9^{th} National Convention on Statistics (NCS)
EDSA Shangri-La Hotel
October 4-5, 2004

Bayesian Estimation Of Mixture Models

by
Cherry T. Nepomuceno

For additional information, please contact:

Author's name:	Cherry T. Nepomuceno
Designation:	Associate Professor
Agency:	Mathematics Department, College of Arts and Sciences, Western Visayas State University
Address:	Iloilo City 5000
Telefax:	(633) 320-0870 to 79
E-mail:	chenep2003@yahoo.com

Bayesian Estimation Of Mixture Models

by

Cherry T. Nepomuceno [1]

ABSTRACT

Assuming component densities of the mixture distribution are normal model parameters including the mixing weights are estimated using Bayesian procedure. Further assuming normal priors for the means, Inverse Gamma distribution for the variances and Dirichlet distribution for the weights, Gibbs sampling was employed. The two-component mixture likelihood used in the illustration resulted to conditional posterior densities that were expressed as mixtures. The model was applied to a study on mean preparation time of teachers for a one-hour lecture.

Keywords: Bayesian estimation, finite mixture models, Gibbs sampling, teaching preparation time

I. Introduction

There is an increasing need for efficient estimation of mixture distribution especially following the explosion in the use of these modeling tools in many applied fields. It is claimed that the descriptions of homogeneity or of the lack of it may be naturally addressed through mixture distributions. Alternatively, even if there is no realistic interpretation of the components, mixtures offer a very flexible modeling environment within a parametric framework. Robert and Mengersen (1999) postulated that parametric mixtures are often sufficient to model the distribution of a given data set while providing a more flexible tool for prediction purposes because of the low dimension of the parameter space. In some direct applications of finite mixture models in addition to the random sample from the mixture distribution models there may also be random samples of observations known to derive from individual underlying categories.

This paper considered finite mixture model for teaching preparation time for one lecture hour. A study conducted among faculty members in a University estimated the length of time spent on preparing for their lecture. A Bayesian approach of estimating this model was used assuming number of mixing components are known. In particular, a Markov Chain Monte Carlo (MCMC) method called Gibbs sampling was implemented. The overall aim of Gibbs sampling is to simulate from a complex (posterior) density by creating a Markov chain with the posterior density as its stationary distribution. This is done by direct successive simulations from the component conditional distributions.

This paper is organized as follows: Section 2 presents the assumptions for the mixture models and prior distributions and the derived form of the posterior distribution; conditional posterior densities where parameters were

[1] Associate Professor, Mathematics Department, College of Arts and Sciences, Western Visayas State University

drawn are presented in Section 3; using Gibbs sampling samples were drawn and used to estimate the parameters and is reported in Section 4 including the comparison of the frequentist approach and the Bayesian method applied on the teaching preparation time study; and Section 5 summarized observations from the estimation procedure employed.

II. Objectives

This study has the following objectives:

1. To model posterior distribution of parameters in a mixture model.
2. To estimate parameters of interest using Gibbs sampling.

III. The model

Let $y = \{y_i\}^n_{i=1}$ be a vector of observations whose distribution is assumed to depend on $z = \{z_i\}^n_{i=1}$ that take values in a discrete space $\{I_1,\ldots,m\}$. The model assumed for y is

$$y_i \sim \sum_{j=1}^{m} \omega_j(z_i)\phi(\bullet_i \; \mu_j \sigma_j^2) \text{ independently for } i = 1,2,\ldots,n$$

(1)

conditional on weights $\omega_j(\bullet)$, means (μ_j), and variances σ_j^2, $j = 1,\ldots m$,

where

$\phi(\bullet : \mu_j, \sigma_j^2)$ is the density of the $N(\mu_j, \sigma_j^2)$ distribution.

Let

$\mu = (\mu_j)^m_{j=1}$ and $\sigma^2 = (\sigma_j^2)^m_{j=1}$

The weights $\omega = (\omega_j)^m_{j=1}$ satisfy $\omega_j(z) \geq 0$ with $\sum_{j=1}^{m} \omega_j(z) = 1 \;\; \forall z$.

