Journal of Korean
Data & Information Science Society
2007, Vol. 18, No. 2, pp. 553 ~ 560

Multiple Comparisons for a Bivariate Exponential Populations Based On Dirichlet Process Priors[1)]

Jang Sik Cho[2)]

Abstract

In this paper, we consider two components system which lifetimes have Freund's bivariate exponential model with equal failure rates. We propose Bayesian multiple comparisons procedure for the failure rates of I Freund's bivariate exponential populations based on Dirichlet process priors(DPP). The family of DPP is applied in the form of baseline prior and likelihood combination to provide the comparisons. Computation of the posterior probabilities of all possible hypotheses are carried out through Markov Chain Monte Carlo(MCMC) method, namely, Gibbs sampling, due to the intractability of analytic evaluation. The whole process of multiple comparisons problem for the failure rates of bivariate exponential populations is illustrated through a numerical example.

Keywords : Bivariate Exponential Population, Dirichlet Process Prior, Gibbs Sampler; Mixture of Dirichlet Processes, Multiple Comparison; Nonparametric Bayes

1. Introduction

In reliability studies of mechanical components, dependence between two components occurs quite often. A system, which functions as long as at least one of the two identical components functions, has a functional correlation between the system components. This dependence among components arises from common environmental stresses and shocks. Freund(1961) 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. For this model, many researches are

---

1) This Research was supported by Kyungsung University Research Grants in 2006.

2) Associate Professor, Department of Informational Statistics, Kyungsung University, Busan, 608-736, Korea.
E-Mail : jscho@ks.ac.kr

studied by authors. Minimum variance unbiased estimator for the system reliability was obtained by Kunchur and Munoli(1994). A estimator of system reliability from stress-strength relationship was derived by Hanagal(1996). Also statistical hypothesis testing procedures for independence and symmetry from a frequentist viewpoint were proposed by Hanagal and Kale(1992). A probability matching priors in Freund's bivariate exponential model was derived by Cho and Baek(2002).

On the other side, the multiple comparison problem(MCP) for I Freund's bivariate exponential populations with equal failure rates $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_I)$ can be viewed as making inferences concerning relationships among the θ's based on observations. This is tantamount to testing the following hypothesis,

$$H_o:\theta_1=\cdots=\theta_I \quad v.s. \quad H_1: \text{ not } H_o.$$

For Freund's bivariate exponential populations, the frequentist approach is not very straightforward. This is partly due to the difficulty in handling the distributional aspects and associated computations. The multiple comparison problem using nonparametric priors in a Bayesian inferential setup was studied by Gopalan and Berry (1998).

In this paper, we propose a Bayesian multiple comparisons procedure based on DPP for the failure rates in I Freund's bivariate exponential populations. The MCMC techniques, in particular, Gibbs sampling is adopted here to evaluate the posterior probabilities of the hypotheses. Reviews on the DPP are presented in Section 2, while Section 3 presents the calculation of posterior probabilities for the hypotheses in MCP. A numerical example illustrating the procedure is presented in Section 4.

2. Preliminaries

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

$$f(x,y:\delta,\delta',\zeta,\zeta')=\begin{cases}\delta\zeta'\exp[-\zeta' y-(\delta+\zeta-\zeta')x], & y>x>0,\\ \delta'\zeta\exp[-\delta' x-(\delta+\zeta-\delta')y], & x>y>0.\end{cases} \tag{1}$$

In this paper, we assume $\delta=\zeta(\equiv\theta)$, $\delta'=\zeta'(\equiv\eta)$ so that the lifetimes of two components are equal failure rates. We assume that $(\underline{\boldsymbol{x}},\underline{\boldsymbol{y}})=(\boldsymbol{x}_1,\boldsymbol{y}_1),\cdots,(\boldsymbol{x}_I,\boldsymbol{y}_I)$ be a set of observations available on I populations, where $(\boldsymbol{x}_i,\boldsymbol{y}_i)=(x_{i1},y_{i1}),\cdots,(x_{In_i},y_{In_i})$ is an $n_i\times 1$ vector of conditionally independent observations on population i, $i=1,\cdots,I$; $j=1,\cdots,n_i$ and $\sum_{i=1}^{I}n_i=n$.

