# Contingency Tables

## 2 x 2 tables

Apply a treatment to 20 mice from strains A and B, and observe survival.

|   | N | Y |   |
|---|---|---|---|
| A | 18 | 2 | 20 |
| B | 11 | 9 | 20 |
|   | 29 | 11 | 40 |

Question:

⟶ Are the survival rates in the two strains the same?

Gather 100 rats and determine whether they are infected with viruses A and B.

|   | I-B | NI-B |   |
|---|---|---|---|
| I-A | 9 | 9 | 18 |
| NI-A | 20 | 62 | 82 |
|   | 29 | 71 | 100 |

Question:

⟶ Is infection with virus A independent of infection with virus B?

# Underlying probabilities

→ Observed data

$$\begin{array}{cc|cc|c} & & \multicolumn{2}{c}{B} & \\ & & 0 & 1 & \\ \cline{3-4} A & 0 & n_{00} & n_{01} & n_{0+} \\ & 1 & n_{10} & n_{11} & n_{1+} \\ \cline{3-4} & & n_{+0} & n_{+1} & n \end{array}$$

→ Underlying probabilities

$$\begin{array}{cc|cc|c} & & \multicolumn{2}{c}{B} & \\ & & 0 & 1 & \\ \cline{3-4} A & 0 & p_{00} & p_{01} & p_{0+} \\ & 1 & p_{10} & p_{11} & p_{1+} \\ \cline{3-4} & & p_{+0} & p_{+1} & 1 \end{array}$$

Model:

$$(n_{00}, n_{01}, n_{10}, n_{11}) \sim \text{Multinomial}(n,\{p_{00}, p_{01}, p_{10}, p_{11}\})$$

or

$$n_{01} \sim \text{Binomial}(n_{0+}, p_{01}/p_{0+}) \quad \text{and} \quad n_{11} \sim \text{Binomial}(n_{1+}, p_{11}/p_{1+})$$

# Conditional probabilities

Underlying probabilities

$$\begin{array}{cc|cc|c} & & \multicolumn{2}{c}{B} & \\ & & 0 & 1 & \\ \cline{3-4} A & 0 & p_{00} & p_{01} & p_{0+} \\ & 1 & p_{10} & p_{11} & p_{1+} \\ \cline{3-4} & & p_{+0} & p_{+1} & 1 \end{array}$$

Conditional probabilities

$\Pr(B = 1 \mid A = 0) = p_{01}/p_{0+}$

$\Pr(B = 1 \mid A = 1) = p_{11}/p_{1+}$

$\Pr(A = 1 \mid B = 0) = p_{10}/p_{+0}$

$\Pr(A = 1 \mid B = 1) = p_{11}/p_{+1}$

→ The questions in the two examples are the same!

They both concern: $p_{01}/p_{0+} = p_{11}/p_{1+}$

Equivalently: $p_{ij} = p_{i+} \times p_{+j}$ for all i,j → think Pr(A and B) = Pr(A) × Pr(B).

# This is a composite hypothesis!

2 x 2 table

$$\begin{array}{cc|cc|c} & & \multicolumn{2}{c}{B} & \\ & & 0 & 1 & \\ \hline A & 0 & p_{00} & p_{01} & p_{0+} \\ & 1 & p_{10} & p_{11} & p_{1+} \\ \hline & & p_{+0} & p_{+1} & 1 \end{array}$$

H$_0$: $p_{ij} = p_{i+} \times p_{+j}$ for all i,j

A different view

$$\boxed{p_{00}\ \ p_{01}\ \ p_{10}\ \ p_{11}}$$

H$_0$: $p_{ij} = p_{i+} \times p_{+j}$ for all i,j

Degrees of freedom = 4 - 2 - 1 = 1

# Expected counts

Observed data

$$\begin{array}{cc|cc|c} & & \multicolumn{2}{c}{B} & \\ & & 0 & 1 & \\ \hline A & 0 & n_{00} & n_{01} & n_{0+} \\ & 1 & n_{10} & n_{11} & n_{1+} \\ \hline & & n_{+0} & n_{+1} & n \end{array}$$

Expected counts

$$\begin{array}{cc|cc|c} & & \multicolumn{2}{c}{B} & \\ & & 0 & 1 & \\ \hline A & 0 & e_{00} & e_{01} & n_{0+} \\ & 1 & e_{10} & e_{11} & n_{1+} \\ \hline & & n_{+0} & n_{+1} & n \end{array}$$

To get the expected counts under the null hypothesis we:

$\longrightarrow$ Estimate $p_{1+}$ and $p_{+1}$ by $n_{1+}/n$ and $n_{+1}/n$, respectively.

