Title
Bayes estimation of the multiple correlation coefficient

Journal
Communication in Statistics- Theory and Methods, 18(4)

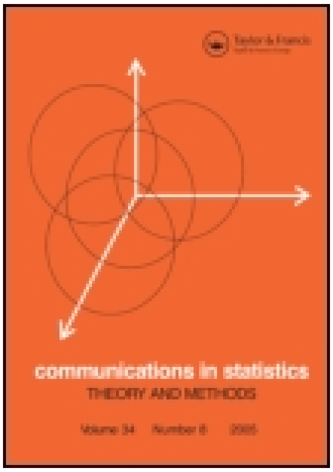
ISSN
0361-0926

Authors
Tiwari, Ram C
Chib, Siddhartha
Jammalamadaka, S Rao

Publication Date
1989

DOI
10.1080/03610928908829974

Peer reviewed



Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Bayes estimation of the multiple correlation coefficient

Ram C. Tiwari [a], Siddhartha Chib [b] & S. Rao Jammalamadaka [c]

[a] Department of Mathematics, University of North Carolina, Charlotte, NC, 28203

[b] Department of Economics, University of Missouri, Columbia, Columbia, MO, 65211

[c] Applied Probability and Statistics Program, University of California, Santa-Barbara, Santa-Barbara, CA, 93106

Published online: 27 Jun 2007.

COMMUN. STATIST.-THEORY METH., 18(4), 1401-1413 (1989)

BAYES ESTIMATION OF THE MULTIPLE CORRELATION COEFFICIENT

RAM C. TIWARI
Department of Mathematics
University of North Carolina, Charlotte
Charlotte, NC 28203

SIDDHARTHA CHIB
Department of Economics
University of Missouri, Columbia
Columbia, MO 65211

S. RAO JAMMALAMADAKA
Applied Probability and Statistics Program
University of California, Santa-Barbara
Santa-Barbara, CA 93106

Key Words and Phrases: Multiple correlation coefficient, sufficient statistic, Bayes estimation; Monte Carlo simulation.

ABSTRACT

Let $\bar{R}$ denote the population multiple correlation coefficient of one variable on the other (m-1), in a m-variate normal distribution. Bayes estimator of $\bar{R}^2$, given only the sample multiple correlation coefficient R^2, is derived with respect to the squared error loss function and a Beta prior distribution. These results are then related to the Bayes estimates of $\bar{R}^2/(1-\bar{R}^2)$, a parameter considered recently by Muirhead (1985). The ideas are illustrated and the effect of various parameters studied through numerical examples. A Monte Carlo study

indicates that the sampling mean squared error of the Bayes estimator is lower than that of R^2, for plausible prior distributions.

1. INTRODUCTION

Let $\underset{\sim}{X} = (X_1,\ldots,X_m)'$ have the m-variate normal $N_m(\underset{\sim}{\mu},\Sigma)$ distribution, where $\underset{\sim}{\mu}$ and Σ are unknown. Let $\bar{R}$ denote the *population* multiple correlation coefficient between X_1 and $\underset{\sim}{X}_2 = (X_2,\ldots,X_m)'$ given by

$$\bar{R} = [1 - \mathrm{Var}(X_1|\underset{\sim}{X}_2)/\mathrm{Var}(X_1)]^{1/2}$$

$$= [(\underset{\sim}{\sigma}_{12}' \Sigma_{22}^{-1} \underset{\sim}{\sigma}_{12})/\sigma_{11}]^{1/2} ,$$

where $\mathrm{Var}(X_1) = \sigma_{11}$, $\mathrm{Cov}(\underset{\sim}{X}_2) = \Sigma_{22}$ and $\underset{\sim}{\sigma}_{12}$ is the $(m-1)\times 1$ vector of covariances between X_1 and each of the variables in $\underset{\sim}{X}_2$. Suppose we observe independent and identically distributed observations from $N_m(\underset{\sim}{\mu},\Sigma)$. Let the i^{th} data vector be $\underset{\sim}{X}_i = (X_{1i},\ldots,X_{mi})'$, $i=1,\ldots,N$. Define

