British Journal of Mathematics & Computer Science
20(6): 1-7, 2017, Article no.BJMCS.31814
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

The Posterior Distributions, the Marginal Distributions and the Normal Bayes Estimators of Three Hierarchical Normal Models

Ying-Ying Zhang[1*], Wen-He Song[1] and Teng-Zhong Rong[1]

[1] *Department of Statistics and Actuarial Science, College of Mathematics and Statistics, Chongqing University, Chongqing, China.*

Authors' contributions

This work was carried out in collaboration between all authors. Author YYZ did theoretical derivations, and wrote the first draft of the manuscript. Author WHS did literature searches and revised the manuscript. Author TZR revised the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2017/31814
Editor(s):
(1) Sergio Serrano, Department of Applied Mathematics, University of Zaragoza, Spain.
(2) H. M. Srivastava, Department of Mathematics and Statistics, University of Victoria, Canada.
Reviewers:
(1) Thomas L. Toulias, Technological Educational Institute of Athens, Greece.
(2) Ismail Olaniyi Muraina, Adeniran Ogunsanya College of Education, Nigeria.
Complete Peer review History: http://www.sciencedomain.org/review-history/18096

***Received:** 25th January 2017*
***Accepted:** 3rd March 2017*
***Published:** 7th March 2017*

Original Research Article

Abstract

We calculate the posterior distributions, the marginal distributions and the normal Bayes estimators of three hierarchical normal models in the same manner. The three models are displayed in increasing complexity. We find that the posterior distributions and the marginal distributions are all normal distributions. We also obtain the normal Bayes estimators under the squared error loss function.

Keywords: Posterior distribution; marginal distribution; normal Bayes estimator; hierarchical normal model; normal distribution.

2010 Mathematics Subject Classification: 62F10, 62F15.

Corresponding author: E-mail: robertzhang@cqu.edu.cn, robertzhangyying@qq.com;

1 Introduction

Statistical inferences are covered in many classical textbooks, see for instance, [1, 2, 3, 4, 5, 6]. Point estimation is an important class of statistical inference. Bayes estimator is a kind of point estimation. The classical Bayesian books are [7, 8, 9, 10, 11, 12, 13, 14, 15].

The hierarchical normal model (also named as the normal mean mixture model or the Gaussian mixture model) has been investigated recently in the literature, for example, [16, 17, 18, 19, 20]. We calculate the posterior distributions, the marginal distributions and the normal Bayes estimators of three Hierarchical Normal Models (HNMs) in the same manner. The three models are displayed in increasing complexity. We find that the posterior distributions and the marginal distributions are all normal distributions. We also obtain the normal Bayes estimators under the squared error loss function (or the quadratic loss function).

The rest of the paper is organized as follows. In the next Section 2, we calculate the posterior distributions, the marginal distributions and the normal Bayes estimators of three HNMs in the same manner. Section 3 concludes.

2 Main Results

The first Hierarchical Normal Model (HNM) (Model I) is

$$\left\{ \begin{array}{l} X|\theta \sim N\left(\theta,\sigma^2\right), \\ \theta \sim N\left(\mu,\tau^2\right), \end{array} \right. \tag{2.1}$$

where σ^2, μ and τ^2 are known constants. In Model I, one θ is drawn from its prior distribution $\pi(\theta) \sim N\left(\mu,\tau^2\right)$. Then given θ, one observation $X|\theta$ is drawn from its sampling distribution $f(x|\theta) \sim N\left(\theta,\sigma^2\right)$. Model I is simple and it has been treated in [3] Example 7.2.16. Our aim is to calculate the posterior distribution of $\theta|X$, the marginal distribution of X and the normal Bayes estimator $\mathrm{E}(\theta|X)$. Although the solution of our aim is already known, we will present our solution and approach here because they will be useful for two other HNMs. The joint distribution of x and θ is

$$f(x,\theta) = f(x|\theta)\,\pi(\theta) = \pi(\theta|x)\,m(x),$$

where $f(x|\theta)$ is the sampling distribution of $x|\theta$ (or the likelihood function), $\pi(\theta)$ is the prior distribution of θ, $\pi(\theta|x)$ is the posterior distribution of $\theta|x$ and $m(x)$ is the marginal distribution of x. It is known from (2.1) that

$$\begin{aligned} f(x,\theta) &= f(x|\theta)\,\pi(\theta) \\ &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right)\frac{1}{\sqrt{2\pi}\tau}\exp\left(-\frac{(\theta-\mu)^2}{2\tau^2}\right). \end{aligned}$$

