*Work supported by the Air Force Office of Scientific Research, A.F.S.C., U.S.A.F., Grant No. AFOSR-7402736.

SOME INVARIANCE PRINCIPLES RELATING TO JACKKNIFING AND THEIR ROLE IN SEQUENTIAL ANALYSIS*

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1055

February 1976

SOME INVARIANCE PRINCIPLES RELATING TO JACKKNIFING AND THEIR ROLE IN SEQUENTIAL ANALYSIS*

Pranab Kumar Sen
University of North Carolina, Chapel Hill

ABSTRACT

For a broad class of jackknife statistics, it is shown that the Tukey estimator of the variance converges almost surely to its population counterpart. Moreover, the usual invariance principles (relating to the Wiener process approximations) usually filter through jackknifing under no extra regularity conditions. These results are then incorporated in providing a bounded-length (sequential) confidence interval and a preassigned-strength sequential test for a suitable parameter based on jackknife estimators.

AMS 1970 Classification No: 60B10, 62E20, 62G35, 62L10

Key Words & Phrases: Almost sure convergence, Brownian motions, bounded-length confidence intervals; invariance principles, preassigned strength sequential tests, jackknife, stopping time, tightness, Tukey estimator of the variance, U-statistics, von Mises' functionals.

*Work supported by the Air Force Office of Scientific Research, A.F.S.C., U.S.A.F., Grant No. AFOSR-74-2736.

1. Introduction

The jackknife estimator, originally introduced for bias reduction by Quenouille and extended by Tukey for robust interval estimation, has been studied thoroughly by a host of workers during the past twenty years; along with some extensive bibliography, detailed studies are made in the recent papers of Arvesen (1969), Schucany, Gray and Owen (1971), Gray, Watkins and Adams (1972) and Miller (1974). One of the major concerns is the asymptotic normality of the studentized form of the jackknife statistics. The purpose of the present investigation is to focus on some deeper asymptotic properties of jackknife estimators and to stress their role in the asymptotic theory of sequential procedures based on jackknifing. Specifically, the almost sure convergence of the Tukey estimator of the variance is established here for a broad class of jackknife statistics and their asymptotic normality results are strengthened to appropriate (weak as well as strong) invariance principles yielding Wiener process approximations for the tail-sequence of jackknife estimators. These results are then incorporated in providing (i) a bounded-length (sequential) confidence interval and (ii) a prescribed-strength sequential test for a suitable parameter based on jackknife estimators.

Section 2 deals with the preliminary notions along with some new interpretations of the jackknife estimator and the Tukey estimator of the variance. For convenience of presentation, in Section 3, we adopt the framework of Arvesen (1969) and present the invariance principles for jackknifing U-statistics. Section 4 displays parallel results for general estimators. The two sequential problems of estimation and testing are treated in the last two sections of the paper.

2. Preliminary Notions

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed random vectors (i.i.d.r.v) with a distribution function F, defined on the $p(\geq 1)$ - dimensional Euclidean space R^p. Let

$$(2.1) \qquad \hat{\theta}_n = T_n(X_1,\ldots,X_n), \quad n \geq 1$$

be a sequence of estimators of a parameter θ, such that

$$(2.2) \qquad E\hat{\theta}_n = \theta + n^{-1}\beta_1 + n^{-2}\beta_2 + \ldots \quad (\Rightarrow E(\hat{\theta}_n - \theta) = O(n^{-1}))$$

where the β_j are unknown constants. Let us denote by

$$(2.3) \qquad \hat{\theta}^i_{n-1} = T_{n-1}(X_1,\ldots,X_{i-1}, X_{i+1},\ldots,X_n), \quad 1 \leq i \leq n,$$

$$(2.4) \qquad \hat{\theta}_{n,i} = n\hat{\theta}_n - (n-1)\hat{\theta}^i_{n-1}, \quad 1 \leq i \leq n,$$

$$(2.5) \qquad \theta^*_n = n^{-1}\textstyle\sum_{i=1}^n \hat{\theta}_{n,i} = n\hat{\theta}_n - (n-1)\left\{n^{-1}\sum_{i=1}^n \hat{\theta}^i_{n-1}\right\}.$$

Then, θ^*_n is termed the *jackknife estimator* of θ. Clearly, by (2.2), (2.3), and (2.5),

$$(2.6) \qquad E\theta^*_n = \theta - \beta_2/n(n-1) + \ldots \quad (\Rightarrow E(\theta^*_n - \theta) = O(n^{-2})).$$

Further let,

$$(2.7) \qquad V^*_n = \frac{1}{n-1}\sum_{i=1}^n [\hat{\theta}_{n,i} - \theta^*_n]^2 = (n-1)\sum_{i=1}^n (\hat{\theta}^i_{n-1} - \frac{1}{n}\sum_{j=1}^n \hat{\theta}^j_{n-1})^2.$$

In support of the Tukey conjecture, various authors have shown that under suitable regularity conditions,

$$(2.8) \qquad n^{\frac{1}{2}}(\theta^*_n - \theta)/[V^*_n]^{\frac{1}{2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \text{ as } n \to \infty.$$

Our contention is to obtain stronger results concerning (i) the almost sure (a.s.) convergence of V_n^* and (ii) Wiener process approximations for the tail-sequence $\{\theta_k^*-\theta;\ k \geq n\}$.

