Pertanika 9(3), 381 – 386 (1986)

# Bayesian Analysis of Normal Scalar Populations with a Common Mean

MAT YUSOFF ABDULLAH
*Mathematics Department,*
*Faculty of Science and Environmental Studies,*
*Universiti Pertanian Malaysia,*
*43400 Serdang, Selangor, Malaysia.*

**Key words: Bayesian analysis; poly-t distribution.**

## ABSTRAK

*Beberapa populasi skalar normal dengan min sepunya $\theta$ dan parameter kepersisan yang berbeza dipertimbangkan. Dengan menggunakan famili konjugat taburan prior, taburan posterior sut bagi $\theta$ adalah k/o poly-t. Analisis posterior menunjukkan taburan bersyarat bagi $\theta$, diberi parameter kepersisan, adalah sentiasa bertaburan normal dan taburan bersyarat bagi parameter kepersisan, diberi min sepunya, adalah gama merdeka. Taburan poly-t ini dapat dinilaikan dengan senang secara plot berangka dan teknik kamiran.*

## ABSTRACT

*A number of normal scalar populations with common mean $\theta$ and distinct precision parameters are considered. Using the conjugate family of prior distributions, the marginal posterior distribution of $\theta$ is a k/o poly-t. The posterior analysis reveals that the conditional distribution of $\theta$ given precision parameters is always a normal distribution and the conditional distribution of the precision parameters given the common mean, is that of independent gamma. The poly-t distribution here can be evaluated easily with numerical plotting and integration techniques.*

## 1. INTRODUCTION

The purpose of the study was to develop the Bayesian analysis of several normal scalar populations each having an unknown mean $\theta$ but distinct precision parameters. The principal goal of a Bayesian analysis is to find the joint posterior distribution of all the parameters in the model and simplify it so that we can determine the marginal posterior distribution and algebraic formulas for the moments of these parameters.

Raiffa and Schlaifer (1961) developed the distribution theory necessary to carry out the Bayesian analysis of the multivariate normal model. Ando and Kaufman (1965) generalised the above treatment by considering the process where neither the mean vector nor the variance coveriance matrix is known. Zellner (1971) and Box and Tiao (1973) have made an extensive contribution to Bayesian inference.

Poly-t densities make its presence known in a Bayesian analysis of quite a number of common statistical models. Dreze (1977) discussed how poly-t densities arise as Bayesian Posterior densities for regression coefficients under a variety of specifications for the prior density and the data generating process and he summarized the results obtained for a number of models.

The poly-t distribution is difficult to work with; no analytical expression exists for the normalizing constant and moments. Several analytical approximations have been proposed.

Lindley and Smith (1972) suggested an approximation using the mode of the posterior distribution in place of the mean, by solving the model equations via an iterative method. The techniques of approximation have also been proposed by Zellner (1971), Box and Tiao (1973), Press (1982) and Broemeling and Abdullah (1984).

## 2. THE PRIOR AND POSTERIOR ANALYSIS

Consider several normal populations $N(\theta, \tau_i)$ for $i = 1, 2, \ldots, k$ where $\theta \in R$ is the common mean and $\tau_i > 0$ is the precision of population i. Let $x_{ij}$, $j = 1, 2, \ldots, n_i$ be a random sample of size $n_i$ from $N(\theta, \tau_i)$, then the likelihood function is

$$L(\theta, \rho | s) \propto \prod_{i=1}^{k} \tau_i^{\frac{n_i}{2}} e^{-\frac{\tau_i}{2}\left[\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + n_i (\theta - \bar{x}_i)^2\right]} \qquad (2.1)$$

where $\rho = (\tau_1, \tau_2, \ldots, \tau_k)$, s is the sample, $\theta \in R$ and $\tau_i > 0$ for all $i = 1, 2, \ldots, k$. We see that the form of the likelihood function suggests.

$$p(\theta, \rho) \propto \prod_{i=1}^{k} \tau_i^{\alpha_i - \frac{1}{2}} e^{-\frac{\tau_i}{2}[2\beta_i + \xi_i (\theta - \mu_i)^2]}, \qquad (2.2)$$

