Biometrika (2014), **xx**, x, *pp*. 1–8

Printed in Great Britain

Posterior expectation based on empirical likelihoods

BY A. VEXLER, G. TAO AND A. D. HUTSON

Department of Biostatistics, University at Buffalo, State University of New York, Buffalo, New York 14214, USA

avexler@buffalo.edu getao@buffalo.edu ahutson@buffalo.edu

SUMMARY

Posterior expectation is widely used as a Bayesian point estimator. In this paper we extend it from parametric models to nonparametric models using empirical likelihood, and provide a nonparametric analog of James–Stein estimation. We use the Laplace method to establish asymptotic approximations to our proposed posterior expectations, and show by simulation that they are often more efficient than corresponding classical nonparametric procedures, especially when the underlying data are skewed.

Some key words: Empirical Bayes methods; Empirical likelihood; James-Stein estimator; Laplace method; Nonparametric estimation; Posterior expectation.

1. INTRODUCTION

Bayesian posterior expectations are commonly used to characterize posterior and predictive distributions (Tierney et al., 1989), and serve as Bayes analogues of frequentist point estimators based on parametric statistical models (Carlin & Louis, 2000). When there are concerns about the appropriateness of a parametric likelihood, Lazar (2003) showed that the empirical likelihood (Owen, 2001) can be used as the basis for robust and accurate Bayesian inference. The key idea of this paper is to develop this idea further by using empirical likelihood-based posterior expectations to provide a robust data-driven alternative to standard Bayesian point estimators.

Tierney & Kadane (1986) developed an easily computable asymptotic approximation for the parametric posterior expectation using the Laplace method. Another key piece of the research developed in this note is the derivation of asymptotic approximations to the proposed nonparametric posterior expectations. We demonstrate the asymptotic propositions are very accurate and have a direct analog to those of parametric posterior-based procedures.

In various Bayesian scenarios, prior functions are known up to a given set of parameters. The empirical Bayes method uses the observed data to estimate the prior's parameters, e.g., by maximizing the marginal distributions (Carlin & Louis, 2000). In this paper, we propose to use empirical likelihoods as substitutes for parametric likelihoods in the empirical Bayesian posterior estimation. The distribution-free estimators obtained via this manner are denoted as double empirical Bayesian point estimators.

In the case of multivariate normally distributed data, Stein (1956) proved that when the dimension of the observed vectors is greater than or equal to three, the maximum likelihood estimators are inadmissible estimators of the corresponding parameters. James & Stein (1961) provided another estimator that yields the frequentist risk, i.e. the mean squared error, which is no larger than that of the corresponding maximum likelihood estimators. Efron & Morris (1972) showed that the James–Stein estimator belongs to a class of parametric empirical Bayes point estimators in the Gaussian/Gaussian model. In this context, we infer and illustrate in this note that the proposed double empirical Bayesian point estimators can lead to nonparametric versions of the James– Stein estimators when normal priors with unknown parameters are utilized.

2. NONPARAMETRIC POSTERIOR EXPECTATIONS

Let $X_1, \ldots, X_n$ be independent and identically distributed observations from a distribution function $F(x|\theta)$, where θ is the parameter to be evaluated. For convenience of exposition and without loss of generality we assume the parameter θ is one-dimensional. The Bayesian point estimator of θ can be defined as the posterior expectation

$$\hat{\theta} = \frac{\int \theta \prod_{i=1}^{n} f(X_i|\theta)\pi(\theta)d\theta}{\int \prod_{i=1}^{n} f(X_i|\theta)\pi(\theta)d\theta}, \quad (1)$$

where f is the density function of X_1 and $\pi(\theta)$ is the prior distribution. The estimator (1) uses the parametric likelihood, $\prod_{i=1}^{n} f(X_i|\theta)$, provided that the form of f is known.

We propose using the relevant empirical likelihood function instead of the parametric likelihood at (1) to obtain the nonparametric posterior expectation. We start with an example of this approach using the mean.

Following the empirical likelihood literature (Owen, 1988; Vexler et al., 2009) we define the log empirical likelihood function with respect to the mean θ of $X_1, \ldots, X_n$ as

$$\ell_1(\theta) = \max_{0<p_1,\cdots,p_n<1} \left\{ \sum_{i=1}^{n} \log p_i : \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i X_i = \theta \right\}.$$

Thus the nonparametric posterior expectation has the form

$$\hat{\theta} = \frac{\int_{X_{(1)}}^{X_{(n)}} \theta e^{\ell_1(\theta)} \pi(\theta)d\theta}{\int_{X_{(1)}}^{X_{(n)}} e^{\ell_1(\theta)} \pi(\theta)d\theta} = \frac{\int_{X_{(1)}}^{X_{(n)}} \theta e^{\ell r_1(\theta)} \pi(\theta)d\theta}{\int_{X_{(1)}}^{X_{(n)}} e^{\ell r_1(\theta)} \pi(\theta)d\theta}, \quad (2)$$

where $X_{(1)} < \cdots < X_{(n)}$ are the order statistics based on the sample $X_1, \ldots, X_n$ and $\ell r_1(\theta) = \ell_1(\theta) + n \log n$ is the log empirical likelihood ratio.

In general, the integrals in (1) are intractable and need to be evaluated numerically. A useful and accurate approximation to integrals necessary for Bayesian calculations can be obtained by assuming that the posterior density is unimodal, or at least dominated by a single mode, such that it is highly peaked about its maximum, which is the posterior mode. If so, we can expand the log-parametric likelihood as quadratic function of θ about the maximum likelihood estimator of θ. This approach yields approximations to the integrands at (1) that have the normal density-type forms. This method is based on the Laplace method (Bleistein & Handelsman, 2010; Tierney & Kadane, 1986).

In this article we show that marginal distributions based on the empirical likelihood approach behave similarly to those based on parametric likelihoods, i.e., $\ell_1(\theta)$ is highly peaked about its maximum value. That is, we can approximate integrals of the form $\int \theta^k \exp\{\ell_1(\theta)\}\pi(\theta)d\theta$, $k = 0, 1$, in a similar manner to the approximations related to the parametric posterior expectations. We refer to the Supplementary Material for technical derivations and proofs.

PROPOSITION 1. *Assume that* $E\left(|X_1|^4\right) < \infty$, $\int |\theta|\pi(\theta)d\theta < \infty$ *and* $\pi(\theta)$ *is twice continuously differentiable in a neighborhood of* $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$. *Then the proposed estimator* (2) *satisfies*

$$\hat{\theta} = \frac{\int \theta \exp\left\{-\frac{n(\bar{X}-\theta)^2}{2\sigma_n^2}\right\} \pi(\theta) d\theta}{\int \exp\left\{-\frac{n(\bar{X}-\theta)^2}{2\sigma_n^2}\right\} \pi(\theta) d\theta} + \frac{M_n^3}{n\,\sigma_n^2} + O_p(n^{-3/2+\varepsilon}),$$

where $\sigma_n^2 = n^{-1}\sum_{i=1}^n (X_i - \bar{X})^2$, *and* $M_n^3 = n^{-1}\sum_{i=1}^n (X_i - \bar{X})^3$, *for all* $\varepsilon > 0$ *as* $n \to \infty$.

COROLLARY 1. *Let* $\pi(\theta) = (2\pi\sigma_\pi^2)^{-1/2}\exp\{-(\theta-\mu_\pi)^2/(2\sigma_\pi^2)\}$, *where* μ_π *and* σ_π^2 *are known hyperparameters, and the conditions of Proposition* 1 *hold. Then the posterior expectation at* (2) *can be approximated as*

$$\hat{\theta} = \tilde{\theta} + \frac{M_n^3}{n\,\sigma_n^2} + O_p(n^{-3/2+\varepsilon}), \quad \tilde{\theta} = \frac{\mu_\pi/\sigma_\pi^2}{1/\sigma_\pi^2 + n/\sigma_n^2} + \frac{n\bar{X}/\sigma_n^2}{1/\sigma_\pi^2 + n/\sigma_n^2}.$$

The estimator $\tilde{\theta}$ is equivalent to the form of the parametric posterior expectation derived under the normal/normal model (Carlin & Louis, 2000). Following the process of the asymptotic evaluation of the parametric posterior expectations we can easily show the following:

COROLLARY 2. *Under the conditions of Proposition* 1*, let* $\pi(\theta)$ *be a prior function with* $|d^3\log(\pi(\theta))/d\theta^3| < \infty$, *for all* θ. *Then, for all* $\varepsilon > 0$, *we have the following result:*

$$\hat{\theta} = \frac{n\bar{X} + \sigma_n^2\{\log\pi(\bar{X})\}' - \sigma_n^2\{\log\pi(\bar{X})\}''\bar{X}}{n - \sigma_n^2\{\log\pi(\bar{X})\}''} + \frac{M_n^3}{n\,\sigma_n^2} + O_p(n^{-3/2+\varepsilon}), \;\; n \to \infty.$$

Now, consider the normal prior, $\pi(\theta)$, when μ_π and σ_π^2 are unknown. According to the empirical Bayes concept the unknown hyperparameters can be estimated by maximizing the respective marginal distributions. This method can be applied to the nonparametric posterior expectation yielding double empirical posterior estimation. In this case, we define

$$\hat{\theta}_E = \frac{\int \theta \exp\{\ell_1(\theta)\}\exp\{-(\theta-\hat{\mu}_\pi)^2/(2\hat{\sigma}_\pi^2)\}d\theta}{\int \exp\{\ell_1(\theta)\}\exp\{-(\theta-\hat{\mu}_\pi)^2/(2\hat{\sigma}_\pi^2)\}d\theta}, \tag{3}$$

where $(\hat{\mu}_\pi, \hat{\sigma}_\pi^2) = \arg\max_{(\mu,\sigma)}[(2\pi\sigma^2)^{-1/2}\int \exp\{\ell_1(\theta)\}\exp\{-(\theta-\mu)^2/2\sigma^2\}d\theta]$. The next result implies a simple asymptotic form of $\hat{\theta}_E$.

COROLLARY 3. *If* $E\left(|X_1|^4\right) < \infty$, *then, for all* $\varepsilon > 0$, *the posterior expectation* (3) *satisfies*

$$\hat{\theta}_E = \frac{\hat{\mu}_\pi\sigma_n^2 + \bar{X}\hat{\sigma}_\pi^2 n}{n\hat{\sigma}_\pi^2 + \sigma_n^2} + \frac{M_n^3}{n\,\sigma_n^2} + O_p(n^{-3/2+\varepsilon}),$$

where $\hat{\mu}_\pi - \bar{X} \to 0$, $\hat{\sigma}_\pi^2 - \max\{0, \sigma_n^2 - \sigma^2\} \to 0$, $\sigma^2 = var(X_1)$, *as* $n \to \infty$.

When $D(\theta)$ defines a function of θ and we denote the nonparametric posterior expectation of $D(\theta)$ as

$$\hat{D} = \int D(\theta)e^{\ell_1(\theta)}\pi(\theta)d\theta\{\int e^{\ell_1(\theta)}\pi(\theta)d\theta\}^{-1},$$

one can present the next result.

PROPOSITION 2. *When* $\int |D(\theta)|\pi(\theta)d\theta < \infty$, $D(\theta) > 0$, $|\{\log D(\theta)\}'''| < \infty$, *and* $|\{\log\pi(\theta)\}'''| < \infty$, *for all* θ, *the nonparametric posterior expectation of* $D(\theta)$, *satisfies,* *for all* $\varepsilon > 0$,

$$\hat{D} = \int D(\theta) e^{-\left(\sum X_i - n\theta\right)^2/(2n\sigma_n^2)} \pi(\theta) d\theta \left\{ \int e^{-\left(\sum X_i - n\theta\right)^2/(2n\sigma_n^2)} \pi(\theta) d\theta \right\}^{-1} + \frac{D'(\bar{X}) M_n^3}{n\,\sigma_n^2} + O_p(n^{-3/2+\varepsilon}),$$

where $\sigma_n^2 = n^{-1}\sum_{i=1}^n (X_i - \bar{X})^2$, $M_n^3 = n^{-1}\sum_{i=1}^n (X_i - \bar{X})^3$ *as* $n \to \infty$.

