4	9,9	8	56	13,7	30	2,2	2,35	9,5	47,4
6	5	11,1	8,2	45	15,4	31	2,9	2,4	9,4	46,8
7	6	9,9	8	50	16	30	2,3	2,8	9,2	46,7
8	7	12	7,5	52	14,5	28	2,3	2,85	10,4	43,3
9	8	7,9	8	45	14,5	32	2,3	2,4	8,1	46,5
10	9	9,9	7,7	40	15,8	31	2,3	2,5	8,5	44,3
11	10	7,7	8,2	40	13,2	31	2,6	2,55	8,8	44,2
12	11	10,4	7,6	50	14,5	34	2,2	2,84	10	45
13	12	7,2	8,2	48	16	30	2,4	2,3	8,4	48,6
14	13	7,5	8,4	39	16	30	2,3	2,4	9	46,2
15	14	10,6	7	49	16	31	2,3	2,65	11	43,6
16	15	8,9	7,8	48	14	32	2,2	2,4	9	44,5
17	16	9,4	8	47	15	30	2,3	2,44	10,5	43,8
18	17	9,9	8,2	49	15	32	2,3	2,15	9,1	44,5
19	18	10,1	7,4	53	14,8	30	2,2	2,85	12,3	48,9
20	19	9,8	8,4	54	15	31	2,5	2,65	8,9	41,3
21	20	9,5	7,4	41	15	30	2,1	2,55	11,8	45,8
22	21	8,5	7,9	47	14,5	31	2,4	2,35	9,4	44,7
23	22	8,55	8,1	48	13,2	33	2,4	2,4	9,2	45,6
24	23	9,1	8	40	14,2	31	2,4	2,5	9,6	51,8
25	24	6,3	8,8	42	16,2	30	2,4	2,1	7,2	49,7
26	25	7	8,3	39	16	30	2,4	2,2	8,2	46,8

# Typologia zbiorów danych

H2									
NIE									
	A	B	C	D	E	F	G	H	I
1	LP	CZAS1	ucięte1	CZAS-placebo	ucięta-placebo	Grupująca	Czas	ucięte XX	
2	1	6	NIE	1	NIE	L	6	NIE	
3	2	6	NIE	1	NIE	L	6	NIE	
4	3	6	NIE	2	NIE	L	6	NIE	
5	4	7	NIE	2	NIE	L	7	NIE	
6	5	10	NIE	3	NIE	L	10	NIE	
7	6	13	NIE	4	NIE	L	13	NIE	
8	7	16	NIE	4	NIE	L	16	NIE	
9	8	22	NIE	5	NIE	L	22	NIE	
10	9	23	NIE	5	NIE	L	23	NIE	
11	10	6	TAK	8	NIE	L	6	TAK	
12	11	9	TAK	8	NIE	L	9	TAK	
13	12	10	TAK	8	NIE	L	10	TAK	
14	13	11	TAK	8	NIE	L	11	TAK	
15	14	17	TAK	11	NIE	L	17	TAK	
16	15	19	TAK	11	NIE	L	19	TAK	
17	16	20	TAK	12	NIE	L	20	TAK	
18	17	25	TAK	12	NIE	L	25	TAK	
19	18	32	TAK	15	NIE	L	32	TAK	
20	19	32	TAK	17	NIE	L	32	TAK	
21	20	34	TAK	22	NIE	L	34	TAK	
22	21	35	TAK	23	NIE	L	35	TAK	
23						P	1	NIE	
24						P	1	NIE	
25						P	2	NIE	
26						P	2	NIE	

# Typologia zbiorów danych

W statystyce praktycznej zbiory danych zwykle stanowią wyniki obserwacji pewnej próby statystycznej.

W statystyce teoretycznej tworzy się też zbiory danych sztucznie (np. w celu sprawdzenia jak dana metoda statystyczna reaguje na różne rozkłady w populacji).

# Typologia zbiorów danych



Author, Year	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Total	Risk of bias
ALTOMONTE et al, 2013 [30]	Y	U	U	Y	Y	Y	Y	Y	U	66,7%	MODERATE
ASCIERTO et al, 2017 [49]	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%	LOW
CAMACHO et al, 2009 [26]	U	Y	U	Y	Y	Y	Y	Y	U	66,7%	MODERATE
DIKA et al, 2017 [50]	U	U	U	U	U	Y	Y	U	U	22,2%	HIGH
EGGERMONT et al, 2016 [45]	Y	Y	U	Y	Y	Y	Y	Y	U	77,7%	LOW
EGGERMONT et al, 2018 [59]	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%	LOW
HAMID et al, 2013 [31]	U	Y	U	Y	Y	Y	Y	Y	U	66,7%	MODERATE
HODI et al, 2016 [46]	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%	LOW
HUA et al, 2016 [47]	U	N	U	N	U	Y	Y	Y	U	33,3%	HIGH
JUNG et al, 2017 [51]	U	U	U	Y	U	Y	Y	Y	Y	55,6%	MODERATE
KU et al, 2010 [27]	U	Y	U	Y	U	Y	Y	Y	U	55,5%	MODERATE
LARKIN et al, 2018 [60]	Y	Y	Y	Y	Y	Y	Y	Y	Y	100%	LOW

Abbreviations: Y= Yes; N= No; U= Unclear; NA= Not/Applicable; Q= Questions.

