

## Nonparametric Bayesian Multiple Comparisons for Geometric Populations

M. Masoom Ali<sup>1)</sup> · Cho, J.S<sup>2)</sup> · Munni Begum<sup>3)</sup>

### Abstract

A nonparametric Bayesian method for calculating posterior probabilities of the multiple comparison problem on the parameters of several Geometric populations is presented. Bayesian multiple comparisons under two different prior/ likelihood combinations was studied by Gopalan and Berry(1998) using Dirichlet process priors. In this paper, we followed the same approach to calculate posterior probabilities for various hypotheses in a statistical experiment with a partition on the parameter space induced by equality and inequality relationships on the parameters of several geometric populations. This also leads to a simple method for obtaining pairwise comparisons of probability of successes. Gibbs sampling technique was used to evaluate the posterior probabilities of all possible hypotheses that are analytically intractable. A numerical example is given to illustrate the procedure.

**Keywords** : Dirichlet Process Prior, Geometric Distribution, Gibbs Sampler, Mixture of Dirichlet Processes, Multiple Comparison, Nonparametric Bayes

### 1. INTRODUCTION

The geometric model has been widely used as a model in areas ranging from studies on the inspection of defective items to research involving statistical quality control. We consider the multiple comparisons problem (MCP) for  $K$  geometric populations with parameters  $\theta = (\theta_1, \dots, \theta_K)$ , the probabilities of successes, to make

---

1) First Author : Department of Mathematical Sciences, Ball State University, Muncie, IN 47306-0490, U.S.A.

2) Department of Informational Statistics, Kyungsoong University, Pusan, Korea, 608-736

3) Department of Mathematical Sciences, Ball State University, Muncie, IN 47306-0490, U.S.A.

inferences on the relationships among the  $\theta_i$ 's based on observations. With equality and inequality relationships among  $\theta_i$ 's we can set up statistical hypotheses as under,

$$\begin{aligned}
 H_0 : \theta_1 = \theta_2 = \dots = \theta_K, \\
 H_1 : \theta_1 \neq \theta_2, \theta_2 = \theta_3 = \dots = \theta_K, \\
 \vdots \\
 H_N : \theta_1 \neq \theta_2 \neq \theta_3 \neq \dots \neq \theta_K.
 \end{aligned} \tag{1}$$

To our knowledge, the multiple comparison problem (MCP) among  $K$  geometric populations has not been studied yet, partly because of the difficulty in handling the distributional aspects and the computations. The Bayesian approach to the MCP for beta/binomial and normal/inverted gamma studied by Gopalan and Berry(1998), is extended here to the geometric populations. In a Bayesian MCP problem, prior probabilities on the hypotheses are elicited through specification of a process prior on the parameters of interest  $\theta = (\theta_1, \dots, \theta_K)$ . Then posterior probabilities are calculated from the posterior distributions induced by the hypotheses. The Dirichlet process prior (DPP) is a typical objective prior specification. The DPP is a prior distribution on the family of distributions, which is dense in the space of distribution functions.

The family of DPPs is introduced by Ferguson(1973) and is extended to mixtures of DPP by Antoniak(1974). The introduction of Markov chain Monte Carlo (MCMC) methods in nonparametric Bayesian modeling was begun with Escobar (1988). Innovations in computations and the development of new MCMC schemes are found in the key contributions by Doss (1994), Bush and MacEachern (1996), Escobar and West (1997), MacEachern and Müller (1998), West, Müller and Escobar (1994). Gopalan and Berry(1998) proposed Bayesian multiple comparisons using Dirichlet process priors for applied normal and binomial data.

In this paper, we consider the Bayesian approach to resolve the multiple comparisons problem for the probabilities of successes among  $K$  geometric populations, based on an hierarchical nonparametric family of Dirichlet process priors. Calculating posterior probabilities for the hypotheses is analytically intractable but can easily be evaluated using Gibbs sampling. In section 2 some reviews on the DPP are given, while section 3 presents the calculation of posterior probabilities for the hypotheses in MCP. A numerical example illustrating the procedure is described in section 4.

