

Testy nieparametryczne



Spis treści

Laboratorium VI: Testy nieparametryczne.....	1
Testy nieparametryczne.....	2
1. Tablica wielorozdzielcza	3
2. Test χ^2 niezależności zmiennych	3
2.1 Test χ^2 niezależności w STATISTICE	4
3. Test U Manna-Whitneya	7
3.1 Test U Manna-Whitneya w STATISTICE	8
4. Testy dla próbek zależnych	9
4.1 Test znaków.....	9
4.2 Test kolejności par Wilcoxa	9
4.3 Testy dla próbek zależnych w STATISTICE	10
5. Test Kruskala-Wallisa i test mediany.....	11
5.1 Test Kruskala-Wallisa i test mediany w STATISTICE	11
Ćwiczenia	12

Testy nieparametryczne

Testy istotności dzielimy na dwa rodzaje: parametryczne i nieparametryczne. Wykorzystanie testów parametrycznych do analizy wyników badań wiąże się z pewnymi ograniczeniami: aby móc przystąpić do testu należy sprawdzić, czy spełnione są założenia, zmienne muszą być mierzalne, mieć rozkład normalny, równość wariancji itd. Przykładami testów parametrycznych, jakie do tej pory poznaliśmy są:

- Testy t porównywania średnich (dla zmiennych zależnych, niezależnych, względem stałej wartości);
- Analiza wariancji;
- Test jednorodności wariancji;

W przypadku, gdy założenia dla testów parametrycznych nie są spełnione, wnioskowanie na ich podstawie jest niepoprawne, traci swoją wiarygodność. Testy te są również całkiem bezużyteczne do analizy zmiennych jakościowych i porządkowych. W przypadku, gdy nie możemy przeprowadzić testu parametrycznego, pomocne stają się testy nieparametryczne. Testy nieparametryczne nie zależą od kształtu rozkładu, mogą być stosowane tam, gdzie niespełnione są założenia dotyczące stosowalności testów parametrycznych. Stosujemy je także w tych sytuacjach, gdy mamy do czynienia ze zmiennymi jakościowymi lub porządkowymi, albo dla grup o bardzo małej liczebności. Moc testów nieparametrycznych (wyliczana jako 1 minus prawdopodobieństwo popełnienia błędu II stopnia) jest jednak niższa niż dla parametrycznych, więc stosujemy je tylko wówczas, gdy nie możemy zastosować parametrycznych.

	Testy parametryczne	Testy nieparametryczne
Testy dla dwóch niezależnych próbek	Test t-Studenta dla zmiennych niepowiązanych	Test serii Walda-Wolfowitza Test U Manna-Whitneya Test Kołmogorowa-Smirnowa
Testy dla dwóch próbek zależnych	Test t-Studenta dla zmiennych powiązanych	Test znaków Test kolejności par Wilcoxon Test McNemara
Testy dla n próbek	Analiza wariancji	Test Kruskala-Wallis Test Friedmana Test Q Cochran

Innymi testami nieparametrycznymi są także:

- Testy korelacji nieparametrycznych (R Spearmana, Tau Kendalla, Test chi-kwadrat);
- Testy zgodności (Test chi-kwadrat, Test Kołmogorowa Smirnowa);
- Testy dopasowywania rozkładów (Test Shapiro-Wilka, Kołmogorowa-Smirnowa, chi-kwadrat);

Dla przypomnienia: aby podjąć decyzję na temat odrzucenia hipotezy zerowej, wyznaczoną w STATISTICE na podstawie statystyki testowej wartość p porównujemy z ustalonym poziomem istotności α :

- jeżeli $p \leq \alpha \Rightarrow$ odrzucamy H_0 przyjmując H_1 ,
- jeżeli $p > \alpha \Rightarrow$ nie ma podstaw odrzucić H_0 .

1. Tablica wielodzielcza

Tablica wielodzielcza przedstawia rozkład obserwacji ze względu na kilka cech jednocześnie:

X \ Y	Y_1	Y_2	...	Y_p	
X_1	n_{11}	n_{12}	...	n_{1p}	$\sum_{j=1}^p n_{1j}$
X_2	n_{21}	n_{22}	...	n_{2p}	$\sum_{j=1}^p n_{2j}$
...
X_k	n_{k1}	n_{k2}	...	n_{kp}	$\sum_{j=1}^p n_{kj}$
	$\sum_{i=1}^k n_{i1}$	$\sum_{i=1}^k n_{i2}$		$\sum_{i=1}^k n_{ip}$	$\sum_{i=1}^k \sum_{j=1}^p n_{ij} = N$

Liczebność n_{ij} oznacza liczbę elementów próby, w której zmienna X przyjęła wartość X_i , a zmienna Y – Y_j . Poniżej przedstawiono przykładowa tablicę wielodzielczą dla danych zestawiających dane na temat palenia papierosów względem płci:

Palenie	Nigdy	Rzadko	Często	Ciągle	
Kobieta	11	8	6	5	30
Mężczyzna	4	4	28	24	60
	15	12	34	29	90

2. Test χ^2 niezależności zmiennych

Jednym z kroków analizy statystycznej danych jakościowych jest sprawdzenie, czy dwie cechy jakościowe populacji są niezależne. Najczęściej stosowanym testem do sprawdzenia zależności pomiędzy cechami jakościowymi jest test chi-kwadrat.

