MULTISAMPLE JACKNIFE STATISTICS

Ibrahim A. Ahmad, University of Petroleum and Minerals

1. INTRODUCTION. The practical aspect of this study grew out of the need to develop a method to reduce the bias in ratio estimation in stratified samples. Several auditing problems and a tax assessment ratio study have benefited from this research in the early seventies.

Bias reduction in parametric estimation was first presented by Quenouille (1949) and his method was called "Jacknife" by Tukey (1958). Miller (1974) gives an excellent review of the subject. Asymptotics of the Jacknife statistics have been recently studied in details by Thorburn (1976, 1977).

Let $X_1, \ldots, X_N$ be a random sample from a distribution function (df) $F(x, \theta)$, where $\theta = \theta(F)$ is an unknown parameter. Suppose that there is a good method to estimate θ, but it is biased, and it is desired to reduce this bias. Divide the N observations into n groups of k each, i.e., $N = nk$. Let $\theta_0 = \hat{\theta}(X_1,\ldots,X_N)$ be the estimate of θ based on all observations, and let $\hat{\theta}_i = \hat{\theta}(X_1,\ldots,X_{(i-1)k}, X_{ik+1}, X_{ik+2},\ldots,X_N)$ be the estimate of θ based on all observations after the deletion of the i-th group, $i = 1,\ldots,n$. Define the "pseudo-values" as follows:

$$\hat{\theta}^i = n\hat{\theta}_0 - (n-1)\hat{\theta}_i,\ i = 1,2,\ldots,n. \qquad (1.1)$$

The "Jacknife" estimate of θ is defined by:

$$\hat{\theta}_J = n^{-1} \Sigma_{i=1}^n \hat{\theta}^i. \qquad (1.2)$$

Simple calculations show that if $E\hat{\theta}_0 = \theta + [a/kn] + [b/(kn)^2] + 0(n^{-3})$, then $E\hat{\theta}_J = \theta - b/kn(n-1) + 0(n^{-3})$. Higher order Jacknife statistics are defined similarly, see Miller (1974).

The purpose of the present investigation is to propose a method to reduce the bias in the multisample case and demonstrate that this method does indeed reduce the bias and eliminates the cross product terms of the bias representation, see equation (2.6) to follow. Cox and Hinkley (1974), Miller (1974), and Jones (1974), independently, proposed a closely related method to reduce the bias of the first order in the multisample case. We shall formulate and discuss this method as we go along and discuss its analogy with our definition.

The paper contains five sections, in Section 2 a formal definition of the multisample Jacknife is presented and an estimate of the variance of the estimate θ_0 based on multisample pseudo-values is developed. In Section 3, the asymptotics of the multisample Jacknife statistics are presented, and in Section 4 the method is applied to ratio estimate from stratified samples. A discussion is presented in Section 5.

2. MULTISAMPLE JACKNIFE STATISTICS. To fix ideas we start by defining Jacknife statistics for the two-sample case. Let $X_{j1},\ldots,X_{jN_j}$, $j=1,2$ be two independent samples from df's F_j, $j=1,2$. Let $\theta = \theta(F_1, F_2)$ be an unknown parameter to be estimated by a reasonably good but biased method. Divide N_j into n_j groups of k_j each, i.e., $N_j = n_j k_j$, $j = 1, 2$. Let $\hat{\theta}_{(0,0)}$ denotes the estimate of θ based on all observations from the two samples, i.e.,

$$\hat{\theta}_{(0,0)} = \hat{\theta}(X_{11},\ldots,X_{1N_1},X_{21},\ldots,X_{2N_2}), \qquad (2.1)$$

let $\hat{\theta}_{(i_1,0)}$ denote the estimate of θ obtained from the two samples after the deletion of the i_1-st group from the first sample, i.e.

$$\hat{\theta}_{(i_1,0)} = (X_{11},\ldots,X_{1[(i_1-1)k_1]}, X_{1(i_1k_1+1)},\ldots,X_{1N_1},X_{21},\ldots,X_{2N_2}). \qquad (2.2)$$

Similarly we define $\hat{\theta}_{(0,i_2)}$ and $\hat{\theta}_{(i_1,i_2)}$ as follows:

$$\hat{\theta}_{(0,i_2)} = \hat{\theta}(X_{11},\ldots,X_{1N_1},X_{21},\ldots,X_{2[(i_2-1)k_2]}, X_{2(i_2k_2+1)},\ldots,X_{2N_2}). \qquad (2.3)$$

and

$$\hat{\theta}_{(i_1,i_2)} = \hat{\theta}(X_{11},\ldots,X_{1[(i_1-1)k_1]},X_{1(i_1k_1+1)},\ldots,X_{1N_1},X_{21},\ldots,X_{2[(i_2-1)k_2]}, x_{2(i_2k_2+1)},\ldots,X_{2N_2}). \qquad (2.4)$$

