Comment

K. Brenda MacGibbon and Glenn E. Shorrock

Professors Maåtta and Casella have written an excellent account of the theory of variance estimation using the approach originally conceived by Stein (1964). We especially appreciate the way in which they follow the historical development of the problem, and their discussion of the conditional properties of frequentist procedures gives new insight into the nature of these procedures. Moreover, their presentation is not purely expository: the prior given in (4.12) is, we believe, new material and is clearly a necessary step in the further derivation of the decision theoretic properties of these intervals.

Invariance seems to play a small role in this paper, while the original work in the field, including that of Stein (1964) and Brewster and Zidek (1974), gave great importance to the invariant group structure of the problem. It's not clear to us whether invariance is just a way of restricting the class of possible estimators to a more manageable subclass or whether it is essential to the problem. Our own experience in the application of this approach to other problems has been mixed; in some cases invariance has been crucial, in others not.

In the estimation of the generalized variance, the invariant structure plays an important role. This is perhaps due to the complicated multivariate structure of the problem. Starting with a multivariate normal linear model in canonical form, a minimal sufficient statistic is (X, S), where X is a normally distributed $p \times k$ matrix with independent columns $X_i \sim N(\xi, \Sigma)$, S is a $p \times p$ Wishart matrix with n degrees of freedom such that $ES = n\Sigma$, X and S are independent, and Σ is positive definite. We seek a point estimate of the determinant $|\Sigma|$ of Σ with the quadratic loss function

$$L\{\varphi(X, S); \Sigma, \xi\} = |\Sigma|^{-2}\{\varphi(X, S) - \Sigma\}^2.$$

This problem is invariant under the transformations

$$X \to AX + B \quad S \to ASA',$$
$$\xi \to A\xi + B \quad \Sigma \to A\Sigma A',$$

where A is any nonsingular $p \times p$ matrix and B is any $p \times k$ matrix. Estimators that are equivariant under this group satisfy

$$\varphi(AX + B, ASA') = |A|^2\varphi(X, S)$$

and have the form $\varphi(S) = c|S|$ where c is a constant. Such estimators have constant risk (expected loss), which is minimized by taking $c = (n - p + 2)!/(n + 2)!$ (Selliah, 1964). Thus, in this problem,

$$\varphi_0(S) = \frac{(n - p + 2)!}{(n + 2)!}|S|,$$

as an estimator of $|\Sigma|$, plays a role analogous to $s^2/(n + 1)$ as an estimator of σ^2.

Using the ideas of Stein (1964) on variance estimation, exploiting zonal polynomials, and searching in a larger class than that of the affine equivariant estimators, Shorrock and Zidek (1976) showed that

$$\varphi(X, S) = \min\left\{\frac{(n-p+2)!}{(n+2)!}|S|, \frac{(n+k+2-p)!}{(n+k+2)!}|S + XX'|\right\}$$

has uniformly smaller risk than $[(n - p + 2)!/(n + 2)!]|S|$. The estimator $\varphi(X, S)$ is the analogue of Stein's estimator of σ^2. Shorrock and Zidek's proof depends heavily on invariance, as does that of Sinha (1976), who extended their results.

A problem in which invariance seems to play a minor role is the estimation of the parameter λ of an Inverse Gaussian distribution. The setup is as follows: Let $X_1, X_2, \cdots, X_n$ be a random sample from an Inverse Gaussian distribution with parameters μ and λ (denoted by $IG(\mu, \lambda)$) and density function

$$f(x; \mu, \lambda) = \sqrt{\lambda/2\pi x^3}\exp\left\{\frac{-\lambda}{2\mu^2 x}(x-\mu)^2\right\}, \quad x, \mu, \lambda > 0.$$

The mean of this distribution is μ and the variance is μ^3/λ. Writing $\bar{X} = (1/n)\sum_{i=1}^n X_i$ and $U = \sum_{i=1}^n (X_i^{-1} - \bar{X}^{-1})$, it is well known that $\bar{X} \sim IG(\mu, n\lambda)$, $U \sim \lambda^{-1}\chi^2_{n-1}$, and $\bar{X}$ and U are independent (Tweedie, 1957). The statistic $V = n\lambda(\bar{X} - 1)^2/\bar{X}$ has a χ^2_1 distribution when $\mu = 1$ (Shuster, 1968), which suggests that the ratio V/U can play a similar role in this problem to that of $Z = \sqrt{n}\bar{X}/S$ in the normal problem; there, a small value of Z indicates that μ is close to 0. In the inverse Gaussian case, a small value of V/U indicates that μ is close to 1. For the loss function $L(\delta, \lambda^{-1}) = (\delta - \lambda^{-1})^2$, the best estimator of λ^{-1}

K. Brenda MacGibbon and Glenn E. Shorrock are Professors, Department of Mathematics and Computer Science, University of Quebec at Montreal, C.P. 8888, Succursale "A", Montreal, Quebec, Canada H3C 3P8.