2.6 Exact Inference for Small Samples

2.6.1 Fisher’s Exact Test

Hypergeometric Distribution

Suppose \( m \) balls are drawn at random without replacement from a box containing \( R \) red balls and \( B \) blue balls. The number of red balls \( X \) in the \( m \) draws has a hypergeometric distribution:

\[
P(X = i) = \frac{\binom{R}{i} \binom{B}{m-i}}{\binom{R+B}{m}}
\]

Here \( \binom{a}{b} = \frac{a!}{b!(a-b)!} \), and \( 0 \leq i \leq R, 0 \leq m-i \leq B \).

The outcome of the draws can be displayed in a 2 \( \times \) 2 table:

<table>
<thead>
<tr>
<th>Color</th>
<th>Drawn</th>
<th>Not Drawn</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>( X )</td>
<td>( R - X )</td>
<td>( R )</td>
</tr>
<tr>
<td>Blue</td>
<td>( m - X )</td>
<td>( B - (m - X) )</td>
<td>( B )</td>
</tr>
<tr>
<td>total</td>
<td>( m )</td>
<td>( R + B - m )</td>
<td>( R + B )</td>
</tr>
</tbody>
</table>

Note the row and column totals are both fixed in advance.

Fisher’s Tea-Tasting Experiment

The lady is told that milk was poured first in 4 cups and tea first in the other 4. Order of tasting is randomized.

- Row and column totals were fixed before the experiment.
- \( n_{11} \) determined the counts in the other 3 cells.
- If the lady had no idea and guessed at random (\( H_0 \)), \( n_{11} \)’d have a hypergeometric distribution

\[
P(n_{11} = i) = \frac{\binom{4}{i} \binom{4}{4-i}}{\binom{8}{4}}, \quad i = 0, 1, 2, 3, 4.
\]

Under \( H_0 \), the lady would be correct for all cups with prob.

\[
P(n_{11} = 4) = \frac{\binom{4}{4} \binom{4}{0}}{\binom{8}{4}} = \frac{1 \times 1}{70} = \frac{1}{70} = 0.014.
\]

Fisher’s Exact Test (for 2 \( \times \) 2 Table)

Under \( H_0: X, Y \) independent

<table>
<thead>
<tr>
<th>( Y = 1 )</th>
<th>( Y = 2 )</th>
<th>margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X = 1 )</td>
<td>( n_{11} )</td>
<td>( n_{12} )</td>
</tr>
<tr>
<td>( X = 2 )</td>
<td>( n_{21} )</td>
<td>( n_{22} )</td>
</tr>
<tr>
<td>margin</td>
<td>( n_{1+} )</td>
<td>( n_{2+} )</td>
</tr>
</tbody>
</table>

Treating the row and column totals as fixed, the exact null distribution of \( n_{ij} \) is the hypergeometric distribution:

\[
P(n_{11}) = \frac{n_{1+} n_{2+}}{n_{11} (n_{1+} - n_{11})} = \frac{\binom{n_{1+}}{n_{11}} \binom{n_{2+}}{n_{21}}}{\binom{n}{n_{11}}} = \frac{\binom{n_{1+}}{n_{11}} \binom{n_{1+} - n_{11}}{n_{21}}}{\binom{n}{n_{11}}}
\]

The two formulas give identical values.

Fisher’s exact test treats the rows and columns symmetrically.
One-Sided Fisher’s Exact Test (for $2 \times 2$ Tables)

For $2 \times 2$ tables,

\[ H_0 : \text{independence} \iff H_0 : \theta = 1 \quad (\theta = \text{odds ratio}) \]

To test \( H_0 : \theta = 1 \) vs \( H_a : \theta > 1 \),

\[ P\text{-value} = P(n_{11} \geq n_{11}^{\text{obs}}) \]

Example (Fisher’s Tea-Tasting Experiment, Cont’d)

The lady guessed correctly on 3 of the milk-first cups and 3 of the tea-first, i.e., \( n_{11} = 3 \).

\[ P\text{-value} = P(n_{11} \geq 3) \]
\[ = P(n_{11} = 3) + P(n_{11} = 4) \]
\[ = 0.229 + 0.014 = 0.243. \]

Very little evidence against \( H_0 \).