Define the two-sample "pseudo-values" as follows:

$$\hat{\theta}^{(i_1,i_2)} = n_1n_2\hat{\theta}_{(0,0)} - (n_1-1)n_2\hat{\theta}_{(i_1,0)} - n_1(n_2-1)\hat{\theta}_{(0,i_2)} + (n_1-1)(n_2-1)\hat{\theta}_{(i_1,i_2)}, \qquad (2.5)$$

for $i_j=1,\ldots,n_j$, $j=1,2$. The two-sample Jacknife statistic of θ is given by:

$$\hat{\theta}_J = (n_1n_2)^{-1}\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}\hat{\theta}(i_1,i_2). \qquad (2.6)$$

LEMMA 2.1. Let $n = \min(n_1,n_2)$ and assume that

$$E\hat{\theta}_{(0,0)} = \theta+[a_1/n_1k_1]+[a_2/n_2k_2]+[b_{12}/n_1n_2k_1k_2] +[c_1/(n_1k_1)^2]+[c_2/(n_2k_2)^2]+0(n^{-3}) \qquad (2.7)$$

Then

$$E\hat{\theta}_J = \theta-[c_1/n_1(n_1-1)k_1^2]-[c_2/n_2(n_2-1)k_2^2]+0(n^{-3}). \qquad (2.8)$$

PROOF. Using expression (2.7) for each of $E\hat{\theta}_{(0,0)}$, $E\hat{\theta}_{(i_1,0)}$, $E\hat{\theta}_{(0,i_2)}$, and $E\hat{\theta}_{(i_1,i_2)}$ we get, ignoring terms of order higher than the second, that

$$
\begin{aligned}
E\hat{\theta}_J = {} & n_1n_2\{\theta+[a_1/n_1k_1]+[a_2/n_2k_2]+[b_{12}/n_1n_2k_1k_2] \\
& +[c_1/(n_1k_1)^2]+[c_2/(n_2k_2)^2]\} \\
& -(n_1-1)n_2\{\theta+[a_1/(n_1-1)k_1]+[a_2/n_2k_2] \\
& +[b_{12}/(n_1-1)n_2k_1k_2]+[c_1/(n_1-1)k_1)^2] \\
& +[c_2/(n_2k_2)^2]\} \\
& -n_1(n_2-1)\{\theta+[a_1/n_1k_1]+[a_2/(n_2-1)k_2] \\
& +[b_{12}/n_1(n_2-1)k_1k_2]+[c_1/(n_1k_1)^2] \\
& +[c_2/((n_2-1)k_2)^2]\} \\
& +(n_1-1)(n_2-1)\{\theta+[a_1/(n_1-1)k_1]+[a_2/(n_2-1)k_2] \\
& +[b_{12}/(n_1-1)(n_2-1)k_1k_2]+[c_1/((n_1-1)k_1)^2] \\
& +[c_2/((n_2-1)k_2^2)^2]\}. \\
= {} & \theta-[c_1/n_1(n_1-1)k_1^2]-[c_2/n_2(n_2-1)k_2^2]. \qquad (2.9)
\end{aligned}
$$

The desired conclusion follows. QED.

REMARKS. (i) Arvesen (1969) proposed a two-sample jacknife estimate, which, in our notation may be written as:

$$
\begin{aligned}
\hat{\hat{\theta}}_J = {} & (n_1+n_2)^{-1}\{(n_1^2+n_2^2)\hat{\theta}_{(0,0)}-(n_1-1)\Sigma_{i_1=1}^{n_1}\hat{\theta}_{(i_1,0)} \\
& -(n_2-1)\Sigma_{i_2=1}^{n_2}\hat{\theta}_{(0,i_2)}\}. \qquad (2.10)
\end{aligned}
$$

Taking expectations, using representation (2.7) we get after simple algebra that

$$
\begin{aligned}
E\hat{\hat{\theta}}_J = {} & \theta+(n_1+n_2)^{-1}[(\frac{n_2}{n_1})(\frac{a_1}{k_1})+(\frac{n_1}{n_2})(\frac{a_2}{k_2})]+0(n^{-2}) \\
= {} & \theta+0(n^{-1}). \qquad (2.11)
\end{aligned}
$$

Thus $\hat{\hat{\theta}}_J$ does not reduce the first order bias.