In order to consider more general cases, we begin with the definition of the log empirical likelihood function presented in the form

$$\ell_2(\theta) = \max\{\sum_{i=1}^n \log p_i : 0 < p_i < 1, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i G(X_i, \theta) = 0\},$$

where we assume for simplicity that $\partial G(u,\theta)/\partial\theta > 0$ or $\partial G(u,\theta)/\partial\theta < 0$ for all u and $E\left(|G(X_1,\theta)|^4\right) < \infty$. In this framework the posterior expectation takes the form

$$\hat{\theta} = \frac{\int_{X_{(1)}}^{X_{(n)}} \theta e^{\ell_2(\theta)} \pi(\theta) d\theta}{\int_{X_{(1)}}^{X_{(n)}} e^{\ell_2(\theta)} \pi(\theta) d\theta} = \frac{\int_{X_{(1)}}^{X_{(n)}} \theta e^{\ell r_2(\theta)} \pi(\theta) d\theta}{\int_{X_{(1)}}^{X_{(n)}} e^{\ell r_2(\theta)} \pi(\theta) d\theta},\ \ell r_2(\theta) = \ell_2(\theta) + n\log n. \tag{4}$$

In the Supplementary Material we prove that $\ell_2(\theta)$ increases and decreases monotonically for $\theta < \theta_M$ and $\theta > \theta_M$, respectively, where θ_M is the root of $n^{-n}\sum_{i=1}^n G(X_i, \theta_M) = 0$. Then, employing the technique used in the proof of Proposition 1, we have the following:

PROPOSITION 3. *If* $\int |\theta| \pi(\theta) d\theta < \infty$ *and* $\pi(\theta)$ *is twice continuously differentiable in a neighborhood of* θ_M*, then the estimator defined at* (4) *has the asymptotic form*

$$\hat{\theta} = \frac{\int \theta \exp\left[-\{\sum_{i=1}^n G(X_i,\theta)\}^2/(2n\sigma_{Gn}^2)\right] \pi(\theta) d\theta}{\int \exp\left[-\{\sum_{i=1}^n G(X_i,\theta)\}^2/(2n\sigma_{Gn}^2)\right] \pi(\theta) d\theta} + \frac{M_{Gn}^3}{n\,\sigma_{Gn}^2} + O_p(n^{-3/2+\varepsilon}),$$

for all $\varepsilon > 0$*, where* $\sigma_{Gn}^2 = n^{-1}\sum_{i=1}^n G(X_i,\theta)^2$, $M_{Gn}^3 = n^{-1}\sum_{i=1}^n G(X_i,\theta)^3$.

Moreover, if $\pi(\theta)$ is a prior distribution function with $|\{\log \pi(\theta)\}'''| < \infty$ and $|\partial^2 G(X_i,\theta)/\partial\theta^2| < \infty$, for all θ, we have that

$$\begin{aligned}\hat{\theta} =& [\{\sum_{i=1}^n \partial G(X_i,\theta_M)/\partial\theta_M\}^2 - \{\log \pi(\theta_M)\}'' \sum_{i=1}^n G(X_i,\theta_M)^2]^{-1} \Bigg[\theta_M \{\sum_{i=1}^n \frac{\partial G(X_i,\theta_M)}{\partial \theta_M}\}^2 \\ &+ \{\log \pi(\theta_M)\}' \sum_{i=1}^n G(X_i,\theta_M)^2 - \sum_{i=1}^n G(X_i,\theta_M) \sum_{i=1}^n \frac{\partial G(X_i,\theta_M)}{\partial\theta_M} \\ &- \{\log \pi(\theta_M)\}'' \sum_{i=1}^n G(X_i,\theta_M)^2 \theta_M \Bigg] + \frac{M_{Gn}^3}{n\,\sigma_{Gn}^2} + O_p(n^{-3/2+\varepsilon}),\end{aligned}$$

for all $\varepsilon > 0$, *as* $n \to \infty$.

Remark 1. The nonparametric posterior expectation of $D(\theta)$ defined earlier and given in the more general form as

$$\hat{D} = \frac{\int_{X_{(1)}}^{X_{(n)}} D(\theta) e^{\ell_2(\theta)} \pi(\theta) d\theta}{\int_{X_{(1)}}^{X_{(n)}} e^{\ell_2(\theta)} \pi(\theta) d\theta}$$

can be analyzed in a similar manner to Propositions 2 and 3.

Now we can define the log empirical likelihood function as

$$\ell_3(\theta_1,\ldots,\theta_K) = \max\{\sum_{i=1}^n \log p_i : 0 < p_i < 1, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i G_k(X_i,\theta_k) = 0, k = 1,\ldots,K\}$$

to propose the posterior estimator

$$\hat{D}_G = \frac{\int\cdots\int D(\theta_1,\ldots,\theta_K) e^{\ell r_3(\theta_1,\ldots,\theta_K)} \pi(\theta_1,\ldots,\theta_K) d\theta_1\cdots d\theta_K}{\int\cdots\int e^{\ell r_3(\theta_1,\ldots,\theta_K)} \pi(\theta_1,\ldots,\theta_K) d\theta_1\cdots d\theta_K},\ \ell r_3 = \ell_3 + n\log n.$$

The Supplementary Material provides the basic technical ingredients to analyze this complex estimator. For example, without loss of generality and for ease of presentation, we consider $K = 2$, $G_k(X_i,\theta_k) = X_i^k - \theta_k$, $k = 1,2$ and

$$\hat{D}_G = \frac{\int_{X_{(1)}}^{X_{(n)}}\int_{V_1}^{V_2} D(\theta_1,\theta_2) e^{\ell r_3(\theta_1,\theta_2)} \pi(\theta_1,\theta_2) d\theta_1 d\theta_2}{\int_{X_{(1)}}^{X_{(n)}}\int_{V_1}^{V_2} e^{\ell r_3(\theta_1,\theta_2)} \pi(\theta_1,\theta_2) d\theta_1 d\theta_2},\ V_1 = \min_{i=1,\ldots,n} X_i^2,\ V_2 = \max_{i=1,\ldots,n} X_i^2.$$

If we assume that $\int\int |D(\theta_1,\theta_2)|\pi(\theta_1,\theta_2) d\theta_1 d\theta_2 < \infty$ exists as well as D and π are twice continuously differentiable in neighborhoods of $\left(\bar{X}, \bar{X^2} \equiv n^{-1}\sum_{i=1}^n X_i^2\right)$, then the following proposition yields the relevant asymptotic result:

PROPOSITION 4. *If $E\left(|X_1|^4\right) < \infty$, then the asymptotic approximation to the proposed posterior expectation of $D(\theta_1,\theta_2)$ is given by*

$$\begin{aligned}\hat{D}_G = \int\int D(\theta_1,\theta_2)\exp\Bigg\{ &- \frac{0.5n(\theta_1-\bar{X})^2}{\sigma_n^2 - (\sigma_{12,n})^2/\sigma_{02,n}^2} - \frac{0.5n(\theta_2-\bar{X^2})^2}{\sigma_{02,n}^2 - (\sigma_{12,n})^2/\sigma_n^2} \\ &+ \frac{n(\theta_1-\bar{X})(\theta_2-\bar{X^2})}{\sigma_{02,n}^2\sigma_n^2/\sigma_{12,n} - \sigma_{12,n}}\Bigg\}\pi(\theta_1,\theta_2)d\theta_1 d\theta_2 \\ &\times\left[\int\int \exp\Bigg\{ - \frac{0.5n(\theta_1-\bar{X})^2}{\sigma_n^2 - (\sigma_{12,n})^2/\sigma_{02,n}^2} - \frac{0.5n(\theta_2-\bar{X^2})^2}{\sigma_{02,n}^2 - (\sigma_{12,n})^2/\sigma_n^2}\right. \\ &\left.+ \frac{n(\theta_1-\bar{X})(\theta_2-\bar{X^2})}{\sigma_{02,n}^2\sigma_n^2/\sigma_{12,n} - \sigma_{12,n}}\Bigg\}\pi(\theta_1,\theta_2)d\theta_1 d\theta_2\right]^{-1} + \frac{J_n}{n} + O_p(n^{-3/2+\varepsilon}),\ J_n = O_p(1),\end{aligned}$$

for all $\varepsilon > 0$, as $n \to \infty$, where $\bar{X^2} = n^{-1}\sum_{i=1}^n X_i^2$, $\sigma_{02,n}^2 = n^{-1}\sum_{i=1}^n (X_i^2 - \bar{X^2})^2$, $\sigma_{12,n} = n^{-1}\sum_{i=1}^n (X_i - \bar{X})(X_i^2 - \bar{X^2})$ and the term J_n has a complicated form presented in the equation (A15) of the Supplementary Material.

3. NONPARAMETRIC ANALOG OF JAMES–STEIN ESTIMATION

The classical James–Stein estimation process assumes that the observations $X_1,\ldots,X_n$ are independent and identically distributed as multivariate normal with corresponding mean vector $\theta = (\theta_1,\ldots,\theta_K)$ and covariance matrix Σ. Efron & Morris (1972) proved that the James–Stein estimator is a parametric empirical Bayes point estimator related to a Gaussian/Gaussian model.

In Section 2, we showed that when $K = 1$ and the prior function is a normal density function the proposed nonparametric posterior expectation is asymptotically equivalent to the parametric posterior expectation derived under assumptions of the normal/normal model. In this section we assume $X_1, \ldots, X_n$ are independent random vectors, $X_i = (X_{i1}, \ldots, X_{iK})^T$, with an unknown distribution and $E\left(|X_{ij}|^4\right) < \infty$, $(j = 1, \ldots, K;\ i = 1, \ldots, n)$. We propose a nonparametric estimator of the mean $(\theta_1, \ldots, \theta_K)^T$ using the double empirical posterior estimation, in the form:

$$\hat{\theta}_{Ej} = \frac{\int \theta \exp\{\ell_{4j}(\theta)\} \exp\{-\theta^2/(2\tilde{\sigma}^2_\pi\})d\theta}{\int \exp\{\ell_{4j}(\theta)\} \exp\{-\theta^2/(2\tilde{\sigma}^2_\pi)\}d\theta}, \tag{5}$$

with $\tilde{\sigma}^2_\pi = \arg\max_{\sigma^2} \sum_{j=1}^K \log[(2\pi\sigma^2)^{-1/2} \int \exp\{\ell_{4j}(\theta)\} \exp\{-\theta^2/(2\sigma^2)\}d\theta]$, where $\ell_{4j}(\theta_j) = \max\{\sum_{i=1}^n \log p_{ij} : 0 < p_{ij} < 1, \sum_{i=1}^n p_{ij} = 1, \sum_{i=1}^n p_{ij}X_{ij} = \theta_j\},\ j = 1,\ \ldots, K$. In the following proposition we show the proposed distribution free estimation is asymptotically equivalent to the parametric version of the James–Stein estimator.

PROPOSITION 5. *For all $\varepsilon > 0$ and as $n \to \infty$ the double empirical posterior estimator* (5) *has the following asymptotic form:*

$$\hat{\theta}_{Ej} = \left\{1 - \frac{(K-2)/n}{\bar{X}^T S^{-1} \bar{X}}\right\} \bar{X}_j + \frac{1}{n} \sum_{r=1}^K \frac{\sum_{i=1}^n (X_{ir} - \bar{X}_r)^3}{\sum_{i=1}^n (X_{ir} - \bar{X}_r)^2} + O_p(n^{-3/2+\varepsilon}), j = 1, \ldots, K, \tag{6}$$

where $\bar{X} = (\bar{X}_1, \ldots, \bar{X}_K)^T$, $\bar{X}_j = n^{-1} \sum_{i=1}^n X_{ij}$ and S is the sample estimator of Σ.

