Asymptotic Accuracy of the Jackknife Variance Estimator for Certain Smooth Statistics

Alex D. Gottlieb

Abstract

We show that that the jackknife variance estimator $v_{jack}$ and the the infinitesimal jackknife variance estimator are asymptotically equivalent if the functional of interest is a smooth function of the mean or a smooth trimmed L-statistic. We calculate the asymptotic variance of $v_{jack}$ for these functionals.

1 Introduction

Let $p$ be a probability measure on a sample space $\mathcal{X}$. Given $n$ samples from $\mathcal{X}$, sampled independently under the probability law $p$, one desires to estimate the value $T(p)$ of some real functional $T$ on the space $\mathcal{P}(\mathcal{X})$ of all probability measures on $\mathcal{X}$. Denote by $\epsilon_n$ the map that converts $n$ data points $x_1, x_2, \ldots, x_n$ into the empirical measure

$$
\epsilon_n(x_1, x_2, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} \delta(x_i)
$$

where $\delta(x_i)$ denotes a point-mass at $x_i$. The plug-in estimate of $T(p)$ given the data $x = (x_1, \ldots, x_n)$ is

$$
T_n = T(\epsilon_n(x)).
$$

Suppose $T_n$ is an asymptotically normal estimator of $T(p)$, so that the distribution of $n^{1/2}(T_n - T(p))$ tends to $\mathcal{N}(0, \sigma^2)$. The jackknife is a computational technique for estimating $\sigma^2$: one transforms the $n$ original data points into $n$ pseudovalues and computes the sample variance of those pseudovalues.

Given the data $x = x_1, x_2, \ldots, x_n$, the jackknife pseudovalues are

$$
Q_{ni} = nT_n(\epsilon_n) - (n-1)T(\epsilon_{ni}) \quad i = 1, 2, \ldots, n
$$

with $\epsilon_n$ as in (1) and

$$
\epsilon_{ni} = \frac{1}{n-1} \sum_{j \neq i} \delta(x_j).
$$
The jackknife variance estimator is

\[ v_{\text{jack}}(x_1, x_2, \ldots, x_n) = \frac{1}{n-1} \sum_{i=1}^{n} (Q_{ni} - \overline{Q}_n)^2 \]  

(4)

where \( \overline{Q}_n = \frac{1}{n} \sum Q_{nj} \). The variance estimator \( v_{\text{jack}} \) is said to be consistent if \( v_{\text{jack}} \rightarrow \sigma^2 \) almost surely as \( n \rightarrow \infty \). Sufficient conditions for the consistency of \( v_{\text{jack}} \) are given in terms of the functional differentiability of \( T \). An early result of this kind states that \( v_{\text{jack}} \) is consistent if \( T \) is strongly Fréchet differentiable [Parr85], and it is now known that \( v_{\text{jack}} \) is consistent even if \( T \) is only continuously Gâteaux differentiable as in Definition [I] below [ST95].

A functional derivative of \( T \) at \( p \), denoted \( \partial T_p \), is a linear functional that best approximates the behavior of \( T \) near \( p \) in some sense. For instance, a functional \( T \) on the space of bounded signed measures \( \mathcal{M}(\mathcal{X}) \) is Gâteaux differentiable at \( p \) if there exists a continuous linear functional \( \partial T_p \) on \( \mathcal{M}(\mathcal{X}) \) such that

\[ \lim_{t \to 0} t^{-1} \left| t^{-1} (T(p + tm) - T(p)) - \partial T_p(m) \right| = 0 \]

for all \( m \in \mathcal{M}(\mathcal{X}) \). More relevant to mathematical statistics is the concept of Hadamard differentiability, for the fluctuations of \( T(\epsilon_n) \) about \( T(p) \) are asymptotically normal if \( T \) is Hadamard differentiable at \( p \). A functional \( T : \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R} \) is Hadamard differentiable at \( p \) if there exists a continuous linear functional \( \partial T_p \) on \( \mathcal{M}(\mathbb{R}) \) such that

\[ \lim_{t \to 0} t^{-1} \left| t^{-1} (T(p + tm_t) - T(p)) - \partial T_p(m) \right| = 0 \]

whenever \( \{m_t\}_{t \in \mathbb{R}} \) is such that \( \lim_{t \to 0} m_t = m \) and \( m_t(\mathbb{R}) = 0 \) for all \( t \), the topology on \( \mathcal{M}(\mathbb{R}) \) being the one induced by the norm \( ||m|| = \sup_{t \in \mathbb{R}} \{|m((\mp \infty, t])|\} \). If \( T \) is Hadamard differentiable at \( p \), the variance of \( n^{1/2}T(\epsilon_n) \) tends to

\[ \sigma^2 = \mathbb{E}_p \phi_p^2 \]

(5)

as \( n \rightarrow \infty \), where \( \phi_p(x) \) is the influence function

\[ \phi_p(x) = \partial T_p(\delta(x) - p) \]

(6)

(this can be shown via the Delta method [vdW98] using Donsker’s theorem).

