The Korean Communications in Statistics
Vol. 13 No. 1, 2006, pp. 39-47

A Bayesian Multiple Testing of Detecting Differentially Expressed Genes in Two-sample Comparison Problem

Hyun Sook Oh[1] and Wan Youn Yang[2]

Abstract

The Bayesian approach to multiple testing procedure for one sample testing problem proposed by Scott and Berger (2003) is extended to two-sample comparison problem in microarray experiments. The prior distribution of each gene's mean for one sample is given conditionally on the corresponding gene's mean for the other sample. Posterior distributions of interesting parameters are derived and estimated based on an importance sampling method. A simulated example is given for illustration.

Keywords : microarray experiment; Bayesian multiple testing; importance sampling method.

1. Introduction

Microarray technologies enable us to simultaneously measure the expression levels of thousands of genes in a biological sample. They are being applied increasingly in biological and medical research for a various kind of problems.

Many statistical approaches have been proposed to analyse the data from microarray experiments (Kerr and Churchill (2000), Tusher et al. (2001), Efron et al. (2001), Yang et al. (2001), Dudoit et al. (2002)). An important and common task in such analyses is to detect genes with differential expression under two experimental conditions such as treatment/control status, two types of tissues, two drug types, etc.

Since a typical microarray experiment measures expression levels for thousands of genes simultaneously, we are faced with an extreme multiple testing problem. There has been many researches on such multiple testing procedures.

In the beginning Westfall and Young (1993) proposed resampling-based $p-value$ adjustment procedure from the typical two-sample t-test statistics to control the family-wise error rate. However, most of multiple testing procedures including Westfall and Young (1993) based on the two-sample t-test are highly relevant to microarray experiments. Also, it is too conservative to control the family-wise error rate in microarray experiments.

The false discovery rate(FDR) developed by Benjamini and Hochberg (1995) has emerged in

1) Associate Professor, Department of Applied Statistics, Kyungwon University, Sungnam 461-701, Korea. Correspondence : hoh@kyungwon.ac.kr
2) Associate Professor, Department of Applied Statistics, Kyungwon University, Sungnam 461-701, Korea.

the context of multiple testing as a practical object to be controlled, as opposed to the family-wise error rate. The FDR was further developed by Storey (2002) with a new concept called positive false discovery rate, which has a Bayesian motivation. Recently multiple testing procedures to control the FDR has been issued mostly in microarray experiments.

On the other hand Scott and Berger (2003) provided a one-sample Bayesian multiple testing procedure in general sense in which the multiple testing is controlled by prior specifications without using the FDR. In this procedure there is a common prior probability that each individual mean is zero and let the data themselves choose this common prior probability under a random effects model or Bayesian hierarchical model for each mean.

The procedure developed by Scott and Berger (2003) is extended to two-sample comparison problem in microarray experiments in the present paper. Each pair of comparing means are different or equal according to their gene responses so the priors for each pair of comparing means are specified conditionally. Multiple testing problem is controlled by the common prior probability that each pair of individual means are equal and let the data themselves choose this common prior probability. In Section 2, the two-comparison model is specified with likelihood function and priors. The posteriors which is our main interests are derived in Section 3. An illustrative simulations are described in Section 4. Finally, Section 5 summarizes and discusses about our procedure.

2. Two-sample comparison model

2.1 The model

We consider a generic situation that for each gene i, $i=1,\cdots,N$, we have expression levels $X_{1i},\cdots,X_{ni}$ from n microarrays under condition 1 and $Y_{1i},\cdots,Y_{ni}$ from n arrays under condition 2.

The goal is to identify genes that $\{X_{1i},\cdots,X_{ni}\}$ and $\{Y_{1i},\cdots,Y_{ni}\}$ have different means. It is assumed X_{ji} and $Y_{j'i}$, $j,j'=1,\cdots n$, are independently $N(X_{ji}\mid\mu_{1i},\sigma^2)$ and $N(Y_{j'i}\mid\mu_{2i},\sigma^2)$, respectively for each i with unknown variance σ^2. It is desired to determine which of the differences, $\mu_{1i}-\mu_{2i}$, are nonzero. Since $\overline{X}_i=\sum_{j=1}^{n}X_{ji}/n$ and $\overline{Y}_i=\sum_{j=1}^{n}Y_{ji}/n$ are sufficient statistics of μ_{1i} and μ_{2i}, they will be used in our testing and for the notational convenience X_i and Y_i will replace them respectively. Also, the variance of $\overline{X}_i$(or $\overline{Y}_i$) is σ^2/n but σ^2 will be used instead.

