Statistica Sinica **12**(2002), 575-598

COMPARISON OF BOOTSTRAP AND JACKKNIFE VARIANCE ESTIMATORS IN LINEAR REGRESSION: SECOND ORDER RESULTS

Arup Bose and Snigdhansu Chatterjee

Indian Statistical Institute and University of Nebraska-Lincoln

Abstract: In an extension of the work of Liu and Singh (1992), we consider resampling estimates for the variance of the least squares estimator in linear regression models. Second order terms in asymptotic expansions of these estimates are derived. By comparing the second order terms, certain generalised bootstrap schemes are seen to be theoretically better than other resampling techniques under very general conditions. The performance of the different resampling schemes are studied through a few simulations.

Key words and phrases: Bootstrap, jackknife, regression, variance comparison.

1. Introduction

Consider the linear regression model given by

$$\mathbf{y}_{(n\times 1)} = \mathbf{X}_{(n\times p)}\beta_{(p\times 1)} + \mathbf{e}_{(n\times 1)}.$$

Suppose $\hat{\beta}$ is the least squares estimate of β calculated from the data. The aim is to estimate $\mathrm{V_n} = E(\hat{\beta}-\beta)(\hat{\beta}-\beta)^T$, the dispersion of the least squares estimator. An estimator $\hat{V}_n$ will be said to be consistent if $n(\mathrm{V_n} - \hat{V}_n) \to 0$.

Let V_{nj} denote the estimator of $\mathrm{V_n}$ obtained using the jth resampling scheme. In an interesting paper, Liu and Singh (1992) showed that

$$T_{nj} = n^{3/2}(V_{nj} - \mathrm{V_n}) = A_{nj} + O_P(n^{-1/2}) \tag{1.1}$$

where A_{nj} is one of two random variables, say Y_n and Z_n (whose expressions will be given later), with $EY_n^2 < EZ_n^2$. Resampling techniques that have Z_n as a lead term are consistent against heteroscedasticity of errors, but those with Y_n as the lead term are not. On the other hand, under homoscedasticity, techniques with lead term Y_n are relatively more efficient. Thus resampling techniques are in two groups: those *robust* against heteroscedasticity (R-class), and those *efficient* under homoscedasiticity (E-class). Among standard resampling techniques, the classical residual bootstrap of Efron (1979) and a weighted jackknife due to Liu and Singh (1992) belong to the E-class, while the paired bootstrap, delete-1 jackknife and external bootstrap belong to the R-class.

The comparison of resampling techniques based on (1.1) is asymptotic in nature. A natural question is whether techniques belonging to the same class have similar small sample performances, *i.e.*, whether results vary greatly from one technique to another in the same class.

As an example, we consider the oxygen uptake data from Rawlings, Pantula and Dickey (1998). This is data on five variables that measure physical fitness and there are thirty-one cases. The authors regress oxygen uptake on these five and have a hypothesis that two of the variables do not influence oxygen uptake. Consider contrasting the the two means. The variance of the estimated contrast is estimated by resampling techniques. We applied three resampling techniques, all of which are consistent under unequal error variances: the delete-1 jackknife, paired bootstrap and Wu's external bootstrap. The variance estimates are, respectively, 30.29, 30.85, 19.80. The differences here are large and call for a closer comparison of different resampling techniques.

In order to broaden the choice of resampling techniques we also include a wide class of weighted bootstraps not considered by Liu and Singh (1992). The weighted bootstrap, or generalized bootstrap, has been an object of study for some time, see Barbe and Bertail (1995) for a review. To illustrate, consider the sample mean $n^{-1}\sum e_i$ from data $e_1, \ldots, e_n$. Notice that Efron's classical bootstrap technique of using simple random sampling yields the bootstrap statistic $n^{-1}\sum w_{ni}e_i$, where $\mathbf{w}_n = (w_{n1}, \ldots, w_{nn}) \sim Multinomial(n; 1/n, \ldots, 1/n)$. *Generalised bootstrap* operates by considering general random vectors $\mathbf{w}_n$, so that the properties of Efron's bootstrap are recovered as a special case. Many other resampling techniques like the Bayesian bootstrap, the m out of n bootstrap, the delete-d jackknives are also special cases, and so the framework of generalised bootstrap serves as a convenient platform for comparison of different resampling schemes. Also, many new techniques may be considered under generalised bootstrap, and it is conceivable that certain nice or desirable properties may be found in some generalised bootstrap schemes that do not have a classical analog.

In the context of regression, generalised bootstrap may be performed either on residuals or on data pairs. The weighted bootstrap (Liu (1988)), the external or wild bootstrap (Wu (1986), Mammen (1993)) are examples of *generalised residual bootstrap* (GBS). We consider a broader class of residual-based resampling techniques that will yield the above two, as well as the classical residual bootstrap of Efron (1979), as special cases. This is done in Section 2. The "uncorrelated weights bootstrap" (UBS) of Chatterjee and Bose (2000) is a class of *generalised paired bootstrap* techniques. We recapitulate essential properties of this in the second half of Section 2.

Using general weights, one has a number of "free parameters" at one's disposal which may be tuned to obtain 'optimal' resampling results. Classical techniques usually correspond to particular choices of such parameters. Bootstrap weights are generally chosen to be exchangeable (but this is not necessary for variance estimation), so that they have the same marginal distribution. One important parameter is the common variance of the weights, σ_n^2. Comparison and 'optimality' results in this paper are based on the fact that in GBS and UBS, one may choose σ_n^2 so that $|ET_{nj}|$ is minimal. Naturally other moments of the weights, including various mixed moments appear, but these can be taken care of by a judicious choice of weights, and often i.i.d. weights suffice. However, note that most classical techniques cannot be realized as resampling with i.i.d. weights, with the exception of the Bayesian bootstrap in some cases.

We extend (1.1) to a second order term:

$$T_{nj} = n^{3/2}(V_{nj} - \mathrm{V_n}) = A_{nj} + n^{-1/2}B_{nj} + O_P(n^{-1}), \tag{1.2}$$

where A_{nj}, and B_{nj} are $O_P(1)$ variables, and A_{nj} has zero mean. We find that $EB_{nj} = b_{nj} + o(1)$, where b_{nj} is a non-zero term, and in most cases $EB_{nj} - b_{nj} = O(n^{-1/2})$.

This implies that ET_{nj} generally has a *bias* term of the order of $O(n^{-1/2})$, and this bias varies according to the resampling technique. This is not a new discovery; for example it is fairly well known that the jackknife variance estimator is biased, although the bias term does not appear in the first order asymptotics. It may be conjectured that the small sample difference in performances is a reflection of the bias terms b_{nj}. Our aim in this paper is threefold.

(a) We report the second order asymptotic expansions (1.2) for all commonly known resampling techniques including various generalized bootstrap techniques.

(b) We present results for GBS which encompasses the classical residual bootstrap of Efron (1979), the external bootstrap of Wu (1986) and the weighted bootstrap of Liu (1988) as special cases. This new class of bootstraps considerably broadens the E-class.

(c) A resampling technique may be preferred to others if its *bias* b_{nj} is 0, or if has a smaller bias term than other techniques. In the E-class, such a 'preferable' or 'optimal' technique belongs to the new GBS class of (b) above; whereas in the R-class such a technique is often a weighted jackknife due to Wu (1986) or a UBS technique. We use the term *second order optimal* for a resampling technique for which $b_{nj} = 0$.

As suggested by the referees, we attempted comparisons for $p = 1$ via

$$ET_{nj}^2 = EA_{nj}^2 + 2n^{-1/2}EA_{nj}(B_{nj} - b_{nj}) + n^{-1}b_{nj}^2 + n^{-1}E(B_{nj} - b_{nj})^2 + o(n^{-1}).$$

From tedious computations it turns out that, very generally, $EA_{nj}(B_{nj} - b_{nj}) = o(n^{-1/2})$ and $E(B_{nj} - b_{nj})^2 = o(1)$. This happens, for example, if the e_i's are symmetric. Thus comparison based on ET_{nj} and ET_{nj}^2 across j tend to agree.

For $p > 1$, the variances (and hence T_{nj}'s) are matrices. For E-class estimators, the second order terms for different schemes differ by scalar multiples of the matrix $(\mathbf{X}^T\mathbf{X})^{-1}$, so that comparison is over scalars. However, for R-class estimators different matrices have to be compared and there is no unique way of doing this. We use the criterion that a random matrix $\mathbf{A}$ is closer to 0 than the random matrix $\mathbf{B}$, if

$$|E\ trace(\mathbf{A})| < |E\ trace(\mathbf{B})|. \tag{1.3}$$

One advantage with this criterion is that both trace and expectation are linear, and hence interchangeable.

For homoscedastic errors where both E-class and R-class techniques are consistent, the best second order performance is obtained for a GBS technique. This does not seem to have appeared in the literature.

Under the additional model condition

$$2\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 {\tau_i}^2 > \sum_{i=1}^{n} x_i^4 \sum_{i=1}^{n} x_i^2 {\tau_i}^2, \tag{1.4}$$

some UBS resampling techniques outperform other R-class techniques. This condition is satisfied if (a) ${\tau_i}^2 = kx_i^{\alpha}$ with $\alpha > 0$; (b) $K_1 \leq {\tau_i}^2 \leq K_2$ and $K_2 < 2K_1$; (c) $K_1 \leq x_i^2 \leq K_2$ and $K_2^3 < 2K_1^3$. It also holds under several other conditions. A weighted jackknife due to Wu (1986) is also observed to perform well in practice.

In the multiple linear regression set-up, optimality for any UBS technique cannot be established without using model assumptions that are difficult to check. On the other hand, Wu's jackknife is difficult to compute for complex problems. However, in Section 5 we see that UBS techniques based on i.i.d. weights, whose variance is not necessarily computed using optimality considerations, perform quite well in practice.

