Rozwiązanie zadania 23 z zestawu nr 3 („na piechotę” i w R)

23. Badano liczby transakcji wsród odsłon aukcji w trzech różnych wersjach. Otrzymano następujące wyniki.

Wersja aukcji	Liczba odsłon	Liczba transakcji
I	100	15
II	120	12
III	115	17

Czy na podstawie tych danych można wnioskować, że wskaźniki konwersji dla tych aukcji są istotnie różne? Przyjmij poziom istotności 0.05.

---

p_i – wskaźnik konwersji dla i-tej wersji aukcji (liczba transakcji/liczba odsłon), i=1, 2, 3

$H_0: p_1 = p_2 = p_3$ (wskaźniki konwersji **nie różnią się istotnie**)

H_1: nie wszystkie p_i są sobie równe (wskaźniki konwersji **różnią się istotnie**)

Próby niezależne	$n_1 = 100$	$k_1 = 15$	$\hat{p}_1 = 0{,}150$
	$n_2 = 120$	$k_2 = 12$	$\hat{p}_2 = 0{,}100$
	$n_3 = 115$	$k_3 = 17$	$\hat{p}_3 = 0{,}148$

Na podstawie tych prób weryfikujemy powyższą hipotezę testem jednorodności chi-kwadrat.

Test jednorodności chi-kwadrat

Niech X_i oznacza cechę badaną w i-tej populacji, w tym zadaniu X_i oznacza wykonanie transakcji na i-tej wersji aukcji ($X_i = 1$ oznacza, że na i-tej wersji aukcji została wykonana transakcja, $X_i = 0$ – że nie została wykonana transakcja), i=1, 2, …, k. Zmienna X_i może przyjmować l wartości, w tym zadaniu przyjmuje dwie wartości ($l = 2$).

Uwaga: Hipoteza H_0 równoważna jest temu, że zmienne X_1, X_2, X_3 mają jednakowy rozkład.

Statystyka testowa: $\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{\left(n_{ij} - \frac{n_{i.} n_{.j}}{n}\right)^2}{\frac{n_{i.} n_{.j}}{n}}$

gdzie n_{ij} – liczba wystąpień wartości x_j w i-tej próbie pochodzącej z rozkładu zmiennej X_i,

$n_{i.} = \sum_{j=1}^{k} n_{ij}, n_{.j} = \sum_{i=1}^{l} n_{ij}, n = \sum_{j=1}^{k} \sum_{i=1}^{l} n_{ij}$

Statystyka testowa przy prawdziwości hipotezy H_0ma rozkład chi-kwadrat z (k-1)(l-1) stopniami swobody.

Obszar krytyczny: $W = [\chi^2_{1-\alpha,(k-1)(l-1)}, +\infty)$

Wyznaczamy wartość statystyki testowej dla danych z zadania 23:

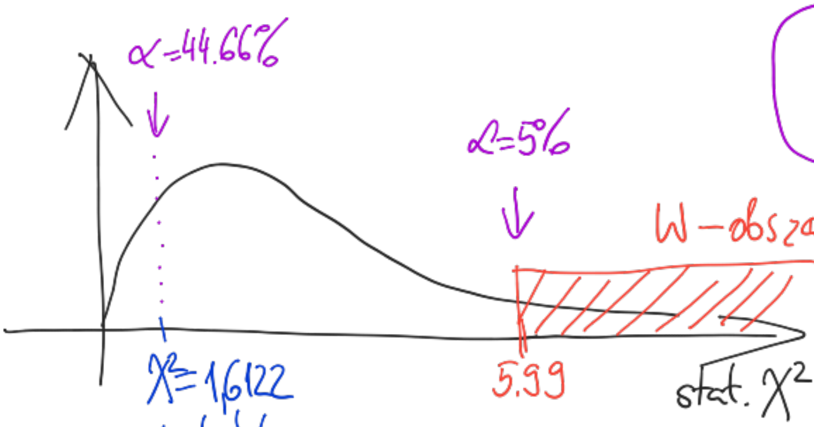
	$x_1 = 1$	$x_2 = 0$	
Próba 1	$n_{11} = 15$	$n_{12} = 85$	$n_{1.} = 100$
Próba 2	$n_{21} = 12$	$n_{22} = 108$	$n_{2.} = 120$
Próba 3	$n_{31} = 17$	$n_{32} = 98$	$n_{3.} = 115$
	$n_{.1} = 44$	$n_{.2} = 291$	$n = 335$

$$\chi^2 = \frac{\left(15 - \frac{100\cdot 44}{335}\right)^2}{\frac{100\cdot 44}{335}} + \ldots + \frac{\left(98 - \frac{115\cdot 291}{335}\right)^2}{\frac{115\cdot 291}{335}} = 1{,}6122$$

(w R z funkcji prop.test(c(15,12,17),c(100,120,115)) ten sam wynik)

$$W = \left[\chi^2_{1-0.05,\,(3-1)(2-1)}, +\infty\right) = \left[\chi^2_{0.95,2}, +\infty\right) = \left[5.99, +\infty\right)$$

uwaga: 1,6122 to jest wartość kwantyla $\left(\chi^2_{1-pv,2}\right)$, gdzie pv = p-value ≈ 0.4466

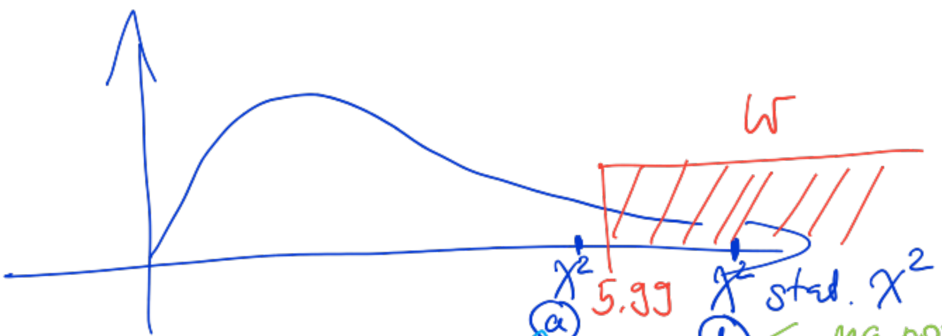


W – obszar krytyczny dla $\alpha = 0{,}05$ (obszar odrzuceń)

wartość stat. testowej $\notin W \Rightarrow$ brak podstaw do odrzucenia H_0, zatem stwierdzamy, że wskaźniki konwersji nie różnią się istotnie (tj. nie ma wystarczających podstaw, by pokazać, że się różnią).

A gdyby było tak....

(a)		(b)		
$k_1 = 15$	15%	$k_1 = 15$	15%	$n_1 = 100$
$k_2 = 30$	25%	$k_2 = 35$	29%	$n_2 = 120$
$k_3 = 17$	15%	$k_3 = 17$	15%	$n_3 = 115$
$\chi^2 = 5.2275$		$\chi^2 = 9.8211$		



(a) na poziomie istotności $\alpha = 0.05$ brak podstaw do odrzucenia H_0

(b) na poziomie istotności $\alpha = 0{,}05$ są podstawy do odrzucenia H_0