

An Evaluation of Robust Linear Mixed Models for Bivariate Datasets with Slash Distributed Residuals ^[1]

Burcu MESTAV *  Kadir KIZILKAYA *

[1] This project was supported by grant TUBITAK TOVAG-1070915 from the Scientific and Technological Research Council of Turkey

* Adnan Menderes University, Faculty of Agriculture, Department of Animal Science, TR-09010 Aydin - TURKEY

Makale Kodu (Article Code): KVFD-2011-5108

Summary

Some symmetric and heavy-tailed distributions, such as Student's-t and Slash, have been suggested for robust inference in linear mixed models. These robust models are characterized by the degrees of freedom of these distributions, and include the normal distribution when the degrees of freedom approach infinity. This simulation study investigated joint estimation of degrees of freedom for the residual and all other genetic and non-genetic parameters in the Slash distributed residual datasets. Bivariate data with heavy-tailed distributed residuals were generated using Slash distributions with 4 or 12 degrees of freedom. Models with bivariate Student's-t, Slash and normal residuals were fitted to each dataset using a hierarchical Bayesian approach. Predictive log-likelihood values strongly favoured the bivariate Student's-t and Slash models over the normal models for simulated heavy-tailed datasets. Posterior mean estimates of degrees of freedom parameters seemed to be accurate and unbiased. Estimates of sire and herd variances were similar, if not identical, across fitted models. Posterior mean and 95% posterior probability interval estimates of error variances in simulated datasets were found to be similar. Reliable estimates of degrees of freedom were obtained in all simulated datasets. The predictive log-likelihood was able to distinguish the correct model among the models fitted to heavy-tailed datasets.

Keywords: Robust model, Slash distribution, Student's-t, MCMC

Hata Terimleri Slash Dağılımı Gösteren Bivaryet Veri Setlerinin Dirençli Doğrusal Karışık Modellerle Değerlendirilmesi

Özet

Student's-t ve Slash gibi bazı simetrik ve kalın kuyruklu dağılımlar, doğrusal karışık modellerde dirençli tahminler için önerilmektedir. Bu dirençli modeller bu dağılımların serbestlik dereceleri ile tanımlanmakta ve serbestlik derecesi sonsuza gittiğinde oluşan normal dağılımı da içermektedir. Bu simülasyon çalışmasında; Slash dağılımı gösteren hata terimli veri setlerinde, hata terimlerine ait serbestlik derecesi ile genetik ve genetik olmayan parametrelerin birlikte tahmini amaçlanmıştır. Kalın kuyruklu hata terimlerine sahip bivaryet veri setleri 4 veya 12 serbestlik dereceli Slash dağılımı kullanılarak türetilmiştir. Hata terimleri bivaryet Student's-t, Slash ve normal olan modeller hiyerarşik Bayesian yaklaşımı kullanılarak her bir veri setine uygulanmıştır. Tahmini log likelihood değerleri, kalın kuyruklu dağılım gösteren veri setleri için Student's-t ve Slash dağılımlı modellerin normal modele göre çok daha uygun olduğunu belirtmiştir. Serbestlik derecesinin ortalamasının yansız ve isabetli tahmin edildiği gözükmektedir. Baba ve sürü varyansları uygulanan modellerce aynı olmasada benzer olarak tahmin edilmiştir. Türetine veri setlerinde tahmin edilen hata varyanslarının ortalama ve %95 olasılık aralıkları da benzer bulundu. Serbestlik derecesinin güvenilir tahmini bütün veri setleri için elde edildi. Tahmini log-likelihood değeri de kalın kuyruklu dağılım gösteren veri setlerine uygulanan modeller arasında doğru modeli doğru şekilde belirledi.

Anahtar sözcükler: Dirençli model, Slash dağılımı, Student's-t, MCMC

INTRODUCTION

Bayesian and likelihood-based methods for prediction and estimation of variance components require definition

of the distributional assumptions relating to the data. The usual definitions involve normality of random factors and



İletişim (Correspondence)



+90 256 7727022/2006



bmestav@adu.edu.tr

residual effects. In practice, field data often exhibit more outliers than would be expected if the true distributions of effects were normal. Such heavy-tailed distributions can result from unknown fixed effects not being included in the model equation, or from distributions of effects being heavy-tailed. Among a number of alternative distributions that exhibit heavy tails, Student's-*t* and Slash distributions are appealing because these are symmetric and converge to normal distributions as their corresponding degrees of freedom exceed 50-100. The objective of this study was to investigate, using simulation, the practicality of jointly estimating the degrees of freedom for the bivariate Slash distributions assumed for the residual effects. The goodness of fit of the normal, Student's-*t* and Slash distributions was investigated for bivariate data simulated using Slash distributions of the residuals. The impact of the distributional assumptions was investigated for prediction of random effects and estimation of variance components.