The model is specified by defining a model index parameter γ to be an N-dimensional vector of 0's and 1's such that

$$\gamma_i=\begin{cases}0 & \text{if } \mu_{1i}=\mu_{2i}\\ 1 & \text{if } \mu_{1i}\neq\mu_{2i}\end{cases}.$$

Then the full likelihood function can be written as

$$f(x,y|\mu_1, \mu_2, \sigma^2, \gamma) = \Pi_{i=1}^{N} \frac{1}{(2\pi\sigma^2)} \exp\left[-\frac{1}{2\sigma^2}\{(x_i - \gamma_i\mu_{1i} - (1-\gamma_i)\mu_{2i})^2 + (y_i - \mu_{2i})^2\} \right] ,$$

where $x = \{x_i, i = 1, \cdots, N\}$, $y = \{y_i, i = 1, \cdots, N\}$, $\mu_1 = \{\mu_{1i}, i = 1, \cdots, N\}$, $\mu_2 = \{\mu_{2i}, i = 1, \cdots, N\}$ and $\gamma = \{\gamma_i, i = 1, \cdots, N\}$.

2.2 Priors

Let us consider the conditional prior distribution of μ_{1i} given μ_{2i} to be as follows;

$$\pi(\mu_{1i}|\mu_{2i}) = \begin{cases} p & \text{if } \mu_{1i} = \mu_{2i} \\ (1-p)N(\mu_{1i}|\mu_{2i}, V) & \text{if } \mu_{1i} \neq \mu_{2i} \end{cases}$$

That is, it has point mass at μ_{2i} and normally distributed centered at μ_{2i} elsewhere. The conjugate prior distribution of μ_{2i} is considered as normal distribution, $N(\mu_{2i}|0, V)$, where V is the unknown hyper parameter.

Thus unknown parameters are $\Theta = (\mu_1, \mu_2, \gamma, p, V, \sigma^2)$ and it is easily verified that

$$\pi(\gamma|p) = \Pi_{i=1}^{N} p^{1-\gamma_i}(1-p)^{\gamma_i} .$$

Then the prior of Θ is

$$\begin{aligned} \pi(\Theta) &= \pi(\mu_1, \mu_2|\gamma, p, V)\, \pi(\gamma|p)\, \pi(p, V, \sigma^2) \\ &= \Pi_{i=1}^{N} \pi(\mu_{1i}|\mu_{2i}, \gamma_i, p, V)\, \pi(\mu_{2i}|\gamma_i, p, V)\, \pi(\gamma_i|p)\, \pi(p, V, \sigma^2) \\ &= \Pi_{i=1}^{N} (N(\mu_{1i}|\mu_{2i}, V))^{\gamma_i} N(\mu_{2i}|0, V) p^{1-\gamma_i}(1-p)^{\gamma_i} \pi(p, V, \sigma^2) . \end{aligned}$$

For prior distribution of $\pi(p, V, \sigma^2)$, let us follow the procedure in Scott and Berger (2003). First, it is assumed that p and (V, σ^2) are independent. For $\pi(p)$, one does have strong prior information about p that p is large($p \approx 1$) in microarray experiments in general. Scott and Berger (2003) suggested a convenient functional form that represent this type of information such as $\pi(p) = (\alpha + 1)p^{\alpha}$, where α is an adjustable parameter allowing one to control how much of $\pi(p)$'s mass is concentrated near 1.