We now give a brief summary of the contents of the other sections. The different weighted jackknives and generalized bootstraps in regression are described in Section 2. We state our results for simple linear regression in Section 3. Section 4 contains E-class and R-class results for multiple regression. Although the delete-1 jackknife is a special case of the UBS, we state the results for it separately.

In Section 5 we take up some examples where the variance of a linear contrast of parameters from linear models is to be estimated, and noise is heteroscedastic. The aim is to study the performance of the different R-class resampling techniques in small and medium sized samples. We took sample sizes 15 and 31 for our study. Although we have results on optimality of UBS in the R-class based on choice of σ_n^2, such choices would require complex calculations using the design matrix and noise variances. Instead, we use two 'rule of thumb' choice of weights. It is seen that these UBS schemes and Wu's weighted jackknife (Wu(1986)) perform well. The paired bootstrap is extremely unstable when sample size is small, presumably due to the variations in the design matrix across resamples. The delete-1 jackknife and the external bootstrap have bias that is noticeable in small and medium sample sizes, and thus are not recommended for variance estimation.

We omit the proofs of theorems stated in this paper, they are lengthy and complicated. Details are in Bose and Chatterjee (1997, 1998).

2. The Different Resampling Techniques in Regression

The model is given by $\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$, where $\mathbf{y}$ is the $(n \times 1)$ observed vector of responses and $\mathbf{X}$ is the $(n \times p)$ size design matrix whose ith row $\mathbf{x}_i^T$ is given by the ith observation of the $(p \times 1)$ vector of explanatory variables. The parameter β is a $(p \times 1)$ vector of unknown constants and $\mathbf{e}^T = (e_1 \ \ e_2 \ \ \ldots \ \ e_n)$ is the noise vector. Suppose $\hat{\beta}$ is the least squares estimate of β as calculated from the data. If $Ee_i^2 = {\tau_i}^2$ then

$$\mathrm{V_R} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{T}\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}, \tag{2.1}$$

where $\mathbf{T} = diag({\tau_i}^2, \quad 1 \leq i \leq n)$. Under homoscedasticity, *i.e.*, when ${\tau_i}^2 = \tau^2, \quad i = 1, \ldots, n$, the expression for the variance of $\hat{\beta}$ simplifies to

$$\mathrm{V_E} = \tau^2(\mathbf{X}^T\mathbf{X})^{-1}. \tag{2.2}$$

Generally for multiple regression problems two common and natural assumptions on the design matrix are

$$d_i/n \in (m, M) \text{ for some } 0 < m \leq M < \infty, \tag{2.3}$$
$$\mathbf{x}_i^T\mathbf{x}_i < K \ < \infty, \qquad i = 1, 2, \ldots, n, \tag{2.4}$$

where $d_i, 1 \leq i \leq p$ are the eigenvalues of $\mathbf{X}^T\mathbf{X}$. Assumption (2.4) appears in Liu and Singh (1992) and essentially eliminates the possibility that an influential design point that can alter the asymptotics. However, in order to make the resampling schemes like the paired bootstrap and the different jackknives feasible on the model, we need slightly stronger assumptions than (2.3). We assume

$$\frac{1}{m}\mathbf{X}^{T*}\mathbf{X}^* > k\mathbf{I} \tag{2.5}$$

for some $k > 0$ and $m \geq n/3$, and all choices of $(m \times p)$ submatrix $\mathbf{X}^*$ that can be formed by choosing any m rows of $\mathbf{X}$. By $\mathbf{A} > \mathbf{B}$ we mean that $\mathbf{A} - \mathbf{B}$ is positive definite. As remarked by Wu ((1986), page 1344), (2.5) seems necessary for most schemes that use data-pairs for resampling. But let us emphasize that (2.5) and (2.4) are only sufficient conditions under which all known resampling schemes work. Individual resampling schemes may work under weaker conditions. It turns out that the best resampling plans do work under (2.3). We assume that the noise terms e_i's are independent with uniformly bounded *eighth* moment. Actual requirements for most resampling schemes may be weaker. Throughout this paper we take p as fixed. For the performance of different resampling techniques as $p \to \infty$ with $n \to \infty$ in the context of estimating the entire distribution of $\hat{\beta}_n$ as well as for variance estimation, see Bickel and Freedman (1983), Mammen (1993), Chatterjee and Bose (2000, 2000a).

2.1. The jackknife variance estimators

The delete-1 jackknife uses the n different $\hat{\beta}_{(-i)}$'s, where $\hat{\beta}_{(-i)}$ is the least squares estimates based on the observations $\{(y_j, \mathbf{x}_j), j = 1, 2, \ldots, i-1, i+1, \ldots, n\}$. The different jackknife estimators use different weighted sums of $(\hat{\beta}_{(-i)} - \hat{\beta})(\hat{\beta}_{(-i)} - \hat{\beta})^T$, with weights possibly depending on the design. Let $\delta_i = \mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_i$. Expressions for the four different jackknife variance estimators are

(i) (Hinkley (1977)) $\mathrm{V}_{\mathrm{JH}} = \frac{n}{n-p}\sum_{i=1}^n (1-\delta_i)^2(\hat{\beta}_{(-i)} - \hat{\beta})(\hat{\beta}_{(-i)} - \hat{\beta})^T$;

(ii) (Wu (1986)) $\mathrm{V}_{\mathrm{JW}} = \sum_{i=1}^n (1-\delta_i)(\hat{\beta}_{(-i)} - \hat{\beta})(\hat{\beta}_{(-i)} - \hat{\beta})^T$;

(iii) (Liu and Singh (1992)) $\mathrm{V}_{\mathrm{JLS}} = \frac{(n-1)\sum_{i=1}^n x_i^2}{n^2}\sum_{i=1}^n \frac{(\hat{\beta}_{(-i)} - \hat{\beta})^2}{x_i^2}$;

(iv) (Quenouille (1949)) $\mathrm{V}_{\mathrm{J}} = \frac{n-1}{n}\sum_{i=1}^n (\hat{\beta}_{(-i)} - \hat{\beta})(\hat{\beta}_{(-i)} - \hat{\beta})^T$.

The last one is the classical delete-1 jackknife. The weighted jackknife due to Liu and Singh (1992) is a scheme for simple regression set-up only, and does not have an extension to higher dimensional models.

The weighted jackknives of Hinkley (1977) and Wu (1986) had been framed with a view to correcting for any unbalanced nature of the design matrix. Our simulations show the merits of these corrections, with Wu's jackknife generally performing very well. However, see the discussion to Wu (1986) for some criticism of his jackknife. In Section 5 we comment in detail on the relative merits of Wu's jackknife and the class of resampling techniques that we suggest.

General delete-d jackknives are also used for resampling, where the pseudovalues are obtained by deleting sets of d observations at a time. For any d, the delete-d jackknife schemes are special cases of UBS, see Chatterjee (1998). The weighted jackknife due to Wu has a delete-d variation which we do not consider.

2.2. The generalised residual bootstrap

This section describes the generalized residual bootstrap (*GBS*) technique. We use the following notations for the rest of the paper: $\mathbf{I}$ is an identity matrix of appropriate order; $\mathbf{J} = n^{-1}\mathbf{1}\mathbf{1}^T$; $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = \beta + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{e}$; $\mathbf{P_x} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, the projection on the column space of $\mathbf{X}$; $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta} = \mathbf{X}\beta + \mathbf{P_x}\mathbf{e}$; $\mathbf{R} = \mathbf{y} - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{P_x})\mathbf{e}$; $\mathbf{R}_0 = (\mathbf{I} - \mathbf{J})(\mathbf{I} - \mathbf{P_x})\mathbf{e}$; $\mathbf{G} = \mathbf{X}^T\mathbf{X}$.

Let $\mathbf{W}$ be an $n\times n$ random matrix, whose ith row is denoted by $\mathbf{w}_i^T$. Suppose $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ are i.i.d. samples from $Multinomial(1, 1/n, 1/n, \ldots, 1/n)$. If we define $\mathbf{y}^* = \mathbf{X}\hat{\beta} + \mathbf{W}\mathbf{R}_0$, we get the usual residual bootstrap. A generalization of this is possible by taking any random $\mathbf{W}$ to be the weight matrix, subject to certain conditions that we discuss later. Also $\mathbf{R}$ or other residuals may be used instead of $\mathbf{R}_0$, but we restrict attention to $\mathbf{R}$ and $\mathbf{R}_0$ in this paper. Using other forms of residuals may lead to a bootstrap that is more robust in the sense that model assumptions can be more relaxed.

If the resampled data set is formed by $\mathbf{y}^* = \mathbf{X}\hat{\beta} + \mathbf{W}\mathbf{R}^*$ where $\mathbf{W}$ is a random matrix and $\mathbf{R}^*$ is either $\mathbf{R}$ or $\mathbf{R}_0$, we call it a *Generalised Residual Bootstrap* (*GBS*). We have $\hat{\beta_{\mathrm{B}}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}^*$ and hence $\hat{\beta_{\mathrm{B}}} - \hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{R}^*$. If $\mathbf{R}^* = \mathbf{R}$ and $\mathbf{W}$ is diagonal with entries that are i.i.d. observations from a random variable with zero mean and unit variance, we have the *external or wild bootstrap.* If $\mathbf{R}^* = \mathbf{R}$ and $\mathbf{W} = (\mathbf{I} - \mathbf{J})\mathbf{W}_1$, where $\mathbf{W}_1$ is diagonal with entries that are i.i.d. observations from a random variable with zero mean and unit variance, we obtain the *weighted bootstrap* of Liu (1988). In general, the behaviour of the *GBS* technique depends on the nature of the matrix $\mathbf{W}$. However, the choice of the residual vector contributes in an interesting way. If $\mathrm{E_B}W = c\mathbf{J}$, define $\mathbf{W}_0 = (W - cJ)(I - J)$. Using the idempotence of $\mathbf{J}$, we have $\mathbf{W}\mathbf{R}_0 = (\mathbf{W} - c\mathbf{J})\mathbf{R}_0 + c\mathbf{J}\mathbf{R}_0 = (\mathbf{W} - c\mathbf{J})(\mathbf{I} - \mathbf{J})\mathbf{R} = \mathbf{W}_0\mathbf{R}$. Thus the choice of $\mathbf{R}^*$ to be $\mathbf{R}$ or $\mathbf{R}_0$ may be offset by the choice of the random weighting matrix.

