

A Note on Improving Inference of Relative Risk

Octavia C. Y. Wong

School of Kinesiology and Health Science, York University, Toronto, Canada

Email: owong3@mymail.yorku.ca

How to cite this paper: Wong, O.C.Y. (2019) A Note on Improving Inference of Relative Risk. *Open Journal of Statistics*, 9, 100-108.

<https://doi.org/10.4236/ojs.2019.91009>

Received: January 11, 2019

Accepted: February 16, 2019

Published: February 19, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Relative risk is a popular measure to compare risk of an outcome in the exposed group to the unexposed group. By applying the delta method and Central Limit Theorem, [1] derives two approximate confidence intervals for the relative risk, and [2] approximates the confidence interval for the relative risk via the likelihood ratio statistic. Both of these approximations require sample size to be large. In this paper, by adjusting the likelihood ratio statistic obtained by [2], a new method is proposed to obtain the confidence interval for the relative risk. Simulation results showed that the proposed method is extremely accurate even when the sample size is small.

Keywords

Bartlett Correction, Confidence Intervals, Coverage Property, Relative Risk

1. Introduction

Consider two groups of subjects: exposure group ($i = 1$) and control group ($i = 2$). Let n_i be the number of subjects in group i with p_i being the risk of a specific outcome in group i . Then the random variable X_i , which is the number of subjects that give the specific outcome in group i , is distributed as Binomial (n_i, p_i). As defined in [3] and [4], the relative risk of the outcome in the exposure group versus the control group is $\theta = p_1/p_2$. Note that θ can be any nonnegative real number. When $\theta < 1$, it suggests that the exposure being considered is associated with a reduction in risk, and $\theta > 1$ suggests that the exposure is associated with an increase in risk. $\theta = 1$ is generally of interest because it suggests that the exposure has no impact on risk. In general, p_i is unknown, but we observed x_i . Then p_i can be estimated by $\hat{p}_i = x_i/n_i$. And, therefore, an estimate of the relative risk based on the observed sample is $\hat{\theta} = \hat{p}_1/\hat{p}_2$.

Relative risk is a popular measure used in biomedical studies because it is easy to compute and interpret, and it is included in standard statistical software

output (e.g., in *R* and *SAS*). [5] gives a detailed discussion on the application of relative risk to failure time data. [6] applies relative risk to study populations with differing disease prevalence. [7] compares relative risk with odds ratio, and absolute risk reduction in comparing the effectiveness of certain treatments.

To illustrate the concept of relative risk, let us consider the following example. [1] examined the Physicians' Health Study, which analyzed whether taking aspirin regularly will reduce cardiovascular disease. Data of the study are reported in **Table 1**.

Out of 11,037 physicians taking aspirin over the course of the study, 104 of them had heart attacks. Similarly 189 of 11,034 physicians in the placebo group had heart attacks. Based on this dataset, the relative risk of having heart attacks among physicians is

$$\hat{\theta} = \frac{(5+99)/(5+99+10933)}{(18+171)/(18+171+10845)} = 0.55.$$

Thus, physicians who took aspirin over the course of the study have 0.55 times the risk of having a heart attack as physicians who were in the placebo group. This suggests that taking aspirin is associated with a reduction in the risk of heart attacks among physicians as they are about half as likely to have a heart attack as physicians who did not take aspirin throughout the study.

Although reporting a point estimate of relative risk is important, it does not provide information about the variations arising from the observed data. Hence, in practice, a $(1-\alpha)100\%$ confidence interval for θ is usually reported and recommended (see [8]). A standard approximated $(1-\alpha)100\%$ confidence interval for θ is given in [1], which is widely implemented in statistical software. [2] proposed an alternate way of approximating a $(1-\alpha)100\%$ confidence interval for θ via the likelihood ratio statistic. It is well-known that both methods are not accurate when the sample size is small. In this paper, by adjusting the likelihood ratio statistic obtained by [2], a new method is proposed to obtain the confidence interval for the relative risk. Simulation results show that the proposed method is extremely accurate even when the sample size is small.