Hipoteza zerowa:

H_0 : cechy X i Y są niezależne;

Hipoteza alternatywna:

H_1 : cechy X i Y są zależne;

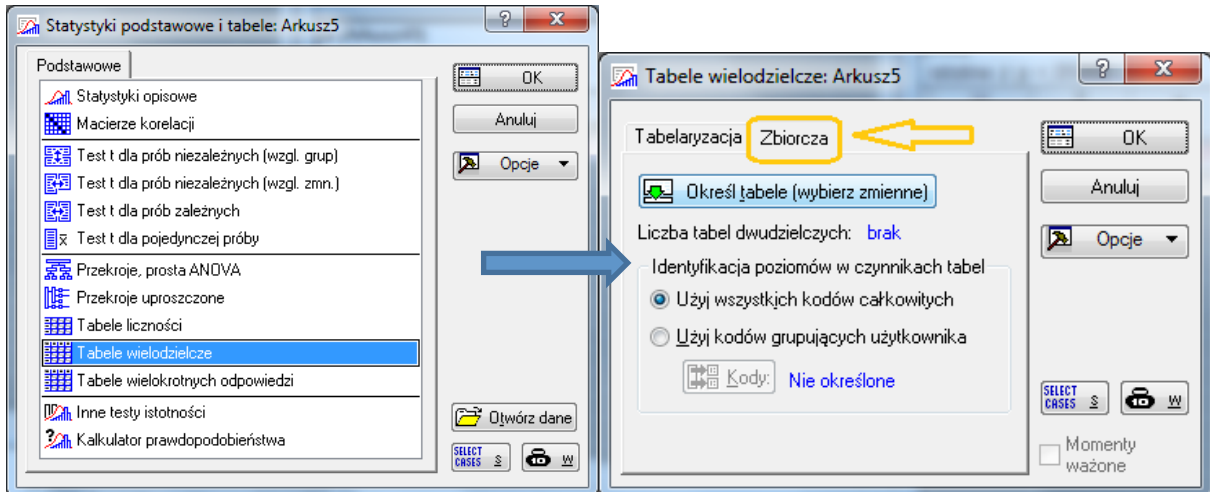
Statystyka testowa:

$$\chi^2 = \sum \frac{(O - E)^2}{E} = \sum_{i=1}^k \sum_{j=1}^p \frac{(n_{ij} - E_{ij})^2}{E_{ij}}$$

Gdzie: E – wartość oczekiwana komórki w tablicy wielodzzielczej, O – obserwowana wartość komórki w tablicy wielodzzielczej. Statystyka ma rozkład χ^2 o $s = (k - 1)(p - 1)$ stopniach swobody.

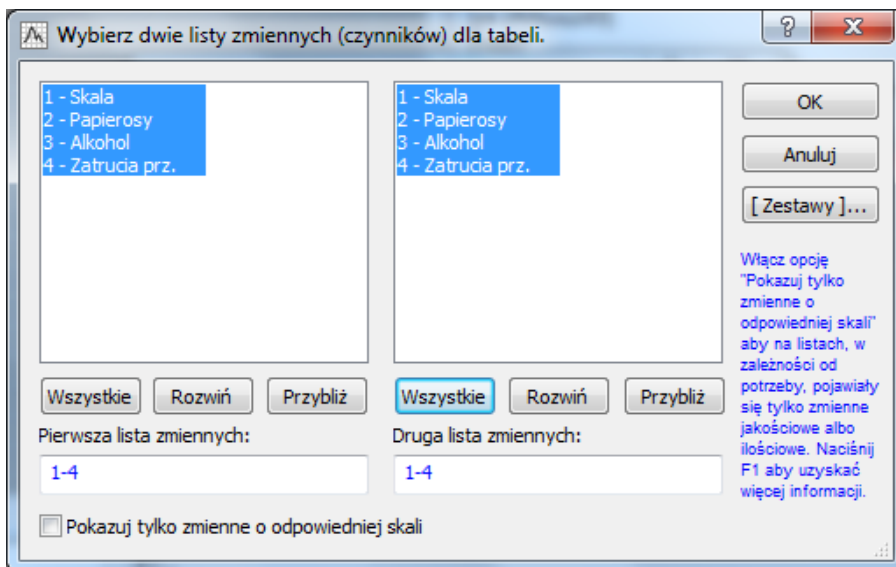
2.1 Test χ^2 niezależności w STATISTICE

Test ten znajdziemy wybierając z menu **głównego Statystyka / Statystyki podstawowe i tabele / Tabele wielodzzielcze**.



