Journal of the Korean Data & Information Science Society
2010, **21**(3), 569–574

한국데이터정보과학회지

Bayesian multiple comparisons in Freund's bivariate exponential populations with type I censored data†

Jang Sik Cho[1]

[1]Basic Science Research Center, Kyungung University

Received 15 March 2010, revised 21 April 2010, accepted 28 April 2010

Abstract

We consider two components system which have Freund's bivariate exponential model. In this case, Bayesian multiple comparisons procedure for failure rates is suggested in K Freund's bivariate exponential populations. Here we assume that the components enter the study at random over time and the analysis is carried out at some prespecified time. We derive fractional Bayes factor for all comparisons under noninformative priors for the parameters and calculate the posterior probabilities for all hypotheses. And we select a hypotheses which has the highest posterior probability as best model. Finally, we give a numerical examples to illustrate our procedure.

Keywords: Bayesian multiple comparison, fractional Bayes factor, noninformative priors, posterior probability.

1. Introduction

Freund (1961), Marshall and Olkin (1967), Block and Basu (1974) and many authors formulated a bivariate extension of the exponential model as a model for a system where the lifetimes of the two components may depend on each other. In particular, the Freund's model has been generalized in the literature in various ways. Some of the generalizations are based on various functional representations of Freund's model obtained by replacing exponential random variables by other random variables. Let (X, Y) be random variables of a Freund's bivariate exponential model with parameters $\Theta = (\alpha, \alpha^{'}, \beta, \beta^{'})$. Then the joint probability density function is given as

$$f(x, y\ : \Theta) = \begin{cases} \alpha\beta^{'} \exp\left[-\beta^{'} y - (\alpha + \beta - \beta^{'})x\right], & y > x > 0, \\ \alpha^{'}\beta \exp\left[-\alpha^{'} x - (\alpha + \beta - \alpha^{'})y\right], & x > y > 0. \end{cases} \tag{1.1}$$

Hanagal (1996) suggested an estimator of system reliability from stress-strength relationship. Cho and Baek (2002) derived a probability matching priors and Cho (2007) suggested

† This research was supported by basic science research center research grants, Kyungsung University in 2009.

[1] Full time researcher, Basic Science Research Center, Kyungsung University, Busan 608-736, Korea. E-mail: jscho@ks.ac.kr

Bayesian test procedure for symmetry. Also Cho (2009) derived fractional Bayes factor for independent test procedure with type I censored data.

On the other hands, we focus on Bayesian multiple comparisons for K Freund's bivariate exponential populations with type I censored data. In Bayesian approach, the model with the highest posterior probability is selected as the best model. Hence we have to compute all the posterior probabilities of the hypotheses under consideration. In many cases, noninformative priors for the parameters are used. But the noninformative priors are typically improper which the priors are only up to arbitrary constants. Berger and Pericchi (1996) and O'Hagan (1995) introduced the intrinsic Bayes factor (IBF) and fractional Bayes factor (FBF), respectively to remove the arbitrariness. Cho *et al.* (2006) and Cho and Cho (2006) suggested Bayesian multiple comparisons procedure for some model using FBF.

In this paper, we suggest a Bayesian multiple comparisons procedure for failure rates in K bivariate exponential populations with type I censored data. And we compute the FBF's for all comparisons and posterior probability for all hypotheses. Also we select the best hypotheses which has the highest posterior probability. Finally, we give a numerical example to illustrate our procedure.

2. Preliminaries

Let $(\boldsymbol{x_i}, \boldsymbol{y_i}) = ((x_{i1}, y_{i1}), \cdots, (x_{\in i}, y_{\in i}))$ be a $n_i \times 1$ vector of independent observations i th population with density $f(x_{ij}, y_{ij}|\theta_i)$ and likelihood function $L_i\left(\theta_i|\boldsymbol{x}_i, \boldsymbol{y}_i\right), i = 1, \cdots, K,\ j = 1, \cdots, n_i$. And let $(\boldsymbol{x}, \boldsymbol{y}) = ((\boldsymbol{x_1}, \boldsymbol{y_1}), \cdots, (\boldsymbol{x_K}, \boldsymbol{y_K}))$. Then multiple comparisons of K populations is to make inferences concerning relationships among the θ_i's based on $(\boldsymbol{x}_i, \boldsymbol{y}_i)$.

