

Here is one problem from simple R:

In an effort to increase student retention, many colleges have tried block programs. Suppose 100 students are broken into two groups of 50 at random. One half are in a block program, the other half not. The number of years in attendance is then measured. We wish to test if the block program makes a difference in retention.

The data is:

Program 1 yr 2 yr. 3 yr 4yr 5+ yrs.

Non-Block 18 15 5 8 4

Block 10 5 7 18 10

Do a test of hypothesis to decide if there is a difference between the two types of programs in terms of retention.

Answer: Here we do chi-square test to see homogeneity for these two group.

```
> Non_Block<- c(18,15,5,8,4)
> Block<- c(10,5,7,18,10)
< : : : : :
> chisq.test(rbind(Non_Block,Block))

      Pearson's Chi-squared test

data:  rbind(Non_Block, Block)
X-squared = 14.037, df = 4, p-value = 0.007179
```

We see P-value = 0.007179 which is less than 0.05, so we reject hypothesis which two groups have same distribution, in this case, we conclude that in 95% confidence level we think block program have increased the retention.