For simplicity, we assume that $p=1$ i.e., the X_i are real valued and $R = (-\infty, \infty)$. For every $n(\geq 1)$, the order statistics corresponding to $X_1,\ldots,X_n$ are denoted by $X_{n,1} \leq \ldots \leq X_{n,n}$. Let $C_n = C(X_{n,1},\ldots,X_{n,n}, X_{n+1},\ldots)$ be the σ-field generated by $(X_{n,1},\ldots,X_{n,n})$ and by X_{n+j}, $j \geq 1$. Then, C_n is non-increasing in $n(\geq 1)$. Note that given C_n, X_{n+j}, $j \geq 1$ are all held fixed while $(X_1,\ldots,X_n)$ are interchangeable and assume all possible permutations of $(X_{n,1},\ldots,X_{n,n})$ with equal conditional probability $(n!)^{-1}$. Hence,

$$(2.9) \qquad E(\hat{\theta}_{n-1}|C_n) = n^{-1}\textstyle\sum_{i=1}^{n} \hat{\theta}_{n-1}^{i} \quad \text{a.e.} ,$$

and, therefore, by (2.5) and (2.9),

$$(2.10) \quad \theta_n^* = n\hat{\theta}_n - (n-1)E(\hat{\theta}_{n-1}|C_n) = \hat{\theta}_n + (n-1)E\{\hat{\theta}_n-\hat{\theta}_{n-1}|C_n\} \quad \text{a.e.}$$

Clearly, if $\{\hat{\theta}_n, C_n\}$ is a reverse-martingale, $\theta_n^* = \hat{\theta}_n$; otherwise, the jackknifing consists in adding up the correction factor

$$(2.11) \qquad \theta_n^* - \hat{\theta}_n = (n-1)\ E\{(\hat{\theta}_n-\hat{\theta}_{n-1})|C_n\} .$$

It follows by similar arguments that

$$(2.12) \qquad V_n^* = n(n-1)\ \mathrm{Var}\{(\hat{\theta}_n-\hat{\theta}_{n-1})|C_n\}$$

$$= n(r-1)\{E[(\hat{\theta}_n-\hat{\theta}_{n-1})^2|C_n] - (E[(\hat{\theta}_n-\hat{\theta}_{n-1})|C_n])^2\} .$$

These interpretations and representations for jackknifing are quite useful for our subsequent results.

For further reduction of bias, higher order jackknife estimators have been proposed by various workers (see [6, 12]). The second order jackknife estimator (see (4.20) of [12]) can be written in our notations as

$$\theta_n^{**} = \tfrac{1}{2}\{n^2\hat{\theta}_n - 2(n-1)^2 E(\hat{\theta}_{n-1}|C_n) + (n-2)^2 E(\hat{\theta}_{n-2}|C_n)\} \tag{2.13}$$

and a similar expression holds for the higher order jackknifing. In fact, we have also a second interpretation for θ_n^*, θ_n^{**} etc. from the weighted least squares point of view. In most of the cases to follow, we shall observe that for some $\nu_1 > 0$ and real ν_2,

$$\mathrm{Var}\{n^{\frac{1}{2}}(\hat{\theta}_n - \theta)\} = \nu_1 + n^{-1}\nu_2 + 0(n^{-2}) \ , \tag{2.14}$$

$$\mathrm{Cov}\{n^{\frac{1}{2}}(\hat{\theta}_n - \theta),\ (n-1)^{\frac{1}{2}}(\hat{\theta}_{n-1} - \theta)\} = \sqrt{\frac{n-1}{n}}\,[\nu_1 + n^{-1}\nu_2 + 0(n^{-2})] \ . \tag{2.15}$$

Also, by (2.2), $E(\hat{\theta}_n - \theta) = n^{-1}\beta_1 + n^{-2}\beta_2 + \ldots$. Thus, neglecting terms of $0(n^{-2})$, the weighted least squares method of estimating $(\theta\ ,\beta_1)$ consists in minimizing

$$\begin{aligned} n^2(\hat{\theta}_n - \theta - \tfrac{1}{n}\beta_1)^2 - 2n(n-1)(\hat{\theta}_n - \theta - \tfrac{1}{n}\beta_1)(\hat{\theta}_{n-1} - \theta - \tfrac{1}{n-1}\beta_1) \\ + n(n-1)(\hat{\theta}_{n-1} - \theta - \tfrac{1}{n}\beta_1)^2 \end{aligned} \tag{2.16}$$

with respect to θ and β_1; the simultaneous equations yield

$$\hat{\theta}_\omega = n\hat{\theta}_n - (n-1)\hat{\theta}_{n-1} \ , \tag{2.17}$$

and our $\theta_n^* = E(\hat{\theta}_\omega|C_n)$. Similarly, by (2.14)-(2.15), on writing

$$Z_k = k\hat{\theta}_k - (k-1)\hat{\theta}_{k-1} \ , \ k \geq 2 \ , \tag{2.18}$$

$$\mathrm{Var}(Z_n) = \nu_1 + 0(n^{-1}) \ , \ \mathrm{Cov}(Z_n, Z_{n-1}) = 0(n^{-1}) \ . \tag{2.19}$$

Thus, Z_n, Z_{n-1} are asymptotically uncorrelated and by (2.2), $EZ_k = \theta - \beta_2/k(k-1) + O(k^{-3})$. Hence, considering

$$[(Z_n - \theta + \beta_2/n(n-1))^2 + (Z_{n-1} - \theta + \beta_2/(n-1)(n-2))^2]/\nu_1 \tag{2.20}$$

and minimizing with respect to (θ, β_2), we obtain the weighted least squares solution

$$\hat{\theta}_\omega = \tfrac{1}{2}(n^2\hat{\theta}_n - 2(n-1)^2\hat{\theta}_{n-1} + (n-2)^2\hat{\theta}_{n-2}) \;. \tag{2.21}$$

In fact, we obtain the same solution by working with $(\hat{\theta}_n, \hat{\theta}_{n-1}, \hat{\theta}_{n-2})$ and applying the weighted least squares method directly on it. Again, $\theta_n^{**} = E(\hat{\theta}_\omega | C_n)$. In general, if we want to reduce the bias to the $O(n^{-k-1})$, for some $k \geq 1$, then we need to work with $(\hat{\theta}_n, \ldots, \hat{\theta}_{n-k})$ and the k-th order jackknife estimator is the conditional expectation (given C_n) of the weighted least squares estimator of θ, neglecting β_{k+j}, $j \geq 1$ in (2.2) and terms $O(n^{-2})$ in (2.14)-(2.15). Thus, we have the following.

