NON-INFORMATIVE BAYESIAN INFERENCE FOR HETEROGENEITY IN A GENERALIZED MARGINAL RANDOM EFFECTS META-ANALYSIS

O. BODNAR

Abstract. In this paper an objective Bayesian inference is proposed for the heterogeneity parameter in a generalized marginal random effects model. Models of this kind are widely used in meta-analysis and in inter-laboratory comparisons. Under the assumption of elliptically contoured distributions, a reference prior for the model parameters is obtained and the analytical expression of the corresponding posterior is derived. We also state necessary conditions for the resulting posterior to be proper and for the existence of its first two moments. The obtained general theoretical results are illustrated for three popular families of elliptically contoured distributions: normal distribution, t-distribution, and Laplace distribution.

Key words and phrases. Non-informative prior, generalized random effects model, meta-analysis, estimation of heterogeneity, elliptically contoured distribution.

2010 Mathematics Subject Classification. 62F15, 62H10, 62H12, 62F10.

1. Introduction

Drawing inferences from data, which themselves are the results of studies, is known as meta-analysis. Recently, meta-analysis has become an important statistical tool in many real-life applications, like the combination of results from clinical trials [5, 39], the determination of fundamental constants [4, 6, 29], or the analysis of interlaboratory comparisons [10]. It provides an effective way to combine information obtained from several sources. Quantitatively, meta-analysis is usually based on two statistical models, namely the fixed effects model (FEM) and the random effects model (REM). The latter approach has increased its popularity recently.

The conventional random effects model is defined under the assumption of normality and it is given by

\[ x_i = \mu + \lambda_i + \varepsilon_i, \quad i = 1, \ldots, n, \quad (1) \]

where \( \mu \) is the common mean, \( \lambda_i \) are the individual effects used to model the additional variability which is not accounted for in the errors \( \varepsilon_i \). Moreover, it is assumed that \( \lambda_i \sim N(0, \tau^2) \), \( \varepsilon_i \sim N(0, \sigma_i^2) \) with known \( \sigma_i^2 \), and \( \lambda_i, \varepsilon_i, \quad i = 1, \ldots, n \), are mutually independent. A direct extension of model (1) violates the assumption of independence in \( \varepsilon_i, \quad i = 1, \ldots, n \) by introducing a covariance matrix into the model that summarizes the correlation coefficients computed for the model residuals. Moreover, integrating over \( \lambda_1, \ldots, \lambda_n \) we get the so-called marginal random effects model given by

\[ x \sim N_n(\mu \mathbf{1}, V + \tau^2 \mathbf{I}), \quad (2) \]

where \( x = (x_1, \ldots, x_n)^T \) is the observation vector, \( V \) is a \( n \times n \) positive definite dispersion matrix of residuals \( \varepsilon_1, \ldots, \varepsilon_n \) that is assumed to be known, \( \mathbf{1} \) is a vector of ones, and \( \mathbf{I} \) denotes the identity matrix of an appropriate order. The unknown model parameters are \( \mu \) and \( \tau \) which have to be estimated when the model is fitted to real data.

Several approaches to infer the common mean \( \mu \) and the extra variability parameter \( \tau \), known as the heterogeneity parameter or the between-study standard deviation, are available in the literature. They are divided into two large groups. The first one uses
the methods of the frequentist statistics, while the second group opts for the Bayesian approach. While many papers treat the problem of estimating the mean parameter \( \mu \), the estimation of \( \tau \) has become a hot topic of research recently.

The one of the mostly used estimator for \( \mu \) is the DerSimonian–Laird estimator (see [9]) which is still a very popular method for meta-analysis [7, 8, 19, 30, 41]. However, this approach completely ignores the estimation error which arises when the heterogeneity parameter is estimated (see [15]) that is needed for the inference about the common mean \( \mu \). As a result, the resulting interval estimator for \( \mu \) appears to be too narrow leading to smaller coverage properties. A finite-sample correction of the interval estimator for \( \mu \) is proposed in [24]. The Mandel–Paule estimator [27, 28, 32] for the between-study variability and the profile likelihood method are other classes of promising alternatives from classical statistics (e.g. [17]).

Bayesian inference for the parameters of the random effects model has been developed in the sequence of papers [5, 18, 20, 21, 25, 31, 33, 38, 40–42], which mainly deal with the estimation of \( \mu \) and do not spend a lot of attention to the problem of estimating \( \tau \). One of the main advantages of the Bayesian approaches with respect to the methods of the frequentist statistics is that it takes automatically the uncertainty about \( \tau \) into account when \( \mu \) is inferred. Also, as an output of a Bayesian procedure, we get the whole posterior distribution of the quantity of interest in contrast to the single point estimator produced in the conventional approach. Also, statistical inference procedures for the heterogeneity parameter alone have been proposed by [13, 35, 36], among others under the assumption of model (2).

In the following we extend the existent results dealing with Bayesian inference for \( \tau \) by weakening the model assumptions imposed on the observation vector \( x \). The new results are obtained by replacing the assumption of normality in (2) by a more general one, namely an elliptically contoured distribution. In this case, the joint density of \( x \) is given by

\[
p(x|\mu, \tau) = \frac{1}{\sqrt{\det(V + \tau^2 I)}} f \left( (x - \mu 1)^T (V + \tau^2 I)^{-1} (x - \mu 1) \right),
\]

(3)

where \( f(.) \) determines a specific family of elliptical distributions. This assertion we denote by \( x|\mu, \tau \sim E_n(\mu 1, V + \tau^2 I, f) \), that is the random vector \( x \) conditionally on \( \mu \) and \( \tau \) has an elliptical distribution with location parameter \( \mu 1 \), dispersion matrix \( V + \tau^2 I \), and density generator \( f \). Similarly, the notation \( E_n(0, I, f) \) denotes the standardized version of the elliptically contoured distribution with density generator \( f \) located around origin and with the identity dispersion matrix. Finally, the density generator should satisfy the condition

\[
\int_0^{\infty} r^{n/2-1} f(r) dr = 1,
\]

(4)

which ensures that the conditional density of \( x \) denoted by \( p(x|\mu, \tau) \) exists. In the rest of the paper we consider only those elliptical distributions for which the density exists, i.e. the density generator \( f(.) \) satisfies (4).

