Bayesian Bivariate Meta-analysis of Diagnostic Test Studies using Penalised Complexity Priors

Jingyi Guo, Håvard Rue, Andrea Riebler

Second Bayesian Young Statisticians Meeting (BAYSM 2014)
Vienna, September 18–19, 2014

Norwegian University of Science and Technology, Department of Mathematical Sciences, Trondheim, Norway
<jingyi.guo, hrue, andrea.riebler>@math.ntnu.no

Abstract

In bivariate meta-analyses the number of studies involved is often low and data are sparse, so that model fitting using likelihood approaches can be problematic and Bayesian approaches are advantageous. Bayesian inference became attractive for routine use after the proposal of integrated nested Laplace approximation (INLA). However, the assignment of suitable prior distributions for the covariance matrix of the bivariate random effects has been still challenging. Here we show how to apply the recently proposed framework of penalised complexity priors to the variance components and the correlation parameter. The priors obtained are more interpretable and can be intuitively specified. The methodology is integrated in the user-friendly R-package `meta4diag` built on top of INLA.

Keywords: bivariate random effects; integrated nested Laplace approximation (INLA); interpretable prior distribution;

1 Introduction

A diagnostic test usually presents two-by-two tables from which pairs of sensitivity and specificity can be computed. A bivariate meta-analysis summarises the results from separately performed studies while keeping the two-dimensionality of the data [5]. Since the number of studies is often small and data may be sparse, maximum likelihood estimation can be challenging. Paul et al. [4] proposed to perform full Bayesian inference using integrated nested Laplace approximations (INLA) [6]. Harbord [2] noted that INLA has considerable promise to be used in routine analysis, but that it is hard to specify suitable prior distributions.

Recently, a new concept for constructing priors was proposed, which uses that many model components are nested within a natural base model [3]. A prior distribution is consequently defined on the distance between the flexible and the base model, and then transformed to the original parameter. Here, we apply this approach to derive sensible and interpretable prior distributions. To make the methodology available to the applied scientist we present a purpose-build R-package on top INLA.

2 Bivariate Model

Let TP, FP, TN and TN denote the number of true positives, false positives, true negatives, and false negatives, respectively. Further, let $\text{Se} = \text{TP}/(\text{TP} + \text{FN})$ be sensitivity and $\text{Sp} = \text{TN}/(\text{TN}+\text{FP})$ specificity. A bivariate model summarises the results of several diagnostic tests $i = 1, \ldots, I$ by modelling sensitivity and specificity jointly:

$$\begin{aligned}
&\text{TP}_i|\text{Se}_i \sim \text{Binomial}(\text{TP}_i + \text{FN}_i, \text{Se}_i), \quad \text{logit}(\text{Se}_i) = \mu + \mathbf{U}_i\alpha + \phi_i, \\
&\text{TN}_i|\text{Sp}_i \sim \text{Binomial}(\text{TN}_i + \text{FP}_i, \text{Sp}_i), \quad \text{logit}(\text{Sp}_i) = \nu + \mathbf{V}_i\beta + \psi_i, \\
&\begin{pmatrix}\phi_i \\ \psi_i\end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}\sigma_\phi^2 & \rho\sigma_\phi\sigma_\psi \\ \rho\sigma_\phi\sigma_\psi & \sigma_\psi^2\end{pmatrix}\right],
\end{aligned}$$

where μ, ν are intercepts for $\text{logit}(\text{Se}_i)$ and $\text{logit}(\text{Sp}_i)$, respectively, and $\mathbf{U}_i$, $\mathbf{V}_i$ are possibly available covariates vectors. The covariance matrix of the random effects parameters ϕ_i and ψ_i is parameterised using between-study variances σ_ϕ^2, σ_ψ^2 and correlation ρ [1]. An equivalent parameterisation uses precisions $\tau_\phi = 1/\sigma_\phi^2$ and $\tau_\psi = 1/\sigma_\psi^2$.

3 Choice of Prior Distributions

Mainly vague or mildly informative priors for τ_ϕ, τ_ψ and ρ, or the whole covariance matrix are used. Habord [2] proposed to use a stronger prior for ρ which is possibly not symmetric around zero and centered around a natural ρ_0 instead of zero. Penalised complexity (PC) priors allow for such a specification [3]. Consider the bivariate random effects model and let

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix}1 & \rho_0 \\ \rho_0 & 1\end{pmatrix} \qquad\qquad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}$$

denote the covariance matrix of the natural base and flexible model, respectively (assuming $\tau_\phi = \tau_\psi = 1$). The increased complexity between $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma_1})$ and $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma_0})$ is measured by the Kullback-Leibler discrepancy (KLD). A constant rate penalisation of the distance $d(\rho) = \sqrt{2\text{KLD}(\rho)}$ results in an exponential prior with parameter λ. The rate λ is determined from knowledge of the scale or some interpretable contrast of ρ, such as $\text{Prob}(|\rho| > U) = 0.01$, where U is a user-provide scale. Figure 1 shows the resulting prior for ρ for different values of λ and ρ_0. Priors for τ_ϕ and τ_ψ follow analogously.

