*Journal of the Korean Statistical Society (2000),* **29**: *2, pp 155–168*

# Bayesian Multiple Comparisons for Normal Variances

Hea-Jung Kim[1]

**Abstract**

Regarding to multiple comparison problem (MCP) of $k$ normal population variances, we suggest a Bayesian method for calculating posterior probabilities for various hypotheses of equality among population variances. This leads to a simple method for obtaining pairwise comparisons of variances in a statistical experiment with a partition on the parameter space induced by equality and inequality relationships among the variances. The method is derived from the fact that certain features of the hierarchical nonparametric family of Dirichlet process priors, in general, make it amenable to solving the MCP and estimating the posterior probabilities by means of posterior simulation, the Gibbs sampling. Two examples are illustrated for the method. For these examples, the method is straightforward for specifying distributionally and to implement computationally, with output readily adapted for required comparison.

*Key Words and Phrases:* Bayesian Multiple comparison; Posterior probabilities of hypotheses; Dirichlet process priors; Hierarchical model; Pairwise posterior probabilities; Gibbs sampler.

## 1. INTRODUCTION

In the literature, multiple comparisons problem (MCP) among $k$ normal means $\theta_1, \theta_2, \ldots, \theta_k$ has been studied by many authors and various procedures have been proposed, including Fisher's least significant difference (LSD), Duncan's multiple range test, Scheffé's test, and so on. (for descriptions of these procedures see Hochberg and Tamhane 1987). The Bayesian approaches to the MCP can be found in Berry (1988) and Gopalan and Berry (1998). Apparently, the MCP of $k$ normal population variances is not discussed by many authors. There are few published papers about the MCP due to Hsu (1977), Inclán (1993), and Chen and Gupta (1997). They considered the MCP under the assumption that $k$ normal population means are common, $\theta_1 = \theta_2 = \ldots = \theta_k$. However, the study about the MCP of $k$ normal population variances under unequal means has not been seen yet, in part because of the difficulty in handling

[1]Department of Statistics, Dongguk University, Seoul 100-715, Korea

the computations. The MCP that is associated with the variance change problem, plays an essential role in finance research to study the nonstationarity in the parameters of the multi-factor market model (cf. Vostrikova 1981 and Chen and Gupta 1997 and references within). In this article, we introduce Bayesian approach to resolving the MCP of $k$ normal population variances having unequal means. As usual, the Bayesian approach is to condition on the observations in updating one's prior probability distribution. But assessing a prior distribution and formulating a likelihood in the presence of large number $N$ of hypotheses ($H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2, \sigma_2^2 = \sigma_3^2 = \ldots = \sigma_k^2$, and so on up to $H_N : \sigma_1^2 \neq \sigma_2^2 \neq \ldots \neq \sigma_k^2$) make the Bayesian approach difficult. As the number $k$ of populations increases, the number of hypotheses increase exponentially. The number of hypotheses as a function of $k$ is given by the Bell exponential number $B_k$ (cf. Berger 1971). The sequence $\{B_k\}$ can be generated by the recursion $B_{k+1} = \sum_{i=1}^{k} {}_kC_i B_i$, $k = 0, 1, 2, \ldots$, where $B_0 = 1$ and $N = B_k - 1$ for $k \geq 2$. Even for a reasonably small number of populations, such as $k = 5$ and $k = 6$, the number of hypotheses $N$ to be considered (52 and 203) is very large. Thus, the MCP of $k$ normal population variances is tedious for moderate $k$ and it is practically impossible for large $k$. To circumvent this problem, in this article, the hierarchical nonparametric family of Dirichlet process priors(DPP) introduced by Ferguson (1973) is applied in the form of baseline prior/likelihood combinations to obtain posterior probabilities for various hypotheses of equality among population variances. Then we develop a numerical technique to calculate the posterior probabilities of the hypotheses based on a hierarchical nonparametric family of DPP.

## 2. MIXTURE OF DIRICHLET PROCESS MODEL

Mixture of Dirichlet Process models have become increasingly popular for modeling when conventional parametric models would impose unreasonably stiff constraints on the distributionally assumptions (such as finite mixture of distributions). A list of applications can be found in MacEachren and Müller (1998). 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 $G \sim D(\alpha G_0)$, where $G \sim D(\alpha G_0)$ refers to $G$ being a random distribution generated by a Dirichlet process with base measure $\alpha G_0$ and total mass parameter

$\alpha$. The more complex models typically require another portion to the hierarchy that allows the introduction of distributions on the hyperparameters $\alpha$ and $G_0$.

