2.6 Exact Inference for Small Samples

2.6.1 Fisher's Exact Test

Hypergeometric Distribution

R red balls, B blue balls

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:**

$$\mathrm{P}(X = i) = \mathrm{P}(i \text{ red},\ m - i \text{ blue}) = \frac{\binom{R}{i}\binom{B}{m-i}}{\binom{R+B}{m}}$$

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

The outcome of the draws can be displayed in a 2×2 table:

Color	Drawn	Not Drawn	total
Red	X	$R - X$	R
Blue	$m - X$	$B - (m - X)$	B
total	m	$R + B - m$	$R + B$

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.

	Lady's Guess		
Poured first	Milk first	Tea first	total
Milk	$n_{11} = ?$		4
Tea			4
total	4	4	8

- 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

$$\mathrm{P}(n_{11} = i) = \binom{4}{i}\binom{4}{4-i}\Big/\binom{8}{4}, \quad i = 0, 1, 2, 3, 4.$$

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

$$\mathrm{P}(n_{11} = 4) = \frac{\binom{4}{4}\binom{4}{0}}{\binom{8}{4}} = \frac{\frac{4!}{4!0!}\frac{4!}{0!4!}}{\frac{8!}{4!4!}} = \frac{4!4!}{8!} = \frac{1}{70} = 0.014.$$

Fisher's Exact Test (for 2×2 Table)

Under H_0: X, Y independent

	$Y = 1$	$Y = 2$	margin
$X = 1$	n_{11}	n_{12}	n_{1+}
$X = 2$	n_{21}	n_{22}	n_{2+}
margin	n_{+1}	n_{+2}	n

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

$$\begin{aligned}\mathrm{P}(n_{11}) &= \binom{n_{1+}}{n_{11}}\binom{n_{2+}}{n_{+1} - n_{11}}\Big/\binom{n}{n_{+1}} \\ &= \binom{n_{+1}}{n_{11}}\binom{n_{+2}}{n_{1+} - n_{11}}\Big/\binom{n}{n_{1+}}.\end{aligned}$$

The two formulas give identical values.
Fisher's exact test treats the rows and columns symmetrically.

One-Sided Fisher's Exact Test (for 2×2 Tables)

For 2×2 tables,

$$H_0 : \text{independence} \quad \Leftrightarrow \quad H_0 : \theta = 1 \quad (\theta = \text{odds ratio})$$

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

$$P\text{-value} = \mathrm{P}(n_{11} \geq n_{11}^{\text{obs}})$$

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.., $n_{11} = 3$.

$$\begin{aligned} P\text{-value} &= \mathrm{P}(n_{11} \geq 3) \\ &= \mathrm{P}(n_{11} = 3) + \mathrm{P}(n_{11} = 4) \\ &= 0.229 + 0.014 = 0.243. \end{aligned}$$

Very little evidence against H_0.

Two-Sided Fisher's Exact Test (for 2×2 Tables)

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

P-value = sum of prob. of outcomes no more likely than obs.

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

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

n_{11}	0	1	2	3	4
$\mathrm{P}(n_{11})$	$\frac{1}{70}$	$\frac{16}{70}$	$\frac{36}{70}$	$\frac{16}{70}$	$\frac{1}{70}$

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-value is thus

$$\begin{aligned} &\mathrm{P}(n_{11} = 0) + \mathrm{P}(n_{11} = 1) + \mathrm{P}(n_{11} = 3) + \mathrm{P}(n_{11} = 4) \\ &= \frac{1}{70} + \frac{16}{70} + \frac{16}{70} + \frac{1}{70} = \frac{34}{70} \approx 0.486 \end{aligned}$$

Fisher's Exact Test In R

```
> TeaTasting = matrix(c(3, 1, 1, 3), nrow = 2,
   dimnames = list(Truth = c("Milk", "Tea"), Guess = c("Milk", "Tea")))
> TeaTasting
      Guess
Truth  Milk Tea
  Milk    3   1
  Tea     1   3
```

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

```
> fisher.test(TeaTasting)

        Fisher's Exact Test for Count Data

data:  TeaTasting
p-value = 0.4857
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
   0.2117329 621.9337505
sample estimates:
odds ratio
  6.408309
```


One-sided Fisher's Exact Test In R

```
> fisher.test(TeaTasting, alternative = "greater")

        Fisher's Exact Test for Count Data

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
```

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).

		Treatment		
		CMT	ECMO	total
Outcome	Die	4	1	5
	Live	6	28	34
	total	10	29	39

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.

$$\mathrm{P}(n_{11} = i) = \frac{\binom{5}{i}\binom{34}{10-i}}{\binom{39}{10}} = \frac{\binom{10}{i}\binom{29}{5-i}}{\binom{39}{5}}, \quad i = 0, 1, 2, 3, 4, 5$$

Example: ECMO — Two-sided P-value

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

$$\mathrm{P}(n_{11} = i) = \frac{\binom{5}{i}\binom{34}{10-i}}{\binom{39}{10}} = \frac{\binom{10}{i}\binom{29}{5-i}}{\binom{39}{5}}, \quad \text{for } i = 0, 1, 2, 3, 4, 5,$$

R can do the computation for us:

```
> 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 0.0105773790
[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

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

Example: ECMO — One-sided P-value

		Treatment		
		CMT	ECMO	total
Outcome	Die	4	1	5
	Live	6	28	34
	total	10	29	39

$\mathrm{H}_0 : \theta \leq 1$ (ECMO is not more effective than CMT) v.s.
$\mathrm{H}_a : \theta > 1$ (ECMO is more effective)

$$\begin{aligned} P\text{-value} &= \mathrm{P}(n_{11} \geq n_{11}^{\text{obs}}) = \mathrm{P}(n_{11} \geq 4) \\ &= \mathrm{P}(n_{11} = 4) + \mathrm{P}(n_{11} = 5) \\ &= \frac{\binom{5}{4}\binom{34}{6}}{\binom{39}{10}} + \frac{\binom{5}{5}\binom{34}{5}}{\binom{39}{10}} \approx 0.0106 + 0.0004 \approx 0.011 \end{aligned}$$

```
> dhyper(4:5,5,34,10)
[1] 0.0105773790 0.0004376846
```



```
> 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
```

```
> 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 $\mathrm{P}(\text{type I error}) < 0.05$ 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].