---

Q1 – Was the sample frame appropriate to address the target population?

Q2- Were study participants recruited in an appropriate way?

Q3- Was the sample size adequate?

Q4- Were the study subjects and setting described in detail?

Q5- Was data analysis conducted with sufficient coverage of the identified sample?

---



## Typologia zbiorów danych

Kolumny reprezentujące zmienne na skali interwałowej lub przedziałowej zawierają w komórkach liczby rzeczywiste.

Kolumny opisujące w komórkach zmienne na skali nominalnej i porządkowej mogą zawierać liczby lub etykiety tekstowe.

Ponadto w dowolnej komórce może także wystąpić przypadek braku danych, który musi być w jakiś sposób możliwy do zidentyfikowania – niekiedy rezerwuje się dla niego specjalną wartość liczbową lub znakową.

# Typologia zbiorów danych

# Typologia zbiorów danych

## zmienne niezależne i zmienne zależne

A1														
	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	Lp.	pleć	wiek	wysokość c. [cm]	ciężar c.[kg]	BMI	BMI>25	c.sk.nl	c.roz.nl	puls.nl	c.sk.pl	c.roz.pl	puls.pl	
2	1	k	54	169	79,0	27,66	TAK	99	71	70	100	65	74	
3	2	k	21	184	98,3	29,03	TAK	135	103	77	135	90	60	
4	3	k	32	171	77,9	26,64	TAK	120	87	59	110	74	54	
5	4	k	42	177	86,7	27,67	TAK	137	103	89	124	99	70	
6	5	k	59	160	74,0	28,91	TAK	153	109	78	140	88	70	
7	6	k	51	180	96,0	29,63	TAK	153	109	78	145	88	75	
8	7	k	52	159	70,0	27,69	TAK	135	99	76	115	80	86	
9	8	k	50	179	79,9	24,94	NIE	179	146	61	140	90	60	
10	9	k	20	173	67,7	22,62	NIE	172	119	84	150	90	73	
11	10	k	28	174	71,4	23,58	NIE	209	162	67	160	100	75	
12	11	k	36	180	69,3	21,39	NIE	156	117	78	142	96	80	
13	12	k	31	167	66,5	23,84	NIE	149	117	100	138	98	80	
14	13	k	59	180	65,0	20,06	NIE	145	101	79	121	74	70	
15	14	k	60	153	48,0	20,5	NIE	135	99	76	115	91	86	
16	15	k	68	160	55,0	21,48	NIE	145	91	84	110	78	70	
17	16	k	21	175	68,0	22,2	NIE	157	120	74	133	98	73	

- Zmienne objaśniane** – zmienne, których wartości są szacowane (wyznaczane) przez model; znajdujące się „po lewej stronie” równości.
- Zmienne objaśniające** – zmienne, za pomocą których wyjaśniane są zm. objaśniane; znajdujące się „po prawej stronie” równości, najczęściej z parametrami strukturalnymi.

# Typologia zbiorów danych

## Skale pomiarowe

Skala pomiarowa jest określonym systemem zapisu danych (symboli) pozwalającym na usystematyzowanie zebranych pomiarów statystycznych.

System zapisu/symbol pozwala scharakteryzować badane przez nas obiekty względem określonej zmiennej.

Wyróżnia się różne skale ze względu na relacje jakie jesteśmy w stanie wyróżnić pomiędzy poszczególnymi danymi.

# Typologia zbiorów danych

## Skale pomiarowe

Wyróżniamy 4 podstawowe rodzaje skal pomiarowych:

- skala nominalna
- skala porządkowa
- skala przedziałowa
- skala ilorazowa

### ***Skala nominalna***

Przykłady: Nazwa miejscowości, płeć, grupa krwi, imiona,

### ***Skala porządkowa***

Przykłady: Poziom wykształcenia, kolejność w rankingu.

### ***Skala przedziałowa/interwałowa***

Przykłady: Temperatura, data urodzenia.

### ***Skala ilorazowa***

Przykłady: Masa, długość, wysokość.