It is remarkable that the class of elliptically contoured distributions includes several well-known families of multivariate distributions, like the multivariate normal distribution, the multivariate \( t \)-distribution, and the multivariate Laplace distribution among others. We will refer to the model (3) as a generalized marginal random effects model. The posterior distribution of \( \mu \) is derived in [6] assuming that the observation vector \( x \) follows (3) and employing the Berger & Bernardo reference prior for model parameters. In the present paper we contribute by deriving the posterior of \( \tau \) and by comparing it with the posteriors obtained by using the existent non-informative and vague priors assigned to \((\mu, \tau)\).
The rest of the paper is organized as follows. In Section 2, the theoretical results of the paper are presented. In particular, the expression of the Fisher information matrix computed under the model (3) is provided in Subsection 2.1, which is then used to derive the Berger&Bernardo reference prior in Subsection 2.2. The posterior under the reference prior together with the posteriors obtained under the existent priors are provided in Subsection 2.3. Section 3 applies the general results to several classes of elliptically contoured distributions, while the results of the numerical study are given in Section 4.

2. Objective Bayesian inference for the heterogeneity

In this section we present the main results of the paper. The objective Bayesian analysis of the heterogeneity parameter \( \tau \) of the model (3) is developed. Throughout this section we treat \( \tau \) as the main parameter of the statistical model, while \( \mu \) stands for the nuisance parameter.

2.1. Fisher information matrix. The starting point of objective Bayesian inference is the Fisher information matrix. From Lemma 1 in [6] we get that the Fisher information matrix for model (3) is given by

\[
F_{\text{rem}} = \begin{pmatrix}
(F_{\text{rem}})_{11} & 0 \\
0 & (F_{\text{rem}})_{22}
\end{pmatrix},
\]

where

\[
(F_{\text{rem}})_{11} = 4I^T(V + \tau^2I)^{-1/2}E\left(ZZ^T\left(\frac{f'(Z^TZ)}{f(Z^TZ)}\right)^2\right)(V + \tau^2I)^{-1/2}1,
\]

\[
(F_{\text{rem}})_{22} = 4\tau^2\text{tr}\left((V + \sigma_1^2I)^{-2}\right)E\left((Z_1^2 - Z_2^2Z_2^2)^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right) + \\
\tau^2\left(\text{tr}\left((V + \sigma_1^2I)^{-1}\right)\right)^2\left(1 + 4E\left(Z_2^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right)^2\right) + 4E\left(Z_2^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right)
\]

with \( Z = (Z_1, \ldots, Z_n)^T \sim E_n(0, I, f) \).

In Theorem 1, the expressions for \((F_{\text{rem}})_{11}\) and \((F_{\text{rem}})_{22}\) are considerably simplified, which makes their applications more easy to the practitioners.

**Theorem 1.** The Fisher information matrix for model (3) is given by (5) with

\[
(F_{\text{rem}})_{11} = \frac{4J_1}{n}1^T(V + \tau^2I)^{-1}1,
\]

\[
(F_{\text{rem}})_{22} = \frac{8J_2\tau^2}{n(n + 2)}\text{tr}\left((V + \tau^2I)^{-2}\right) + \tau^2\left(\text{tr}\left((V + \tau^2I)^{-1}\right)\right)^2\left(\frac{4J_2}{n(n + 2)} - 1\right),
\]

where

\[
J_i = E\left((Z^TZ)^i\left(\frac{f'(Z^TZ)}{f(Z^TZ)}\right)^2\right) \quad \text{and} \quad Z \sim E_n(0, I, f).
\]

**Proof.** First, we calculate the three expectations, which are present in \((F_{\text{rem}})_{22}\). It holds that

\[
E\left(Z_1^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right)^2 = E\left(Z_1^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right)E\left(Z^TZ\frac{f'(Z^TZ)}{f(Z^TZ)}\right),
\]

\[
E\left(Z_2^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right)^2 = E\left(Z_2^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right)E\left((Z^TZ)^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right),
\]

and

\[
E\left(Z_1^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right) = E\left(Z_1^2\frac{f'(Z^TZ)}{f(Z^TZ)}\right).
\]
\[ E\left( Z_1^4 - Z_2^4 \right) = E\left( \frac{Z_1^4}{(Z_1^2)^2} \right) \]

where we use the properties of spherical distributions, i.e. the fact that \( Z/\sqrt{Z^T Z} \) and \( Z^T Z \) are independent (see, e.g. Theorem 2.15 in [16]).

Next, we calculate the expectations in (7), (8) and (9), where the term \( f'(Z^T Z) \) is not present. Here, we use that \( Z/\sqrt{Z^T Z} \) is uniformly distributed on the unit sphere (cf. Theorem 2.15 in [16]), i.e. the distribution of \( Z/\sqrt{Z^T Z} \) does not depend on the type of elliptical distribution. Consequently, it coincides with the distribution of \( Z_N/\sqrt{Z_N^T Z_N} \) where \( Z_N \sim N_n(0, I) \). Hence,

\[ E\left( \frac{Z_1^2}{Z^T Z} \right) = E\left( \frac{Z_{N,1}^2}{Z_{N,2}} \right) = \frac{1}{E(Z_N^T Z_N)} E\left( \frac{Z_{N,1}^2}{Z_{N,2}} \right) = \frac{1}{n} E(Z_{N,1}) = \frac{1}{n}. \]