## 2. PRELIMINARIES

The Dirichlet process prior  $G$  is determined by two parameters: a distribution function  $G_0(\cdot)$  and a positive scalar precision parameter  $\alpha$ .  $G_0(\cdot)$  defines the location of the DPP. So  $G_0(\cdot)$  is sometimes called prior "guess" or baseline prior. The precision parameter  $\alpha$  determines the concentration of the prior for  $G$  around the prior guess  $G_0$ , and therefore measures the strength of belief in  $G_0$ . Then the DPP is denoted as  $G \sim D(G | G_0, \alpha)$ . For large values of  $\alpha$ , a sampled  $G$  is very likely to be close to  $G_0$ . For small values of  $\alpha$ , a sampled  $G$  is likely to put most of its probability mass on just a few atoms.

Mixture of Dirichlet Process models have become increasingly popular for modeling when conventional parametric models would impose unreasonably stiff constraints on the distributional assumptions (such as finite mixture of distributions). Despite of the large variety of applications, the core of the mixture of Dirichlet process model can basically be thought of as a simple Bayes model given by the likelihood and prior with added uncertainty about the prior distribution  $G \sim D(G | G_0, \alpha)$ . The more complex models typically require another portion to the hierarchy that allows the introduction of distributions on the hyperparameters  $\alpha$  and  $G_0$ .

Consider  $K$  geometric populations with the probabilities of success  $\theta = (\theta_1, \theta_2, \dots, \theta_K)$ . Observations  $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_K)$  are available on these populations, where  $\mathbf{Y}_i = (y_{i1}, \dots, y_{in_i})$  is  $n_i \times 1$  vector of conditionally independent observations on population  $i$ ,  $i = 1, 2, \dots, K$ ;  $j = 1, 2, \dots, n_i$  and  $\sum_{i=1}^K n_i = n$ . Then the probability density function of  $y_{ij}$  is

$$f(y_{ij} | \theta_i) = \theta_i \cdot (1 - \theta_i)^{y_{ij}}, y_{ij} = 0, 1, 2, \dots \tag{2}$$

We assume that the  $\theta_i$ 's come from  $G$ , and that  $G \sim D(G | G_0, \alpha)$ . This structure results in a posterior distribution which is a mixture of Dirichlet processes (Antoniak 1974). Using the Polya urn representation of the Dirichlet process (Blackwell and MacQueen 1973), the joint posterior distribution has the form

$$\theta_i | \mathbf{Y} \propto \prod_{i=1}^K f(\mathbf{y}_i | \theta_i) \frac{\alpha G_0(\theta_i) + \sum_{k \neq i} \delta(\theta_i | \theta_k)}{\alpha + i - 1}, \quad (3)$$

where  $\delta(\theta_i | \theta_k)$  is the distribution which has a point mass on  $\theta_k$ . For each  $i = 1, \dots, K$ , the conditional posterior distribution of  $\theta_i$  is given by

$$\theta_i | \theta_k, k \neq i, \mathbf{Y} \propto q_0 G_b(\theta_i | y_i) + \sum_{k \neq i} q_k \delta(\theta_i | \theta_k), \quad (4)$$

where  $G_b(\theta_i | bfy_i)$  is the baseline posterior distribution,

$$q_0 \propto \alpha \int f(bfy_i | \theta_i) dG_0(\theta_i), \quad q_k \propto f(\mathbf{y}_k | \theta_k) \quad \text{and} \quad 1 = q_0 + \sum_{k \neq i} q_k.$$

