Bayesian Analysis for the Difference of Exponential Means

Sang Gil Kang[1] · Dal Ho Kim[2] · Woo Dong Lee[3]

Abstract

In this paper, we develop the noninformative priors for the exponential models when the parameter of interest is the difference of two means. We develop the first and second order matching priors. We reveal that the second order matching priors do not exist. It turns out that Jeffreys' prior does not satisfy a first order matching criterion. The Bayesian credible intervals based on the first order matching meet the frequentist target coverage probabilities much better than the frequentist intervals of Jeffreys' prior.

Keywords : Matching Prior; Difference of Two Means; Exponential Distributions.

1. Introduction

The exponential distribution plays an important role in the field of reliability. The reasons for using the exponential distribution assumption in reliability applications can be found in the early work of Davis (1952), Epstein and Sobel (1953), and others. Further justification, in the form of theoretical arguments to support the use of the exponential distribution as the failure law of complex equipment, is presented in the book by Barlow and Proschan (1975) and Lawless (2003).

The present paper focuses on noninformative priors for the difference of two exponential means. We consider Bayesian priors such that the resulting credible intervals for the difference of two exponential means have coverage probabilities equivalent to their frequentist counterparts. Although this matching can be justified only asymptotically, our simulation results indicate that this is indeed achieved for small or moderate sample sizes as well.

This matching idea goes back to Welch and Peers (1963). Interest in such priors revived with the work of Stein (1985) and Tibshirani (1989). Among others, we may cite the work of Mukerjee and Dey (1993), DiCiccio and Stern (1994), Datta and Ghosh (1995a,b, 1996), Mukerjee and Ghosh (1997).

On the other hand, Ghosh and Mukerjee (1992), and Berger and Bernardo (1989,1992) extended Bernardo's (1979) reference prior approach, giving a general algorithm to derive a reference prior by splitting the parameters into several groups according to their order of inferential importance. This approach is very successful in various practical problems. Quite often reference priors satisfy the matching criterion described earlier.

The problem of comparison for two exponential means has been investigated by many authors. For the comparison of two exponential distributions, most of the studies are the ratio of means. However there is a little work in the interval estimation for the

1) Department of Applied Statistics, Sangji University, Wonju, 220-712, Korea

2) Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea

3) Department of Asset Management Science, Daegu Haany University, Kyungsan, 712-240, Korea

difference between two exponential means. Akahira (2002) proposed a systematic method of the construction of a confidence interval for the difference between two means in the exponential distributions. This construction of a confidence interval is as follows:

Suppose that $X_1, \cdots, X_n$ are independent and identically distributed random variables according to the exponential distribution with mean μ_1 and $Y_1, \cdots, Y_n$ are independent and identically distributed random variables according to the exponential distribution with mean μ_2. A condition is given by

$$p(\delta) = E_\delta\left[F_V\left(\frac{n(1-\delta)}{U\tilde{a}\left(\dfrac{U}{U+\delta(1-U)}\right)}\right)\right]\chi_{(0,1]}(\delta) + E_\delta\left[F_V\left(\frac{n(1-\frac{1}{\delta})}{U\tilde{a}\left(\dfrac{1-\frac{1}{\delta}}{U+\frac{1}{\delta}(1-U)}\right)}\right)\right]\chi_{(1,\infty)}(\delta) \le \alpha,$$

where $\chi_A(\cdot)$ is the indicator of a set A, $F(\cdot)$ is the cumulative distribution function, U is a random variable with beta distribution $B(n,n)$ and V is a random variable with gamma distribution $gamma(2n,1)$. Using $\tilde{a}(\cdot)$ satisfying the above condition uniformly in δ, we have

$$P_{\mu_1,\mu_2}[-a(Z_y, Z_x) \le n(\mu_1 - \mu_2) \le a(Z_x, Z_y)] \ge 1-\alpha,$$

where $Z_x = \sum_{i=1}^n X_i$ and $Z_y = \sum_{i=1}^n Y_i$. So $Z[-a(Z_y, Z_x), a(Z_x, Z_y)]$ is the confidence interval for $n(\mu_1 - \mu_2)$ at level $1-\alpha$, where $a(z_1, z_2) = z_1\tilde{a}(z_1/(z_1+z_2))$. But the proposed method only apply for the case of equal sample sizes. For a given significance level and sample size, some function satisfying the above condition will always be computed with respect to the level and sample size, and even the function may not be unique. Also, there is a little work in this problem from the viewpoint of Bayesian framework.