These are the MLEs under H$_0$!

$\longrightarrow$ Turn these into estimates of the $p_{ij}$.

$\longrightarrow$ Multiply these by the total sample size, n.

# The expected counts

The expected count (assuming $H_0$) for the “11” cell is the following:

$$\begin{aligned} e_{11} &= n \times \hat{p}_{11} \\ &= n \times \hat{p}_{1+} \times \hat{p}_{+1} \\ &= n \times (n_{1+}/n) \times (n_{+1}/n) \\ &= (n_{1+} \times n_{+1})/n \end{aligned}$$

The other cells are similar.

⟶ We can then calculate a $\chi^2$ or LRT statistic as before!

# Example 1

Observed data

| | N | Y | |
|---|---|---|---|
| A | 18 | 2 | 20 |
| B | 11 | 9 | 20 |
| | 29 | 11 | 40 |

Expected counts

| | N | Y | |
|---|---|---|---|
| A | 14.5 | 5.5 | 20 |
| B | 14.5 | 5.5 | 20 |
| | 29 | 11 | 40 |

$$X^2 = \frac{(18-14.5)^2}{14.5} + \frac{(11-14.5)^2}{14.5} + \frac{(2-5.5)^2}{5.5} + \frac{(9-5.5)^2}{5.5} = 6.14$$

$$\text{LRT} = 2 \times \left[18 \log(\tfrac{18}{14.5}) + \ldots + 9 \log(\tfrac{9}{5.5})\right] = 6.52$$

P-values (based on the asymptotic $\chi^2(\text{df} = 1)$ approximation):

1.3% and 1.1%, respectively.

# Example 2

Observed data

| | I-B | NI-B | |
|---|---|---|---|
| I-A | 9 | 9 | 18 |
| NI-A | 20 | 62 | 82 |
| | 29 | 71 | 100 |

Expected counts

| | I-B | NI-B | |
|---|---|---|---|
| I-A | 5.2 | 12.8 | 18 |
| NI-A | 23.8 | 58.2 | 82 |
| | 29 | 71 | 100 |

$$X^2 = \frac{(9-5.2)^2}{5.2} + \frac{(20-23.8)^2}{23.8} + \frac{(9-12.8)^2}{12.8} + \frac{(62-58.2)^2}{58.2} = 4.70$$

$$\text{LRT} = 2 \times \left[9 \log(\tfrac{9}{5.2}) + \ldots + 62 \log(\tfrac{62}{58.2})\right] = 4.37$$

P-values (based on the asymptotic $\chi^2$(df = 1) approximation):

3.0% and 3.7%, respectively.

# Fisher's exact test

Observed data

| | N | Y | |
|---|---|---|---|
| A | 18 | 2 | 20 |
| B | 11 | 9 | 20 |
| | 29 | 11 | 40 |

- Assume the null hypothesis (independence) is true.
- Constrain the marginal counts to be as observed.
- What's the chance of getting this exact table?
- What's the chance of getting a table at least as "extreme"?

# Hypergeometric distribution

- Imagine an urn with K white balls and N – K black balls.
- Draw n balls without replacement.
- Let x be the number of white balls in the sample.
- x follows a hypergeometric distribution (w/ parameters K, N, n).

In urn

| | white | black | |
|---|---|---|---|
| sampled | x | | n |
| not sampled | | | N – n |
| | K | N – K | N |

# Hypergeometric probabilities

Suppose X $\sim$ Hypergeometric (N, K, n).

No. of white balls in a sample of size n, drawn without replacement from an urn with K white and N – K black.

$$\Pr(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

Example:

In urn

| | 0 | 1 | |
|---|---|---|---|
| sampled | 18 | | 20 |
| not | | | 20 |
| | 29 | 11 | 40 |

N = 40, K = 29, n = 20

$$\Pr(X = 18) = \frac{\binom{29}{18}\binom{40-29}{20-18}}{\binom{40}{20}} \approx 1.4\%$$

# The hypergeometric in R

```
dhyper(x, m, n, k)
phyper(q, m, n, k)
qhyper(p, m, n, k)
rhyper(nn, m, n, k)
```

In R, things are set up so that

`m` = no. white balls in urn

`n` = no. black balls in urn

`k` = no. balls sampled (without replacement)