Similarly, we get

\[ E\left( \frac{Z_1^4 - Z_2^4 Z_2^2}{(Z^T Z)^2} \right) = \frac{1}{E(Z_N^T Z_N)^2} E\left( \frac{Z_{N,1}^4 - Z_{N,1}^2 Z_{N,2}^2}{Z_{N,2}^2} \right) = \frac{1}{n(n+2)} E(Z_{N,1}^4 - Z_{N,1}^2 Z_{N,2}^2) = \frac{2}{n(n+2)} \]

and

\[ E\left( \frac{Z_1^2 Z_2^2}{(Z^T Z)^2} \right) = \frac{1}{E(Z_N^T Z_N)^2} E\left( \frac{Z_{N,1}^2 Z_{N,2}^2}{Z_{N,2}^2} \right) = \frac{1}{n(n+2)} E(Z_{N,1}^2 Z_{N,2}^2) = \frac{1}{n(n+2)}. \]

Furthermore, since \( f(.) \) is the density generator of \( Z \), we get that the density of \( R^2 = Z^T Z \) is given by \( f_{R^2}(r) = r^{n/2-1} f(r) \) (cf. Theorem 2.16 in [16]) and

\[ E\left( \frac{Z^T Z f'(Z^T Z)}{f(Z^T Z)} \right) = E\left( \frac{R^2 f'(R^2)}{f(R^2)} \right) = \int_0^\infty r f'(r) r^{n/2-1} f(r) dr = \int_0^\infty r^{n/2} f(r) dr = \frac{n}{2} \int_0^\infty r^{n/2-1} f(r) dr = -\frac{n}{2}, \]

where we use that \( \int_0^\infty r^{n/2-1} f(r) dr = 1 \) because \( r^{n/2-1} f(r) \) is the density of \( R^2 \), which also implies that \( r^{n/2-1} f(r) = o(r^{-1}) \) as \( r \to \infty \), i.e. \( r^{n/2} f(r) = o(1) \) as \( r \to \infty \).

Using the above results we get

\[ (F_{rem})_{22} = \frac{8 J_2 T^2}{n(n+2)} \left( (V + \tau^2 I)^{-2} \right) + \tau^2 \left( \text{tr}((V + \tau^2 I)^{-1}) \right)^2 \left( \frac{4 J_2}{n(n+2)} - 1 \right). \]
Similar calculations are performed in case of \((F_{\text{rem}})_{11}\). For the \((i, j)\)th element of
\[
E\left(ZZ^T \left( \frac{f''(Z^T Z)}{f'(Z^T Z)} \right) \right)^2
\]
we get
\[
E\left(Z_i Z_j \frac{f''(Z^T Z)}{f'(Z^T Z)} \right)^2 = E\left(Z_i Z_j \frac{f''(Z^T Z)}{f'(Z^T Z)} \right)^2 E\left(Z^T Z \left( \frac{f''(Z^T Z)}{f'(Z^T Z)} \right) \right)^2.
\]
Since
\[
E\left(\frac{Z_i Z_j}{Z^T Z} \right) = E\left(\frac{Z_{N,i} Z_{N,j}}{Z_N^T Z_N} \right) = \frac{1}{n} E(Z_{N,i} Z_{N,j}) = \begin{cases} 0 & \text{for } i \neq j, \\ \frac{1}{n} & \text{for } i = j, \end{cases}
\]
we obtain
\[
(F_{\text{rem}})_{11} = \frac{4J_1}{n} 1^T (V + \tau^2 I)^{-1} 1.
\]

The results of Theorem 1 show that the Fisher information matrix is finite if
\[
E\left(Z^T Z \left( \frac{f''(Z^T Z)}{f'(Z^T Z)} \right) \right)^2 < \infty \quad \text{and} \quad E\left(Z^T Z^2 \left( \frac{f''(Z^T Z)}{f'(Z^T Z)} \right)^2 \right) < \infty.
\] (10)
The conditions in (10) depend on the generating function \(f(.)\) only, i.e. on the type of the elliptically contoured distribution. Consequently, throughout the paper, we assume that the generating function is chosen such that the expectations in (10) are finite.

2.2. Reference prior. In many practical situations, no prior information or only vague information about the model parameters is present. In such situations, the application of non-informative priors is preferable since they would reduce the effect of an incorrect prior on the resulting posterior. Several non-informative priors exist in Bayesian statistics. Historically, the first non-informative prior is suggested by P. S. Laplace in [26], who proposes to assign a constant (probably improper) prior to the parameters of statistical models. This approach, however, does not obviously lead to desirable properties of the posterior. For instance, this prior is not invariant to transformations of model parameters. To deal with this problem, H. Jeffreys suggests a new approach in [22] by determining a non-informative prior as a square root of the determinant of the Fisher information matrix. In the case of the model (3) it is given by
\[
\pi_J(\mu, \tau) \propto \sqrt{\text{det}(F_{\text{rem}})} = \sqrt{(F_{\text{rem}})_{11} \sqrt{(F_{\text{rem}})_{22}}},
\] (11)
which appears to be a function of \(\tau\) only.

J. Berger and J. Bernardo extend the Jeffreys’ prior to the multi-parameter problems in [2] by proposing the so-called reference prior, which has become one of the mostly used non-informative priors, recently. In the derivation of the reference prior the model parameter are divided into several groups. Then, the conditional and marginal priors are assigned to each group starting with the least relevant group of parameters. In the paper, we specify \(\tau\) as the main model parameter and \(\mu\) as a nuisance parameter, since the goal is to infer \(\tau\). The expression of the reference prior is presented in Theorem 2.