The outline of the remaining sections is as follows. In Section 2, we develop first order and second order probability matching priors for the difference of two exponential means. We revealed that the second order matching prior does not exist. It turns out that the Jeffreys' prior does not satisfy a first order matching criterion. We provide that the propriety of the posterior distribution for the first order matching priors in Section 3. In Section 4, simulated frequentist coverage probabilities under the proposed priors are given. The Bayesian credible intervals based on the first order matching prior meet the frequentist target coverage probabilities much better than the frequentist intervals of the Jeffreys' prior.

2. The Noninformative Priors

For a prior π, let $\theta_1^{1-\alpha}(\pi; Y)$ denote the $(1-\alpha)$th percentile of the posterior distribution of θ_1, that is,

$$P^{\pi}[\theta_1 \le \theta_1^{1-\alpha}(\pi; \boldsymbol{Y}) \mid \boldsymbol{Y}] = 1-\alpha,$$

where $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_t)^T$ and θ_1 is the parameter of interest. We want to find priors π for which

$$P[\theta_1 \le \theta_1^{1-\alpha}(\pi; \boldsymbol{Y}) \mid \boldsymbol{\theta}] = 1-\alpha + o(n^{-u}). \tag{1}$$

for some $u > 0$, as n goes to infinity. Priors π satisfying (1) are called matching priors. If $u = 1/2$, then π is referred to as a first order matching prior, while if $u=1$, π is referred to as a second order matching prior.

Consider that $X_1, \cdots, X_{n_1}$ are independent and identically distributed random variables according to the exponential distribution with mean μ_1 and $Y_1, \cdots, Y_{n_2}$ are independent and identically distributed random variables according to the exponential distribution with mean μ_2. Then the likelihood function is

$$L(\mu_1, \mu_2) \propto \left(\frac{1}{\mu_1}\right)^{n_1}\left(\frac{1}{\mu_2}\right)^{n_2} \exp\left(-\sum_{i=1}^{n_1}\frac{x_i}{\mu_1} - \sum_{i=1}^{n_2}\frac{y_i}{\mu_2}\right), \tag{2}$$

where the $\mu_1 > 0$ and the $\mu_2 > 0$.

In order to find matching priors π, let

$$\theta_1 = \mu_1 - \mu_2 \text{ and } \theta_2 = \frac{n_1}{\mu_1} + \frac{n_2}{\mu_2}.$$

With this parametrization, the likelihood function of parameters (θ_1, θ_2) for the likelihood (2) is given by

$$\begin{aligned} L(\theta_1, \theta_2) \ \propto\ & \theta_2^{n_1+n_2}\{n_1+n_2+\theta_1\theta_2 + [(n_2-n_1+\theta_1\theta_2)^2 + 4n_1n_2]^{1/2}\}^{-n_1} \\ \times\ & \{n_1+n_2-\theta_1\theta_2 + [(n_2-n_1+\theta_1\theta_2)^2 + 4n_1n_2]^{1/2}\}^{-n_2} \\ \times\ & exp\left(-\sum_{i=1}^{n_1}\frac{2\theta_2 x_i}{n_1+n_2+\theta_1\theta_2+[(n_2-n_1+\theta_1\theta_2)^2+4n_1n_2]^{1/2}}\right. \\ -\ & \left.\sum_{i=1}^{n_2}\frac{2\theta_2 y_i}{n_1+n_2-\theta_1\theta_2+[(n_2-n_1+\theta_1\theta_2)^2+4n_1n_2]^{1/2}}\right). \end{aligned} \tag{3}$$

