Adv. & Appl. in Stat. **7**(3) (2007), 441-449

MULTIPLE COMPARISONS FOR FREUND'S BIVARIATE EXPONENTIAL POPULATIONS BASED ON FRACTIONAL BAYES FACTOR

JANG SIK CHO

Department of Informational Statistics
Kyungsung University
Busan, 608-736, Korea
e-mail: jscho@ks.ac.kr

Abstract

In this paper, we suggest a Bayesian multiple comparisons procedure for failure rates in K Freund's bivariate exponential populations based on a noninformative prior for the parameters. We compute fractional Bayes factor for all comparisons. Also, we calculate the posterior probabilities for all models and select the model with highest posterior probability as best model. Finally we give a numerical example to illustrate our procedure.

1. Introduction

Let us consider a life testing experiment in which two components are put on test. In many cases of life testing and reliability analysis, two components are assumed to have independent lifetime distributions. However, it is more realistic to assume some form of positive dependence among components. Freund [4], Marshall and Olkin [7], and Block and Basu [2] formulated a bivariate extension of the exponential model as a model for a system where the lifetimes of the two components may

2000 Mathematics Subject Classification: 62F15.

Keywords and phrases: Bayesian multiple comparison, bivariate exponential model, fractional Bayes factor, noninformative priors, multiple comparison, posterior probability.

Received May 10, 2007

depend on each other. In particular, many authors studied for Freund's bivariate exponential model. Kunchur and Munoli [6] obtained minimum variance unbiased estimator for the system reliability. Hanagal [5] suggested an estimator of system reliability from stress-strength relationship. Cho and Baek [3] derived a probability matching prior for Freund's bivariate exponential model.

In this paper, we focus on Bayesian multiple comparisons problem for K Freund's bivariate exponential populations. The test of equality of the failure rates more than two populations relies on likelihood ratio test statistic which is distributed as approximately χ^2-distribution. And classical tests only decide whether the null hypothesis will be rejected or not. When the null hypothesis is rejected, we do not know which hypothesis is best for describing the equality of parameters.

Bayesian approach to resolve the multiple comparisons problem selects the model with the highest posterior probability. And we can compute all the posterior probabilities of the hypotheses under consideration. In many cases, noninformative priors for the parameters are used. Since noninformative priors are typically improper, the priors are only up to arbitrary constants which affects the values of Bayes factors. Berger and Pericchi [1] and O'Hagan [8] introduced the intrinsic Bayes factor (IBF) and fractional Bayes factor (FBF), respectively, to remove the arbitrariness. These approaches have shown to be quite useful in several statistical areas.

In this paper, we propose a Bayesian multiple comparisons procedure for failure rates in K bivariate exponential models based on FBF. We compute the FBF for all comparisons and posterior probability for all models as best model. Finally, we give a numerical example to illustrate our procedure.

2. Preliminaries

Let us consider K Freund's bivariate exponential populations with parameters $\Theta = (\theta_1, \theta_2, \ldots, \theta_K)$. And let $M_1, \ldots, M_N$ be models under consideration. The random sample $(x, y) = ((x_1, y_1), \ldots, (x_n, y_n))$ have a likelihood function $L_i(\theta|x, y)$ under model M_i, $i = 1, \ldots, N$. Let $(x_i, y_i) = ((x_{i1}, y_{i1}), \ldots, (x_{in_i}, y_{in_i}))$ be an $n_i \times 1$ vector of independent observations on θ_i with density $f(x_{ij}, y_{ij} | \theta_i)$, $i = 1, \ldots, K$, $j = 1, \ldots, n_i$. The parameter vectors θ are unknown. Let $\pi_i(\theta)$ be a prior distribution of model M_i, and let p_i be the prior probabilities of model M_i. Then the posterior probability that the model M_i is true is given as

$$P(M_i | x, y) = \left(\sum_{j=1}^{N} \frac{p_j}{p_i} B_{ji} \right)^{-1},$$

where B_{ij} is the Bayes factor of model M_j to model M_i defined by

$$B_{ji} = \frac{m_j(x, y)}{m_i(x, y)} = \frac{\int_{\Theta_j} L_j(\theta | x, y) \pi_j(\theta) d\theta}{\int_{\Theta_i} L_i(\theta | x, y) \pi_i(\theta) d\theta}. \tag{1}$$

The computation of B_{ji} needs specification of the prior distribution $\pi_i(\theta)$ and $\pi_j(\theta)$. Usually, one can use the noninformative prior which is improper. Let π_i^N be the noninformative prior for model M_i. Then the use of improper prior $\pi_i^N(\cdot)$ in (1) causes the B_{ji} to contain arbitrary constants.