### 2.1. Dirichlet Process Prior

Consider $k$ normal populations with parameters $(\theta_1, \tau_1), (\theta_2, \tau_2), \ldots, (\theta_k, \tau_k)$, respectively, where $\tau_i = 1/\sigma_i^2$, $i = 1, \ldots, k$. Let $\mathbb{X} = \{X_1, X_2, \ldots, X_k\}$ are observations available on these populations, where $X_i = (x_{i1}, x_{i2}, \ldots, x_{in_i})'$ is a $n_i \times 1$ vector of conditionally independent observations on population $i = 1, \ldots, k$; $j = 1, \ldots, n_i$. The multiple comparisons problem (MCP) of $k$ variances is to make inferences concerning relationships among the $\tau$'s based on $\mathbb{X}$. Let $\Omega = \{\tau = (\tau_1, \tau_2, \ldots, \tau_k) : \tau_i \in R^+,\ i = 1, 2, \ldots, k\}$ be the $k$-dimensional parameter space. Equality and inequality relationships among the $\tau_i$'s induce statistical hypotheses that subsets of $\Omega$: $H_0 : \Omega_0 = \{\tau_i : \tau_1 = \tau_2 = \ldots = \tau_k\}$, $H_1 : \Omega_1 = \{\tau_i : \tau_1 \neq \tau_2, \tau_2 = \tau_3 = \ldots = \tau_k\}$, and so on up to $H_N : \Omega_N = \{\tau_i : \tau_1 \neq \tau_2 \neq \ldots \neq \tau_k\}$. The hypotheses $(H_r : \Omega_r; r = 1, 2, \ldots, N)$, are disjoint, and $\Omega = \cup_{r=0}^{N} \Omega_r$.

For the prior distribution of $k$ normal population precisions, $\tau_i$'s, we use the family of DPPs introduced by Ferguson (1973) and extended to mixtures of DPP by Antoniak (1974). Certain features of this family make it amenable to applications in the MCP.

Antoniak (1974) defines a $C$ class and show that the DPP imposed on $\Omega = \{\tau_1, \tau_2, \ldots, \tau_k\}$ has the property of assigning prior probabilities to the hypotheses, which are subsets of the parameter space, induced by equality and inequality relationships among the $\tau_i$'s.

**Definition 1 (C class).** Let $\tau_1$, $\tau_2, \ldots,$ $\tau_k$ be a sample of size $k$ from a DPP. We will say that the sample belongs to the class $C(m_1, m_2, \ldots, m_k)$, and write $(\tau_1,\ \tau_2, \ldots,\ \tau_k) \in C(m_1, m_2, \ldots, m_k)$, if there are $m_1$ distinct values of $\tau$ that occur only once, $m_2$ that occur exactly twice, $\ldots$, $m_k$ that occur exactly $k$ times.

Two immediate consequences of this definition are that $k = \sum_{i=1}^{k} i m_i$, and the total number of distinct values that occur is $p = \sum_{i=1}^{k} m_i$. As an example of this notation we note that $H_0$ in the preceding discussion belongs to the class $C(0, 0, \ldots, 1)$. The prior probability of a hypothesis of interest in terms of its $C$ class is given by the following proposition.

**Proposition 1** (Antoniak, 1974). Let $G \sim D(\alpha G_0)$ be a DPP on a standard Borel space $(\Omega, A)$, with concentration parameter $\alpha > 0$. Let $\tau_1$, $\tau_2, \ldots,$ $\tau_k$ be a sample of size $k$ from $D(\alpha G_0)$. Then

$$P\{(\tau_1,\tau_2,\ldots,\tau_k)\in C(m_1,m_2,\ldots,m_k)\} = \frac{k!}{\prod_{i=1}^k i^{m_i}(m_i!)}\frac{\alpha^{\sum_{i=1}^k m_i}}{\prod_{i=1}^k(\alpha+i-1)}. \tag{2.1}$$

The proposition yields prior probability of each hypothesis. For example, $P(H_0)$ and $P(H_N)$ are $\alpha(k-1)!/\prod_{i=1}^k(\alpha+i-1)$ and $\alpha^k/\prod_{i=1}^k(\alpha+i-1)$, respectively. This is a key feature of the model structure, and of its analysis, relates to the discreteness of $G(\cdot)$ under the Dirichlet process assumption. Briefly, in any sample $\Omega=\{\tau_1,\tau_2,\ldots,\tau_k\}$ of size $k$ from $G(\cdot)$ there is positive probability of coincident values. Using the DPP for Bayesian inference requires choosing $\alpha$. Gopalan and Berry (1998) suggests a method for choosing $\alpha$ based on the quantities of $P(H_0)$ and $P(H_N)$. Another consequence of this prior is that the prior full conditional distributions of $\tau_i$ can be expressed as follows (cf. Antoniak 1974 and Ferguson 1973)

$$\tau_i|\Omega^{(i)} \sim \alpha a_{k-1}G_0(\tau_i)+a_{k-1}\sum_{j\neq i}\delta_j(\tau_i), \tag{2.2}$$