2. Methodology

Let X_i , $i = 1, 2$, be independent random variables distributed as Binomial (n_i, p_i) . Then the relative risk is defined as $\theta = p_1/p_2$. A standard estimator of θ is $\hat{\Theta} = (X_1/n_1)/(X_2/n_2)$. With realizations x_1 and x_2 , a standard estimate of θ is $\hat{\theta} = (x_1/n_1)/(x_2/n_2)$. [1] considered the parameter $\psi = \ln \theta$. The corresponding estimator of ψ is $\hat{\Psi} = \ln \hat{\Theta}$. By applying the delta method, we have

Table 1. Cross-classification of aspirin use and heart attack.

Group	Heart Attack		
	Fatal	Non-fatal	No attack
Placebo	18	171	10,845
Aspirin	5	99	10,933

$$E(\hat{\Psi}) = E(\ln \hat{\Theta}) \approx \ln \theta = \psi \quad \text{and} \quad \text{var}(\hat{\Psi}) = \text{var}(\ln \hat{\Theta}) \approx \sum_{i=1}^2 \left(\frac{1}{n_i p_i} - \frac{1}{n_i} \right).$$

Therefore, an estimate of ψ is $\hat{\psi} = \ln \hat{\theta}$, and the estimated variance of $\hat{\Psi}$ is

$$\widehat{\text{var}}(\hat{\Psi}) \approx \sum_{i=1}^2 \left(\frac{1}{n_i \hat{p}_i} - \frac{1}{n_i} \right) = \sum_{i=1}^2 \left(\frac{1}{x_i} - \frac{1}{n_i} \right).$$

Hence, when n_1 and n_2 are large, by the Central Limit Theorem, an approximate $(1-\alpha)100\%$ confidence interval for $\psi = \ln \theta$ is:

$$\left(\ln \hat{\theta} - z_{\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\Psi})}, \ln \hat{\theta} + z_{\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\Psi})} \right)$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)100^{\text{th}}$ percentile of the standard normal distribution. Since ψ and θ are one-one correspondence, we have an approximate $(1-\alpha)100\%$ confidence interval for θ is

$$\left(e^{\ln \hat{\theta} - z_{\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\Psi})}}, e^{\ln \hat{\theta} + z_{\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\Psi})}} \right).$$

The above interval is directly available from R using the riskratio() command.

Since $\hat{\Theta}$ is a biased estimator of θ , [1] suggests using a modified estimator for θ , which takes the form

$$\tilde{\Theta} = \frac{(X_1 + 0.5)/(n_1 + 0.5)}{(X_2 + 0.5)/(n_2 + 0.5)} \quad \text{and} \quad \tilde{\Psi} = \ln \tilde{\Theta}.$$

The estimated variance of $\tilde{\Psi} = \ln \tilde{\Theta}$ is

$$\widehat{\text{var}}(\tilde{\Psi}) \approx \sum_{i=1}^2 \left(\frac{1}{x_i + 0.5} - \frac{1}{n_i + 0.5} \right).$$

Thus, the corresponding approximate $(1-\alpha)100\%$ confidence interval for θ is

$$\left(e^{\ln \tilde{\theta} - z_{\alpha/2} \sqrt{\widehat{\text{var}}(\tilde{\Psi})}}, e^{\ln \tilde{\theta} + z_{\alpha/2} \sqrt{\widehat{\text{var}}(\tilde{\Psi})}} \right).$$

[2] proposed to construct an approximate $(1-\alpha)100\%$ confidence interval for θ based on the likelihood ratio statistic. Since X_i are independently distributed as Binomial (n_i, p_i) , the joint log-likelihood function is

$$l(p_1, p_2) = \sum_{i=1}^2 [x_i \ln p_i + (n_i - x_i) \ln(1 - p_i)].$$

The point (\hat{p}_1, \hat{p}_2) that maximizes the log-likelihood function is known as the maximum likelihood estimate (MLE) of (p_1, p_2) , which can be obtained by solving

$$\frac{\partial l(p_1, p_2)}{\partial p_1} = 0 \quad \text{and} \quad \frac{\partial l(p_1, p_2)}{\partial p_2} = 0.$$

In this case, the MLE of (p_1, p_2) is $(\hat{p}_1, \hat{p}_2) = \left(\frac{x_1}{n_1}, \frac{x_2}{n_2} \right)$. Moreover, for a given