`x` = no. white balls in sample

`nn` = no. of observations

# Back to Fisher's exact test

Observed data

|   | N | Y |   |
|---|---|---|---|
| A | 18 | 2 | 20 |
| B | 11 | 9 | 20 |
|   | 29 | 11 | 40 |

- Assume the null hypothesis (independence) is true.
- Constrain the marginal counts to be as observed.
- Pr(observed table $|$ $H_0$) = Pr(X=18)

  X $\sim$ Hypergeometric (N=40, K=29, n=20)

# Fisher's exact test

1. For all possible tables (with the observed marginal counts), calculate the relevant hypergeometric probability.

2. Use that probability as a statistic.

3. P-value (for Fisher's exact test of independence):

$\longrightarrow$ The sum of the probabilities for all tables having a probability equal to or smaller than that observed.

# An illustration

The observed data

|   | N | Y |   |
|---|---|---|---|
| A | 18 | 2 | 20 |
| B | 11 | 9 | 20 |
|   | 29 | 11 | 40 |

All possible tables (with these marginals):

$\begin{array}{|cc|}\hline 20 & 0 \\ 9 & 11 \\ \hline\end{array} \rightarrow 0.00007$

$\begin{array}{|cc|}\hline 19 & 1 \\ 10 & 10 \\ \hline\end{array} \rightarrow 0.00160$

$\begin{array}{|cc|}\hline 18 & 2 \\ 11 & 9 \\ \hline\end{array} \rightarrow 0.01380$

$\begin{array}{|cc|}\hline 17 & 3 \\ 12 & 8 \\ \hline\end{array} \rightarrow 0.06212$

$\begin{array}{|cc|}\hline 16 & 4 \\ 13 & 7 \\ \hline\end{array} \rightarrow 0.16246$

$\begin{array}{|cc|}\hline 15 & 5 \\ 14 & 6 \\ \hline\end{array} \rightarrow 0.25994$

$\begin{array}{|cc|}\hline 14 & 6 \\ 15 & 5 \\ \hline\end{array} \rightarrow 0.25994$

$\begin{array}{|cc|}\hline 13 & 7 \\ 16 & 4 \\ \hline\end{array} \rightarrow 0.16246$

$\begin{array}{|cc|}\hline 12 & 8 \\ 17 & 3 \\ \hline\end{array} \rightarrow 0.06212$

$\begin{array}{|cc|}\hline 11 & 9 \\ 18 & 2 \\ \hline\end{array} \rightarrow 0.01380$

$\begin{array}{|cc|}\hline 10 & 10 \\ 19 & 1 \\ \hline\end{array} \rightarrow 0.00160$

$\begin{array}{|cc|}\hline 9 & 11 \\ 20 & 0 \\ \hline\end{array} \rightarrow 0.00007$

# Fisher's exact test: example 1

Observed data

|   | N | Y |   |
|---|---|---|---|
| A | 18 | 2 | 20 |
| B | 11 | 9 | 20 |
|   | 29 | 11 | 40 |

P-value $\approx$ 3.1%

In R: `fisher.test()`

Recall:

→ $\chi^2$ test: P-value = 1.3%

→ LRT: P-value = 1.1%

# Fisher's exact test: example 2

Observed data

|   | I-B | NI-B |   |
|---|---|---|---|
| I-A | 9 | 9 | 18 |
| NI-A | 20 | 62 | 82 |
|   | 29 | 71 | 100 |

P-value $\approx$ 4.4%

Recall:

→ $\chi^2$ test: P-value = 3.0%

→ LRT: P-value = 3.7%

# Summary

Testing for independence in a 2 x 2 table:

- A special case of testing a composite hypothesis in a one-dimensional table.
- You can use either the LRT or $\chi^2$ test, as before.
- You can also use Fisher's exact test.
- If Fisher's exact test is computationally feasible, do it!

# Paired data

Gather 100 rats and determine whether they are infected with viruses A and B.

| | I-B | NI-B | |
|---|---|---|---|
| I-A | 9 | 9 | 18 |
| NI-A | 20 | 62 | 82 |
| | 29 | 71 | 100 |

Underlying probabilities

| | | B | | |
|---|---|---|---|---|
| | | 0 | 1 | |
| A | 0 | $p_{00}$ | $p_{01}$ | $p_{0+}$ |
| | 1 | $p_{10}$ | $p_{11}$ | $p_{1+}$ |
| | | $p_{+0}$ | $p_{+1}$ | 1 |

$\longrightarrow$ Is the rate of infection of virus A the same as that of virus B?

