An inevitable overview: what this is about.

- Log odds ratios
- Log likelihood ratio
Introduction

An inevitable overview: what this is about.

- Log odds ratios
- Log likelihood ratio

More importantly, it’s going to be about the connections between these techniques.
Background

Why this topic?

- Kremers’ 2005 presentation “Test of independence for two-way contingency tables: Application of log likelihood ratio to child acquisition data.”
- In that talk Kremers posed questions about using log odds ratios and the log likelihood ratio statistic.
Background

Why this topic?

- Kremers’ 2005 presentation “Test of independence for two-way contingency tables: Application of log likelihood ratio to child acquisition data.”
- In that talk Kremers posed questions about using log odds ratios and the log likelihood ratio statistic.
- First, a whistle-stop tour of log odds ratios.
A fictitious example:

Forty-six speakers interviewed; eighteen from Exeter, twenty-eight from Plymouth. Conversations were recorded and transcribed. Phrases or words sought in pairs:

Tourist/visitor ↔ grokele/emmet; e.g. "There be many grokeles down from up country dis year"

Dreckly ↔ so on/when possible; e.g. "I'll come dreckly"

Be ↔ [present tense conjugations of to be]; e.g. "I be going up ter Lunnen to da y"

Conversations were guided so that one of these three were observed per conversation. One conversation was conducted with each subject.
A fictitious example:

- Does the place of birth/residence of a person affect the dialect he speaks?
- Forty-six speakers interviewed; eighteen from Exeter, twenty-eight from Plymouth.
- Conversations were recorded and transcribed.
A fictitious example:

- Does the place of birth/residence of a person affect the dialect he speaks?
- Forty-six speakers interviewed; eighteen from Exeter, twenty-eight from Plymouth.
- Conversations were recorded and transcribed.
- Phrases or words sought in pairs:
  - Tourist/visitor $\leftrightarrow$ grockle/emmet; e.g. "There be many grockles down from up country dis year"
  - Dreckly $\leftrightarrow$ soon/when possible; e.g. "I’ll come dreckly"
  - Be $\leftrightarrow$ [present tense conjugations of to be]; e.g. "I be going up ter Lunnen today"

- Conversations were guided so that one of these three were observed per conversation. One conversation was conducted with each subject.
## Log Odds Ratios

What is a success? The odds of success for a particular row is given by the probability of success divided by the probability of failure:

\[ \frac{p_1}{p_2} = \frac{p_1}{1-p_1} \]

### Table

<table>
<thead>
<tr>
<th></th>
<th>dialect ($p_{1.}$)</th>
<th>SBE ($p_{2.}$)</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plymouth</td>
<td>23 (0.82)</td>
<td>5 (0.18)</td>
<td>28</td>
</tr>
<tr>
<td>Exeter</td>
<td>7 (0.39)</td>
<td>11 (0.61)</td>
<td>18</td>
</tr>
<tr>
<td>total</td>
<td>30</td>
<td>16</td>
<td>46</td>
</tr>
</tbody>
</table>
### Log Odds Ratios

What is a *success*?

<table>
<thead>
<tr>
<th>Location</th>
<th>Dialect ((p_1^o))</th>
<th>SBE ((p_2^o))</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plymouth</td>
<td>23 (0.82)</td>
<td>5 (0.18)</td>
<td>28</td>
</tr>
<tr>
<td>Exeter</td>
<td>7 (0.39)</td>
<td>11 (0.61)</td>
<td>18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>30</strong></td>
<td><strong>16</strong></td>
<td><strong>46</strong></td>
</tr>
</tbody>
</table>
What is a *success*?
The *odds of success* for a particular row is given by

\[
\frac{\text{probability of success}}{\text{probability of failure}} = \frac{p_2^o}{p_1^o}.
\]
So the *odds of success* for Plymothians is

\[
O_{\text{Ply}} = \frac{p_{12}^o}{p_{11}^o} = \frac{0.18}{0.82} = 0.22
\]