Theorem 2.1. *Under assumptions* (2.14)-(2.15), *the jackknife estimators (of different orders) are the conditional expectations (given* C_n*) of the weighted least squares estimators obtained from the original (biased) estimators for the successive sample sizes.*

Since in Section 3, we shall be concerned with jackknifing functions of U-statistics, we find it convenient to introduce the following notations at this stage. Let $\phi(X_1, \ldots, X_m)$, symmetric in its m arguments, be a Borel measurable *kernel* of *degree* $m(\geq 1)$ and consider the *regular functional* (estimable parameter)

$$\xi = \xi(F) = \int_{R^{pm}} \cdots \int \phi(x_1, \ldots, x_m) dF(x_1) \cdots dF(x_m) \;, \quad F \in \mathcal{F} \tag{2.22}$$

where $F = \{F: |\xi(F)| < \infty\}$. Then, for $n \geq m$, the U-statistics corresponding to ξ is defined by

(2.23) $$U_n = \binom{n}{m}^{-1} \sum_{C_{n,m}} \phi(X_{i_1}, \ldots, X_{i_m}) \; ; \; C_{n,m} = \{1 \leq i_1 < \ldots < i_m \leq n\} .$$

Note that $EU_n = \xi(F)$, $\forall\, n \geq m$. Further, let

(2.24) $$\zeta_h = \mathrm{Var}\{\phi_h(X_1, \ldots, X_h)\};$$

$$\phi_h(x_1, \ldots, x_h) = E\phi(x_1, \ldots, x_h, X_{h+1}, \ldots, X_m)$$

for $h = 0, \ldots, m$, where $\zeta_0 = 0$ and $\phi_0 = \xi$. We assume that

(2.25) $$0 < \zeta_1 , \; \zeta_m < \infty \quad (\text{where } \zeta_1 \leq m^{-1}\zeta_m) .$$

3. Invariance Principles Relating to Jackknifing U-Statistics

We shall be concerned here mainly with the following two types of estimators:

(i) Let g, defined on R, have a bounded second derivative in some neighborhood of ξ, and

(3.1) $$\hat{\theta}_n = g(U_n) , \quad \forall\, n \geq m .$$

(ii) For some positive integer q, we have

(3.2) $$\hat{\theta}_n = \sum_{s=0}^{q} \alpha_{n,s} U_n^{(s)} , \quad n \geq m ,$$

where $U_n^{(0)} = U_n$ is an unbiased estimator of $\theta = \xi(F)$,

(3.3) $$\alpha_{n,0} = 1 + n^{-1}c_{0,1} + n^{-2}c_{0,2} + O(n^{-3}) ,$$

$U_n^{(1)},\ldots,U_n^{(q)}$ are appropriate U-statistics with expectations $\theta_1,\ldots,\theta_q$ (unknown but finite) and

(3.4) $$\alpha_{n,h} = n^{-h}c_{h,0} + O(n^{-h-1}) , \quad h \geq 1 ;$$

the $c_{s,j}$ are real constants; possibly, some being equal to 0. The classical von Mises' (1947) differentiable statistical function (corresponding to $\xi(F)$) is a special case of (3.2) with $q = m$ and $c_{0,1} = -\binom{m}{2}$.

First, we consider the following.

Theorem 3.1. *For* $\{\hat{\theta}_n\}$ *defined by* (3.1) *or* (3.2)-(3.4),

(3.5) $$V_n^* \to \gamma^2 \quad a.s., \text{ as } \quad n \to \infty ,$$

where

(3.6) $$\gamma^2 = \begin{cases} [g'(\xi)]^2 m^2 \zeta_1 , & \text{for } (3.1) \\ m^2 \zeta_1 , & \text{for } (3.2) . \end{cases}$$

Proof. In the context of weak convergence of Rao-Blackwell estimator of distribution functions, Bhattacharyya and Sen (1974) have shown that under (2.25),

(3.7) $$n(n-1)E[(U_{n-1}-U_n)^2|C_n] \to m^2\zeta_1 \quad \text{a.s., as} \quad n \to \infty .$$

On the other hand, as in Section 2,

(3.8) $$n(n-1)E[(U_{n-1}-U_n)^2|C_n] = (n-1)\sum_{i=1}^n [U_{n-1}^i - U_n]^2$$

where the U_{n-1}^i are defined as in (2.3) with T_{n-1} being replaced by U_{n-1}. Hence, from (3.7) and (3.8), we obtain that

$$(3.9)\qquad \max_{1\le i\le n} (U^i_{n-1}-U_n)^2 = O(n^{-1}) \quad \text{a.s., as} \quad n\to\infty .$$

Further, $\{U_n, C_n, n\ge m\}$ is a reverse martingale, so that $U_n \to \xi(F)$ a.s., as $n\to\infty$, and hence, by (3.9),

$$(3.10)\qquad \max_{1\le i\le n} |U^i_{n-1}-\xi(F)| \to 0 \quad \text{a.s., as} \quad n\to\infty .$$

First, consider the case of (3.1). Then, we have

$$(3.11)\qquad \hat\theta_{n-1}-\hat\theta_n = g(U_{n-1})-g(U_n)$$

$$= g'(U_n)[U_{n-1}-U_n] + \tfrac{1}{2}g''(hU_n+(1-h)U_{n-1})[U_{n-1}-U_n]^2, \quad 0<h<1 .$$

Note that $E[U_{n-1}|C_n] = U_n$ a.e. and further by (3.7), (3.8), (3.10) and the boundedness of g'' (in a neighborhood of ξ), we have

$$(3.12)\qquad |E\{g''(hU_n+(1-h)U_{n-1})[U_{n-1}-U_n]^2|C_n\}|$$

$$\le \max_{1\le i\le n} |g''(hU_n+(1-h)U^i_{n-1})|\{\tfrac{1}{n}\textstyle\sum_{i=1}^n [U^i_{n-1}-U_n]^2\}$$

$$= O(n^{-2}) \quad \text{a.s., as} \quad n\to\infty .$$

Hence, we obtain from (3.11) and (3.12) that

$$(3.13)\qquad E(\hat\theta_{n-1}-\hat\theta_n|C_n) = O(n^{-2}) \quad \text{a.s., as} \quad n\to\infty .$$

