The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 2076–2100

JACKKNIFING *U*-STATISTICS[1]

BY JAMES N. ARVESEN

Purdue University

1. Introduction. Several recent papers have shown the versatility of a relatively new estimation procedure. The *jackknife* was originally proposed by Quenouille in [17], and later expanded on by the same author in [18]. Subsequently, Tukey [15] proposed the use of it to obtain approximate t-statistics to be used in testing or construction of confidence intervals. Two papers of Miller [13], and [14] are rigorous justifications of situations where Tukey's method proves valid (as well as situations where it is grossly invalid).

Let θ be an unknown parameter, and let $X_1, \cdots, X_N$ be N independent, identically distributed observations from the cdf F_θ. The essence of the jackknife is to divide the N observations into n groups of k observations each $(N = nk)$. Let $\hat{\theta}_n{}^0$ be the estimate of θ based on all N observations, and $\hat{\theta}_{n-1}^i$, $i = 1, \cdots, n$, denote the estimate obtained after deletion of the ith group of observations.

Let

$$(1) \qquad \hat{\theta}_i = n\hat{\theta}_n{}^0 - (n-1)\hat{\theta}_{n-1}^i, \qquad i = 1, \cdots, n.$$

These are called pseudo-values by Tukey. Then the jackknife estimate of θ is the average of the $\hat{\theta}_i$, $i = 1, \cdots, n$,

$$(2) \qquad \hat{\theta} = n^{-1} \sum_{i=1}^n \hat{\theta}_i .$$

When originally proposed, Quenouille considered $n = 2$, and found that the technique eliminated the $1/N$ term from any bias. This result holds for all values of n, for if

$$(3) \qquad E(\hat{\theta}_n{}^0) = \theta + a/kn + b/(kn)^2 + \cdots$$

one can show that

$$(4) \qquad E(\hat{\theta}) = \theta - b/[k^2 n(n-1)] + \cdots .$$

Perhaps more importantly, Tukey has proposed that in most instances $\hat{\theta}_1, \cdots, \hat{\theta}_n$ can be treated as n approximately *independent*, identically distributed observations from which an approximate t_{n-1} statistic can be constructed (note that if $X_1, \cdots, X_N$ are independent, identically distributed random variables, then $\hat{\theta}_1, \cdots, \hat{\theta}_n$ are interchangeable random variables for each n). Equivalently, Tukey conjectured that

$$(5) \qquad n^{\frac{1}{2}}(\hat{\theta} - \theta)/((n-1)^{-1} \sum_{i=1}^n (\hat{\theta}_i - \hat{\theta})^2)^{\frac{1}{2}}$$

Received 3 February 1969.

[1] Adapted from the author's doctoral dissertation at Stanford University. The research was supported in part by Public Health Service Grant USPHS-2T1 GM 25-11.

is approximately distributed as a t random variable with $n - 1$ degrees of freedom.

One might feel that a reduction of bias as indicated in (4) might be accompanied by an increase in the mean square error of the estimate. However, in [3], Durbin exhibited a class of problems in ratio estimation where this is not the case. Both the bias and mean square error were shown to decrease under a reasonable model.

What follows is essentially an application of Tukey's idea. The original motivation was to obtain robust procedures in Model II ANOVA. However, it was found this could best be done by studying U-statistics (see Hoeffding [6]), and functions of several U-statistics. This more general setting leads to other results as well.

In Section 2, the jackknife technique will be applied to U-statistics and functions of several U-statistics. With suitable regularity conditions, it will be shown that Miller's results of [13] can be extended to these cases. In Section 3, these results are applied to obtaining robust procedures in a one-way Model II ANOVA layout. Also included are applications to ratio problems, and correlation coefficients. Section 4 considers extensions to two-sample problems. Some previous work of Cornfield and Tukey [2], and Hooke [7] and [8], dealing with two-way Model II ANOVA layouts is also examined. It is shown that the jackknife procedure can be profitably used in the two-way layout.

2. Jackknifing U-statistics.

(a) *Functions of means.* In the construction of the jackknife estimate, it is necessary to divide the original N observations into n groups each containing k elements ($N = nk$). Temporarily, we will assume $k = 1$, and hence $N = n$.

Recall the definition of the jackknife estimate of a parameter as given in the introduction. Consider the case where $\hat{\theta}_n^0$ is an unbiased estimate of a parameter θ, and $g(\hat{\theta}_n^0)$ is a transformation of the statistic $\hat{\theta}_n^0$. Then in general, $g(\hat{\theta}_n^0)$ is a biased estimate of $g(\theta)$ due to the nonlinearity of the transformation. In many cases, it will be biased of order $1/n$. Hence the jackknife might lend itself to consideration in this case.

Let $\hat{\theta}_i = ng(\bar{X}) - (n-1)g(\bar{X}^i)$, $X = n^{-1}\sum_{j=1}^n X_j$, $\bar{X}^i = (n-1)^{-1}\sum_{j\neq i} X_j$ for $i = 1, \cdots, n$, $\hat{\theta} = n^{-1}\sum_{i=1}^n \hat{\theta}_i$, and $\theta = g(\mu)$. Then Miller [13] proved the following two theorems.

THEOREM 1. *Let $\{X_i\}$ be a sequence of independent, identically distributed random variables with mean μ and variance $0 < \sigma^2 < +\infty$. Let g be a function defined on the real line, which in neighborhood of μ has a bounded second derivative. Then as $n \to \infty$, $n^{\frac{1}{2}}(\hat{\theta} - \theta)$ is asymptotically normally distributed with mean zero and variance $\sigma^2(g'(\mu))^2$.*

THEOREM 2. *Let $\{X_i\}$ be a sequence of independent, identically distributed random variables with mean μ, variance $0 < \sigma^2 < +\infty$. Let g be a function with a continuous first derivative near μ. Then as $n \to \infty$, $s_g^2 = (n-1)^{-1}\sum_{i=1}^n(\hat{\theta}_i - \hat{\theta})^2 \to_P \sigma^2(g'(\mu))^2$.*

Note the stronger assumption on g in Theorem 1. The unjackknifed estimate has an asymptotic normal distribution with only the assumptions on g of Theorem 2. Also note that the stronger assumptions of Theorem 1 yield Tukey's conjecture (5).

(b) *U-statistics.* Miller's two theorems can be extended to the consideration of U-statistics. Let $X_1, \cdots, X_n$ be independent identically distributed random variables, and $f(X_1, \cdots, X_m)$ be an unbiased estimate of some parameter η, where m is the smallest number of observations needed to estimate η. Then there exists a symmetric form of $f(X_1, \cdots, X_m)$, given by

$$f^*(X_1, \cdots, X_m) = (m!)^{-1} \textstyle\sum_{P_m} f(X_{\beta_1}, \cdots, X_{\beta_m}) \tag{6}$$

where P_m indicates that the sum is over the $m!$ permutations of the subscripts.

Then the U-statistic for the parameter n can be written in the form

$$U(X_1, \cdots, X_n) = \binom{n}{m}^{-1} \textstyle\sum_{C_n} f^*(X_{\alpha_1}, \cdots, X_{\alpha_m}) \tag{7}$$

where C_n indicates that the summation is over all combinations $\alpha_1, \cdots, \alpha_m$ of m integers chosen from $1, \cdots, n$.

A U-statistic is unbiased, and hence,

$$E\{U(X_1, \cdots, X_n)\} = \eta. \tag{8}$$

Let $f_c^*(x_1, \cdots, x_c) = E\{f^*(X_1, \cdots, X_c, X_{c+1}, \cdots, X_m) \mid X_1 = x_1, \cdots, X_c = x_c\}$ and let

$$\zeta_c = \text{Var}\,\{f_c^*(X_1, \cdots, X_c)\}, c = 1, \cdots, m, \qquad \zeta_0 = 0. \tag{9}$$

Then Hoeffding [6] shows:

THEOREM 3. *Let $X_1, \cdots, X_n$ be independent identically distributed real random variables. If $f^*(X_1, \cdots, X_m)$ is a real-valued symmetric statistic with expectation η and finite second moment $E[f^*(X_1, \cdots, X_m)]^2 < \infty$, then as $n \to \infty$, the limiting distribution of $n^{\frac{1}{2}}[U - \eta]$ is normal with mean zero and variance $m^2\zeta_1$.*

As the following theorem indicates, one is also able to obtain a.s. convergence of U-statistics.

THEOREM 4. *Let $X_1, \cdots, X_n, \cdots$ be independent identically distributed real random variables. If $f^*(X_1, \cdots, X_m)$ is a real-valued symmetric statistic with expectation η and $E|f^*(X_1, \cdots, X_m)| < \infty$, then $U X_1, \cdots, X_n) \to_{\text{a.s.}L_1} \eta$ as $n \to \infty$.*

PROOF. The proof is given in Berk [1], and is given here for future reference.

Let $Y_n = \{X_{(1)}, \cdots, X_{(n)}\}$ be the set of order statistics, and let $\mathfrak{F}_n = \mathcal{B}(Y_n, X_{n+1}, X_{n+2}, \cdots)$ be the σ-field generated by $Y_n, X_{n+1}, X_{n+2}, \cdots$. Letting $Z_n = E[f^*(X_1, \cdots, X_m) \mid \mathfrak{F}_n] = U(X_1, \cdots, X_n)$ a.s., one obtains

$$E[Z_n \mid \mathfrak{F}_{n+1}] = E[f^*(X_1, \cdots, X_m) \mid \mathfrak{F}_{n+1}] = Z_{n+1} \text{ a.s.}$$

since $\mathfrak{F}_n \supset \mathfrak{F}_{n+1}$. Hence $\{Z_n, \mathfrak{F}_n\}$ is a reverse martingale, and by the martingale convergence theorem, there exists Z_∞ such that $U(X_1, \cdots, X_n) = Z_n \to_{\text{a.s.}L_1} Z_\infty$.

Given $\epsilon > 0$, let

$$A = [U(X_1, \cdots, X_n) > \eta + \epsilon \text{ infinitely often}].$$

Since any finite permutation of the sequence $X_1, \cdots, X_n, \cdots$ leaves the event A invariant, by the Hewitt-Savage theorem, $P(A) = 0$ or 1. Since $E(Z_\infty) = E(Z_m) = \eta$, $P(A) = 0$. A similar argument works with $B = [U(X_1, \cdots, X_n) < \eta - \epsilon$ infinitely often] to yield the result.

Theorems 1 and 2 can be extended to U-statistics. First, the following notation will be helpful:

$$U_n^{\,0} = \binom{n}{m}^{-1} \sum\nolimits_{C_n} f^*(X_{\alpha_1}, \cdots, X_{\alpha_m}),$$
$$U_{n-1}^i = \binom{n-1}{m}^{-1} \sum\nolimits_{C_{n-1}^i} f^*(X_{\beta_1^i}, \cdots, X_{\beta_m^i}), \tag{10}$$

where C_n is as in (7) and C_{n-1}^i indicates that the summation is over all combinations $(\beta_1^{\,i}, \cdots, \beta_m^{\,i})$ of m integers chosen from $(1, \cdots, i-1, i+1, \cdots, n)$, and $f^*(X_1, \cdots, X_m)$ is as in (6). Let g be a real-valued function, and

$$\hat\theta_n^{\,0} = g(U_n^{\,0}), \qquad \hat\theta_{n-1}^i = g(U_{n-1}^i),$$
$$\hat\theta_i = n\hat\theta_n^{\,0} - (n-1)\hat\theta_{n-1}^i \quad \text{for} \quad i = 1, \cdots, n, \tag{11}$$
$$\hat\theta = n^{-1} \sum\nolimits_{i=1}^n \hat\theta_i, \qquad \theta = g(\eta),$$
$$s_g^{\,2} = (n-1)^{-1} \sum\nolimits_{i=1}^n (\hat\theta_i - \hat\theta)^2.$$

THEOREM 5. *Let $X_1, \cdots, X_n$ be independent, identically distributed random variables, and let $f^*(X_1, \cdots, X_m)$ be a real-valued symmetric statistic with expectation η, and finite second moment $E[f^*(X_1, \cdots, X_m)]^2 < +\infty$. Let g be a function defined on the real line, which in a neighborhood of η has a bounded second derivative. Then as $n \to \infty$, $n^{\frac12}(\hat\theta - \theta) \to_{\mathcal{L}} \mathfrak{N}(0, m^2\zeta_1[g'(\eta)]^2)$.*

PROOF. The proof follows that of Miller [13]. Without loss of generality, let $\eta = 0$. Let $I = (-3\Delta, +3\Delta)$, $\Delta > 0$, be any neighborhood of zero in which g'' is bounded. Then, as $n \to \infty$, $U_n^{\,0} \to_P 0$, and hence

$$\Pr\{U_n^{\,0} \,\varepsilon\, (-\Delta, \Delta)\} \to 1. \tag{12}$$

Note that

$$U_{n-1}^i = \binom{n}{m}\binom{n-1}{m}^{-1} U_n^{\,0} - \binom{n-1}{m}^{-1} \sum\nolimits_{D_{n-1}^i} f^*(X_i, X_{\alpha_1^i}, \cdots, X_{\alpha_{m-1}^i}) \tag{13}$$

where $\sum_{D_{n-1}^i}$ indicates that the sum is over all combinations of $m-1$ integers $(\alpha_1^{\,i}, \cdots, \alpha_{m-1}^i)$ chosen from $(1, \cdots, i-1, i+1, \cdots, n)$. Let $Y_i = \binom{n-1}{m-1}^{-1} \cdot \sum_{D_{n-1}^i} f^*(X_i, X_{\alpha_1^i}, \cdots, X_{\alpha_{m-1}^i})$, $i = 1, \cdots, n$. First note that for n sufficiently large, $\operatorname{Var} Y_i \leqq \zeta_1 + 1$, $i = 1, \cdots, n$. Then using the fact that the Y's are a sequence of interchangeable random variables, and the Chebyshev inequality

$$\Pr\{\max_{1 \leqq i \leqq n}\{|mY_1|/(n-m), \cdots, |mY_n|/(n-m)\} > \Delta\}$$
$$\leqq n \Pr\{|mY_1|/(n-m) < \Delta\} \leqq nm^2(\zeta_1 + 1)/((n-m)^2\Delta^2) \to 0 \tag{14}$$

as $n \to \infty$. From (12), (13) and (14), it follows that

$$\Pr\{U_n^{\,0}, U_{n-1}^1, \cdots, U_{n-1}^n \,\varepsilon\, I \text{ simultaneously}\} \to 1. \tag{15}$$

Also, since if $P\{E_n\} \to 1$, $\lim P\{A_n\} = \lim P\{A_nE_n\}$, one may tacitly assume (15) in what follows.

