Introduction to Statistics for Clinical Trials

Proportions, chi-squared tests and odds ratios

Martin Bland

Professor of Health Statistics

University of York

http://www-users.york.ac.uk/~mb55/

Analyses for qualitative data

Also called nominal, categorical.

Only two categories: dichotomous, attribute, quantal, binary.

Methods:

- Chi-squared test for association
- Fisher's exact test
- Chi-squared test for trend
- Risk ratio, relative risk, rate ratio
- Odds ratio

Contingency tables

Cross tabulation of two categorical variables:

Time of delivery by housing tenure

Housing tenure	Premature	Term	Total
Owner-occupier	50	849	899
Council tenant	29	229	258
Private tenant	11	164	175
Lives with parents	6	66	72
Other	3	36	39
Total	99	1344	1443

This kind of crosstabulation of frequencies is also called a **contingency table** or **cross classification**.

Want to test the null hypothesis that there is no relationship or association between the two variables.

Contingency tables

Cross tabulation of two categorical variables:

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted	Rejected	Total
Married	71	415	486
Living w. partner	41	181	222
Single	15	35	50
Div./wid./sep.	7	23	30
Total	134	654	788

Meadows J, Jenkinson S, Catalan J. (1994) Who chooses to have the HIV antibody test in the antenatal clinic? *Midwifery* **10**, 44-48.

Contingency tables

Cross tabulation of two categorical variables:

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted	Rejected	Total
Married	71	415	486
Living w. partner	41	181	222
Single	15	35	50
Div./wid./sep.	7	23	30
Total	134	654	788

This kind of cross-tabulation of frequencies is also called a **contingency table** or **cross classification**.

Called 4 by 2 table or 4×2 table.

In general, $r \times c$ table.

Contingency tables

Cross tabulation of two categorical variables:

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted	Rejected	Total
Married	71	415	486
Living w. partner	41	181	222
Single	15	35	50
Div./wid./sep.	7	23	30
Total	134	654	788

Want to test the null hypothesis that there is no relationship or association between the two variables.

If the sample is large, we can do this by a chi-squared test.

If the sample is small, we must use Fisher's exact test.

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted	Rejected	Total
Married	71	415	486
Living w. partner	41	181	222
Single	15	35	50
Div./wid./sep.	7	23	30
Total	134	654	788

Null hypothesis: no association between the two variables.

Alternative hypothesis: an association of some type.

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted	Rejected	Total
Married	82.6		486
Living w. partner			222
Single			50
Div./wid./sep.			30
Total	134	654	788

Proportion who accepted = 134/788

Out of 486 married, expect 486 × 134/788 = 82.6
to accept if the null hypothesis were true.

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted	Rejected	Total
Married	82.6	403.4	486
Living w. partner			222
Single			50
Div./wid./sep.			30
Total	134	654	788

Proportion who refused = 654/788

Out of 486 married, expect 486 × 654/788 = 403.4
to refuse if the null hypothesis were true.

Note that 82.6 + 403.4 = 486.

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test Accepted	Rejected	Total
Married	82.6	403.4	486
Living w. partner	37.8	184.2	222
Single			50
Div./wid./sep.			30
Total	134	654	788

Out of 222 living with partner, expect 222 × 134/788 = 37.8 to accept if the null hypothesis were true.

Out of 222 living with partner, expect 222 × 654/788 = 184.2 to refuse if the null hypothesis were true.

Note that 37.8 + 184.2 = 222.

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test Accepted	Rejected	Total
Married	82.6	403.4	486
Living w. partner	37.8	184.2	222
Single	8.5	41.5	50
Div./wid./sep.	5.1	24.9	30
Total	134	654	788

Note that 82.6 + 37.8 + 8.5 + 5.1 = 134,

403.4 + 184.2 + 41.5 + 24.9 = 654.

Observed and expected frequencies have the same row and column totals.

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test Accepted	Rejected	Total
Married	82.6	403.4	486
Living w. partner	37.8	184.2	222
Single	8.5	41.5	50
Div./wid./sep.	5.1	24.9	30
Total	134	654	788

Expected frequency if null hypothesis true =

$$\frac{\text{row total} \times \text{column total}}{\text{grand total}}$$

The chi-squared test for association

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test: Accepted		Rejected		Total
Married	71	82.6	415	403.4	486
Living w. partner	41	37.8	181	184.2	222
Single	15	8.5	35	41.5	50
Div./wid./sep.	7	5.1	23	24.9	30
Total	134		654		788

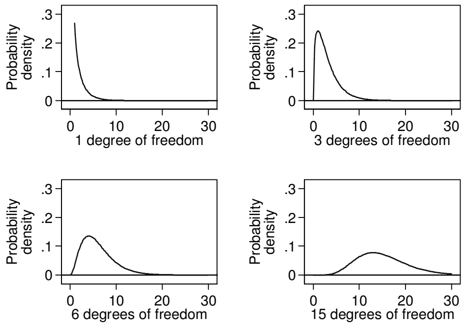
Compare the observed and expected frequencies.

Add (observed – expected)^2/expected for all cells = 9.15.

