Contingency Tables

Much of this should be a straight forward review. Let's just do a quick example:

Seat belts in Florida (data from 1988):

<table>
<thead>
<tr>
<th>Safety equipment</th>
<th>Injury</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fatal</td>
<td>Non-fatal</td>
<td>Total</td>
</tr>
<tr>
<td>None</td>
<td>1,601</td>
<td>165,527</td>
<td>167,128</td>
</tr>
<tr>
<td>Seat belt</td>
<td>510</td>
<td>412,368</td>
<td>412,878</td>
</tr>
</tbody>
</table>

H₀: \( p₁ = p₂ \)

\( p₁ \) = proportion of fatalities for people not wearing a seat belt (estimated by \( \hat{p}_₁ \)).

\( p₂ \) = proportion of fatalities for people wearing a seat belt (estimated by \( \hat{p}_₂ \)).

H₁: \( p₁ > p₂ \)

We calculate the expected values:

Remember:

\[
\text{Expected value} = \frac{(\text{Row total}) \times (\text{Column total})}{(\text{Grand total})}
\]

<table>
<thead>
<tr>
<th>Safety equipment</th>
<th>Injury</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fatal</td>
<td>Non-fatal</td>
<td>Total</td>
</tr>
<tr>
<td>None</td>
<td>608.28</td>
<td>166,519.72</td>
<td>167,128</td>
</tr>
<tr>
<td>Seat belt</td>
<td>1,502.72</td>
<td>411,375.28</td>
<td>412,878</td>
</tr>
<tr>
<td>Total</td>
<td>2,111</td>
<td>577,895</td>
<td>580,006</td>
</tr>
</tbody>
</table>

And our we calculate our \( \chi^2 \) using:

\[
\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}
\]
\[
\chi^2* = \frac{(1601-608.28)^2}{608.28} + \frac{(165,527-166,519.72)^2}{166,519.72} + \\
\frac{(510-1502)^2}{1502.72} + \frac{(412,368-411,375)^2}{411,375.28} = 2284.25
\]

If we look up our critical value of \(\chi^2\) in the table (1 d.f., and \(\alpha = .05\)), we get:

\[
\chi^2_{\text{table}} = \chi^2_{.05,1} = 2.71
\]

So we reject \(H_0\) and conclude that seat belts do affect the outcome of a traffic accident.

Somewhere we should have made sure the data deviate from the null hypothesis in the direction of your alternative (or we would have stopped).

In other words, we should have made sure that \(\hat{p}_1 > \hat{p}_2\).

We note that \(\hat{p}_1 = 0.00958, \hat{p}_2 = 0.00124\), so we did okay.

Some comments:

The chi-square contingency test can actually be used to test two different hypotheses. We either:

1) Compare proportions (like in the seatbelt example)

- or -

2) Establish independence/dependence

And we would phrase our hypotheses accordingly:

1) \(H_0: p_1 = p_2\), where the \(p\)'s are proportions,  
   \(H_1: p_1 \neq p_2\) (or \(p_1 > p_2\), etc.)

2) \(H_0\): Factor 1 and factor 2 are independent  
   \(H_1\): Factor 1 and factor 2 are dependent

The math and our decision rule is identical(!)

\(R \times K\) tables work the same way (no changes to the math).
So what about the assumptions?

i) random data

ii) smallest expected value $\geq 5$

(There is some debate as to how serious this restriction is, but we'll stick with it).

So what can you do if you violate the assumptions?

**Fisher's exact test:**

Let's start with an example that doesn’t come from biology, but a little history of statistics doesn’t hurt (this is a famous example from statistics).

Way back when, Fisher came across a lady who claimed to be able to tell if tea or milk were poured into a cup first (the British like to drink milk with their tea). So he put her to the test with eight cups, four of which had milk added first, the other four tea first. He got the following results:

<table>
<thead>
<tr>
<th>Lady’s guess</th>
<th>Milk</th>
<th>Tea</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poured first</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Tea</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

Looking at the results, one can see that the lady seems to have done a little better than “random guessing”, but is this result explainable by chance?

Can we use a chi-squared statistic? NO! (all categories have expected values < 5)

But we can generate exact probabilities!

But before we do:

H$_0$: The lady can not tell the difference between pouring tea or milk first.

H$_1$: The lady can tell if milk or tea is poured first (note that this is actually a directional hypothesis)

$\alpha = .05$
What we do is calculate the probability of having gotten that result or “worse” if our null hypothesis is that the lady can’t tell the difference. Our marginal totals are fixed for this test (they can't change - more on that later).

What is “worse”?

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 4 | 0 | 0 | 4 | (the lady did perfectly)

So now we just need to calculate the probability of getting this and the actual result we got.