Similarly,

$$(3.14)\qquad \mathrm{Var}\{(\hat\theta_{n-1}-\hat\theta_n)|C_n\} = E\{(\hat\theta_{n-1}-\hat\theta_n)^2|C_n\} + O(n^{-4}) \quad \text{a.s.}$$

$$= n^{-1}\textstyle\sum_{i=1}^n [g(U^i_{n-1}) - g(U_n)]^2 + O(n^{-4}) \quad \text{a.s., as} \quad n\to\infty .$$

Again as in (3.12), for some $0 < h < 1$

$$(3.15)\quad |n^{-1}\sum_{i=1}^{n}[g(U_{n-1}^{i})-g(U_n)]^2-[g'(U_n)]^2 n^{-1}\sum_{i=1}^{n}(U_{n-1}^{i}-U_n)^2|$$

$$\le \left\{\max_{1\le i\le n}|g'(hU_n+(1-h)U_{n-1}^{i})-g'(U_n)|^2\right\}\left\{\frac{1}{n}\sum_{i=1}^{n}[U_{n-1}^{i}-U_n]^2\right\}$$

$$=\{o(1)\text{ a.s.}\}\{O(n^{-2})\text{ a.s.}\} = o(n^{-2})\text{ a.s., as } n\to\infty,$$

where by (3.9), (3.8) and the a.s. convergence of U_n to $\xi(F)$,

$$(3.16)\quad [g'(U_n)]^2(n-1)\sum_{i=1}^{n}[U_{n-1}^{i}-U_n]^2 \to m^2\zeta_1[g'(\xi)]^2 \text{ a.s., as } n\to\infty.$$

Hence, from (3.14)-(3.16), we obtain that

$$(3.17)\qquad V_n^* = (n-1)\sum_{i=1}^{n}[g(U_{n-1}^{i}) - g(U_n)]^2$$

$$= n(n-1)\mathrm{Var}\{(\hat{\theta}_{n-1}-\hat{\theta}_n)|C_n\}$$

$$\to m^2\zeta_1[g'(\xi)]^2 = \gamma^2 \text{ a.s., as } n\to\infty.$$

For the case of (3.2), we note that

$$(3.18)\qquad \alpha_{n-1,0}U_{n-1}^{(0)}-\alpha_{n,0}U_n^{(0)}=(U_{n-1}-U_n)+c_{0,1}\{(n-1)^{-1}U_{n-1}-n^{-1}U_n\}$$

$$+O(n^{-2})\text{ a.s.},$$

$$(3.19)\quad \alpha_{n-1,h}U_{n-1}^{(h)}-\alpha_{n,h}U_n^{(h)}=\alpha_{n-1,h}(U_{n-1}^{(h)}-U_n^{(h)})+O(n^{-h-1})U_n^{(h)},\ h\ge 1,$$

and hence, the proof of (3.5) follows on parallel lines. Q.E.D.

Remark 1. From (2.11) and (3.13), we obtain that for every $\varepsilon > 0$,

$$(3.20)\qquad n^{1-\varepsilon}|\theta_n^*-\hat{\theta}\ | \to 0 \text{ a.s., as } n\to\infty.$$

The last equation is of fundamental importance to the main results of this section.

Remark 2. By virtue of (3.14)-(3.16), V_n^* is asymptotically equivalent to

$$[g'(U_n)]^2 s_n^2 \quad \text{where} \quad s_n^2 = (n-1)\sum_{i=1}^{n}[U_{n-1}^{i}-U_n]^2 \; ; \tag{3.21}$$

in case (3.2), (3.21) holds with $g'(U_n) \equiv 1$. Let us also denote by

$$V_{ni} = \binom{n-1}{m-1}^{-1}\sum_{n,i}\phi(X_i, X_{i_2}, \ldots, X_{i_m}) \, , \quad 1 \le i \le n \, , \tag{3.22}$$

where the summation $\sum_{n,i}$ extends over all $1 \le i_2 < \ldots < i_m \le n$ with $i_j \ne i$ for $2 \le j \le m$. Then, $U_n = n^{-1}\sum_{i=1}^{n} V_{ni}$. Further, let

$$V_n = (n-1)^{-1}\sum_{i=1}^{n}[V_{ni}-U_n]^2 \; . \tag{3.23}$$

Sen (1960) has shown that V_n is a distribution-free estimator of ζ_1. It is interesting to note that by definition

$$\binom{n-1}{m}U_{n-1}^{i} + \binom{n-1}{m-1}V_{ni} = \binom{n}{m}U_n \, , \quad \forall \; 1 \le i \le n \, , \tag{3.24}$$

and, as a result, it follows by routine steps that

$$s_n^2 = m^2(n-1)^2(n-m)^{-2}V_n \, , \quad \forall \; n > m \; . \tag{3.25}$$

Hence, the a.s. convergence of s_n^2 (to $m^2\zeta_1$) insures the same for V_n (to ζ_1). However, from the computational point of view, the labor involved in the computation of V_n is $O(n^m)$ whereas for s_n^2, it is $O(n^{m+1})$. Hence, V_n should be preferable to s_n^2. (3.25) will be of use in Section 5.

By virtue of (3.5) and (3.20) and the invariance principles for U-statistics, studied by Loynes (1970), Miller and Sen (1972) and Sen (1974b), we are in a position to present the following results (without derivation):

(i) Consider a sequence $\{W_n^*\}$ of stochastic processes, where

$$W_n^* = \{W_n^*(t) = n^{\frac{1}{2}}[\theta_{k_n(t)}^* - \theta]/\gamma \, , \; 0 \le t \le 1\} \, , \quad n > m \, , \tag{3.26}$$

and $k_n(t) = \min\{k: n/k \le t\}$, $0 \le t \le 1$. Note that $W_n^*(0)$ is equal to 0 with probability 1 and for every $n(\ge m)$, W_n^* belongs to the $D[0,1]$ space with which we associate the J_1-topology. Further, let $W^* = \{W^*(t), 0 \le t \le 1\}$ be a standard Brownian motion on $[0,1]$. Then, as $n \to \infty$,

$$W_n^* \xrightarrow{\mathcal{D}} W^*, \text{ in the } J_1\text{-topology on } D[0,1] . \tag{3.27}$$

Finally, if in (2.36), we replace γ^2 by V_n^* and denote the corresponding process by $\hat{W}_n^*$, then (3.27) also holds for $\{\hat{W}_n^*\}$.

