

Test t-Studenta

Trochę formalnych wzorów 😞

Ale nie za wiele 😊

Test t dla dwóch niezależnych prób

$H_0: \mu_1 = \mu_2$ (nie ma różnicy w średnich wartościach cechy pomiędzy dwoma grupami (populacjami))

$H_1: \mu_1 \neq \mu_2$ (są różnice w średnich wartościach cechy pomiędzy dwoma grupami (populacjami))

Test t Studenta

Statystyka testowa, w zależności od wielkości próby

W obu grupach co najmniej 30 ($n_1 \geq 30$ i $n_2 \geq 30$) statystyka Z

$$Z = \frac{|x_1 - x_2|}{\sqrt{\frac{sd_1^2}{n_1} + \frac{sd_2^2}{n_2}}}$$

gdzie x_1, x_2 średnie w 1 oraz 2 grupie,

sd_1 oraz sd_2 odpowiednio odchylenia standardowe

Test t Studenta

Statystyka testowa, w zależności od wielkości próby

W przynajmniej jednej z grup mniej niż 30 ($n_1 < 30$ lub $n_2 < 30$) statystyka t (rozkład t Studenta, z odpowiednią liczbą stopni swobody)

$$t = \frac{|x_1 - x_2|}{\sqrt{\frac{(n_1 - 1)sd_1^2 + (n_2 - 1)sd_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Stopnie swobody

gdzie x_1, x_2 średnie w 1 oraz 2 grupie,

sd_1 oraz sd_2 odpowiednio odchylenia standardowe

Test t dla prób zależnych

$H_0: \check{d} = 0$ (średnia wartość różnicy pomiędzy parami obserwacji nie różni się od zera)

$H_1: \check{d} \neq 0$ (średnia wartość różnicy pomiędzy parami obserwacji różni się od zera)

Statystyka testowa, w zależności od wielkości próby

$n \geq 30$ – statystyka Z (rozkład normalny)

$$Z = \frac{d}{\sqrt{sd}} \sqrt{n}$$

$n < 30$ – statystyka t (rozkład t Studenta)

$$t = \frac{d}{\sqrt{sd}} \sqrt{n - 1}, \text{ liczbą stopni swobody } df = n - 1$$

gdzie d – średnia wartość różnicy pomiędzy parami obserwacji dla danej cechy

Test t jednej próby

Czy średnia wartość cechy w naszej grupie (populacji) różni się od „ogólnej” średniej populacyjnej.

$H_0: \mu_p = \mu$ (średnia wartość w grupie nie różni się od średniej w populacji)

$H_1: \mu_p \neq \mu$ (średnia wartość w grupie różni się od średniej w populacji)

Statystyka testowa, w zależności od wielkości próby

$n \geq 30$ – statystyka Z (rozkład normalny)

$$Z = \frac{x - \mu}{\sqrt{sd}} \sqrt{n}$$

$n < 30$ – statystyka t (rozkład t Studenta)

$$t = \frac{x - \mu}{\sqrt{sd}} \sqrt{n - 1}, \text{ liczbą stopni swobody } df = n - 1$$

gdzie x – średnia wartość cechy w naszej próbie

Testy nieparametryczne:

Porównujemy rozkłady, nie parametry (średnia, sd)

Zastosowanie:

Rozkład cechy odbiega od rozkładu normalnego

Mamy cechy zmierzone co najmniej na skali porządkowego (skala Likerta)

Test oparty na rangach

Test U Manna-Whitneya

Statystyka testowa:

$$U_1 = (n_1)(n_2) + \frac{n_1(n_1 + 1)}{2} - \sum R_1$$

$$U_2 = (n_1)(n_2) + \frac{n_2(n_2 + 1)}{2} - \sum R_2$$

$$T = \min\{U_1, U_2\}$$

Gdy mała liczebność, rozkład dokładny, tablice
N>20, aproksymacja rozkładem normalnym

$$T = \frac{U - m}{\sigma}, \quad m = \frac{n_1 n_2}{2}, \quad \sigma = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}, \quad T \sim N(0, 1)$$

$$\sum R_1$$

Suma rang w pierwszej grupie

$$\sum R_2$$

Suma rang w drugiej grupie

n_1, n_2 - liczebności w grupach

Rangi - przykład

Poniżej podano wartości pewnej skali w dwóch grupach wartości pewnej skali

grupa 1

grupa 2

24

28

18

42

45

63

57

57

12

90

30

68

Rangujmy obserwacje (w całej próbie)

group 1

group 2

24

28

18 2

42

45

63

57

57

12 1

90

30

68

grupa 1

24 3

18 2

45

57

12 1

30

grupa 2

28 4

42

63

57

90

68

grupa 1

grupa 2

24 3

28 4

18 2

42 6

45 7

63

57

57

12 1

90

30 5

68

Rangi wiązane, średnia
wartość rang:

$$\frac{8 + 9}{2} = 8,5$$

group 1

24 3

18 2

45 7

57 8.5

12 1

30 5

group 2

28 4

42 6

63

57 8.5

90

68

grupa 1

24 3

18 2

45 7

57 8.5

12 1

30 5

grupa 2

28 4

42 6

63 10

57 8.5

90 12

68 11

Liczmy sumę rang dla dwóch grup

grupa 1

grupa 2

24 3

28 4

18 2

42 6

45 7

63 10

57 8.5

57 8.5

12 1

90 12

30 5

68 11

26.5

51.5

grupa 1

24 3

18 2

45 7

57 8.5

12 1

30 5

26.5

grupa 2

28 4

42 6

63 10

57 8.5

90 12

68 11

51.5

Sprawdzamy sumę rang

$$\sum R = \frac{n(n+1)}{2} = \frac{12 \cdot 13}{2} = 78$$

$$51,5 + 26,5 = 78$$

OK 😊

Wartość statystyki testowej U_1

$$U_1 = (n_1)(n_2) + \frac{n_1(n_1 + 1)}{2} - \sum R_1$$

$$U_1 = (6)(6) + \frac{6(7)}{2} - 26.5$$

$$U_1 = 36 + 21 - 26.5$$

$$U_1 = 30.5$$

Wartość statystyki U_2

$$U_2 = (n_1)(n_2) + \frac{n_2(n_2 + 1)}{2} - \sum R_2$$

$$U_2 = (6)(6) + \frac{6(7)}{2} - 51.5$$

$$U_2 = 36 + 21 - 51.5$$

$$U_2 = 5.5$$

$$T = \min\{30,5; 5,5\} = 5,5$$

Wyniki z programu Statistica

Test U Manna-Whitneya (z poprawką na ciągłość) (Arkusz6)										
Względem zmiennej: gr										
Zaznaczone wyniki są istotne z $p < ,05000$										
Zmienna	Sum.rar g Grupa 1	Sum.rar g Grupa 2	U	Z	p	Z popraw.	p	N ważn. Grupa 1	N ważn. Grupa 2	2*1str. dokł. p
Wynik	26,5000	51,5000	5,50000	-1,9215	0,05466	-1,9249	0,05424	6	6	0,04112