The definition of the *GBS* estimator depends heavily on the nature of the weights used. We present three different versions to incorporate some of the existing and related estimators in our study. Let σ_n^2 be the variance of the bootstrap weights,

$$\mathrm{V_{GB1}} = \frac{1}{n\sigma_n^2}\mathrm{E_B}(\hat{\beta_{\mathrm{B}}} - \hat{\beta})(\hat{\beta_{\mathrm{B}}} - \hat{\beta})^T, \tag{2.6}$$

$$\mathrm{V_{GB2}} = \mathrm{E_B}(\hat{\beta_{\mathrm{B}}} - \hat{\beta})(\hat{\beta_{\mathrm{B}}} - \hat{\beta})^T, \tag{2.7}$$

$$\mathrm{V_{GB3}} = \frac{1}{\sigma_n^2}\mathrm{E_B}(\hat{\beta_{\mathrm{B}}} - \hat{\beta})(\hat{\beta_{\mathrm{B}}} - \hat{\beta})^T. \tag{2.8}$$

These are all scaled versions of $\mathrm{E_B}(\hat{\beta_{\mathrm{B}}} - \hat{\beta})(\hat{\beta_{\mathrm{B}}} - \hat{\beta})^T$, the scaling depending on σ_n^2 and can be viewed as natural extension of the variance estimates used in Barbe and Bertail (1995) for generalised bootstrapping of the sample mean.

$\mathrm{V}_{\mathrm{GB2}}$ and $\mathrm{V}_{\mathrm{GB3}}$ are, respectively, of the same form as the variance estimate under residual bootstrap and external bootstrap. The consistency of $\mathrm{V}_{\mathrm{GB2}}$ requires $n\sigma_n^2 \to 1$ and, under this condition, it is essentially the same as $\mathrm{V}_{\mathrm{GB3}}$. However, it deserves separate attention for two reasons: this is the variance expression for the classical residual bootstrap, for which $\sigma_n^2 = (n-1)/n^2$; we obtain a result below showing that the 'best under homoscedasticity' resampling scheme has a variance expression of this form.

Under certain conditions $\mathrm{V}_{\mathrm{GB1}}$ and $\mathrm{V}_{\mathrm{GB2}}$ belong to the E-class. Under some other conditions, $\mathrm{V}_{\mathrm{GB3}}$ belongs to the R-class. These results are discussed in details later. The behaviour of GBS is largely governed by whether $\mathbf{W}$ is a diagonal matrix or not. Generally speaking, if off-diagonal entries are non-zero, the scheme is consistent only under homoscedasticity.

2.3. The uncorrelated weights bootstrap UBS

This generalized bootstrap scheme was introduced in Chatterjee and Bose (2000). Let $\{w_{i:n}; 1 \le i \le n, n \ge 1\}$ be a triangular array of nonnegative random weights. *We will drop the suffix n from the notation.* The resampling scheme is carried out by weighting each data point $(y_i, \mathbf{x}_i)$ with the random weight $\sqrt{w_i}$, then computing the statistic of interest and taking expectation of the random weight vector.

This is a direct generalization of the *paired bootstrap*, where the random weights $\{w_i; 1 \le i \le n\}$ are $Multinomial(n; 1/n, \ldots 1/n)$. The different delete-$d$ jackknives variance estimators can also be viewed as special cases of this, see Chatterjee (1998) for details. Other examples of UBS techniques include the Bayesian bootstrap, the weighted likelihood bootstrap, the m out of n bootstrap, and several variations of these, some of which are available in Praestgaard and Wellner (1993). Weights which form an i.i.d. sample from a suitable distribution may also be used as UBS weights.

The weights $w_1, \ldots, w_n$ used for UBS resampling satisfy certain conditions that we now state. Let $V(w_i) = \sigma_n^2$. Assume that the quantities $E((w_a-1)/\sigma_n)^i$ $((w_b-1)/\sigma_n)^j((w_c-1)/\sigma_n)^k \ldots$ for distinct $a, b, c \ldots$ are functions of the powers $i, j, k \ldots$ only, and not of the indices $a, b, c \ldots$. Thus we can write $c_{ijk\ldots} = E((w_a-1)/\sigma_n)^i((w_b-1)/\sigma_n)^j((w_c-1)/\sigma_n)^k \ldots$. Note that if the weights are assumed to be exchangeable, then the above condition follows. Also let $\mathcal{W}$ be the set on which at least m_0 of the weights are greater than some fixed constant $k_2 > 0$. The value of m_0 is $\ge n/3$. Throughout the notations k, K are generic for constants. The weights are required to satisfy

$$E(w_i) = 1, \tag{2.9}$$

$$\mathrm{P_B}[Kn \geq \sum_{i=1}^{n} w_i \geq kn, \quad K > k > 0] = 1, \tag{2.10}$$

$$\mathrm{P_B}[\mathcal{W}] = 1 - O(n^{-1}), \tag{2.11}$$

$$K > \sigma_n^2 > 0, \tag{2.12}$$

$$c_{11} = O(n^{-1}), \tag{2.13}$$

$$\forall \quad i_1, i_2, \ldots, i_k \text{ satisfying } \sum_{j=1}^{k} i_j = 3, \quad c_{i_1 i_2 \ldots i_k} = O(n^{-k+1}\sigma_n{}^{-1}), \tag{2.14}$$

$$\forall \quad i_1, i_2, \ldots, i_k \text{ satisfying } \sum_{j=1}^{k} i_j = 4, \quad c_{i_1 i_2 \ldots i_k} = O(n^{-k+2}), \tag{2.15}$$

$$\forall \quad i_1, i_2, \ldots, i_k \text{ satisfying } \sum_{j=1}^{k} i_j = 6, \quad c_{i_1 i_2 \ldots i_k} = O(n^{-k+3}), \tag{2.16}$$

$$\forall \quad i_1, i_2, \ldots, i_k \text{ satisfying } \sum_{j=1}^{k} i_j = 8, \quad c_{i_1 i_2 \ldots i_k} = O(n^{-k+4}). \tag{2.17}$$

3. Second Order Results for $p = 1$

Throughout this section we assume $p = 1$ and one dimensional versions of (2.3)-(2.5). This is equivalent to assuming that for all i, $0 < k < |x_i| < K < \infty$. We also assume that the errors are independent with uniformly bounded eighth moments. This section is for an easier understanding of the results to come for general p, and also to relate to the first order results of Liu and Singh (1992). Hence we do not consider the GBS scheme here but in the next section, and we separately compute the results for the delete-1 jackknife, which is a special case of UBS.

We use the notation $\mathrm{L_n} = \sum_{i=1}^{n} x_i^2$. If $Ee_i^2 = \tau_i{}^2$, then $\mathrm{V_R} = \mathrm{L_n^{-2}} \sum_{i=1}^{n} x_i^2 \tau_i{}^2$. Under homoscedasticity of the noise terms, *i.e.*, when $\tau_i{}^2 = \tau^2$, $i = 1, \ldots, n$, the expression for variance of $\hat{\beta}$ simplifies to $\mathrm{V_E} = \mathrm{L_n^{-1}}\tau^2$. We assume the existence of the following limits.

$$\alpha = \lim_{n\to\infty} (\frac{n}{\mathrm{L_n}})^{1/2}; \quad \alpha_0 = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} x_i^2 \tau_i{}^2; \quad \alpha_1 = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} x_i^4 \tau_i{}^2;$$

$$\alpha_2 = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} x_i^6 \tau_i{}^2; \ \gamma_1 = \lim_{n\to\infty} \frac{\sum x_i}{n}; \ \gamma_2 = \lim_{n\to\infty} \frac{\sum x_i^4}{n}; \ \gamma_3 = \lim_{n\to\infty} \frac{\sum x_i^6}{n}$$

Let $Y_n = n^{1/2} \frac{1}{\mathrm{L_n}} \sum_{i=1}^{n} (e_i^2 - \tau^2)$, $Z_n = n^{3/2} \frac{1}{\mathrm{L_n^2}} \sum_{i=1}^{n} x_i^2 (e_i^2 - \tau_i{}^2)$, and let C_{nj} denote a quantity which satisfies $EC_{nj} = 0$ and $\sup_n EC_{nj}^2 < \infty$.

