Moment matching priors for non-regular models

Shintaro Hashimoto
Department of Mathematics, Hiroshima University

March 28, 2019

Abstract

This paper presents moment matching priors for non-regular models whose supports depend on an unknown parameter. Both one-parameter and multi-parameter cases are considered. The resulting priors are given by matching the posterior mean and bias-adjusted maximum likelihood estimator up to the higher order. Some examples of proposed priors are also given.

Keywords: Maximum likelihood estimator; Moment matching priors; Non-regular distributions; Objective priors

1 Introduction

In Bayesian inference, the selection of priors has been an important and much discussed problem. When we have a little prior information or it is required the objectivity of data analysis, we often use 'objective priors' or 'default priors'. Then we are often faced with a problem of the selection of an appropriate objective prior in a given context. For a regular family of distributions, the Jeffreys prior which is proportional to the positive square root of the Fisher information number is widely used as the objective prior. It is also known that the Jeffreys prior is invariant under smooth one-to-one transformation. Another class of objective priors is the reference prior, which was proposed by Bernardo (1979). The reference prior is defined by maximizing the expected Kullback-Leibler divergence between the prior and the posterior under some regularity conditions. The probability matching prior proposed by Welch and Peers (1963) is also known as an objective prior (see also Tibshirani (1989) and Datta and Mukerjee (2004)). These priors match the Bayesian credible intervals with the corresponding frequentist coverage probabilities, either exactly or approximately. Since the situations in which there exist exact probability matching priors are very limited, we often focus on approximating them based on the asymptotic theory of the maximum likelihood estimator under some regularity conditions. Recently, Ghosh and Liu (2011) derived the priors which are based on the moment matching criterion for regular one-parameter and multi-parameter family of distributions. Moment matching criterion leads the prior which is higher order matching of the moment of the posterior and the maximum likelihood estimator (MLE), and such priors are called the moment matching priors. Therefore the moment matching prior leads to the posterior mean which shares the asymptotic optimality of the MLE's up to the higher order. As stated in Ghosh and Liu (2011), if one is interested in asymptotic bias or mean squared error reduction of the MLE's through some adjustment, the same adjustment applies directly to the posterior means. In this sense, it is possible to achieve Bayesian-frequentist synthesis of point estimates. Interestingly, Ghosh and Liu (2011) showed that the moment matching prior is different from the Jeffreys or probability matching priors in regular cases.

However, these objective priors strongly depend on the regularity of statistical models and cannot be applied for non-regular distributions which does not satisfy regularity conditions such as the models with parameter-dependent supports. These non-regular models have been appeared in many practical situations. For examples, the auction and search models in structural econometric models have a jump in the density and the jump is very informative about the parameters. In such non-regular cases, for example, the asymptotic normality of the posterior distribution does not hold. In non-regular cases, the reference priors in the sense of Bernardo (1979) were obtained by Ghosal and Samanta (1997a) for one-parameter case and by Ghosal (1997) for multi-parameter case in the presence of nuisance parameter. Ghosal (1999) derived the probability matching prior for both one-parameter and multi-parameter non-regular cases and made comparison with the corresponding reference priors. Furthermore, Wang and Sun (2012) derived the objective prior for another type of non-regular model.

In this paper, we deal with the same non-regular models as those of Ghosal (1999) and derive the moment matching priors which match the posterior mean and bias-adjusted MLE in such models. Both one-parameter and multi-parameter cases are considered. The resulting priors are given by solving certain differential equations. Further, we show some properties of the resulting prior and make comparison with the corresponding reference or probability matching priors for non-regular case.

This paper is organized as follows: In Section 2, we give the moment matching prior for one-parameter case by using the higher order asymptotic expression of posterior. In Section 3, we extend the result in Section 2 to multi-parameter case in the presence of nuisance parameter. In Section 4, we give some examples of the proposed priors.

2 Moment matching prior for one parameter non-regular model

Let $X_1, \ldots, X_n$ be independent and identically distributed observations from a density $f(x;\theta)$ ($\theta \in \Theta \subset \mathbb{R}$) with respect to the Lebesgue measure. We assume that for all $\theta \in \Theta$, $f(x;\theta)$ is strictly positive in a closed interval $S(\theta) := [a_1(\theta), a_2(\theta)]$ depending on unknown parameter θ and is zero outside $S(\theta)$. It is permitted that one of the endpoints is free from θ and may be plus or minus infinity. In order to ensure the validity of the expansion of the posterior density, we assume the conditions (A1)–(A5) and (A6) with $r = 1$ in Ghosal and Samanta (1997b) on $a_1(\cdot)$, $a_2(\cdot)$ and $f(x;\theta)$. Some examples which belong to this family are given in Section 4.

Let π be a prior density of θ, and we assume that π is twice differentiable. Without loss of generality, we assume that $S(\theta)$ is monotone decreasing with respect to θ, that is, we assume that $a_1(\cdot)$ is monotone increasing and $a_2(\cdot)$ is monotone decreasing. Indeed, the case where $S(\theta)$ increases with θ may be reduced to the case where $S(\theta)$ decreases by the reparametrization $\theta \mapsto (-\theta)$. By the assumption of $S(\theta)$, the MLE of θ is given by

$$\hat{\theta}_n := \min\{a_1^{-1}(X_{(1)}), a_2^{-1}(X_{(n)})\},$$

where $X_{(1)} := \min_{1\le i\le n} X_i$ and $X_{(n)} := \max_{1\le i\le n} X_i$. Note that it holds that $\hat{\theta}_n - \theta = O_p(n^{-1})$ $(n \to \infty)$, rather than $O_p(n^{-1/2})$ as in regular cases. Hereafter, we often omit the argument '$n \to \infty$' for simplicity. Define

$$\sigma := \frac{1}{n}\sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(X_i; \hat{\theta}_n).$$