$$A = \sum_{i=1}^{N} (\underset{\sim}{X}_i-\bar{\underset{\sim}{X}})(\underset{\sim}{X}_i-\bar{\underset{\sim}{X}})',$$

where $\bar{\underset{\sim}{X}} = \frac{1}{N}\sum_{i=1}^{N} \underset{\sim}{X}_i$ is the sample mean vector. Partition A as

$$A = \begin{bmatrix} a_{11} & \underset{\sim}{a}_{12}' \\ \underset{\sim}{a}_{12} & A_{22} \end{bmatrix}$$

where A_{22} is $(m-1)\times(m-1)$. The *sample* multiple correlation coefficient between X_1 and $\underset{\sim}{X}_2$ is defined as

$$R = [(\underset{\sim}{a}_{12}' A_{22}^{-1} \underset{\sim}{a}_{12})/a_{11}]^{1/2} .$$

The sampling distribution of R^2 has been studied extensively (cf. e.g. Anderson, 1984, pp. 143-146, and Muirhead, 1982, pp. 171-177), and is provided below in equation (2.2).

Surprisingly, there is only a limited literature on the Bayes estimation of $\bar{R}^2$. Under the assumption of a diffuse prior, Geisser (1965) derived the posterior distribution of $\bar{R}^2$. In the regression context with $\underset{\sim}{X}_2$ non-random, Press and Zellner (1978), study the posterior distributions of $\bar{R}^2$ using diffuse and natural conjugate prior distributions. In this paper, we extend the original work of Geisser (1965) and show how an informative, Beta prior analysis of $\bar{R}^2$ can be conducted.

The plan of the paper is as follows. In Section 2, the posterior probability density function (pdf) of $\bar{R}^2$, and the Bayes estimator of $\bar{R}^2$ under a squared error loss is derived. We also discuss the Bayes estimation of the related parameter $\theta = \bar{R}^2/(1-\bar{R}^2)$ that is considered recently by Muirhead (1985). Finally, in Section 3, some numerical results are provided including plots of the posterior distribution of $\bar{R}^2$ for different parameter values. We also carry out a Monte Carlo simulation to compare the sampling properties of the Bayes estimator and R^2. A word about the notation. Throughout we take liberty with the commonly used notation and employ the same symbol for a random variable and its realization. For example, R^2 is used for the random variable as well as its sample realization.

2. BAYES ESTIMATION

Let R be the sample moment multiple correlation coefficient between X_1 and $\underset{\sim}{X}_2$ based on a sample $\underset{\sim}{X}_i = (X_{1i},\ldots,X_{mi})$, $i=1,\ldots,N$, of size $N=n+1$ from $N_m(\underset{\sim}{\mu},\Sigma)$. The parameter of interest is the population multiple correlation coefficient $\bar{R}$.

The distribution of $\bar{R}^2$ can be obtained through the following approach (see Muirhead, 1985). Let K, V_1 and V_2 be random variables such that

(i) K has a negative binomial distribution with parameters n/2 and $\bar{R}^2$; the probability function of K being

$$P(K=k|\bar{R}^2) = [kB(k,\tfrac{n}{2})]^{-1}(\bar{R}^2)^k(1-\bar{R}^2)^{\frac{n}{2}}, \qquad k = 0,1,\ldots \quad (2.1)$$

where $B(\alpha,\beta) = \Gamma(\alpha).\Gamma(\beta)/\Gamma(\alpha+\beta)$, $\alpha,\beta \geq 0$,

(ii) The conditional distribution of V_1, given K=k and $\bar{R}^2$ is a chi-squared with (m-1+2k) degrees of freedom, independent of $\bar{R}^2$.

(iii) The random variable V_2 is independent of (K,V_1), and V_2 has a chi-squared distribution with (n-m+1) degrees of freedom.