where $\Omega^{(i)}=\{\tau_1,\ldots,\tau_{i-1},\tau_{i+1},\ldots,\tau_k\}$, $\delta_j(\tau_i)$ denotes a unit point mass at $\tau_i=\tau_j$ and $a_r=1/(\alpha+r)$ for positive integers $r$. The elements of $\Omega$ themselves behave as described by (2.2) and so with positive probability, they will reduce to some $p\le k$ distinct values. Let $\{\tau_1^*,\ldots,\tau_p^*\}$ denote the set of distinct $\tau_i$'s, where $p\le k$ is the number of distinct elements in the vector $\Omega$. Let $S=(S_1,S_2,\ldots,S_k)$ denote the vector of configuration indicators defined by $S_i=j$ iff $\tau_i=\tau_j^*$, $i=1,\ldots,k$, and let $I_j$ be the index set for those occurrences, $I_j=\{i: S_i=j\}$, $j=1,\ldots,p$. As an illustration, let $k=5$ and $S=\{1,2,1,2,3\}$. Then $p=3$, $I_1=\{1,3\}$,$I_2=\{2,4\}$ and $I_3=\{5\}$. We will term the "cluster" to refer to the set of all observation $X_i$, or just the indexes $i$, or the corresponding $\tau_i$'s, with identical configuration indicators $S_i$. Let $n_j^*$ is the size of the $j$th cluster: $n_j^*=|\{i: S_i=j\}|$ with $\sum_{j=1}^p n_j^*=k$. Suppose that there are $n_j^*$ occurrences of $\tau_i$'s in group $j$ that share the common parameter value $\tau_j^*$. Then (2.2) reduces to the mixture of fewer components,

$$\tau_i|\Omega^{(i)},S^{(i)},p^{(i)} \sim \alpha a_{k-1}G_0(\tau_i)+a_{k-1}\sum_{j=1}^{p^{(i)}} n_j^{*(i)}\delta_i(\tau_j^{*(i)}), \tag{2.3}$$

where $S^{(i)}$ denotes the configuration of $\Omega^{(i)}$ into $p^{(i)}$ distinct values with $n_j^{*(i)}$ of them taking the common value $\tau_j^{*(i)}$, and $\delta_i(\tau_j^{*(i)})$ is a unit point mass at

$\tau_i = \tau_j^{*(i)}$. The decomposition of (2.3) demonstrates the moderating effect of the DPP, as new value of $\tau$ only occur with probability $\alpha a_{k-1}$.

### 2.2. Mixture of Dirichlet Process

Let $\mathbb{X}_j = \{X_j : S_i = j\}$ be the corresponding group of $n_{I_j} = \sum_{i \in I_j} n_i$ observations. To proceed, we need to specify the prior mean $G_0(\cdot)$ of $G(\cdot)$. A convenient form is the normal/gamma conjugate to the normal sampling model; thus, under $G_0$, we assume $\tau_i \sim Gamma(s/2, V/2)$, a gamma prior with shape $s/2$ and scale $V/2$, so that $dG_0(\tau_i) \propto \tau_i^{s/2-1} e^{-V\tau_i/2}$ and $(\theta_i|\tau_i) \sim N(\mu_i, (\beta\tau_i)^{-1})$, for some mean $\mu_i$, $i = 1, \ldots, k$, and scale factor $\beta \sim Gamma(\phi,$
$\psi)$. Thus we have the following hierarchy of prior distributions:

$$\begin{aligned}
\theta_i|\tau_i &\sim N(\mu_i, (\beta\tau_i)^{-1}),\\
G_0(\tau_i) &\overset{iid}{\sim} Gamma(s/2, V/2),\\
G|G_0 &\sim D(\alpha G_0),\\
\tau_1, \tau_2, \ldots, \tau_k|G &\overset{iid}{\sim} G,
\end{aligned}$$

and the hyperparameter $\beta$ has gamma distribution

$$\beta \sim Gamma(\phi, \psi).$$

The likelihood function is

$$L(\theta, \Omega|\mathbb{X}) \propto \prod_{i=1}^{k} \tau_i^{1/2} \exp\left[-\frac{\tau_i}{2}\sum_{j=1}^{n_i}(x_{ij} - \theta_i)^2\right].$$

A similar conditional (2.3) can be obtained a posteriori: given all other parameters and observations $\mathbb{X} = (X_1, X_2, \ldots, X_k)$, the new value $\tau_i$ is equal to $\tau_h$, $h \neq i$ with probability $q_{i,h} \propto p(X_i|\theta_i, \tau_h)$, or probability $q_{i,0} \propto \alpha \int p(X_i|\theta_i,$
$\tau_i)dG_0(\tau_i)$ is a draw from $dG_i(\tau_i) \propto dG_0(\tau_i)p(X_i|\theta_i, \tau_i)$. The distribution $G_i$ is the posterior in a simple Bayes model given by likelihood $X_i \sim p(X_i|\theta_i, \tau_i)$. Combining identical $\tau_h$'s, and redefining $q_{i,j} \propto n_j^{*(i)} p(X_i|\theta_i, \tau_j^*)$, the conditional posterior of $\tau_i$ is given by