4. SIMULATIONS

We evaluate the performance of the proposed estimation and the asymptotic results presented in Section 2 via a simulation study, where we generate samples of size $n = 10, 20, 30, 50$ and 75 from both a $N(1, 1)$ and $LogN(0, 1)$ distribution. A more detailed explanation can be found in the Supplementary Material. Define $\mu = 1$, if $X_1 \sim N(1, 1)$ and $\mu = \exp(1/2)$, if $X_1 \sim LogN(0, 1)$. We consider different scenarios regarding the prior function selections to illustrate the following situations: a) The prior distribution π is supposed to contain no correct information about the true values of θ. For example, π corresponds to the $N(0, 1)$ or $N(0, 0.5^2)$ distribution, these distribution functions are not centered around the true values of the parameter $\theta = E(X_1)$; b) the priors $N(\mu, \sigma^2_\pi)$ and $U[\mu - 0.5, \mu + 0.5]$ are centered near the true values of the parameter; c) the $0.5\{N(-\mu, \sigma^2_\pi) + N(\mu, \sigma^2_\pi)\}$– type prior distributions reflect information that the target parameter can be μ units $+/-$ within the standard deviation σ^2_π; d) The hyperparameter $\sigma_\pi = 0.5, 0.1, 0.05$ represents our relative confidence with respect to the prior information. In these cases we compare $\hat{\theta}$ defined in (2) with its corresponding asymptotic forms stated in Proposition 1 and Corollary 1, the classical nonparametric estimator $\bar{X}$ and the corresponding maximum likelihood estimators θ_{MLE} of $\theta = E(X_1)$. For the simulation results presented in Fig. 1, we define $\Delta_n(q_n) = \{V(q_n) - V(\bar{X})\}/V(\bar{X})$, $B_{1\,n} = n\,A(\hat{\theta} - \hat{\theta}_{P1})$ and $B_{2\,n} = n\,A(\hat{\theta} - \hat{\theta}_{C1})$, where the operators $V(q_n)$ and $A(q_n)$ are the Monte Carlo variance and mean of an estimator q_n of θ, respectively. The notations $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ denote the approximations to $\hat{\theta}$ obtained in Proposition 1 and Corollary 1, respectively. The extensive Monte Carlo study shows that the corresponding variances of the proposed estimation procedures, which incorporate non-informative priors or priors presenting no correct information about true values of estimated parameters, are generally smaller than those of the traditional nonparametric estimator $\bar{X}$. This is especially true when the underlying data distributions are skewed, e.g., when the data follows a log-normal distribution. When priors are selected based on information that accurately reflects the true values of the underlying parameter, e.g., prior distributions centered around the

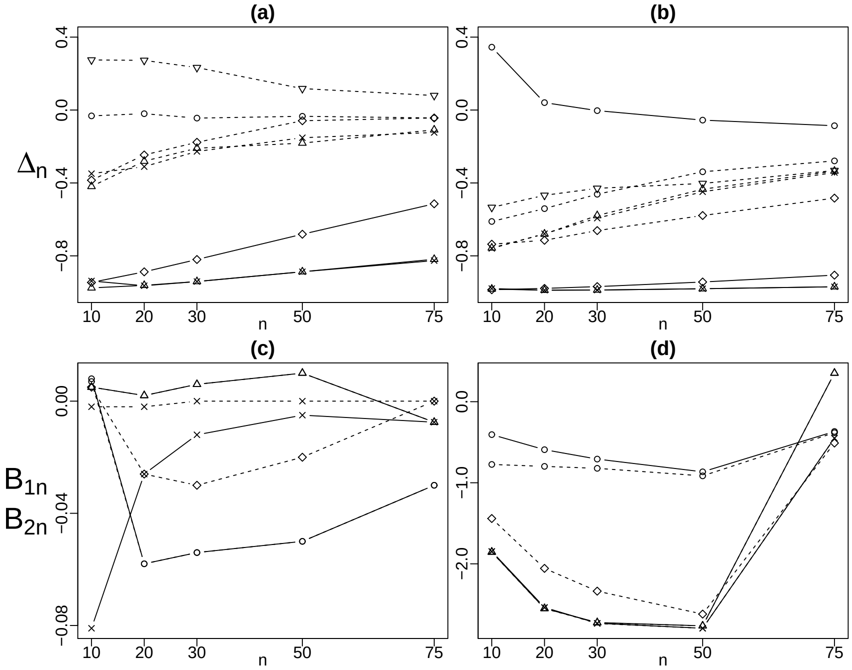

Fig. 1: Simulation results based on $X_i \sim N(1,1)$ and $X_i \sim LogN(0,1)$, $i = 1,\ldots,n$, presented in graphs (a),(c) and (b),(d), respectively. Plots (a) and (b) show the relative efficiency $\Delta_n(\hat{\theta})$ of the proposed estimation with the priors: $N(0,1)$ (curve - -○- -) , $N(\mu, 0.5^2)$ (- -△- -), $N(\mu, 0.1^2)$ (–△–), $N(\mu - 1, 0.5^2)$ (- -▽- -), $\{N(-\mu, 0.5^2) + N(\mu, 0.5^2)\}/2$, (- -×- -), $\{N(-\mu, 0.1^2) + N(\mu, 0.1^2)\}/2$ (–×–), $U[0, \mu + 0.5]$ (- -◇- -) and $U[\mu - 0.25, \mu + 0.25]$ (–◇–). Curve (–○–) depicts $\Delta_n\left(\theta_{\text{MLE}}\right)$ in (b). Plots (c) and (d) show the following biases: $B_{1\,n}$ via curves (- -○- -), (- -△- -), (- -×- -), (- -◇- -) and $B_{2\,n}$ via curves (–○–) , (–△–), (–×–), where the plotting symbols ○, △, ×, ◇ correspond to the priors $N(0,1)$, $N(\mu, 0.5^2)$, $\{N(-\mu, 0.5^2) + N(\mu, 0.5^2)\}/2$, $U[0, \mu + 0.5]$, respectively.

true values of the estimated parameters, the empirical likelihood-based posterior expectation has variances smaller than those of their parametric frequentist maximum likelihood counterparts.

In the Supplementary Material we compare the nonparametric James–Stein estimator $\hat{\theta}_{Ej}$ to the classical nonparametric estimator $\bar{X}$. Figure 2 shows the Monte Carlo results related to the simulations based on data from $MVN\{(1,1,1)^T, I\}$ and $MVLogN\{(1,1,1)^T, I\}$ distributions. The Monte Carlo study in the multivariate setting confirmed that the proposed nonparametric James–Stein estimator has smaller variances than the classical nonparametric estimator $\bar{X}$ for data generated from multivariate normal and multivariate log-normal distributions.

ACKNOWLEDGMENT

This research was supported by the National Institutes of Health, U.S.A. We are grateful to the editor, and reviewers for helpful comments.

SUPPLEMENTARY MATERIAL

Supplementary Material available at *Biometrika* online includes detailed explanations of the Monte Carlo study and its outputs as well as details of the technical derivations and proofs corresponding to the theoretical results presented in this paper.

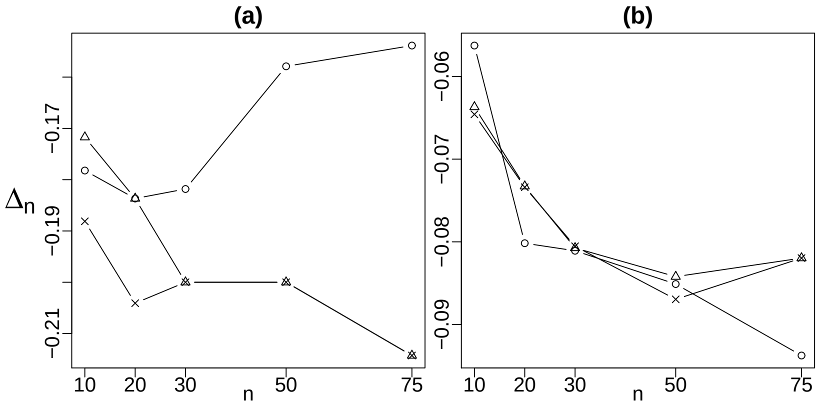

Fig. 2: Monte Carlo evaluations of the nonparametric James–Stein estimator. (a): $(X_1, X_2, X_3)^T \sim MVN\{(1,1,1)^T, I\}$; (b): $(X_1, X_2, X_3)^T \sim MVLogN\{(1,1,1)^T, I\}$. Curves (–○–, –△–, –×–) correspond to values of $\Delta_n = \{V(\hat{\theta}_{Ej}) - V(\bar{X}_j)\}/V(\bar{X}_j)$ for $j = 1, 2, 3$, respectively.

REFERENCES

BLEISTEIN, N. & HANDELSMAN, R. A. (2010). *Asymptotic Expansions of Integrals*. New York: Courier Dover Publications.

CARLIN, B. P. & LOUIS, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. New York: Chapman and Hall/CRC.

EFRON, B. & MORRIS, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators- Part II: The empirical bayes case. *Journal of the American Statistical Association* **67**, 130–139.

JAMES, W. & STEIN, C. (1961). Estimation with quadratic loss. In *Proc. 4rd Berkeley Symp. Math. Statist. Prob.*, vol. 1, pp. 361–379. University of California Press.

LAZAR, N. A. (2003). Bayesian empirical likelihood. *Biometrika* **90**, 319–326.

OWEN, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–249.

OWEN, A. B. (2001). *Empirical likelihood*. New York: Chapman and Hall/CRC.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proc. 3rd Berkeley Symp. Math. Statist. Prob.*, vol. 1, pp. 197–206. University of California Press.

TIERNEY, L. & KADANE, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association* **81**, 82–86.

TIERNEY, L., KASS, R. E. & KADANE, J. B. (1989). Fully exponential laplace approximations to expectations and variances of nonpositive functions. *Journal of the American Statistical Association* **84**, 710–716.

VEXLER, A., LIU, S., KANG, L. & HUTSON, A. D. (2009). Modifications of the empirical likelihood interval estimation with improved coverage probabilities. *Communications in Statistics – Simulation and Computation* **38**, 2171–2183.

Biometrika (2014), **xx**, x, *pp.* 1–13

Printed in Great Britain

Posterior expectation based on empirical likelihoods: Supplementary Material

BY A. VEXLER, G. TAO AND A. D. HUTSON

Department of Biostatistics, University at Buffalo, State University of New York, Buffalo, New York 14214, USA

avexler@buffalo.edu getao@buffalo.edu ahutson@buffalo.edu

1. MONTE CARLO RESULTS

We begin our simulation study with numerical comparisons between the posterior expectation $\hat{\theta}$ defined in (2), its asymptotic forms stated in Proposition 1 and Corollary 1, the classical nonparametric estimator $\bar{X}$ and the corresponding maximum likelihood estimators of $\theta = E(X_1)$. Towards this end 15,000 samples of size $n = 10, 20, 30, 50$ and 75 were generated from both a $N(1,1)$ and $LogN(0,1)$ distribution. The maximum likelihood estimator, say θ_{MLE}, of $\theta = E(X_1)$ is $\bar{X}$ or $\exp(\bar{X} + \sigma_n^2/2)$ when data follow a normal or lognormal distribution, respectively.

Let μ be equal to 1 or $\exp(1/2)$ when normal or lognormal samples are used, respectively. We considered the following prior distributions to be employed in this study: $N(0,1)$, $N(\mu, \sigma_\pi^2)$, $N(\mu \pm 1, \sigma_\pi^2)$, $\{N(-\mu, \sigma_\pi^2) + N(\mu, \sigma_\pi^2)\}/2$, $U[0, \mu + 0.5]$, $U[\mu - 0.5, \mu + 0.5]$, $U[\mu - 0.25, \mu + 0.25]$, where $\sigma_\pi = 0.5, 0.1, 0.05$ represents different scenarios depicting our relative confidence with respect to our prior information pertaining to the unknown parameter. Note that, to examine posterior distributions based on empirical likelihood functions, Lazar (2003) used priors of $N(\mu, 1/n)$-type forms for Monte Carlo evaluations, where n denotes the corresponding sample size.

The Monte Carlo estimates of the means and variances for the estimators $\bar{X}$ and θ_{MLE} are presented in Table 1. Table 2 provides the estimated mean and variance values for $\hat{\theta}$ and its corresponding asymptotic forms $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ from Proposition 1 and Corollary 1, respectively.

Table 2 shows that when the data are normally distributed, and the $N(0,1)$-prior distribution is used, then the variances of $\hat{\theta}$ are comparable to those of $\theta_{\text{MLE}} = \bar{X}$. The variances of $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ are very close to those of $\hat{\theta}$ even for the moderately small sample size setting at $n = 20$. When using the prior $N(1, 0.5^2)$, the proposed estimator $\hat{\theta}$ performs significantly better than $\bar{X}$. When $n = 10$, the variance of $\hat{\theta}$ is 42% smaller than that of $\bar{X}$. As n increases, the variance of $\hat{\theta}$ becomes close to that of $\bar{X}$. The above conclusions are magnified when the prior $N(1, 0.1^2)$ is used. When using the improper prior, $N(0, 0.5^2)$, the variance of $\hat{\theta}$ is about 27% greater than that of $\bar{X}$ for samples of size $n = 10$. However, when n is large, the variances of $\hat{\theta}$ are comparable to those of $\bar{X}$. When the prior distribution $\{N(-1, 0.5^2) + N(1, 0.5^2)\}/2$ is applied, the variance of $\hat{\theta}$ is about 35% smaller than that of $\bar{X}$ for samples of size $n = 10$. As the sample size increases the variance of $\hat{\theta}$ is comparable to that of $\bar{X}$. These conclusions, relative to the gains in efficiency, are confirmed when $\{N(-1, 0.1^2) + N(1, 0.1^2)\}/2$ is used as the prior distribution. When a non-informative uniform prior distribution is used, e.g., $U[0, 1.5]$, the variance of $\hat{\theta}$ is about 38% smaller than that of $\bar{X}$ for samples of size $n = 10$. When n increases, the variance of $\hat{\theta}$ becomes close to that of $\bar{X}$. When an uniform prior centered near the true parameter value is used, e.g., $U[0.75, 1.25]$, the variance of $\hat{\theta}$ is about 94% smaller than that of $\bar{X}$ for $n = 10$.