Note that it holds that $\sigma - c(\theta) = O_p(n^{-1})$ (Lemma 2.1 in Ghosal and Samanta (1997b)), where

$$c(\theta) := \mathrm{E}\left[\frac{\partial}{\partial\theta} \log f(X_i; \theta)\right] = a_1'(\theta)f(a_1(\theta); \theta) - a_2'(\theta)f(a_2(\theta); \theta).$$

is the expectation of the score function. Since $S(\theta)$ is monotone decreasing, we note that it holds $c(\theta) > 0$. By Theorem 3.1 of Ghosal and Samanta (1997b), the posterior density of the normalized random variable $u = n\sigma(\theta - \hat{\theta}_n)$ given $\boldsymbol{X} = (X_1, \ldots, X_n)$ has the stochastic expansion

$$\pi(u|\boldsymbol{X}) = e^u \left[1 + \frac{1}{n}\left\{\frac{\pi'(\hat{\theta}_n)}{\sigma\pi(\hat{\theta}_n)}(u+1) + \frac{c_2}{\sigma^2}(u^2 - 2)\right\} + O_p(n^{-2})\right] \tag{2.1}$$

for $u < 0$, where

$$c_2 := \frac{1}{2n}\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \log f(X_i, \hat{\theta}_n).$$

From (2.1) we note that the first order asymptotic posterior is the exponential distribution. We find a prior such that the difference between the resulting posterior mean and the bias-adjusted maximum likelihood estimator converges to zero up to the order of $O_p(n^{-3})$. From (2.1) we have the following theorem.

Theorem 2.1. *Let $\hat{\theta}^B_{n,\pi}$ be the posterior mean of θ under the prior $\pi(\theta)$ and $\hat{\theta}^*_n = \hat{\theta}_n - \{1/(\sigma n)\}$ be the bias-adjusted MLE of θ. Then we have*

$$\hat{\theta}^B_{n,\pi} - \hat{\theta}^*_n = \frac{1}{n^2}\left\{\frac{1}{\sigma^2}\left(\frac{\pi'(\hat{\theta}_n)}{\pi(\hat{\theta}_n)} - \frac{4c_2}{\sigma}\right)\right\} + O_p\left(\frac{1}{n^3}\right). \tag{2.2}$$

The proof of this theorem is omitted because the proof can be make the use of the moment of the exponential distribution and transformation of variable. As we mentioned before, it holds that $\hat{\theta}_n - \theta = O_p(n^{-1})$ and $\sigma - c(\theta) = O_p(n^{-1})$. Also, by the law of large number and the condition (A6) with $r = 1$ in Ghosal and Samanta (1997b), we can show that $c_2 - (d(\theta)/2) =$

$O_p(n^{-1})$. By using these facts and continuity of the prior density $\pi(\theta)$, we have

$$n^2(\hat{\theta}^B_{n,\pi} - \hat{\theta}^*_n) \xrightarrow{p} \frac{1}{c(\theta)^2}\left(\frac{\pi'(\theta)}{\pi(\theta)} - \frac{2d(\theta)}{c(\theta)}\right), \tag{2.3}$$

where $d(\theta) := \mathrm{E}[(\partial^2/\partial\theta^2)\log f(X_i;\theta)]$ and we assume that $d(\theta) < \infty$. If we now choose the prior as

$$\pi(\theta) = \exp\left[2\int^\theta \frac{d(t)}{c(t)}\mathrm{d}t\right], \tag{2.4}$$

the right-hand side of (2.3) equals zero. Hence, we obtain $\hat{\theta}^B_{n,\pi} - \hat{\theta}^*_n = O_p(n^{-3})$ under the prior (2.4). In other words, this prior leads to the posterior mean which matches with the bias-adjusted MLE's up to the order of $O_p(n^{-3})$. We will denote this prior as $\pi_M(\theta)$ and call it the moment matching prior for θ. We note that it is not clear whether the prior $\pi_M(\theta)$ is proper or not in general. Proprieties of the prior and the corresponding posterior are discussed in Section 4 through some specific examples. The prior $\pi_M(\theta)$ is different from the reference prior $\pi(\theta) \propto c(\theta)$ for non-regular case given by Ghosal and Samanta (1997a) or the probability matching prior $\pi(\theta) \propto c(\theta)$ for non-regular case given by Ghosal (1999). We note that if $S(\theta)$ is monotone increasing, these priors are $\pi(\theta) \propto |c(\theta)|$. We have the following theorem concerning the one-to-one transformation of θ.

Theorem 2.2. *Let $\pi_M(\theta)$ be a moment matching prior for θ given by* (2.4) *and η is a one-to-one function of θ. Then we have*

$$\pi^*_M(\eta) = \pi_M(\theta)\left(\frac{\mathrm{d}\theta}{\mathrm{d}\eta}\right)^2. \tag{2.5}$$

Proof. Let $g(x;\eta)$ be a reparametrized density. Define

$$c(\eta) = \mathrm{E}\left[\frac{\partial}{\partial\eta}\log g(X_i;\eta)\right] \quad \text{and} \quad d(\eta) = \mathrm{E}\left[\frac{\partial^2}{\partial\eta^2}\log g(X_i;\eta)\right].$$

From (2.4) we obtain

$$\pi_M(\eta) = \exp\left[2\int \frac{d(\eta)}{c(\eta)}\mathrm{d}\eta\right].$$

We note that it holds that $c(\eta) = c(\theta)(\mathrm{d}\theta/\mathrm{d}\eta)$. By the chain rule of differentiation, we have

$$d(\eta) = d(\theta)\left(\frac{\mathrm{d}\theta}{\mathrm{d}\eta}\right)^2 + c(\theta)\left(\frac{\mathrm{d}^2\theta}{\mathrm{d}\eta^2}\right).$$

Hence, we have

$$\pi^*_M(\eta) = \exp\left[2\int \frac{d(\eta)}{c(\eta)}\mathrm{d}\eta\right]$$

$$= \exp\left[2\int \frac{d(\theta)(\mathrm{d}\theta/\mathrm{d}\eta)^2}{c(\theta)(\mathrm{d}\theta/\mathrm{d}\eta)}\mathrm{d}\eta\right] \cdot \exp\left[2\int \frac{\mathrm{d}^2\theta/\mathrm{d}\eta^2}{\mathrm{d}\theta/\mathrm{d}\eta}\mathrm{d}\eta\right]$$
$$= \pi_M(\theta)\exp\left[2\log\left(\frac{\mathrm{d}\theta}{\mathrm{d}\eta}\right)\right] = \pi_M(\theta)\left(\frac{\mathrm{d}\theta}{\mathrm{d}\eta}\right)^2.$$