If \( T \) is smooth enough then \( n^{1/2}(v_{\text{jack}} - \sigma^2) \) is also asymptotically normal. In this note we calculate the asymptotic variance of \( v_{\text{jack}} \) (i.e., the limit as \( n \rightarrow \infty \) of the variance of \( n^{1/2}v_{\text{jack}} \)) for two very well behaved functionals \( T \): smooth functions of the mean \( T(p) = g(\overline{p}) \) and smooth trimmed L-functionals. In these cases, the asymptotic variance of \( v_{\text{jack}} \) equals that of \( \mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 \), the estimator of \( \sigma^2 \) obtained from (5) by substituting the empirical measure for \( p \). This is known as the infinitesimal jackknife estimator [ST95, p 48]. We are tempted to conjecture that \( v_{\text{jack}} \) and the infinitesimal jackknife variance estimator are asymptotically equivalent for sufficiently regular functionals \( T \), but we have no general results in this direction.
The literature does not address the accuracy of \( v_{\text{jack}} \) adequately. In fact, [ST95, Section 2.2.3] gets it wrong, conjecturing that the asymptotic variance of \( v_{\text{jack}} \) should equal \( \text{Var} \phi_p^2 \) for sufficiently regular functionals! However, Theorem 2 of [Ber84] does contain a general formula for the variance of \( v_{\text{jack}} \) which is valid when the functional \( T \) has a kind of second-order functional derivative. The theorem there applies to the trimmed L-functionals we discuss in Section 4, and to many other functionals besides, but it is hampered by the hypothesis that \( p \) have bounded support. We recommend Theorem 2 of [Ber84] for its generality and its revelation of the role of second-order differentiability, but our particular results cannot be derived from it directly.

The text [ST95, p 43] purports to prove that the asymptotic variance of \( n^{1/2} (v_{\text{jack}} - \sigma^2) \) equals \( \text{Var} \phi_p^2 \) when \( T \) is of the form (7), but there is a mistake there. We paraphrase the following definition from [ST95, p 43]: For probability measures \( p \) and \( q \) on the line, let \( \rho(p,q) \) denote the \( L^\infty \) distance between the cdf’s of \( p \) and \( q \). A functional \( T : \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R} \) is \( \rho \)-Lipschitz differentiable at \( q \) if

\[
T(p_k) - T(q_k) - \partial T_q(p_k - q_k) = O(\rho(p_k, q_k)^2)
\]

for all sequences \( \{p_k\} \) and \( \{q_k\} \) such that \( \rho(p_k, q) \) and \( \rho(q_k, q) \) converge to 0. Assuming that \( \text{Var} \phi_p^2 < \infty \) and \( T \) is \( \rho \)-Lipschitz differentiable, the authors prove (correctly) that \( n^{1/2} (v_{\text{jack}} - \sigma^2) \) is asymptotically normal with variance \( \text{Var} \phi_p^2 \). They go on to assert that smooth trimmed L-functionals are \( \rho \)-Lipschitz differentiable, but this is false (it is not difficult to construct counterexamples).

A close look at the definition of \( \rho \)-Lipschitz differentiability leads one to wonder whether there are any functionals (besides trivial, linear ones) that satisfy the definition. The problem is that \( q \) appears on the left hand side of (7) but not on the right; it is easy to imagine \( p_k \) and \( q_k \) that are close to one another in the \( \rho \) metric, yet far enough from \( q \) that \( \partial T_q(p_k - q_k) \) badly approximates \( T(p_k) - T(q_k) \). Replacing \( \partial T_q(p_k - q_k) \) by \( \partial T_{q_k}(p_k - q_k) \) in the left-hand-side of (7) might result in a more useful characteristic of smoothness for a functional \( T \). Indeed, it was this observation that guided our calculations in Sections 3 and 4.

In this note we work with modified pseudovalues

\[
Q'_n(x_1, x_2, \ldots, x_n) = (n - 1) [T(\epsilon_n) - T(\epsilon_{ni})].
\]

Substituting \( Q'_n \) for \( Q_{ni} \) and \( \overline{x}_n = \frac{1}{n} \sum Q'_{nj} \) for \( \overline{Q}_n = \frac{1}{n} \sum Q_{nj} \) in (7) does not change the value of \( v_{\text{jack}} \), so one may compute \( v_{\text{jack}} \) by the same formula using the \( Q'^{'}_n \). Using the modified pseudovalues \( Q'_{ni} \) makes it easier to take advantage of the magic formula \((n - 1)(\epsilon_n - \epsilon_{ni}) = \delta_{x_i} - \epsilon_n \).

2 Using pseudovalues to estimate the variance of \( \phi_p^2 \)

One aim of this letter is to emphasize that \( \text{Var} \phi_p^2 \) is typically not the asymptotic variance of \( n^{1/2} (v_{\text{jack}} - \sigma^2) \), contrary to the assertion of [ST95, p 42]. However, should one desire an estimate of \( \text{Var} \phi_p^2 \) for some reason, the pseudovalues can be used to this end. Once one has already computed \( v_{\text{jack}} \), the variance of \( \phi_p^2 \) is easy to estimate with very little additional labor: just compute the sample variance of the squares of the pseudovalues. We prove this, assuming that the functional
$T$ is \textit{continuously Gâteaux differentiable} and $\phi_p$ is bounded (trimmed L-functionals satisfy these requirements, for instance). This section is an interlude whose results will not be invoked in Sections 3 and 4, the main part of this note.

Continuous Gâteaux differentiability is introduced in [ST95] as a sufficient condition for the strong consistency of the jackknife variance estimator.