$T = \min\{30,5, 5,5\} = 5,5$

Test sprawdzające założenia

Zgodność z rozkładem normalnym

Test Shapiro-Wilka

H0: Rozkład danej cechy jest zgodny z rozkładem normalnym (nie odbiega od rozkładu normalnego)

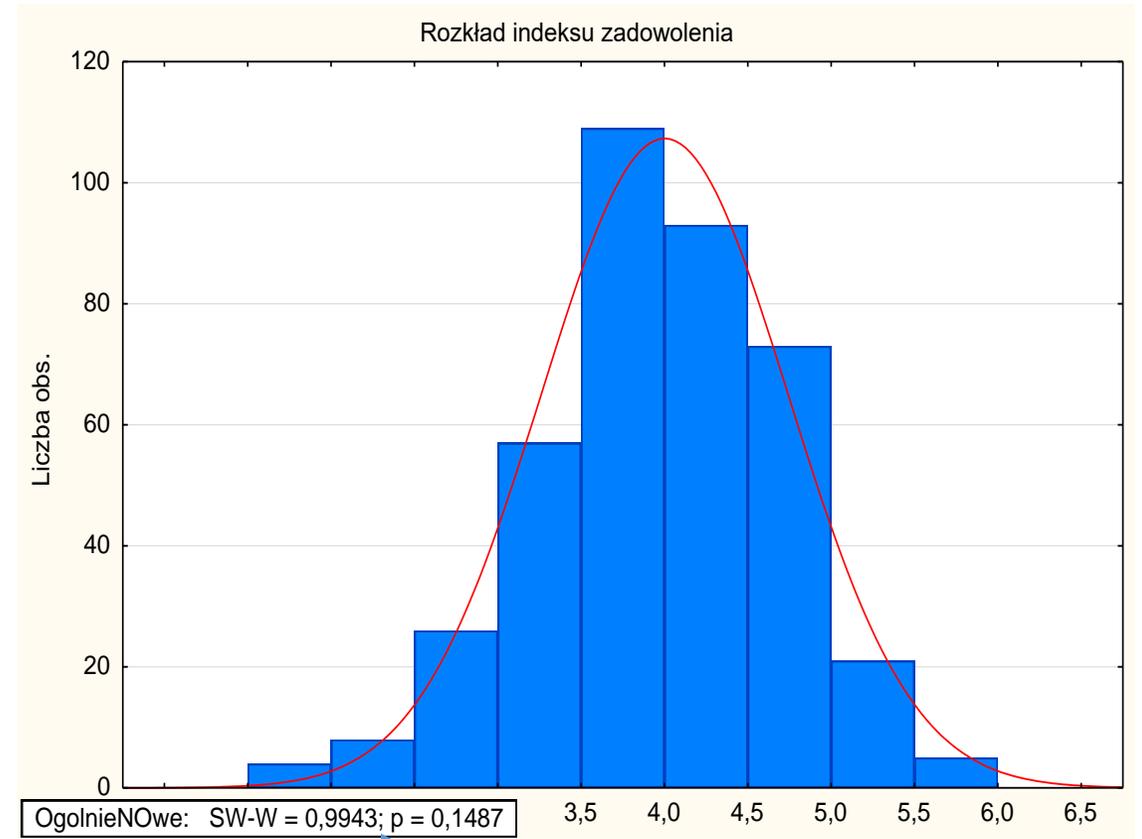
H1: Rozkład danej cechy odbiega (różni się od rozkładu normalnego) od rozkładu normalnego.

Zgodność z rozkładem normalnym – przykład

W pewnym badaniu przeprowadzonym w Polsce na temat satysfakcji personelu medycznego, przeprowadzono ankietę i uzyskano indeks opisujący zadowolenie z pracy (1- zła, 6 – bardzo dobra).

H0: Rozkład indeksu zadowolenia **jest zgodny** z rozkładem normalnym

H1: Rozkład indeksu zadowolenia **odbiega od rozkładu normalnego**.



$p > \alpha$ ($0,15 > 0,05$), brak podstaw do odrzucenia hipotezy zerowej
Rozkład indeksu zadowolenia jest zgodny z rozkładem normalnym

Test Shapiro-Wilka - Statistica

The screenshot shows the Statistica software interface with the 'Histogramy 2W' dialog box open. The dialog box is titled 'Histogramy 2W' and has tabs for 'Podstawowe', 'Więcej', 'Wygląd', 'Kategoryzowane', 'Opcje 1', and 'Opcje 2'. The 'Podstawowe' tab is active, showing the following settings:

- Zmienna: 61
- Rodzaj wykresu: Zwykły
- Dopasuj: Wylaczone
- Sposob wyświetlania: Standardowe
- Wykres Pareto:
- Odstępy między kolumnami:
- Pokaż procenty:
- Pokaż liczby:
- Oś Y: N
- Przedziały: Zmienna: OgolnieNOWe
- Tryb całkowity: Automatyczny
- Unikalne wartości: Niesort. Ros. Mal.
- Kategorie: 10
- Granice: brak
- Kody: brak
- Podzbiory wielowarunkowe:
- Zmień zmienną:
- Statystyki: Test Shapiro Wilka, Całkowita liczność, Statystyki opisowe, Test Kolmogorowa-Smimowa

The background shows a histogram of the variable 'OgolnieNOWe' with a y-axis labeled 'Liczba obs.' ranging from 0 to 120. The x-axis is labeled 'OgolnieNOWe'. The data values are listed in a table on the right side of the dialog box:

OgolnieNOWe
4,4
4,0
4
4
4,0
3
.
4,0
4
4
3,8
3
4,5
>

The bottom of the screen shows the taskbar with the text 'Aby uzyskać pomoc naciśnij F1.' and the system tray with the text 'ZwykEUniver CAP NUM REC' and the time '14:23'.