This completes the proof. □

Unfortunately, the moment matching prior $\pi_M(\theta)$ is not invariant under smooth one-to-one transformation. However, in the model considered in this paper, we are mainly interested in the estimation of θ not other parametrizations, e.g., θ^2, $\sqrt{\theta}$, and so on. If we are interested in the estimation of θ^2 or $\sqrt{\theta}$, then we have to derive other moment matching prior for θ^2 or $\sqrt{\theta}$. Although the invariance of smooth one-to-one transformation is desirable property for the prior distribution, there exists objective priors which are not necessary to have invariant property. For example, Datta and Ghosh (1996) showed that the reverse reference prior is not invariant under smooth one-to-one transformation (for details, see Datta and Ghosh (1996)). In a regular parametric model, Ghosh and Liu (2011) derived the moment matching prior

$$\pi_M(\theta) = \exp\left[-\frac{1}{2}\int^{\theta} \frac{g_3(t)}{I(t)}\mathrm{d}t\right], \tag{2.6}$$

where $I(\theta) = -d(\theta) < \infty$ and $g_3(\theta) = \mathrm{E}[(\partial^3/\partial\theta^3)\log f(X_i;\theta)] < \infty$. They also showed that the prior (2.6) is not invariant under smooth one-to-one transformation. It may be interesting to find conditions which the moment matching prior corresponds to the reference (or probability matching) prior $\pi(\theta) \propto c(\theta)$ for non-regular case or the uniform prior $\pi(\theta) \propto$ constant. The former holds if and only if $d(\theta) = c'(\theta)/2$, while the later holds if and only if $d(\theta) = 0$. Some examples of the moment matching prior by (2.4) are given in Section 4.

3 Moment matching priors for multi-parameter non-regular model

We now consider an additional parameter φ, and consider the parametric model $f(x;\theta,\varphi)$. We suppose that φ is the regular parameter, that is, we assume that the model is regular parametric family when the non-regular parameter θ is known. For simplicity, we give the result in the case of scalar φ. The multi-dimensional extension of φ may also be treated in the same manner. Let $\pi(\theta,\varphi)$ be the joint prior density of (θ,φ), and we assume that $\pi(\theta,\varphi)$ is piecewise differential in θ and φ up to the third order. Further, we assume that $f(x;\theta,\varphi)$ is piecewise differentiable in θ and φ up to the forth order. Let $\hat{\varphi}_n$ be a solution of the the modified likelihood equation

$$\sum_{i=1}^{n} \frac{\partial}{\partial\varphi}\log f(X_i;\hat{\theta}_n,\hat{\varphi}_n) = 0. \tag{3.1}$$

Smith (1985) showed the consistency for the special case when θ is a location parameter, but the argument can easily be generalized. Hence, we may assume that $(\hat{\theta}_n,\hat{\varphi}_n)$ is consistent.

We put

$$\sigma := \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i;\hat{\theta}_n,\hat{\varphi}_n), \quad b^2 := -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\varphi^2}\log f(X_i;\hat{\theta}_n,\hat{\varphi}_n)$$

and we note that $\sigma \to c(\theta,\varphi)$ and $b^2 \to \lambda^2(\theta,\varphi)$ almost surely, where

$$c(\theta,\varphi) := \mathrm{E}\left[\frac{\partial}{\partial\theta}\log f(X;\theta,\varphi)\right], \quad \lambda^2(\theta,\varphi) := \mathrm{E}\left[-\frac{\partial^2}{\partial\varphi^2}\log f(X;\theta,\varphi)\right].$$

When $S(\theta)$ is monotone decreasing, we can show that $c(\theta,\varphi) > 0$. Hereafter, we may assume that $c(\theta,\varphi) > 0$. Let $u := n\sigma(\theta - \hat{\theta}_n)$ and $v := \sqrt{n}b(\varphi - \hat{\varphi}_n)$ be normalized random variables of θ and φ, respectively. From Appendix in Ghosal (1999) the joint posterior density of (u,v) given $\boldsymbol{X} = (X_1,\ldots,X_n)$ has the stochastic expansion up to the order $O_p(n^{-3/2})$

$$\pi(u,v|\boldsymbol{X}) = \frac{1}{\sqrt{2\pi}}e^{u-(v^2/2)}\left\{1 + \frac{1}{\sqrt{n}}D_1 + \frac{1}{n}D_2 + O_p(n^{-3/2})\right\} \tag{3.2}$$

for $u < 0$, where

$$\begin{aligned}
D_1 =& \frac{\hat{\pi}_{01}}{\hat{\pi}_{00}b}v + \frac{2a_{11}}{\sigma b}uv + \frac{a_{03}}{b^3}v^3, \\
D_2 =& \frac{\hat{\pi}_{10}}{\hat{\pi}_{00}\sigma}(u+1) + \frac{\hat{\pi}_{02}}{2\hat{\pi}_{00}b^2}(v^2-1) + \frac{a_{20}}{\sigma^2}(u^2-2) \\
&+ \frac{2(\hat{\pi}_{01}/\hat{\pi}_{00})a_{11} + 3a_{12}}{\sigma b^2}(uv^2+1) + \frac{\hat{\pi}_{01}a_{03}}{\hat{\pi}_{00}b^4}(v^4-3) \\
&+ \frac{2a_{11}^2}{\sigma^2 b^2}(u^2v^2-2) + \frac{2a_{11}a_{03}}{\sigma b^4}(uv^4+3) + \frac{a_{03}^2}{b^6}(v^6-15)
\end{aligned}$$

