The Annals of Statistics
1975, Vol. 3, No. 5, 1122–1134

ASYMPTOTICALLY ROBUST TESTS IN UNBALANCED VARIANCE COMPONENT MODELS

BY JAMES N. ARVESEN[1] AND MAXWELL W. J. LAYARD

Pfizer Pharmaceuticals, Inc., and Columbia University; National Cancer Institute

Spjøtvoll [1967] has obtained a test associated with an unbalanced one-way layout for Model II ANOVA. Under the assumption of normality, his test possesses several optimum properties. Without the normality assumption, the significance level is (in general) highly nonrobust. An attempt to remedy this situation, using a test based on the jackknife technique, appears in Arvesen [1969]. The present paper proposes as an alternative a jackknifed version of Spjøtvoll's test. The new test is not sensitive to departures from normality, and Monte Carlo sampling and asymptotic efficiency results suggest that it is more powerful than Arvesen's test. The paper also includes some general results for use of the jackknife technique with nonidentically distributed random variables.

1. Introduction. We obtain in this paper an asymptotically robust test for hypothesis $\Delta \leqq \Delta_0$ against $\Delta > \Delta_0$, where Δ is the variance ratio in an unbalanced one-way layout for Model II ANOVA. The test is based on an extension of Arvesen (1969) using the jackknife. Theoretical and Monte Carlo results show the robustness of the proposed test for nonnormal data, and that it performs similarly to Spjøtvoll's (1967) test if the data are normal. It is also possible to obtain a confidence interval for Δ using the proposed test. Section 2 discusses the basic model, while Section 3 digresses to discuss some general results concerning use of the jackknife with nonidentically distributed random variables. Section 4 applies the results of the previous section to the variance component problem, while Sections 5 and 6 discuss asymptotic efficiency results and Monte Carlo results respectively.

2. The model. The basic model assumed in an unbalanced one-way layout for Model II ANOVA is

$$Y_{ij} = \mu + a_i + e_{ij}\,, \qquad i = 1, \cdots, n\,, \quad j = 1, \cdots, J_i\,, \tag{2.1}$$

where μ is an unknown constant, $\{a_i\}$ and $\{e_{ij}\}$ are all mutually independent normal random variables with zero means and variances σ_A^2 and σ_e^2 respectively. If we let $\Delta = \sigma_A^2/\sigma_e^2$, one hypothesis of interest is

$$H_0: \Delta \leqq \Delta_0 \quad \text{vs.} \quad H_A: \Delta > \Delta_0\,. \tag{2.2}$$

One reason for considering such an hypothesis is that we may be interested in

Received October 1971; revised February 1974.

[1] This author's research was supported by the Purdue Research Foundation at Purdue University.

AMS 1970 *subject classifications*. Primary 62G32; Secondary 62E20, 62E25.

Key words and phrases. Variance components, jackknife, *U*-statistics.

the proportion of the total variation of the Y_{ij} which is due to the random effects a_i. This proportion is $\sigma_A^2/(\sigma_A^2 + \sigma_e^2)$, which quantity is also the intraclass correlation coefficient.

For a specified alternative $\Delta = \Delta_1$, Spjøtvoll (1967) has obtained the most powerful invariant similar α-level test of H_0. The value Δ_1 enters into the test statistic. He also proposes an alternate test letting $\Delta_1 \to \infty$, tantamount to achieving high power for distant alternatives. It is this test we now consider. Letting

$$(2.3) \qquad T = \sum_{i=1}^{n} J_i(\Delta_0 J_i + 1)^{-1}(\bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet})^2 (\sum_{i=1}^{n} \sum_{j=1}^{J_i} (Y_{ij} - \bar{Y}_{i\bullet})^2)^{-1}$$

where

$$Y_{i\bullet} = J_i^{-1} \sum_{j=1}^{J_i} Y_{ij}\,, \qquad \text{and}$$
$$Y_{\bullet\bullet} = (\sum_{i=1}^{n} J_i(\Delta_0 J_i + 1)^{-1})^{-1} \sum_{i=1}^{n} J(\Delta_0 J_i + 1)^{-1}\bar{Y}_{i\bullet}\,,$$
$$N^* = \sum_{i=1}^{n} J_i\,,$$

one rejects H_0 at the α-level if

$$(2.4) \qquad (N^* - n)(n-1)^{-1}T > F_{\alpha;n-1,N^*-n}\,,$$

where $F_{\alpha;\nu_1,\nu_2}$ denotes the upper α point of an F distribution with ν_1, ν_2 degrees of freedom.

Note that when $J_i = J$, the test given by (2.4) is the same as the standard F-test given in Scheffé (1959). It is well-known (see Scheffé (1959)) that this standard F-test is not robust if the observations are nonnormal, especially the random effect terms. The significance levels are invalid except in the case $\Delta_0 = 0$. Spjøtvoll also obtains a confidence interval for Δ, although it is subject to the same criticism as the test concerning its nonrobust character.

In the balanced case, a competitor based on the jackknife has been proposed in Arvesen (1969), and its moderate sample size properties were examined by a Monte Carlo computer simulation in Arvesen and Schmitz (1970). Also, in the unbalanced case, a test based on the jackknife was proposed in Arvesen (1969). This test will be further discussed in Sections 5 and 6, where evidence is presented suggesting that in terms of the power it is inferior to the test proposed in this paper, which is based on jackknifing the logarithm of Spjøtvoll's statistic (2.3). The computation of the proposed test is described in Section 4.

3. The jackknife for nonidentically distributed random variables.

(a) Background. First let us describe the jackknife procedure. For a more detailed discussion the reader is referred to Miller (1964). Let $X_1, \cdots, X_N$ be independent identically distributed observations from the cdf F_θ. Partition these N observations into n groups with k observations in each group ($N = nk$). Then if $\hat{\theta}_n^0$ is some estimate based on all n groups of observations (all N observations), let $\hat{\theta}_{n-1}^i$, $i = 1, \cdots, n$ denote the estimate obtained after deletion of the ith group of observations.