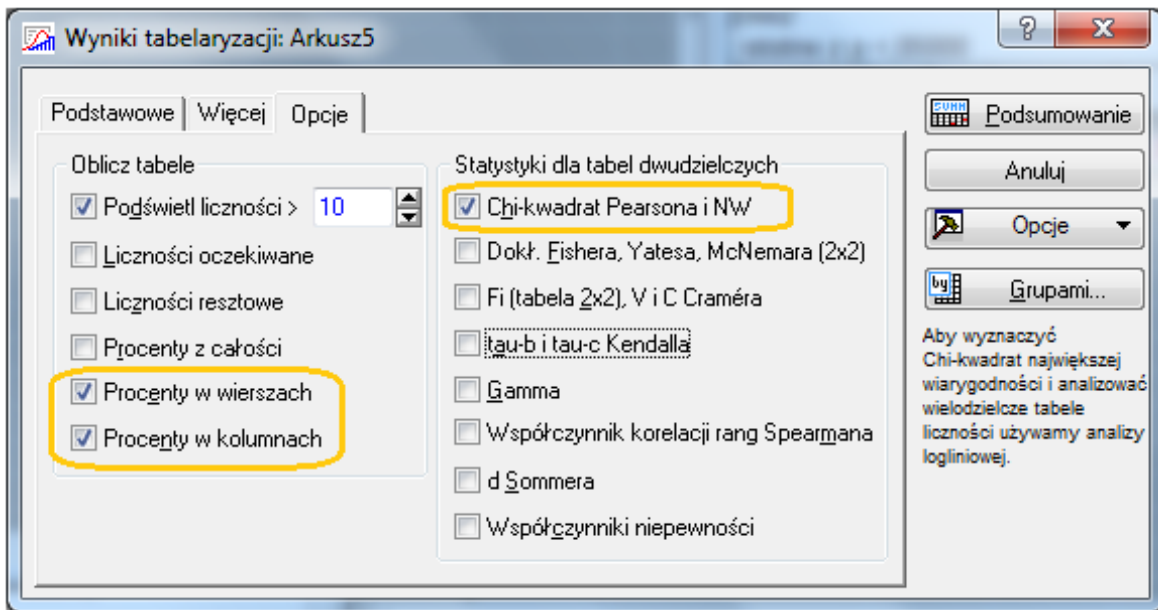
Rys.1 Okno Statystyki podstawowe i tabele (po lewej) i Tabele wielodzzielcze (po prawej).

W oknie Tabele wielodzzielcze należy dokonać wyboru zmiennych w zakładce Zbiorcza. Wybierając te same zmiennej w pierwszej i drugiej liście zmiennych:



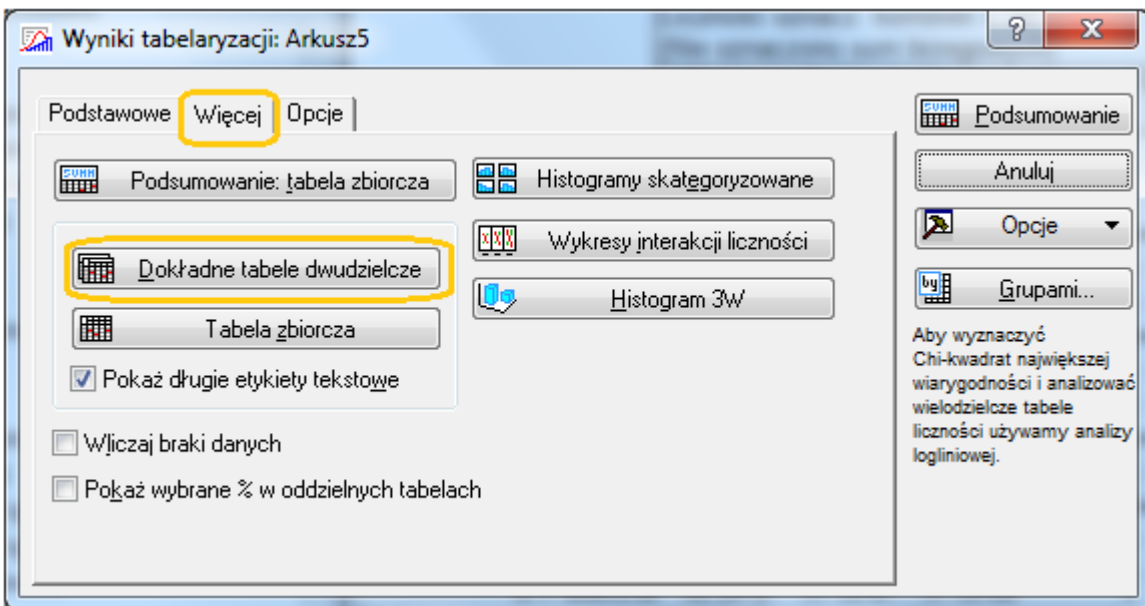
Rys.2 Wybór list zmiennych przy konstruowaniu tablicy wielodzzielczej.