Now a distribution function $G_o(\cdot)$ and a positive scalar precision parameter α together determine the DPP G. Here $G_o(\cdot)$ that defines the location of the DPP, is sometimes called prior "guess" or baseline prior. The precision parameter α determines the concentration of the prior for G around the prior guess G_o, and therefore measures the strength of belief in G_o. The DPP is usually denoted by $G \sim D(G \mid G_o, \alpha)$.

We assume that the θ_i's come from G, and that $G \sim D(G \mid G_o, \alpha)$ as stated above. This structure results in a posterior distribution which is a mixture of Dirichlet processes (Antoniak, 1974). Now following the Polya urn representation of the Dirichlet process (Blackwell and MacQueen, 1973), the joint posterior distribution can be written as,

$$\theta_i \mid \underline{x}, \underline{y} \propto \prod_{i=1}^{I} f(x_i, y_i \mid \theta_i) \times \frac{\alpha G_o(\theta_i) + \sum_{k<i} \delta(\theta_i \mid \theta_k)}{\alpha + i - 1} \tag{2}$$

where $\delta(\theta_i \mid \theta_k)$ is the distribution putting a point mass on θ_k. For each $i = 1, \cdots, I$, the conditional posterior distribution of θ_i is given by,

$$\theta_i \mid \theta_k, k \neq i, \underline{x}, \underline{y} \propto q_o G_b(\theta_i \mid x_i, y_i) + \sum_{k \neq i} q_k \delta(\theta_i \mid \theta_k), \tag{3}$$

where $G_b(\theta_i \mid x_i, y_i)$ is the baseline posterior distribution, $q_o \propto \alpha \int f(x_i, y_i \mid \theta_i) dG_o(\theta_i)$, $q_k \propto f(x_i, y_i \mid \theta_k)$, and $1 = q_o + \sum_{k \neq i} q_k$.

Let $\boldsymbol{\Theta} = \{\boldsymbol{\theta} = (\theta_1, \cdots, \theta_I) : \theta_i \in R, i = 1, \cdots, I\}$ be the I-dimensional parameter space. Equality and inequality relationships among θ's induce statistical hypotheses that are subsets of $\boldsymbol{\Theta}$. Thus the MCP becomes testing the following hypotheses.

$$H_o : \boldsymbol{\theta_o} = \{\theta_i : \theta_1 = \theta_2 = \cdots = \theta_I\},\ H_1 : \boldsymbol{\theta_1} = \{\theta_i : \theta_1 \neq \theta_2 = \theta_3 = \cdots = \theta_I\},\ \cdots,$$
$$H_N : \boldsymbol{\theta_N} = \{\theta_i : \theta_1 \neq \theta_2 \neq \cdots \neq \theta_I\}. \tag{4}$$

The hypotheses $H_r : \boldsymbol{\theta_r}, r = 0, 1, 2, \cdots, N$ are disjoint, and $\cup_{r=0}^{N} \boldsymbol{\theta_r} = \boldsymbol{\Theta}$. The elements of $\boldsymbol{\Theta}$ themselves behave as described by (4) and so with positive probability, they will reduce to some $p < I$ distinct values. Let superscript $*$ denotes distinct values of the parameters. Then any realization of I parameters θ_i generated from G lies in a set of $p < I$ distinct values, denoted by $\boldsymbol{\theta}^* = (\theta_1^*, \cdots, \theta_p^*)$. The computation of posterior probabilities for different hypotheses through Gibbs algorithm becomes manageable using the notion of Configuration as termed by Gopalan and Berry(1998). Their definition of Configuration is restated here,

Definition(*Configuration*): The set of indices $S=\{S_1,\cdots,S_I\}$ determines a classification of the data $\Theta=(\theta_1,\cdots,\theta_I)$ into I^* distinct groups or clusters; the $n_l=\#\{S_i=l\}$ observations in group l share the common parameter value θ_l^*. Now, define I_l^* as the set of indices of observations in group l. That is, $I_l^*=\{i:S_i=l\}$. Let $(X,Y)_{(l)}=\{(X_i,Y_i):S_i=l\}$ be the corresponding group of $n_{I_l^*}=\sum_{i\in I_l^*}n_i$ observations.

