Statistics and Probability Letters 85 (2014) 91–97

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On the independence Jeffreys prior for skew-symmetric models

Francisco Javier Rubio [a,*], Brunero Liseo [b]

[a] *University of Warwick, Department of Statistics, Coventry, CV4 7AL, UK*
[b] *MEMOTEF, Sapienza Università di Roma, Italy*

ARTICLE INFO

Article history:
Received 24 August 2013
Received in revised form 21 November 2013
Accepted 22 November 2013
Available online 2 December 2013

Keywords:
Posterior existence
Scale mixtures of normals
Skewness

ABSTRACT

We study the Jeffreys prior of the skewness parameter of a general class of scalar skew-symmetric models. We show that this prior is symmetric, proper, and with tails $O(|\lambda|^{-3/2})$ under mild regularity conditions. We also calculate the independence Jeffreys prior for the case with unknown location and scale parameters, and investigate conditions for the propriety of the corresponding posterior distribution.

1. Introduction

The need for modelling data presenting departures from symmetry has fostered the development of distributions that can capture skewness. A popular method to produce this sort of distributions consists of adding a parameter that controls skewness to a symmetric distribution. In this line, Azzalini (1985) proposed a transformation to produce an asymmetric normal density, termed *skew-normal*, as follows

$$\text{sn}(y;\mu,\sigma,\lambda) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\lambda\frac{y-\mu}{\sigma}\right), \tag{1}$$

where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$, $\lambda \in \mathbb{R}$, ϕ is the standard normal probability density function (PDF), and Φ is the standard normal cumulative distribution function (CDF). The parameter λ is often interpreted as a skewness parameter given that the density (1) is asymmetric for $\lambda \neq 0$, and it reduces to the normal PDF for $\lambda = 0$. Subsequently, Wang et al. (2004) showed that, in particular, this method can be extended to any continuous symmetric density f, with support on $\mathbb{R}$ and mode at 0, through the transformation

$$\text{ss}(y;\mu,\sigma,\lambda) = \frac{2}{\sigma}f\left(\frac{y-\mu}{\sigma}\right)\pi\left(\lambda\frac{y-\mu}{\sigma}\right), \tag{2}$$

where π, termed the *skewing function*, is a function that satisfies $0 \le \pi(y) \le 1$, and $\pi(-y) = 1-\pi(y)$. It follows, then, that any symmetric CDF can be used as a skewing function. Several choices for f and π have been explored in the literature, such as the power exponential distribution with power $\delta \in \mathbb{R}_+$ (Azzalini, 1986), the Student-t distribution with $\nu \in \mathbb{R}_+$ degrees

* Corresponding author.
E-mail addresses: Francisco.Rubio@warwick.ac.uk (F.J. Rubio), brunero.liseo@uniroma1.it (B. Liseo).

0167-7152/$ – see front matter
http://dx.doi.org/10.1016/j.spl.2013.11.012

of freedom (Azzalini and Capitanio, 2003), the logistic distribution (Nadarajah, 2009), among others. Distributions obtained by means of this method are called *skew-symmetric* distributions. These distributions are widely used nowadays in several contexts such as binary regression (Bazán et al., 2010), meta-analysis (Guolo, 2013), data fitting (Branco et al., 2012), among many others.

It has been found that several skew-symmetric models present inferential issues. For instance, Azzalini (1985) showed that the Fisher information matrix of the parameters (μ, σ, λ) is singular at $\lambda = 0$ for the skew-normal sampling model. In addition, the maximum likelihood estimator of the parameter λ can be ∞ with positive probability. The cases with infinite estimators are more commonly found in small and moderate samples. These inferential issues are present in other skew-symmetric models (Hallin and Ley, 2012). Some authors have proposed the use of the Bayesian approach in order to avoid these inferential problems (Liseo and Loperfido, 2006; Branco et al., 2012). In Bayesian practice it is often of interest to employ *noninformative* priors given that they typically produce posterior inference with appealing frequentist properties. However, due to the singularity of the Fisher information matrix at $\lambda = 0$ of some skew-symmetric models, the use of the Jeffreys-rule prior, which is defined as the square root of the determinant of the Fisher information matrix, has been avoided in this kind of models. In addition, the calculation of this sort of prior is typically cumbersome. Reference priors, which are another kind of noninformative priors, have been studied for the skew-normal and the skew Student-t models in Liseo and Loperfido (2006) and Branco et al. (2012). An alternative noninformative prior is the independence Jeffreys prior. This prior is constructed as the product of the Jeffreys priors for each parameter, while treating the remaining parameters as fixed.