Theorem 3.1. *Suppose the error terms have the same variance.*
(i) *The residual bootstrap variance estimate satisfies*

$$n^{3/2}(\mathrm{V_{GB2}} - \mathrm{V_E}) = Y_n - \frac{1}{n^{1/2}}\left[\alpha^2(\frac{\sum e_i}{n^{1/2}})^2 - 2\alpha^3\gamma_1(\frac{\sum e_i}{n^{1/2}})(\frac{\sum x_i e_i}{(\sum x_i^2)^{1/2}}) + (\alpha^2 + \alpha^4\gamma_1^2)(\frac{\sum x_i e_i}{(\sum x_i^2)^{1/2}})^2\right] + n^{-1}C_{n1}.$$

(ii) Liu and Singh's *weighted jackknife variance estimate satisfies*

$$n^{3/2}(\mathrm{V_{JLS}} - \mathrm{V_E}) = Y_n - n^{-1/2}\left[\alpha^2(\frac{\sum_{i=1}^n x_i e_i}{(\sum_{i=1}^n x_i^2)^{1/2}})^2 - \alpha^2\tau^2\right] + n^{-1}C_{n2}.$$

A simple computation shows that the second term in the expansion for the weighted jackknife of Liu and Singh (1992) has zero expectation. However, it is only available for $p = 1$.

We introduce some notation to present the results for the R-class. Let $\eta_1 = \sum x_i e_i/(\sum x_i^2)^{1/2}$, $\eta_2 = \sum x_i^3 e_i/(\sum x_i^6)^{1/2}$ and $\eta = (\eta_1\ \eta_2)^T$, with

$$\mathbf{H} = \begin{pmatrix} \alpha^6\gamma_2 & -\alpha^5\gamma_3^{1/2} \\ -\alpha^5\gamma_3^{1/2} & 0 \end{pmatrix}, \qquad \mathbf{Q} = \eta^T\mathbf{H}\eta.$$

Theorem 3.2.
(i) *The delete-*1 *jackknife variance estimate satisfies*

$$n^{3/2}(\mathrm{V_J} - \mathrm{V_R}) = Z_n + n^{-1/2}\left[\mathbf{Q} - \alpha^4\alpha_0 + 2\alpha^6\alpha_1\right] + n^{-1}C_{n3}.$$

(ii) *The external bootstrap and weighted bootstrap variance estimate satisfies*

$$n^{3/2}(\mathrm{V_{GB3}} - \mathrm{V_R}) = n^{3/2}(\mathrm{V_{WB}} - \mathrm{V_n}) = Z_n + n^{-1/2}\mathbf{Q} + n^{-1}C_{n4}.$$

(iii) *Hinkley's weighted jackknife variance estimate satisfies*

$$n^{3/2}(\mathrm{V_{JH}} - \mathrm{V_R}) = Z_n + n^{-1/2}\left[\mathbf{Q} + \alpha^4\alpha_0\right] + n^{-1}C_{n5}.$$

(iv) *Wu's weighted jackknife variance estimate satisfies*

$$n^{3/2}(\mathrm{V_{JW}} - \mathrm{V_R}) = Z_n + n^{-1/2}\left[\mathbf{Q} + \alpha^6\alpha_1\right] + n^{-1}C_{n6}.$$

(v) *The UBS variance estimate under* (2.9)-(2.17) *satisfies*

$$n^{3/2}(\mathrm{V_{UBS}} - \mathrm{V_R}) = Z_n + n^{-1/2}T_n + n^{-1}C_{n7},$$

where the second order term T_n is given by

$$T_n = \mathbf{Q} + 3\sigma_n^2 c_{22}\alpha^8\alpha_0\gamma_2 - 2\sigma_n c_3\alpha^6\alpha_1 + n(3\sigma_n^2 c_{112} - 2\sigma_n c_{12} - c_{11})\alpha^4\alpha_0$$

The various mixed-moments determine the term T_n in Theorem 3.2 (v), and hence the exact second order value depends on the sequence of random weights used in UBS. For particular choice of weights, the above expression simplifies considerably, and we present some important special cases as a corollary.

Corollary 3.1. (a) *The paired bootstrap variance estimate satisfies*

$$n^{3/2}(\mathrm{V}_{\mathrm{UBS}} - \mathrm{V}_{\mathrm{R}}) = Z_n + n^{-1/2}\left[\mathbf{Q} + 3\alpha^8\alpha_0\gamma_2 - 2\alpha^6\alpha_1\right] + n^{-1}C_{n8}.$$

(b) *The UBS variance estimate with i.i.d. weights satisfies*

$$n^{3/2}(\mathrm{V}_{\mathrm{UBS}} - \mathrm{V}_{\mathrm{R}}) = Z_n + n^{-1/2}\left[\mathbf{Q} + 3\sigma_n^2\alpha^8\alpha_0\gamma_2 - 2\sigma_n c_3\alpha^6\alpha_1\right] + n^{-1}C_{n9}.$$

In particular, if the weights are symmetric about their mean, then $c_3 = 0$ and the second order term is $\mathbf{Q} + 3\sigma_n^2\alpha^8\alpha_0\gamma_2$.

Note that all the second order terms in Theorem 3.2 and Corollary 3.1 are of the form $\mathbf{Q}+\nu_j$. The minimum bias is achieved when $E\mathbf{Q}+\nu_j = 0$. A resampling scheme is *second order optimal* if this happens. If $\gamma_2 = \alpha = 1$, then the delete-1 jackknife is second order optimal. If $\alpha_1 = \alpha^2\alpha_0\gamma_2$, then the paired bootstrap is second order optimal. Note that the second order term for UBS depends on σ_n^2 and c_3. So the performance of any UBS scheme can be characterised by the second and third moments of the weights. For a UBS resampling scheme with i.i.d. weights to be optimal, the following relation must be satisfied.

$$\alpha^2\alpha_0\gamma_2(1 + 3\sigma_n^2) = 2\alpha_1(1 + \sigma_n c_3). \tag{3.1}$$

Suppose the additional model condition (1.4) holds. If we assume that the weights come from a distribution that is not skewed, and with variance given by

$$\sigma_n^2 = \frac{2\alpha_1 - \alpha^2\alpha_0\gamma_2}{3\alpha^2\alpha_0\gamma_2}, \tag{3.2}$$

then the resulting UBS is second order optimal. The additional model condition (1.4) comes from requiring that $2\alpha_1 > \alpha^2\alpha_0\gamma_2$. Thus it is seen that under very general conditions, a UBS resampling technique exists that is superior to other techniques. Under other less general model conditions we have shown that the paired bootstrap is also optimal. The delete-1 jackknife is also optimal under certain other restrictions on the model.

Consider the problem of estimation of the common unknown mean if all error variances are equal. Then all the x_i's are equal to 1 and $Y_n = Z_n =$

$n^{-1/2}\sum(e_i^2-\sigma_i^2)$. In this particular model the distinction between E-class and R-class does not remain and all the resampling schemes have the same first order behaviour. The second order term for residual bootstrap, weighted bootstrap, and external bootstrap is $-n^{-1}(\sum e_i)^2$. For the delete -1 jackknife, weighted jackknife (Liu and Singh), weighted jackknife (Hinkley), weighted jackknife (Wu) and paired bootstrap, the second order term is $-n^{-1}(\sum e_i)^2+n^{-1}\sum e_i^2$.

All resampling schemes with second order term equal to $-n^{-1}(\sum e_i)^2+n^{-1}\sum e_i^2$ are optimal. This, in particular proves the superiority of the delete -1 jackknife over the residual bootstrap for the restricted model under consideration.

In this special case (1.4) is satisfied, and the variance requirement in (3.2) is $\sigma_n^2=1/3$. Consider the discrete distribution supported with equal probability on $1\pm\sqrt{2/3}$. A UBS with i.i.d. weights from this distribution is also optimal.

4. Second Order Results for $p>1$

We give results for the E class and the R class separately.

4.1. Results for E-class bootstrap, $p>1$

In this section we discuss the conditions under which a GBS belongs to the E-class. Assume that the elements W_{ij} of $\mathbf{W}_0$ satisfy

$$E\ W_{ij}=0,\quad 1\le i,j\le n, \tag{4.1}$$
$$Var\ W_{ij}=\sigma_n^2>0,\quad 1\le i,j\le n, \tag{4.2}$$
$$Corr\ (W_{ij},W_{ik})=c_{11}=O(n^{-1}),\quad j\ne k,\quad 1\le i\le n, \tag{4.3}$$
$$Corr\ (W_{ij},W_{lk})=0,\quad 1\le i\ne l\le n,\ \forall\ j,k. \tag{4.4}$$

For the residual bootstrap, observe that $\mathbf{W}_0=\mathbf{W}-\mathbf{J}$. Also, rows of $\mathbf{W}$ are independent. So the above conditions are satisfied with

$$\mathrm{E_B}\mathbf{W}=\mathbf{J} \tag{4.5}$$
$$\sigma_n^2=\frac{n-1}{n^2}\qquad 1\le i,j\le n, \tag{4.6}$$
$$c_{11}=-\frac{1}{n-1},\quad j\ne k,\ \forall\ 1\le i\le n. \tag{4.7}$$

Theorem 4.1 shows the consistency of the GBS and establishes the second order results under the above assumptions. Let us define $d_1=\tau^2[p+c_{11}\sum_{i\ne j}((\mathbf{P_x}))_{ij}]$, $d_2=\tau^2[p+c_{11}\sum_{i\ne j}((\mathbf{P_x}))_{ij}-(n^2\sigma_n^2-n)]$. Let

$$g=c_{11}\mathbf{e}^T[n\mathbf{J}-\mathbf{I}]\mathbf{e}+nc_{11}\mathbf{e}^T\mathbf{P_x}\mathbf{J}\mathbf{P_x}\mathbf{e}-2nc_{11}\mathbf{e}^T\mathbf{J}\mathbf{P_x}\mathbf{e}-(1-2c_{11})\mathbf{e}^T\mathbf{P_x}\mathbf{e},$$
$$r=(n\sigma_B^2-1)\mathbf{e}^T\mathbf{e}+n\sigma_B^2 g.$$

Theorem 4.1. *Assume conditions* (2.3)-(2.5) *and that the errors are independent with uniformly bounded eighth moment. If the weights satisfy* (4.1)-(4.4) *and the residual* $\mathbf{R}$ *is used, then for the two variance estimators given in* (2.6) *and* (2.7),

$$n^{3/2}(\mathrm{V_{GB1}} - \mathrm{V_E}) = \left\{n^{-1/2} \sum (e_i^2 - \tau^2) + n^{-1/2} g\right\} \{n\mathbf{G}^{-1}\},$$
$$n^{3/2}(\mathrm{V_{GB2}} - \mathrm{V_E}) = \left\{n^{-1/2} \sum (e_i^2 - \tau^2) + n^{-1/2} r\right\} \{n\mathbf{G}^{-1}\}.$$

Thus $\mathrm{V_{GB1}}$ *is consistent, and* $\mathrm{V_{GB2}}$ *is consistent if* $\lim_{n\to\infty} n\sigma_n^2 = 1$. *Further,* $Eg = d_1$ *and* $Er = d_2$.