Then, the random variable R^2 is distributed as $V_1/(V_1+V_2)$. To put it differently, the two experiments (of observing) R^2 and $V_1/(V_1+V_2)$ are equivalent. Thus, we may say that there is an underlying random quantity K such that conditional on K=k, R^2 is distributed as Beta (type I) distribution with parameters $(\frac{m-1+2k}{2})$ and $(\frac{n+1-m}{2})$, that is with pdf

$$g(R^2|\bar{R}^2,K=k) = [B(\tfrac{m-1+2k}{2},\tfrac{n+1-m}{2})]^{-1} R^{2^{\frac{m-1+2k}{2}-1}} (1-R^2)^{\frac{n+1-m}{2}-1}, \quad (2.2)$$

independent of $\bar{R}^2$. From (2.2) we observe that the distribution of R^2, given K=k and $\bar{R}^2$, depends on the parameter of interest $\bar{R}^2$, only through this underlying random quantity K. This is equivalent to saying that K would be a sufficient statistic for $\bar{R}^2$ if the data were (K,R^2). Since K has the distribution given in (2.1), it follows that the family of Beta (type I) distributions is a conjugate family of priors for $\bar{R}^2$.

Thus, to proceed with the Bayesian analysis of $\bar{R}^2$, we can employ a Beta (type I) prior distribution for $\bar{R}^2$ with pdf

$$\pi(\bar{R}^2|K=k,n) = [B(k+1, \tfrac{n}{2}+1)]^{-1}(\bar{R}^2)^{k/2}(1-\bar{R}^2)^{n/2}, \qquad 0 < \bar{R}^2 < 1$$

or more generally,

$$\pi(\bar{R}^2) = [B(\alpha,\beta)]^{-1}(\bar{R}^2)^{\alpha-1}(1-\bar{R}^2)^{\beta-1}, \quad 0 \leq \bar{R}^2 \leq 1, \quad \alpha,\beta > 0, \quad (2.3)$$

where α and β are hyperparameters that are assumed known. Value of α and β can be selected to represent different prior beliefs about $\bar{R}^2$.

Now from (2.2), the likelihood function of $\bar{R}^2$, given R^2, is obtained by averaging over the distribution of K, which yields

$$\ell(\bar{R}^2|R^2) = [B(\tfrac{m-1}{2},\tfrac{n-1+m}{2})]^{-1} \{(R^2)^{\frac{m-1}{2}-1} (1-R^2)^{\frac{n-m+1}{2}-1}\} \quad (2.4)$$

$$X\ (1-\bar{R}^2)^{n/2}\ {}_2F_1(\tfrac{n}{2},\ \tfrac{n}{2};\ \tfrac{m-1}{2};\ (R\bar{R})^2)\ ,$$

where, for positive integers p, q, and real z,

$${}_pF_q(\alpha_1,\ldots,\alpha_p;\ \beta_1,\ldots,\beta_q;\ z)$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma(\alpha_1+r)\ldots\Gamma(\alpha_p+r)\Gamma(\beta_1)\ldots\Gamma(\beta_q)}{r!\Gamma(\alpha_1)\ldots\Gamma(\alpha_p)\Gamma(\beta_1+r)\ldots\Gamma(\beta_q+r)}\ z^r \quad (2.5)$$

is the generalized hypergeometric function. In (2.5), if $p=q+1$, the series converges for $|z|<1$ and diverges for $|z|>1$. Hence, from (2.3), (2.4) and the usual Bayes formula, the posterior pdf of $\bar{R}^2$, given R^2, is

$$\pi(\bar{R}^2|R^2) = \ell(\bar{R}^2|R^2)(\bar{R}^2)^{\alpha-1}(1-\bar{R}^2)^{\beta-1}/\int_0^1 \ell(\bar{R}^2|R^2)(\bar{R}^2)^{\alpha-1}(1-\bar{R}^2)^{\beta-1}d\bar{R}^2$$

$$= [B(\alpha,\tfrac{n}{2}+\beta)]^{-1}\left\{(\bar{R}^2)^{\alpha-1}(1-\bar{R}^2)^{\frac{n}{2}+\beta-1}\right\} .$$

$${}_2F_1(\tfrac{n}{2},\tfrac{n}{2};\tfrac{m-1}{2};(R\bar{R}^2)/{}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha;\tfrac{m-1}{2},\tfrac{n}{2}+\alpha+\beta;R^2),\quad 0 \le \bar{R}^2 < 1. \quad (2.6)$$