The jackknife estimate of θ is

$$\hat{\theta} = n^{-1} \sum_{i=1}^{n} \hat{\theta}_i , \tag{3.1}$$

where

$$\hat{\theta}_i = n\hat{\theta}_n{}^0 - (n-1)\hat{\theta}_{n-1}^i , \quad i = 1, \cdots, n . \tag{3.2}$$

If

$$s_{\hat{\theta}}^2 = (n-1)^{-1} \sum_{i=1}^{n} (\hat{\theta}_i - \hat{\theta})^2 , \tag{3.3}$$

it is interesting to conjecture that if n is held fixed,

$$t = n^{\frac{1}{2}}(\hat{\theta} - \theta)/s_{\hat{\theta}} \rightarrow_{\mathscr{L}} t_{n-1} \qquad N \rightarrow \infty . \tag{3.4}$$

Miller (1964), (1968), and Arvesen (1969) give a large class of situations where this conjecture proves valid. In what follows, we will assume $k = 1$, and hence $N = n$, and the convergence in (3.4) is to a standard normal distribution.

If one starts with $X_1, \cdots, X_n$ independent but not necessarily identically distributed, the situation becomes more complicated. Some results for this case were given in Arvesen (1969), but those results required $\hat{\theta}_n{}^0$ to be of a very special restrictive form. In the notation of that paper, let $f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})$ be a symmetric kernel with the same expectation $E[f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})] = \eta$ for all $\alpha_1, \cdots, \alpha_m$. The U-statistic (see Hoeffding (1948)) for estimating η is

$$U_n{}^0 = \binom{n}{m}^{-1} \sum_{C_n} f^*(X_{\alpha_1}, \cdots, X_{\alpha_m}) \tag{3.5}$$

where C_n indicates the summation is over all combinations $\alpha_1, \cdots, \alpha_m$ of m integers chosen from $1, \cdots, n$. Theorems 10 and 11 of Arvesen (1969) show that under mild regularity conditions, the conjecture of (3.4) is valid with $\hat{\theta}_n{}^0 = g(U_n{}^0)$, $\theta = g(\eta)$. However for many purposes, including those to be discussed in Section 4 below, the requirement that the kernels have the same expectation is too restrictive.

(b) A modified jackknife estimate. To modify this restriction, let $X_1, \cdots, X_n$ be independent (not necessarily identically distributed) random variables, and assume

$$E[f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})] = \eta_{\alpha_1 \cdots \alpha_m} \mu \tag{3.6}$$

where $\eta_{\alpha_1 \cdots \alpha_m}$ is a *known* constant, and μ is an *unknown* parameter. Also, let $\eta_n{}^0 = \binom{n}{m}^{-1} \sum_{C_n} \eta_{\alpha_1 \cdots \alpha_m}$, $\bar{\eta} = \lim_{n \to \infty} \eta_n{}^0$ (which we assume exists, is finite and nonzero), and $\eta_{n-1}^i = \binom{n-1}{m}^{-1} \sum_{C_{n-1}^i} \eta_{\beta_1{}^i \cdots \beta_m{}^i}$ where $\sum_{C_{n-1}^i}$ indicates the sum is over all combinations of m integers $(\beta_1{}^i, \cdots, \beta_m{}^i)$ chosen from $(1, \cdots, i-1, i+1, \cdots, n)$. In what follows, definitions of symbols not explicitly given may be found in Arvesen (1969).

Let

$$\hat{\theta}_n{}^0 = g(U_n{}^0/\eta_n{}^0) , \qquad \hat{\theta}_{n-1}^i = g(U_{n-1}^i/\eta_{n-1}^i) , \qquad i = 1, \cdots, n , \tag{3.7}$$
$$\theta = g(\mu) .$$

Then following (3.1), (3.2) and (3.3), we will be interested in the statistic $n^{\frac{1}{2}}(\hat{\theta} - \theta)/s_{\hat{\theta}}$, and conditions under which it is asymptotically standard normal.

We begin with two lemmas.

LEMMA 1. *Let* $X_1, \cdots, X_n$ *be independent random variables, and assume for some* $0 < A < \infty$ *that* $|n_{\alpha_1 \cdots \alpha_m}| < A$, *and for some* $0 < B < \infty$, *and for some* $\delta > 0$ *that* $E|f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})|^{2+\delta} < B$. *Let*

$$h_{1(\nu)}(X_\nu) = \binom{n-1}{m-1}^{-1} \sum_{\neq \nu} (f^*_{1;\beta_1,\cdots,\beta_{m-1}}(X_\nu) - \eta_{\nu\beta_1\cdots\beta_{m-1}}\mu) \tag{3.8}$$

where the sum is over all sets $(\beta_1, \cdots, \beta_{m-1})$ *chosen from the first n integers excluding the integer* ν. *Then if* $\zeta_{1,n} \to \zeta_1$ *as* $n \to \infty$, $0 < \zeta_1 < +\infty$, $Z_n = n^{-1} \sum_{i=1}^n (h_{1(i)}(X_i))^2 \to_P \zeta_1$.

PROOF. First note that

$$E(Z_n) = n^{-1}\binom{n-1}{m-1}^{-2} \sum_{i=1}^n \sum_{(i)} \zeta_{i;(i)\beta_1,\cdots,\beta_{m-1};\gamma_1,\cdots,\gamma_{m-1}}$$

where $\sum_{(i)}$ denotes the sum is over all combinations $(\beta_1, \cdots, \beta_{m-1})$ of $m - 1$ integers chosen form $(1, \cdots, i-1, i+1, \cdots, n)$ and all combinations $(\gamma_1, \cdots, \gamma_{m-1})$ of $m - 1$ integers chosen from $(1, \cdots, i-1, i+1, \cdots, n)$. However, since $E|f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})|^{2+\delta} < B$, one obtains

$$E(Z_n) = \binom{n-m}{m-1}\binom{n-1}{m-1}^{-1} n^{-1} \sum_{i=1}^n \zeta_{1(i)} + o(1) \to \binom{n-m}{m-1}\binom{n-1}{m-1}^{-1}\zeta_{1,n} \to \zeta_1 .$$

From Theorem $B(i)$, page 275 of Loeve (1963), the result now follows. (Note that since we are dealing with a doubly indexed array, we can only obtain convergence in probability).