$\tau_i > 0, \xi_i > 0, \alpha_i > 0, \beta_i > 0, \mu_i \in R$ for every $i = 1, 2, \ldots, k$ and $\theta \in R$ as the conjugate prior density, which is the product of k normal-gamma densities, one for each population. By completing the square on $\theta$ in (2.2), the equation can be written as

$$p(\theta, \rho) \propto \left( \left( \sum_{i=1}^{k} \tau_i \xi_i \right)^{1/2} e^{-\frac{1}{2}\left[\theta - \left(\sum_{i=1}^{k} \tau_i \xi_i\right)^{-1}\left(\sum_{i=1}^{k} \tau_i \xi_i \mu_i\right)\right]^2 \left(\sum_{i=1}^{k} \tau_i \xi_i\right)} \right) \left( \prod_{i=1}^{k} \tau_i^{\alpha_i - \frac{1}{2}} e^{-\tau_i \beta_i} \frac{e^{-\frac{1}{2}\left[\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \tau_i \tau_j \xi_i \xi_j (\mu_i - \mu_j)^2\right]}}{\left(\sum_{i=1}^{k} \tau_i \xi_i\right)^{1/2}} \right)$$

$$\propto p_1(\theta | \rho)\, p_2(\rho).$$

The function inside the first set of brackets if regarded as a function of $\theta$, must be proportional to the conjugate density function of $\theta$ given $\rho$, since $\theta$ does not appear in the second set of brackets. This function is proportional to a normal distribution with mean $\left(\sum_{i=1}^{k} \tau_i \xi_i\right)^{-1}\left(\sum_{i=1}^{k} \tau_i \xi_i \mu_i\right)$ and precision $\sum_{i=1}^{k} \tau_i \xi_i$. It follows that the function inside the second set of brackets is proportional to the marginal prior density of $\rho$ which is not that of independent gammas.

Of course, if one lets $\xi_i \to 0$ for $i = 1, 2, \ldots, k$ in (2.2), one is assuming $\theta$ has a constant density over R, then the conjugate density is the product of k gamma densities, namely

$$p(\theta, \rho) \propto \prod_{i=1}^{k} \tau_i^{\alpha_i - \frac{1}{2}} e^{-\tau_i \beta_i}, \quad \theta \in R, \tau_i > 0,$$

which is improper. Also from (2.2), the marginal prior density of $\theta$ is a k/o poly-t distribution.

$$p(\theta) \propto \prod_{i=1}^{k} [2\beta_i + \xi_i (\theta - \mu_i)^2]^{-(2\alpha_i + 1)/2}, \quad \theta \in R \qquad (2.3)$$

and in the special case when $\beta_i = \beta$, $\xi_i = \xi$ and $\mu_i = \mu$ for all $i = 1, 2, \ldots, k$,

$$p(\theta) \propto [2\beta + \xi(\theta - \mu)^2]^{-\sum_{i=1}^{k}(2\alpha_i + 1)/2}, \quad \theta \in R$$

which is a t density with $\sum_{i=1}^{k} 2\alpha_i + k - 1$ degrees of freedom, location parameter $\mu$ and the precision

$$\frac{\xi\left(\sum_{i=1}^{k} 2\alpha_i + k - 1\right)}{2\beta}$$

By combining the likelihood function (2.1) and the conjugate density (2.2), we have the posterior density of the parameters.

$$p(\theta, \rho \mid s) \propto \prod_{i=1}^{k} [\tau_i^{\frac{n_i + 2\alpha_i + i}{2} - 1} e^{-\frac{\tau_i}{2}(2\beta_i + \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + n_i(\theta - \bar{x}_i)^2 + \xi_i(\theta - \mu_i)^2)}] \quad (2.4)$$

Completing the square on $\theta$ and intergrating with respect to $\tau_i, \tau_2, \ldots, \tau_k$, we have the marginal posterior density of $\theta$ which is a k/o poly-t distribution.

$$p(\theta \mid s) \propto \sum_{i=1}^{k} \left[2\beta_i + \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \frac{n_i \xi_i (\bar{x}_i - \mu_i)^2}{(n_i + \xi_i)} + (\theta - (n_i + \xi_i)^{-1}(n_i \bar{x}_i + \xi_i \mu_i))^2 (n_i + \xi_i)\right]^{\frac{-(n_i + 2\alpha_i + 1)}{2}} \quad (2.5)$$