In the case where data have an assumed lognormal distribution and the proposed estimator is based on the prior distribution, $N(0,1)$, we have that the variance of $\hat{\theta}$ is about 61% less than that of $\bar{X}$ and about 71% less than that of $\theta_{\rm MLE}$ when based on samples of size $n=10$. When using the prior $N(\exp(1/2),0.5^2)$, the proposed estimator $\hat{\theta}$ performs much better than $\bar{X}$ and $\theta_{\rm MLE}$, e.g., when $n=10$, the variance of $\hat{\theta}$ is about 76% smaller than that of $\bar{X}$ and about 82% less than that of $\theta_{\rm MLE}$. As n increases, the variances of $\hat{\theta}$ become close to the variance of $\bar{X}$ and $\theta_{\rm MLE}$. The asymptotic forms $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$, perform similarly to $\hat{\theta}$, even when $n=20$. These results are highlighted when the prior distribution $N(\mu,0.1^2)$ is used. When using an improper prior distribution, e.g., $N(\mu-1,0.5^2)$, the performance of the proposed estimator is still better than that of $\bar{X}$ and $\theta_{\rm MLE}$, e.g., when $n=10$, the variance of $\hat{\theta}$ is about 53% smaller than that of $\bar{X}$ and about 65% smaller than that of $\theta_{\rm MLE}$. When n is large, the variances of $\hat{\theta}$ are comparable to those of $\bar{X}$ and $\theta_{\rm MLE}$. The estimators $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ are very close to the proposed estimator $\hat{\theta}$. When we have information that the target parameter can be equal to μ or $-\mu$ and we employ the $\{N(-\mu,0.5^2)+N(\mu,0.5^2)\}/2$ prior, the variance of $\hat{\theta}$ is about 76% smaller than that of $\bar{X}$ and about 82% less than that of $\theta_{\rm MLE}$, for samples of size $n=10$. When a non-informative uniform prior is used, e.g., $U[0,\mu+0.5]$, the variance of $\hat{\theta}$ is about 74% smaller than that of $\bar{X}$ and about 80% less than that of $\theta_{\rm MLE}$, for samples of size $n=10$. When an uniform prior centered near the true parameter value is used, e.g., $U[\mu-0.25,\mu+0.25]$, the variance of $\hat{\theta}$ is about 99% smaller than that of $\bar{X}$ and $\theta_{\rm MLE}$ for samples of size $n=10$.

One can also note that the use of the normal prior distributions with mean 0 when estimating the parameter $\theta=1$ or $\theta=\exp(0.5)$ leads to negative biases of the estimators. These biases are relatively small and vanish when the sample size increases.

-TABLES 1, 2-

From the Monte Carlo study based on data sampled from a lognormal distribution we observe that the proposed estimator outperforms $\bar{X}$ and $\theta_{\rm MLE}$, even when using priors that are misspecified relative to the true values of θ. The efficiency of the proposed estimator is clearly demonstrated in the case of skewed data. It has been discussed in the literature that the traditional estimation of the mean of a lognormal distribution is inaccurate due to the non-quadratic and asymmetric shape of the likelihood profile (Wu et al. 2003). In this case the proposed approach can serve as a valid alternative to the traditional techniques.

The performance of the asymptotic forms $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ of $\hat{\theta}$ are observed to be similar to that of $\hat{\theta}$ across a wide range of scenarios. Note that in additional Monte Carlo evaluations, which were omitted from our Supplementary Material, we consistently observed that the estimators $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ provided accurate approximations to $\hat{\theta}$. We also numerically evaluated the double empirical Bayesian estimator $\hat{\theta}_E$ given at equation (3) and the corresponding asymptotic form from Corollary 3. We concluded that these proposed estimators are comparable to $\bar{X}=\theta_{\rm MLE}$, when data are normally distributed, e.g., when $n=20$, the variances of $\bar{X}$ and $\hat{\theta}_E$ were 0.050 and 0.049, respectively. However, when data were generated from a lognormal distribution the proposed estimator demonstrated an improvement in efficiency as compared with the nonparametric estimator $\bar{X}$. For example, when $n=75$, the variance of $\hat{\theta}_E$ was approximately 10% smaller than that of $\bar{X}$.

Monte Carlo evaluations of the nonparametric James–Stein estimator: We carried out numerical evaluations of the nonparametric James–Stein estimator $\hat{\theta}_{Ej}$ and compared it to the classical nonparametric estimator $\bar{X}_j$ in terms of relative bias and efficiency. For simplicity and without loss of generality we assumed the dimension $K=3$ for the underlying multivariate distributions used within this study. The independent samples were generated from either a

$MVN\{(1,1,1)^T, \Sigma\}$ or a $MVLogN\{(0,0,0)^T, \Sigma\}$, where we used the covariance structures

$$\Sigma = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}.$$

The Monte Carlo variance estimates for the respective estimators are defined as $V(\hat{\theta}_{Ej}) = \sum_{i=1}^{T}(\hat{\theta}_{Eij} - \theta_{ij})^2/T$ and $V(\bar{X}_j) = \sum_{i=1}^{T}(\bar{X}_{ij} - \theta_{ij})^2/T$, respectively, where $j = 1, 2, 3$ and $T = 15000$ is the number of the Monte Carlo replications. The experimental means and variances of the estimators are presented in Tables 3-6.

-TABLES 3-6-

In the cases where samples were generated from a $MVN\{(1,1,1)^T, I\}$ distribution the proposed James–Stein estimator was more efficient than $\bar{X} = \theta_{\text{MLE}}$ for small sample sizes, e.g. when $n = 10$ the variances of $(\hat{\theta}_{E1}, \hat{\theta}_{E2}, \hat{\theta}_{E3})$ were $(0.062, 0.065, 0.064)$, respectively, while the variances of $\bar{X}$ were $(0.096, 0.100, 0.099)$. As the sample size increased, the variance of the nonparametric James–Stein estimators were observed to be close to those of each component of $\bar{X}$. In the case where samples were generated from $MVLogN\{(0,0,0)^T, I\}$ the James–Stein estimator had a smaller component-wise variance as compared to the corresponding estimators for each element of $\bar{X}$. For example, when $n = 10$, the variance of each element of $(\hat{\theta}_{E1}, \hat{\theta}_{E2}, \hat{\theta}_{E3})$ was $(0.398, 0.531, 0.395)$, while the variance of each element of $\bar{X}$ was $(0.476, 0.591, 0.485)$. In the case where correlated data was generated, we observed similar results in that the performance of $(\hat{\theta}_{E1}, \hat{\theta}_{E2}, \hat{\theta}_{E3})$ was better than that of $\bar{X}$ in the sense of the relative efficiency. We conclude that for fixed sample sizes, ranging from small to large, the nonparametric James–Stein estimator consistently outperforms the classical nonparametric estimator, $\bar{X}$, in the multivariate setting.

A. Appendix: Theoretical Derivations and Proofs

Lemma 1

The following lemma illustrates a similarity between behaviors of empirical and parametric likelihood functions.

LEMMA A1. *Define θ_M to be a root of the equation $n^{-1}\sum_{i=1}^{n} G(X_i, \theta_M) = 0$, where $\partial G(X_i, \theta)/\partial\theta < 0$ (or $\partial G(X_i, \theta)/\partial\theta > 0$), for all $i = 1, \ldots, n$. Then the argument θ_M is a global maximum of the function*

$$W(\theta) = \max\{\prod_{i=1}^{n} p_i : 0 < p_i < 1, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i G(X_i, \theta) = 0\}$$

that increases and decreases monotonically for $\theta < \theta_M$ and $\theta > \theta_M$, respectively.

For example, when $G(u, \theta) = u - \theta$ we obtain $\theta_M = \bar{X} = n^{-1}\sum_{i=1}^{n} X_i$ and the function $W(\theta) = \exp\{\ell_1(\theta)\}$. Now, we can obtain the following results that are analogues to the asymptotic propositions that are well addressed in the parametric literature (DasGupta, 2008; Carlin & Louis, 2000). Lemma 1 can be considered as a non-asymptotic alternative to Lemma 1 presented in Qin & Lawless (1994).

Proof of Lemma 1. It is clear that the argument θ_M, a root of $n^{-1}\sum_{i=1}^{n} G(X_i, \theta_M) = 0$, maximizes the function $W(\theta)$. It follows from the fact that $W(\theta_M) = n^{-n}$ with $p_i = n^{-1}$, $i = 1, \ldots, n$, which maximize $\prod_{i=1}^{n} p_i$ given the sole constraint $\sum_{i=1}^{n} p_i = 1$. Using the Lagrange method, one can represent $W(\theta)$ as

$$W(\theta) = \prod_{i=1}^{n} p_i,\ 0 < p_i = (n + \lambda G(X_i, \theta))^{-1} < 1,\ i = 1, \ldots, n,$$

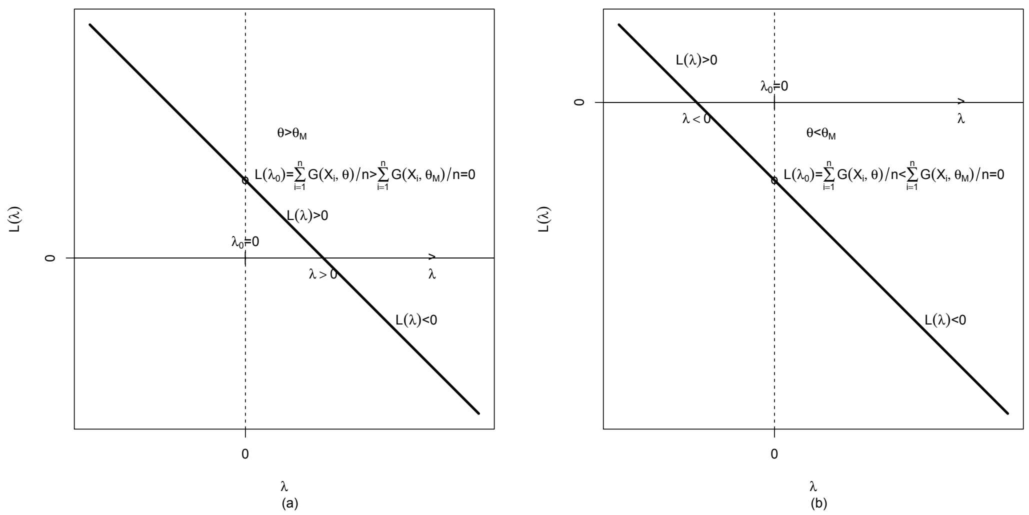

Fig. 1: The schematic behaviors of $L(\lambda)$ plotted against λ (the axis of abscissa), when (a): $\theta > \theta_M$ and (b): $\theta < \theta_M$, respectively.

where the Lagrange multiplier λ is a root of the equation $\sum G(X_i,\theta)\{n+\lambda G(X_i,\theta)\}^{-1} = 0$ (Owen, 2001). This then yields the following expression

$$\frac{d\log\{W(\theta)\}}{d\theta} = -\lambda\sum_{i=1}^{n}\frac{\partial G(X_i,\theta)/\partial\theta}{n+\lambda G(X_i,\theta)} - \sum_{i=1}^{n}\frac{G(X_i,\theta)}{n+\lambda G(X_i,\theta)}\frac{\partial\lambda}{\partial\theta} = -\lambda\sum_{i=1}^{n}\frac{\partial G(X_i,\theta)/\partial\theta}{n+\lambda G(X_i,\theta)}, \quad \text{(A1)}$$

where without loss of generality we assume $\partial G(X_i,\theta)/\partial\theta > 0$, $i = 1,\ldots,n$.