LEMMA 2. *Let* $X_1, \cdots, X_n$ *be independent random variables, and assume for some* $\delta > 0$ *and some* $0 < A < \infty$,

$$\begin{aligned} &0 < |\eta_{\alpha_1\cdots\alpha_m}| < A , \\ &0 < E|f^*(X_{\alpha_1}, \cdots, X_{\alpha_m})|^{2+\delta} < A \qquad \textit{for all } (\alpha_1, \cdots, \alpha_m) , \end{aligned} \tag{3.9}$$

$$E|h_{1(\nu)}(X_\nu)|^3 < \infty \quad \textit{for} \quad \nu = 1, \cdots, n , \qquad \textit{and} \tag{3.10}$$

$$\lim_{n\to\infty} \sum_{\nu=1}^n E\{|h_{1(\nu)}(X_\nu)|^3\}/[\sum_{\nu=1}^n E\{h_{1(\nu)}(X_\nu)\}^2]^{\frac{3}{2}} = 0 . \tag{3.11}$$

If $\zeta_{1,n} \to \zeta_1$ *as* $n \to \infty$, $0 < \zeta_1 < \infty$, *then* $(n-1) \sum_{i=1}^n (U_{n-1}^i - (\eta_{n-1}^i/\eta_n^0)U_n^0)^2 \to_P m^2\zeta_1$.

PROOF. Let

$$\begin{aligned} &T_n = (n-1) \sum_{i=1}^n (U_{n-1}^i - (\eta_{n-1}^i/\eta_n^0)U_n^0)^2 , \quad \text{and since} \\ &E(U_{n-1}^i) = \eta_{n-1}^i \mu , \quad \text{we may assume} \quad \mu = 0 . \end{aligned} \tag{3.12}$$

Let $\alpha_i = (n\eta_n^0 - (n-1)\eta_{n-1}^i)/\eta_n^0$, then

$$\begin{aligned} T_n &= (n-1) \sum_{i=1}^n (U_{n-1}^i - (n - a_i)U_n^0/(n-1))^2 \\ &= (n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0 + (a_i - 1)U_n^0/(n-1))^2 \\ &= (n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 + (U_n^0)^2 \sum_{i=1}^n (a_i - 1)^2/(n-1) \\ &\quad + X\text{-product term.} \end{aligned} \tag{3.13}$$

It will be shown that the first term of (3.13) converges in probability to $m^2\zeta_1$, the second term converges in probability to zero, and hence by the Cauchy-Schwarz inequality, the cross-product term converges to zero in probability.

Recall that if $A_n \to_P c$, and $E(A_n - B_n)^2 \to 0$ then $B_n \to_P c$. Hence if

$$V_n = n^{-1} \sum_{i=1}^n (mh_{1(i)}(X_i) - (n(n-1))^{\frac{1}{2}}(U_n^0 - U_{n-1}^i))^2 , \tag{3.14}$$

and $E(V_n) \to 0$, this fact the Lemma 1 suffice to show that the first term of (3.13) converges in probability to $m^2\zeta_1$.

Let

$$\begin{aligned} S_n &= (n-1) \sum_{i=1}^n (U_{n-1}^i - U_n^0)^2 \\ &= (n-1)\{\sum_{i=1}^n (U_{n-1}^i)^2 - n(U_n^0)^2\} \\ &= (n-1)n^{-1}\binom{n-1}{m}^{-2} \sum_{c=0}^m (cn - m^2) \\ &\qquad \times \sum_c f^*(X_{\alpha_1}, \cdots, X_{\alpha_m}) f^*(X_{\beta_1}, \cdots, X_{\beta_m}) \end{aligned} \tag{3.15}$$

as in (21) of Arvesen (1969) 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. But now, as in the expression immediately before (47) of Arvesen (1969),

$$E(S_n) = m^2\zeta_{1,n} + o(1) = m^2\zeta_{1,n} + o(1) . \tag{3.16}$$

We are done with the first term of (3.13) if it can be shown that

$$E[\sum_{i=1}^n h_{1(i)}(X_i)(U_n^0 - U_{n-1}^i)] = m\zeta_1 + o(1) . \tag{3.17}$$

To this end, note that

$$\begin{aligned} U_n^0 - U_{n-1}^i = \binom{n}{m}^{-1} &\sum_{D_{n-1}^i} f^*(X_i, X_{\alpha_1^i}, \cdots, X_{\alpha_{m-1}^i}) \\ &- [\binom{n}{m}^{-1} - \binom{n-1}{m}^{-1}] \sum_{C_{n-1}^i} f^*(X_{\beta_1^i}, \cdots, X_{\beta_m^i}) \end{aligned}$$

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 $(1, \cdots, i-1, i+1, \cdots, n)$. Hence

$$\begin{aligned} &E[h_{1(i)}(X_i)(U_n^0 - U_{n-1}^i)] \\ &\quad = \binom{n-1}{m-1}^{-1}\binom{n}{m}^{-1} E[(\sum_{\neq i} f^*_{1;\beta_1\cdots,\beta_{m-1}}(X_i)) \\ &\qquad \times (\sum_{D_{n-1}^i} f^*(X_i, X_{\alpha_1^i}, \cdots, X_{\alpha_{m-1}^i}))] \\ &\quad = \binom{n-1}{m-1}^{-1}\binom{n}{m}^{-1} \sum_{(i)} \zeta_{1;(i)\beta_1,\cdots,\beta_{m-1};\gamma_1,\cdots,\gamma_{m-1}} . \end{aligned} \tag{3.18}$$

Summing (3.18) over i, and using (3.9) one obtains (3.17), and hence $E(V_n) \to 0$.

