TWO-TAILS APPROXIMATE CONFIDENCE INTERVALS FOR THE RATIO OF PROPORTIONS.

A. Martín Andrés and M. Álvarez Hernández.

Biostatistics. Faculty of Medicine. University of Granada.

INTRODUCTION

Comparing two independent binomial proportions:

- Difference: $d=p_2-p_1$ (has received much attention)
- Ratio: $R=p_2/p_1$
- Odds-ratio: $O=p_2 q_1/p_1 q_2$

→ two-tails approximate inferences about R.

- Exact point of view: computationally very intensive.
  not feasible for large sample size.
- Approximate point of view: researchers have devoted great attention

Objective:

- Propose new approximate methods.
- Comparison between new and classic methods from the literature.

CONFIDENCE INTERVALS AND HYPOTHESIS TESTS

Agresti & Min (2001):

Obtaining the two-tailed exact CI through the inversion of the two-tailed test $H: R=\rho$

Statistical inference coherent:

1. Perspective of the test or perspective of the CI.
2. Evaluating a CI method is equivalent to evaluating its associated test method (to the same nominal error α).

* CI → Real coverage and average length.	Coverage + Error =1.
* Test → Real error and power.	Greater power, lower length.

Consequently:
The comparative evaluation will be made with reference to the test that defines them.

PROCEDURES BASED ON THE Z STATISTIC

Let $x_i \sim B(n_i, p_i)$ two independent binomial random variables.

Let $\overline{R} = \overline{p}_2 / \overline{p}_1$ sample estimator of R with $\overline{p}_i = x_i / n_i$.

To contrast H: $R=\rho$ vs. K: $R\neq \rho$ (where $0< \rho <+\infty$), the most common is:

Z statistic by Katz et al (1978): $$z_Z^2 = \frac{\left(\overline{p}_2 - \rho \overline{p}_1\right)^2}{\rho^2 \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

It is necessary to estimate the unknown proportions p_i.

- <u>Test:</u> comparing the value z_{exp}^2 which is obtained with $z_{\alpha/2}^2$
- <u>Confidence interval:</u> solving through ρ the equation $z_{exp}^2 = z_{\alpha/2}^2$

ESTIMATORS

- Classic estimator:

$p_{iW} = \overline{p}_i = x_i / n_i$ (estimator **W**) → ZW (Wald procedure).

- Estimator of Newcombe:

$$\begin{aligned} p_{1N} &= u_1, p_{2N} = l_2 \quad \text{if} \quad \overline{R} > \rho \\ p_{1N} &= l_1, p_{2N} = u_2 \quad \text{if} \quad \overline{R} < \rho \end{aligned}$$

(estimator **N**) → ZN procedure.

$$\left(l_i; u_i\right) = \frac{x_i + \frac{z_{\alpha/2}^2}{2} \pm z_{\alpha/2}\sqrt{\frac{z_{\alpha/2}^2}{4} + \frac{x_i y_i}{n_i}}}{n_i + z_{\alpha/2}^2}$$

CI of p_i by Wilson (1927).

Estimators restricted by H: $p_2 = \rho p_1$ only parameter to be estimated is p_1.

- Conditioned point of view ($x_1+x_2=a_1$):

$$p_{1C} = Min\left\{1; \frac{a_1}{n_1 + n_2\rho}\right\}, p_{2C} = Min\left\{1; \frac{\rho a_1}{n_1 + n_2\rho}\right\}$$

(estimator **C**)

→ ZC procedure by Farrington & Manning (1990)

- Unconditioned point of view:

$$p_{1E}=\frac{(n_1+x_2)+(n_2+x_1)\rho-\sqrt{\{(n_1+x_2)+(n_2+x_1)\rho\}^2-4na_1\rho}}{2n\rho},\ p_{2E}=p_{1E}\rho \quad \text{(estimator } \mathbf{E}\text{)}$$

→ ZE procedure (score method) by Koopman (1984).

CI can be obtained through a cubic equation (Nam, 1995).

- Unconditioned approximately estimator:

$$p_{1A}=Min\left\{1;\frac{x_2+x_1\rho}{n\rho}\right\},\ p_{2A}=Min\left\{1;\frac{x_2+x_1\rho}{n}\right\}$$ (estimator **A**) → ZA procedure.

