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 ϵ_n the map that converts n data points $x_1, x_2, \ldots, x_n$ into the empirical measure

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

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

$$T_n \; = \; T(\epsilon_n(\mathbf{x})). \tag{2}$$

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 σ^2: one transforms the n original data points into n pseudovalues and computes the sample variance of those pseudovalues.

Given the data $\mathbf{x} = x_1, x_2, \ldots, x_n$, the *jackknife pseudovalues* are

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

with ϵ_n as in (1) and

$$\epsilon_{ni} \; = \; \frac{1}{n-1}\sum_{j \neq i} \delta(x_j). \tag{3}$$

The *jackknife variance estimator* is

$$v_{jack}(x_1, x_2, \ldots, x_n) \; = \; \frac{1}{n-1}\sum_{i=1}^{n} \left(Q_{ni} - \overline{Q_n} \right)^2 \tag{4}$$

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

A functional derivative of T at p, denoted ∂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 ∂T_p on $\mathcal{M}(X)$ such that

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

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 $T : \mathcal{P}(\mathbb{R}) \longrightarrow \mathbb{R}$ is *Hadamard differentiable* at p if there exists a continuous linear functional ∂T_p on $\mathcal{M}(\mathbb{R})$ such that

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

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 $\|m\| = \sup_{t\in\mathbb{R}}\left\{\left|m((-\infty,t])\right|\right\}$. If T is Hadamard differentiable at p, the variance of $n^{1/2}T(\epsilon_n)$ tends to

$$\sigma^2 \;=\; \mathbb{E}_p\phi_p^2 \tag{5}$$

as $n \longrightarrow \infty$, where $\phi_p(x)$ is the *influence function*

$$\phi_p(x) \;=\; \partial T_p(\delta(x) \;-\; p) \tag{6}$$

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

If T is smooth enough then $n^{1/2}\left(v_{jack} - \sigma^2\right)$ is also asymptotically normal. In this note we calculate the asymptotic variance of v_{jack} (i.e., the limit as $n \longrightarrow \infty$ of the variance of $n^{1/2}v_{jack}$) for two very well behaved functionals T: smooth functions of the mean $T(p) \;=\; g\left(\overline{p}\right)$ and smooth trimmed L-functionals. In these cases, the asymptotic variance of v_{jack} equals that of $\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n}$, the estimator of σ^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_{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_{jack} adequately. In fact, [ST95, Section 2.2.3] gets it wrong, conjecturing that the asymptotic variance of v_{jack} should equal Var ϕ_p^2 for sufficiently regular functionals! However, Theorem 2 of [Ber84] does contain a general formula for the variance of v_{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 $n^{1/2}\left(v_{jack} - \sigma^2\right)$ equals Var ϕ_p^2 when T is of the form (14), 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^∞ distance between the cdf's of p and q. A functional $T : \mathcal{P}(\mathbb{R}) \longrightarrow \mathbb{R}$ is ρ-**Lipschitz differentiable at** q if*

$$T(p_k) - T(q_k) \;-\; \partial T_q(p_k - q_k) \;=\; O\left(\rho(p_k, q_k)^2\right) \tag{7}$$

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

A close look at the definition of ρ-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 ρ 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

$$Q'_{ni}(x_1, x_2, \ldots, x_n) \;=\; (n-1)\left[T(\epsilon_n) - T(\epsilon_{ni})\right]. \tag{8}$$

Substituting Q'_{ni} for Q_{ni} and $\overline{Q'_n} = \frac{1}{n}\sum Q'_{nj}$ for $\overline{Q_n} = \frac{1}{n}\sum Q_{nj}$ in (4) does not change the value of v_{jack}, so one may compute v_{jack} by the same formula using the Q'_{ni}. Using the modified pseudovalues Q'_{ni} makes it easier to take advantage of the magic formula $(n-1)\left(\epsilon_n - \epsilon_{ni}\right) \;=\; \delta_{x_i} - \epsilon_n$.