Next, note that

$$(16)\quad \begin{aligned}\hat\theta_i &= ng(U_n^0) - (n-1)g(U_{n-1}^i)\\ &= ng(U_n^0) - (n-1)\{g(U_n^0)\\ &\qquad + (U_{n-1}^i - U_n^0)g'(U_n^0) + (U_{n-1}^i - U_n^0)^2 g''(\xi_i)/2\}\end{aligned}$$

where ξ_i lies between U_n^0 and U_{n-1}^i, $i = 1, \cdots, n$. Hence,

$$(17)\quad \begin{aligned}\hat\theta &= n^{-1}\textstyle\sum_{i=1}^n \hat\theta_i\\ &= g(U_n^0) - (n-1)n^{-1}\textstyle\sum_{i=1}^n (U_{n-1}^i - U_n^0)g'(U_n^0)\\ &\qquad - (n-1)n^{-1}\textstyle\sum_{i=1}^n g''(\xi_i)(U_{n-1}^i - U_n^0)^2/2\\ &= g(U_n^0) - (n-1)n^{-1}\textstyle\sum_{i=1}^n g''(\xi_i)(U_{n-1}^i - U_n^0)^2/2\end{aligned}$$

since $n^{-1}\sum_{i=1}^n U_{n-1}^i = U_n^0$. Rewriting (17),

$$(18)\quad n^{\frac12}(\hat\theta - \theta) = n^{\frac12}(g(U_n^0) - g(0)) - (n-1)n^{-\frac12}\textstyle\sum_{i=1}^n g''(\xi_i)(U_{n-1}^i - U_n^0)^2/2.$$

For $U_n^0 \varepsilon I$, $g(U_n^0) = g(0) + U_n^0 \cdot g'(\xi)$ where ξ lies between U_n^0 and 0. Asymptotically $n^{\frac12}U_n^0$ is $\mathfrak{N}(0, m^2\zeta_1)$ by Theorem 3, and $g'(\xi) \to_P g'(0)$ (with ξ defined arbitrarily if $U_n^0 \not\varepsilon I$). By Slutsky's theorem, the first term on the right of (18) is $\mathfrak{N}(0, m^2\zeta_1(g'(0))^2)$. Applying Slutsky's theorem again, one only need show that

$$(19)\quad (n-1)n^{-\frac12}\textstyle\sum_{i=1}^n g''(\xi_i)(U_{n-1}^i - U_n^0)^2/2 \to_P 0.$$

Consider the expression

$$(20)\quad \begin{aligned}&(n-1)\textstyle\sum_{i=1}^n (U_{n-1}^i - U_n^0)^2\\ &= (n-1)\{\textstyle\sum_{i=1}^n (U_{n-1}^i)^2 - n(U_n^0)^2\}\\ &= (n-1)[\binom{n-1}{m}^{-2}\textstyle\sum_{i=1}^n\{\sum_i f^*(X_{\alpha_1^i}, \cdots, X_{\alpha_m^i})f^*(X_{\beta_1^i}, \cdots, X_{\beta_m^i})\}\\ &\qquad - n\binom{n}{m}^{-2}\textstyle\sum f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})f^*(X_{\beta_1}, \cdots, X_{\beta_m})]\end{aligned}$$

where $\sum_i$ is over all combinations $(\alpha_1^i, \cdots, \alpha_m^i)$ of m integers from $(1, \cdots, i-1, i+1, \cdots, n)$ and all combinations $(\beta_1^i, \cdots, \beta_m^i)$ of m integers from $(1, \cdots, i-1, i+1, \cdots, n)$, and $\sum$ is over all combinations $(\alpha_1, \cdots, \alpha_m)$ of m integers from $(1, \cdots, n)$ and all combinations $(\beta_1, \cdots, \beta_m)$ of m integers from $(1, \cdots, n)$.

Collecting terms, (20) becomes,

$$(21)\quad \begin{aligned}&(n-1)\{\binom{n-1}{m}^{-2}\textstyle\sum_{c=0}^m (n-2m+c)\sum_c f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})\\ &\qquad \cdot f^*(X_{\beta_1}, \cdots, X_{\beta_m}) - n\binom{n}{m}^{-2}\textstyle\sum_{c=0}^m\sum_c f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})\\ &\qquad \cdot f^*(X_{\beta_1}, \cdots, X_{\beta_m})\}\\ &= (n-1)n^{-1}\binom{n-1}{m}^{-2}\\ &\qquad \cdot\textstyle\sum_{c=0}^m (cn - m^2)\sum_c\{f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})f^*(X_{\beta_1}, \cdots, X_{\beta_m})\}\end{aligned}$$

where $\sum_c$ indicates that the sum is over all combinations $(\alpha_1, \cdots, \alpha_m)$ of m integers from $(1, \cdots, n)$ and all combinations $(\beta_1, \cdots, \beta_m)$ of m integers from $(1, \cdots, n)$ having exactly c common members.

Let

$$(22) \quad U_c = \binom{n}{c}^{-1}\binom{n-c}{m-c}^{-1}\binom{n-m}{m-c}^{-1} \sum f^*(X_{\alpha_1}, \cdots, X_{\alpha_c}, X_{\beta_1}, \cdots, X_{\beta_{m-c}}) \cdot f^*(X_{\alpha_1}, \cdots, X_{\alpha_c}, X_{\gamma_1}, \cdots, X_{\gamma_{m-c}})$$

where $\sum$ is over all disjoint sets $(\alpha_1, \cdots, \alpha_c)$, $(\beta_1, \cdots, \beta_{m-c})$, $(\gamma_1, \cdots, \gamma_{m-c})$ of distinct integers chosen from $(1, \cdots, n)$. Note that U_c is a U-statistic with symmetric kernel

$$K^*(X_{\alpha_1}, \cdots, X_{\alpha_c}, X_{\beta_1}, \cdots, X_{\beta_{m-c}}, X_1, \cdots, X_{\gamma_{m-c}}) = \binom{2m-c}{c\ (m-c)\ (m-c)}^{-1} \sum\nolimits_{\Pi_m} f^*(X_{\alpha_1}, \cdots, X_{\alpha_c}, X_{\beta_1}, \cdots, X_{\beta_{m-c}}) \cdot f^*(X_{\alpha_1}, \cdots, X_{\alpha_c}, X_{\gamma_1}, \cdots, X_{\gamma_{m-c}})$$

whose expected value is ζ_c, and where $\sum_{\Pi_m}$ indicates the sum over all of the $\binom{2m-c}{c\ (m-c)\ (m-c)}$ possible permutations of $(\alpha_1, \cdots, \alpha_c, \beta_1, \cdots, \beta_{m-c}, \gamma_1, \cdots, \gamma_{m-c})$ that are not permutations only within the sets $(\alpha_1, \cdots, \alpha_c)$, $(\beta_1, \cdots, \beta_{m-c})$, or $(\gamma_1, \cdots, \gamma_{m-c})$. Note that U_c can also be expressed in the form

$$(23) \quad U_c = \binom{n}{2m-c}^{-1} \cdot \sum\nolimits_{C_{2m-c}} K^*(X_{\alpha_1}, \cdots, X_{\alpha_c}, X_{\beta_1}, \cdots, X_{\beta_{m-c}}, X_{\gamma_1}, \cdots, X_{\gamma_{m-c}})$$

where $\sum_{C_{2m-c}}$ indicates that the summation is over all combinations of $(2m - c)$ integers $(\alpha_1, \cdots, \alpha_c, \beta_1, \cdots, \beta_{m-c}, \gamma_1, \cdots, \gamma_{m-c})$ chosen from $(1, \cdots, n)$. Since by assumption $E(f^*(X_1, \cdots, X_m))^2 < \infty$, use of Theorem 4 shows that $U_c \to \zeta_c$ a.s. and hence

$$(24) \qquad U_c \to_P \zeta_c \quad \text{as} \quad n \to \infty.$$

Hence substituting (22) into (21), and using (24), one obtains

$$(25) \quad (n-1) \sum\nolimits_{i=1}^n (U_{n-1}^i - U_n^0)^2 = (n-1)n^{-1}\binom{n-1}{m}^{-2} \cdot \sum\nolimits_{c=0}^m (cn - m^2)\binom{n}{c}\binom{n-c}{m-c}\binom{n-m}{m-c} U_c \to_P m^2\zeta_1 \quad \text{as} \quad n \to \infty.$$

If (15) holds, $|g''(\xi_i)| < M$, $i = 1, \cdots, n$ for some $0 < M < +\infty$. Since (25) holds, the $n^{-\frac{1}{2}}$ term makes (19) $\to_P 0$ with or without $U_n^0, U_{n-1}^1, \cdots, U_{n-1}^n \in I$ (ξ_i arbitrary if $U_n^0, U_{n-1}^i \notin I$). Hence the result follows.

THEOREM 6. *Let $X_1, \cdots, X_n$ be independent, identically distributed random variables, and $f^*(X_1, \cdots, X_m)$ a real-valued symmetric statistic with expectation η, and finite second moment $E\{[f^*(X_1, \cdots, X_m)]^2\} < +\infty$. Let g be a function defined on the real line, which in a neighborhood of η has a continuous first derivative. Then, as $n \to \infty$, $s_g^2 \to_P [g'(\eta)]^2 m^2 \zeta_1$.*

PROOF. The proof essentially follows that of Theorem 2. Assume $\eta = 0$ without loss of generality. Again one can show that (15) holds. In this case $g(U_{n-1}^i) =$

$g(U_n^0) - (U_{n-1}^i - U_n^0)g'(\tau_i)$ for $i = 1, \cdots, n$ where τ_i lies between U_n^0 and U_{n-1}^i. Hence

$$s_g^2 = (n-1)^{-1} \sum_{i=1}^n (\hat{\theta}_i - \hat{\theta})^2 = (n-1) \sum_{i=1}^n (\hat{\theta}_{n-1}^i - \hat{\theta}_{n-1}^{\cdot})^2$$

where

$$\hat{\theta}_{n-1}^{\cdot} = n^{-1} \sum_{i=1}^n \hat{\theta}_{n-1}^i,$$

$$\begin{aligned} s_g^2 &= (n-1) \sum_{i=1}^n (g(U_{n-1}^i) - n^{-1} \sum_{j=1}^n g(U_{n-1}^j))^2 \\ &= (n-1) \sum_{i=1}^n ((U_{n-1}^i - U_n^0)g'(\tau_i) - n^{-1} \sum_{j=1}^n (U_{n-1}^j - U_n^0)g'(\tau_j))^2 \\ &= (n-1) \sum_{i=1}^n \{(U_{n-1}^i - U_n^0)g'(0) + (U_{n-1}^i - U_n^0)(g'(\tau_i) - g'(0)) \\ &\quad - n^{-1} \sum_{j=1}^n (U_{n-1}^j - U_n^0)g'(\tau_j)\}^2 \\ &= [g'(0)]^2(n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 \\ &\quad + (n-1) \sum_{i=1}^n \{(U_{n-1}^i - U_n^0)(g'(\tau_i) - g'(0)) \\ &\quad - n^{-1} \sum_{j=1}^n (U_{n-1}^j - U_n^0)(g'(\tau_j) - g'(0))\}^2 + X - \text{product term.} \end{aligned}$$

The first term above, $[g'(0)]^2(n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 \to_P [g'(0)]^2 m^2 \zeta_1$ by (25). Hence the proof is complete if the second term can be shown $\to_P 0$, since then the X-product will $\to_P 0$ by the Cauchy-Schwarz inequality. Let $h(x) = g'(x) - g'(0)$, then if it can be shown that

$$(26) \qquad (n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 (h(\tau_i))^2 \to_P 0 \quad \text{as} \quad n \to \infty,$$

the second term will $\to_P 0$. Since g' is continuous near zero, for $\epsilon > 0$ there exists $\Delta_\epsilon > 0$ such $|h(x)| < \epsilon$ for $x \in I_\epsilon = (-\Delta_\epsilon, \Delta_\epsilon)$. For $U_n^0, U_{n-1}^1, \cdots, U_{n-1}^n \in I_\epsilon$,

$$(27) \quad (n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 (h(\tau_i))^2 \leqq \epsilon^2 (n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 \to_P \epsilon^2 m^2 \zeta_1.$$

Since ϵ is arbitrary, the left side $\to_P 0$, and the result follows.