# Typologia zbiorów danych

## Skale pomiarowe

### *Skala nominalna*

Przykłady: Nazwa miejscowości, płeć, grupa krwi, imiona,

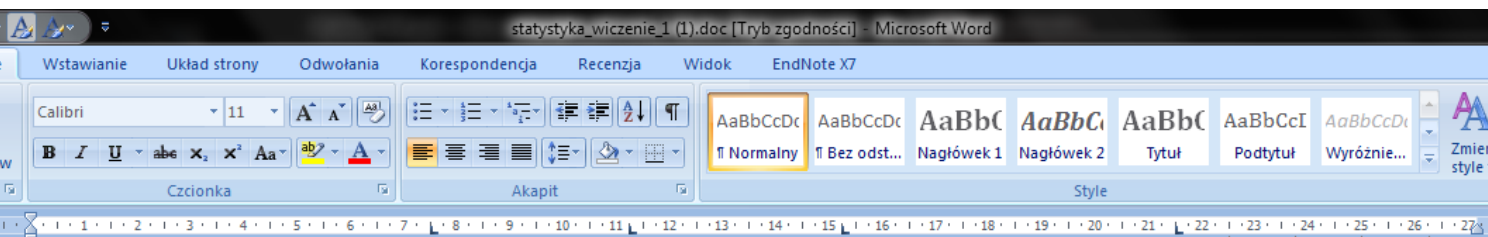
### *Skala ilorazowa*

Przykłady: Masa, długość, wysokość.

Nazwa uniwersytetu (cecha nominalna)	Poziom edukacji (cecha porządkowa)	Data rozpoczęcia sesji letniej (cecha interwałowa)	Ilość studentów (cecha ilorazowa)
Szkoła Handlowa Główna	Wysoki	14. czerwca 2019	32300
Uniwersytet Ekonomiczny	Bardzo wysoki	21. Czerwca 2019	12760
Politechnika Techniczna	Bardzo wysoki	13. Czerwca 2019	18710
SGWG	Dostateczny	28. Czerwca 2019	21290

Rejestracja, porządkowanie  
i przetwarzanie danych w celu ich  
dalszej analizy.

# Przykłady „na żywo”



tablica 1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

tablica 2

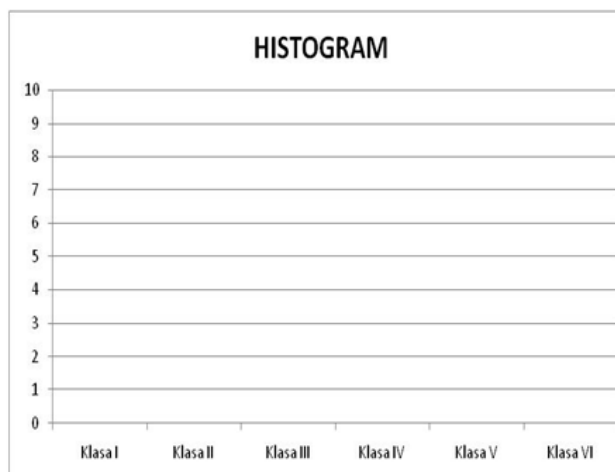
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

tablica 3

Opis	Wartość
Minimum	
Kwartyl 1	
Mediana	
Kwartyl 3	
Maksimum	

tablica 4

Klasa	Od	Do	Ilość	Środek klasy
I	Min			
II				
III				
IV				
V				
VI		Maks.		

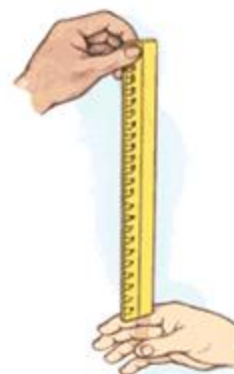


Nr	Rzut piłką	Szybkość	Moc	Zwinność	Wytrzymałość	Bieg 20m	skok w dal
1	9,50	8,60	42	16,50	30	2,5	2,10
2	6,20	8,10	49	15,00	34	2,2	2,40
3	7,00	9,00	45	16,00	32	2,3	2,20
4	9,90	8,00	56	13,70	30	2,2	2,35
5	11,10	8,20	45	15,40	31	2,9	2,40
6	9,90	8,00	50	16,00	30	2,3	2,80
7	12,00	7,50	52	14,50	28	2,3	2,85
8	7,90	8,00	45	14,50	32	2,3	2,40
9	9,90	7,70	40	15,80	31	2,3	2,50
10	7,70	8,20	40	13,20	31	2,6	2,55
11	10,40	7,60	50	14,50	34	2,2	2,84
12	7,20	8,20	48	16,00	30	2,4	2,30
13	7,50	8,40	39	16,00	30	2,3	2,40
14	10,60	7,00	49	16,00	31	2,3	2,65
15	8,90	7,80	48	14,00	32	2,2	2,40
16	9,40	8,00	47	15,00	30	2,3	2,44
17	9,90	8,20	49	15,00	32	2,3	2,15
18	10,10	7,40	53	14,80	30	2,2	2,85



## BADANIE CZASU REAKCJI PROSTEJ – LINIJKA

Badanie czasu reakcji prostej przeprowadzono przy pomocy testu „LINIJKA”. Test składał się z trzech prób. Prowadzący badania w dowolnym momencie wypuszczał linijkę (foto poniżej), a badany musiał jak najszybciej ją uchwycić. Wynik w mierzony w centymetrach z dokładnością do 0,5 cm zapisywano w arkuszu badań dla tegoż testu (patrz załącznik wyników badań). Wyniki następnie przeliczono na milisekundy ze wzoru na spadek swobodny ciała (patrz wzór).