MATERIAL and METHODS

Statistical Model

A bivariate linear mixed model with normal/independent errors for animal *i* is

$$\mathbf{y}_i = \mathbf{X}_i \mathbf{b} + \mathbf{Z}_i \mathbf{s} + \mathbf{W}_i \mathbf{h} + \boldsymbol{\varepsilon}_i \quad (1)$$

where \mathbf{Y}_i is a vector of observations for two traits from animal *i*, \mathbf{X}_i , \mathbf{Z}_i and \mathbf{W}_i are design matrices for animal *i*, corresponding to the *p*-dimensional vector of fixed effects (e.g., gender), q_s -dimensional vector of random sire effects and q_h -dimensional vector of uncorrelated random herd effects.

The error term $\boldsymbol{\varepsilon}_i$ is uncorrelated between animals and bivariate heavy-tailed or normal for any particular animal

$$\begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}_i = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}_i \lambda_i^{-\frac{1}{2}} = \mathbf{e}_i \lambda_i^{-\frac{1}{2}} \quad (2)$$

Here, \mathbf{e}_i is a bivariate normally distributed vector with mean zero and variance \mathbf{R}_0 . The scalar λ_i is a random positive variable with a density function of $p(\lambda | \nu)$ with ν degrees of freedom (df). Note that in this model the scalar λ_i applies to the residuals of both traits so that values of λ_i approaching 0 produce heavy-tailed residuals for both traits.

Prior and Posterior Distributions

The Bayesian approach requires specifying prior distributions for the parameters of interest. A totally uninformative or flat prior was assumed for the fixed effects (\mathbf{b}). A bivariate normal prior

$$\mathbf{s} | \mathbf{G}_0, \mathbf{A} \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G}_0) \quad (3)$$

was assigned for random sire effects $[\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{q_s}]$ with $\mathbf{s}_i = [s_{i1}, s_{i2}]'$ representing the two traits for sire *i*. The variance-covariance matrix is a kronecker product of the numerator relationship matrix \mathbf{A} , representing covariances between the sires, and a 2x2 variance-covariance matrix, $\mathbf{G}_0 = \begin{bmatrix} \sigma_{s1}^2 & \sigma_{s12} \\ \sigma_{s21} & \sigma_{s2}^2 \end{bmatrix}$, for the sire effects.

A similar bivariate normal prior

$$\mathbf{h} | \mathbf{H}_0 \sim N(\mathbf{0}, \mathbf{I} \otimes \mathbf{H}_0) \quad (4)$$

was assigned from herd effects, except that these effects between herds were uncorrelated as evident by the use of the identity matrix \mathbf{I} in the kronecker product and the effects between traits within herd were uncorrelated as

$$\text{shown by diagonal } \mathbf{H}_0 = \begin{bmatrix} \sigma_{h1}^2 & 0 \\ 0 & \sigma_{h2}^2 \end{bmatrix}.$$

Inverted Wishart (IW_2) distributions were assigned for \mathbf{G}_0 , \mathbf{H}_0 and \mathbf{R}_0 :

$$\boldsymbol{\Psi} | \mathbf{K}_\phi, \kappa_\phi \sim IW_2 \left((\kappa_\phi - 2 - 1) \mathbf{K}_\phi, \kappa_\phi \right) \quad (5)$$

where $\mathbf{K}_\phi = E(\boldsymbol{\Psi} | \mathbf{K}_\phi, \kappa_\phi)$ and κ_ϕ are degrees of freedoms of IW_2 for $\boldsymbol{\Psi} = \mathbf{G}_0$, \mathbf{H}_0 or \mathbf{R}_0 , and $\phi = s, h$ or e .

In equation (2), λ_i is an unknown "weight" variable, and is applied to both traits of animal *i*. There are many possible distributional assumptions on λ_i , each yielding a different specification on the distribution of the residuals $\boldsymbol{\varepsilon}_i$ relative to the use of the Gaussian distribution^{1,2}. Bivariate Slash (BS) distributions as an alternative to the bivariate Gaussian or normal distribution are specifically considered in this study.