That of Exonians is

\[
O_{\text{Ex}} = \frac{p_{22}^o}{p_{12}^o} = \frac{0.61}{0.39} = 1.56
\]

(Almost there!)
The log odds ratio with respect to Plymothians ($\theta_{\text{Ply}}$) is thus:

$$\theta_{\text{Ply}} = \frac{O_{\text{Ply}}}{O_{\text{Ex}}} = \frac{0.22}{1.56} = 0.14$$

But how do we interpret this ratio $\theta$? If $\theta_{\text{Ply}}$ were equal to 1, then the odds of success for both Plymothians and Exonians would be equal. If $\theta_{\text{Ply}}$ is greater than 1, Plymothians would be more likely to succeed than Exonians. If $\theta_{\text{Ply}}$ is less than 1, Exonians would be more likely to succeed than Plymothians. Since $\theta_{\text{Ply}} < 1$, Plymothians are (much) less likely to succeed than Exonians.
The *log odds ratio with respect to Plymothians* ($\theta_{Ply}$) is thus:

$$\theta_{Ply} = \frac{O_{Ply}}{O_{Ex}} = \frac{0.22}{1.56} = 0.14$$

But how do we interpret this ratio $\theta$?
Log Odds Ratios

The log odds ratio with respect to Plymothians ($\theta_{Ply}$) is thus:

$$\theta_{Ply} = \frac{O_{Ply}}{O_{Ex}} = \frac{0.22}{1.56} = 0.14$$

But how do we interpret this ratio $\theta$?

If $\theta_{Ply}$ were equal to 1, then the odds of success for both Plymothians and Exonians would be equal.

If $\theta_{Ply}$ is greater than 1, Plymouthians would be more likely to succeed than Exonians.

If $\theta_{Ply}$ is less than 1, Exonians would be more likely to succeed than Plymouthians.
The *log odds ratio with respect to Plymothians* \( (\theta_{\text{Ply}}) \) is thus:

\[
\theta_{\text{Ply}} = \frac{O_{\text{Ply}}}{O_{\text{Ex}}} = \frac{0.22}{1.56} = 0.14
\]

But how do we interpret this ratio \( \theta \)?

If \( \theta_{\text{Ply}} \) were equal to 1, then the odds of success for both Plymothians and Exonians would be equal.

If \( \theta_{\text{Ply}} \) is *greater* than 1, Plymothians would be more likely to succeed than Exonians.

If \( \theta_{\text{Ply}} \) is *less* than 1, Exonians would be more likely to succeed than Plymothians.

\( \theta_{\text{Ply}} < 1 \), so Plymothians are (much) less likely to “succeed” than Exonians.
The odds ratio as it is now is *skewed*. Why?
The odds ratio as it is now is *skewed*. Why?

This is solved by dealing with the *logarithm* of the ratio instead.

\[
\log(\theta_{\text{Ply}}) = \log(0.14) = -1.22
\]

Compare this with \(\theta_{\text{Ex}}\) and \(\log(\theta_{\text{Ex}})\):

\[
\begin{align*}
\theta_{\text{Ex}} &= 7.09 \\
\log(\theta_{\text{Ex}}) &= 1.96
\end{align*}
\]
Log Odds Ratios: However!

- The odds ratio as it is now is skewed. Why?
- This is solved by dealing with the logarithm of the ratio instead.

So

\[ \log(\theta_{Ply}) = \log(0.14) = -1.96 \]

Compare this with \( \theta_{Ex} \) and \( \log \theta_{Ex} \):

\[
\begin{align*}
\theta_{Ex} & \quad 7.09 \\
\log \theta_{Ex} & \quad 1.96
\end{align*}
\]
Kremers presented a significance test called the *log likelihood ratio*, given by the following:

\[ G^2 = 2 \sum_{n_{ij}} n_{ij} \log \frac{n_{ij}}{\mu_{ij}} \]

Where \( n_{ij} \) is the *observed* frequency and \( \mu_{ij} \) the *expected* frequency.
Kremers presented a significance test called the log likelihood ratio, given by the following:

\[ G^2 = 2 \sum n_{ij} \log \frac{n_{ij}}{\mu_{ij}} \]