Priors on V and σ^2 are given as

$$\begin{aligned} \pi(V, \sigma^2) &= \pi(V|\sigma^2)\, \pi(\sigma^2) \\ &= \frac{1}{\sigma^2}(1 + V/\sigma^2)^{-2} \frac{1}{\sigma^2} \\ &= \frac{1}{(V + \sigma^2)^2} . \end{aligned}$$

For the hyper parameter V a vague proper prior $(1 + V/\sigma^2)^{-2}/\sigma^2$ is considered conditioned on σ^2 and $\pi(\sigma^2) = 1/\sigma^2$ is a well-known Jeffrey's noninformative prior.

3. Posteriors

Under the modeling assumptions described in section 2, the posterior density of

$\Theta = (\mu_1, \mu_2, \gamma, p, V, \sigma^2)$ given (x,y) is

$$\pi(\Theta | x, y) \propto f(x, y | \mu_1, \mu_2, \gamma, \sigma^2)\, \pi(\Theta) \quad . \tag{1}$$

Lemma 1. The posterior density in (1) is proper.

Proof. Let $\rho = \sigma^2 V/(\sigma^2 + V)$. Then it is easy to show that

$$\sum_{\gamma_i=0}^{1} \int\int f(x_i, y_i | \mu_{1i}, \mu_{2i}, \sigma^2, \gamma_i) \mathrm{N}(\mu_{1i} | \mu_{2i}, \mathrm{V})^{\gamma_i} \mathrm{N}(\mu_{2i} | 0, \mathrm{V}) \mathrm{d}\mu_{1i}\, \mathrm{d}\mu_{2i}$$

$$= N(y_i | 0, \sigma^2 + V) \left\{ p N\left(x_i | \frac{V y_i}{\sigma^2 + V},\ \sigma^2 + \rho\right) + (1-p) N\left(x_i | \frac{V y_i}{\sigma^2 + V},\ \sigma^2 + V + \rho\right) \right\}. \tag{2}$$

Now, to prove our lemma it is sufficient to show that

$$\int_0^1 \int_0^\infty \int_0^\infty \prod_{i=1}^{N} (2) \times \pi(V, \sigma^2)\, \pi(p)\, dV d\sigma^2 dp \tag{3}$$

is finite. Let $u = 1/(\sigma^2 + V)$. Then by change of variables,

$$(3) = K \int_0^1 \int_0^\infty \int_0^{\sigma^{-2}} \prod_{i=1}^{N} \sqrt{u}\, e^{-\frac{u y_i^2}{2}} g_i(p, u, \sigma^2)\, \pi(p) du\, d\sigma^2 dp$$

for some constant K, where

$$g_i(p,u,\sigma^2) = \frac{p}{\sigma\sqrt{2-u\sigma^2}} \exp\left(- \frac{(x_i - (1-u\sigma^2) y_i)^2}{2\sigma^2(2-u\sigma^2)} \right)$$

$$+ \frac{1-p}{\sqrt{u + \sigma^2(1-u\sigma^2)}} \exp\left(- \frac{(x_i - (1-u\sigma^2) y_i)^2}{u + \sigma^2(1-u\sigma^2)} \right).$$

Since $0 < u < 1/\sigma^2$, for all p, u, σ^2, $g_i(p, u, \sigma^2) < c$ for some constant c. Thus

$$(3) < K c^N \int_0^1 \int_0^\infty \int_0^{\sigma^{-2}} u^{N/2} e^{-u|y|^2/2}\, \pi(p)\, du\, d\sigma^2 dp$$

$$= K c^N \int_0^1 \pi(p)\, dp \int_0^\infty \int_0^{\frac{1}{u}} u^{N/2} e^{-u|y|^2/2}\, d\sigma^2 du < \infty,$$

where $|y|^2 = \sum_{i=1}^{N} y_i^2$. □

Now, the main goal of the problem is to derive the posterior probability of γ_i(or $1-\gamma_i$) for each gene i. Let

$$P_i = P(\gamma_i = 0 | x, y).$$

Then

$$P_i = \int P(\gamma_i = 0 | x_i, y_i, p, V, \sigma^2)\, \pi(p, V, \sigma^2 | x_i, y_i)\, dp\, dV d\sigma^2 .$$