Po wyborze zmiennych należy wcisnąć **OK** w oknie przedstawionym na Rys.2 oraz **OK** w oknie z Rys.1. W wyniku tego działania otwiera się okno **Wyniki tabelaryzacji**. W zakładce **Opcje** należy zaznaczyć, że chcemy przeprowadzić test chi-kwadrat:



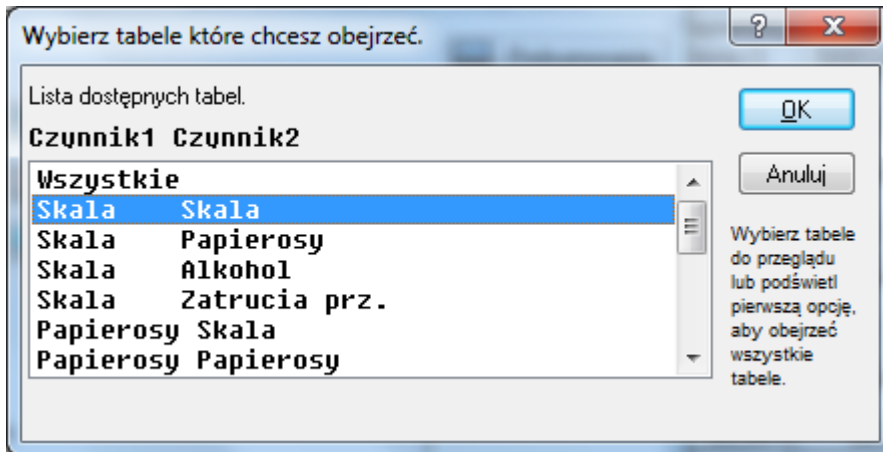
Rys.3 Okno Wyniki tabelaryzacji. Wybór testu chi-kwadrat.

Dobrze zaznaczyć też opcje: procenty w wierszach i kolumnach. Następnie w tym samym oknie, ale w zakładce **Więcej** należy wcisnąć przycisk **Dokładne Tabele dwudzielcze**:



Rys.4 Okno Wyniki tabelaryzacji, zakładka Więcej.

Pojawia się możliwość wyboru tablicy, dla której zostanie przeprowadzony test chi-kwadrat:



Rys.5 Okno wyboru par zmiennych do testu zależności chi-kwadrat.

Po wciśnięciu OK pojawia się wynik testu chi-kwadrat oraz tabela wielodzzielcza:

A.

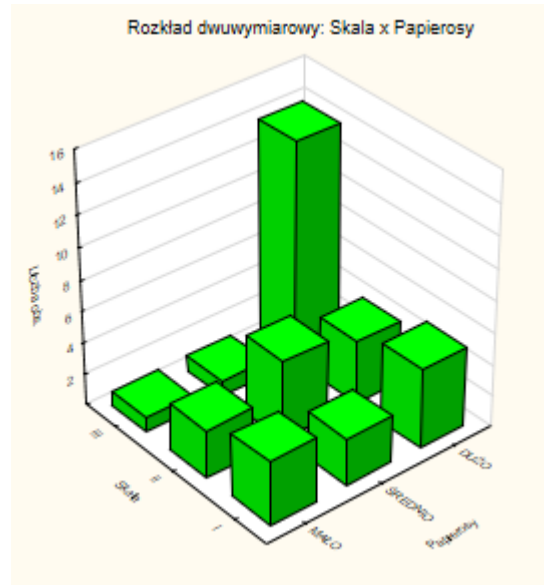
		Statystyka: Skala(3) x Papierosy(3) (Arkusz5)			
statystyka		Chi-kwadr.	df	p	
Chi ² Pearsona		10,84591	df=4	p=,02835	
Chi ² NW		11,50458	df=4	p=,02144	

B.

Podsumowująca tabela dwudzzielcza: częstości obserwowane Liczność oznacz. komórek > 10				
Skala	Papierosy MAŁO	Papierosy ŚREDNIO	Papierosy DUŻO	Wiersz Razem
I	4	3	5	12
%kolumny	50,00%	33,33%	21,74%	
%wiersza	33,33%	25,00%	41,67%	
II	3	5	4	12
%kolumny	37,50%	55,56%	17,39%	
%wiersza	25,00%	41,67%	33,33%	
III	1	1	14	16
%kolumny	12,50%	11,11%	60,87%	
%wiersza	6,25%	6,25%	87,50%	
Ogół	8	9	23	40

Rys.6 A - wynik testu chi-kwadrat, B – tablica wielodzzielcza.

w którym oprócz wyliczonej wartości statystyki wyświetla się poziom prawdopodobieństwa p na podstawie którego dokonujemy decyzji o odrzuceniu lub nie odrzuceniu hipotezy zerowej. Wyniki testu można też zilustrować 3-wymiarowym histogramem (dostępnym pod przyciskiem **Histogram 3W**, w oknie z Rys.4).



Rys.7 Trójwymiarowy histogram ilustrujący zależność dwóch cech jakościowych.

3. Test U Manna-Whitneya

Nieparametryczna odmiana testu t-Studenta dla prób niezależnych. Zbiór danych, które mają zostać poddane temu testowi powinien zawierać: zmienną niezależną (grupującą) i zmienną zależną, która została zmierzona w skali porządkowej lub ilościowej. Test ten weryfikuje hipotezę zerową, która mówi, że dwie losowo wybrane próby pochodzą z tej samej populacji. Przed przystąpieniem do testu **wykonywanego bez oprogramowania statystycznego** należy dodać do danych nową zmienną, która przypisze do obserwacji rangi. Przypisanie rang wykonuje się w następujący sposób:

- 1) Należy posortować rosnąco wartości obydwu prób;
- 2) Następnie do posortowanych obserwacji, rozpoczynając od najmniejszej, przypisujemy kolejne liczby naturalne;
- 3) W przypadku, gdy występują przynajmniej dwie takie same wartości, przyporządkowujemy im tzw. rangi wiązane – czyli średnią arytmetyczną z rang, które miały być im przypisane.