Theorem 2. The Berger & Bernardo reference prior \(\pi(\mu, \tau)\) for the generalized marginal random effects model (3) and grouping \(\{\tau, \mu\}\) (i.e. with \(\mu\) as the nuisance parameter) equals the reference prior for model (3) and grouping \(\{\mu, \tau\}\) (i.e. with \(\tau\) as the nuisance parameter). Moreover, it is given by
\[
\pi_R(\mu, \tau) \propto \sqrt{(F_{\text{rem}})_{22}},
\] (12)
where \((F_{\text{rem}})_{22}\) is given by (6).
Proof. The reference prior for model (3) and grouping \{µ, τ\} is derived in [6] and it is given by (12). Next, we prove that it coincides with the reference prior for model (3) and grouping \{τ, µ\}. Just as for grouping \{µ, τ\} the reference prior turns out to be improper and is determined as the limit of proper priors on compact subsets. In using subsets of the form \(Ω^l_μ \times Ω^l_τ\), \(l = 1, 2, \ldots\), where \(Ω^1_μ \subset Ω^2_μ \subset \ldots\) with \(\bigcup Ω^l_μ = (−∞, ∞)\), and \(Ω^1_τ \subset Ω^2_τ \subset \ldots\) with \(\bigcup Ω^l_τ = (0, ∞)\), we get from Proposition 4 in [3] that

\[
π^{(l)}(μ|τ) = \frac{∫_{Ω^l_μ} (F_{rem})_{11} Ω^l_τ}{∫_{Ω^l_μ} (√(F_{rem})_{11}) dμ} = Ω^l_μ \frac{(F_{rem})_{22}}{|Ω^l_μ|} \sqrt{Ω^l_τ}
\]

and, consequently, the reference prior \(π(τ, μ)\) is obtained as the limit

\[
π(τ, μ) = \lim_{l→∞} π^{(l)}(τ, μ) = \frac{∫_{Ω^l_μ} π^{(l)}(μ|τ) log(√(F_{rem})_{22}) dμ}{∫_{Ω^l_μ} π^{(l)}(μ|τ) log(√(F_{rem})_{22}) dτ} \frac{1}{Ω^l_μ} \times √(F_{rem})_{22} \frac{1}{Ω^l_μ},
\]

where \(\hat{τ}\) and \(\hat{μ}\) denote some fixed values of τ and μ. □

The derivation of both the Jeffreys’ prior and the reference prior is based on the Fisher information matrix (5), which does not depend on the nuisance parameter μ. As a result, it is hold that \(π_j(τ) = π_j(μ, τ)\) and \(π_R(τ) = π_R(μ, τ)\), while only the second diagonal element of \(F_{rem}\) is used in the computation of \(π_R(τ)\).

2.3. Posterior. The objective Bayesian inference for the parameters of the generalized marginal random effects model (3) are derived in [6]. In particular, it is shown that the marginal posterior \(π_R(τ|x)\) obtained for the reference prior (12) is expressed as

\[
π_R(τ|x) ∝ C(τ) \frac{(F_{rem})_{22}}{det(V + τ^2I) (1^T(V + τ^2I)^{-1}1)}, \tag{13}
\]

where \(F_{22}\) is given in (6) and

\[
C(τ) = ∫_∞^-∞ f(x^TR(τ)x + u^2) du \tag{14}
\]

with

\[
R(τ) = (V + τ^2I)^{-1} - \frac{(V + τ^2I)^{-1}11^T(V + τ^2I)^{-1}}{1^T(V + τ^2I)^{-1}1}. \tag{15}
\]

This result can be extended to any prior for \(μ\) and τ suggested for model (3) such that \(π(μ, τ) = π(τ)\), i.e. the joint prior is a function of τ only. Such priors are widely used in the literature (see, e.g., [13, 25]) and are motivated by the fact that \(μ\) is the location parameter of the model and, consequently, the constant prior should be employed for it.

Theorem 3. The marginal posterior \(π(τ|x)\) obtained for the prior \(π(μ, τ) = π(τ)\) is given by

\[
π(τ|x) ∝ C(τ) \frac{π(τ)}{sqrt{det(V + τ^2I) (1^T(V + τ^2I)^{-1}1)}}, \tag{16}
\]

where \(C(τ)\) is given in (14).

Proof. The joint posterior for \(μ\) and τ under the generalized marginal random effects model (3) is given by

\[
π(μ, τ|x) ∝ π(μ, τ) f((x - μ1)^T(V + τ^2I)^{-1}(x - μ1)) \tag{17}
\]

\[
\sqrt{det(V + τ^2I)} \]

where \(C(τ)\) is given in (14).
It holds that
\[
(x - \mu)\big((V + \tau^2 I)^{-1}\big) = x^T R(\tau) x + I^T (V + \tau^2 I)^{-1} 1 \left( \mu - \frac{I^T (V + \tau^2 I)^{-1} x}{I^T (V + \tau^2 I)^{-1} 1} \right)^2,
\]
where \( R(\tau) \) is given in (15).

Hence,
\[
\pi(\mu, \tau | x) \propto \pi(\tau) \frac{f\left( x^T R(\tau) x + I^T (V + \tau^2 I)^{-1} 1 \left( \mu - \frac{I^T (V + \tau^2 I)^{-1} x}{I^T (V + \tau^2 I)^{-1} 1} \right)^2 \right)}{\sqrt{\det(V + \tau^2 I)}}.
\]

Integrating out \( \mu \) we get the statement of Theorem 3.

\[\square\]

2.4. **Propriety.** In this subsection we investigate under which condition imposed on \( \pi(\tau) \), the marginal posterior for \( \tau \) as given in Theorem 3 is proper. It is remarkable that the necessary conditions for propriety depend only on the sample size and the behaviour of the prior \( \pi(\tau) \) at infinity. Furthermore, similar conditions also ensure the existence of the posterior mean and variance.

**Theorem 4.** Let \( \pi(\tau) \) have no singularity in 0 and let \( \pi(\tau) \approx O(\tau^{-m}) \) at infinity. Then the posterior \( \pi(\tau | x) \) obtained in Theorem 3 is proper if \((n + m) > 2\). The according marginal posterior \( \pi(\tau | x) \) mean or variance exist if \((n + m) > 3\) or \((m + n) > 4\), respectively.

**Proof.** First, we note that since \( f(.) \) is a density generator of an elliptically contoured distribution then \( \int r^{n-1} f(r^2)dr < \infty \) converges (cf. Theorem 2.7. in [16]) and, consequently, \( \int v^{n/2-1} f(v)dv < \infty \) converges.

