class include the UMVUE ($S_e^2/I(J-1)$) and the

$$\text{MLE}\left(\min\left(\frac{S_e^2}{I(J-1)}, \frac{S_e^2+S_a^2}{IJ}\right)\right).$$

KMZ use a version of Stein's method to show that the MLE dominates the UMVUE. The authors go on to show, again using Stein's method, that any equivariant estimator of σ_e^2 which is greater than the sum of squares of all observations divided by $IJ + 2$ (i.e., $[IJY^2 + S_a^2 + S_e^2]/(IJ + 2)$) with positive probability, is inadmissible. This last result implies that there is additional information about σ_e^2 in the overall mean Y even though the variance of Y is a multiple of $\sigma_e^2 + J\sigma_a^2$ and not of σ_e^2.

Estimation of σ_a^2, of course, involves the additional wrinkle that the

$$\text{UMVUE } E\left(\frac{1}{J}\left(\frac{S_a^2}{I-1} - \frac{S_a^2}{I(J-1)}\right)\right)$$

is negative with positive probability. KMZ use Stein's method to investigate dominance relations among several estimators and show that the overall mean can sometimes be used to construct improvements. Portnoy (1971) and others have constructed Bayes equivariant estimators with good sampling properties. Loh (1986) has studied the problem of estimating σ_a^2/σ_e^2 using similar methods.

In this setup, since there is only one degree of freedom for the grand mean, the likely improvement is small (once one has selected a good equivariant estimator). It is possible that larger gains could occur in higher way mixed models where several degrees of freedom are available for the mean vector.

Presumably extensions of Brown's and Zidek's methods can be applied in these models and improved confidence intervals can be constructed as well.

ADDITIONAL REFERENCES

KLOTZ, J. H., MILTON, R. C. and ZACKS, S. (1969). Mean square efficiency of estimators of variance components. *J. Amer. Statist. Assoc.* **64** 1383–1402.

LINDLEY, D. V. and SMITH, A. F. M. (1972). Bayes estimates for the linear model (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 1–41.

LOH, W.-Y. (1986). Improved estimators for ratios of variance components. *J. Amer. Statist. Assoc.* **81** 699–702.

PORTNOY, S. (1971). Formal Bayes estimation with application to a random effects model. *Ann. Math. Statist.* **42** 1379–1402.

Rejoinder

Jon M. Maatta and George Casella

To begin, we thank all discussants for their kind remarks and stimulating comments. This project was started to enhance our understanding of the topic, but also helped to improve our knowledge and perspective. As mentioned by several discussants, the scope of our work was limited. This work was an intentional decision, because our relatively narrow focus presented a reasonable size task, and allowed us a fuller understanding of one part of this complicated subject. Many of the discussants had similar concerns, and we will structure our rejoinder to respond to the major topics mentioned.

PRACTICAL CONSIDERATIONS

Though somewhat surprising to us, much concern was expressed over the magnitude of possible improvement. A major point was that the possible improvements in variance estimation seem small when compared to those possible in the estimation of means. This is true, but we feel that the improvement here is still worthy of consideration.

Berger expresses concerns about this and, in his inimitable way, anticipates some of our rejoinder. While the magnitude of improvement is small (as demonstrated by other discussants), it does increase in the generalized linear model case, which we do not consider "less realistic," but useful in practice. Very interesting calculations are provided by both Hwang and Rukhin, showing the limiting amount of improvement possible, approximately 25% in practical cases. Rather than interpreting these findings in the pessimistic way of Professor Hwang, however, we find more hope for future improvements (although we certainly agree that greater improvement seems possible in the estimation of means).

Some of our optimism is supported, and Hwang's pessimism negated, by the comments of George and Strawderman. They suggest that we have not yet fully exploited the structure of the problem. The risk (or interval length) improvement in variance estimation obtains when the means are close to the point to which we are shrinking. George and Strawderman each point out ways to shrink toward subspaces, and, further, George suggests that we can shrink toward multiple subspaces. Such estimators may provide substantial practical gains, since the region of improvement will be expanded. Another interesting possibility pointed out by George is that of using improved variance estimators in Stein-type estimators of a mean. Here, the lack of independence of mean and variance estimates is challenging theoretically, and there might be practical gains to be had. We find all of these suggestions exciting.

All concerns over practical improvement need to be evaluated in the light of the interesting comment of Strawderman, on the imperfect parallel between the mean and variance problems. Perhaps this is saying that we are expecting too much improvement in this one-dimensional problem, and comparisons to a multidimensional problem are somehow unfair.

LOSS FUNCTIONS

As in any decision-theory problem, the loss function is of great concern, and in the variance estimation problem we have great sensitivity of the solution to the loss function. In particular, quadratic loss has come under attack since Stein's 1964 paper, and Brown shows us quite convincingly that this loss is inappropriate. We agree that a more appropriate loss is Stein's loss (Brown's equation 1) and, as Rukhin points out, greater improvements (than those under squared error) are possible. We are also fascinated with the (formal) reasoning of Rukhin and the (intuitive) reasoning of Brown that both lead to the conclusion that the "usual" estimator of variance is too big. Bravo.