(ii) If one is only interested in reducing the bias of the first order in representation (2.7), a simpler formulation of the Jacknife estimate would suffice, viz.,

$$
\begin{aligned}
\tilde{\theta}_J = {} & (n_1+n_2-1)\hat{\theta}_{(0,0)}-[(n_1-1)/n_1]\Sigma_{i_1=1}^{n_1}\hat{\theta}_{(i_1,0)} \\
& -[(n_2-1)/n_2\ \Sigma_{i_2=1}^{n_2}\hat{\theta}_{(0,i_2)}. \qquad (2.12)
\end{aligned}
$$

It is easy to see that the first order bias in (2.7) is eliminated. The above extension is due to Cox and Hinkley (1974), p. 264 and was investigated in the context of stratified sampling from finite populations by Jones (1974). It was also suggested by Miller (1974). Note also that Jones (1974) suggests a second order Jacknife in the two-sample case but he, unfortunately, overlooked the fact that the third term on the right-hand side of his (3.2) is also of second order whenever $a_g=0(a_h)$, $g\neq h$, which invalidates his conclusion that his $\tilde{X}^{(2)}$ is unbiased to the third order.

(iii) It is possible to use the pseudo-values $\hat{\theta}^{(i_1,i_2)}$ to provide a Jacknife-based estimate of the variance of $\hat{\theta}_{(0,0)}$ as follows:

$$
\hat{\sigma}_J^2=(n_1n_2)^{-1}\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}(\hat{\theta}^{(i_1,i_2)}-\hat{\theta}_J)^2. \qquad (2.13)
$$

Note that another estimate based on $\tilde{\theta}_J$ is also possible, let the pseudo-values of $\tilde{\theta}_J$ be given by

$$
\begin{aligned}
\tilde{\theta}^{(i_1,i_2)} = {} & (n_1+n_2-1)\hat{\theta}_{(0,0)}-(n_1-1)n_2\hat{\theta}_{(i_1,0)} \\
& -n_1(n_2-1)\hat{\theta}_{(0,i_2)}, \qquad (2.14)
\end{aligned}
$$

for $i_j=1,\dots,n_j$, $j=1,2$. Thus we can write

$$
\tilde{\theta}_J=(n_1n_2)^{-1}\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}\tilde{\theta}(i_1,i_2)
$$

and hence

$$
\begin{aligned}
\tilde{\sigma}_J^2 = {} & (n_1n_2)^{-1}\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_1}(\tilde{\theta}^{(i_1,i_2)}-\tilde{\theta}_J)^2 \\
= {} & n_1^{-1}\Sigma_{i_1=1}^{n_1}(\hat{\theta}_{(i_1,0)}-n_1^{-1}\Sigma_{i_1=1}^{n_1}\hat{\theta}_{(i_1,0)})^2 \\
& +n_2^{-1}\Sigma_{i_2=1}^{n_2}(\hat{\theta}_{(0,i_2)}-n_2^{-1}\Sigma_{i_2=1}^{n_2}\hat{\theta}_{(0,i_2)})^2. \qquad (2.15)
\end{aligned}
$$

Note that $\tilde{\sigma}_J^2$ was proposed by Ahmad (1973) and for stratified samples from finite populations by Jones (1974). It is easy to see that $\tilde{\sigma}_J^2 \le \tilde{\sigma}_J^2$.

The multisample Jacknife estimates may be expressed in an analogous way to that of the two-sample case. Let $\underline{0}'=(0,\dots,0)$, $(i_j,\underline{0})=(0,\dots,0,i_j,0,\dots,0)$ and, similarly define $(i_{j1},i_{j2},\underline{0}),\dots,$ until $\underline{i}'=(i_1,\dots,i_c)$. Let $\{X_{j1},\dots,X_{jN_j}\}$, $j=1,\dots,c$ be c independent samples from df's $F_1,\dots,F_c$, respectively. Denote by $\hat{\theta}_{\underline{0}}$ the estimate of a parameter $\theta=\theta(F_1,\dots,F_c)$ based on all observations from the c-samples, $\hat{\theta}_{(i_j,\underline{0})}$ the estimate of θ based on all observations from the c-samples after the deletion of the i-th group from the j-th samples, $j=1,\dots,c$, $\hat{\theta}_{(i_{j1},i_{j2},\underline{0})},\dots,$ $\hat{\theta}_{(i_1,\dots,i_c)}$ are defined similarly. The c-sample pseudo-values are defined by:

$$
\begin{aligned}
\hat{\theta}^{(i_1,\dots,i_c)} = {} & (\prod_{j=1}^{c}n_j)\hat{\theta}_{\underline{0}}-\Sigma_{j=1}^{c}(n_j-1)\prod_{\ell\ j}^{c}n_\ell\hat{\theta}_{(i_j,\underline{0})} \\
& +\Sigma_{j_1=1}^{c}\Sigma_{j_2=1}^{c}(n_{j1}-1)(n_{j2}-1)\prod_{\ell\neq j1,j2} \\
& n_\ell\hat{\theta}_{(i_{j1},i_{j2},\underline{0})}-\dots(-1)^c\hat{\theta}_{(i_1,\dots i_c)}. \qquad (2.16)
\end{aligned}
$$