Thus a one-to-one correspondence between hypotheses and configurations follows. And the required computations are reduced by the fact that the distinct θ_i's are typically reduced to fewer than I due to the clustering of the θ_i's inherent in the Dirichlet process. Hence, (3) can be rewritten as:

$$\theta_i \mid \theta_k, k\neq i, \underline{\boldsymbol{x},\boldsymbol{y}} \propto q_o G_b(\theta_i \mid \boldsymbol{x_i},\boldsymbol{y_i}) + \sum_{i\neq k} n_k q_k^* \delta(\theta_i \mid \theta_k^*), \tag{5}$$

with $q_k^* \propto f(\boldsymbol{x_i},\boldsymbol{y_i} \mid \theta_k^*)$, and $1=q_0+\sum_{k\neq i} n_k q_k^*$. In addition to the simplification of notations, the cluster structure of the θ_i also improves the efficiency of the algorithm.

3. Posterior Sampling In Dirichlet Process Mixtures

A gamma distribution with parameters (α_{oi},β_{oi}) is considered as baseline prior G_o. This implies that $\theta_1,\cdots,\theta_I$ are i.i.d. from G_o. Then a hierarchical set up for the Dirichlet process analysis as outlined above becomes,

$$\boldsymbol{x_i},\boldsymbol{y_i} \mid \theta_i,\eta_i \sim BVE(\boldsymbol{x_i},\boldsymbol{y_i} \mid \eta_i,\theta_i), \tag{6}$$

$$\theta_i \mid G \sim G(\theta_i), \tag{7}$$

$$G \mid G_o,\alpha \sim D(G \mid G_o,\alpha), \tag{8}$$

$$G_o \mid \alpha_{oi},\beta_{oi} \sim Gam(\alpha_{oi},\beta_{oi}), \tag{9}$$

$$\eta_i \mid \alpha_{1i},\beta_{1i} \sim Gam(\alpha_{1i},\beta_{1i}). \tag{10}$$

Here, BVE and Gam stand for Freund's bivariate exponential and gamma distributions, respectively. Also we consider a gamma prior for α with a shape parameter a and scale parameter b, that is, $\alpha \sim Gam(a,b)$. Thus the $Gam(a,b)$ becomes the reference prior if $a\to 0$ and $b\to 0$. And we have access to a neat data augmentation device for sampling α by Escobar and West(1995).

The configuration notation is more convenient to use in describing the Gibbs sampling algorithm as the full conditionals can be written in closed form as under:

$$\theta_i \mid \underline{x}, \underline{y}, \theta_k, k \neq i, \alpha \sim q_o Gam\left(n_i + \alpha_{oi}, 2\left(\sum_{i,j \in D_1} x_{ij} + \sum_{i,j \in D_2} y_{ij}\right) + \beta_{oi}\right) + \sum_{k \neq i} q_k \delta(d\theta_i \mid \theta_k), \quad (11)$$

$$\eta_i \mid \underline{x}, \underline{y}, \theta_i, \alpha \sim Gam\left(n_i + \alpha_{1i}, \sum_{i,j \in D_1} (y_{ij} - x_{ij}) + \sum_{i,j \in D_2} (x_{ij} - y_{ij}) + \beta_{1i}\right), \quad (12)$$

$$\theta_l^* \mid \underline{x}, \underline{y}, S \sim Gam\left(\sum_{i \in I_l^*} n_i + \alpha_{oi}^*, \ 2\sum_{i \in I_l^*}\left(\sum_{i,j \in D_1} x_{ij} + \sum_{i,j \in D_2} y_{ij}\right) + \beta_{oj}^*\right), \quad (13)$$

$$\alpha \mid \zeta, I^* \sim \pi_\zeta Gam(\alpha + I^*, \ b - \log(\zeta)) + (1 - \pi_\zeta) Gam(\alpha + I^* - 1, \ b - \log(\zeta)), \quad (14)$$

