

Analiza danych jakościowych

Lista 6

1. **Dokładny test Fishera.** Na n karteczkach umieszczamy losowo i *niezależnie* dwa rodzaje symboli: najpierw na n_{1+} karteczkach umieszczamy losowo kółka (O) i na pozostałych n_{2+} karteczkach ($n_{1+}+n_{2+} = n$) umieszczamy trójkąty (∇). Potem, niezależnie od umieszczania tych symboli na n_{+1} karteczkach umieszczamy losowo kropki ($\bullet$) i na pozostałych n_{+2} karteczkach ($n_{+1}+n_{+2} = n$) umieszczamy serduszka ($\heartsuit$). Pokaż, że prawdopodobieństwo, iż na n_{11} karteczkach znajdują się jednocześnie symbole O i $\bullet$ wynosi:

$$\frac{\binom{n_{1+}}{n_{11}} \binom{n_{2+}}{n_{+1}-n_{11}}}{\binom{n}{n_{+1}}} \quad (1)$$

Wzór ten można wyrazić inaczej: prawdopodobieństwo otrzymania tablicy:

		Y		
		$\bullet$	$\heartsuit$	
X	O	n_{11}	n_{12}	n_{1+}
	∇	n_{21}	n_{22}	n_{2+}
		n_{+1}	n_{+2}	n

gdy zmienne X i Y są niezależne wynosi

$$\frac{n_{1+}!n_{2+}!n_{+1}!n_{+2}!}{n_{11}!n_{12}!n_{21}!n_{22}!n!} \quad (2)$$

Pokaż, że wzory (1) i (2) są równoważne.

Zastanów się, jak sformułować to zadanie dla przypadku *jednorodności* zmiennej Y w grupach oznaczonych O i ∇ . Pokaż, że dla tego przypadku wzory (1) i (2) obowiązują.

2. W populacji prawdopodobieństwa wystąpienia wartości zmiennych X i Y są następujące:

		Y		
		$\bullet$	$\heartsuit$	
X	O	p_{11}	p_{12}	p_{1+}
	∇	p_{21}	p_{22}	p_{2+}
		p_{+1}	p_{+2}	1

Rozważ wszystkie tablice $T(u)$ takie, że sumy licznosci w wierszach i sumy licznosci w kolumnach są stałe, a w lewym górnym rogu tabeli licznosc wynosi u . W n -elementowej próbie zaobserwowano występowanie tych symboli o licznosciach, opisanych w tabeli $T(n_{11})$.

Zastępując prawdopodobieństwa przez ich estymatory, pokaż, że $\hat{\theta} = 1 \Leftrightarrow n_{11} = \hat{n}_{11}$,

$$\hat{\theta} > 1 \Leftrightarrow n_{11} > \hat{n}_{11}, \quad \hat{\theta} < 1 \Leftrightarrow n_{11} < \hat{n}_{11}, \quad \text{gdzie } \hat{n}_{11} = \frac{n_{1+}n_{+1}}{n}.$$

Wskazówka. Skorzystaj z zadania 3.1

3. Niech $\hat{\theta}(u)$ będzie ilorazem krzyżowym dla macierzy $T(u)$. Pokaż, że jeśli $u \geq u^*$ oraz $\hat{\theta}(u^*) > 1$ to $\hat{\theta}(u) \geq \hat{\theta}(u^*) > 1$. Niech n_{ij}^* są elementami macierzy $T(u^*)$, a n_{ij} elementami

$$\text{macierzy } T(u). \text{ Wtedy } \frac{n_{11}}{n_{1+}} - \frac{n_{21}}{n_{2+}} \geq \frac{n_{11}^*}{n_{1+}^*} - \frac{n_{21}^*}{n_{2+}^*}$$

$$\text{Wskazówka. Pokaż, że } (\hat{\theta}(u) - 1)n_{12}n_{21} \geq (\hat{\theta}(u^*) - 1)n_{12}^*n_{21}^*$$

4. Sformułuj i udowodnij analogiczny wynik dla przypadku $\hat{\theta}(u^*) < 1$