Then, no singularity is present at \( \tau = 0 \) since
\[
C(\tau = 0) = \int_{-\infty}^{\infty} f(x^T R(0) x + u^2) \, du = \int_{0}^{\infty} u^{-1/2} f(x^T R(0) x + u) \, du < \infty,
\]
where the last inequality follows from Lemma 9 in [14].

At infinity we get
\[
\frac{\pi(\tau)}{\sqrt{\det(V + \tau^2 I)}} \frac{1}{\sqrt{I^T (V + \tau^2 I)^{-1} 1}} \approx \tau^{-(m+n-1)}
\]
and
\[
\lim_{\tau \to \infty} C(\tau) = \int_{-\infty}^{\infty} \lim_{\tau \to \infty} f(x^T R(\tau) x + u^2) \, du = \int_{-\infty}^{\infty} f(u^2) \, du < \infty,
\]
where the last inequality follows from Lemma 9 in [14].

Hence,
\[
\pi(\tau | x) = O(\tau^{-(m+n-1)})
\]
(17)
as \( \tau \to \infty \), and the posterior is proper if and only if \( m + n > 2 \).

For the mean of the marginal posterior \( \pi(\tau | x) \) we note that the following two integrals
\[
\int_{-\infty}^{\infty} \tau \pi(\tau | x) \, d\tau
\]
(18)
and
\[
\int_{-\infty}^{\infty} \tau^2 \pi(\tau | x) \, d\tau
\]
(19)
are finite for \((n + m) > 3\) in case of (18) and for \((n + m) > 4\) in case of (19), respectively.

\[\square\]
The explicit formula for the marginal reference posterior $\pi(\tau|x)$ given in (16), can be utilized in the numerical calculation of marginal posterior mean and standard deviation,

$$E(\tau|x) = \int_0^\infty \tau \pi(\tau|x) \, d\tau,$$

and

$$\text{Var}(\tau|x) = \int_0^\infty \tau^2 \pi(\tau|x) \, d\tau - E(\tau|x)^2.$$

The shortest 95% credible interval can be obtained by minimizing over $\beta \in (0, 0.05)$ the length of the interval

$$[a_{0.05-\beta}, a_{1-\beta}],$$

where $a_\gamma$ is the solution of

$$\gamma = \int_0^{a_\gamma} \pi(\tau|x) \, d\tau.$$

(20)

3. Results for several families of elliptically contoured distributions

Next, we apply the obtained theoretical findings to three families of elliptical distributions: multivariate normal distribution, multivariate $t$-distribution, and multivariate Laplace distribution. We derive the reference priors for each of the considered classes of the elliptically contoured distributions as well as the corresponding posteriors. Furthermore, it is noted that similar expressions can also be obtained by applying other priors of the type $\pi(\mu, \tau) = \pi(\tau)$, for example the Jeffreys’ prior $\pi_J(\tau)$ as given in (11).

3.1. Normal distribution. In this subsection, we provide some explicit results under the additional assumption of normality, i.e. it is assumed that $f(u) = \exp(-u/2)/(2\pi)^{n/2}$. The results in Proposition 1 are taken from [6].

Proposition 1. The Berger & Bernardo reference prior $\pi(\mu, \tau)$ for the normal random effects model (i.e. model (3) with $f(u) = \exp(-u/2)/(2\pi)^{n/2}$) is given by

$$\pi_R(\tau) = \pi_R(\mu, \tau) \propto \sqrt{\tau^2 \cdot \text{tr}((V + \tau^2 I)^{-2})}.$$  

Similarly, the Jeffreys’ prior is obtained and it is given by

$$\pi_J(\tau) = \pi_J(\mu, \tau) \propto \sqrt{1^T(V + \tau^2 I)^{-1}1} \sqrt{\tau^2 \cdot \text{tr}((V + \tau^2 I)^{-2})}.$$  

From the findings of Subsection 2.2 and Proposition 1, we obtain the marginal posterior $\pi(\tau|x)$ as

$$\pi(\tau|x) \propto \frac{\pi(\tau)}{\sqrt{\det(V + \tau^2 I)\sqrt{1^T(V + \tau^2 I)^{-1}1}}} \exp\left(-\frac{1}{2}x^T R(\tau)x\right),$$  

(21)

where $R(\tau)$ is defined in Subsection 2.2 and $\pi(\tau)$ is a prior for $(\mu, \tau)$ that depends only on $\tau$.

Unfortunately, no analytical expressions are available for the posterior $\pi(\tau|x)$ mean and variance, and numerical means have to be applied. Markov chain Monte Carlo methods (cf. [34]) may be used or, since only one parameter is involved in the problem, one-dimensional numerical integration (see, e.g., [12]) is expected to provide a better result. The findings reported in Section 4 are obtained by the latter approach.
3.2. \textit{t}-distribution. In this subsection, we derive the closed-form expressions for the reference prior and the corresponding posterior in the case of the multivariate \textit{t}-distribution, i.e. when

\[ f(u) = K_{n,d}(1 + u/d)^{-(n+d)/2} \text{ with } K_{n,d} = (\pi d)^{n/2} \Gamma((d + n)/2)/\Gamma(d/2). \]

The expression of the reference prior is presented in Theorem 5.