(2)

Suppose $\omega_j = \mathbf{P}(z_i = j)$ for any i.

The number of components m is given while μ, σ^2, and ω are subject to inference.

Let $x_{ij} = \begin{cases} 1 \text{ if ith unit is drawn from the jth mixture component} \\ 0 \quad \text{otherwise} \end{cases}$

Hence $\omega_j = \mathbf{P}(z_i = j) = \mathbf{P}(x_{ij} = 1)$.

The density of the vector observation y is

$$f\left(y/x,\mu,\sigma^2,\omega\right)=\prod_{i=1}^{n}\sum_{j=1}^{m}\left[\omega_j\left(x_{ij}\right)\phi\left(y_i/\mu_j,\sigma_j^{\,2}\right)\right]^{x_{ij}}$$

(3)

Assuming the following prior distribution for unknown parameters i.e. for j = 1,…,m

$$\mu_j \sim \mathrm{N}\left(\theta,\tau^2\right)$$

$$\sigma_j^{\,2} \sim \text{Inv Gamma }(\alpha,\beta)$$

$$\left(\omega_1,\ldots,\omega_m\right) \sim \text{Dirichlet }\left(\psi_1,\ldots,\psi_m\right). \text{ Let } \omega_j\left(x_{ij}\right)=\omega_j.$$

Now the joint prior can be expressed as

$$\pi\left(\mu,\sigma^2,\omega/\theta,\tau^2,\alpha,\beta,\psi\right)=\pi(\omega/\psi)\pi\left(\mu/\sigma,\tau^2\right)\pi\left(\sigma^2/\alpha,\beta\right)$$

The posterior

$$f\left(\mu,\sigma^2,\omega/y,x,\theta,\tau^2,\alpha,\beta,\psi\right)\propto\pi\left(\mu,\sigma^2,\omega/\sigma,\tau^2,\alpha,\beta,\psi\right)f\left(y/x,\mu,\sigma^2,\omega\right)$$

(4)

$$=\frac{\Gamma(\psi_1+\ldots+\psi_m)}{\Gamma(\psi_1)\ldots\Gamma(\psi_m)}\omega_1^{\,\psi_1-1}\ldots\omega_m^{\,\psi_m-1}\left[\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi\tau^2}}\exp\left\{-\frac{1}{2\tau^2}\left(\mu_j-\theta\right)^2\right\}\right]\left[\prod_{j=1}^{m}\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\sigma_j^{\,2}\right)^{-(\alpha+1)}e^{-\beta/\sigma_j^{\,2}}\right]$$

$$\bullet\sum_{j=1}^{m}\left[\omega_j\frac{1}{\sqrt{2\pi\sigma_j^{\,2}}}\exp\left\{-\frac{1}{2\sigma_j^{\,2}}\left(y_i-\mu_j\right)^2\right\}\right]^{\sum_{j=1}^{n}x_{ij}}$$

$$=\frac{\Gamma(\psi_1+\ldots+\psi_m)}{\Gamma(\psi_1)\ldots\Gamma(\psi_m)}\omega_1^{\,\psi_1-1}\ldots\omega_m^{\,\psi_m-1}\frac{1}{\left(2\pi\tau^2\right)^{m/2}}\exp\left\{-\frac{1}{2\tau^2}\sum_{j=1}^{m}\left(\mu_j-\theta\right)^2\right\}\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right)^m\left[\prod_{j^m}^{m}\left(\sigma_j^{\,m}\right)^{-(\alpha+1)}\right]e^{-\beta\sum_{j=1}^{m}\frac{1}{\sigma_j^{\,2}}}$$

$$\bullet\left[\sum_{j=1}^{m}\omega_j^{\,n_j}\frac{1}{\left(2\pi\sigma_j^{\,2}\right)^{n_j/2}}\exp\left\{-\frac{1}{2\sigma_j^{\,2}}\sum_{i=1}^{n}x_{ij}\left(y_i-\mu_j\right)^2\right\}\right]$$

