

Statystyka i teoria obsługi masowej

Laboratorium 4

Dawid Polak

Zad 1.

Przeliczyłem wartości zmiennych w arkuszu.

Poniżej otrzymany uzupełniony w wartości arkusz:

	1 Z1	2 Z2	3 Z3	4 X	5 Grupa
1	23,38	15,15	24,27	26,78	0
2	20,51	12,17	20,06	16,02	0
3	22,64	32,55	16,17	15,77	1
4	19,83	12,03	21,37	18,95	0
5	22,28	15,79	19,98	26,26	1
6	20,12	17,04	20,53	19,04	0
7	21,95	25,95	21,34	14,2	0
8	21,71	12,58	15,77	15,29	0
9	17,2	15,35	20,64	21,14	1
10	18,82	12,49	14,48	17,71	1
11	19,37	13,78	18,52	14,9	1
12	17,32	14,88	23,84	20,61	1
13	22,36	21,74	21,78	18,67	0
14	17,26	13,27	29,54	16,67	0
15	18,78	19	10,57	24,21	1
16	19,61	14,21	29,64	14,77	0
17	20,29	14,96	18,58	14,51	1
18	19,25	18,95	20,83	19,84	0
19	18,54	31,06	16,84	17,69	1
20	20,33	22,4	14,77	15,89	1

Zad 2.1.

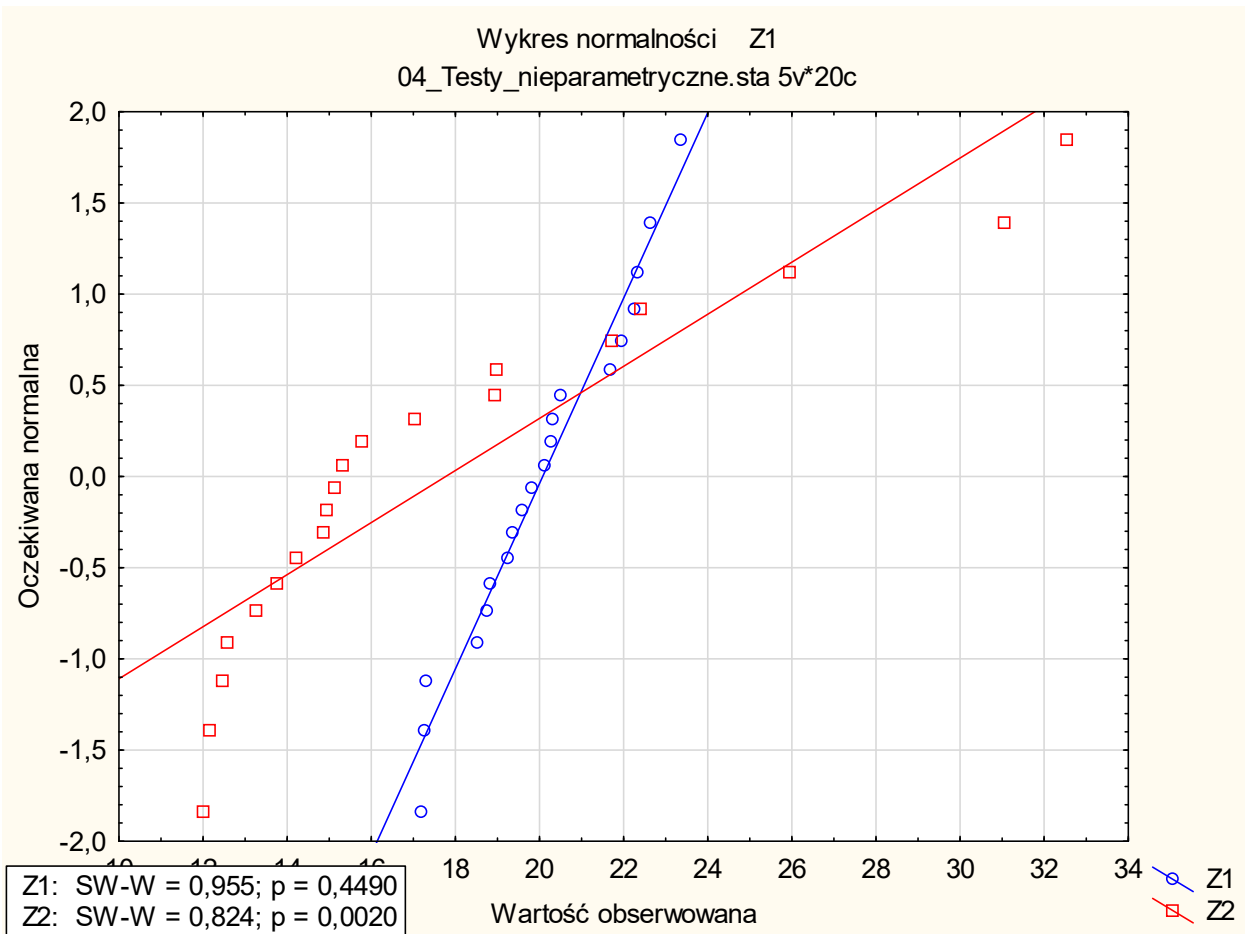
Dane:

poziom istotności $\alpha = 0.09$

Z1, Z2 - zmienne przedstawiające wyniki pomiarów dla tych samych obiektów w kilku warunkach.

H0 - nie nastąpiła istotna zmiana między wynikiem pomiaru Z1 a Z2

Najpierw musimy sprawdzić hipotezę o normalności rozkładów.



Pomiar Z1 może mieć rozkład normalny (nie ma podstaw do odrzucenia założenia), lecz pomiar Z2 ma za małe p względem przyjętego poziomu istotności - **mamy podstawy do stwierdzenia, że rozkład tego pomiaru nie jest normalny**, więc do porównania tych dwóch wartości nie możemy wykorzystać **testu parametrycznego**.

Klikam zakładkę Statystyka, moduł "Nieparametryczne". Wartości dotyczą tego samego obiektu, więc wykonujemy test porównujący dwie próby zależne.

Wykonuję test znaków, który zlicza ilość dodatnich i ujemnych różnic zmiennych i je zestawia ze sobą.

Para zmiennych	Test znaków (04_Testy_nieparametryczne.sta)			
	Zaznaczone wyniki są istotne z $p < ,09000$			
	Liczba Niewiąz.	Procent $v < V$	Z	p
Z1 & Z2	20	25,00000	2,012461	0,044171

Parametr p wyszedł mniejszy od założonego poziomu istotności, więc hipotezę zerową należy odrzucić - **różnice pomiarów są istotne statystycznie.**

Dodatkowo wykonuję Test kolejności par Wilcoxona (oprócz samych znaków dodatkowo brana pod uwagę jest wartość bezwzględna różnic wartości zmiennych), który potwierdza **odrzućcie hipotezę zerową.**

Para zmiennych	Test kolejności par Wilcoxona (04_Testy_nieparametryczne.sta) Zaznaczone wyniki są istotne z $p < ,09000$			
	N Ważnych	T	Z	p
Z1 & Z2	20	54,00000	1,903966	0,056915

Zad 2.2.