\textbf{Theorem 5.} The Berger & Bernardo reference prior \(\pi(\mu, \tau)\) for the \textit{t}-distributed random effects model (i.e. model (3) with \(f(u) = K_{n,d}(1 + u/d)^{-(n+d)/2}\)) is given by

\[ \pi_R(\tau) = \pi_R(\mu, \tau) \propto \sqrt{\frac{2(n+d)}{n+d+2}} \tau^2 \cdot \text{tr}(\mathbf{V} + \tau^2 \mathbf{I})^{-2} - \frac{2}{n+d+2} \tau^2 (\text{tr}((\mathbf{V} + \tau^2 \mathbf{I})^{-1}))^2. \]

\textbf{Proof.} In case of \(f(u) = K_{n,d}(1 + u/d)^{-(n+d)/2}\), we get

\[ \frac{f'(u)}{f(u)} = -\frac{n + d}{2} \frac{K_{n,d}(1 + u/d)^{-(n+d)/2 - 1/d}}{K_{n,d}(1 + u/d)^{-(n+d)/2}} = -\frac{n + d}{2d} \cdot (1 + u/d)^{-1}. \]

Let \(C_{n,d}(q) = (q)^{n/2}/B(n/2, d/2)\). Using that \(F = \mathbf{Z}^T \mathbf{Z}/d \sim F_{n,d}\) (\(F\)-distribution with \(n\) and \(d\) degrees of freedom), we get

\[ \frac{4d^2}{(n + d)^2} J_2 = E \left( \left( \frac{\mathbf{Z}^T \mathbf{Z}}{1 + \mathbf{Z}^T \mathbf{Z}/d} \right)^2 \right) = n^2 \cdot \frac{\Gamma^2}{(1 + n/d F)^2} - \frac{2}{n + d} \cdot \frac{\Gamma}{\Gamma(d/2)} \]

\[ = n^2 C_{n,d}(n/d) \int_0^\infty x^n \left( 1 + \frac{n}{d} x \right)^{-n/2} \frac{(n + 4) + d}{2} \frac{\Gamma(n + 4 + d/2)}{\Gamma(n/2) \Gamma(d/2)} dx = \]

\[ = n^2 \frac{C_{n,d}(n/d)}{C_{n+4,d}(n/d)} \frac{(n + 4 + d/2) \Gamma(n + 4 + d/2)}{\Gamma(n/2) \Gamma(d/2)} \]

\[ = \frac{n^2 d^2}{n^2} \frac{\left( \frac{n + 4 + d}{2} \right)}{n + d + 1} \frac{n + 4}{2} = \frac{d^2 n(n + 2)}{(n + d)(n + 2 + d)}. \]

Hence, from Theorem 5 we get

\[ (F_{\text{rem}})_{22} = 2 \frac{n + d}{n + d + 2} \tau^2 \text{tr}((\mathbf{V} + \tau^2 \mathbf{I})^{-2}) - \frac{2}{n + d + 2} \tau^2 (\text{tr}((\mathbf{V} + \tau^2 \mathbf{I})^{-1}))^2. \]

Similar, under the \textit{t}-distributed random effects model, the Jeffreys’ prior is expressed as

\[ \pi_J(\tau) = \pi_J(\mu, \tau) \propto \sqrt{\frac{1}{n^T (\mathbf{V} + \tau^2 \mathbf{I})^{-1} \mathbf{1}}} \times \]

\[ \times \sqrt{\frac{2(n+d)}{n + d + 2} \tau^2 \cdot \text{tr}((\mathbf{V} + \tau^2 \mathbf{I})^{-2}) - \frac{2}{n + d + 2} \tau^2 (\text{tr}((\mathbf{V} + \tau^2 \mathbf{I})^{-1}))^2}. \]
Next, we derive the expression of the posterior by using Proposition 1 in [6]. It holds that

\[ \pi(\mu|\tau, x) \propto \left( 1 + \frac{1}{d} x^T R(\tau) x + \frac{1}{d} I^T (V + \tau^2 I)^{-1} 1 \right) \left( \mu - \frac{1}{I^T (V + \tau^2 I)^{-1} 1} \right)^2 \]

\[ \times \frac{\sqrt{1 + x^T R(\tau) x/d}}{\sqrt{1^T (V + \tau^2 I)^{-1} 1}} \times \frac{\sqrt{1^T (V + \tau^2 I)^{-1} 1}}{\sqrt{1 + x^T R(\tau) x/d}} \times \left( 1 + \frac{1}{n + \tau^2} \binom{n + d}{d} \right) \]

\[ \times \left( 1 + \frac{1}{n + \tau^2} \frac{1 + x^T R(\tau) x/d}{1 + x^T R(\tau) x/d} \right) \]

\[ \propto \left( 1 + \frac{1}{d} x^T R(\tau) x \right)^{-\frac{n + d + 1}{2}} \left\{ \frac{1}{\sqrt{\det(V + \tau^2 I)(I^T (V + \tau^2 I)^{-1} 1)}} \right\} \]

\[ \times \left( 1 + \frac{1}{n + \tau^2} \binom{n + d}{d} \right) \]

i.e., the conditional posterior of \( \mu \) given \( \tau \) and \( x \) has a univariate \( t \)-distribution with \((n + d - 1)\) degrees of freedom, location parameter

\[ \frac{I^T (V + \tau^2 I)^{-1} x}{I^T (V + \tau^2 I)^{-1} 1}, \]

and scale parameter

\[ \frac{1}{n + \tau^2} \frac{d + x^T R(\tau) x}{d} \]

where \( R(\tau) \) is defined in Subsection 2.3. Moreover, from the above derivation we immediately get

\[ \pi(\tau|x) \propto \pi(\tau) \left( 1 + \frac{1}{d} x^T R(\tau) x \right)^{-\frac{n + d + 1}{2}} \left\{ \frac{1}{\sqrt{\det(V + \tau^2 I)(I^T (V + \tau^2 I)^{-1} 1)}} \right\} \]

\[ \times \left( 1 + \frac{1}{n + \tau^2} \binom{n + d}{d} \right) \]

(22)

for a prior \( \pi(\tau) \) that is independent of \( \mu \).

Similarly to the case of normal distribution, no analytical expressions for the posterior \( \pi(\tau|x) \) mean and variance are available and numerical means have to be applied, like Markov chain Monte Carlo methods or numerical integration. The results reported in Section 4 are obtained by the latter approach. We further point out that the unconditional marginal posterior of \( \mu \) is a mixture of \( t \)-distributions.