\textbf{Definition 1.} A functional $T$ is \textbf{continuously Gâteaux differentiable} at $p$ if it has Gâteaux derivative $\partial T_p$ at $p$ and if

$$\lim_{k \to \infty} \sup_{x \in \mathbb{R}} \left\{ \frac{\left| T(p_k + t_k(\delta(x) - p_k)) - T(p_k) - \partial T_p(\delta(x) - p_k) \right|}{t_k} \right\} = 0$$

for any sequence of probability measures $p_k$ whose cdf’s converge uniformly to that of $p$ and for any sequence of real numbers $t_k$ that converges to 0.

The proof in [ST95] that continuous Gâteaux differentiability implies strong consistency of the jackknife [ST95, Theorem 2.3] also serves to prove the following proposition.

\textbf{Proposition 1.} Suppose that $T : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ is continuously Gâteaux differentiable at $p$, with influence function $\phi_p(x) = \partial T_p(\delta(x) - p)$ satisfying

$$\int |\phi_p(x)| p(dx) < \infty \quad \int \phi_p(x)p(dx) = 0.$$  

If the data $X_1, X_2, X_3, \ldots$ are iid $p$ then the empirical measures of the jackknife pseudovalues obtained from the data converge almost surely to $p \circ \phi_p^{-1}$:

$$\epsilon_n(Q'_{n,1}, Q'_{n,2}, \ldots, Q'_{n,n}) \to p \circ \phi_p^{-1} \quad \text{a.s.}$$

\textbf{Proof:} Omitted, but cf. the proof of Theorem 2.3 in [ST95]. \hfill \Box

Now, suppose that $T : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ has a bounded influence function and satisfies the conditions of Proposition 1. Given iid $p$ data $X_1, X_2, \ldots, X_n$ compute the jackknife pseudovalues

$$Q'_{n,1}, Q'_{n,2}, \ldots, Q'_{n,n}$$

and the jackknife estimate $v_{jack}$ based on these pseudovalues. Set

$$\text{sq}(x) = \min\{x^2, \|\phi_p\|_{\infty}^2\},$$

and

$$\tau^2 = \frac{1}{n} \sum_{j=1}^{n} \left( \text{sq}(Q'_{n,j}) - \frac{1}{n} \sum_{j=1}^{n} \text{sq}(Q'_{n,j}) \right)^2.$$  

By Proposition 1, the empirical measure of the jackknife pseudovalues converges almost surely in $\mathcal{P}(\mathbb{R})$ to $p \circ \phi_p^{-1}$. It follows that $\tau^2 \to \text{Var} \phi_p^2$ almost surely.
One may also estimate $\text{Var } \phi_p^2$ by applying the bootstrap to the pseudovalues themselves, just as if the pseudovalues were actually iid. To bootstrap, sample $n$ times with replacement from the empirical measure of the pseudovalues $Q_{n,1}, \ldots, Q_{n,n}$, to produce a bootstrap sample

$$Q_{n,1}^*, Q_{n,2}^*, \ldots, Q_{n,n}^*$$

and compute

$$\frac{1}{n^{1/2}} \sum_{i=1}^n (\text{sq}(Q_{n,i}^*) - \text{sq}(Q_{n,i}')).$$

(10)

Given a triangular array of pseudovalues $Q_{n,i,j}$ having the property that $\epsilon_n(Q_{n,1}', \ldots, Q_{n,n}') \rightarrow p \circ \phi_p^{-1}$ as $n \rightarrow \infty$, one may define $Y_{n,i} = \text{sq}(Q_{n,i}') - \frac{1}{n} \sum_j \text{sq}(Q_{n,j}')$ and apply the Lindeberg-Feller Central Limit Theorem to the array $\{Y_{n,i}\}_{n,i}$ to show that (10) converges in distribution to $\mathcal{N}(0, \text{Var } \phi_p^2)$. But $\epsilon_n(Q_{n,1}, \ldots, Q_{n,n})$ almost surely converges to $p \circ \phi_p^{-1}$ by Proposition 1. It follows that, almost surely, (10) converges in distribution to $\mathcal{N}(0, \text{Var } \phi_p^2)$.

3 Functions of the mean

When $q$ is a measure, we denote $\int xq(dx)$ by $\overline{q}$ if the integral is defined. Let $g \in C^1(\mathbb{R})$ and let

$$T(m) = g(\overline{m})$$

be defined for all finite signed measures $m$ with finite first moment. The functional derivative at $m$ of $T$, evaluated at $q$, is $\partial T_n(q) = g'(\overline{m})\overline{q}$; the influence function (11) is $\phi_m(x) = g'(\overline{m}) (x - \overline{m})$. Suppose that $x_1, x_2, \ldots$ are iid $p$, and $p$ has a finite second moment. Let $T_n$ denote the plug-in estimator defined in (11). Then the asymptotic variance of $n^{1/2}(T_n - T(p))$ is

$$\sigma^2 = g'(\overline{p})^2 \left\{ \int x^2 p(dx) - \overline{p}^2 \right\}.$$

(11)

Let $v_{\text{jack}}$ denote the jackknife variance estimator for $\sigma^2$.

Proposition 2. If $g'$ is (globally) Hölder continuous of order $h > 1/2$ and $p$ has a finite moment of order $2(1+h)$ then $n^{1/2}(v_{\text{jack}} - \sigma^2)$ and $n^{1/2}(\mathbb{E}_n \phi_{\epsilon_n}^2 - \sigma^2)$ have the same limit in distribution, if any.

