

Metody estymacji i Wnioskowania Statystycznego

Ćwiczenia nr 4

Zadanie 1

X – rzut, $X \sim \text{dwumianowy}(120; 0,5)$

Zmienna losowa X może przyjmować tylko dwie wartości (dwie klasy), gdyż wynik rzutu może przyjąć jedną z dwu wartości – konterfekt lub reszkę.

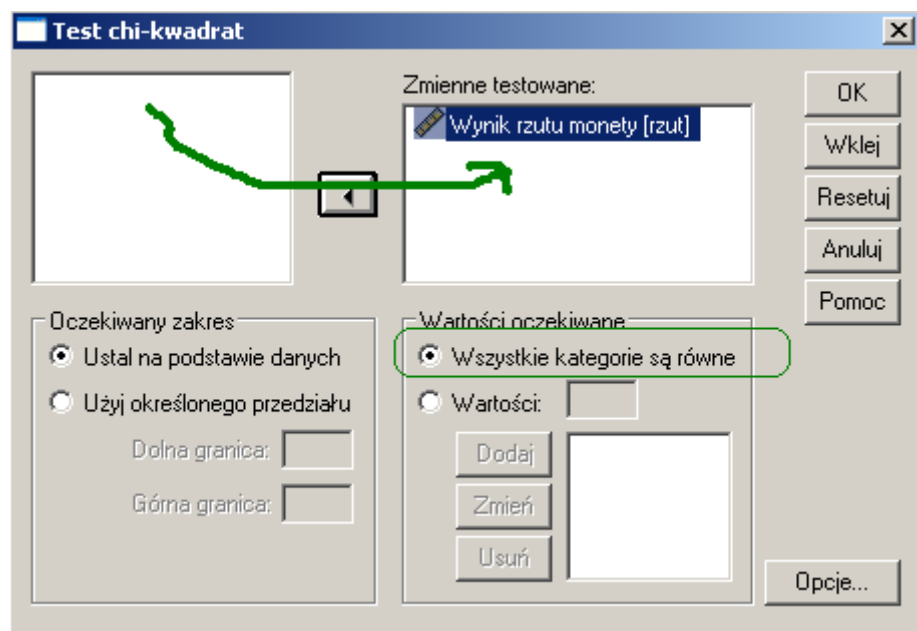
Sprawdzamy, czy moneta jest symetryczna, czyli stawiamy hipotezę:

H: $p_1 = 0,5$, $p_2 = 0,5$

K: $\neg H$

SPSS: Plik >> Otwórz >> Dane >> *moneta.sav*

Analiza >> Testy nieparametryczne >> Chi-kwadrat



Test chi-kwadrat

Tabela częstości

Wynik rzutu monety

	Obserwowane N	Oczekiwane N	Reszty
konterfekt	62	60,0	2,0
reszka	58	60,0	-2,0
Ogółem	120		

Statystyki testu

	Wynik rzutu monety
Chi-kwadrat ^a	,133
df	1
Istotność asymptotyczna	,715

a. 0 komórek (,0%) ma liczebność oczekiwaną mniejszą od 5. Minimalna liczebność oczekiwana w komórce wynosi 60,0.

Istotność asymptotyczna = $0,715 > 0,05 \Rightarrow$ nie ma podstaw do odrzucenia hipotezy zerowej H_0

Odp. Wnioskujemy, że moneta jest symetryczna.

Zadanie 2

X – kolor kwiatów

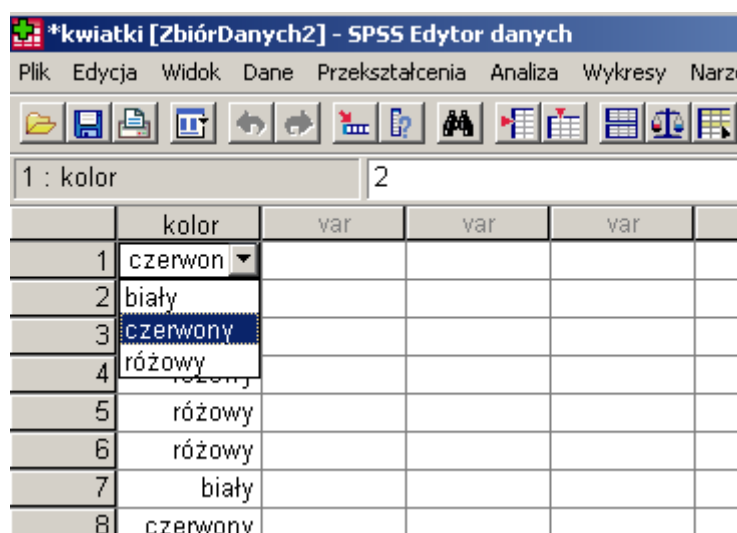
Zmienna losowa X może przyjmować trzy wartości (3 klasy): 1-biały, 2-czerwony, 3-różowy.

