Comment

William E. Strawderman

This article was a pleasure for me to read. I thank the authors for giving the time and effort required to present a thorough and efficient development of the key decision-theoretic aspects of variance estimation.

I would like to elaborate somewhat on the authors' comments concerning practical improvements by drawing some parallels with the problem of estimating a mean vector. Additionally, I would like to briefly discuss the problem of estimating variance components.

Consider the model of the second paragraph of Section 5. For simplicity I'll assume $\sigma^2 = 1$ when discussing estimation of μ with loss equal to $\| \delta - \mu \|^2$. In both problems (estimating μ and estimating σ^2), the potential gains in risk over the best equivariant estimator increase with the dimension p of μ. For the case of estimating μ, the James–Stein estimator

$$\left(1 - \frac{p-2}{X_2' X_2}\right) X_2$$

(here $p \geq 3$) has minimum risk equal to 2 when $\mu = 0$ rising as a function of $\| \mu \|^2$ to p as $\| \mu \|^2 \to \infty$. Hence the maximum relative savings is $(p-2)/p$ which increases from ⅓ to 1 as p increases. The larger the dimension p, the more one is able to "borrow strength" across coordinates. In this case, as well as for the estimator (5.1) of the variance, meaningful gains (for large p) will result only if our prior guess (of $\mu = 0$) for the true mean vector is accurate. In either case, if our guess is poor and $\| \mu \|$ is large, $X_2' X_2$ will be large with high probability. Hence both the James–Stein estimator and (5.1) will, with high probability, be equal (or close) to the respective best equivariant estimators and little improvement will result.

It is still possible to achieve meaningful gains in the two problems (for large p) even if we are unable to accurately guess the true value of μ. For example, if we (accurately) guess that all of the components are equal but can't guess what the common value is, the Lindley–Smith estimator of μ (see Lindley and Smith, 1972) will work well. This estimator shrinks all components of X_2 toward the average of these components and results in an estimator of μ with risk equal to 3 when all components of μ are exactly equal and increases to p as the variance of $\{\mu_1, \cdots, \mu_p\}$ tends to ∞. Similarly in the problem of estimating σ^2 an estimator similar to (5.1) but with

$$Y^2 = \| X_2 - \bar{X}_2 1 \|^2$$

and

$$\phi_S(Z) = \min\left(\frac{1}{\nu + 2}, \frac{1 + Z^2}{\nu + p + 1}\right)$$

will achieve meaningful gains for p large relative to ν provided the coordinates of μ are all equal. Again gains will not be substantial unless the prior guess is accurate.

Similar parallel estimators are possible in the two problems based for example on guesses that the mean vector μ lies in a given subspace and the relative gains will be large provided the co-dimension of the subspace is large and the prior guess is good.

Of course, the parallels in the two problems are imperfect. In the problem of estimating the mean vector, the dimensions of the estimator and estimand are also growing as is the minimax risk. In the variance estimation problem, the corresponding values are fixed. In the variance estimation problem if ν is fixed, $p \to \infty$, and $\mu = (0, \cdots, 0)$ the estimator (5.1) $\to \min(S^2/(\nu + 2), \sigma^2)$, and hence the relative savings in risk do not approach 1 even when the prior guess for μ is correct.

I believe it is also worthwhile to mention some work in variance component estimation that is directly related to Stein's method. Klotz, Milton and Zacks (1969) (KMZ) studied estimation of σ_e^2 and σ_a^2 for squared error loss in the balanced random effects one way layout.

Sufficiency reduces the problem to consideration of

$$Y \sim N(\mu, \sigma_e^2 + J\sigma_a^2/IJ),$$

$$S_a^2 \sim (\sigma_e^2 + J\sigma_a^2)\chi_{I-1}^2$$

and

$$S_e^2 \sim \sigma_e^2 \chi_{I(J-1)}^2,$$

where Y, S_a^2 and S_e^2 are independent. One aspect of this problem that is of interest as it contrasts with the fixed effects model of Section 5 is that location and scale invariance are not enough to produce a best equivariant estimator.

The fully invariant estimators of the σ_e^2 are of the form $S_e^2 g(S_a^2/S_e^2)$. "Reasonable" estimators in this

William E. Strawderman is Professor and Chair, Department of Statistics, Hill Center, Busch Campus, Rutgers University, New Brunswick, New Jersey 08903.