L-STATISTICS WITH SMOOTH WEIGHT
FUNCTIONS JACKKNIFE WELL
by
William C. Parr
William R. Schucany

Technical Report No. 138
Department of Statistics ONR Contract
May 13, 1980

Research Sponsored by the Office of Naval Research
Contract N00014-75-C-0439/
Project NR 042-280

Reproduction in whole or in part is permitted
for any purpose of the United States Government

This document has been approved for public release
and sale; its distribution is unlimited

DIVISION OF MATHEMATICAL SCIENCES
Department of Statistics/
Southern Methodist University
Dallas, Texas 75275
L-STATISTICS WITH SMOOTH WEIGHT FUNCTIONS JACKKNIFE WELL

by

William C. Parr and William R. Schucany

ABSTRACT

The behavior of linear combinations of order statistics (L-statistics) under jackknifing is discussed. The asymptotic behavior of a jackknifed L-statistic is produced, and the pseudo-value based variance estimate is shown to be consistent under moderate smoothness and a trimming condition of the weight function. The results extend easily to functions of L-statistics and thus answer a question posed in Miller (1974). Monte Carlo results support small-sample applicability of the large-sample results for the construction of approximate confidence intervals.

Key Words: Confidence intervals; Jackknife; Linear combinations of order statistics; Variance estimation.

William C. Parr is Assistant Professor, Institute of Statistics, Texas A&M University, College Station, Texas 77843. William R. Schucany is Professor of Statistics, Southern Methodist University, Dallas, Texas 75275. Research of the second author was supported in part by ONR Contract N00014-75-C-0439. The authors would like to thank Emanuel Parzen for helpful conversations.
1. INTRODUCTION, NOTATION, AND DEFINITIONS

Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a distribution with distribution function $F$, and let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ denote the associated order statistics. For a fixed weight function $J(u)$ defined for $0 < u < 1$, we define the L-statistic

$$S_n = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{i:n}. \quad (1.1)$$

Other definitions, typically asymptotically equivalent to that above, include $T_n = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n}\right) X_{i:n}$ and $U_n = \frac{1}{n} \sum_{i=1}^{n} c_{i,n} X_{i:n}$

where $c_{i,n} = \int_{(i-1)/n}^{i/n} J(u)du$. $S_n$ is chosen for study in this paper as being typical of actual L-statistics used in practice (use of $U_n$, in which the integration "smoothes" $J$, might result in fewer conditions on the weight function).

L-statistics of the form of $S_n$ are often used in estimation problems since they are typically computationally simple and (at least for location and scale problems) asymptotically efficient given the proper choice of the weight function $J$. Thus, they are often good choices a) as estimates for their own sake, b) as good starting values for iterative estimation
procedures, and c) as quick and consistent estimators of nuisance parameters (such as unknown scale in regression problems) to minimize the number of parameters being simultaneously estimated via an iterative procedure. Herein we consider primarily a), where $L$-statistics are used to make parametric inferences on their own. However, $S_n$ is often biased as an estimator of $S_0 = \int Q(u) J(u) du$ where $Q(u) = \inf\{ x : F(x) > u \}$, as is $g(S_n)$ biased for $g(S_0)$ for many choices of $g$. Also, there seems to be a dearth of procedures for consistent nonparametric estimation of the variance of $\sqrt{n} (S_n - S_0)$. (But see Sen (1979) in this regard.) For both of these problems, reduction of bias and consistent nonparametric variance estimation, the jackknife is a natural choice as a possibly non-optimal rough-and-ready tool.

The ordinary jackknife of $S_n$ may be written as

\[
S_n = \frac{1}{n} \sum_{i=1}^{n} S_{i,n}
\]

\[
= \frac{n}{n+1} \sum_{i=1}^{n} X_{i,n} - \frac{1}{n} \sum_{i=1}^{n} \left[ (1 - 1)J\left(\frac{i-1}{n}\right) + (n - 1)J\left(\frac{i}{n}\right) \right] X_{i,n}
\]

\[
= n S_n - \frac{n-1}{n} \sum_{i=1}^{n} S(i)
\]
where the $i$th pseudo-value is

$$S_{i,n} = \frac{1}{n} \sum_{j=1}^{n} J_{n} \left( \frac{i - 1}{n} \right) x_{j,n} - \frac{1}{n} \sum_{j=1}^{n} J_{n} \left( \frac{j - 1}{n} \right) x_{j,n} \quad (1.4)$$

$S_{n-1}^{(1)}$ is the same $L$-statistic computed after deleting $x_{i,n}$ from the sample, and $I_{h(i)}$ is understood to be defined to be zero if $a > b$. (See Miller (1974) for basic definitions of the jackknife and pseudo-values in general.) Note that the definition of $J$ at 0 and 1 is completely arbitrary since the associated terms will cancel out in (1.3).

We further define the sample variance of the pseudo-values as

$$s_{p}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (S_{i,n} - \bar{S}_{n})^2 \quad (1.5)$$

2. RESULTS ON JACKKNIFING $L$-STATISTICS

**Theorem 1:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a distribution $F$ with $E_{F}|X|^p < \infty$ for some $\frac{1}{2} < p < 1$.

Further, let $S_n$ and $\bar{S}_n$ be as defined by (1.1) and (1.3). If

a) $p = 1$ and $J'$ satisfies a Hölder condition with $a > \frac{1}{2}$, or

b) $\frac{1}{2} < p < 1$ and $J'$ satisfies a Hölder condition with
\[ a = \frac{1}{p} - \frac{1}{2}, \text{ then } \sqrt{n}(S_n - S_n) \to 0 \text{ with probability } 1 \text{ as } n \to \infty. \]

**Proof:** From elementary algebra, \( S_n - S_n = \sum_{i=1}^{n} D_j(i,n)X_i,n' \), where

\[ D_j(i,n) = \frac{n - 1}{n} J\left(\frac{i}{n+1}\right) - \frac{i - 1}{n} J\left(\frac{i - \frac{1}{n}}{n}\right) - \frac{n - \frac{i}{n}}{n} J\left(\frac{i}{n}\right), \]

\[ 1 \leq i \leq n. \]

Since \( J' \) obeys a Hölder condition with exponent \( \alpha \) we have (with \( \frac{1}{n+1} < \xi_{in1} < \frac{i}{n} \))

\[ J\left(\frac{i}{n}\right) = J\left(\frac{i}{n+1}\right) + J'\left(\frac{i}{n+1}\right)\left(\frac{i}{n} - \frac{i}{n+1}\right) \]

\[ + \left( J'\left(\xi_{in1}\right) - J'\left(\frac{i}{n+1}\right)\left(\frac{i}{n} - \frac{i}{n+1}\right) \right) \]

\[ = J\left(\frac{i}{n+1}\right) + J'\left(\frac{i}{n+1}\right) \frac{1}{n(n+1)} + B_{in1} \left| \frac{i}{n(n+1)} \right|^{1+\alpha} \]

and (with \( \frac{1}{n} < \xi_{in2} < \frac{i}{n+1} \))
\[
J\left(\frac{i - 1}{n}\right) = J\left(\frac{i}{n + 1}\right) + J'\left(\frac{i}{n + 1}\right)\left[\frac{i - 1}{n} - \frac{i}{n + 1}\right]
+ (J'\xi_{\text{in}2}) - J'\left(\frac{i}{n + 1}\right)\left(\frac{i - 1}{n} - \frac{i}{n + 1}\right)