Let $\Theta = \{(\theta_1, \theta_2, \cdots, \theta_K) :\ \theta_i \in R,\ i = 1, 2, \cdots, K\}$ be the K-dimensional parameter space. Equality and inequality relationships among the θ_i's induce statistical hypotheses H_i such that subsets of Θ, that is, $H_1 : \ \Theta_1 = \{\theta_i : \theta_1 = \theta_2 = \cdots = \theta_K\},\ H_2 : \Theta_2 = \{\theta_i : \theta_1 \neq \theta_2 = \cdots = \theta_K\}$ and so on up to $H_N : \ \Theta_N = \{\theta_i : \theta_1 \neq \theta_2 \neq \cdots \neq \theta_K\}$. The hypotheses $H_r : \Theta_r, r = 1, 2, \cdots, N$, are disjoint, and $\cup_{r=1}^{N} \Theta_r = \Theta$.

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

Definition (Configuration) The set of indices $S = \{S_1, \cdots, S_K\}$ determines a classification of the data $\Theta = \{\theta_1, \cdots, \theta_K\}$ 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, $K_j = \{i : S_i = j\}$. There is a one to one correspondence between hypotheses and configurations. Therefore the Bayes factor for multiple comparisons can easily compute by this configuration notation.

On the other hands, let $\pi_i(\theta_i)$ and p_i be a prior distribution and the prior probabilities of hypotheses H_i, respectively. Then the posterior probability that the hypotheses H_i is true is given as $P(H_i|\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{j=1}^{N} (p_j/p_i)\, B_{ji}\right)^{-1}$, where B_{ij} is the Bayes factor of hypotheses

H_j to hypotheses H_i defined by

$$B_{ji} = \frac{m_j(\boldsymbol{x}, \boldsymbol{y})}{m_i(\boldsymbol{x}, \boldsymbol{y})} = \frac{\int_{\Theta_j} L\left(\theta_j | \boldsymbol{x}, \boldsymbol{y}\right) \pi_j(\theta_j) d\theta_j}{\int_{\Theta_i} L\left(\theta_i | \boldsymbol{x}, \boldsymbol{y}\right) \pi_i(\theta_i) d\theta_i}. \tag{2.1}$$

The computation of B_{ji} needs specification of the prior distributions $\pi_i(\theta_i)$ and $\pi_j(\theta_j)$. In this paper, we use the noninformative prior which has improper distribution. Let π_i^N be the noninformative prior for hypotheses H_i. Then the use of improper priors $\pi_i^N(\cdot)$ in (2.1) causes the B_{ji} to contain arbitrary constants.

By O'Hagan (1995), the FBF of hypotheses H_j versus hypotheses H_i is given as

$$B_{ji}^F = \frac{q_j(b, \boldsymbol{x}, \boldsymbol{y})}{q_i(b, \boldsymbol{x}, \boldsymbol{y})}, \tag{2.2}$$

where $q_i(b, \boldsymbol{x}, \boldsymbol{y}) = \left(\int_{\Theta_i} L\left(\theta_i | \boldsymbol{x}, \boldsymbol{y}\right) \pi_i^N(\theta_i) d\theta_i\right) / \left(\int_{\Theta_i} L^b\left(\theta_i | \boldsymbol{x}, \boldsymbol{y}\right) \pi_i^N(\theta_i) d\theta_i\right)$ and b specifies a fraction of the likelihood which is to be used as a prior density.

Suppose that a hypotheses is classified r distinct groups. Then the likelihood function is given by $L\left(\theta_1^*, \cdots, \theta_r^* | \boldsymbol{x}, \boldsymbol{y}\right) = \prod_{t=1}^r \prod_{\{i \in K_t\}} \prod_{j=1}^{n_i} f(x_{ij}, y_{ij} | \theta_t)$. If we use noninformative prior for the hypotheses $\pi_r^N(\theta_1^*, \cdots, \theta_r^*)$, then the FBF is given by

$$q(b, \boldsymbol{x}, \boldsymbol{y}) = \frac{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L\left(\theta_1^*, \cdots, \theta_r^* | \boldsymbol{x}, \boldsymbol{y}\right) \cdot \pi_r^N\left(\theta_1^*, \cdots, \theta_r^*\right) d\theta_1^* \cdots d\theta_r^*}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L^b\left(\theta_1^*, \cdots, \theta_r^* | \boldsymbol{x}, \boldsymbol{y}\right) \cdot \pi_r^N\left(\theta_1^*, \cdots, \theta_r^*\right) d\theta_1^* \cdots d\theta_r^*}. \tag{2.3}$$