Factor the exponent in the above mathematical expression, we obtain

$$-\frac{(x-\theta)^2}{2\sigma^2} - \frac{(\theta-\mu)^2}{2\tau^2} = -\frac{(\theta-\delta_1(x))^2}{2v_1^2} - \frac{(x-\mu)^2}{2\left(\tau^2+\sigma^2\right)}, \tag{2.2}$$

where

$$\delta_1(x) = \frac{\tau^2 x + \sigma^2\mu}{\tau^2+\sigma^2} \text{ and } v_1^2 = \frac{\tau^2\sigma^2}{\tau^2+\sigma^2}.$$

We remark that the proof of (2.2) is elementary but tedious. To obtain (2.2), the following completion of the square in θ is useful:

$$a\theta^2 + b\theta + c = a\left(\theta + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}.$$

Let $N(a,b)$ denote the pdf of a normal distribution with mean a and variance b. The factorization (2.2) shows that

$$\begin{aligned} f(x,\theta) &= f(x|\theta)\,\pi(\theta) = \pi(\theta|x)\,m(x) \\ &= N\left(\theta,\sigma^2\right) \times N\left(\mu,\tau^2\right) = N\left(\delta_1(x), v_1^2\right) \times N\left(\mu,\tau^2+\sigma^2\right), \end{aligned} \tag{2.3}$$

where the posterior distribution of $\theta|x$ is $\pi(\theta|x) \sim N\left(\delta_1(x), v_1^2\right)$ and the marginal distribution of x is $m(x) \sim N\left(\mu,\tau^2+\sigma^2\right)$. Thus, the normal Bayes estimator is $\mathrm{E}(\theta|X) = \delta_1(X)$.

The second HNM (Model II) is

$$\begin{cases} X_i|\theta \overset{iid}{\sim} N\left(\theta,\sigma^2\right), \ i = 1,2,\ldots,n, \\ \theta \sim N\left(\mu,\tau^2\right), \end{cases} \tag{2.4}$$

where σ^2, μ and τ^2 are known constants. In Model II, one θ is drawn from its prior distribution $\pi(\theta) \sim N\left(\mu,\tau^2\right)$. Then given θ, a sample $X_1|\theta$, $X_2|\theta$, $\ldots$, $X_n|\theta$ is drawn from its sampling distribution $f(x_i|\theta) \sim N\left(\theta,\sigma^2\right)$, $i = 1,2,\ldots,n$. Model II has been treated in [3] Exercise 7.22. For completeness, we present the solution here. Model II can be changed to Model IIa below:

$$\begin{cases} \bar{X}|\theta \sim N\left(\theta,\sigma^2/n\right), \\ \theta \sim N\left(\mu,\tau^2\right). \end{cases} \tag{2.5}$$

To compute the posterior distribution $\pi(\theta|\bar{x})$ and the marginal distribution $m(\bar{x})$, we follow the route of Model I. The joint distribution of $\bar{x}$ and θ is

$$f(\bar{x},\theta) = f(\bar{x}|\theta)\,\pi(\theta) = \pi(\theta|\bar{x})\,m(\bar{x}),$$

where $f(\bar{x}|\theta)$ is the sampling distribution of $\bar{x}|\theta$ (or the likelihood function), $\pi(\theta)$ is the prior distribution of θ, $\pi(\theta|\bar{x})$ is the posterior distribution of $\theta|\bar{x}$ and $m(\bar{x})$ is the marginal distribution of $\bar{x}$. It is known from (2.5) that

$$\begin{aligned} f(\bar{x},\theta) &= f(\bar{x}|\theta)\,\pi(\theta) \\ &= \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}} \exp\left(-\frac{(\bar{x}-\theta)^2}{2\sigma^2/n}\right) \frac{1}{\sqrt{2\pi}\tau} \exp\left(-\frac{(\theta-\mu)^2}{2\tau^2}\right). \end{aligned}$$

Factor the exponent in the above mathematical expression, we get

$$-\frac{(\bar{x}-\theta)^2}{2\sigma^2/n} - \frac{(\theta-\mu)^2}{2\tau^2} = -\frac{(\theta-\delta_2(\bar{x}))^2}{2v_2^2} - \frac{(\bar{x}-\mu)^2}{2\left(\tau^2+\sigma^2/n\right)}, \tag{2.6}$$

where

$$\delta_2(\bar{x}) = \frac{\tau^2\bar{x} + \left(\sigma^2/n\right)\mu}{\tau^2+\sigma^2/n} \text{ and } v_2^2 = \frac{\tau^2\sigma^2/n}{\tau^2+\sigma^2/n}.$$

