

An objective Bayesian procedure for meta-analysis of binomial data

E. Moreno, F.J. Vázquez-Polo and M.A. Negrín

Abstract This paper presents a Bayesian model for meta-analysis of sparse discrete binomial data which are out of the scope of the usual hierarchical normal random effect models. To assure coherency between the marginal and conditional prior distributions utilized in the analysis, the crucial linking distribution between the experimental effectiveness and the unconditional effectiveness (meta-effectiveness) is constructed from specific bivariate classes of distributions with given marginals. Applications to real multicenter data are given and compared with previous meta-analysis.

Key words: Intrinsic link distribution, Farlie-Gumbel-Morgenstern link distribution, Fréchet class, Testing on meta-parameters.

1 Introduction

The usual scenario for Bayesian meta-analysis of medical treatment effectiveness is as follows. A given treatment is applied to k samples of patients of sizes $\{n_i, i = 1, \dots, k\}$, and the sample effectiveness $\mathbf{x}_i = (x_{i1}, \dots, x_{in_i})$ at the i^{th} center is assumed to be normally distributed $N_{n_i}(\mathbf{x}_i | \theta_i \mathbf{1}_{n_i}, \sigma_i^2 \mathbf{I}_{n_i})$, where θ_i represents the treatment effectiveness conditional on center i , and σ_i^2 the variance. A non-observable meta-variable x is implicitly considered and assumed to be $N(x | \theta, \sigma^2)$ distributed, where the meta-parameter θ represents the unconditional treatment effectiveness, and σ^2 the meta-variance.

E. Moreno
University of Granada, e-mail: emoreno@ugr.es

F.J. Vázquez-Polo
University of Las Palmas de G.C., e-mail: fjvpolo@dmc.ulpgc.es

M.A. Negrín
University of Las Palmas de G.C., e-mail: mnegrin@dmc.ulpgc.es

In many practical situations the sample variables are discrete, for instance 0 – 1 random variables, and in that case the above normal hierarchical model is applied to the logit transformation $y_i = \log[x_i/(n_i - x_i)]$ with the reparametrization $\log[\theta_i/(1 - \theta_i)]$, where x_i denotes the number of successes at the i^{th} center, (Morris and Normand, 1992; Carlin, 1992; Sutton et al., 2000; Lambert et al., 2005; Nissen and Wolski, 2007, among many others). Unfortunately, when the multicenter binomial sample contains zeros this normal approximation does not work properly, even if a continuity correction to the original data be applied (Sweeting *et al.* 2004).

This paper proposes a hierarchical Bayesian model for meta-analysis of multicenter binomial data that may contain zeros. The crucial step of choosing a linking distribution, the distribution of θ_i conditional on θ , is addressed by imposing two constraints on it. The linking distribution has to be coherent with the specified marginals, and has to be able to accommodate different degrees of heterogeneity between the experimental trials.

The rest of the paper is organized as follows. In Section 2 we give the linking distributions and study some of their properties. In Section 3 a Bayesian procedure for testing the equality of the meta-effectiveness of two treatments is proposed. In Section 4 the Bayesian meta-analysis procedure is applied to real data from Nissen and Wolski (2007). Section 5 contains some concluding remarks.

2 Binomial Bayesian models for sparse data

Let M_i be the Bayesian model for the observational study in center i

$$M_i : \{Ber(x_i|\theta_i), \pi(\theta_i)\}, i = 1, \dots, k,$$

where $Ber(x_i|\theta_i) = \theta_i^{x_i}(1 - \theta_i)^{1-x_i}$, x_i denotes a 0 – 1 random variable the meaning of which is $x_i = 1$ if a patient receiving the treatment at center i successes and $x_i = 0$ otherwise, and θ_i denotes the probability of success of the treatment conditional on center i . Uniform or Jeffreys' priors are widely accepted as objective priors for the Bernoulli parameter, although for sparse data some arguments favoring the uniform over the Jeffreys have been given by Tuly *et al.* (2008).

Likewise, the Bayesian meta-model M_0 for a 0 – 1 unobservable meta-treatment effectiveness x is given by the pair of distributions $M_0 : \{Ber(x|\theta), \pi(\theta)\}$, where the meta-parameter θ means the unconditional probability of success of the treatment, and $\pi(\theta)$ is the uniform.