= J\left(\frac{i}{n + 1}\right) + J'\left(\frac{i}{n + 1}\right)\frac{i - n - 1}{n(n + 1)}
+ B_{\text{in}2}\left|\frac{i - n - 1}{n(n + 1)}\right|^{1+\alpha},
\]

where the $B_{\text{in}1}$ and $B_{\text{in}2}$ are all uniformly bounded in absolute value by some finite positive constant $B$. Hence,

\[
D_j(i,n) = -J'\left(\frac{i}{n + 1}\right) \frac{n - 2i + 1}{n^2(n + 1)} - \frac{n - i}{n} B_{\text{in}1}\left(\frac{i}{n(n + 1)}\right)^{1+\alpha}
- \frac{i - 1}{n} B_{\text{in}2}\left|\frac{i - n - 1}{n(n + 1)}\right|^{1+\alpha}.
\]

Thus,

\[
\sqrt{n}|\bar{S}_n - S_n| \leq \sqrt{n} \sum_{i=1}^{n} |D_j(i,n)||X_{i,n}|
\leq n^{1+\alpha} \sup_{1 \leq i \leq n} |D_j(i,n)| \sum_{i=1}^{n} \frac{|X_i|}{n(1+2\alpha)/2}
= Z_n.
\]
If \(0 < \frac{2}{1 + 2\alpha} < 1\) and \(E_{\xi} |X|^{\frac{2}{1 + 2\alpha}} < \infty\), with \(J'\) satisfying a Hölder condition with exponent \(\alpha\), \(Z_n \to 0\) with probability one, by Marcinkiewicz' theorem (Loeve (1977), p.254). If \(E_{\xi} |X| < \infty\) and \(J'\) satisfies a Hölder condition with exponent \(\alpha > \frac{1}{2}\), \(Z_n \to 0\) with probability one, using the strong law of large numbers.

Several comments are in order:

1) This result is the univariate analogue for L-statistics of result J.1 of Reeds (1978) for M-estimators. The extension to the multivariate case is immediate and hence omitted.

2) The moment condition on \(F\) is seen by an inspection of the proof to be superfluous if \(S\) "trims", i.e., if \(J(u) = 0\) for \(u \notin (0,\varepsilon) \cup (1 - \varepsilon, 1)\) for some \(0 < \varepsilon < \frac{1}{2}\).

3) If further \(\sqrt{n}(S_n - S_0) \xrightarrow{d} N(0, \sigma^2)\), where

\[
g^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(x))J(F(y))(F(\min(x,y)) - F(x)F(y)) \, dx \, dy > 0 ,
\]

then \(\sqrt{n}(S_n - S_0) \xrightarrow{d} N(0, \sigma^2)\) likewise. This is true if the conditions of Theorem 1 hold with \(p = 1\) and \(J'\) is of bounded
variation (using the result of D. S. Moore (1968), with $\sigma^2 < \infty$).

4) Since $J''$ bounded on $(0,1)$ implies that $J'$ satisfies a Hölder condition with $\alpha = 1$, the condition on $J$ and $F$ in the theorem may be replaced by the stronger but intuitively clearer condition that $J''$ is bounded and $E_F|X|^{2/3} < \infty$.

5) A similar theorem was stated under much stronger conditions by Thorburn (1976).

6) Theorem 1 yields a law of the iterated logarithm for the jackknife of a linear function of order statistics from the corresponding law for the original statistics. Using Theorem 4 (Example 1) of Wellner (1977b), we obtain that if $E_F|X|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$ (and $\alpha > \frac{1}{2}$)

$$\limsup_n \frac{\sqrt{n} |\tilde{S}_n - \int_0^1 Q(u)J_n(u)du|}{\sqrt{2 \sigma^2 \log \log n}} = 1$$

with probability one, where $J_n(u) = J\left(\frac{1}{n+1}\right)\frac{1}{n}$ for $\frac{1}{n+1} < u \leq \frac{1}{n}$, $1 \leq i \leq n$, and $J_i(0) = J\left(\frac{1}{n+1}\right)$.

7) Similarly, a Berry-Esseen rate for $\tilde{S}_n$ follows directly from that of Helmers (1977). If $E_F|X|^3 < \infty$, $\int_0^1 |J'(u)|dQ(u) \leq 0$, and $\sigma^2 > 0$, we quickly obtain that

$$\sup_x |F_n^*(x) - \Phi(x)| = O(n^{-1/2})$$
where \( F_n^*(x) \) is the cumulative distribution of \( (\bar{S}_n - E[S_n])/(\text{Var}S_n)^{1/2} \) 
and \( \phi(\cdot) \) is the standard normal cumulative.

The following theorem gives conditions under which the 
jackknife provides a consistent estimator of the asymptotic 
variance of \( \sqrt{n}(\bar{S}_n - S_o) \) and \( \sqrt{n}(\tilde{S}_n - S_o) \), and hence makes 
possible the construction of asymptotically pivotal quantities.

**Theorem 2:** Let \( X_1, \ldots, X_n \) be a random sample of size \( n \) from 
a distribution \( F \), and let \( s_p^2 \) be defined by (1.5). If there 
exist positive numbers \( \varepsilon \) and \( \delta \) such that \( 0 < \delta < \varepsilon < \frac{1}{2} \), 
(let \( C(\varepsilon) = [0, \varepsilon) \cup (1 - \varepsilon, 1] \); \( J(u) = 0 \) for \( u \in C(\varepsilon) \), 
\( fQ(u) \geq B > 0 \) for \( u \in C(\varepsilon - \delta) \), and \( J' \) 
meets a Hölder condition with exponent \( \alpha > 0 \), then 
\( s_p^2 \overset{p}{\to} \sigma^2 \), with \( \sigma^2 \) given by (2.1).