Recall it has been assumed in the grouping $N = nk$, that $k = 1$. The above proofs can be repeated with only slight modifications if $n \to \infty$ with $k > 1$. Hence Tukey's conjecture, (5), holds in this case, and

$$n^{\frac{1}{2}}(\hat{\theta} - \theta)/((n-1)^{-1} \sum_{i=1}^n (\hat{\theta}_i - \hat{\theta})^2)^{\frac{1}{2}}$$

has an asymptotic standard normal distribution as $n \to \infty$. The case where n remains finite when $N \to \infty$ remains to be shown. The notation is as in (11), except $U_{n-1}^i = \binom{(n-1)k}{m}^{-1} \sum_{C_{n-1}^i} f^*(X_{\beta_1^i}, \cdots, X_{\beta_m^i})$ where C_{n-1}^i indicates the summation is over all combinations $\beta_1^i, \cdots, \beta_m^i$ of m integers chosen from $1, \cdots, (i-1)k, ik+1, \cdots, N$.

THEOREM 7. *Let $X_1, \cdots, X_N$ be N independent identically distributed random variables, and let $f^*(X_1, \cdots, X_m)$ be a real-valued symmetric statistic with expectation η, and finite second moment $E[f^*(X_1, \cdots, X_m)]^2 < +\infty$. Let g be a function* *defined on the real line, which in a neighborhood of η has a continuous first derivative. Then as $N \to \infty$,*

$$n^{\frac{1}{2}}(\hat{\theta} - \theta)/s_g \to_{\mathcal{L}} t_{n-1},$$

where t_{n-1} denotes the Student-t distribution with $(n - 1)$ degrees of freedom.

PROOF. Assume $n = 2$; the proof extends with slight modification for finite n. Also without loss of generality let $\eta = 0$.

Then

$$\hat{\theta}_1 = 2g(U_2^0) - g(U_1^1) = g(U_1^2) + \delta_1$$
$$\hat{\theta}_2 = 2g(U_2^0) - g(U_1^2) = g(U_1^1) + \delta_2$$

where $\delta_1 = \delta_2 = 2g(U_2^0) - g(U_1^1) - g(U_1^2)$. But $\delta_i = o_p(1)$, $i = 1, 2$, since

$$\text{Var } (\delta_i) \to (g'(0))^2 \text{Var}(2U_2^0 - U_1^1 - U_1^2) = (g'(0))^2(4 \text{ Var } (U_2^0) + \text{Var } (U_1^1) + \text{Var } (U_1^2) - 4 \text{ Cov } (U_2^0, U_1^1) - 4 \text{ Cov } (U_2^0, U_1^2)).$$

But

$$\text{Var } (U_2^0) = \binom{N}{m}^{-1} \sum_{c=1}^{n} \binom{m}{c}\binom{N-m}{m-c}\zeta_c,$$
$$\text{Var } (U_1^1) = \text{Var } (U_1^2) = \binom{N/2}{m}^{-1} \sum_{c=1}^{n} \binom{m}{c}\binom{\frac{1}{2}N-m}{m-c}\zeta_c, \quad \text{and}$$
$$\text{Cov } (U_2^0, U_1^1) = \text{Cov } (U_2^0, U_1^2) = \binom{N}{m}^{-1}\binom{\frac{1}{2}N}{m}^{-1} \sum_{c=1}^{m} \binom{\frac{1}{2}N}{m}\binom{\frac{1}{2}N-c}{c-c}\binom{N-m}{m-c}\zeta_c.$$

Combining the above expressions with Chebyshev's inequality, $\delta_i = o_p(1)$, $i = 1, 2$.

Hence, $\hat{\theta}_1$, $\hat{\theta}_2$ tend to be independent $\mathfrak{N}(0, m^2\zeta_1(g'(0))^2)$ random variables as $N \to \infty$. Applying Theorem 5 in Mann-Wald [12], one obtains the desired result.

Note the weaker assumptions on the function $g(\cdot)$ than in Theorem 5. In all the following results, we will assume in the grouping $N = nk$ that $k = 1$, and hence $N = n$. One could repeat the above argument and obtain convergence to a t distribution in the case where n remains finite.

(c) *Functions of several U-statistics.* It is possible to generalize Theorems 5 and 6 to vector U-statistics. To this end, assume $X_1, \cdots, X_n$ are n independent, identically distributed random vectors of p components. Let $U^{(1)}, \cdots, U^{(q)}$ be such that

$$U^{(j)} = \binom{n}{m_j}^{-1} \sum_{C_n} f^{*(j)}(X_{\alpha_1}, \cdots, X_{\alpha_{m_j}}), \qquad j = 1, \cdots, q, \tag{28}$$

where C_n as in (7), and $f^{*(j)}$ is a *real-valued* symmetric kernel based on m_j observations, and is an unbiased estimate of $\eta^{(j)}$. Let g be a real-valued function of q arguments, and

$$\hat{\theta}_n^0 = g(U^{(1)}, \cdots, U^{(q)}) \tag{29}$$

$$\theta_{n-1}^i = g(U_{\sim i}^{(1)}, \cdots, U_{\sim i}^{(q)})$$

where $$U_{\sim i}^{(j)} = \binom{n-1}{m_j}^{-1} \sum_{C_{n-1}} f^{*(j)}(X_{\beta_1^i}, \cdots, X_{\beta_{m_j}^i})$$

and C^i_{n-1}, $(\beta_1^{\ i}, \cdots, \beta^i_{m_j})$ as in (10),

$$\hat{\theta}_i = n\hat{\theta}_n^{\ 0} - (n-1)\hat{\theta}^i_{n-1}, \qquad i = 1, \cdots, n.$$
$$\hat{\theta} = n^{-1}\textstyle\sum_{i=1}^n \hat{\theta}_i,$$
$$\theta = g(\eta^{(1)}, \cdots, \eta^{(q)}),$$
$$s_g^{\ 2} = (n-1)^{-1}\textstyle\sum_{i=1}^n (\hat{\theta}_i - \hat{\theta})^2 \quad \text{and}$$
$$g_k = \partial g(t^{(1)}, \cdots, t^{(q)})/\partial t^{(k)} \mid_{(\eta^{(1)}, \cdots, \eta^{(q)})},$$
$$g_{kl} = \partial g(t^{(1)}, \cdots, t^{(q)})/\partial t^{(k)}\partial t^{(l)} \mid_{(\eta^{(1)}, \cdots, \eta^{(q)})},$$

for $k, l = 1, \cdots, q$.

THEOREM 8. *Let $X_1, \cdots, X_n$ be independent, identically distributed random vectors of p components. Let $f^{*(j)}(X_1, \cdots, X_{m_j})$ be a real-valued symmetric statistic with expectation $\eta^{(j)}$ and finite second moment $E[f^{*(j)}(X_1, \cdots, X_{m_j})]^2 < +\infty$, $j = 1, \cdots, q$. Let g be a real-valued function defined on R^q, which in a neighborhood of $(\eta^{(1)}, \cdots, \eta^{(q)})$ has bounded second partial derivatives. Then, as $n \to \infty$, $n^{\frac{1}{2}}(\hat{\theta} - \theta)$ is asymptotically normally distributed with mean zero and variance $\sum_{i=1}^q \sum_{j=1}^q m_i m_j g_i g_j \zeta_1^{(i,j)}$ where*

$$\zeta_c^{(i,j)} = \operatorname{Cov}(f_c^{*(i)}(X_1, \cdots, X_c), f_c^{*(j)}(X_1, \cdots, X_c)).$$

PROOF. Extending (15), one can readily show for any interval I_j containing $\eta^{(j)}$, in which g has bounded second partial derivatives, that

$$\Pr\{U^{(j)}, U^{(j)}_{\sim h}, \cdots, U^{(j)}_{\sim n} \ \varepsilon\ I_j \text{ simultaneously}\} \to 1 \tag{30}$$

for $j = 1, \cdots, q$. Also,

$$\begin{aligned} \theta_i &= ng(U^{(1)}, \cdots, U^{(q)}) - (n-1)g(U^{(1)}_{\sim i}, \cdots, U^{(q)}_{\sim i}) \\ &= ng(U^{(1)}, \cdots, U^{(q)}) - (n-1)\{g(U^{(1)}, \cdots, U^{(q)}) \\ &\quad + \textstyle\sum_{j=1}^q (U^{(j)}_{\sim i} - U^{(j)})g_j \\ &\quad + \textstyle\sum_{j=1}^q \sum_{k=1}^q (U^{(j)} - U^{(j)})(U^{(k)}_{\sim i} - U^{(k)})g_{jk}(\xi_i)/2\} \end{aligned} \tag{31}$$

where ξ_i indicates that the partial derivative is evaluated on the line segment between $(U^{(1)}, \cdots, U^{(q)})$ and $(U^{(1)}_{\sim i}, \cdots, U^{(q)}_{\sim i})$. Hence

$$\begin{aligned} \hat{\theta} &= n^{-1}\textstyle\sum_{i=1}^n \hat{\theta}_i \\ &= g(U^{(1)}, \cdots, U^{(q)}) - (n-1)n^{-1}\textstyle\sum_{j=1}^n \{\sum_{j=1}^q \sum_{k=1}^q g_{jk}(\xi_i) \\ &\quad \cdot (U^{(j)}_{\sim i} - U^{(j)})(U^{(k)}_{\sim i} - U^{(k)})/2\}. \end{aligned} \tag{32}$$

Rewriting (32), one obtains

$$\begin{aligned} n^{\frac{1}{2}}(\hat{\theta} - \theta) &= n^{\frac{1}{2}}(g(U^{(1)}, \cdots, U^{(q)}) - g(\eta^{(1)}, \cdots, \eta^{(q)})) \\ &\quad - (n-1)n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n \{\sum_{j=1}^q \sum_{k=1}^q g_{jk}(\xi_i)(U^{(j)}_{\sim i} - U^j) \\ &\quad \cdot (U^{(k)}_{\sim i} - U^{(k)})/2\}. \end{aligned} \tag{33}$$

When the events described in (30) hold,

$$g(U^{(1)}, \cdots, U^{(q)}) = g(\eta^{(1)}, \cdots, \eta^{(q)}) + \sum_{j=1}^{q} (U^{(j)} - \eta^{(j)}) g_j(\xi)$$

where ξ is some point on the line segment between $(U^{(1)}, \cdots, U^{(q)})$ and $(\eta^{(1)}, \cdots, \eta^{(q)})$. Combining Hoeffding's Theorem 7.1 [6], and the multivariate form of Slutsky's theorem,

$$n^{\frac{1}{2}} \sum_{j=1}^{q} (U^{(j)} - \eta^{(j)}) g_j(\xi) \to_{\mathcal{L}} \mathfrak{N}(0, \sum_{j=1}^{q} \sum_{k=1}^{q} m_j m_k g_j g_k \zeta_1^{(j,k)})$$

as $n \to \infty$.

Returning to (33), using Slutsky's theorem it only remains to show that

$$(34) \quad (n-1) n^{-\frac{1}{2}} \sum_{i=1}^{n} \{ \sum_{j=1}^{q} \sum_{k=1}^{q} g_{jk}(\xi_i)(U_{\sim i}^{(j)} - U^{(j)}) \cdot (U_{\sim i}^{(k)} - U^{(k)})/2 \} \to_P 0.$$

However, when the events of (30) hold, $|g_{jk}(\xi_i)| < M, i = 1, \cdots, n; j, k = 1, \cdots, q$ for some $0 < M < +\infty$. But (25) shows that $(n-1)n^{-\frac{1}{2}} M \sum_{i=1}^{n} \sum_{j=1}^{q} (U_{\sim i}^{(j)} - U^{(j)})^2/2 \to_P 0$, and the cross product terms can be handled by the Cauchy-Schwarz inequality. Hence (34) holds.