To solve this problem, O'Hagan [8] proposed the procedure for Bayesian model selection problem based on FBF as follows. B_{ij} based on noninformative prior $\pi_i^N(\cdot)$ is given as

$$B_{ji}^N = \frac{m_j^N(x, y)}{m_i^N(x, y)} = \frac{\int_{\Theta_j} L_j(\theta | x, y) \pi_j^N(\theta) d\theta}{\int_{\Theta_i} L_i(\theta | x, y) \pi_i^N(\theta) d\theta}. \tag{2}$$

Hence the FBF of model M_j versus model M_i is given as

$$B_{ji}^F = \frac{q_j(b, x, y)}{q_i(b, x, y)}, \tag{3}$$

where $q_i(b, x, y) = \dfrac{\int_{\Theta_i} L_i(\theta | x, y)\pi_i^N(\theta)d\theta}{\int_{\Theta_i} L_i^b(\theta | x, y)\pi_i^N(\theta)d\theta}$ and b specifies a fraction of the likelihood which is to be used as a prior density. One frequently suggested choice is $b = m/n$, where m is the size of the minimal training sample.

The multiple comparisons of K populations is to make inferences concerning relationships among the θ_i's based on (X, Y).

Let $\Omega = \{(\theta_1, \theta_2, \ldots, \theta_K) : \theta_i \in R, i = 1, 2, \ldots, K\}$ be the K-dimensional parameter space. Equality and inequality relationships among the θ_i's induce statistical hypotheses such that subsets of Θ, that is, $M_1 : \Omega_1 = \{\theta_i : \theta_1 = \theta_2 = \cdots = \theta_K\}$, $M_2 : \Omega_2 = \{\theta_i : \theta_1 \neq \theta_2 = \cdots = \theta_K\}$ and so on up to $M_N : \Omega_N = \{\theta_i : \theta_1 \neq \theta_2 \neq \cdots \neq \theta_K\}$. The hypotheses $M_r : \Omega_r$, $r = 1, 2, \ldots, N$, are disjoint, and $\bigcup_{r=1}^{N} \Omega_r = \Omega$.

The elements of Θ themselves with positive probability, will reduce to some $r \leq K$ distinct values. That is, the model can be classified $r(r = 1, \ldots, K)$ distinct groups. Let superscript $*$ be distinct values of the parameters and let $\theta_1^*, \ldots, \theta_r^*$ denote the set of distinct θ_i's. We need to define the configuration notation.

Definition (Configuration)**.** The set of indices $S = \{S_1, \ldots, S_K\}$ determines a classification of the data $\Theta = \{\theta_1, \ldots, \theta_K\}$ into r distinct groups or clusters; the n_j be number of observations in group j share the common parameter value θ_j^*. Now, we define K_j as the set of indices of observations in group j; that is, $K_j = \{i : S_i = j\}$. Let $|K_j| = \sum_{i \in K_j} n_i$ be total number of observations in group j.

There is a one to one correspondence between hypotheses and configurations. Therefore the Bayes factor for multiple comparisons can easily be computed by this configuration notation.