2 Using pseudovalues to estimate the variance of ϕ_p^2

One aim of this letter is to emphasize that Var ϕ_p^2 is typically *not* the asymptotic variance of $n^{1/2}\big(v_{jack} \;-\; \sigma^2\big)$, contrary to the assertion of [ST95, p 42]. However, should one desire an estimate of Var ϕ_p^2 for some reason, the pseudovalues can be used to this end. Once one has already computed v_{jack}, the variance of ϕ_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 *continuously Gâteaux differentiable* and ϕ_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.

Definition 1. *A functional T is* **continuously Gâteaux differentiable** *at p if it has Gâteaux derivative ∂T_p at p and if*

$$\lim_{k\to\infty}\sup_{x\in\mathbb{R}}\left\{\left|\frac{T(p_k+t_k(\delta(x)-p_k))-T(p_k)}{t_k} \;-\; \partial T_p(\delta(x)-p_k)\right|\right\} \;=\; 0 \tag{9}$$

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.

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

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

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 $p\circ\phi_p^{-1}$:

$$\epsilon_n(Q'_{n1}, Q'_{n2}, \ldots, Q'_{nn}) \;\longrightarrow\; p\circ\phi_p^{-1} \qquad \text{a.s.}$$

Proof: Omitted, but cf. the proof of Theorem 2.3 in [ST95]. □

Now, suppose that $T : \mathcal{P}(\mathbb{R}) \longrightarrow \mathbb{R}$ 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

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

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 $\tau^2 \longrightarrow \text{Var }\phi_p^2$ almost surely.

One may also estimate Var ϕ_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

$$\frac{1}{n^{1/2}}\sum_{i=1}^{n}\left(\text{sq}(Q^*_{n,i}) - \text{sq}(Q'_{n,i})\right). \tag{10}$$

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

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

$$T(m) \; = \; g\left(\overline{m}\right)$$

be defined for all finite signed measures m with finite first moment. The functional derivative at m of T, evaluated at q, is $\partial T_m(q) = g'\left(\overline{m}\right)\overline{q}$; the influence function (6) is $\phi_m(x) = g'\left(\overline{m}\right)(x - \overline{m})$. 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 (2). Then the asymptotic variance of $n^{1/2}\left(T_n - T(p)\right)$ is

$$\sigma^2 \; = \; g'(\overline{p})^2\Big\{\int x^2 p(dx) - \overline{p}^2\Big\}. \tag{11}$$

Let v_{jack} denote the jackknife variance estimator for σ^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 $n^{1/2}(v_{jack} - \sigma^2)$ and $n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} - \sigma^2\right)$ have the same limit in distribution, if any.*

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

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

whence

$$n^{1/2}\left(v_{jack} - \sigma^2\right) \; = \; n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} - \sigma^2\right) + \frac{n^{1/2}}{n-1}\,\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} \; + \; \frac{n^{3/2}}{n-1}\left\{\frac{2}{n}\sum_{i=1}^n \phi_{\epsilon_n}(x_i)\Delta_{ni} \; + \; \frac{1}{n}\sum_{i=1}^n \Delta^2_{ni}\right\}.$$

To prove that $n^{1/2}(v_{jack} - \sigma^2)$ and $n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} - \sigma^2\right)$ have the same limit in distribution (if any) it suffices to show that

$$\frac{n^{1/2}}{n-1}\,\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} \; + \; \frac{n^{3/2}}{n-1}\left(\frac{1}{n}\sum_{i=1}^n \phi_{\epsilon_n}(x_i)\Delta_{ni} \; + \; \frac{1}{n}\sum_{i=1}^n \Delta^2_{ni}\right) \tag{12}$$

converges almost surely to 0.

