Econ 4130

HG Nov. 2017

Job satisfaction example (A. Agresti, "Categorical Data Analysis", Wiley 1990, p. 50)

	$O_{ij} = X_{ij}$ = Obs(ij)	C (Job satisfaction) Very dissatisfied	Little dissatisfied	Moderately satisfied	Very satisfied	Total
R (Income)	<6000	20	24	80	82	206
	6000-15000	22	38	104	125	289
	15000-25000	13	28	81	113	235
	>25000	7	18	54	92	171
	Total	62	108	319	412	n = 901

$$H_0 : R \text{ and } C \text{ independent} \quad \Leftrightarrow \quad H_0 : P(R = i \cap C = j) = P(R = i) \cdot P(C = j) \text{ for all } (i, j) \quad \Leftrightarrow$$

$$\Leftrightarrow \quad H_0 : p_{ij} = p_{i+} \cdot p_{+j} \quad \text{for all } (i, j)$$

$$E_{ij} = \text{ expected frequency under } H_0 \text{ (estimated)} = \frac{X_{i+} \cdot X_{+j}}{n} \qquad (n = 901)$$

E_{ij} =Expected frequency under H_0	Very dissatisfied	Little dissatisfied	Moderately satisfied	Very satisfied	Total
<6000	14.2	24.7	72.9	94.2	206
6000-15000	19.9	34.6	102.3	132.2	289
15000-25000	16.2	28.2	83.2	107.5	235
>25000	11.8	20.5	60.5	78.2	171
Total	62	108	319	412	901

$$Q_{ij} = \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \qquad Q = \sum_{\text{all } i,j} Q_{ij}, \qquad df = (r-1)(c-1) = (4-1)(4-1) = 9$$

Q_{ij}	Very dissatisfied	Little dissatisfied	Moderately satisfied	Very satisfied	sum
<6000	2.4	0.0	0.7	1.6	4.7
6000-15000	0.2	0.3	0.0	0.4	1.0
15000-25000	0.6	0.0	0.1	0.3	1.0
>25000	1.9	0.3	0.7	2.4	5.4

Q = sum Q_{ij} 12.0

$$P - \text{value} = P_{H_0}(Q > 12.0) \approx P_{\chi^2_9}(Q > 12) = 0.214$$