Thus if a hypotheses H_i is classified r_i distinct groups and a hypotheses H_j is classified r_j distinct groups then the FBF of H_j versus H_i is given by $B_{ji}^F = q_j(b, \boldsymbol{x}, \boldsymbol{y}) / q_i(b, \boldsymbol{x}, \boldsymbol{y})$, where

$$q_i(b, \boldsymbol{x}, \boldsymbol{y}) = \frac{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L\left(\theta_1^*, \cdots, \theta_{r_i}^* | \boldsymbol{x}, \boldsymbol{y}\right) \cdot \pi_r^N\left(\theta_1^*, \cdots, \theta_{r_i}^*\right) d\theta_1^* \cdots d\theta_{r_i}^*}{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L^b\left(\theta_1^*, \cdots, \theta_{r_i}^* | \boldsymbol{x}, \boldsymbol{y}\right) \cdot \pi_r^N\left(\theta_1^*, \cdots, \theta_{r_i}^*\right) d\theta_1^* \cdots d\theta_{r_i}^*}.$$

3. Bayesian multiple comparisons

We let t_i be type I censoring time for i th observation which is fixed constant. For $j = 1, 2;\ k = 1, 2, 3;\ i = 1, 2, \cdots, n$, and we let $G_{1i} = I(X_i > t_i)$, $G_{2i} = I(Y_i > t_i)$, $G_{ji}^o = 1 - G_{ji},\ j = 1, 2,\ R_i = I(X_i < Y_i),\ R_i^o = 1 - R_i$. Then i th observed lifetime (x_i, y_i) is observed as follows;

(1) $(x_i, y_i) = (x_i, y_i)$, if $x_i < t_i,\ y_i < t_i$ (2) $(x_i, y_i) = (t_i, y_i)$, if $x_i > t_i,\ y_i < t_i$
(3) $(x_i, y_i) = (x_i, t_i)$, if $x_i < t_i,\ y_i > t_i$ (4) $(x_i, y_i) = (t_i, t_i)$, if $x_i > t_i,\ y_i > t_i$

In this paper, we assume $\alpha = \beta (\equiv \psi)$, $\alpha' = \beta' (\equiv \eta)$ so that the lifetimes of two components are equal failure rates. And suppose that the k th hypotheses H_k classified r_k distinct groups,

$k = 1, \cdots, K$.

$$L\left((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)|(\boldsymbol{x}, \boldsymbol{y})\right) = \prod_{t=1}^{r_k} \left\{ {\psi_t^*}^{n_{1t}+n_{2t}+n_{4t}+n_{5t}} \right\} \cdot {\eta_t^*}^{n_{1t}+n_{2t}}$$
$$\cdot \exp\left[-2\psi_t^* \left(\sum_{i\in S_{14};j\in K_t} x_{ij} + \sum_{i\in S_{25};j\in K_t} y_{ij} + \sum_{i\in S_6;j\in K_t} (x_{ij}+y_{ij}) \right)\right] \tag{3.1}$$
$$\cdot \exp\left[-\eta_t^* \left(\sum_{i\in S_{25};j\in K_t} (x_{ij}-y_{ij}) + \sum_{i\in S_{14};j\in K_t} (y_{ij}-x_{ij}) \right)\right],$$

where K_t is the set of indices of t th group in the k th hypotheses. $n_{1t} = \sum_{i\in K_t} R_i G_{1i}^o G_{2i}^o$, $n_{2t} = \sum_{i\in K_t} R_i^o G_{1i}^o G_{2i}^o$, $n_{4t} = \sum_{i\in K_t} R_i G_{1i}^o G_{2i}$, $n_{5t} = \sum_{i\in K_t} R_i^o G_{1i} G_{2i}^o$, $t = 1, \cdots, r_k$, $k = 1, \cdots, K$.

On the other hand, we assume that the noninformative prior for $\left((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)\right)$ is given by

$$\pi_k^N\left((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)\right) \propto \frac{1}{(\psi_1^* \cdot \eta_1^*) \cdots (\psi_{r_k}^* \cdot \eta_{r_k}^*)},$$
$$0 < \psi_1^*, \cdots, \psi_{r_k}^*, \eta_1^*, \cdots, \eta_{r_k}^* < \infty. \tag{3.2}$$