If null hypothesis true and samples are large enough, this is an observation from a chi squared distribution, often written χ^2.

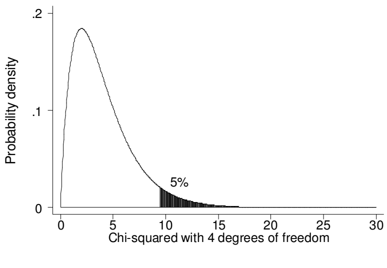
The Chi-squared distribution

Family of distributions, one parameter, called the **degrees of freedom**.



The Chi-squared distribution

Family of distributions, one parameter, called the **degrees of freedom**.



Percentage points of the Chi-squared Distribution

Degrees of freedom	Probability that the tabulated value is exceeded: 10% 0.10	5% 0.05	1% 0.01	0.1% 0.001
1	2.71	3.84	6.63	10.83
2	4.61	5.99	9.21	13.82
3	6.25	7.81	11.34	16.27
4	7.78	9.49	13.28	18.47
5	9.24	11.07	15.09	20.52
6	10.64	12.59	16.81	22.46
7	12.02	14.07	18.48	24.32
8	13.36	15.51	20.09	26.13
9	14.68	16.92	21.67	27.88
10	15.99	18.31	23.21	29.59
.	.	.	.	.
.	.	.	.	.

The chi-squared test for association

Time of delivery by housing tenure

Housing tenure	Premature	Term	Total
Owner-occupier	50	849	899
Council tenant	29	229	258
Private tenant	11	164	175
Lives with parents	6	66	72
Other	3	36	39
Total	99	1344	1443

For a contingency table, the degrees of freedom are given by:

(number of rows – 1) × (number of columns – 1).

We have (5 – 1) × (2 – 1) = 4 degrees of freedom.

$\chi^2 = 10.5$, 4 d.f., $P < 0.05$. Using a computer, $P = 0.03$.

The chi-squared test for association

The chi-squared statistic is not an index of the strength of the association.

If we double the frequencies, this will double chi-squared, but the strength of the association is unchanged.

The chi-squared test for association

The test statistic follows the Chi-squared Distribution provided the expected values are large enough.

This is a large sample test.

The smaller the expected values become, the more dubious will be the test.

The conventional criterion for the test to be valid is this: the chi-squared test is valid if at least 80% of the expected frequencies exceed 5 and all the expected frequencies exceed 1.

Also known as the **Pearson chi-squared test**.

Fisher's exact test

Also called the **Fisher-Irwin exact test**.

Works for any sample size.

Used to be used only for small samples in 2 by 2 tables, because of computing problems.

Calculate the probability of every possible table with the given row and column totals.

Sum the probabilities for all the tables as or less probable than the observed.

Fisher's exact test

Acceptance of HIV test grouped by marital status

Marital status	Acceptance of HIV test Accepted	Rejected	Total
Married	71	415	486
Living w. partner	41	181	222
Single	15	35	50
Div./wid./sep.	7	23	30
Total	134	654	788

χ^2 = 9.15, 3 d.f., P = 0.027.

Fishers' exact test: P = 0.029.

Yates' correction

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

(Callam *et al.*, 1992)

Fisher's exact test: P = 0.0049.

Chi-squared test: chi-squared = 8.87, P = 0.0029.

Callam MJ, Harper DR, Dale JJ, Brown D, Gibson B, Prescott RJ, Ruckley CV. (1992) Lothian Forth Valley leg ulcer healing trial—part 1: elastic versus non-elastic bandaging in the treatment of chronic leg ulceration. *Phlebology* 7: 136-41.

Yates' correction

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

(Callam *et al.*, 1992)

Fisher's exact test: P = 0.0049.

Chi-squared test: chi-squared = 8.87, P = 0.0029.

As expected frequencies get smaller, chi-squared and Fisher's exact disagree.

Fisher's produces the 'correct' P value.

Chi-squared produces a P value which is too small.

Yates' correction

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

(Callam *et al.*, 1992)

Fisher's exact test: P = 0.0049.

Chi-squared test: chi-squared = 8.87, P = 0.0029.

Yates introduced a modified chi-squared test for a 2 by 2 table which adjusts for this.

Also called the **continuity correction**.

Yates' correction

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

(Callam *et al.*, 1992)

Fisher's exact test: P = 0.0049.

Chi-squared test: chi-squared = 8.87, P = 0.0029.

Chi-squared with Yates' correction:
chi-squared = 7.84, P = 0.0051.

Yates' correction now obsolete as we can always do the exact test.

The chi-squared test for trend

Assessment of radiological appearance at six months as compared with appearance on admission (MRC 1948)

Radiological assessment	Streptomycin	Control
Considerable improvement	28	4
Moderate or slight improvement	10	13
No material change	2	3
Moderate or slight deterioration	5	12
Considerable deterioration	6	6
Deaths	4	14
Total	55	52

Association: chi-squared = 26.97, 5 d.f., P = 0.0001.

Does not take the ordering of the categories into account.

Trend: chi-squared = 17.93, 1 d.f., P < 0.0001.

About trend: chi-squared = 9.04, 4 d.f., P = 0.06.

