

(4)  $\chi^2$ -test for independence

100 students from major mathematics of three Viennese universities were randomly chosen and asked which lecture, either a: calculus, b: algebra, or c: probability, they enjoyed most. The frequencies are given in the following table:

	Uni A	Uni B	Uni C	
calculus	10	5	5	20
algebra	10	20	10	40
probability	20	5	15	40
	40	30	30	100

Perform a  $\chi^2$ -test to test whether the preference for a lecture is independent from the university, on a 5% significance level.

- (a) Only use the following table which gives the 95%-quantile  $q$  of the  $\chi^2$ -distribution with  $df$  degrees of freedom.

$df$	1	2	3	4	5	6	7	8	9
$q$	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92

- (b) Solve the previous exercise using R.

Let us choose  $r = 15$

a) we define the Matrix  $O := \begin{pmatrix} 10 & 5 & 5 \\ 10 & 20 & 10 \\ 20 & 5 & 15 \end{pmatrix}$

The total number of students that prefer calculus is 20 and which is  $\frac{20}{100}$  of all students. Since 40 of these students study at Uni A we expect  $e_{11} := 40 \cdot \frac{20}{100}$  of the students of Uni A to prefer calculus.

Similarly, we define  $e_{ij} := \sum_{k=1}^3 o_{kj} \cdot \frac{1}{100} \sum_{l=1}^3 o_{il}$  and obtain

$$E = \frac{1}{100} \begin{pmatrix} 20 \cdot 40 & 20 \cdot 30 & 20 \cdot (5+15) \\ 40 \cdot 20 & 40 \cdot 30 & 40 \cdot (5+15) \\ (25+15) \cdot 40 & (25+15) \cdot 30 & (25+15) \cdot (5+15) \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 40 & 30 & 30 \\ 80 & 60 & 60 \\ 80 & 60 & 60 \end{pmatrix} = \begin{pmatrix} 8 & 6 & 6 \\ 16 & 12 & 12 \\ 16 & 12 & 12 \end{pmatrix}$$

$$\sum_{i=1}^3 \sum_{j=1}^3 \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \approx 14.58$$

This sum is approximately  $\chi^2(m)$  distributed, where  $m = 9 - 1 = 8$

The 95% quantile of the  $\chi^2(8)$  distribution is 15.51 which is larger than 14.58, hence we fail to reject the hypothesis, that the preference does not depend on the university.

b)  $p\text{-value} = 0.068$