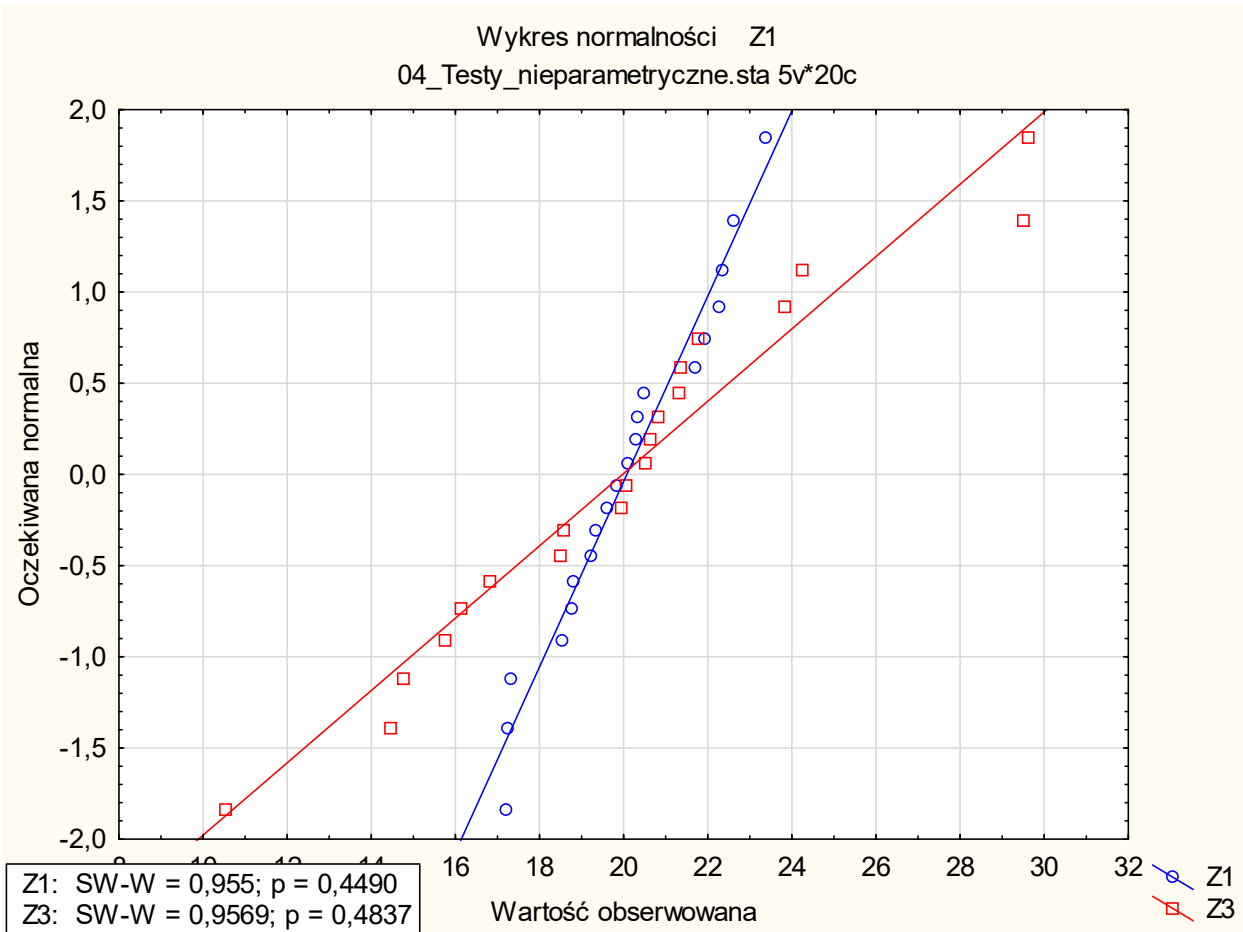
Dane:

poziom istotności $\alpha = 0.09$

Z1, Z3 - zmienne przedstawiające wyniki pomiarów dla tych samych obiektów w kilku warunkach.

H0 - nie nastąpiła istotna zmiana między wynikiem pomiaru Z1 a Z3.

Najpierw musimy sprawdzić hipotezę o normalności rozkładów.



Pomiar Z1 oraz Z3 może mieć rozkład normalny (**nie ma podstaw do odrzucenia założenia**). Możemy więc użyć testów parametrycznych, które mają większą moc i dają nam większą pewność, że nasze wyniki będą wiarygodne.

Wykonałem więc Test t dla prób zależnych, jako że wartości dotyczą pomiarów tego samego obiektu.

Zmienna	Test T dla prób zależnych (04_Testy_nieparametryczne.sta)						
	Zaznaczone różnice są istotne z $p < ,09000$						
	Średnia	Odch.st.	Ważnych	Różnica	Odch.st. Różnica	t	df
Z1	20,07750	1,848453					
Z3	19,97600	4,693976	20	0,101500	5,223433	0,086901	19

Zmienna	Test T dla prób zależnych (04_Testy_nieparametryczne.sta)		
	p	Ufność -95,000%	Ufność +95,000%
Z1			
Z3	0,931659	-2,34314	2,546142

Dla przyjętego poziomu istotności **nie mamy podstaw do odrzucenia hipotezy 0**, mówiącej o tym, że wystąpiły istotne zmiany między wartością pomiaru Z1 a Z3.