The distribution of λ_i in equation (2) for BS is a $Beta(\nu, 1)$ distribution with density function

$$p(\lambda_i | \nu) = \nu \lambda_i^{\nu-1} \quad (6)$$

where $0 < \lambda_i \leq 1$, $i=1, 2, \dots, n$ and $\nu > 0$ is ν parameter. That results in a BS distribution specification for $p(\boldsymbol{\varepsilon}_i | \mathbf{R}_0, \nu)$, with parameter \mathbf{R}_0 and df ν such that the marginal variance of $\boldsymbol{\varepsilon}_i$ is $\mathbf{R}_E = \frac{\nu}{\nu-1} \mathbf{R}_0$. A truncated Gamma prior (α, β), with small positive values of α and β ($\beta \ll \alpha$), is adopted for ν in (6), so that the prior density is then

$$p(v | \alpha, \beta) \propto v^{\alpha-1} \exp\{-\beta v\} I\{v > 1\} \quad (7)$$

with v greater than 1.

The joint posterior density of all unobservables is then:

with parameters $\omega = n + \kappa_e$ and

$$\mathbf{\Omega}_e = (\kappa_e - 2 - 1)\mathbf{K}_e + \sum_{i=1}^m \lambda_i (\mathbf{y}_i - \boldsymbol{\mu}_i)(\mathbf{y}_i - \boldsymbol{\mu}_i)'$$

$$p(\mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{G}_0, \mathbf{H}_0, \mathbf{R}_0, \boldsymbol{\lambda}, v | \mathbf{y}) \propto \prod_{i=1}^n \left(\frac{1}{(2\pi)^{\frac{1}{2}} \left| \frac{\mathbf{R}_0}{\lambda_i} \right|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (\mathbf{y}_i - \boldsymbol{\mu}_i)' \left(\frac{\mathbf{R}_0}{\lambda_i} \right)^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i)\right\} \right) \times$$

$$|\mathbf{R}_0|^{-\left(\frac{\kappa_e + 2 + 1}{2}\right)} \exp\left\{-\frac{1}{2} \text{tr}(\mathbf{K}_e \mathbf{R}_0^{-1})\right\} \times |\mathbf{G}_0|^{-\left(\frac{q_s + \kappa_s + 2 + 1}{2}\right)} \exp\left\{-\frac{1}{2} \text{tr}([\mathbf{K}_s + \mathbf{Q}_s] \mathbf{G}_0^{-1})\right\} \times$$

$$|\mathbf{H}_0|^{-\left(\frac{q_h + \kappa_h + 2 + 1}{2}\right)} \exp\left\{-\frac{1}{2} \text{tr}([\mathbf{K}_h + \mathbf{Q}_h] \mathbf{H}_0^{-1})\right\} \times p(\boldsymbol{\lambda} | v) \times p(v | \eta) \quad (8)$$

where $\boldsymbol{\lambda}' = (\lambda_1, \lambda_2, \dots, \lambda_n)$.

Inferences on parameters of interest can be made using Gibbs sampling^{3,4}. The fully conditional posterior distributions of each of the unknown parameters are used to generate proposal samples from the target distribution (the joint posterior), which are accepted with probability equal to 1. The fully conditional posterior distribution of \mathbf{b} is a Gaussian. Likewise, the conditional posterior distributions of sire effects (\mathbf{s}) and herd effects (\mathbf{h}) are Gaussian.

Combining terms associated with \mathbf{G}_0 and \mathbf{H}_0 from the joint posterior density (8), the fully conditional posterior densities of \mathbf{G}_0 and \mathbf{H}_0 are:

$$p(\mathbf{G}_0 | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{H}_0, \mathbf{R}_0, \boldsymbol{\lambda}, v, \mathbf{y}) \propto |\mathbf{G}_0|^{-\left(\frac{q_s + \kappa_s + 2 + 1}{2}\right)} \exp\left\{-\frac{1}{2} \text{tr}([\mathbf{K}_s - 2 - 1]\mathbf{K}_s + \mathbf{Q}_s] \mathbf{G}_0^{-1})\right\} \quad (9)$$

