

19.1 Since there are two categorical variables, we will go ahead with chi-square distribution.

$H_0$ : There is no relationship between gender and level of education

$H_a$ : ~~There~~ There is a relationship between gender and level of education.

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

$$\text{Expected} = \frac{\text{Row totals} \times \text{column totals}}{N}$$

By applying above formula, below is the table of expected counts

	Highschool	Bachelors	Masters	ph.d	Total
male	50.886	49.868	50.377	49.868	201
female	49.114	48.132	48.623	48.132	194
total	100	98	99	98	395

$$\begin{aligned} \text{So now } \chi^2 &= \frac{(60 - 50.886)^2}{50.886} + \frac{(40 - 49.114)^2}{49.114} + \frac{(54 - 49.868)^2}{49.868} \\ &+ \frac{(44 - 48.132)^2}{48.132} + \frac{(46 - 50.377)^2}{50.377} + \frac{(53 - 48.623)^2}{48.623} + \frac{(41 - 49.868)^2}{49.868} \\ &+ \frac{(57 - 48.132)^2}{48.132} \end{aligned}$$

$$= 8.006$$

Degrees of freedom = (No. of classes - 1) = 4 - 1 = 3

critical value with  $\chi^2 = 8.006$  &  $df = 3$  and  $\alpha = 0.05$

7.815. Now that  $\chi^2$  critical is <sup>less</sup> ~~greater~~ than  $\chi^2$ , we

will reject  $H_0$ . So the conclusion is

There is relationship between gender and level of education.