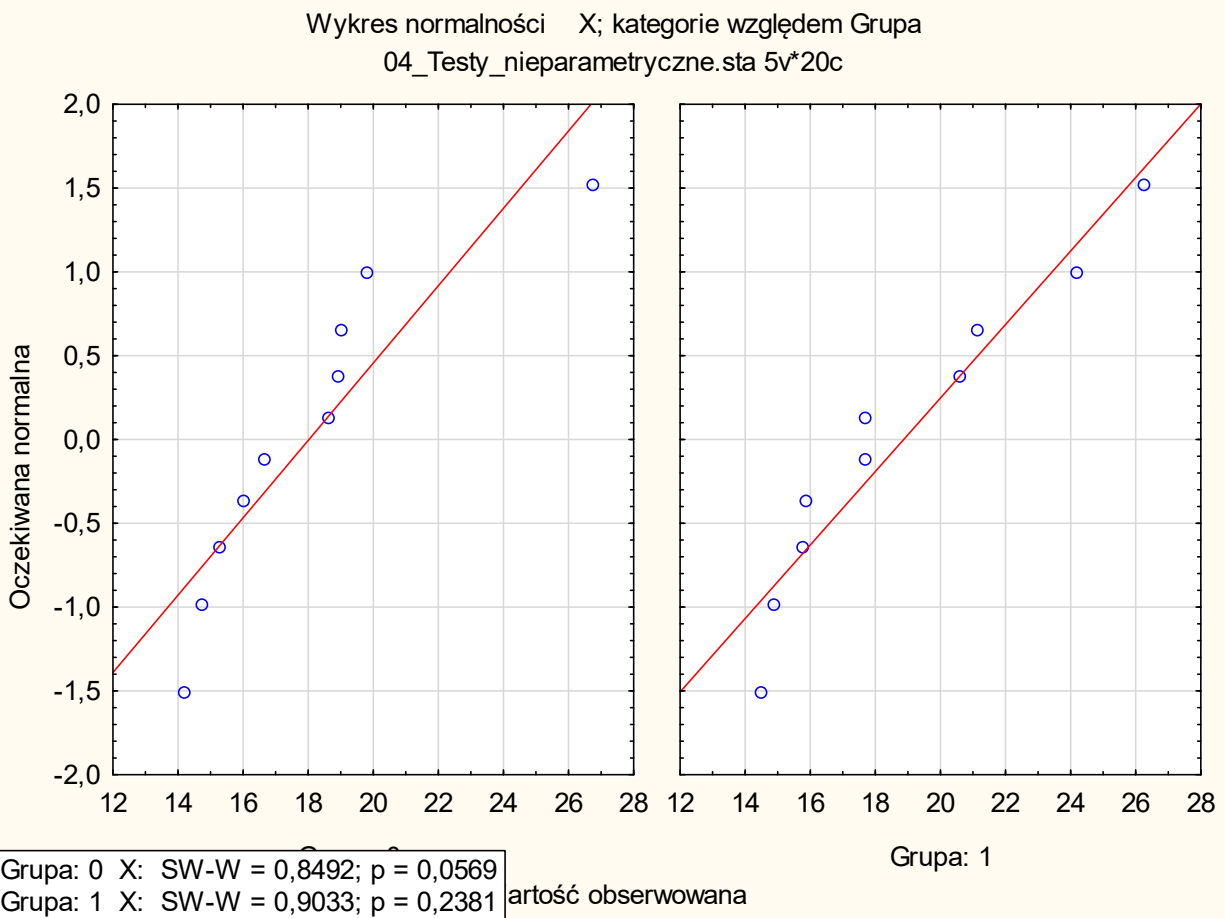
Zad 3.

Dane:

poziom istotności $\alpha = 0.1$

H0 - nie nastąpiła istotna zmiana w wartości zmiennej X pomiędzy grupami wyznaczonymi przez zmienną Grupa - rozkład zmiennej X w Grupie 0 jest taki sam jak rozkład zmiennej X w Grupie 1

Wykonuje test Shapiro-Wilka który jest testem normalności. W zakładce Skateryzowane klikam "włącz"



Na podstawie wyników testów przy przyjętym poziomie istotności mamy podstawy by stwierdzić, że **grupa 0 nie spełnia założeń rozkładu normalnego**, więc muszę użyć test nieparametryczny.

W takim razie użyłem test nieparametryczny - porównanie dwóch zmiennych niezależnych (grup) - Test Walda Wolfowitza

Zmienna	Test serii Walda-Wolfowitza (04_Testy_nieparametryczne.sta)						
	Względem zmiennej: Grupa						
	Zaznaczone wyniki są istotne z $p < 0,10000$						
	N ważn. Grupa 1	N ważn. Grupa 2	Średnia Grupa 1	Średnia Grupa 2	Z	p	Z skoryg.
X	10	10	18,02300	18,86900	0,00	1,000000	-0,229734

Zmienna	Test serii Walda-Wolfowitza (04_Testy_nieparametryczne.sta)		
	Względem zmiennej: Grupa		
	p	Liczba serii	Liczba wiązanych
X	0,818299	11	0

Przy założonym poziomie istotności **nie mamy podstaw do odrzucenia hipotezy zerowej**.
Jako że mamy mało liczne próby, dodatkowo bierzemy pod uwagę z skorygowane i
odpowiadającą jemu wartość p .