Theorem 4.1 shows the consistency of both variance estimators using the *GBS* in homoscedastic linear regression models. Note that the variance estimator $\mathrm{V_{GB1}}$ is very general and is valid for all choices of σ_n^2. However $\mathrm{V_{GB2}}$ requires $n\sigma_n^2 \to 1$. We now state a corollary of Theorem 4.1 to take care of the important special case of the residual bootstrap. This corollary uses the identification of the usual residual bootstrap as a *GBS* with the variance estimator being $\mathrm{V_{GB2}}$. Let $d_3 = \tau^2[p + 1 - n^{-1} \sum_{i\neq j} ((\mathbf{P_x}))_{ij}]$, $d_{13} = \tau^2[2 - n^{-1}(\sum x_i)^2 / \sum x_i^2]$.

Corollary 4.1. *For the classical residual bootstrap the following expansions hold for the multiple and simple linear regression, respectively:*

$$n^{3/2}(\mathrm{V_{GB2}} - \mathrm{V_E}) = \left\{n^{-1/2} \sum (e_i^2 - \tau^2) - n^{-1/2}(d_3 + C_{n10})\right\} \{n\mathbf{G}^{-1}\},$$
$$n^{3/2}(\mathrm{V_{GB2}} - \mathrm{V_E}) = \left\{n^{-1/2} \sum (e_i^2 - \tau^2) - n^{-1/2}(d_{13} + C_{n11})\right\} \frac{n}{\sum x_i^2}.$$

The second part of Corollary 4.1 matches the corresponding calculations in Section 3. We now discuss a special *GBS* scheme that is always second order optimal, in the sense that the expectation of the second order term is identically zero. Suppose $\{W_{ij},\ 1 \le i, j \le n\}$ are i.i.d. random variables with mean zero and variance $\sigma_n^2 = (n + p)/n^2$. These random variables are easily seen to satisfy (4.1)-(4.4).

Theorem 4.2. *Under the conditions of Theorem* 4.1, *for a GBS scheme with i.i.d. mean zero weights and* $\sigma_n^2 = n^{-2}(n + p)$, *the bootstrap variance estimator* $\mathrm{V_{GB2}}$ *is second order optimal and satisfies*

$$n^{3/2}(\mathrm{V_{GB2}} - \mathrm{V_E}) = \{n^{-1/2} \sum (e_i^2 - \tau^2) + C_{n12}\}\{n\mathbf{G}^{-1}\}.$$

Within the E-class, Theorem 4.2 shows that i.i.d. weights *GBS* is the better choice. Apart from efficiency considerations, let us also emphasise the significant computational advantage of using i.i.d. weights over other resampling methods. Still, much depends on the behaviour of the random variable C_{n12}. A little algebra shows that this term generally contributes only more lower order terms, and is thus unlikely to influence even small sample performance.

4.2. Results for R class bootstrap $p > 1$.

Liu and Singh (1992) showed that the paired bootstrap, the delete-1 jackknife, the external or wild bootstrap, and the weighted bootstrap of Liu (1988) belong to the R-class. In Section 3 we showed that the weighted jackknives proposed by Hinkley (1977) and Wu (1986) also belong to the R-class. From Chatterjee and Bose (2000) it is known that all the different delete-d jackknives belong to the R-class. We have remarked earlier that the paired bootstrap and the delete-d jackknives are special cases of the UBS. Recall that the external bootstrap is a GBS technique for which $\mathbf{W}$ is diagonal with entries that are i.i.d. observations from a random variable with zero mean and finite variance. It has been observed earlier (Liu and Singh (1992), and Section 3 here) that the variance estimator for the weighted bootstrap of Liu (1988) is same as that of the external bootstrap, so we mention the result for the external bootstrap only.

Let us introduce the following notation: $\mathbf{T} = diag(\tau_i^2,\ 1 \le i \le n)$; $\mathbf{T}_n = diag(e_i^2,\ 1 \le i \le n)$; $\mathbf{D} = diag(-2\tau_i^2((\mathbf{P_x}))_{ii} + \sum_{a=1}^n \tau_a^2((\mathbf{P_x}))_{ia}^2,\ \ 1 \le i \le n)$ $\mathbf{C}_E = diag(-2e_i \sum_{j=1}^n((\mathbf{P_x}))_{ij}e_j + \sum_{a,b=1}^n((\mathbf{P_x}))_{ia}((\mathbf{P_x}))_{ib}e_ae_b,\ \ 1 \le i \le n)$; $\delta_i = ((\mathbf{P_x}))_{ii} = \mathbf{x}_i^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_i$; $\mathbf{D}_{Wu} = diag(\delta_i{\tau_i}^2,\ \ 1 \le i \le n)$; $\mathbf{F} = \mathbf{X}^T(\mathbf{T}_n - \mathbf{T})\mathbf{X}$; $\mathbf{S} = \mathbf{X}^T\mathbf{D}\mathbf{X}$; and $\mathbf{A}_n = \{n\mathbf{G}^{-1}\}\{n^{-1/2}\mathbf{F}\}\{n\mathbf{G}^{-1}\}$.

Theorem 4.3. *Assume* (2.3)-(2.5) *and that the errors are independent with uniformly bounded eighth moment. Then the following expansions hold for the variance estimates.*

(i) *External Bootstrap:*

$$n^{3/2}(\mathrm{V_{GB3}} - \mathrm{V_R}) = \mathbf{A}_n + n^{-1/2}\{n\mathbf{G}^{-1}\}\mathbf{X}^T\mathbf{C}_E\mathbf{X}\{n\mathbf{G}^{-1}\}.$$

Further, $E\mathbf{C}_E = D$.

(ii) *Weighted Jackknife* (*Hinkley*) :

$$n^{3/2}(\mathrm{V_{JH}} - \mathrm{V_R}) = \mathbf{A}_n + n^{-1/2}\{n\mathbf{G}^{-1}\}\mathbf{X}^T[\frac{p}{n-p}\mathbf{T}_n - 2\mathbf{D}_2 + \mathbf{D}_3]\mathbf{X}\{n\mathbf{G}^{-1}\},$$

where $\mathbf{D}_2 = diag(e_i\mathbf{x}_i^T(\hat\beta - \beta),\ \ 1 \le i \le n)$*;* $\mathbf{D}_3 = diag([\mathbf{x}_i^T(\hat\beta - \beta)]^2,\ \ 1 \le i \le n)$*, and further* $E[(p/(n-p))\mathbf{T}_n - 2\mathbf{D}_2 + \mathbf{D}_3] = (p/(n-p))\mathbf{T} + \mathbf{D}$.

(iii) *Weighted Jackknife* (*Wu*) :

$$n^{3/2}(\mathrm{V_{JW}} - \mathrm{V_R}) = \mathbf{A}_n + n^{-1/2}\{n\mathbf{G}^{-1}\}\mathbf{X}^T\mathbf{D}_4\mathbf{X}\{n\mathbf{G}^{-1}\},$$

where $\mathbf{D}_4 = diag([\delta_i\mathbf{T}_{ni} - 2\mathbf{D}_{2i} + \mathbf{D}_{3i}] + \delta_i^2(1-\delta_i)^{-1})\mathbf{T}_{ni} + [\delta_i + \delta_i^2(1-\delta_i)^{-1})][-2\mathbf{D}_{2i} + \mathbf{D}_{3i}],\ \ 1 \le i \le n)$*. Further,* $E\mathbf{D}_4 = \mathbf{D}_{Wu} + \mathbf{D}$.