From (2.6) it is clear that the posterior pdf of $\bar{R}^2$, given R^2, is a weighted sum of Beta (type I) densities. The Bayes estimator of $\bar{R}^2$ with respect to the squared error loss function is

$$\bar{R}_n^2 = E(\bar{R}^2|R^2)$$

$$= \alpha(\alpha + \beta + \tfrac{n}{2})^{-1} \; {}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha + 1; \tfrac{m-1}{2}, \tfrac{n}{2} + \alpha + \beta + 1; R^2)$$

$$/{}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha; \tfrac{m-1}{2},\tfrac{n}{2} + \alpha + \beta; R^2) \qquad (2.7)$$

and the variance of $\bar{R}^2$, given R^2, is

$$V(\bar{R}^2|R^2) = \alpha(\alpha+1)\{(\alpha + \beta + \tfrac{n}{2})(\alpha + \beta + \tfrac{n}{2} + 1)\}^{-1} \, .$$

$${}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha + 2;\tfrac{m-1}{2},\tfrac{n}{2} + \alpha + \beta + 2;R^2)/{}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha;\tfrac{m-1}{2},\tfrac{n}{2} + \alpha +\beta;R^2)$$

$$- \bar{R}_n^2 \, . \qquad (2.8)$$

It should be noted that (2.7) and (2.8) require the computation of (three) ${}_3F_2$ functions. In our simulation exercises, some of which are reported in Section 3, we have found that the expression in (2.5) converges fairly quickly and often no more than 150 terms need to be included in the sum.

Remark 1. The results developed thus far can be readily adapted to provide the Bayes estimator of the parameter $\theta = \bar{R}^2/(1-\bar{R}^2)$. Muirhead (1985) considers the classical estimation of θ and showed that the best estimators of θ, including the unique minimum variance unbiased estimator, are linear functions of $Y=R^2/(1-R^2)$. The sampling distribution of Y has been considered by Gurland (1968) and Muirhead (1982). The sampling distribution can also be derived by using the fact that Y has the same distribution as V_1/V_2 and that there is a one-to-one correspondence between $\bar{R}^2$ and θ. Now, noticing that a Beta type I prior (2.3) on $\bar{R}^2$ corresponds to a Beta type II prior for θ with pdf

$$\pi(\theta) = [B(\alpha,\beta)]^{-1} \, \theta^{\alpha-1} \, (1 + \theta)^{-(\alpha+\beta)} \, , \qquad 0 \le \theta \le \infty \, , \quad (2.9)$$

the posterior pdf of θ (by a change of variable in equation (2.6) or directly) is

$$\pi(\theta|Y) = [B(\alpha,\tfrac{n}{2}+\beta)]^{-1}\{\theta^{\alpha-1}(1+\theta)^{-(\frac{n}{2}+\alpha+\beta)}\} \cdot$$

$$_2F_1(\tfrac{n}{2},\tfrac{n}{2};\tfrac{m-1}{2};(\tfrac{\theta}{1+\theta}))/{}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha;\tfrac{m-1}{2},\tfrac{n}{2}+\alpha+\beta;(\tfrac{Y}{1+Y})),$$

$$0 \le \theta < \infty \quad . \tag{2.10}$$

From (2.10), the Bayes estimator of θ, given Y, with respect to the squared error loss function, is

$$\hat{\theta}_n = \alpha(\tfrac{n}{2}+\beta-1)^{-1}\ {}_3F_2(\tfrac{n}{2},\tfrac{n}{2},\ \alpha+1;\ \tfrac{m-1}{2},\ \tfrac{n}{2}+\alpha+\beta;\ (\tfrac{Y}{1+Y}))\ / \tag{2.11}$$

$$_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha;\ \tfrac{m-1}{2},\ \tfrac{n}{2}+\alpha+\beta;\ (\tfrac{Y}{1+Y}))$$

