ESTIMATION OF A RESTRICTED VARIANCE RATIO

BY

JAN E. GELFAND

TECHNICAL REPORT NO. 410
DECEMBER 14, 1986

PREPARED UNDER CONTRACT
NGO014-86-C-0156 (NR-042-267)
FOR THE OFFICE OF NAVAL RESEARCH

Reproduction in whole or in part is permitted for any purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

88 12 28 079
ESTIMATION OF A RESTRICTED VARIANCE RATIO

BY

ALAN E. GELFAND

TECHNICAL REPORT NO. 410
DECEMBER 14, 1988

Prepared Under Contract
N00014-86-K-0156 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted for any purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ABSTRACT

The estimation of a variance ratio $\hat{\tau} = \tau_2/\tau_1$ is studied under restrictions $\hat{\tau} \geq \hat{\tau}_0$ or $\hat{\tau} \leq \hat{\tau}_0$. We assume first that we observe $U_i \sim \chi^2_{n_i}$, independent. A Bayesian viewpoint is taken. We then assume additional information on $\tau_2$ is available in the form of an independent observation from a noncentral $\chi^2$ distribution. A natural application arises when the $W_i$ are sums of squares in a variance components model.

1. INTRODUCTION

In this paper we consider the estimation of a variance ratio over a truncated parameter space. Suppose that $U_i \sim \chi^2_{n_i}$, $i = 1, 2$, independent, $n_1 \geq 5$, $n_2 \geq 3$. Let $\hat{\tau} = \tau_2/\tau_1$. Under the scale invariant quadratic loss function

$$L(\hat{\tau}, \tau) = (\hat{\tau} - \tau)^2$$

the estimator

$$\hat{\tau}_0 = \frac{n_1 - 4}{n_2 + 2} \frac{U_2}{U_1}$$

is best invariant in the class based upon $U_2/U_1$. In fact, using the approach of Brown and Fox (1974), it is straightforward to establish that $\hat{\tau}_0$ is admissible under (1.1). (See Gelfand and Dey (1986) for details.) Suppose we restrict $\hat{\tau} = \hat{\tau}_0$ or $\hat{\tau} > \hat{\tau}_0$. Such restrictions arise naturally when, for example, the $U_i$ are sums of squares in a variance components model. Then $\hat{\tau}_0$
is no longer admissible. A usual approach to dominating $\delta_0$ is to restrict $\delta_0$ in the same fashion that $\theta$ is restricted. Such estimators are no longer smooth, hence inadmissible. From a Bayesian viewpoint, such an approach is essentially taking the posterior mode resulting from a prior over the restricted space. An alternative is to take the posterior mean. Such estimators will be admissible, but exhibit strange behavior. In Section 2 we look into all of these issues drawing upon ideas of Hill (1965) and recent work of Loh (1986).

A broader setting presumes that we have additional information about the $\tau_i$ in the form of $V_i \sim \tau_i X_i^2, \lambda_i^2$, or $V_i \sim (\tau_i + \zeta_i) \lambda_i^2$, $i = 1, 2$, $V_i$ independent of $U_i$ and of each other. Now $\delta_0$ is no longer admissible for $\theta$ under (1.1). An example is that of $X_{1i}$, $\sim N(\mu_{1i}, \tau_{1i})$, $i = 1, 2$, $j = 1, \ldots, n_i$, with $U_i = \sum(X_{1i} - \bar{X}_i)^2$. Then (1.2) is not admissible for $\theta$. In fact, $(\bar{X}_1, \bar{X}_2, U_1, U_2)$ is a version of the complete, sufficient statistic and the $\bar{X}_i$ (hence $\bar{X}_1^2$) contain information about $\tau_i$. (Gelfand and Dey (1986) discuss this example at length.) In this broader setting, we again seek to estimate a restricted $\delta$. This problem is the issue of Section 3, drawing upon ideas dating to Stein (1964), Brown (1968) and Klotz, Milton and Zacks (1969).

The seminal paper by Katz (1961) on admissibility for estimators of restricted parameters is inapplicable here since the distribution of $W = U_2/U_1$ does not belong to the exponential family.

2. THE BAYESIAN APPROACH

In the spirit of variance components models, let $\tau_1 = a_1 + b_2 \eta_2$, $\tau_2 = c_1 + d_2 \eta_2$ where $\eta_i > 0$, $a, b, c, d > 0$ and $r = ad - bc \neq 0$. We necessarily have $\theta_1 \leq \theta \leq \theta_2$ where $\theta_1 = \min(a^{-1}c, b^{-1}d)$, $\theta_2 = \max(a^{-1}c, b^{-1}d)$. If $a$ or $b = 0$ ($c$ or $d = 0$), we obtain a one-sided restriction below (above).