The multiple comparison problem of  $K$  geometric populations is to make inferences concerning relationships among the  $\theta$ 's based on  $\mathbf{Y}$ . Let  $\Theta = \theta = (\theta_1, \theta_2, \dots, \theta_K) : \theta_i \in R, i = 1, 2, \dots, K$  be the  $K$ -dimensional parameter space. Equality and inequality relationships among the  $\theta$ 's induce statistical hypotheses that are subsets of  $\Theta$ , *i.e.*,  $H_0 : \theta_0 = \theta_i : \theta_1 = \theta_2 = \dots = \theta_K$ ,  $H_1 : \theta_1 = \theta_i : \theta_1 \neq \theta_2 = \dots = \theta_K$  and so on up to  $H_N : \theta_N = \theta_i : \theta_1 \neq \theta_2 \neq \dots \neq \theta_K$ . The hypotheses  $H_r : \theta_r, r = 0, 1, 2, \dots, N$ , are disjoint, and  $\cup_{r=0}^N \theta_r = \Theta$ .

The elements of  $\Theta$  themselves behave as described by (3) and so with positive probability, they will reduce to some  $p \leq K$  distinct values. We denote the distinct values of the parameters by putting a superscript  $*$  on them. Then any realization of  $K$  parameters  $\theta_i$  generated from  $G$  lies in a set of  $p \leq K$  distinct values, denoted by  $\theta^* = (\theta_1^*, \theta_2^*, \dots, \theta_p^*)$ . The Gibbs sampling for MCP problem can be described through what is termed as Configuration by Gopalan and Berry (1998). Their definition of *Configuration* is restated here,

**Definition:** The configuration  $S = S_1, S_2, \dots, S_l$  determines a classification of  $\theta$  into  $p$  distinct groups or clusters;  $n_j^\theta =$  number of  $\theta_i$ 's in group  $j$  that share the common parameter value  $\theta_j^*$ . Write  $K_j$  for the set of indices of parameters in group  $j$ ,  $K_j = i : S_i = j$ . Let  $\mathbf{X}_j = \{\mathbf{X}_j : S_i = j\}$  be the corresponding group of

$$n_{K_j} = \sum_{i \in K_j} n_i \text{ observations}$$

There is a one-to-one correspondence between hypotheses and configurations.

And the required computations are reduced by the fact that the distinct  $\theta_i$ 's are typically reduced to fewer than  $K$  due to the clustering of the  $\theta_i$ 's inherent in the Dirichlet process. Hence, (4) can be rewritten as:

$$\theta_i \mid \theta_{k,k} \neq i, \mathbf{y} \propto q_0 G_b(\theta_i \mid \mathbf{y}_i) + \sum_{k=1}^{K^*} n_k q_k^* \delta(\theta_i \mid \theta_k^*), \tag{5}$$

with  $q_k^* \propto f(\mathbf{y}_i \mid \theta_k^*)$ , and  $1 = q_0 + \sum_k n_k q_k^*$ . In addition to simplifying notation, the cluster structure of the  $\theta_i$  can also be used to improve the efficiency of the algorithm.

### 3. POSTERIOR SAMPLING IN DIRICHLET PROCESS MIXTURES

We take a beta distribution as baseline prior  $G_0$  with parameters  $(\alpha_{oi}, \beta_{oi})$  and  $\theta_1, \theta_2, \dots, \theta_K$  are iid from  $G_0$ . Extending to a Dirichlet process analysis as outline in the above description results

$$\mathbf{y}_i \mid \theta_i \sim Ge(\mathbf{y}_i \mid \theta_i), \tag{6}$$

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

$$G \mid G_0, \alpha \sim D(G \mid G_0, \alpha), \tag{8}$$

$$G_0 \mid (\alpha_{oi}, \beta_{oi}) \sim B(\alpha_{oi}, \beta_{oi}). \tag{9}$$

Here,  $Ge$  indicates geometric distribution. Now the choice of the precision parameter  $\alpha$  in Dirichlet process is extremely important for the model. Here, we consider the gamma prior for  $\alpha$  with a shape parameter  $a$  and scale parameter  $b$ , that is,  $\alpha \sim \Gamma(a, b)$ . Then the  $\Gamma(a, b)$  is to be the reference prior by  $a \rightarrow 0$  and  $b \rightarrow 0$ . And we have access to a neat data augmentation device for sampling  $\alpha$  by Escobar and West(1995).