Proof: Set $\Delta_{ni} = (Q_{ni}' - \overline{Q}_n') - \phi_{\epsilon_n}(x_i)$ and note that

$$v_{\text{jack}} = \frac{1}{n-1} \sum_{i=1}^n (Q_{ni}' - \overline{Q}_n')^2 = \frac{n}{n-1} \left\{ \mathbb{E}_n \phi_{\epsilon_n}^2 + \frac{1}{n} \sum_{i=1}^n \phi_{\epsilon_n}(x_i) \Delta_{ni} + \frac{1}{n} \sum_{i=1}^n \Delta_{ni}^2 \right\},$$

whence

$$n^{1/2} (v_{\text{jack}} - \sigma^2) = n^{1/2} (\mathbb{E}_n \phi_{\epsilon_n}^2 - \sigma^2) + \frac{n^{3/2}}{n-1} \mathbb{E}_n \phi_{\epsilon_n}^2 + \frac{n^{3/2}}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^n \phi_{\epsilon_n}(x_i) \Delta_{ni} + \frac{1}{n} \sum_{i=1}^n \Delta_{ni}^2 \right\}.$$
To prove that $n^{1/2}(v_{\text{jack}} - \sigma^2)$ and $n^{1/2} \left( \mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2 \right)$ have the same limit in distribution (if any) it suffices to show that

$$
\frac{n^{1/2}}{n-1} \mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 + \frac{n^{3/2}}{n-1} \left( \frac{1}{n} \sum_{i=1}^{n} \phi_{\epsilon_n}(x_i) \Delta_{ni} + \frac{1}{n} \sum_{i=1}^{n} \Delta_{ni}^2 \right)
$$

(12)

converges almost surely to 0.

Recall the notation $\epsilon_n$ and $\epsilon_{ni}$ of (1) and (3). The first term in (12) converges almost surely to 0 since

$$
\mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 = \frac{1}{n} \sum_{i=1}^{n} \phi_{\epsilon_n}(x_i) = \frac{1}{n} \sum_{i=1}^{n} g'((\epsilon_n)^2 (x_i - \bar{\epsilon}_n)^2
$$

converges almost surely to $\sigma^2$.

To show that the other terms tend to zero we need a bound on $\Delta_{ni}$. Since $g$ is differentiable, $g(\bar{\epsilon}_{nj}) - g(\bar{\epsilon}_{ni}) = g'( \eta_{ji} ) (\bar{\epsilon}_{nj} - \bar{\epsilon}_{ni})$ for some $\eta_{ji}$ between $\bar{\epsilon}_{ni}$ and $\bar{\epsilon}_{nj}$, so that

$$
Q'_{ni} - \bar{Q}_n = \frac{n-1}{n} \sum_{j=1}^{n} \left( g(\bar{\epsilon}_{nj}) - g(\bar{\epsilon}_{ni}) \right) = \frac{n-1}{n} \sum_{j=1}^{n} g'( \eta_{ji} ) (\bar{\epsilon}_{nj} - \bar{\epsilon}_{ni}).
$$

Therefore, since $\phi_{\epsilon_n}(x_i) = g'(\epsilon_n)(x_i - \bar{\epsilon}_n) = \frac{1}{n} \sum_j g'(\epsilon_n)(x_i - x_j)$,

$$
\Delta_{ni} = (Q'_{ni} - \bar{Q}_n) - \phi_{\epsilon_n}(x_i) = \frac{n-1}{n} \sum_{j=1}^{n} g'( \eta_{ji} ) (\bar{\epsilon}_{nj} - \bar{\epsilon}_{ni}) - \frac{1}{n} \sum_{j=1}^{n} g'(\epsilon_n)(x_i - x_j)
$$

$$
= \frac{1}{n} \sum_{j=1}^{n} \left( g'( \eta_{ji} ) - g'(\epsilon_n) \right) (x_i - x_j).
$$

But $g'$ is Hölder continuous of order $h$ and $|\eta_{ji} - \bar{\epsilon}_n| < \max\{|\bar{\epsilon}_{nj} - \epsilon_n|, |\epsilon_n - \bar{\epsilon}_n|\}$, so

$$
|g'( \eta_{ji} ) - g'(\epsilon_n)| \leq C(|\bar{\epsilon}_{nj} - \epsilon_n|^h + |\epsilon_n - \bar{\epsilon}_n|^h) \leq C(n-1)^{-h}(|\epsilon_n - x_j|^h + |\epsilon_n - x_i|^h),
$$

where $C$ is a global Hölder constant for $g'$. It follows that

$$
|\Delta_{ni}| = C(n-1)^{-h} \frac{1}{n} \sum_{j=1}^{n} (|\epsilon_n - x_j|^h + |\epsilon_n - x_i|^h)(|\epsilon_n - x_j| + |\epsilon_n - x_i|).
$$

With this bound on $\Delta_{ni}$, and assuming that $p$ has a finite moment of order $2(1 + h)$, it may be shown that

$$
\frac{1}{n} \sum_{i=1}^{n} \Delta_{ni}^2 = O_s(n^{-2h}),
$$
and then, by the Cauchy-Schwartz inequality, that

$$\left| \frac{1}{n} \sum_{i=1}^{n} \phi_{\epsilon_n}(x_i) \Delta_{ni} \right| = O_s(n^{-h}).$$

The preceding estimates and the assumption that $h > 1/2$ imply that the last two terms in (12) converge to almost surely to 0. Thus, $n^{1/2} (v_{\text{jack}} - \sigma^2)$ and $n^{1/2} (\mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2)$ have the same limit in distribution, if any.