**Proof:** Note that \( \sigma^2 = \text{Var}(H(U)) \) where \( U \sim u(0,1) \) and 
\( H(u) = Q(u)J(u) - \int_0^1 Q(t)J'(t)[t - I(u \leq t)]dt, \) translating
the results of Boos (1979, eq. 3.3) into the quantile domain and then integrating by parts. Now, from (1.5), it will suffice for the desired result if we show that

$$\sup_{1 \leq i < n} |S_{i,n} - H\left(\frac{i}{n+1}\right)| \xrightarrow[p]{} 0, \quad \text{as } n \to \infty. \quad (2.2)$$

If (2.2) holds, then

$$s_p^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left( H\left(\frac{i}{n+1}\right) - \frac{1}{n} \sum_{j=1}^{n} H\left(\frac{j}{n+1}\right) \right)^2 + o(p).$$

Furthermore, \( \frac{1}{n-1} \sum_{i=1}^{n} \frac{1}{n} H^2\left(\frac{i}{n+1}\right) \) and \( \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n} H\left(\frac{i}{n+1}\right) \) will converge to \( \int H^2(u)du \) and \( \int H(u)du \) respectively, giving

$$s_p^2 \xrightarrow[p]{} \sigma^2. \quad \text{By definition,}$$

$$S_{i,n} = \sum_{j=1}^{n} J\left(\frac{j}{n+1}\right) X_{j,n} - \sum_{j=1}^{i-1} J\left(\frac{j}{n}\right) X_{j,n} = \sum_{j=i+1}^{n} J\left(\frac{j-1}{n}\right) X_{j,n}$$

$$= J\left(\frac{i}{n+1}\right) X_{i,n} - \sum_{j=1}^{i-1} (J\left(\frac{j}{n}\right) - J\left(\frac{j}{n+1}\right)) X_{j,n}$$

$$- \sum_{j=i+1}^{n} (J\left(\frac{j-1}{n}\right) - J\left(\frac{j}{n+1}\right)) X_{j,n}$$

$$= J\left(\frac{i}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^{i-1} J\left(\frac{j}{n+1}\right) X_{j,n} - \frac{1}{n} \sum_{j=i+1}^{n} J\left(\frac{j}{n+1}\right) X_{j,n}$$

$$= \left(\frac{1}{n+1} - 1\right) X_{i,n} + R_{ni1}$$
\[ n_i(T_{n-1}, n) = \frac{1}{n+1} J\left(\frac{1}{n+1}\right) X_1, n \]

\[ + \frac{1}{n} \sum_{j=1}^{n} J\left(\frac{1}{n+1}\right) \left(\frac{1}{n+1} - \frac{1}{n} \leq \frac{1}{n+1}\right) X_j, n \]

using Taylor expansions and the Hölder condition on \( J' \). Also using the trimming condition on \( J \), the remainder term is such that

\[ \sup_{1 \leq i \leq n} |R_{n2i}| = o_p(1) \text{ as } n \to \infty. \]

From the continuity and boundedness of \( J \) and the fact that \( fQ(u) > B > 0 \) whenever \( J(u) \neq 0 \),

\[ J\left(\frac{1}{n+1}\right) Q\left(\frac{1}{n+1}\right) + R_{n3i}, \text{ with } \sup_{1 \leq i \leq n} |R_{n3i}| = o_p(1). \]

Further

\[ \frac{1}{n} \sum_{j=1}^{n} J\left(\frac{1}{n+1}\right) \left(\frac{1}{n+1} - \frac{1}{n} \leq \frac{1}{n+1}\right) X_j, n \to \int_0^1 Q(t) J'(t) F(t) dt \text{ with probability } 1 \]

one from Corollary 2 to Theorem 1 of Wellner (1977).

Then, for fixed \( K \) let \( u_i = \frac{i}{K+1}, i = 0, 1, \ldots, K + 1 \).

It follows easily that

\[ L_i = \frac{1}{n} \sum_{j=\lfloor nu_i \rfloor}^{n} J\left(\frac{1}{n+1}\right) X_j, n \to \int_{u_i}^{1} J'(t) Q(t) dt \]

\[ = \int_0^1 J'(t) I(u_i \leq t) Q(t) dt. \]
Since K is finite, the convergence in probability is uniform in i, i = 0, 1, ..., K + 1. Furthermore, letting \( \inf_{i=1,...,K} |u_i - u| \) be achieved at index value \( i^* \),

\[
\sup_{0 \leq u \leq 1} |L_{i^*} - \frac{1}{n} \sum_{j=[nu]}^{n} X_j/n| \leq \sup_{0 \leq u \leq 1} |J'(u)| \frac{|F_n^{-1}(\varepsilon)| + |F_n^{-1}(1 - \varepsilon)|}{K + 1},
\]

which may be made less than any specified positive number in probability through a sufficiently large choice of K. Hence, (2.2) holds and \( s^2 \xrightarrow{p} \sigma^2 \), by the uniform continuity of \( \int J'(t)Q(t)dt \).

Some pertinent comments follow:

1) This result provides a method for consistent variance estimation for L-statistics, being the univariate analogue for L-statistics of result J.2 of Reeds (1978) for M-estimators.

2) Finiteness of \( \sigma^2 \) is clearly implied by the trimming and boundedness conditions on J.

3) This result makes possible the construction of nonparametric approximate confidence intervals for

\[
S_0 = \int_0^1 Q(u)J(u)du,
\]

using as pivotal quantities

\[
\sqrt{n} \left( \frac{S_n - S_0}{s_p} \right) \quad \text{or} \quad \sqrt{n} \left( \frac{S_n - S_0}{s_p} \right).
\]
This is, to the knowledge of the authors, the only nonparametric method of consistent variance estimation for L-statistics (i.e., in the absence of a specified parametric form for the unknown density) other than that of Sen (1979) discussed below. For a specific parametric family, a consistent estimator would of course typically be provided by

\[ \hat{\sigma}^2 = \int \int J(F_\theta(x))J(F_\theta(y))[F_\theta(\min(x,y)) \\
- F_\theta(x)F_\theta(y)] \, dx \, dy, \]  \hspace{1cm} (2.3)

where \( F_\theta, \theta \in \Omega \) is the parametric family of densities and \( \hat{\theta} \) is a weakly consistent estimator of \( \theta \), possibly multivariate.

4) It would be interesting to compare the properties of this estimator with that proposed by Sen (1979), which is essentially

\[ \tilde{\sigma}^2 = \int \int J(F_n(x))J(F_n(y))[F_n(x,y) - F_n(x)F_n(y)] \, dx \, dy, \]  \hspace{1cm} (2.4)

obtained by substituting the empirical distribution function \( F_n \) for \( F \) in (2.1).

5) Relaxation of the trimming condition on \( J \), which would require either moment conditions on \( F \) or joint conditions governing \( J(u) \) and \( Q(u) \) as \( u \) approaches 0 and 1 is questionable from
the standpoint of robust inference and hence Theorem 2 is stated as being of interest in its own right. Theorem 3, however, obtains strong convergence of $s_p^2$, while dropping the trimming condition, at the cost of moment assumptions.

