



Ćwiczenie: Test χ^2 i miary na nim oparte.

Zadanie (MS EXCEL)

Czy istnieje zależność między płcią a paleniem papierosów?

1. W arkuszu Excel utworzyć dwie tabele

		Tabela		
		Liczebności obserwowane		
		płeć		
		kobiety	mężczyźni	suma
palenie	0			
	1			
suma				

		Tabela		
		Liczebności oczekiwane		
		płeć		
		kobiety	mężczyźni	suma
palenie	0			
	1			
suma				

2. Uzupełnić wartości w tabeli z danymi obserwowanymi

		Tabela		
		Liczebności obserwowane		
		płeć		
		kobiety	mężczyźni	suma
palenie	0	58	11	69
	1	13	18	31
suma		71	29	100

3. Na podstawie przedstawionych poniżej wzorów obliczyć liczebności oczekiwane

		Tabela		
		Liczebności obserwowane		
		płeć		
		kobiety	mężczyźni	suma
palenie	0	a	b	
	1	c	d	
suma				

$$f_a = \frac{(a+b)(a+c)}{n}$$

$$f_b = \frac{(a+b)(b+d)}{n}$$

		Tabela		
		Liczebności oczekiwane		
		płeć		
		kobiety	mężczyźni	suma
palenie	0	48.99	20.01	69
	1	22.01	8.99	31
suma		71	29	100

$$f_c = \frac{(c+d)(a+c)}{n}$$

$$f_d = \frac{(c+d)(b+d)}{n}$$

4. Za pomocą funkcji statystycznej sprawdzić wartość testu χ^2 . Podaj odpowiedź.



=TEST.CHI(R8:S9;R18:S19)

P	Q	R	S	T
p dla chi2=TEST.CH				
Tabela Liczebności obserwowane				
palenie		płeć		
		kobiety	mężczyźni	suma
	0	58	11	69
1	13	18	31	
suma		71	29	100
Tabela Liczebności oczekiwane				
palenie		płeć		
		kobiety	mężczyźni	suma
	0	48.99	20.01	69
1	22.01	8.99	31	
suma		71	29	100

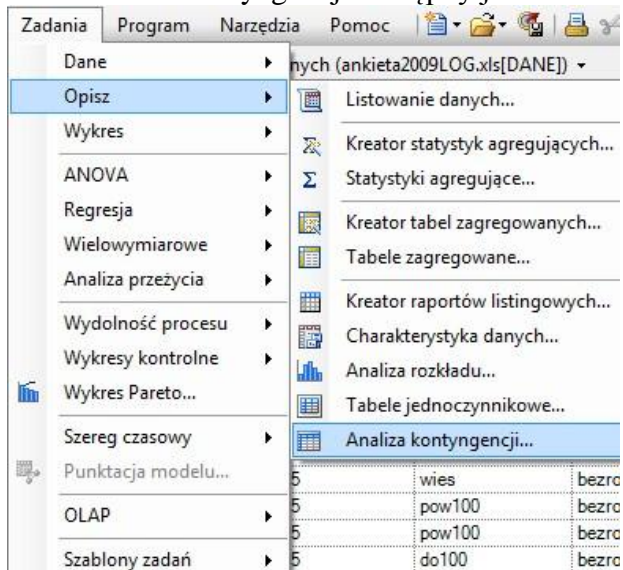
Przykład (SAS EG) – ankietaX.xls:

Sprawdź czy istnieje zależność między płcią a udzielonymi w ankiecie odpowiedziami (kwiek, papierosy, miejsce).