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

$$\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} \;=\; \frac{1}{n}\sum_{i=1}^{n}\phi^2_{\epsilon_n}(x_i) \;=\; \frac{1}{n}\sum_{i=1}^{n} g'\left(\overline{\epsilon_n}\right)^2\left(x_i-\overline{\epsilon_n}\right)^2$$

converges almost surely to σ^2.

To show that the other terms tend to zero we need a bound on Δ_{ni}. Since g is differentiable, $g\left(\overline{\epsilon_{nj}}\right) - g\left(\overline{\epsilon_{ni}}\right) \;=\; g'\left(\eta_{ji}\right)\left(\overline{\epsilon_{nj}}-\overline{\epsilon_{ni}}\right)$ for some η_{ji} between $\overline{\epsilon_{ni}}$ and $\overline{\epsilon_{nj}}$, so that

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

Therefore, since $\phi_{\epsilon_n}(x_i) = g'\left(\overline{\epsilon_n}\right)\left(x_i-\overline{\epsilon_n}\right) = \frac{1}{n}\sum_j g'\left(\overline{\epsilon_n}\right)\left(x_i-x_j\right)$,

$$\begin{aligned}\Delta_{ni} \;=\; \left(Q'_{ni}-\overline{Q'_n}\right)-\phi_{\epsilon_n}(x_i) \;&=\; \frac{n-1}{n}\sum_{j=1}^{n} g'\left(\eta_{ji}\right)\left(\overline{\epsilon_{nj}}-\overline{\epsilon_{ni}}\right) \;-\; \frac{1}{n}\sum_{j=1}^{n} g'\left(\overline{\epsilon_n}\right)\left(x_i-x_j\right)\\ &=\; \frac{1}{n}\sum_{j=1}^{n}\Big(g'\left(\eta_{ji}\right)-g'\left(\overline{\epsilon_n}\right)\Big)\left(x_i-x_j\right).\end{aligned}$$

But g' is Hölder continuous of order h and $|\eta_{ji}-\overline{\epsilon_n}| < \max\{|\overline{\epsilon_{nj}}-\overline{\epsilon_n}|, |\overline{\epsilon_{ni}}-\overline{\epsilon_n}|\}$, so

$$\Big|g'\left(\eta_{ji}\right)-g'\left(\overline{\epsilon_n}\right)\Big| \;\le\; C\Big(|\overline{\epsilon_{nj}}-\overline{\epsilon_n}|^h+|\overline{\epsilon_{ni}}-\overline{\epsilon_n}|^h\Big) \;\le\; C(n-1)^{-h}\Big(|\overline{\epsilon_n}-x_j|^h+|\overline{\epsilon_n}-x_i|^h\Big),$$

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

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

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

$$\frac{1}{n}\sum_{i=1}^{n}\Delta^2_{ni} \;=\; O_s\big(n^{-2h}\big),$$

and then, by the Cauchy-Schwartz inequality, that

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

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, $n^{1/2}\left(v_{jack}-\sigma^2\right)$ and $n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n}-\sigma^2\right)$ 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 $n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n}-\sigma^2\right)$. 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

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

so that

$$\begin{aligned}\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} \;&=\; \frac{1}{n}\sum_{i=1}^{n}\phi^2_{\epsilon_n}(x_i) \;=\; \left[g'\left(\overline{p}\right) + g''\left(\overline{p}\right)\left(\overline{\epsilon_n} - \overline{p}\right)\right]^2 \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{\epsilon_n}\right)^2 \;+\; O_s\big(n^{-(r+1)/2}\big) \\ &=\; \left[g'\left(\overline{p}\right)^2 + 2g'\left(\overline{p}\right)g''\left(\overline{p}\right)\left(\overline{\epsilon_n} - \overline{p}\right)\right]\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{\epsilon_n}\right)^2 \;+\; O_s\big(n^{-(r+1)/2}\big).\end{aligned}$$