Of course from a Bayesian standpoint, all we have to do is plot $p(\theta \mid s)$ vs. $\theta$, compute the mean, mode, median, etc., and these must be done numerically. Note that there are no formulas for the intergrating constants and moments of $p(\theta / s)$ (except in special cases). Also note that if

$n_i + \xi_i = a$ for all $i = 1, 2, \ldots, k$ (choose $\xi_i$'s)

$n_i \bar{x}_i + \xi_i \mu_i = b$ for all $i = 1, 2, \ldots, k$ (choose $\mu_i$'s)

and $2\beta_i + \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + \frac{n_i \xi_i (x_i - \mu_i)^2}{(n_i + \xi_i)}$

$= c$ for all

$i = 1, 2, \ldots, k$

(choose $\beta_i$'s), then

$$p(\theta \mid s) \propto [c + (\theta - a^{-1} b) a]^{-\sum_{i=1}^{k} \frac{(n_i + 2\alpha_i + 1)}{2}}, \theta \in R$$

which is a t density.

To find the marginal distribution of $\rho = (\tau_1, \tau_2, \ldots, t_k)$ we have to complete the square on $\theta$ from the exponent in (2.4), then integrate with respect to $\theta$. The result is

$$p(\rho \mid s) \propto \left(\prod_{i=1}^{k} \tau_i^{\frac{n_i + 2\alpha_i}{2} - 1} e^{-\frac{\tau_i}{2}[2\beta_i + \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2}\right)\left(\frac{e^{-\frac{A}{2}}}{[\sum_{i=1}^{k} \tau_i (n_i + \xi_i)]^{½}}\right) \quad (2.6)$$

where $\tau_i > 0$ for all $i = 1, 2, \ldots, k$ and

$$A = \Big[\sum_{i=1}^{k} \tau_i^2 n_i \xi_i (\bar{x}_i - \mu_i)^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \tau_i \tau_j n_i n_j (\bar{x}_i - \bar{x}_j)^2 + n_i \xi_j (\bar{x}_i - \mu_j)^2 + n_j \xi_i (\bar{x}_j - \mu_i)^2 + \xi_i \xi_j (\mu_i - \mu_j)^2)\Big] \left[\sum_{i=1}^{k} \tau_i (n_i + \xi_i)\right]^{-1}$$

which is a very complex distribution. The first part of (2.6) is just the density of k independent gammas, and the second part is a factor involving the squares and cross products of $\tau_i$ and $\tau_j$, $i, j = 1, 2, \ldots, k$. Thus the marginal density of $\rho$ is not that of independent gammas.

By (2.4) we see that given $\theta$, the conditional distribution of $\rho$ is that of independent gammas, that is

$$(\tau_i \mid \theta) \sim G\left[\frac{n_i + 2\alpha_i + 1}{2}, \frac{2\beta_i + \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + n_i(\theta - \bar{x}_i)^2 + \xi_i(\theta - \mu_i)^2}{2}\right]$$

for all $i = 1, 2, \ldots k$, and given $\rho$, it can be shown that $\theta$ is distributed as normal with

and

$$E(\theta \mid \rho, s) = \left[\sum_{i=1}^{k} \tau_i (n_i + \xi_i)\right]^{-1}$$

$$\left[\sum_{i=1}^{k} \tau_i (n_i \bar{x}_i + \xi_i \mu_i)\right] \text{ and}$$

$$V(\theta \mid \rho, s) = \left(\sum_{i=1}^{k} \tau_i (n_i + \xi_i)\right]^{-1}$$

The modes of the conditional distributions are

$$M(\tau_i \mid \tau_j, s, \theta)$$
$$j \neq i$$

$$= \frac{n_i + 2\alpha_i - 1}{2\beta_i + \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 + n_i(\theta - \bar{x}_i)^2 + \xi_i(\theta - \mu_i)^2}$$

for every $i = 1, 2, \ldots, k$ and

$$M(\theta \mid \rho) = \frac{\sum_{i=1}^{k} \tau_i (n_i \bar{x}_i + \xi_i \mu_i)}{\sum_{i=1}^{k} \tau_i (n_i + \xi_i)} \qquad (2.7)$$

The latter is a convex combination of $x_i$'s and $\mu_i$'s.