with

$$\hat{\pi}_{rs} = \frac{\partial^{r+s}}{\partial\theta^r\partial\varphi^s}\pi(\hat{\theta}_n,\hat{\varphi}_n), \quad a_{rs} = \frac{1}{(r+s)!\,n}\sum_{i=1}^{n}\frac{\partial^{r+s}}{\partial\theta^r\partial\varphi^s}\log f(X_i;\hat{\theta}_n,\hat{\varphi}_n)$$

for $r,s = 0,1,2,\ldots$, and note that $a_{rs} \to A_{rs}(\theta,\varphi)$ almost surely, where

$$A_{rs}(\theta,\varphi) = \frac{1}{(r+s)!}\mathrm{E}\left[\frac{\partial^{r+s}}{\partial\theta^r\partial\varphi^s}\log f(X_i;\theta,\varphi)\right]$$

for $r,s = 0,1,2,\ldots$. Note that $c(\theta,\varphi) = A_{10}$ and $\lambda^2(\theta,\varphi) = -2A_{02}$. From (3.2) we can find that the random variables u and v are the first order asymptotic independent and their first order asymptotic marginal posterior distributions are the exponential and the normal distributions, respectively. From (3.2) we can obtain the second order asymptotic marginal posterior densities $\pi(u|\boldsymbol{X})$ and $\pi(v|\boldsymbol{X})$. The second order asymptotic marginal posterior density of u is given by

$$\pi(u|\boldsymbol{X}) = e^{u}\left[1+\frac{1}{n}\left\{\left(\frac{\hat{\pi}_{10}}{\hat{\pi}_{00}\sigma}+\frac{2(\hat{\pi}_{01}/\hat{\pi}_{00})a_{11}+3a_{12}}{\sigma b^2}+\frac{6a_{11}a_{03}}{\sigma b^4}\right)(u+1)\right.\right.$$
$$\left.\left.+\left(\frac{a_{20}}{\sigma^2}+\frac{2a_{11}^2}{\sigma^2 b^2}\right)(u^2-2)\right\}+O_p\left(n^{-2}\right)\right] \tag{3.3}$$

for $u<0$, while that of v is given by

$$\pi(v|\boldsymbol{X}) = \frac{1}{\sqrt{2\pi}}e^{-v^2/2}\left[1+\frac{1}{\sqrt{n}}\left\{\left(\frac{\hat{\pi}_{01}}{\hat{\pi}_{00}b}-\frac{2a_{11}}{\sigma b}\right)v+\frac{a_{03}}{b^3}v^3\right\}+O_p\left(n^{-1}\right)\right]. \tag{3.4}$$

First, we assume that θ is the parameter of interest and φ is the nuisance parameter. From the asymptotic expansion of marginal posterior (3.3) we have the following theorem.

Theorem 3.1. *Let $\hat{\theta}_n^* = \hat{\theta}_n - \{1/(\sigma n)\}$ be the bias-adjusted MLE of θ. The marginal posterior mean $\hat{\theta}_{n,\pi}^B$ under the prior $\pi(\theta,\varphi)$ is expressed by*

$$\hat{\theta}_{n,\pi}^B - \hat{\theta}_n^* = \frac{1}{n^2}\left\{\left(\frac{\hat{\pi}_{10}}{\hat{\pi}_{00}\sigma}+\frac{2(\hat{\pi}_{01}/\hat{\pi}_{00})a_{11}+3a_{12}}{\sigma b^2}+\frac{6a_{11}a_{03}}{\sigma b^4}\right)\right.$$
$$\left.-4\left(\frac{a_{20}}{\sigma^2}+\frac{2a_{11}^2}{\sigma^2 b^2}\right)\right\}+O_p\left(n^{-3}\right)\Bigg]. \tag{3.5}$$

The proof of theorem is omitted for the same reason as Theorem 2.1. From (3.5), by the law of large number and consistency of $(\hat{\theta}_n,\hat{\varphi}_n)$, we have

$$n^2(\hat{\theta}_{n,\pi}^B - \hat{\theta}_n^*) \xrightarrow{p} \frac{1}{c(\theta,\varphi)}\frac{\partial}{\partial\theta}\log\pi(\theta,\varphi)+\frac{2A_{11}(\theta,\varphi)}{c(\theta,\varphi)\lambda^2(\theta,\varphi)}\frac{\partial}{\partial\varphi}\log\pi(\theta,\varphi)$$
$$+\frac{3A_{12}(\theta,\varphi)}{c(\theta,\varphi)\lambda^2(\theta,\varphi)}+\frac{6A_{11}(\theta,\varphi)A_{03}(\theta,\varphi)}{c(\theta,\varphi)\lambda^4(\theta,\varphi)}$$
$$-4\left(\frac{A_{20}(\theta,\varphi)}{c^2(\theta,\varphi)}+\frac{2A_{11}^2(\theta,\varphi)}{c^2(\theta,\varphi)\lambda^2(\theta,\varphi)}\right). \tag{3.6}$$

Then the moment matching prior $\pi_M^{\theta}(\theta,\varphi)$ when θ is the parameter of interest is the solution of the partial differential equation

$$\frac{1}{c(\theta,\varphi)}\frac{\partial}{\partial\theta}\log\pi(\theta,\varphi)+\frac{2A_{11}(\theta,\varphi)}{c(\theta,\varphi)\lambda^2(\theta,\varphi)}\frac{\partial}{\partial\varphi}\log\pi(\theta,\varphi)$$
$$=4\left(\frac{A_{20}(\theta,\varphi)}{c^2(\theta,\varphi)}+\frac{2A_{11}^2(\theta,\varphi)}{c^2(\theta,\varphi)\lambda^2(\theta,\varphi)}\right)-\frac{3A_{12}(\theta,\varphi)}{c(\theta,\varphi)\lambda^2(\theta,\varphi)}-\frac{6A_{11}(\theta,\varphi)A_{03}(\theta,\varphi)}{c(\theta,\varphi)\lambda^4(\theta,\varphi)}. \tag{3.7}$$

If we choose the prior as the solution of (3.7), the right-hand side of (3.6) equals zero. Hence, we obtain $\hat{\theta}_{n,\pi}^B - \hat{\theta}_n^* = O_p(n^{-3})$ under the prior as the solution of (3.7).