For given p, V, σ^2,

$$P(\gamma_i = 0 | x_i, y_i) = \frac{f(x_i, y_i | \gamma_i = 0)\, \pi(\gamma_i = 0)}{f(x_i, y_i | \gamma_i = 0)\, \pi(\gamma_i = 0) + m_1(x_i, y_i | \gamma_i = 1)\, \pi(\gamma_i = 1)}, \tag{4}$$

where

$$m_1(x_i, y_i|\gamma_i = 1) = \int_{\mu_{1i} \neq \mu_{2i}} f(x_i, y_i|\mu_{1i}, \mu_{2i})\,\pi(\mu_{1i}, \mu_{2i}) \ .$$

Since X_i and Y_i are independent and each is normally distributed with mean μ_{2i} and variance σ^2 for given $\gamma_i = 0$ and the prior distribution of μ_{2i} is $N(\mu_{2i}|0, V)$, in the formula (4),

$$f(x_i, y_i|\gamma_i = 0) = N(y_i|0, \sigma^2 + V)\, N\left(x_i|\frac{Vy_i}{\sigma^2 + V},\ \sigma^2 + \rho\right),$$

where ρ is defined as in the proof of the above lemma.

When $\gamma_i = 1$, X_i and Y_i are independent and normally distributed with mean μ_{1i} and μ_{2i}, respectively. Combining the prior distribution of μ_{1i} and μ_{2i} defined in section 2.2,

$$m_1(x_i, y_i|\gamma_i = 1) = N(y_i|\ \sigma^2 + V)\, N\left(x_i|\frac{Vy_i}{\sigma^2 + V},\ \sigma^2 + V + \rho\right). \tag{5}$$

Thus

$$P(\gamma_i = 0|x, y, p, V, \sigma^2)$$
$$= \left[1 + \frac{1-p}{p}\sqrt{\frac{\sigma^2 + V + \rho}{\sigma^2 + \rho}}\ \exp\left\{\frac{1}{2}\left(x_i - \frac{Vy_i}{\sigma^2 + V}\right)^2\left(\frac{1}{\sigma^2 + \rho} - \frac{1}{\sigma^2 + V + \rho}\right)\right\}\right]^{-1}. \tag{6}$$

The joint posterior distribution of μ_{1i} and μ_{2i} given p, V, σ^2 when $\gamma_i = 1$ can be derived as follows; For given p, V, σ^2,

$$\pi(\mu_{1i}, \mu_{2i}|\gamma_i = 1, x, y) = \frac{f(x_i, y_i|\mu_{1i}, \mu_{2i}, \gamma_i = 1)\pi(\mu_{1i}, \mu_{2i}|\gamma_i = 1)}{m_1(x_i, y_i|\gamma_{i=1})}$$
$$= N\left(\mu_{1i}|\frac{\sigma^2\mu_{2i} + Vx_i}{\sigma^2 + V},\ \rho\right) N\left(\mu_{2i}|\frac{Vy_i + \rho x_i}{\sigma^2 + V + \rho},\ \frac{\sigma^2 V}{\sigma^2 + V + \rho}\right). \tag{7}$$

Since (6) and (7) are the posterior distributions of $\mu_{1i} = \mu_{2i}$ and (μ_{1i}, μ_{2i}) conditioned on p, V, σ^2, we need to integrate out these conditioned parameters by multiplying of $\pi(p, V, \sigma^2|x, y)$ to each (6) and (7). Let us derive $\pi(p, V, \sigma^2|x, y)$.

$$\pi(p, V, \sigma^2|x, y) = \int \pi(\mu_1, \mu_2, \gamma, p, V, \sigma^2|x, y)\, d\mu_1\, d\mu_2\, d\gamma$$
$$\propto \int \Pi_{i=1}^{N}[N(x_i|\gamma_i\mu_{1i} + (1-\gamma_i)\mu_{2i}, \sigma^2)N(y_i|\mu_{2i}, \sigma^2)(N(\mu_{1i}|\mu_{2i}, V))^{\gamma_i}$$
$$\times N(\mu_{2i}|0, V)p^{1-\gamma_i}(1-p)^{\gamma_i}]\pi(p, V, \sigma^2)d\mu_1 d\mu_2 d\gamma$$

Hence

$$\pi(p, V, \sigma^2|x, y) = C^{-1}\Pi_{i=1}^{N}\left[p\,N(y_i|0, \sigma^2 + V)\,N\left(x_i\frac{Vy_i}{\sigma^2 + V}, \sigma^2 + \rho\right)\right.$$
$$\left. + (1-p)\,N(y_i|0, \sigma^2 + V)N\left(x_i|\frac{Vy_i}{\sigma^2 + V},\ \sigma^2 + V + \rho\right)\right]\pi(p, V, \sigma^2),$$

where C^{-1} is the normalizing constant.