The multisample Jacknife statistics is defined by:

$$
\hat{\theta}_J=(\prod_{j=1}^{c}n_j)^{-1}\Sigma_{i=1}^{n}\dots\Sigma_{i_c=1}^{n_c}\hat{\theta}^{(i_1,\dots,i_c)}. \qquad (2.17)
$$

A Jacknife-based estimate of the variance of $\hat{\theta}_{\underline{0}}$ is also given by:

$$
\hat{\sigma}_J^2=(\prod_{j=1}^{c}n_j)^{-1}\Sigma_{i_1=1}^{n_1}\dots\Sigma_{i_c=1}^{n_c}(\hat{\theta}^{(i_1,\dots,i_c)}-\hat{\theta}_J)^2. \qquad (2.18)
$$

Note that it is also possible to extend the Cox and Hinkley-Miller-Jones estimate $\tilde{\theta}_J$ to multisample cases and propose an estimate of the variance of $\hat{\theta}_{\underline{0}}$ as follows:

$$
\tilde{\theta}_J=(\prod_{j=1}^{c}n_j)^{-1}\Sigma_{i_1=1}^{n_1}\dots\Sigma_{i_c=1}^{n_c}\tilde{\theta}^{(i_1,\dots,i_c)} \qquad (2.19)
$$

and

$$\tilde{\sigma}_J^2=\Sigma_{j=1}^c n_j^{-1} n_j \Sigma_{i_j=1}^{n_j} (\tilde{\theta}(i_j,\underline{0})-n_j^{-1}\Sigma_{i_j=1}^{n_j}\tilde{\theta}(i_j,\underline{0}))^2, \tag{2.20}$$

where

$$\tilde{\theta}^{(i_1,\ldots,i_c)}=(\Sigma_{j=1}^c n_j+1)\hat{\theta}_{\underline{0}}-\Sigma_{j=1}^c (n_j-1)\Pi_{\ell\neq j} n_\ell \hat{\theta}(i_j,\underline{0}),$$

$$i_j=1,\ldots,n_j,\ j=1,\ldots,c. \tag{2.21}$$

3. ASYMPTOTIC PROPERTIES. In this section some asymptotic properties of $\hat{\theta}_J$ and $\hat{\sigma}_J^2$ are presented. In the spirit of the recent work of Thorburn (1976, 1977) we find sufficient conditions such that $\hat{\theta}_{\underline{0}}$ and $\hat{\theta}_J$ are asymptotically equivalent in distribution and that $\hat{\sigma}_J^2$ is consistent. Let $n = \min_{1\leq j\leq n} n_j$ and assume that $n_j = O(n)$, $j=1,\ldots,c$. The following conditions are needed to establish Theorems 3.1 and 3.2 below. Assume throughout this section that $k_j=1$, $j=1,\ldots,c$.

Condition (I). Assume that $\text{Var}\ \hat{\theta}_{\underline{0}}=\sigma^2/n+\varepsilon_n$, where (i) $\varepsilon_n=O(n^{-1})$, (ii) $\varepsilon_n-\varepsilon_{n-1}=o(n^{-2})$, and (iii) $\varepsilon_n-2\varepsilon_{n-1}+\varepsilon_{n-2}=o(n^{-3})$.

Condition (II). Let $\hat{\theta}_{\underline{0},(n_j+1)}$ denote the estimate $\hat{\theta}_{\underline{0}}$ with the j-th sample of size (n_j+1), $j=1,\ldots,c$. Assume that $E(\hat{\theta}_{\underline{0},(n_j+1)} | X_{11},\ldots,X_{1n_1},\ldots,X_{j1},\ldots,X_{jn_j},\ldots,X_{c1},\ldots,X_{cn_c}) = \frac{n_j}{n_j+1}\hat{\theta}_{\underline{0}}+r_{nj}$, where (iv) $\text{Var}(r_{nj})=o(n^{-3})$, (v) $\text{cov}(r_{nj},\hat{\theta}_{\underline{0}})=\delta_{nj}=o(n^{-2})$, and $\delta_{nj}-\delta_{n(n-1)} = o(n^{-3})$, $j=1,\ldots,c$.

Note that the above conditions are multisample versions of those used by Thorburn (1976, 1977) where he discusses cases where the conditions are satisfied and also presents the interrelations between them (see p. 30 of Thorburn (1976)).

