Test of independence for two-way contingency tables

Application of log likelihood ratio to child acquisition data

Contents

- Odds ratio and log odds ratio
- Log likelihood and chi-square
- Application:

  acquisition of relative clause barrierhood in wh-questions and quantifiers

Open questions marked red!

Odds ratio

X / Y	Yes	No
Male	P1	(1-P1)
Female	P2	(1-P2)

Odds = P1/(1-P1)

Odds ratio (θ)= odds1/odds2 = $\dfrac{P1/(1-P1)}{P2/(1-P2)}$

Inference from odds ratio

Baseline for comparison = 1 = independent

$0 < \theta < 1$: success less likely in row 1 than in row 2

$1 < \theta < \partial$: success more likely in row 1 than in row 2

In general:

further away from 1 in any direction = stronger association /less independent

Log odds ratio

- Problem: small – moderate sample sizes odds ratio = skewed
- Solution: apply natural logarithm (log)

Odds ratio	**Log odds ratio**
1	0
2	0.7
0.5	-0.7

Significance test

- Log likelihood ratio (G^2)

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

Question: what is the relation between log odds ratio and (this formula of) log likelihood ratio?

Independence = 0: larger = less independent

Question: can the outcome also be –x?

P-value: estimates significance

Comparison X^2 and log likelihood

$$X^2 = \sum (n_{ij} - \mu_{ij})^2 / \mu_{ij}$$

$$G^2 = 2 \sum n_{ij} \log \left(n_{ij}/\mu_{ij}\right)$$

X2	- overestimates effect in large sample size
	- misses effect in small sample size
	- observations must be independent
Log odds ratio	- independent of sample size
	- invariant of marginal distribution
	- invariant of row/column order

Introduction study (1)

Relative clause barrierhood:

Hoe zei Kees [dat de jongen gevallen was]?
How did Kees say the boy fell?
→ LD (how – fell) and SD (how – say)

Hoe hielp de sterke Indiaan de Indiaan [die naar zijn wigwam ging]?
How did the strong Indian help the Indian who went to his wigwam?
→ only SD (how – help)

Introduction study (2)

Relative clause barrierhood:

Alle cowboys zitten op een paard.

Every cowboy is sitting on a horse.

→distributive or collective

Er is een paard [waar alle cowboys op zitten].

There is a horse that every cowboy is sitting on.

→collective only

Results

	Quantifiers		
WH	yes	no	**Total**
yes	9	7	**16**
no	0	4	**4**
Total	**9**	**11**	**20**

Application (1)

- Analyze – descriptive statitstics – crosstabs:

wh * quant Crosstabulation

			quant		Total
			ja	nee	
wh	ja	Count	9	7	16
		Expected Count	7,2	8,8	16,0
	nee	Count	0	4	4
		Expected Count	1,8	2,2	4,0
Total		Count	9	11	20
		Expected Count	9,0	11,0	20,0

Application (2)

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4,091[b]	1	,043		
Continuity Correction[a]	2,134	1	,144		
Likelihood Ratio	5,595	1	,018		
Fisher's Exact Test				,094	,068
N of Valid Cases	20				

a. Computed only for a 2x2 table

b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 1,80.

Question:

1. what exactly is likelihood ratio?
2. If not Log likelihood ratio then where to find log likelihood ratio in SPSS?

Application (3)

Observations adjusted:

wh * quant Crosstabulation

			quant		Total
			ja	nee	
wh	ja	Count	9	7	16
		Expected Count	7,6	8,4	16,0
	nee	Count	10	14	24
		Expected Count	11,4	12,6	24,0
Total		Count	19	21	40
		Expected Count	19,0	21,0	40,0

Application (4)

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,819[b]	1	,366	,520	,281
Continuity Correction[a]	,338	1	,561		
Likelihood Ratio	,820	1	,365	,520	,281
Fisher's Exact Test				,520	,281
N of Valid Cases	40				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 7,60.

Conclusions

- With large enough samples chi-square and log likelihood give same results

- Log likelihood is independent of sample size, marginal distribution and row/column order and therefore often more reliable