From formula (11) for σ^2 we see that

$$\begin{aligned}n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} - \sigma^2\right) \;&=\; g'\left(\overline{p}\right)^2 n^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{\epsilon_n}\right)^2 \;-\; \left\{\int x^2 p(dx) - \overline{p}^2\right\}\right) \\ &\quad + 2g'\left(\overline{p}\right)g''\left(\overline{p}\right)n^{1/2}\left(\overline{\epsilon_n} - \overline{p}\right)\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{\epsilon_n}\right)^2 \;+\; O_s\left(n^{-r/2}\right). \qquad (13)\end{aligned}$$

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

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

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 $n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} - \sigma^2\right)$ is asymptotically normal with variance $(a,b)\Gamma(a,b)^{tr}$, where $(a,b) = \left(g'(\overline{p})^2,\; 2g'(\overline{p})g''(\overline{p})\right)$ and Γ 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 $n^{1/2}\left(v_{jack} - \sigma^2\right)$ equals $(a,b)\Gamma(a,b)^{tr}$. In contrast, under the same conditions on p and g it may be shown that Var $\phi^2_p = a^2\Gamma_{1,1}$.

4 Trimmed L-statistics

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

$$L(p) \;=\; \int_0^1 P^{-1}(s)\ell(s)ds. \qquad (14)$$

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 ℓ discards outliers.) L-statistics are good for robust estimation of a location parameter.

Now assume that ℓ is continuous. Then L is Hadamard differentiable (and the L-statistics are asymptotically normal) at all $p \in \mathcal{P}(\mathbb{R})$ [vdW98, Lemma 22.10]. Assuming that P is continuous, i.e., that p has no point masses, the functional derivative at p of L, evaluated at a bounded signed measure m, is

$$\partial L_p(m) \;=\; -\int \ell(P(x))M(x)dx$$

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

$$\sigma^2 \;=\; \int\int \ell(P(y))\Gamma(y,z)\ell(P(z))dydz,$$

where

$$\Gamma(y,z) \;=\; P(y) \wedge P(z) \;-\; P(y)P(z). \tag{15}$$

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

$$\Gamma(s,t) \;=\; \mathbb{E}_p\left[\mathbf{B}(s)\mathbf{B}(t)\right] \;=\; P(s) \wedge P(t) - P(s)P(t). \tag{16}$$

Finally, the influence function is

$$\phi_p(x) \;=\; \partial L_p(\delta(x) - p) \;=\; -\int \ell(P(y))(H_x - P)(y)dy, \tag{17}$$

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

$$\mathbb{E}_{\epsilon_n}\phi_{\epsilon_n}^2 \;=\; \int\int \ell(P_n(y))\left[P_n(y) \wedge P_n(z) - P_n(y)P_n(z)\right]\ell(P_n(z))dydz.$$

Let v_{jack} denote the jackknife variance estimator for σ^2. We find that the v_{jack} is asymptotically equivalent to $\mathbb{E}_{\epsilon_n}\phi_{\epsilon_n}^2$ and asymptotically normal:

Proposition 3. *Suppose p has no point masses and ℓ' is Hölder continuous of order $h > 1/2$. Then*

$$n^{1/2}\left(v_{jack} - \sigma^2\right) \;=\; n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi_{\epsilon_n}^2 - \sigma^2\right) \;+\; O_s\big(n^{1/2-h}\big) \tag{18}$$

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

$$\begin{aligned} Y &= \int\int \ell(P(y))\left\{\mathbf{B}(y\wedge z) - P(y)\mathbf{B}(z) - \mathbf{B}(y)P(z)\right\}\ell(P(z))dydz \\ Z &= 2\int\int \ell'(P(y))\mathbf{B}(y)\Gamma(y,z)\ell(P(z))dydz \end{aligned} \tag{19}$$

and $\mathbf{B}$ denotes the Brownian Bridge (16).