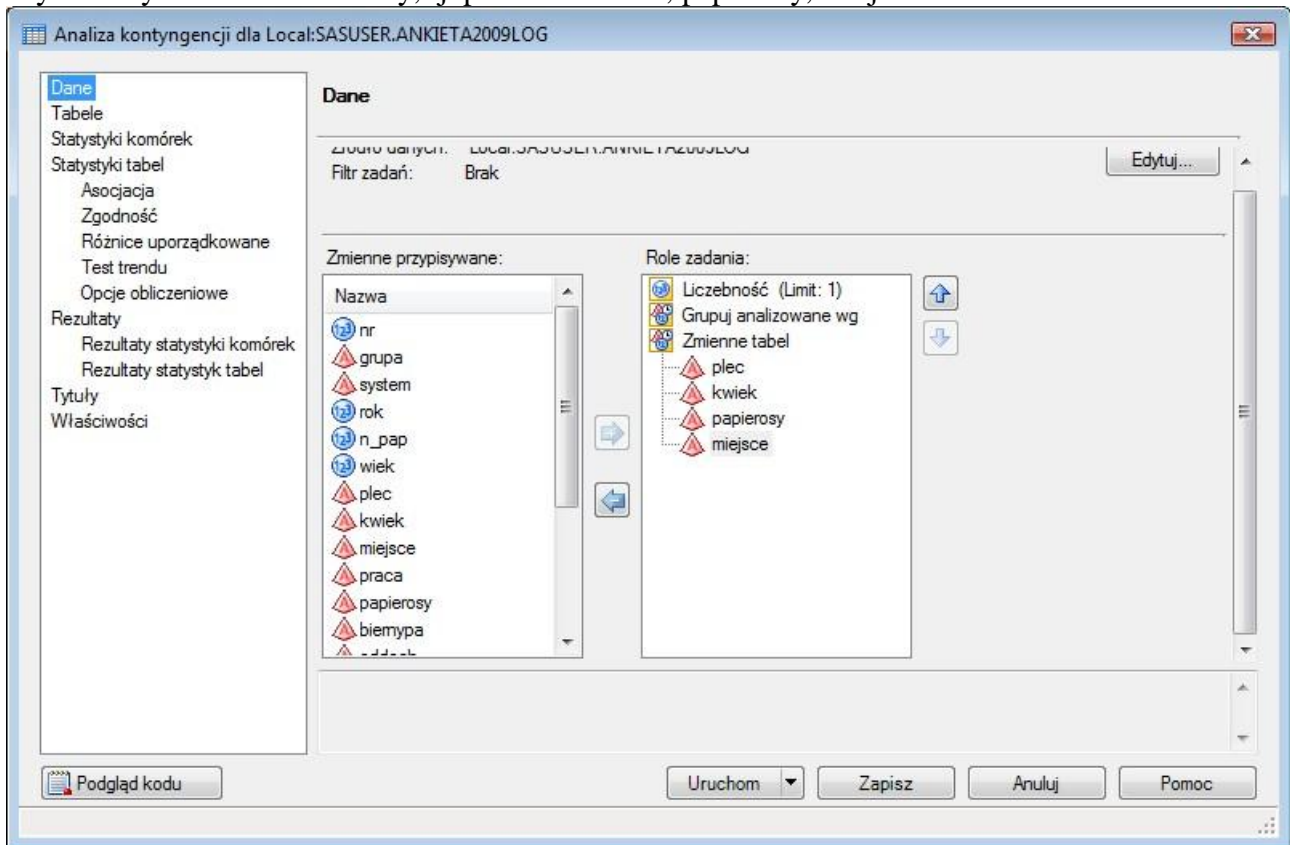
H_0 : nie istnieje zależność między płcią osób a treścią udzielanych odpowiedzi

H_1 : istnieje zależność między płcią osób a treścią udzielanych odpowiedzi

Kreator tabel kontyngencji dostępny jest w menu Zadania => Analiza kontyngencji

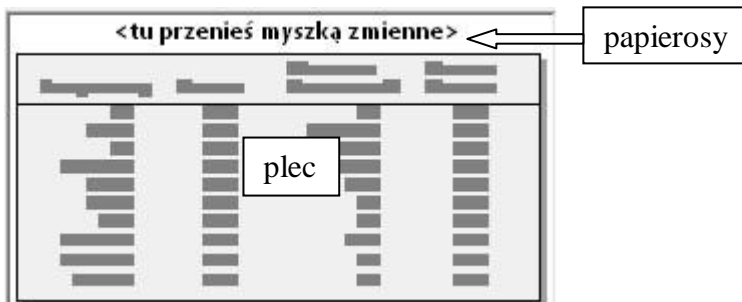


Wybieramy zmienne do analizy, tj. plec oraz kwiek, papierosy, miejsce.