The configuration notation is more convenient to use in describing the Gibbs sampling algorithm. This results in sampling from the following conditional posterior distributions:

$$(\theta_i | Y, \theta_k, k \neq i, \alpha, \alpha_{oi}, \beta_{oi}) \sim q_0 B(n_i + \alpha_{oi}, \sum_{j=1}^{n_i} y_{ij} + \beta_{oi}) + \sum_{k \neq i} q_k \delta(d\theta_i | \theta_k), \quad (10)$$

$$(\theta_j^* | Y, S, \alpha_{oi}, \beta_{oi}) \sim B(\sum_{k \in J_j} n_k + \alpha_{oj}^*, \sum_{k \in J_j} \sum_{l=1}^{n_i} y_{kl} + \beta_{oj}^*), \quad (11)$$

$$(\alpha | \eta, K^*) \sim \pi_\eta \Gamma(a + K^*, b - \log(\eta)) + (1 - \pi_\eta) \Gamma(a + K^{*-1}, b - \log(\eta)), \quad (12)$$

$$(\eta | \alpha, K^*) \sim B(\alpha + 1, K). \quad (13)$$

where

$$q_0 \propto \alpha \frac{\Gamma(\alpha_{oi} + \beta_{oi})}{\Gamma(\alpha_{oi}) \cdot \Gamma(\beta_{oi})} \cdot \frac{\Gamma(n_i + \alpha_{oi}) \cdot \Gamma(\sum_{j=1}^{n_i} y_{ij} + \beta_{oi})}{\Gamma(n_i + \sum_{j=1}^{n_i} y_{ij} + \alpha_{oi} + \beta_{oi})},$$

$$q_k \propto \theta_k^{n_i} (1 - \theta_k)^{\sum_{j=1}^{n_i} y_{ij}}.$$

Gibbs sampling proceeds by simply iterating through (10) – (13) in order, sampling at each stage based on current values of all the conditioning variables.

The configuration gives the equality and inequality relationships among the  $\theta$ 's, which correspond to the partitions on the parameter space  $\Theta$  and in turn to the hypotheses of interest. To estimate the posterior probability of a hypothesis  $H_r$  from a large number ( $L$ ) of sample draws, we use

$$P(H_r | \mathbf{Y}) \approx \frac{1}{L} \sum_{l=1}^L \delta_{S_l}(H_r),$$

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

The probability of equality for any two  $\theta$ 's can be calculated from the posterior distributions on hypotheses,  $P(H_r | \mathbf{Y})$ ,  $r = 1, 2, \dots, N$ . This can be achieved by adding probabilities of those hypotheses in which  $\theta_i$  and  $\theta_j$  are equal. For example,

$$P(\theta_i = \theta_j | \mathbf{Y}) \approx \frac{1}{L} \sum_{l=1}^L \delta_{S_l}(\theta_i = \theta_j) = \sum_{r=1}^N P(H_r | \mathbf{Y}) \delta_{H_r}(\theta_i = \theta_j), \quad i \neq j, \quad (15)$$

where  $\delta_{S_l}(\theta_i = \theta_j)$  and  $\delta_{H_r}(\theta_i = \theta_j)$  denote unit point mass for the case where  $S_l$  and  $H_r$  indicate  $\theta_i = \theta_j$ , respectively.