- Pekun estimator:

$$p_{1P}=Min\left\{1;\frac{n_1+n_2\rho}{2n\rho}\right\},\ p_{2P}=Min\left\{1;\frac{n_1+n_2\rho}{2n}\right\}$$ (estimator **P**)

based on the criteria of Sterne (1954): z_Z^2 will be significant when it is for any value of p_1.

→ ZP procedure by Martín & Herranz (2010).

PROCEDURES BASED ON THE L STATISTIC

Another quite common is:

L statistic by Woolf (1955): $$z_L^2 = \frac{ln^2\left(\bar{R}/\rho\right)}{\dfrac{q_1}{n_1 p_1} + \dfrac{q_2}{n_2 p_2}}$$

Once again we have to estimate the values of p_i.

- Classic estimator was proposed by Woolf → LW procedure.
- Proceeding in a similar way with: N, C, E and A

→ Procedures LN, LC, LE and LA (LC and LE were proposed by Martín & Herranz (2010)

CI of LC, LE and LA are obtained through iterative methods.

PROCEDURES BASED ON THE A STATISTIC

Herranz & Martín (2008), in the context of the case of the difference:

$$\textbf{A}\text{ statistic: } z_A^2 = \frac{4n_1n_2\left(\overline{d}' - \delta'\right)^2}{n_1 + n_2} \qquad \begin{cases} \overline{d}' = sin^{-1}\sqrt{\overline{p}_2} - sin^{-1}\sqrt{\overline{p}_1} \\ \delta' = sin^{-1}\sqrt{p_2} - sin^{-1}\sqrt{p_1} \end{cases}$$

Proceeding as in the previous sections: C, E and A estimator

→ Procedures AC, AE and AA.

Inferences are derived from classic mode.

SAMPLE DATA TO BE USED

LW procedure performs badly (Woolf, 1955; Koopman, 1984).

The traditional improvement: original data increased by a quantity h_i .

Case 0: h_i=0.

Case 1: h_i=0.5 (Woolf, 1955).

Case 2: h_i=1 (Dann & Koch, 2005).

ZW procedure also performs very badly (Katz *et al.*, 1978), the same increases can also be applied to it.

Other possibilities in a more general text (Martín *et al.*, 2010)

Case 3: $h_i = z_{\alpha/2}^2 \big/ 4$

Case 4:
$$h_i = \begin{cases} \dfrac{z_{\alpha/2}^2\left(1+2I_i\right)}{4} \text{ where } I_1 = \begin{cases} 1 \text{ si } \overline{p}_1 = 0 \\ 0 \text{ si } \overline{p}_1 \neq 0 \end{cases}, \ I_2 = \begin{cases} 1 \text{ si } \overline{p}_2 = 1 \\ 0 \text{ si } \overline{p}_2 \neq 1 \end{cases} & \text{if } \overline{R} > \rho \\ \dfrac{z_{\alpha/2}^2\left(1+2S_i\right)}{4} \text{ where } S_1 = \begin{cases} 1 \text{ si } \overline{p}_1 = 1 \\ 0 \text{ si } \overline{p}_1 \neq 1 \end{cases}, \ S_2 = \begin{cases} 1 \text{ si } \overline{p}_2 = 0 \\ 0 \text{ si } \overline{p}_2 \neq 0 \end{cases} & \text{if } \overline{R} < \rho \end{cases}$$

Case 5: h_i=3/8 (Anscombe transformation).

PROCEDURE TO OBTAIN THE RESULTS

1. Selecting one of the errors α=1%, 5% or 10%.
2. Selecting one of the values ρ=0.01, 0.1, 0.2, 0.5, 0.8, 1, 1.25, 2, 5, 10 and 100.
3. Selecting one of the pairs (n_1, n_2) with $n_1 \le n_2$ and n_i=40, 60, 100.