Thus the posterior distribution of $\mu_{1i} = \mu_{2i}$ is

$$P(\gamma_i = 0 | x,y) = \int h_i(p, V,\sigma^2)\, \pi(p, V,\sigma^2 | x,y) dpd\, Vd\sigma^2 = E^{\pi(p, V,\sigma^2|x,y)}[h_i(p, V,\sigma^2)], \quad (8)$$

where $h_i(p, V,\sigma^2)$ is the formula (6) which is the posterior probability of $\gamma_i = 0$ given p, V,σ^2. Similarly,

$$\pi(\mu_{1i},\mu_{2i} | \gamma_i = 1, x, y) = \int f_i(p, V,\sigma^2)\pi(p, V,\sigma^2 | x,y) dpd\, Vd\sigma^2 = E^{\pi(p, V,\sigma^2|x,y)}[f_i(p, V,\sigma^2)], \quad (9)$$

where $f_i(p, V,\sigma^2)$ is the formula (7) which is the posterior distribution of (μ_{1i},μ_{2i}) given p, V,σ^2 when $\mu_{1i} \neq \mu_{2i}$.

Since we have computational difficulties for deriving (8) and (9), the "Importance Sampling Method" can be used. In this case the multivariate t-distribution is generally used as the common important function and sampling is based on this given t-distribution for $i = 1,\cdots,N$.

For removing the restrictions on p, V,σ^2, let $\lambda = \log(p/1-p)$, $\xi = \log V$ and $\eta = \log\sigma^2$. Then

$$\pi^*(\lambda, \xi, \eta | x,y) = \pi((1+e^{-\lambda})^{-1}, e^{\xi}, e^{\eta} | x,y) e^{\xi+\eta-\lambda}(1+e^{-\lambda})^{-2}.$$

Random samples $\{(\lambda_k, \xi_k, \eta_k) | 1 \leq k \leq m\}$ are drawn from the t-distribution and let w_k be the ratio of the given t-distribution density function to the π^* at $(\lambda_k, \xi_k, \eta_k)$, $k = 1,\cdots,m$. Then $P_i = P(\gamma_i = 0 | x,y)$ is evaluated as the weighted mean of $h_i((1+e^{-\lambda_k})^{-2}, e^{\xi_k}, e^{\eta_k})$ with respect to w_k;

$$\hat{P}_i = \sum_{k=1}^{m} h_i((1+e^{-\lambda_k})^{-1}, e^{\xi_k}, e^{\eta_k})\, w_k / \sum_{l=1}^{m} w_l \,. \quad (10)$$

In the similar way $\pi(\mu_{1i}, \mu_{2i} | \gamma_i = 1, x, y)$ can be estimated by using f_i instead of h_i in the right side of (10).

Specifically, from the formula (7), each marginal posterior distributions of μ_{1i}, μ_{2i} can be derived and the posterior density functions of $\mu_{1i} - \mu_{2i}$ given $\gamma_i = 1$ for given p, V,σ^2 are derived as follows.

$$\begin{aligned}\pi(\mu_{1i} - \mu_{2i} | \gamma_{i=1}, p, V,\sigma^2, x, y) &= \frac{1}{\sqrt{2\pi}\sqrt{v_1 v_2 (v_3 + v_4)}} exp[-\frac{1}{2}\frac{1}{v_1 v_2 (v_3 + v_4)}(t - B/A)^2] \\ &\equiv q_i(p, V,\sigma^2), \quad say,\end{aligned}$$

where $v_1 = \rho$, $v_2 = \sigma^2 V/(\sigma^2 + V + \rho)$ $v_3 = V/\{\sigma^2(\sigma^2 + V)\}$, $B = x_i - (Vy_i + \rho x_i)/(\sigma^2 + V + \rho)$ and $A = (\sigma^2 + V)/V$.