and

$$p(\mathbf{H}_0 | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{G}_0, \mathbf{R}_0, \boldsymbol{\lambda}, v, \mathbf{y}) \propto |\mathbf{H}_0|^{-\left(\frac{q_h + \kappa_h + 2 + 1}{2}\right)} \exp\left\{-\frac{1}{2} \text{tr}([\mathbf{K}_h - 2 - 1]\mathbf{K}_h + \mathbf{Q}_h] \mathbf{H}_0^{-1})\right\} \quad (10)$$

which are IW_2 . Similarly, the fully conditional posterior distribution for \mathbf{R}_0 is also inverted Wishart

$$p(\mathbf{R}_0 | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{G}_0, \mathbf{H}_0, \boldsymbol{\lambda}, v, \mathbf{y}) \propto |\mathbf{R}_0|^{-\left(\frac{\omega + 2 + 1}{2}\right)} \exp\left\{-\frac{1}{2} \text{tr}(\mathbf{\Omega}_e \mathbf{R}_0^{-1})\right\} \quad (11)$$

For each element of the unobservable vector $\boldsymbol{\lambda}' = (\lambda_1, \lambda_2, \dots, \lambda_n)$, the fully conditional posterior distribution of λ_i for animal i is independent of that of λ_j for animal j . Thus, the fully conditional posterior distribution of λ_i ($i = 1, 2, \dots, n$) for the BS model is

$$p(\lambda_i | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{G}_0, \mathbf{H}_0, \mathbf{R}_0, \boldsymbol{\lambda}_{-i}, v, \mathbf{y}_{-i}) \propto (\lambda_i)^v \exp\left\{-\frac{\lambda_i}{2} (\mathbf{y}_i - \boldsymbol{\mu}_i)' \mathbf{R}_0^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i)\right\} \quad (12)$$

for $0 < \lambda_i < 1$, which is a Truncated - Gamma distribution

$$\lambda_i | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{R}_0, \boldsymbol{\lambda}_{-i}, v, \mathbf{y}_{-i} \sim \text{Truncated-Gamma}(\alpha, \beta) \quad (13)$$

$$0 < \lambda_i < 1$$

here $\alpha = v + 1$ and $\beta = \frac{1}{2} (\mathbf{y}_i - \boldsymbol{\mu}_i)' \mathbf{R}_0^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i)$

The fully conditional posterior distribution of v is:

$$p(v | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{G}_0, \mathbf{H}_0, \mathbf{R}_0, \boldsymbol{\lambda}, \mathbf{y}) \propto v^{n+a-1} \exp\left\{-v \left(b - \sum_{i=1}^n \log \lambda_i\right)\right\} \quad (14)$$

This is a Gamma distribution

$$v | \mathbf{b}, \mathbf{s}, \mathbf{h}, \mathbf{G}_0, \mathbf{H}_0, \mathbf{R}_0, \boldsymbol{\lambda}, \mathbf{y} \sim \text{Gamma}(n + a, b - \sum_{i=1}^n \log \lambda_i)^2$$

Simulation Study

A simulation study was carried out using two models for simulating the bivariate data. We refer to the model used to simulate the data as the true model. These two models were the Slash distributions with $\nu = 4$ (BS-4) or $\nu = 12$ (BS-12) degrees of freedom.

Two replicated datasets were generated for each of the two true models (BS-4 and BS-12). Phenotypes of 2500 progeny from 50 unrelated sires for two traits were simulated using Equation (1).

The model included a gender effect in \mathbf{b} ,

$$\mathbf{G}_0 = \begin{bmatrix} \sigma_{s1}^2 = 2.0 & \sigma_{s12} = 1.5 \\ \sigma_{s21} = 1.5 & \sigma_{s2}^2 = 4.0 \end{bmatrix}, \mathbf{h} \text{ comprising 100 herds}$$

with $\mathbf{H}_0 = \begin{bmatrix} \sigma_{h1}^2 = 1.5 & 0 \\ 0 & \sigma_{h2}^2 = 6.0 \end{bmatrix}$ and residuals

$$\mathbf{e}_i \sim N(\mathbf{0}, \mathbf{R}_0) \text{ where } \mathbf{R}_0 = \begin{bmatrix} \sigma_{e1}^2 = 15.0 & \sigma_{e12} = 4.0 \\ \sigma_{e21} = 4.0 & \sigma_{e2}^2 = 20.0 \end{bmatrix}.$$

For each animal i , λ_i was generated from $p(\lambda_i | \nu) = \nu \lambda_i^{\nu-1}$ for BS-4 and BS-12. Offspring were randomly distributed to herd and gender groups by using uniform distribution.