$$H: p_1 = 0,25, p_2 = 0,25, p_3 = 0,5$$

$$K: \neg H$$

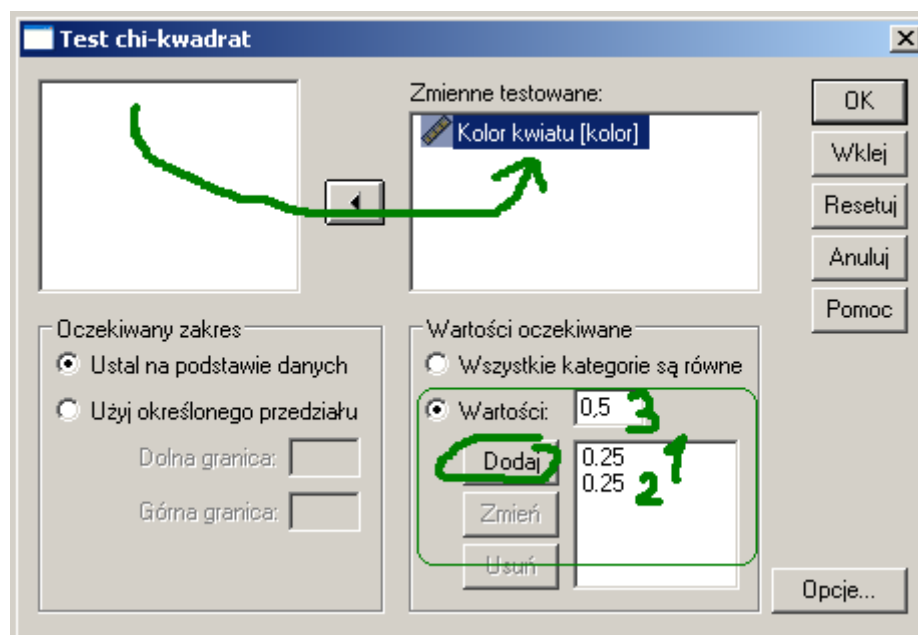
SPSS: Plik >> Otwórz >> Dane >> *kwiatki.sav*

Widok >> Etykiety wartości



	kolor	var	var	var
1	czerwony			
2	biały			
3	czerwony			
4	różowy			
5	różowy			
6	różowy			
7	biały			
8	czerwony			

Analiza >> Testy nieparametryczne >> Chi-kwadrat



Test chi-kwadrat

Zmienne testowane:
Kolor kwiatu [kolor]

Oczekiwany zakres:
☒ Ustal na podstawie danych
☐ Użyj określonego przedziału
Dolna granica:
Górna granica:

Wartości oczekiwane:
☐ Wszystkie kategorie są równe
☒ Wartości: 0,5
Dodaj 0,25 0,25
Zmień
Usuń

OK
Wklej
Resetuj
Anuluj
Pomoc
Opcje...

Test chi-kwadrat

Tabela częstości

Kolor kwiatu

	Obserwowane N	Oczekiwane N	Reszty
biały	31	30,0	1,0
czerwony	29	30,0	-1,0
różowy	60	60,0	,0
Ogółem	120		

Statystyki testu

	Kolor kwiatu
Chi-kwadrat ^a	,067
df	2
Istotność asymptotyczna	,967

a. 0 komórek (,0%) ma liczebność oczekiwaną mniejszą od 5. Minimalna liczebność oczekiwana w komórce wynosi 30,0.