θ value, the point $(\tilde{p}_1, \tilde{p}_2)$ that maximized the log-likelihood function subject

to the constraint $\frac{\tilde{p}_1}{\tilde{p}_2} = \theta$ is known as the constrained MLE of (p_1, p_2) . [2]

gives a numerical algorithm to obtain $(\tilde{p}_1, \tilde{p}_2)$. However, by applying the Lagrange multiplier technique, we have the explicit closed form of the constrained MLE:

$$\tilde{p}_1 = \theta \tilde{p}_2$$

and

$$\tilde{p}_2 = \frac{[(n_1 + x_2)\theta + (n_2 + x_1)] - \sqrt{[(n_1 + x_2)\theta + (n_2 + x_1)]^2 - 4(x_1 + x_2)(n_1 + n_2)\theta}}{2(n_1 + n_2)\theta}$$

The observed likelihood ratio statistic is

$$w(\theta) = 2[l(\hat{p}_1, \hat{p}_2) - l(\tilde{p}_1, \tilde{p}_2)].$$

With the regularity conditions given in [9], the Wilks Theorem can be applied, and hence, $W(\theta)$ is asymptotically distributed as the chi-square distribution with 1 degree of freedom, χ_1^2 . Therefore, the approximate $(1-\alpha)100\%$ confidence interval for θ obtained in [2] is

$$\{\theta : w(\theta) < \chi_{1,\alpha}^2\}$$

where $\chi_{1,\alpha}^2$ is the $(1-\alpha)100^{\text{th}}$ percentile of the χ_1^2 distribution.

It is well-known that the above methods are not very accurate when the sample size is small. Although $W(\theta)$ is asymptotically distributed as χ_1^2 distribution, except in special cases, $E[W(\theta)] \neq 1$, which is the mean of the χ_1^2 distribution. [10] proposed a scale transformation of $W(\theta)$, such that the mean of the transformed statistics is the mean of the χ_1^2 distribution. This transformed statistic is known as the Bartlett corrected likelihood ratio statistic. Mathematically, let the Bartlett corrected likelihood ratio statistic be

$$W^*(\theta) = \frac{W(\theta)}{E[W(\theta)]}.$$

Then $W^*(\theta)$ is asymptotically distributed as χ_1^2 distribution and $E[W^*(\theta)] = 1$. However, the explicit form of $E[W(\theta)]$ is only available in a few well-defined problems.

In this paper, I propose to use the following algorithm to approximate $E[W(\theta)]$ and hence the observed Bartlett corrected likelihood ratio statistic $w^*(\theta)$.

Algorithm to approximate the mean of the likelihood ratio statistic

- Given:* An observed data set x_1, n_1, x_2 , and n_2 .
- Step 1:* From the observed data set, obtain the overall MLE (\hat{p}_1, \hat{p}_2) , and the log-likelihood function evaluated at the overall MLE, $l(\hat{p}_1, \hat{p}_2)$.
- Step 2:* For a given θ value, obtain the constrained MLE $(\tilde{p}_1, \tilde{p}_2)$, and the log-likelihood function evaluated at the constrained MLE, $l(\tilde{p}_1, \tilde{p}_2)$.
- Step 3:* From Step 1 and Step 2, obtain the observed log-likelihood ratio statistic $w(\theta)$.
- Step 4:* Simulate M independent samples of x_i from Binomial(n_i, \tilde{p}_i) and repeat Step 1 to Step 3 for each of the simulated samples. Hence, by the end of Step 4, we have $w_1(\theta), \dots, w_M(\theta)$.

- Step 5:* $E[W(\theta)]$ can be approximated by using $\bar{w}(\theta) = \frac{\sum_{i=1}^M w_i(\theta)}{M}$. Thus, the observed Bartlett corrected likelihood ratio statistic is $w^*(\theta) = \frac{w(\theta)}{\bar{w}(\theta)}$.
- Step 6:* The approximated $(1 - \alpha)100\%$ confidence interval based on the Bartlett corrected likelihood ratio statistic is $\{\theta: w^*(\theta) < \chi_{1,\alpha}^2\}$.