$$=\frac{\Gamma(\psi_1+\ldots+\psi_m)}{\Gamma(\psi_1)\ldots\Gamma(\psi_m)}\omega_1^{\,\psi_1-1}\ldots\omega_m^{\,\psi_m-1}\frac{1}{\left(2\pi\tau^2\right)^{m/2}}\left[\frac{\beta^\alpha}{\Gamma(\alpha)}\right]^m\left[\prod_{j=1}^{m}\sigma_j^{\,-(\alpha+1)}\right]e^{-\beta\sum_{j=1}^{m}\frac{1}{\sigma_j^{\,2}}}$$

$$\bullet\exp\left\{-\frac{1}{2\tau^2}\sum_{j=1}^{m}\left(\mu_j-\theta\right)^2\right\}\sum_{j=1}^{m}\omega_j^{\,n_j}\frac{1}{\left(2\pi\sigma_j^{\,2}\right)^{n_j/2}}\exp\left\{-\frac{1}{2\sigma_j^{\,2}}\sum_{i=1}^{n}x_{ij}\left(y_i-\mu_j\right)^2\right\}$$

Hence the posterior density is some kind of a mixture model.
Assume

$m = 2, \sigma_1^{\,2} = \sigma_2^{\,2} = \sigma_,^{\,2} \qquad n_1 + n_2 = n_, \qquad \mu' = (\mu_{1,}\mu_2)$

$\omega' = (\omega_{1,}\omega_2) \qquad \omega_2 = 1 - \omega_1$

$\psi' = (\psi_{1,}\psi_2)$

$$f(\mu, \sigma^2, \omega / y, x, \theta, \tau^2, \alpha, \beta, \psi)$$

$$\propto \frac{1}{(2\pi\tau^2)^{2/2}} \exp\left\{-\frac{1}{2\tau^2}\sum_{j=2}^{2}(\mu_j - \theta)^2\right\}\left[\frac{\beta^\alpha}{\Gamma(\alpha)}\right]^2 (\sigma^2)^{-2(\alpha+1)} \exp\left\{-\beta\frac{2}{\sigma^2}\right\}$$

$$\frac{\Gamma(\psi_1 + \psi_2)}{\Gamma(\psi_1)\Gamma(\psi_2)}\omega_1^{\psi_1 - 1}\omega_2^{\psi_2 - 1}\left[\omega_1^{n_1}\frac{1}{(2\pi\sigma^2)^{n_1/2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_1}x_{i1}(y_i - \mu_1)^2\right\}\right.$$

(5)

$$\left. + \omega_2^{n_2}\frac{1}{(2\pi\sigma^2)^{n_2/2}}\exp\left\{ \quad \frac{1}{2\sigma^2}\sum_{i=1}^{n_2}x_{i2}(y_i \quad \mu_2)^2\right\}\right]$$

IV. Estimation of the mixture model by Gibbs sampling.

For notational convenience, let

$$f(\mu, \sigma^2, \omega / y, x, \theta, \tau^2, \alpha, \beta, \psi) = f(\mu_1, \mu_2, \sigma^2, \omega / y, x)$$

(6)

$$\propto \frac{1}{2\pi\tau^2}\left[\frac{\beta^\alpha}{\Gamma(\alpha)}\right]^2 (\sigma^2)^{-2(\alpha+1)}\frac{\Gamma(\psi_1 + \psi_2)}{\Gamma(\psi_1)\Gamma(\psi_2)}\omega^{\psi_1 - 1}(1-\omega)^{\psi_2 - 1}\exp\left\{-\frac{\theta}{2\tau^2}\right\}\exp\left\{-\frac{2\beta}{\sigma^2}\right\}$$