Example: In the balanced one-way ANOVA, i.e., $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, $i = 1, \ldots, I$, $j = 1, \ldots, J$, with $\alpha_i \sim N(0, \sigma_\alpha^2)$, $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$ all independent, we have $U_1 = \sum(Y_{ij} - \bar{Y}_i)^2$, $U_2 = \sum(Y_{ij} - \bar{Y})^2$, $n_1 = I(J-1)$, $n_2 = I-1$, $\eta_1 = \sigma_\epsilon^2$, $\eta_2 = \sigma_\alpha^2$, $a = 1$, $b = 0$, $c = 1$, $d = J$ and $\theta > 1$. 

2
We now develop the relevant distribution theory. We assume a prior over \( \eta_1, \eta_2 \) of the form \( \pi(\eta_1) \cdot \pi(\eta_2) \) with \( \eta_i^{-1} \) having a gamma distribution, i.e.,

\[
\pi(\eta_i) = \frac{-\lambda_i}{\eta_i}, \quad i = 1, 2.
\] (2.1)

Hill (1965, p. 811) argues for the plausibility of the independence assumption in the context of variance components. The resulting prior for \( \tau_1, \tau_2 \) on the domain \( \theta_1 \tau_1 \leq \tau_2 \leq \theta_2 \tau_2 \) is

\[
\pi(\tau_1, \tau_2) = \frac{-(k_1 + k_2 + 2)}{(d_{11} - b_{12})} \cdot \frac{-(k_1 + 1)}{(a_{12} - c_{11})} \cdot \exp \left\{ -\frac{\lambda_1 (a_{12} - c_{11}) + \lambda_2 (d_{11} - b_{12})}{(d_{11} - b_{12}) (a_{12} - c_{11})} \right\}
\] (2.2)

The noninformative prior on \( \eta_1 \) arises as the limiting case \( k_1 = 0, \lambda_1 = 0 \) in (2.1) and induces \( [(d_{11} - b_{12}) \cdot (a_{12} - c_{11})]^{-1} \) as the prior for \( \tau_1, \tau_2 \). This differs from the noninformative prior of Box and Tiao (1973, p. 253), \( (\tau_1, \tau_2)^{-1} \) which cannot arise from independent \( \eta_i \) unless the transformation from \( \eta_i \) to \( \tau_i \) is trivial and has been criticized in the variance components case for its dependence in the \( \eta \) space upon the sample size \( J \).

From (2.2)

\[
\pi(\theta) \propto \frac{(a-c)}{(d-b \theta)} \cdot \frac{a_{12} - c_{11}}{(a_{12} - c_{11})} \cdot \frac{k_1 + k_2 - 2}{k_1 - 1} \cdot \frac{k_2 - 1}{k_2 - k_1 - k_2 + 2} \cdot \frac{\lambda_1 (a_{12} - c_{11}) + \lambda_2 (d_{11} - b_{12})}{(d_{11} - b_{12}) (a_{12} - c_{11})} \cdot \theta^k_1 \cdot \frac{(\theta - \theta_0)}{(1 + \lambda (\theta - \theta_0)^2)}
\] (2.3)

i.e., \( \theta \) follows a generalized Beta distribution. Two cases of (2.3) which we study in greater detail are:

\[
\bar{\pi}(\theta) \propto \frac{(\theta - \theta_0)}{(1 + \lambda (\theta - \theta_0)^2)} \cdot k_2^{-1} \cdot \theta^{k_1-1}, \quad 0 \leq \theta \leq \theta_0.
\] (2.4)

i.e., \( \lambda(\theta - \theta_0) \) follows a nonstandardized F distribution or limit thereof and

\[
\pi(\theta) \propto \frac{(\theta - \theta_0)}{(1 + \lambda (\theta - \theta_0)^2)} \cdot \theta^{k_1-1} \cdot (\theta - \theta_0)^{-k_2+2}, \quad 0 \leq \theta \leq \theta_0.
\] (2.5)

i.e., \( \theta/\theta_0 \) follows a Beta distribution or limit thereof.