THEOREM 3.1. *Suppose that conditions (I) and (II) are satisfied. If the pseudo-values* $\hat{\theta}^{(i_1,\ldots,i_c)}$ *converge in the second mean to random variables* $\Theta^{(i_1,\ldots,i_2)}$ *for every* $i_1,\ldots,i_c$, *as* $n\to\infty$, *then* $n^{1/2}(\hat{\theta}_J-\theta)$ *converges in distribution to a normal random variable with 0 mean and variance* σ^2, *as* $n\to\infty$.

PROOF. Since this resembles very closely that of Theorem 3.2 of Thorburn, it shall only be sketched. Again for notational convenience let c=2. Thus

$$\text{Var}\ (\hat{\theta}^{(i_1,i_2)}$$
$$=(n_1n_2)^2\text{Var}\hat{\theta}_{(0,0)}+(n_1-1)^2n_2^2\text{Var}\hat{\theta}_{(i_1,0)})$$
$$+n_2^2(n_2-1)^2\text{Var}\hat{\theta}_{(0,i_2)}$$
$$\pm 2 \text{ all the covariances.} \tag{3.1}$$

Using Conditions (I) and (II) and after some algebra and cancellation it can be shown that $\text{Var}(\hat{\theta}^{(i_1,i_2)})=\sigma^2+o(1)$, $i_j=1,\ldots,n_j$, $j=1,2$. In similar fashion we may show that

$$\text{cov}(\hat{\theta}^{(i_1,i_2)},\hat{\theta}^{(i_1,i_2^*)})=\text{cov}(\hat{\theta}^{(i_1,i_2)},\hat{\theta}^{(i_1^*,i_2)})$$
$$=\text{cov}(\hat{\theta}^{(i_1,i_2)},\hat{\theta}^{(i_1^*,i_2^*)})=O(n^{-1}), \tag{3.2}$$

for all $i_1\neq i_1^*$ and $i_2\neq i_2^*$, $1\leq i_j,\ i_j^*\leq n_j$, $j=1,2$. Hence $\text{Var}(\hat{\theta}_J-\Theta_{\underline{0}})=o(n^{-1})$, and thus it follows from Theorem 2.2 of Thorburn (1977) that $\Theta^{(i_1,i_2)}$ are independent random variables. Assume without loss of generality that $E\hat{\theta}^{(i_1,i_2)} = E\Theta^{(i_1,i_2)}=0$ for all $1\leq i_j\leq n_j$, $j=1,2$. Let

$$\Theta_J = (n_1n_2)^{-1}\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}\Theta^{(i_1,i_2)}. \tag{3.3}$$

It is possible, using methods similar to those of Thorburn, Theorem 3.2, to show that $\text{Var}(\hat{\theta}_J-\Theta_J) =o(n^{-1})$, but Θ_J is equivalent to a U-statistic (Hoeffding (1948)) and is easily seen from the results of Hoeffding (1948) to be asymptotically normal with mean 0 and variance σ^2. The result follows. QED

THEOREM 3.2. *Under the conditions of Theorem 3.1,* $\hat{\sigma}_J^2$ *is a consistent estimate of* σ^2.

PROOF. Again we let c=2. Note that it is not difficult to see that

$$\sigma_n^2=(n_1n_2)^{-1}\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}(\Theta^{(i_1,i_2)}-\Theta_J)^2 \to \sigma^2 \text{ in probability as } n\to\infty. \tag{3.4}$$

But also we can establish, using arguments outlined in Theorem 3.1 above except obviously more tedious calculations are needed, that $\hat{\sigma}_J^2-\sigma_n^2\to 0$ in probability as $n\to\infty$. QED

REMARKS. (i) It follows from Theorems 3.1 and 3.2 that $n^{1/2}(\hat{\theta}_J-\theta)/\hat{\sigma}_J$ is asymptotically standard normal. Thus approximate Jacknife confidence interval for θ may be given by $\hat{\theta}_J\pm z_{\alpha/2}\hat{\sigma}_J$ where $z_{\alpha/2}$ is such that $P[|Z|\leq z_{\alpha/2}] = 1-\alpha$ where Z denotes the standard normal variate.

(ii) If $\theta=f(\theta_1,\ldots,\theta_c)$ where θ_j admits an estimate $\hat{\theta}_j$, $j=1,\ldots,c$. Thus $\hat{\theta}_{\underline{0}}=f(\hat{\theta}_1,\ldots,\hat{\theta}_c)$ and $\hat{\theta}_J$ is the multisample Jacknife estimate of θ. If f admits bounded partial derivatives up to the fourth order and if the conditions of Theorem 3.1, above are satisfied then $n^{1/2}(\hat{\theta}_J-\theta)$ is asymptotically normal with mean 0 and variance $\sigma^2=\Sigma_{j=1}^c\sigma_j^2 f_j(\theta_1,\ldots,\theta_c)$, where $f_j(\theta_1,\ldots,\theta_c) = \partial f(\theta_1,\ldots,\theta_c)/\partial\theta_j$, $j=1,\ldots,c$. An application of this result is the case of means $\mu_j=EX_{j1}=\theta_j$, $j=1,\ldots,c$ in which case all conditions of Theorem 3.1 are satisfied if f has bounded fourth derivative and we obtain an extension to multisample of a result of Miller (1964). A direct proof of this special case based on Miller's argument appears in Ahmad (1973).

