Journal of Korean
Data & Information Science Society
2003, Vol. 14, No.4 pp. 1023~1030

Bayesian Test for Equality of Coefficients of Variation in the Normal Distributions

Hee Choon Lee · Sang Gil Kang[1] · Dal Ho Kim[2]

Abstract

When X and Y have independent normal distributions, we develop a Bayesian testing procedure for the equality of two coefficients of variation. Under the reference prior of the coefficient of variation, we propose a Bayesian test procedure for the equality of two coefficients of variation using fractional Bayes factor. A real data example is provided.

Keywords : Fractional Bayes Factor; Reference Prior; Normal Distribution; Coefficients of Variations.

1. INTRODUCTION

The coefficient of variation is an important parameter in many physical, biological and medical sciences. In general, it measures the consistency or uniformity of a set of observations on a random variable. Since the coefficient of variation is the standard deviation per unit mean, it represents a measure of relative variability. Groups can have the same relative variability even if the means and variances of the variable of interest are different.

The present paper focuses on Bayesian testing procedure for the equality of two coefficients of variation. In Bayesian testing problem, the Bayes factor under proper priors or informative priors have been very successful. However, limited information and time constraints often require the use of noninformative priors. Since noninformative priors such as Jeffreys' priors or reference priors (Berger and Bernardo, 1989, 1992) are typically improper so that such priors are only defined up to arbitrary constants which affects the values of Bayes factors. Spiegalhalter and Smith (1982), O'Hagan (1995) and Berger and Pericchi (1996)

1) Department of Applied Statistics, Sangji University, Wonju, 220-702, Korea
2) Department of Statistics, Kyungpook National University, Taegu, 702-701, Korea

have made efforts to compensate for that arbitrariness.

Berger and Pericchi (1996) introduced the intrinsic Bayes factor using a data-splitting idea, which would eliminate the arbitrariness of improper priors. O'Hagan (1995) proposed the fractional Bayes factor. For removing the arbitrariness he used to a portion of the likelihood with a so-called the fraction b. These approaches have shown to be quite useful in many statistical areas.

For testing the equality of coefficients of variation (CV), Miller and Karson (1977) presented a test for the equality of two CVs. Doornbos and Dijkstra (1983) developed a likelihood ratio test and a non central t test for the case of k normal samples of possible unequal sizes. The likelihood ratio test involves an algebraically unsolvable equation when more than two populations are considered. So Gupta and Ma (1996) provided a better method of solving this equation numerically than the on suggested by Doornbs and Dijkstra (1983) and developed a new test, the so called score test. Rao and Vidya (1992) provided a Wald test for testing the equality of CVs in two populations with equal sample sizes.

Almost all the work mentioned above is the analysis based on the classical point of view, there is a little work on this problem from the viewpoint of Bayesian framework. And we feel a strong necessity to develop objective Bayesian procedure for dealing this problem. So we want to develop the Bayesian test procedure for the equality of two CVs using Bayes factor. Using the noninformative priors developed previously, we calculate the posterior probabilities of the hypotheses using the fractional Bayes factor of O'Hagan (1995). Our testing for the equality of CVs will imply that the two means are of equal sign. Thus we can, without loss of generality, assume that the two means are positive (Sinha, Rao and Clement, 1978; Gupta, Ramakrishnan and Zhou, 1999).

The outline of the remaining sections is as follows. In Section 2, using the reference priors, we provide the Bayesian testing procedure based on the fractional Bayes factor for the testing equality of two coefficients of variation. In Section 3, a real example is given.

2. BAYESIAN TEST USING THE FRACTIONAL BAYES FACTOR

2.1 Preliminaries

Models (or Hypotheses) H_1, $H_2, \cdots$, H_q are under consideration, with the data $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ having probability density function $f_i(\boldsymbol{x} \mid \boldsymbol{\theta}_i)$ under model $H_i, i = 1, 2, \cdots, q$. The parameter vectors $\boldsymbol{\theta}_i$ are unknown. Let $\pi_i(\boldsymbol{\theta}_i)$ be the prior distribution of model H_i, and let p_i be the prior probabilities of model H_i,