(ii) Let $S = \{S(t), t \in [0,\infty)\}$ be a random process defined by

$$S(t) = \begin{cases} 0, & 0 \le t < m+1 , \\ k[\theta_k^* - \theta]/\gamma, & k \le t < k+1,\ k \ge m+1 , \end{cases} \tag{3.28}$$

and let $W = \{W(t), t \in [0,\infty)\}$ be a standard Wiener process on $[0,\infty)$. Further, we assume that for some $r > 2$

$$E|\phi(X_1,\ldots,X_m)|^r < \infty . \tag{3.29}$$

Then, we have the following

$$S(t) = W(t) + o(t^{\frac{1}{2}}) \text{ a.s., as } t \to \infty . \tag{3.30}$$

(iii) In (3.1), we have considered $\hat{\theta}_n = g(U_n)$. It is possible to take $\hat{\theta}_n = g(U_n^{(1)},\ldots,U_n^{(k)})$, for some $k \ge 1$, where g has bounded second order partial derivatives in a neighborhood of the point $(EU_n^{(1)},\ldots,EU_n^{(k)})$ $(\in R^k)$. The proof follows as a straight-forward extension of what has been done before, and hence, for intended brevity, the details are omitted.

(iv) Jackknifing functions of generalized U-statistics has been considered by Arvesen (1969). Here also, as in Sen (1974c), we may consider the product sigma-field formed by the individual sample sequence $\{C_n\}$ and express the usual jackknife estimator as the conditional expectation of a linear combination of original estimators for adjacent sample sizes. Further, a result parallel to (3.20) holds in this case. Hence, by virtue of Theorem 2.2 of Sen (1974a), we are in a position to derive a similar invariance principle for the jackknife estimators. Further, by virtue of (3.19)-(3.23) of Sen (1974a), it can be shown that (3.30) extends to a multi-parameter Gaussian process. For intended brevity, the details are omitted again.

4. Invariance Principles for General $\{\theta_n^*\}$

Structural properties of U-statistics have enabled us to study the invariance principles in Section 3 without having any extra regularity conditions. If $\hat{\theta}_n$ is not a function of U-statistics, we need, however, a few extra regularity conditions to derive similar results. These will be studied here.

Concerning the original sequence of estimators $\{\hat{\theta}_n\}$, we assume that

(4.1) $$\hat{\theta}_n \to \theta \text{ a.s., as } n \to \infty ,$$

(4.2) $$\delta_n^2 = \text{Var}(\hat{\theta}_n) \downarrow 0 \text{ as } n \to \infty ; \lim_{n\to\infty} n\delta_n^2 = \delta^2 , \quad 0 < \delta < \infty .$$

Let use also define

(4.3) $$Y_n = n(n-1)(\hat{\theta}_{n-1}-\hat{\theta}_n)^2 , \quad n \geq 2 ,$$

and assume that

(4.4) $$E[Y_n | C_n] \to \delta^2 \text{ a.s., as } n \to \infty ,$$

(4.5) $$Y_n \text{ is uniformly (in } n\text{) integrable },$$

(4.6) $$\sum_{n\geq N} |E(\hat{\theta}_n - \hat{\theta}_{n+1} | C_{n+1})| = o(N^{-\frac{1}{2}}) \quad \text{as} \quad N \to \infty ,$$

Consider now a sequence of stochastic processes $\{W_n\}$, where

(4.7) $$W_n = \{W_n(t) = \delta_n^{-1}(\hat{\theta}_{k_n(t)} - \theta) , \quad 0 \leq t \leq 1\} ;$$

(4.8) $$k_n(t) = \min\{k: \delta_k^2/\delta_n^2 \leq t\} , \quad 0 \leq t \leq 1 .$$

Then, we have the following.

Theorem 4.1. *Under the assumptions made above, as* $n \to \infty$

(4.9) $$W_n \overset{\mathcal{D}}{\to} W^*, \quad \textit{in the } J_1\textit{-topology on } D[0,1] ,$$

where W^* *is a standard Brownian motion on* $[0,1]$.

Outline of the proof. Let us write $Q_k = \hat{\theta}_k - \hat{\theta}_{k+1}$, $k \geq 1$. Then, by (4.1),

(4.10) $$\hat{\theta}_N - \theta = \sum_{k\geq N} Q_k \quad \text{a.s.,} \quad \forall\, N \geq 1 .$$

Also, by (4.2), (4.4) and (4.8), for every $t \in [0,1]$

(4.11) $$\delta_n^{-2} \sum_{k\geq k_n(t)} E[Q_k^2 | C_{k+1}] \to t \quad \text{a.s., as} \quad n \to \infty .$$

Further, by (4.2) and (4.6),

(4.12) $$\delta_n^{-1} \sum_{k\geq n} |E(Q_k | C_{k+1})| \to 0 \quad \text{a.s., as} \quad n \to \infty .$$

Finally, by (4.5), for every $\varepsilon > 0$,

$$(4.13)\qquad \delta_n^{-2}\sum_{k\geq n} E\{Q_k^2 I(Q_k^2 > \varepsilon\delta_n^2)\}$$

$$= \delta_n^{-2}\sum_{k\geq n} \frac{1}{k(k+1)} E\{Y_k^2 I(Y_k^2 > \varepsilon k(k+1)\sigma_n^2)\}$$

$$= (\delta_n^{-2})(o(1))(\sum_{k\geq n} \frac{1}{k(k+1)}) = o(1) ,$$

and hence

$$(4.14)\qquad \delta_n^{-2}\sum_{k\geq n} E\{Q_k^2 I(Q_k^2 > \varepsilon\delta_n^2)|C_{k+1}\} \overset{p}{\to} 0 \quad \text{as} \quad n \to \infty .$$

Reversing now the order of the index set $\{k: k \geq n\}$ (to $\{k: -\infty < k \leq -n\}$), the rest of the proof follows by an appeal to the general functional central limit theorem of McLeish (1974) where (4.11), (4.12) and (4.14) insure the satisfaction of his underlying regularity conditions. Q.E.D.