If we put $\alpha_i = 0$, $\beta_i = 0$, $\xi_i = 0$, $i = 1, 2, \ldots, k$ in the above results, we will get the analysis under improper prior.

## 3. INFERENCES FOR THE PARAMETERS

We have seen that the posterior distributions for the parameters give complex distribution; however, the marginal distribution of $\theta$, (2.5) is a k/o poly-t distribution. For $k = 2$, it is a 2/0 poly-t (one-dimensional), hence all we have to do is plot this density and calculate numerically the posterior mean, variance and other moments. On the other hand, the marginal distribution of ($\tau_1$, $\tau_2$) is quite complex since $\tau_1$ and $\tau_2$ are not independent. Hence one must numerically plot the contours of the joint density (2.6) for $k = 2$, over the region where $\tau_1 > 0$ and $\tau_2 > 0$.

Another approach is to find the mode of the joint density of ($\theta$, $\tau_1$, $\tau_2$). Using Lindley and Smith's (1972) idea, we can consider the conditional modes (2.7) for $k = 2$ and attempt to solve these equations using an interative algorithm. So starting with some estimate of $\tau_1$ and $\tau_2$, say the sample precision, estimate $\theta$; then solve for $\tau_1$ and $\tau_2$ and continue the procedure until the estimate stabilizes. The solution, perhaps, is the mode of the joint distribution, however, the joint density may be multi-modal and the solution may not converge to the joint mode, but to a local maximum or minimum.

## 4. NUMERICAL STUDY AND RESULTS

The numerical study here is based on the iterative method (Lindley and Smith, 1972) applied to univariate two factor poly-t. Table 1 presents the comparison for the mode of the marginal density of $\theta$ (equation (2.5) for $k = 2$) by numerical integration, the mode by interation from equation (2.7) for $k = 2$, and the marginal mean of by numerical integration.

TABLE 1
Comparison of the mode (from marginal distribution), mode (by iteration) and the mean

| $\alpha_1$ | $\alpha_2$ | $\beta_1$ | $\beta_2$ | $\xi_1$ | $\xi_2$ | $\mu_1$ | $\mu_2$ | Mode (Marginal) | Mode (Iterative) | Mean |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2261 |
| 0 | 0 | 0 | 0 | 0 | 0 | 4 | 4 | .2458 | .2418 | .2261 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2261 |
| 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | .2060 | .1995 | .1867 |
| 0 | 0 | 0 | 0 | 4 | 4 | 0 | 0 | .1661 | .1698 | .1589 |
| 0 | 0 | 0 | 0 | 6 | 6 | 0 | 0 | .1528 | .1478 | .1384 |
| 0 | 0 | 0 | 0 | 10 | 10 | 0 | 0 | .1130 | .1173 | .1099 |
| 0 | 0 | 0 | 0 | 14 | 14 | 0 | 0 | .0997 | .0973 | .0911 |
| 0 | 0 | 0 | 0 | 18 | 18 | 0 | 0 | .0864 | .0831 | .0779 |
| 0 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | .1894 | .1989 | .1867 |
| 0 | 0 | 2 | 6 | 0 | 0 | 0 | 0 | .1528 | .1679 | .1592 |
| 0 | 0 | 3 | 9 | 0 | 0 | 0 | 0 | .1528 | .1443 | .1389 |
| 0 | 0 | 5 | 15 | 0 | 0 | 0 | 0 | .1162 | .1113 | .1105 |
| 0 | 0 | 7 | 21 | 0 | 0 | 0 | 0 | .0797 | .0891 | .0915 |
| 0 | 0 | 9 | 27 | 0 | 0 | 0 | 0 | .0797 | .0733 | .0778 |
| 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2305 |
| 4 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2330 |
| 6 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2346 |
| 10 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2365 |
| 14 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | .2458 | .2418 | .2376 |
| 2 | 2 | 1 | 1 | 3 | 3 | −8 | −8 | −1.7741 | −1.7634 | −1.7636 |
| 2 | 2 | 1 | 1 | 3 | 3 | −4 | −4 | −.8173 | −.8225 | −.8245 |
| 2 | 2 | 1 | 1 | 3 | 3 | 0 | 0 | .1794 | .1745 | .1664 |
| 2 | 2 | 1 | 1 | 3 | 3 | 4 | 4 | 1.0565 | 1.0572 | 1.0520 |
| 2 | 2 | 1 | 1 | 3 | 3 | 8 | 8 | 1.9336 | 1.9512 | 1.9489 |
| 2 | 2 | 1 | 1 | 0 | 0 | 1 | 1 | .2359 | .2299 | .2191 |
| 2 | 2 | 1 | 1 | 3 | 3 | 1 | 1 | .4186 | .4078 | .3996 |
| 2 | 2 | 1 | 1 | 6 | 6 | 1 | 1 | .5382 | .5189 | .5124 |
| 2 | 2 | 1 | 1 | 9 | 9 | 1 | 1 | .5781 | .5949 | .5895 |
| 2 | 2 | 1 | 1 | 12 | 12 | 1 | 1 | .6578 | .6502 | .6455 |
| 2 | 2 | 1 | 1 | 15 | 15 | 1 | 1 | .6977 | .6922 | .6881 |
| 2 | 2 | 1 | 1 | 18 | 18 | 1 | 1 | .7375 | .7252 | .7215 |
| 2 | 2 | 1 | 3 | 2 | 2 | 2 | 2 | .4983 | .4947 | .4877 |
| 4 | 4 | 2 | 6 | 4 | 4 | 4 | 4 | 1.2425 | 1.2501 | 1.2477 |
| 6 | 6 | 3 | 9 | 6 | 6 | 6 | 6 | 2.3322 | 2.3318 | 2.3301 |
| 8 | 8 | 4 | 12 | 8 | 8 | 8 | 8 | 3.6346 | 3.6236 | 3.6229 |