3.3. Laplace distribution. In the case of the multivariate Laplace distribution, we get (cf. [11])

\[ f(u) = \frac{1}{(2\pi)^{n/2}} \int_0^\infty z^{-n/2} \exp\left( -\frac{u}{2z} - z \right) \, dz = \]

\[ = \frac{2(2\pi)^{-n/2}}{(\sqrt{u/2})^{n/2-1}} K_{n/2-1}(\sqrt{2u}), \]

(23)

where \( K_\alpha(x) \) denotes the modified Bessel function of the second kind (see, e.g., [1]). It holds that

\[ f'(u) = -\frac{1}{(2\pi)^{n/2}} \int_0^\infty z^{-n/2-1} \exp\left( -\frac{u}{2z} - z \right) \, dz = \]

\[ = -\frac{(2\pi)^{-n/2}}{(\sqrt{u/2})^{n/2}} K_{n/2}(\sqrt{2u}). \]

Hence,

\[ J_2 = \mathbb{E} \left( (Z^T Z)^2 f'(Z^T Z) \right)^2 \]

\[ = \frac{1}{2} \mathbb{E} \left( Z^T Z \left( \frac{K_{n/2}(\sqrt{2Z^T Z})}{K_{n/2-1}(\sqrt{2Z^T Z})} \right)^2 \right), \]

(24)
where the last expectation can be evaluated via simulations. Then, the reference prior for \((\mu, \tau)\) is given by

\[
\pi_R(\tau) = \pi_R(\mu, \tau) \propto \sqrt{(F_{\text{rem}})_{22}},
\]

where \((F_{\text{rem}})_{22}\) is given by (6) with \(J_2\) as in (24), and the Jeffreys’ prior we get

\[
\pi_J(\tau) = \pi_J(\mu, \tau) \propto \sqrt{1^T(\mathbf{V} + \tau^2\mathbf{I})^{-1}1 \sqrt{(F_{\text{rem}})_{22}}}.
\]

Next, we calculate marginal posterior \(\pi(\tau|x)\). It holds that

\[
C(\tau) = \int_{-\infty}^{\infty} f(x^T \mathbf{R}(\tau)x + u^2) \, du = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2}} \int_{0}^{\infty} z^{-n/2} \exp \left( -\frac{x^T \mathbf{R}(\tau)x + u^2}{2z} - z \right) \, dz \, du = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{(n-1)/2}} z^{-(n-1)/2} \exp \left( -\frac{x^T \mathbf{R}(\tau)x}{2z} - z \right) \left( \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{u^2}{2z} \right) \, du \right) \, dz = \frac{1}{(2\pi)^{(n-1)/2}} \int_{-\infty}^{\infty} z^{-(n-1)/2} \exp \left( -\frac{x^T \mathbf{R}(\tau)x}{2z} - z \right) \, dz = \frac{2(2\pi)^{-(n-1)/2}}{(\sqrt{x^T \mathbf{R}(\tau)x/2})^{(n-3)/2}} K_{(n-3)/2}\left(\sqrt{2x^T \mathbf{R}(\tau)x}\right). (25)
\]

Hence, the marginal posterior \(\pi(\tau|x)\) is obtained from (16) with \(C(\tau)\) as in (25).

3.4. Comparison of the Jeffreys’ prior and the reference prior. We perform a preliminary comparison of the Jeffreys’ prior and the reference prior derived under the assumption of the normal distribution, of the \(t\)-distribution, and of the Laplace distribution. In each figure the observation vector \(x\) with \(n \in \{5, 10\}\) is drawn from the generalized marginal random effects model (3) with the density generator \(f(.)\) corresponding to the considered family of elliptically contoured distributions. For each run, the simulation procedure of Subsection 4.1 is used. Although the results are present for one simulated sample in each plot of Figure 1, a similar behaviour is also observed for other simulated samples from the corresponding distribution. Finally, we note that Jeffreys’ prior is proper for any elliptically contoured distribution, while the reference prior is improper for all elliptically contoured distributions. For that reason, we normalized the reference prior by the constant obtained by the integration of the reference prior over a closed interval, here the interval \([0, 20]\) is used. Although the resulting curve is not a true density function, proceeding in this way allows us to assess the shape of both priors.

The plots in Figure 1 show that the Jeffreys’ prior has always a larger probability mass around its center. The picks on the prior curves corresponding to the Jeffreys’ prior are considerably larger and are located slightly closer to zero than those obtained for the reference prior. Both the derived posteriors by employing the Jeffreys’ prior and the reference prior are proper following the results of Theorem 4. As expected from the Bernstein-von-Mises theorem, the posteriors become more concentrated around their modes and the difference between them is smaller when the sample size \(n\) is larger. Besides that, we also observe that the posteriors obtained under the normality assumption have smaller posterior variance as those obtained under the \(t\)-distribution and the Laplace distribution, which is due to the heavy-tailed behaviour of their densities. Finally, larger tails are present in the case of Laplace distribution meaning that extremely large values of \(\tau\) can appear with higher probability than for the normal distribution and for the \(t\)-distribution.
Figure 1. Jeffreys’ prior and reference prior together with the corresponding posteriors for simulated data with \( n \in \{5, 10\} \) from the normal distribution, from the \( t \)-distribution, and from the Laplace distribution.

4. Frequentist Properties

The inferential properties of the posteriors obtained by employing the Jeffreys’ prior and the reference prior under the generalized marginal random effects model (3) are studied in this section via Monte Carlo study. We also compare the findings with three vague priors, which are commonly used in the random effects meta-analysis. The comparison is performed in terms of the coverage probabilities of 95% credible intervals determined by repeatedly simulating and analyzing sets of data for several families of elliptically contoured distributions.
4.1. Design of simulation study. The data are simulated in each run from model (3) under the assumption of the normal random effects model, of the $t$ random effects model, and of the Laplace random effects model. Without loss of generality we set $\mu = 0$ and $n \in \{5, 10\}$. The heterogeneity parameter $\tau$ is assumed to take one of the values from $\{0, 0.1, 0.2, 0.3, 0.5, 1, 1.5, 2\}$. We also take the effect of correlated model errors by specifying $V = CD$, where $C = (c_{i,j})_{i,j=1,...,n}$ is a correlation matrix of autoregressive type, i.e. $c_{i,j} = \rho^{|i-j|}$ with $\rho = 0.5$. The matrix $D$ is diagonal with entries $u_i$, which are the standard errors. In order to capture different situations, the $u_i$ are chosen differently for each single simulated data set. Specifically, the $u_i$ are drawn randomly from a uniform distribution on the interval $[0.0, 0.5]$.