How to hold the ruler

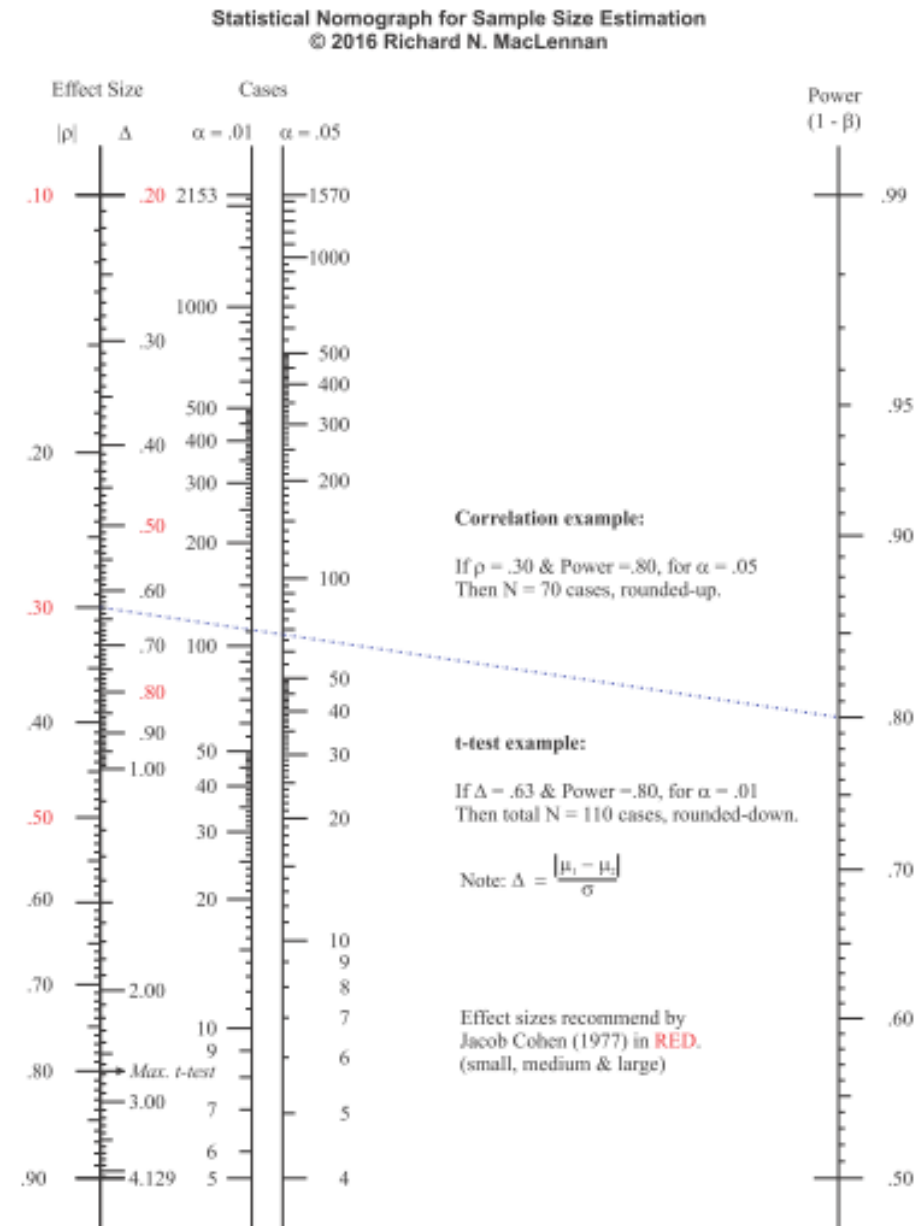
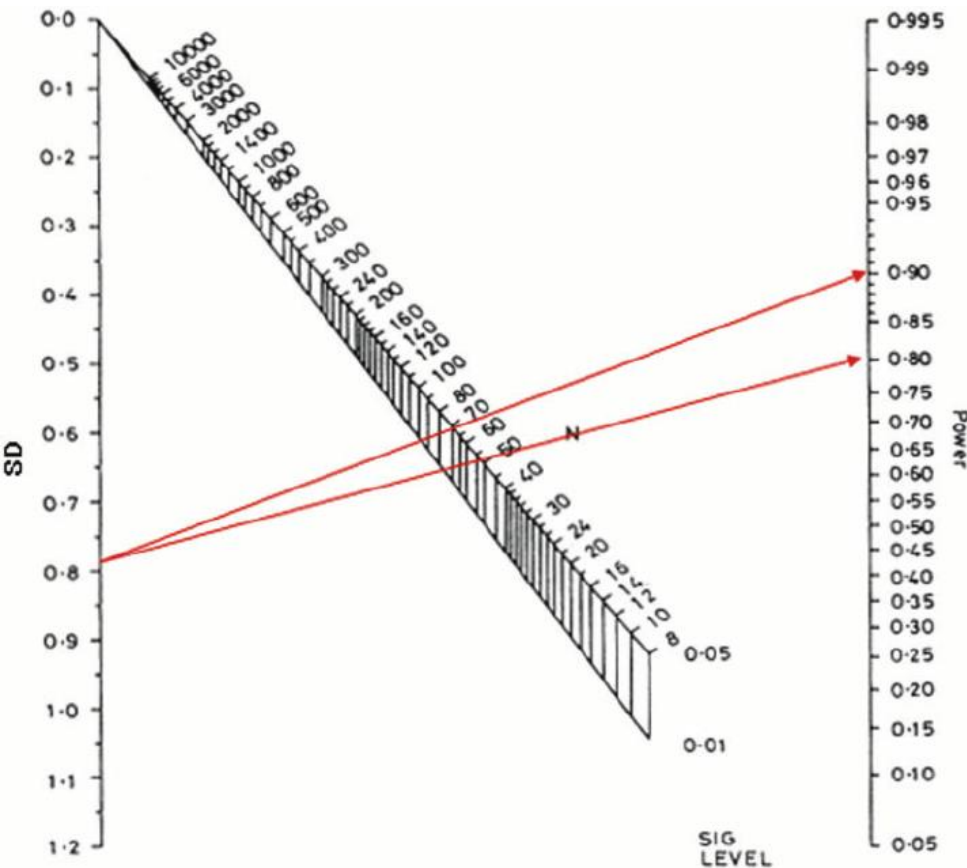
WZOR:

$$t = \sqrt{\frac{2h}{g}}, \text{ gdzie}$$



Nomogram Altmana do obliczeń  
wielkości próby.

# Nomogramy do obliczeń wielkości próby.



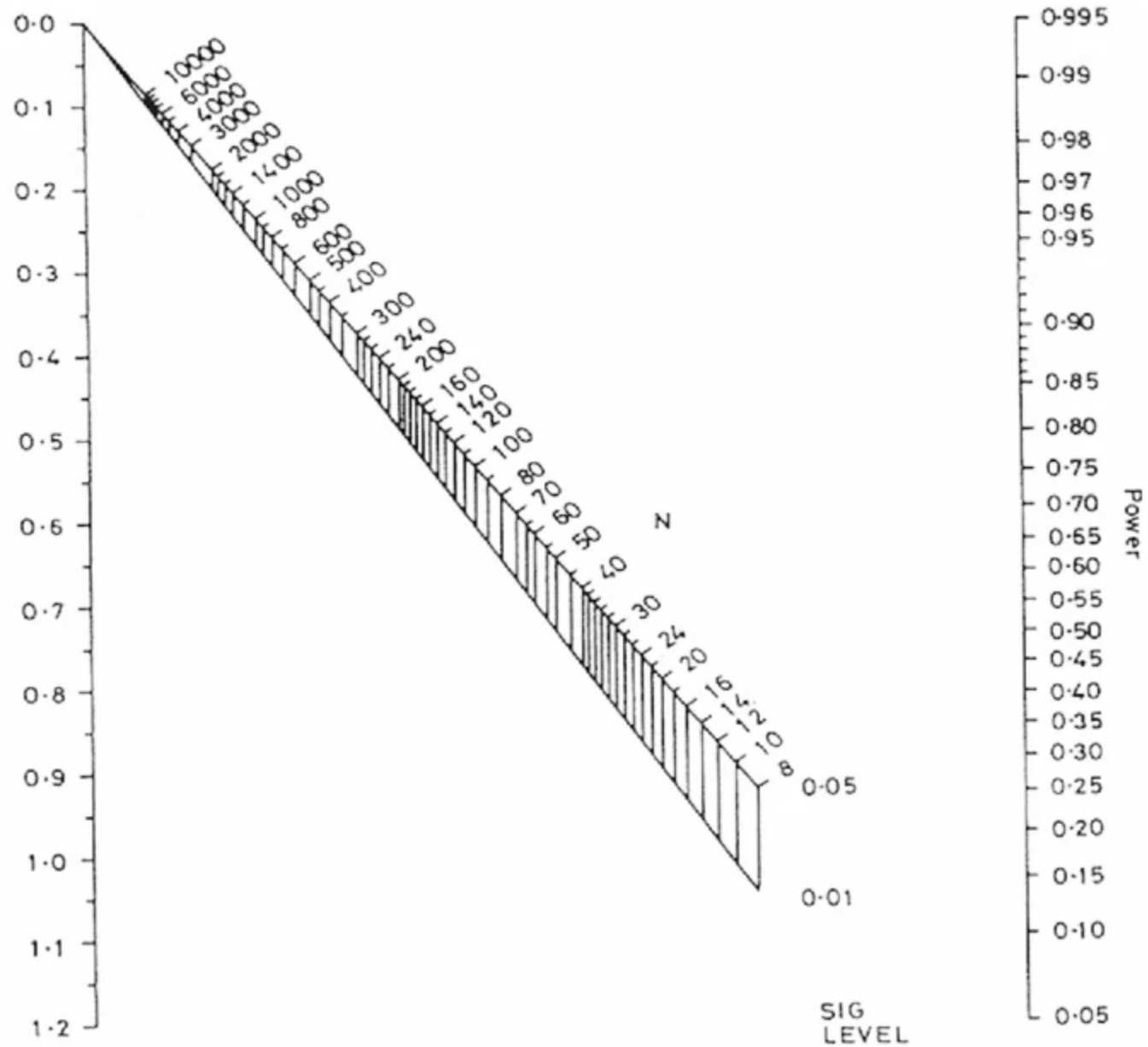
# Nomogram Altmana

Badania naukowe są prowadzone z myślą o wielu różnych celach. Można przeprowadzić badanie w celu