The mode obtained through the iterative method is the mode of the joint density of all the parameters in the model and in general, it is not the mode of the marginal density (O'Hagan, A. 1976). Sample values used are $n_1 = n_2 = 10$,

$$\sum_{j=1}^{n_1} (x_{ij} - \bar{x}_1)^2 = 31.7882 \quad , \quad \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2$$

$= 9.9999$, $\bar{x}_1 = .1499$ and $\bar{x}_2 = .3611$. From the table we observe that, if $\beta$ 's and $\xi$ 's are equal to zero, changing the values of $\mu$'s would not change the modes and mean of $\theta$ and this is obvious from equations (2.4) and (2.5). The joint mode obtained through iteration is quite close to the mode of marginal by numerical integration whatever the values of $\alpha$ , $\beta$ , $\xi$ and $\mu$'s. These modes can be used to estimate means because the distribution is unimodal and $n_1$ and $n_2$ are fairly large.

We can conclude that the iterative method is a good approximation for estimating the mode of the joint distribution which is nearly symmetric. In the bimodal case, the modal value will converge to a local mode, depending on the starting value.

## REFERENCES


ANDO, A. and G.M. KAUFMAN (1965): Bayesian Analysis of the Independent Multinormal Process Neither Mean nor Precision Known. *J. Am. Statis. Assoc.* **60:** 347 – 385.

BOX, G.E.P. and G.C. TIAO (1973): Bayesian Inference in Statistical Analysis. Reading, Massachusetts: Addison-Wesley.

BROEMELING, L. and M.Y. ABDULLAH (1984): An Approximation to the Poly-t Distribution. *Commun. Statis. Theor. Meth.* **13(11):** 1407 – 1422.

DREZE, J.H. (1977): Bayesian Regression Analysis Using Poly-T Densities. *J. Econometrics,* **6:** 329 – 354.

LINDLEY, D.V. and A.F.M. SMITH (1972): Bayes Estimates for the Linear Model. *J. Roy. Statist. Soc., B,* **34:** 1 – 18.

O'HAGAN, A. (1976): On Posterior Joint and Marginal Modes. *Biometrika,* **63(2):** 329 – 333.

PRESS, J.S. (1982): Applied Multivariate Analysis, Robert E. Kreiger Publishing Co., Malabar, Florida.

RAIFFA, H. and R. SCHLAIFER (1961): Applied Statistical Decision Theory. Boston, Massachusetta: Harvard University Press.

ZELLNER, A. (1971): An Introduction to Bayesian Inference in Econometrics, John Wiley and Sons, Inc. New York.


*(Received 25 February, 1985)*