Since (4.4) corresponds to (3.7), virtually repeating the proof of Theorem 3.1, it follows that under the same conditions on g, as in Section 3,

$$(4.15)\qquad |\theta_n^* - \hat{\theta}_n| = (n-1)|E\{(\hat{\theta}_{n-1} - \hat{\theta}_n)|C_n\}| = o(n^{-\frac{1}{2}}) \quad \text{a.s.} ,$$

by (4.6), and

$$(4.16)\qquad V_n^* \to \delta^2 \quad \text{a.s., as} \quad n \to \infty .$$

Hence, if in (4.7), we replace $\{\hat{\theta}_k\}$ by $\{\theta_k^*\}$ and denote the corresponding process by W_n^*, then (4.9) holds for $\{W_n^*\}$ as well. Further, δ_n^{-1} may also be replaced by $n^{\frac{1}{2}}(V_n^*)^{-\frac{1}{2}}$.

The conditions (4.1), (4.2), (4.4), (4.5) and (4.6) are most conveniently verifiable if $\hat{\theta}_n$ can be expressed as

$$(4.17)\qquad \hat{\theta}_n = m_n + r_n ,$$

where $\{m_n, C_n\}$ is a reverse martingale and $|r_{n-1} - r_n| = o(n^{-3/2})$ a.s., as $n \to \infty$.

5. Asymptotic Sequential Confidence Intervals Based on Jackknifing

Tukey proposed the use of (2.8) for a robust confidence intervals for θ. By virtue of our invariance principles, we are in a position to consider the following robust sequential interval estimation problem.

Suppose θ, γ^2, $\hat{\theta}_n$, θ_n^* and V_n^* are defined as before. The underlying df F, and hence, θ and γ^2 being unknown, it is desired to determine (sequentially) a confidence interval for θ having a maximum-width $2d$, $d > 0$ being predetermined, and a preassigned confidence coefficient $1 - \alpha$, $0 < \alpha < 1$. For every $n \geq 1$ and $d > 0$, let

$$I_n(d) = \{\theta\colon \theta_n^* - d \leq \theta \leq \theta_n^* + d\}\ , \tag{5.1}$$

and let τ_α be the upper $100\alpha\%$ point of the standard normal df. Finally, let $n_0(=n_0(d))$ be the initial sample size. Then, we consider a *stopping variable* $N(=N(d))$, defined by

$$N = \text{Smallest integer } n(\geq n_0) \text{ such that } V_n^* \leq nd^2/\tau_{\alpha/2}^2\ ; \tag{5.2}$$

if no such n exist, we let $N = \infty$. Whenever $N < \infty$, the proposed confidence interval for θ is $I_N(d)$, defined by (5.1) for $n = N = N(d)$. The above procedure is a direct adaptation of the Chow-Robbins (1965) procedure under our jackknifing setup.

Theorem 5.1. *Under the hypothesis of Theorem* 3.1 (*or* 4.1),

$$\lim_{d\to 0} P\{\theta \in I_{N(d)}(d)\} = 1 - \alpha\ , \tag{5.3}$$

$$\lim_{d\to 0}\{(N(d)d^2)/(\tau_{\alpha/2}^2\gamma^2)\} = 1 \quad \text{a.s.} \tag{5.4}$$

If, in addition $E\{\sup_n V_n^*\} < \infty$, *then*

(5.5) $$\lim_{d\to 0}\{(d^2 EN(d)/(\tau^2_{\alpha/2}\gamma^2)\} = 1 \ .$$

Outline of the proof. We follow the line of attack of Chow and Robbins (1965). We need to show that (a) $V^*_n \to \gamma^2$ a.s., as $n \to \infty$, (b) (2.8) holds and (c) for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta (0 < \delta < 1)$ and an n^*, such that for $n \geq n^*$,

(5.6) $$P\left\{\max_{n':|n-n'|\leq\delta n} n^{\frac{1}{2}}|\theta^*_{n'}-\theta^*_n| > \varepsilon\right\} < \eta \ .$$

Now (a) has already been proved, (b) is a direct consequence of (3.27) and finally (c) follows from the *tightness* property of $\{W^*_n\}$ which, in turn, is insured by (3.27). Q.E.D.

The condition that $E\left(\sup_n V^*_n\right) < \infty$, needed for (5.5), however, does not follow from the hypothesis of Theorem 3.1 (or 4.1); nor it is a very readily verifiable one. It is possible to obtain (5.5) under a somewhat different condition which is more easily verifiable.

Theorem 5.2. *If* $\{\hat{\theta}_n\}$ *is defined by* (3.1) *or* (3.2)-(3.4) *and the hypothesis of Theorem* 3.1 *holds, then* $E\{|\phi(X_1,\ldots,X_m)|^r\} < \infty$ *for some* $r > 4$, *insures* (5.5).

Proof. Consider the estimator V_n, defined by (3.23). It follows from Sproule (1969) that V_n can be expressed as a linear combination of several U-statistics whose moments of the order $q(>0)$ exists whenever $E|\phi|^{2q} < \infty$. As such, using Theorem 1 of Sen (1974c), it follows that for every $\varepsilon > 0$, there exist a positive $K_\varepsilon(< \infty)$ and an $n_0(\varepsilon)$, such for $n \geq n_0(\varepsilon)$,

(5.7) $$P\{|V_n-\zeta_1| > \varepsilon/2\} \leq K_\varepsilon n^{-s}, \ s = r/2 > 1 \ .$$

Further, $g'(t)$ has a bounded derivative in a neighborhood of $t=\xi$, and hence, by Theorem 1 of Sen (1974c), again,

$$P\{|g'(U_n)-g'(\xi)|>\varepsilon/2\} \le K_\varepsilon n^{-r}, \ \forall\ n \ge n_0(\varepsilon) \ . \tag{5.8}$$

From (3.15), (3.21), (3.25), (5.7) and (5.8), it follows by some standard steps that for every $\eta>0$, there exist a constant $K_\eta(<\infty)$ and an $n_0(\eta)$, such that for $n \ge n_0(\eta)$,

$$P\{|V_n^*-\gamma^2|>\eta\} \le K_\eta n^{-s}; \ s = r/2 > 1 \ . \tag{5.9}$$

Having established this, we may proceed as in the proof of Theorem 3.1 of Sen and Ghosh (1971) [namely, as in their (5.16)-(5.19)], and complete the proof of (5.5) by using (5.3) and (5.9). For brevity, the details are omitted. Q.E.D.