Test jednorodności wariancji

$H_0: \sigma_1^2 = \sigma_2^2$ (wariancje badanej cechy nie różnią się pomiędzy populacjami)

$H_1: \sigma_1^2 \neq \sigma_2^2$ (wariancje badanej cechy różnią się pomiędzy populacjami)

Test F Fishera

$$F = \frac{\max\{sd_1^2, sd_2^2\}}{\min\{sd_1^2, sd_2^2\}}$$

Zastosowanie

Testy parametryczne

Testy t – Studenta (wnioskowanie o różnicy w średnich wartościach)

1. **Testy t dla jednej próby**
2. Test t dla prób niezależnych
3. Test t dla prób zależnych (pary powiązane)

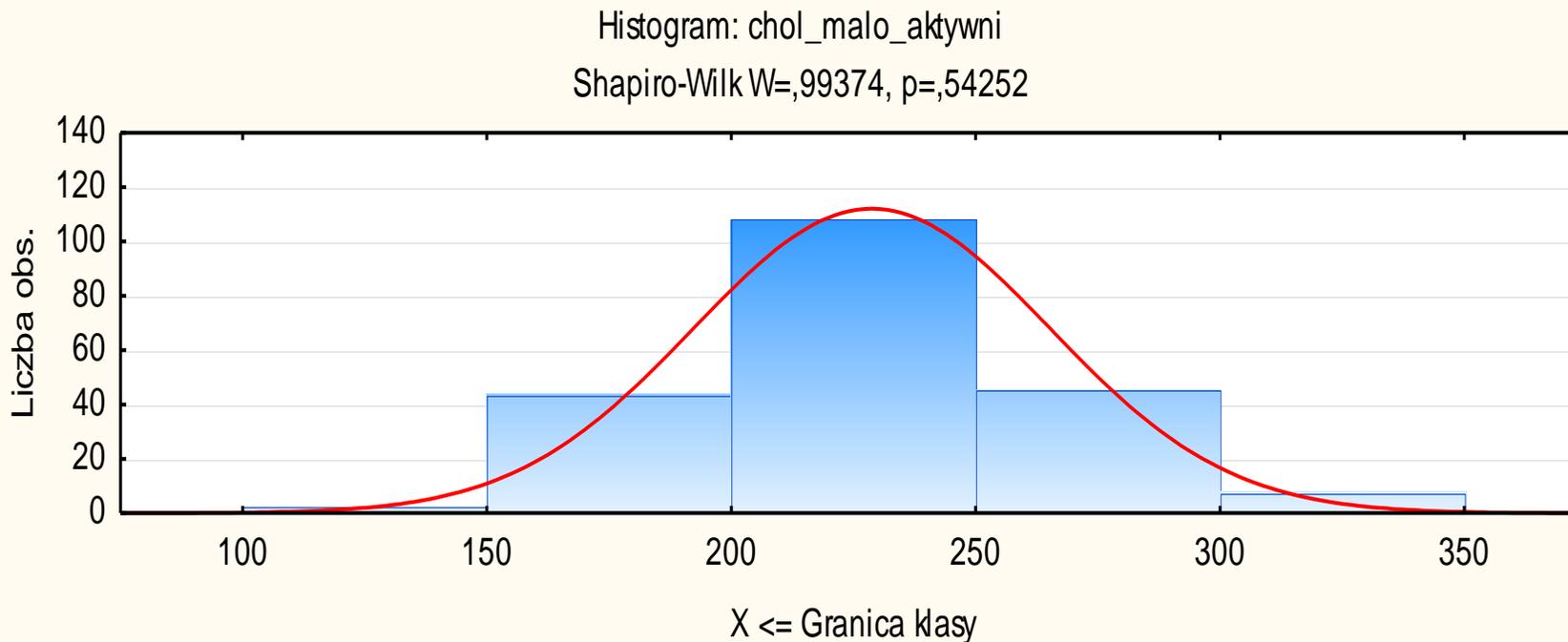
Główne założenie:

Rozkład badanej cechy jest zgodny z **rozkładem normalnym**

Test t dla jednej próby

Z populacji mężczyzn w wieku 25-44 lat zamieszkujących pewne województwo wylosowano próbę 205 mężczyzn wykazujących się niską aktywnością fizyczną (AF). Sprawdźmy, czy średnia wartość poziomu cholesterolu w badanej grupie różni się od średniej wartości poziomu cholesterolu w populacji mężczyzn w tym wieku - 220 mg/dl

Statystyki opisowe									
Zmienna	N	Średnia	Mediana	Min	Maks	Q ₁	Q ₃	sd	Skośność
chol_malo_aktywni	205	228,74	228,43	140,31	336,79	202,76	251,41	36,50	0,23



Sprawdzanie założenia
Rozkład analizowanej cechy jest zgodny z rozkładem normalnym?
Test Shapiro-Wilka

Test t dla jednej próby

H0: **Średnia** wartość stężenia cholesterolu w populacji mężczyzn o niskiej AF wynosi 220

H1: **Średnia** wartość stężenia cholesterolu w populacji mężczyzn o niskiej AF różni się od 220

Test średnich względem stałej wartości odniesienia										
Zmienna	Średnia	sd	N	Bł. std.	Ufność -95%	Ufność +95%	Odniesienie - Stała	t	df	p
chol	228,74	36,5	205	2,549	223,72	233,77	220	3,4291	204	0,00073

Decyzja: Odrzucamy H₀, przyjmujemy H₁

Średnia wartość stężenia cholesterolu w populacji mężczyzn o niskiej AF różni się od 220 mg/dl

Statistica

The screenshot displays the Statistica software interface. The main window title is "Statistica - [SKOROSZYT] - TEST U Manna-Whitneya (z poprawką na ciągłość) (BazaDOKTORATPRZESZCICENIA I Statystyka 1)". The "Statystyka" menu is highlighted in the top ribbon. The "Statystyki podstawowe i tabele: BazaDoktorat..." dialog box is open, showing a list of statistical tests. The "Test t dla prób zależnych" option is selected and circled in red. The background shows a data table with columns for "Zmienna", "Sum.rang Grupa 1", and "N ważn. Grupa 1".

Statystyki podstawowe i tabele: BazaDoktorat...

Podstawowe

- Statystyki opisowe
- Macierze korelacji
- Test t dla prób niezależnych (wzgl. grup)
- Test t dla prób niezależnych (wzgl. zm.)
- Test t dla prób zależnych**
- Test t dla pojedynczej próby
- Przekroje, prosta ANOVA
- Przekroje uproszczone
- Tabele licznosci
- Tabele wielodzzielcze
- Tabele wielokrotnych odpowiedzi
- Inne testy istotności
- Kalkulator prawdopodobieństwa

OK
Anuluj
Opcje

Otwórz dane

SELECT CASES

Zmienna	Sum.rang Grupa 1	N ważn. Grupa 1	N ważn. Grupa 2
% desat. <90%	21970.50	137	155

Testy parametryczne

Testy t – Studenta (wnioskowanie o różnicy w średnich wartościach)

1. Testy t dla jednej próby
2. **Test t dla prób niezależnych**
3. Test t dla prób zależnych (pary powiązane)

Główne założenie:

Rozkład badanej cechy jest zgodny z **rozkładem normalnym**

Test t dla prób niezależnych

H0: **Średnie** wartości badanej cechy **nie różnią** się pomiędzy dwoma populacjami

