In closing I note that the bootstrap can be an inconsistent procedure when it is employed to correct potentially large biases of some nonparametric techniques. See Section 11.7 of Breiman et al. (1984).

REFERENCES

ATHREYA, K. B. (1987). Bootstrap of the mean in the infinite variance case. To appear in *Ann. Statist.*

BABU, G. J. and SINGH, K (1983). Inference on means using the bootstrap. *Ann. Statist.* **11** 999–1003.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A. and STONE, C. J. (1984). *Classification and Regression Trees*. Wadsworth, Belmont, Calif.

EFRON, B. and TIBSHIRANI, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist. Sci.* **1** 54–75, 77.

GHOSH, M., PARR, W. C., SINGH, K. and BABU, G. J. (1984). A note on bootstrapping the sample median. *Ann. Statist.* **12** 1130–1135.

HARTIGAN, J. A. (1986). Comment on "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy" by Efron and Tibshirani. *Statist. Sci.* **1** 75–77.

OLSHEN, R. A., GILPIN, E. A., HENNING, H., LEWINTER, M. L., COLLINS, D. and ROSS, J., JR. (1985). Twelve month prognosis following myocardial infarction: classification trees, logistic regression, and stepwise linear discrimination. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.) **1** 245–267. Wadsworth, Monterey, Calif.

STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LA JOLLA, CALIFORNIA 92093

J. N. K. RAO AND N. G. N. PRASAD

Carleton University and University of Alberta

1. General comments. This important paper is a major contribution to jackknife methodology. A major strength of the proposed weighted jackknife method is its ready extensibility to nonlinear situations, including the important generalized linear models with uncorrelated errors briefly discussed in Section 8.

In the case of a linear regression model with uncorrelated errors, the delete-1 jackknife variance estimator, $v_{J(1)}$, is shown to be exactly unbiased for $\mathrm{Var}(\hat{\beta})$ under $\mathrm{Var}(e) = \sigma^2 I$, and approximately unbiased (as $n \to \infty$) under $\mathrm{Var}(e) = \mathrm{diag}(\sigma_i^2)$. However, $v_{J(1)}$ seems to have no special advantage over the MINQUE (minimum norm quadratic unbiased estimator) of $\mathrm{Var}(\hat{\beta})$ (Rao (1973)) under the criterion of bias robustness since the latter estimator is exactly unbiased under $\mathrm{Var}(e) = \mathrm{diag}(\sigma_i^2)$ unlike $v_{J(1)}$. It may also be noted that the MINQUE of $\mathrm{Var}(\hat{\beta})$ seldom becomes negative definite even though the MINQUE of individual σ_i^2 may assume negative values. If $\theta = g(\beta)$, then the linearization technique can be used to get a MINQUE-based estimator of the variance of $\hat{\theta} = g(\hat{\beta})$. This variance estimator should be satisfactory since Wu's simulation study shows that the "linearization method is a winner."

Apart from bias robustness, one should also consider efficiency robustness to arrive at a suitable variance estimator. One of us (Rao (1973)) derived the variance of MINQUE in the case of linear regression with one independent variable. It would be useful to compare the efficiencies of alternative variance estimators relative to the usual variance estimator $\hat{v} = \hat{\sigma}^2(X^TX)^{-1}$ that is "best" under $\mathrm{Var}(e) = \sigma^2 I$.

2. Variance components. We have recently used a weighted jackknife method to obtain robust confidence intervals on a smooth function $\theta = g(\sigma_e^2, \sigma_v^2)$ of variance components σ_v^2 and σ_e^2 in the nested error regression model,

$$y_{ij} = \alpha + \beta x_{ij} + v_i + e_{ij} = \alpha + \beta x_{ij} + u_{ij}, \tag{1}$$

with $E(v_i) = E(e_{ij}) = 0$ and $\mathrm{Var}(v_i) = \sigma_v^2$, $\mathrm{Var}(e_{ij}) = \sigma_e^2$, $j = 1, \dots, n_i$, $i = 1, \dots, t$. Arvesen (1969) considered a similar problem for the special case of a one-way ANOVA model $y_{ij} = \mu + v_i + e_{ij}$. Taking the usual Henderson unbiased estimator $\hat{\sigma}_e^2$ of σ_e^2 and a slightly modified unbiased estimator $\tilde{\sigma}_v^2$ of σ_v^2 (similar to Arvesen's estimator of σ_v^2 for the ANOVA model), we expressed $\hat{\sigma}_e^2$ and $\tilde{\sigma}_v^2$ as

$$\hat{\sigma}_e^2 = (t-1)^{-1} \sum_{i=1}^{t} a_i \hat{\sigma}_e^2(-i) \quad \text{and} \quad \tilde{\sigma}_v^2 = (t-1)^{-1} \sum_{i=1}^{t} b_i \tilde{\sigma}_v^2(-i). \tag{2}$$

Here a_i and b_i are positive constants depending on the sample sizes n_i and the x_{ij}-values, and $\hat{\sigma}_e^2(-i)$ and $\tilde{\sigma}_v^2(-i)$ are obtained by omitting the "Henderson residuals" $\hat{e}_{ij}$ and $\hat{u}_{ij}$ $(j = 1, \dots, n_i)$ for the ith group and then constructing unbiased estimators of σ_e^2 and σ_v^2, respectively. It may be noted that the representation (2) involves two different weights, a_i and b_i, unlike the single weight w_s in Wu's problem.