$$[\tau_i|\Omega^{(i)}, S^{(i)}, p^{(i)}, \theta, \mathbb{X}] \sim \sum_{j=1}^{p^{(i)}} q_{i,j}\delta_i(\tau_j^{*(i)}) + q_{i,0}G_i(\tau_i), \tag{2.4}$$

where $q_{i,0} + \sum_{j=1}^{p^{(i)}} q_{i,j} = 1$ and

$$q_{i,0} \propto \alpha \int \tau_i^{(n_i+s)/2-1} \exp\{-\tau_i[V + \sum_{\ell=1}^{n_i}(x_{i\ell} - \theta_i)^2/2]\}d\tau_i,$$

$$\propto \frac{\alpha\Gamma\{(n_i+s)/2\}(V/2)^{s/2}}{(2\pi)^{n_i/2}\{(V+\sum_{\ell=1}^{n_i}(x_{i\ell}-\theta_i)^2)/2\}^{(n_i+s)/2}\Gamma\{(s/2)\}}, \tag{2.5}$$

$$q_{i,j} \propto (2\pi)^{-n_i/2} n_j^{*(i)} \tau_j^{*n_i/2} \exp\{-\tau_j^* \sum_{\ell=1}^{n_i}(x_{i\ell}-\theta_i)^2/2\},\ j=1,\ldots,p^{(i)}. \tag{2.6}$$

Note that sampling $\tau_i$ implicitly samples a new configuration indicator $S_i$ such that equation (2.4) implies conditional posterior probabilities for the configuration indicators as

$$P(S_i=j|\Omega^{(i)},S^{(i)},p^{(i)},\theta) = q_{i,j}, \tag{2.7}$$

where $\theta=\{\theta_1,\ldots,\theta_k\}$.

By use of the definition above, we can easily see that

$$G_i(\tau_i) = Gamma\left(\frac{n_i+s}{2}, \frac{V+\sum_{\ell=1}^{n_i}(x_{i\ell}-\theta_i)^2}{2}\right). \tag{2.8}$$

The conditional distributions (2.4) imply that $\tau_i$ is a new parameter with probability $q_{i,0}$ and is equal to another parameter otherwise. Therefore, once the Gibbs sequence of simulation (described in the next section) from (2.4) has been run for all $i$, and $p$ different values of $\tau_i$ have been obtained, the actual distribution of the sample $\mathbf{X}$ is indeed mixture of $k$ normal distributions consisting of $p$ clusters of equal variance, although the generating component of each $X_i$ is well known. This feature makes quite different with the usual mixture approach, as $p$ varies at each simulation step.

## 3. GIBBS SAMPLING SCHEME

MCMC implementation to estimate the mixture of Dirichlet process model discussed in the previous section may be presented in terms of Steps 1 through 3.

Step1. Given current values of $\Omega^*=\{\tau_1^*,\ldots,\tau_p^*\}$, $p\le k$, $\theta$ and $S$, generate a new configuration by sequentially sampling indicators from the posterior probabilities (2.7), successively simulating and substituting $S_1,\ldots,S_k$ for all index $i$ such that when $S_i=0$ draw a new $\tau_i$ from $G_i(\tau_i)$ in (2.8).

Step 2. Given $p$, $S$ and $\Omega^*$, resample $\theta_i$'s from the posterior distribution

$$[\theta_i|p,S,\Omega^*,\beta,\mathbf{X}] \sim N\left(\frac{n_i\bar{X}_i+\beta\mu_i}{n_i+\beta}, \frac{1}{(n_i+\beta)\tau_j^*}\right) \quad \text{for } i\in I_j,$$

where $\bar{X}_i = \sum_{\ell=1}^{n_i} x_{i\ell}$. Then generate a new set of parameters $\Omega^*$ by sampling each new $\tau_j^*$ from the relevant component posterior

$$[\tau_j^*|p, S, \theta, \mathbf{X}] \sim Gamma\left(\frac{n_{I_j}+s}{2}, \frac{V+w_j}{2}\right),$$

where $w_j = \sum_{r\in I_j}\sum_{\ell=1}^{n_r}(x_{r\ell}-\theta_r)^2$.

Step 3. Generate the hyperparameter $\beta$ from

$$[\beta|p, S, \Omega^*, \theta, \mathbf{X}] \sim Gamma\left(\phi+\frac{k}{2}, \psi+\frac{\sum_{j=1}^{p}(\tau_j^*\sum_{i\in I_j}(\theta_i-\mu_i)^2)}{2}\right).$$

Return to Step 1, and proceed iteratively until convergence.