$$\left[\frac{\omega^{n_1}}{(2\pi\sigma^2)^{n_1/2}}\left\{-\frac{1}{2\tau^2}[\mu_2 - \theta]^2\right\}\exp\left\{-\frac{1}{2}\left[\frac{1}{\tau^2} + \frac{n_1}{\sigma^2}\right]\left[\mu_1 - \frac{\frac{\theta}{\tau^2} + \sum_{i=1}^{n_1}\frac{x_{i1}y_i}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n_1}{\sigma^2}}\right]^2\right\}\right.$$

$$\bullet \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}x_{i1}y_i^{\;2} - \frac{\left(\frac{\theta\sigma^2}{\tau^2} + \sum_{i=1}^{n_1}x_{i1}y_i\right)^2}{\frac{\sigma^2}{\tau^2} + n_1}\right]\right\}$$

$$+ \frac{(1-\omega)^{n_2}}{(2\pi\sigma^2)^{n_2/2}}\exp\left\{-\frac{1}{2\tau^2}(\mu_1 - \theta)^2\right\}\exp\left\{-\frac{1}{2}\left(\frac{1}{\tau^2} + \frac{n_2}{\sigma^2}\right)\left[\mu_2 - \frac{\frac{\theta}{\tau^2} + \sum_{i=1}^{n_2}\frac{x_{i2}y_i}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n_2}{\sigma^2}}\right]^2\right\}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_2} x_{i2} y_i^{\ 2} - \frac{\left(\frac{\theta\sigma^2}{\tau^2} + \sum_{i=1}^{n_2} x_{i2} y_i\right)^2}{\frac{\sigma^2}{\tau^2} + n_2}\right]\right\}.$$

Evaluate the conditional densities

$$f\left(\mu_1/\mu_2, \sigma^2, \omega, y, x\right) = \frac{f\left(\mu_1, \mu_2, \sigma^2, \omega/y, x\right)}{\int f\left(\mu_1, \mu_2, \sigma^2, \omega/y, x\right) d\mu_1}.$$

Note that

$$\mathrm{A}\left(\sigma^2, \omega\right) = \frac{1}{2\pi\tau^2}\left[\frac{\beta^\alpha}{\Gamma(\alpha)}\right]^2 \left(\sigma^2\right)^{-2(\alpha+1)} \frac{\Gamma(\psi_1 + \psi_2)}{\Gamma(\psi_1)\Gamma(\psi_2)} \omega^{\psi_1 - 1} (1-\omega)^{\psi_2 - 1} \exp\left\{-\frac{\theta^2}{2\tau^2} - \frac{2\beta}{\sigma^2}\right\}\left[2\pi\left(\frac{1}{\frac{1}{\tau^2} + \frac{n_1}{\sigma^2}}\right)\right]^{1/2}$$

is a factor in the posterior distribution that does not involve μ_1 and μ_2 hence will banish when the conditional density is evaluated. Let

$$C_1\left(\mu_2, \sigma^2, \omega\right) = \frac{\omega^{n_1}}{\left(2\pi\tau^2\right)^{n_1/2}} \exp\left\{-\frac{1}{2\tau^2}(\mu_2 - \theta)^2\right\} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1} x_{i1} y_i^{\ 2} - \frac{\left(\frac{\theta\sigma^2}{\tau^2} + \sum_{i=1}^{n_1} x_{i1} y_i\right)^2}{\frac{\sigma^2}{\tau^2} + n_1}\right]\right\}$$

$$C_2\left(\mu_2, \sigma^2, \omega\right) = \frac{(1-\omega)^{n_2}}{\left(2\pi\sigma^2\right)^{n_2/2}} \exp\left\{-\frac{1}{2}\left[\frac{1}{\tau^2} + \frac{n_2}{\sigma^2}\right]\left[\mu_2 - \frac{\frac{\theta}{\tau^2} + \sum_{i=1}^{n_2} \frac{x_{i2} y_i}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n_2}{\sigma^2}}\right]^2\right\}\left(2\pi\tau^2\right)^{1/2}$$

$$\bullet \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_2} x_{i2} y_i^{\ 2} - \frac{\left(\frac{\theta\sigma^2}{\tau^2} + \sum_{i=1}^{n_2} x_{i2} y_i\right)^2}{\frac{\sigma^2}{\tau^2} + n_2}\right]\right\}.$$