If we strengthen the smoothness assumption on $g$ and the moment assumption on $p$ then we can calculate the limit in distribution of $n^{1/2} (\mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2)$. Suppose that $g''$ is bounded (so that $g'$ is globally Lipschitz) and Hölder continuous of order $r > 0$, and suppose that $p$ has a finite fourth moment. Then

$$\phi_{\epsilon_n}(x_i) = g'(\epsilon_n) (x_i - \epsilon_n) = \left[ g'(\overline{\epsilon}) + g''(\overline{\epsilon}) (\epsilon_n - \overline{\epsilon}) + O_s(n^{-(r+1)/2}) \right] (x_i - \epsilon_n),$$

so that

$$\mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 = \frac{1}{n} \sum_{i=1}^{n} \phi_{\epsilon_n}^2(x_i) = \left[ g'(\overline{\epsilon}) + g''(\overline{\epsilon}) (\epsilon_n - \overline{\epsilon}) \right]^2 \frac{1}{n} \sum_{i=1}^{n} (x_i - \epsilon_n)^2 + O_s(n^{-(r+1)/2})$$

$$= \left[ g'(\overline{\epsilon})^2 + 2g'(\overline{\epsilon}) g''(\overline{\epsilon}) (\epsilon_n - \overline{\epsilon}) \right] \frac{1}{n} \sum_{i=1}^{n} (x_i - \epsilon_n)^2 + O_s(n^{-(r+1)/2}).$$

From formula (11) for $\sigma^2$ we see that

$$n^{1/2} (\mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2) = g'(\overline{\epsilon})^2 n^{1/2} \left( \frac{1}{n} \sum_{i=1}^{n} (x_i - \epsilon_n)^2 - \left\{ \int x^2 p(dx) - \overline{\epsilon}^2 \right\} \right)$$

$$+ 2g'(\overline{\epsilon}) g''(\overline{\epsilon}) n^{1/2} (\epsilon_n - \overline{\epsilon}) \frac{1}{n} \sum_{i=1}^{n} (x_i - \epsilon_n)^2 + O_s \left( n^{-r/2} \right). \quad (13)$$

Set $Z_n = n^{1/2} (\epsilon_n - \overline{\epsilon})$ and

$$Y_n = n^{1/2} \left( \frac{1}{n} \sum_{i=1}^{n} (x_i - \epsilon_n)^2 - \left\{ \int x^2 p(dx) - \overline{\epsilon}^2 \right\} \right).$$

Since $p$ has a finite fourth moment, the random vector $(Y_n, Z_n)$ has a Gaussian limit by the Central Limit Theorem. Equation (13) shows that $n^{1/2} (\mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2)$ is asymptotically normal with variance $(a, b) \Gamma(a, b)^{tr}$, where $(a, b) = (g'(\overline{\epsilon})^2, 2g'(\overline{\epsilon}) g''(\overline{\epsilon}))$ and $\Gamma$ denotes the asymptotic covariance matrix for $(Y_n, Z_n)$.

In view of Proposition 2, we find that if $g''$ is bounded and Hölder continuous of order $r > 0$, and if $p$ has a finite fourth moment, then the asymptotic variance of $n^{1/2} (v_{\text{jack}} - \sigma^2)$ equals $(a, b) \Gamma(a, b)^{tr}$. In contrast, under the same conditions on $p$ and $g$ it may be shown that $\text{Var} \phi_p^2 = a^2 \Gamma_{1,1}$. 

7
4 Trimmed L-statistics

Suppose that \( \ell : (0, 1) \to \mathbb{R} \) is supported on \([\alpha, 1 - \alpha]\) for some \(0 < \alpha < 1/2\), and let

\[
L(p) = \int_0^1 P^{-1}(s) \ell(s) ds.
\]

(14)

Here \(P^{-1}\) denotes the quantile function for \(p\), i.e., \(P^{-1}(s) = \min\{x : P(x) \geq s\}\) for \(0 < s < 1\) where \(P\) denotes the cdf of \(p\). A plug-in estimate for \(L\) is called a trimmed L-statistic, or a trimmed linear combination of quantiles. (It is called trimmed because the restricted support of \(\ell\) discards outliers.) L-statistics are good for robust estimation of a location parameter.

Now assume that \(\ell\) is continuous. Then \(L\) is Hadamard differentiable (and the L-statistics are asymptotically normal) at all \(p \in P(\mathbb{R})\) [vdW98, Lemma 22.10]. The functional derivative at \(p\) of \(L\), evaluated at a bounded signed measure \(m\), is

\[
\frac{\partial L}{\partial p}(m) = -\int \ell(P(x)) M(x) dx
\]

where \(M(x) = m((\infty, x])\). The asymptotic variance of the L-statistics is

\[
\sigma^2 = \int \int \ell(P(y)) \Gamma(y, z) \ell(P(z)) dy dz,
\]

where

\[
\Gamma(y, z) = P(y) \wedge P(z) - P(y) P(z).
\]

(15)

This formula is obtained via Donsker’s Theorem: Let \(P_n\) denote the cdf of \(\epsilon_n\), a random bounded function. Then \(n^{1/2}(P_n(t) - P(t))\) converges in law to a Gaussian process \(\{B(t)\}_{t \in \mathbb{R}}\) with covariance

\[
\Gamma(s, t) = \mathbb{E}_p [B(s) B(t)] = P(s) \wedge P(t) - P(s) P(t).
\]

(16)