Istotność = 0,967 > 0,05 \Rightarrow nie ma podstaw do odrzucenia hipotezy zerowej H_0

Odp. Wnioskujemy, że wnioski profesora są słuszne.

Zadanie 3

X – rejon wydobywania, Y – jakość surowca

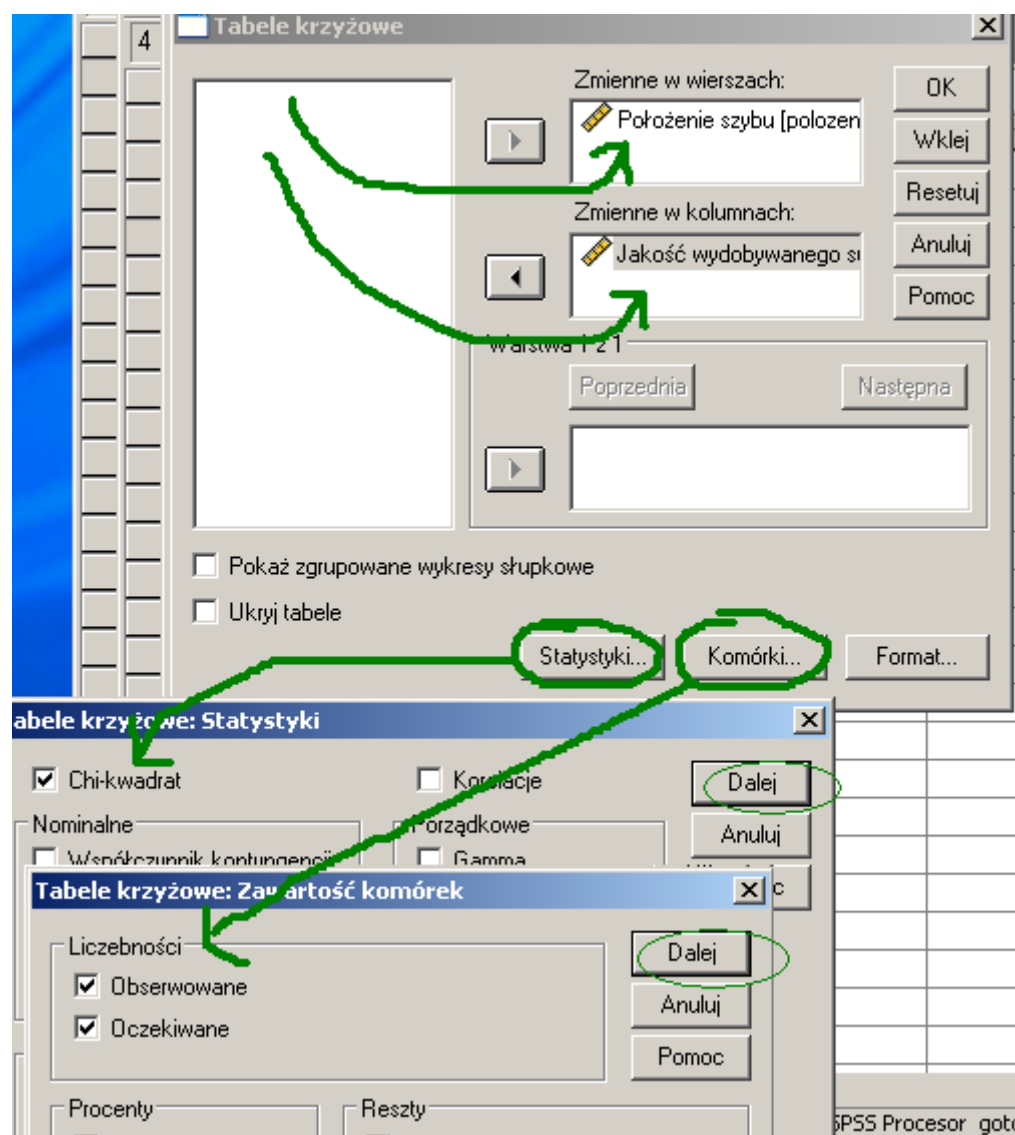
H: X i Y są niezależne

K: $\neg H$ (istnieje zależność między X i Y)