To complete the formulation of the Bayesian meta-analysis we need a linking distribution $\pi(\theta_i|\theta)$ to which we impose to be compatible with the marginal $\pi(\theta_i)$ and $\pi(\theta)$ given above. This implies that the bivariate distribution $\pi(\theta_i, \theta) = \pi(\theta_i|\theta)\pi(\theta)$ must satisfy the integral equations

$$\int_0^1 \pi(\theta_i, \theta) d\theta_i = \pi(\theta), \int_0^1 \pi(\theta_i, \theta) d\theta = \pi(\theta_i). \quad (1)$$

For the case where $\pi(\theta) = 1_{(0,1)}(\theta)$ and $\pi(\theta_i) = 1_{(0,1)}(\theta_i)$, a class of bivariate distributions that satisfies the integral equations (1) is the following Farlie–Gumbel–Morgenstern class, a parametric class of probability distributions given by

$$\pi^{FGM}(\theta_i, \theta | \kappa) = 1 + \kappa(2\theta_i - 1)(2\theta - 1), \quad (2)$$

for $0 \leq \kappa \leq 1$ (Morgenstern, 1956). The question here is which is the between center heterogeneity captured by this class. It follows that $Var^{FGM}(\theta_i | \theta)$ is a decreasing function of the correlation coefficient ρ . We note that $0 \leq \rho \leq 1/3$ as κ varies in $(0, 1)$ so that the *FGM* family is unable to model a linear correlation coefficient between θ_i and θ larger than $1/3$.

To extend the FGM class we consider the intrinsic-uniform (IU) class of priors

$$\pi^{IU}(\theta_i, \theta | t) = \sum_{z=0}^t \binom{t}{z} \frac{(\theta_i \theta)^z [(1 - \theta_i)(1 - \theta)]^{t-z}}{Be(z+1, t-z+1)} \quad (3)$$

where $t = 1, 2, \dots$, and $Be(a, b)$ is the Beta function. For each t this bidimensional prior is a beta–binomial mixture, which is derived using the intrinsic method (Berger and Pericchi 1996; Moreno *et al.* 1998; Casella and Moreno 2005, among others).

It follows that the variance of θ_i is a decreases function of ρ for any θ . We note that ρ varies in $(1/3, 1)$ as t ranges in $\{1, 2, \dots\}$.

To eliminate the hyperparameters, we consider integrate out κ to obtain the unconditional FGM bivariate prior

$$\pi^{FGM}(\theta_i, \theta) = 1 + \frac{1}{2}(2\theta_i - 1)(2\theta - 1). \quad (4)$$

Similarly, we obtain the bivariate unconditional intrinsic–uniform prior

$$\pi^{IU}(\theta_i, \theta) = \sum_{t=1}^{\infty} \pi^{IU}(\theta_i, \theta | t) \frac{3}{(t+2)(t+3)}. \quad (5)$$

2.1 Likelihood and posterior distribution of the meta-parameter

For a given link $\pi(\theta_i | \theta)$, the likelihood of the meta-parameter θ for the data (x_i, n_i) drawn from center i is given by

$$\Pr(x_i | n_i, \theta) = \int_0^1 \binom{n_i}{x_i} \theta_i^{x_i} (1 - \theta_i)^{n_i - x_i} \pi(\theta_i | \theta) d\theta_i. \quad (6)$$

If we replace $\pi(\theta_i | \theta)$ by the associated linking distribution given in (4) and (5) we obtain the likelihoods

$$\Pr(x_i | n_i, \theta, FGM) = \frac{1}{n_i + 1} \left(1 + \frac{1}{2}(2\theta - 1) \left(\frac{2x_i - n_i}{n_i + 2} \right) \right),$$

$$\Pr(x_i|n_i, \theta, IU) = \sum_{t=1}^{\infty} (1+t)(1-\theta)^t \binom{n_i}{x_i} \text{Beta}(1+n_i+t-x_i, 1+x_i) \\ \times {}_3F_2\left(a, b, \frac{\theta}{\theta-1}\right) \frac{3}{(t+2)(t+3)},$$

where ${}_3F_2(a, b, \frac{\theta}{\theta-1})$ denotes the generalized hypergeometric function with $a = (1-t, -t, 1+x_i)$ and $b = (1, -n_i-t+x_i)$.