In the comparison study additionally to the Jeffreys’ prior and the reference prior, we also use three vague priors, which are popular in the literature [13, 23, 37]. These priors are

1. Exponential prior: $\pi_E(\tau) = \pi_E(\mu, \tau) \propto \exp(-\tau)$,
2. Half-Cauchy prior: $\pi_{HC}(\tau) = \pi_{HC}(\mu, \tau) \propto \frac{1}{1+\tau^2}$,
3. Uniform prior: $\pi_U(\tau) = \pi_U(\mu, \tau) \propto 1$.

4.2. Numerical results. In Figure 2 the resulting credible intervals at 95% significance level obtained by using the exponential prior, the half-Cauchy prior, the Jeffreys’ prior, the reference prior, and the uniform prior are present in the case of the normal distribution, the $t$-distribution, and the Laplace distribution. Very good performance of all priors is obtained under the normal distribution. Here, we observe perfect coverage properties for the reference prior when $\tau \geq 0.5$ for $n = 5$ and when $\tau \geq 0.3$ for $n = 10$. The application of the Jeffreys’ prior leads to slightly smaller values of coverage probabilities for both considered sample sizes, while the underestimation is present for the uniform prior when $n = 5$. In contrast, the credible intervals under the exponential prior and the half-Cauchy prior are slightly too wide with the resulting coverage probability to be larger than the desired level of 95% in all of the considered cases when $n = 5$ and they are at 95% level of significance for $n = 10$.

Under the $t$ random effects model, the coverage probabilities for all considered priors are below 95% with the reference prior showing the smallest departures from this level for both $n = 5$ and $n = 10$ and also being above 95% when $n = 5$ and $\tau \leq 0.5$. In the case of $n = 5$, the Jeffreys’ prior show the worst performance in terms of the coverage probability, while for $n = 10$ it is ranked on the second place following by the uniform prior. It is remarkable, that for large values of $\tau$, the coverage probabilities obtained for the exponential prior and for the half-Cauchy prior are considerably smaller than 95% and they are close to 70% when $n = 10$ and $\tau \geq 1$.

Finally, it is remarkable that only the application of the reference prior leads to the reliable results in the case of the multivariate Laplace distribution. The coverage probabilities are always slightly larger than 95% under this prior. On the second place, we have the Jeffreys’ prior with the coverage probabilities to be larger 90% in almost all of the considered cases. On the other side, very bad performance is observed for the exponential prior, the half-Cauchy prior, and the uniform prior, especially when $n = 10$ with the values of the coverage probability being close to 60% for $\tau \geq 0.3$.

Acknowledgments

This research was supported by National Institute of Standards and Technology (NIST) Exchange Visitor Program. The author is grateful to the Statistical Engineering Division and especially Antonio Possolo for providing an excellent and inspiring environment for research.
The author would like to thank the Editor and the Reviewer for the thoughtful and constructive suggestions that led to an improvement of the paper.

Figure 2. Coverage probabilities of 95% credible intervals obtained by employing the exponential prior, the half-Cauchy prior, the Jeffreys’ prior, the reference prior, and the uniform prior for simulated data with $n \in \{5, 10\}$ from the normal distribution, from the $t$-distribution, and from the Laplace distribution.

UNIT OF STATISTICS, SCHOOL OF BUSINESS, ÖREBRO UNIVERSITY, FAKULTETSGATAN 1, 70182 ÖREBRO, SWEDEN

E-mail address: olha.bodnar@mdh.se

Received 28.02.2019

НЕИФОРМАТИВНЫЙ БАЙЕССИАННЫЙ АНАЛИЗ ПОДРОБНОЙ НЕОДНОРОДНОСТИ У МЕТААНАЛИЗА УЗАГАЛЬНЕННЫХ МОДЕЛЕЙ ИЗ МАРГИНАЛЬНЫМИ ВИПАДКОВЫМИ ЭФФЕКТАМИ

О. БОДНАР

Анотация. Запрограммирован об'єктивний байесівський аналіз для параметра неоднорідності в узагальненій моделі з маргінальними випадковими ефектами, яка широко використовується в метаналізі та в міжлабораторних порівняльних дослідженнях. За припинення еліптичного розподілу, введено референтний априорний розподіл параметрів моделі та отримано аналітичний вигляд для відповідного апостеріорного розподілу. Наведено необхідні умови обґрунтованості апостеріорного розподілу, а також їх використання його перших двох моментів. Отримані теоретичні результати застосовуються до трьох поширенних сімей еліптичних розподілів: нормального розподілу, й розподілу Лапласа.
НЕИНФОРМАТИВНЫЙ БАЙЕСОВСКИЙ АНАЛИЗ НЕОДНОРОДНОСТИ
В МЕТААНАЛИЗЕ ОБОБЩЕННЫХ МОДЕЛЕЙ С МАРГИНАЛЬНЫМИ
СЛУЧАЙНЫМИ ЭФФЕКТАМИ
О. БОДНАР

Аннотация. Предложен объективный байесовский анализ для параметра неоднородности в обобщенной модели с маргинальными случайными эффектами, которая широко используется в метаанализе и в межлабораторных сравнительных исследованиях. При условии эллиптичности распределений выведено референтное априорное распределение параметров модели и получена аналитическое выражение для соответствующего апостериорного распределения. Сформулированы необходимые условия обоснованности апостериорного распределения, а также существования его первых двух моментов. Полученные теоретические результаты применяются к трём распространённым семействам эллиптических распределений: нормальному распределению, t-распределению и распределению Лапласа.