Note that the key step of the algorithm is Step 4 where we simulate new data from the Binomial distribution where the parameter is chosen to be the constrained MLE obtained in Step 2. The reason is that we are trying to obtain a sampling distribution of the likelihood ratio statistic $W(\theta)$, which is a function of the θ value given in Step 2. Hence, constrained MLE is used in Step 4.

As a final note in this section, the method by [2] is a computationally intensive method because, to obtain the required confidence limits, we need to find the smallest θ value and also the largest θ value such that $w(\theta) = \chi_{1,\alpha}^2$. The same needs to be done for the proposed method. However, the [1] methods have a closed form expression of the confidence limits, so they are easier to calculate and are available in statistical software.

3. Results

Our first example is to revisit the dataset discussed in previous section. **Table 2** recorded the 95% confidence interval for the relative risk obtained by the method discussed in this paper. Since the sample sizes are very large, it is not surprising that all the intervals are very close to each other.

As for our second example, the number of divorces during 2006 in a random sample of Army Reserve and Army Guard couples is reported in [11]. The data are presented in **Table 3**.

The estimated relative risk is $\hat{\theta} = \frac{12/324}{7/286} = 1.51$, which indicates that the divorce rate for Army Reserve personnel is higher than the divorce rate of the

Table 2. 95% confidence interval for the relative risk of having heart attacks among physicians taking aspirin versus physicians taking a placebo.

Method	95% confidence interval for θ
Agresti's method without adjustment	(0.4337, 0.6978)
Agresti's method with adjustment	(0.4348, 0.6990)
Zhou's method	(0.4323, 0.6961)
Proposed method	(0.4333, 0.6955)

Table 3. The number of divorces during 2006 in a random sample of Army Reserve and Army Guard couples.

Personal	Number of couples	Number of divorces
Reserve	324	12
Guard	286	7

Army Guard. **Table 4** recorded the 95% confidence interval for the relative risk obtained by the method discussed in this paper. Despite the sample sizes being relatively large, the results are still quite different.

For this example, we also calculated the probability that the true relative risk is as extreme or more extreme than the estimated relative risk by the four methods discussed in this paper. The results are plotted in **Figure 1(a)** for small true relative risk and **Figure 1(b)** for large true relative risk. The plots clearly showed that the four methods give different results especially when the true relative risk is large.

Hence, it is important to investigate which method is more accurate when sample size is small. The following simulation studies were performed.

- Set up:* Choose n_1, p_1, n_2 , and p_2 . Hence, the true relative risk is $\theta_0 = \frac{p_1}{p_2}$.
- Step 1:* Randomly generate x_1 from Binomial (n_1, p_1) , and x_2 from Binomial (n_2, p_2) .
- Step 2:* Obtain the 95% confidence intervals for θ using the four methods discussed in this paper.
- Step 3:* Record if the true relative risk, θ_0 , is smaller than the lower confidence limit, between the lower and upper confidence limits, or larger than the upper confidence limit of the 95% confidence interval calculated in *Step 2*.
- Step 4:* Repeat Step 1 to Step 3 N times.

Note that the proportion of samples with θ_0 less than the lower confidence limit is known as the lower error proportion, the proportion of samples with θ_0 larger than the upper confidence limit is known as the upper error proportion, and the proportion of samples with θ_0 falling within the confidence interval is known as the central coverage proportion. Moreover, the average absolute bias is defined as

$$\frac{|\text{lower error proportion} - 0.025| + |\text{upper error proportion} - 0.025|}{2},$$

which is a measure of bias of the 95% confidence interval. The nominal values for the lower error proportion, central coverage proportion, upper error proportion, and average absolute bias are 0.025, 0.95, 0.025, and 0, respectively.

Table 5 records the lower error proportion, central coverage proportion, and upper error proportion for a sample of simulation studies that I have performed. Results for other combinations of n_1, p_1, n_2 and p_2 are very similar and are available upon request.