Based on (3), the Fisher information matrix is given by

$$\boldsymbol{I} = \begin{pmatrix} I_{11} & 0 \\ 0 & I_{22} \end{pmatrix},$$

where $g \equiv g(\theta_1, \theta_2) = (n_2 - n_1 + \theta_1\theta_2)^2 + 4n_1n_2$,

$$I_{11} = \frac{8n_1n_2\theta_2^2[(n_1+n_2)g(\theta_1,\theta_2)^{1/2} + (n_1+n_2)^2 + (n_2-n_1)\theta_1\theta_2]}{g(\theta_1,\theta_2)^{1/2}[n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]^2[n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]^2}$$

and

$$I_{22} = \frac{2\,[(n_1+n_2)^2 + (n_2-n_1)\theta_1\theta_2 + (n_1+n_2)g(\theta_1,\theta_2)^{1/2}]^3}{\theta_2^2\, g(\theta_1,\theta_2)^{1/2}[n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]^2[n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]^2}.$$

From the above Fisher information matrix $\boldsymbol{I}$, θ_1 is orthogonal to θ_2 in the sense of Cox and Reid (1987). Following Tibshirani (1989), the class of first order probability matching prior is characterized by

$$\pi_m^{(1)}(\theta_1,\theta_2) \propto \frac{\theta_2\,[(n_1+n_2)g(\theta_1,\theta_2)^{1/2} + (n_1+n_2)^2 + (n_2-n_1)\theta_1\theta_2]^{\frac{1}{2}}}{g(\theta_1,\theta_2)^{1/4}[n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}][n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]}\, d(\theta_2), \tag{4}$$

where $d(\theta_2) > 0$ is an arbitrary function differentiable in its argument.

The class of prior given in (4) can be narrowed down to the second order probability matching priors as given in Mukerjee and Ghosh (1997).

A second order probability matching prior is of the form (4), and also d must satisfy an additional differential equation (cf (2.10)) of Mukerjee and Ghosh (1997), namely

$$\frac{1}{6}d(\theta_2)\frac{\partial}{\partial\theta_1} I_{11}^{-\frac{3}{2}} L_{1,1,1} + \frac{\partial}{\partial\theta_2} I_{11}^{-\frac{1}{2}} L_{112} I^{22} d(\theta_2) = 0, \tag{5}$$

where

$$\begin{aligned} L_{1,1,1} &= E\left[\left(\frac{\partial \log L}{\partial\theta_1}\right)^3\right] \\ &= \frac{h_1(\theta_1,\theta_2)}{g(\theta_1,\theta_2)^{3/2}(n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2})^6(n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2})^6} \\ L_{112} &= E\left[\frac{\partial^3 \log L}{\partial\theta_1^2\partial\theta_2}\right] \\ &= \frac{h_2(\theta_1,\theta_2)}{g(\theta_1,\theta_2)(n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2})^3(n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2})^3}. \end{aligned}$$

Here

$$\begin{aligned} h_1(\theta_1,\theta_2) = &-512 n_1 n_2 \theta_2^3\,[(n_1+n_2)^2 + (n_2-n_1)\theta_1\theta_2 + (n_1+n_2)g(\theta_1,\theta_2)^{1/2}]^3 \\ &\times \{(n_1-n_2)(n_1+n_2)^4 - \theta_1\theta_2\,[3\,(n_1+n_2)^2(n_1^2-n_1n_2+n_2^2) \\ &+ 3\,(n_2^3-n_1^3)\theta_1\theta_2 + (n_2^2+n_1^2)\theta_1^2\theta_2^2\,] + (n_1+n_2)[(n_1-n_2)(n_1+n_2)^2 \\ &- (2n_1^2-n_1n_2+2n_2^2)\theta_1\theta_2 + (n_1-n_2)\theta_1^2\theta_2^2\,]g(\theta_1,\theta_2)^{1/2}\} \end{aligned}$$

and

$$\begin{aligned} h_2(\theta_1,\theta_2) = &-64 n_1 n_2 (n_1+n_2)\theta_2\,[(n_1+n_2)^4 + 2\,(n_2-n_1)(n_1+n_2)^2\theta_1\theta_2 \\ &+ (n_1^2-n_1n_2+n_2^2)\theta_1^2\theta_2^2\,] - 64 n_1 n_2 \theta_2\,[(n_1+n_2)^4 \\ &+ 2\,(n_2-n_1)(n_1+n_2)^2\theta_1\theta_2 + (n_1^2+n_1n_2+n_2^2)\theta_1^2\theta_2^2\,]g(\theta_1,\theta_2)^{1/2}. \end{aligned}$$