When φ is the parameter of interest we consider the moment matching prior in a similar way to the above. From the asymptotic expansion of marginal posterior (3.3) we have the following theorem.

Theorem 3.2. *Let $\hat{\varphi}_n$ be the solution of the modified likelihood equation* (3.1)*. The marginal posterior mean $\hat{\varphi}^B_{n,\pi}$ under the prior $\pi(\theta,\varphi)$ is expressed by*

$$\hat{\varphi}^B_{n,\pi} - \hat{\varphi}_n = \frac{1}{nb^2}\left(\frac{\hat{\pi}_{01}}{\hat{\pi}_{00}} - \frac{2a_{11}}{\sigma} + \frac{3a_{03}}{b^2}\right) + O_p\left(\frac{1}{n\sqrt{n}}\right). \tag{3.8}$$

The proof of theorem is omitted for the same reason as Theorem 2.1. From (3.8), by the law of large number and consistency of $(\hat{\theta}_n, \hat{\varphi}_n)$, we have

$$n(\hat{\varphi}^B_{n,\pi} - \hat{\varphi}_n) \xrightarrow{p} \frac{1}{\lambda^2(\theta,\varphi)}\left(\frac{\partial}{\partial\varphi}\log\pi(\theta,\varphi) - \frac{2A_{11}(\theta,\varphi)}{c(\theta,\varphi)} + \frac{3A_{03}(\theta,\varphi)}{\lambda^2(\theta,\varphi)}\right). \tag{3.9}$$

If we now choose the prior as

$$\pi(\theta,\varphi) = \exp\left[\int^{\varphi}\left(\frac{2A_{11}(\theta,t)}{c(\theta,t)} - \frac{3A_{03}(\theta,t)}{\lambda^2(\theta,t)}\right)\mathrm{d}t\right], \tag{3.10}$$

the right-hand side of (3.9) equals zero. Hence, we obtain $\hat{\varphi}^B_{n,\pi} - \hat{\varphi}_n = O_p(n^{-3/2})$ under the prior (3.10). We will denote this prior as $\pi^{\varphi}_M(\theta,\varphi)$ and we call it the moment matching prior when φ is the parameter of interest. In contrast to the case θ is the parameter of interest, the moment matching prior is given by explicit form when φ is the parameter of interest.

4 Examples

In this Section, we consider some examples of proposed priors, and discuss the corresponding posterior propriety through some specific examples. First, we give examples for one-parameter case.

Example 4.1 (Location family)**.** Let f_0 be a strictly positive density on $[0,\infty)$. Consider the location family of distribution $f(x,\theta) = f_0(x-\theta)$. In particular, the shifted exponential distribution $f(x;\theta) = e^{-(x-\theta)}$ $(x > \theta)$ belongs to this location family. In this case, we have

$$c(\theta) = f_0(0+), \quad d(\theta) = \iota_1,$$

where $\iota_1 = \int_0^{\infty}\{f_0''(t) - (f_0'(t)^2/f_0(t))\mathrm{d}t\}$. Hence, the moment matching prior of θ is given by

$$\pi_M(\theta) \propto \exp(\theta^{2\tau_1}),$$

where $\tau_1 = \iota_1/f_0(0+)$. For the shifted exponential distribution, we have $\tau_1 = 0$ because $f_0(0+) = 1$ and $\iota_1 = 0$. Therefore we have $\pi_M(\theta) \propto$ constant, so this is the uniform prior. In this case, the marginal density of $\boldsymbol{X} = (X_1,\ldots,X_n)$ $(n \geq 1)$ is given by

$$\begin{aligned} m(\boldsymbol{x}) = \int f(\boldsymbol{x};\theta)\pi(\theta)\mathrm{d}\theta \propto \int_{-\infty}^{x_{(1)}} e^{-\sum_{i=1}^n (x_i-\theta)}\mathrm{d}\theta =& e^{-\sum_{i=1}^n x_i}\int_{-\infty}^{x_{(1)}} e^{n\theta}\mathrm{d}\theta \\ =& \frac{1}{n}e^{-\sum_{i=1}^n (x_i - x_{(1)})} < \infty, \end{aligned}$$

where $x_{(1)} := \min_{1\le i\le n} x_i$. Although the prior $\pi(\theta) \propto 1$ is improper, the posterior distribution is proper for $n \ge 1$ in the shifted exponential case. The posterior density of θ given $\boldsymbol{X} = (X_1, \dots, X_n)$ is $\pi(\theta|\boldsymbol{X}) = n \exp\{n \sum_{i=1}^n (\theta - x_(1))\}$ $(-\infty < \theta < x_{(1)})$, and the corresponding posterior mean is given by $\hat{\theta}^B_{n,\pi} = E(\theta|\boldsymbol{X}) = X_{(1)} - (1/n)$. We can find that $\hat{\theta}^B_{n,\pi}$ is the same as the bias adjusted MLE $\hat{\theta}^*_n = \hat{\theta}_n - \{1/(\sigma n)\} = X_{(1)} - (1/n)$, and $\hat{\theta}^*_n$ is also the unique uniformly minimum variance unbiased (UMVU) estimator. Since the moment matching prior $\pi(\theta) \propto 1$ is also the probability matching prior in Ghosal (1999) and the reference prior in Ghosal and Samanta (1997a), we note that the same result is obtained under these priors.