$i=1,2,\cdots,q$. Then the posterior probability that the model H_i is true is

$$P(H_i \mid x)=\left(\sum_{j=1}^{q}\frac{p_j}{p_i}\cdot B_{ji}\right)^{-1}, \tag{1}$$

where B_{ji} is the Bayes factor of model H_j to model H_i defined by

$$B_{ji}=\frac{m_j(x)}{m_i(x)}=\frac{\int f_j(x \mid \theta_j)\pi_j(\theta_j)d\theta_j}{\int f_i(x \mid \theta_i)\pi_i(\theta_i)d\theta_i}. \tag{2}$$

The B_{ji} interpreted as the comparative support of the data for the model j to i. The computation of B_{ji} needs specification of the prior distribution $\pi_i(\theta_i)$ and $\pi_j(\theta)$. Usually, one can use the noninformative prior, often improper, such as uniform prior, Jeffreys prior, reference prior or probability matching prior. Denote it as π_i^N. The use of improper priors $\pi_i^N(\cdot)$ in (2) causes the B_{ji} to contain unspecified constants. To solve this problem, O'Hagan (1995) proposed the fractional Bayes factor for Bayesian testing and model selection problem as follow.

When the $\pi_i^N(\theta_i)$ is noninformative prior under H_i, equation (2) becomes

$$B_{ji}^N=\frac{\int f_j(x \mid \theta_j)\pi_j^N(\theta_j)d\theta_j}{\int f_i(x \mid \theta_i)\pi_i^N(\theta_i)d\theta_i}.$$

Then the fraction Bayes factor (FBF) of model H_j versus model H_i is

$$B_{ji}^F=\frac{q_j(b,x)}{q_i(b,x)},$$

where

$$q_i(b,x)=\frac{\int f_i(x \mid \theta_i)\pi_i^N(\theta_i)d\theta_j}{\int f_i^b(x \mid \theta_i)\pi_i^N(\theta_i)d\theta_i},$$

and $f_i(x \mid \theta_i$ is the likelihood function and b specifies a fraction of the likelihood which is to be used as a prior density. He proposed three ways for the choice of the fraction b. One frequently suggested choice is $b=m/n$, where m is the size of the minimal training sample, assuming this is well defined. (see O'Hagan, 1995 and the discussion by Berger and Mortera of O'Hagan, 1995).

2.2 Bayesian Test

Suppose that $X=(X_1,\cdots,X_{n_1})$ is a random sample of size n_1 from a normal population with mean μ_1 and variance $\mu_1^2\gamma_1^2$ and $Y=(Y_1,\cdots,Y_{n_2})$ is a random sample of size n_2 from a normal population with mean μ_2 and variance $\mu_2^2\gamma_2^2$. Here γ_1 and γ_2 are the CV for each population. Then the joint probability density function is

$$f(x,y \mid \mu_1,\mu_2,\gamma_1,\gamma_2) = (2\pi)^{-(n_1+n_2)/2}\gamma_1^{-n_1}\gamma_2^{-n_2}\mu_1^{-n_1}\mu_2^{-n_2} \times \exp\{-\frac{1}{2\gamma_1^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{1}{2\gamma_2^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\},$$

where $\mu_1>0$, $\mu_2>0$, $\gamma_1>0$ and $\gamma_2>0$.

We want to test the hypotheses $H_1: \gamma_1=\gamma_2$ vs. $H_2: \gamma_1\neq\gamma_2$. The hypothesis H_1 indicate the common CV. Our interest is to develop a Bayesian test for H_1 vs. H_2 which is an alternative to the classical tests.

Under the hypothesis H_1, one-at-a-time reference prior for $\gamma(\equiv\gamma_1=\gamma_2)$, μ_1 and μ_2 is

$$\pi_{H_1}(\gamma,\mu_1,\mu_2)=\mu_1^{-1}\mu_2^{-1}\gamma^{-1}(1+2\gamma^2)^{-1/2},\ \ \mu_1,\mu_2,\gamma>0.$$

This reference prior developed by Lee and Kang (2003). Also they proved that the posterior density under this reference prior is proper. The likelihood function under H_1 is

$$L(\gamma,\mu_1,\mu_2 \mid x,y) = (2\pi)^{-(n_1+n_2)/2}\mu_1^{-n_1}\mu_2^{-n_2}\gamma^{-(n_1+n_2)} \times \exp\{-\frac{1}{2\gamma^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{1}{2\gamma^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}.$$