It remains to show that the second term of (3.13) convergences to zero in probability. First note that by Hoeffding's (1948) U-statistic Central Limit Theorem for the nonidentically distributed case, one obtains

$$(U_n^0)^2 \to_P 0 . \tag{3.19}$$

Also note that since $|\eta_{\alpha_1\cdots\alpha_m}| < A$ for all $(\alpha_1, \cdots, \alpha_m)$

$$\begin{aligned} |\varepsilon_i - 1| &= (n-1)|\binom{n}{m}^{-1} \sum_{D_{n-1}^i} \eta_{i\alpha_1^i\cdots\alpha_{m-1}^i} + (\binom{n}{m}^{-1} - \binom{n-1}{m}^{-1}) \sum_{C_{n-1}^i} \eta_{\beta_1^i\cdots\beta_m^i}|/|\eta_n^0| \\ &< (n-1)[\binom{n}{m}^{-1}\binom{n-1}{m-1}A + (\binom{n-1}{m}^{-1} - \binom{n}{m}^{-1})\binom{n-1}{m}A]/|\eta_n^0| \\ &< 2Am/|\eta_n^0| . \end{aligned}$$

Since η_n^0 converges to nonzero $\bar{\eta}$,

$$\lim_{n\to\infty} (n-1)^{-1} \sum_{i=1}^n (a_1 - 1)^2 \leqq A^2(2m)^2/\bar{\eta}^2 . \tag{3.20}$$

Combining (3.9) and (3.20), the second term of (3.13) converges to zero in probability, and the lemma follows.

THEOREM 1. *Let* $X_1, \cdots, X_n$ *be independent random variables, and assume that* (3.9), (3.10), *and* (3.11) *hold. Let* g *be a function defined on the real line, which in a neighborhood of* μ *has a bounded second derivative. Let* $\hat{\theta}$, *the jackknife estimate of* $\theta = g(\mu)$ *be based on* $\hat{\theta}_n^0 = g(U_n^0/\eta_n^0)$ *as in* (3.8). *Define* $\hat{\theta}_{n-1}^i$ *as in* (3.8), *and follow the procedures of* (3.1)—(3.3). *Then if*

$$\zeta_{1,n} \to \zeta_1 \quad \text{as} \quad n \to \infty\,, \quad 0 < \zeta_1 < \infty\,, \tag{3.21}$$

the distribution of $(\hat{\theta} - \theta)/g'(\mu)(\operatorname{Var}(U_n^0/\eta_n^0))^{\frac{1}{2}})$ *is asymptotically normal with mean zero and variance one.*

PROOF. Without loss of generality, let $\mu = 0$, and also let $Y_i = \binom{n-1}{m-1}^{-1} \sum_{D_{n-1}^i} f^*(X_i, X_{\alpha_1 i}, \cdots, X_{\alpha_m i})$. Noting that $E(Y_i^2) = \operatorname{Var}(Y_i) \leqq \zeta_{1(i)} + 1$ for n sufficiently large, the proof follows from that of Theorem 10 of Arvesen (1969) until we expand terms in a power series to obtain

$$\begin{aligned}(\hat{\theta} - \theta) &= (ng(U_n^0/\eta_n^0) - (n-1)n^{-1} \textstyle\sum_{i=1}^n g(U_{n-1}^i/\eta_{n-1}^i) - g(0)) \\ &= [g(U_n^0/\eta_n^0) - g(0)] - (n-1)n^{-1} \\ &\quad \times [g'(U_n^0/\eta_n^0) \textstyle\sum_{i=1}^n (U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0) \\ &\quad + \textstyle\sum_{i=1}^n (U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0)^2 g''(\xi_i)/2]\end{aligned} \tag{3.22}$$

where ξ_i lies between U_{n-1}^i/η_{n-1}^i and U_n^0/η_n^0. First note that

$$\begin{aligned}(n-1)\{&\textstyle\sum_{i=1}^n (\eta_n^0)^2 (U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0)^2 \\ &- \textstyle\sum_{i=1}^n (U_{n-1}^i - (\eta_{n-1}^i/\eta_n^0)U_n^0)^2\} \to_P 0\end{aligned} \tag{3.23}$$

since $((\eta_n^0)^2 - (n_{n-1}^i)^2)/(\eta_{n-1}^i)^2 = (\eta_n^0 - \eta_{n-1}^i)(\eta_n^0 + \eta_{n-1}^i)/(n_{n-1}^i)^2 < 2mA(n-1)^{-1}(\eta_n^0 + \eta_{n-1}^i)/(\eta_{n-1}^i)^2$ by (3.20). Also (3.9) assures there is a $0 < M < \infty$ such that $|\eta_n^0 + \eta_{n-1}^i|/(\eta_{n-1}^i)^2 < M$, and using Lemma 2, (3.23) follows. Since $\lim_{n\to\infty} \eta_n^0 = \bar{\eta} \neq 0$, the second and third terms of (3.22) converge to zero using the proof of Arvesen (1969) (note that the Cauchy–Schwarz inequality handles the second term).

THEOREM 2. *Let* $X_1, \cdots, X_n$ *be independent random variables, and assume the hypotheses of Theorem* 1. *Then*

$$s_{\hat{\theta}}^2 \to_P m^2 \zeta_1 (g'(\mu))^2/(\bar{\eta})^2 \tag{3.24}$$

where $s_{\hat{\theta}}^2$ *is given by* (3.3).