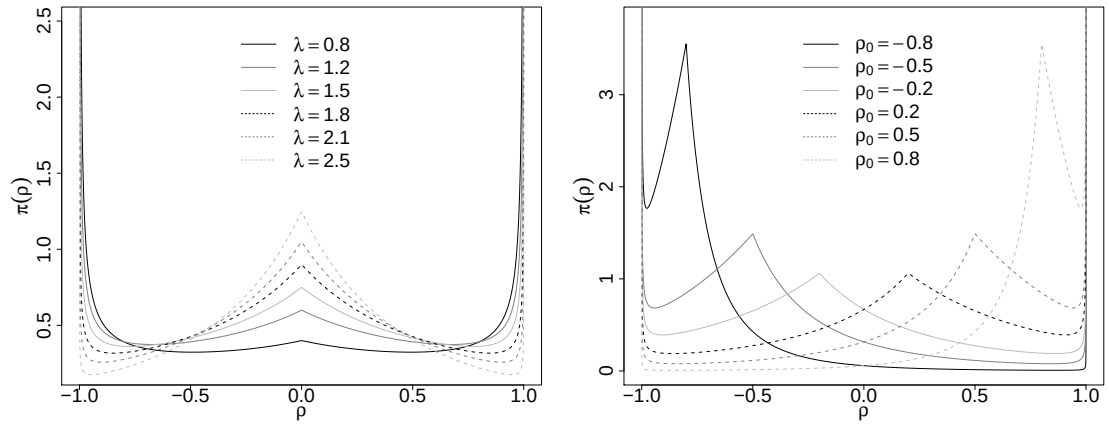

Figure 1: Prior density for ρ. Left: PC prior with $\rho_0 = 0$ and different values of λ. Right: PC prior with different values for ρ_0 and $\lambda = 2$.

4 R Package `meta4diag`

With our R-package `meta4diag` model and prior specification are straightforward, and standard outputs are directly available. Figure 2 shows an exemplary summary plot of a fitted model.

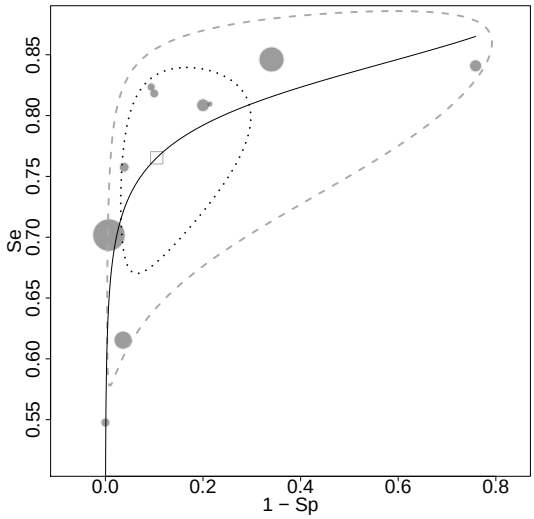

Figure 2: Summary plot from `meta4diag`: Gray circles denote study based estimates with diameter proportional to the study size. The square is the summary estimate and the black solid, black dashed and gray dashed line represent a summary ROC curve, 95% credible region and 95% prediction region, respectively.

5 Discussion

We proposed to use the novel PC priors for bivariate meta-analysis. These priors are easily interpretable and can be intuitively specified. First results of a simulation study indicate better performance of the new priors compared to alternative specifications. The whole methodology will be available via our novel R-package `meta4diag`.

References

[1] Chu, H., Cole S. R. (2006). "Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach." *Journal of Clinical Epidemiology*, **59**(12):1331–1333.

[2] Harbord, R. M. (2011). "Commentary on Multivariate metaanalysis: potential and promise. " *Statistics in Medicine*, **30**(20), 2507–2508.

[3] Martins, T. G., Simpson, D. P., Riebler, A., Rue, H., and Sørbye, S. H. (2014). "Penalising model component complexity: A principled, practical approach to constructing priors." *arXiv preprint arXiv:1403.4630*.

[4] Paul, M., Riebler, A., Bachmann, L. M., Rue, H., and Held, L. (2010). "Bayesian bivariate metaanalysis of diagnostic test studies using integrated nested Laplace approximations". *Statistics in Medicine*, **29**(12), 1325–1339.

[5] Reitsma J. B., Glas A. S., Rutjes A. W. S., Scholten R. J. P. M., Bossuyt P. M., Zwinderman A. H. (2005). "Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews". *Journal of Clinical Epidemiology*, **58**(10):982–990

[6] Rue, H., Martino, S., and Chopin, N. (2009). "Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion)" *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **71**(2), 319–392.