We remark that the proof of (2.6) is elementary but tedious. The factorization (2.6) shows that

$$\begin{aligned} f(\bar{x},\theta) &= f(\bar{x}|\theta)\,\pi(\theta) = \pi(\theta|\bar{x})\,m(\bar{x}) \\ &= N\left(\theta,\frac{\sigma^2}{n}\right) \times N\left(\mu,\tau^2\right) = N\left(\delta_2(\bar{x}), v_2^2\right) \times N\left(\mu,\tau^2+\frac{\sigma^2}{n}\right), \end{aligned} \tag{2.7}$$

where the posterior distribution of $\theta|\bar{x}$ is $\pi(\theta|\bar{x}) \sim N\left(\delta_2(\bar{x}), v_2^2\right)$ and the marginal distribution of $\bar{x}$ is $m(\bar{x}) \sim N\left(\mu,\tau^2+\frac{\sigma^2}{n}\right)$. Therefore, the normal Bayes estimator is $\mathrm{E}\left(\theta|\bar{X}\right) = \delta_2\left(\bar{X}\right)$. In fact, we can get (2.7) from (2.3) directly by noting that (2.5) has the same structure with (2.1). In (2.5), compared to (2.1), $\bar{x}$ replaces x and σ^2/n replaces σ^2, other things being equal. Note that Model II is a generalization of Model I. Take $n = 1$ in Model II, then it reduces to Model I. Similarly, Model

IIa is also a generalization of Model I.

The third HNM (Model III) is

$$\begin{cases} X_i|\theta_i \overset{\text{independent}}{\sim} N\left(\theta_i, \sigma^2\right), \ i = 1, 2, \ldots, n, \\ \theta_i \overset{iid}{\sim} N\left(\mu, \tau^2\right), \ i = 1, 2, \ldots, n, \end{cases} \tag{2.8}$$

where σ^2, μ and τ^2 are known constants. In Model III, a sample $\theta_1, \theta_2, \ldots, \theta_n$ is drawn from its prior distribution $\pi(\theta_i) \sim N\left(\mu, \tau^2\right), \ i = 1, 2, \ldots, n$. Then given $\theta_i, \ i = 1, 2, \ldots, n$, a random variable $X_i|\theta_i$ is drawn from its sampling distribution $f(x_i|\theta_i) \sim N\left(\theta_i, \sigma^2\right)$. The resulting random variables $X_1|\theta_1, X_2|\theta_2, \ldots, X_n|\theta_n$ are independent but certainly not identically distributed. Model III is not new. It has been treated in [3] Exercise 7.25. In Exercise 7.25, it has been shown that the marginal distribution of X_i is $N\left(\mu, \tau^2 + \sigma^2\right)$ and that, marginally, $X_1, X_2, \ldots, X_n$ are iid. However, the posterior distribution and the normal Bayes estimator is not given. Model III can be changed to Model IIIa below:

$$\begin{cases} \bar{X}|\bar{\theta} \sim N\left(\bar{\theta}, \sigma^2/n\right), \\ \bar{\theta} \sim N\left(\mu, \tau^2/n\right). \end{cases} \tag{2.9}$$

Now we derive the distribution of $\bar{X}|\bar{\theta}$ in (2.9). We have

$$\bar{X}|\bar{\theta} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)|\bar{\theta} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i|\bar{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i|\theta_i\right),$$

since $\bar{\theta}$ has the same information of $(\theta_1, \theta_2, \ldots, \theta_n)$ and X_i depends only on θ_i. Since $X_i|\theta_i, \ i = 1, 2, \ldots, n$, are independent and they are normal, their linear combination is also normal. We have

$$\mathrm{E}\left(\bar{X}|\bar{\theta}\right) = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i|\theta_i\right)\right] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left(X_i|\theta_i\right) = \frac{1}{n}\sum_{i=1}^{n}\theta_i = \bar{\theta}$$

and

$$\begin{aligned} \mathrm{Var}\left(\bar{X}|\bar{\theta}\right) &= \mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i|\theta_i\right)\right] = \frac{1}{n^2}\mathrm{Var}\left[\sum_{i=1}^{n}\left(X_i|\theta_i\right)\right] \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\left(X_i|\theta_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}. \end{aligned}$$