Then draw the (l+1)st μ_1 from

$$f\left(\mu_1/\mu_2^{(l)}, \sigma^{2(l)}, \omega^{(l)}, y, x\right) \propto \frac{C_1\left(\mu_2, \sigma^2, \omega\right) N_{\mu_1}\left(\frac{\frac{\theta}{\tau^2} + \sum_{i=1}^{n_1} \frac{x_{i1} y_i}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n_1}{\sigma^2}}, \frac{1}{\frac{1}{\tau^2} + \frac{n_1}{\sigma^2}}\right) + C_2\left(\mu_2, \sigma^2, \omega\right) N_{\mu_1}\left(\theta, \tau^2\right)}{C_1\left(\mu_2, \sigma^2, \omega\right) + C_2\left(\mu_2, \sigma^2, \omega\right)} \quad (7)$$

To sample μ_2, evaluate

$$f\left(\mu_2/\mu_1,\sigma^2,\omega,y,x\right)=\frac{f\left(\mu_1,\mu_2,\sigma^2,\omega/y,x\right)}{\int f\left(\mu_1,\mu_2,\sigma^2,\omega/y,x\right)d\mu_2}.$$

Let

$$C_1\left(\mu,\sigma^2,\omega\right)=\frac{\omega^{n_1}}{\left(2\pi\sigma^2\right)^{n_1/2}}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_1}x_{i1}y_i^{\;2}-\frac{\left(\frac{\theta\sigma^2}{\tau^2}+\sum_{i=1}^{n_1}x_{ij}y_i\right)^2}{\frac{\sigma^2}{\tau^2}+n_1}\right]\right\}$$

$$.\left(2\pi\tau^2\right)^{1/2}\exp\left\{-\frac{1}{2}\left[\frac{1}{\tau^2}+\frac{n_1}{\sigma^2}\right]\left[\mu_1-\frac{\frac{\theta}{\tau^2}+\sum_{i=1}^{n_1}\frac{x_{i1}y_i}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n_1}{\sigma^2}}\right]^2\right\}$$

$$C_2\left(\mu_1,\sigma^2,\omega\right)=\frac{(1-\omega)^{n_2}}{\left(2\pi\sigma^2\right)^{n_2/2}}\exp\left\{-\frac{1}{2\tau^2}\left(\mu_1-\theta\right)^2\right\}\left[2\pi\frac{1}{\frac{1}{\tau^2}+\frac{n_2}{\sigma^2}}\right]^{1/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n_2}x_{i2}y_i^{\;2}-\frac{\left(\frac{\theta\sigma^2}{\tau^2}+\sum_{i=1}^{n_2}x_{i2}y_i\right)^2}{\frac{\sigma^2}{\tau^2}+n_2}\right]\right\}.$$

Then draw the (l+1)st μ_2 from

$$f\left(\mu_2/\mu_1^{(l+1)},\sigma^{2(l)},\omega^{(l)},y,x\right)\propto\frac{C_1\left(\mu_1,\sigma^2,\omega\right)N_{\mu_2}\left(\theta,\tau^2\right)+C_2\left(\mu_1,\sigma^2,\omega\right)N_{\mu_2}\left(\frac{\frac{\theta}{\tau^2}+\sum_{i=1}^{n_2}\frac{x_{i2}y_i}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n_2}{\sigma^2}},\frac{1}{\frac{1}{\tau^2}+\frac{n_2}{\sigma^2}}\right)}{C_1\left(\mu_1,\sigma^2,\omega\right)+C_2\left(\mu_1,\sigma^2,\omega\right)}\quad(8)$$