In other words: Is $p_{1+} = p_{+1}$? Equivalently, is $p_{10} = p_{01}$?

# McNemar’s test

$H_0$: $p_{01} = p_{10}$

Under $H_0$, e.g. if $p_{01} = p_{10}$, the expected counts for cells 01 and 10 are both equal to $(n_{01} + n_{10})/2$.

The $\chi^2$ test statistic reduces to $X^2 = \dfrac{(n_{01} - n_{10})^2}{n_{01} + n_{10}}$

For large sample sizes, this statistic has null distribution that is approximately a $\chi^2$(df = 1).

For the example: $X^2 = (20 - 9)^2 / 29 = 4.17$ $\longrightarrow$ P = 4.1%.

# An exact test

Condition on $n_{01} + n_{10}$.

Under $H_0$, $n_{01} \mid n_{01} + n_{10} \sim \text{Binomial}(n_{01} + n_{10}, 1/2)$.

In R, use the function `binom.test`.

$\longrightarrow$ For the example, P = 6.1%.

# Paired data

Paired data

| | I-B | NI-B | |
|---|---|---|---|
| I-A | 9 | 9 | 18 |
| NI-A | 20 | 62 | 82 |
| | 29 | 71 | 100 |

→ P = 6.1%

Unpaired data

| | I | NI | |
|---|---|---|---|
| A | 18 | 82 | 100 |
| B | 29 | 71 | 100 |
| | 47 | 153 | 200 |

→ P = 9.5%

⟶ Taking appropriate account of the “pairing” is important!

# r x k tables

| | **Blood type** | | | | |
|---|---|---|---|---|---|
| **Population** | A | B | AB | O | |
| Florida | 122 | 117 | 19 | 244 | 502 |
| Iowa | 1781 | 1351 | 289 | 3301 | 6721 |
| Missouri | 353 | 269 | 60 | 713 | 1395 |
| | 2256 | 1737 | 367 | 4258 | 8618 |

⟶ Same distribution of blood types in each population?

# Underlying probabilities

Observed data

$$\begin{array}{c|cccc|c}
 & 1 & 2 & \cdots & k & \\
\hline
1 & n_{11} & n_{12} & \cdots & n_{1k} & n_{1+} \\
2 & n_{21} & n_{22} & \cdots & n_{2k} & n_{2+} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
r & n_{r1} & n_{r2} & \cdots & n_{rk} & n_{r+} \\
\hline
 & n_{+1} & n_{+2} & \cdots & n_{+k} & n
\end{array}$$

Underlying probabilities

$$\begin{array}{c|cccc|c}
 & 1 & 2 & \cdots & k & \\
\hline
1 & p_{11} & p_{12} & \cdots & p_{1k} & p_{1+} \\
2 & p_{21} & p_{22} & \cdots & p_{2k} & p_{2+} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
r & p_{r1} & p_{r2} & \cdots & p_{rk} & p_{r+} \\
\hline
 & p_{+1} & p_{+2} & \cdots & p_{+k} & 1
\end{array}$$

$H_0$: $p_{ij} = p_{i+} \times p_{+j}$ for all i,j.

# Expected counts

Observed data

| | A | B | AB | O | |
|---|---|---|---|---|---|
| F | 122 | 117 | 19 | 244 | 502 |
| I | 1781 | 1351 | 289 | 3301 | 6721 |
| M | 353 | 269 | 60 | 713 | 1395 |
| | 2256 | 1737 | 367 | 4258 | 8618 |

Expected counts

| | A | B | AB | O | |
|---|---|---|---|---|---|
| F | 131 | 101 | 21 | 248 | 502 |
| I | 1759 | 1355 | 286 | 3321 | 6721 |
| M | 365 | 281 | 59 | 689 | 1395 |
| | 2256 | 1737 | 367 | 4258 | 8618 |

Expected counts under $H_0$: $e_{ij} = n_{i+} \times n_{+j}/n$ for all i,j.