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

Define

$$\begin{aligned} Y_n &= n^{1/2}\Big(\sum_{i=1}^n \phi_p(x_i)^2 - \sigma^2\Big) \\ Z_n &= -2n^{-1/2}\sum_{i=1}^n \phi_p(x_i)\int \ell'(P(y))(P_n - P)(y)\left(H_{x_i} - P_n\right)(y)dy. \end{aligned} \tag{20}$$

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

$$\begin{aligned} Y_n &= \int\int \ell(P(y))n^{1/2}\Big(\frac{1}{n}\sum_{i=1}^n H_{x_i}(y)H_{x_i}(z) \;-\; P(y)\wedge P(z)\Big)\ell(P(z))dydz \\ &\quad - \int\int \ell(P(y))P(y)n^{1/2}(P_n - P)(z)\ell(P(z))dydz \\ &\quad\quad - \int\int \ell(P(y))n^{1/2}(P_n - P)(y)P(z)\ell(P(z))dydz \\ Z_n &= 2n^{-1/2}\sum_{i=1}^n\int\int \ell'(P(y))(P_n - P)(y)\left(H_{x_i} - P_n\right)(y)\ell(P(z))(H_{x_i} - P)(z)dydz \\ &= 2\int\int \ell'(P(y))n^{1/2}(P_n - P)(y)\Big(\frac{1}{n}\sum_{i=1}^n H_{x_i}(y)H_{x_i}(z) - P_n(y)P_n(z)\Big)\ell(P(z))dydz. \end{aligned}$$

Note that $\frac{1}{n}\sum H_{x_i}(y)H_{x_i}(z) - P_n(y)P_n(z)$ in the expression for Z_n converges almost surely to $\Gamma(y,z)$ of (15). Also, in the expression for Y_n,

$$n^{1/2}\Big(\frac{1}{n}\sum_{i=1}^n H_{x_i}(y)H_{x_i}(z) - P(y)\wedge P(z)\Big) \;=\; n^{1/2}\left(P_n(y)\wedge P_n(z) - P(y)\wedge P(z)\right)$$

converges in law to the Gaussian process $\mathbf{B}(y\wedge z)$. Writing $M_{ni} = H_{x_i} - P_n$, we find that

$$\begin{aligned} \phi_{\epsilon_n}(x_i) &= -\int \Big\{\ell(P(y) + \ell'(P(y))(P_n - P)(y) + O_s(n^{-h})\Big\} M_{ni}(y)dy \\ &= \phi_p(x_i) \;-\; \int \ell'(P(y))(P_n - P)(y)M_{ni}(y)dy \;+\; O_s\big(n^{-h}\big). \end{aligned} \tag{21}$$

Equations (21) and (20) imply that

$$\begin{aligned} n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n} - \sigma^2\right) \;=\; Y_n \;+\; Z_n \;&+\; n^{-1/2}\sum_{i=1}^n\Big(\int \ell'(P(y))(P_n(y) - P(y))M_{ni}(y)dy\Big)^2 \\ &+\; O_s\big(n^{1/2-h}\big). \end{aligned}$$

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

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left(\int \ell'(P(y))(P_n(y)-P(y))M_{ni}(y)dy\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\int\int \ell'(P(y))(P_n-P)(y)\ell'(P(z))(P_n-P)(z)M_{ni}(y)M_{ni}(z)dydz \\
&= \int\int \ell'(P(y))(P_n-P)(y)\ell'(P(z))(P_n-P)(z)\frac{1}{n}\sum_{i=1}^{n}M_{ni}(y)M_{ni}(z)dydz
\end{aligned}
$$

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,

$$n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n}-\sigma^2\right) \;=\; Y_n \;+\; Z_n \;+\; O_s(n^{1/2-h}),$$

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

It remains to establish (18). To this end it suffices to show that

$$\max_{1\le i\le n}\left\{\left|Q'_{ni}-\overline{Q'_n}-\phi_{\epsilon_n}(x_i)\right|\right\} = O_s(n^{-h}), \tag{22}$$