Although not explicitly included in model defined in Subsection 2.2, we would include a hyperprior on the total mass $\alpha$ of $G$ and other parameters such as $\mu_i$'s. For example sampling of $\alpha$ was described by Escobar and West (1995). See also Liu (1996) for an alternative approach based on sequential imputation. Sampling of other hyperparameters is typically straightforward, because, as in Step 2 and 3, conditioning on the configuration $S$ reduces the problem to a conventional hierarchical model. The configurations $S$'s, resulted from successive implementations of the algorithm, give the equality and inequality relation among the $\tau_i$'s which correspond to the hypotheses for the MCP of $k$ normal variances. To estimate the posterior probability of a hypothesis $H_r$ from a large number $(L)$ of Gibbs samples, use

$$P(H_r|\mathbf{X}) = \frac{1}{L}\sum_{\ell=1}^{L}\delta_{S_\ell}(H_r), \tag{3.1}$$

where $\delta_{S_\ell}(H_r)$ denotes unit point mass for the case where $\ell$th draw of $S$, i.e. $S_\ell$, corresponds to $H_r$. For the paired comparison among $\tau_1, \tau_2, \ldots, \tau_k$, we may use the approach to the MCP, suggested by Berry (1988), as an alternative to classical procedures: The probability of equality for any two $\tau$'s (equivalently any two $\sigma^2$'s) can be calculated from the posterior distributions on hypotheses, $P(H_r|\mathbf{X})$, $r = 1, \ldots, N$. This can be achieved by adding probabilities of those hypotheses in which the two $\tau$'s are equal. That is

$$P(\tau_i=\tau_j|\mathbf{X}) = \frac{1}{L}\sum_{\ell=1}^{L}\delta_{S_\ell}(\tau_i=\tau_j) = \sum_{r=1}^{N}P(H_r|\mathbf{X})\delta_{H_r}(\tau_i=\tau_j), \quad i\neq j, \tag{3.2}$$

where $\delta_{S_\ell}(\tau_i = \tau_j)$ and $\delta_{H_r}(\tau_i = \tau_j)$ denote unite point mass for the case where $S_\ell$ and $H_r$ indicate $\tau_i = \tau_j$, respectively.

Convergence of the foregoing Gibbs sampling scheme can be easily established as follows. Conditioning on the configuration $S$ and $p$, the model in this article reduces to the standard normal/gamma hierarchical model. Because the number of configurations $S$'s is finite, proofs for the consistency of Markov chain Monte Carlo estimates of the posterior $p(\Omega, \theta, \beta|\mathbf{X})$ and the posterior probability of $H_r$ would be a simple extension of the proofs for the standard normal/gamma hierarchical model. For the extension, we may use proofs similar to those contained in Theorem 2 through Theorem 5 of Escobar and West (1995). The extension shows that the above algorithm converges to a true posterior distribution for almost all starting values.

## 4. ILLUSTRATIVE EXAMPLES

We consider two examples illustrating the suggested MCP method for $k$ normal variances. For the first example we use an artificial data set in order to examine the performance of the Bayesian MCP method. Then a real data set is applied for the second example. In both cases, via SAS/IML, we checked the Gibbs sampler for convergence using parallel chains, as suggested by Gelman and Rubin (1992) and Cowles and Carlin (1996). Convergence was achieved using 8 parallel chains with 1,000 burn-in iterations. We stored 1,500 iterations after burn-in.

We used a number of values for $\alpha$ to help show the sensitivity of different prior probabilities on the hypotheses $H_0$ and $H_N$. The value of $\alpha$ are obtained from the formula of the prior probability of each hypothesis in Proposition 1. Thus, for given $k$ and $P(H_0)$, we calculate corresponding value of $\alpha$ form (2.1) and, in turn, we obtain $P(H_N)$ from (2.1) using the value of $\alpha$.

### 4.1. An Artificial Data Example

To highlight the performance of the Bayesian MCP method, we consider the following case for each set of five univariate normal distribution parameters: (i) Population 1 : $N(0,1)$; (ii) Population 2 : $N(2,5)$; (iii) Population 3: $N(-2,5)$; (iv) Population 4: $N(3,10)$; (v) Population 5 : $N(-3,10)$, so that true hypothesis is $H_{(true)} : \tau_1 \neq \tau_2 = \tau_3 \neq \tau_4 = \tau_5$. Under the five distributions artificial data set $\{X_1, X_2, X_3, X_4, X_5\}$ were generated and the Bayesian MCP method was applied. For the Bayesian MCP method, we assume that each $\tau_i,\ i = 1, \ldots, 5$ follows in priori $Gamma(.05, .05)$ to reflect vagueness of the prior knowledge and take $\mu_1 = -1$, $\mu_2 = 1$, $\mu_3 = -1$, $\mu_4 = 1$, $\mu_5 = -1$, $\phi = .05$, and $\psi = .05$.

Table 1, summarizing the result of the simulation with $n_1 =, \dots, = n_5 = 20$, presents posterior probabilities for each 18 possible hypotheses (remaining 34 hypotheses, having zero posterior probabilities, are deleted from the table). As expected, the hypothesis $H_{(true)} : \tau_1 \neq \tau_2 = \tau_3 \neq \tau_4 = \tau_5$ attains large posterior probability (ranging from .423 to .706) compared to that of the other hypotheses. This suggests that the data lend the greatest support to two equal pairs $(\tau_2, \tau_3)$ and $(\tau_4, \tau_5)$ with $\tau_1$ being different from others. Thus this example shows good performance of the Bayesian MCP method for several normal variances.