Assuming that the center samples $\{(x_i, n_i), i = 1, \dots, k\}$ are independent, conditional on θ , we obtain the likelihood of the meta-parameter θ for the whole data set as the product

$$\Pr(\mathbf{x}|\mathbf{n}, \theta, w) = \prod_{i=1}^k \Pr(x_i|n_i, \theta, w), \quad (7)$$

where $\mathbf{x} = (x_1, \dots, x_k)$, $\mathbf{n} = (n_1, \dots, n_k)$ and $w = FGM$ and IU .

Let us assume that *a priori* the links have probabilities $\pi(FGM) = \pi(IU) = 1/2$. Then, the posterior model probabilities are given by

$$\Pr(M^{FGM}|\mathbf{x}, \mathbf{n}) = \frac{m(\mathbf{x}|\mathbf{n}, FGM)}{m(\mathbf{x}|\mathbf{n}, FGM) + m(\mathbf{x}|\mathbf{n}, IU)} = 1 - \Pr(M^{IU}|\mathbf{x}, \mathbf{n}) \quad (8)$$

where $m(\mathbf{x}|\mathbf{n}, FGM) = \int_0^1 \Pr(\mathbf{x}|\mathbf{n}, \theta, FGM) d\theta$ and $m(\mathbf{x}|\mathbf{n}, IU) = \int_0^1 \Pr(\mathbf{x}|\mathbf{n}, \theta, IU) d\theta$ are the marginals of the data $(\mathbf{x}|\mathbf{n})$ under model FGM and IU, respectively.

Finally, the posterior distribution of the meta-parameter θ is given by the mixture of the posterior distribution of θ conditional on FGM and on IU , and the weights of the mixture are given by their posterior model probabilities.

3 Testing the equality of treatment effectiveness

Treatments effectiveness comparison based on multiple studies is one of the most important chapter in meta-analysis. We assume that there are available samples $(\mathbf{x}_1, \mathbf{n}_1)$ of the effectiveness of a given treatment in k_1 healthcare centers, and $(\mathbf{x}_2, \mathbf{n}_2)$ for another treatment in k_2 healthcare centers. We want testing $H_0 : \zeta = \xi$ versus the unrestricted alternative $H_1 : (\zeta, \xi) \in (0, 1) \times (0, 1)$.

This problem is equivalent to the model selection problem between the Bayesian model M_0

$$M_0 : \{\Pr(\mathbf{x}_1|\mathbf{n}_1, \theta) \Pr(\mathbf{x}_2|\mathbf{n}_2, \theta), \pi(\theta)\},$$

and M_1

$$M_1 : \{\Pr(\mathbf{x}_1|\mathbf{n}_1, \zeta) \Pr(\mathbf{x}_2|\mathbf{n}_2, \xi), \pi(\zeta, \xi)\},$$

where we assumed that the data $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent, conditional on the meta-parameters. The optimal solution to this decision problem under a 0–1 loss function is to reject the null H_0 if its posterior probability is smaller than $1/2$, that is if $\Pr(M_0|\mathbf{x}_1, \mathbf{n}_1, \mathbf{x}_2, \mathbf{n}_2) < 1/2$.

Assuming that *a priori* $\Pr(M_0) = \Pr(M_1) = 1/2$, the posterior probability of M_0 is given by

$$\Pr(M_0|\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{1 + B_{10}(\mathbf{x}_1, \mathbf{x}_2)}, \quad (9)$$

where B_{10} , the Bayes factor to compare M_1 and M_0 , is

$$B_{10}(\mathbf{x}_1, \mathbf{x}_2) = \frac{\int_0^1 \int_0^1 \Pr(\mathbf{x}_1|\mathbf{n}_1, \zeta) \Pr(\mathbf{x}_2|\mathbf{n}_2, \xi) \pi(\zeta, \xi) d\zeta d\xi}{\int_0^1 \Pr(\mathbf{x}_1|\mathbf{n}_1, \theta) \Pr(\mathbf{x}_2|\mathbf{n}_2, \theta) \pi(\theta) d\theta}.$$

An empirical evaluation of the Bayesian testing procedure given above is provided in Moreno *et al.* (2014), suggesting that they are quite balanced.