CONFIDENCE INTERVALS

There is less debate about loss functions in interval (or set) estimation than in point estimation, and we have been relatively happy with the loss functions described by Cohen. Cohen's table, showing the relationship between admissibility of both point and interval estimators under various losses, is very interesting. The table shows almost a complete dichotomy between "testing-type" losses, which use false-coverage probability, and "estimation-type" losses, which use length or distance measures. This should lead us to rethink our criteria for these estimation problems, and to realize that testing and estimation are very different procedures. In particular, we could be interested in *estimating confidence* in the set estimation problem, and a "testing-type" loss would be inappropriate.

Even though loss functions tend to move to the background in interval estimation problems, their consideration is still important. The importance of considering a criterion like false coverage is demonstrated by Brown. (The example given by Brown is discussed in detail by Madansky, 1962, as a response to the work of Pratt, 1961.) The intervals of Goutis, which are constructed starting from the minimum-length interval, can also be constructed from the shortest-unbiased interval.

THE ROLE OF INVARIANCE

The comments of MacGibbon and Shorrock, suggest some interesting questions. Invariance does result in a computational reduction so, for example, it can help in complicated problems like estimation of covariance matrices or generalized variances. Invariance is also exploited by Rukhin, who shows some interesting representations of Bayes estimators starting from invariant priors. He also describes a fascinating connection to the problem of estimating a positive mean, a problem dating back at least to Katz (1961).

Yet the role of invariance seems to be only a means to an end, giving us a defined structure within which to work. It is unclear where a search for improved estimators that totally ignore invariance considerations would lead. Furthermore, the gains from such estimators are not likely to be great, given the results of Proskin (1985). Brown's conjecture, that Proskin's admissibility results for point estimation can be extended to the invariant intervals of Shorrock and Goutis, is probably correct. That is, we believe that these intervals are admissible within the class of all intervals. The proof is quite difficult.

CONDITIONAL INFERENCE

Checking conditional properties of improved estimators is of supreme importance. Using such powerful tools as the Stein effect, and the Brewster–Zidek construction, constructing improved estimators is almost too easy. Conditional concerns help to keep these improvements "honest."

Our conditional beliefs are somewhere between those of Berger and Brown. We believe that negative bias is bad (that is, overstatement of actual confidence), and are in strong agreement with Brown's observations about the nature of the conditioning set. More precisely, we are quite concerned with procedures that allow "obvious" negatively biased conditioning sets (a sentiment expressed in Casella (1988), and labeled as "crucial" by Brown). With respect to positive bias (understatement of confidence), we are concerned with it, but not so much as Berger.

We are most happy with an interval estimator like $\langle I_\tau, \gamma_\tau(\bar{X}, S^2)\rangle$ (see the discussion surrounding equation (4.15)). Such an estimator has good conditional properties, being free from major conditional defects (both positive and negative bias). Whether it is *frequency valid*, as defined by Berger (1988), is of less concern to us, since the conditional behavior is quite good. In fact, the concept of *frequency validity*, which requires the reported confidence to be below the actual confidence, can lead to interval estimators that are positively biased.

We disagree with Berger about the value of the betting interpretation of confidence intervals. The concept of confidence (whether frequentist or Bayesian) can be quite elusive. When it is illustrated by betting, in telling an experimenter that it would cost \$95 if wrong and \$5 if right, the elusive concept of confidence is put on a common and understandable scale.

The estimated confidence approach discussed by Brown, attributed to Kiefer and Robinson, and also promoted by Berger, is very exciting. It represents a way to report confidence that should be acceptable to a frequentist and a conditionalist. The confidence procedures $\langle I_\tau, \gamma_\tau(\bar{X}, S^2)\rangle$ or $\langle I_{\mathrm{SU}}, \gamma_{\mathrm{SU}}(S^2)\rangle$ are both examples of such procedures.

OTHER ESTIMATION PROBLEMS

As previously mentioned, the scope of the paper was intentionally limited to review the estimation of one normal variance. Some related problems not covered in our review, and mentioned by several discussants are: estimation of covariance matrices, estimation of components of variance, and estimation of quantiles. These topics are all important, and were only excluded so that our review could be finished in finite time!

The variance component problem, in particular, has had little decision theoretic treatment. The comments of Strawderman, suggesting that practical improvements may be possible, are very interesting.

ACKNOWLEDGMENT

This research was supported by National Science Foundation Grant DMS-89-0039.

ADDITIONAL REFERENCES

KATZ, M. W. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Ann. Math. Statist.* **32** 136–142.

MADANSKY, A. (1962). More on length of confidence intervals. *J. Amer. Statist. Assoc.* **57** 586–589.