Then (5) simplifies to

$$\frac{1}{6}d(\theta_2)\frac{\partial}{\partial\theta_1}\{w_1(\theta_1,\theta_2)\}+\frac{\partial}{\partial\theta_2}\{w_2(\theta_1,\theta_2)d(\theta_2)\}=0, \tag{6}$$

where

$$w_1(\theta_1,\theta_2)$$
$$=\frac{(8n_1n_2)^{-3/2}\theta_2^{-3}h_1(\theta_1,\theta_2)[(n_1+n_2)g^{1/2}+(n_1+n_2)^2+(n_2-n_1)\theta_1\theta_2]^{-3/2}}{g(\theta_1,\theta_2)^{3/4}(n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2})^3(n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2})^3}$$

and

$$w_2(\theta_1,\theta_2)=\frac{(8n_1n_2)^{-1/2}\theta_2h_2(\theta_1,\theta_2)}{2g(\theta_1,\theta_2)^{1/4}[(n_1+n_2)g^{1/2}+(n_1+n_2)^2+(n_2-n_1)\theta_1\theta_2]^{7/2}}.$$

However there can be no solution to (6) unless the d is the function of θ_1 and θ_2. Thus the second order matching prior does not exist.

From the Fish information matrix I, the Jeffreys′ prior is given by

$$\pi_J(\theta_1,\theta_2)$$
$$\propto\frac{[(n_1+n_2)g(\theta_1,\theta_2)^{1/2}+(n_1+n_2)^2+(n_2-n_1)\theta_1\theta_2]^2}{g(\theta_1,\theta_2)^{1/2}[n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]^2[n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]^2}. \tag{7}$$

Remark 1. In the original parameterization (μ_1,μ_2), the first order matching prior is given by

$$\pi_m^{(1)}(\mu_1,\mu_2)\propto\mu_1^{-1}\mu_2^{-1}\left(\frac{n_1}{\mu_1^2}+\frac{n_2}{\mu_2^2}\right)^{1/2}d(\theta_2(\mu_1,\mu_2)).$$

And the Jeffreys′ prior is given by

$$\pi_J(\mu_1,\mu_2)\propto\mu_1^{-1}\mu_2^{-1}. \tag{8}$$

Notice that the matching priors (4) include many different matching priors because of the arbitrary selection of the function d. However every function is not permissible in the construction of priors. For instance, we consider any function of the form θ_2^{-a}. If a is negative integer, then the posterior distribution under function of the form θ_2^{-a} is proper. But the condition of propriety in this form strongly depend on the a. Moreover the posterior under this form is complex. Also there does not seem to be any improvement in the coverage probabilities with this posterior distribution. So we have chosen d to be a constant function. The resulting prior is given by

$$\pi_m^{(1)}(\theta_1,\theta_2)$$
$$\propto\frac{\theta_2[(n_1+n_2)g(\theta_1,\theta_2)^{1/2}+(n_1+n_2)^2+(n_2-n_1)\theta_1\theta_2]^{1/2}}{g(\theta_1,\theta_2)^{1/4}[n_1+n_2-\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}][n_1+n_2+\theta_1\theta_2+g(\theta_1,\theta_2)^{1/2}]}. \tag{9}$$

Thus in the original parameterization (μ_1,μ_2), the first order matching prior is given by

$$\pi_m^{(1)}(\mu_1,\mu_2)\propto\mu_1^{-1}\mu_2^{-1}\left(\frac{n_1}{\mu_1^2}+\frac{n_2}{\mu_2^2}\right)^{1/2}. \tag{10}$$

3. Implementation of the Bayesian Procedure

We investigate the propriety of posteriors for a general priors which include Jeffreys' prior (8) and the first order matching prior (10). We consider the class of priors

$$\pi_g(\mu_1, \mu_2) \propto \mu_1^{-a} \mu_2^{-b} \left(\frac{n_1}{\mu_1^2} + \frac{n_2}{\mu_2^2} \right)^c, \tag{11}$$

where $a > 0, b > 0$ and $c \geq 0$. The following theorem can be proved.