Założenia testu:

Dysponujemy dwoma próbkami o liczebności n_1 i n_2 , pobranymi z populacji, w której cecha ma rozkład typu ciągłego. Dane są przedstawione w skali mierzalnej lub porządkowej. $F(x), G(x)$ to dystrybuanty rozpatrywanych populacji.

Hipoteza zerowa:

$$H_0: F(x) = G(x)$$

Hipoteza alternatywna:

$$H_1: F(x) \neq G(x)$$

Statystyka testowa:

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

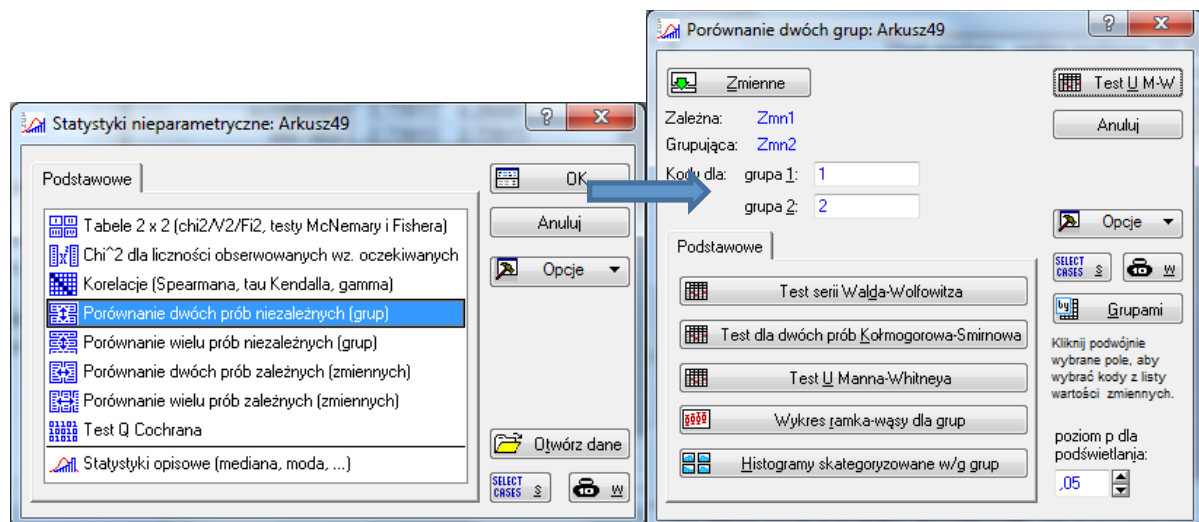
Gdzie: R_1 - suma rang przypisanych do wartości pierwszej próby. Statystyka ta ma skomplikowany rozkład podawany w tablicach statystycznych. Gdy liczebność każdej z prób jest większa niż 20, można skorzystać ze statystyki, która ma w przybliżeniu rozkład normalny:

$$Z = \frac{R_1 - R_2 - (n_1 - n_2)(n + 1)/2}{\sqrt{n_1 n_2 (n + 1)/3}}$$

Gdzie: $n = n_1 + n_2$ – całkowita liczba obserwacji, natomiast R_1 - suma rang przypisanych do wartości drugiej próby.

3.1 Test U Manna-Whitneya w STATISTICE

W ten sposób przeprowadzałoby się ten test pisemnie, albo w prostym arkuszu kalkulacyjnym. W STATISTICE wykonuje się go w następujący sposób: w menu głównym wybieramy **Statystyka / Statystyki nieparametryczne / Porównanie dwóch prób niezależnych**:



Rys.8 Okna statystyk nieparametrycznych (po lewej) oraz testu porównywania grup (po prawej).

W oknie **Porównanie dwóch prób niezależnych** po wyborze odpowiedniej zmiennej zależnej i grupującej można wykonać test U Manna-Whitneya, naciskając przycisk z nazwą tego testu. Wynik wyświetla się w tabeli:

Test U Manna-Whitneya (Arkus49) Względem zmiennej: Zmn2 Zaznaczone wyniki są istotne z $p < ,05000$										
Zmienna	Sum.rang Grupa 1	Sum.rang Grupa 2	U	Z	p	Z popraw.	p	N ważn. Grupa 1	N ważn. Grupa 2	2*1str. dokł. p
Zmn1	169,0000	107,0000	29,00000	2,246420	0,024678	2,246420	0,024678	11	12	0,022500

Liczebności grup
 dokładna wartość prawdopodobieństwa p dla małych liczebności, < 20
 poziom istotności dla Z popr.
 wartość testu skorygowanego ze względu na wagi związane
 poziom istotności dla testu dla wartości testu Z
 wartość testu Manna-Whitneya stosowana, gdy liczebność obu grup jest większa od 20
 wartość testu Manna-Whitneya stosowana dla małych liczebności < 20
 suma rang dla Grupy 2
 suma rang dla Grupy 1

Rys.9 Wynik testu U Manna-Whitneya.