Finally, the influence function is

\[
\phi_p(x) = \frac{\partial L_p}{\partial p}(\delta(x) - p) = -\int \ell(P(y))(H_x - P)(y) dy,
\]

(17)

where \(H_x\) denotes the cdf of \(\delta(x)\). Note that \(\sigma^2 = \mathbb{E}_p \phi_p^2\) and

\[
\mathbb{E}_p \phi_p^2 = \int \int \ell(P_n(y)) [P_n(y) \wedge P_n(z) - P_n(y) P_n(z)] \ell(P_n(z)) dy dz.
\]

Let \(v_{\text{jack}}\) denote the jackknife variance estimator for \(\sigma^2\). We find that the \(v_{\text{jack}}\) is asymptotically equivalent to \(\mathbb{E}_p \phi_p^2\) and asymptotically normal:
Proposition 3. Suppose $p$ has no point masses and $\ell'$ is Hölder continuous of order $h > 1/2$. Then

$$n^{1/2} \left( v_{\text{jack}} - \sigma^2 \right) = n^{1/2} \left( \mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2 \right) + O_s(n^{1/2-h})$$

and converges in law to the Gaussian random variable $Y + Z$, where

$$Y = \int \int \ell(P(y)) \{ B(y \wedge z) - P(y)B(z) - B(y)P(z) \} \ell(P(z)) dydz$$

$$Z = 2 \int \int \ell'(P(y))B(y)\Gamma(y, z)\ell(P(z)) dydz$$

and $B$ denotes the Brownian Bridge (16).

Proof: We prove first that $n^{1/2} \left( \mathbb{E}_{\epsilon_n} \phi_{\epsilon_n}^2 - \sigma^2 \right)$ converges in law to $Y + Z$, and afterwards we establish (18).

Define

$$Y_n = n^{1/2} \left( \sum_{i=1}^{n} \phi_p(x_i)^2 - \sigma^2 \right)$$

$$Z_n = -2n^{-1/2} \sum_{i=1}^{n} \phi_p(x_i) \int \ell'(P(y))(P_n - P)(y)(H_{x_i} - P_n)(y) dy.$$ (20)

We claim that $Y_n$ converges in law to $Y$ and $Z_n$ converges in law to $Z$. To see this, substitute (17) for $\phi_p$ in the definitions of $Y_n$ and $Z_n$, and apply Donkser’s Theorem. Substituting (17) for $\phi_p$ yields

$$Y_n = \int \int \ell(P(y))n^{1/2} \left( \frac{1}{n} \sum_{i=1}^{n} H_{x_i}(y)H_{x_i}(z) - P(y) \wedge P(z) \right) \ell(P(z)) dydz$$

$$- \int \int \ell(P(y))P(y)n^{1/2}(P_n - P)(z)\ell(P(z)) dydz$$

$$- \int \int \ell(P(y))n^{1/2}(P_n - P)(y)P(z)\ell(P(z)) dydz$$

$$Z_n = 2n^{-1/2} \sum_{i=1}^{n} \int \ell'(P(y))(P_n - P)(y)(H_{x_i} - P_n)(y)\ell(P(z))(H_{x_i} - P)(z) dydz$$

$$= 2 \int \ell'(P(y))n^{1/2}(P_n - P)(y) \left( \frac{1}{n} \sum_{i=1}^{n} H_{x_i}(y)H_{x_i}(z) - P_n(y)P_n(z) \right) \ell(P(z)) dydz.$$ (19)

Note that $\frac{1}{n} \sum_{i=1}^{n} H_{x_i}(y)H_{x_i}(z) - P_n(y)P_n(z)$ in the expression for $Z_n$ converges almost surely to $\Gamma(y, z)$ of (13). Also, in the expression for $Y_n$,

$$n^{1/2} \left( \frac{1}{n} \sum_{i=1}^{n} H_{x_i}(y)H_{x_i}(z) - P(y) \wedge P(z) \right) = n^{1/2} \left( P_n(y) \wedge P_n(z) - P(y) \wedge P(z) \right)$$
converges in law to the Gaussian process \( \mathbf{B}(y \wedge z) \). Writing \( M_{ni} = H_{x_i} - P_n \), we find that

\[
\phi_{\epsilon_n}(x_i) = - \int \left\{ \ell(P(y) + \ell(P(y))(P_n - P)(y) + O_s(n^{-h}) \right\} M_{ni}(y) dy = \phi_{\epsilon_n}(x_i) - \int \ell'(P(y))(P_n - P)(y)M_{ni}(y)dy + O_s(n^{-h}).
\]  
(21)

Equations (21) and (20) imply that

\[
n^{1/2} (\mathbb{E}_n \phi_{\epsilon_n}^2 - \sigma^2) = Y_n - Z_n + n^{-1/2} \sum_{i=1}^{n} \left( \int \ell'(P(y))(P_n - P)(y)M_{ni}(y)dy \right)^2 + O_s(n^{1/2-h}).
\]

But the third term on the right hand side of the last equation is \( O_s(n^{-1/2}) \), since

\[
\frac{1}{n} \sum_{i=1}^{n} \left( \int \ell'(P(y))(P_n - P)(y)M_{ni}(y)dy \right)^2 = \frac{1}{n} \sum_{i=1}^{n} \int \ell'(P(y))(P_n - P)(y)\ell'(P(y))(P_n - P)(y)M_{ni}(y)M_{ni}(z)dydz
\]
\[
= \int \int \ell'(P(y))(P_n - P)(y)\ell'(P(y))(P_n - P)(z)M_{ni}(y)M_{ni}(z)dydz
\]

and

\[
\frac{1}{n} \sum_{i=1}^{n} M_{ni}(y)M_{ni}(z) = \frac{1}{n} \sum_{i=1}^{n} H_{x_i}(y)H_{x_i}(z) - P_n(y)P_n(z),
\]

converges almost surely to \( \Gamma(y, z) \). Thus,

\[
n^{1/2} (\mathbb{E}_n \phi_{\epsilon_n}^2 - \sigma^2) = Y_n - Z_n + O_s(n^{-h}),
\]

so that \( n^{1/2} (\mathbb{E}_n \phi_{\epsilon_n}^2 - \sigma^2) \) converges in law to \( Y + Z \), a Gaussian random variable.