Therefore,

$$\bar{X}|\bar{\theta} \sim N\left(\bar{\theta}, \frac{\sigma^2}{n}\right).$$

To compute the posterior distribution $\pi\left(\bar{\theta}|\bar{x}\right)$ and the marginal distribution $m(\bar{x})$, we follow the route of Model I. The joint distribution of $\bar{x}$ and $\bar{\theta}$ is

$$f\left(\bar{x}, \bar{\theta}\right) = f\left(\bar{x}|\bar{\theta}\right)\pi\left(\bar{\theta}\right) = \pi\left(\bar{\theta}|\bar{x}\right)m(\bar{x}),$$

where $f\left(\bar{x}|\bar{\theta}\right)$ is the sampling distribution of $\bar{x}|\bar{\theta}$ (or the likelihood function), $\pi\left(\bar{\theta}\right)$ is the prior distribution of $\bar{\theta}$, $\pi\left(\bar{\theta}|\bar{x}\right)$ is the posterior distribution of $\bar{\theta}|\bar{x}$ and $m(\bar{x})$ is the marginal distribution of $\bar{x}$. It is known from (2.9) that

$$\begin{aligned} f\left(\bar{x}, \bar{\theta}\right) &= f\left(\bar{x}|\bar{\theta}\right)\pi\left(\bar{\theta}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\exp\left(-\frac{\left(\bar{x}-\bar{\theta}\right)^2}{2\sigma^2/n}\right)\frac{1}{\sqrt{2\pi}\tau/\sqrt{n}}\exp\left(-\frac{\left(\bar{\theta}-\mu\right)^2}{2\tau^2/n}\right). \end{aligned}$$

Factor the exponent in the above mathematical expression, we have

$$-\frac{\left(\bar{x}-\bar{\theta}\right)^2}{2\sigma^2/n}-\frac{\left(\bar{\theta}-\mu\right)^2}{2\tau^2/n}=-\frac{\left(\bar{\theta}-\delta_3\left(\bar{x}\right)\right)^2}{2v_3^2}-\frac{\left(\bar{x}-\mu\right)^2}{2\left(\tau^2/n+\sigma^2/n\right)}, \tag{2.10}$$

where

$$\delta_3\left(\bar{x}\right)=\frac{\left(\tau^2/n\right)\bar{x}+\left(\sigma^2/n\right)\mu}{\tau^2/n+\sigma^2/n}=\frac{\tau^2\bar{x}+\sigma^2\mu}{\tau^2+\sigma^2} \text{ and } v_3^2=\frac{\left(\tau^2/n\right)\left(\sigma^2/n\right)}{\tau^2/n+\sigma^2/n}=\frac{\tau^2\sigma^2}{n\left(\tau^2+\sigma^2\right)}.$$

We remark that the proof of (2.10) is elementary but tedious. The factorization (2.10) shows that

$$\begin{aligned} f\left(\bar{x},\bar{\theta}\right)&=f\left(\bar{x}|\bar{\theta}\right)\pi\left(\bar{\theta}\right)=\pi\left(\bar{\theta}|\bar{x}\right)m\left(\bar{x}\right)\\ &=N\left(\bar{\theta},\frac{\sigma^2}{n}\right)\times N\left(\mu,\frac{\tau^2}{n}\right)=N\left(\delta_3\left(\bar{x}\right),v_3^2\right)\times N\left(\mu,\frac{\tau^2}{n}+\frac{\sigma^2}{n}\right), \end{aligned} \tag{2.11}$$

where the posterior distribution of $\bar{\theta}|\bar{x}$ is $\pi\left(\bar{\theta}|\bar{x}\right)\sim N\left(\delta_3\left(\bar{x}\right),v_3^2\right)$ and the marginal distribution of $\bar{x}$ is $m\left(\bar{x}\right)\sim N\left(\mu,\frac{\tau^2}{n}+\frac{\sigma^2}{n}\right)$. Consequently, the normal Bayes estimator is $\mathrm{E}\left(\bar{\theta}|\bar{X}\right)=\delta_3\left(\bar{X}\right)$. In fact, we can get (2.11) from (2.3) directly by noting that (2.9) has the same structure with (2.1). In (2.9), compared to (2.1), $\bar{x}$ replaces x, $\bar{\theta}$ replaces θ, σ^2/n replaces σ^2 and τ^2/n replaces τ^2, other things being equal. Note that Model III is a generalization of Model I. Take $n=1$ in Model III, then it reduces to Model I. Similarly, Model IIIa is also a generalization of Model I. Note also that Model III is a generalization of Model II. In Model III, when

$$\theta_1=\theta_2=\cdots=\theta_n=\theta\sim N\left(\mu,\tau^2\right), \tag{2.12}$$

then Model III reduces to Model II. However,

$$P\left(\theta_1=\theta_2=\cdots=\theta_n\right)=0.$$

Since Model III generalizes Model II, we can get (2.11) from (2.7) directly by noting that (2.9) has the same structure with (2.5). Under (2.12), we have

$$\bar{\theta}=\frac{1}{n}\sum_{i=1}^{n}\theta_i=\frac{1}{n}\sum_{i=1}^{n}\theta=\theta\sim N\left(\mu,\tau^2\right).$$