To determine the conditional posterior of σ^2 evaluate

$$f\left(\sigma^2/\mu_1,\mu_2,\omega,y,x\right)=\frac{f\left(\mu_1,\mu_2,\sigma^2,\omega/y,x\right)}{\int f\left(\mu_1,\mu_2,\sigma^2,\omega/y,x\right)\mathrm{d}\sigma^2}$$

that contains

$$\mathrm{A}\left(\mu_1,\mu_2,\omega\right)=\frac{1}{2\pi\tau^2}\left[\frac{\beta^\alpha}{\Gamma(\alpha)}\right]^2\frac{\Gamma\left(\psi_1+\psi_2\right)}{\Gamma\left(\psi_1\right)\Gamma\left(\psi_2\right)}\omega^{\psi_1-1}\left(1-\omega\right)^{\psi_2-1}\exp\left\{-\frac{1}{2\tau^2}\left[\sum_{j=1}^{2}\left(\mu_j-\theta\right)^2\right]\right\}$$

which will banish when the ratio is evaluated.

Hence, letting

$$C_1(\mu_1,\mu_2,\omega) = \omega^{n_1} \frac{1}{(2\pi)^{n_1/2}} \frac{\Gamma\left(\frac{n_1}{2}+2\alpha+1\right)}{\left[2\beta+\sum_{i=1}^{n_1} x_{i1}(y_i - \mu_1)^2\right]^{\frac{n_1}{2}+2\alpha+1}}$$

$$C_2(\mu_1,\mu_2,\omega) = (1-\omega)^{n_2} \frac{1}{(2\pi)^{n_2/2}} \frac{\Gamma\left(\frac{n_2}{2}+2\alpha+1\right)}{\left[2\beta+\sum_{i=1}^{n_2} x_{i2}(y_i - \mu_2)^2\right]^{\frac{n_2}{2}+2\alpha+1}},$$

draw the (l+1)st σ^2 from

$$f\left(\sigma^2/\mu_1^{(l+1)},\mu_2^{(l+1)},\omega^{(l)},y,x\right) \propto \frac{C_1(\mu_1,\mu_2,\omega) InvGamma\left(\frac{n_1}{2}+2\alpha+1, 2\beta+\sum_{i=1}^{n_1} x_{i1}(y_i-\mu_1)^2\right) + C_2(\mu_1,\mu_2,\omega) InvGamma\left(\frac{n_2}{2}+2\alpha+1, 2\beta+\sum_{i=1}^{n_2} x_{i2}(y_i-\mu_2)^2\right)}{C_1(\mu_1,\mu_2,\omega)+C_2(\mu_1,\mu_2,\omega)} \quad (9)$$

Finally, the conditional posterior of the mixing weight ω is given by

$$f\left(\omega/\mu_1^{(l+1)},\mu_2^{(l+1)},\sigma^{2(l+1)},y,x\right) \propto \frac{C_1(\mu_1,\mu_2,\sigma^2)\beta_\omega(\psi_1+n_1,\psi_2)+C_2(\mu_1,\mu_2,\sigma^2)\beta_\omega(\psi_1,\psi_2+n_2)}{C_1(\mu_1,\mu_2,\sigma^2)+C_2(\mu_1,\mu_2,\sigma^2)} \quad (10)$$

where

$$C_2(\mu_1,\mu_2,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n_1/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_1} x_{i1}(y_i-\mu_1)^2\right\} \frac{\Gamma(\psi_1+n_1)\Gamma(\psi_2)}{\Gamma(\psi_1+n_1+\psi_2)}$$

$$C_1(\mu_1,\mu_2,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n_2/2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_2} x_{i2}(y_i-\mu_2)^2\right\} \frac{\Gamma(\psi_2+n_2)\Gamma(\psi_1)}{\Gamma(\psi_1+n_2+\psi_2)}$$

where the (l+1)st ω is drawn.

The above models revealed that the conditional posterior of the parameters from a mixture likelihood are also mixtures whose mixing weights are the Cj's divided by the sum of Cj's, j = 1, 2, respectively. The component densities of the mixture posteriors belong to the corresponding family of prior densities with posterior parameters being functions of prior parameters and the data set.

