To Logarithmity and Beyond!

Livi Ruffle

26th April 2006

Introduction

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.

Log Odds Ratios

A fictitious example:

Log Odds Ratios

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.

Log Odds Ratios

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 ↔ grockle/emmet; e.g. *"There be many grockles down from up country dis year"*
  - Dreckly ↔ soon/when possible; e.g. *"I'll come dreckly"*
  - Be ↔ [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

	dialect ($p^o_{1.}$)	SBE ($p^o_{.2}$)	total
Plymouth	23 (0.82)	5 (0.18)	28
Exeter	7 (0.39)	11 (0.61)	18
total	30	16	46

Log Odds Ratios

	dialect ($p^o_{1.}$)	SBE ($p^o_{.2}$)	total
Plymouth	23 (0.82)	5 (0.18)	28
Exeter	7 (0.39)	11 (0.61)	18
total	30	16	46

What is a *success*?

Log Odds Ratios

	dialect ($p^o_{1.}$)	SBE ($p^o_{.2}$)	total
Plymouth	23 (0.82)	5 (0.18)	28
Exeter	7 (0.39)	11 (0.61)	18
total	30	16	46

What is a *success*?

The *odds of success* for a particular row is given by

$$\frac{\text{probability of success}}{\text{probability of failure}} = \frac{p^o_{.2}}{p^o_{1.}}$$

Log Odds Ratios

So the *odds of sucess for Plymothians* is

$$\begin{aligned} O_{Ply} &= \frac{p^o_{12}}{p^o_{11}} \\ &= \frac{0.18}{0.82} \\ &= 0.22 \end{aligned}$$

That of Exonians is

$$\begin{aligned} O_{Ex} &= \frac{p^o_{22}}{p^o_{12}} \\ &= \frac{0.61}{0.39} \\ &= 1.56 \end{aligned}$$

(Almost there!)

Log Odds Ratios

The *log odds ratio with respect to Plymothians* (θ_{Ply}) is thus:

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

Log Odds Ratios

The *log odds ratio with respect to Plymothians* (θ_{Ply}) is thus:

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

But how do we interpret this ratio θ?

Log Odds Ratios

The *log odds ratio with respect to Plymothians* (θ_{Ply}) is thus:

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

But how do we interpret this ratio θ?
If θ_{Ply} were equal to 1, then the odds of success for both Plymouthians and Exonians would be equal.
If θ_{Ply} is *greater* than 1, Plymouthians would be more likely to succeed than Exonians.
If θ_{Ply} is *less* than 1, Exonians would be more likely to succeed than Plymouthians.

Log Odds Ratios

The *log odds ratio with respect to Plymothians* (θ_{Ply}) is thus:

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

But how do we interpret this ratio θ?
If θ_{Ply} were equal to 1, then the odds of success for both Plymouthians and Exonians would be equal.
If θ_{Ply} is *greater* than 1, Plymouthians would be more likely to succeed than Exonians.
If θ_{Ply} is *less* than 1, Exonians would be more likely to succeed than Plymouthians.
$\theta_{Ply} < 1$, so Plymouthians are (much) less likely to "succeed" than Exonians.

Log Odds Ratios: However!

- The odds ratio as it is now is *skewed*. Why?

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.

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 θ_{Ex} and $\log\theta_{Ex}$:

θ_{Ex}	7.09
$\log\theta_{Ex}$	1.96

Log likelihood statistic

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 μ_{ij} the *expected* frequency.

Log likelihood statistic

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 μ_{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"?

Log likelihood statistic

A reminder of our data:

	dialect (expected)	SBE (expected)	total
Plymouth	23 (18)	5 (10)	28
Exeter	7 (12)	11 (6)	18
total	30	16	46

Log likelihood statistic

A reminder of our data:

	dialect (expected)	SBE (expected)	total
Plymouth	23 (18)	5 (10)	28
Exeter	7 (12)	11 (6)	18
total	30	16	46

Calculating the log likelihood statistic gives:

$$\begin{aligned}
G^2 &= 2\sum_{n_{ij}} n_{ij}\log\left(\frac{n_{ij}}{\mu_{ij}}\right) \\
&= 2\left(23\log\frac{23}{18} + 5\log\frac{5}{10} + 7\log\frac{7}{12} + 11\log\frac{11}{6}\right) \\
&= 2(5.638 + -3.466 + -3.773 + 6.667) \\
&= 2(5.066) \\
&= 10.132
\end{aligned}$$

Log likelihood statistic

How can we interpret this result? Remember:

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

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:

Answering questions:

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

Rounding up

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?

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 (!)

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"?

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.

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?