Dodatkowo analogicznie do ćwiczeń posortowałem sobie wartości X rosnąco, by zobaczyć
sobie jak rozkładają się serie.

	1 Z1	2 Z2	3 Z3	4 X	5 Grupa
1	21,95	25,95	21,34	14,2	0
2	20,29	14,96	18,58	14,51	1
3	19,61	14,21	29,64	14,77	0
4	19,37	13,78	18,52	14,9	1
5	21,71	12,58	15,77	15,29	0
6	22,64	32,55	16,17	15,77	1
7	20,33	22,4	14,77	15,89	1
8	20,51	12,17	20,06	16,02	0
9	17,26	13,27	29,54	16,67	0
10	18,54	31,06	16,84	17,69	1
11	18,82	12,49	14,48	17,71	1
12	22,36	21,74	21,78	18,67	0
13	19,83	12,03	21,37	18,95	0
14	20,12	17,04	20,53	19,04	0
15	19,25	18,95	20,83	19,84	0
16	17,32	14,88	23,84	20,61	1
17	17,2	15,35	20,64	21,14	1
18	18,78	19	10,57	24,21	1
19	22,28	15,79	19,98	26,26	1
20	23,38	15,15	24,27	26,78	0

Natępnie wykonuję test Kołomogorowa-Smirnowa

Zmienna	Test Kołmogorowa-Smirnowa (04_Testy_nieparametryczne.sta) Względem zmiennej: Grupa Zaznaczone wyniki są istotne z $p < ,10000$						
	Maks.uj. Różnica	Maks.dod Różnica	p	Średnia Grupa 1	Średnia Grupa 2	Odch.std Grupa 1	Odch.std Grupa 2
X	-0,300000	0,100000	$p > .10$	18,02300	18,86900	3,663757	4,045671

Zmienna	Test Kołmogorowa-Smirnow a	
	N ważn. Grupa 1	N ważn. Grupa 2
X	10	10

$p > 0.1$, więc **nie mamy podstaw do odrzucenia hipotezy 0**, która mówi o tym, że rozkłady w obu grupach są jednakowe.

Następnie wykonuję ostatni test dostępny z grupy testów, który jest domyślny i jest z nich najmocniejszy.

Z tego co mi się wydaje (ale nie jestem pewny) powinienem odznaczyć poprawkę na ciągłość. Poniżej wynik:

Zmienna	Test U Manna-Whitneya (04_Testy_nieparametryczne.sta) Względem zmiennej: Grupa Zaznaczone wyniki są istotne z $p < ,10000$					
	Sum.rang Grupa 1	Sum.rang Grupa 2	U	Z	p	Z popraw.
X	100,0000	110,0000	45,00000	-0,377964	0,705457	-0,377964

Zmienna	Test U Manna-Whitneya (04_Testy_nieparametryczne.sta) Względem zmiennej: Grupa			
	p	N ważn. Grupa 1	N ważn. Grupa 2	2*1str. dokł. p
X	0,705457	10	10	0,739364

Dla prób zawierających więcej niż 20 pomiarów możemy wykorzystać statystykę Z - podstawowa wersja testu.

Rangi nam się nie powtarzają, więc Z poprawione jest takie samo - uwzględniona poprawka na rangi wiązane - powtarzające się rangi w próbie).
Ostatnie p jest dla prób niewielkich (mniej niż 20 elementów).

Na podstawie otrzymanego wyniku przy założonym poziomie istotności można stwierdzić, że **nie ma podstaw do odrzucenia hipotezy 0** mówiącej, że nastąpiła istotna zmiana w wartości zmiennej X pomiędzy grupami wyznaczonymi przez zmienną Grupa - rozkład zmiennej X w Grupie 0 jest taki sam jak rozkład zmiennej X w Grupie 1