Now define the function $L(\lambda) = \sum_{i=1}^{n} G(X_i,\theta)\{n+\lambda G(X_i,\theta)\}^{-1}$. Since $dL(\lambda)/d\lambda < 0$, the function $L(\lambda)$ decreases with respect to λ and has just one root relative to solving $L(\lambda) = 0$. Consider the scenario with $\theta > \theta_M$. In this case when $\lambda_0 = 0$ we can conclude that

$$L(\lambda_0) = \sum_{i=1}^{n} G(X_i,\theta)(n)^{-1} \geq \sum_{i=1}^{n} G(X_i,\theta_M)(n)^{-1} = 0,$$

since $G(X_i,\theta)$ increases with respect to θ ($\partial G(X_i,\theta)/\partial\theta > 0$).

Note that the function $L(\lambda)$ decreases. This implies that the root of $L(\lambda) = 0$ should be located on the right side from $\lambda_0 = 0$ and then this root is positive. For a graphical representation of this case see Figure 1(a) presented in the Supplementary Material below. Thus, by virtue of (A1), we prove that the function $W(\theta)$ decreases when $\theta > \theta_M$.

Taking the same approach, one can show that the root of $L(\lambda) = 0$ should be to the left of $\lambda_0 = 0$ when $\theta < \theta_M$. For a graphical representation of this case see Figure 1(b). This result combined with (A1) completes the proof of Lemma 1.

Proof of Proposition 1

To prove the proposition, we first show that

$$\int_{X_{(1)}}^{X_{(n)}} \theta^{\upsilon} e^{\ell r_1(\theta)}\pi(\theta)d\theta \cong \int_{\bar{X}-\varphi_n n^{-1/2}}^{\bar{X}+\varphi_n n^{-1/2}} \theta^{\upsilon} e^{\ell r_1(\theta)}\pi(\theta)d\theta,\ \upsilon = 0,1,$$

where a positive sequence $\varphi_n \to \infty$, $\varphi_n n^{-1/2} \to 0$, as $n \to \infty$. This approximation allows us to analyze the numerator ($\upsilon = 1$) and the denominator ($\upsilon = 0$) defined at (2). Let us define the Lagrangian,

$$\Lambda = \sum_{i=1}^{n} \log p_i + \lambda_1 (1 - \sum_{i=1}^{n} p_i) + \lambda_2 (\theta - \sum_{i=1}^{n} p_i X_i),$$

where λ_1 and λ_2 are Lagrange multipliers. Maximizing Λ, one can show that $p_i = \{n + \lambda(X_i - \theta)\}^{-1}$, where λ is a root of the equation $\sum_{i=1}^{n} (X_i - \theta)/\{n + \lambda(X_i - \theta)\} = 0$.

Now define the function

$$L(\lambda) = \sum_{i=1}^{n} (X_i - \theta)/\{n + \lambda(X_i - \theta)\}. \tag{A2}$$

According to Lemma 1, when $\theta < \bar{X}$ then the function $\ell r_1(\theta)$ is strictly increasing and when $\theta > \bar{X}$ then the function $\ell r_1(\theta)$ is strictly decreasing. This implies that the function $\ell r_1(\theta)$ is maximized at the point $\theta = \bar{X}$. Now, denote $a = \bar{X} - \varphi_n n^{-1/2}$ and $b = \bar{X} + \varphi_n n^{-1/2}$, where $\varphi_n = n^{1/6-\beta}$ and $\beta \in (0, 1/6)$. Then it follows that

$$\int_{X_{(1)}}^{X_{(n)}} e^{\ell r_1(\theta)} \pi(\theta) d\theta = \int_{X_{(1)}}^{a} e^{\ell r_1(\theta)} \pi(\theta) d\theta + \int_{a}^{b} e^{\ell r_1(\theta)} \pi(\theta) d\theta + \int_{b}^{X_{(n)}} e^{\ell r_1(\theta)} \pi(\theta) d\theta.$$

By virtue of the above considerations we can bound the remainder term

$$\int_{X_{(1)}}^{a} e^{\ell r_1(\theta)} \pi(\theta) d\theta \leq e^{\ell r_1(a)} \int_{X_{(1)}}^{X_{(n)}} \pi(\theta) d\theta \leq e^{\ell r_1(a)}.$$

In order to arrive at an expression for the value of $\ell r_1(a)$, taking into account the definition of ℓr_1 in (2), we evaluate (A2) at $\theta = a$ such that

$$\begin{aligned} L(\lambda) &= \sum_{i=1}^{n} \frac{(X_i - \bar{X} + \varphi_n n^{-1/2})}{n + \lambda(X_i - \bar{X} + \varphi_n n^{-1/2})} \\ &= \frac{1}{n} \{ \sum_{i=1}^{n} (X_i - \bar{X} + \varphi_n n^{-1/2}) - \lambda n^{-1} \sum_{i=1}^{n} \frac{(X_i - \bar{X} + \varphi_n n^{-1/2})^2}{1 + \lambda n^{-1} (X_i - \bar{X} + \varphi_n n^{-1/2})} \}. \end{aligned} \tag{A3}$$

Defining $\lambda_c = n^{2/3} \tau_n^{-1}$, where $\tau_n = n^{\gamma}$, $0 < \gamma < \beta < 1/6$, and substituting it into (A3) yields

$$n^{1/2} L(\lambda_c) = \varphi_n - n^{1/2} \frac{n^{2/3-1}}{\tau_n} \frac{1}{n} \sum_{i=1}^{n} \frac{(X_i - \bar{X} + \varphi_n n^{-1/2})^2}{1 + n^{-1/3} \tau_n^{-1} (X_i - \bar{X} + \varphi_n n^{-1/2})},$$

Since $(X_i - \bar{X})/(n^{1/3} \tau_n) = O_p(1)$ (Owen, 1988), we have

$$n^{1/2} L(\lambda_c) = \varphi_n - \frac{n^{1/6}}{\tau_n} \frac{1}{n} \sum_{i=1}^{n} \frac{(X_i - \bar{X} + \varphi_n n^{-1/2})^2}{1 + O_p(1)}.$$

Now, it follows that $n^{1/2} L(\lambda_c) \to -\infty$, as $n \to \infty$. In a similar manner, $n^{1/2} L(-\lambda_c) \to \infty$, as $n \to \infty$. Thus, the solution, λ_0, of equation $\sqrt{n} L(\lambda_0) = 0$ belongs to the interval $(-\lambda_c, \lambda_c)$, i.e. $\lambda_0 = O_p(n^{2/3} \tau_n^{-1})$.

Let us now derive the approximate value corresponding to λ_0 as $n \to \infty$. Since $L(\lambda_0) = 0$,

$$\sum_{i=1}^{n} (X_i - \bar{X} + \varphi_n n^{-1/2}) \frac{1}{1 + \lambda_0 n^{-1} (X_i - \bar{X} + \varphi_n n^{-1/2})} = 0. \tag{A4}$$

Applying a Taylor series expansion to (A4) we then obtain

$$\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})\{1-\lambda_0 n^{-1}(X_i-\bar{X}+\varphi_n n^{-1/2})+\frac{\lambda_0^2 n^{-2}(X_i-\bar{X}+\varphi_n n^{-1/2})^2}{(1+\omega_i)^2}\}=0, \tag{A5}$$

where $0<\omega_i<\lambda_0 n^{-1}(X_i-\bar{X}+\varphi_n n^{-1/2})$. Since $\lambda_0=O_p(n^{2/3}\tau_n^{-1})$, we can re-express (A5) as

$$\sum_{i=1}^{n}(X_i-\bar{X}+\frac{\varphi_n}{n^{1/2}})-\frac{\lambda}{n}\sum_{i=1}^{n}(X_i-\bar{X}+\frac{\varphi_n}{n^{1/2}})^2+\frac{O(n^{1/3})}{\tau_n^2}\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}+\frac{\varphi_n}{n^{1/2}})^3=0. \tag{A6}$$

Then it follows that the approximate solution based on solving (A6) is given by

$$\lambda_0=\frac{\varphi_n n^{1/2}}{n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2}+\frac{O(n^{1/3})}{\tau_n^2}. \tag{A7}$$

Applying a Taylor series expansion to $\ell r_1(\theta)$ by (2) with $\theta=a$ yields the following expression

$$\begin{aligned}\ell r_1(a)= \quad & -\sum_{i=1}^{n}\log\{1+\frac{\lambda_0}{n}(X_i-\bar{X}+\varphi_n n^{-1/2})\} = -\sum_{i=1}^{n}\frac{\lambda_0}{n}(X_i-\bar{X}+\varphi_n n^{-1/2})\\ & +\frac{1}{2}\sum_{i=1}^{n}\frac{\lambda_0^2}{n^2}(X_i-\bar{X}+\varphi_n n^{-1/2})^2-\frac{1}{3}\sum_{i=1}^{n}\frac{\lambda_0^3}{n^3}\frac{(X_i-\bar{X}+\varphi_n n^{-1/2})^3}{(1+\omega_i^*)^3},\end{aligned}$$

where $0<\omega_i^*<\lambda_0 n^{-1}(X_i-\bar{X}+\varphi_n n^{-1/2})$. By virtue of (A7) and $\lambda_0=O(n^{2/3}/\tau_n)$ we then have

$$\begin{aligned}\ell r_1(a)&=-\frac{\lambda}{n}\varphi_n n^{1/2}+\frac{1}{2}\sum_{i=1}^{n}\frac{\lambda^2}{n^2}(X_i-\bar{X}+\varphi_n n^{-1/2})^2-O(n^{-3\gamma})\\ &=-\frac{\varphi_n^2 n}{nn^{-1}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2}-\frac{O(n^{4/3})}{\tau_n^2 n^2}\varphi_n n^{1/2}\\ &+\frac{1}{2}\left[\frac{\varphi_n^2 n}{\{n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2\}^2}+2\frac{O(n^{4/3})}{\tau_n^2 n}\frac{\varphi_n^2 n^{1/2}}{n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2}\right.\\ &\left.+\frac{O(n^{8/3})}{\tau_n^4 n^2}\right]\frac{1}{n^2}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2-O(n^{-3\gamma})\\ &=\frac{-\varphi_n^2}{2n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2}-O\left(n^{4/3-2-2\gamma+1/6-\beta+1/2}\right)+O\left(n^{4/3-1-2\gamma+1/6-\beta+1/2-1}\right)\\ &+O\left(n^{8/3-2-4\gamma-1}\right)-O(n^{-3\gamma})\\ &=-\frac{1}{2}\frac{\varphi_n^2}{n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}+\varphi_n n^{-1/2})^2}-O\left(n^{-3\gamma}\right)\to\infty\end{aligned}$$

as $n\to\infty$, where $\varphi_n^2=n^{1/3-2\beta}\to\infty$ and $0<\gamma<\beta<1/6$. Thus, we arrive at the result that $\int_{X_{(1)}}^{a}\exp\{\ell r_1(\theta)\}\pi(\theta)d\theta\le\exp\{\ell r_1(a)\}=O\left(\exp(-wn^{1/3-2\beta})\right)\to 0$ as $n\to\infty$, where w is a positive constant. It follows similarly that $\int_{b}^{X_{(n)}}\exp\{\ell r_1(\theta)\}\pi(\theta)d\theta\le\exp\{\ell r_1(b)\}=O\left(\exp(-w_1 n^{1/3-2\beta})\right)\to 0$ as well as $\int_{X_{(1)}}^{a}\theta\exp\{\ell r_1(\theta)\}\pi(\theta)d\theta\le O\left(\exp(-w_2 n^{1/3-2\beta})\right)\to 0$, $\int_{b}^{X_{(n)}}\theta\exp\{\ell r_1(\theta)\}\pi(\theta)d\theta\le O\left(\exp(-w_3 n^{1/3-2\beta})\right)\to 0$, where w_1,w_2,w_3 are positive constants and $n\to\infty$.