Chapter 2C - 5

Fisher’s Exact Test In R

\[ \text{TeaTasting} = \text{matrix}(c(3, 1, 1, 3), \text{nrow} = 2, \text{dimnames} = \text{list(Truth} = c(“Milk”, “Tea”), \text{Guess} = c(“Milk”, “Tea”))) \]

\[ \text{TeaTasting} \]

\begin{array}{ccc}
\text{Truth} & \text{Milk} & \text{Tea} \\
\text{Milk} & 3 & 1 \\
\text{Tea} & 1 & 3 \\
\end{array}

Without any specification, R performs the two-sided test.

\[ \text{fisher.test(TeaTasting)} \]

\begin{tabular}{l}
Fisher’s Exact Test for Count Data \\
\hline \\
\text{data: TeaTasting} \\
p-value = 0.4857 \\
alternative hypothesis: true odds ratio is not equal to 1 \\
95 percent confidence interval: \\
0.2117329 6.219337505 \\
sample estimates: \\
odds ratio \\
6.408309 \\
\end{tabular}

Chapter 2C - 7

Two-Sided Fisher’s Exact Test (for $2 \times 2$ Tables)

To test \( H_0 : \theta = 1 \) vs \( H_a : \theta \neq 1 \),

\[ P\text{-value} = \text{sum of prob. of outcomes no more likely than obs.} \]

Example (Fisher’s Tea-Tasting Experiment, Cont’d)

Under \( H_0 \), we know \( P(n_{11} = i) = \binom{4}{i}\binom{4}{4-i}/\binom{8}{4} \), for \( i = 0, 1, 2, 3, 4 \), from which we can compute the prob. for all possible values of \( n_{11} \):

\[ \begin{array}{ccccc}
\hline
\text{n}_{11} & 0 & 1 & 2 & 3 & 4 \\
\hline
\text{P(n}_{11}) & 1/70 & 16/70 & 36/70 & 16/70 & 1/70 \\
\hline
\end{array} \]

The set of outcomes “no more likely than the observed (\( n_{11} = 3 \))” includes \( n_{11} = 0, 1, 3, \) and 4. The two-sided \( P\text{-value} \) is thus

\[ P(n_{11} = 0) + P(n_{11} = 1) + P(n_{11} = 3) + P(n_{11} = 4) \]
\[ = \frac{1}{70} + \frac{16}{70} + \frac{16}{70} + \frac{1}{70} = \frac{34}{70} \approx 0.486 \]

Chapter 2C - 6

One-sided Fisher’s Exact Test In R

\[ \text{fisher.test(TeaTasting, alternative = “greater”) \[ \]

\begin{tabular}{l}
Fisher’s Exact Test for Count Data \\
\hline \\
data: TeaTasting \\
p-value = 0.2429 \\
alternative hypothesis: true odds ratio is greater than 1 \\
95 percent confidence interval: \\
0.3135693 Inf \\
sample estimates: \\
odds ratio \\
6.408309 \\
\end{tabular}

Chapter 2C - 8
**Example: ECMO**

Extracorporeal membrane oxygenation (ECMO) is a potentially life-saving procedure for treating newborn babies suffering from severe respiratory failure. An experiment was conducted in which 29 babies were treated with ECMO and 10 babies treated with conventional medical therapy (CMT).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>CMT</th>
<th>ECMO</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Die</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Live</td>
<td>6</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>29</td>
<td>39</td>
</tr>
</tbody>
</table>

Though the row totals (5, 34) are not fixed in advance, Fisher’s exact test can be applied. The conditional distribution of $n_{ij}$ given the row totals is still hypergeometric.

$$P(n_{11} = i) = \frac{\binom{5}{i} \binom{34}{10-i}}{\binom{39}{10}}, \quad i = 0, 1, 2, 3, 4, 5$$

**Example: ECMO — One-sided $P$-value**

To find the one-sided $P$-value, we calculate the probability for all outcomes not more likely than the observed ($n_{11} = 4$).

$$P(n_{11} = i) = \frac{\binom{5}{i} \binom{34}{10-i}}{\binom{39}{10}}, \quad i = 0, 1, 2, 3, 4, 5,$$

R can do the computation for us:

```r
> cbind(0:5, dhyper(0:5,5,34,10))
[,1] [,2]
[1,]  0 0.2062588905
[2,]  1 0.4125177809
[3,]  2 0.2855892330
[4,]  3 0.0846190320
[5,]  4 0.0105773900
[6,]  5 0.0004376846
```

The set of outcomes “no more likely than the observed ($n_{11} = 4$)” only includes $n_{11} = 4$ and 5. The two-sided $P$-value is thus equal to the one-sided $P$-value

$$P(n_{11} = 4) + P(n_{11} = 5) \approx 0.0106 + 0.0004 \approx 0.011$$

**Example: ECMO — Two-sided $P$-value**

To find the 2-sided $P$-value, we have to compute the prob. for all possible outcomes.