$$\zeta \mid \alpha, I^* \sim Beta(\alpha + 1, I^*), \quad (15)$$

where $D_1 = \{(x_{ij}, y_{ij}) \mid x_{ij} < y_{ij}, \ i = 1, \cdots, I, j = 1, \cdots, n_i\}$,
$D_2 = \{(x_{ij}, y_{ij}) \mid x_{ij} > y_{ij}, \ i = 1, \cdots, I, j = 1, \cdots, n_i\}$,

$$q_o = \alpha \frac{\beta_{oi}^{\alpha_{oi}}}{\Gamma(\alpha_{oi})} \frac{\beta_{1i}^{\alpha_{1i}}}{\Gamma(\alpha_{1i})} \frac{\Gamma(n_i + \alpha_{oi})}{\left[2\left(\sum_{i,j \in D_1} x_{ij} + \sum_{i,j \in D_2} y_{ij}\right) + \beta_{oi}\right]^{n_i + \alpha_{oi}}} \times \frac{\Gamma(n_i + \alpha_{1i})}{\left[\sum_{i,j \in D_1} (y_{ij} - x_{ij}) + \sum_{i,j \in D_2} (x_{ij} - y_{ij}) + \beta_{1i}\right]^{n_i + \alpha_{1i}}}$$

and

$$q_k \propto \theta_k^{n_k} \eta_k^{n_k} \exp\left(-\eta_k\left(\sum_{i,j \in D_1} (y_{ij} - x_{ij}) - \sum_{i,j \in D_2} (x_{ij} - y_{ij})\right)\right) \cdot \exp\left(-2\theta_k\left(\sum_{i \in D_1} \sum_{j=1}^{n_i} x_{ij} + \sum_{i \in D_2} \sum_{j=1}^{n_i} y_{ij}\right)\right).$$

Gibbs sampling proceeds by simply iterating through (11) – (15) in order, sampling at each stage based on the current values of all the conditioning variables.

The configuration induces the equality and inequality relationships among the θ's, that corresponds to the partitions on the parameter space $\boldsymbol{\Theta}$ and in turn to the hypotheses of interest. In order to estimate the posterior probability of a hypothesis H_r from a large number (L) of sample draws, we take

$$P(H_r \mid \underline{x}, \underline{y}) \approx \frac{1}{L} \sum_{l=1}^{L} \delta_{S_l}(H_r) \quad , \quad (16)$$

where $\delta_{S_l}(H_r)$ denotes unit point mass for the case where lth draw of S, S_l corresponds to H_r.

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 $n_i = 10$, $i = 1,\cdots,4$ 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 $H_{True} : \theta_1 = \theta_2 \neq \theta_3 = \theta_4$. The simulated data are given as follows.

Table 1 The simulated data for each populations

populations	simulated data
1	(0.1720, 0.3284), (0.1735, 0.2636), (3.0748, 0.8220), (0.1875, 0.8710), (1.0869, 0.5975), (0.6868, 0.7684), (0.6802, 0.5826), (0.6041, 0.5079), (0.7245, 1.6152), (0.0135, 0.7087)
2	(0.6259, 0.3833), (0.1061, 0.3369), (0.1098, 0.5392), (0.4286, 0.2595), (0.0794, 0.2518), (0.2511, 0.0414), (0.4345, 0.7155), (1.8190, 1.3002), (0.2433, 0.1433), (0.9515, 0.9345)
3	(0.2415, 0.1285), (0.1644, 0.4884), (0.2619, 1.3812), (0.1134, 0.5902), (0.0161, 0.2061), (0.3743, 0.0263), (0.1445, 0.1614), (0.1757, 0.1729), (0.6590, 0.0129), (0.9584, 0.3516)
4	(0.0108, 0.1512), (0.3300, 0.1720), (0.4603, 0.0485), (0.5986, 0.4157), (0.1134, 0.0358), (0.7188, 0.2094), (0.6368, 0.7092), (0.0462, 0.4763), (0.1965, 0.0376), (0.0757, 0.3016)

For the precision parameter α, we consider Gamma priors with parameters (a,b) = (0.01, 0.01) in order to have equal mean 1 and variance 100 that the prior be fairly noninformative. We also set a priori that each $\theta_i, i = 1,\cdots,4$ follows a gamma distribution with parameters $\alpha_{oi} = \alpha_{1i} = 2$ and $\beta_{oi} = \beta_{1i} = 0.001$ to reflect vagueness of the prior knowledge.