Risk ratio

Wound healing by type of bandage

Bandage	Healed		Did not heal		Total	
Elastic	35	53.8%	30	46.2%	65	100%
Inelastic	19	28.4%	48	71.6%	67	100%
Total	54		78		132	

Want an estimate of the size of the treatment effect.

Difference between proportions: 0.538 – 0.284 = 0.254
or 53.8% – 28.4% = 25.4 percentage points.

Proportion who heal is called the **risk** of healing for that population.

Risk ratio = 53.8/28.4 = 1.89.

Also called **relative risk**, **rate ratio**, **RR**.

Risk ratio

Wound healing by type of bandage

Bandage	Healed		Did not heal		Total	
Elastic	35	53.8%	30	46.2%	65	100%
Inelastic	19	28.4%	48	71.6%	67	100%
Total	54		78		132	

Risk ratio = 53.8/28.4 = 1.89.

Because risk ratio is a ratio, it has a very awkward distribution.

If we take the log of the risk ratio, we have something which is found by adding and subtracting log frequencies.

The distribution becomes approximately Normal.

Provided frequencies are not small, simple standard error.

Risk ratio

Wound healing by type of bandage

Bandage	Healed		Did not heal		Total	
Elastic	35	53.8%	30	46.2%	65	100%
Inelastic	19	28.4%	48	71.6%	67	100%
Total	54		78		132	

Risk ratio, RR = 53.8/28.4 = 1.89.

$\log_e$(RR) = 0.6412.

SE for $\log_e$(RR) = 0.2256.

95% CI for $\log_e$(RR)
= 0.6412 – 1.96×0.2256 to 0.6412 + 1.96 ×0.2256
= 0.1990 to 1.0834.

95% CI for RR = exp(0.1990) to exp(1.0834) = 1.22 to 2.95.

Risk ratio

Wound healing by type of bandage

Bandage	Healed		Did not heal		Total	
Elastic	35	53.8%	30	46.2%	65	100%
Inelastic	19	28.4%	48	71.6%	67	100%
Total	54		78		132	

$\log_e$(RR) = 0.6412, 95% CI = 0.1990 to 1.0834.

Risk ratio, RR = 53.8/28.4 = 1.89, 95% CI = 1.22 to 2.95.

RR is not in the middle of its confidence interval.

The interval is symmetrical on the log scale, not the natural scale.

Odds

	Healed	Did not heal	Total
Elastic	35 53.8%	30 46.2%	65 100%

Risk of healing = 35/65 = 0.538

Odds of healing = 35/30 = 1.17

Risk = number experiencing event divided by number who could.

Odds = number experiencing event divided by number who did not experience event.

Risk: for every person treated, 0.538 people heal,
for every 100 people treated, 53.8 people heal.

Odds: for every person who do not heal, 1.17 people heal,
for every 100 people who do not heal, 117 people heal.

Odds ratio

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

Odds of healing given elastic bandages: 35/30 = 1.17.

Odds of healing given inelastic bandages: 19/48 = 0.40.

Odds ratio = (35/30)/(19/48) = 1.17/0.40 = 2.95.

For every person who does not heal, 2.95 times as many will heal with elastic bandages as will heal with inelastic bandages.

Odds ratio

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

Odds ratio, OR = (35/30)/(19/48) = 2.95.

Like RR, OR has an awkward distribution. We use the log odds ratio.

The distribution becomes approximately Normal.

Provided frequencies are not small, simple standard error.

Odds ratio

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

Odds ratio, OR = (35/30)/(19/48) = 2.95.

$\log_e(\text{OR}) = 1.0809$.

SE $\log_e(\text{OR}) = 0.3679$

95% CI for $\log_e(\text{OR})$
= 1.0809 – 1.96 × 0.3679 to 1.0809 + 1.96 × 0.3679
= 0.3598 to 1.8020.

95% CI for OR = exp(0.3598) to exp(1.8020) = 1.43 to 6.06.

Odds ratio

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

$\log_e(\text{OR}) = 1.0809$, 95% CI = 0.3598 to 1.8020.

Odds ratio, OR = 2.95, 95% CI = 1.43 to 6.06.

OR is not in the middle of its confidence interval.

The interval is symmetrical on the log scale, not the natural scale.

Odds ratio

Wound healing by type of bandage

Bandage	Healed	Did not heal	Total
Elastic	35	30	65
Inelastic	19	48	67
Total	54	78	132

Odds ratio for healing: OR = (35/30)/(19/48) = 2.95.

Doesn’t matter which way round we do it.

Odds ratio for treatment: OR = (35/19)/(30/48) = 2.95.

Both OR = (35×48)/(30 ×19).

Ratio of cross products.

Odds ratio

Wound healing by type of bandage

Bandage	Did not heal	Healed	Total
Elastic	30	35	65
Inelastic	48	19	67
Total	78	54	132

Switching the rows or columns inverts the odds ratio.

Odds ratio for not healing given elastic bandage:
OR = (30/35)/(48/19) = 0.339 = 1/2.95.

There are only two possible odds ratios.

On the log scale, equal and opposite.

$\log_e(2.95) = 1.082$, $\log_e(0.339) = -1.082$.