6. Sequential Tests based on Jackknife Estimators

The embedding of Wiener processes in (3.30) and the strong convergence of V_n^* (to γ^2) in (3.5) enable us to construct the following type of asymptotic sequential tests; for further motivation of this type of procedures, we may refer to Sen (1973) and Sen and Ghosh (1974).

Consider a suitable parameter θ (for which the sequences $\{\hat{\theta}_n\}$ and $\{\theta_n^*\}$ of estimators are available sequentially), and suppose that we desire to test

$$H_0: \theta = \theta_0 \ \text{ vs. } \ H_1: \theta = \theta_1 = \theta_0 + \Delta, \ \Delta > 0 \ , \tag{6.1}$$

where θ_0 and Δ are specified and we like the test to have the prescribed strength (α,β). Since the df F is not known, no fixed sample size test sounds feasible and we take recourse to the following sequential procedure:

Suppose that $0 < \alpha, \beta < \frac{1}{2}$ and consider two positive numbers (A,B): $0 < B < 1 < A < \infty$, where $\beta/(1-\alpha) \le B$ and $(1-\beta)/\alpha \ge A$. Starting with an initial sample of size $n_0 (= n_0(\Delta))$, continue drawing observations one by one as long as

$$(6.2) \qquad bV_m^* < m\Delta[\theta_m^* - (\theta_0+\theta_1)/2] < aV_m^*, \quad m \ge n_0(\Delta) \ ,$$

where $a = \log A$, $b = \log B$ ($\Longrightarrow -\infty < b < 0 < a < \infty$), θ_m^* is the jack-knife estimator of θ based on $X_1,\ldots,X_m$ [viz., (2.5)] and V_m^* is defined as in (2.7). If, for the first time, (6.2) is vitiated for $m = N$ and $\Delta[\theta_N^* - (\theta_0+\theta_1)/2]$ is $\le bV^*$ (or $\ge aV_N^*$), accept H_0 (or H_1); the *stopping variable* N is denoted by $N(\Delta)$.

Since $m^{-\frac{1}{2}}V_m^* \to 0$ a.s., as $m \to \infty$ (by (3.5)) and $m^{\frac{1}{2}}(\theta_m^* - \theta)$ is asymptotically normal with mean 0 and variance γ^2, it is easy to see that for every fixed θ and Δ,

$$(6.3) \qquad P_\theta\{N(\Delta) > n\} \le P\{n^{-\frac{1}{2}}bV_n^* < \Delta n^{\frac{1}{2}}[\theta_n^* - \tfrac{1}{2}(\theta_0+\theta_1)] < n^{-\frac{1}{2}}aV_n^*\}$$
$$\to 0 \text{ as } n \to \infty \ ,$$

and hence, the proposed test terminates with probability one. For the OC and ASN function, as in Sen (1973) and Sen and Ghosh (1974), we consider the asymptotic situation where we let $\Delta \to 0$ (comparable to $d \to 0$ in Section 5) and set

$$(6.4) \qquad \theta = \theta_0 + \phi\Delta \quad \text{where} \quad \phi \in \Phi = \{\phi\colon |\phi| < K < \infty\} \ ,$$

$$(6.5) \qquad \lim_{\Delta\to 0} n_0(\Delta) = \infty \quad \text{but} \quad \lim_{\Delta\to 0} \Delta^2 n_0(\Delta) = 0 \ ,$$

(6.6) $$e^{a} = A = (1-\beta)/\alpha\ , \quad e^{b} = B = \beta/(1-\alpha)\ , \quad 0 < \alpha\ \beta < \tfrac{1}{2}\ .$$

Finally, let us denote by $L_F(\phi,\Delta)$ the OC (i.e. probability of accepting H_0 when actually $\theta = \theta_0 + \phi\Delta$) of the test based on (6.2). Then, we have the following:

<u>Theorem 6.1</u>. *Under* (6.4)-(6.6) *and the hypothesis of Theorem* 3.1 (*or* 4.1)

(6.7) $$\lim_{\Delta\to 0} L_F(\phi,\Delta) = P(\phi) = \begin{cases} (A^{1-2\phi}-1)/(A^{-12\phi}-B^{1-2\phi})\ , & \phi \neq \tfrac{1}{2}\ , \\ a/(a-b)\ , & \phi = \tfrac{1}{2}\ , \end{cases}$$

and hence, asymptotically the OC *does not depend on* F. *Further*

(6.8) $$P(0) = 1-\alpha \quad \text{and} \quad P(1) = \beta\ ,$$

so that the test has asymptotic strength (α,β).

<u>Proof</u>. Let us choose a sequence $\{n^* = n^*(\Delta)\}$ such that

(6.9) $n^*(\Delta) \sim K\Delta^{-2}$ as $\Delta \to 0$, where $K(<\infty)$ is arbitrarily large .

Then, by (3.27), (6.4), (6.5) and (6.9), for $\theta = \theta_0 + \phi\Delta$, defining $U^*_\Delta = \{U_\Delta(t) = \Delta n[\theta^*_n - \tfrac{1}{2}(\theta_0+\theta_1)]/\gamma\ ,\ \Delta^2 n \le t < \Delta^2(n+1)\ ,\ n_0(\Delta) \le n < n^*(\Delta)\}$, it follows that as $\Delta \to 0$,

(6.10) $$U^*_\Delta \overset{\mathcal{D}}{\to} \{W(t) + (\phi-\tfrac{1}{2})t/\gamma\ ,\ 0 < t \le K\}$$

where $\{W(t)\ ,\ t > 0\}$ is a standard Wiener process on $[0,\infty)$. Also, by (3.5), $\sup\{|V^*_n/\gamma^2 - 1| : n_0(\Delta) \le n \le n^*(\Delta)\} \overset{p}{\to} 0$ as $\Delta \to 0$. Finally, by (6.3), for every $\eta > 0$, there exists a $K = K_\eta\ (< \infty)$, such that defining $n^*(\Delta)$ by (6.9) with $K = K_\eta$, we have $P\{N(\Delta) > n^*(\Delta)\} < \eta$. Hence, using (6.10), (6.2) and the classical result of Dvoretzky, Kiefer and Wolfowitz (1953) on the boundary crossing probability for Wiener processes, (6.7) follows readily.