Table 2 gives the posterior probabilities for equality of pairs of precisions. We see that the highly plausible equality pairs $(\tau_2, \tau_3)$ and $(\tau_4, \tau_5)$ exactly match the design of our simulation. As is evident in this table, posterior probabilities can depend greatly on $\alpha$, and thus careful assessment of $\alpha$ is important.

**Table 1. Directly Calculated $P(H_0)$ and $P(H_N)$ and Posterior Probabilities**

| | | | | $\alpha$ | | | |
|---|---|---|---|---|---|---|---|
| | .372 | .831 | 1.595 | 2.213 | 4.588 | 9.996 | 13.00 |
| | | | Prior | | Prob. | | |
| $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5$ | .500 | .250 | .100 | ..055 | .010 | .001 | <.001 |
| $H_N : \tau_1 \neq \tau_2 \neq \tau_3 \neq \tau_4 \neq \tau_5$ | <.001 | .005 | .026 | .055 | .184 | .416 | .500 |
| Hypothesis | | | Posterior | | Prob. | | |
| $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5$ | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| $\tau_1 = \tau_2 = \tau_3 \neq \tau_4 \neq \tau_5$ | .013 | .015 | .014 | .012 | .005 | .003 | .004 |
| $\tau_1 = \tau_2 = \tau_3 \neq \tau_4 = \tau_5$ | .320 | .178 | .100 | .063 | .016 | .006 | .007 |
| $\tau_1 \neq \tau_2 = \tau_3 \neq \tau_4 \neq \tau_5$ | .017 | .053 | .101 | .115 | .075 | .112 | .221 |
| $\tau_1 \neq \tau_2 = \tau_3 \neq \tau_4 = \tau_5$ | .590 | .616 | .706 | .664 | .508 | .457 | .423 |
| $\tau_1 \neq \tau_2 \neq \tau_3 \neq \tau_4 = \tau_5$ | .036 | .110 | .046 | .132 | .223 | .377 | .291 |
| $\tau_1 \neq \tau_2 \neq \tau_3 = \tau_5 \neq \tau_4$ | .003 | .005 | .005 | .002 | .007 | .001 | .005 |
| $\tau_1 \neq \tau_2 = \tau_4 \neq \tau_3 \neq \tau_5$ | .000 | .000 | .000 | .000 | .000 | .000 | .001 |
| $\tau_1 \neq \tau_2 = \tau_5 \neq \tau_3 = \tau_4$ | .000 | .000 | .000 | .000 | .001 | .005 | .001 |
| $\tau_1 \neq \tau_2 = \tau_3 = \tau_5 \neq \tau_4$ | .001 | .001 | .002 | .002 | .001 | .001 | .003 |
| $\tau_1 = \tau_2 \neq \tau_3 \neq \tau_4 = \tau_5$ | .001 | .004 | .002 | .001 | .003 | .002 | .002 |
| $\tau_1 = \tau_3 \neq \tau_2 \neq \tau_4 \neq \tau_5$ | .002 | .000 | .004 | .000 | .004 | .002 | .002 |
| $\tau_1 = \tau_3 \neq \tau_2 \neq \tau_4 = \tau_5$ | .013 | .020 | .011 | .009 | .010 | .005 | .005 |
| $\tau_1 = \tau_5 \neq \tau_2 \neq \tau_3 \neq \tau_4$ | .000 | .000 | .000 | .002 | .001 | .001 | .001 |
| $\tau_1 = \tau_5 \neq \tau_2 = \tau_3 \neq \tau_4$ | .001 | .000 | .000 | .000 | .002 | .001 | .001 |
| $\tau_1 = \tau_3 = \tau_5 \neq \tau_2 \neq \tau_4$ | .000 | .001 | .001 | .000 | .001 | .002 | .002 |
| $\tau_1 \neq \tau_2 \neq \tau_3 \neq \tau_4 \neq \tau_5$ | .007 | .001 | .011 | .003 | .147 | .031 | .041 |

**Table 2. Pairwise Posterior Probabilities**

| | $\alpha$ | | | | | | |
|---|---|---|---|---|---|---|---|
| Pair | .372 | .831 | 1.595 | 2.213 | 4.588 | 9.996 | 13.00 |
| $(\tau_1, \tau_2)$ | .333 | .193 | .115 | .075 | .024 | .009 | .011 |
| $(\tau_1, \tau_3)$ | .346 | .211 | .129 | .082 | .038 | .014 | .016 |
| $(\tau_1, \tau_4)$ | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| $(\tau_1, \tau_5)$ | .001 | .001 | .001 | .000 | .001 | .001 | .001 |
| $(\tau_2, \tau_3)$ | .941 | .861 | .923 | .855 | .614 | .585 | .657 |
| $(\tau_2, \tau_4)$ | .000 | .000 | .000 | .000 | .000 | .001 | .001 |
| $(\tau_2, \tau_5)$ | .001 | .001 | .002 | .002 | .002 | .006 | .003 |
| $(\tau_3, \tau_4)$ | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| $(\tau_3, \tau_5)$ | .003 | .001 | .009 | .004 | .004 | .006 | .008 |
| $(\tau_4, \tau_5)$ | .957 | .926 | .865 | .866 | .766 | .842 | .726 |