PROOF. The proof follows Theorem 11 of Arvesen (1969). However note that

$$\begin{aligned}
s_{\hat\theta}^2 &= (n-1)^{-1} \textstyle\sum_{i=1}^n (\hat\theta_i - \hat\theta)^2 \\
&= (n-1) \textstyle\sum_{i=1}^n (g(U_{n-1}^i/\eta_{n-1}^i) - n^{-1} \sum_{j=1}^n g(U_{n-1}^j/\eta_{n-1}^j))^2 \\
&= (n-1) \textstyle\sum_{i=1}^n [(U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0) g'(\tau_i) \\
&\qquad - n^{-1} \textstyle\sum_{j=1}^n (U_{n-1}^j/\eta_{n-1}^j - U_n^0/\eta_n^0) g'(\tau_j)]^2 \\
&= (n-1) \textstyle\sum_{i=1}^n [(U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0) g'(0) \qquad (3.25) \\
&\qquad + (U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0) g'(\tau_i) - g'(0)) \\
&\qquad - n^{-1} \textstyle\sum_{j=1}^n (U_{n-1}^j/\eta_{n-1}^j - U_n^0/\eta_n^0) g'(\tau_j)]^2 \\
&= (n-1) \textstyle\sum_{i=1}^n (U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0)^2 (g'(0))^2 \\
&\qquad + (n-1)[\textstyle\sum_{i=1}^n (U_{n-1}^i/\eta_{n-1}^i - U_n^0/\eta_n^0)(g'(\tau_i) - g'(0)) \\
&\qquad - n^{-1} \textstyle\sum_{j=1}^n (U_{n-1}^j/\eta_{n-1}^j - U_n^0/\eta_n^0) g'(\tau_j)]^2 + X\text{-product term}
\end{aligned}$$

where τ_i lies between U_{n-1}^i/η_{n-1}^i and U_n^0/η_n^0. Now from (3.23), and Lemma 2, the first term of (3.25) converges to $m^2\zeta_1(g'(0))^2/(\bar\eta)^2$. The second term may also be readily shown to converge to zero in probability. Hence the result follows.

Combining Theorems 1 and 2, one obtains the result that $n^{\frac{1}{2}}(\hat\theta - \theta)/s_{\hat\theta}$ is asymptotically standard normal. In the original grouping $N = nk$, if n remains finite as $N \to \infty$, Theorem 7 of Arvesen (1969) can be readily extended to obtain convergence to a t distribution with $n-1$ degrees of freedom· Again, in what follows, we will assume $k = 1$. The generalization of Theorems 1 and 2 to functions of several U-statistics is straightforward, proceeding along the lines of Theorems 12 and 13 of Arvesen (1969).

4. An asymptotically robust test.

The results of the previous section will now be used to obtain an asymptotically robust test of (2.2). Consider the model specified in (2.1) without the normality assumption, but assuming moments of order at least six.

Temporarily let us assume that we are on the boundary of H_0 as given in (2.2), that is $\Delta_0 = \sigma_A^2/\sigma_e^2$. Let

$$X_i = \begin{pmatrix} \bar Y_{i\cdot} \\ \sum_{j=1}^{J_i} (Y_{ij} - \bar Y_{i\cdot})^2 \end{pmatrix}, \qquad w_i = J_i(\Delta_0 J_i + 1)^{-1}, \qquad i = 1, \cdots, n,$$

$$f^{*(1)}(X_{\alpha_1}, X_{\alpha_2}) = w_{\alpha_1} w_{\alpha_2} (\bar Y_{\alpha_1\cdot} - \bar Y_{\alpha_2\cdot})^2/2,$$

$$f^{*(2)}(X_{\alpha_1}) = \textstyle\sum_{j=1}^{J_{\alpha_1}} (Y_{\alpha_1 j} - \bar Y_{\alpha_1\cdot})^2 \qquad \text{and} \qquad W = \sum_{i=1}^n w_i .$$

If

$$U^{(1)} = \tbinom{n}{2}^{-1} \textstyle\sum_{\alpha_1<\alpha_2} f^{*(1)}(X_{\alpha_1}, X_{\alpha_2}) = \tbinom{n}{2}^{-1} W (\sum_{i=1}^n w_i (\bar Y_{i\cdot} - \bar Y_{\cdot\cdot})^2/2),$$

$$\bar Y_{\cdot\cdot} = W^{-1} \textstyle\sum_{i=1}^n w_i \bar Y_{i\cdot}, \quad \text{and}$$

$$U^{(2)} = \tbinom{n}{1}^{-1} \textstyle\sum_{\alpha_1} f^*(X_{\alpha_1}) = n^{-1} \sum_{i=1}^n \sum_{j=1}^{J_i} (Y_{ij} - \bar Y_{i\cdot})^2,$$

we find that

$$E(U^{(1)}) = W\sigma_e^2/n, \qquad \text{and} \qquad E(U^{(2)}) = (N^* - n)\sigma_e^2/n. \tag{4.1}$$

Also, note that $\mu^{(1)} = \mu^{(2)} = \sigma_e^2$, and

$$\eta^{(1)} = \binom{n}{2}^{-1} \sum_{\alpha_1<\alpha_2} w_{\alpha_1} w_{\alpha_2}(\Delta_0 + (J_{\alpha_1}^{-1} + J_{\alpha_2}^{-1})/2) = W/n\,, \qquad (4.2)$$
$$\eta^{(2)} = (N^* - n)/n\,.$$

Hence letting

$$g(U^{(1)}, U^{(2)}) = \frac{U^{(1)}/\eta^{(1)}}{U^{(2)}/\eta^{(2)}}\,, \qquad (4.3)$$

the hypotheses of the generalization of Theorems 1 and 2 to functions of several U-statistics are satisfied as long as σ_e^2 is nonzero and $\max(J_1, \cdots, J_n)$ remains bounded as $n \to \infty$. Note that the test statistic in (2.4) is identical to (4.3).