Now we consider the main term $\int_a^b\theta\exp\{\ell r_1(\theta)\}\pi(\theta)d\theta$ of the marginal distribution defined at (2). This integral consists of $\ell r_1(\theta)$ that, by virtue of the Taylor theorem and (A1), is

$$\ell r_1(\theta)=\ell r_1(\bar{X})+(\theta-\bar{X})\lambda(\bar{X})+\frac{1}{2}(\theta-\bar{X})^2\left(\frac{d\lambda(u)}{du}\Big|_{u=\bar{X}}\right) \tag{A8}$$

$$+\frac{1}{6}(\theta-\bar{X})^3\left(\frac{d^2\lambda(u)}{du^2}\Big|_{u=\bar{X}}\right)+\frac{1}{24}(\theta-\bar{X})^4\left(\frac{d^3\lambda(u)}{du^3}\Big|_{u=\theta+\varpi(\bar{X}-\theta)}\right),\varpi\in(0,1).$$

Since the function $\lambda(u)$ is defined by $\sum(X_i-u)/\{n+\lambda(u)(X_i-u)\}$, one can show that

$$\frac{d\lambda(\theta)}{d\theta}=-\frac{n\sum_{i=1}^n p_i^2}{\sum_{i=1}^n(X_i-\theta)^2p_i^2},$$
$$\frac{d^2\lambda(\theta)}{d\theta^2}=\frac{2(d\lambda(\theta)/d\theta)^2\sum_{i=1}^n(X_i-\theta)^3p_i^3+4n(d\lambda(\theta)/d\theta)\sum_{i=1}^n(X_i-\theta)p_i^3-2n\lambda(\theta)\sum_{i=1}^n p_i^3}{\sum_{i=1}^n(X_i-\theta)^2p_i^2},$$
$$\frac{d^3\lambda(\theta)}{d\theta^3}=\left\{\sum_{i=1}^n(X_i-\theta)^2p_i^2\right\}^{-1}\left[6\frac{d\lambda(\theta)}{d\theta}\frac{d^2\lambda(\theta)}{d\theta^2}\sum_{i=1}^n(X_i-\theta)^3p_i^3+6n\frac{d^2\lambda(\theta)}{d\theta^2}\sum_{i=1}^n(X_i-\theta)p_i^3\right.$$
$$-6\left(\frac{d\lambda(\theta)}{d\theta}\right)^3\sum_{i=1}^n(X_i-\theta)^4p_i^4-18n\left(\frac{d\lambda(\theta)}{d\theta}\right)^2\sum_{i=1}^n(X_i-\theta)^2p_i^4+12n\left(\frac{d\lambda(\theta)}{d\theta}\right)\sum_{i=1}^n p_i^3$$
$$\left.-\left\{18n^2\left(\frac{d\lambda(\theta)}{d\theta}\right)+6n(\lambda(\theta))^2\right\}\sum_{i=1}^n p_i^4\right],\quad p_i=\frac{1}{n+\lambda(\theta)(X_i-\theta)}.$$

Noting that $\bar{X}=\arg\max_\theta \ell r_1(\theta)$, $\ell r_1(\bar{X})=0$ and $\lambda(\bar{X})=0$, we have

$$\frac{d\lambda(\theta)}{d\theta}\Big|_{\theta=\bar{X}}=-\frac{n}{n^{-1}\sum_{i=1}^n(X_i-\bar{X})^2}=-\frac{n}{\sigma_n^2},\ \frac{d^2\lambda(\theta)}{d\theta^2}\Big|_{\theta=\bar{X}}=\frac{2n\sum_{i=1}^n(X_i-\bar{X})^3/n}{\left\{n^{-1}\sum_{i=1}^n(X_i-\bar{X})^2\right\}^3}=\frac{2nM_n^3}{(\sigma_n^2)^3}$$

as well as $d^3\lambda(\theta)/d\theta^3=O(n)$, for $\theta\in(a,b)$. This result follows using the same techniques applied to the previous proofs and employing results found in Owen (1988) and Lazar & Mykland (1998) where one can derive the following expressions:

$$\lambda=\frac{\sum_{i=1}^n(X_i-\bar{X})}{n^{-1}\sum_{i=1}^n(X_i-\bar{X})^2}+\frac{O(n^{1/3})}{\tau_n^2}=O(n^{2/3-\beta}),\ \frac{\lambda(\theta)}{n}(X_i-\theta)=O(1),$$
$$p_i=\frac{1}{n}\left\{1+\frac{\lambda(\theta)}{n}(X_i-\theta)\right\}^{-1}=O(n^{-1}),$$

when $|\bar{X}-\theta|\le\varphi_n n^{-1/2}=n^{-1/3-\beta}$, $0<\beta<1/6$. The above asymptotic results, (A8) and a Taylor expansion imply

$$\int_a^b e^{\ell r_1(\theta)}\pi(\theta)d\theta=\int_a^b\exp\left\{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2+\frac{nM_n^3}{3(\sigma_n^2)^3}(\theta-\bar{X})^3+O(n)(\theta-\bar{X})^4\right\}\pi(\theta)d\theta$$
$$=\int\exp\left\{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2\right\}\pi(\theta)d\theta+\frac{nM_n^3}{3(\sigma_n^2)^3}\int(\theta-\bar{X})^3\exp\left\{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2\right\}\pi(\theta)d\theta$$
$$+O(n)\int_a^b(\theta-\bar{X})^4\exp\left\{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2\right\}\pi(\theta)d\theta. \qquad \text{(A9)}$$

It follows similarly that

$$\int_a^b(\theta-\bar{X})e^{\ell r_1(\theta)}\pi(\theta)d\theta=\int(\theta-\bar{X})\exp\left\{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2\right\}\pi(\theta)d\theta+\frac{nM_n^3}{3(\sigma_n^2)^3} \qquad \text{(A10)}$$
$$\times\int(\theta-\bar{X})^4\exp\left\{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}\right\}\pi(\theta)d\theta+O(n)\int_a^b(\theta-\bar{X})^5\exp\left\{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}\right\}\pi(\theta)d\theta.$$

By virtue of the definition (2), the nonparametric posterior expectation $\hat{\theta}$ can be represented in the form of

$$\hat{\theta} = \frac{\int \theta \exp\left\{-\frac{n}{2\sigma_n^2}(\theta - \bar{X})^2\right\}\pi(\theta)d\theta}{\int \exp\left\{-\frac{n}{2\sigma_n^2}(\theta - \bar{X})^2\right\}\pi(\theta)d\theta} + Q_n,$$

where

$$Q_n \equiv \frac{\int_a^b \theta e^{\ell r_1(\theta)}\pi(\theta)d\theta \int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta - \int \theta e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int_a^b e^{\ell r_1(\theta)}\pi(\theta)d\theta}{\int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int_a^b e^{\ell r_1(\theta)}\pi(\theta)d\theta}$$
$$= \left\{\int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int_a^b e^{\ell r_1(\theta)}\pi(\theta)d\theta\right\}^{-1}\left\{\int_a^b (\theta - \bar{X})e^{\ell r_1(\theta)}\pi(\theta)d\theta\right.$$
$$\left.\int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta - \int (\theta-\bar{X})e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int_a^b e^{\ell r_1(\theta)}\pi(\theta)d\theta\right\}.$$

It is clear that, taking into account the results (A9), (A10), the facts $\pi(\theta) = \pi(\bar{X}) + (\theta - \bar{X})\pi'(\bar{X}) + 0.5(\theta - \bar{X})^2\pi''(\bar{X} + q(\theta - \bar{X}))$, $q \in (0,1)$, $\int (\theta - \bar{X})e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}d\theta = 0$ and $b - a = n^{1/6-\beta}/n^{1/2}$, we obtain

$$Q_n = \left\{\frac{nM_n^3}{3(\sigma_n^2)^3}\int (\theta-\bar{X})^4 e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta\right.$$
$$+ O(n)\int_a^b (\theta-\bar{X})^5 e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta$$
$$- \frac{nM_n^3}{3(\sigma_n^2)^3}\int (\theta-\bar{X})^3 e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int (\theta-\bar{X}) e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta$$
$$\left. - O(n)\int_a^b (\theta-\bar{X})^4 e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int (\theta-\bar{X}) e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta\right\}$$
$$\times\left[\left\{\int e^{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}}\pi(\theta)d\theta\right\}^2 + \frac{nM_n^3}{3(\sigma_n^2)^3}\int (\theta-\bar{X})^3 e^{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}}\pi(\theta)d\theta \int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta\right.$$
$$\left. + O(n)\int_a^b (\theta-\bar{X})^4 e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta \int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}\pi(\theta)d\theta\right]^{-1}$$
$$= \left\{\frac{nM_n^3}{3(\sigma_n^2)^3}\int (\theta-\bar{X})^4 e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}d\theta \int e^{-\frac{n}{2\sigma_n^2}(\theta-\bar{X})^2}d\theta\right.$$
$$\left. + O(n)\int_a^b (\theta-\bar{X})^5 e^{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}}d\theta \int e^{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}}d\theta\right\}\left\{\int e^{-\frac{n(\theta-\bar{X})^2}{2\sigma_n^2}}d\theta\right\}^{-2} + O(n^{-1/2-6\beta}).$$

Computing the definite integrals written above, we deduce that

$$\hat{\theta} = \frac{\int \theta \exp\left\{-\frac{n}{2\sigma_n^2}(\theta - \bar{X})^2\right\}\pi(\theta)d\theta}{\int \exp\left\{-\frac{n}{2\sigma_n^2}(\theta - \bar{X})^2\right\}\pi(\theta)d\theta}$$
$$+ \frac{\frac{2nM_n^3}{3(\sigma_n^2)^3}\frac{4!(\pi)^{1/2}(2\sigma_n^2)^{5/2}}{2!n^{5/2}2^5}\frac{(2\pi\sigma_n^2)^{1/2}}{n^{1/2}} + O\left(n(n^{1/6-\beta-1/2})^6\right)\frac{(2\pi\sigma_n^2)^{1/2}}{n^{1/2}}}{(2\pi\sigma_n^2)n^{-1}} + O(n^{-1/2-6\beta}).$$

Making use of $\beta = 1/6 - \varepsilon/6$, $\varepsilon > 0$ completes the proof of Proposition 1.

Proof of Corollary 1

Corollary 1 can be proven by directly applying the result of Proposition 1.

Proof of Corollary 2

To prove this corollary, we can use the result

$$\hat{\theta} = \frac{\int_a^b \theta \exp\left\{-\frac{1}{2n}\frac{(\sum X_i - n\theta)^2}{\sigma_n^2}\right\}\pi(\theta)d\theta}{\int_a^b \exp\left\{-\frac{1}{2n}\frac{(\sum X_i - n\theta)^2}{\sigma_n^2}\right\}\pi(\theta)d\theta} + \frac{M_n^3}{n\,\sigma_n^2} + O_p(n^{-3/2+\varepsilon}),$$

where $a = \bar{X} - \varphi_n n^{-1/2}$, $b = \bar{X} + \varphi_n n^{-1/2}$, $\varphi_n = n^{1/6-\beta}$, $0 < \gamma < \beta < 1/6$, $\beta = 1/6 - \varepsilon/6$, $\varepsilon > 0$.

This approximation was obtained above via the process related to the proof of Proposition 1. Applying the Taylor expansion $\pi(\theta) = \pi(\bar{X}) + (\theta - \bar{X})\pi'(\bar{X}) + (\theta - \bar{X})^2\pi''(\bar{X})/2 + (\theta - \bar{X})^3\pi'''(\tilde{X})/6$, $\tilde{X} \in (\theta, \bar{X})$ to the asymptotic form of $\hat{\theta}$, and in a similar manner of the Laplace method (Bleistein & Handelsman, 2010, p.180), we compete the proof.

Proof of Corollary 3

The proof of Corollary 3 is technical and follows directly from the proof scheme of Proposition 1. Thus the proof is omitted.

Proof of Proposition 2

The proof of this proposition is similar to that of Proposition 1.

Proof of Proposition 3

The proof of this proposition is similar to that of Proposition 1 and Corollary 2.

Proof of Proposition 4

We begin with the asymptotic analysis related to the numerator of the definition of the proposed posterior expectation $\hat{D}_G$. To approximate the double integral $\int\int D(\theta_1,\theta_2)e^{\ell r_3(\theta_1,\theta_2)}\pi(\theta_1,\theta_2)d\theta_1 d\theta_2$, we first show that the main term of the integral is

$$\int_a^b \int_{a_1}^{b_1} D(\theta_1,\theta_2)e^{\ell r_3(\theta_1,\theta_2)}\pi(\theta_1,\theta_2)d\theta_1 d\theta_2,$$

where $a = \bar{X} - \varphi_n n^{-1/2}$, $b = \bar{X} + \varphi_n n^{-1/2}$, $a_1 = \overline{X^2} - \varphi_n n^{-1/2}$, $b_1 = \overline{X^2} + \varphi_n n^{-1/2}$, $\varphi_n = n^{1/6-\beta}$, $0 < \beta < 1/6$ and $\overline{X^2} = \sum_{i=1}^n X_i^2/n$. Since

$$\begin{aligned}&\int_{X_{(1)}}^a \int_{V_1}^{V_2} D(\theta_1,\theta_2)e^{\ell r_3(\theta_1,\theta_2)}\pi(\theta_1,\theta_2)d\theta_1 d\theta_2 \\ &\le \int_{X_{(1)}}^a \int_{V_1}^{V_2} D(\theta_1,\theta_2)\pi(\theta_1,\theta_2)d\theta_2 e^{\ell r_1(\theta_1)}d\theta_1, \quad V_1 = \min_{i=1,\ldots,n} X_i^2, V_2 = \max_{i=1,\ldots,n} X_i^2,\end{aligned}$$

in a similar manner to the proof of Proposition 1, we conclude

$$\begin{aligned}&\int_{X_{(1)}}^a \int_{V_1}^{V_2} D(\theta_1,\theta_2)\pi(\theta_1,\theta_2)d\theta_2 e^{\ell r_1(\theta_1)}d\theta_1 \\ &\le \int_{X_{(1)}}^a \int_{V_1}^{V_2} D(\theta_1,\theta_2)\pi(\theta_1,\theta_2)d\theta_2 d\theta_1 e^{\ell r_1(a)} = O\left(e^{-wn^{1/3-2\beta}}\right) \to 0,\end{aligned}$$

where w is a positive constant and $n \to \infty$.