**Theorem 3:** Let $X_1, \ldots, X_n$ be a random sample of size $n$ from a distribution $F$ with $E|X|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. If $J'$ is continuous on $[0,1]$, then $s_p^2 \rightarrow \sigma^2$ with probability one.

**Proof:** We proceed as in the proof of Theorem 2 to write

$$S_{1,n} = J \left( \frac{1}{n+1} \right) X_{1,n} - \frac{1}{n} \sum_{j=1}^{n} J' \left( \frac{1}{n+1} \right) \left( I(j \leq i) X_{j,n} - \frac{1}{n} \sum_{j=1}^{n} (J'(n_{j|i}) - J' \left( \frac{1}{n+1} \right)) (I(j \leq i) X_{j,n}),
\right.$$  

where $\frac{1}{n} \leq n_{j|i} \leq \frac{1}{n}$ and in particular $n_{i|i} = \frac{1}{n+1}$. The proof is then concluded by showing i) $\frac{1}{n} \sum_{i=1}^{n} S_i + \int_0^1 Q(u)J(u)du$, and ii) $\frac{1}{n} \sum_{i=1}^{n} S_{1,n} \rightarrow \sigma^2 + (\int_0^1 Q(u)J(u)du)^2$, both convergences being with probability one. Proceeding with the first part,
\[
\frac{1}{n} \sum_{i=1}^{n} S_{i,n} = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{i}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^{n} J'\left(\frac{1}{n+1}\right) \left(\frac{1}{n+1} - \frac{i}{n}\right) X_{j,n}
\]

\[
- \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n} \sum_{j=1}^{n} (J'(\eta_{nj}) - J'\left(\frac{1}{n+1}\right)) \left(\frac{1}{n+1} - I(j \leq i)\right) X_{j,n}.
\]

The first term converges to \(\int_{0}^{1} Q(u)J(u)\,du\) with probability one by Theorem 4, example 1 of Wellner (1977a).

The second term is \(-\frac{1}{n} \sum_{i=1}^{n} (J'(\eta_{nj}) - J'\left(\frac{1}{n+1}\right)) \left(\frac{1}{n+1} - I(j \leq i)\right) X_{j,n}\) which converges to 0 with probability one by the same result (since \(\frac{1}{n} \sum_{i=1}^{n} J'(\frac{1}{n+1}) \left(\frac{1}{n+1} - X_{j,n}\right)\) converges to \(\int_{0}^{1} Q(u)uJ'(u)\,du\) with probability one). Lastly,

\[
\left| \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n} \sum_{j=1}^{n} (J'(\eta_{nj}) - J'\left(\frac{1}{n+1}\right)) \left(\frac{1}{n+1} - I(j \leq i)\right) X_{j,n}\right|
\]

\[
\leq \sup_{i,j} \left| J'(\eta_{nj}) - J'\left(\frac{1}{n+1}\right) \right| \frac{1}{n} \sum_{i=1}^{n} |X_{j,n}| \frac{1}{n} \sum_{j=1}^{n} |X_{j,n}|
\]

which converges to zero with probability one using the uniform continuity of \(J'\) and the first moment assumption on \(F\). For the second part,
\[
\frac{1}{n} \sum_{i=1}^{n} S_{i,n}^2 = \frac{1}{n} \sum_{i=1}^{n} \left[ J \left( \frac{1}{n+1} \right) X_{i,n} - \frac{1}{n} \sum_{j=1}^{n} J' \left( \frac{1}{n+1} \right) \left( \frac{1}{n+1} - I(j \leq i) \right) X_{j,n} \right] \\
- \frac{1}{n} \sum_{j=1}^{n} \left( J' \left( \eta_{nj} \right) - J' \left( \frac{1}{n+1} \right) \left( \frac{1}{n+1} - I(j \leq i) \right) \right) X_{j,n}^2 \\
+ \frac{1}{n} \sum_{i=1}^{n} \left[ J \left( \frac{1}{n+1} \right) X_{i,n} - \frac{1}{n} \sum_{j=1}^{n} J' \left( \frac{1}{n+1} \right) \left( \frac{1}{n+1} - I(j \leq i) \right) X_{j,n} \right] \\
- \frac{1}{n} \sum_{j=1}^{n} \left( J' \left( \eta_{nj} \right) - J' \left( \frac{1}{n+1} \right) \left( \frac{1}{n+1} - I(j \leq i) \right) \right) X_{j,n}^2 \\
+ \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{1}{n} \sum_{j=1}^{n} \left[ J' \left( \eta_{nj} \right) - J' \left( \frac{1}{n+1} \right) \left( \frac{1}{n+1} - I(j \leq i) \right) \right] X_{j,n} \right\}^2 \\
= A_{3n} + A_{2n} + A_{1n} + A_{3n} ' .
\]

The third term, \( A_{3n} ' \), is less than or equal to (in absolute value)
\[
\left( \frac{1}{n} \sup_{i,j} \left| J' \left( \eta_{nj} \right) - J' \left( \frac{1}{n+1} \right) \sum_{i=1}^{n} \left| X_i \right| \right| \right)^2 ,
\]
which converges with probability one to zero via the strong law of large numbers and the continuity condition on \( J' \). The
second term, \( A_{2n} \), is disposed of in similar fashion. The first term, \( A_{1n} \), requires a more extended analysis.

\[
A_{1n} = \frac{1}{n} \sum_{i=1}^{n} j \left( \frac{i}{n+1} \right) j_{i,n}^2 + \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n+1} \sum_{j=1}^{n} j \left( \frac{j}{n+1} \right) \left( \frac{j}{n+1} - I(j \leq n) \right) j_{j,n}^2 
\]

\[
- \frac{2}{n} \sum_{i=1}^{n} j \left( \frac{i}{n+1} \right) j_{i,n} + \frac{1}{n} \sum_{j=1}^{n} j \left( \frac{j}{n+1} \right) j_{j,n} - I(j \leq n) j_{j,n}^2 
\]

\[
= B_{1n} + B_{2n} + B_{3n}
\]

\[
B_{1n} \to \int_0^1 q^2(u) j^2(u) du
\]

with probability one by Wellner's Theorem 4.

By a similar argument and rearrangement of terms,

\[
B_{2n} \to ( \int_0^1 q(u) j'(u) du )^2 - 2 \int_0^1 u j'(u) q(u) du \cdot \int_0^1 j'(u) (1-u) q(u) du
\]

\[
+ \int_0^1 \left( \int_0^u q(v) j'(v) dv \right)^2 du
\]

and

\[
B_{3n} \to -2 \int_0^1 q(u) j(u) du \cdot \int_0^1 q(u) j'(u) du
\]

\[
- 2 \int_0^1 q(u) j(u) \left( \int_0^u q(v) j'(v) dv \right) du
\]
both convergences holding with probability one. The result
then follows by integrating the expression for $\sigma^2$ given by (2.1)
by parts and observing that the quantity to which $s_p^2$ converges
with probability one is indeed $\sigma^2$. (Note that the appeals to
Wellner's result are actually to a slight modification of it
allowing random $J_n$ which satisfy the boundedness and convergence
conditions with probability one.)