4. AN APPLICATION. In this section we show that the multisample Jacknife estimates defined in Section 2 not only do they reduce the bias but also they reduce the variance when estimating the ratio of means from stratified samples taken from normal populations. Let $\alpha \in [0, 1]$ be any real number, and let F_1 and F_2 be two independent d.f.'s. Further, assume that $F=\alpha F_1+(1-\alpha)F_2$, and that F is two dimensional such that $(X,Y) \sim F$. We shall be concerned with estimating $R=EY/EX$. If $(X_j,Y_j) \sim F_j$, $j=1,2$, then we can write

$$R=[\alpha EY_1+(1-\alpha)EY_2]/[\alpha EX_1+(1-\alpha)EX_2]. \tag{4.1}$$

Let $(X_{j1},Y_{j1}),\ldots,(X_{jn_j},Y_{jn_j})$ be a random sample from $F_j(x,y)$, $j=1,2$, then the usual estimate of R is given by

$$R_0=[\alpha\bar{Y}_1+(1-\alpha)\bar{Y}_2]/[\alpha\bar{X}_1+(1-\alpha)\bar{X}_2]. \tag{4.2}$$

Now assume that $\bar{X}_j \sim N(1,h_j)$, where $h_j=0(n_j^{-1})$, $j=1,2$ and that $Y_j=\gamma_j+\beta X_j+e_j$, where $Ee_j=0$ and $\mathrm{cov}(X_j,e_j)=0$, $j=1,2$, in which case we have

$$R=\beta+[\alpha\gamma_1+(1-\alpha)\gamma_2]/[\alpha EX_1+(1-\alpha)EX_2], \tag{4.3}$$

and hence

$$\hat{R}_0=\beta+\{[\alpha\gamma_1+(1-\alpha)\gamma_2 + \alpha\bar{e}_1+(1-\alpha)\bar{e}_2]\}/[\alpha\bar{X}_1+(1-\alpha)\bar{X}_2]. \tag{4.4}$$

Thus

$$E\hat{R}_0=\beta+[\alpha\gamma_1+(1-\alpha)\gamma_2]E[\alpha\bar{X}_1+(1-\alpha)\bar{X}_2]^{-1}. \tag{4.5}$$

To obtain the bias in $E\hat{R}_0$ we proceed as follows: let $\xi_1=1-X_1$, and $\xi_2=1-X_2$, thus $\alpha\bar{X}_1+(1-\alpha)\bar{X}_2=1-(\alpha\bar{\xi}_1+(1-\alpha)\bar{\xi}_2)$, where $\bar{\xi}_j=n_j^{-1}\Sigma_{i=1}^{n_j}\xi_{ji}$, $j=1,2$, and hence

$$E(\alpha\bar{X}_1+(1-\alpha)\bar{X}_2)^{-1}=E[1-(\alpha\bar{\xi}_1+(1-\alpha)\bar{\xi}_2)]^{-1}$$
$$= 1+E(\alpha\bar{\xi}_1+(1-\alpha)\ \bar{\xi}_2)^2+E(\alpha\bar{\xi}_1+(1-\alpha)\bar{\xi}_2)^4$$
$$+ E(\alpha\bar{\xi}_1+(1-\alpha)\ \bar{\xi}_2)^6+0(n^{-4}),$$

where $n=\min(n_1,n_2)$. Obtaining the above expected values and simplifying we get

$$\mathrm{Bias}(\hat{R}_0)=E(\hat{R}_0-R)$$
$$=[\alpha\gamma_1+(1-\alpha)\gamma_2][\alpha^2h_1+(1-\alpha)^2h_2+3\alpha^4h_1^2$$
$$+3(1-\alpha)^4h_2^2+6\alpha^2(1-\alpha)^2h_1h_2+15\alpha^6h_1^3$$
$$+15(1-\alpha)^6h_2^3+36\alpha^2(1-\alpha)^4h_1h_2$$
$$+36\alpha^4(1-\alpha)^2h_1^2h_2+0(n^{-4})]. \tag{4.7}$$