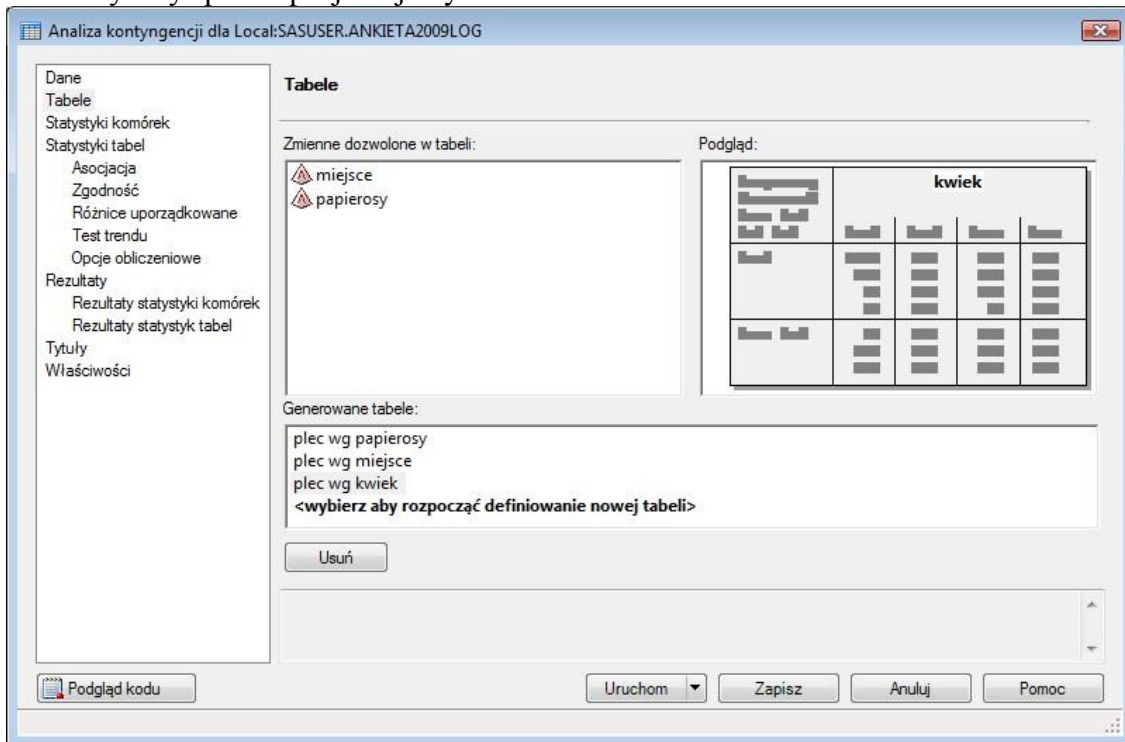


W kolejnym kroku projektujemy tabele wynikowe – 3 zestawienia: plec-papierosy, plec-miejsce, plec-kwiek. Tworzenie nowej tabeli uruchamiamy klikając na tekst **<wybierz aby rozpocząć tworzenie nowej tabeli>**. Zakładamy, że w ostatecznych tabelach poziomy płci umieszczone będą w kolejnych wierszach. Z kolei odpowiedzi dotyczące palenia papierosów, miejsca zamieszkania czy wieku będą stanowiły kolumny w tabeli.