Theorem 1. The posterior distribution of (μ_1, μ_2) under the prior (11) is proper if $n_1 + a - 1 > 0$ and $n_2 + b - 1 > 0$.

Proof. Under the prior (11), the joint posterior for μ_1 and μ_2 given $\boldsymbol{x}$ and $\boldsymbol{y}$ is

$$\pi(\mu_1, \mu_2 \mid \boldsymbol{x}, \boldsymbol{y}) \propto \left(\frac{1}{\mu_1}\right)^{n_1 + a} \left(\frac{1}{\mu_2}\right)^{n_2 + b} \left(\frac{n_1}{\mu_1^2} + \frac{n_2}{\mu_2^2} \right)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right).$$

For $1 \leq \mu_1 < \infty$ and $1 \leq \mu_2 < \infty$,

$$\int_1^\infty \int_1^\infty \pi(\mu_1, \mu_2 \mid \boldsymbol{x}, \boldsymbol{y}) d\mu_1 d\mu_2$$

$$\leq \int_1^\infty \int_1^\infty \left(\frac{1}{\mu_1}\right)^{n_1 + a} \left(\frac{1}{\mu_2}\right)^{n_2 + b} (n_1 + n_2)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right) d\mu_1 d\mu_2$$

$$\leq \int_0^\infty \int_0^\infty \left(\frac{1}{\mu_1}\right)^{n_1 + a} \left(\frac{1}{\mu_2}\right)^{n_2 + b} (n_1 + n_2)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right) d\mu_1 d\mu_2 < \infty,$$

if $n_1 + a - 1 > 0$ and $n_2 + b - 1 > 0$. For $0 < \mu_1 < 1$ and $1 \leq \mu_2 < \infty$,

$$\int_1^\infty \int_0^1 \pi(\mu_1, \mu_2 \mid \boldsymbol{x}, \boldsymbol{y}) d\mu_1 d\mu_2$$

$$\leq \int_1^\infty \int_0^1 \left(\frac{1}{\mu_1}\right)^{n_1 + a + 2c} \left(\frac{1}{\mu_2}\right)^{n_2 + b} (n_1 + n_2)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right) d\mu_1 d\mu_2$$

$$\leq \int_0^\infty \int_0^\infty \left(\frac{1}{\mu_1}\right)^{n_1 + a + 2c} \left(\frac{1}{\mu_2}\right)^{n_2 + b} (n_1 + n_2)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right) d\mu_1 d\mu_2$$

$$< \infty,$$

if $n_1 + a + 2c - 1 > 0$ and $n_2 + b - 1 > 0$. For $0 < \mu_1 < 1$ and $0 < \mu_2 < 1$,

$$\int_0^1 \int_0^1 \pi(\mu_1, \mu_2 \mid \boldsymbol{x}, \boldsymbol{y}) d\mu_1 d\mu_2$$

$$\leq \int_0^1 \int_0^1 \left(\frac{1}{\mu_1}\right)^{n_1 + a + 2c} \left(\frac{1}{\mu_2}\right)^{n_2 + b + 2c} (n_1 + n_2)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right) d\mu_1 d\mu_2$$

$$\leq \int_0^\infty \int_0^\infty \left(\frac{1}{\mu_1}\right)^{n_1 + a + 2c} \left(\frac{1}{\mu_2}\right)^{n_2 + b + 2c} (n_1 + n_2)^c \exp\left(- \sum_{i=1}^{n_1} \frac{x_i}{\mu_1} - \sum_{i=1}^{n_2} \frac{y_i}{\mu_2} \right) d\mu_1 d\mu_2$$

$$< \infty,$$

if $n_1 + a + 2c - 1 > 0$ and $n_2 + b + 2c - 1 > 0$. This completes the proof.