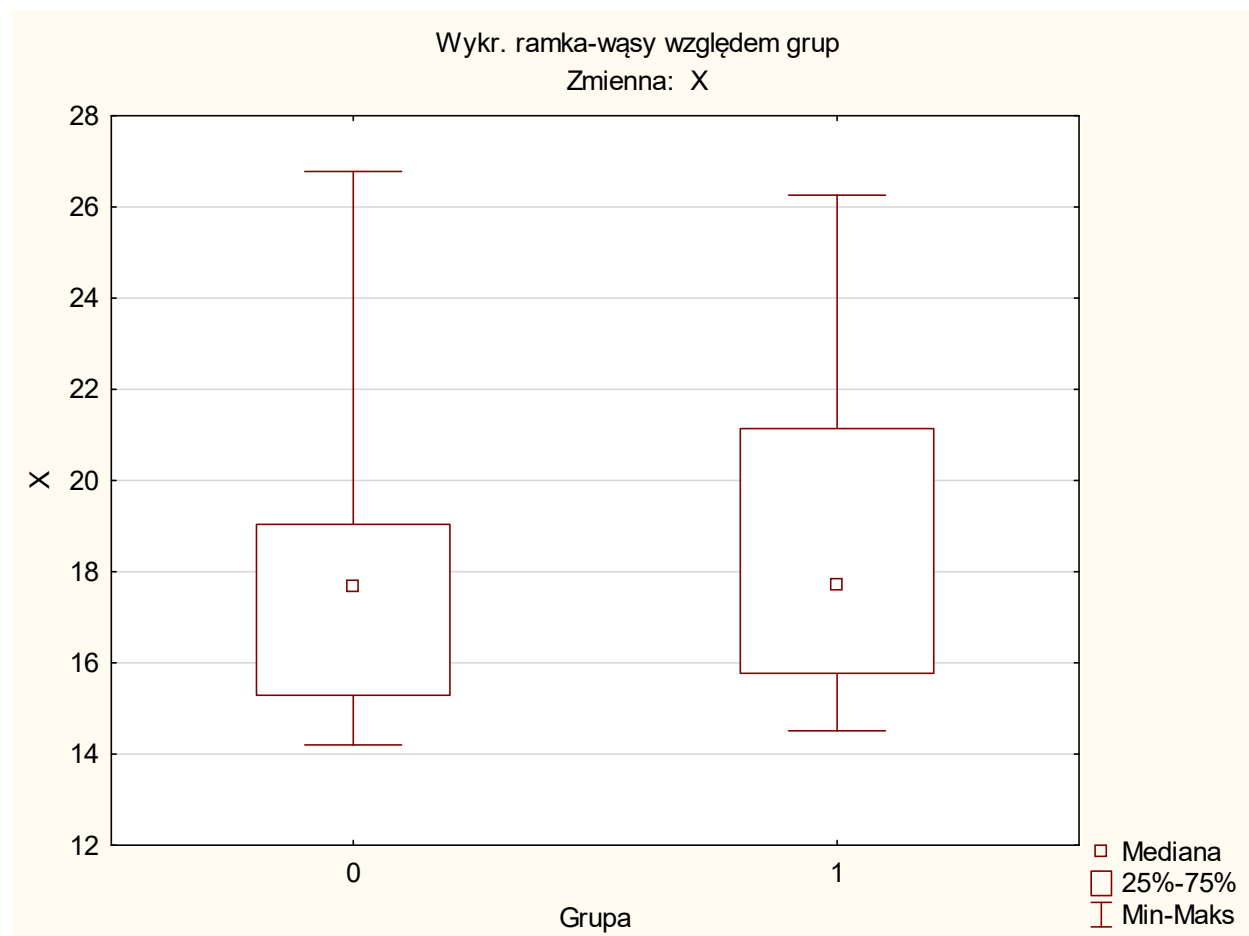
Na wszelki wypadek wykonałem również test z zaznaczoną poprawką na ciągłość, na podstawie którego również nie ma podstawy do odrzucenia hipotezy 0 (nie ma kolosalnych różnic).

Zmienna	Test U Manna-Whitneya (z poprawką na ciągłość) (04_Testy_nieparametryczne.sta) Względem zmiennej: Grupa					
	Sum.rang Grupa 1	Sum.rang Grupa 2	U	Z	p	Z popraw.
X	100,0000	110,0000	45,00000	-0,340168	0,733730	-0,340168

Zmienna	Test U Manna-Whitneya (z poprawką na ciągłość) (04_Testy_nieparametryczne.sta) Względem zmiennej: Grupa			
	p	N ważn. Grupa 1	N ważn. Grupa 2	2*1str. dokł. p
X	0,733730	10	10	0,739364

Dodatkowo stworzyłem wykresy w celu wizualizacji badanego przypadku.

Wykres ramka wąsy względem grup:



Skategoryzowany histogram (możemy podjąć wartości średnie):

Skategoryzowany histogram

Zmienna: X

Grupa: 0 $X = 10 \cdot 2 \cdot \text{Normal}(\text{Średnia}=18,023; \text{Sigma}=3,6638)$

Grupa: 1 $X = 10 \cdot 2 \cdot \text{Normal}(\text{Średnia}=18,869; \text{Sigma}=4,0457)$