Under (2.1) the posterior distribution of \( \theta \) given the data is
\[ \xi(\varepsilon | U_1, U_2) = \frac{\frac{n_1 + n_2}{2} + k_2 - 1}{(d - b\varepsilon) (a\varepsilon - c)} \frac{1}{\frac{n_1 + n_2}{2} + k_1 - 1} \frac{k_1 + k_2 - 2}{r} \]

\[ \frac{n_2^2}{\varepsilon^2} \left[ \frac{(U_2 - U_1)(d - b\varepsilon)(a\varepsilon - c)}{2} + \lambda_1 (a\varepsilon - c) + \lambda_2 (d - b\varepsilon) \right] \frac{n_1 + n_2}{2} + k_1 + k_2 \]

(2.6)

Since (2.6) is analytically intractable, we consider instead the simpler posterior of \( \theta | W = U_1/U_2 \). The distribution of \( W | \theta \) is a nonstandardized \( F \), i.e.,

\[ f(w | \theta) = \frac{n_1 + n_2}{\Gamma(\frac{n_1}{2}) \Gamma(\frac{n_2}{2})} \frac{n_1}{\theta^2} \frac{n_2}{(1 + \theta w)^{n_1 + n_2}} \]

whence under (2.1)

\[ \xi(\theta | w) \propto \frac{k_1 - 1}{(a\varepsilon - c)} \frac{k_2 - 1}{(d - b\varepsilon)} \frac{n_1/2}{k_1 + k_2 - 2} \frac{\theta^r}{1 + \theta w} \frac{n_1 + n_2}{2} \]

In the special case (2.4) with \( \lambda = \varepsilon_0^{-1} \), we obtain

\[ \xi(\varepsilon | w) \propto [\lambda_1 (a\varepsilon - c) + \lambda_2 (d - b\varepsilon)] (1 + \theta w)^{-2} \]

(2.7)

(2.8)

In the special case (2.5), we obtain

\[ \xi(\varepsilon | w) \propto \frac{n_1}{\varepsilon_0^{-2}} - (k_1 + k_2) \frac{(\varepsilon - \varepsilon_0)}{\varepsilon_0^{-2}} \frac{n_1 + n_2}{2}, \quad 0 < \varepsilon < \varepsilon_0. \]

(2.9)

The case \( k_1 = 1, k_2 = 0 \) in (2.9) or \( k_1 = 0, k_2 = 1 \) in (2.10) produces the posterior associated with the Box-Tiao prior.

Turning to estimation of \( \theta \), we first consider the posterior mode in (2.9) and (2.10). If \( k_1 < 1 \) in (2.9) or \( k_2 < 1 \) in (2.10), the mode occurs at \( \varepsilon_0 \). If \( k_1 > 1 \) in (2.9) or \( k_2 > 1 \) in (2.10), the posterior need not be unimodal. However, if in (2.9), \( k_1 = 1 \) and \( k_2 < (n_1 - 2)/2 \), we obtain a unique mode at
\[ d_{k_2, \theta_0} = \max(\theta_0, \frac{1}{\theta_0^2 n_2^2 + 2(k_2 + 1)}). \]  \hspace{1cm} (2.11)

If in (2.10), \( k_2 = 1 \) and \( k_1 < \frac{n_2 + 2}{2} \), we obtain a unique mode at

\[ d_{k_1, \theta_0} = \min(\theta_0, \frac{1}{\theta_0^2 n_2^2 + 2(k_1 - 1)}). \]  \hspace{1cm} (2.12)

Note that \( k_1 = 1 \) in (2.4), \( k_2 = 1 \) in (2.5) asserts prior information concentrated near \( \theta_0 \).

As remarked earlier (2.11) and (2.12) are not admissible. In particular for \( k_2 < \frac{(n_1 + n_2 - 2)(n_2 + 2)}{2} \), \( \max(\theta_0, \delta_0) \), \( \delta_0 \) as in (1.2) dominates (2.11) using the fact that \( \delta_0 \) is best invariant in the unrestricted problem and Lemma 1 of the appendix. However, \( \min(\epsilon_0, \epsilon_0^*) \) does not dominate (2.12) by using the same argument as in Loh (1986, p. 700).

Consider now formal Bayes rules extending the loss (1.1) to

\[ L(\theta, a) = \theta^c(\theta - a) \]  \hspace{1cm} (2.13)

where \( c \) is arbitrary.