   H: $R=\rho$ and H': $1/R = 1/\rho$ are equivalent.
4. Constructing the critical region (CR): $RC = \left\{ \left(x_1, x_2 \right) | \, z_{exp}^2 \geq z_{\alpha/2}^2 \right\}$
5. Calculating the real error (test size) $\alpha^*=\text{máx}_p \sum_{CR} \text{P}(x_1, x_2 | H)$ with

$$P\left(x_1, x_2 \,|\, H\right) = C\left(n_1, x_1\right) \times C\left(n_2, x_2\right) \rho^{x_2} p^{a_1} \left(1-p\right)^{y_1} \left(1-\rho p\right)^{y_2}$$

   and the increase $\Delta\alpha=\alpha-\alpha^*$.
6. Calculating the value of "power":

$$\theta=100\times(\text{nº of points of the CR set})/[(n_1+1)(n_2+1)]\ \%$$

7. Determining if the method "fails": $\Delta\alpha \leq -1\%$, -2% o -4% for $\alpha=1\%$, 5% or 10%.
8. Calculating the total number of failures (F) and the average values of $\Delta\alpha$ and of θ for $0.1 \leq \rho \leq 10$ on the one hand, and for $\rho=0.01$ and 100 on the other hand.

ANALYZE THE RESULTS

a) Reject methods with an excessive number of failures.
b) Choose those which have a $\overline{\Delta\alpha}$ closest to 0, showing a preference for conservative methods ($\overline{\Delta\alpha}>0$).
c) Prefer those with the greatest $\bar{\theta}$
d) Prefer the method that is the most simple to apply.

SELECTION OF THE OPTIMA METHOD

Methods selected and proposed for the literature (α=5%)

$0.1 \leq \rho \leq 10$				ρ=0.01 and ρ=100			
Method	*F*	$\overline{\Delta\alpha}$	$\bar{\theta}$	*Method*	*F*	$\overline{\Delta\alpha}$	$\bar{\theta}$
ZW4	0	0.06	79.71	ZA1	1	2.34	95.70
LE3	0	0.06	79.69	AE1	0	3.55	95.00
ZA1	0	-0.27	80.50	AE5	0	4.10	94.79
LC3	1	-0.21	80.07	ZW4	0	4.74	93.82
AE5	0	0.43	79.48	LC3	0	4.98	86.84
AE1	0	0.73	79.41	LE3	0	5.00	84.28
ZP0	0	1.47	76.04	ZP0	0	5.00	87.91
LW0	9	-0.82	77.45	LW0	11	-3.49	95.16
ZC0	16	-1.17	80.84	ZC0	12	-6.03	96.57
ZE0	22	-1.42	80.81	ZE0	12	-6.28	97.00
LW1	19	-1.56	79.90	LW1	12	-39.08	97.56
LW2	21	-3.24	80.07	ZW0	12	-84.84	93.86
ZW0	44	-16.91	80.61	LC0	12	-89.44	87.03
LE0	54	-54.72	80.32	LE0	12	-89.44	84.39
LC0	54	-54.73	80.69	LW2	12	-91.67	98.42

Methods selected: ZW4, LE3, ZA1, LC3, AE5, AE1.

Methods proposed for the literature: ZP0, LW0, ZC0, ZE0, LW1, LW2, ZW0, LE0, LC0.

- The methods proposed for the literature have a lot of failures, or in the case of ZP0, it is very conservative method and has a little power.
- All of the methods selected are reliable as they have at most one failure.
- Methods AE1 and AE5 (for a moderate ρ) and LE3 and LC3 (for an extreme ρ) can be rejected: the most conservative and/or the lowest power .
- From the rest: for a moderate ρ is preferable ZW4, for a extreme ρ the best performance is ZA1 (it’s the least conservative and the highest power).

Regarding the previous statement (and the results obtained for 1% and 10%):

- ❑ The best method in general is ZA1
- ❑ A good alternative is ZW4 (simpler method)

EXAMPLE

MAXWELL (1961)

Infection		YES	NO	Total
Inoculated	YES	11	35	46
	NO	48	54	102
Total		59	89	148

Sample estimation: $\overline{R} = (48/102)/(11/46) = 1{,}97$

Confidence intervals (95%) for *R*:

ZA1 → (1,1659; 3,5082)

ZW4 → (1,1887; 3,7853)

ZE0 → (1,1768; 3,4976)

LW1 → (1,1187; 3,3104)

ZE0 exact → (1,1705; 3,6164)

CONCLUSIONS

In Medicine, it's common a two-tailed confidence interval for the ratio *R*.