H1: **Średnie** wartości badanej cechy **różnią się** pomiędzy dwoma populacjami

Test t dla prób niezależnych

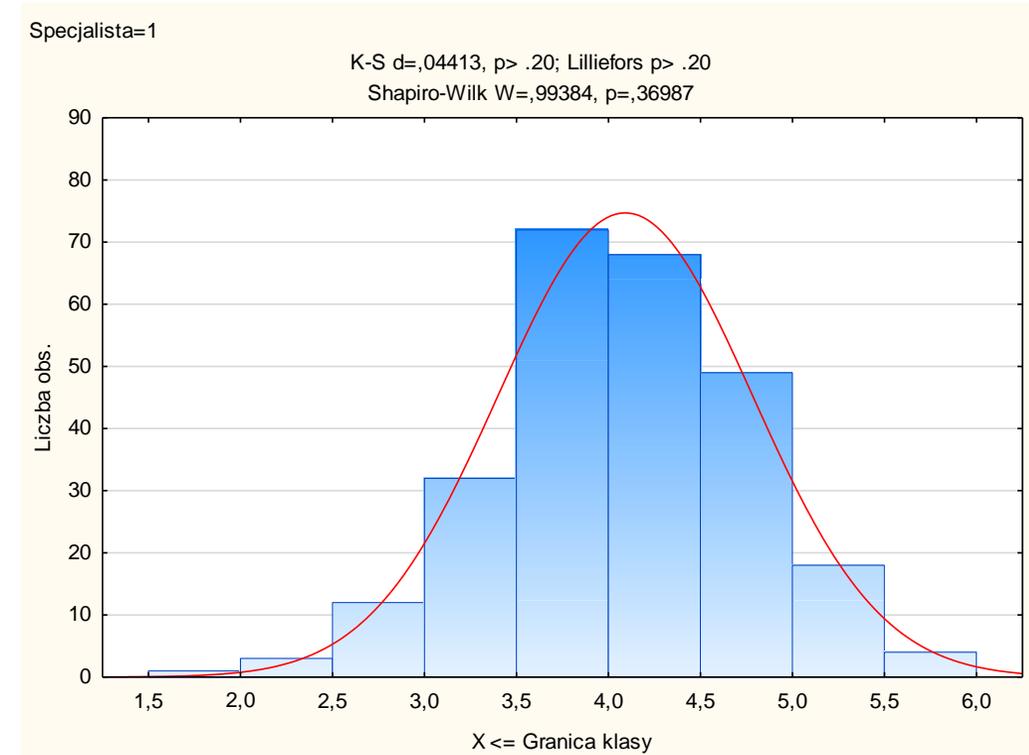
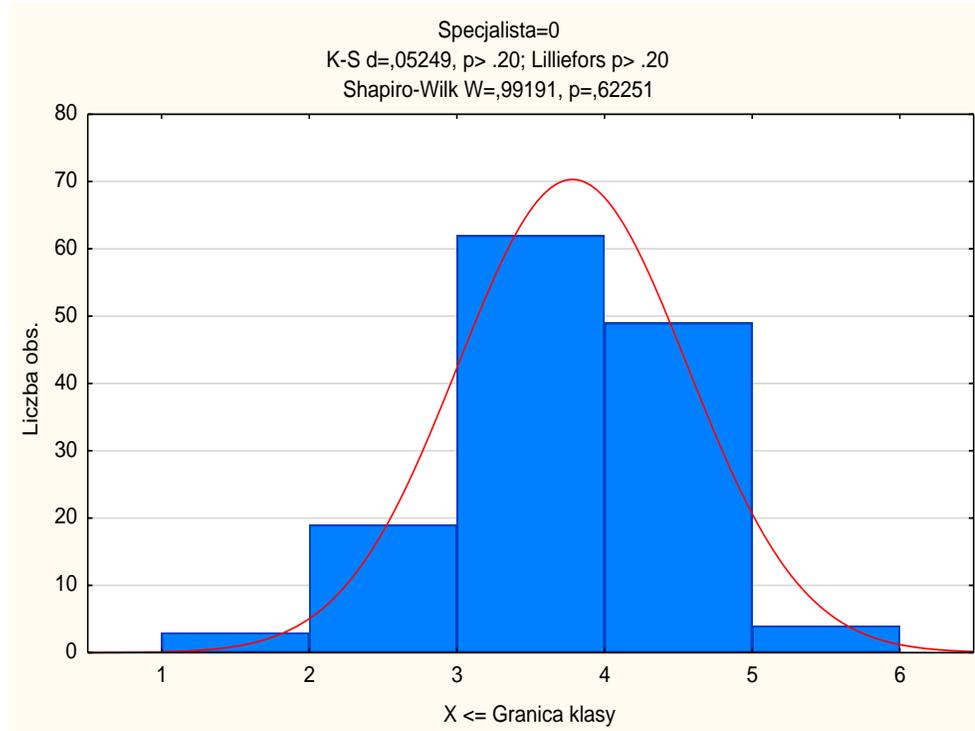
Przykład

W pewnym badaniu postanowiono ocenić, czy wartość satysfakcji wśród lekarzy różni się pomiędzy badanymi posiadającymi specjalizację oraz bez specjalizacji. Poniżej znajdują się statystyki opisowe.

Specjalizacja	n	Średnia	SD
nie	137	3,8	0,78
tak	259	4,1	0,69

Założenia

Rozkład badanej cechy zgodny z rozkładem normalnym w obu grupach



Czy mogę wykonać test t ?

Założenia!

1. Rozkład indeksu satysfakcji w dwóch grupach zgodny z rozkładem normalnym (test Shapiro-Wilka, w obu grupach rozkład cechy musi być zgodny z rozkładem normalnym)

2. Jednorodność wariancji – test F Fishera

(H0: Wariancje indeksu satysfakcji są jednorodne (równe))

(H1: Wariancje indeksu satysfakcji nie są jednorodne)

Brak podstaw do odrzucenia H0,
wariancje indeksu satysfakcji są równe (jednorodne)
– założenie jednorodności
wariancji spełnione

Zmienna	Średnia		t	df	p	Nważnych		Odch.std		iloraz F Wariancje	p Wariancje
	0	1				0	1	0	1		
OgólnieNOwe	3,78359	4,09152	-4,0354	394	0,00006	137	259	0,77727	0,69158	1,26317	0,11189

Założenia spełnione!