Finally, note that under the assumption of normality in (2.1) the numerator and denominator of (4.3) are independent random variables, each distributed as a constant times a chi-square random variable. This was shown by Spjøtvoll. Since the logarithm of a chi-square variable is approximately normally distributed, we propose testing H_0 by using the jackknife in conjunction with the log transformation of (4.3), that is by jackknifing

$$\hat{\theta}_n^{\,0} = \log\left\{\frac{U^{(1)}/\eta^{(1)}}{U^{(2)}/\eta^{(2)}}\right\}. \qquad (4.4)$$

Note that the variances of $U^{(1)}$, $U^{(2)}$ go to zero as $n \to \infty$ as shown by Tukey (1957). Moreover, for σ_A^2, σ_e^2 arbitrary, we find that

$$E[U^{(1)}] = W\sigma_e^2/n + (\Delta - \Delta_0)(W^2 - \textstyle\sum_{i=1}^n w_i^2)\sigma_e^2/(n(n-1))\,, \qquad (4.5)$$

and hence under H_A, $U^{(1)}$ converges in probability to a quantity greater than $W\sigma_e^2/n$, on the boundary of H_0 and H_A, $U^{(1)}$ converges in probability to $W\sigma_e^2/n$, and in the interior of H_0, $U^{(1)}$ converges in probability to a quantity less than $W\sigma_e^2/n$. Thus, from the generalization of Theorems 1 and 2, one obtains an asymptotically robust and unbiased test of (2.2) by rejecting H_0 at the α-level if

$$n^{\frac{1}{2}}\hat{\theta}/s_{\hat{\theta}} > Z_\alpha\,, \qquad (4.6)$$

where Z_α is the upper α point of a standard normal distribution.

Finally, in practice (and to be conservative), one might replace the cutoff point in (4.6) by the upper α point of a t distribution with $(n-1)$ degrees of freedom. This point will be discussed again in Section 6.

Unfortunately, it is not possible to readily use (4.6) to obtain a lower confidence bound for Δ since the boundary value Δ_0 of (2.2) appears in the statistic $\hat{\theta}$. However, note that in testing

$$H_0: \Delta \leqq \Delta^*$$
$$H_A: \Delta > \Delta^*$$

at the α-level using the proposed jackknife technique, if $\Delta_{Acc} = \{\Delta^*: H_0$ is accepted$\}$, and $\Delta_L = \inf \Delta_{Acc}$, then $\Delta_L < \Delta$ forms an asymptotic lower $(1-\alpha) \times 100\%$ confidence bound for Δ. But the actual computation of such a bound may be quite difficult. Spjøtvoll's proposed confidence interval also has this unpleasant property. Note that in theory this technique can be readily used to obtain a two-sided confidence interval, or an upper confidence bound for Δ.

5. Asymptotic efficiency results. Section 4 of this paper suggests that $H_0: \Delta \leqq \Delta_0$ be tested by applying the jackknife procedure to the logarithm of Spjøtvoll's F statistic ($F = (U^{(1)}/\eta^{(1)})/(U^{(2)}/\eta^{(2)})$ in the notation of Section 4). Section 3(c) of Arvesen (1969) proposed the use of the jackknife with the statistic $H = \text{MSA}/\text{MSE}$, where

$$\text{MSA} = (n-1)^{-1} \sum_{i=1}^{n} (Y_{i\cdot} - n^{-1} \sum_{i'=1}^{n} Y_{i'\cdot})^2,$$
$$\text{MSE} = n^{-1} \sum_{i=1}^{n} (J_i - 1)^{-1} \sum_{j=1}^{J_i} (Y_{ij} - Y_{i\cdot})^2,$$

and $Y_{i\cdot} = J_i^{-1} \sum_{j=1}^{J_i} Y_{ij}$. The fact that for normally distributed variables MSA and MSE are inefficient estimators, relative to the numerator and denominator of F, suggests that F may be a superior test statistic to H in terms of power. (This presumption is also raised by the manner of Spjøtvoll's derivation of F.) In this section we find that the Pitman ARE of F versus H, when all effects are assumed to be normally distributed, is greater than 1 when the J_i are not all equal. The log transformation and jackknifing do not affect the ARE, so the result holds for the comparison of tests based on the jackknifed versions of $\log F$ and $\log H$ respectively (these tests were used in Monte Carlo study discussed in the next section).

Recalling that F is the ratio of independent random variables, each distributed as a constant times a χ^2 variable (in particular, for $\Delta = \Delta_0$, $F \sim F_{n-1, N^*-n}$), and using (4.5), we readily find that the Pitman efficacy of F is

$$n \frac{\lim (n-1)^{-2}(W - \sum_{i=1}^{n} w_i^2/W)^2}{\lim 2\bar{J}/(\bar{J}-1)} \tag{5.1}$$

where $\bar{J} = n^{-1} \sum_{i=1}^{n} J_i$.
The Pitman efficacy of H can be obtained as

$$n/\lim 2[(\Delta_0 + \bar{J})^2(1 + n^{-1} \sum_{i=1}^{n} (J_i - 1)^{-1} + n^{-1} \sum_{i=1}^{n} (J_1^{-1} - \bar{J})^2)]. \tag{5.2}$$

If all $J_i = J$, both (5.1) and (5.2) reduce to $nJ(J-1)/2(\Delta_0 J + 1)^2$, and the Pitman ARE of F versus H is 1. If the J_i are not all equal (and $\lim n^{-1} \sum_{i=1}^{n} (J_i^{-1} - \bar{J})^2 > 0$), then the ARE is > 1, after applying Jensen's inequality. Suppose that the J_i have values 2, 3, and 4 in equal proportions, and that $\Delta_0 = 1$ (which is the case in the Monte Carlo simulation of Section 6). Then the ARE of F vs. H is 1.1.

