Assignment 18.1 : Problem Statement

Is gender independent of education level? A random sample of 395 people were surveyed and each person was asked to report the highest education level they obtained. The data that resulted from the survey is summarized in the following table:

	High School	Bachelors	Masters	Ph.d.	Total
Female	60	54	46	41	201
Male	40	44	53	57	194
Total	100	98	99	98	395

Question: Are gender and education level dependent at 5% level of significance? In other words, given the data collected above, is there a relationship between the gender of an individual and the level of education that they have obtained?

Solution :

Chi-Square Test Statistic : $\chi^2 = \sum(O-E)^2 / E$

where O represents the observed frequency.
E is the expected frequency under the null hypothesis and computed as follows :

E= row total × column total / sample size

E(Female , High School) = 201* 100 /395 = 50.886 , Similarly computing for rest as in below table

Here's the table of expected counts:

	High School	Bachelors	Masters	Ph.d.	Total
Female	50.886	49.868	50.377	49.868	201
Male	49.114	48.132	48.623	48.132	194
Total	100	98	99	98	395

We will compare the value of the test statistic to the critical value of χ^2_α with degree of freedom = (r - 1) (c - 1), and reject the null hypothesis if $\chi^2 > \chi^2_\alpha$

So, working this out, $\chi 2=[(60-50.886)2/ 50.886]+\cdots+](57-48.132)2/ 48.132]= 8.006$

The critical value of $\chi 2\alpha$ with 3 degree of freedom is 7.815. Since 8.006 > 7.815, therefore we reject the null hypothesis and conclude that the education level depends on gender at a 5% level of significance