# $\chi^2$ and LRT statistics

Observed data

| | A | B | AB | O | |
|---|---|---|---|---|---|
| F | 122 | 117 | 19 | 244 | 502 |
| I | 1781 | 1351 | 289 | 3301 | 6721 |
| M | 353 | 269 | 60 | 713 | 1395 |
| | 2256 | 1737 | 367 | 4258 | 8618 |

Expected counts

| | A | B | AB | O | |
|---|---|---|---|---|---|
| F | 131 | 101 | 21 | 248 | 502 |
| I | 1759 | 1355 | 286 | 3321 | 6721 |
| M | 365 | 281 | 59 | 689 | 1395 |
| | 2256 | 1737 | 367 | 4258 | 8618 |

$$X^2 \text{ statistic} = \sum \frac{(\text{obs}-\text{exp})^2}{\text{exp}} = \cdots = 5.64$$

$$\text{LRT statistic} = 2 \times \sum \text{obs} \ln(\text{obs}/\text{exp}) = \cdots = 5.55$$

# Asymptotic approximation

If the sample size is large, the null distribution of the $\chi^2$ and likelihood ratio test statistics will approximately follow a

$$\chi^2 \text{ distribution with } (r-1) \times (k-1) \text{ d.f.}$$

In the example, df = $(3-1) \times (4-1) = 6$

$X^2 = 5.64 \quad \longrightarrow \quad P = 0.46.$

$\text{LRT} = 5.55 \quad \longrightarrow \quad P = 0.48.$

# Fisher’s exact test

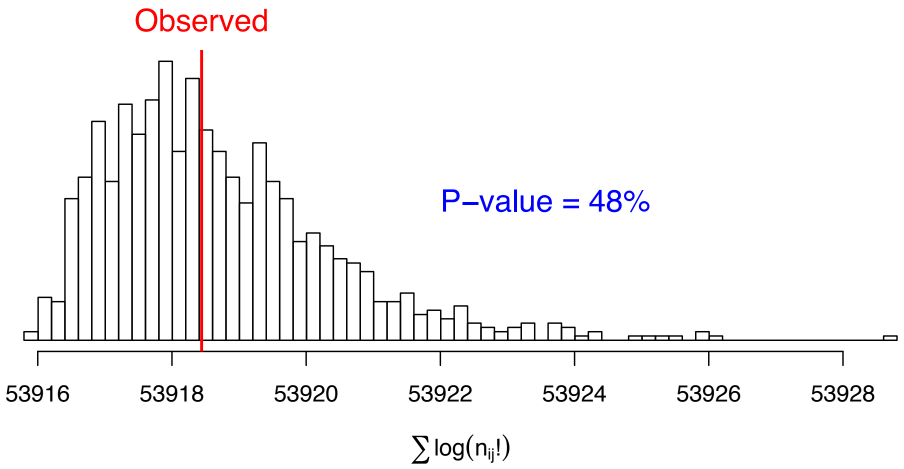
Observed data

$$\begin{array}{c|cccc|c} & 1 & 2 & \cdots & k & \\ \hline 1 & n_{11} & n_{12} & \cdots & n_{1k} & n_{1+} \\ 2 & n_{21} & n_{22} & \cdots & n_{2k} & n_{2+} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ r & n_{r1} & n_{r2} & \cdots & n_{rk} & n_{r+} \\ \hline & n_{+1} & n_{+2} & \cdots & n_{+k} & n \end{array}$$

- Assume $H_0$ is true.
- Condition on the marginal counts
- Then $\Pr(\text{table}) \propto 1/\prod_{ij} n_{ij}!$

- Consider all possible tables with the observed marginal counts
- Calculate Pr(table) for each possible table.
- P-value = the sum of the probabilities for all tables having a probability equal to or smaller than that observed.

# Fisher’s exact test: the example



⟶ Since the number of possible tables can be very large, we often must resort to computer simulation.

# Another example

Survival following treatment in five mouse strains:

| | Survive | |
|---|---|---|
| **Strain** | No | Yes |
| A | 15 | 5 |
| B | 17 | 3 |
| C | 10 | 10 |
| D | 17 | 3 |
| E | 16 | 4 |

⟶ Is the survival rate the same for all strains?

# Results

Observed

| | Survive | |
|---|---|---|
| **Strain** | No | Yes |
| A | 15 | 5 |
| B | 17 | 3 |
| C | 10 | 10 |
| D | 17 | 3 |
| E | 16 | 4 |

Expected under $H_0$

| | Survive | |
|---|---|---|
| **Strain** | No | Yes |
| A | 15 | 5 |
| B | 15 | 5 |
| C | 15 | 5 |
| D | 15 | 5 |
| E | 15 | 5 |

$X^2 = 9.07$ ⟶ P = 5.9% (how many df?)