A weighted jackknife estimator of the variance of $\tilde{\theta} = g(\hat{\sigma}_e^2, \tilde{\sigma}_v^2)$, similar to Wu's in the regression case, is given by

$$v_J(\tilde{\theta}) = \sum_{i=1}^{t} \left[g\{\hat{\sigma}_e^2 + \hat{\sigma}_1^2(i), \tilde{\sigma}_v^2 + \tilde{\sigma}_2^2(i)\} - g(\sigma_e^2, \sigma_v^2)\right]^2, \tag{3}$$

where

$$\hat{\sigma}_1^2(i) = a_i^{1/2}\{\hat{\sigma}_e^2(-i) - \hat{\sigma}_e^2\}, \qquad \tilde{\sigma}_2^2(i) = b_i^{1/2}\{\tilde{\sigma}_v^2(-i) - \tilde{\sigma}_v^2\}. \tag{4}$$

We have shown that $v_J(\tilde{\theta})$ is consistent as $t \to \infty$, without normality assumption. Hence, $v_J(\tilde{\theta})$ provides robust $(1-\alpha)$-level confidence intervals on θ: $\tilde{\theta} \pm t_{\alpha/2, t-1}\{v_J(\tilde{\theta})\}^{1/2}$, where $t_{\alpha/2, t-1}$ is the upper $\alpha/2$-point of a t-distribution with $t-1$ degrees of freedom. In the linear case $\theta = \sigma_e^2$, it can be shown that $E[v_J(\hat{\sigma}_e^2)] = K\ \mathrm{Var}(\hat{\sigma}_e^2)$, where K is a function only of the x_{ij}'s and converges to 1 as t increases. This result is similar to Wu's in the regression case with $\mathrm{Var}(e) = \sigma^2 I$. We are now developing results similar to (3) using Henderson's estimator $\hat{\sigma}_v^2$ instead of $\tilde{\sigma}_v^2$.

The jackknife method can also be used to obtain an approximately unbiased estimator $\tilde{\theta}_J$ of $\theta = g(\sigma_e^2, \sigma_v^2)$, i.e., $E(\tilde{\theta}_J) - \theta = o(t^{-1})$ for large t, without normality assumption. The estimator $\tilde{\theta}_J$ can be used in small area estimation to get approximately unbiased estimators of the weights in the best predictors. It may be noted that in the empirical Bayes literature (e.g., Morris (1983)), the weights are unbiasedly estimated under normality assumption in the balanced case, $n_i = m$.

Details of these results will be reported in a separate paper.

REFERENCES

ARVESEN, J. N. (1969). Jackknifing U-statistics. *Ann. Statist.* **40** 2076–2100.

MORRIS, C. N. (1983). Parametric empirical Bayes inference: theory and applications. *J. Amer. Statist. Assoc.* **78** 47–65.

RAO, J. N. K. (1973). On the estimation of heteroscedastic variances. *Biometrics* **29** 11–24.

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
OTTAWA, ONTARIO K1S 5B6
CANADA

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA T6G 2G1
CANADA

JUN SHAO

University of Wisconsin-Madison

Applications of the jackknife and other resampling methods to regression analysis have been thoroughly discussed in Professor Wu's paper. One interesting and stimulating aspect of his approach is the use of a weighting scheme that takes into account the unbalanced nature of regression data. He has provided a fundamental tool for handling very general non-i.i.d. problems for which the classical jackknife method may not work well. In Section 8, he considered extensions of his method to several non-i.i.d. situations. More research is needed and is being done in this area.

In this discussion, I would like to (A) propose another weighted resampling scheme that gives an interpretation of Wu's weighted jackknife and provides an alternative resampling estimation procedure, (B) discuss the use of Tukey's pseudovalue, and (C) obtain the stochastic order of the weighted jackknife bias estimator.

In the following, all notation will be the same as that of Wu.

(A) Another weighted resampling scheme. In the regression situation, the information contained in different subsets of data may be quite different. The idea of my proposed weighted resampling scheme is to take account of the unbalanced nature of the data in the resampling process. That is, the probability of selecting a subset of data is not a constant as is usually done, but is proportional to the determinant of the Fisher information matrix of the corresponding subset model with i.i.d. errors. We will see that the bias and variance