H0: **Średnie** wartości indeksu satysfakcji **nie różnią** się pomiędzy lekarzami bez specjalizacji i ze specjalizacją

H1: **Średnie** wartości indeksu satysfakcji **różnią się** pomiędzy lekarzami bez specjalizacji i ze specjalizacją

Decyzja: odrzucamy H0

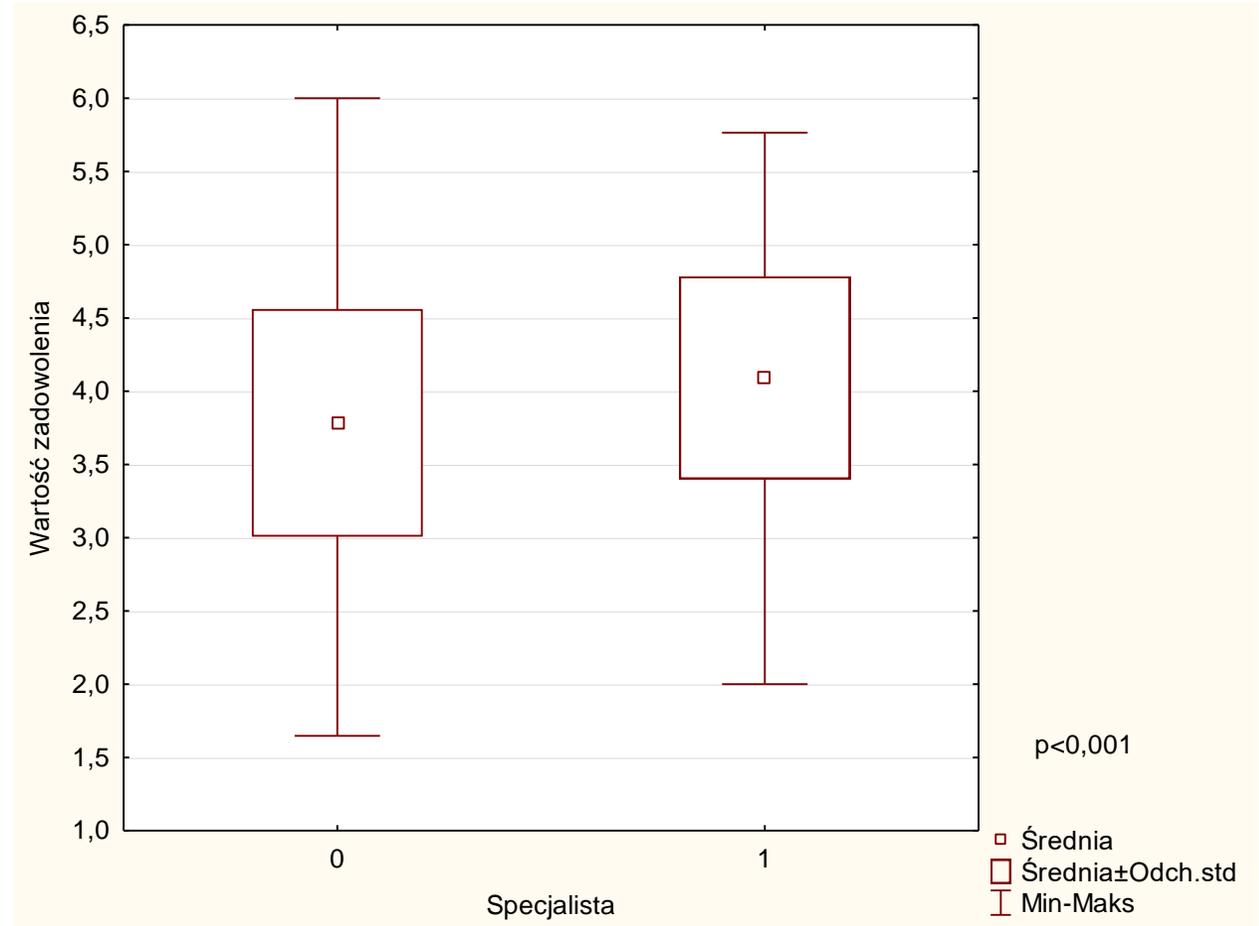
Wniosek: Średnie wartości indeksu satysfakcji różnią się pomiędzy tymi dwoma populacjami.

Zmienna	Grupa 1: 0 Grupa 2: 1 Warunek uwzględniania: Uniwersytecki=1										
	Średnia 0	Średnia 1	t	df	p	Nważnych 0	Nważnych 1	Odch.std 0	Odch.std 1	iloraz F Wariacje	p Wariacje
OgólnieNOwe	3,783598	4,091529	-4,03544	394	0,000065	137	259	0,777271	0,691582	1,263176	0,111897

Test t – prezentacja wyników

Tabela. Porównanie średnich wartości satysfakcji pomiędzy lekarzami ze specjalizacją i bez specjalizacji.

Specjalizacja	n	Średnia	SD	p
nie	137	3,8	0,78	<0,001
tak	259	4,1	0,69	



Rycina. Porównanie średnich wartości satysfakcji pomiędzy lekarzami ze specjalizacją i bez specjalizacji

 Statystyki podstawowe
 Regresja wieloraka
 ANOVA
 Nieparametryczne
 Dopasowanie rozkładu
 Rozkłady i symulacja

Podstawowe

 Modele zaawansowane
 Sieci neuronowe
 Wielowymiarowe
 Analiza mocy testu
 PLS, PCA, ...
 VEPAC

Zaawansowane i wielowymiarowe

 Karty kontrolne
 Wielowymiarowe
 Predykcyjne
 Analiza procesu
 DOE
 Sześć Sigma

Statystyki przemysłowe

 STATISTICA VB
 Grupami
 Kalkulatory
 Statystyki bloku

Narzędzia

Plan: 2**(4-0) (composit.sta)

	1 A	2 B	3 C	4 D
1	-1	-1	-1	-1
2	1	-1	-1	-1
3	-1	1	-1	-1
4	1	1	-1	-1
5	-1	-1	1	-1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	-1
9	-1	-1	-1	1
10	1	-1	-1	1
11	-1	1	-1	1
12	1	1	-1	1
13	-1	-1	1	1
14	1	-1	1	1
15	-1	1	1	1
16	1	1	1	1

Statystyki podstawowe i tabele: 2(4-0) ? x

Podstawowe

- Statystyki opisowe
- Macierze korelacji
- Test t dla prób niezależnych (wzgl. grup)**
- Test t dla prób niezależnych (wzgl. zm.)
- Test t dla prób zależnych
- Test t dla pojedynczej próby

- Przekroje, prosta ANOVA
- Przekroje uproszczone
- Tabele licznosci
- Tabele wielozdzielcze
- Tabele wielokrotnych odpowiedzi