We recall that if \( 0 < a < b \)

\[ \int_0^{\theta_0} \frac{\theta^{a-1}}{(1 + \theta)^{b}} d\theta = \frac{\Gamma(a)\Gamma(b-a)}{\Gamma(b)} \eta_0 \]  \hspace{1cm} (2.14)

where \( \eta_0 \) is the incomplete Beta function evaluated at

\[ \eta_0 = (1 + \theta_0)^{-1} \theta_0 \theta. \]

Consider first the case \( \theta \geq \theta_0 \). Under (2.9) with loss (2.13), let \( m_j = n_1/2 + c + 1 - (k_1 + k_2) + j \) and \( \epsilon_j = (n_1 + n_2)/2 - m_j \). If \( k_1 \) is a positive integer and \( m_0 > 0 \) using (2.14), we can obtain the unique Bayes rule as

\[ \tilde{d}_{k_1, k_2, c, \theta_0} = \sum_{j=0}^{k_1 - 1} \frac{\Gamma(m_j + 1)\Gamma(\epsilon_j - 1)I_{\eta_0}(\epsilon_j, m_j)}{\Gamma(m_j)\Gamma(\epsilon_j)I_{\eta_0}(\epsilon_j, m_j)} \]  \hspace{1cm} (2.15)

Similarly, if \( 0 < \theta \leq \theta_0 \), under (2.10) with loss (2.13), let \( m_j = n_1/2 + c + k_1 + j \) and \( \epsilon_j = (n_1 + n_2)/2 - m_j \). If \( k_2 \) is a positive
integer using (2.14), we can obtain the unique Bayes rule as

$$
\delta_{k_1, k_2, c, \theta_0} \left( \begin{array}{c}
\frac{k_2-1}{k_2} \frac{k_2-1}{k_2} \\
\sum_{i=0}^{k_2-1} i \Gamma(m_i + 1) \Gamma(l_i + 1) \Gamma(m_i + l_i + 1) \\
\end{array} \right) 
$$

(2.16)

**Remark 1:** Expression (2.15) depends upon $c$ and $k_2$ through $c - k_2 = \gamma_2$ so we can denote it by \( \delta_{k_1, \gamma_2, \theta_0} \); expression (2.16) depends upon $c$ and $k_1$ through $c + k_1 = \gamma_1$ so we can denote it by \( \delta_{\gamma_1, k_2, \theta_0} \).

**Remark 2:** Computation of (2.15) and (2.16) requires calculation of the incomplete Beta functions only for, say, the denominator by using the well-known relationship (see, e.g., Abramovitz and Stegun, 1965)

$$
I_x(c+1, d-1) = I_x(c, d) - \frac{\Gamma(c+d)}{\Gamma(c+1)\Gamma(d)} x^c (1-x)^{d-1}.
$$

**Remark 3:** Using Lemma 2 of the appendix, we may show that

$$
\lim_{w \to \infty} \delta_{k_1, \gamma_2, \theta_0} = 0
$$

In this sense the behavior of \( \delta \), \( \delta \) differs strikingly from that of \( \delta, \delta \) as \( W \) approaches the extremes of its domain. The latter have positive probability of equaling \( \theta_0 \). Loh's Theorem 3.1 shows a special case of this.

**Remark 4:** Paralleling (2.11) and (2.12) when \( k_1 = 1 \), we obtain
The estimators (2.15) and (2.16) are admissible under (1.1) within the class of rules based upon \( W \). However, admissibility in the larger class based upon \( U_1, U_2 \) (equivalently \( W, U_1 \)) is a more difficult question. The approach of Brown and Fox (1974) mentioned after (1.2) is not applicable to a restricted parameter space.

### 3. ADDITIONAL INFORMATION ABOUT \( \gamma_1 \)

Recalling the notation of Section 1, suppose \( V_i \sim \gamma_i \frac{2}{\gamma_i + 1} \). We have the following results whose proof is essentially contained in Stein (1964) or Strawderman (1974).

**Result 1:** In estimating \( \gamma_2 \) under squared error loss,
\[
\min\{(n_2+2)^{-1} U_2, (m_2+n_2+2)^{-1} (U_2^* + V_2^*)\} \text{ dominates } (n_2+2)^{-1} U_2.
\]

**Result 2:** In estimating \( \gamma_1 \) under squared error loss,
\[
\max\{(n_1-4) U_1^{-1}, (m_1+n_1-4)(U_1^* + V_1^*)^{-1}\} \text{ dominates } (n_1-4) U_1^{-1}.
\]
If instead $V_i \sim (\zeta_i + \eta_i)^2 \chi^2_{\lambda_i}$, Results 1 and 2 still hold. That is, we may think of $V_i$ arising from $V_i | Z_i = z_i \sim \zeta_i \chi^2_{\lambda_i}, Z_i$ where $Z_i \sim (\zeta_i/2\eta_i)^2 \chi^2_m$. Since Results 1 and 2 hold regardless of $Z_i$, they hold unconditionally. Klotz, Milton and Zacks (1969, p. 1392) allude to this idea in a special case.