SPSS: Plik >> Otwórz >> Dane >> wydobywanie.sav

Widok >> Etykiety wartości

Analiza >> Opis statystyczny >> Tabele krzyżowe



Informacja o analizowanych danych

	Obserwacje					
	Uwzględnione		Wykluczone		Ogółem	
	N	Procent	N	Procent	N	Procent
Położenie szybu * Jakość wydobywanego surowca	109	100,0%	0	,0%	109	100,0%

Tabela krzyżowa Położenie szybu * Jakość wydobywanego surowca

			Jakość wydobywanego surowca				Ogółem
			słaba	przeciętna	dobra	znakomita	
Położenie szybu	północ	Liczebność	5	9	6	8	28
		Liczebność oczekiwana	5,9	7,7	8,0	6,4	28,0
	południe	Liczebność	6	6	12	6	30
		Liczebność oczekiwana	6,3	8,3	8,5	6,9	30,0
	wschód	Liczebność	5	10	7	5	27
		Liczebność oczekiwana	5,7	7,4	7,7	6,2	27,0
	zachód	Liczebność	7	5	6	6	24
		Liczebność oczekiwana	5,1	6,6	6,8	5,5	24,0
Ogółem	Liczebność	23	30	31	25	109	
	Liczebność oczekiwana	23,0	30,0	31,0	25,0	109,0	

Testy Chi-kwadrat

	Wartość	df	Istotność asymptotyczna (dwustronna)
Chi-kwadrat Pearsona	5,923 ^a	9	,748
Iloraz wiarygodności	5,741	9	,765
Test związku liniowego	,430	1	,512
N Ważnych obserwacji	109		

a. ,0% komórek (0) ma liczebność oczekiwaną mniejszą niż 5.
Minimalna liczebność oczekiwana wynosi 5,06.

Istotność = 0,748 > 0,05 \Rightarrow nie ma podstaw do odrzucenia hipotezy H

Odp. Wnioskujemy, że jakość ropy nie zależy od miejsca jej wydobywania.