THEOREM 9. *Let $X_1, \cdots, X_n$ be independent, identically distributed random vectors of p components. Let $f^{*(j)}(X_1, \cdots, X_m)$ be a real-valued symmetric statistic with expectation $\eta^{(j)}$, and finite second moment $E[f^{*(j)}(X_1, \cdots, X_{m_j})]^2 < +\infty, j = 1, \cdots, q$. Let g be a real-valued function defined on R^q, which in a neighborhood of $(\eta^{(1)}, \cdots, \eta^{(q)})$ has continuous first partial derivatives. Then as $n \to \infty$,*

$$s_g^2 \to_P [\sum_{i=1}^{q} \sum_{j=1}^{q} m_i m_j g_i g_j \zeta_1^{(i,j)}].$$

PROOF. As in the previous theorem, one can show the events of (30) hold for intervals $\{I_j\}_{j=1}^{q}$ about $(\eta^{(1)}, \cdots, \eta^{(q)})$ in which g has continuous first partial derivatives. Then if (30) holds,

$$g(U_{\sim i}^{(1)}, \cdots, U_{\sim i}^{(q)}) = g(U^{(1)}, \cdots, U^{(q)}) + \sum_{j=1}^{q} (U_{\sim i}^{(j)} - U^{(j)}) g_j(\tau_i)$$

where τ_i lies between $(U^{(1)}, \cdots, U^{(q)})$ and $(U_{\sim i}^{(1)}, \cdots, U_{\sim i}^{(q)})$. Hence,

$$\begin{aligned}
s_g^2 &= (n-1)^{-1} \sum_{i=1}^{n} (\hat{\theta}_i - \theta)^2 = (n-1) \sum_{i=1}^{n} (\hat{\theta}_{n-1}^{i} - \hat{\theta}_{n-1}^{\cdot})^2 \\
&= (n-1) \sum_{i=1}^{n} (g(U_{\sim i}^{(1)}, \cdots, U_{\sim i}^{(q)}) - n^{-1} \sum_{k=1}^{n} g(U_{\sim k}^{(1)}, \cdots, U_{\sim k}^{(q)}))^2 \\
&= (n-1) \sum_{i=1}^{n} (\sum_{j=1}^{q} (U_{\sim i}^{(j)} - U^{(j)}) g_j(\tau_i) \\
&\qquad - n^{-1} \sum_{k=1}^{n} \sum_{j=1}^{q} (U_{\sim k}^{(j)} - U^{(j)}) g_j(\tau_k))^2 \\
(35) \quad &= (n-1) \sum_{i=1}^{n} \{ \sum_{j=1}^{q} (U_{\sim i}^{(j)} - U^{(j)}) g_j + \sum_{j=1}^{q} (U_{\sim i}^{(j)} - U^{(j)}) \\
&\qquad \cdot (g_j(\tau_i) - g_j) - n^{-1} \sum_{k=1}^{n} \sum_{j=1}^{q} (U_{\sim k}^{(j)} - U^{(j)})(g_j(\tau_k) - g_j) \}^2 \\
&= (n-1) \sum_{i=1}^{n} \sum_{j=1}^{q} \sum_{k=1}^{q} (U_{\sim i}^{(j)} - U^{(j)})(U_{\sim i}^{(k)} - U^{(k)}) g_j g_k \\
&\quad + (n-1) \sum_{i=1}^{n} \{ \sum_{j=1}^{q} (U_{\sim i}^{(j)} - U^{(j)})(g_j(\tau_i) - g_j) - n^{-1} \\
&\qquad \cdot \sum_{k=1}^{n} \sum_{j=1}^{q} (U_{\sim k}^{(j)} - U^{(j)})(g_j(\tau_k) - g_j) \}^2 + X - \text{product term.}
\end{aligned}$$

Analogous to expressions (20), (21), it can be readily shown that

$$(n-1)\sum_{i=1}^{n}(U_{\sim i}^{(j)}-U^{(j)})(U_{\sim i}^{(k)}-U^{(k)})$$

$$(36)\quad = (n-1)n^{-1}\binom{n-1}{m_j}^{-1}\binom{n-1}{m_k}^{-1}\sum_{c=0}^{m_{jk}}(cn-m_jm_k)\sum_c f^{*(j)}(X_{\alpha_1},\cdots,X_{\alpha_{m_j}})$$

$$\cdot f^{*(k)}(X_{\beta_1},\cdots,X_{\beta_{m_k}}) \qquad m_{jk}=\min\{m_j,m_k\}, j,k=1,\cdots,q,$$

where $\sum_c$ indicates that the sum is taken over all combinations $(\alpha_1,\cdots,\alpha_{m_j})$ of m_j integers from $(1,\cdots,n)$ and all combinations $(\beta_1,\cdots,\beta_{m_k})$ of m_k integers chosen from $(1,\cdots,n)$ having exactly c common members.

Next, let $U^{(j,k)} = [\binom{n}{m_j}\binom{m_j}{c}\binom{n-m_j}{m_k-c}]^{-1}\sum_c f^{*(j)}(X_{\alpha_1},\cdots,X_{\alpha_{m_j}})\cdot f^{*(k)}\cdot(X_{\beta_1},\cdots,X_{\beta_{m_k}})$ (there are exactly $\binom{n}{m_j}\binom{m_j}{c}\binom{n-m_j}{m_k-c}=\binom{n}{c}\binom{n-c}{m_j-c}\binom{n-m_j}{m_k-c}$ pairs of sets $(\alpha_1,\cdots,\alpha_{m_j})$, $(\beta_1,\cdots,\beta_{m_k})$ having exactly c common integers). By an argument analogous to that of (25),

$$(37)\qquad (n-1)\sum_{i=1}^{n}(U_{\sim i}^{(j)}-U^{(j)})(U_{\sim i}^{(k)}-U^{(k)})\rightarrow_P m_jm_k\zeta_1^{(j,k)}.$$

Moreover, since q is finite, the first term of (35) $\rightarrow_P \sum_{i=1}^{q}\sum_{j=1}^{q}m_im_jg_ig_j\zeta_1^{(i,j)}$.

The remainder of (35) is shown to $\rightarrow_P 0$ by an argument exactly as that in Theorem 5.

(d) *Non-identically distributed case.* Let $X_1,\cdots,X_n$ be independent (not necessarily identically distributed) random variables, and let $f^*(X_{\alpha_1},\cdots,X_{\alpha_m})$ be a symmetric kernel such that $E[f^*(X_{\alpha_1},\cdots,X_{\alpha_m})]=\eta$ for all $\alpha_1,\cdots,\alpha_m$.

Let

$$U(X_1,\cdots,X_n)$$
$$=\binom{n}{m}^{-1}\sum_{C_n}f^*(X_{\alpha_1},\cdots,X_{\alpha_m});\ f^*_{c;\beta_1,\cdots,\beta_{m-c}}(x_1,\cdots,x_c)$$
$$=E\{f^*(X_1,\cdots,X_c,X_{\beta_1},\cdots,X_{\beta_{m-c}})\mid X_1=x_1,\cdots,X_c=x_c\};$$

$$\zeta_{c;(\alpha_1,\cdots,\alpha_c)\beta_1,\cdots,\beta_{m-c};\gamma_1,\cdots,\gamma_{m-c}}$$
$$=\mathrm{Cov}\,\{f^*_{c;\beta_1,\cdots,\beta_{m-c}}(X_{\alpha_1},\cdots,X_{\alpha_c}),f^*_{c;\gamma_1,\cdots,\gamma_{m-c}}(X_{\alpha_1},\cdots,X_{\alpha_c})\},$$

and

$$\zeta_{c(\alpha_1,\cdots,\alpha_c)}=[\binom{n-c}{m-c}\binom{n-m}{m-c}]^{-1}\sum\zeta_{c;(\alpha_1,\cdots,\alpha_c)\beta_1,\cdots,\beta_{m-c};\gamma_1,\cdots,\gamma_{m-c}}$$

where the sum is extended over all disjoint sets $\{\beta_1,\cdots,\beta_{m-c}\}$, $\{\gamma_1,\cdots,\gamma_{m-c}\}$ chosen from $\{1,\cdots,n\}$ excluding $\{\alpha_1,\cdots,\alpha_c\}$, and finally $\zeta_{c,n}=\binom{n}{c}^{-1}\cdot\sum\zeta_{c(\alpha_1,\cdots,\alpha_c)}$ where the sum is over all combinations $\{\alpha_1,\cdots,\alpha_c\}$ of c integers chosen from $\{1,\cdots,n\}$. Then one can show $\mathrm{Var}\,U=\binom{n}{m}^{-1}\sum_{c=1}^{m}\binom{m}{c}\binom{n-m}{m-c}\zeta_{c,n}$.

Let

$$(38)\qquad h_{1(\nu)}(X)=\binom{n-1}{m-1}^{-1}\sum_{\neq\nu}(f^*_{1;\beta_1,\cdots,\beta_{m-1}}(X)-\eta),$$

where the sum is over all sets $(\beta_1,\cdots,\beta_{m-1})$ chosen from the first n integers excluding the integer ν.

THEOREM 10. *Let $X_1,\cdots,X_n$ be n independent random variables, and assume* *for some* $0 < A < \infty$

(39) $$E[f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})]^4 < A \quad \text{for all} \quad (\alpha_1, \cdots, \alpha_m),$$

(40) $$E|h_{1_{(\nu)}}(X_\nu)|^3 < \infty \quad \text{for} \quad \nu = 1, \cdots, n,$$

and

(41) $$\lim_{n\to\infty} \sum_{\nu=1}^n E\{|h_{1_{(\nu)}}(X_\nu)|^3\}/[\sum_{\nu=1}^n E\{h_{1_{(\nu)}}(X_\nu)\}^2]^{\frac{3}{2}} = 0.$$

Let g be a function defined on the real line, which in a neighborhood of η has a bounded second derivative. Let $\hat{\theta}$, the jackknife estimate of $g(\eta)$ be defined as in (11). *Then if*

(42) $$\zeta_{1,n} \to \zeta_1 \quad \text{as} \quad n \to \infty, \qquad 0 < \zeta_1 < +\infty,$$

the distribution of $(\hat{\theta} - \theta)/(g'(\eta)(\text{Var } U)^{\frac{1}{2}})$ is asymptotically normal with mean zero and variance one.

PROOF. Without loss of generality, let $\eta = 0$, and let U_n^0, U_{n-1}^i, $i = 1, \cdots, n$ be as in (10). Let $I = (-3\Delta, 3\Delta)$, $\Delta > 0$, be any neighborhood of zero in which g'' is bounded. As $n \to \infty$, (42) implies that $U_n^0 \to_P 0$, and hence

(43) $$\Pr\{U_n^0 \,\varepsilon\, (-\Delta, \Delta)\} \to 1.$$

Let $Y_i = \binom{n-1}{m-1}^{-1} \sum_{D_{n-1}^i} f^*(X_i, X_{\alpha_1 i}, \cdots, X_{\alpha_{m-1}^i})$, where $\sum_{D_{n-1}^i}$ indicates that the sum is over all combinations of $m - 1$ integers $(\alpha_1, \cdots, \alpha_{-m1})$ chosen from $(1, \cdots, i-1, i+1, \cdots, n)$, $i = 1, \cdots, n$. Note Var $Y_i \leqq \zeta_{1(i)} + 1$ for n sufficiently large, $i = 1, \cdots, n$. Hence using the Chebyshev inequality and (42),

$$\Pr\{\max_{1\leqq i\leqq n}\{|mY_1|/(n-m), \cdots, |mY_n|/(n-m)\} > \Delta\}$$

(44) $$\leqq \sum_{i=1}^n \Pr\{|mY_i|/(n-m) > \Delta\}$$

$$\leqq m^2(\sum_{i=1}^n \zeta_{1(i)} + 1)/(\Delta^2(n-m)^2) = m^2 n(\zeta_{1,n} + 1)/(\Delta^2(n-m)^2) \to 0.$$

Using (13), (43) and (44), one obtains

(45) $$\Pr\{U_n^0, U_{n-1}^1, \cdots, U_{n-1}^n \,\varepsilon\, I \text{ simultaneously}\} \to 1.$$

Again, (45) may be tacitly assumed in what follows.

After expanding terms in a power series, one again obtains the expression

$$(\hat{\theta} - \theta) = (g(U_n^0) - g(0)) - (n-1)n^{-1} \sum_{i=1}^n g''(\xi_i)(U_{n-1}^i - U_n^0)^2/2$$

where ξ_i lies between U_{n-1}^i and U_n^0, $i = 1, \cdots, n$. This can be written to obtain

(46) $$(\hat{\theta} - \theta)/(g'(0)(\text{Var } U_n^0)^{\frac{1}{2}}) = (g'(0)(\text{Var } U_n^0)^{\frac{1}{2}})^{-1}\{(g(U_n^0) - g(0)) - (n-1)n^{-1} \sum_{i=1}^n g''(\xi_i)(U_{n-1}^i - U_n^0)^2/2\}.$$

For $U_n^0 \,\varepsilon\, I$, $g(U_n^0) = g(0) + U_n^0 \cdot g'(\xi_{U_n^0})$, where $|\xi_{U_n^0}| \leqq |U_n^0|$. By (39), (40), (41) and Hoeffding's U-statistic theorem for non-identically distributed random variables, $U_n^0/(\text{Var } U_n^0)^{\frac{1}{2}}$ is asymptotically normally distributed with mean zero and variance one. Since $g'(\xi_{U_n^0})/g'(0) \to_P 1$, the first term on the right of (46) is asymptotically $\mathfrak{N}(0, 1)$ by Slutsky's theorem. It remains to show the second term $\to_P 0$.

Consider the expression

$$T_n = (n-1)\sum_{i=1}^{n} (U_{n-1}^{i} - U_n^{0})^2$$
$$= (n-1)n^{-1}\binom{n-1}{m}^{-2}\{\sum_{c=0}^{m} (cn - m^2)\sum_c f^*(X_{\alpha_1}, \cdots, X_{\alpha_m}) \cdot f^*(X_{\beta_1}, \cdots, X_{\beta_m})\}$$

where the notation is the same as in (21). Applying (42), one obtains

$$E(T_n) = m^2\zeta_{1,n} + 0(n^{-1}) \to m^2\zeta_1, \quad \text{as} \quad n \to \infty.$$

Also, $E(T_n - m^2\zeta_1)^2 \to 0$ as $n \to \infty$, using (39), (42) and some messy but straightforward algebra. Hence,

$$(47) \qquad T_n \to_P m^2\zeta_1 \quad \text{as} \quad n \to \infty.$$

When (45) holds, $|g''(\xi_i)| < M$, $i = 1, \cdots, n$ for some $0 < M < \infty$. This fact and (47) imply that

$$(n-1)n^{-1}\sum_{i=1}^{n} g''(\xi_i)(U_{n-1}^{i} - U_n^{0})^2/(2g'(0)(\operatorname{Var} U_n^{0})^{\frac{1}{2}}$$
$$\to Mm\zeta_1^{\frac{1}{2}}/(2n^{\frac{1}{2}}g'(0)) \to 0 \quad \text{as} \quad n \to \infty$$

using (42). Hence the result follows.