Likewise, we have $\int_b^{X_{(n)}} \int_{V_1}^{V_2} D(\theta_1,\theta_2)e^{\ell r_3(\theta_1,\theta_2)}\pi(\theta_1,\theta_2)d\theta_1 d\theta_2 = O(e^{-wn^{1/3-2\beta}}) \to 0$, as $n \to \infty$.

Now, we define $\ell r_5(\theta) = n\log n + \max\limits_{0<p_1,\ldots,p_n<1}\{\sum_{i=1}^n \log p_i : \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i X_i^2 = \theta\}$ It is clear that $\ell r_5(\theta) \geq \ell r_3(\theta_1,\theta)$ for all (θ_1,θ) and hence

$$\int_a^b \int_{V_1}^{a_1} D(\theta_1,\theta_2) e^{\ell r_3(\theta_1,\theta_2)} \pi(\theta_1,\theta_2) d\theta_1 d\theta_2 \leq \int_a^b \int_{V_1}^{a_1} D(\theta_1,\theta_2)\pi(\theta_1,\theta_2) d\theta_1 e^{\ell r_5(\theta_2)} d\theta_2$$
$$\leq \int_a^b \int_{V_1}^{a_1} D(\theta_1,\theta_2)\pi(\theta_1,\theta_2) d\theta_1 d\theta_2 e^{\ell r_5(a_1)} = O(e^{-wn^{1/3-2\beta}}) \to 0,$$

as well as $\int_a^b \int_{b_1}^{V_2} D(\theta_1,\theta_2) e^{\ell r_3(\theta_1,\theta_2)} \pi(\theta_1,\theta_2) d\theta_1 d\theta_2 \to 0$, $n\to\infty$.

In order to apply almost directly the proof scheme of Proposition 1, we note that

$$\ell r_3(\theta_1,\theta_2) = -\sum_{i=1}^n \log\{1 + \frac{\lambda_1}{n}(X_i-\theta_1) + \frac{\lambda_2}{n}(X_i^2-\theta_2)\},$$

where the Lagrange multipliers λ_1 and λ_2 satisfy

$$L_1(\theta_1,\theta_2) \equiv \sum_{i=1}^n (X_i-\theta_1)p_i = 0 \text{ and} \tag{A11}$$
$$L_2(\theta_1,\theta_2) \equiv \sum_{i=1}^n (X_i^2-\theta_2)p_i = 0, \quad p_i = \left(n + \lambda_1(X_i-\theta_1) + \lambda_2(X_i^2-\theta_2)\right)^{-1}$$

(Owen, 2001). Since (A11), one can show that

$$\frac{\partial \ell r_3(\theta_1,\theta_2)}{\partial\theta_1} = \lambda_1(\theta_1,\theta_2) \text{ and } \frac{\partial \ell r_3(\theta_1,\theta_2)}{\partial\theta_2} = \lambda_2(\theta_1,\theta_2) \tag{A12}$$

Then the fact $\lambda_1(\bar{X},\overline{X^2}) = 0$, $\lambda_2(\bar{X},\overline{X^2}) = 0$ and a Taylor expansion argument yield

$$\ell r_3(\theta_1,\theta_2) = \frac{1}{2}(\theta_1-\bar{X})^2 \frac{\partial\lambda_1}{\partial\theta_1}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} + (\theta_1-\bar{X})(\theta_2-\overline{X^2})\frac{\partial\lambda_1}{\partial\theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} \tag{A13}$$
$$+\frac{1}{2}(\theta_2-\overline{X^2})^2\frac{\partial\lambda_2}{\partial\theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} + \frac{1}{3!}\Bigg\{(\theta_1-\bar{X})^3\frac{\partial^2\lambda_1}{\partial\theta_1^2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}}$$
$$+3(\theta_1-\bar{X})^2(\theta_2-\overline{X^2})\frac{\partial^2\lambda_1}{\partial\theta_1\partial\theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} + 3(\theta_1-\bar{X})(\theta_2-\overline{X^2})^2\frac{\partial^2\lambda_2}{\partial\theta_1\partial\theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}}$$
$$+(\theta_2-\overline{X^2})^3\frac{\partial^2\lambda_2}{\partial\theta_2^2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}}\Bigg\} + O(n^{-1/3-4\beta}),$$

when $\theta_1\in(a,b)$, $\theta_2\in(a_1,b_1)$, $0<\beta<1/6$ and where

$$\frac{\partial\lambda_1}{\partial\theta_1}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} = -\frac{n}{\sigma_n^2-(\sigma_{12,n})^2/\sigma_{02,n}^2},\ \frac{\partial\lambda_2}{\partial\theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} = -\frac{n}{\sigma_{02,n}^2-(\sigma_{12,n})^2/\sigma_n^2}, \tag{A14}$$
$$\frac{\partial\lambda_1}{\partial\theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} = \frac{\partial\lambda_2}{\partial\theta_1}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} = \frac{n}{\sigma_{02,n}^2\sigma_n^2/\sigma_{12,n}-\sigma_{12,n}},$$
$$\frac{\partial^2\lambda_1}{\partial\theta_k^2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} = \frac{2n^{-2}\sigma_{12,n}\sum_{i=1}^n(X_i^2-\overline{X^2})\Psi_{ki} - \sigma_{02,n}^2\sum_{i=1}^n(X_i-\bar{X})\Psi_{ki}}{(\sigma_{12,n})^2-\sigma_{02,n}^2\sigma_n^2},$$
$$\frac{\partial^2\lambda_2}{\partial\theta_k^2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} = \frac{2n^{-2}\sigma_n^2\sum_{i=1}^n(X_i^2-\overline{X^2})\Psi_{ki} - \sigma_{12,n}\sum_{i=1}^n(X_i-\bar{X})\Psi_{ki}}{\sigma_{02,n}^2\sigma_n^2-(\sigma_{12,n})^2},$$
$$\frac{\partial^2\lambda_1}{\partial\theta_1\partial\theta_2} = \frac{\partial^2\lambda_2}{\partial\theta_1^2} \text{ and } \frac{\partial^2\lambda_2}{\partial\theta_1\partial\theta_2} = \frac{\partial^2\lambda_1}{\partial\theta_2^2}$$

that can be obtained by utilizing (A11) and (A12) with the definitions

$$\Psi_{ki} = \left\{ \frac{\partial \lambda_1}{\partial \theta_k}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} (X_i - \bar{X}) + \frac{\partial \lambda_2}{\partial \theta_k}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} (X_i^2 - \overline{X^2}) \right\}^2, \ k = 1, 2,$$

$$\sigma_{02,n}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - \overline{X^2})^2, \ \sigma_{12,n} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(X_i^2 - \overline{X^2}).$$

The validity of the Proposition 4 follows by arguments similar to those of the proof of Proposition 1 (see the proof scheme from (A8) to the end of the Proposition 1's proof) where the Taylor expansion for

$$D(\theta_1,\theta_2) = D(\bar{X},\overline{X^2}) + (\theta_1 - \bar{X})\frac{\partial D(\theta_1,\theta_2)}{\partial \theta_1}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} + (\theta_2 - \overline{X^2})\frac{\partial D(\theta_1,\theta_2)}{\partial \theta_2}\Big|_{\theta_1=\bar{X},\theta_2=\overline{X^2}} + ...$$

is applied evaluating a Q_n-type remainder term (see the remainder term Q_n and its analysis in the proof of Proposition 1). In this case, we present the remainder term, J_n, which appears in the expansion of Proposition 4, in the integral form

$$J_n = \Bigg[\int\int \left\{ (\theta_1 - \bar{X})\frac{\partial D(t_1,t_2)}{\partial t_1}\Big|_{t_1=\bar{X},t_2=\overline{X^2}} + (\theta_2 - \overline{X^2})\frac{\partial D(t_1,t_2)}{\partial t_2}\Big|_{t_1=\bar{X},t_2=\overline{X^2}} \right\} \tag{A15}$$

$$\times \frac{1}{6}\Bigg\{ (\theta_1 - \bar{X})^3 \frac{\partial^2 \lambda_1}{\partial t_1^2}\Big|_{t_1=\bar{X},t_2=\overline{X^2}} + 3(\theta_1 - \bar{X})^2(\theta_2 - \overline{X^2})\frac{\partial^2 \lambda_1}{\partial t_1 \partial t_2}\Big|_{t_1=\bar{X},t_2=\overline{X^2}}$$

$$+ 3(\theta_1 - \bar{X})(\theta_2 - \overline{X^2})^2 \frac{\partial^2 \lambda_2}{\partial t_1 \partial t_2}\Big|_{t_1=\bar{X},t_2=\overline{X^2}} + (\theta_2 - \overline{X^2})^3 \frac{\partial^2 \lambda_2}{\partial t_2^2}\Big|_{t_1=\bar{X},t_2=\overline{X^2}} \Bigg\}$$

$$\times e^{-\frac{0.5n(\theta_1-\bar{X})^2}{\sigma_n^2-(\sigma_{12,n})^2/\sigma_{02,n}^2} - \frac{0.5n(\theta_2-\overline{X^2})^2}{\sigma_{02,n}^2-(\sigma_{12,n})^2/\sigma_n^2} + \frac{n(\theta_1-\bar{X})(\theta_2-\overline{X^2})}{\sigma_{02,n}^2\sigma_n^2/\sigma_{12,n}-\sigma_{12,n}}} \pi(\theta_1,\theta_2)d\theta_1 d\theta_2 \Bigg]$$

$$\times \left\{ \int\int e^{-\frac{0.5n(\theta_1-\bar{X})^2}{\sigma_n^2-(\sigma_{12,n})^2/\sigma_{02,n}^2} - \frac{0.5n(\theta_2-\overline{X^2})^2}{\sigma_{02,n}^2-(\sigma_{12,n})^2/\sigma_n^2} + \frac{n(\theta_1-\bar{X})(\theta_2-\overline{X^2})}{\sigma_{02,n}^2\sigma_n^2/\sigma_{12,n}-\sigma_{12,n}}} \pi(\theta_1,\theta_2)d\theta_1 d\theta_2 \right\}^{-1},$$

where the corresponding derivatives of λ_1 and λ_2 are defined in (A14).

Proof of Proposition 5

The proof of Proposition 5 is technical and follows directly from the steps used to prove Propositions 1 and 4, respectively. Thus the proof is omitted.

REFERENCES

BLEISTEIN, N. & HANDELSMAN, R. A. (2010). *Asymptotic Expansions of Integrals*. New York: Courier Dover Publications.

CARLIN, B. P. & LOUIS, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. New York: Chapman and Hall/CRC.

DASGUPTA, A. (2008). *Asymptotic theory of statistics and probability*. New York: Springer.

LAZAR, N. & MYKLAND, P. A. (1998). An evaluation of the power and conditionality properties of empirical likelihood. *Biometrika* **85**, 523–534.

LAZAR, N. A. (2003). Bayesian empirical likelihood. *Biometrika* **90**, 319–326.

OWEN, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–249.

OWEN, A. B. (2001). *Empirical likelihood*. New York: Chapman and Hall/CRC.

QIN, J. & LAWLESS, J. (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics* **22**, 300–325.

WU, J., WONG, A. & JIANG, G. (2003). Likelihood-based confidence intervals for a log-normal mean. *Statistics in medicine* **22**, 1849–1860.

Table 1: Monte Carlo means and variances of $\bar{X}$ and θ_{MLE}.