- Inne testy istotności
- Kalkulator prawdopodobieństwa

OK

Anuluj

Opcje

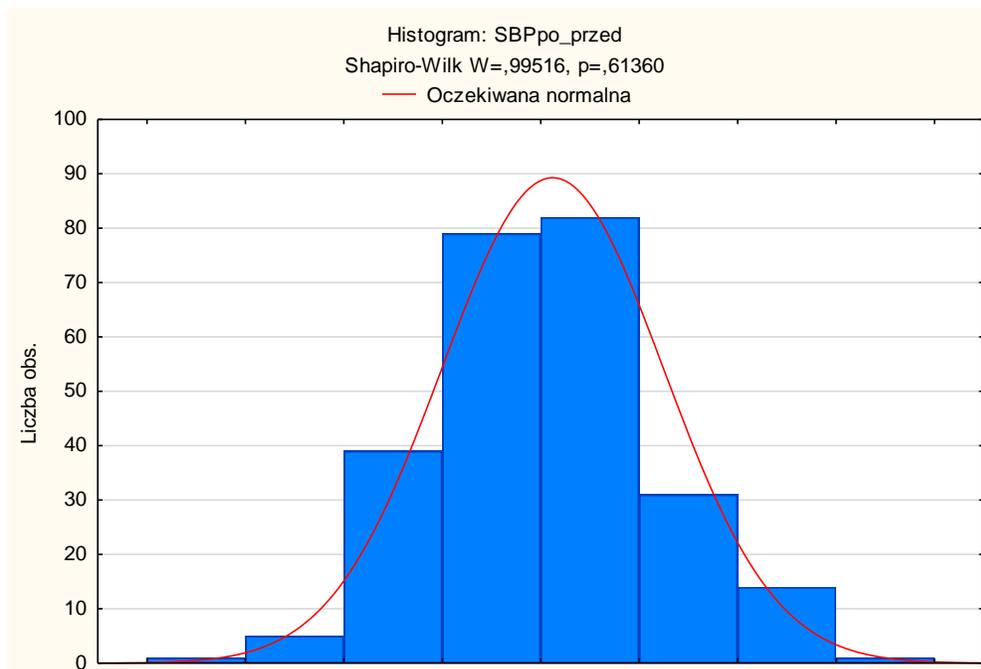
Otwórz dane

SELECT CASES S W

Test t dla prób zależnych

W pewnym projekcie porównano ciśnienie skurczowe krwi (SBP) przed i po terapii. Poniżej znajdują się statystyki opisowe oraz sprawdzanie założeń do testu t- studenta dla par zależnych.

Zmienna	Statystyki opisowe (Arkusz1 w Dane do Testow t Studenta)				
	Nważnych	Średnia	Minimum	Maksimum	Odch.std
sbp_przed	305	138,586	91,0000	236,5000	21,9247
sbp_po	305	134,013	86,0000	230,0000	22,8789



Założenia

Rozkład różnicy dla poszczególnych par jest zgodny z rozkładem normalnym?

Spełnione!

Hipotezy

H_0 : średnia wartość różnicy SBP przed i po terapii dla poszczególnych par obserwacji równa jest zero.

H_1 : średnia wartość różnicy SBP przed i po terapii dla poszczególnych par obserwacji różni się od zera.

Test T dla prób zależnych (Arkusz 1 w Dane do Testow t Studenta)										
Zaznaczone różnice są istotne z $p < ,05000$										
Zmienna	Średnia	Odch.s	Ważnyc	Różnic	Odch.s Różnic	t	df	p	Ufność -95,000	Ufność +95,000
sbp_prze	138,58	21,924								
sbp_po	134,01	22,878	30	4,5737	5,6818	14,058	30	0,0	3,9338	5,2138

Decyzja: Odrzucamy H_0

Wniosek:

Zaobserwowaliśmy istotną różnicę w średnich wartościach SBP przed i po terapii.

Statistica - Dane do Testow t Studenta* - [Arkusz1]

Plik Podstawowe Edycja Widok Format Statystyka Data Mining Wykresy Narzędzia Dane Skoroszyt

Statystyki podstawowe Regresja wieloraka ANOVA Nieparametryczne Dopasowanie rozkładu Rozkłady i symulacja

Modele zaawansowane Sieci neuronowe Karty kontrolne Analiza procesu Statistica VB

Wielowymiarowe PLS, PCA, ... Wielowymiarowe DOE Grupami

Analiza mocy testu VEPAC Predykcyjne Sześć Sigma Kalkulatory Statystyki bloku

Podstawowe Zaawansowane i wielowymiarowe Statystyki przemysłowe Narzędzia

Dane do Testow t Studenta* - Arkusz1

Arkusze

Arkusze

Podst. statystyki/Tabele

- Statystyki opisowe
- Histogram: R
- Statystyki opisowe
- Histogram: SI
- Statystyki opisowe

	gender	chol	sbp	chol_malo	aktywni	poj_pluc.sport	poj.pluc_niesport	sbp_przed	sbp_po	SBPpo_pr
1	1	5,80000019	137,5		248,32809	4085,80671	3377,28546	164	160	-4
2	1	5,17000008	151		263,141021	4074,0004	3150,66244	137	132	-5
3	1	5,0999999	179		219,765252	4084,54282	4209,63692	189,5	177	-12,5
4	1	5,23999977	137		209,392272	4077,1468	3539,48995	165	162	-3
5	2	5,40999985	132,5		255,8852	4081,51137	2699,60352	104	98	-6
6	1	6,07999992	126		196,504544	4078,61021	3675,72049	134	128	-6
7	1	5,34000015	176		218,463103	4068,7398				
8	2	6,71999979	122,5		229,604793	4073,24785				
9	2	6,57000017	164,5		254,300351	4079,4707				
10	1				221,497085	4084,77063				
11	2	3,93000007	121		229,895925	4077,36782				
12	1	5,13000011	138,5		186,450795	4080,08197				
13	2	4,21999979	131		224,2247	4083,59238				
14	1	4,09000015	154,5		239,562388	4082,38299				
15	2	8,38000011	144,5		244,831123	4083,06752				
16	1		4	148,5	289,913574	4087,80428				
17	1	4,11000013	169		204,22798	4077,22502				
18	2	6,92000008	145		217,995319	4092,24709				
19	2	5,76999998			255,45337	4077,98449				
20	2	6,15999985	102		242,649294	4082,12851				
21	2	5,30000019	151		178,550187	4082,47935				
22	2	7,73000002	108		247,13324	4076,6494				
23	2	6,9000001	140		235,398193	4079,06424				
24	1	7,57999992	136,5		309,20825	4083,38977				
25	2	5,94999981	132,5		216,482414	4081,94376				
26	2	6,69999981	150,5		189,873024	4085,1134				

Statystyki podstawowe i tabele: Arkusz1 w Dane do Te... ? X

Podstawowe

- Statystyki opisowe
- Macierze korelacji
- Test t dla prób niezależnych (wzgl. grup)
- Test t dla prób niezależnych (wzgl. zmn.)
- Test t dla prób zależnych**
- Test t dla pojedynczej próby
- Przekroje, prosta ANOVA
- Przekroje uproszczone
- Tabele licznosci
- Tabele wielodzzielcze
- Tabele wielokrotnych odpowiedzi
- Inne testy istotnosc
- Kalkulator prawdopodobienstwa