### 4.2. A Real Data Example

Steel and Torrie (1981) reported an experiment measuring nitrogen content in milligrams of red clover plants innoculated with cultures of *Rhizobium trifolli* plus a composite of five *Rhizobium meliloti* strains. The treatments were each of five red clover cultures *R. trifolli* tested individually with a composite of five alfalfa strains (Treatment 1, ..., Treatment 5), *R. meliloti,* and a composite of red clover strains also tested with a composite of the alfalfa strains, making six in all. The experiment was conducted in a greenhouse using completely randomized design with five pots per treatment. The objective is to compare variances of the nitrogen for different treatments. Table 3 gives the data.

**Table 3. Rhizobium Data**

| Treatments | treat. 1 | treat. 2 | treat. 3 | treat. 4 | treat. 5 | composite |
|---|---|---|---|---|---|---|
| | 14.3 | 17.0 | 20.7 | 17.7 | 19.4 | 17.3 |
| | 14.4 | 19.4 | 21.0 | 24.8 | 32.6 | 19.4 |
| | 11.8 | 9.1 | 20.5 | 27.9 | 27.0 | 19.1 |
| | 11.6 | 11.9 | 18.8 | 25.2 | 32.1 | 16.9 |
| | 14.2 | 15.8 | 18.6 | 24.3 | 33.0 | 20.8 |
| Mean | 13.26 | 14.64 | 19.92 | 23.98 | 28.82 | 18.70 |
| SD | 1.43 | 4.12 | 1.13 | 1.60 | 3.78 | 5.80 |

We assume that each $\tau_i$, $i = 1, \ldots, 6$, follows in priori $Gamma(.05, .05)$ to reflect vagueness of the prior knowledge. We take $\mu_1 = 14$, $\mu_2 = 15$, $\mu_3 = 20$, $\mu_4 = 24$, $\mu_5 = 29$, $\mu_6 = 19$, $\phi = .05$, and $\psi = .05$. The number of possible hypotheses for the MCP is 203. Thus, to save the space, we notes the prior probabilities of $H_0$ and $H_N$ and posterior probabilities for five highly plausible hypotheses across different values of $\alpha$ in Table 4.

**Table 4. Posterior Probabilities of Five Highly Plausible Hypotheses**

| | | | | $\alpha$ | | | |
|---|---|---|---|---|---|---|---|
| | .334 | 1.373 | 1.956 | 2.605 | 3.462 | 4.909 | 19.88 |
| | | | Prior | | Prob. | | |
| $H_0$ | .500 | .100 | .050 | .026 | .012 | .004 | <.001 |
| $H_N$ | <.001 | .004 | .012 | .026 | .050 | .100 | .500 |
| Hypothesis | | | Posterior | | Prob. | | |
| $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \tau_6$ | .739 | .459 | .340 | .234 | .192 | .157 | .021 |
| $\tau_1 = \tau_2 = \tau_3 = \tau_5 = \tau_6 \neq \tau_4$ | .015 | .035 | .020 | .024 | .023 | .033 | .033 |
| $\tau_1 = \tau_2 = \tau_5 = \tau_6 \neq \tau_3 = \tau_4$ | .012 | .021 | .018 | .032 | .027 | .041 | .023 |
| $\tau_1 = \tau_3 = \tau_4 \neq \tau_2 = \tau_5 = \tau_6$ | .168 | .294 | .410 | .440 | .494 | .457 | .262 |
| $\tau_1 = \tau_4 \neq \tau_2 = \tau_3 = \tau_5 = \tau_6$ | .021 | .028 | .055 | .040 | .049 | .046 | .029 |