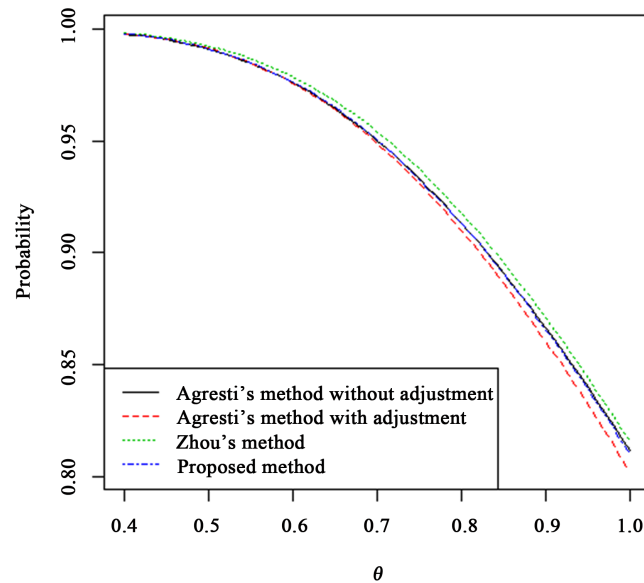
Table 4. 95% confidence interval for the relative risk of divorce in the Army Reserve versus the Army Guard.

Method	95% confidence interval for θ
Agresti's method without adjustment	(0.6040, 3.7908)
Agresti's method with adjustment	(0.6035, 3.5881)
Zhou's method	(0.6174, 4.0177)
Proposed method	(0.6244, 3.8329)

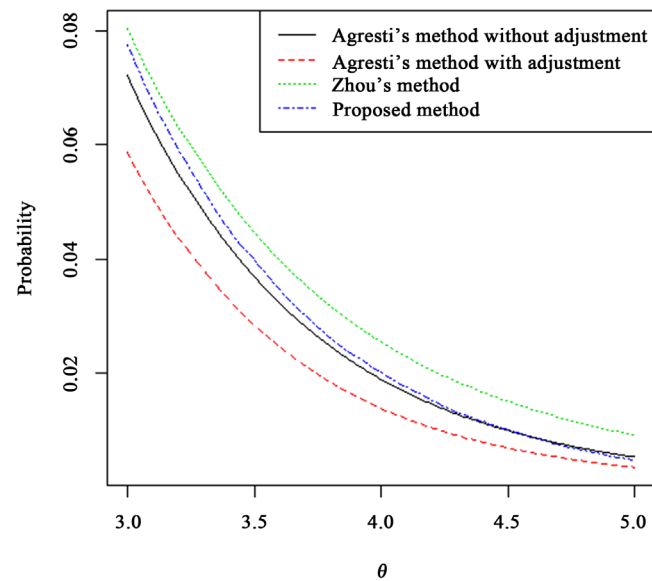
Table 5. Lower error proportion (le), central coverage proportion (cc), upper error proportion (ue), and absolute average bias (aab) of the 95% confidence interval for θ with $N = 10,000$ and $M = 200$.

n_1	p_1	n_2	p_2	Method	le	cc	ue	aab
10	0.4	10	0.6	1	0.0324	0.9676	0	0.0162
				2	0.0324	0.9676	0	0.0162
				3	0.0246	0.9555	0.0199	0.0028
				4	0.0280	0.9488	0.0232	0.0024
10	0.5	10	0.5	1	0.0070	0.9868	0.0062	0.0184
				2	0.0070	0.9868	0.0062	0.0184
				3	0.0237	0.9527	0.0236	0.0013
				4	0.0244	0.9508	0.0248	0.0004
20	0.3	25	0.6	1	0.0367	0.9605	0.0028	0.0170
				2	0.0447	0.9535	0.0018	0.0214
				3	0.0287	0.9433	0.0280	0.0034
				4	0.0247	0.9485	0.0268	0.0011
20	0.8	25	0.8	1	0.0196	0.9678	0.0126	0.0089
				2	0.0196	0.9678	0.0126	0.0089
				3	0.0200	0.9582	0.0218	0.0041
				4	0.0242	0.9466	0.0292	0.0025
20	0.7	25	0.3	1	0.0054	0.9612	0.0334	0.0140
				2	0.0030	0.9605	0.0365	0.0168
				3	0.0234	0.9535	0.0231	0.0018
				4	0.0270	0.9486	0.0244	0.0013
50	0.2	50	0.6	1	0.0321	0.9556	0.0123	0.0099
				2	0.0409	0.9501	0.0090	0.0160
				3	0.0222	0.9498	0.0280	0.0029
				4	0.0230	0.9495	0.0275	0.0023
50	0.7	100	0.4	1	0.0220	0.9492	0.0288	0.0034
				2	0.0207	0.9505	0.0288	0.0041
				3	0.0253	0.9454	0.0293	0.0023
				4	0.0260	0.9465	0.0275	0.0018