LRT = 8.41 ⟶ P = 7.8%

Fisher's exact test: P = 8.7%

# All pairwise comparisons

$$\begin{array}{c|cc|} & N & Y \\ \hline A & 15 & 5 \\ B & 17 & 3 \\ \hline \end{array} \longrightarrow P=69\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline A & 15 & 5 \\ C & 10 & 10 \\ \hline \end{array} \longrightarrow P=19\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline A & 15 & 5 \\ D & 17 & 3 \\ \hline \end{array} \longrightarrow P=69\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline A & 15 & 5 \\ E & 16 & 4 \\ \hline \end{array} \longrightarrow P=100\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline B & 17 & 3 \\ C & 10 & 10 \\ \hline \end{array} \longrightarrow P=4.1\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline B & 17 & 3 \\ D & 17 & 3 \\ \hline \end{array} \longrightarrow P=100\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline B & 17 & 3 \\ E & 16 & 4 \\ \hline \end{array} \longrightarrow P=100\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline C & 10 & 10 \\ D & 17 & 3 \\ \hline \end{array} \longrightarrow P=4.1\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline C & 10 & 10 \\ E & 16 & 4 \\ \hline \end{array} \longrightarrow P=9.6\%$$

$$\begin{array}{c|cc|} & N & Y \\ \hline D & 17 & 3 \\ E & 16 & 4 \\ \hline \end{array} \longrightarrow P=100\%$$

Is this a good thing to do?

# Two-locus linkage in an intercross

| | BB | Bb | bb |
|---|---|---|---|
| AA | 6 | 15 | 3 |
| Aa | 9 | 29 | 6 |
| aa | 3 | 16 | 13 |

Are these two loci linked?

# General test of independence

Observed data

| | BB | Bb | bb |
|---|---|---|---|
| AA | 6 | 15 | 3 |
| Aa | 9 | 29 | 6 |
| aa | 3 | 16 | 13 |

Expected counts

| | BB | Bb | bb |
|---|---|---|---|
| AA | 4.3 | 14.4 | 5.3 |
| Aa | 7.9 | 26.4 | 9.7 |
| aa | 5.8 | 19.2 | 7.0 |

$\chi^2$ test: $X^2 = 10.4 \longrightarrow$ P = 3.5% (df = 4)

LRT test: LRT = 9.98 $\longrightarrow$ P = 4.1%

Fisher's exact test: P = 4.6%

# A more specific test

Observed data

| | BB | Bb | bb |
|---|---|---|---|
| AA | 6 | 15 | 3 |
| Aa | 9 | 29 | 6 |
| aa | 3 | 16 | 13 |

Underlying probabilities

| | BB | Bb | bb |
|---|---|---|---|
| AA | $\frac{1}{4}(1-\theta)^2$ | $\frac{1}{2}\theta(1-\theta)$ | $\frac{1}{4}\theta^2$ |
| Aa | $\frac{1}{2}\theta(1-\theta)$ | $\frac{1}{2}[\theta^2+(1-\theta)^2]$ | $\frac{1}{2}\theta(1-\theta)$ |
| aa | $\frac{1}{4}\theta^2$ | $\frac{1}{2}\theta(1-\theta)$ | $\frac{1}{4}(1-\theta)^2$ |

$H_0$: $\theta = 1/2$ versus $H_a$: $\theta < 1/2$

Use a likelihood ratio test!

$\longrightarrow$ Obtain the general MLE of $\theta$.

$\longrightarrow$ Calculate the LRT statistic = $2 \ln \left\{ \frac{\Pr(\text{data} \mid \hat{\theta})}{\Pr(\text{data} \mid \theta{=}1/2)} \right\}$

$\longrightarrow$ Compare this statistic to a $\chi^2(\text{df} = 1)$.

# Results

|  | BB | Bb | bb |
|---|---|---|---|
| AA | 6 | 15 | 3 |
| Aa | 9 | 29 | 6 |
| aa | 3 | 16 | 13 |

MLE: $\hat{\theta} = 0.359$

LRT statistic: LRT = 7.74 ⟶ P = 0.54% (df = 1)

⟶ Here we assume Mendelian segregation, and that deviation from $H_0$ is “in a particular direction.”

⟶ If these assumptions are correct, we’ll have greater power to detect linkage using this more specific approach.