6. Monte Carlo simulation. To obtain some information about the small sample behavior of the tests discussed in the preceding sections, a Monte Carlo simulation study was made. The program was run on the CDC6500 at Purdue University using procedures described by Rubin (1971). The model selected to test the jackknife procedure was:

$$Y_{ij} = a_i + e_{ij}, \quad i = 1, \cdots, 15, \quad j = 1, \cdots, J_i, \tag{6.1}$$
$$J_1 = \cdots = J_5 = 2, \quad J_6 = \cdots = J_{10} = 3, \quad J_{11} = \cdots = J_{15} = 4,$$

and the $\{a_i\}$, $\{e_{ij}\}$ are mutually independent random variables with mean zero, variance σ_A^2, σ_e^2 respectively. As in Arvesen and Schmitz (1970), three distributions were considered for the $\{a_i\}$, $\{e_{ij}\}$: both sets normal random variables, both sets double exponential random variables (kurtosis of 3), and both sets uniform random variables (kurtosis of -1.2).

The Monte Carlo study compares the empirical power functions of the Spjøtvoll F test to the jackknife procedure in testing

$$H_0 : \Delta = \sigma_A^2/\sigma_e^2 \leqq 1 \quad \text{vs.} \quad \Delta > 1 . \tag{6.2}$$

The jackknife was used with

$$\hat{\theta}_{15}^0(1) = \log\left(\frac{U^{(1)}/\eta^{(1)}}{U^{(2)}/\eta^{(2)}}\right) \tag{6.3}$$

where $U^{(i)}$, $\eta^{(i)}$, $i = 1, 2$ are given in Section 4, and with

$$\hat{\theta}_{15}^0(2) = \log(\text{MSA}/\text{MSE}) \quad \text{(see Section 5)} . \tag{6.4}$$

Of course $n = 15$, as stated above. In terms of the decomposition $N = nk$, $k = 1$ for this study.

There were 1000 sets of $\{a_i\}$, $\{e_{ij}\}$ generated according to the three distributions. They were first generated with $\Delta = 1$, and then scaled so that $\Delta = .5$, 1.5, 2.5, 4, 6, 9. Hence 180,000 pseudo-random numbers were generated in all. As mentioned in Section 4, there is some question as to whether the t_{n-1} distribution (t_{14} in this case) or the standard normal should be used in practice with moderate samples. The latter seems to be preferable for reasonable significance levels as the results in Table 1 demonstrate. Results are given separately for $\alpha = .10$, $\alpha = .05$, $\alpha = .01$. Finally, $J(\hat{\theta}_{15}^0(1))$ $J(\hat{\theta}_{15}^0(2))$ denotes the jackknife procedure using (6.3) and (6.4) respectively, and they are used either with the t_{14} distribution (w/t) or with the standard normal distribution (w/z) to obtain critical values. Note that the Monte Carlo accuracy is readily obtained from the binomial standard error. For example, the $\alpha = .10$ case yields a standard error of approximately $(.10 \times .90/1000)^{\frac{1}{2}} = .01$.

Examination of Table 1 produces several interesting results.

(i) The definite nonrobustness of the significance level of the Spjøtvoll F test is readily apparent. Comparison with an even more leptokurtic distribution than the double exponential would further emphasize this fact.

(ii) It was felt that the jackknife would work well at $\alpha = .10$, giving poorer results in the tails at $\alpha = .01$. Actually, at $\alpha = .10$, the jackknife using $J(\hat{\theta}_{15}^0(1))(w/z)$ is an excellent competitor to the Spjøtvoll F test even if the data are normal, and gives a more appropriate empirical significance level if the data are double exponential or uniform.

(iii) At $\alpha = .10$, $J(\hat{\theta}_{15}^0(1))$ appears to be slightly more powerful than $J(\hat{\theta}_{15}^0(2))$ using either t or z critical values. Of course, this is also essentially shown by Spjøtvoll and Section 5.

(iv) At $\alpha = .10$, the use of z critical values appear to be recommended as t

TABLE 1

Values of the Monte Carlo power function for testing (6.2)