and the variance of θ, given Y, is

$$V(\theta|Y) = \alpha(\alpha+1)\{\tfrac{n}{2}+\beta-1)(\tfrac{n}{2}+\beta-2)\}^{-1} \cdot$$

$$_3F_2(\tfrac{n}{2},\ \tfrac{n}{2},\ \alpha+2;\ \tfrac{m-1}{2},\ \tfrac{n}{2}+\alpha+\beta;\ (\tfrac{Y}{1+Y}))\ /$$

$$_3F_2(\tfrac{n}{2},\tfrac{n}{2},\alpha;\ \tfrac{m-1}{2},\ \tfrac{n}{2}+\alpha+\beta;\ (\tfrac{Y}{1+Y})) - [\hat{\theta}_n]^2\ , \tag{2.12}$$

where $\hat{\theta}_n$ is given by (2.11).

<u>Remark 2</u>: The assessment of the hyperparameters α and β. Since the parameters α and β of a Beta (type I) prior for $\bar{R}^2$ are unknown, one estimates α and β using the past data on R^2, say, by the method of moments or as suggested by a referee, by the method of maximum likelihood. Since the latter method involves a difficult maximization, we restrict attention to the method of moments approach. Since the expressions for the first two moments of R^2 given $\bar{R}^2$, are complicated, we shall instead estimate α and β from the first two moments of Y, given θ. These

are

$$E(Y|\theta) = \frac{n}{n-m-1}\theta + \frac{m-1}{n-m-1}$$

and

$$E(Y^2|\theta) = \frac{n(n+2)\theta^2 + n(2m+1)\theta + m(m-1)}{(n-m-1)(n-m-3)} .$$

Hence, the first two moments of the unconditional distribution of Y are

$$\mu_1 = E(Y) = \frac{n\alpha}{(n-m-1)(\beta-1)} + \frac{m-1}{n-m-1} \tag{2.13}$$

and

$$\mu_2 = E(Y^2) = \frac{1}{(n-m-1)(n-m-3)} \{\frac{n(n+2)\alpha(\alpha+1)}{(\beta-1)(\beta-2)} + \frac{n(2m+1)\alpha}{\beta-1} + m(m-1)\} . \tag{2.14}$$

From (2.13) and (2.14)

$$\alpha = \frac{\beta-1}{n} L \tag{2.15}$$

and from (2.15) and (2.15)

$$\beta = \frac{(n+2)L(L+n)}{\{\mu_2 nm-(n+2)L^2-2n(m-1)L-m(m-1)\}} + 2 , \tag{2.16}$$

where

$$L = \mu_1(n-m-1) - (m-1)$$

and

$$M = (n-m-1)(n-m-3) \tag{2.17}$$

Let $R_1^2,\ldots,R_t^2$ be the values of R^2 based on the past t independent samples, each of size $N = n + 1$, on $\underline{X}$. Let $Y_i = R_i^2/(1-R_i^2)$, $i=1,\ldots,t$, and

$$m_1 = \frac{1}{t}\sum_{i=1}^{t} Y_i \,, \qquad \text{and} \qquad m_2 = \frac{1}{2}\sum_{i=1}^{t} Y_i^2 \tag{2.18}$$

be the first and second sample moments of Y. Then, using m_1 and m_2 as estimators of μ_1 and μ_2, the estimates of α and β from (2.15) and (2.16) are

$$\hat{\alpha} = \frac{(\hat{\beta}-1)}{n}\hat{L} \tag{2.19}$$

and

$$\hat{\beta} = \frac{(n+2)\hat{L}(\hat{L}+n)}{\{m_2 nM-(n+2)\hat{L}^2-2n(m-1)\hat{L}-m(m-1)\}} + 2 \quad , \tag{2.20}$$

where from (2.17) and (2.18) $\hat{L} = m_1(n-m-1)-(m-1)$. Thus, when α and β are unknown, the empirical Bayes estimator of $\bar{R}^2$ at the (t+1)th stage, based on $R_1^2,\ldots,R_t^2$, is given by