Fitted Model

Data generated from each of the 2 true models (BS-4, BS-12), were analyzed in 3 alternative fitted bivariate models. These were the model (BN) assuming normally distributed residuals, and the models assuming heavy-tailed distributions according to the Student's- t (BSt) or Slash (BS) distributions with ν treated as an unknown. Inferences on parameters are made from their marginal posteriors estimated using the Markov Chain Monte Carlo (MCMC) sampling process described previously.

MCMC Implementation

Bayesian inference was based on 50,000 post burn-in MCMC chains following 50,000 cycles of burn-in for each of the 3 fitted models applied to each replicate of the data simulated from the 2 true models. The length of the burn-in was judged by inspection of the plots of samples across rounds to ensure convergence. Every post-burn-in successive sample was retained for each replicate, and inferences were based on the pool of 100,000 samples taken from the two replicates fitted in each scenario involving a fitted and true model combination. Posterior means of the parameters were obtained from their respective marginal posterior densities. Interval estimates were determined as posterior probability intervals (95% PPI) obtained from the 2.5 and 97.5 percentiles of each posterior density.

Model Comparison

Distinguishing the goodness of fit of the 3 fitted models applied to each of the 2 true models was done by comparing predictive log-likelihoods (PLL) for different fitted models. Predictive log-likelihood over all observations (n) under Model M_k ($k = \text{BN, BSt or BS}$) was obtained as:

$$PLL_k = \sum_{i=1}^n \log p(\mathbf{y}_i | \mathbf{y}_{(-i)}, M_k) = \sum_{i=1}^n \log \left(\frac{1}{G} \sum_{j=1}^G p^{-1}(\mathbf{y}_i | \boldsymbol{\theta}^{(j)}, M_k) \right)^{-1} \quad (15)$$

where $\left(\frac{1}{G} \sum_{j=1}^G p^{-1}(\mathbf{y}_i | \boldsymbol{\theta}^{(j)}, M_k) \right)^{-1}$ is the harmonic mean of $p(\mathbf{y}_i | \boldsymbol{\theta}^{(j)}, M_k)$ across G MCMC samples⁵. Differences exceeding 2 were assumed to be significant⁶.

The impact of alternative models on selection was quantified by computing the correlations ($r_{\hat{\mathbf{s}}, \mathbf{s}}$) between the simulated true (\mathbf{s}) and predicted ($\hat{\mathbf{s}}$) sire effects in each of the three fitted models. Further, the prediction error variance (PEV) ($V(\mathbf{s} - \hat{\mathbf{s}})$) of the sire effects was calculated to provide an informative comparative assessment of model prediction performance. Higher correlations and lower prediction error variances will be associated with fitted models that are better at predicting breeding values than models with low correlations and high prediction error variance. Some fitted models might be significantly better than others from a likelihood framework, yet have little impact on selection response if they do not markedly change correlations. Minimizing the prediction error variances is important when investment decisions depend upon the magnitude of the sire predictions, not just the ranking of the sires.

RESULTS

Model Comparison

The predictive log-likelihood values in [Table 1](#) were computed for BSt, BS and BN models fitted to the simulated BS-4 and BS-12 datasets. The fitted models with heavy tails (BSt and BS) for BS-4 datasets were significantly better than the normal model (BN), with hardly any difference between BSt and BS. However, the models BSt, BS and BN fit the BS-12 datasets almost equally well, which means that $\nu = 12$ in Slash distribution provides almost normally distributed datasets. Also, PLL results in [Table 1](#) indicated that the difference in PLL between the fitted models with heavy tails and the normal model was inversely related to the degrees of freedom of the simulated residuals. In summary, regardless of how the residuals were simulated, PLL was the similar for the heavy-tailed fitted models, where the ν was treated as an unknown. This indicates that both (BSt and BS) models fit the data equally well.

The normal fitted model, however, did not perform well for the datasets having low degrees of freedom ($\nu < 4$).

Posterior distributions of degrees of freedom from BSt or BS fitted models, and their posterior mean (M) and 95% posterior probability intervals (L and U) were given in Fig. 1. Posterior distributions of ν were reasonably symmetric when the fitted model was the true model. The posterior means of ν were 4.7 and 16 from the fitted model BS for BS-4 and BS-12 and agreed well with the true values of ν . As seen in Fig. 1, posterior mean estimate of ν seems unbiased and sharp, and 95% posterior probability interval concentrated on low ν values for BS-4 datasets.