Theorem 2. The marginal posterior density of θ_1 under the matching prior (9) is given by

$$\begin{aligned}
&\pi_m(\theta_1 \mid \boldsymbol{x}, \boldsymbol{y}) \\
&\propto \int_0^\infty \theta_2^{n_1+n_2+1}\,[n_1+n_2+\theta_1\theta_2+g^{1/2}]^{-n_1-1}[n_1+n_2-\theta_1\theta_2+g^{1/2}]^{-n_2-1} \\
&\times\; g^{-1/4}[(n_1+n_2)g^{1/2}+(n_1+n_2)^2+(n_2-n_1)\theta_1\theta_2]^{\frac{1}{2}} \\
&\times\; exp\left(-\sum_{i=1}^{n_1}\frac{2\theta_2 x_i}{n_1+n_2+\theta_1\theta_2+g^{1/2}}-\sum_{i=1}^{n_2}\frac{2\theta_2 y_i}{n_1+n_2-\theta_1\theta_2+g^{1/2}}\right)d\theta_2.
\end{aligned}$$

And the marginal posterior density of θ_1 under the Jeffreys$'$ prior (7) is given by

$$\begin{aligned}
&\pi_J(\theta_1 \mid \boldsymbol{x}, \boldsymbol{y}) \\
&\propto \int_0^\infty \theta_2^{n_1+n_2}\,[n_1+n_2+\theta_1\theta_2+g^{1/2}]^{-n_1-2}[n_1+n_2-\theta_1\theta_2+g^{1/2}]^{-n_2-2} \\
&\times\; g^{-1/2}[(n_1+n_2)g^{1/2}+(n_1+n_2)^2+(n_2-n_1)\theta_1\theta_2]^{2} \\
&\times\; exp\left(-\sum_{i=1}^{n_1}\frac{2\theta_2 x_i}{n_1+n_2+\theta_1\theta_2+g^{1/2}}-\sum_{i=1}^{n_2}\frac{2\theta_2 y_i}{n_1+n_2-\theta_1\theta_2+g^{1/2}}\right)d\theta_2.
\end{aligned}$$

Actually the normalizing constant for the marginal density of θ_1 requires a two dimensional integration. Therefore we have the marginal posterior density of θ_1, and so it is to compute the marginal moment of θ_1. In Section 4, we investigate the frequentist coverage probabilities for the Jeffreys$'$ prior π_J and first order matching prior π_m, respectively.

4. Numerical Studies and Discussion

We evaluate the frequentist coverage probability by investigating the credible interval of the marginal posteriors density of θ_1 under the noninformative prior π given in Section 3 for several configurations (μ_1, μ_2) and (n_1, n_2). That is to say, the frequentist coverage of a $(1-\alpha)th$ posterior quantile should be close to $(1-\alpha)$. This is done numerically. Table 1 gives numerical values of the frequentist coverage probabilities of 0.05 (0.95) posterior quantiles for the our priors. The computation of these numerical values is based on the following algorithm for any fixed true (μ_1, μ_2) and any prespecified probability value α. Here α is 0.05 (0.95). Let $\theta_1^\pi(\alpha \mid \boldsymbol{X}, \boldsymbol{Y})$ be the posterior α-quantile of θ_1 given $\boldsymbol{X}$ and $\boldsymbol{Y}$. That is to say, $F(\theta_1^\pi(\alpha \mid \boldsymbol{X}, \boldsymbol{Y}) \mid \boldsymbol{X}, \boldsymbol{Y}) = \alpha$, where $F(\cdot \mid \boldsymbol{X}, \boldsymbol{Y})$ is the marginal posterior distribution of θ_1. Then the frequentist coverage probability of this one sided credible interval of θ_1 is

$$P_{(\mu_1,\mu_2)}(\alpha;\theta_1) = P_{(\mu_1,\mu_2)}(0 < \theta_1 \le \theta_1^{\pi}(\alpha \mid \boldsymbol{X}, \boldsymbol{Y})).$$

The estimated $P_{(\mu_1,\mu_2)}(\alpha;\theta_1)$ when $\alpha = 0.05$ (0.95) is shown in Table 1.