Therefore, $\pi(\mu_{1i} - \mu_{2i} | \gamma_i = 1, x, y) = E^{\pi(p, V,\sigma^2|x,y)}[q_i(p, V,\sigma^2)]$, which can be estimated by the importance sampling method as above.

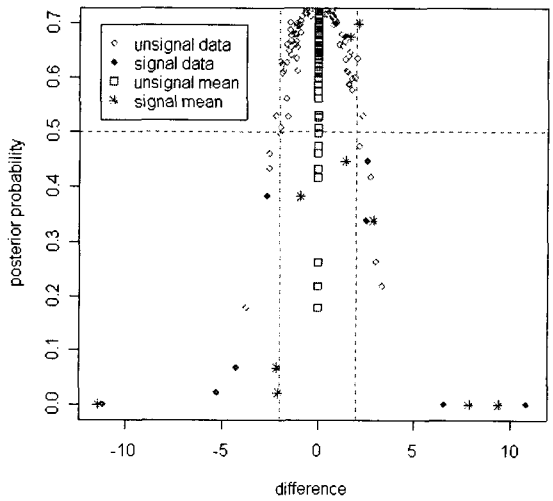
4. Simulations

For y_i, $i = 1,\cdots,110$, 110 samples are generated based on $N(\mu_{2i}, 1)$ where μ_{2i}'s are random samples from $N(0,9)$. Then the each value of x_i is generated from $N(\mu_{2i}, 1)$ for $i = 1,\cdots,100$ and the other 10 values of x_i are from $N(\mu_{1i}, 1)$ where μ_{1i} is a random sample from

$N(\mu_{2i},9)$ for $i=101,\cdots,110$. Thus we have $\mu_{1i}=\mu_{2i}$ for $i=1,\cdots,100$ and $\mu_{1i}\neq\mu_{2i}$ for $i=101,\cdots,110$ in which 7% of $\{\mu_{1i}-\mu_{2i}\}$ are non zero.

With this simulated example, the posterior probabilities of $\mu_{1i}=\mu_{2i}$ for $i=1,\cdots 110$ are estimated by the importance sampling method described in the previous section. For the multivariate t-distribution of $(p,\ V,\ \sigma^2)$ as the common importance distribution, the parameters of the t-distribution are selected as follows; 5 degree of freedom is chosen which ensures a heavy tailed distribution. The mean vector can be chosen subjectively. In general, the best guess of the parameters are used. For example, if $p=0.9$ would be the best guess then let the median of $\pi(p)$ be 0.9 so that $\pi(p)=5.58p^{5.58}$. The covariance matrix is chosen as big such as $5I$ so that the t-distribution is flat.

<Figure 1> shows the posterior probabilities of $\mu_{1i}=\mu_{2i}$ given x_i-y_i and $\mu_{1i}-\mu_{2i}$, respectively when $\pi(p)=11p^{10}$ is used as the prior distribution of p of which the median is about 0.94. The figure shows that 8 signal data values have lower posterior probabilities of $\mu_{1i}=\mu_{2i}$. However, 2 signal values in which the observed values are indistinguishable with the noise(unsignal) data have posterior probabilities of $\mu_{1i}=\mu_{2i}$ bigger than 0.5 even though their true mean differences are not zero.



<Figure 1> Posterior probabilities of $\mu_{1i}=\mu_{2i}$ with respect to x_i-y_i, $\mu_{1i}-\mu_{2i}$ when $\pi(p)\propto 11p^{10}$

