Is there a relationship between handedness and gender? An investigator gathered the following data from a sample of 100 people.

| | Left-handed | Right-handed |
|---|---|---|
| Female | 3 | 47 |
| Male | 4 | 46 |

Test the hypothesis of no relationship between handedness and gender using:

1. Fisher's exact test;
2. The chi-square test of independence;
3. A likelihood ratio test.

In each case, compute a test statistic and indicate what calculation is needed to find a p-value.

Solution:

1. **Fisher's exact test.** The expected totals in each cell under the hypothesis of independence are

| | Left-handed | Right-handed |
|---|---|---|
| Female | 3.5 | 46.5 |
| Male | 3.5 | 46.5 |

Focussing on the upper left cell, the absolute difference between the observed and expected counts is 0.5. We can think of the test statistic as being this absolute difference 0.5. The p-value is the probability of observing an absolute difference of 0.5 or greater under a hypergeometric probability distribution with the given fixed marginal totals. Since the absolute difference will be at least 0.5 for *any* possible table under this distribution, the p-value is one. Formally,

$$p = \sum_{k:|k-3.5|\geq 0.5} \frac{\binom{50}{k}\binom{50}{7-k}}{\binom{100}{7}} = \sum_{k=0}^{7} \frac{\binom{50}{k}\binom{50}{7-k}}{\binom{100}{7}} = 1$$

2. **Chi-square test of indepedence.** Let $\hat{\mu}_{ij}$ be the expected count in cell $ij$ under the model of independence and let $f_{ij}$ be the observed count. Then, the chi-square test statistic is

$$\begin{aligned} X^2 &= \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(f_{ij} - \hat{\mu_{ij}})^2}{\hat{\mu_{ij}}} \\ &= \frac{(3-3.5)^2}{3.5} + \frac{(4-3.5)^2}{3.5} + \frac{(46-46.5)^2}{46.5} + \frac{(47-46.5)^2}{46.5} \\ &\doteq 0.1536 \end{aligned}$$

The p-value is then the area to the right of 0.1536 under a chi-square density with one degree of freedom, which is 0.6951.

An alternative way to compute the p-value here would be to simulate many tables from the likelihood model maximized under the null hypothesis of independence, compute the test statistic for each, and see what proportion of the test statistics equal or exceed that from the original data. Code for such a simulation is in the file `practiceCategory.R`, and the result is 0.7089, which is very similar (the simulation standard error is about 0.005).

3. **Likelihood ratio test.** The probability of being in cell $ij$ is $\theta_{ij}$. The estimates under the full model are

| | Left-handed | Right-handed |
|---|---|---|
| Female | 0.03 | 0.47 |
| Male | 0.04 | 0.46 |

and under the null model are

| | Left-handed | Right-handed |
|---|---|---|
| Female | 0.035 | 0.465 |
| Male | 0.035 | 0.465 |

The full model estimates are the observed counts divided by $n$, the null model estimates are the expected counts divided by $n$. The likelihood ratio test statistic, $-2\log\Lambda$ is

$$\begin{aligned} -2\log\Lambda &= 2(3\log(0.03) + 4\log(0.04) + 47\log(0.47) + 46\log(0.46) \\ &\quad -3\log(0.035) + 4\log(0.035) + 47\log(0.465) + 46\log(0.465) \\ &\doteq 0.1541 \end{aligned}$$

The p-value from the chi-square approximation is 0.6946.