OK Anuluj Opcje

Otwórz dane

SELECT CASES S ID W

chol	sbp	chol_malo_aktywni	poj_pluc.sport	poj_pluc_niesport	sbp_przed	sbp_po	SBPpo_pri
5,80000019	137,5	248,32809	4085,80671	3377,28546	164	160	-4
5,17000008	151	263,141021	4074,0004	3150,66244	137	132	-5
5,09999999	179	219,765252	4084,54282	4209,63692	189,5	177	-12,5
5,23999977	137	209,392272	4077,1468	3539,48995	165	162	-3
5,40999985	132,5	255,8852	4081,51137	2699,60352	104	98	-6
6,07999992	126	196,504544	4078,61021	3675,72049	134	128	-6
5,34000015	176	218,463103	4068,7398	4769,49222	154	144	-10
6,71999979	122,5	229,604793	4073,24785	3400,40000	122,5	122,5	0
6,57000017	164,5	254,300351	4079,4707	3400,40000	164,5	164,5	0
		221,497085	4084,77063	3400,40000			
3,93000007	121	229,895925	4077,36782	3400,40000			
5,13000011	138,5	186,450795	4080,08197	3400,40000			
4,21999979	131	224,2247	4083,59238	3400,40000			
4,09000015	154,5	239,562388	4082,38299	3400,40000			
8,38000011	144,5	244,831123	4083,06752	3400,40000			
4	148,5	289,913574	4087,80428	3400,40000			
4,11000013	169	204,22798	4077,22502	3400,40000			
6,92000008	145	217,995319	4092,24709	3400,40000			
5,76999998		255,45337	4077,98449	3400,40000			
6,15999985	102	242,649294	4082,12851	3400,40000			
5,30000019	151	178,550187	4082,47935	3400,40000			
7,73000002	108	247,13324	4076,6494	3400,40000			
6,9000001	140	235,398193	4079,06424	3400,40000			
7,57999992	136,5	309,20825	4083,38977	3400,40000			
5,94999981	132,5	216,482414	4081,94376	3400,40000			
6,69999981	150,5	189,873024	4085,1134	3400,40000			

Test t dla prób zależnych: Arkusz1 w Dane do T... ? X

Zmienne:

Pierw. lista: sbp_przed
 Druga lista: sbp_po

Podstawowe | Więcej

Podsumowanie: testy t
 Wykresy ramka-wąsy

Podsumowanie

Anuluj

Opcje

Grupami...

SELECT CASES S W

Momenty ważone

DF =
 W-1 N-1

Usuwanie BD
 Przypadkami
 Parami

Testy nieparametryczne

Testy nieparametryczne

Dwie próby niezależne

Przykład

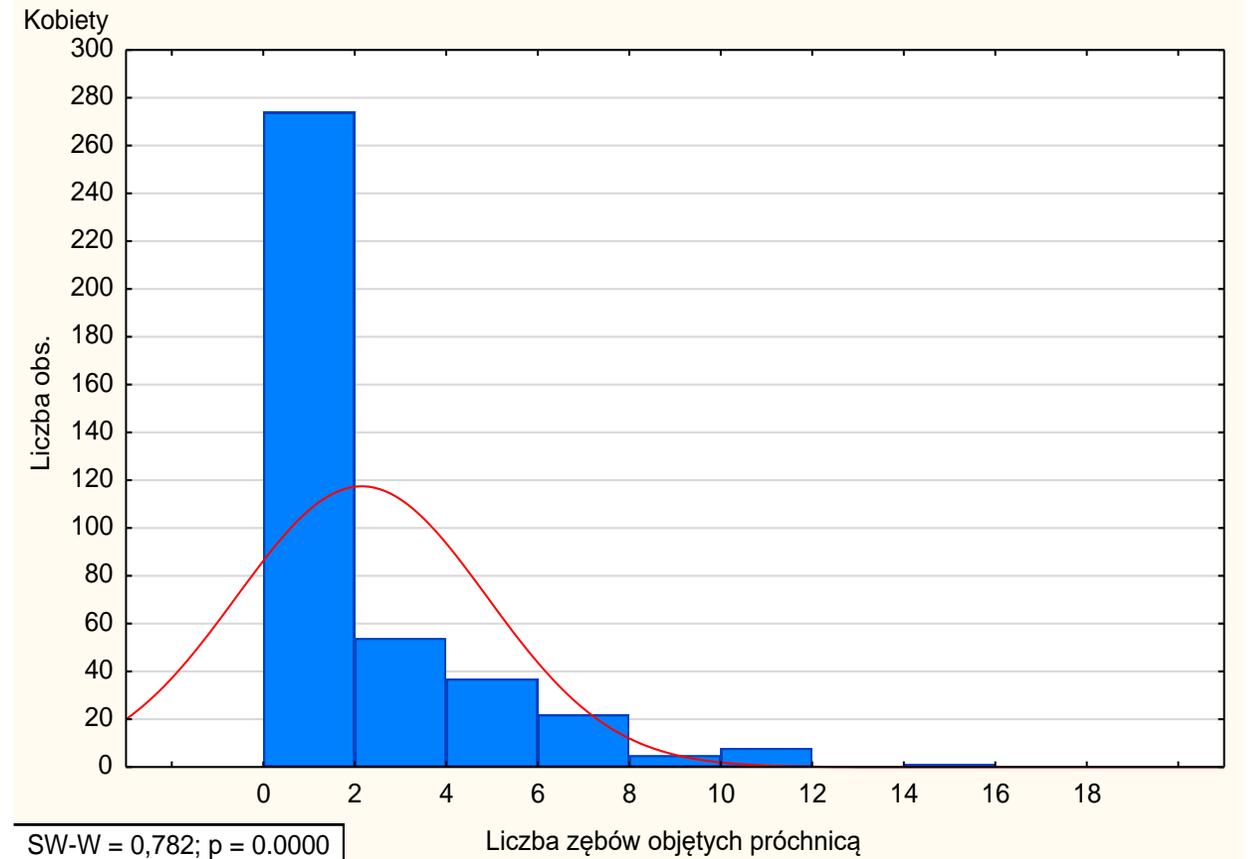
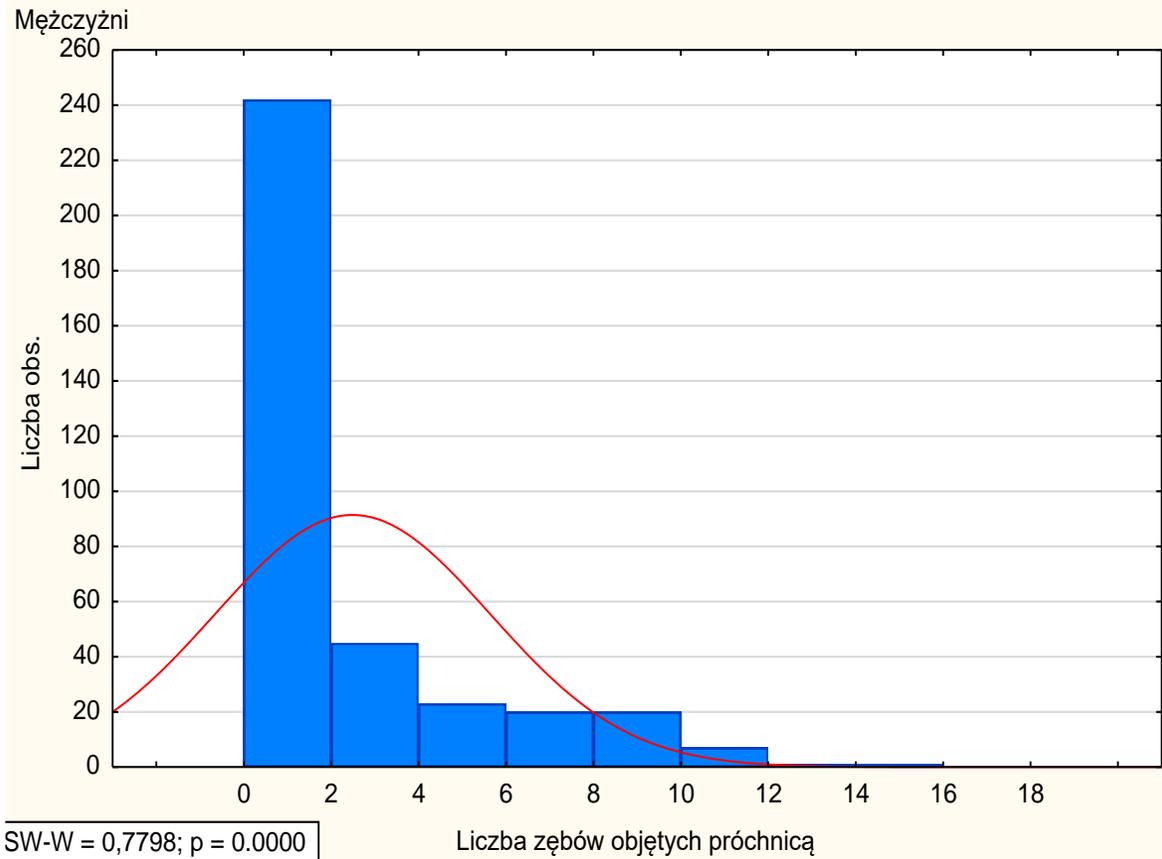
Celem pewnego badania było porównanie liczby zębów objętych próchnicą pomiędzy mężczyznami a kobietami.