**Table 5. Pairwise Posterior Probabilities**

| | | | | $\alpha$ | | | |
|---|---|---|---|---|---|---|---|
| Pair | .334 | 1.373 | 1.956 | 2.605 | 3.462 | 4.909 | 19.88 |
| $(\tau_1, \tau_2)$ | .786 | .563 | .434 | .361 | .301 | .279 | .125 |
| $(\tau_1, \tau_3)$ | .950 | .884 | .868 | .828 | .832 | .781 | .638 |
| $(\tau_1, \tau_4)$ | .957 | .894 | .894 | .848 | .851 | .786 | .638 |
| $(\tau_1, \tau_5)$ | .792 | .572 | .442 | .376 | .310 | .302 | .144 |
| $(\tau_1, \tau_6)$ | .775 | .544 | .408 | .332 | .280 | .250 | .113 |
| $(\tau_2, \tau_3)$ | .792 | .592 | .476 | .401 | .328 | .295 | .168 |
| $(\tau_2, \tau_4)$ | .768 | .536 | .394 | .315 | .256 | .209 | .069 |
| $(\tau_2, \tau_5)$ | .982 | .955 | .940 | .916 | .928 | .900 | .736 |
| $(\tau_2, \tau_6)$ | .891 | .932 | .930 | .904 | .912 | .887 | .708 |
| $(\tau_3, \tau_4)$ | .947 | .878 | .866 | .831 | .841 | .792 | .630 |
| $(\tau_3, \tau_5)$ | .802 | .601 | .490 | .418 | .343 | .316 | .192 |
| $(\tau_3, \tau_6)$ | .786 | .563 | .454 | .370 | .300 | .270 | .138 |
| $(\tau_4, \tau_5)$ | .774 | .540 | .408 | .336 | .267 | .228 | .071 |
| $(\tau_4, \tau_6)$ | .758 | .510 | .372 | .284 | .233 | .181 | .041 |
| $(\tau_5, \tau_6)$ | .977 | .936 | .923 | .908 | .910 | .878 | .732 |

Table 5 gives the pairwise posterior probabilities for all pairs. In case we take a natural choice, i.e. the same prior probabilities of $H_0$ and $H_N$ with $\alpha = 2.605$, Table 4 notes that the hypothesis $H : \tau_1 = \tau_3 = \tau_4 \neq \tau_2 = \tau_5 = \tau_6$ is the most plausible model among 203 hypotheses. Table 5 also gives the same implication that each of precision pairs $(\tau_1, \tau_4)$, $(\tau_2, \tau_5)$, $(\tau_2, \tau_6)$, and $(\tau_5, \tau_6)$ is equal with high posterior probability. The tables notes that the smaller value of $\alpha$ leads to the larger probability of homogeneity among 6 variances and vice versa. This sensitiveness of the Bayesian MCP method is also revealed in Table 1 and Table 2.

## 5. CONCLUDING REMARKS

We have considered the problem of developing a Bayesian multiple comparison for variances (or precisions) of $k$ normal populations with unequal means. As an alternative to a formal Bayesian analysis of a mixture model that usually leads to intractable calculations, the Dirichlet prior process is used to provide a nonparametric Bayesian method for obtaining posterior probabilities for various hypotheses of equality among population variances. Finding posterior distributions is analytically intractable so that we solve the computational difficulty by developing a Gibbs sampler algorithm. As usual, the suggested Bayesian method allows for direct probability calculations of hypotheses of equality and inequality among population variances. The method that we propose has some flexibility in assignment of prior probabilities, because we can easily assess them via equation (2.1) with given any number of populations $k$ and $P(H_0)$. If we don't have information about $P(H_0)$, we may set up more complex model which require another portion to the hierarchy that allows the introduction of distributions on the hyperparameters $\alpha$.

An extension of the method to the multiple comparison problem for the multivariate normal populations would be accomplished straightforwardly. 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.

## ACKNOWLEDGEMENTS

The author would like to thank the editor and the refrees whose comments helped him a great deal in the revision. This paper was supported by a grant from Dongguk University, Korea.

## References

Antoniak, C. E. (1974). Mixtures of Dirichlet processes with application to Bayesian nonparametric problems, *The Annals of Statistics,* Vol. 2, 1152-1174.

Berry, D. A. (1988). Multiple comparisons, multiple tests and data dredging: a Bayesian perspective, *Bayesian Statistics 3,* Eds. by Bernardo, J. M. et al., Oxford University Press.

Chen, J. and Gupta, A. K. (1997). Testing and locating variance change points with application to stock prices, *Journal of the American Statistical Association,* Vol. 92, 739-747.

Cowles, M. K. and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review, *Journal of the American Statistical Association,* 91, 883-904.

Escorbar, M. D. and West, M. (1995). Bayesian density estimation and inferences using mixtures, *Journal of the American Statistical Association,* Vol. 90, 577-588.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems, *The Annals of Statistics,* Vol. 1, 209-230.

Gopalan, R. and Berry, D. A. (1998). Bayesian multiple comparisons using Dirichlet process prior, *Journal of the American Statistical Association,* Vol. 93, 1130-1139.

Hochberg, T. and Tamhane, A. C. (1987). *Multiple Comparison Procedures,* New York: Wiley.

Hsu, D. A. (1977). Tests of variance shifts at an unknown time point, *Applied Statistics,* Vol. 26, 179-184.

Inclán, C. (1993). Detection of multiple changes of variance using posterior odds, *Journal of Business and Economic Statistics,* Vol. 11, 189-300.

Liu, J. (1996). Nonparametric hierarchical Bayes via sequential imputations, *The Annals of Statistics,* Vol. 24, 911-930.

MacEachern, S. N. and Müller, P. (1998). Estimating mixture Dirichlet process models, *Journal of Computational and Graphical Statistics,* Vol. 7, 223-238.

Vostrikova, L. J. (1981). Detecting disorder in multidimensional random processes, *Soviet Mathematics Doklady,* Vol. 24, 55-59.