Then the elements of FBF for hypotheses H_k is computed as follows;

$$\int_0^\infty \cdots \int_0^\infty L_k\left((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)|(\boldsymbol{x}, \boldsymbol{y})\right) \pi_k^N\left((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)\right) d\psi_1^* \cdots d\eta_{r_k}^*$$
$$= c \cdot \prod_{t=1}^{r_k} \left[\frac{\Gamma(n_{1t}+n_{2t})}{\left\{ \sum_{i\in S_{25};j\in K_t} (x_{ij}-y_{ij}) + \sum_{i\in S_{14};j\in K_t} (y_{ij}-x_{ij}) \right\}^{n_{1t}+n_{2t}}} \right]$$
$$\cdot \left[\frac{\Gamma(n_{1t}+n_{2t}+n_{4t}+n_{5t})}{\left\{ 2\left(\sum_{i\in S_{14};j\in K_t} x_{ij} + \sum_{i\in S_{25};j\in K_t} y_{ij} + \sum_{i\in S_6;j\in K_t} (x_{ij}+y_{ij}) \right) \right\}^{n_{1t}+n_{2t}+n_{4t}+n_{5t}}} \right] (\equiv S_{k1})$$

and

$$\int_0^\infty \cdots \int_0^\infty L_k^b\left((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)|(\boldsymbol{x}, \boldsymbol{y})\right) \pi_k^N((\psi_1^*, \eta_1^*), \cdots, (\psi_{r_k}^*, \eta_{r_k}^*)) d\psi_1^* \cdots d\eta_{r_k}^*$$
$$= c \cdot \prod_{t=1}^{r_k} \left[\frac{\Gamma(b(n_{1t}+n_{2t}))}{\left\{ b\left(\sum_{i\in S_{25};j\in K_t} (x_{ij}-y_{ij}) + \sum_{i\in S_{14};j\in K_t} (y_{ij}-x_{ij}) \right) \right\}^{b(n_{1t}+n_{2t})}} \right]$$
$$\cdot \left[\frac{\Gamma(b(n_{1t}+n_{2t}+n_{4t}+n_{5t}))}{\left\{ 2b\left(\sum_{i\in S_{14};j\in K_t} x_{ij} + \sum_{i\in S_{25};j\in K_t} y_{ij} + \sum_{i\in S_6;j\in K_t} (x_{ij}+y_{ij}) \right) \right\}^{b(n_{1t}+n_{2t}+n_{4t}+n_{5t})}} \right] (\equiv S_{k2}).$$

Here c is a constant which is independent of ψ_i^* and η_i^*, $i = 1, \cdots, r_k$. Hence $q_k(b, \boldsymbol{x}, \boldsymbol{y}) = S_{k1}/S_{k2}$. If a hypotheses H_i is classified r_i distinct groups and a hypotheses H_j is classified r_j distinct groups then the FBF of H_j versus H_i is given by

$$B_{ji}^F = \frac{q_j(b, \boldsymbol{x}, \boldsymbol{y})}{q_i(b, \boldsymbol{x}, \boldsymbol{y})} = \frac{S_{j1}/S_{j2}}{S_{i1}/S_{i2}}. \tag{3.3}$$

Hence the FBF for all comparisons can be computed. Using these FBF, we can calculate the posterior probability for hypothesis $H_i,\ i = 1, \cdots, K$. 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 Freund's bivariate exponential models each with size $n_i = 15,\ i = 1, \cdots, 4$ and $(2.0, 2.5)$ for (ψ_1, η_1) and (ψ_2, η_2), $(3.0, 3.5)$ for (ψ_3, η_3) and (ψ_4, η_4), respectively. Then the numbers of possible hypotheses for multiple comparisons are 15. And we note that the true hypothesis may be $H_{true} : \xi_1 = \xi_2 \neq \xi_3 = \xi_4$, where $\xi_i = (\psi_i, \eta_i),\ i = 1, \cdots, 4$. The simulated data are given by table 4.1. Where * means censored data.

Table 4.1 The simulated data

K	simulated data
1	(.5816, 1.2363*), (.3502, .4268), (.4255, .0565), (.0920, .3000), (1.4248*, .0710), (.2170, .0522), (.1020, 1.0493*), (1.7308*, .4147), (1.3584*, .0860), (1.1769*, .3275), (.6405, .1859), (.9745*, .5463), (.1777, .2779), (.6424, 1.0402*), (.1025, .3591)
2	(.7250, .1051), (1.4029*, 1.4571*), (.6764, .0546), (.0258, .0509), (.2930, 1.6875*), (.4207, .7319), (.6826, .0365), (.3081, .0469), (.0517, .3927), (1.2268*, .4150), (1.8848*, 1.6212*), (.8865, .1947), (.3597, .0249), (.4326, 1.2948*), (.2384, .3544)
3	(.5242, .4362), (.2365, .1277), (.0954, .0418), (.3916, .2962), (.3299, .0977), (.0868, .9382*), (.1293, .0641), (.6393, .2353), (.4079, .0922), (.0513, .3902), (.0460, .3064), (.6475, .2611), (.2921, .3063), (.2005, .1349), (.7147, .6151)
4	(.1820, .2271), (.9506*, .0673), (.0875, .3407), (.1657, .0446), (.2119, .2696), (.0637, .0858), (.1034, .3916), (.0141, .1058), (.5176, .1923), (.3092, .2285), (.2002, .5951), (.1679, .1160), (.1518,1.3960*), (.1592, .1992), (.2836, .2827)

Also calculated posterior probabilities for all possible hypotheses is given by table 4.2.