**Example 1:** The sample mean, for which $J(u) \equiv 1$, clearly
satisfies the conditions of Theorem 1 (but not the trimming
conditions of Theorem 2) if $E_P|X|^{2/3} < \infty$. In fact, $S_n \equiv S_n$
\[ n \sum (X_i - \bar{X})^2 \]
in this case, and
\[ s_p^2 = \frac{\sum_{i=1}^{n-1} X_i^2}{n-1} \xrightarrow{P} \sigma_P^2 \text{ if } E_P|X|^2 < \infty. \]
Theorem 3 requires $E_P|X|^{2+\epsilon} < \infty$ for some $\epsilon > 0$. Thus, the
usual strong law for the sample variance "just fails" to
be a corollary of Theorem 3.

**Example 2:** While the ordinary trimmed means do not satisfy
the score conditions of Theorems 1 and 2, a smoothed version
causing $J$ to return to zero in such a way that it is differentiable with $J'$ obeying a Hölder condition with $\alpha > \frac{1}{2}$ would
satisfy those conditions. We assume that the modified score
is also zero on $C(\epsilon)$. 
Example 3: Gini's mean difference \((J(u) = u - \frac{1}{2})\) and the optimal score for location estimation for a logistic population \((J(u) = 6u(1 - u))\) clearly satisfy the conditions of Theorem 1 if \(E_F(X) < \sigma\) and those of Theorem 3 if \(E|X|^{2+\epsilon} < \sigma\), but violate the trimming conditions of Theorem 2.

In most instances, Theorems 2 or 3 will be of primary interest, providing methods for the construction of approximate confidence intervals based upon L-statistics (and one-to-one functions thereof). The bias of an L-statistic will often be small (but see Section 3 in this regard), and hence \(S_n\) will be of limited practical use for the purposes of estimation with reduced bias. However, the end goal of an analysis may be to estimate or construct approximate tests or confidence intervals for \(g(S_o)\). For estimation, \(g(S_n)\) may suffer from severe bias if \(g\) is highly non-linear near \(S_o\) (recall that
\[
E_F[g(S_n)] = g(S_0) + \frac{g''(S_0)}{2} \text{Var}_F(S_n)\]
Hence jackknifing \(g(S_n)\) would be of interest in such cases for reduction of bias. If \(g\) is non-monotone, confidence intervals for \(g(S_n)\) obtained by finding a confidence interval for \(S_0\) using \(\sqrt{n}(S_n - S_o)/s_p\) as an approximately standard normal pivotal quantity and taking the image of such an interval under \(g\) may well result in longer intervals than would be obtained by pivoting about \(g(S_n)\).
Thus, it seems to be of independent interest to study the behavior of $g(S_n)$ under jackknifing. The following theorem parallels Theorem 1, establishing that $g(S_n)$ and its jackknife have the same limiting distribution.

**Theorem 4:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a distribution $F$ with $E[|X|^{4/3}] < \infty$. Let $J'$ obey a Hölder condition with $\alpha > \frac{1}{2}$. If $g$ is a function with a bounded second derivative in a neighborhood of $S_0$, and $\tilde{g}(S_n) = \frac{n-1}{n} \sum_{i=1}^{n} g(S^{(i)}_n)$, the jackknife of $g(S_n)$, then

$$\sqrt{n}(g(S_n) - \tilde{g}(S_n)) \rightarrow 0 \text{ with probability one.}$$

**Proof:**

$$\sqrt{n}(g(S_n) - \tilde{g}(S_n)) = \sqrt{n} \frac{n-1}{n} \sum_{i=1}^{n} [g'(S_n)(S^{(i)}_n - S_n)$$

$$+ \frac{1}{2} g''(\xi_{in})(S^{(i)}_n - S_n)^2]$$
\[ \frac{n-1}{n} g'(S_n) \sqrt{n} (S_n - S_n) \]

\[ + \frac{n-1}{2n} \sqrt{n} \sum_{i=1}^{n} (g''(\xi_{i,n})(S_{n-1}^{(i)} - S_n)^2, \]

where \( \xi_{i,n} \) is between \( S_n \) and \( S_{n-1}^{(i)} \). To justify this expansion, we need to prove

\[ \sup_{1 \leq i \leq n} |S_{n-1}^{(i)} - S_n| \rightarrow 0 \text{ with probability one.} \]

Now,

\[ S_{n-1}^{(i)} - S_n = \sum_{j=1}^{i-1} \left[ \frac{1}{n-1} J_{j,n}^{(i)} - \frac{1}{n} J_{j,n}^{(i+1)} \right] X_{j,n} \]

\[ + \sum_{j=i+1}^{n} \left[ \frac{1}{n-1} J_{j,n}^{(i)} - \frac{1}{n} J_{j,n}^{(i+1)} \right] X_{j,n} \]

\[ - \frac{1}{n} J_{j,n}^{(i+1)} X_{i,n} \]

\[ = \sum_{j=1 \atop j \neq i}^{n} \left[ \frac{1}{n-1} \left( J_{j,n}^{(i)} + J_{j,n}^{(i+1)} \right) \frac{1}{n(n+1)} \right] + (J'(n_{i,n}) - J'(\frac{i}{n+1}) \frac{i}{n(n+1)}) \]

\[ - \frac{1}{n} J_{j,n}^{(i+1)} X_{j,n} - \frac{1}{n} J_{j,n}^{(i+1)} X_{i,n} \]
\[
\begin{align*}
&= \sum_{j=1}^{n} \left\{ \frac{1}{n(n-1)} \left[ J(\frac{1}{n+1}) + \frac{1}{n(n+1)(n-1)} J'(\frac{1}{n+1}) \right] \\
&\quad + \frac{1}{n(n+1)(n-1)} (J'(\eta_j) - J'(\frac{1}{n+1})) \right\} X_{j,n} \\
&\quad - \frac{1}{n} J\left( \frac{1}{n+1} \right) X_{1,n} 
\end{align*}
\]

So

\[
\sup_{1 \leq i \leq n} |S^{(i)} - S_n| \leq \sum_{j=1}^{n} \frac{n}{n^2} |X_{j,n}| + \frac{1}{n} \sup_{0 \leq u \leq 1} |J(u)|||X_{1,n}||X_{n,n}|)
\]

with \( M \) determined by the bounds on \( J \) and \( J' \). Finally, this expression converges to zero with probability one by the strong law of large numbers and the fact that \( \max(|X_{1,n}|, |X_{n,n}|)/n \to 0 \) with probability one if \( E|X| < \infty \).