However, the posterior means of ν were 32.7 and 142.5 from the fitted model BSt for BS-4 and BS-12, and was higher than the true value. This indicates that BSt with ν degrees of freedom agrees with BS with degrees of freedom less than ν .

Average correlation between true and estimated sire effects and average PEV from two replicates using BSt, BS and BN fitted models are presented in Table 2. When the true model was BS, the both criteria (correlation and PEV) indicate that the heavy-tailed fitted models seem not to be superior on normal model, especially when the true value of $\nu=4$. This is not surprising as from Fig. 1 we can observe that BS-4 corresponds to a BSt- ν with ν about 30.

Estimation of Variance Components

Table 3 summarizes inferences on sire, herd and error variances based on the replicated datasets from the two different populations (BS-4 and BS-12), comparing BSt, BS and BN fitted models. Posterior mean estimates of (co) variances from three fitted models were found to be similar within each of two population; however, they were different across populations. Posterior mean estimates of σ_{h1}^2 for BS-12 and σ_{h2}^2 for BS-4 were sharp and unbiased in herd variances. Estimates of σ_{h1}^2 for BS-4 and σ_{h2}^2 for BS-12 seemed to be upward biased. The 95% posterior probability intervals for herd variance components from the three fitted models widely overlapped and included the true parameter values. Heritabilities of simulated bivariate traits were $h_1^2 = 0.36$ and $h_2^2 = 0.44$ for BS-4 and $h_1^2 = 0.40$ and $h_2^2 = 0.52$ for BS-12. Posterior mean estimates of sire (co)variances from BSt, BS and BN fitted models were identical in BS-4 and BS-12 populations. As the posterior mean of σ_{s2}^2 for BS-4 was over estimated, posterior mean estimates of other sire (co)variances for BS-4 and BS-12 populations were found to be sharp and seemingly unbiased. True parameters of sire (co)variances were included by 95% equal-tailed posterior probability interval of parameters (Table 3). The marginal error (co)variance

Table 1. Predictive log-likelihood¹ (PLL) comparisons between bivariate Student's-t (BSt), Slash (BS) and normal (BN) fitted models (in column) for BS-4 and BS-12 datasets

Tablo 1. BS-4 ve BS-12 veri setleri için uygulanan bivaryet Student's-t (BSt), Slash (BS) ve normal (BN) modellerine ait tahmini log-likelihood (PLL) değerlerinin karşılaştırması

True Model ² - DF	Fitted Model ³					
	BSt		BS		BN	
	Rep1	Rep2	Rep1	Rep2	Rep1	Rep2
BS-4	-957	-930	-957	-931	-961	-939
BS-12	-446	-531	-445	-531	-445	-532

¹ Predictive log-likelihood values were reported after adding 14000, ² Used to simulate data, ³ Used in analysis of simulated data

Fig 1. Marginal posterior densities of degrees of freedom estimated by true and fitted models

M represents posterior mean, L represents the 2.5th percentiles of the posterior density, U represent 97.5th percentiles of the posterior density

Şekil 1. Gerçek ve uygulanan modeller ile tahmin edilen serbestlik derecelerine ait marjinal son yoğunluklar