Wynik testu interpretuje się porównując odpowiednią wartość prawdopodobieństwa p (zwracając uwagę na liczebności poszczególnych próbek) z zadanyim poziomem istotności, na jakim test był przeprowadzany.

4. Testy dla próbek zależnych

Są to testy nieparametryczne odpowiadające testowi t dla zmiennych powiązanych. Test ten służy do sprawdzenia istotności różnic pomiędzy zależnymi próbkami (pomiar „PRZED” i „PO”). Może wyróżnić dwa takie testy:

- Test znaków;
- Test kolejności par Wilcoxon

Hipotezy zerowe mówią o tym, że wyniki obydwu próbek są jednakowe. Jedyne założenia, jakie trzeba wziąć pod uwagę przed przystąpieniem do tych testów to ciągłość rozkładu analizowanej zmiennej i możliwość przedstawienia danych przy pomocy skali porządkowej lub ilościowej. W sytuacji, gdy możemy obliczyć różnicę pomiędzy pomiarami stosujemy test kolejności par Wilcoxon, ponieważ jest to mocniejszy test od testu znaków.

4.1 Test znaków

Wykorzystuje znaki porównań pomiędzy odpowiednimi parami pomiarów „przed” i „po”. Jeśli $x_{przed} > x_{po}$, to otrzymujemy „+”, w przeciwnym wypadku „-”. W teście nie bierzemy w ogóle pod uwagę przypadków, dla których $x_{przed} = x_{po}$. Hipoteza zerowa zakłada, że liczba „plusów” jest równa liczbie „minusów”, czyli prawdopodobieństwa ich wystąpienia są takie same: $P(+)=P(-)=0,5$. Test ten stosujemy w szczególności dla zmiennych jakościowych, gdzie „+” mówi o występowaniu, a „-” o nieobecności danego zjawiska.

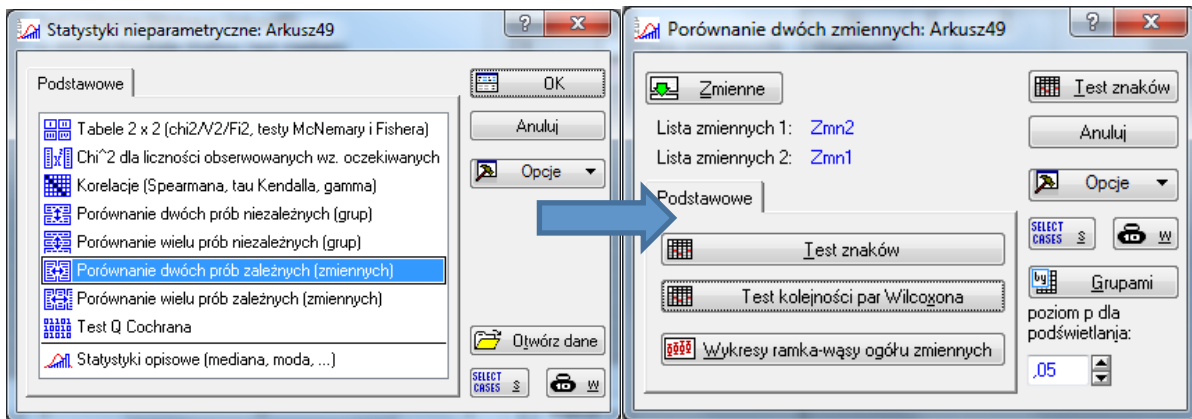
4.2 Test kolejności par Wilcoxon

Poza znakiem różnic, wykorzystuje też ich wartości oraz kolejność, a więc nie nadaje się do testów zmiennej jakościowej. Różnice sortuje się w kolejności rosnącej i przypisuje się im rangi. Osobno sumowane są rangi dla różnic dodatnich, a osobno dla ujemnych. Następnie tak otrzymane dwie sumy są porównywane, mniejsza z nich to wartość testu Wilcoxon. Wartość tą należy porównać z odpowiednią wartością tablicową i na podstawie porównania zdecydować na temat odrzucenia hipotezy zerowej. Test ten jest znacznie mocniejszy od testu znaków, dlatego, właśnie test kolejności

par Wilcoxon'a stosujemy w przypadkach, gdy z powodu nie spełnienia założenia o normalności nie możemy zastosować testu t.

4.3 Testy dla próbek zależnych w STATISTICE

Obydwa testy dostępne są w STATISTICE z menu głównego **Statystyka / Statystyki nieparametryczne / Porównanie dwóch prób zależnych**.



Rys.10 Okno Statystyk nieparametrycznych (po lewej) oraz okno porównywania dwóch zmiennych zależnych (po prawej).