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

$$
\begin{aligned}
n^{1/2}\left(v_{jack}-\sigma^2\right) &= n^{1/2}\Big(\frac{1}{n-1}\sum_{i=1}^{n}\phi^2_{\epsilon_n}(x_i) \;-\; \sigma^2\Big) \;+\; O_s(n^{1/2-h}) \\
&= n^{1/2}\left(\mathbb{E}_{\epsilon_n}\phi^2_{\epsilon_n}-\sigma^2\right) \;+\; O_s(n^{1/2-h}).
\end{aligned}
$$

Let P_{ni} denote the cdf of ϵ_{ni}. Integration by parts of (17) shows that

$$\phi_{\epsilon_n}(x_i) \;=\; \int x d\left[\ell(P_n)(H_{x_i}-P_n)(y)\right] \tag{23}$$

(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

$$
\begin{aligned}
\phi_{\epsilon_n}(x_i) &= x_i\ell(P_n(x_i)) \;+\; \sum_{j:\;x_j>x_i} x_j\left\{\ell(P_n(x_j)) \;-\; \ell\big(P_n(x_j)-1/n\big)\right\} \\
&\quad - \sum_{j=1}^{n} x_j\left\{\ell(P_n(x_j))P_n(x_j) \;-\; \ell\big(P_n(x_j)-1/n\big)\big(P_n(x_j)-1/n\big)\right\},
\end{aligned}
$$

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

$$
\begin{aligned}
A &= -\frac{1}{n}\sum_{j=1}^{n} x_j \ell\big(P_n(x_j) - 1/n\big) \\
B_i &= x_i \ell(P_n(x_i)) \\
C_i &= -\sum_{j:x_j \le x_i} x_j \left\{\ell(P_n(x_j)) - \ell(P_n(x_j) - 1/n)\right\} P_n(x_j) \\
D_i &= \sum_{j:x_j > x_i} x_j \left\{\ell(P_n(x_j)) - \ell(P_n(x_j) - 1/n)\right\} (1 - P_n(x_j))\,.
\end{aligned} \tag{24}
$$

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

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

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

$$
\begin{aligned}
\zeta_{nk}(x) - \zeta_{ni}(x) &= (n-1)\int_{P_{ni}(x)}^{P_{ni}(x)+\frac{1}{n-1}} \ell(s) - \ell\big(s - 1/(n-1)\big)ds \qquad \text{if} \quad x_i < x < x_k \\
\zeta_{nk}(x) - \zeta_{ni}(x) &= -(n-1)\int_{P_{ni}(x)-\frac{1}{n-1}}^{P_{ni}(x)} \ell(s) - \ell\big(s - 1/(n-1)\big)ds \qquad \text{if} \quad x_k < x < x_i.
\end{aligned} \tag{25}
$$

Thus $L(\epsilon_{ni}) = \frac{1}{n-1}\sum\limits_{j:j\ne i} x_j \zeta_{ni}(x_j)$ and

$$
\begin{aligned}
Q'_{ni} - \overline{Q'_n} &= -\sum_{j:j\ne i} x_j\zeta_{ni}(x_j) \;+\; \frac{1}{n}\sum_{k=1}^{n}\sum_{j:j\ne k} x_j\zeta_{nk}(x_j) \\
&= -\frac{1}{n}\sum_{k=1}^{n} x_k\zeta_{ni}(x_k) \;+\; \frac{1}{n}\sum_{k=1}^{n} x_i\zeta_{nk}(x_i) \;+\; \frac{1}{n}\sum_{k=1}\sum_{j:j\ne k,i}^{n} x_j\left\{\zeta_{nk}(x_j) - \zeta_{ni}(x_j)\right\} \\
&= -\frac{1}{n}\sum_{k=1}^{n} x_k\zeta_{ni}(x_k) \;+\; \frac{1}{n}\sum_{k=1}^{n} x_i\zeta_{nk}(x_i) \;+\; \frac{1}{n}\sum_{j:x_j<x_i}\;\sum_{k:x_k<x_j} x_j\left(\zeta_{nk}(x_j) - \zeta_{ni}(x_j)\right) \\
&\quad + \frac{1}{n}\sum_{j:x_j>x_i}\;\sum_{k:x_k>x_j} x_j\left(\zeta_{nk}(x_j) - \zeta_{ni}(x_j)\right).
\end{aligned}
$$