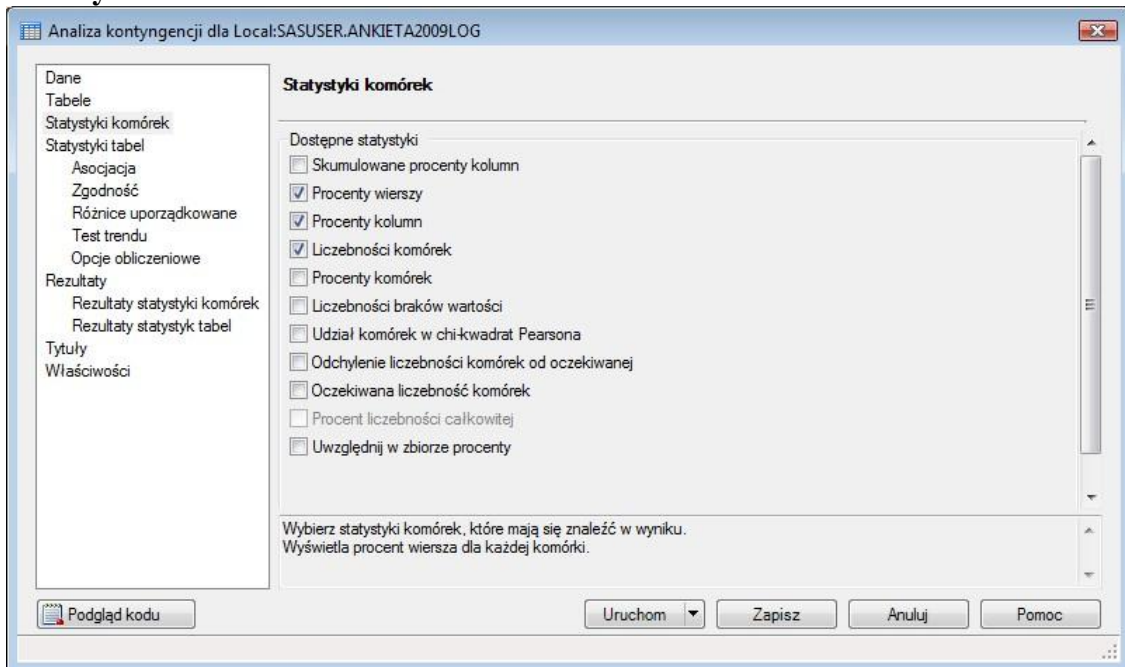
Założony cel osiągniemy przeciągając „papierosy” na pole <tu przenieś myszką zmienne>, zaś „plec” przeciągamy na środek umieszczonego ekranu.



W identyczny sposób projektujemy dalsze tabele.

