Testing Independence in a 2-d table with R

```r
> math =
read.table("http://www.utstat.toronto.edu/~brunner/312f12/code_n_data/mathcat.data")
> math[1:5,]
  hsgpa hsengl hscalc  course passed outcome
1  78.0    80  Yes Mainstrm  No  Failed
2  66.0    75  Yes Mainstrm  Yes  Passed
3  80.2    70  Yes Mainstrm  Yes  Passed
4  81.7    67  Yes Mainstrm  Yes  Passed
5  86.8    80  Yes Mainstrm  Yes  Passed
> attach(math) # Variable names are now global
> calcpass = table(hscalc,passed); calcpass
  passed
  hscalc  No Yes
    No   17  4
    Yes 141 232
> help(chisq.test)
starting httpd help server ... done
> chisq.test(calcpass,correct=F) # Don't correct for continuity

  Pearson's Chi-squared test

data:  calcpass
  X-squared = 15.4111, df = 1, p-value = 8.648e-05

> test1 = chisq.test(calcpass,correct=F) # Save the test object

> test1$expected
  passed
  hscalc       No   Yes
    No 8.42132 12.57868
    Yes 149.57868 223.42132
> n = sum(calcpass); (17+4)*(17+141)/n # Checking expected freq
[1] 8.42132
> test1$stdres # Standardized residuals
  passed
  hscalc       No   Yes
    No 3.925692 -3.925692
    Yes -3.925692  3.925692
```
# Likelihood ratio test uses the same expected frequencies
freq = calcpass; muhat = test1$expected
G2 = 2*sum(freq*log(freq/muhat)); G2  # Compare X^2 = 15.41
[1] 15.54447
pval = 1-pchisq(G2,1); pval
[1] 8.058708e-05

courseout = table(course,outcome); courseout

<table>
<thead>
<tr>
<th>course</th>
<th>Disappeared</th>
<th>Failed</th>
<th>Passed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch-up</td>
<td>21</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Elite</td>
<td>2</td>
<td>5</td>
<td>24</td>
</tr>
<tr>
<td>Mainstrm</td>
<td>74</td>
<td>50</td>
<td>204</td>
</tr>
</tbody>
</table>

test2 = chisq.test(courseout)
Warning message:
In chisq.test(courseout, correct = F):
  Chi-squared approximation may be incorrect
test2

Pearson's Chi-squared test
data:  courseout
X-squared = 32.4943, df = 4, p-value = 1.516e-06

test2$expected

<table>
<thead>
<tr>
<th>course</th>
<th>Disappeared</th>
<th>Failed</th>
<th>Passed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch-up</td>
<td>8.616751</td>
<td>5.418782</td>
<td>20.96447</td>
</tr>
<tr>
<td>Elite</td>
<td>7.631980</td>
<td>4.799492</td>
<td>18.56853</td>
</tr>
<tr>
<td>Mainstrm</td>
<td>80.751269</td>
<td>50.781726</td>
<td>196.46701</td>
</tr>
</tbody>
</table>

test2$stdres

<table>
<thead>
<tr>
<th>course</th>
<th>Disappeared</th>
<th>Failed</th>
<th>Passed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch-up</td>
<td>5.0901853</td>
<td>0.2845216</td>
<td>-4.6841760</td>
</tr>
<tr>
<td>Elite</td>
<td>-2.4462856</td>
<td>0.1037180</td>
<td>2.0736869</td>
</tr>
<tr>
<td>Mainstrm</td>
<td>-2.1142521</td>
<td>-0.2915432</td>
<td>2.0735694</td>
</tr>
</tbody>
</table>

Compared to students in the Elite and Mainstream courses, students in the Catch-up course were more likely to disappear and less likely to pass.