Po naciśnięciu przycisku **Test znaków** w oknie **Porównywanie dwóch zmiennych** otrzymamy wynik testu znaków:

		Test znaków (Arkusz49)			
		Zaznaczone wyniki są istotne z $p < ,05000$			
Para zmiennych		Liczba Niewiąz.	Procent $v < V$	Z	p
Zmn2	& Zmn1	24	100,0000	4,694855	0,000003

Liczebność grup
 % liczebności zmiennych, dla których różnica ma wartość ujemną
 wartość testu znaków
 poziom prawdopodobieństwa p

Rys.11 Wynik testu znaków.

Po naciśnięciu przycisku **Test kolejności par Wilcoxon'a**, otrzymamy wynik tego testu:

		Test kolejności par Wilcoxon'a (Arkusz49)			
		Zaznaczone wyniki są istotne z $p < ,05000$			
Para zmiennych		N Ważnych	T	Z	p
Zmn2	& Zmn1	24	0,00	4,285714	0,000018

Liczebność grup
 wartość testu Wilcoxa dla grup o liczebności > 25
 wartość testu Wilcoxa dla grup o liczebności < 25
 prawdopodobieństwo p

Rys.12 Wynik testu kolejności par Wilcoxon'a.

Wynik testów interpretuje się porównując odpowiednią wartość prawdopodobieństwa p z zadaniem poziomem istotności, na jakim test był przeprowadzany.

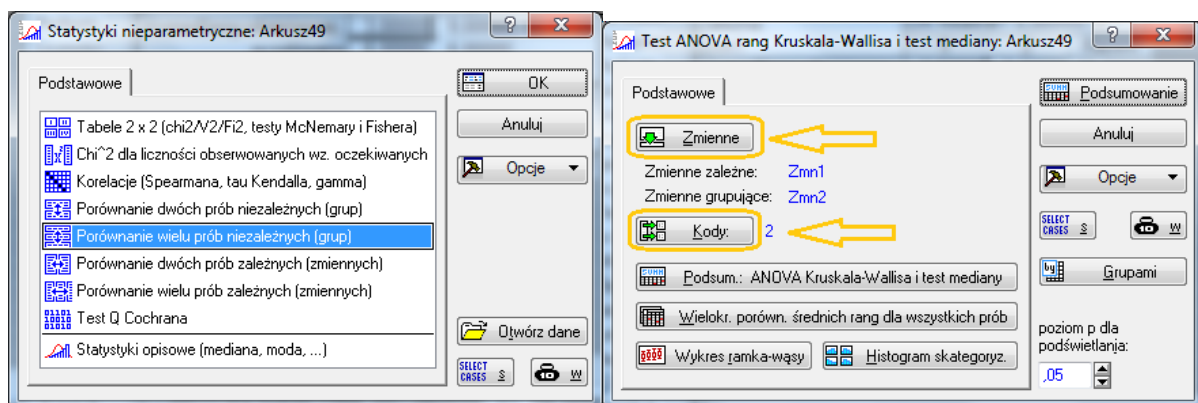
5. Test Kruskala-Wallisa i test mediany

Jest to nieparametryczny odpowiednik dla jednoczynnikowej analizy wariancji (ANOVA). Zakłada, że zmienna jest przedstawiona przy pomocy skali przedziałowej. Za pomocą tego testu oceniamy, czy n niezależnych próbek pochodzi z tej samej populacji. Liczebność prób może być różna. Jednocześnie możemy porównywać do 10 prób na raz. Interpretacja wyników w tym teście jest identyczna jak w przypadku analizy ANOVA, z tym, że zamiast wartości średnich wykorzystuje rangi. W przypadku negatywnego wyniku testu Kruskala-Wallisa, można wykonać odpowiedni test Post-Hoc, aby sprawdzić, które próbki spowodowały odrzucenie hipotezy zerowej.

5.1 Test Kruskala-Wallisa i test mediany w STATISTICE

Podczas wykonywania tego testu w STATISTICE, uzyskuje się również wyniki testu mediany, w którym hipoteza zerowa zakłada, że wszystkie próbki pochodzą z populacji o takiej samej wartości mediany. Do zweryfikowania takiej hipotezy zerowej STATISTICA sprawdza ile przypadków w każdej z prób wypada poniżej oraz powyżej mediany. Do zweryfikowania hipotez wykorzystuje statystykę χ^2 dla wyników zawartych w tablicy kontyngencji $2 \times n$. Test jest szczególnie pomocny w przypadkach, gdy skala pomiarowa zawiera sztuczne ograniczenia i wiele przypadków znajduje się na końcach skali.

W Statistice test Kruskala-Wallisa dostępny jest z menu głównego: **Statystyka / Statystyki nieparametryczne / Porównanie wielu prób niezależnych (grup)**.



Rys.13 Okno Statystyk nieparametrycznych (po lewej) oraz testu

Wyboru zmiennych dokonuje się klikając na przycisk **Zmienne**. Należy wskazać odpowiednią zmienną zależną i odpowiednią grupującą. Następnie klikając przycisk **Kody** można wskazać, o które grupy nam chodzi (jeśli interesują nas wszystkie grupy, to w oknie z Kodami należy wcisnąć przycisk **Wszystko**).