Several methods have been evaluated, obtaining this conclusions:

- None of the classic methods (including ZE0 score method) are reliable as they are excessively liberal.
- The best of them is LW1 which is only valid for large sample and moderate ρ.

LW1: 1) Increasing all the data from both groups (successes and failures) in 0.5.

2) Using the statistic based on logarithmic transformation:

$$z_{LW}^2 = \frac{ln^2\left(\bar{R}/\rho\right)}{\frac{\bar{q}_1}{n_1\bar{p}_1}+\frac{\bar{q}_2}{n_2\bar{p}_2}} = \frac{ln^2\left(\bar{R}/\rho\right)}{\frac{1}{x_1}+\frac{1}{x_2}-\frac{n}{n_1 n_2}} \quad \text{and} \quad R \in \bar{R} \times exp\left\{\pm z_{\alpha/2}\sqrt{\frac{y_1}{n_1 x_1}+\frac{y_2}{n_2 x_2}}\right\}$$

- In general, the best method is **ZA1:**

1) Increasing all the data from both groups (successes and failures) in 0.5.

2) Using the approximation of the ZE0 statistic, the test is given by:

$$z_{ZA}^2 = \frac{\left(\overline{p}_2 - \rho \overline{p}_1\right)^2}{\rho^2 \frac{p_{1A}\left(1-p_{1A}\right)}{n_1} + \frac{p_{2A}\left(1-p_{2A}\right)}{n_2}} \quad \text{with } p_{1A} = Min\left\{1; \frac{x_2 + x_1\rho}{n\rho}\right\}, p_{2A} = Min\left\{1; \frac{x_2 + x_1\rho}{n}\right\}$$

3) If the objective is the CI, solve:

$$\left(\rho_L, \rho_U\right) = \frac{n x_1 x_2 + \frac{z_{\alpha/2}^2\left(n_1 x_1 + n_2 x_2 - 2x_1 x_2\right)}{2} \pm z_{\alpha/2}\sqrt{n^2 x_1 x_2\left(a_1 - n\overline{p}_1\overline{p}_2\right) + \left\{\frac{z_{\alpha/2}\left(n_2 x_2 - n_1 x_1\right)}{2}\right\}^2}}{x_1\left\{n n_2 \overline{p}_1 - z_{\alpha/2}^2\left(n_2 - x_1\right)\right\}}$$

the two solutions (ρ_L, ρ_U) must verify that $x_2/\left(n-x_1\right) \le \rho_L, \rho_U \le \left(n-x_2\right)/x_1$,

If the boundary that fails is ρ_L, obtain the value:

$$\rho_L = \frac{1}{n_2\overline{p}_1^2 + z_{\alpha/2}^2}\left\{x_2\overline{p}_1 + \frac{z_{\alpha/2}^2}{2} - z_\alpha\sqrt{\frac{z_{\alpha/2}^2}{4} + x_2\left(\overline{p}_1 - \overline{p}_2\right)}\right\}$$

If the boundary that fails is ρ_U, obtain the value:

$$\rho_U \in \frac{1}{n_1\overline{p}_1^2}\left\{x_1\overline{p}_2 + \frac{z_{\alpha/2}^2}{2} + z_\alpha\sqrt{\frac{z_{\alpha/2}^2}{4} + x_1\left(\overline{p}_2 - \overline{p}_1\right)}\right\}$$

- Alternative, we can use the even simpler method **ZW4**:

1) Increasing all the data from both groups (successes and failures) in

$$h_i = \begin{cases} \dfrac{z_{\alpha/2}^2 (1+2I_i)}{4} & \text{with } I_1 = \begin{cases} 1 & \text{if } \bar{p}_1 = 0 \\ 0 & \text{if } \bar{p}_1 \neq 0 \end{cases}, I_2 = \begin{cases} 1 & \text{if } \bar{p}_2 = 1 \\ 0 & \text{if } \bar{p}_2 \neq 1 \end{cases} \text{ if } \bar{R} > \rho \\ \dfrac{z_{\alpha/2}^2 (1+2S_i)}{4} & \text{with } S_1 = \begin{cases} 1 & \text{if } \bar{p}_1 = 1 \\ 0 & \text{if } \bar{p}_1 \neq 1 \end{cases}, S_2 = \begin{cases} 1 & \text{if } \bar{p}_2 = 0 \\ 0 & \text{if } \bar{p}_2 \neq 0 \end{cases} \text{ if } \bar{R} < \rho \end{cases}$$