Example 4.2 (Scale family)**.** Let f_0 be a strictly positive density on $[0, 1]$. Consider the scale family of distribution $f(x;\theta) = \theta^{-1} f_0(x/\theta)$ $(\theta > 0)$. In this case, we have

$$c(\theta) = -\frac{\iota_2}{\theta}, \quad d(\theta) = \frac{\iota_3}{\theta^2},$$

where $\iota_2 = 1 + \int_0^1 t f_0'(t)\mathrm{d}t$ and $\iota_3 = \int_0^1 [1 + t^2\{(f_0''(t)/f_0(t)) - (f_0'(t)/f_0(t))^2\} + t(f_0'(t)/f_0(t))] f_0(t)\mathrm{d}t$ are constant numbers which does not depend on θ. Hence, the moment matching prior of θ is given by

$$\pi_M(\theta) \propto \exp[(\log\theta)^{-2\tau_2}] = \theta^{-2\tau_2},$$

where $\tau_2 = \iota_3/\iota_2$. In the case of the uniform distribution $U(0, \theta)$ $(\theta > 0)$, we have $\tau_2 = 1$. Since the support of $U(0, \theta)$ is monotone increasing, we may consider the reparametrization $\theta \leftrightarrow -\theta$ as we mentioned in Section 2. Therefore the moment matching prior is given by $\pi_M(\theta) \propto \theta^{-2}$. In this case, the marginal density of $\boldsymbol{X} = (X_1, \dots, X_n)$ $(n \ge 1)$ is given by

$$m(\boldsymbol{x}) = \int f(\boldsymbol{x};\theta)\pi(\theta)\mathrm{d}\theta \propto \int_{x_{(n)}}^{\infty} \frac{1}{\theta^n} \cdot \frac{1}{\theta^2}\mathrm{d}\theta = \frac{1}{n+1} x_{(n)}^{-(n+1)} < \infty,$$

where $x_{(n)} = \max_{1\le i\le n} x_i$. Although the prior $\pi(\theta) \propto \theta^{-2}$ is improper, the posterior distribution is proper for $n \ge 1$ in the uniform case. The posterior density of θ given $\boldsymbol{X} = (X_1, \dots, X_n)$ is $\pi(\theta|\boldsymbol{X}) = (n+1) X_{(n)}^{n+1} \theta^{-(n+2)}$ $(x_{(n)} < \theta < \infty)$, and the corresponding posterior mean is given by $\hat{\theta}^B_{n,\pi} = E(\theta|\boldsymbol{X}) = (1 + n^{-1}) X_{(n)}$. We can find that $\hat{\theta}^B_{n,\pi}$ is the same as the bias adjusted MLE $\hat{\theta}^*_n = \hat{\theta}_n - \{1/(\sigma n)\} = (1 + n^{-1}) X_{(n)}$, and $\hat{\theta}^*_n$ is also the unique uniformly minimum variance unbiased (UMVU) estimator. On the other hand, the prior $\pi(\theta) \propto \theta^{-1}$ is the probability matching prior in Ghosal (1999) and is also the reference prior in Ghosal and Samanta (1997a). The posterior mean under the prior $\pi(\theta) \propto \theta^{-1}$ is given by $nX_{(n)}/(n-1)$ $(n \ge 2)$, and this is not UMVU estimator.

Example 4.3 (Truncation family)**.** Let $g(x)$ be a strictly positive density on $(0, \infty)$ and let $f(x;\theta) = g(x)/\bar{G}(\theta)$ $(x > \theta)$, where $\bar{G}(\theta) = \int_\theta^\infty g(t)\mathrm{d}t$. In this case, we have

$$c(\theta) = \frac{g(\theta)}{\bar{G}(\theta)}, \quad d(\theta) = \frac{g'(\theta)}{\bar{G}(\theta)} - \frac{g(\theta)\bar{G}'(\theta)}{\bar{G}(\theta)^2}.$$

Hence, the moment matching prior of θ is given by

$$\pi_M(\theta) = \exp\left[2\int\left(\frac{g'(\theta)}{g(\theta)} - \frac{\bar{G}'(\theta)}{\bar{G}(\theta)}\right)\mathrm{d}\theta\right] \propto \exp\left[2\log\left(\frac{g(\theta)}{\bar{G}(\theta)}\right)\right] = \left(\frac{g(\theta)}{\bar{G}(\theta)}\right)^2.$$

In particular, this family corresponds to the shifted exponential distribution when $g(x) = e^{-x}$. In this case, the moment matching prior of θ is

$$\pi_M(\theta) \propto \left(\frac{g(\theta)}{\bar{G}(\theta)}\right)^2 = \left(\frac{e^{-\theta}}{\int_\theta^\infty e^{-t}\mathrm{d}t}\right)^2 = 1.$$

Hence, $\pi_M(\theta)$ is the uniform prior which is the same as that of Example 4.1. Other examples of the truncation family are discussed in Example 4.4 and Example 4.5.

Next, we give some examples for multi-parameter case in the presence of a nuisance parameter. In the following examples, we consider the set-up of Example 4.3 where the density g also involves an additional regular parameter φ.

Example 4.4 (Shifted exponential distribution with scale)**.** Consider the shifted exponential distribution with scale parameter $\varphi \in (0,\infty)$ with the density function $f(x;\theta,\varphi) = \varphi^{-1}e^{-(x-\theta)/\varphi}$ $(x > \theta)$. In this case, we have $c(\theta,\varphi) = \varphi^{-1}$, $\lambda^2(\theta,\varphi) = \varphi^{-2}$, $A_{11}(\theta,\varphi) = -1/(2\varphi^2)$, $A_{12}(\theta,\varphi) = 1/(3\varphi^3)$, $A_{20}(\theta,\varphi) = 0$ and $A_{03}(\theta,\varphi) = 2/(3\varphi^3)$. If θ is the parameter of interest, the moment matching prior $\pi_M^\theta(\theta,\varphi)$ of (θ,φ) is given by solving the partial differential equation

$$\varphi\frac{\partial}{\partial\theta}\log\pi(\theta,\varphi) - \varphi\frac{\partial}{\partial\varphi}\log\pi(\theta,\varphi) = 3. \tag{4.1}$$