As for the variance of $\hat{R}_0$ we obtain after some calculations, that

$$\mathrm{Var}\ \hat{R}_0=[\alpha\gamma_1+(1-\alpha)\gamma_2]^2[\alpha^2h_1+(1-\alpha)^2h_2+8\alpha^4h_1^2$$
$$+8(1-\alpha)^4h_2^2+16\alpha^2(1-\alpha)^2h_1h_2$$
$$+69\alpha^6h_1^3+69(1-\alpha)^6h_2^3+168\alpha^2(1-\alpha)^4h_1h_2^2$$
$$+168\alpha^4(1-\alpha)^2h_1^2h_2+0(n^{-4})]$$
$$+[\alpha^2\delta_1+(1-\alpha)^2\delta_2][3\alpha^2h_1+3(1-\alpha)^2h_2+15\alpha^6h_1^2$$
$$+15(1-\alpha)^4h_2^2+30\alpha^2(1-\alpha)^2h_1h_2$$
$$+105\alpha^6h_1^3+105(1-\alpha)^6h_2^3+252\alpha^2(1-\alpha)^4h_1h_2^2$$
$$+252\alpha^4(1-\alpha)^2h_1^2h_2+0(n^{-4})], \tag{4.8}$$

where $\delta_j=E\bar{e}_j^2$, $i=1,2$. Now, to obtain the multisample Jacknife estimate of R we use definition (2.6) of Section 2 and get

$$\hat{R}_J=\beta+[\alpha\gamma_1+(1-\alpha)\gamma_2]\{[n_1n_2/(\alpha\bar{X}_1+(1-\alpha)\bar{X}_2)]$$
$$-[(n_1-1)n_2/n_1]\Sigma_{i_1=1}^{n_1}(\alpha\bar{X}_1^{i_1}+(1-\alpha)\bar{X}_2)^{-1}$$
$$-[n_1(n_2-1)/n_2]\Sigma_{i_2=1}^{n_2}(\alpha\bar{X}_1+(1-\alpha)\bar{X}_2^{i_2})^{-1}$$
$$+[n_1-1)(n_2-1)/n_1n_2]\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}(\alpha\bar{X}_1^{i_1}$$
$$+(1-\alpha)\bar{X}_2^{i_2})^{-1}\}$$
$$+\{[n_1n_2(\alpha\bar{e}_1+(1-\alpha)\bar{e}_2)/(\alpha\bar{X}_1+(1-\alpha)\bar{X}_2)]$$
$$-[(n_1-1)n_2/n_1]\Sigma_{i_1=1}^{n_1}[(\alpha\bar{e}_1^{i_1}+(1-\alpha)\bar{e}_2)/(\alpha\bar{X}_1^{i_1}$$
$$+(1-\alpha)\bar{X}_2)]$$
$$-[n_1(n_2-1)/n_2]\Sigma_{i_2=1}^{n_2}[(\alpha\bar{e}_1+(1-\alpha)\bar{e}_2^{i_2})/(\alpha\bar{X}_1$$
$$+(1-\alpha)\bar{X}_2^{i_2})]+[(n_1-1)(n_2-1)/n_1n_2]\Sigma_{i_1=1}^{n_1}\Sigma_{i_2=1}^{n_2}$$
$$[(\alpha\bar{e}_1^{i_1}+(1-\alpha)\bar{e}_2^{i_2})/(\alpha\bar{X}_1^{i_1}+(1-\alpha)\bar{X}_2^{i_2})]\}. \tag{4.9}$$

Straightforward but lengthy algebra leads to the following expression of the bias and the variance of $\hat{R}_J$:

$$\mathrm{Bias}(\hat{R}_J)=E(\hat{R}_J-R)$$
$$=[\alpha\gamma_1+(1-\alpha)\gamma_2]\{-3\alpha^4h_1^2[n_1/(n_1-1)]$$
$$-3(1-\alpha)^4h_2^2[n_2/(n_2-1)]$$
$$-15\alpha^6h_1^3[n_1(2n_1-1)/(n_1-1)^2]$$
$$-15(1-\alpha)^6h_2^3[n_2(2n_2-1)/(n_2-1)^2]+0(n^{-4})\}.$$

It is evident from (4.7) and (4.10) that the bias of $\hat{R}_J$ is smaller than the bias of $\hat{R}_0$, and the bias of $\hat{R}_J$ attains its minimum at $n_j=N_j$ (i.e., $k_j=1$), $j=1,2$. Next, we obtain the variance of $\hat{R}_J$ as follows:

$$\mathrm{Var}\hat{R}_J=[\alpha\gamma_1+(1-\alpha)\gamma_2]^2\{\alpha^2h_1+(1-\alpha)^2h_1+$$
$$+[2n_1/(n_1-1)]\alpha^4h_1^2+[2n_2/(n_2-1)](1-\alpha)^4h_2^2$$
$$-[6n_1(4n_1^2-14n_1+11)/(n_1-1)^3]\alpha^6h_1^3$$
$$-[6n_2(4n_2^2-14n_2+11)/(n_2-1)^3](1-\alpha)^6h_2^3\}$$

