Second-order accurate likelihood inference for meta-analysis of comparative studies with binary outcomes

Ruggero Bellio[1], Annamaria Guolo[2], Cristiano Varin[3]

[1] University of Udine, Italy
[2] University of Padova, Italy
[3] Ca' Foscari University of Venice, Italy

E-mail for correspondence: ruggero.bellio@uniud.it

Abstract: We apply modern likelihood asymptotics to one-step methods for meta-analysis of comparative studies with binary outcomes. This approach requires a complete statistical model for the original data, rather than just a model to combine the summary statistics of individual studies as in two-step methods. We illustrate the advantages of calculating accurate confidence intervals in the common case where meta-analysis combines a limited number of studies and thus ordinary first-order accurate likelihood methods may yield incorrect inferential conclusions.

Keywords: Binomial data; High-order likelihood asymptotics; Meta-analysis; Random effects.

1 Introduction

Meta-analysis comprises various methods for combining the outcomes of distinct studies. Here, we focus on the meta-analysis of n clinical trials with a binary outcome comparing two experimental treatments. The statistic of interest for each study may be the risk difference, the risk ratio or the odds ratio. Using the terminology of Guo et al. (2023), combining the studies' information can be done with two-step methods based on approximate modelling of the study summary statistics or one-step methods that instead require modelling the individual data within the studies.

In the case of studies with a binary outcome, the observed data is given by a set of 2×2 tables, for which a binomial distribution is assumed, namely $Y_{1i} \sim \text{Binomial}(n_{1i}, \pi_{1i})$ for the first treatment and $Y_{2i} \sim \text{Binomial}(n_{2i}, \pi_{2i})$ for the second treatment or the control group, $i = 1, \ldots, n$. The rest of the model

This paper was published as a part of the proceedings of the 39th International Workshop on Statistical Modelling (IWSM), Limerick, Ireland, 13–18 July 2025.

specification corresponds to suitable assumptions on the study success rates π_{1i} and π_{2i}. Two examples of such assumptions are discussed in Section 3.

2 Likelihood asymptotics at work

Modern frequentist likelihood-based theory provides accurate inference methods that improve standard large-sample approximations. Such methods are promising for meta-analysis, often based on a relatively small number of studies or studies with a small sample size. Both situations may make standard first-order asymptotics unreliable. A review of likelihood asymptotics is provided in Pierce and Bellio (2017), where the accompanying `R` package `likelihoodAsy` is illustrated. Let θ denote the vector of model parameters and let $\psi(\theta)$ denote a scalar general parametric function of interest. Pierce and Bellio (2017) highlight the r^* formula, which allows for second-order accurate computation of tail probabilities for the signed likelihood root for $\psi(\theta)$, overcoming the computation based on the large-sample normal approximation. An `R` package for likelihood asymptotics in two-step models for meta-analysis is illustrated in Guolo and Varin (2012). This work instead focuses on a one-step meta-analysis, which has the advantage over two-step methods in that it does not introduce approximations of the distribution of the study summary statistics, which could lead to underestimating important sources of variability in the data.

3 Two models for meta-analysis of binary outcomes

3.1 The binomial-normal hierarchical model

A classical approach to accounting for the study heterogeneity is that of hierarchical models with normal random effects. This approach is used among others in control-risk regression. Following the notation in Smith, Spiegelhalter, and Thomas (1998), if index 1 denotes the treatment of interest and index 2 denotes the control, the binomial-normal model is

$$\text{logit}(\pi_{1i}) = \mu_i + \delta_i/2\,, \qquad \text{logit}(\pi_{2i}) = \mu_i - \delta_i/2\,,$$

where μ_i are study-specific fixed intercepts, and $\delta_i \sim \text{Normal}(\delta, \tau^2)$ are random effects. The parameter of interest is the log-odds ratio, denoted by δ. The likelihood function is computed by integrating the random effects from the joint distribution of the observations and the random effects. We employ the adaptive Gaussian quadrature of Liu and Pierce (1994) to obtain a highly accurate approximation of the likelihood integrals. We use the data from Lu et al. (2014) on diabetes and Parkinson's disease as an illustrative example. There are 14 studies, with sample sizes between 140 and 82 140.

Figure 1 displays the confidence limits at all levels for δ based on first- and second-order methods. Two-sided confidence intervals can be read off the plot corresponding to standard normal quantiles. The traditional 95% confidence interval based on standard first-order asymptotics is $(-0.58, -0.018)$ with the Wald statistic and $(-0.61, -0.002)$ with the likelihood ratio statistic, both suggesting a weakly significant effect. The accurate second-order confidence interval is $(-0.64, 0.029)$ and thus leads to the conclusion that there is no significant effect, though the adjustment is indeed small.