The posterior probabilities for all possible hypotheses are approximated by the Gibbs sampling algorithm using 20,000 iterations with 10,000 burn out and 5 replications and are presented in Table 2.

Table 2 Calculated posterior probabilities for each hypothesis

H_r	$P(H_r \mid \underline{x}, \underline{y})$	H_r	$P(H_r \mid \underline{x}, \underline{y})$	H_r	$P(H_r \mid \underline{x}, \underline{y})$
$\theta_1 = \theta_2 = \theta_3 = \theta_4$	0.0139	$\theta_1 = \theta_3 = \theta_4 \neq \theta_2$	0.0083	$\theta_1 \neq \theta_2 = \theta_3 = \theta_4$	0.0492
$\theta_1 = \theta_2 = \theta_3 \neq \theta_4$	0.0252	$\theta_1 = \theta_3 \neq \theta_2 = \theta_4$	0.0072	$\theta_1 \neq \theta_2 = \theta_3 \neq \theta_4$	0.0350
$\theta_1 = \theta_2 = \theta_4 \neq \theta_3$	0.0331	$\theta_1 = \theta_3 \neq \theta_2 \neq \theta_4$	0.0113	$\theta_1 \neq \theta_2 = \theta_4 \neq \theta_3$	0.0467
$\theta_1 = \theta_2 \neq \theta_3 = \theta_4$	0.3531	$\theta_1 = \theta_4 \neq \theta_2 = \theta_3$	0.0067	$\theta_1 \neq \theta_2 \neq \theta_3 = \theta_4$	0.1745
$\theta_1 = \theta_2 \neq \theta_3 \neq \theta_4$	0.1483	$\theta_1 = \theta_4 \neq \theta_2 \neq \theta_3$	0.0141	$\theta_1 \neq \theta_2 \neq \theta_3 \neq \theta_4$	0.0733

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 0.3531, 0.1745 and 0.1483, respectively. Thus the data lend greatest support to equalities for $\theta_1 = \theta_2$ and $\theta_3 = \theta_4$ being different from the others.

The Bayesian approach using nonparametric Dirichlet process priors facilitates studying the problem of multiple comparisons in a number of different distributions. So far, the MCP was carried out for a bivariate distribution. The method can be extended to a multivariate distribution as well, with moderate effort.

References

1. Antoniak, C.E.(1974), Mixtures of Dirichlet Processes with Applications to Nonparametric Problems, *The Annals of Statistics*, 2, 1152-1174. 1129-1140.
2. Blackwell, D. and MacQueen, J. B.(1973), Ferguson Distribution via Polya Urn Schemes, *The Annals of Statistics*, 1, 353-355.
3. 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(3), 263-268.
4. Escobar, M. D. and West, M.(1995), Bayesian Density Estimation and Inference using Mixtures, *Journal of the American Statistical Association*, 90, 577-588.
5. Freund, J. E.(1961). A Bivariate Extension of the Exponential Distribution, *Journal of American Statistical Association*, 56, 971-977.
6. Gopalan, R. and Berry, D.A.(1998), Bayesian Multiple Comparisons Using Dirichlet Process Priors, *Journal of the American Statistical Association*, 90, 1130 - 1139.
7. Hanagal, D. D.(1996). Estimation of System Reliability From Stress-Strength Relationship, *Communication in Statistics-Theory and Methods*, 25(8), 1783-1797.
8. Hanagal, D. D. and Kale, B. K.(1992). Large Sample Tests for Testing

Symmetry and Independence in Some Bivariate Exponential Model, *Communication in Statistics-Theory and Methods*, 21(9), 2625-2643.

9. Kunchur, S. H. and Munoli, S. B.(1994). Estimation of Reliability in Freund Model for Two Component System, *Communication in Statistics-Theory and Methods*, 23(11), 3273-3283.

[received date : Apr. 2007, accepted date : May. 2007]