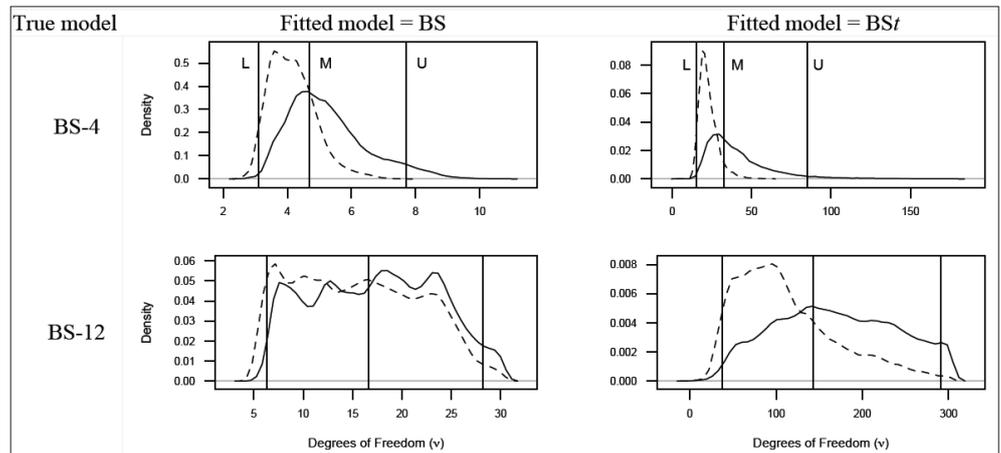


Table 2. Average correlations (*r*) between true and predicted sire effects from two replicates and prediction error variance (PEV) of sire effects using the bivariate Student's-t (BSt), Slash (BS) and normal (BN) fitted models with different residual degrees of freedom (DF)

Table 2. Bivaryet Student's-t (BSt), Slash (BS) ve normal (BN) modelleri uygulanarak tahmin edilen baba etkileri ile gerçek değerleri arasındaki ortalama korelasyon ve baba etkilerine ait tahmin hata varyansları

True Model ¹ - DF	Fitted Model ²											
	Trait1						Trait2					
	BSt		BS		BN		BSt		BS		BN	
	r	PEV	r	PEV	r	PEV	r	PEV	r	PEV	r	PEV
BS-4	0.92	0.35	0.92	0.35	0.92	0.35	0.94	0.48	0.94	0.48	0.94	0.48
BS-12	0.93	0.29	0.93	0.29	0.93	0.29	0.95	0.37	0.95	0.37	0.95	0.37

¹ Used to simulate data, ² Used in analysis of simulated data

Table 3. Posterior inference on sire, herd and marginal error (co)variances using the bivariate Student's-t (BSt), Slash (BS) and normal (BN) fitted models

Table 3. Bivaryet Student's-t (BSt), Slash (BS) ve normal (BN) modelleri kullanılarak elde edilen baba, sürü ve marjinal hata varyans-kovaryans değerleri

True Parameters	True Model ¹ - DF	Fitted Model ²					
		BSt		BS		BN	
		PM ³	95% PPI ⁴	PM	95% PPI	PM	95% PPI
$\sigma_{S1}^2 = 2.0$	BS-4	1.94	[1.13, 3.18]	1.94	[1.13, 3.18]	1.97	[1.15, 3.21]
	BS-12	2.28	[1.34, 3.70]	2.28	[1.34, 3.69]	2.28	[1.34, 3.69]
$\sigma_{S12}^2 = 1.5$	BS-4	1.56	[0.11, 3.57]	1.56	[0.13, 3.56]	1.56	[0.11, 3.57]
	BS-12	1.23	[0.29, 2.49]	1.24	[0.30, 2.49]	1.24	[0.30, 2.50]
$\sigma_{S2}^2 = 4.0$	BS-4	5.18	[2.30, 9.79]	5.18	[2.29, 9.84]	5.12	[2.26, 9.72]
	BS-12	3.78	[2.33, 5.95]	3.79	[2.33, 5.99]	3.79	[2.34, 6.00]
$\sigma_{h1}^2 = 1.5$	BS-4	1.98	[1.29, 2.90]	1.98	[1.29, 2.89]	1.98	[1.29, 2.90]
	BS-12	1.66	[0.81, 2.80]	1.66	[0.81, 2.80]	1.66	[0.81, 2.80]
$\sigma_{h2}^2 = 6.0$	BS-4	6.02	[4.11, 8.55]	6.04	[4.13, 8.56]	6.12	[4.20, 8.64]
	BS-12	7.12	[5.05, 9.84]	7.11	[5.04, 9.86]	7.10	[5.03, 9.86]
$\sigma_{E1}^2 = 20.0$	BS-4	19.16	[17.49, 20.92]	19.17	[17.48, 20.95]	19.15	[17.55, 20.86]
$\sigma_{E1}^2 = 16.4$	BS-12	16.61	[15.09, 18.25]	16.61	[15.09, 18.24]	16.61	[15.09, 18.24]
$\sigma_{E12}^2 = 5.3$	BS-4	5.73	[4.48, 6.99]	5.74	[4.51, 6.99]	5.73	[4.51, 6.98]
$\sigma_{E12}^2 = 4.4$	BS-12	4.71	[3.72, 5.74]	4.70	[3.72, 5.73]	4.71	[3.72, 5.74]
$\sigma_{E2}^2 = 26.7$	BS-4	27.22	[25.27, 29.36]	27.24	[25.28, 29.38]	27.19	[25.35, 29.14]
$\sigma_{E2}^2 = 21.8$	BS-12	21.60	[20.38, 22.89]	21.59	[20.37, 22.87]	21.58	[20.37, 22.86]