It remains to establish (21). To this end it suffices to show that

\[
\max_{1 \leq i \leq n} \{ |Q'_n - Q'_n - \phi_{\epsilon_n}(x_i)| \} = O_s(n^{-h}),
\]

for then, since \( v_{jack} = (n - 1)^{-1} \sum (Q'_n - Q'_n)^2 \), it would follow that

\[
n^{1/2} (v_{jack} - \sigma^2) = n^{1/2} \left( \frac{1}{n - 1} \sum_{i=1}^{n} \phi_{\epsilon_n}^2(x_i) - \sigma^2 \right) + O_s(n^{1/2-h}) = n^{1/2} (\mathbb{E}_n \phi_{\epsilon_n}^2 - \sigma^2) + O_s(n^{1/2-h}).
\]
Let $P_{ni}$ denote the cdf of $\epsilon_{ni}$. Integration by parts of (17) shows that

$$
\phi_{\epsilon_n}(x_i) = \int x d[\ell(P_n)(H_{x_i} - P_n)(y)]
$$

(23)

(the boundary term vanishes because (14) is trimmed). Suppose $x_1, x_2, x_3, \ldots$ are distinct (we are assuming that $p$ has no point masses, so this is the case almost surely). Then (23) becomes

$$
\phi_{\epsilon_n}(x_i) = x_i \ell(P_n(x_i)) + \sum_{j: x_j > x_i} x_j \{\ell(P_n(x_j)) - \ell(P_n(x_j) - 1/n)\}
$$

$$
- \sum_{j=1}^n x_j \{\ell(P_n(x_j))P_n(x_j) - \ell(P_n(x_j) - 1/n)(P_n(x_j) - 1/n)\},
$$

which we rewrite as $\phi_{\epsilon_n}(x_i) = A + B_i + C_i + D_i$ with

$$
A = -\frac{1}{n} \sum_{j=1}^n x_j \ell(P_n(x_j) - 1/n)
$$

$$
B_i = x_i \ell(P_n(x_i))
$$

$$
C_i = -\sum_{j: x_j \leq x_i} x_j \{\ell(P_n(x_j)) - \ell(P_n(x_j) - 1/n)\} P_n(x_j)
$$

$$
D_i = \sum_{j: x_j > x_i} x_j \{\ell(P_n(x_j)) - \ell(P_n(x_j) - 1/n)\} (1 - P_n(x_j)).
$$

(24)

For $1 \leq i \leq n$, let

$$
\zeta_{ni}(x) = (n-1) \int_{P_n(x)-\frac{1}{n-1}}^{P_n(x)} \ell(s) ds.
$$

Observe that $\ell(\zeta_{ni}(x)) = \ell(\zeta_{nk}(x))$ if $x < \min\{x_i, x_k\}$ or if $x > \max\{x_i, x_k\}$, and

$$
\zeta_{nk}(x) - \zeta_{ni}(x) = (n-1) \int_{P_n(x)}^{P_n(x)+\frac{1}{n-1}} \ell(s) - \ell(s-1/(n-1)) ds \quad \text{if } x_i < x < x_k
$$

$$
\zeta_{nk}(x) - \zeta_{ni}(x) = -(n-1) \int_{P_n(x)-\frac{1}{n-1}}^{P_n(x)} \ell(s) - \ell(s-1/(n-1)) ds \quad \text{if } x_k < x < x_i.
$$

(25)
Thus $L(\epsilon_{ni}) = \frac{1}{n-1} \sum_{j \neq i} x_j \zeta_{ni}(x_j)$ and

$$Q'_{ni} - \bar{Q}'_n = - \sum_{j \neq i} x_j \zeta_{ni}(x_j) + \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} x_j \zeta_{nk}(x_j)$$

$$= - \frac{1}{n} \sum_{k=1}^{n} x_k \zeta_{ni}(x_k) + \frac{1}{n} \sum_{k=1}^{n} x_k \zeta_{nk}(x_k) + \frac{1}{n} \sum_{k=1}^{n} \sum_{j \neq k} x_j \{\zeta_{nk}(x_j) - \zeta_{ni}(x_j)\}$$

$$= - \frac{1}{n} \sum_{k=1}^{n} x_k \zeta_{ni}(x_k) + \frac{1}{n} \sum_{k=1}^{n} x_k \zeta_{nk}(x_k) + \frac{1}{n} \sum_{j: x_j < x_i} x_j \{\zeta_{nk}(x_j) - \zeta_{ni}(x_j)\}$$

$$+ \frac{1}{n} \sum_{j: x_j > x_i} \sum_{k: x_k > x_j} x_j \{\zeta_{nk}(x_j) - \zeta_{ni}(x_j)\}.$$ 