Then the element of FBF under H_1 is given by

$$\begin{aligned}&\int_0^\infty\int_0^\infty\int_0^\infty L^b(\gamma,\mu_1,\mu_2 \mid x,y)\pi_1(\gamma,\mu_1,\mu_2)d\gamma d\mu_1 d\mu_2\\ &= (2\pi)^{-b(n_1+n_2)/2}\int_0^\infty\int_0^\infty\int_0^\infty \mu_1^{-bn_1-1}\mu_2^{-bn_2-1}\gamma^{-b(n_1+n_2)-1}(1+2\gamma^2)^{-\frac{1}{2}}\\ &\times \exp\{-\frac{b}{2\gamma^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{b}{2\gamma^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}d\gamma d\mu_1 d\mu_2.\end{aligned}$$

Let

$$S(x,y) = \int_0^\infty\int_0^\infty\int_0^\infty \mu_1^{-n_1-1}\mu_2^{-n_2-1}\gamma^{-(n_1+n_2)-1}(1+2\gamma^2)^{-1/2} \times \exp\{-\frac{1}{2\gamma^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{1}{2\gamma^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}d\gamma d\mu_1 d\mu_2$$

and

$$S^b(x,y) = \int_0^\infty\int_0^\infty\int_0^\infty \mu_1^{-bn_1-1}\mu_2^{-bn_2-1}\gamma^{-b(n_1+n_2)-1}(1+2\gamma^2)^{-1/2} \times \exp\{-\frac{b}{2\gamma^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{b}{2\gamma^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}d\gamma d\mu_1 d\mu_2.$$

Then

$$q_1(b,\ x,y)=\frac{(2\pi)^{-\frac{(n_1+n_2)}{2}}S(x,y)}{(2\pi)^{-\frac{b(n_1+n_2)}{2}}S^b(x,y)}.$$

For the H_2, one-at-a-time reference prior for μ_1,μ_2,γ_1 and γ_2 is

$$\begin{aligned}&\pi_{H_2}(\mu_1,\mu_2,\gamma_1,\gamma_2)\\ &=\ \pi(\mu_1,\gamma_1)\pi(\mu_2,\gamma_2)\\ &=\ \mu_1^{-1}\mu_2^{-1}\gamma_1^{-1}\gamma_2^{-1}(1+2\gamma_1^2)^{-1/2}(1+2\gamma_2^2)^{-1/2},\ \mu_1,\mu_2,\gamma_1,\gamma_2>0.\end{aligned}$$

Note that the propriety of the posterior distribution under this reference prior is given in Appendix 1. The likelihood function under H_2 is

$$\begin{aligned}L(\mu_1,\mu_2,\gamma_1,\gamma_2\mid x,y)\ &=\ (2\pi)^{-(n_1+n_2)/2}\mu_1^{-n_1}\mu_2^{-n_2}\gamma_1^{-n_1}\gamma_2^{-n_2}\\ &\times\ \exp\{-\frac{1}{2\gamma_1^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{1}{2\gamma_2^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}.\end{aligned}$$

Thus the element of FBF under H_2 gives as follows.

$$\begin{aligned}&\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty L^b(\mu_1,\mu_2,\gamma_1,\gamma_2\mid x,y)\pi_2(\mu_1,\mu_2,\gamma_1,\gamma_2)d\mu_1 d\mu_2 d\gamma_1 d\gamma_2\\ &=\ (2\pi)^{-\frac{b(n_1+n_2)}{2}}\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty\mu_1^{-bn_1-1}\mu_2^{-bn_2-1}\gamma_1^{-bn_1-1}\gamma_2^{-bn_2-1}(1+2\gamma_1^2)^{-1/2}(1+2\gamma_2^2)^{-1/2}\\ &\times\ \exp\{-\frac{b}{2\gamma_1^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{b}{2\gamma_2^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}d\mu_1 d\mu_2 d\gamma_1 d\gamma_2.\end{aligned}$$