$$\hat{\bar{R}}^2_{n,t+1} = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}+\frac{n}{2}} \; \frac{{}_3F_2(\frac{n}{2},\frac{n}{2},\hat{\alpha}+1;\frac{m-1}{2},\frac{n}{2}+\hat{\alpha}+\hat{\beta}+1;R^2_{t+1})}{{}_3F_2(\frac{n}{2},\frac{n}{2},\hat{\alpha};\frac{m-1}{2},\frac{n}{2}+\hat{\alpha}+\hat{\beta};R^2_{t+1})} \,, \tag{2.21}$$

where R_{t+1} denotes the sample correlation coefficient based on a sample of size N on $\underline{X}$ at (t+1)th stage. The procedures described earlier can be used to find the variance of the empirical Bayes estimator $\bar{R}^2_{n,t+1}$, and the confidence intervals for $\bar{R}^2$, given $R_1^2,\ldots,R_{t+1}^2$. We may remark here that merely plugging the estimates $\hat{\alpha}$ and $\hat{\beta}$ in (2.21) gives a naive empirical Bayes estimator, which quite frequently under estimates the corresponding variance.

3. NUMERICAL EVALUATION

As with all Bayes procedures, it is important to examine the sensitivity of the posterior distribution and the Bayes estimator, to the choice of prior parameters, (α,β), and the sample size, N. We picked two sets of values (α,β) namely (2,6) and (6,2). For the choice of $R^2 = 0.6$, $m = 3$ and $n = 10, 20, 30, 40, 50$, we provide in Tables 3.1 and 3.2 some summary characteristics of the posterior pdf of $\bar{R}^2$.

TABLE 3.1: SUMMARY CHARACTERISTICS OF $\pi(\bar{R}^2|R^2)$

	$R^2=0.6$	$m=3$	$\alpha=2$	$\beta=6$	
n / Posterior measure	10	20	30	40	50
mean	0.3082	0.3897	0.4405	0.4730	0.4950
mode	0.2800	0.3900	0.4500	0.4900	0.5100
variance	0.0201	0.0167	0.0130	0.0103	0.0084

TABLE 3.2: SUMMARY CHARACTERISTICS OF $\pi(\bar{R}^2|R^2)$

	$R^2=0.6$	$m=3$	$\alpha=6$	$\beta=2$	
n / Posterior measure	10	20	30	40	50
mean	0.6450	0.6278	0.6198	0.6153	0.6125
mode	0.6800	0.6500	0.6400	0.6300	0.6300
variance	0.0147	0.0107	0.0084	0.0069	0.0059

Note: The mean and variance are computed using expressions (2.7) and (2.8). For the given parameter values, it was found necessary to include only 140 terms in the evaluation of the ${}_3F_2$ function appearing in (2.6). To avoid an overflow, the typical term of (2.6) was first logged and then exponentiated. The mode was computed through a global grid search of the posterior pdf (2.6) as $\bar{R}^2$ varies from .01 to .99 in increments of .01. A finer grid was not thought necessary for the point being made.

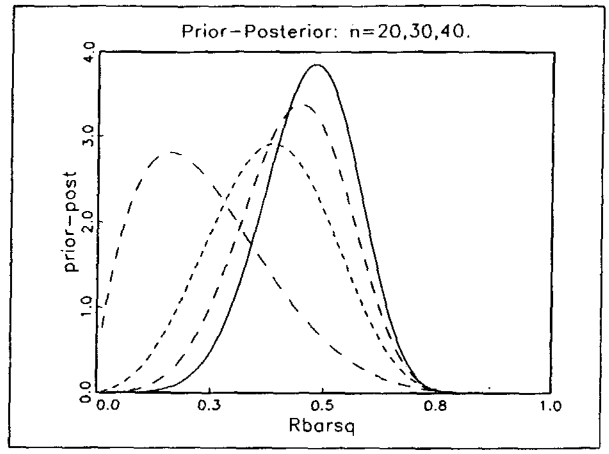


Figure 3.1. Examples of posterior densities of $\bar{R}^2$ for $m=3, R^2=.6, \alpha=2, \beta=6$ with $n=20,30,40$. Legend: n=20,----- ; n=30,'-'-'-; n=40,—______.