In this paper, we study the independence Jeffreys prior associated to the class of skew-symmetric distributions obtained by using a CDF as a skewing function in (2). In Section 2, we analyse the Jeffreys prior of the skewness parameter λ in skew-symmetric models without location and scale parameters. We show that this prior is proper, symmetric about 0, and with tails $O(|\lambda|^{-\frac{3}{2}})$ under rather mild regularity conditions. Using these results, we construct the independence Jeffreys prior for the general model with location and scale parameters. In Section 3 we obtain easy to check sufficient conditions for the propriety of the posterior distribution when the sampling model f in (2) belongs to the family of scale mixtures of normal distributions. The case of samples containing censored observations is covered as well. In Section 4, we present the use of these results on the skew-logistic distribution. We conclude with some discussion and extensions of this work in Section 5.

2. Independence Jeffreys prior for univariate skew-symmetric models

Throughout we focus on the study of skew-symmetric models of the type

$$s(y; \mu, \sigma, \lambda) = \frac{2}{\sigma} f\left(\frac{y-\mu}{\sigma}\right) G\left(\lambda \frac{y-\mu}{\sigma}\right), \tag{3}$$

where f is a continuous symmetric density function with support on $\mathbb{R}$, and G is a CDF with continuous symmetric density g with support on $\mathbb{R}$. This structure covers many cases of practical interest such as the skew-normal distribution (Azzalini, 1985), the skew-t distribution (Azzalini and Capitanio, 2003), the skew-logistic distribution (Nadarajah, 2009), among many others.

Consider first the particular skew-symmetric model (3) without location and scale parameters, this is, assuming that $\mu = 0$ and $\sigma = 1$. Recall that the Fisher information of the parameter λ is defined as

$$I(\lambda) = \int_{\mathbb{R}} \left[\frac{\partial \log s(y; 0, 1, \lambda)}{\partial \lambda}\right]^2 s(y; 0, 1, \lambda) dy.$$

The Jeffreys prior of the parameter λ is defined, up to a proportionality constant, as the square root of the Fisher information $I(\lambda)$, this is, $\pi(\lambda) \propto \sqrt{I(\lambda)}$. The following result characterises the cases where this prior is well-defined at $\lambda = 0$.

Remark 1. The Fisher information of λ associated to model (3), and consequently the Jeffreys prior of λ, is well-defined at $\lambda = 0$ if and only if the second moment of f exists.

Proof. See Appendix.

Particular cases of Remark 1 have already been reported in the literature. For instance, Branco et al. (2012) report the presence of a pole at $\lambda = 0$ in the Jeffreys prior of λ for the skew Student-t model with $\nu \leq 2$ degrees of freedom. Remark 1 shows that this feature is present in many other skew-symmetric models, and that this sort of singularity is linked to the existence of the moments of the underlying symmetric density f.

Liseo and Loperfido (2006) and Branco et al. (2012) show that the Jeffreys priors of the parameter λ, for the cases where f and g are normal or Student-t distributions, are proper, decreasing in $|\lambda|$, and with tails $O(|\lambda|^{-\frac{3}{2}})$. Their proofs rely upon basic properties of these models, which suggests that there may be other models that lead to a Jeffreys prior of λ with the same properties under some reasonable regularity conditions. In order to establish this result, we introduce the following set of sufficient conditions.