### 4. ILLUSTRATIVE EXAMPLE

In this section, we use hypothetical data to illustrate the multiple comparisons for the parameters of geometric populations. Here, we consider 4 geometric populations and sample size of 50 from each geometric distributions. In this paper, we consider multiple comparisons for two cases, so that true hypothesis are  $H_{true} : \theta_1 = \theta_2 = \theta_3 \neq \theta_4$  for case I and  $H_{true} : \theta_1 = \theta_2 \neq \theta_3 = \theta_4$  for case II, respectively.

**Table 1** The observed summary statistics for each populations

Populations	Case I				Case II			
	1	2	3	4	1	2	3	4
$y_i = \sum_{j=1}^{n_i} y_{ij}$	30	31	27	76	30	31	76	71
$n_i$	15	15	15	15	15	15	15	15

The observed summary statistics for each case are given as Table 1. And the numbers of possible hypothesis are 15 for each case. For the precision parameter  $\alpha$ , we consider three Gamma priors with parameters  $(a, b) = (1.0, 1.0)$ ,  $(0.1, 0.1)$  and  $(0.01, 0.01)$  such that same means 1 and different variances 1, 10, and 100, respectively. Then latter prior is fairly noninformative, giving reasonable mass to both high and low values of  $\alpha$ . But, the  $\Gamma(1.0, 1.0)$  prior favors relatively low values of  $\alpha$ . Also we set that each  $\theta_i, i = 1, \dots, 4$  follows in priori beta with parameter  $\alpha_{oi} = \beta_{oi} = 1.0$  to reflect vagueness of the prior knowledge. All the calculated posterior probabilities for all possible hypotheses are approximated by the Gibbs sampling algorithm using 40,000 iterations with 20,000 burn-in iterations. Table 2 and Table 3 give the calculated posterior probabilities for all possible hypotheses of case I and case II, respectively.

**Table 2** Calculated posterior probabilities for each hypothesis with three cases of  $(a, b)$  in Case I

Hypothesis	(1.0, 1.0)	(0.1, 0.1)	(0.01, 0.01)
$\theta_1 = \theta_2 = \theta_3 = \theta_4$	0.0064	0.0225	0.1883
$\theta_1 = \theta_2 = \theta_4 \neq \theta_3$	0.7370	0.6427	0.4412
$\theta_1 = \theta_2 = \theta_4 \neq \theta_3$	0.0011	0.0011	0.0006
$\theta_1 = \theta_2 \neq \theta_3 = \theta_4$	0.0014	0.0014	0.0005
$\theta_1 = \theta_2 \neq \theta_3 = \theta_4$	0.0711	0.0786	0.0633
$\theta_1 = \theta_2 \neq \theta_3 \neq \theta_4$	0.0007	0.0004	0.0001
$\theta_1 = \theta_3 = \theta_4 \neq \theta_2$	0.0064	0.0053	0.0036
$\theta_1 = \theta_3 \neq \theta_2 = \theta_4$	0.0206	0.0239	0.0238
$\theta_1 = \theta_3 \neq \theta_2 \neq \theta_4$	0.0048	0.0032	0.0019
$\theta_1 = \theta_2 = \theta_3 = \theta_4$	0.0011	0.0011	0.0007
$\theta_1 = \theta_4 \neq \theta_2 = \theta_3$	0.0015	0.0010	0.0011
$\theta_1 \neq \theta_2 = \theta_3 = \theta_4$	0.1159	0.1503	0.1319
$\theta_1 \neq \theta_2 = \theta_4 \neq \theta_3$	0.0013	0.0012	0.0009
$\theta_1 \neq \theta_2 \neq \theta_3 = \theta_4$	0.0004	0.0001	0.0001
$\theta_1 \neq \theta_2 \neq \theta_3 \neq \theta_4$	0.0303	0.0672	0.1420

It is evident from Table 2, that the hypotheses for " $\theta_1 = \theta_2 = \theta_3 \neq \theta_4$ " have the largest posterior probabilities 0.7370, 0.6427, and 0.4412 for all priors of the precision parameter  $\alpha$ , respectively. This suggests that the data lend greatest support to equalities for  $\theta_1 = \theta_2 = \theta_3$  and  $\theta_4$  being different from the others. Thus this example shows good performance of the Bayesian multiple comparisons method for several geometric parameters.