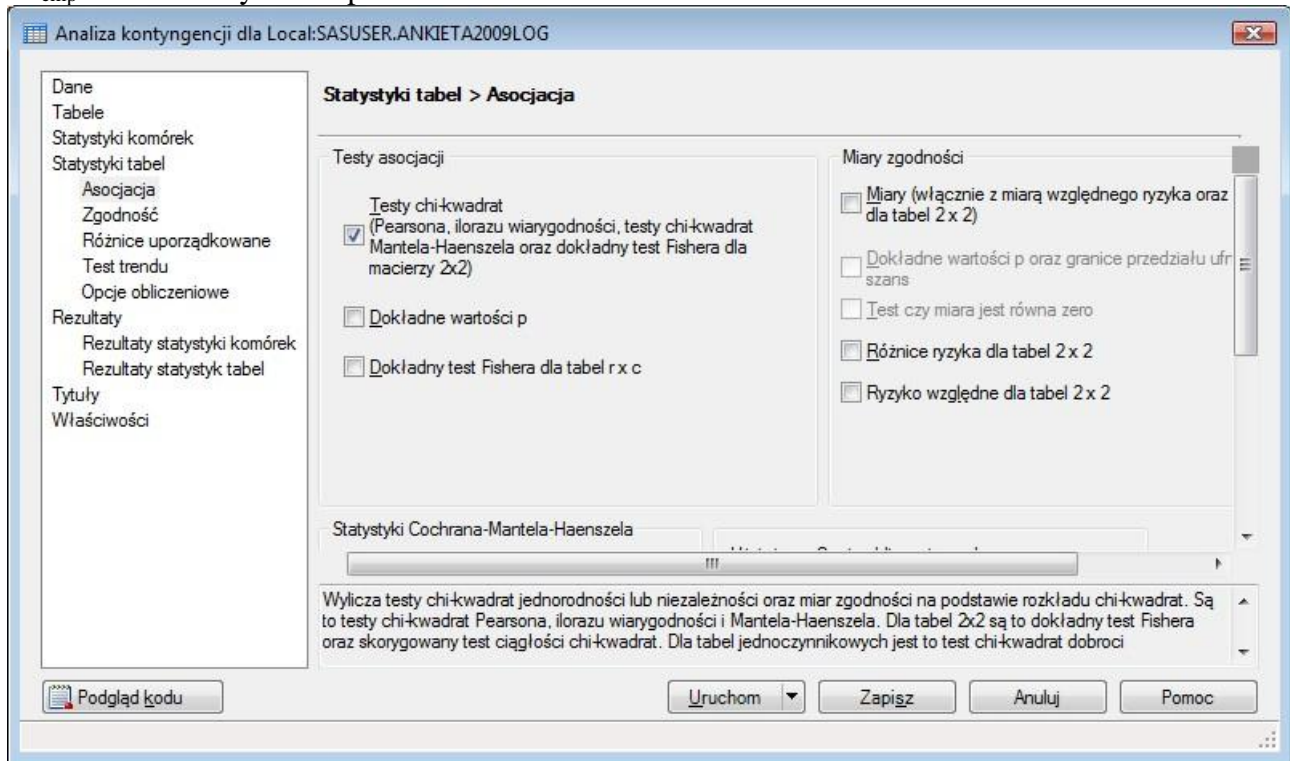


W zadaniach statystyki komórek zaznaczone powinny być wskazane poniżej pola: **Procenty wierszy i kolumn** oraz **Liczebności komórek**.





W celu zbadania zależności między zmiennymi plec a papierosy, miejsce, kwiek wykorzystamy X^2_{emp} . Zaznaczamy zatem pole: Test chi-kwadrat.



Klikając na **Uruchom** otrzymujemy niżej umieszczone wyniki:

Tabela plec wg papierosy				
		papierosy		Razem
		nie	tak	
plec	Liczebność	193	53	246
	Proc. wier.	78.46	21.54	
	Proc. kol.	69.93	72.60	
mezczyzn	Liczebność	83	20	103
	Proc. wier.	80.58	19.42	
	Proc. kol.	30.07	27.40	
Razem	Liczebność	276	73	349

Liczebność braków danych = 3

Statystyki dla tabeli plec wg papierosy

Statystyka	St. sw.	Wartość	Prawd.
Chi-kwadrat	1	0.1986	0.6558
Chi-kw. ilorazu wiarygodn.	1	0.2007	0.6541
Poprawka uciagl. chi-kwadrat	1	0.0908	0.7631
Chi-kwadrat Mantela-Haenszela	1	0.1980	0.6563
Współczynnik FI		-0.0239	
Współczynnik wielodzielczości		0.0238	
V Cramera		-0.0239	



Prawdopodobieństwo, jakie wynika z przeprowadzonego testu X^2 (0,6558) nie pozwala odrzucić hipotezy zerowej i stwierdzić, że istnieje zależność istotna między płcią osób a odpowiedzią na pytanie: czy palisz papierosy? Zestawione w tabeli wyniki pozwala stwierdzić, że wśród kobiet udział palących wyniósł 21,54% i był około 2 j. p. niższy niż wśród mężczyzn (19,428%).

Zadanie (SAS EG) – roztocze.xls

Poddaj ocenie zależność między rodzajem zakładu przemysłowego a wybranymi gatunkami mechowców:

C_sch_r	Chamobates schuetzi
E_obl_r	Eremaeus oblongus
E_tor_r	Eupelops torulosus

Wcześniej jednak dokonaj przekodowania powyższych kolumn, tzn. utwórz 3 nowe kolumny zawierające „1” i „0”. „1” wystąpi w sytuacji, gdy została zarejestrowana obecność w próbie danego gatunku mechowców, zaś „0”, gdy był on nieobecny.

Zadanie (SAS EG) – skora.xls

Sprawdź czy istnieje zależność między paleniem papierosów (palenie) a niektórymi dolegliwościami skórnymi oraz płcią osób.

Przykład (SAS EG) – ankieta2011Z.xls:

Sprawdź czy istnieje zależność między płcią a udzielonymi w ankiecie odpowiedziami (kwiek, papierosy, miejsce). Jest to ankieta wzbogacona o dane dotyczące studentów z obecnego semestru.