Taking these ideas further to the estimation of an unrestricted variance ratio using essentially the proof of Theorem 3.1 in Gelfand and Dey (1986), we can show

**Result 3:** In estimating $\delta$ under squared error loss,

$$\zeta = \min(\delta, \frac{(n_1-4)(U_2+V_2)}{U_1} \cdot \frac{m_2+n_2+2}{m_1+n_1+2}) \leq \delta_0 \text{ dominates } \delta_0$$

$$\overline{\zeta} = \max(\delta, \frac{U_2}{n_2+2} \cdot \frac{m_1+n_1-4}{U_1+V_1}) \geq \delta_0 \text{ dominates } \delta_0.$$

By Lemma 1 of the appendix, we have immediately that with squared error loss under the restriction $\theta \geq \theta_0$,

$$\max(\zeta, \delta_0) \text{ dominates } \max(\delta, \delta_0) \quad (3.1)$$

and under the restriction $0 \leq \theta \leq \theta_0$,

$$\min(\overline{\zeta}, \delta_0) \text{ dominates } \min(\delta, \delta_0). \quad (3.2)$$

In Gelfand and Dey other estimators (e.g., using ideas of Brown, 1968) which dominate $\delta_0$ in the unrestricted problem are given. These estimators may be used to obtain results similar to (3.1) and (3.2). We omit the details.

**APPENDIX**

**Lemma 1:** Under squared error loss

(a) If $T$ dominates $U$ on $\theta_0 \leq \theta$ and $T \leq U$, then $\max(T, \delta_0)$ dominates $\max(U, \delta_0)$ on $\theta_0 \leq \theta$.

(b) If $T$ dominates $U$ on $0 \leq \theta \leq \theta_0$ and $T \geq U$, then $\min(T, \delta_0)$ dominates $\min(U, \delta_0)$ on $0 \leq \theta \leq \theta_0$.

**Proof:** The proof is essentially that of a lemma in Klotz, Milton and Zacks (1969, p. 1394).

**Lemma 2:** In the notation of (2.14),

$$\lim_{w \to 0} w^{-a} I_n(a, b-a) = \frac{\Gamma(b)}{\Gamma(a+1)\Gamma(b-a)} \delta^a_0 \quad (A.1)$$
or equivalently
\[ \lim_{\omega \to 0} \omega^a I_{1-a}(a,b-a) = \frac{\Gamma(b)}{\Gamma(a+1)\Gamma(b-a)} \theta_0^{-a}. \] 

(A.2)

**Proof:** Since the limit as \( \omega \to 0 \) of the left-hand side in (2.14) is \( \theta_0^a/a \), we obtain (A.1). But (A.2) follows by replacing \( \omega \) with \( \omega^{-1} \) and \( \theta_0 \) with \( \theta_0^{-1} \) in (A.1).

**REFERENCES**


**Title:** Estimation of a Restricted Variance Ratio

**Authors:** Alan E. Gelfand

**Performing Organization:** Department of Statistics, Stanford University, Stanford, CA 94305

**Controlling Office:** Office of Naval Research, December 14, 1988

**Contract or Grant Number:** N00014-86-K-0156

**Program Element, Project, Task, Area, Work Unit Numbers:** NR-042-267

**Number of Pages:** 12

**Security Classification:** UNCLASSIFIED

**Distribution Statement:** Approved for public release: distribution unlimited.

**Keywords:** Truncated parameter space, Admissibility, Bayesian viewpoint.

**Abstract:**

Please see following page.
20. ABSTRACT

The estimation of a variance ratio $\theta = \tau_2 / \tau_1$ is studied under restrictions $\theta \geq \theta_0$ or $\theta \leq \theta_0$. We assume first that we observe $U_i \sim \tau_1 \chi^2_{n_i}$, independent. A Bayesian viewpoint is taken. We then assume additional information on $\tau_1$ is available in the form of an independent observation from a noncentral $\chi^2$ distribution. A natural application arises when the $U_i$ are sums of squares in a variance components model.