Expression (39) can probably be weakened to some extent. In showing (7), one needs an extension of Markov's theorem to U-statistics. Such an extension seems likely, but is not in the literature.

THEOREM 11. *Let $X_1, \cdots, X_n$ be a sequence of independent random variables, and $f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})$ a real-valued symmetric kernel with expectation η for all $\{\alpha_1, \cdots, \alpha_m\}$, and for some $0 < A < \infty$,*

$$(48) \qquad E[f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})]^4 < A \quad \text{for all} \quad \alpha_1, \cdots, \alpha_m.$$

Assume (42) holds, and that g is a function defined on the real line, which in a neighborhood of $\theta(=g(\eta))$ has a continuous first derivative. Then as $n \to \infty$,

$$s_g^2 \to [g'(\eta)]^2 m^2\zeta_1 \qquad \text{(see (11))}.$$

PROOF. The proof is identical with that of Theorem 5, except to note that by (47),

$$(n-1)\sum_{i=1}^{n} (U_{n-1}^{i} - U)^2 \to_P m^2\zeta_1 \quad \text{as} \quad n \to \infty.$$

Thus with the stronger conditions of Theorem 3, one obtains $n^{\frac{1}{2}}(\hat{\theta} - \theta)/s_g \to \mathfrak{N}(0, 1)$ as $n \to \infty$.

Once more it is possible to generalize these results to vector U-statistics. Assume $X_1, \cdots, X_n$ are n independent random vectors of p components. Let $U^{(1)}, \cdots, U^{(q)}$ be as defined in the discussion preceding Theorem 6. Let

$$\zeta_{c;(\alpha_1,\cdots\alpha_c)\beta_1,\cdots,\beta_{m-c};\gamma_1,\cdots,\gamma_{m-c}}^{(i,j)}$$
$$(49) \qquad = \operatorname{Cov}\{f_{c;\beta_1,\cdots,\beta_{m-c}}^{*(i)}(X_{\alpha_1}, \cdots, X_{\alpha_c}), f_{c;\gamma_1,\cdots,\gamma_{m-c}}^{*(j)}(X_{\alpha_1}, \cdots, X_{\alpha_c})\},$$
$$\zeta_{c;(\alpha_1,\cdots,\alpha_c)}^{(i,j)}$$
$$= [\binom{n-c}{m-c}\binom{n-m}{m-c}]^{-1}\sum \zeta_{c;(\alpha_1,\cdots,\alpha_c)\beta_1,\cdots,\beta_{m-c};\gamma_1,\cdots,\gamma_{m-c}}^{(i,j)}$$

where the sum is extended over all disjoint sets $\{\beta_1, \cdots, \beta_{m-c}\}$, $\{\gamma_1, \cdots, \gamma_{m-c}\}$ chosen from $\{1, \cdots, n\}$ excluding $\{\alpha_1, \cdots, \alpha_c\}$, and $\zeta_{c,n}^{(i,j)} = \binom{n}{c}^{-1} \sum \zeta_{c;(\alpha_1,\cdots,\alpha_c)}$ where the sum is over all combinations $\{\alpha_1, \cdots, \alpha_c\}$ of c integers chosen from $\{1, \cdots, n\}$.

THEOREM 12. *Let $X_1, \cdots, X_n$ be independent random vectors of p components. Let $f^{*(j)}(X_{\alpha_1}, \cdots, X_{\alpha_{m_j}})$ be a real-valued symmetric statistic with expectation $\eta^{(j)}$ for all $(\alpha_1, \cdots, \alpha_{m_j})$, $j = 1, \cdots, q$. Assume (38)–(41) hold for each $f^{*(j)}(X_1, \cdots, X_{m_j})$, $j = 1, \cdots, q$, and that*

$$\zeta_{1,n}^{(i,i)} \to \zeta_1^{(i,i)} > 0 \quad \text{for} \quad i = 1, \cdots, q, \quad \text{and} \tag{50}$$

$$\zeta_{1,n}^{(i,j)} \to \zeta_1^{(i,j)} \quad \text{for} \quad i, j = 1, \cdots, q \quad \text{as} \quad n \to \infty.$$

Let g be a real-valued function defined on R^q, which in a neighborhood of $(\eta^{(1)}, \cdots, \eta^{(q)})$ has bounded second partial derivatives.

Then as $n \to \infty$, $n^{\frac{1}{2}}(\hat{\theta} - \theta)$ is asymptotically normally distributed with mean zero and variance $\sum_{i=1}^q \sum_{j=1}^q m_i m_j g_i g_j \zeta_1^{(i,j)}$.

PROOF. The proof is the same as that of Theorem 8 (with the slight modifications of Theorem 10), except that

$$(n-1) \sum_{i=1}^n (U_{\sim i}^{(j)} - U^{(j)})(U_{\sim i}^{(k)} - U^{(k)}) \to_P m_j m_k \zeta_1^{(j,k)}$$

because of the extension of assumption (39).

THEOREM 13. *Let $X_1, \cdots, X_n$ be independent random vectors of p components. Let $f^{*(j)}(X_{\alpha_1}, \cdots, X_{\alpha_{m_j}})$ be a real-valued symmetric kernel with expectation $\eta^{(j)}$, for all $\{\alpha_1, \cdots, \alpha_{m_j}\}$ and $j = 1, \cdots, q$. Also assume that $E[f^{*(j)}(X_{\alpha_1}, \cdots, X_{\alpha_{m_j}})]^4 < A$ for some $0 < A < \infty$, and all $\{\alpha_1, \cdots, \alpha_m\}$ and $j = 1, \cdots, q$, and that (50) holds. Let g be a real-valued function defined on R^q, which in a neighborhood of $(\eta^{(1)}, \cdots, \eta^{(q)})$ has continuous first partial derivatives. Then as $n \to \infty$, $s_g^2 \to_P \sum_{i=1}^q \sum_{j=1}^q m_i m_j g_i g_j \zeta_1^{(i,j)}$.*

PROOF. It is the same as for Theorem 9 with the comments to the proof of Theorem 12.

Once more Tukey's conjecture is valid with the stronger assumptions of Theorem 12. Now some applications of these results will be presented.

3. Applications.

(a) *Variance component models.* The following assumptions are characteristic of a one-way layout in Model II analysis of variance:

$$Y_{ij} = \mu + a_i + e_{ij}, \qquad i = 1, \cdots, I, \qquad j = 1, \cdots, J; \tag{51}$$

where the $I + IJ$ random variables $\{a_i\}$, $\{e_{ij}\}$ are completely independent, the $\{a_i\}$ are $\mathfrak{N}(0, \sigma_A^2)$, the $\{e_{ij}\}$ are $\mathfrak{N}(0, \sigma_e^2)$.

If one wishes to test

$$H_0: \theta = \sigma_A^2/\theta_e^2 \leqq \theta_0 \quad \text{vs.} \quad H_A: \theta > \theta_0 \tag{52}$$

then the usual F-test is both UMP invariant and UMP unbiased (Lehmann [10]). One can also obtain confidence intervals for $\theta = \sigma_A^2/\sigma_e^2$ using this F-distribution. Spjøtvoll [21] exhibited a test depending on the alternative θ_1, which in the unbalanced case is UMP invariant and a maxmin test.

However, it is well known (Scheffé [20]) that these procedures are not robust against non-normality, especially of the random effects. More precisely, it can be shown that even if σ_e^2 is known, non-zero kurtosis of the random effects invalidates confidence coefficients, and the probability of both types of error in the testing case. The only exception to this statement is that the probability of type I is not affected in testing

$$(53) \qquad H_0:\sigma_A^2 = 0 \quad \text{vs} \quad H_A:\sigma_A^2 > 0.$$

Two commonly employed procedures to obtain confidence intervals for σ_A^2 rely heavily on the assumptions of normality in (51). One is due to Satterthwaite [19]. The other is due to Bulmer, and is described in detail in Scheffé [20]. Let the kurtosis of the random effects be defined by $\gamma_A = E(a^4)/\sigma_A^4 - 3$. Then one can show that if $\gamma_A < 0$, these two procedures yield actual confidence coefficients greater than the stated value. Conversely, if $\gamma_A > 0$, the actual confidence coefficients will be less than the stated value. Recall if the $\{a_i\}$ are normal random variables, $\gamma_A = 0$.

Both of these techniques yield approximate confidence intervals for the variance component. The jackknife also yields approximate intervals which do not suffer when the kurtosis is non-zero. To see this, let us return to the model described in (60). Without the restriction of normality, the following is a reasonable model:

$$(54) \qquad Y_{ij} = \mu + a_i + e_{ij}, \qquad i = 1, \cdots, I, \qquad j = 1, \cdots, J,$$

where μ is some constant and the $\{a_i\}$ are independent random variables with zero means, variance σ_A^2, and finite fourth moments, the $\{e_{ij}\}$ are independent random variables with zero means, variance σ_e^2, and finite fourth moments, and the $I + IJ$ random variables $\{a_i\}$, $\{e_{ij}\}$ are independent.

Suppose one wants a confidence interval for σ_A^2 and is not willing to assume normality. Then at least asymptotically, a confidence interval can be obtained using Theorems 8 and 9. Note that the procedure also enables us to make tests on σ_A^2. The procedure is asymptotic in the sense that $I \to \infty$ with J infinite.

Using the notation of Theorems 8 and 9, let

$$X_i = \begin{pmatrix} Y_{i\cdot} \\ (J-1)^{-1} \sum_{j=1}^{J} (Y_{ij} - Y_{i\cdot})^2 \end{pmatrix}, \qquad i = 1, \cdots, I,$$

be I independent 2×1 vectors. Let $f^{*(1)}(X_1, X_2) = (Y_{1\cdot} - Y_{2\cdot})^2/2$ be a symmetric kernel for estimating $\sigma_A^2 + \sigma_e^2/J$, with corresponding U-statistic

$$(55) \qquad U^{(1)} = (I-1)^{-1} \sum_{i=1}^{I} (Y_{i\cdot} - Y_{\cdot\cdot})^2.$$

Let $f^{*(2)}(X_1) = (J-1)^{-1} \sum_{j=1}^{J} (Y_{1j} - Y_{1\cdot})^2$ be a symmetric kernel for estimating σ_e^2, with corresponding U-statistic,

$$(56) \qquad U^{(2)} = (I(J-1))^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - Y_{i\cdot})^2.$$

Let $g(U^{(1)}, U^{(2)}) = U^{(1)} - U^{(2)}/J$. Then in forming the jackknife version of this estimate, one obtains

$$\begin{aligned}\hat{\theta}_k = {} & (I-1)^{-1} \textstyle\sum_{i=1}^{I} (Y_{i.} - Y_{..})^2 \\ & + (I/(I-2))\{(Y_{k.} - Y_{..})^2 - I^{-1} \textstyle\sum_{i=1}^{I} (Y_{i.} - Y_{..})^2\} \\ & \cdot (J(J-1))^{-1} \textstyle\sum_{j=1}^{J} (Y_{kj} - Y_{k.})^2, \qquad k = 1, \cdots, I,\end{aligned}$$

and the jackknife estimate is $\hat{\theta} = I^{-1} \sum_{k=1}^{I} \hat{\theta}_k = \hat{\sigma}_A^{\,2}$, the usual estimate. Next, note that

$$\begin{aligned} s_g^{\,2} & = (I-1)^{-1} \textstyle\sum_{k=1}^{I} (\hat{\theta}_k - \hat{\theta})^2 \\ & = (I-1)^{-1} \textstyle\sum_{k=1}^{I} \{(I/(I-2))\,(Y_{k.} - Y_{..})^2 - I^{-1} \sum_{i=1}^{I} (Y_{i.} - Y_{..})^2 \\ & \quad - J^{-1}(J-1)^{-1} \textstyle\sum_{j=1}^{J} (Y_{kj} - Y_{k.})^2 \\ & \quad + (I(J-1)J)^{-1} \textstyle\sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - Y_{i.})^2\}^2. \end{aligned} \tag{57}$$

Then, by Theorem 8, $I^{\frac{1}{2}}(\hat{\theta} - \sigma_A^{\,2}) \to_{\mathfrak{L}} \mathfrak{N}(0, \sigma^2)$ as $I \to \infty$, with $\sigma^2 = 4\zeta_1^{(1,1)} - 4\zeta_1^{(1,2)}/J + \zeta_1^{(2,2)}/J^2$, with $\zeta_1^{(i,j)}$ $i, j = 1, 2$ given as in Theorem 8. It is not necessary to compute σ^2 in order to use Theorem 9, and from Theorem 9, $s_g^{\,2} \to_P \sigma^2$ as $I \to \infty$. Hence, one can conclude that $I^{\frac{1}{2}}(\hat{\theta} - \sigma_A^{\,2})/s_g \to_{\mathfrak{L}} \mathfrak{N}(0, 1)$ as $I \to \infty$.