ustalenia różnicy lub odwrotnie, podobieństwa między dwiema grupami zdefiniowanymi pod względem określonego czynnika ryzyka lub schematu leczenia.

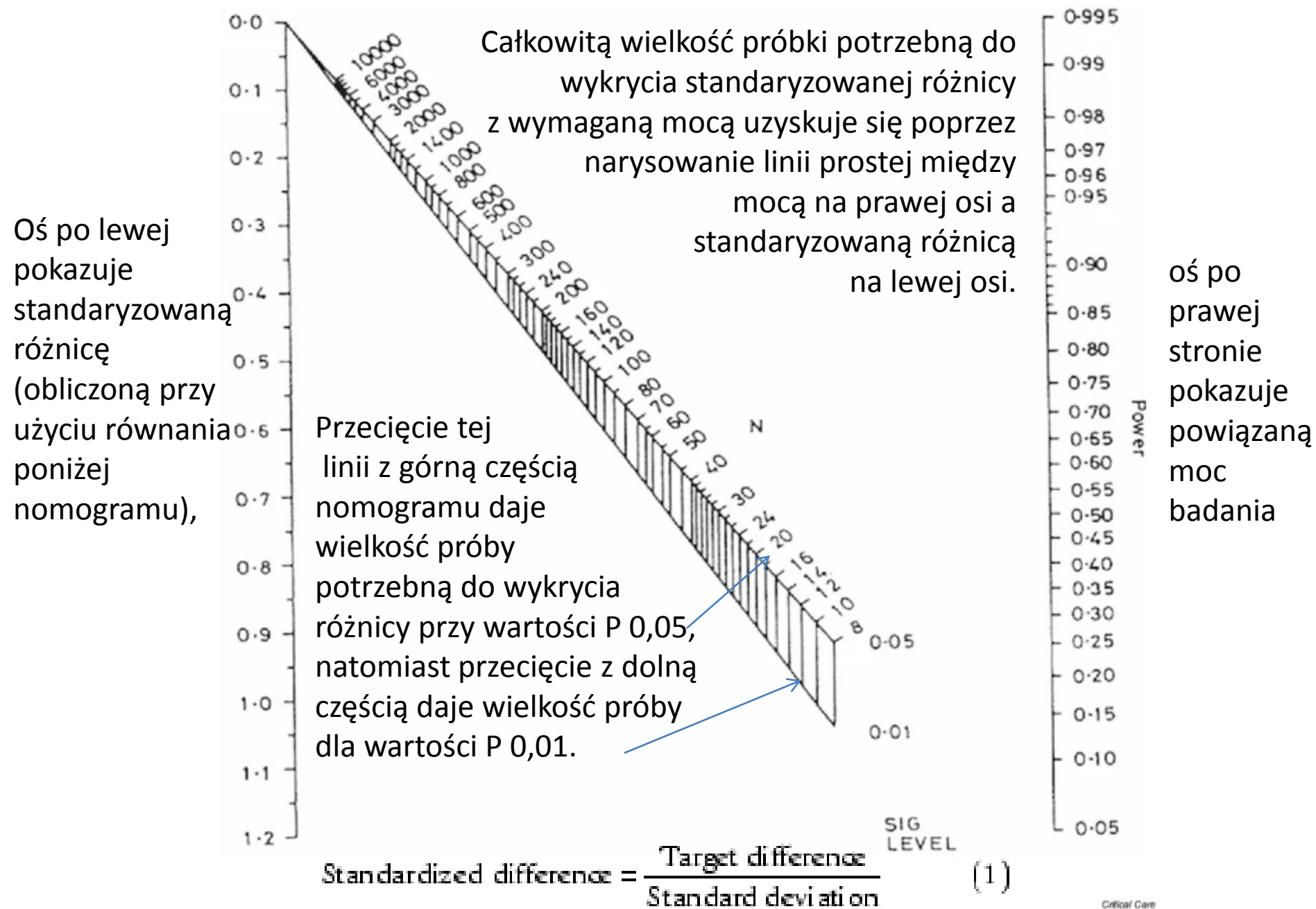
Bez względu na motywację badania, istotne jest, aby miało ono odpowiednią wielkość, aby osiągnąć zamierzone cele.