The prior $\pi(\theta,\varphi) \propto \varphi^{-3}$ is a solution of the partial differential equation (4.1) and $\pi^\theta(\theta,\varphi) \propto \varphi^{-3}$ is a moment matching prior when θ is the parameter of interest. This prior also satisfies the equation (3.10) which is the moment matching prior when φ is the parameter of interest. Hence, both cases lead to the same moment matching prior. The marginal density of $\boldsymbol{X} = (X_1,\ldots,X_n)$ $(n \geq 2)$ under the prior $\pi(\theta,\varphi) \propto \varphi^{-3}$ is given by

$$\begin{aligned} m(\boldsymbol{x}) &= \int\int f(\boldsymbol{x};\theta,\varphi)\pi(\theta,\varphi)\mathrm{d}\theta\mathrm{d}\varphi \\ &\propto \int_0^\infty\int_{-\infty}^{x_{(1)}} \varphi^{-n}e^{-(1/\varphi)\sum_{i=1}^n(x_i-\theta)}\varphi^{-3}\mathrm{d}\theta\mathrm{d}\varphi \\ &= \frac{\Gamma(n+1)}{n\{\sum_{i=1}^n(x_i - x_{(1)})\}^{n+1}} < \infty, \end{aligned}$$

where $x_{(1)} = \min_{1\leq i\leq n} x_i$ and $\Gamma(k)$ $(k > 0)$ is the gamma function defined by $\Gamma(k) = \int_0^\infty x^{k-1}e^{-x}\mathrm{d}x$. Although the prior $\pi_M^\theta(\theta,\varphi) = \pi_M^\varphi(\theta,\varphi) \propto \varphi^{-3}$ is improper, the posterior distribution is proper for $n \geq 2$. The joint posterior density of (θ,φ) given $\boldsymbol{X} = (X_1,\ldots,X_n)$

under the prior $\pi(\theta,\varphi) \propto \varphi^{-3}$ is given by

$$\pi(\theta,\varphi|\boldsymbol{x}) = \frac{1}{m(\boldsymbol{x})}\varphi^{-(n+3)}\exp\left\{-\frac{1}{\varphi}\sum_{i=1}^{n}(x_i-\theta)\right\}$$

for $-\infty < \theta < x_{(1)}$ and $0 < \varphi < \infty$. Hence, the marginal posterior density of θ and φ are given by

$$\begin{aligned}\pi(\theta|\boldsymbol{x}) =& n(n+1)\frac{\{\sum_{i=1}^{n}(x_i-x_{(1)})\}^{n+1}}{\{\sum_{i=1}^{n}(x_i-\theta)\}^{n+2}} \quad (-\infty<\theta<x_{(1)}),\\ \pi(\varphi|\boldsymbol{x}) =& \frac{\{\sum_{i=1}^{n}(x_i-x_{(1)})\}^{n+1}}{\Gamma(n+1)}\varphi^{-(n+2)}\exp\left\{-\frac{1}{\varphi}\sum_{i=1}^{n}(x_i-x_{(1)})\right\},\end{aligned}$$

respectively. They are the same as the bias-adjusted MLE $\hat{\theta}_n^* = \hat{\theta}_n - \{1/(\sigma n)\}$ for θ and the MLE $\hat{\varphi}_n$ of φ. In particular, the marginal posterior distribution of $\varphi|\boldsymbol{X}$ is the inverse gamma distribution with the shape $n+1$ and the scale $\sum_{i=1}^{n}(x_i-x_{(1)})$, that is, $\mathrm{IG}(n+1,\sum_{i=1}^{n}(x_i-x_{(1)}))$. Further, the marginal posterior means of θ and φ are expressed by

$$\begin{aligned}E(\theta|\boldsymbol{X}) =& X_{(1)} - \frac{1}{n^2}\sum_{i=1}^{n}(X_i-X_{(1)}),\\ E(\varphi|\boldsymbol{X}) =& \frac{\sum_{i=1}^{n}(X_i-X_{(1)})}{(n+1)-1} = \frac{1}{n}\sum_{i=1}^{n}(X_i-X_{(1)}),\end{aligned}$$

respectively. Next, we consider the marginal posterior means of θ and φ under the probability matching prior $\pi(\theta,\varphi)\propto\varphi^{-2}$ in Ghosal (1999) (or the reference prior in Ghosal and Samanta (1997a)). In a similar way to the case of the moment matching prior, we have

$$\begin{aligned}E(\theta|\boldsymbol{X}) &= X_{(1)} - \frac{1}{n(n-1)}\sum_{i=1}^{n}(X_i-X_{(1)}),\\ E(\varphi|\boldsymbol{X}) &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i-X_{(1)}).\end{aligned}$$

We note that the UMVU estimators of θ and φ are given by $\hat{\theta}_{\mathrm{UMVU}} = X_{(1)} - n^{-1}(n-1)^{-1}\sum_{i=1}^{n}(X_i-X_{(1)})$ and $\hat{\varphi}_{\mathrm{UMVU}} = (n-2)^{-1}\sum_{i=1}^{n}(X_i-X_{(1)})$, respectively.

Example 4.5 (Truncated Weibull distribution)**.** Consider the truncated Weibull distribution with the scale parameter $\varphi>0$ and the shape parameter $\alpha>0$ with the density function $f(x;\theta,\varphi) = \alpha\varphi^{\alpha}x^{\alpha-1}\exp\{-\varphi^{\alpha}(x^{\alpha}-\theta^{\alpha})\}$ $(x>\theta)$. We assume that the shape parameter $\alpha>0$ is known and $\theta>0$ in this example, and consider moment matching priors for (θ,φ). In this case, we have $c(\theta,\varphi)=\alpha\varphi^{\alpha}\theta^{\alpha-1}$, $\lambda^2(\theta,\varphi)=\alpha^2/\varphi^2$, $A_{11}(\theta,\varphi)=(1/2)\alpha^2\varphi^{\alpha-1}\theta^{\alpha-1}$, $A_{12}(\theta,\varphi)=(1/6)\alpha^2(\alpha-1)\varphi^{\alpha-2}\theta^{\alpha-1}$, $A_{20}(\theta,\varphi)=(1/2)\alpha(\alpha-1)\varphi^{\alpha}\theta^{\alpha-2}$ and $A_{03}(\theta,\varphi)=-\alpha^2(\alpha-3)/(6\varphi^3)$. If θ is the parameter of interest, the moment matching prior $\pi_M^{\theta}(\theta,\varphi)$ of