Note that this technique provides asymptotic confidence intervals for variance components, and yields symmetric intervals. To obtain the more reasonable intervals skewed to the right, one might want to get a confidence interval for $\log \sigma_A^{\,2}$.

Recall that in the grouping for the jackknife estimate, $N = nk$, it has been assumed that $k = 1$. As stated before, Theorems 8 and 9 remain valid with arbitrary values of k (with $n \to \infty$). However, when $k = 1$, the preceding paragraph can be obtained in a more straight forward fashion.

Note that when $k = 1$, the jackknife estimate of $\sigma_A^{\,2}$ and the unjackknifed estimate (54), are identical. Moreover,

$$\hat{\sigma}^2 = \operatorname{Var}(\hat{\theta}_A^{\,2}) = J^{-2}(\operatorname{Var}(MS_A) - 2\operatorname{Cov}(MS_A, MS_e) + \operatorname{Var}(MS_e)), \tag{58}$$

and one would assume that (58) could be estimated by

$$\sigma^2 = J^{-2}(I-1)^{-1} \textstyle\sum_{m=1}^{I} ((Z_m - \bar{Z}) - (W_m - \bar{W}))^2 \tag{59}$$

where $Z_m = J(Y_{m.} - Y_{..})^2, \quad \bar{Z} = (J/I) \sum_{i=1}^{I} (Y_{i.} - Y_{..})^2,$

$$W_m = (J - I)^{-1} \textstyle\sum_{j=1}^{J} (Y_{mj} - Y_{m.})^2, \qquad \bar{W} = (I(J-1))^{-1} \textstyle\sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - Y_{i.})^2.$$

Note that $s_g^{\,2}$ in (57) is essentially the same as $\hat{\theta}^2$ in (59).

These remarks appear to hold whenever the function g of Theorems 8 and 9 is linear. However, the jackknife seems to make a real contribution if one wants to make a test or confidence interval for $\theta = \sigma_A^{\,2}/\sigma_e^{\,2}$. For now $g(U^{(1)}, U^{(2)}) = (U^{(1)} - U^{(2)}/J)/U^{(2)}$ and the usual estimate is biased. Not only will the jackknife estimate reduce the bias, but the sum of squares will provide a consistent estimate of a rather messy quantity.

(b) *Jackknifing an admissible estimate.* The above estimates of σ_A^2 all have the undesirable property that with finite samples, there is positive probability of obtaining negative estimates (see Scheffé [20]). Recently, Portnoy [16] has obtained an estimate without this undesirable property. Of more interest is the fact that under the normal theory model of (51), his estimate is minmax, and admissible, with respect to squared error loss, among all estimates invariant under the transformation

$$Y_{ij} \to \alpha Y_{ij} + \beta, \qquad i = 1, \cdots, I, \qquad j = 1, \cdots, J, \alpha, \beta \quad \text{real numbers.}$$

Let
$$S_1 = \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - Y_{i.})^2, \qquad S_2 = J \sum_{i=1}^I (Y_{i.} - Y_{..})^2,$$
$$c = (IJ + 1)/2, \qquad d = (I + 1)/2, \qquad A = S_1/(S_1 + S_2),$$
$$\beta(a, b) = \Gamma(a) \cdot \Gamma(b)/\Gamma(a + b), \qquad \Gamma(a) = \int_0^\infty x^{s-1} e^{-X}\, dX,$$
$$I_A(a, b) = (\beta(a, b))^{-1} \int_0^A t^{a-1}(1 - t)^{b-1}\, dt,$$

the incomplete beta integral, and $C(m, n) = \binom{m}{n}$. Then Portnoy's estimate can be expressed in either of the following two equivalent ways:

$$\text{(i)} \quad \hat{\sigma}_{A_p}^2 = (2J)^{-1}\{S_2(1 - I_A(c - d, d))/(d(1 - I_A(c - d, d + 1))) - S_1(1 - I_A(c - d - 1, d + 1))/ ((c - d - 1)(1 - I_A(c - d, d + 1)))\} \quad \text{or}$$

$$\text{(ii)} \quad \hat{\sigma}_{A_p}^2 = (J)^{-1}\{S_2/(I + 1) - S_1/(I(J - 1) - 2) + (c - 1)(S_1 + S_2)A^{c-d}/(2cd(c - d - 1)H(S_1, S_2))\}$$

where $H(S_1, S_2)$

$$= ((S_1 + S_2)/S_2)^{d+1}(1 - I_A(c - d, d + 1))\beta(c - d, d + 1)$$
$$= \sum_{k=0}^\infty (-1)^k C(c - d - 1, k)(S_2/(S_1 + S_2))^k/(k + d + 1).$$

Portnoy proposes his estimate as a point estimate when the normal theory model of (51) holds. However, under the weaker assumptions of (54), one can also jackknife his estimate and obtain tests or confidence intervals for σ_A^2. The assumptions of Theorems 8 and 9 are satisfied if $\sigma_e^2 > 0$.

One might prefer to jackknife Portnoy's estimate, rather than the standard estimate, when the possibility of a negative estimate is considerable. Also note that Portnoy's estimate is biased, and the jackknife will reduce this bias (at the expense of losing the admissibility property).

(c) *Unbalanced case.* The results of Theorems 10 and 11 can be applied to the unbalanced one-way layout. Assume

$$(60) \quad Y_{ij} = \mu + a_i + e_{ij}, \qquad i = 1, \cdots, I, \qquad j = 1, \cdots, J_i(J_i \geqq 2),$$

where μ is a constant and the $\{a_i\}$ are independent random variables with mean zero, variance σ_A^2, and moments of all order, and the $\{e_{ij}\}$ are independent random variables with means zero, variances σ_e^2, and moments of all order, and the $I + \sum_{i=1}^I J_i$ random variables $\{a_i\}$ and $\{e_{ij}\}$ are independent.

Let $$X_i = \begin{pmatrix} Y_{i.} \\ (J_i - 1)^{-1} \sum_{j=1}^{J_i} (Y_{ij} - Y_{i\cdot})^2 \end{pmatrix}, \qquad i = 1, \cdots, I,$$

be I independent 2×1 vectors. Let $f^{*(1)}(X_{\alpha_1}, X_{\alpha_2}) = [(Y_{\alpha_1\cdot} - Y_{\alpha_2\cdot})^2 - \sigma_e^2(J_{\alpha_1}^{-1} + J_{\alpha_2}^{-1})]/2$ be a symmetric kernel for estimating σ_A^2, with corresponding U-"statistic", $U^{(1)} = (I-1)^{-1} \sum_{i=1}^I (Y_{i.} - Y_{..})^2 - (I^{-1} \sum_{i=1}^I J_i^{-1})\sigma_e^2$. Note that $U^{(1)}$ is not a statistic since it depends on the unknown σ_e^2. This is for mathematical convenience, and $U^{(2)}$ will correct this. Let $f^{*(2)}(X\alpha_1) = (J_{\alpha_1} - 1)^{-1} \sum_{j=1}^{J_{\alpha_1}} (Y_{\alpha_1 j} - Y_{\alpha_1\cdot})^2 - \sigma_e^2$ be a symmetric kernel to estimate zero, with corresponding U-"statistic," $U^{(2)} = I^{-1} \sum_{i=1}^I [(J_i - 1)^{-1} \sum_{j=1}^{J_i} (Y_{ij} - Y_{i.})^2] - \sigma_e^2$. Assume $K_I = I^{-1} \sum_{i=1}^I J_i^{-1} \to K$ as $I \to \infty$, with $0 \leqq K \leqq \frac{1}{2}$, since $J_i \geqq 2$, $i = 1, \cdots, I$.

Let $g(U^{(1)}, U^{(2)}) = U^{(1)} - KU^{(2)}$ be an estimate of σ_A^2. Note that $g(\cdot\,,\cdot)$ is a statistic, and Theorems 10 and 11 are satisfied if σ_e^2 is non-zero and $\max(J_1, \cdots, J_I)$ remains finite as $I \to \infty$. Hence one obtains the result that $I^{\frac{1}{2}}(\hat\theta - \sigma_A^2)/s_g \to_{\mathcal{L}} \mathfrak{N}(0, 1)$ as $I \to \infty$, with $\hat\theta$ the jackknife estimate of σ_A^2, and s_g^2 the sum-of-squares of the pseudo-values.

Note that if I is actually finite as in practice, replace K by K_I, but do not alter K_I during the jackknifing. In this case

$$(61)\quad \hat\theta = (I-1)^{-1} \sum_{i=1}^I (Y_{i.} - Y_{..})^2 - I^{-2}(\textstyle\sum_{i=1}^I J_i^{-1}) \sum_{i=1}^I \sum_{j=1}^{J_i} (Y_{ij} - Y_{i.})^2/(J_i - 1).$$

Tukey [22] shows that there are circumstances in which these weights are to be preferred over the more standard estimate in which rows are weighted proportionally to the number of elements. The weights in (61) are usually preferred if σ_A^2 is larger than σ_e^2.

Unfortunately, the most general form of the estimate for σ_A^2 which Tukey describes (arbitrary weights for rows, and for mean squares), does not seem compatible with the jackknife technique.

As before, a confidence interval or test for σ_A^2/σ_e^2 is possible using these methods. Moreover, the jackknife can be used on nested designs, both in the balanced and unbalanced cases. The extension is straightforward using the above techniques.

(d) *Jackknife applied to ratios and correlation coefficients.*

(i) *Ratio estimation.* Recall the work of Durbin [3] mentioned in the introduction. The problem he considered is a special case of the following. Consider the sequence of independent, identically distributed bivariate random variables

$$(62)\quad \binom{X_1}{Y_1}, \cdots, \binom{X_n}{Y_n} \text{ where } EX = \mu, \quad \operatorname{Var} X = \sigma^2, \quad EY = \eta \neq 0, \quad \operatorname{Var} Y = \tau^2, \quad \operatorname{Corr}(X, Y) = \rho,$$

$0 < \sigma^2, \tau^2 < \infty$, $-1 < \rho < 1$. Suppose one wants a test or confidence interval for

$\theta = \mu/\eta$, and all parameters are unknown. Then use of Theorems 8 and 9 on $g(\bar{X}, \bar{Y}) = \bar{X}/\bar{Y}$ enables one to state that $n^{\frac{1}{2}}(\hat{\theta} - \theta)/s_g \rightarrow_{\mathfrak{L}} \mathfrak{N}(0, 1)$ as $n \rightarrow \infty$ where $\hat{\theta}$ is the jackknife estimate of θ, and s_g as in (11).

(ii) *Correlation coefficients.* Consider the sequence of bivariate random variables presented in (62) without the stipulation that $\eta \neq 0$. Suppose one wants a test or confidence interval for ρ, the correlation coefficient. Then three U-statistics enter consideration. They are

$$U^{(1)} = (n-1)^{-1}\{\textstyle\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})\} \quad \text{with kernel}$$

$$f^{*(1)}(\binom{X_1}{Y_1}, \binom{X_2}{Y_2}) = (X_1 - X_2)(Y_1 - Y_2)/2,$$

$$U^{(2)} = (n-1)^{-2}\textstyle\sum_{i=1}^{n} (X_i - \bar{X})^2 \quad \text{with kernel}$$

$$f^{*(2)}(\binom{X_1}{Y_1}, \binom{X_2}{Y_2}) = (X_1 - X_2)^2/2, \quad \text{and}$$

$$U^{(3)} = (n-1)^{-1}\textstyle\sum_{i=1}^{N} (Y_i - \bar{Y})^2/2 \quad \text{with kernel}$$

$$f^{*(3)}(\binom{X_1}{Y_1}, \binom{X_2}{Y_2}) = (Y_1 - Y_2)^2/2.$$

Then use of Theorems 8 and 9 on $g(U^{(1)}, U^{(2)}, U^{(3)}) = U^{(1)}/(U^{(2)} \cdot U^{(3)})^{\frac{1}{2}}$ enables one to state that $n^{\frac{1}{2}}(\hat{\theta} - \theta)/s_g \rightarrow_{\mathfrak{L}} \mathfrak{N}(0, 1)$ as $n \rightarrow \infty$, where $\hat{\theta}$ is the jackknife estimate of θ, and s_g as in (11).

This result was previously obtained by Layard [9]. He also shows that the standard procedures for tests and confidence intervals on ρ are not robust against non-normality when $\rho \neq 0$.

The jackknife can also be used to obtain asymptotic tests and confidence intervals for partial and multiple correlation coefficients. The extension of the above is straightforward.