Zadanie 4

X – ocena z filozofii, Y – ocena z etyki

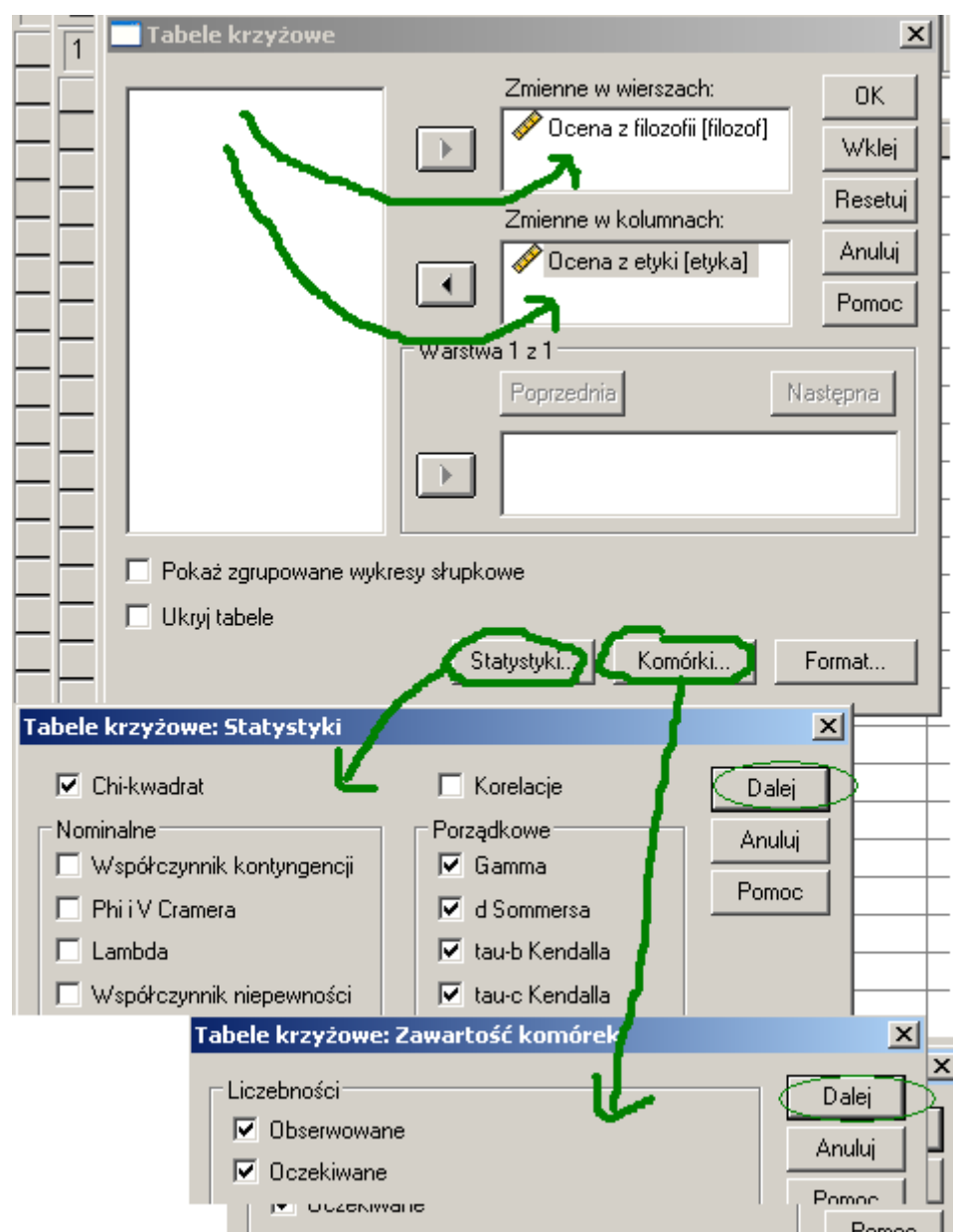
H: X i Y są niezależne

K: $\neg H$ (istnieje zależność między X i Y)

SPSS: Plik >> Otwórz >> Dane >> *oceny2.sav*

Widok >> Etykiety wartości

Analiza >> Opis statystyczny >> Tabele krzyżowe



Informacja o analizowanych danych

	Obserwacje					
	Uwzględnione		Wykluczone		Ogółem	
	N	Procent	N	Procent	N	Procent
Ocena z filozofii * Ocena z etyki	136	100,0%	0	,0%	136	100,0%

Tabela krzyżowa Ocena z filozofii ^ Ocena z etyki

			Ocena z etyki				Ogółem
			2	3	4	5	
Ocena z filozofii	2	Liczebność	17	5	0	0	22
		Liczebność oczekiwana	5,7	5,2	5,8	5,3	22,0
	3	Liczebność	11	17	7	2	37
		Liczebność oczekiwana	9,5	8,7	9,8	9,0	37,0
	4	Liczebność	6	8	14	8	36
		Liczebność oczekiwana	9,3	8,5	9,5	8,7	36,0
	5	Liczebność	1	2	15	23	41
		Liczebność oczekiwana	10,6	9,6	10,9	9,9	41,0
Ogółem	Liczebność	35	32	36	33	136	
	Liczebność oczekiwana	35,0	32,0	36,0	33,0	136,0	

Testy Chi-kwadrat

	Wartość	df	Istotność asymptotyczna (dwustronna)
Chi-kwadrat Pearsona	84,976 ^a	9	,000
Iloraz wiarygodności	92,743	9	,000
Test związku liniowego	66,571	1	,000
N Ważnych obserwacji	136		

a. ,0% komórek (0) ma liczebność oczekiwaną mniejszą niż 5.
Minimalna liczebność oczekiwana wynosi 5,18.

Istotność = 0 < 0,05 \Rightarrow odrzucamy hipotezę zerową H_0 , czyli istnieje zależność pomiędzy ocenami z obu przedmiotów.

W zadaniu mamy dane porządkowe (poszczególne kategorie są uporządkowane). Do badania zależności pomiędzy tego rodzaju danymi wykorzystywane są testy:

- gamma,
- d Sommersa,
- tau-b Kendalla,
- tau-c Kendalla.