Where \( n_{ij} \) is the observed frequency and \( \mu_{ij} \) the expected frequency. She posed three questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio?
- Can the value of \( G^2 \) be negative?
- Is the “likelihood ratio” value found in SPSS the same as the “log likelihood ratio”?
A reminder of our data:

<table>
<thead>
<tr>
<th></th>
<th>dialect (expected)</th>
<th>SBE (expected)</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plymouth</td>
<td>23 (18)</td>
<td>5 (10)</td>
<td>28</td>
</tr>
<tr>
<td>Exeter</td>
<td>7 (12)</td>
<td>11 (6)</td>
<td>18</td>
</tr>
<tr>
<td>total</td>
<td>30</td>
<td>16</td>
<td>46</td>
</tr>
</tbody>
</table>

Calculating the log likelihood statistic gives:

\[
G^2 = 2 \sum n_{ij} \log \left( \frac{n_{ij}}{\mu_{ij}} \right) = 2 \left( 5.638 - 3.466 + 3.773 + 6.667 \right) = 2 \times 5.066 = 10.132
\]
A reminder of our data:

<table>
<thead>
<tr>
<th></th>
<th>dialect (expected)</th>
<th>SBE (expected)</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plymouth</td>
<td>23 (18)</td>
<td>5 (10)</td>
<td>28</td>
</tr>
<tr>
<td>Exeter</td>
<td>7 (12)</td>
<td>11 (6)</td>
<td>18</td>
</tr>
<tr>
<td>total</td>
<td>30</td>
<td>16</td>
<td>46</td>
</tr>
</tbody>
</table>

Calculating the log likelihood statistic gives:

\[
G^2 = 2 \sum_{n_{ij}} n_{ij} \log \left( \frac{n_{ij}}{\mu_{ij}} \right)
\]

\[
= 2(23 \log \frac{23}{18} + 5 \log \frac{5}{10} + 7 \log \frac{7}{12} + 11 \log \frac{11}{6})
\]

\[
= 2(5.638 - 3.466 - 3.773 + 6.667)
\]

\[
= 2(5.066)
\]

\[
= 10.132
\]

Log likelihood statistic
How can we interpret this result? Remember:

$$G^2 = 2 \sum_{n_{ij}} n_{ij} \log \left( \frac{n_{ij}}{\mu_{ij}} \right) = 10.132$$
Log likelihood statistic

How can we interpret this result? Remember:

$$G^2 = 2 \sum_{n_{ij}} n_{ij} \log\left(\frac{n_{ij}}{\mu_{ij}}\right) = 10.132$$

- SPSS gives a value of $G^2 = 9.107$ and $p = 0.005$
- This looks awfully like....
Rounding up

Answering questions:

What is the relation between log odds ratio and (this formula of) log likelihood ratio?

Taking logs.

Can the value of $G^2$ be negative?

No (!)

Is the likelihood ratio value found in SPSS the same as the log likelihood ratio?

Yes.

Why are logs used in statistics?

To Logarithmity and Beyond!
Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio?
Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio? Taking logs.
Rounding up

Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio? Taking logs.
- Can the value of $G^2$ be negative?
Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio? Taking logs.
- Can the value of $G^2$ be negative? No (!)
Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio? Taking logs.
- Can the value of $G^2$ be negative? No (!)
- Is the “likelihood ratio” value found in SPSS the same as the “log likelihood ratio”?
Rounding up

Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio? Taking logs.
- Can the value of $G^2$ be negative? No (!)
- Is the “likelihood ratio” value found in SPSS the same as the “log likelihood ratio”? Yes.
Answering questions:

- What is the relation between log odds ratio and (this formula of) log likelihood ratio? Taking logs.
- Can the value of $G^2$ be negative? No (!)
- Is the “likelihood ratio” value found in SPSS the same as the “log likelihood ratio”? Yes.

Why are logs used in statistics?