Put

$$\begin{aligned}T(x,y)=&\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty\mu_1^{-n_1-1}\mu_2^{-n_2-1}\gamma_1^{-n_1-1}\gamma_2^{-n_2-1}(1+2\gamma_1^2)^{-\frac{1}{2}}(1+2\gamma_2^2)^{-\frac{1}{2}}\\ &\times\exp\{-\frac{1}{2\gamma_1^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{1}{2\gamma_2^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}d\mu_1 d\mu_2 d\gamma_1 d\gamma_2\end{aligned}$$

and

$$\begin{aligned}T^b(x,y)=&\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty\mu_1^{-bn_1-1}\mu_2^{-bn_2-1}\gamma_1^{-bn_1-1}\gamma_2^{-bn_2-1}(1+2\gamma_1^2)^{-\frac{1}{2}}(1+2\gamma_2^2)^{-\frac{1}{2}}\\ &\times\exp\{-\frac{b}{2\gamma_1^2\mu_1^2}\sum_{i=1}^{n_1}(x_i-\mu_1)^2-\frac{b}{2\gamma_2^2\mu_2^2}\sum_{i=1}^{n_2}(y_i-\mu_2)^2\}d\mu_1 d\mu_2 d\gamma_1 d\gamma_2.\end{aligned}$$

Then

$$q_2(b,\ x,y)=\frac{(2\pi)^{-\frac{(n_1+n_2)}{2}}T(x,y)}{(2\pi)^{-\frac{b(n_1+n_2)}{2}}T^b(x,y)}.$$

Therefore the FBF of H_2 versus H_1 is given by

$$B_{21}^F(x,y)=\frac{T(x,y)S^b(x,y)}{T^b(x,y)S(x,y)}.$$

Note that the element of FBF under H_1 requires a three dimensional integration and the element of FBF under H_2 requires a two dimensional integration. Therefore we have the value of the FBF of H_2 versus H_1. In Section 3, we investigate our testing procedure and the classical test statistic.

3. AN EXAMPLE

The data in Table 1 are taken from Nelson (1990) and represent the hours to failure of 20 motorettes with a new class-H insulation run at $240\ ^{\circ}C$ and $220\ ^{\circ}C$. It has been observed by Nelson (1990) that lognormal distribution adequately fits at the two temperatures. Note that X and Y denote the nature logarithm of the failure times in the Table 1. We thus assume that the below data come from independent normal distributions.

Table 1: Failure Times at the Two Temperatures

X ($240\ ^{\circ}C$)	7.0690	7.0690	7.3271	7.3582	7.3883
	7.4176	7.4176	7.4460	7.4736	7.5771
Y ($220\ ^{\circ}C$)	7.4753	7.7981	7.7981	7.7981	7.7981
	7.7981	8.0417	8.0417	8.0417	8.0417

For the equality of the CV, the score test developed by Gupta and Ma (1996). That is, under H_1, the test statistics is given by

$$Z = \frac{\hat{\gamma}^2(1+2\hat{\gamma}^2)}{2}\left(\frac{a_1^2}{n_1}+\frac{a_2^2}{n_2}\right),$$

where

$$a_1=\sum_{i=1}^{n_1}\frac{(x_i-\hat{\mu}_1)^2}{\hat{\mu}_1^2\hat{\gamma}^3}-\frac{n_1}{\hat{\gamma}},\ a_2=\sum_{i=1}^{n_2}\frac{(y_i-\hat{\mu}_2)^2}{\hat{\mu}_2^2\hat{\gamma}^3}-\frac{n_2}{\hat{\gamma}}$$

and $\widehat{\mu_1}, \widehat{\mu_2}, \hat{\gamma}$ is maximum likelihood estimator of μ_1, μ_2, γ. Under H_1, Z has a chi-square distribution with 1 degree of freedom. Since $\hat{\mu}_1=7.354238$, $\hat{\mu}_2=7.863381$ and $\hat{\gamma}=0.021644$, the observed value of the test statistic Z is 0.0113. Hence H_1 not rejected and the p-value of the test is almost one (Gupta, Ramakrishnan and Zhou, 1999).

The value of fractional Bayes factor of H_2 versus H_1 is $B_{21}^F=0.2939$. We assume that the prior probabilities are equal. Then the posterior probability for H_1 is 0.7729. Thus there are strong evidence for H_1 in terms of the posterior probability.

Therefore both of the classical method and Bayes factor give reasonable answers in this example.