4 Illustration

We consider data from 42 comparative studies (Nissen and Wolski, 2007) for analyzing the possible cardiovascular toxicity of the rosiglitazone, a type 2 diabetes drug, *versus* the control. They conducted searches of the published literature, the Web site of the Food and Drug Administration and a clinical-trials registry maintained by the drug manufacturer (GlaxoSmithKline), and the criteria for inclusion of the study for doing meta-analysis included a study duration of more than 24 weeks, the use of a randomized control group not receiving rosiglitazone, and the availability of outcome data for myocard infarction from cardiovascular causes.

For those data we test the null hypothesis M_0 that the efficacy of rosiglitazone and placebo are equal. As a previous step we compute the weights of the FGM and IU models under uniform marginals. The values obtained are $\Pr(M^{FGM}|\mathbf{x}_1, \mathbf{x}_2) = 0$ and $\Pr(M^{IU}|\mathbf{x}_1, \mathbf{x}_2) = 1$, respectively. These values show that the data favour models for large homogeneity.

Then we compute the posterior probabilities of the null model $\Pr^U(M_0|\mathbf{x}_1, \mathbf{x}_2) = 0.995$. This posterior probability indicates that there is strong empirical evidence in favour of the equality of myocard infarction for patients under rosiglitazone and control. This conclusion is in disagreement with that obtained by Nissen and Wolski (2007, p. 2459).

5 Discussion

The misuse and abuse of the different continuity correction methods in meta-analysis for sparse data was presented in Sweeting *et al.* (2004). We agree with these criticisms and in this paper we abandon the idea of correcting the data, and propose an objective Bayesian meta-analysis procedure for binomial data which is valid for analyzing binomial sparse data.

We have proposed classes of bivariate prior distributions for (θ_i, θ) such that they have given marginals, and can be parametrized in term of the Pearson' correlation coefficient between these parameters, so that the choice of the linking distributions have been driven by the linear correlation coefficient. For these classes we have seen that the higher the correlation coefficient the smaller the between center heterogeneity. Further, the linking and marginal distributions of θ_i and θ are coherent, and the elusive heterogeneity notion of the studies to be pooled is described by the Pearson' correlation coefficient, a quantity that varies in a bounded real set.

We have proposed the FGM and the intrinsic-uniform classes of distributions, a pair that cover Pearson's correlation coefficient values from 0 to 1, for testing the equality of meta-parameters. Since we use two different linking distributions, we need to averaging over the models and hence the final testing result is obtained as a mixture of the tests for each of the links. The weights of the mixture are given by the posterior probabilities of the models involved, and these weights inform us about the linking distribution favoured by the data.

Acknowledgements Research partially funded by grants MTM2011-28945 and ECO2009-14152 (MICINN, Spain).

References

1. Berger JO, Pericchi LR, (1996). The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, 91: 109-122.
2. Carlin JB. Meta-analysis for 2×2 tables: a Bayesian approach, (1992). *Statistics in Medicine*, 11: 141-159.
3. Casella G, Moreno E, (2005). Intrinsic meta analysis of contingency tables. *Statistic in Medicine*, 24: 583-604.
4. Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR. (2005). How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine*, 24: 2401-2428.
5. Moreno E, Bertolino F, Racugno W, (1998). An intrinsic limiting procedure for model selection and hypothesis testing. *Journal of the American Statistical Association*, 93: 145-1460.
6. Moreno E, Vázquez-Polo FJ and Negrín MA, (2014). Objective Bayesian meta-analysis for sparse discrete data. *Statistics in Medicine*, (to appear).
7. Morris CN, Normand SL (1992). Hierarchical models for combining information and for meta-analyses. In *Bayesian Statistics 4* (J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, eds.), pp. 321-344. Clarendon Press: Oxford.
8. Morgenstern D, (1956). Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsblatt für Mathematische Statistik*, 8: 234-235.
9. Nissen SE, Wolski K, (2007). Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *New England Journal of Medicine*, 356: 2457-2471.
10. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F, (2000). *Methods for meta-analysis in medical research*. Wiley: New York, 2000.
11. Sweeting MJ, Sutton AJ, Lambert PC, (2004). What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine*, 23: 1351-1375.
12. Tuyl F, Gerlach R, Mengersen K, (2008). A comparison of Bayes-Laplace, Jeffreys, and other priors: the case of zero events. *The American Statistician*, 62 (1): 40-44