Suppose that a model is classified r distinct groups. Then the likelihood function is given by

$$L(\theta_1^*, \ldots, \theta_r^* \mid x, y) = \prod_{t=1}^{r} \prod_{\{i: i \in K_t\}} \prod_{j=1}^{n_i} f(x_{ij}, y_{ij} \mid \theta_t). \tag{4}$$

Since the noninformative prior for the model is $\pi_r^N(\theta_1^*, \ldots, \theta_r^*)$, the FBF is given by

$$q(b, x, y) = \frac{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L(\theta_1^*, \ldots, \theta_r^* \mid x, y) \cdot \pi_r^N(\theta_1^*, \ldots, \theta_r^*) d\theta_1^* \cdots d\theta_r^*}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L^b(\theta_1^*, \ldots, \theta_r^* \mid x, y) \cdot \pi_r^N(\theta_1^*, \ldots, \theta_r^*) d\theta_1^* \cdots d\theta_r^*}. \tag{5}$$

Thus if a model M_i is classified r_i distinct groups and a model M_j is classified r_j distinct groups, then the FBF of M_j versus M_i is given by $B_{ji}^F = \frac{q_j(b, x, y)}{q_i(b, x, y)}$, where

$$q_i(b, x, y) = \frac{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_i(\theta_1^*, \ldots, \theta_{r_i}^* \mid x, y) \cdot \pi_r^N(\theta_1^*, \ldots, \theta_{r_i}^*) d\theta_1^* \cdots d\theta_{r_i}^*}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_i^b(\theta_1^*, \ldots, \theta_{r_i}^* \mid x, y) \cdot \pi_r^N(\theta_1^*, \ldots, \theta_{r_i}^*) d\theta_1^* \cdots d\theta_{r_i}^*}.$$

3. Bayesian Multiple Comparisons

Let (X, Y) be random variables of Freund's bivariate exponential model with parameters $(\delta, \delta', \zeta, \zeta')$. Then the joint probability density function is given as

$$f(x, y : \delta, \delta', \zeta, \zeta') = \begin{cases} \delta\zeta' \exp[-\zeta' y - (\delta + \zeta - \zeta')x], & y > x > 0, \\ \delta'\zeta \exp[-\delta' x - (\delta + \zeta - \delta')y], & x > y > 0. \end{cases} \tag{6}$$

In this paper, we assume $\delta = \zeta (\equiv \theta)$, $\delta' = \zeta' (\equiv \eta)$ so that the lifetimes of two components are equal failure rates.

Suppose that a model M_k is classified r_k distinct groups. Then the noninformative prior for $((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*))$ is given by

$$\pi_k^N((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*)) \propto \frac{1}{(\theta_1^* \cdot \eta_1^*) \cdots (\theta_{r_k}^* \cdot \eta_{r_k}^*)},$$

$$0 < \theta_1^*, \ldots, \theta_{r_k}^*, \eta_1^*, \ldots, \eta_{r_k}^* < \infty. \tag{7}$$

And likelihood function is given by

$$L_k((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*) \mid x, y)$$

$$= \prod_{t=1}^{r_k} \left\{ (\theta_t^* \cdot \eta_t^*)^{|K_t|} \cdot \exp\left(-\eta_t^* \sum_{i \in K_t} \sum_{j=1}^{n_i} (y_{ij} \cdot I(A) + x_{ij} \cdot I(B)) \right) \right.$$

$$\left. \cdot \exp\left(-(2\theta_t^* \cdot \eta_t^*) \sum_{i \in K_t} \sum_{j=1}^{n_i} (x_{ij} \cdot I(A) + y_{ij} \cdot I(B)) \right) \right\}, \tag{8}$$

where $|K_t|$ is the number of the set of indices K_t, $I(A) = \{(i, j) \mid x_{ij} < y_{ij}\}$ and $I(B) = \{(i, j) \mid x_{ij} > y_{ij}\}$. Then the elements of FBF for model M_k are computed as follows:

$$\int_0^\infty \cdots \int_0^\infty L_k((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*)) \pi_k^N((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*)) d\theta_1^* \cdots d\eta_{r_k}^*$$

$$= c \cdot \prod_{t=1}^{r_k} \left[\frac{\Gamma(|K_t|)}{\left\{ \sum_{i \in K_t} \sum_{j=1}^{n_i} (x_{ij}(I(B) - I(A)) + y_{ij}(I(A) - I(B))) \right\}^{|K_t|}} \right]$$

$$\cdot \left[\frac{\Gamma(|K_t|)}{\left\{ 2 \sum_{i \in K_t} \sum_{j=1}^{n_i} (x_{ij} I(A) + y_{ij} I(B)) \right\}^{|K_t|}} \right] (\equiv S_{k1})$$