**Table 3** Calculated posterior probabilities for each hypothesis with three cases of  $(a, b)$  in Case II

Hypothesis	(1.0, 1.0)	(0.1, 0.1)	(0.01, 0.01)
$\theta_1 = \theta_2 = \theta_3 = \theta_4$	0.0079	0.0470	0.2795
$\theta_1 = \theta_2 = \theta_4 \neq \theta_3$	0.0006	0.0010	0.0004
$\theta_1 = \theta_2 = \theta_4 \neq \theta_3$	0.0000	0.0000	0.0000
$\theta_1 = \theta_2 \neq \theta_3 = \theta_4$	0.6763	0.5672	0.3444
$\theta_1 = \theta_2 \neq \theta_3 = \theta_4$	0.1177	0.1192	0.0914
$\theta_1 = \theta_2 \neq \theta_3 \neq \theta_4$	0.0072	0.0050	0.0038
$\theta_1 = \theta_3 = \theta_4 \neq \theta_2$	0.0000	0.0000	0.0000
$\theta_1 = \theta_3 \neq \theta_2 = \theta_4$	0.0008	0.0011	0.0009
$\theta_1 = \theta_3 \neq \theta_2 \neq \theta_4$	0.0000	0.0000	0.0000
$\theta_1 = \theta_2 = \theta_3 = \theta_4$	0.0001	0.0002	0.0002
$\theta_1 = \theta_4 \neq \theta_2 = \theta_3$	0.0118	0.0095	0.0071
$\theta_1 \neq \theta_2 = \theta_3 = \theta_4$	0.0016	0.0015	0.0009
$\theta_1 \neq \theta_2 = \theta_4 \neq \theta_3$	0.0001	0.0001	0.0002
$\theta_1 \neq \theta_2 \neq \theta_3 = \theta_4$	0.1391	0.1487	0.1103
$\theta_1 \neq \theta_2 \neq \theta_3 \neq \theta_4$	0.0368	0.0995	0.1609

Table 3 indicates that the hypotheses for " $\theta_1 = \theta_2 \neq \theta_3 = \theta_4$ " have the largest posterior probabilities 0.6763, 0.5672, and 0.3444 for all priors of the precision parameter  $\alpha$ , respectively. This suggests that the data lend greatest support to equalities for  $\theta_1 = \theta_2$  and  $\theta_3 = \theta_4$  being different from the others. Also, Thus this example shows good performance of the Bayesian multiple comparisons method for several geometric parameters.

**Table 4** Pairwise Posterior Probabilities with three cases of  $(a, b)$  in Case I

Hypothesis	(1.0, 1.0)	(0.1, 0.1)	(0.01, 0.01)
$\theta_1 = \theta_2$	0.8170	0.7463	0.6939
$\theta_1 = \theta_3$	0.7711	0.6948	0.6570
$\theta_1 = \theta_4$	0.0141	0.0283	0.1916
$\theta_2 = \theta_3$	0.8656	0.8197	0.7644
$\theta_2 = \theta_4$	0.0167	0.0311	0.1945
$\theta_3 = \theta_4$	0.0104	0.0254	0.1901

Table 4 shows the pairwise posterior probabilities for the equalities of pairs of  $\theta$ 's in case I. By table 4, the equalities of  $(\theta_1 = \theta_2)$  and  $(\theta_2 = \theta_3)$  have the largest posterior probabilities (0.8170, 0.7463, 0.6939) and (0.8656, 0.8197, 0.7644) for three cases of  $(a, b)$ , respectively. This suggests that there is strong evidence in the equalities  $\theta_1 = \theta_2$  and  $\theta_2 = \theta_3$ .