Następnie wciskając przycisk: **Podsum: ANOVA Kruskala-Wallisa i test mediany** oprócz wyniku testu Kruskala-Wallisa otrzymujemy też wynik testu mediany. W oknie dla tego testu dostępne są też przyciski umożliwiające wykonanie Wykresów ramka-wąsy lub skategoryzowanych histogramów. W przypadku, gdy test Kruskala-Wallisa da negatywny wynik, wciskając przycisk **Wielokr. Porówn. Średnich rang** dla wszystkich prób, uzyskujemy wynik testu Post-hoc – polega on na porównaniu średnich rang przypisanych do poszczególnych przypadków przez oprogramowanie. Wykonując test

Post-hoc, możemy sprawdzić, które próbki odpowiadają za negatywny wynik tego testu, a więc istotnie różnią się od pozostałych. Pomocą w interpretacji danych mogą służyć wykres ramka-wąsy i histogramy.

Ćwiczenia

UWAGA: staraj się dbać o to, co znajduje się w skoroszybie, poszczególne foldery nazywaj numerami zadań (patrz rysunek na końcu instrukcji). Zostaw w skoroszybie tylko potrzebne wyniki, wszystko, co nieistotne usuń.

Zad.1 Otwórz dane **dane6_zad1.sta**. Dane zawierają wyniki ankiety przeprowadzonej na pewnej grupie zawodowej wraz ze skalą natężenia choroby. W danych znajdują się informacje na temat ilości wypalanych papierosów, ilości wypijanego alkoholu oraz stopnia narażenia na substancje toksyczne w miejscu pracy. Wartości skali natężenia choroby odpowiadają I – najłżejszej postaci choroby, III – najcięższej postaci choroby. Na poziomie istotności $\alpha = 0,05$ przeprowadź test zależności chi-kwadrat dla par zmiennych: Skala i Papierosy, Skala i Alkohol, Skala i Zatrucia_przemysłowe.

Wskazówka: wybieraj kolejno odpowiednie pary w oknie przedstawionym na rys.

- Od jakich czynników zależy stopień nasilenia choroby? Jako ilustrację do wyników testów wykonaj histogramy 3D.
- Na podstawie tabeli wielodzielczej odpowiedz na pytania: jaki procent osób, które cierpią na najłżejszą postać choroby stanowią osoby w ogóle nie pijące alkoholu? Jaki procent wśród osób, które w ogóle nie piją alkoholu stanowią Ci, którzy mają najłżejszą postać choroby?

(patrz rozdz. 2)

Zad.2 Otwórz **dane6_zad2.sta**. W dwóch ogrodach zoologicznych przeprowadzono pomiary długości zółwi w różnym wieku. Na poziomie istotności $\alpha = 0,05$ zweryfikuj hipotezę zerową mówiącą o tym, że zółwie z obydwu ogrodów zoologicznych są tego samego gatunku. Sprawdź, czy do weryfikacji tej hipotezy możesz zastosować test t dla zmiennych niepowiązanych weryfikujący hipotezę zerową o równości średnich (sprawdź, czy zachodzi założenie o normalności rozkładu).

Wskazówka: w celu sprawdzenia normalności rozkładu wykorzystaj test Shapiro-Wilka – Statystyka/Statystyki podstawowe i tabele/Statystyki opisowe, zakładka Normalność).

Jeśli dane z prób nie mają rozkładu normalnego, to znaczy, że nie można przeprowadzić testu t. W takiej sytuacji wykonaj test U Manna-Whitneya.

(patrz rozdz. 3)

Zad.3 Otwórz dane **dane6_zad3.sta**. Dane zawierają pomiary tętna dla tej samej grupy osób przed i po 4 minutowym wysiłku na bieżni. Na poziomie istotności $\alpha = 0,05$ zweryfikuj hipotezę mówiącą o tym, że wysiłek fizyczny wpływa na wartość tętna. Sprawdź, czy możesz wykonać test t dla zmiennych zależnych (sprawdź, czy zachodzi założenie o normalności rozkładu). Jeśli pomiary nie mają rozkładu normalnego, zamiast testu t wykonaj test kolejności par Wilcoxon. Wykonaj też test znaków, porównaj wyniki z wynikami testu kolejności par Wilcoxon. Gdyby testy dały różne wyniki, to wyników którego z nich powinniśmy się trzymać?

(patrz rozdz. 4)

Zad.4 Otwórz dane **dane6_zad4.sta**. Dane zawierają informację na temat działania różnych metod treningu biegowego u 20 osób. Każda metoda została oceniona przez trenera w umownej skali (od 0 do 80 punktów). Na poziomie istotności $\alpha = 0,05$ zweryfikuj hipotezę mówiąca o tym, że każdy sposób treningu daje takie same wyniki. Wiedząc, że założenia analizy wariancji nie są spełnione, przeprowadź test Kruskala-Wallisa. W razie negatywnego wyniku testu przeprowadź test Post-hoc, sporządź również wykres ramka wąsy dla Punktów uzyskanych przez różne metody. Analizując wyniki testu Post-hoc oraz wykres pudełkowy odrzuć tą z metod, która najbardziej odstaje od pozostałych i dla tak zmodyfikowanego zbioru danych ponownie przeprowadź test Kruskala-Wallisa.

(patrz rozdz. 5)