and

$$\int_0^\infty \cdots \int_0^\infty L_k^b((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*)|x, y)\pi_k^N((\theta_1^*, \eta_1^*), \ldots, (\theta_{r_k}^*, \eta_{r_k}^*))d\theta_1^* \cdots d\eta_{r_k}^*$$

$$= c \cdot \prod_{t=1}^{r_k} \left[\frac{\Gamma(b|K_t|)}{\left\{ b \sum_{i \in K_t} \sum_{j=1}^{n_i} (x_{ij}(I(B) - I(A)) + y_{ij}(I(A) - I(B))) \right\}^{b|K_t|}} \right]$$

$$\cdot \left[\frac{\Gamma(b|K_t|)}{\left\{ 2b \sum_{i \in K_t} \sum_{j=1}^{n_i} (x_{ij} I(A) + y_{ij} I(B)) \right\}^{b|K_t|}} \right] (\equiv S_{k2}).$$

Hence, $q(b, x, y)$ is given as $q_k(b, x, y) = \dfrac{S_{k1}}{S_{k2}}$.

Thus if a model M_i is classified r_i distinct groups and a model M_j is classified r_j distinct groups then the FBF of M_j versus M_i is given by

$$B_{ji}^F(x, y) = \frac{q_j(b, x, y)}{q_i(b, x, y)} = \frac{S_{j1}/S_{j2}}{S_{i1}/S_{i2}}. \tag{9}$$

Hence the FBF for all comparisons can be computed by equation (3). Using these FBF, we can calculate the posterior probability for hypothesis M_i, $i = 1, \ldots, K$ by (1). Thus, we can select the hypothesis with highest posterior probability in Bayesian multiple comparisons based on FBF.

4. A Numerical Example

A numerical example of the multiple comparisons for the failure rates in Freund's bivariate exponential populations is presented in this section using simulated data. We consider 4 bivariate exponential populations each with size $n_i = 15, i = 1, \ldots, 4$ and $(2.0, 2.5)$ for (θ_1, η_1) and (θ_2, η_2), $(3.0, 3.5)$ for (θ_3, η_3) and (θ_4, η_4), respectively. Then the numbers of possible hypotheses for multiple comparisons are 15. And we note that the true hypothesis may be $M_{\text{True}} : \theta_1 = \theta_2 \neq \theta_3 = \theta_4$. The simulated data are given by Table 1. Also the calculated posterior probabilities for all possible models are given by Table 2.

Table 1. The simulated data

populations	simulated data			
I	(0.6125, 0.3652)	(0.2797, 0.0347)	(2.0817, 0.3510)	(1.2561, 0.0734)
	(1.5567, 1.6094)	(1.3048, 0.1797)	(0.8251, 0.1291)	(0.0016, 0.0162)
	(0.2420, 0.2950)	(0.5699, 0.1534)	(0.3935, 0.1765)	(0.6988, 0.4840)
	(0.1087, 1.0277)	(2.0381, 0.0883)	(0.0748, 0.4610)	
II	(0.7826, 0.3742)	(0.0859, 0.2121)	(0.0795, 1.0365)	(0.0476, 0.6160)
	(0.0283, 0.4952)	(0.9406, 0.2076)	(0.0619, 0.7031)	(0.6819, 0.3618)
	(0.0606, 1.5147)	(0.3477, 0.6982)	(0.4719, 0.4616)	(0.6665, 1.0410)
	(0.0711, 0.1559)	(0.3404, 0.5933)	(0.7818, 0.1696)	
III	(0.2617, 0.0605)	(0.1521, 0.0761)	(0.0480, 0.0488)	(0.6542, 0.3424)
	(0.0328, 0.5424)	(0.6271, 0.1671)	(0.0788, 0.2292)	(0.5153, 0.7975)
	(0.7732, 0.1334)	(0.0391, 0.1262)	(0.3260, 1.1081)	(0.0199, 0.0224)
	(0.4709, 0.2846)	(0.1697, 0.3694)	(0.0459, 0.0915)	
IV	(0.2368, 0.0854)	(0.0626, 0.4331)	(0.0733, 0.0213)	(0.4786, 0.3526)
	(1.3102, 0.5169)	(0.2856, 0.0031)	(0.0528, 0.2206)	(0.0478, 0.0358)
	(0.1723, 0.0237)	(0.0885, 0.3408)	(0.1272, 0.3380)	(0.0470, 0.2693)
	(0.0484, 0.1448)	(0.0837, 0.1519)	(0.7934, 0.4022)	