(iv) *For* UBS *satisfying* (2.9)-(2.17):

$$n^{3/2}(\mathrm{V_{UBS}}-\mathrm{V_R})=\mathbf{A}_n+n^{-1/2}\{n\mathbf{G}^{-1}\}\mathbf{X}^T\mathbf{M}\mathbf{X}\{n\mathbf{G}^{-1}\}+n^{-1}C_{n13},$$

where

$$\begin{aligned}
\mathbf{M}_{ii}&=\sigma_n^2[3c_4e_i^2((\mathbf{P_x}))_{ii}^2+c_{22}\sum_{a\neq i}e_a^2((\mathbf{P_x}))_{ia}^2]+(\sum_{a=1}^n e_a(\mathbf{P_x})_{ia}^2-2\sigma_nc_3e_i^2(\mathbf{P_x})_{ii}^2\\
&\quad-2e_i\sum_{a=1}^n e_a(\mathbf{P_x})_{ia}+2\sigma_n^2c_{31}e_i^2\sum_{a\neq i}((\mathbf{P_x}))_{ia}^2,\\
\mathbf{M}_{ij}&=\sigma_n^2[2c_{31}\{e_i^2((\mathbf{P_x}))_{ii}((\mathbf{P_x}))_{ij}+e_j^2((\mathbf{P_x}))_{jj}((\mathbf{P_x}))_{ij}\}\\
&\quad+c_{211}\sum_{a\neq i,j}e_a^2((\mathbf{P_x}))_{ia}((\mathbf{P_x}))_{ja}+c_{211}(e_i^2+e_j^2)\sum_{a\neq i,j}((\mathbf{P_x}))_{ia}((\mathbf{P_x}))_{ja}\\
&\quad+c_{22}\{e_j^2((\mathbf{P_x}))_{ii}((\mathbf{P_x}))_{ij}+e_i^2((\mathbf{P_x}))_{jj}((\mathbf{P_x}))_{ij}\}]+c_{11}e_ie_j\\
&\quad+c_{11}(\sum_{a=1}^n\sum_{1\leq b\leq n}e_ae_b(\mathbf{P_x})_{ia}(\mathbf{P_x})_{jb})-c_{11}[e_j\sum_{a=1}^n e_a(\mathbf{P_x})_{ia}+e_i\sum_{a=1}^n e_a(\mathbf{P_x})_{ja}]\\
&\quad+\sigma_n[c_{21}(e_i^2+e_j^2)(\mathbf{P_x})_{ij}+c_{21}e_ie_j(\mathbf{P_x})_{ii}+\mathbf{P_x})_{jj})],\\
E\mathbf{M}_{ii}&=\sigma_n^2[3c_4{\tau_i}^2((\mathbf{P_x}))_{ii}^2+c_{22}\sum_{a\neq i}{\tau_a}^2((\mathbf{P_x}))_{ia}^2]+\sum_{a=1}^n{\tau_a}^2(\mathbf{P_x})_{ia}^2-2\sigma_nc_3{\tau_i}^2(\mathbf{P_x})_{ii}^2\\
&\quad-2{\tau_i}^2(\mathbf{P_x})_{ii}+2\sigma_n^2c_{31}{\tau_i}^2\sum_{a\neq i}((\mathbf{P_x}))_{ia}^2,\\
E\mathbf{M}_{ij}&=\sigma_n^2[2c_{31}\{{\tau_i}^2((\mathbf{P_x}))_{ii}((\mathbf{P_x}))_{ij}+{\tau_i}^2((\mathbf{P_x}))_{jj}((\mathbf{P_x}))_{ij}\}\\
&\quad+c_{211}\sum_{a\neq i,j}{\tau_a}^2((\mathbf{P_x}))_{ia}((\mathbf{P_x}))_{ja}+c_{211}({\tau_i}^2+{\tau_j}^2)\sum_{a\neq i,j}((\mathbf{P_x}))_{ia}((\mathbf{P_x}))_{ja}\\
&\quad+c_{22}\{{\tau_j}^2((\mathbf{P_x}))_{ii}((\mathbf{P_x}))_{ij}+{\tau_i}^2((\mathbf{P_x}))_{jj}((\mathbf{P_x}))_{ij}\}]\\
&\quad+c_{11}(\sum_{a=1}^n{\tau_a}^2(\mathbf{P_x})_{ia}(\mathbf{P_x})_{ja})-c_{11}[{\tau_j}^2+{\tau_i}^2](\mathbf{P_x})_{ij}+\sigma_n[c_{21}({\tau_i}^2+{\tau_j}^2)(\mathbf{P_x})_{ij}].
\end{aligned}$$

Exact expression for the delete-d jackknives and the paired bootstrap schemes can be obtained by inserting the exact values of the different moments involved.

Comparison between the external bootstrap, Hinkley's jackknife and Wu's jackknife is difficult without further assumptions. The second term in the expansion for these three schemes involves a diagonal matrix, and with a little algebra it can be seen that in invoking criterion (1.3) we may compare the expected trace of these diagonal matrices. If ${\tau_i}^2\equiv\tau^2$, then the ith terms in the expectations are respectively $-\tau^2(\mathbf{P_x})_{ii}$, $-\tau^2[(\mathbf{P_x})_{ii}+p/(n-p)]$ and 0 for the external bootstrap, Hinkley's jackknife and Wu's jackknife. Hence the respective traces work out to be $-p\tau^2$, $p^2\tau^2/(n-p)$ and 0. This shows that Wu's jackknife may be expected to perform better than Hinkley's jackknife, which in turn may be expected to perform better than the external bootstrap. Indeed, Wu's jackknife shows good performance in simulations.

If we use UBS with i.i.d., symmetric weights, whenever there is an odd number in the index of $c_{ijk\ldots}$, that term disappears. Then the expectation of the second order term is $n^{-1/2}\{n\mathbf{G}^{-1}\}\mathbf{X}^T[\mathbf{D}+\sigma_n^2\mathbf{D}_{UBS}]\mathbf{X}\{n\mathbf{G}^{-1}\}$, where $(\mathbf{D}_{UBS})_{ii} = 3c_4{\tau_i}^2((\mathbf{P}_{\mathbf{x}}))^2_{ii} + \sum_{a\neq i}{\tau_a}^2((\mathbf{P}_{\mathbf{x}}))^2_{ia}$, $(\mathbf{D}_{UBS})_{ij} = {\tau_j}^2((\mathbf{P}_{\mathbf{x}}))_{ii}((\mathbf{P}_{\mathbf{x}}))_{ij} + {\tau_i}^2((\mathbf{P}_{\mathbf{x}}))_{jj}((\mathbf{P}_{\mathbf{x}}))_{ij}$. Since this is not a diagonal matrix, comparison with other resampling techniques is not easy. Observe that by choosing σ_n^2 such that $trace\left[\mathbf{X}^T[\mathbf{D}+\sigma_n^2\mathbf{D}_{UBS}]\mathbf{X}\right]=0$, a UBS technique can be second order optimal. This naturally requires additional model conditions, comparable to (1.4), so as to ensure the compatibility with $\sigma_n^2>0$. Details of such conditions may be found in Bose and Chatterjee (1998). We do not discuss the 'optimality' of UBS in details here, since the choice of σ_n^2 is more likely to be based on practical convenience. Results in Section 5 show that such practical (and perhaps sub-optimal) choices of σ_n^2 also lead to good resampling performance.

5. Results from Simulation Experiments

As an illustration we consider the design matrix from the oxygen uptake data from Rawlings, Pantula, Dickey ((1998), page 124). Since the design matrix is fixed in our results, we take the design matrix from this data as fixed. We fix a set of five parameter values, between 1.79 and 2.31, and vary the noise process to generate different sets of data, then estimate the parameters using least squares. We study the distribution of the estimated variance of the contrast given by $(-0.25,-0.25,-0.25,-0.25,1)$. Extensive simulations with other linear combinations of parameters and different design matrices yield similar conclusions.

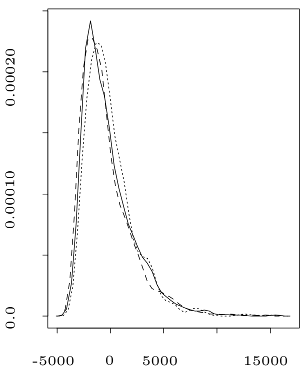
With noise variances fixed, the actual variance of the least squares estimate ($\mathrm{V_n}$) is known, and for every noise sequence and every resampling plan this is estimated. Say for a particular resampling plan the variance estimator is called V_{nj}. From our theorems, we know that $n^{3/2}(V_{nj}-\mathrm{V_n}) = A_n + n^{-1/2}B_{nj} + O_p(n^{-1})$ holds, and we want to check whether EB_{nj} differs significantly across j (resampling techniques). We present the density plots of $n^{3/2}(V_{nj}-\mathrm{V_n})$ for three different kinds of independent noise random variables with sample size 15 and 31.

Figures 1, 2 and 3 use the first fifteen design points from the oxygen uptake data. Figures 4, 5 and 6 consider the full design matrix based on 31 observations.

Noise is generated as follows. Let $\tau_i = 2\log(i+2)$. For Figures 1 and 4, the ith noise is $Normal(0,{\tau_i}^2)$. For Figures 2 and 5, the ith noise is $Exponential$ with mean τ_i, left shifted so that the mean becomes zero. For Figures 3 and

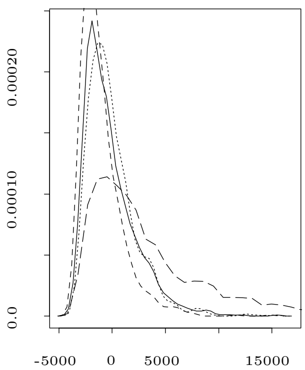
6, the ith noise is *Gamma* with parameters 10 and τ_i, also left shifted so that the mean becomes zero. Note that the given value of τ_i implies that the noise variance ranges from moderate to quite high values.