**Table 5** Pairwise Posterior Probabilities with three cases of  $(a, b)$  in Case II

Hypothesis	(1.0, 1.0)	(0.1, 0.1)	(0.01, 0.01)
$\theta_1 = \theta_2$	0.8025	0.7344	0.7157
$\theta_1 = \theta_3$	0.0165	0.0541	0.2846
$\theta_1 = \theta_4$	0.0152	0.0522	0.2835
$\theta_2 = \theta_3$	0.0219	0.0590	0.2879
$\theta_2 = \theta_4$	0.0198	0.0566	0.2868
$\theta_3 = \theta_4$	0.8423	0.7774	0.7451

Table 5 demonstrates the pairwise posterior probabilities for equality of pairs of  $\theta$ 's in case II. By table 5, the equalities of  $(\theta_1 = \theta_2)$  and  $(\theta_3 = \theta_4)$  are most large posterior probability (0.8025, 0.7344, 0.7157) and (0.8423, 0.7774, 0.7451) for three cases of  $(a, b)$ , respectively. This suggests that there is strong evidence in the equalities  $\theta_1 = \theta_2$  and  $\theta_3 = \theta_4$ .

Up to this point, we have considered the problem of developing a Bayesian multiple comparisons for means of  $K$  geometric populations. As an alternative to a formal Bayesian analysis of a mixture model that usually leads to intractable

calculations, the DPP is used to provide a nonparametric Bayesian method for obtaining posterior probabilities for various hypotheses of equality among population means.

Extension of the above approach to the multiple comparison problems for the another population is straightforward. The research topics pertaining to the extension of the method and the examination of its performance are worthy to study and are left as a future subject of research.

**Acknowledgement** : The authors are thankful to the two referees for their careful reading of the manuscript and for their helpful comments.

## REFERENCES

1. Antoniak, C.E. (1974), Mixtures of Dirichlet Processes with Applications to Nonparametric Problems, *The Annals of Statistics*, 2, 1152-1174.
2. Blackwell, D. and MacQueen, J.B. (1973), Ferguson Distribution via Polya Urn Schemes, *The Annals of Statistics*, 1, 353-355.
3. Bush, C.A. and MacEachern, S.N. (1996), A Semi-parametric Bayesian Model for randomized Block Designs, *Biometrika*, 83, 275-285.
4. Doss, H. (1994), Bayesian Nonparametric Estimation for Incomplete Data via Successive Substitution Sampling, *The Annals of Statistics*, 22, 1763-1786.
5. Escobar, M. D. (1988), Estimating the Means of Several Normal Populations by Nonparametric Estimation of the Distribution of the Means. Unpublished dissertation, Yale University.
6. Escobar, M. D. and West, M. (1995), Bayesian Density Estimation and Inference using Mixtures, *Journal of the American Statistical Association*, 90, 577-588.
7. Escobar, M. D. and West, M. (1997), Computing Nonparametric Hierarchical Models, *ISDS Discussion Paper #97-15*, Duke University.
8. Ferguson, T.S. (1973), A Bayesian Analysis of Some Nonparametric Problems, *The Annals of Statistics*, 1, 209-230.
9. Gopalan, R. and Berry, D.A. (1998), Bayesian Multiple Comparisons Using Dirichlet Process Priors, *Journal of the American Statistical Association*, 90, 1130 - 1139.
10. MacEachern, S.N. and Müller, P. (1998), Estimating Mixture of Dirichlet Process Models, *Journal of Computational and Graphical Statistics*, 7, 223-239.
11. West, M., Müller, P. and Escobar, M.D. (1994), *Hierarchical Priors and Mixture Models, with Application in Regression and Density*

*Estimation, Aspects of Uncertainty* : A Tribute to D. V. Lindley (eds: A.F.M. Smith and P.R. Freeman), London : John Wiley and Sons, 363-386.

[ received date : Sep. 2005, accepted date : Nov. 2005 ]