Miary kierunkowe

			Wartość	Asymptotyczny błąd standardowy ^a	Przybliżone T ^b	Istotność przybliżona
Porządkowa przez Porządkowa	d Somersa	Symetryczna	,616	,045	13,528	,000
		Zmienna zależna: Ocena z filozofii	,611	,045	13,528	,000
		Zmienna zależna: Ocena z etyki	,620	,044	13,528	,000

a. Nie zakładając hipotezy zerowej.

b. Użyto asymptotycznego błędu standardowego, przy założeniu hipotezy zerowej.

Miary symetryczne

		Wartość	Asymptotyczny błąd standardowy ^a	Przybliżone T ^b	Istotność przybliżona
Porządkowa przez Porządkowa	tau-b Kendalla	,616	,045	13,528	,000
	tau-c Kendalla	,611	,045	13,528	,000
	Gamma	,775	,047	13,528	,000
N Ważnych obserwacji		136			

a. Nie zakładając hipotezy zerowej.

b. Użyto asymptotycznego błędu standardowego, przy założeniu hipotezy zerowej.

Odp. Zależność pomiędzy ocenami z filozofii i etyki istnieje i jak wynika z wielkości miar zależności jest ona dodatnia i umiarkowana.

Zadanie 5

p_1 – prawdopodobieństwo popełnienia błędu przez sekretarkę dyrektora FirMy

p_2 - prawdopodobieństwo popełnienia błędu przez asystenta dyrektorki FacTorki

W celu sprawdzenia, czy obie osoby piszą tak samo bezbłędnie, weryfikujemy hipotezę:

$$\mathbf{H: } p_1 = p_2$$

$$\mathbf{K: } p_1 \neq p_2$$

za pomocą statystyki

$$T = (n_2 k_1 - n_1 k_2) \sqrt{\frac{n_1 + n_2}{n_1 n_2 (k_1 + k_2) (n_1 + n_2 - k_1 - k_2)}},$$

gdzie n_1 – liczność pierwszej próby ($n_1 = 1205$)

n_2 - liczność drugiej próby ($n_2 = 1506$)

k_1 – liczba elementów wyróżnionych w pierwszej próbie ($k_1 = 10$)

k_2 – liczba elementów wyróżnionych w drugiej próbie ($k_2 = 25$)

$|T| < z_\beta \Rightarrow$ nie ma podstaw do odrzucenia hipotezy zerowej H

$T < -z_\beta \Rightarrow$ odrzucamy hipotezę zerową H na korzyść hipotezy K: $p_1 < p_2$

$T > z_\beta \Rightarrow$ odrzucamy hipotezę zerową H na korzyść hipotezy K: $p_1 > p_2$,

gdzie z_β jest kwantylem rzędu β standardowego rozkładu normalnego

$$\alpha = 0,1 = 1 - \beta \Rightarrow \beta = 0,9 \Rightarrow z_\beta = \text{ROZKLAD.NORMALNY.ODW}(0,9;0;1) = 1,281551$$

$$T = 1,9026 > z_{0,9} \Rightarrow \text{odrzucaamy hipotezę zerową H na korzyść hipotezy K: } p_1 > p_2$$

Odp. Asystent dyrektorki FacTorki popełnia mniej błędów niż sekretarka dyrektora FirMy.

Zadanie 6

X – sprzedaż

Zmienna losowa X może przyjmować trzy wartości (trzy klasy): 1-klient kupi produkt na pewno, 2-klient kupi produkt z dużym prawdopodobieństwem, 3-klient nie kupi produktu.

Weryfikujemy hipotezę:

$$H: p_1 = 0,2, p_2 = 0,3, p_3 = 0,5$$

$$K: \neg H$$

za pomocą statystyki:

$$\chi^2 = \sum \frac{(n_i - np_i)^2}{np_i}$$

nr klasy i	n_i	p_i	$n p_i$
1	34	0,2	24,8
2	45	0,3	37,2
3	45	0,5	62,0
	$n = 124$	$p_1 + p_2 + p_3 = 1$	$np_1 + np_2 + np_3 = 124$

Po wykonaniu obliczeń, dostajemy $\chi^2 = 9,71$.

$$\chi^2_{1-\alpha, k-1} = \chi^2_{0,95, 2} = 5,99$$

Mamy więc $\chi^2 = 9,71 > 5,99 \Rightarrow$ odrzucamy hipotezę zerową H.

Odp. Wnioski analityków nie były słuszne.