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

$$\begin{aligned}
A'_i &= -\frac{1}{n}\sum_{j=1}^{n} x_j \zeta_{ni}(x_j) \\
B'_i &= \frac{1}{n}\sum_{j=1}^{n} x_i \zeta_{nj}(x_i) \\
C'_i &= (n-1)\sum_{j:x_j<x_i} x_j\,(P_n(x_j)-1/n)\int_{P_{ni}(x_j)-\frac{1}{n-1}}^{P_{ni}(x_j)} \ell(s)-\ell\big(s-1/(n-1)\big)ds \\
D'_i &= (n-1)\sum_{j:x_j>x_i} x_j\,(1-P_n(x_j))\int_{P_{ni}(x_j)}^{P_{ni}(x_j)+\frac{1}{n-1}} \ell(s)-\ell\big(s-1/(n-1)\big)ds. \qquad (26)
\end{aligned}$$

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 ℓ 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_i, D_i, A'_i, C'_i and D'_i in (24) and (26) may be replaced with sums over j such that $|x_i|>M$. Thus

$$\begin{aligned}
|A'_i - A| &\le M\frac{n-1}{n}\sum_{j=1}^{n}\int_{P_{ni}(x_j)-\frac{1}{n-1}}^{P_{ni}(x_j)} \big|\ell(s)-\ell\big(P_n(x_j)-1/n\big)\big|\,ds \\
|B'_i - B_i| &\le M\frac{n-1}{n}\sum_{j=1}^{n}\int_{P_{nj}(x_i)-\frac{1}{n-1}}^{P_{nj}(x_i)} |\ell(s)-\ell(P_n(x_i))|\,ds
\end{aligned}$$

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

$$\begin{aligned}
|C'_i - C_i| \le\ & M\big|\ell'(t_n(P_n(x_i)))\big|P_n(x_i) + \frac{M}{n}\sum_{j:x_j<x_i}\int_{P_{ni}(x_j)-\frac{1}{n-1}}^{P_{ni}(x_j)} \big|\ell'(t_{n-1}(s))\big|ds \\
&+ M\sum_{j:x_j<x_i} P_n(x_j)\int_{P_{ni}(x_j)-\frac{1}{n-1}}^{P_{ni}(x_j)} \Big|\ell'(t_{n-1}(s))-\ell'(t_n(P_n(x_j)))\Big|ds \\
|D'_i - D_i| \le\ & M\sum_{j:x_j>x_i}(1-P_n(x_j))\int_{P_{ni}(x_j)}^{P_{ni}(x_j)+\frac{1}{n-1}} \Big|\ell'(t_{n-1}(s))-\ell'(t_n(P_n(x_j)))\Big|ds.
\end{aligned}$$

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

$$\big|Q'_{ni} - \overline{Q'_n} - \phi_{\epsilon_n}(x_i)\big| \le |A'_i - A| + |B'_i - B_i| + |C'_i - C_i| + |D'_i - D_i| = O_s\big(n^{-h}\big)$$

uniformly in i, establishing (22). □

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 when P is continuous. An argument similar to the one above shows that the asymptotic variance of $n^{1/2}\left(v_{jack} - \sigma^2\right)$ equals Var $(Y + Z)$ with Y and Z as in (19). On the other hand, one can show that Var $\phi_p^2 = $ Var Y. This is contrary to [ST95, p 43], where it is asserted that 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.