This uses something called the “hypergeometric distribution”, which is a bit similar to the binomial. Here’s how you make it work:

The number of ways the lady could guess milk first in 3 out of 4 cups with milk poured first is:

\[
\binom{4}{3} = \frac{4!}{3!1!} = 4
\]

The number of ways the lady could guess milk first in 1 out of 4 cups with tea poured first is:

\[
\binom{4}{1} = \frac{4!}{1!3!} = 4
\]

And the number of ways she can guess milk first in four out of eight cups:

\[
\binom{8}{4} = \frac{8!}{4!4!} = 70
\]

Notice that what we’re doing is calculating the probability of getting the result she got and then dividing by the total number of possible outcomes one can get.

So now we just multiply the first two quantities and divide by the last:

\[4 \times 4 = 16 \quad \frac{16}{70} = 0.229\]

Pr{of getting the above result} = 0.229
Now we repeat all this for the table with the “worse result”:

\[ \text{Pr\{of getting even worse result\}} = 0.014 \]

add these up and we get \( 0.229 + 0.014 = 0.243 \) (\( = p \)).

So the probability of the lady guessing is 0.243, or a little less than \( \frac{1}{4} \). We conclude that we do not have any evidence that she can tell the difference (our \( p > \alpha \)).

Some more details:

Here’s a formula (for 2 x 2 tables):

\[
\text{Probability of getting result in column 1 total} = \frac{\text{# of ways of getting result in first cell} \times \text{# of ways of getting result in second cell (row wise)}}{\text{# of ways of getting result in first cell}} \times \frac{\text{# of ways of getting result in second cell (row wise)}}{\text{# of ways of getting result in column totals}}
\]

where:

\[
\text{# of ways of getting result in first cell} = \text{Binom (row 1 total, first cell count)}
\]

\[
\text{# of ways of getting result in second cell} = \text{Binom (row 2 total, second cell count)}
\]

\[
\text{# of ways of getting result in column totals} = \text{Binom(grand total, column 1 totals)}
\]

(you’re basically going row by row, using the first column and the total column to get your binomial coefficients).

As you can see, this gets messy real fast. Not only do you have to calculate the probability of getting the table you got (the above formula), but also the probability of getting each worse table.

Fisher’s exact test is almost always done on a computer, particularly for 3 x 3 tables and higher (the formula above will need to be modified).

You should know how to do a very simple test, sort of like the Tea-drinker example. See also example 24.15, p. 544 in our text (Fisher's exact test is also described quite well in the 214 text).
As presented, Fisher’s test is actually directional. One can use Fisher’s test in a non-directional setting, but then we’d have to calculate all the tables in the “opposite direction” of our result or worse. For our tea drinker, we’d have to include tables:

\[
\begin{array}{cc}
1 & 3 \\
3 & 1
\end{array}
\quad \text{and} \quad
\begin{array}{cc}
0 & 4 \\
4 & 0
\end{array}
\]

(add up the probabilities of these to our total)

See example 24.19 on page 550 of our text for an example.

Not all statistical packages will not perform Fisher’s exact test (but many other packages (e.g. R) do). Incidentally, any problems you get will be very simple.

Let’s do one more example (since this is sort of confusing):

An exercise from the 214 text (fictional data):

<table>
<thead>
<tr>
<th>Treatment</th>
<th>A</th>
<th>B</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Die</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Live</td>
<td>10</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

H₀: Treatment A and B are the same
H₁: Treatment B is better (directional)

\[ \alpha = .05 \]

Which tables are worse than our result? First, remember that our marginal totals are fixed. Why? because we picked 14 people for treatment A and 16 people for group B. Also, our H₀ implies that the proportions in column a and b are the same, so that our row totals can’t change (if there is no effect, then the probability of dying is the same regardless of column, and we can use the row total to estimate that => this means that row totals do not change.

So here are the tables that would be worse than what we got.

\[
\begin{array}{ccc}
5 & 1 & 6 \\
9 & 15 & 8
\end{array}
\]
So now all we have to do (ha, ha) is figure out the probability of getting these three tables and add these probabilities up.

Probability of getting the result we got:

\[
\frac{\binom{6}{4} \frac{24}{10}}{\binom{30}{14}} = 0.20229
\]

(we could stop here - why??)

Probability of getting the two worse tables:

\[
\frac{\binom{6}{5} \frac{24}{9}}{\binom{30}{14}} = 0.0539 \quad \text{and} \quad \frac{\binom{6}{6} \frac{24}{8}}{\binom{30}{14}} = 0.00506
\]

Adding up all these probabilities, we get:

\[
0.20229 + 0.0539 + 0.00506 = 0.26129
\]

Since our \( p \)-value is > \( \alpha \), we conclude that we have no evidence that treatment B is better than treatment A and the observed result is due to chance.