4. Two-sample problems.

(a) *Background.* An extension of Hoeffding's theorem to the two-sample case is possible. Let $X_1, \cdots, X_{n_1}$ be n_1 independent, identically distributed random variables, let $Y_1, \cdots, Y_{n_2}$ be n_2 independent, identically distributed random variables, and let the X's and Y's be independent. Let

$$\begin{aligned} f^*_{c_1c_2}&(x_1, \cdots, x_{c_1}; y_1, \cdots, y_{c_2}) \\ &= E\{f^*(X_1, \cdots, X_{m_1}; Y_1, \cdots, Y_{m_2}) \mid X_1 = x_1, \cdots, X_{c_1} = x_{c_1}; \\ &\quad Y_1 = y_1, \cdots, Y_{c_2} = y_{c_2}\}, \end{aligned}$$

$$\zeta_{c_1c_2} = \operatorname{Var}\{f^*_{c_1c_2}(X_1, \cdots, X_{c_1}; Y_1, \cdots, Y_{c_2})\}, \quad \text{and}$$

$$\sigma^2 = m_1^2\zeta_{10} + m_2^2(\lim (n_1/n_2))\zeta_{01}.$$

Then the following theorem is proved in Lehmann [11].

THEOREM 14. *Let $f^*(X_1, \cdots, X_{m_1}; Y_1, \cdots, Y_{m_2})$ be a real-valued statistic symmetric in the X's and symmetric in the Y's, with expectation η and with finite second moment. Let*

$$U = [\binom{n_1}{m_1}\binom{n_2}{m_2}]^{-1} \textstyle\sum_C f^*(X_{\alpha_1}, \cdots, X_{\alpha_{m_1}}; Y_{\beta_1}, \cdots, Y_{\beta_{m_2}}), \tag{63}$$

where $\sum_C$ *indicates summation over all combinations* $(\alpha_1, \cdots, \alpha_{m_1})$ *from* $(1, \cdots, n_1)$ *and all combinations* $(\beta_1, \cdots, \beta_{m_2})$ *from* $(1, \cdots, n_2)$. *Then if* $n_1 \leqq n_2$, *and* $n_1 \to \infty$ *such that* lim (n_1/n_2) *exists*, $n_1^{\frac{1}{2}}(U - \eta) \to_{\mathcal{L}} \mathfrak{N}(0, \sigma^2)$.

Again it is possible to obtain an a.s. convergence result.

THEOREM 15. *Let* $f^*(X_1, \cdots, X_{m_1}; Y_1, \cdots, Y_{m_2})$ *be a real-valued statistic symmetric in the X's and symmetric in the Y's, with expectation* η *and* $E|f^*(X_1, \cdots, X_{m_1}; Y_1, \cdots, Y_{m_2})| < \infty$. *Let U be the U-statistic defined in* (63), *then if* $n_1 \leqq n_2$, $U \to_{\text{a.s.}L_1} \eta$ *as* $n_1 \to \infty$.

PROOF. The proof of Theorem 4 is readily extended. One only needs to make a straightforward generalization of the Hewitt-Savage theorem (see Feller [4]) to the two-sample case.

It is possible, in a natural way, to extend the jackknife technique to two-sample problems. As with the one-sample jackknife, split the two sets of observations into groups. Thus suppose one obtains $X_1, \cdots, X_{N_1}$ from the first population, and $Y_1, \cdots, Y_{N_2}$ from the second population. Next, let $N_1 = n_1 k_1$, $N_2 = n_2 k_2$ (all integers) and split the X's into n_1 groups of k_1 observations each, and the Y's into n_2 groups of k_2 observations each. Let $\hat{\theta}^0_{n_1,n_2}$ be the estimate of θ based on all the observations, and let $\hat{\theta}^{i,\cdot}_{n_1-1,\cdot}$ denote the estimate obtained after deletion of the ith group of X's, $i = 1, \cdots, n_1$, and let $\hat{\theta}^{\cdot,j}_{\cdot,n_2-1}$ denote the estimate obtained after deletion of the jth group of Y's, $j = 1, \cdots, n_2$. Next, let

$$(64)\qquad \begin{aligned} \hat{\theta}_{i,\cdot} &= n_1 \hat{\theta}^0_{n_1,n_2} - (n_1 - 1)\hat{\theta}^{i,\cdot}_{n_1-1,\cdot}, \qquad & i &= 1, \cdots, n_1, \\ \hat{\theta}_{\cdot,j} &= n_2 \hat{\theta}^0_{n_1,n_2} - (n_2 - 1)\hat{\theta}^{\cdot,j}_{\cdot,n_2-1}, \qquad & j &= 1, \cdots, n_2, \end{aligned}$$

and define the jackknife estimate of θ to be

$$(65)\qquad \hat{\theta} = \left(\textstyle\sum_{i=1}^{n_1} \hat{\theta}_{i,\cdot} + \sum_{j=1}^{n_2} \hat{\theta}_{\cdot,j}\right)/(n_1 + n_2).$$

Let ${}_1\hat{\theta}_{\cdot\cdot} = n_1^{-1} \sum_{i=1}^{n_1} \hat{\theta}_{i,\cdot}$ and ${}_2\hat{\theta}_{\cdot\cdot} = n_2^{-1} \sum_{j=1}^{n_2} \hat{\theta}_{\cdot,j}$. Then a sum of squares can be defined by

$$(66)\qquad s_g^2 = n_1\left((n_1(n_1 - 1))^{-1} \textstyle\sum_{i=1}^{n_1} (\hat{\theta}_{i,\cdot} - {}_1\hat{\theta}_{\cdot\cdot})^2 + (n_2(n_2 - 1))^{-1} \sum_{j=1}^{n_2} (\hat{\theta}_{\cdot,j} - {}_2\hat{\theta}_{\cdot\cdot})^2\right).$$

The following theorem extends Theorems 8 and 9 to the two-sample case. As before, only the case where $k_1 = k_2 = 1$ will be considered.

THEOREM 16. *Let* $X_1, \cdots, X_{n_1}$ *be* n_1 *independent identically distributed random vectors of p components, let* $Y_1, \cdots, Y_{n_2}$ *be* n_2 *independent identically distributed random vectors of p components, and let the X's and Y's be independent. Let* $f^{*(j)}(X_1, \cdots, X_{m_j^1}; Y_1, \cdots, Y_{m_j^2})$ *be a real-valued statistic symmetric in the X's and symmetric in the Y's, with expectation* $\eta^{(j)}$ *and finite second moment for* $j = 1, \cdots, q$. *Let* $U^{(j)}$ *be as in* (63) *for* $j = 1, \cdots, q$, *and let g be a real-valued function defined on* R^q, *which in a nighborhood of* $(\eta^{(1)}, \cdots, \eta^{(q)})$ *has bounded second partial derivatives. Let the jackknife estimate,* $\hat{\theta}$, *be defined as in* (64), *with* $\hat{\theta}^0_{n_1,n_2} = g(U^{(1)}, \cdots, U^{(q)})$. *Then if* $n_1 \leqq n_2$, *and* $n_1 \to \infty$ *such that* lim (n_1/n_2) *exists*, $n_1^{\frac{1}{2}}(\hat{\theta} - \theta)$ *is asymptotically normally distributed with mean zero and variance.*

$$(67)\qquad \sigma^2 = \lim_{n_1\to\infty} \left\{\textstyle\sum_{i=1}^{q} \sum_{j=1}^{q} g_i\, g_j \{m_i^1 m_j^1 \zeta_{10}^{(i,j)} + (n_1/n_2) m_i^2 m_j^2 \zeta_{01}^{(i,j)}\}\right\}$$

where $\zeta_{c_1c_2}^{(i,j)} = \text{Cov}\,(f_{c_1c_2}^{*(i)}(X_1, \cdots, X_{c_1}; Y_1, \cdots, Y_{c_2}), f_{c_1c_2}^{*(j)}(X_1, \cdots, X_{c_1}; Y_1, \cdots, Y_{c_2}))$, *and* g_i, $i = 1, \cdots, q$ *as in* (27). *In addition*, $s_g^2 \to_P \sigma^2$.

PROOF. Use of Theorem 14, combined with Theorem 8 proves the asymptotic normality almost immediately. (30) is readily extended, and one only needs to note that

$$(68)\quad S_{n_1}^{(k,l)} = (n_1 - 1)\sum_{i=1}^{n_1} (U_{\sim i}^{(k)}. - U^{(k)})(U_{\sim i}^{(l)}. - U^{(l)}) \to_P \zeta_{10}^{(k,l)} m_k^1 m_l^1,$$

and

$$(69)\quad S_{n_2}^{(k,l)} = (n_2 - 1)\sum_{j=1}^{n_2} (U_{\cdot,\sim j}^{(k)} - U^{(k)})(U_{\cdot,\sim j}^{(l)} - U^{(\ )}) \to_P \zeta_{01}^{(k,l)} m_k^2 m_l^2,$$

where the notation is an obvious extension of (29). This follows from the assumption of finite second moments on $f^{*(j)}(X_1, \cdots, X_{m_j1}; Y_1, \cdots, Y_{m_j2})$ and Theorem 15.

Use of (68), (69), and Theorem 7 proves the consistency of s_g^2.

(b) *Application to* ANOVA. The model appropriate to a Model II ANOVA two-way layout is given by

$$(70)\quad Y_{ijk} = \mu + a_i + b_j + c_{ij} + e_{ijk}, \qquad i = 1, \cdots, I \qquad j = 1, \cdots, J, \qquad k = 1, \cdots, K,$$

where the $\{a_i\}$, $\{b_j\}$, $\{c_{ij}\}$, and $\{e_{ijk}\}$ are independently normal, with zero means and respective variances σ_A^2, σ_B^2, σ_{AB}^2, σ_e^2. The drawbacks of relying too heavily on the normality have already been discussed in Section 3. Another strong objection can be raised against the assumptions on the interactions in the above model. That the interaction c_{ij} between the ith factor at the first level and the jth factor at the second level is independent of the ith factor or the jth factor is contrary to the intuitive notion of interaction.

In light of this discussion, a more appropriate model for a Model II ANOVA two-way layout would be

$$(71)\quad Y_{ijk} = \mu + a(u_i) + b(v_j) + c(u_i, u_j) + e(u_i, v_j)_k \qquad i = 1, \cdots, I, \quad j = 1, \cdots, J, \quad k = 1, \cdots, K,$$

where

(i) the u_i's, $i = 1, \cdots, I$ are an independent sample from some infinite population,

(ii) the v_j's, $j = 1, \cdots, J$ are an independent sample from some infinite population, and

(iii) the two populations are independently sampled (u_i and v_j are independent for all (i, j)).

The $a(\cdot)$, $b(\cdot)$, $c(\cdot, \cdot)$ represent the two main and interaction random effects, and the $\{e(u_i, v_j)_k\}$ form the error variance, while μ is the overall mean. Using (i), (ii), and (iii), the following are reasonable additional assumptions. The $\{a(u_i)\}$, $\{b(v_j)\}$, and $\{e(u_i, v_j)_k\}$ are independent random variables with zero means, respective variances σ_A^2, σ_B^2, σ_e^2, and finite fourth moments. In addition, the $\{c(u_i, v_j)\}$ are uncorrelated random variables, are uncorrelated with the other random variables, have zero mean, variance σ_{AB}^2, and finite fourth moments. The $\{c(u_i\ v_j)\}$ are also independent of all $\{e(u_i, v_j)_k\}$.

Note that this is essentially the model described in Scheffé [20] up to the point where he makes the assumption of normality. The latter assumption implies independence of the interactions, a highly suspect assumption as mentioned above. For example, if the interactions were of the form $c_{ij} = a_i \cdot b_j$, $i = 1, \cdots, J$ $j = 1, \cdots, J$, they would satisfy model (71) but not model (70).

The model of (71) was treated by Cornfield and Tukey [2]. They obtained the expected values of the usual mean squares under (71). Also, Hooke [7] and [8] studied this model using bipolykays. This technique seems quite difficult to utilize in practice. Use of the jackknife technique will be shown to adequately handle the problem.

First note that if no interactions are present in (71), then construction of tests or confidence intervals for σ_B^2 or σ_B^2/σ_e^2 can be accomplished using Theorems 8 and 9. Let

$$X_j = \begin{pmatrix} Y_{\cdot j \cdot} \\ (I(K-1))^{-1} \sum_{i=1}^{I} \sum_{k=1}^{K} (Y_{ijk} - Y_{ij\cdot})^2 \end{pmatrix}, \qquad j = 1, \cdots, J,$$

and

$$f^{*(1)}(X_1, X_2) = (Y_{\cdot 1 \cdot} - Y_{\cdot 2 \cdot})^2/2$$

with corresponding U-statistic,

$$U^{(1)} = (J-1)^{-1} \sum_{j=1}^{J} (Y_{\cdot j \cdot} - Y_{\cdots})^2.$$

Let

$$f^{*(2)}(X_1) = (I(K-1))^{-1} \sum_{i=1}^{I} \sum_{k=1}^{K} (Y_{ijk} - Y_{ij\cdot})^2$$

with corresponding U-statistic,

$$U^{(2)} = (IJ(K-1))^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (Y_{ijk} - Y_{ij\cdot})^2.$$

Then if $g(U^{(1)}, U^{(2)}) = U^{(1)} - U^{(2)}/(IK)$, $J^{\frac{1}{2}}(\hat{\theta} - \sigma_B^2) \rightarrow_{\mathcal{L}} \mathfrak{N}(0, \sigma^2)$, as $J \rightarrow \infty$, and $s_g^2 \rightarrow_P \sigma^2$ where $\hat{\theta}$ is the jackknife estimate of σ_B^2 as in (10), and s_g^2 is the sum of squares as in (11). Similar remarks hold for σ_B^2/σ_e^2.