In (2.9), compared to (2.5), $\bar{\theta}$ replaces θ and τ^2/n replaces τ^2, other things being equal.

3 Conclusion

We calculate the posterior distributions, the marginal distributions and the normal Bayes estimators of three hierarchical normal models in the same manner. The three models are displayed in increasing complexity in the sense that Model II generalizes Model I, and Model III generalizes Models I and II. We find that the posterior distributions and the marginal distributions are all normal distributions. We also obtain the normal Bayes estimators under the squared error loss function (or the quadratic loss function).

Acknowledgement

The authors gratefully acknowledge the constructive comments offered by the referees. Their comments improve the quality of the paper significantly. The research was supported by the Fundamental Research Funds for the Central Universities (CQDXWL-2012-004 and CDJRC10100010), China Scholarship Council (201606055028), and the MOE project of Humanities and Social Sciences on the west and the border area (14XJC910001).

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Lehmann EL, Casella G. Theory of Point Estimation [M]. Springer, New York. 2nd edition; 1998.

[2] Stuart A, Ord JK, and Arnold S. Advanced Theory of Statistics, Volume 2A: Classical Inference and the Linear Model [M]. Oxford University Press, London. 6th edition; 1999.

[3] Casella G, Berger RL. Statistical Inference [M]. Duxbury, USA. 2nd edition; 2002.

[4] Shao J. Mathematical Statistics [M]. Springer, New York. 2nd edition; 2003.

[5] Lehmann EL, Romano JP. Testing Statistical Hypotheses [M]. Springer, New York. 3rd edition; 2005.

[6] Shi NZ, Tao J. Statistical Hypothesis Testing: Theory and Methods [M]. World Scientific Publishing, Singapore; 2008.

[7] Berger JO. Statistical decision theory and bayesian analysis [M]. Springer, New York. 2nd edition; 1985.

[8] Box GE, Tiao GC. Bayesian inference in statistical analysis [M]. Wiley, New York; 1992.

[9] Bernardo JM, Smith AFM. Bayesian Theory [M]. Wiley, New York; 1994.

[10] Chen MH, Shao QM, Ibrahim JG. Monte carlo methods in bayesian computation [M]. Springer, New York; 2000.

[11] Robert CP, Casella G. Monte carlo statistical methods [M]. Springer, New York. 2nd edition; 2005.

[12] Robert CP. The bayesian choice: From decision-theoretic motivations to computational implementation [M]. Springer, New York. 2nd paperback edition; 2007.

[13] Albert J. Bayesian computation with R (Use R!) [M]. Springer, New York. 2nd edition; 2009.

[14] Robert CP, Casella G. Introducing monte carlo methods with r (Use R!) [M]. Springer, New York; 2009.

[15] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis [M]. Chapman & Hall, London. 3rd edition; 2013.

[16] Conlon EM. A bayesian mixture model for metaanalysis of microarray studies [J]. Functional & integrative genomics. 2008;8:43-53.

[17] Gonzalez JEI, Thompson PM, Zhao AS. Modeling diffusion-weighted mri as a spatially variant gaussian mixture: Application to image denoising [J]. Medical Physics. 2011;38:4350-4364.

[18] Ji ZX, Xia Y, Sun QS, Chen Q, Xia DS, Feng DG. Fuzzy local gaussian mixture model for brain mr image segmentation [J]. IEEE Transactions on Information Technology in Biomedicine. 2012;16:339-347.

[19] Parent E, Rivot E. The basic normal model: A bayesian comparison of means [C]. In Introduction to Hierarchical Bayesian Modeling for Ecological Data. Chapman & Hall CRC; 2013.

[20] Moatti M, Chevret S, Zohar S, Rosenberger WF. A bayesian hybrid adaptive randomisation design for clinical trials with survival outcomes [J]. Methods of Information in Medicine. 2016;55:4-13.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/18096