APPENDIX 1. Propriety of Posterior Distribution

Under the reference prior $\pi(\mu,\gamma)=\mu^{-1}\gamma^{-1}(1+2\gamma^2)^{-1/2}$, the joint posterior for μ,γ given x is

$$\pi(\mu,\gamma\mid x)\propto\mu^{-(n+1)}\gamma^{-(n+1)}(1+2\gamma^2)^{-1/2}\exp\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\gamma^2\mu^2}\}.$$

Since $(1+2\gamma^2)^{-1/2}\leq 2^{-1/2}\gamma^{-1}$, thus

$$\pi(\mu,\gamma\mid x)\leq C_1\mu^{-(n+1)}\gamma^{-(n+2)}\exp\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\gamma^2\mu^2}\},\qquad(3)$$

where C_1 is a constant. Integrating with respect to γ in (3), then

$$C_2[\sum_{i=1}^{n}(x_i-\mu)^2]^{-\frac{n+1}{2}}=C_3[1+\frac{n(\bar{x}-\mu)^2}{S^2}]^{-\frac{n+1}{2}},$$

where $S^2=\sum_{i=1}^{n}(x_i-\bar{x})^2$ and C_2 and C_3 are a constant. Thus the above form has a Student- t with n degrees of freedom. Thus the posterior distribution is proper. This completes the proof. □

REFERENCES

1. Berger, J. O. and Bernardo, J. M. (1989). Estimating a Product of Means : Bayesian Analysis with Reference Priors. *Journal of the American Statistical Association*, 84, 200-207.
2. Berger, J. O. and Bernardo, J. M. (1992). On the Development of Reference Priors (with discussion). *Bayesian Statistics IV*, J. M. Bernardo, et. al., Oxford University Press, Oxford, 35-60.
3. Berger, J. O. and Pericchi, L. R. (1996). The Intrinsic Bayes Factor for Model Selection and Prediction, *Journal of the American Statistical Association*, 91, 109-122.
4. Doornbos, R. and Dijkstra, J. B. (1983). A Multi Sample Test for the Equality of Coefficients of Variation in Normal Populations. *Communication in Statistics-Simulation and Computation*, 12, 147-158.
5. Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press, New York.
6. Lohrding, R. K. (1969). A Test of Equality of Two Normal Means

Assuming Homogeneous Coefficients of Variation. *The Annals of Mathematical Statistics*, 40, 1374-1385.
7. Gupta, R. C. and Ma, S. (1996). Testing the Equality of Coefficient of Variation in k Normal Populations. *Communication in Statistics-Theory and Methods*, 25, 115-132.
8. Gupta, R. C., Ramakrishnan, S. and Zhou, X. (1999). Point and Interval Estimation of $P(X < Y)$: The Normal Case with Common Coefficient of Variation. *Annals of the Institute of Statistical Mathematics*, 51, 571-584.
9. Lee, H. C. and Kang, S. G. (2003). Reference Priors in the Normal Distributions with Common Coefficient of Variation. *Journal of Korean Data & Information Science Society*, 14, 697-705.
10. Miller, E. C. and Karson, M. J. (1977). Testing Equality of Two Coefficients of Variation. *American Statistical Association: Proceedings of the Business and Economics Section, Part I*, 278-283.
11. Nelson, W. (1990). *Accelerated Testing*, Wiley, New York.
12. O' Hagan, A. (1995). Fractional Bayes Factors for Model Comparison (with discussion), *Journal of Royal Statistical Society*, B, 57, 99-118.
13. Sinha, B. K., Rao, B. R. and Clement, B. (1978). Behrens-Fisher Problem under the Assumption of Homogeneous Coefficients of Variation. *Communication in Statistics-Theory and Methods*, 7, 637-656.
14. Rao, K. A. and Vidya, R. (1992). On the Performance of a Test for Coefficient of Variation. *Calcutta Statistical Association Bulletin*, 42, March and June, 165-166, 87-95.
15. Spiegelhalter, D. J. and Smith, A. F. M. (1982). Bayes Factors for Linear and Log-Linear Models with Vague Prior Information, *Journal of Royal Statistical Society*, B, 44, 377-387.

[received date : Sep. 2003, accepted date : Nov. 2003]