(θ, φ) is given by solving the partial differential equation

$$\varphi^{-\alpha}\theta^{1-\alpha}\frac{\partial}{\partial\theta}\log\pi(\theta,\varphi)+\varphi\frac{\partial}{\partial\varphi}\log\pi(\theta,\varphi)=2(\alpha-1)\varphi^{-\alpha}\theta^{-\alpha}+2\alpha-1. \tag{4.2}$$

The prior $\pi(\theta, \varphi) \propto \theta^{2(\alpha-1)}\varphi^{2\alpha-1}$ is a solution of the partial differential equation (4.2) and $\pi_M^\theta(\theta, \varphi) \propto \theta^{2(\alpha-1)}\varphi^{2\alpha-1}$ is a moment matching prior when θ is the parameter of interest. In this case, the marginal density of $\boldsymbol{X} = (X_1, \ldots, X_n)$ is given by

$$\begin{aligned}
m(\boldsymbol{x}) =& \int\int f(\boldsymbol{x};\theta,\varphi)\pi^\theta(\theta,\varphi)\mathrm{d}\theta\mathrm{d}\varphi \\
\propto& \int_0^\infty\int_0^{x_{(1)}}\alpha^n\varphi^{n\alpha}\left(\prod_{i=1}^n x_i^{\alpha-1}\right)\exp\left[-\varphi^\alpha\sum_{i=1}^n x_i^\alpha\right]\exp[n\varphi^\alpha\theta^\alpha] \\
&\times\theta^{2(\alpha)-1}\varphi^{2\alpha-1}\mathrm{d}\theta\mathrm{d}\varphi \\
<& \frac{C_1}{2\alpha-1}x_{(1)}^{2\alpha-1}\alpha^{n-1}\left(\prod_{i=1}^n x_i^{\alpha-1}\right)\frac{\Gamma(n+2)}{(\sum_{i=1}^n x_i^\alpha)^{n+2}}<\infty
\end{aligned}$$

for $\alpha \geq 1$, where $x_{(1)} = \min_{1\leq i\leq n} x_i$ and $C_1 > 0$ is a constant. For $\alpha < 1$, we note that it holds $m(\boldsymbol{x}) = \infty$. Although the prior $\pi^\theta(\theta, \varphi) \propto \theta^{2(\alpha-1)}\varphi^{2\alpha-1}$ is improper, the posterior distribution is proper for $\alpha \geq 1$ and $n \geq 1$.

On the other hand, if φ is the parameter of interest, from (3.10) the moment matching prior $\pi_M^\varphi(\theta, \varphi)$ of (θ, φ) is given by $\pi_M^\varphi(\theta, \varphi) \propto \varphi^{3(\alpha-1)/2}$. In this case, the marginal density of $\boldsymbol{X} = (X_1, \ldots, X_n)$ is given by

$$\begin{aligned}
m(\boldsymbol{x}) =& \int\int f(\boldsymbol{x};\theta,\varphi)\pi^\varphi(\theta,\varphi)\mathrm{d}\theta\mathrm{d}\varphi \\
\propto& \int_0^\infty\int_0^{x_{(1)}}\alpha^n\varphi^{n\alpha}\left(\prod_{i=1}^n x_i^{\alpha-1}\right)\exp\left[-\varphi^\alpha\sum_{i=1}^n x_i^\alpha\right]\exp[n\varphi^\alpha\theta^\alpha] \\
&\times\varphi^{3(\alpha-1)/2}\mathrm{d}\theta\mathrm{d}\varphi \\
<& C_2 x_{(1)}\alpha^{n-1}\left(\prod_{i=1}^n x_i^{\alpha-1}\right)\frac{\Gamma((2n+\alpha-3)/(2\alpha))}{(\sum_{i=1}^n x_i^\alpha)^{(2n+\alpha-3)/(2n)}}<\infty
\end{aligned}$$

for $\alpha > \max\{0, 3-2n\}$, where $x_{(1)} = \min_{1\leq i\leq n} x_i$ and $C_2 > 0$ is a constant.. Although the prior $\pi_M^\varphi(\theta, \varphi) \propto \varphi^{3(\alpha-1)/2}$ is improper, the posterior distribution is proper for $\alpha > \max\{0, 3-2n\}$ and $n \geq 1$. In this case, since the posterior distribution is intractable, we may compute the posterior mean by using Markov chain Monte Carlo method. However, we do not discuss it here.

Acknowledgments

This work was supported by Grant-in-Aid for Young Scientists (B), Japan Society for the Promotion of Science (JSPS), under Contract Number 17K14233.

References

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion). *Journal of the Royal Statistical Society*. Ser. B, **41**, 113–147.

Datta, G. S. and Ghosh, M. (1996). On the invariance of noninformative priors. *The Annals of Statistics*, **24**, 141–159.

Datta, G. S. and Mukarjee, R. (2004). *Probability Matching Priors: Higher Order Asymptotics*. Springer, New York.

Ghosal, S. (1997). Reference priors in multiparameter nonregular cases. *Test*. **6**, 159–186.

Ghosal, S. (1999). Probability matching priors for non-regular cases. *Biometrika*. **86**, 956–964.

Ghosal, S. and Samanta, T. (1997a). Expansion of Bayes risk for entropy loss and reference prior in non-regular cases. *Statistics & Decisions*. **15**, 129–140.

Ghosal, S. and Samanta, T. (1997b). Asymptotic expansion of posterior distributions in non-regular case. *Annals of the Institute of Statistical Mathematics*. **49** (1), 181–197.

Ghosh, M. and Liu, R. (2011). Moment matching priors. *Sankhyā*. **73**-A, 185–201.

Tibshirani, R. J. (1989). Noninformative priors for one parameter of many. *Biometrika*. **76**, 604–608.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*. **72**, 67–90.

Wang, H. and Sun, D. (2012). Objective Bayesian analysis for a truncated model. *Statistics and Probability Letters*. **82**, 2125–2135.

Welch, B. L. and Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *Journal of the Royal Statistical Society*. Ser. B, **25**, 318–329.