Note: Method 1 = Agresti's method without adjustment, Method 2 = Agresti's method with adjustment, Method 3 = Zhou's method, and Method 4 = Proposed method.



(a)



(b)

Figure 1. (a) and (b) show probability that the true relative risk is as extreme or more extreme than the estimated relative risk.

From **Table 5**, the two methods by [1] do not give satisfactory results. While one can argue that they have decent central coverage proportion when the sample sizes are large, they also have asymmetric tail errors. Moreover, although the aim of the adjusted method in [1] is a bias adjustment to the standard point estimator, it has little effect on the central coverage proportion, and it has adverse effect on the tail errors proportion. [2] method gives good central coverage proportion, but the tail errors are asymmetric. The proposed method outperformed the other three methods discussed in this paper regardless of the sample sizes.

4. Conclusion

In this paper, we demonstrated via simulations that the two methods discussed in [1], which are implemented in most standard statistical software, do not have good central coverage properties and the tail errors are extremely asymmetric, particularly when the sample sizes are small. Thus, practitioners should interpret confidence intervals obtained from standard statistical software with caution, especially when the sample sizes are small. The likelihood ratio method proposed in [2] has good central coverage, but the tail errors are asymmetric, which is still an improvement over [1] methods. In comparison, the proposed modification of the likelihood ratio method outperforms the other three methods in terms of both central coverage and tail error symmetry even when the sample sizes are small.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Agresti, A. (2012) *Categorical Data Analysis*. 3rd Edition, Wiley, New York.
- [2] Zhou, M. (2018) Confidence Intervals for Relative Risk by Likelihood Ratio Test. *Biostatistics and Biometrics Open Access Journal*, **6**, 1-3. <https://doi.org/10.19080/BBOAJ.2018.06.555700>
- [3] Siström, C.L. and Garvan, C.W. (2004) Proportions, Odds, and Risk. *Radiology*, **230**, 12-19. <https://doi.org/10.1148/radiol.2301031028>
- [4] Di Lorenzo, L., Coco, V., Forte, F., Trinchese, G.F., Forte, A.M. and Pappagallo, M. (2014) The Use of Odds Ratio in the Large Population-Based Studies: Warnings to Readings. *Muscles Ligaments and Tendons Journal*, **4**, 90-92. <https://doi.org/10.11138/mltj/2014.4.1.090>
- [5] Kalbfleisch, J.D. and Prentice, R.L. (2011) *The Statistical Analysis of Failure Time Data*. 2nd Edition, Wiley, New York.
- [6] Tenny, S. and Bhimji, S.S. (2017) *Relative Risk*. *StatPearls*. <https://www.ncbi.nlm.nih.gov/books/NBK430824>
- [7] Schechtman, E. (2002) Odds Ratio, Relative Risk, Absolute Risk Reduction, and the Number Needed to Treat—Which of These Should We Use? *Value Health*, **5**, 431-436. <https://doi.org/10.1046/J.1524-4733.2002.55150.x>
- [8] Gardner, M.J. and Altman, D.G. (1986) Confidence Intervals Rather Than P-Values: Estimation Rather Than Hypothesis Testing. *British Medical Journal (Clinical Research Edition)*, **292**, 746-750. <https://doi.org/10.1136/bmj.292.6522.746>
- [9] Shao, J. (2009) *Mathematical Statistics*. 2nd Edition, Springer, New York.
- [10] Bartlett, M.S. (1937) Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **160**, 268-282.
- [11] Kokoska, S. (2011) *Introductory Statistics: A Problem-Solving Approach*. Freeman, New York.