Table 4.2 Calculated posterior probabilities for each hypotheses

$\mathbf{H_r}$	$\mathbf{P(H_r\|\boldsymbol{x},\boldsymbol{y})}$	$\mathbf{H_r}$	$\mathbf{P(H_r\|\boldsymbol{x},\boldsymbol{y})}$	$\mathbf{H_r}$	$\mathbf{P(H_r\|\boldsymbol{x},\boldsymbol{y})}$
$\xi_1 = \xi_2 = \xi_3 = \xi_4$	.0336	$\xi_1 = \xi_3 = \xi_4 \neq \xi_2$	.0603	$\xi_1 \neq \xi_2 = \xi_3 = \xi_4$	.0198
$\xi_1 = \xi_2 = \xi_3 \neq \xi_4$	.0711	$\xi_1 = \xi_3 \neq \xi_2 = \xi_4$	.0146	$\xi_1 \neq \xi_2 = \xi_3 \neq \xi_4$	.0298
$\xi_1 = \xi_2 = \xi_4 \neq \xi_3$	.0320	$\xi_1 = \xi_3 \neq \xi_2 \neq \xi_4$	.0597	$\xi_1 \neq \xi_2 = \xi_4 \neq \xi_3$	.0148
$\xi_1 = \xi_2 \neq \xi_3 = \xi_4$	.2892	$\xi_1 = \xi_4 \neq \xi_2 = \xi_3$	.0164	$\xi_1 \neq \xi_2 \neq \xi_3 = \xi_4$	.1246
$\xi_1 = \xi_2 \neq \xi_3 \neq \xi_4$	.1405	$\xi_1 = \xi_4 \neq \xi_2 \neq \xi_3$	.0332	$\xi_1 \neq \xi_2 \neq \xi_3 \neq \xi_4$	.0605

From table 4.2, it is to be noted that the hypotheses $\xi_1 = \xi_2 \neq \xi_3 = \xi_4$, $\xi_1 \neq \xi_2 \neq \xi_3 = \xi_4$ and $\xi_1 = \xi_2 \neq \xi_3 \neq \xi_4$ have the large posterior probabilities 0.3513, 0.1693 and 0.1566, respectively. Thus the data lend greatest support to equalities for $\xi_1 = \xi_2$ and $\psi_3 = \psi_4$ being different from the others.

So far, the multiple comparisons procedure was carried out for K Freund's bivariate exponential populations with type I censored data 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

Berger, J. O. and Pericchi, L. R. (1996). The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, **91**, 109-122.

Block, H. W. and Basu, A. P. (1974). A continuous bivariate exponential extension. *Journal of the American Statistical Association*, **69**, 1031-1037.

Cho, J. S. and Baek, S. U. (2002). Probability matching priors in Freund's bivariate exponential model. *Journal of the Korean Data Analysis Society*, **4**, 263-268.

Cho, J. S., Cho, K. H. and Choi, S. B. (2006). Bayesian multiple comparison of bivariate exponential populations based on fractional Bayes factor. *Journal of the Korean Data & Information Science Society*, **17**, 843-850.

Cho, J. S. and Cho, K. H. (2006). Bayesian multiple comparisons for the ratio of the failure rates in two components system. *Journal of the Korean Data & Information Science Society*, **17**, 647-655.

Cho, J. S. (2007). Fractional Bayes factor approach for homogeneity testing in a bivariate exponential model. *Far East Journal of Theoretical Statistics*, **22**, 217-224.

Cho, J. S. (2009). Bayesian testing procedure for independence in Freund's bivariate exponential model with type I censored data: Fractional Bayes factor approach. *Advances and Applications in Mathematical Sciences*, **2**, 149-157.

Freund, J. E. (1961). A bivariate extension of the exponential distribution. *Journal of American Statistical Association*, **56**, 971-977.

Hanagal, D. D. (1996). Estimation of system reliability from stress-strength relationship. *Communication in Statistics-Theory and Methods*, **25**, 1783-1797.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, **62**, 30-44.

O' Hagan, A. (1995). Fractional Bayes factors for model comparison (with discussion). *Journal of Royal Statistical Society*, **56**, 99-118.