The first term in (2.5) converges to zero with probability one by Theorem 1. The second term, denoted by \( R_n \), can be bounded as follows,

\[
|R_n| = \sum_{i=1}^{n} x_i (\xi_i)^{(i)} (S^{(i)} - S) \leq L \sum_{i=1}^{n} (S^{(i)} - S)^2 \]

where \( |g''(a)| \leq L \) for \( a \) in some neighborhood of \( S_0 \), for \( n \) sufficiently large.
\[ |R_n| \leq \frac{L(n-1)}{2/n} \sum_{i=1}^{n} \left( \sum_{j=1}^{n} \frac{|X_{i,j,n}|}{n^2} \right)^2 + \frac{|J\left(\frac{x}{n+1}\right)|X_{1,n}|^2}{n} \]

\[ \leq \frac{L(n-1)}{2/n} \left( \frac{M^2}{n^3} \left( \sum_{j=1}^{n} |X_{j,n}| \right)^2 \right) + \frac{1}{n^2} \sum_{i=1}^{n} \left( J\left(\frac{x}{n+1}\right)X_{i,n} \right)^2 \]

\[ 2M \sup_{0 \leq u \leq 1} |J(u)| + \frac{1}{n^3} \left( \sum_{j=1}^{n} |X_{j,n}| \right)^2 \rightarrow 0 \text{ with probability one} \]

by the moment condition on $F$. Hence, the theorem is established.

Naturally, the moment condition may be omitted if $J$ trims.

The analogous result on a consistent variance estimator follows.

**Theorem 5:** Let $X_1, \ldots, X_n$ be a random sample of size $n$ from a distribution $F$ obeying the conditions of Theorem 2. Let $J$ satisfy the conditions of Theorem 2 and $g$ have a bounded second
derivative in a neighborhood of $S_0$. If

$$s^2_{p(g)} = (n-1) \sum_{i=1}^{n} \left( g(S_{n-1}^{(i)}) - g(S_n) \right)^2,$$

(2.6)

then $s^2_{p(g)} \xrightarrow{P} (g'(S_0))^2 \sigma^2$, with $\sigma^2$ given by (2.1).

**Proof:** The proof proceeds by second-order Taylor expansion of the $g(S_{n-1}^{(i)})$ about $S_n$, and then follows the method of Theorem 2.

It should be noted that Theorem 2 (1) is (is not) a special case of Theorem 5 (4) with $g(\cdot)$ the identity, due to the identical (additional) conditions imposed on $F$ and $J$ in the latter theorem.

If the conditions for both Theorems 2 and 4 are satisfied, asymptotically pivotal quantities for the construction of confidence intervals for $g(S_0)$ include

$$Z_{1n} = \frac{\sqrt{n}(g(S_n) - g(S_0))}{\sigma_p(g)}$$

and

$$Z_{2n} = \frac{\sqrt{n}(g(S_n) - g(S_0))}{g'(S_n)s_p}.$$
If conditions for Theorems 1 and 3 hold, \( g(S_n) \) could be replaced by \( \hat{g}(S_n) \) and \( g'(S_n) \) by \( g'(\hat{S}_n) \) (or even \( \hat{g}'(S_n) \)) if \( g \) has a bounded third derivative in a neighborhood of \( S_0 \). The question naturally arises as to which choices would be best in moderate-size sample applications of the above results. The issue of what quantity to jackknife, that is, whether to use

\[
\frac{\sqrt{n}(S_n - S_0)}{s_p} \quad \text{or} \quad \frac{\sqrt{n}(g(S_n) - g(S_0))}{s_p(g)}
\]

has not been addressed in a systematic fashion in the literature. However, the common suggestions of Miller (1974), p.12) and Efron (1972, p.1920) that the function to be jackknifed should be variance stabilized are reasonable, for example \( \tanh^{-1} r \) does jackknife more satisfactorily than \( r \), the ordinary sample correlation coefficient. Using this advice, when estimating a function \( g \) of a location parameter \( S_g \), the location parameter estimate \( S_n \) should itself be jackknifed to produce a variance estimator. In other words, \( Z_{2n} \) is recommended.

3. TRIAL BY NUMBERS

While the above results give a large-sample justification for use of the jackknife method for the creation of approximate confidence intervals, they leave unaddressed questions
regarding appropriateness of the technique in small-sample situations. Accordingly, a modest Monte Carlo study is in order both i) to relate the large-sample theory to samples of a size likely to be encountered in practice, and ii) to explore the behavior of the jackknife for L-statistics whose score functions violate one or more of the regularity conditions of Theorems 1-4. All computations were performed on the AMDAHL 470 V/6 at Texas A&M University.

Only location parameter estimation is considered, that is \( J(u) > 0 \) for \( 0 < u < 1 \) and \( \int_0^1 J(u) du = 1 \). The distributions considered are i) \( N(5,1) \), a normal with mean 5 and variance one. ii) a logistic with mean 5 and scale parameter 1, with

\[
F(x) = \left[1 + e^{-(x-5)} \right]^{-1}, \quad -\infty < x < \infty,
\]

and iii) \( u(4,6) \), a uniform distribution on the interval \( (4,6) \). Five hundred random samples of sizes 5, 10, 20 and 40 from each of the above three distributions were examined. The score functions considered are
\[ J_1(u) = \begin{cases} 0 & 0 \leq u < .05 \\ 23.53 \ (u - .05) & .05 \leq u < .10 \\ 1.1765 & .10 \leq u < .90 \\ 23.53 \ (.95 - u) & .90 \leq u < .95 \\ 0 & .95 \leq u \leq 1 \end{cases} \]

a "smoothly" trimmed mean designed to obey the trimming conditions of the theorems but to fail to be differentiable at some points in \([0,1]\);

\[ J_2(u) = \begin{cases} 4u & 0 \leq u < .5 \\ 4(1 - u) & .5 \leq u \leq 1.0, \end{cases} \]

a "triangular" weight function neither trimming nor being everywhere differentiable; and

\[ J_3(u) = 6u(1 - u) \quad 0 \leq u \leq 1, \]

a score function meeting all differentiability requirements but failing to trim.