Return to the model described in (71) where interactions are present. In what follows, U-statistics will be based on the unobservable random effects, however the U-statistics themselves will be functions of $\{Y_{ijk}\}$, and hence statistics. Let

$$\begin{aligned} f^{*(1)}&(u_{\alpha_1}, u_{\alpha_2}; v_{\beta_1}, v_{\beta_2}) \\ &= [e(u_{\alpha_2}, v_{\beta_1}) + e(u_{\alpha_1}, v_{\beta_1})_\cdot - c(u_{\alpha_1}, v_{\beta_2}) - e(u_{\alpha_1}, v_{\beta_2})_\cdot - c(u_{\alpha_2}, v_{\beta_1}) \\ &\quad - e(u_{\alpha_1}, v_{\beta_1})_\cdot + c(u_{\alpha_2}, v\beta_2)_\cdot + e(u_{\alpha_2}, v_{\beta_2})_\cdot]^2/4 \\ &= [(Y_{\alpha_1\beta_1} - Y_{\alpha_1\beta_2\cdot}) - (Y_{\alpha_2\beta_1\cdot} - Y_{\alpha_2\beta_2\cdot})]^2/4, \end{aligned}$$

which generates the U-statistic,

$$\begin{aligned} U^{(1)} = {} & [(I-1)(J-1)]^{-1} \\ & \cdot \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij\cdot} - Y_{i\cdot\cdot} - Y_{\cdot j \cdot} + Y_{\cdots})^2. \end{aligned}$$

Let

$$\begin{aligned} f^{*(2)}(u_{\alpha_1}; v_{\beta_1}) &= (K-1)^{-1} \sum_{k=1}^{K} (e(u_{\alpha_1}, v_{\beta_1})_k - e(u_{\alpha_1}, v_{\beta_1})_\cdot)^2 \\ &= (K-1)^{-1} \sum_{k=1}^{K} (Y_{\alpha_1\, \beta_1\, k} - Y_{\alpha_1\, \beta_1 \cdot})^2 \end{aligned}$$

which generates the U-statistic

$$U^{(2)} = [IJ(K-1)]^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (Y_{ijk} - Y_{ij\cdot})^2.$$

Let $g(U^{(1)}, U^{(2)}) = U^{(1)} - U^{(2)}/K$ be an estimate of σ_{AB}^2, and let $\hat{\theta}$ be the jackknife estimate and s_g^2 the sum of squares as defined in (66). Then by Theorem 16, if $I \leqq J$, $I^{\frac{1}{2}}(\hat{\theta} - \sigma_{AB}^2) \rightarrow_{\mathcal{L}} \mathfrak{N}(0, \sigma^2)$, and $s_g^2 \rightarrow_P \sigma^2$ as $I \rightarrow \infty$, where $\sigma^2 = 4\zeta_{10}^{(1,1)} + 4 \lim (I/J)\zeta_{01}^{(1,1)}$ (the other terms of (67) are found to be zero).

One should check that $0 < \sigma^2 < +\infty$. Note that

$$f_{10}^{*(1)}(u_{\alpha_1}) = [2E(c^2(u_{\alpha_1}, v_{\beta_1}) \mid u_{\alpha_1}) + 2\sigma_{AB}^2 + 4\sigma_e^2/K]/4.$$

Hence one needs to assume that $E(c^2(u_{\alpha_1}, v_{\beta_1}) \mid u_{\alpha_1})$ is a random variable with positive variance. This is a reasonable assumption if one feels that the interactions somehow depend on the factors. Note that this is not the case if the interactions are independent. A similar argument treats $\zeta_{01}^{(1,1)}$, and hence the jackknife works well.

Similar arguments hold if one wishes to estimate σ_{AB}^2/σ_e^2 by $g(U^{(1)}, U^{(2)}) = U^{(1)}/U^{(2)} - K^{-1}$.

To apply this technique to obtain tests or confidence intervals for σ_B^2 in (71) is somewhat more difficult. A natural way would be to let

$$\begin{aligned} f^{*(3)}&(u_1, u_2, \cdots, u_I; v_{\beta_1}, v_{\beta_2}) \\ &= \{I^{-1} \textstyle\sum_{i=1}^{I} (b(v_{\beta_1}) + c(u_i, v_{\beta_1}) + e(u_i, v_{\beta_1})\cdot) \\ &\quad - I^{-1} \textstyle\sum_{i=1}^{I} (b(v_{\beta_2}) + c(u_i, v_{\beta_2}) + e(u_i, v_{\beta_2})\cdot)\}^2/2 \\ &= \{I^{-1} \textstyle\sum_{i=1}^{I} Y_{i\beta_1\cdot} - I^{-1} \textstyle\sum_{i=1}^{I} Y_{i\beta_2\cdot}\}^2/2, \end{aligned}$$

an $I \times 2$ kernel with U-statistic

$$U^{(3)} = (J-1)^{-1} \sum_{j=1}^{J} (Y_{\cdot j\cdot} - Y_{\cdots})^2.$$

However, it is tacitly assumed in Theorem 16 that the kernel used in constructing the U-statistic is of finite size. The theorem can be extended to consider kernels like $f^{*(3)}(\cdot\,;\,\cdot)$, however the following is a simpler approach.

Let $$\begin{aligned} f^{*(4)}(u_{\alpha_1}; v_{\beta_1}) &= (\mu + a(u_{\alpha_1}) + b(v_{\beta_1}) + c(u_{\alpha_1}, v_{\beta_1}) + e(u_{\alpha_1}, v_{\beta_1})\cdot)^2 \\ &= Y_{\alpha_1\beta_1}^2. \end{aligned}$$

be a scalar kernel with U-statistic,

$$U^{(4)} = (IJ)^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij\cdot}^2.$$

Let

$$\begin{aligned} f^{*(5)}(u_{\alpha_1}, u_{\alpha_2}; v_{\beta_1}) &= (\mu + a(u_{\alpha_1}) + b(v_{\beta_1}) + c(u_{\alpha_1}, v_{\beta_1}) + e(u_{\alpha_1}, v_{\beta_1})\cdot) \\ &\quad \cdot (\mu + a(u_{\alpha_2}) + b(v_{\beta_1}) + c(u_{\alpha_2}, v_{\beta_1}) + e(u_{\alpha_2}, v_{\beta_1})\cdot) \\ &= Y_{\alpha_1\beta_1\cdot} Y_{\alpha_1\beta_1\cdot}. \end{aligned}$$

be a 2×1 kernel with U-statistic,

$$U^{(5)} = 2(I(I-1)J)^{-1} \sum_{i<l} \sum_{j=1}^{J} Y_{ij\cdot} \cdot Y_{lj\cdot}\,.$$

Let $f^{*(6)}(u_{\alpha_1}; v_{\beta_1}, v_{\beta_2}) = (\mu + a(\mu_{\alpha_1}) + b(v_{\beta_1}) + c(u_{\alpha_1}, v_{\beta_1}) + e(u_{\alpha_1}, v_{\beta_1})\cdot)$
$\cdot (\mu + a\ (u_{\alpha_1}) + b(v_{\beta_2}) + c(u_{\alpha_1}, v_{\beta_2}) + e(u_{\alpha_1}, v_{\beta_2})\cdot)$
$= Y_{\alpha_1\beta_1\cdot} \cdot Y_{\alpha_1\beta_2\cdot}$

be a 1×2 kernel with U-statistic,

$$U^{(6)} = 2(IJ(J-1))^{-1} \sum_{i=1}^{I} \sum_{j<l} Y_{ij\cdot} \cdot Y_{il\cdot}\,.$$

Let $f^{*(7)}(u_{\alpha_1}, u_{\alpha_2}; v_{\beta_1}, v_{\beta_2}) = (\mu + a(u_{\alpha_1}) + b(v_{\beta_1}) + c(u_{\alpha_1}, v_{\beta_1}) + e(u_{\alpha_1}, v_{\beta_1})\cdot)$
$\cdot (\mu + a(u_{\alpha_2}) + b(v_{\beta_2}) + c(u_{\alpha_2}, v_{\beta_2}) + e(u_{\alpha_2}, v_{\beta_2})\cdot)$
$= Y_{\alpha_1\beta_1\cdot} \cdot Y_{\alpha_2\beta_2\cdot}$

be a 2×2 kernel with U-statistic,

$$U^{(7)} = 4(I(I-1)J(J-1))^{-1} \sum_{i<l} \sum_{j<k} Y_{ij\cdot} \cdot Y_{lk\cdot}\,.$$

Then note that

$$U^{(3)} = I^{-1}U^{(4)} + (I-1)I^{-1}U^{(5)} - I^{-1}U^{(6)} - I(I-1)^{-1}U^{(7)},$$

and as an estimate of σ_B^2,

$$\hat{\theta}_B^2 = g(U^{(1)}, U^{(4)}, U^{(5)}, U^{(6)}, U^{(7)}) = U^{(3)} - U^{(1)}/I. \tag{72}$$

Note that this is the usual estimate (see Scheffé [20]). Then by Theorem 16, if $I \leqq J$, $I^{\frac{1}{2}}(\hat{\theta} - \sigma_B^2) \to_{\mathfrak{L}} \mathfrak{N}(0, \sigma^2)$, $s_g^2 \to_P \sigma^2$ as $I \to \infty$, with $\hat{\theta}$ and s_g^2 as in (66). Moreover one can show that

$$\sigma^2 = 4\zeta_{10}^{(5,5)} + \lim\ (I/J)\zeta_{01}^{(5,5)}.$$

Again one should check that $0 < \sigma^2 < +\infty$. To this end, note that

$$f_{10}^{*(5)}(u_{\alpha_1}) = \mu^2 + \sigma_B^2 + \sigma_e^2/K + E\{c(u_{\alpha_1}, v_{\beta_1}) \cdot c(u_{\alpha_2}, v_{\beta_1}) \mid u_{\alpha_1}\},$$

and $f_{01}^{*(5)}(v_{\beta_1}) = b^2(v_{\beta_1}) + \sigma_e^2/K$.

But since

$$E\{c(u_{\alpha_1}, v_{\beta_1}) \cdot c(u_{\alpha_2}, v_{\beta_1}) \mid u_{\alpha_1}\} = E\{E\{c(u_{\alpha_1}, v_{\beta_1}) \cdot c(u_{\alpha_2}, v_{\beta_1}) \mid u_{\alpha_1}, v_{\beta_1}\} \mid u_{\alpha_1}\} = 0,$$

a different normalization is needed. That is, $J^{\frac{1}{2}}(\hat{\theta} - \sigma_B^2) \to_{\mathfrak{L}} \mathfrak{N}(0, \sigma^2)$ and $s_g^2 \to_P \sigma^2$, where $\sigma^2 = \zeta_{01}^{(5,5)} = \text{Var}\ (b^2(v_{\beta_1}))$. However, note that one needs $I \to \infty$ for this result to hold in general. Again one can obtain tests or confidence intervals for the quantity σ_B^2/σ_e^2.

Acknowledgment. The author would like to express his sincere appreciation to Professor R. G. Miller, Jr. for his inspiration and guidance during the course of this investigation.

REFERENCES

[1] Berk, R. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.* **37** 51–58.

[2] Cornfield, J. and Tukey, J. W. (1956). Average values of mean squares in factorials. *Ann. Math. Statist.* **27** 907–949.

[3] Durbin, J. (1959). A note on the application of Quenoille's method of bias reduction to the estimation of ratios. *Biometrika* **46** 477–480.

[4] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

[5] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics*. Wiley, New York.

[6] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[7] Hooke, R. (1956a). Symmetric functions of a two-way array. *Ann. Math. Statist.* **27** 55–79.

[8] Hooke, R. (1956). Some applications of bipolykays to the estimation of variance components and their moments. *Ann. Math. Statist.* **27** 80–98.

[9] Layard, M. (1968). Unpublished thesis. Stanford University.

[10] Lehmann, E. L. (1959). *Testing Statistical Hypothesis*. Wiley, New York.

[11] Lehmann, E. L. (1951). Consistency and unbiasedness of certain non-parametric tests. *Ann. Math. Statist.* **22** 165–179.

[12] Mann, H. B. and Wald, A. (1943). On stochastic limit and order relationships. *Ann. Math. Statist.* **14** 217–226.

[13] Miller, R. G., Jr. (1964). A trustworthy jackknife. *Ann. Math. Statist.* **35** 1594–1605.

[14] Miller, R. G., Jr. (1968). Jackknifing variances. *Ann. Math. Statist.* **39** 567–582.

[15] Mosteller, F. and Tukey, J. W. (1965). Data analysis, including statistics. *Handbook of Social Psychology*.

[16] Portnoy, S. (1968). Formal Bayes estimation with application to a random effects analysis of variance model. Technical Report No. 42 Stanford Univ.

[17] Quenouille, M. H. (1949). Approximate tests of correlation in time-series. *J. Roy. Statist. Soc. Ser. B.* **11** 68–84.

[18] Quenouille, M. H. (1956). Notes on bias in estimation, *Biometrika* **43** 353–360.

[19] Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, **2** 110–114.

[20] Scheffe, H. (1959). *Analysis of Variance*. Wiley, New York.

[21] Spjøtvoll, E. (1967). Optimal invariant tests in unbalanced variance components models. *Ann. Math. Statist.* **38** 422–429.

[22] Tukey, J. W. (1957). Variances of variance components: II. The unbalanced single classification. *Ann. Math. Statist.* **28** 43–56.