Jaki test?

Test t-studenta (czy mogę zastosować?)

Sprawdzamy założenia

Zgodność rozkładu liczby zębów objętych próchnicą z rozkładem normalnym.



Rozkład liczby zębów objętych próchnicą **odbiega od rozkładu normalnego**

Test nieparametryczny, test U Manna-Whitneya.

Hipotezy

H_0 : **Rozkład liczby zębów objętych próchnicą nie różni się** pomiędzy mężczyznami a kobietami

H_1 : **Rozkład liczby zębów objętych próchnicą różni się** pomiędzy mężczyznami a kobietami

Wyniki testu U Manna-Whitney'a

Zmienna	Względem zmiennej: gender Zaznaczone wyniki są istotne z p <,05000								
	Sum.rang Grupa 1	Sum.rang Grupa 2	U	Z	p	Rangi wiązane		N ważn. Grupa 1	N ważn. Grupa 2
C_PROC	149007,0	140173,0	68406,00	-1,18253	0,236997	Z popraw.	p	401	359

Decyzja: Brak podstaw do odrzucenia hipotezy zerowej.

Wniosek:

Nie zaobserwowano istotnych różnic w rozkładzie liczby zębów pomiędzy mężczyznami a kobietami.

Przedstawienie wyników

Porównanie rozkładu liczby zębów objętych próchnicą pomiędzy mężczyznami i kobietami				
Płeć	n	Mediana	Q1-Q3	P
Mężczyzna	359	1	0-4	
Kobieta	401	1	0-3	0.22

STATISTICA - PlusDolaczonyAlkoholOstatteczny

Plik Podstawowe Edycja Widok Format **Statystyka** Data Mining Wykresy Narzędzia Dane Zestaw Medyczny Analizy marketingowe i rynkowe

Statystyki podstawowe Regresja wieloraka ANOVA Nieparametryczne Podstawowe Dopasowanie rozkładu Rozkłady i symulacja

Modele zaawansowane Sieci neuronowe Karty kontrolne Analiza procesu STATISTICA VB

Wielowymiarowe PLS, PCA, ... Wielowymiarowe DOE Grupami

Analiza mocy testu VEPAC Predykcyjne Sześć Sigma Kalkulatory Statystyki bloku

Zaawansowane i wielowymiarowe Statystyki przemysłowe Narzędzia

Feature Finder Opcje

Dane: PlusDolaczonyAlkoholOstatteczny (190 zm * 909 prz)

	56 zeb20<	57 zeb4Nowe	58 bezzeb	59 brakZeb	67 NowaZm	68 id	69 h21	70 p55	71 p55a	72 spotkania_wsk	73 p57	74 p57
1	1	0	1	32								
2	0	3	0	12								
3	0	3	0	10								
4	1	2	0	23								
5	1	2	0	24								
6	0	3	0	4								
7	1	2	0	13								
8	1	0	1	32								
9	1	2	0	21								
10	1	1	0	28								
11	1	2	0	17								
12	1	2	0	13								
13	0	3	0	12								
14	1	0	1	32								
15	1	2	0	22								
					3	3	1	1	0	0	0	0

Statystyki nieparametryczne: PlusDolaczonyAlk...

Podstawowe

- Tabele 2 x 2 (chi²/V2/Fi2, testy McNemary i Fishera)
- Chi² dla licznosci obserwowanych wz. oczekiwanych
- Korelacje (Spearmana, tau Kendalla, gamma)
- Porównanie dwóch prób niezależnych (grup)**
- Porównanie wielu prób niezależnych (grup)
- Porównanie dwóch prób zależnych (zmiennych)
- Porównanie wielu prób zależnych (zmiennych)
- Test Q Cochran
- Statystyki opisowe (mediana, moda, ...)

OK Anuluj Opcje Otwórz dane