Finally, (6.8) follows from (6.7) by substituting $\phi = 0$ and 1. Q.E.D.

As in Theorem 5.2, for the study of the ASN (i.e., $E\{N(\Delta)|\theta = \theta_0 + \phi\Delta\}$ for $\Delta \to 0$) function, the weak (or a.s.) convergence results of Section 3 (or 4) are not enough and we need some analogous moment convergence results which, in turn, may demand more restrictive conditions on the df F. Suppose that as in Theorem 5.2, we assume $\hat{\theta}_n = g(U_n)$ and that

(6.11) $$E|\phi(X_1,\ldots,X_m)|^r < \infty \quad \text{for some} \quad r > 4 .$$

Then, not only, we have (5.9), but also, it can be shown by steps similar to those in Section 3 that

(6.12) $$P\{|\theta_n^* - \theta| > \varepsilon\} \le C_\varepsilon n^{-s}, \quad s > 1 ,$$

for n sufficiently large. Further, for $n_0(\Delta) \le n \le n^*(\Delta)$, we may write

(6.13) $$n\Delta[\theta_n^* - \tfrac{1}{2}(\theta_0+\theta_1)] = \Delta Z_n + n\Delta^2(\phi - \tfrac{1}{2}) + R_n^\Delta ,$$

where for every $\varepsilon > 0$,

(6.14) $$P\left\{\max_{n_0(\Delta)\le n\le n^*(\Delta)} |R_n^\Delta| > \varepsilon\right\} \le C_\varepsilon n^{-s}, \quad \text{for} \ \forall\ \Delta:\ 0 < \Delta \le \Delta_0 ,$$

and where $\{Z_n, \mathcal{B}_n;\ n \ge 1\}$ is a martingale; $\mathcal{B}_n$ being the σ-field generated by $X_1,\ldots,X_n$, $n \ge 1$ ($\Rightarrow \mathcal{B}_n$ is $\nearrow$ in n). As such the method of attack of Sen (1973) and Ghosh and Sen (1976) can directly be adapted to arrive at the following.

Theorem 6.2. *Under the assumptions made earlier, for every* $\phi \in \Phi$,

(6.15) $$\lim_{\Delta\to 0}\{\Delta^2 E[N(\Delta)|\theta = \theta_0 + \phi\Delta]\} = \psi(\phi,\gamma)$$

where

(6.16)
$$\psi(\phi,\gamma) = \begin{cases} \{bP(\phi)+a[1-P(\phi)]\}\{\gamma^2/(\phi-\tfrac{1}{2})\}\ , & \phi \neq \tfrac{1}{2} \\ -\gamma^2 ab\ , & \phi = \tfrac{1}{2} \end{cases}$$

and γ^2 *is defined by* (3.6).

REFERENCES

[1] Arvesen, J.M. (1969). Jackknifing U-statistics. *Ann. Math. Statist.* 40, 2076-2100.

[2] Bhattacharyya, B.B. and Sen, P.K. (1974). Weak convergence of Rao-Blackwell estimator of distribution functions. *Inst. Statist. Univ. North Carolina Mimeo Report No.* 942.

[3] Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential intervals for the mean. *Ann. Math. Statist.* 36, 457-462.

[4] Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1953). Sequential decision problems for proceeses with continuous time parameter: testing hypotheses. *Ann. Math. Statist.* 24, 254-264.

[5] Ghosh, M. and Sen, P.K. (1976). Sequential rank tests for regression. *Sankhyā, Ser. A.* (to appear).

[6] Gray, H.L., Watkins, T.A. and Adams, J.E. (1972). On the jackknife statistic, its extensions and its relation to e_n-transformations. *Ann. Math. Statist.* 43, 1-30.

[7] Loynes, R.M. (1970). An invariance principle for reverse martingales. *Proc. Amer. Math. Soc.* 25, 56-64.

[8] McLeish, D.L. (1974). Dependent central limit theorems and invariance principles. *Ann. Prob.* 2, 620-628.

[9] Miller, R.G., Jr. (1974). An unbalanced jackknife. *Ann. Statist.* 2, 880-891.

[10] Miller, R.G., Jr. and Sen, P.K. (1972). Weak convergence of U-statistics and von Mises' differentiable statistical functions. *Ann. Math. Statist.* 43, 31-41.

[11] Mises, R.V. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* 18, 309-348.

[12] Schucany, W.R., Gray, H.L. and Owen, D.B. (1971). Bias reduction in estimation. *J. Amer. Statist. Assoc.* 66, 524-533.

[13] Sen, P.K. (1960). On some convergence properties of U-statistics. *Calcutta Statist. Assoc. Bull.* 10, 1-18.

[14] Sen, P.K. (1973). Asymptotic sequential tests for regular functionals of distribution functions. *Theory Probability Appl.* 18, 235-249.

[15] Sen, P.K. (1974a). Weak convergence of generalized U-statistics. *Ann. Probability* 2, 90-102.

[16] Sen, P.K. (1974b). Almost sure behavior of U-statistics and von Mises' differentiable statistical functions. *Ann. Statist* 2, 837-395.

[17] Sen, P.K. (1974c). On L^p-convergence of U-statistics. *Ann. Inst. Statist. Math.* 26, 55-60.

[18] Sen, P.K. and Ghosh, M. (1971). On bounded length sequential confidence intervals based on one-sample rank order statistics. *Ann. Math. Statist.* 42, 189-203.

[19] Sen, P.K. and Ghosh, M. (1974). Sequential rank tests for location. *Ann. Statist.* 2, 540-552.

[20] Sproule, R.N. (1969). A sequential fixed-width confidence interval for the mean of a U-statistic. Ph.D. dissertation, Univ. North Carolina, Chapel Hill.