book leans more heavily toward CMA than does the latter, but neither one really exemplifies MMA.

SPECIFIC COMMENTS ON THE SCHERVISH REVIEW

1. Where Schervish discusses the development of power functions in multivariate analysis, it would have helped a bit if power were discussed in terms of how power functions are normally used in multivariate statistical practice, namely, from the viewpoint of someone trying to determine sample size for an experiment involving multiple (correlated) outcomes. How is this sample size selection problem best solved? There is not much discussion of this kind of question in the two books reviewed.

2. The author refers to Anderson's discussion of the Scheffé procedure (it was extended to the multivariate case by Bennett) for dealing with the multivariate Behrens-Fisher problem of testing equality of means in two normal populations with unequal variances, when he says, "Data is discarded with a vengeance." The issue here is that if we have M observations on one population, and N observations on another, and $M < N$, Scheffé suggests that we randomly match M observations from the two populations and discard the remaining $(N - M)$ in the matching process. Actually, all of the observations in each of the populations should be used to estimate each of the variances. If M and N are large there is of course no problem in ignoring $(N - M)$ observations in the testing. The only occasion when a problem arises is in the case of M and N small, and $M \ll N$. From the Bayesian point of view these types of issues never arise, at the tradeoff cost of having to develop prior information for the parameters.

3. Schervish suggests that "one other unfortunate feature of Section 5.5 is ... This test is simply not another example of the type of test proposed for the Behrens-Fisher problem." Here, Anderson suggests that we can test the hypothesis that two normal subvectors have equal means (with unequal covariance matrices, so that it is a Behrens-Fisher type problem) by forming the difference in the sample mean vectors, "and this statistic is most relevant to $(\mu^{(1)} - \mu^{(2)})$" (Anderson, page 178). This is a special case of the Scheffé/Bennett approach discussed in Item 2, above, for the case of $M = N$, where the two-sample problem is reduced to a one-sample problem by subtraction of the sample means.

4. Schervish's suggested alternative to a second principle of classical inference is a bit harsh. Although "unbiasedness" is not a particularly relevant property for situations in which we are going to have to estimate only once, or only a few times, and although unbiasedness is a property that violates the "likelihood principle," I believe that most any reasonable statistician who is in the position of having to recommend an estimator that will be close to the true value on the average, over many estimations of the same quantity, would find unbiasedness a compelling property when taken in conjunction with the requirement of small variance.

Summary

In summary, the review of these important books on multivariate analysis by Mark Schervish not only provides a helpful perspective from which to appreciate these contributions to our field, but also, is refreshing and entertaining.

Rejoinder

Mark J. Schervish

I wish to thank the discussants for taking the time to carefully read the review and offer their own views on the topics covered. They have each made it more informative and useful for the interested reader. Because some of the comments of the authors of the two books reviewed are in the way of rejoinder to my review, I will refrain from offering further commentary on those remarks. Much of the discussion provides the readers with brief overviews of areas that I failed to mention in my review. As my review already is a comment, at great length, on the work of many people, I will keep my comments on the discussion brief.

Because Professor Anderson's comments are almost entirely concerned with my review of his text, I will let him have the last word on the matter. I will thank him, however, for bringing to my attention part (b) of Problem 3 in Chapter 11 of his book, which indeed does suggest the predictive interpretation of principal components. A further suggestion of this interpretation appears in the paper of Kettenring (1971), whom I also thank for the reference.

Some of the discussants mention projection pursuit as a computationally intensive multivariate method that I did not discuss. Professor Goldstein remedies