Unfortunately, even with the above three symmetric score functions which integrate to one and symmetric parents, biases can result due to the fact that

\[ \frac{1}{n} \prod_{i=1}^{n} \frac{1}{n + 1} \neq 1. \]

Table 1 gives values of \( \frac{1}{n} \prod_{i=1}^{n} \frac{1}{n + 1} \).
for \( n = 19, 20, 39, \) and 40 and \( K = 1, 2, 3. \) Note that a value of 1 corresponds to no bias.

\[(\text{TABLE 1 ABOUT HERE})\]

In fact, often such biases will die out at the rate of \( O(1/n^2) \)
(obtained by viewing \( \frac{1}{n} \sum_{i=1}^{n} J \left( \frac{i}{n+1} \right) \) as an application of the trapezoidal rule in approximating the integral \( \int_{0}^{1} J(u) \, du \) - if \( J'' \)
is bounded and continuous the error is \( O(1/n^2) \)). Hence, the ordinary jackknife would be of little or no use in dealing with these biases. Also, a practitioner using L-statistics would doubtlessly use the modified L-statistic

\[
S^*_n = \frac{\sum_{i=1}^{n} J \left( \frac{i}{n+1} \right) X_{i,n}}{\sum_{i=1}^{n} J \left( \frac{i}{n+1} \right)}
\]

to guarantee unbiasedness. All of Theorems 1-5 continue to hold for these modified \( S^* \) if they hold for \( S_n \), so long as

\[
\lim_{n \to \infty} n^{3/2} \left( 1 - \frac{1}{n} \sum_{i=1}^{n} J \left( \frac{i}{n+1} \right) \right) = 0.
\]

Based upon the 500 samples for each sample size and distribution combination the following are estimated:
1. Bias Factors for L-Statistics

<table>
<thead>
<tr>
<th>n</th>
<th>$J_1$</th>
<th>$J_2$</th>
<th>$J_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>1.053</td>
<td>1.053</td>
<td>1.050</td>
</tr>
<tr>
<td>20</td>
<td>1.048</td>
<td>1.048</td>
<td>1.048</td>
</tr>
<tr>
<td>39</td>
<td>1.026</td>
<td>1.026</td>
<td>1.026</td>
</tr>
<tr>
<td>40</td>
<td>1.024</td>
<td>1.024</td>
<td>1.024</td>
</tr>
</tbody>
</table>
a) Variance of $S_n^*$

b) Variance of $S_n^*$, the jackknife of $S_n$

c) Mean of $s_p^2/n$, the jackknife variance estimator

d) Mean of $s_p/n$, the jackknife standard deviation estimator

e-f) Percent coverages of approximate $100(1 - \alpha)$% confidence intervals for $\int Q(u)J(u)du$, obtained as

$$S_n^* + t \frac{s_p}{\sqrt{n}},$$

where $t_{1-\alpha/2, n-1}$ is the $100(1 - \alpha/2)$ percentile point of a t-distribution with $n - 1$ degrees of freedom, for

e) $\alpha = .10$, and f) $\alpha = .05$.

and g-h) Percent coverages for confidence intervals identical to these above, but centered on $S_n^*$, for

g) $\alpha = .10$, and h) $\alpha = .05$.

Tables 2-5 present the results. It may be seen, even for samples of size 5, that the approximate confidence intervals maintain actual confidence coefficients close to the nominal 90 and 95 percent levels. Typically, the worst cases of undercoverage seem to occur for, curiously enough, the uniform parent. Intervals centered on $S_n^*$ seem slightly better in this regard than those centered on $S_n^*$, although the
### 2. Results for $n = 5$

<table>
<thead>
<tr>
<th>Normal Parent Score Function</th>
<th>Logistic Parent Score Function</th>
<th>Uniform Parent Score Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_1$</td>
<td>$J_2$</td>
<td>$J_3$</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>a) Variance of $S_n^*$</td>
<td>.2033</td>
<td>.2089</td>
</tr>
<tr>
<td>b) Variance of $S_n^*$</td>
<td>.2033</td>
<td>.2375</td>
</tr>
<tr>
<td>c) Mean of $s^2/n_p$</td>
<td>.2069</td>
<td>.2170</td>
</tr>
<tr>
<td>d) Mean of $s_p/\sqrt{n}$</td>
<td>.4271</td>
<td>.4319</td>
</tr>
<tr>
<td>e) Coverage for 90% CI centered on $S_n^*$</td>
<td>.9020</td>
<td>.8840</td>
</tr>
<tr>
<td>f) Coverage for 95% CI centered on $S_n^*$</td>
<td>.9440</td>
<td>.9400</td>
</tr>
<tr>
<td>g) Coverage for 90% CI centered on $S_n^*$</td>
<td>.9020</td>
<td>.8840</td>
</tr>
<tr>
<td>h) Coverage for 95% CI centered on $S_n^*$</td>
<td>.9440</td>
<td>.9320</td>
</tr>
</tbody>
</table>

** Denotes estimated coverage is at least two standard errors different from the nominal.
3. Results for Sample Size $n = 10$

<table>
<thead>
<tr>
<th></th>
<th>Normal Parent Score Function</th>
<th>Logistic Parent Score Function</th>
<th>Uniform Parent Score Function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$J_1$</td>
<td>$J_2$</td>
<td>$J_3$</td>
</tr>
<tr>
<td>a) Variance of $\bar{S}_n$</td>
<td>.0927</td>
<td>.1001</td>
<td>.0970</td>
</tr>
<tr>
<td>b) Variance of $\bar{S}_n$</td>
<td>.1107</td>
<td>.1058</td>
<td>.1029</td>
</tr>
<tr>
<td>c) Mean of $s^2/n_p$</td>
<td>.1032</td>
<td>.1094</td>
<td>.1054</td>
</tr>
<tr>
<td>d) Mean of $s/\sqrt{n}$</td>
<td>.3119</td>
<td>.3176</td>
<td>.3138</td>
</tr>
<tr>
<td>e) Coverage for 90% CI centered on $\bar{S}_n$</td>
<td>.9240</td>
<td>.8940</td>
<td>.9080</td>
</tr>
<tr>
<td>f) Coverage for 95% CI centered on $\bar{S}_n$</td>
<td>.9580</td>
<td>.9480</td>
<td>.9540</td>
</tr>
<tr>
<td>g) Coverage for 90% CI centered on $\bar{S}_n$</td>
<td>.9020</td>
<td>.8880</td>
<td>.9000</td>
</tr>
<tr>
<td>h) Coverage for 95% CI centered on $\bar{S}_n$</td>
<td>.9460</td>
<td>.9420</td>
<td>.9460</td>
</tr>
</tbody>
</table>

** Denotes estimated coverage is at least two standard errors different from the nominal.
4. Results for Sample Size $n = 20$