In particular, for fixed (μ_1, μ_2), we take 5,000 independent random samples of $\boldsymbol{X}$ and $\boldsymbol{Y}$ from the model (2). For the cases presented in Table 1, we see that the first order matching prior meets very well the target coverage probabilities for small and moderate values of n_1 and n_2. Note that the Jeffreys' prior does not satisfy the first order matching criterion but it meets the target coverage probabilities well.

Example. The following data, given by Proschan (1963), are time intervals of successive failures of the air conditioning equipment in Boeing 720 aircraft. For aircraft 1, the Bayes estimate of μ_1 under Jeffrey's prior is 69.95. And the Kolmogorov-Smirnov test statistic is 0.1143 and its p-value is 0.88. For aircraft 2, the Bayes estimate of μ_2 under Jeffreys' prior is 94.36. Also the Kolmogorov-Smirnov test statistic is 0.1791 and its p-value is 0.62. So we can assume that the time between successive failures for each plane is exponentially distributed.

Aircraft 1	50	44	102	72	22	39	3	15	197	188	79	88	46	5	5	36
	22	139	210	97	30	23	13	14								
Aircraft 2	102	209	14	57	54	32	67	59	134	152	27	14	230	66	61	34

Under the Jeffreys' prior and the matching prior, the Bayes estimates and the 95% Bayesian credible intervals of the θ_1 are −21.15 (−82.08, 27.38) and −21.54 (−81.84, 26.34), respectively. Bayes estimates under two priors have similar values and the length of the confidence interval under the matching prior is shorter than the Jeffreys' prior.

5. Conclusion

In the two exponential distributions, we have found a prior which is a first order matching prior for the difference of means. It turns out that the second order matching prior does not exist. And this first order matching prior possesses good frequentist properties in that the coverage probabilities of credible intervals for the difference of means based on this prior match their frequentist counterpart very closely even for small and moderate sample sizes. Also the Jeffreys' prior does not satisfy the first order matching criterion. From our simulation results and example, we recommend to use the first order matching prior for the Bayesian inference of the difference of two exponential means.

Table 1: Frequentist Coverage Probabilities of 0.05 (0.95) Posterior Quantiles for θ_1

μ_2	μ_1	n_1	n_2	π_J	π_m
1	0.1	5	5	0.053 (0.957)	0.053 (0.951)
		5	10	0.054 (0.961)	0.051 (0.951)
		10	10	0.055 (0.955)	0.053 (0.951)
		10	20	0.049 (0.956)	0.046 (0.952)
	0.5	5	5	0.056 (0.963)	0.065 (0.941)
		5	10	0.055 (0.960)	0.061 (0.937)
		10	10	0.055 (0.953)	0.058 (0.939)
		10	20	0.057 (0.953)	0.059 (0.942)
	3	5	5	0.034 (0.942)	0.047 (0.935)
		5	10	0.040 (0.951)	0.047 (0.949)
		10	10	0.046 (0.949)	0.055 (0.949)
		10	20	0.045 (0.049)	0.049 (0.950)
	5	5	5	0.040 (0.941)	0.052 (0.941)
		5	10	0.045 (0.950)	0.051 (0.950)
		10	10	0.042 (0.948)	0.048 (0.950)
		10	20	0.045 (0.946)	0.048 (0.947)
	10	5	5	0.045 (0.945)	0.049 (0.948)
		5	10	0.047 (0.950)	0.048 (0.951)
		10	10	0.049 (0.950)	0.054 (0.952)
		10	20	0.050 (0.948)	0.052 (0.949)
10	0.1	5	5	0.050 (0.952)	0.050 (0.952)
		5	10	0.052 (0.948)	0.052 (0.948)
		10	10	0.050 (0.958)	0.050 (0.958)
		10	20	0.045 (0.954)	0.045 (0.953)
	1	5	5	0.056 (0.956)	0.054 (0.950)
		5	10	0.056 (0.955)	0.053 (0.948)
		10	10	0.053 (0.953)	0.052 (0.949)
		10	20	0.055 (0.954)	0.051 (0.950)
	5	5	5	0.054 (0.957)	0.065 (0.939)
		5	10	0.055 (0.961)	0.061 (0.938)
		10	10	0.054 (0.961)	0.057 (0.950)
		10	20	0.055 (0.957)	0.056 (0.943)
	30	5	5	0.038 (0.944)	0.053 (0.938)
		5	10	0.043 (0.948)	0.051 (0.945)
		10	10	0.039 (0.948)	0.046 (0.948)
		10	20	0.042 (0.950)	0.046 (0.950)
	100	5	5	0.039 (0.945)	0.045 (0.947)
		5	10	0.048 (0.947)	0.052 (0.947)
		10	10	0.046 (0.951)	0.049 (0.953)
		10	20	0.048 (0.945)	0.049 (0.946)