These tables illustrate in a concise way the effect of the prior information and sample size on the posterior of $\bar{R}^2$. When $\alpha=2$ and $\beta=6$, the posterior mean of $\bar{R}^2$ increases towards the sample value 0.6 with n, whereas when $\alpha=6$ and $\beta=2$, the posterior mean of $\bar{R}^2$ decreases towards 0.6 with n. As the sample size n increases, the mode of the posterior gets closer to the mean indicating a tendency towards symmetry. Also, as n increases, the curves become more peaked and concentrated towards the center as corroborated by the fact that the variance of the posterior distribution approaches zero as n tends to infinity. These facts are also revealed by the plots of the prior and posterior pdfs in Figure 3.1.

We also report some simulation results related to the sampling distribution of R^2 and $\bar{R}_n^2$. The true data was generated from the trivariate normal distribution with mean zero and variance

$$\Sigma = \begin{bmatrix} 1 & .6 & .36 \\ .6 & 1 & .6 \\ .36 & .6 & 1 \end{bmatrix}.$$

TABLE 3.3. A SAMPLING COMPARISON: N=20, M=3, 100 REPLICATIONS

	R^2	$\bar{R}_n^2$			
		$\alpha=2,\beta=6$	$\alpha=3,\beta=3$	$\alpha=4,\beta=3$	$\alpha=6,\beta=2$
Bias	0.0309	-0.0774	0.0462	0.0858	0.1739
Variance	0.0274	0.0064	0.0083	0.0071	0.0057
MSE	0.0284	0.0124	0.0105	0.0145	0.0360

Note: The sampling bias and variance are computed using 100 replications for each pair of α,β values. The sampling MSE is the bias squared plus the variance.

Thus the population multiple correlation, $\bar{R}^2$, is .36. 100 replications of size N=20 are drawn from this distribution. The sampling bias, variance, and sampling mean square error (MSE) of the two estimators of $\bar{R}^2$ is reported below in Table 3.3.

The table above shows that with prior information represented by $\alpha = 2$ and $\beta = 6$, the Bayes estimate is downward biased. This is quite reasonable given that the prior mean of 0.25 is below the true $\bar{R}^2$ value of 0.36. Further, the Bayes estimate, which is more biased than the classical estimate, possesses a lower sampling mean square error than R^2. Similar observations can be made about the other cases shown above. Interestingly, when the prior information is implausible relative to the true value of $\bar{R}^2$ (i.e., when $\alpha = 6$, $\beta = 2$), the Bayes estimate is considerably biased. This increases the MSE of $\bar{R}_n^2$ above that of R^2. We can conclude that as long as R^2 is near $E(\bar{R}^2)$, the prior expectation, the Bayes estimate $\bar{R}_n^2$ will possess a lower MSE than the estimator R^2.

ACKNOWLEDGEMENTS

We would thank John Gurland, and two anonymous referees, for suggestions that have substantially improved the paper.

BIBLIOGRAPHY

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. Second Edition, John Wiley, New York.

Geisser, S. (1965). Bayesian estimation in Multivariate Analysis. *Ann. Math. Statis.* 36, 150-159.

Gurland, J. (1968). A relative simple form of the distribution of the multiple correlation coefficient. *Biometrika* 24, 823-834.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. John Wiley, New York.

Muirhead, R.J. (1985). Estimating a particular function of the multiple correlation coefficient. *J. Amer. Statist. Assoc.* 80, 923-925.

Olkin, I. and Pratt, J.W. (1958). Unbiased estimation of certain correlation coefficients. *Ann. Math. Statist.* 29, 201-211.

Press, S.J. and Zellner, A. (1978). Posterior distribution for the multiple correlation coefficient with fixed regressors. *J. Econometrics* 8, 307-321.

Received by Editorial Board Member April 1988; Revised November 1988.

Recommended by John Gurland, University of Wisconsin, Madison, WI.

Refereed Anonymously.