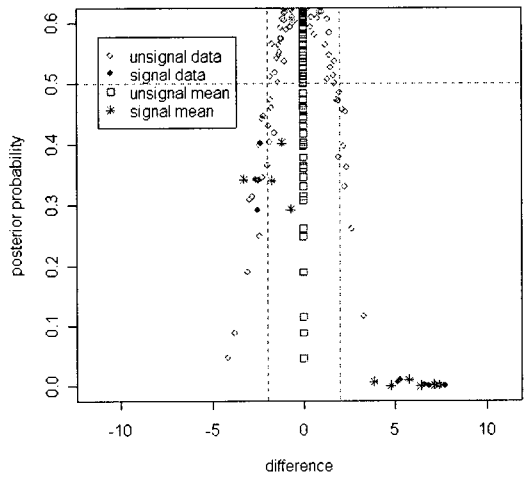
<Figure 2> is the plot of the posterior probabilities of $\mu_{1i}=\mu_{2i}$ given x_i-y_i and $\mu_{1i}-\mu_{2i}$, respectively when $\pi(p)=6p^5$ is used as the prior distribution of p of which the median is about 0.89. All 10 signal values have posterior probabilities of $\mu_{1i}=\mu_{2i}$ less than 0.4 but there are also more false positive values. If we use the uniform prior for p $(\pi(p)=1,\ 0<p<1)$, the

posterior probabilities are smaller than the above cases. Thus the posterior probabilities of $\mu_{1i}-\mu_{2i}$ depend on the prior of p.

5. Summary and Discussion

In this article, the Bayesian approach to multiple testing procedure proposed by Scott and Berger (2003) has been applied to two-sample comparison problem in microarray experiments.

The prior distribution of each gene's mean for one sample has been given conditionally on the corresponding gene's mean for the other sample. It has been assumed that few genes are changed by the different condition.



<Figure 2> Posterior probabilities of $\mu_{1i}=\mu_{2i}$ with respect to x_i-y_i, $\mu_{1i}-\mu_{2i}$ when $\pi(p)\propto 6p^5$

The multiple testing problem has been controlled by the common prior probability that each pair of individual means are equal and let the data themselves choose this common prior probability. Prior distributions for the other parameters have been given objectively(vague prior) as possible.

Posterior distributions of the parameters of interest have been derived such as the posterior probability that two means are equal. To get around the complexity of the posterior distributions, the importance sampling method has been proposed.

The proposed method has developed based on the one-sample Bayesian multiple testing procedure by Scott and Berger (2003). In two-sample comparison problem, it is possible to transform to one-sample testing problem by the pairwise comparison. However, the pairwise comparison is appropriate only when each sample data is obtained from the same experimental condition. If two sample have been applied to separate arrays, the pairwise comparison by

their differences is not appropriate.

Furthermore, in general, one sample data is for the control and the other is for the treatment so it is of interest to detect which genes are changed by the treatment sample. This motivated to use the conditional prior for the mean of each gene for the treatment sample. Also, the marginal posterior distributions of the parameters such as the posterior distribution of the mean of each gene in each sample can be derived and estimated in the proposed method.

References

[1] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A pratical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B,* Vol. 57, 289-300.

[2] Dudoit, S, Yang, Y., Callow, M., Speed, T. (2002). Statistical methods for identifying genes with differential expression in replicated cDNA microarray experiments. *Statistica Sinica,* Vol. 12, 111-139.

[3] Efron, B., Tibshirani, R., Storey, J., Tusher, V. (2000). Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association,* Vol. 96, 1151-1160.

[4] Kerr, M. and Churchill, G. (2000). Analysis of variance for gene expression microarray data. *Journal of Computational Biology,* Vol. 7, 819-837.

[5] Scott, J. and Berger, J. (2003). *An exploration of aspects of Bayesian multiple testing.* Technical Report, Institute of Statistics and Decision Sciences, Duke University, Durham, Available from: http://www.stat.duke.edu/~berger/papers/multcomp.pdf

[6] Storey, J. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B,* Vol. 64, 479-498.

[7] Tusher, V., Tibshirani, R. and Chu, C. (2001). Significance analysis of microarrays applied to transcriptional responses to ionizing radiation. *Proceedings of the National Academy of Sciences,* Vol. 98, 5116-5121.

[8] Westfall, P. and Young, S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for P-value Adjustment,* Wiley, New York.

[9] Yang, Y., Buckley, M. and Speed, T. (2001). Analysis of cDNA microarray images, *Briefings in Bioinformatics,* Vol. 2, 341-349.

[Received July 2005, Accepted December 2005]