Table 2. Calculated posterior probabilities for each model

M_r	$P(M_r \mid x, y)$	M_r	$P(M_r \mid x, y)$	M_r	$P(M_r \mid x, y)$
$\theta_1 = \theta_2 = \theta_3 = \theta_4$	0.0112	$\theta_1 = \theta_3 = \theta_4 \neq \theta_2$	0.0066	$\theta_1 \neq \theta_2 = \theta_3 = \theta_4$	0.0472
$\theta_1 = \theta_2 = \theta_3 \neq \theta_4$	0.0464	$\theta_1 = \theta_3 \neq \theta_2 = \theta_4$	0.0064	$\theta_1 \neq \theta_2 = \theta_3 \neq \theta_4$	0.0604
$\theta_1 = \theta_2 = \theta_4 \neq \theta_3$	0.0138	$\theta_1 = \theta_3 \neq \theta_2 \neq \theta_4$	0.0189	$\theta_1 \neq \theta_2 = \theta_4 \neq \theta_3$	0.0256
$\theta_1 = \theta_2 \neq \theta_3 = \theta_4$	0.3513	$\theta_1 = \theta_4 \neq \theta_2 = \theta_3$	0.0048	$\theta_1 \neq \theta_2 \neq \theta_3 = \theta_4$	0.1693
$\theta_1 = \theta_2 \neq \theta_3 \neq \theta_4$	0.1566	$\theta_1 = \theta_4 \neq \theta_2 \neq \theta_3$	0.0060	$\theta_1 \neq \theta_2 \neq \theta_3 \neq \theta_4$	0.0755

From Table 2, it is to be noted that the hypotheses $\theta_1 = \theta_2 \neq \theta_3 = \theta_4$, $\theta_1 \neq \theta_2 \neq \theta_3 = \theta_4$ and $\theta_1 = \theta_2 \neq \theta_3 \neq \theta_4$ have the large posterior probabilities 0.3513, 0.1693 and 0.1566, respectively. Thus the data lend greatest support to equalities for $\theta_1 = \theta_2$ and $\theta_3 = \theta_4$ being different from the others.

So far, the multiple comparisons procedure was carried out for K Freund's bivariate exponential populations based on FBF. Also, the method can be extended to a bivariate exponential populations with incomplete data or multivariate exponential populations as well, with moderate effort.

References

[1] J. O. Berger and L. R. Pericchi, The intrinsic Bayes factor for model selection and prediction, J. Amer. Statist. Assoc. 91 (1996), 109-122.

[2] H. W. Block and A. P. Basu, A continuous bivariate exponential extension, J. Amer. Statist. Assoc. 69 (1974), 1031-1037.

[3] J. S. Cho and S. U. Baek, Probability matching priors in Freund's bivariate exponential model, J. Korean Data Anal. Soc. 4(3) (2002), 263-268.

[4] J. E. Freund, A bivariate extension of the exponential distribution, J. Amer. Statist. Assoc. 56 (1961), 971-977.

[5] D. D. Hanagal, Estimation of system reliability from stress-strength relationship, Comm. Statist. Theory Methods 25(8) (1996), 1783-1797.

[6] S. H. Kunchur and S. B. Munoli, Estimation of reliability in Freund model for two component system, Comm. Statist. Theory Methods 23(11) (1994), 3273-3283.

[7] A. W. Marshall and I. Olkin, A multivariate exponential distribution, J. Amer. Statist. Assoc. 62 (1967), 30-44.

[8] A. O'Hagan, Fractional Bayes factors for model comparison (with discussion), J. Roy. Statist. Soc. 56 (1995), 99-118.

■