V. Applications

A study on teaching preparation time per lecture hour was conducted among faculty members in a University. When classified by length of teaching experience the results is presented in Table 1.

Table 1. Mean preparation time and the variance of teachers by length of teaching Experience.

Length of Teaching Experience	n	Teaching Preparation Time	
		Mean	Variance
Short	34	2.971	5.439
Long	234	1.937	1.258
Total	268	2.068	1.889

Hence using the frequentist procedure and improving assumptions used in Bayesian analysis, that is, ${\sigma_1}^2 = {\sigma_2}^2 = \sigma^2$ and $\omega_j = P(\chi_{ij} = 1)$, the estimates for mean preparation time for two groups, their common variance, and the mixing weight are 2.971, 1.937, 1.889, and 0.127 $\left(= 34/268\right)$, respectively.

When homogeneity test for variances and test for mean difference were performed, significant results for both tests were obtained. A mixture density was then considered to model the distribution of the teaching preparation time. Bayesian approach was employed with prior distributions specified in Section 2. Gibbs sampling was employed but since conditional posterior densities are not of closed form as derived in Section 3, sampling procedure was performed using SPlus. Parameters estimated were mean preparation time for teachers with short teaching experience, mean preparation time for teachers with longer teaching experience, the assumed common variance and the mixing weights.

Results showed that for any initial vector value for the parameters of interest, Gibbs sampling gave a value of 1 for all the parameters. Still for any assumption on hyperprior parameters, same estimates were obtained. For any initial value, the mixture density always gives very low value hence elements of the sample seem to be trapped values. Even for long chains, sampled values seem not to move out of its present state.

VI. Concluding Remarks

While mixture distributions seem to be a generalization of some forms of distributions, the complexity of adding more parameters to be estimated can result to poor estimation. The decision on what assumptions to impose is arbitrary hence can be a source of problem of nonconvergence of chains. In many statistical situations, mixtures models seem plausible, but the complications in estimation may not warrant mixtures as a better alternative. For pedagogical reasons, it may be worth exploring.

VII. Acknowledgement

The author is most grateful to the administration of the West Visayas State University. Special thanks is extended to Dr. Adolfo de Guzman of UP School of Statistics for his encouragement to come up with a paper. Study sessions with Prof. Gen Sante of UP School of Statistics have been intellectually motivating. Antonette Uyaco was efficient in doing typing assignment.

References

Carlin, B. P. and Louis T. A. (1996). Bayes and empirical Bayes methods for data analysis. London: Chapman & Hall.

Congdon, P. (2003). Applied Bayesian modeling. Chichester: John Wiley & Sons.

Doss, H. and Huffer, F. W. (2003). *Monte Carlo methods for Bayesian analysis of survival data using mixtures of Dirichlet process priors*. J. of Computational and Graphical Statistics, Vol. 12, 2, 282 – 307.

Harn, M., Justel, A., and Rodert, C. P. (2003). *Estimating mixtures of regressions* J. of Computational & Graphical Statistics, Vol. 12, 1, 55 – 79.

McLachlan, G. J. and Basford, K. E. (1987). Mixture Models: Inference and application to clustering. New York: Marcel Dakker Inc.

Robert, C. P. and Mengersen, K. L. (1999). *Reparameterization issues in mixture modeling and their bearing on MCMC algorithms.* Computational Statistics & Data Analysis, Vol 29, 325 – 343.

Scaccia, L. and Green, P. J. (2003). *Bayesian growth curves using normal mixtures with nonparametric weights*. J. of Computational & Graphical Statistics. Vol. 12, 2, 308 – 331.

Schnatter, S. (2001). *Markov chain Monte Carlo estimation of classical & dynamic switching and mixture models*, J. Amer. Statistical Assn., Vol. 96, 453, 194 – 209.

Silverman, B. W. (1986). Density estimation for statistics and data analysis. London: Chapman & Hall.

Titterington, D. M., Smith, A. F. M. & Makov, U. E. (1985). Statistical analysis of finite mixture distributions. Chichester: John Wiley & Sons.