Najczęstszym celem jest prawdopodobnie określenie pewnej różnicy między dwiema grupami

# Nomogram Altmana



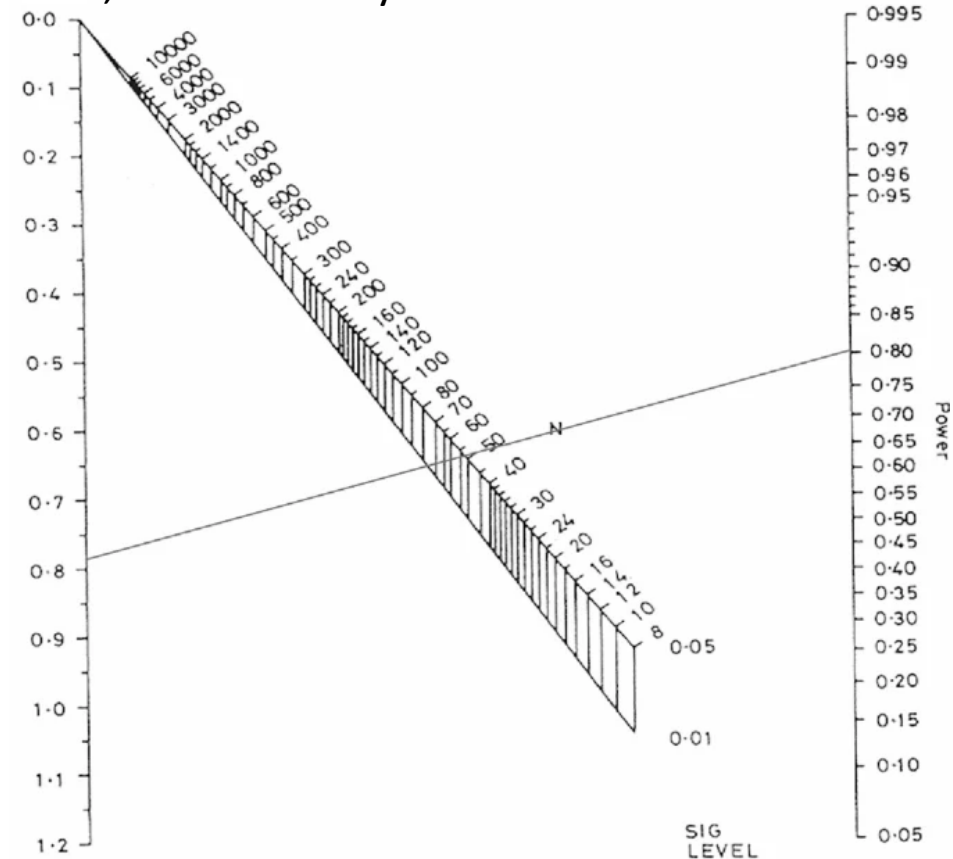
# Nomogram Altmanna



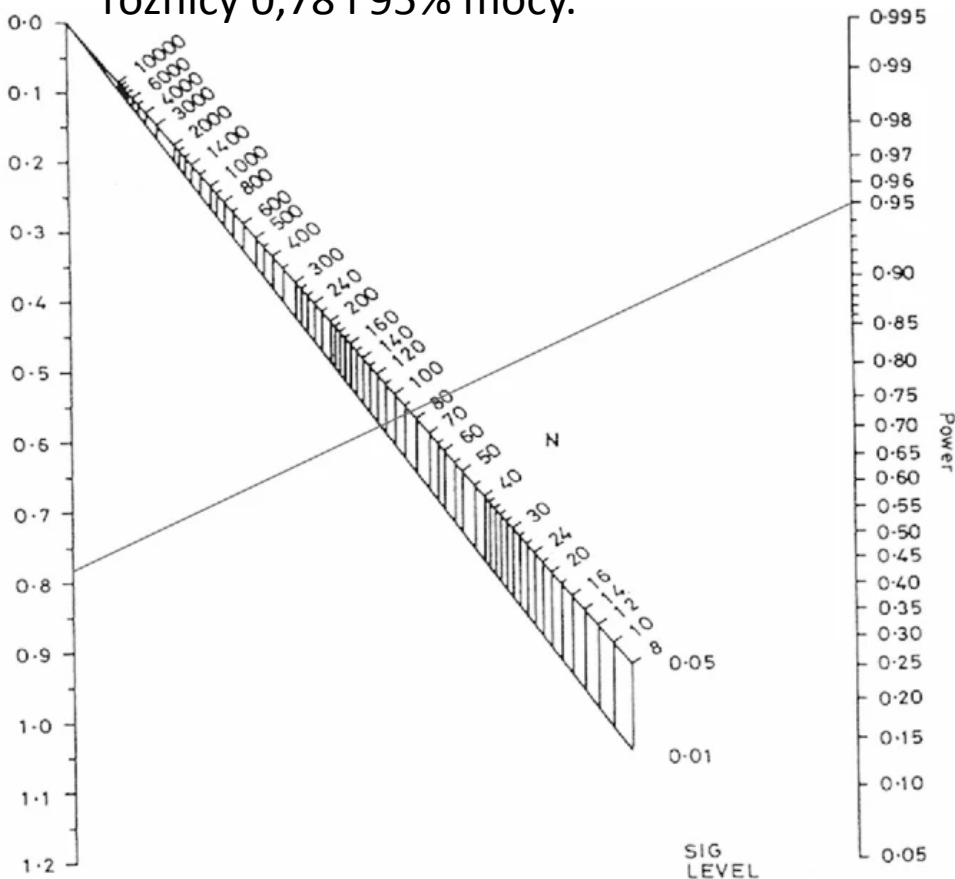


# Nomogram Altmanna

Nomogram przedstawiający obliczenie wielkości próby dla znormalizowanej różnicy 0,78 i 80% mocy.



Nomogram przedstawiający obliczenie wielkości próby dla znormalizowanej różnicy 0,78 i 95% mocy.



Powszechnie używane wartości  $c_p$  , moc

P	Moc (%)			
	50	80	90	95
0,05	3.8	7.9	10,5	13.0
0,01	6.6	11.7	14,9	17.8

# Nomogram Altmana

Przykładowo badacze wzięli średnie ciśnienie tętnicze 6 godzin po rozpoczęciu terapii.

Średnie ciśnienie tętnicze wynosiło 95 i 81 mmHg odpowiednio w grupach leczonych wczesną terapią celową i tradycyjną, co odpowiada różnicy 14 mmHg ( $95-81=14$ ).

$$\text{Standardized difference} = \frac{\text{Target difference}}{\text{Standard deviation}} \quad (1)$$

W bieżącym przykładzie odchylenie standardowe średniego ciśnienia tętniczego wynosiło około 18 mmHg, więc standaryzowana różnica do wykrycia, obliczona przy użyciu równania 1, wynosiła  $14/18 = 0,78$ .



# Nomogram Altmana

Odpowiedni wzór na porównanie średnich w dwóch równych grupach jest następujący:

$$n = \frac{2}{d^2} \times c_{p, power} \quad (2)$$

gdzie n jest wymaganą liczbą badanych w każdej grupie, d jest standaryzowaną różnicą, a  $c_p$ , moc jest stałą określoną przez wartości wybrane dla wartości P i mocy.

$$\begin{aligned} n &= \frac{2}{0.78^2} \times c_{0.05, 80\%} \\ &= \frac{2}{0.6084} \times 7.9 \\ &= 2.39 \times 7.9 \\ &= 26.0 \end{aligned}$$

	Moc (%)			
P	50	80	90	95
0,05	3.8	7.9	10,5	13.0
0,01	6.6	11.7	14,9	17.8




W związku z tym wymaganych jest 26 uczestników w każdym ramieniu badania, co zgadza się z szacunkami dostarczonymi przez nomogram


# https://kiero.net edu@kiero.net


Przejdź do treści ZALOGUJ

**Kiero.net** | *lokujemy wyłącznie na Twoją obsługę 4 pliki cookie* | *stosujemy dyrektywę RODO*

[START](#) [CZAT](#)

AMiSNS 2022/2023		Tematów	Postów	Ostatni post
	Informatyka i Biostatystyka 2-semestr Lekarski <b>Moderowane przez:</b> Administrator	3	3	01-03-2023 11:30 przez kiero
	Statystyka medyczna Pielęgniarstwo	4	4	02-03-2023 20:58 przez kiero
	Informatyka i Biostatystyka Ratownictwo	3	3	02-03-2023 20:58 przez kiero

 - Nowe treści od Twojej ostatniej wizyty.

 - Brak nowych treści od Twojej ostatniej wizyty.

*Dziękuję za uwagę*