$$P(n_{11} = i) = \frac{\binom{5}{i} \binom{34}{10-i}}{\binom{39}{10}}, \quad i = 0, 1, 2, 3, 4, 5,$$

R can do the computation for us:

```r
> ECMOdata = matrix(c(4, 6, 1, 28), nrow = 2, dimnames = list(Outcome = c("Die", "Live"), Treatment = c("CMT", "ECMO")))
> ECMOdata
          Treatment
    Outcome CMT ECMO
Die        4   1
Live       6  28
> fisher.test(ECMOdata, alternative="greater")
Fisher's Exact Test for Count Data

data: ECMOdata
p-value = 0.01102
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
  1.833681 Inf
sample estimates:
  odds ratio
16.78571
```

**Chapter 2C - 9**

**Example: ECMO**

Extracorporeal membrane oxygenation (ECMO) is a potentially life-saving procedure for treating newborn babies suffering from severe respiratory failure. An experiment was conducted in which 29 babies were treated with ECMO and 10 babies treated with conventional medical therapy (CMT).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>CMT</th>
<th>ECMO</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Die</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Live</td>
<td>6</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>29</td>
<td>39</td>
</tr>
</tbody>
</table>

Though the row totals (5, 34) are not fixed in advance, Fisher’s exact test can be applied. The conditional distribution of $n_{ij}$ given the row totals is still hypergeometric.

$$P(n_{11} = i) = \frac{\binom{5}{i} \binom{34}{10-i}}{\binom{39}{10}}, \quad i = 0, 1, 2, 3, 4, 5$$

**Example: ECMO — One-sided $P$-value**

To find the one-sided $P$-value, we calculate the probability for all outcomes not more likely than the observed ($n_{11} = 4$).

$$P(n_{11} = i) = \frac{\binom{5}{i} \binom{34}{10-i}}{\binom{39}{10}}, \quad i = 0, 1, 2, 3, 4, 5,$$

R can do the computation for us:

```r
> cbind(0:5, dhyper(0:5,5,34,10))
[,1] [,2]
[1,]  0 0.2062588905
[2,]  1 0.4125177809
[3,]  2 0.2855892330
[4,]  3 0.0846190320
[5,]  4 0.0105773900
[6,]  5 0.0004376846
```

The set of outcomes “no more likely than the observed ($n_{11} = 4$)” only includes $n_{11} = 4$ and 5. The two-sided $P$-value is thus equal to the one-sided $P$-value

$$P(n_{11} = 4) + P(n_{11} = 5) \approx 0.0106 + 0.0004 \approx 0.011$$

**Example: ECMO — Two-sided $P$-value**

To find the 2-sided $P$-value, we have to compute the prob. for all possible outcomes.

$$P(n_{11} = i) = \frac{\binom{5}{i} \binom{34}{10-i}}{\binom{39}{10}}, \quad i = 0, 1, 2, 3, 4, 5,$$

R can do the computation for us:

```r
> ECMOdata = matrix(c(4, 6, 1, 28), nrow = 2, dimnames = list(Outcome = c("Die", "Live"), Treatment = c("CMT", "ECMO")))
> ECMOdata
          Treatment
    Outcome CMT ECMO
Die        4   1
Live       6  28
> fisher.test(ECMOdata, alternative="greater")
Fisher’s Exact Test for Count Data

data: ECMOdata
p-value = 0.01102
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
  1.833681 Inf
sample estimates:
  odds ratio
16.78571
```

**Chapter 2C - 10**
R code:

```r
> fisher.test(ECMOdata)

Fisher's Exact Test for Count Data

data: ECMOdata
p-value = 0.01102
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  1.366318 944.080411
sample estimates:
  odds ratio
  16.78571
```

Remarks About Fisher’s Exact Test

- Fisher’s exact test is conservative. If one rejects $H_0$ when $P$-value $\leq \alpha = 0.05$, actual $P$ (type I error) $< 0.05$ because of discreteness (see section 2.6.3 in [ICDA]).
- There are more than one way to compute the two-sided $P$-values. See Section 3.5.3 in [CDA].
- Small sample confidence intervals for odds ratio can be also constructed. See Section 3.5.4 and 16.6.4 in [CDA].
- Fisher’s exact tests for $I \times J$ tables exist but are computationally intensive. See section 16.5 in [CDA].