¹ Used to simulate data, ² Used in analysis of simulated data, ³ Posterior mean, ⁴ 95% equal-tailed posterior probability interval based on the 2.5th and 97.5th percentiles of the posterior density

components were estimated by using $\hat{\mathbf{R}}_E = \frac{\hat{\mathbf{v}}}{\hat{\mathbf{v}} - 1} \hat{\mathbf{R}}_0$

within BS-4 and BS-12 populations. As seen from the formula, marginal error (co)variances are the function of degrees of freedom and true marginal error (co)variances are different across populations (BS-4 and BS-12). Inferences on marginal error (co)variance components using the BSt, BS and BN fitted models were found to be similar within BS-4 and BS-12 populations and true marginal error parameter values were covered by 95% equal-tailed PPI of parameters (Table 3).

DISCUSSION

Bayesian techniques are capable of fitting models where residuals have a heavy-tailed (Student's-t or Slash) distribution with unknown degrees of freedom. Although Student's-t distribution within heavy-tailed densities is used as a viable alternative to normal distribution in linear mixed effects models, there is not enough study about Slash distribution as a robust alternative to normal distribution. This article discusses a MCMC Bayesian implementation of bivariate linear mixed effects model

with Slash distributed errors as a robust alternative to normally or Student's-*t* distributed linear models. The objectives were to obtain inference on robustness parameter (degrees of freedom) and variance components. Our approach was illustrated with Slash distributed bivariate datasets simulated based on 4 or 12 degrees of freedom and the goodness of fit of the normal, Student's-*t* and Slash distributions was investigated for these datasets.

Model comparisons, using predictive log likelihood (PLL), typically favoured the heavy-tailed models over the normal model for BS-4 datasets, but not for BS-12 datasets. There was little difference in PLL between Student's-*t* and Slash approaches.

Inferences on degrees of freedom parameter from BSt fitted model in BS-4 and BS-12 populations indicated that the values of 4 and 12 of degrees of freedom in BS model corresponded to the values of 32 and 142 of degrees of freedom in BSt model. These results showed that heavy-tailed datasets could be simulated using lower degrees of freedom than four in BS models, and there were no difference between BS and BN models when the degrees of freedom approached to 12 in BS models.

Despite the difference between the fitted models in terms of PLL, the benefits of heavy-tailed models were only modest in improving the correlation between true and predicted sire merit. However, the simulated data represented better-balanced data in terms of the distribution of offspring by gender and herd than would often be the case in field data. Further, all sires had equal numbers of offspring, and these numbers were sufficient for reliable assessment of sires in all models. It would be

expected that the superiority of heavy-tailed methods would increase in applications to field data.

Estimates of marginal error (co)variance components using the BS fitted model were biased up in some situations. The extra effort to treat degrees of freedom as an unknown was negligible. These methods need to be further explored with field data, and in other animal breeding applications such as QTL detection and genomic selection.

ACKNOWLEDGEMENTS

This project was supported by grant TUBITAK TOVAG-1070915 from the Scientific and Technological Research Council of Turkey (Project coordinator: Dr. Kadir KIZILKAYA). We are grateful to I. Misztal for making available Sparsem90 and Fspak90.

REFERENCES

1. **Lange K, Sinsheimer JS:** Normal/independent distributions and their applications in robust regression. *J Am Stat Assoc*, 2, 175-198, 1993.
2. **Rosa GJM, Padovani CR, Gianola D:** Robust linear mixed models with normal/ independent distributions and Bayesian MCMC implementation. *Biom J*, 45, 573-590, 2003.
3. **Geman D, Geman S:** Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell*, 6, 721-741, 1984.
4. **Gelfand AE, Smith AFM:** Sampling based approaches to calculating marginal densities. *J Am Stat Assoc*, 85, 398-409, 1990.
5. **Gelfand AE:** Model determination using sampling-based methods. **In**, Markov Chain Monte Carlo in Prac-tice. pp. 145-161, Chapman & Hall, London, U.K, 1996.
6. **Raftery AE:** Hypothesis testing and model selection. **In**, Markov Chain Monte Carlo in Prac-tice. pp. 163-187, Chapman & Hall, London, U.K, 1996.