Using (25) we find that $Q'_{ni} - \bar{Q}'_n = A'_i + B'_i + C'_i + D'_i$ with

$$A'_i = - \frac{1}{n} \sum_{j=1}^{n} x_j \zeta_{ni}(x_j)$$

$$B'_i = \frac{1}{n} \sum_{j=1}^{n} x_j \zeta_{nj}(x_i)$$

$$C'_i = (n-1) \sum_{j: x_j < x_i} x_j (P_n(x_j) - 1/n) \int_{P_n(x_j)}^{P_n(x_j) + \frac{1}{n-1}} \ell(s) - \ell(s - 1/(n-1)) ds$$

$$D'_i = (n-1) \sum_{j: x_j > x_i} x_j (1 - P_n(x_j)) \int_{P_n(x_j)}^{P_n(x_j) + \frac{1}{n-1}} \ell(s) - \ell(s - 1/(n-1)) ds. \quad (26)$$

The sequence $\{P_n\}$ converges almost surely to $P$ and hence it is almost surely tight. Thus there exists a (random) bound $M > 0$ such that $P_n(x) < \alpha/2$ if $x < M$ and $P_n(x) > 1 - \alpha/2$ if $x > M$. Since $\ell$ vanishes off of $[\alpha, 1 - \alpha]$, it follows that $B_i = 0$ if $|x_i| > M$, and $B'_i = 0$ if $|x_i| > M$ and $1/(n-1) < \alpha/4$. Similarly, if $n$ is sufficiently large, the sums defining $A', C', D', A'_i, C'_i$ and $D'_i$ in (24) and (25) may be replaced with sums over $j$ such that $|x_i| > M$. Thus

$$|A'_i - A| \leq M \frac{n-1}{n} \sum_{j=1}^{n} \int_{P_n(x_j)}^{P_n(x_j) + \frac{1}{n-1}} |\ell(s) - \ell(P_n(x_j) - 1/n)| ds$$

$$|B'_i - B| \leq M \frac{n-1}{n} \sum_{j=1}^{n} \int_{P_n(x_j)}^{P_n(x_j) + \frac{1}{n-1}} |\ell(s) - \ell(P_n(x_i))| ds$$

are both $O_s(1/n)$ since $\ell$ is differentiable. For $n > 1N$ and $s \in [1/n, 1]$, let $t_n(s)$ be a number between $s - 1/n$ and $s$ such that $\ell'(t_n(s)) = n (\ell(s) - \ell(s - 1/n))$. (The functions $t_n$ may be chosen to be
continuous, since $\ell'$ is continuous.) We now have

$$|C'_i - C_i| \leq M |\ell'(t_n(P_n(x_i)))| P_n(x_i) + \frac{M}{n} \sum_{j : x_j < x_i} \int_{P_n(x_j) - \frac{1}{n^h}}^{P_n(x_j)} |\ell'(t_n-1(s))| ds \leq M \sum_{j : x_j < x_i} P_n(x_j) \int_{P_n(x_j) - \frac{1}{n^h}}^{P_n(x_j)} |\ell'(t_n-1(s))| ds,$$

$$|D'_i - D_i| \leq M \sum_{j : x_j > x_i} (1 - P_n(x_j)) \int_{P_n(x_j) + \frac{1}{n^h}}^{P_n(x_j)} |\ell'(t_n-1(s))| ds.$$

But $\ell'(t_n-1(s)) - \ell'(t_n(P_n(x_j))) = O(n^{-h})$ throughout the interval of integration because of the Hölder continuity of $\ell'$, and so $|C'_i - C_i|$ and $|D'_i - D_i|$ are both $O_s(n^{-h})$ uniformly in $i$. The preceding estimates show that

$$|Q'_{ni} - Q_n - \phi_{e_n}(x_i)| \leq |A'_i - A| + |B'_i - B| + |C'_i - C_i| + |D'_i - D_i| = O_s(n^{-h})$$

uniformly in $i$, establishing (22). \hfill \Box

Proposition 3 is also true as stated for $L(p) = \int x \ell(P(x)) p(dx)$, which is not exactly the same as the L-functional (14) but has the same functional derivative. An argument similar to the one above shows that the asymptotic variance of $n^{1/2} \left( v_{jack} - \sigma^2 \right)$ equals $\text{Var}(Y + Z)$ with $Y$ and $Z$ as in (14). On the other hand, one can show that $\text{Var} \phi_p = \text{Var} Y$. This is contrary to [ST95, p 43], where it is asserted that $\text{Var} Y$ is the asymptotic variance of $n^{1/2} \left( v_{jack} - \sigma^2 \right)$.

5 Acknowledgments

Thanks to Steve Evans for his advice and encouragement. Thanks to Rudolf Beran. The author is supported by the Austrian START project *Nonlinear Schrödinger and quantum Boltzmann equations*.

6 References

[Ber84] R. Beran. Jackknife approximations to bootstrap estimates. *The Annals of Statistics* 12 (1): 101 - 118, 1984.

[Parr85] W. Parr. Jackknifing differentiable statistical functions. *Journal of the Royal Statistical Society B* 47 (1): 56 - 66, 1985.

[ST95] J. Shao and D. Tu. *The Jackknife and Bootstrap*. Springer-Verlag, New York, 1995.

[vdW98] A.W. van der Waart. *Asymptotic Statistics*. Cambridge University Press, 1998.