Figure 1A



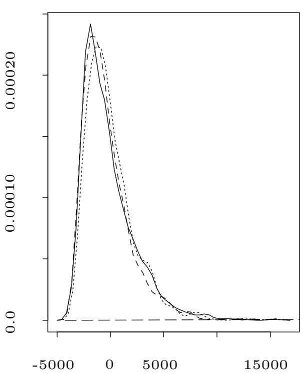
Simulated, Discrete, Wu

Figure 1B



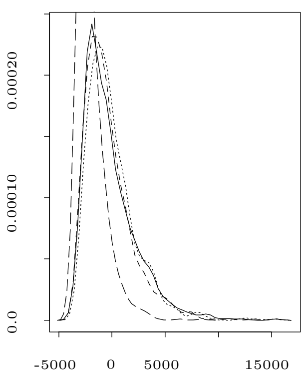
Simulated, Discrete, Hinkley, Jackknife

Figure 1C



Simulated, Discrete, Uniform, Paired

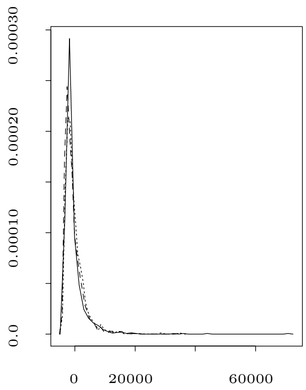
Figure 1D



Simulated, Discrete, Uniform, External

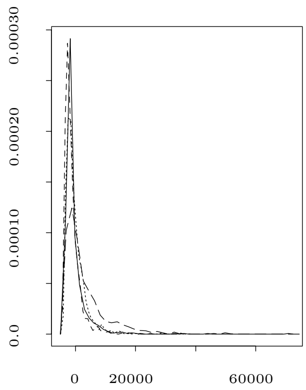
Figure 1. Density plots of $n^{3/2}(V_{nj} - \mathrm{V_n})$ for different bootstrap variance estimators of the least squares estimator variance from $y_i = \mathbf{x}_i^T\beta + e_i$, where $e_i \sim N(0, 4[\log(i+2)]^2)$, data size $n = 15$. In all figures, the bold line is the simulated true distribution, the dotted line is the UBS estimator with i.i.d. weights $w_{ni} = 0.15(1.85)$ with probability 0.5(0.5). The dashed line is Wu's weighted jackknife estimator in A, Hinkley's weighted jackknife in B, and the UBS estimator with i.i.d. weights $w_{ni} \sim U(0, 2)$ in C and D. Longer dashed lines are the usual unweighted jackknife, paired bootstrap and external bootstrap estimators in B, C and D, respectively.

Figure 2A



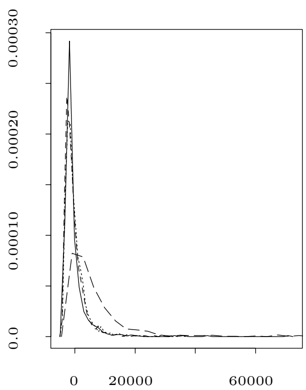
Simulated, Discrete, Wu

Figure 2B



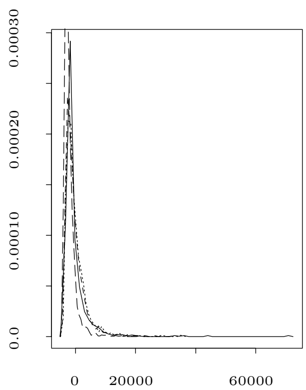
Simulated, Discrete, Hinkley, Jackknife

Figure 2C



Simulated, Discrete, Uniform, Paired

Figure 2D

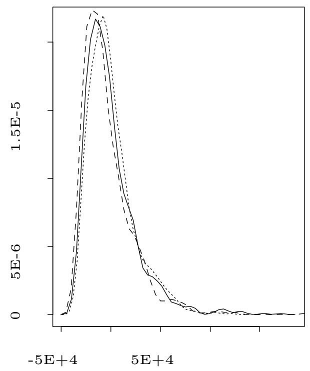


Simulated, Discrete, Uniform, External

Figure 2. Density plots of $n^{3/2}(V_{nj} - \mathrm{V_n})$ for different bootstrap variance estimators of the least squares estimator variance from $y_i = \mathbf{x}_i^T \beta + e_i$, where $e_i \sim exponential(2\log(i+2)) - 2\log(i+2)$, data size $n = 15$. In all figures, the bold line is the simulated true distribution, the dotted line is the UBS estimator with i.i.d. weights $w_{ni} = 0.15(1.85)$ with probability 0.5(0.5). The dashed line is Wu's weighted jackknife estimator in A, Hinkley's weighted jackknife in B, and the UBS estimator with i.i.d. weights $w_{ni} \sim U(0, 2)$ in C and D. Longer dashed lines are the usual unweighted jackknife, paired bootstrap and external bootstrap estimators in B, C and D, respectively.

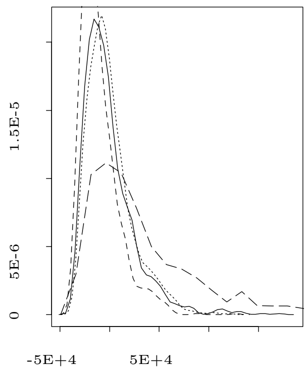
Although an optimal UBS depends on the choice of σ_n^2 (the variance of the weights) based on design parameters, we ignore this fact and work with practical selections of i.i.d. weights, which are easy to generate. The two i.i.d. UBS

Figure 3A



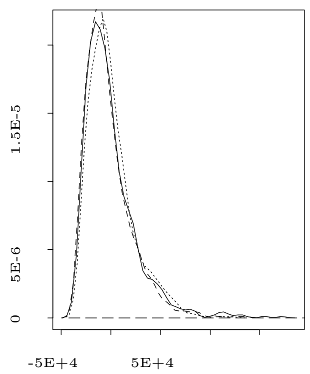
Simulated, Discrete, Wu

Figure 3B



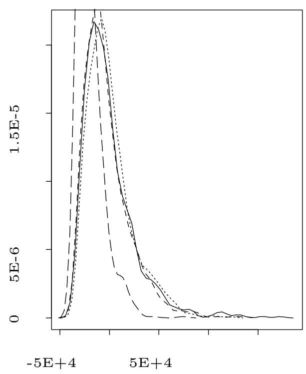
Simulated, Discrete, Hinkley, Jackknife

Figure 3C



Simulated, Discrete, Uniform, Paired

Figure 3D

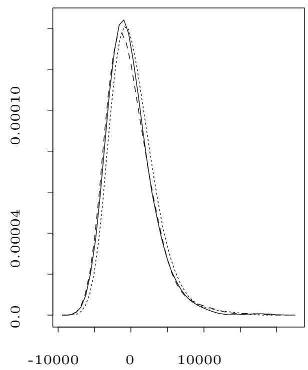


Simulated, Discrete, Uniform, External

Figure 3. Density plots of $n^{3/2}(V_{nj} - \mathrm{V_n})$ for different bootstrap variance estimators of the least squares estimator variance from $y_i = \mathbf{x}_i^T\beta + e_i$, where $e_i \sim gamma(10, 2\log(i+2)) - 20\log(i+2)$, data size $n = 15$. In all figures, the bold line is the simulated true distribution, the dotted line is the UBS estimator with i.i.d. weights $w_{ni} = 0.15(1.85)$ with probability 0.5(0.5). The dashed line is Wu's weighted jackknife estimator in A, Hinkley's weighted jackknife in B, and the UBS estimator with i.i.d. weights $w_{ni} \sim U(0, 2)$ in C and D. Longer dashed lines are the usual unweighted jackknife, paired bootstrap and external bootstrap estimators in B, C and D, respectively.

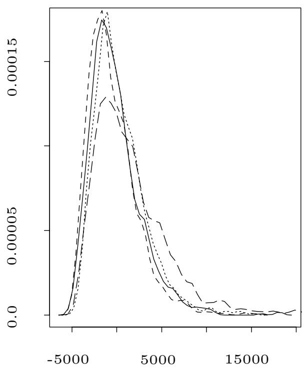
schemes chosen for simulation are as follows. The first is a two-point distribution, putting equal masses on 0.15 and 1.85, called the *discrete UBS*. This is a slightly modified version of the optimal UBS for the variance of the sample mean in the one-dimensional case. The other UBS scheme uses weights from a mean 1, variance 1/3 $Uniform$ distribution, and is called the *Uniform UBS*. We also include the paired bootstrap and delete-1 jackknife in our study, also UBS techniques.

Figure 4A



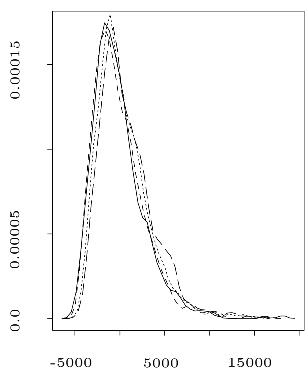
Simulated, Discrete, Wu

Figure 4B



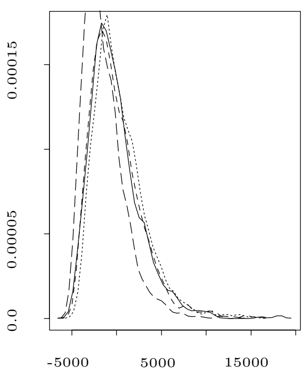
Simulated, Discrete, Hinkley, Jackknife

Figure 4C



Simulated, Discrete, Uniform, Paired

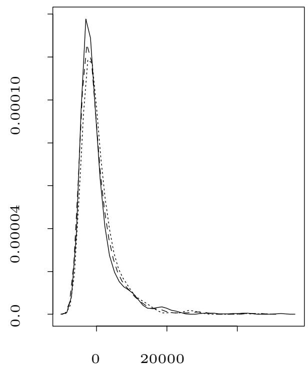
Figure 4D



Simulated, Discrete, Uniform, External

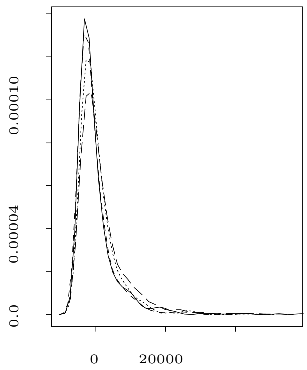
Figure 4. Density plots of $n^{3/2}(V_{nj} - \mathrm{V_n})$ for different bootstrap variance estimators of the least squares estimator variance from $y_i = \mathbf{x}_i^T\beta + e_i$, where $e_i \sim N(0, 4[\log(i+2)]^2)$, data size $n = 31$. In all figures, the bold line is the simulated true distribution, the dotted line is the UBS estimator with i.i.d. weights $w_{ni} = 0.15(1.85)$ with probability 0.5(0.5). The dashed line is Wu's weighted jackknife estimator in A, Hinkley's weighted jackknife in B, and the UBS estimator with i.i.d. weights $w_{ni} \sim U(0, 2)$ in C and D. Longer dashed lines are the usual unweighted jackknife, paired bootstrap and external bootstrap estimators in B, C and D, respectively.