$\Delta = \sigma_A{}^2/\sigma_e{}^2$	.5	1.0	1.5	2.5	4	6	9
			Normal distribution				
$\alpha = .10$							
Spjøtvoll F test	.010	.109	.305	.631	.889	.977	.995
$J(\hat{\theta}^0_{15}(1))(w/t)$	.009	.091	.276	.595	.852	.959	.990
$J(\hat{\theta}^0_{15}(2))(w/t)$	.012	.093	.280	.596	.833	.952	.985
$J(\hat{\theta}^0_{15}(1))(w/z)$	.010	.109	.300	.619	.869	.966	.993
$J(\hat{\theta}^0_{15}(2))(w/z)$	.013	.109	.299	.585	.884	.956	.986
$\alpha = .05$							
Spjøtvoll F test	.001	.053	.184	.516	.813	.948	.990
$J(\hat{\theta}^0_{15}(1))(w/t)$	.001	.039	.154	.449	.743	.913	.980
$J(\hat{\theta}^0_{15}(2))(w/t)$	.002	.043	.147	.433	.706	.885	.970
$J(\hat{\theta}^0_{15}(1))(w/z)$	.003	.056	.178	.488	.775	.927	.982
$J(\hat{\theta}^0_{15}(2))(w/z)$	.004	.052	.181	.471	.739	.906	.974
$\alpha = .01$							
Spjøtvoll F test	.000	.006	.051	.282	.600	.849	.965
$J(\hat{\theta}^0_{15}(1))(w/t)$	.000	.008	.041	.191	.457	.707	.888
$J(\hat{\theta}^0_{15}(2))(w/t)$	.000	.009	.038	.185	.435	.664	.837
$J(\hat{\theta}^0_{15}(1))(w/z)$	.000	.015	.066	.272	.550	.788	.927
$J(\hat{\theta}^0_{15}(2))(w/z)$	.000	.014	.064	.264	.532	.743	.906
			Double exponential distribution				
$\alpha = .10$							
Spjøtvoll F test	.023	.143	.300	.559	.799	.916	.977
$J(\hat{\theta}^0_{15}(1))(w/t)$	.012	.092	.199	.430	.681	.831	.934
$J(\hat{\theta}^0_{15}(2))(w/t)$	.018	.088	.210	.437	.668	.826	.928
$J(\hat{\theta}^0_{15}(1))(w/z)$	.016	.100	.220	.464	.695	.843	.946
$J(\hat{\theta}^0_{15}(2))(w/z)$	.025	.093	.229	.466	.683	.838	.939
$\alpha = .05$							
Spjøtvoll F test	.012	.084	.219	.461	.725	.861	.958
$J(\hat{\theta}^0_{15}(1))(w/t)$	.004	.045	.108	.289	.536	.718	.868
$J(\hat{\theta}^0_{15}(2))(w/t)$	.006	.046	.107	.297	.527	.713	.860
$J(\hat{\theta}^0_{15}(1))(w/z)$	.007	.055	.128	.330	.574	.749	.886
$J(\hat{\theta}^0_{15}(2))(w/z)$	.011	.057	.135	.331	.566	.748	.877
$\alpha = .01$							
Spjøtvoll F test	.002	.031	.094	.287	.522	.763	.895
$J(\hat{\theta}^0_{15}(1))(w/t)$	.001	.010	.038	.101	.248	.433	.617
$J(\hat{\theta}^0_{15}(2))(w/t)$	.000	.012	.032	.100	.257	.424	.608
$J(\hat{\theta}^0_{15}(1))(w/z)$	.000	.018	.055	.154	.320	.536	.712
$J(\hat{\theta}^0_{15}(2))(w/z)$	.001	.020	.052	.149	.335	.525	.703
			Uniform distribution				
$\alpha = .10$							
Spjøtvoll F test	.005	.060	.230	.672	.939	.993	1.000
$J(\hat{\theta}^0_{15}(1))(w/t)$	.005	.074	.277	.728	.952	.994	.998
$J(\hat{\theta}^0_{15}(2))(w/t)$	.007	.084	.279	.693	.930	.989	.998
$J(\hat{\theta}^0_{15}(1))(w/z)$	.006	.088	.299	.749	.956	.995	.999
$J(\hat{\theta}^0_{15}(2))(w/z)$	.008	.097	.299	.719	.934	.992	.999

TABLE 1 CONTINUED

$\Delta = \sigma_A^2/\sigma_e^2$	.5	1.0	1.5	2.5	4	6	9
$\alpha = .05$							
Spjøtvoll F test	.003	.023	.119	.519	.881	.982	.999
$J(\hat{\theta}_{15}^0(1))(w/t)$	.004	.026	.161	.572	.892	.983	.997
$J(\hat{\theta}_{15}^0(2))(w/t)$	.004	.031	.161	.546	.853	.976	.996
$J(\hat{\theta}_{15}^0(1))(w/z)$	.004	.040	.190	.612	.916	.986	.998
$J(\hat{\theta}_{15}^0(2))(w/z)$	.005	.044	.184	.584	.882	.980	.997
$\alpha = .01$							
Spjøtvoll F test	.001	.004	.021	.201	.655	.917	.987
$J(\hat{\theta}_{15}^0(1))(w/t)$	.000	.005	.036	.280	.676	.898	.977
$J(\hat{\theta}_{15}^0(2))(w/t)$	.000	.006	.042	.251	.620	.860	.976
$J(\hat{\theta}_{15}^0(1))(w/z)$	.002	.010	.066	.361	.760	.947	.990
$J(\hat{\theta}_{15}^0(2))(w/z)$	.001	.009	.067	.338	.722	.917	.984

critical values are too conservative. Note that at $\alpha = .01$, t values appear to be recommended, but then the power of the jackknife procedure is too low to recommend it as a competitor to Spjøtvoll's test. Of course, a larger sample size would correct this situation. Thus, there appears to be an interesting question as to the connection between sample size and the general asymptotic results.

In conclusion, it appears that if the jackknife works well, it should be used with $J(\hat{\theta}_n^0(1))(w/z)$. A researcher will have to be careful to see that n is large enough to use the normal approximation. In the Monte Carlo study given for $n = 15$, $\alpha = .10$ results are excellent, $\alpha = .05$ results are good, $\alpha = .01$ results are poor.

7. Acknowledgments. The first author would like to express appreciation to the Purdue Research Foundation for support. Both authors would like to acknowledge assistance from the referee and Professor Herbert Robbins.

REFERENCES

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

[2] ARVESEN, J., and SCHMITZ, T. (1970). Robust procedures for variance component problems using the jackknife. *Biometrics.* **26** 677–686.

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

[4] LOÈVE, M. (1963). *Probability Theory.* D. van Nostrand Company, Princeton.

[5] MILLER, R. (1964). A trustworthy jackknife. *Ann. Math. Statist.* **35** 1594–1605.

[6] MILLER, R. (1968). Jackknifing variances. *Ann. Math. Statist.* **39** 567–582.

[7] RUBIN, H. (1971). Some fast methods of generating random variables with preassigned distributions. Purdue University Technical Report.

[8] SCHEFFÉ, H. (1959). *Analysis of Variance.* Wiley, New York.

[9] SPJØTVOLL, E. (1967). Optimum invariant tests in unbalanced variance components models. *Ann. Math. Statist.* **38** 422–429.

[10] TUKEY, J. (1957). Variance components: II. The unbalanced single classification. *Ann. Math. Statist.* **28** 43-56.

PFIZER PHARMACEUTICALS, INC.
235 E. 42nd ST.
NEW YORK, NEW YORK 10017

MATHEMATICAL STATISTICS SECTION
BIOMETRY BRANCH
NATIONAL CANCER INSTITUTE
BETHESDA, MARYLAND 20014