<table>
<thead>
<tr>
<th>Normal Parent</th>
<th>Logistic Parent</th>
<th>Uniform Parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score Function</td>
<td>Score Function</td>
<td>Score Function</td>
</tr>
<tr>
<td>$J_1$</td>
<td>$J_2$</td>
<td>$J_3$</td>
</tr>
<tr>
<td>a) Variance of $S_n^*$</td>
<td>0.0487</td>
<td>0.0506</td>
</tr>
<tr>
<td>b) Variance of $\tilde{S}_n^*$</td>
<td>0.0500</td>
<td>0.0520</td>
</tr>
<tr>
<td>c) Mean of $\frac{s^2}{n}$</td>
<td>0.0527</td>
<td>0.0546</td>
</tr>
<tr>
<td>d) Mean of $\frac{s}{\sqrt{n}}$</td>
<td>0.2258</td>
<td>0.2290</td>
</tr>
<tr>
<td>e) Coverage for 90% CI centered on $S_n^*$</td>
<td>0.9100</td>
<td>0.9040</td>
</tr>
<tr>
<td>f) Coverage for 95% CI centered on $S_n^*$</td>
<td>0.9560</td>
<td>0.9520</td>
</tr>
<tr>
<td>g) Coverage for 90% CI centered on $\tilde{S}_n^*$</td>
<td>0.9060</td>
<td>0.8980</td>
</tr>
<tr>
<td>h) Coverage for 95% CI centered on $\tilde{S}_n^*$</td>
<td>0.9540</td>
<td>0.9540</td>
</tr>
</tbody>
</table>

** Denotes estimated coverage is at least two standard errors different from the nominal.
5. Results for Sample Size $n = 40$

<table>
<thead>
<tr>
<th>Normal Parent Score Function</th>
<th>Logistic Parent Score Function</th>
<th>Uniform Parent Score Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_1$</td>
<td>$J_2$</td>
<td>$J_3$</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>a) Variance of $S_n$ *</td>
<td>.0274</td>
<td>.0283</td>
</tr>
<tr>
<td>b) Variance of $\tilde{S}_n$ *</td>
<td>.0277</td>
<td>.0287</td>
</tr>
<tr>
<td>c) Mean of $s^2/n$</td>
<td>.0262</td>
<td>.0276</td>
</tr>
<tr>
<td>d) Mean of $s/\sqrt{n}$</td>
<td>.1605</td>
<td>.1643</td>
</tr>
<tr>
<td>e) Coverage for 90% CI centered on $S_n$ *</td>
<td>.8880</td>
<td>.8820</td>
</tr>
<tr>
<td>f) Coverage for 95% CI centered on $S_n$ *</td>
<td>.9420</td>
<td>.9460</td>
</tr>
<tr>
<td>g) Coverage for 90% CI centered on $\tilde{S}_n$ *</td>
<td>.8840</td>
<td>.8820</td>
</tr>
<tr>
<td>h) Coverage for 95% CI centered on $\tilde{S}_n$ *</td>
<td>.9420</td>
<td>.9460</td>
</tr>
</tbody>
</table>

** Denotes estimated coverage is at least two standard errors different from the nominal.
observed differences are of the same order as the standard errors of the empirical confidence coefficients. This may be related to the typically slightly larger standard error of $S_n^*$. Consistent with the result of Efron and Stein (1979), the jackknife variance estimator appears to be typically positively biased as an estimator of the variance of $S_n^*$. For 28 of the 36 combinations of sample size, score, and parent population, the estimated mean of the jackknife variance estimator was greater than or equal to the estimated variance of $S_n^*$. Out of the 36 combinations, $S_n^*$ had a larger estimated variance than $S_n^*$ 27 times, with 3 ties (to 4 decimal places). However, the increase was typically small relative to the size of the estimates themselves. Interestingly, the estimated bias of $s_p/\sqrt{n}$ as an estimate of the standard error of $S_n^*$ is small, perhaps indicating that while the variance estimate may suffer from a positive bias, the standard error estimate is relatively better off. (Parenthetically, if $s_p/\sqrt{n}$ were exactly unbiased for $(\text{Var}(S_n^*))^{1/2}$, then $s_p^2/n$ would have a bias of order $1/n^2$, assuming $\text{Var}(S_n^*) = O(1/n)$.)

4. APPLICATIONS

The large-sample results of Section 2, bolstered by the favorable Monte Carlo results in Section 3, provide a metho-
dology for robust inference in linear models, in particular for completely randomized designs with multiple observations per treatment. We consider the model

\[ Y_{ij} = \mu + a_i + \epsilon_{ij}, \]

where \( a_i \) represents the "effect" of the \( i \)th treatment. Note that we do not rule out a factorial structure for the \( t \) treatments. If, instead of the usual assumptions that the \( \epsilon_{ij} \) are normally and independently distributed with mean 0 and common variance, we merely assume that the \( \epsilon_{ij} \) are independently and identically distributed, symmetrically about 0, we can pursue an L-statistic approach to analysis of variance (the symmetry assumption is merely convenient - not necessary). If we assume \( \min(n_1, n_2, \ldots, n_t) = \) and \( \min(n_1, n_2, \ldots, n_t)/t \) \( \sum n_i > a > 0 \), then for a symmetric score function \( J(\cdot) \) meeting the conditions of the appropriate theorems in Section 2, we denote

\[ s^i = \frac{1}{n_1} \sum_{j=1}^{n_1} \frac{1}{(n + 1)} X_j, n_1 \]
and \( s_i^2 \) the corresponding variance estimate, where \( X_{j,n_i} \) is the 
\( j \)th order statistic among those receiving treatment \( i \). We then 
pool our variance estimates by

\[
\hat{\sigma}^2 = \frac{t}{\sum \frac{(n_i - 1)s_i^2}{\sum (n_i - 1)}}
\]

mimicking ordinary analysis of variance. Then, an approximate 

\[
\text{test for a contrast } \sum C_i \alpha_i \text{ is provided by rejecting } \sum C_i \alpha_i = 0 \text{ if and only if}
\]

\[
\left( \sum \frac{C_i \alpha_i}{\sum s_i^2} \right) > \frac{Z_{1-\alpha/2}^2 \hat{\sigma}^2 \sum \frac{C_i^2}{\sum s_i^2}}{n}
\]

with \( Z_{1-\alpha/2} \) the \( 1 - \frac{\alpha}{2} \) quantile for the standard normal. An 
obvious modification in the critical point would permit 
Scheffé-type procedures. Similarly, "robust" multiple 
comparisons could be done. Robustness of these procedures would, 
of course, depend upon the robustness and convergence rates 
of the associated L-statistic for location – a well-studied 
topic. Note that the "sums of squares" for this type of analysis
can be simply computed by separately computing $S^i$ and $s_i^2$ for each treatment, and then inputting these into any standard analysis of variance package which will accept treatment means and variances as input.

5. SUMMARY

The ordinary jackknife is a computationally simple means for the construction of large-sample confidence intervals for functionals of the form $S_0 = \int_0^1 Q(u)J(u)du$. Simulation results indicate that the technique is effective for small samples.
REFERENCES


The behavior of linear combinations of order statistics (L-statistics) under jackknifing is discussed. The asymptotic behavior of a jackknifed L-statistic is produced, and the pseudo-value based variance estimate is shown to be consistent under moderate smoothness and a trimming condition of the weight function. The results extend easily to functions of L-statistics and thus answer a question posed in Miller (1974). Monte Carlo results support small-sample applicability of the large-sample results for the construction of approximate confidence intervals.