Figure 5A



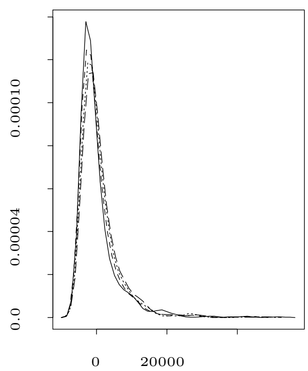
Simulated, Discrete, Wu

Figure 5B



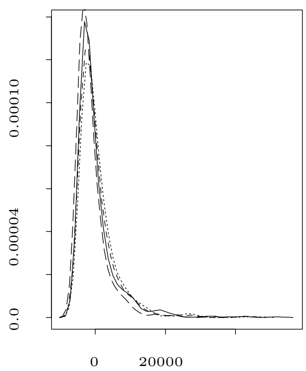
Simulated, Discrete, Hinkley, Jackknife

Figure 5C



Simulated, Discrete, Uniform, Paired

Figure 5D



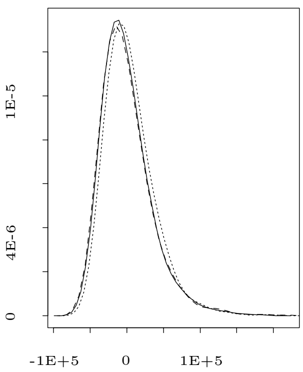
Simulated, Discrete, Uniform, External

Figure 5. Density plots of $n^{3/2}(V_{nj} - \mathrm{V_n})$ for different bootstrap variance estimators of the least squares estimator variance from $y_i = \mathbf{x}_i^T\beta + e_i$, where $e_i \sim exponential(2\log(i+2)) - 2\log(i+2)$, data size $n = 31$. In all figures, the bold line is the simulated true distribution, the dotted line is the UBS estimator with i.i.d. weights $w_{ni} = 0.15(1.85)$ with probability 0.5(0.5). The dashed line is Wu's weighted jackknife estimator in A, Hinkley's weighted jackknife in B, and the UBS estimator with i.i.d. weights $w_{ni} \sim U(0, 2)$ in C and D. Longer dashed lines are the usual unweighted jackknife, paired bootstrap and external bootstrap estimators in B, C and D, respectively.

The estimated densities are denoted as follows: (1) Wu's jackknife: dashed line (Figure A, 1-6); (2) Hinkley's jackknife: dashed line (Figure B, 1-6); (3) Standard delete-1 jackknife: longer dashed line (Figure B, 1-6); (4) *Uniform*

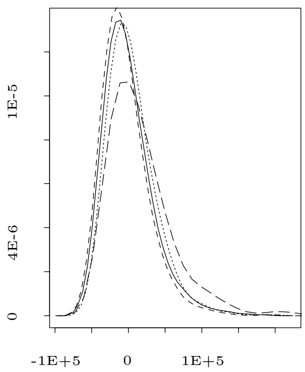
UBS: dashed line (Figure C, D, 1-6); (5) Paired bootstrap: longer dashed line (Figure C, 1-6); (6) External bootstrap: longer dashed line (Figure D, 1-6); (7) *Discrete UBS*: dotted line (all figures); (8) Simulated data density: solid line (all figures).

Figure 6A



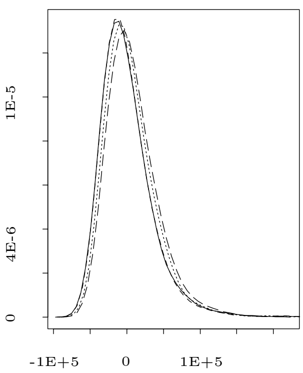
Simulated, Discrete, Wu

Figure 6B



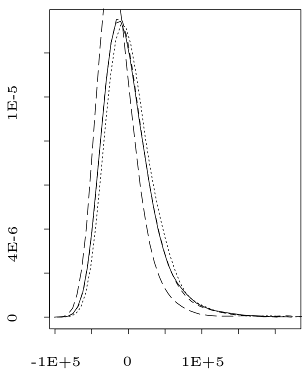
Simulated, Discrete, Hinkley, Jackknife

Figure 6C



Simulated, Discrete, Uniform, Paired

Figure 6D



Simulated, Discrete, Uniform, External

Figure 6. Density plots of $n^{3/2}(V_{nj} - \mathrm{V_n})$ for different bootstrap variance estimators of the least squares estimator variance from $y_i = \mathbf{x}_i^T\beta + e_i$, where $e_i \sim gamma(10, 2\log(i+2)) - 20\log(i+2)$, data size $n = 31$. In all figures, the bold line is the simulated true distribution, the dotted line is the UBS estimator with i.i.d. weights $w_{ni} = 0.15(1.85)$ with probability $0.5(0.5)$. The dashed line is Wu's weighted jackknife estimator in A, Hinkley's weighted jackknife in B, and the UBS estimator with i.i.d. weights $w_{ni} \sim U(0,2)$ in C and D. Longer dashed lines are the usual unweighted jackknife, paired bootstrap and external bootstrap estimators in B, C and D, respectively.

The figures show that certain resampling techniques work well even when sample size is small and noise variance is high. We have tried other noise variances to get similar results. The *Discrete UBS*, *Uniform UBS* and *Wu's jackknife* estimate the densities remarkably well, even in small samples. *Hinkley's jackknife* also does reasonably well, although there is some underestimation of the right tail. The *paired bootstrap* performs very badly with small samples, as expected, but does reasonably well with moderate sized samples. The usual *delete*-1 *jackknife* and *external bootstrap* perform poorly. The external bootstrap substantially underestimates the right tail, even for $n = 31$. We have checked with data sets of size 80 and found that the bias in delete-1 jackknife and external bootstrap persists.

It may be mentioned in the context of weighted jackknives, that the use of the diagonal elements of $\mathbf{P_x}$ is special to least squares estimation and linear regression. The use of weighted jackknives may be difficult with other estimators or other models. However the performance of the weighted jackknives may improve further if instead of a delete-1 jackknife, a delete-d jackknife is used, but then computational difficulties and data storage problems must be considered.

In conclusion, the weighted jackknife of Wu (1986) and some UBS schemes seem to do well in simulations, and bear further study. The relative advantage seems to lie with the UBS technique, which can be proven to be optimal under certain conditions, and is perhaps less difficult to implement than any weighted jackknife.

Acknowledgement

We are grateful to the Referees for careful reading of the manuscript and highly constructive suggestions. This has led to a better presentation and discussion of the results and to a much improved manuscript. We thank the Editor for encouraging us to work on a revision.

References

Barbe, P. and Bertail, P. (1995). *The Weighted Bootstrap.* Lecture Notes in Statistics, Volume **98**. Springer - Verlag, New York.

Bickel, P. J. and Freedman, D. A. (1983). Bootstrapping regression models with many parameters. In *Festschrift for Erich L. Lehmann* (Edited by P. Bickel, K. Doksum and J. Hodges, Jr.), 28-48. Wadsworth, Belmont, Calif.

Bose, Arup and Chatterjee, S. (1997). Second order comparison of different resampling schemes in linear regression. Tech. Report No. 15/97, Stat-Math Unit, Indian Statistical Institute, Calcutta. Available from http://www.math.unl.edu/~schatterjee.

Bose, Arup and Chatterjee, S. (1998). Second order comparison of different bootstrap and jackknife methods in multiple regression. Tech. Report No. 10/98, Stat-Math Unit, Indian Statistical Institute, Calcutta. Available from http://www.math.unl.edu/~schatterjee.

Chatterjee, S. and Bose, Arup (2000). Variance estimation in high dimensional regression models. *Statist. Sinica* **10**, 497-515.

Chatterjee, S. and Bose, Arup(2000a). Dimension asymptotics for the uncorrelated weights bootstrap in linear regression. To appear in *Ann. Inst. Math. Statist.*

Chatterjee, S. (1998). Another look at the jackknife: further examples of generalised bootstrap. *Statist. Probab. Lett.* **40**, 307-319.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7**, 101-118.

Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics* **19**, 285-292.

Liu, R. Y. (1988). Bootstrap procedures under some non i.i.d. models. *Ann. Statist.* **16**, 1697-1708.

Liu, R. Y. and Singh, K. (1992). Efficiency and robustness in resampling. *Ann. Statist.* **20**, 370-384.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Ann. Statist.* **21**, 255-285.

Praestgaard, J. and Wellner, J. A. (1993). Exchangeably weighted bootstrap of the general empirical process. *Ann. Probab.* **21**, 2053-2086.

Quenouille, M. (1949). Approximation tests of correlation in time series. *J. Roy. Statist. Soc. Ser. B* **11**, 18-84.

Rawlings, J. O., Pantula, S. G. and Dickey, D. A. (1998). *Applied regression analysis: a research tool*, 2nd edition. Springer-Verlag, New York.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *Ann. Statist.* **14**. 1261-1350.

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 203 B. T. Road, Calcutta 700 035, India.

E-mail: abose@isical.ac.in

Department of Mathematics and Statistics, University of Nebraska-Lincoln, Oldfather Hall, Lincoln, NE 68588, U.S.A.

E-mail: schatterjee@math.unl.edu

(Received May 1999; accepted September 2001)