2) Using the classic Wald procedure:

$$z_{ZW}^2 = \frac{(\bar{p}_2 - \rho \bar{p}_1)^2}{\rho^2 \dfrac{\bar{p}_1 \bar{q}_1}{n_1} + \dfrac{\bar{p}_2 \bar{q}_2}{n_2}} \quad \text{and} \quad R \in \frac{\bar{R}}{1 - z_{\alpha/2}^2 \dfrac{y_1}{n_1 x_1}} \left\{ 1 \pm z_{\alpha/2} \sqrt{\frac{y_1}{n_1 x_1} + \frac{y_2}{n_2 x_2} - z_{\alpha/2}^2 \frac{y_1}{n_1 x_1} \frac{y_2}{n_2 x_2}} \right\}$$

REFERENCES

Agresti, A. (2003). Dealing with discreteness: making 'exact' confidence intervals for proportions, differences of proportions, and odds ratios more exact. Statistical Methods in Medical Research 12, 3-21.

Agresti, A. and Min, Y. (2001). On small-sample confidence intervals for parameters in discrete distributions. Biometrics 57, 963-971.

Barnard, G.A. (1947). Significance tests for 2×2 tables. Biometrika 34, 123-138.

Chan, I.S.F. (1998). Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies. Statistics in Medicine 17, 1403-1413.

Dann, R.S. and Koch, G.G. (2005). Review and evaluation of methods for computing confidence intervals for the ratio of two proportions and considerations for non-inferiority clinical trials. Journal of Biopharmaceutical Statistics 15, 85-107.

Farrington C.P. and Manning, G. (1990). Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. Statistics in Medicine 9, 1447-1454.

Gart, J.J. and Nam, J. (1990). Approximate interval estimation of the difference in binomial parameters: Correction for skewness and extension to multiple tables. Biometrics 46, 637- 643.

Herranz Tejedor, I. and Martín Andrés, A. (2008). A numerical comparison of several unconditional exact tests in problems of equivalence based on the difference of proportions. Journal of Statistical Computation and Simulation 78 (11), 969-981.

Katz, D.; Baptista, J.; Azen, S.P. and Pike, M.C. (1978). Obtaining confidence intervals for the risk ratio in cohort studies. Biometrics 34, 469-474.

Koopman, P.A.R. (1984). Confidence intervals for the ratio of two binomial proportions. Biometrics 40, 513-517.

Martín Andrés, A. and Herranz Tejedor, I. (2010). Asymptotic inferences about a linear combination of two proportions. JP Journal of Biostatistics 4(3), 253-277.

Martín Andrés, A., Álvarez Hernández, M. and Herranz Tejedor, I. (2010). Inferences about a linear combination of proportions. Statistical Methods in Medical Research. Prepublished March 11, 2010,.

Miettinen, O. and Nurminen, M. (1985). Comparative analysis of two rates. Statistics in Medicine 4, 213-226.

Nam, Jun-Mo (1995). Confidence limits for the ratio of two binomial proportions based on likelihood scores: Non-iterative method. Biometrical Journal 37(3), 375-379.

Newcombe, R.G. (1998). Interval estimation for the difference between independent proportions: comparison of eleven methods. Statistics in Medicine 17, 873-890.

Peskun, P.H. (1993). A new confidence interval method based on the normal approximation for the difference of two binomial probabilities. Journal of the American Statistical Association 88 (422), 656-661.

Price, R. M. and Bonett, D. G. (2008). Confidence intervals for a ratio of two independent binomial proportions. Statistics in Medicine 27, 5497-5508.

Sterne, T.E. (1954). Some remarks on confidence of fiducial limits. Biometrika 41 (1/2), 275-278.

Wilson, E.B. (1927). Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association 22, 209- 212.

Woolf, B. (1955). On estimating the relation between blood group and disease. Annals of Human Genetics 19 (4), 251-352.

Zou, G. and Donner, A. (2008). Construction of confidence limits about effect measures: A general approach. Statistics in Medicine 27, 1693-1702.