$$+[\alpha^2\delta_1+(1-\alpha)^2\delta_2]\{1+[n_1/(n_1-1)]\alpha^2h_1$$
$$+[n_2/(n_2-1)](1-\alpha)^2h_2-[2n_1(2n_1^2-9n_1+8)/$$
$$(n_1-1)^3]\alpha^4h_1^2$$
$$-[2n_2(2n_2^2-9n_2+8)/(n_2-1)^3](1-\alpha)^4h_1^2$$
$$-[3n_1(28n_1^4-149n_1^3+260n_1^2-173n_1+32)/$$
$$(n_1-1)^5\alpha^6h_1^3-[3n_2(28n_2^4-149n_2^3+260n_2^2$$
$$-173n_2+32)/(n_2-1)^5](1-\alpha)^6h_2^3+0(n^{-4})\}. \quad (4.11)$$

Note that when $\alpha=0$ or $\alpha=1$ expressions (4.10) and (4.11) reduce to the one sample case of Rao (1965). It is clear from (4.8) and (4.11) that the multisample Jacknife estimate in this case reduces the variance.

To illustrate the difference between $\hat{\theta}_J$ and $\tilde{\theta}_J$ defined, respectively, in (2.6) and (2.12) we calculated each one of them and their corresponding variance estimates $\hat{\sigma}_J^2$ and $\tilde{\sigma}_J^2$ as given in (2.13) and (2.15), respectively, using the data given in the illustration of Jones (1974). The results are as follows: $\hat{\theta}_J=2.834$, $\tilde{\theta}_J=2.911$, $\hat{\sigma}_J^2=0.5788$, and $\tilde{\sigma}_J^2=0.7630$. Since the usual estimate of R tends to be positively biased then the bias reduction of $\hat{\theta}_J$ is better due to the elimination of the cross product term. Note also that $\hat{\sigma}_J^2$ is smaller than $\tilde{\sigma}_J^2$.

5. DISCUSSION. From previous sections we have seen that when extending Quenoulli's method of bias reduction to multisample case there are more than one way to do so. Two methods have been discussed in the present investigation. Since the bias representation would include cross-product terms, the method we proposed here eliminates the first order bias as well as the cross product term, while in the definition of Cox and Hinkley (1974), and Jones (1974) only the bias of the first order is eliminated. The variance estimate based on our definition is smaller than that based on the definition of the above mentioned authors.

We did not attempt to answer all the questions pertaining to all aspects of theory and applications of the proposed multisample Jacknife and evidently not a few open questions are left which will, hopefully, stimulate some more interest in the subject. A few open questions that are obvious include: (i) How can higher order multisample Jacknife estimates be obtained? (ii) How can one obtain multisample generalized Jacknife statistics (for generalized Jacknife see Section 2.2 of Miller (1974))? Can the multisample Jacknife be applied to the maximum likelihood estimation in a rigorous way? Computational aspects of the subject need also be investigated. This writer is now working on a general solution to calculate the multisample Jacknife with arbitrary group sizes.

In Remark (ii) following Theorems 3.1 and 3.2 we pointed out that an application of Theorem 3.1 would be to estimate functions of populations means by sample means. As done in Arvesen (1969), it is possible to extend this remark to cover functions of parameters that are estimable by U-statistics. The interested reader may find it useful exercise to put down the details.

In closing, the review of Miller (1974) carries very many areas of potential applications and further research for multivariate Jacknife statistics.

REFERENCES

AHMAD, I.A. (1973). Multisample Jacknife statistics. Unpublished manuscript.

ARVESEN, J.N. (1969). Jacknifing U-statistics. Ann. Math. Statist. 40, 2076-90.

COX, D.R. and HINKLEY, D.V. (1974). Theoretical Statistics. Chapman and Hall, London.

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, 293-25.

JONES, H.L. (1974). Jacknife estimation of functions of stratum means. Biometrika. 61, 243-48.

MILLER, R.J. (1974). The Jacknife - a review. Biometrika 61, 1-15.

MILLER, R.J. (1964). A trustworthy Jacknife. Ann. Math. Statist. 39, 567-82.

QUENOUILLE, M.H. (1949). Approximate tests of correlation in time series. J.R. Statist. Soc. B 11, 68-84.

RAO, J.N.K. (1965). A note on bias reductions in estimation. Biometrika 43, 353-60.

THORBURN, D. (1977). On the asymptotic normality of the Jacknife. Scand. J. Statist. 4, 113-18.

THORBURN, D. (1976). Some asymptotic properties of Jacknife statistics. Biometrika 63, 305-13.

TUKEY, J. (1958). Bias and confidence in not quite large samples (abstract). Ann. Math. Statist. 29, 614.