	$X_1,\ldots,X_n \sim N(1,1)$		$X_1,\ldots,X_n \sim LogN(0,1)$			
	$\bar{X}$		$\bar{X}$		θ_{MLE}	
n	mean	var	mean	var	mean	var
10	1.003	0.100	1.651	0.456	1.796	0.614
20	0.997	0.050	1.645	0.238	1.706	0.248
30	0.999	0.034	1.653	0.154	1.688	0.153
50	1.002	0.020	1.652	0.092	1.673	0.087
75	0.999	0.014	1.659	0.063	1.665	0.057

Table 2: Monte Carlo means and variances of the estimator $\hat{\theta}$ by (2) and its asymptotic forms $\hat{\theta}_{P1}$ and $\hat{\theta}_{C1}$ obtained by Proposition 1 and Corollary 1, respectively.

	$X_1,\ldots,X_n \sim N(1,1)$						$X_1,\ldots,X_n \sim LogN(0,1)$					
	Prior: $\pi \sim N(0,1)$						Prior: $\pi \sim N(0,1)$					
	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$	
n	mean	var	mean	var	mean	var	mean	var	mean	var	mean	var
10	0.916	0.096	0.916	0.095	0.916	0.095	1.411	0.177	1.238	0.238	1.201	0.276
20	0.949	0.049	0.952	0.048	0.952	0.048	1.528	0.109	1.371	0.136	1.361	0.145
30	0.962	0.032	0.964	0.032	0.964	0.032	1.564	0.083	1.434	0.097	1.430	0.100
50	0.981	0.020	0.981	0.020	0.981	0.020	1.600	0.061	1.503	0.063	1.502	0.064
75	0.987	0.013	0.987	0.013	0.987	0.013	1.620	0.045	1.547	0.045	1.547	0.045
	Prior: $\pi \sim N(1,0.5^2)$						Prior: $\pi \sim N(\exp(1/2),0.5^2)$					
	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$	
n	mean	var	mean	var	mean	var	mean	var	mean	var	mean	var
10	1.000	0.058	1.000	0.057	1.000	0.057	1.579	0.111	1.514	0.111	1.514	0.111
20	1.000	0.036	1.000	0.036	1.000	0.036	1.643	0.076	1.573	0.073	1.573	0.073
30	1.000	0.027	1.000	0.027	1.000	0.027	1.659	0.065	1.593	0.060	1.593	0.060
50	1.000	0.017	1.000	0.017	1.000	0.017	1.675	0.052	1.615	0.047	1.615	0.047
75	1.000	0.012	1.000	0.012	1.000	0.012	1.686	0.042	1.633	0.036	1.633	0.036
	Prior: $\pi \sim N(1,0.1^2)$						Prior: $\pi \sim N(\exp(1/2),0.1^2)$					
	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$	
n	mean	var	mean	var	mean	var	mean	var	mean	var	mean	var
10	1.000	0.002	1.000	0.002	1.000	0.002	1.619	0.008	1.612	0.008	1.612	0.008
20	1.000	0.002	1.000	0.002	1.000	0.002	1.634	0.003	1.621	0.005	1.621	0.005
30	1.000	0.002	1.000	0.002	1.000	0.002	1.638	0.002	1.625	0.004	1.625	0.004
50	0.999	0.002	0.999	0.002	0.999	0.002	1.642	0.002	1.630	0.003	1.630	0.003
75	1.001	0.002	1.001	0.003	1.001	0.003	1.644	0.002	1.633	0.003	1.633	0.003
	Prior: $\pi \sim N(0,0.5^2)$						Prior: $\pi \sim N(\exp(1/2)-1,0.5^2)$					
	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$	
n	mean	var	mean	var	mean	var	mean	var	mean	var	mean	var
10	0.755	0.127	0.752	0.126	0.752	0.126	1.250	0.213	1.094	0.338	1.079	0.356
20	0.834	0.063	0.843	0.061	0.843	0.061	1.361	0.127	1.210	0.219	1.207	0.223
30	0.877	0.042	0.884	0.040	0.884	0.040	1.424	0.088	1.290	0.152	1.289	0.154
50	0.923	0.023	0.927	0.022	0.927	0.022	1.495	0.055	1.390	0.088	1.390	0.089
75	0.948	0.015	0.950	0.015	0.950	0.015	1.528	0.042	1.447	0.060	1.447	0.060
	Prior: $\pi \sim \{N(-1,0.5^2)+N(1,0.5^2)\}/2$						Prior: $\pi \sim \{N(-e^{1/2},0.5^2)+N(e^{1/2},0.5^2)\}/2$					

Table 3: The Monte Carlo means and variances of $\bar{X}_j$ and the estimator $\hat{\theta}_{Ej}$ by (6) based on $MVN\{(1,1,1)^T, I\}$.

n	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$V(\bar{X}_1)$	$V(\bar{X}_2)$	$V(\bar{X}_3)$	$\hat{\theta}_{E1}$	$\hat{\theta}_{E2}$	$\hat{\theta}_{E3}$	$V(\hat{\theta}_{E1})$	$V(\hat{\theta}_{E2})$	$V(\hat{\theta}_{E3})$
10	(0.995	0.995	0.989)	(0.096	0.100	0.099)	(0.994	0.995	0.990)	(0.062	0.065	0.064)
20	(1.002	0.993	1.000)	(0.052	0.050	0.048)	(1.001	0.995	0.999)	(0.033	0.032	0.031)
30	(0.999	1.002	1.000)	(0.033	0.034	0.034)	(0.999	1.001	1.000)	(0.021	0.021	0.022)
50	(0.999	1.000	1.000)	(0.020	0.020	0.020)	(0.999	1.000	1.000)	(0.013	0.013	0.013)
75	(1.003	1.000	0.999)	(0.014	0.014	0.014)	(1.003	1.000	1.000)	(0.009	0.009	0.009)

	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$	
n	mean	var	mean	var	mean	var	mean	var	mean	var	mean	var
10	0.990	0.065	0.991	0.064	0.999	0.058	1.582	0.110	1.516	0.108	1.516	0.108
20	0.995	0.034	0.995	0.035	0.996	0.034	1.640	0.077	1.570	0.073	1.570	0.073
30	1.000	0.026	1.000	0.026	1.001	0.026	1.663	0.063	1.597	0.058	1.597	0.058
50	0.995	0.017	0.995	0.017	0.995	0.017	1.672	0.051	1.613	0.046	1.613	0.046
75	0.998	0.012	0.998	0.012	0.998	0.012	1.686	0.041	1.634	0.036	1.634	0.036

Prior: $\pi \sim \{N(-1, 0.1^2) + N(1, 0.1^2)\}/2$ — Prior: $\pi \sim \{N(-e^{1/2}, 0.1^2) + N(e^{1/2}, 0.1^2)\}/2$

	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}_{C1}$	
n	mean	var	mean	var	mean	var	mean	var	mean	var	mean	var
10	0.996	0.006	0.998	0.004	1.000	0.002	1.618	0.009	1.610	0.009	1.611	0.009
20	0.999	0.002	0.999	0.002	0.999	0.002	1.634	0.003	1.620	0.005	1.620	0.005
30	1.000	0.002	1.000	0.002	1.000	0.002	1.639	0.002	1.627	0.003	1.627	0.003
50	1.000	0.002	1.000	0.002	1.000	0.002	1.643	0.002	1.631	0.003	1.631	0.003
75	1.000	0.002	0.999	0.002	0.999	0.002	1.643	0.002	1.632	0.003	1.632	0.003

Prior: $\pi \sim U[0, 1.5]$ — Prior: $\pi \sim U[0, \exp(1/2) + 0.5]$

	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$	
n	mean	var	mean	var	mean	var	mean	var
10	0.942	0.061	0.942	0.061	1.469	0.120	1.363	0.154
20	0.974	0.038	0.976	0.038	1.556	0.068	1.462	0.086
30	0.991	0.028	0.992	0.028	1.595	0.052	1.516	0.063
50	1.000	0.019	1.000	0.019	1.634	0.039	1.571	0.043
75	0.999	0.013	0.999	0.013	1.651	0.032	1.600	0.034

Prior: $\pi \sim U[0.75, 1.25]$ — Prior: $\pi \sim U[\exp(1/2) - 0.25, \exp(1/2) + 0.25]$

	$\hat{\theta}$		$\hat{\theta}_{P1}$		$\hat{\theta}$		$\hat{\theta}_{P1}$	
n	mean	var	mean	var	mean	var	mean	var
10	0.999	0.005	0.999	0.005	1.619	0.006	1.607	0.007
20	1.001	0.006	1.001	0.006	1.629	0.005	1.612	0.007
30	0.999	0.006	0.999	0.006	1.634	0.005	1.616	0.006
50	0.999	0.006	0.998	0.007	1.639	0.005	1.622	0.006
75	1.001	0.007	1.001	0.007	1.641	0.006	1.624	0.007

Table 4: The Monte Carlo means and variances of $\bar{X}_j$ and the estimator $\hat{\theta}_{Ej}$ by (6) based on $MVLogN\{(0,0,0)^T, I\}$.

n	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$V(\bar{X}_1)$	$V(\bar{X}_2)$	$V(\bar{X}_3)$	$\hat{\theta}_{E1}$	$\hat{\theta}_{E2}$	$\hat{\theta}_{E3}$	$V(\hat{\theta}_{E1})$	$V(\hat{\theta}_{E2})$	$V(\hat{\theta}_{E3})$
10	(1.630	1.656	1.638)	(0.446	0.581	0.445)	(1.631	1.655	1.638)	(0.398	0.531	0.395)
20	(1.657	1.654	1.654)	(0.235	0.232	0.239)	(1.657	1.654	1.654)	(0.206	0.204	0.211)
30	(1.644	1.654	1.640)	(0.161	0.162	0.148)	(1.645	1.653	1.640)	(0.140	0.141	0.128)
50	(1.650	1.642	1.640)	(0.093	0.088	0.085)	(1.650	1.642	1.640)	(0.080	0.076	0.073)
75	(1.649	1.645	1.653)	(0.061	0.059	0.063)	(1.649	1.645	1.653)	(0.052	0.051	0.054)

Table 5: The Monte Carlo means and variances of $\bar{X}_j$ and the estimator $\hat{\theta}_{Ej}$ by (6) based on $MVN\{(1,1,1)^T, \Sigma\}$.

n	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$V(\bar{X}_1)$	$V(\bar{X}_2)$	$V(\bar{X}_3)$	$\hat{\theta}_{E1}$	$\hat{\theta}_{E2}$	$\hat{\theta}_{E3}$	$V(\hat{\theta}_{E1})$	$V(\hat{\theta}_{E2})$	$V(\hat{\theta}_{E3})$
10	(0.995	0.993	0.993)	(0.101	0.099	0.101)	(0.994	0.993	0.994)	(0.083	0.082	0.082)
20	(0.996	0.997	0.998)	(0.049	0.049	0.049)	(0.997	0.997	0.998)	(0.040	0.040	0.039)
30	(1.001	1.000	1.000)	(0.033	0.035	0.035)	(1.001	1.000	1.000)	(0.027	0.028	0.028)
50	(1.000	1.003	0.997)	(0.019	0.020	0.020)	(0.999	1.002	0.998)	(0.016	0.016	0.016)
75	(1.002	0.998	0.999)	(0.013	0.014	0.014)	(1.001	0.999	1.000)	(0.011	0.011	0.011)

Table 6: The Monte Carlo means and variances of $\bar{X}_j$ and the estimator $\hat{\theta}_{Ej}$ by (6) based on $MVLogN\{(0,0,0)^T, \Sigma\}$.

n	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$V(\bar{X}_1)$	$V(\bar{X}_2)$	$V(\bar{X}_3)$	$\hat{\theta}_{E1}$	$\hat{\theta}_{E2}$	$\hat{\theta}_{E3}$	$V(\hat{\theta}_{E1})$	$V(\hat{\theta}_{E2})$	$V(\hat{\theta}_{E3})$
10	(1.663	1.637	1.650)	(0.551	0.471	0.480)	(1.663	1.637	1.650)	(0.520	0.441	0.449)
20	(1.651	1.658	1.652)	(0.237	0.232	0.218)	(1.651	1.658	1.652)	(0.218	0.215	0.202)
30	(1.651	1.652	1.641)	(0.148	0.161	0.149)	(1.650	1.651	1.642)	(0.136	0.148	0.137)
50	(1.652	1.648	1.644)	(0.094	0.095	0.092)	(1.652	1.648	1.645)	(0.086	0.087	0.084)
75	(1.648	1.648	1.645)	(0.064	0.061	0.061)	(1.648	1.648	1.645)	(0.058	0.056	0.056)