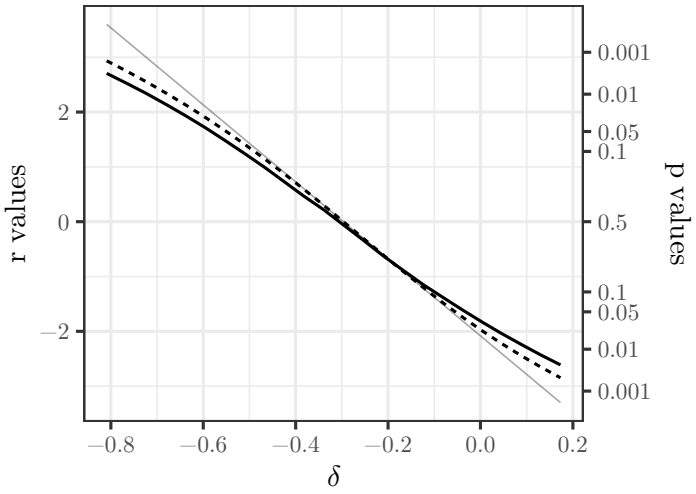

FIGURE 1. Lu et al. (2014) data on diabetes and Parkinson's disease: upper and lower one-sided confidence limits at all levels, using first-order (dashed line) and second-order (solid line) likelihood methods. The gray straight line corresponds to the Wald statistic values $(\widehat{\delta} - \delta)/\mathrm{SE}(\widehat{\delta})$.

3.2 The bivariate beta-binomial model

Another option is to account for the association between the two arms with the beta-binomial model revised by Guo et al. (2023). This model assumes a Sarmanov bivariate beta model for the success probabilities,

$$\begin{pmatrix} \pi_{1i} \\ \pi_{2i} \end{pmatrix} \sim \left[1 + \rho \prod_{k=1}^{2} (\pi_{ki} - \mu_k)\right] \prod_{k=1}^{2} \frac{\pi_{ki}^{\alpha_k - 1} (1 - \pi_{ki})^{\beta_k - 1}}{B(\alpha_k, \beta_k)},$$

where (α_k, β_k), $k = 1, 2$, are the parameters of the marginal distribution of π_{ki}, which is Beta(α_k, β_k), $\mu_k = \mathrm{E}(\pi_{ki}) = \alpha_k/(\alpha_k + \beta_k)$, $B(\cdot)$ is the beta function and ρ is a correlation parameter. This is a more parsimonious model than that of Section 3.1, with no study-specific parameters. The population log-odds ratio is given by $\delta = \log\{(\alpha_1/\beta_1)/(\alpha_2/\beta_2)\}$.

When applied to the data of Lu et al. (2014), we find a nearly perfect agreement between the results based on first-order and second-order accuracy methods, both in close agreement with the Wald-test inference. The second-order 95% confidence interval for δ is $(-0.97, 0.26)$, wider than that based on the binomial-normal hierarchical model. In this case, second-order accurate likelihood inference has a diagnostic value because it allows us to conclude that the results based on the first-order approximation are still reliable despite the low number of studies.

The difference between the results obtained with the binomial-normal hierarchical model and those with the beta-binomial model shows how important the assumptions made in meta-analysis models are, assumptions that are also difficult to evaluate when there are few studies. This suggests that meta-analyses should possibly be conducted with different models and assumptions to assess the robustness of inferential conclusions.

Acknowledgments: This publication was produced with funding from the

Italian Ministry of University and Research as part of the Call for Proposals for the scrolling of the final rankings of the PRIN 2022 - Project title "MEMIMR: Measurement Errors and Missing Information in Meta-Regression" - Project No. 2022FZY9PM - CUP C53C24000740006.

References

Günhan, B.K., Röver C. and Friede, T. (2020). Random-effects meta-analysis of few studies involving rare events. *Research Synthesis Methods*, **11**, 74 – 90.

Guo, J., Xiao, M., Chu, H. and Lin, L. (2023). Meta-analysis methods of risk difference: a comparison of different models. *Statistical Methods in Medical Research*, **32**, 3 – 21.

Guolo, A. and Varin, C. (2012). The R package metaLik for likelihood inference in meta-analysis. *Journal of Statistical Software*, **50**, 1 – 14.

Liu, Q. and Pierce, D.A. (1994). A note on Gauss-Hermite quadrature. *Biometrika*, **81**, 624 – 629.

Lu, L., Fu, D-l., Li, H-q., Liu, A-j., Li, J-h. and Zheng, G-q. (2014). Diabetes and risk of Parkinson's disease: an updated meta-analysis of case-control studies. *PLoS ONE*, **9(1)**, e85781.

Pierce, D.A. and Bellio, R. (2017). Modern likelihood-frequentist inference. *International Statistical Review*, **85**, 15 – 51.

Smith, T.C., Spiegelhalter, D.J. and Thomas, A. (1995). Bayesian approaches to random-effects meta-analysis: a comparative study. *Statistics in Medicine*, **14**, 2685 – 2699.