References

1. Barlow, R.E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing*. Holt, Reinhart and Winston, New York.
2. Berger, J.O. and Bernardo, J.M. (1989). Estimating a Product of Means : Bayesian Analysis with Reference Priors. *Journal of the American Statistical Association*, 84, 200-207.
3. Berger, J.O. and Bernardo, J.M. (1992a). Reference Priors in a Variance Components Problem. *Bayesian Analysis in Statistics and Econometrics*, P. Goel and N.S. Iyengar(eds.). New York: Springer Verlag.
4. Berger, J.O. and Bernardo, J.M. (1992b). On the Development of Reference Priors (with discussion). *Bayesian Statistics IV*, J.M. Bernardo, et. al., Oxford University Press, Oxford, 35-60.

5. Bernardo, J.M. (1979). Reference Posterior Distributions for Bayesian Inference (with discussion). *Journal of Royal Statistical Society*, B, 41,
6. Cox, D.R. and Reid, N. (1987). Orthogonal Parameters and Approximate Conditional Inference (with discussion). *Journal of Royal Statistical Society*, B, 49, 1-39.
7. Datta, G.S. and Ghosh, J.K. (1995a). On Priors Providing Frequentist Validity for Bayesian Inference. *Biometrika*, 82, 37-45.
8. Datta, G.S. and Ghosh, M. (1995b). Some Remarks on Noninformative Priors. *Journal of the American Statistical Association*, 90, 1357-1363.
9. Datta, G.S. and Ghosh, M. (1996). On the Invariance of Noninformative Priors. *The Annal of Statistics*, 24, 141-159.
10. Davis, D.J. (1952). An Analysis of Some Failure Data. *Journal of the American Statistical Association*, 47, 113-150.
11. DiCiccio, T.J. and Stern, S.E. (1994). Frequentist and Bayesian Bartlett Correction of Test Statistics based on Adjusted Profile Likelihood. *Journal of Royal Statistical Society*, B, 56, 397-408.
12. Epstein, B. and Sobel, M. (1953). Life Testing. *Journal of the American Statistical Association*, 48, 486-502.
13. Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, Inc., Hoboken, New Jersey.
14. Masafumi Akahira (2002). Confidence Intervals for the Difference of Means: Application to the Behrens-Fisher Type Problem. *Statistical Papers*, 43, 273-284.
15. Mukerjee, R. and Dey, D.K. (1993). Frequentist validity of Posterior Quantiles in the Presence of a Nuisance Parameter : Higher Order Asymptotics. *Biometrika*, 80, 499-505.
16. Mukerjee, R. and Ghosh, M. (1997). Second Order Probability Matching Priors. *Biometrika*, 84, 970-975.
17. Proschan, F. (1963). Theoretical Explanation of Observed Decreasing Failure Rate. *Technometrics*, 5, 375-383.
18. Stein, C. (1985). On the Coverage Probability of Confidence Sets based on a Prior Distribution. *Sequential Methods in Statistics*, Banach Center Publications, 16, 485-514.
19. Tibshirani, R. (1989). Noninformative Priors for One Parameter of Many. *Biometrika*, 76, 604-608.
20. Welch, B.L. and Peers, H.W. (1963). On Formulae for Confidence Points based on Integrals of Weighted Likelihood. *Journal of Royal Statistical Society*, B, 25, 318-329.