Approximate Tests of Independence in Contingency Tables from Complex Stratified Cluster Samples

Several approximate tests based on half-sample estimates are proposed for testing hypotheses in contingency tables from complex stratified cluster samples. Monte Carlo methods are used to evaluate the power and expected significance level of each of these tests.

DHEW Publication No. (HSM) 73-1327

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service

Health Services and Mental Health Administration
National Center for Health Statistics
Rockville, Md. March 1973

Vital and Health Statistics-Series 2-No. 53

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402
Price 40 cents domestic postpaid or 30 cents GPO Bookstore

NATIONAL CENTER FOR HEALTH STATISTICS

OFFICE OF STATISTICAL METHODS

Vital and Health Statistics-Series 2-No. 53

DHEW Publication No. (HSM) 73-1327
Library of Congress Catalog Card Number 72-600132

FOREWORD

The analysis of data collected by the National Center for Health Statistics presents difficult problems because the classical tests of statistical hypotheses are often based on assumptions that are not satisfied when applied to data based on complex sample surveys. Consequently the Center sponsored a contract with the Statistics Department, Hebrew University in Israel, to develop tests of hypotheses suitable for the analysis of data collected in the Center's sample surveys. The contract produced several interesting and useful reports, including this one by Dr. Gad Nathan of the Hebrew University and the Israel Central Bureau of Statistics. The report was completed by Dr. Nathan while on leave of absence at the Department of Biostatistics, University of North Carolina at Chapel Hill.

Dr. Bernard Greenberg, Dean, School of Public Health, University of North Carolina, served as project officer, and Dr. Reuben Gabriel, Chairman, Department of Statistics, Hebrew University, was the project director for this contract. Dr. Gary Koch, Department of Biostatistics, University of North Carolina, and Dr. Paul Levy, Office of Statistical Methods, National Center for Health Statistics, reviewed drafts of Dr. Nathan's manuscript and made helpful suggestions. Dr. Levy also assumed responsibility for working with the editorial staff in preparing this report for publication.

MONROE G. SIRKEN

CONTENTS

APPROXIMATE TESTS OF INDEPENDENCE IN CONTINGENCY TABLES FROM COMPLEX STRATIFIED CLUSTER SAMPLES

Gad Nathan, *Hebrew University and Central Bureau of Statistics, Jerusalem*

1. Introduction and Summary

For simple random sampling within strata, approximate methods for testing overall independence in a contingency table have been proposed by Bhapkar[1] and by Garza-Hernandez.[2] In this case, the maximum likelihood ratio can be approximated as closely as required, as shown by Nathan.[3] This is not so when the contingency table is obtained from a complex stratified cluster sample. Chapman[4] and McCarthy[5] have proposed using statistics based on the replicated balanced half-sample method of McCarthy,[6] published in *Vital and Health Statistics*, Series 2, No. 14. These statistics are the differences between cell estimates obtained from one half sample and the product of the relevant marginal estimates obtained from the complementary half sample. Chapman's test procedure, based on the signs of these statistics, relies on assumptions of (1) zero expectations of the statistics under the null hypothesis, (2) independence between statistics from different sets of half samples, and (3) fixed covariances between statistics from the same pair of half samples. As will be shown, an exact evaluation of the relevant expectations and covariances of Chapman's statistics indicates that these assumptions are not always tenable.

Instead we propose here to use half-sample estimates to obtain some modified statistics, which have exactly expectation zero under the null hypothesis and for which the relevant covariances can be evaluated approximately. Test procedures—some based on the large sample statistics and others on Hotelling's T^2—are then obtained on the basis of sample estimates of the covariance approximations.

For a numerical example it is shown that the effect of the various assumptions and approximations made on the values of the statistics is very small.

2. The Model and Notation

In each of L strata two primary sampling units (PSU's) are selected with equal probabilities and without replacement. Second-stage sampling (within PSU's) can be by any method which ensures the following two conditions:

(a) If P_{ijh} is the probability of being classified in cell (i, j) of the contingency table $(i=1, \ldots, r;\ j=1, \ldots, c)$, conditional on being in stratum h $(h=1, \ldots, L)$, then an unbiased estimate, $\hat{P}_{ijha}$, of P_{ijh} is available from each of the selected PSU's $(a=1, 2)$.

(b) $\hat{P}_{ijh1}$ and $\hat{P}_{ijh2}$ are independent within stratum h.

Weights W_h $(h=1,\ldots, L)$—the probability of inclusion in stratum h—are assumed as known; and it follows that $P_{ij}=\sum_{h=1}^{L} W_h P_{ijh}$ is the overall unconditional probability of being in cell (i, j).

Let $P_{i\cdot}=\sum_{j=1}^{c} P_{ij}$ and $P_{\cdot j}=\sum_{i=1}^{r} P_{ij}$ be the marginal unconditional probabilities. Then the null hypothesis to be tested is that of overall independence, i.e.,

$$H_0\colon P_{ij}=P_{i\cdot}P_{\cdot j}\ (i=1, \ldots, r;\ j=1, \ldots, c). \tag{2.1}$$

In order to obtain statistics for which variances and covariances can be estimated to test this hypothesis, a set of K-balanced half samples is defined by McCarthy's technique.[6] Each of these half samples consists of a selection of one of the PSU's originally selected in each stratum. Therefore each half sample and its complement are simple stratified samples with one PSU per stratum. In addition, the estimates based on any half sample and its complement are independent. The half-sample selection defines indicator functions as follows:

$$\alpha_h^{(k)} = \begin{cases} 1 \text{ if PSU 1 is selected in the } h\text{th stratum for the } k\text{th half sample} \\ 0 \text{ otherwise} \end{cases}$$

$$(h=1, \ldots, L;\ k=1, \ldots, K). \tag{2.2}$$

The two unbiased estimates of the probability P_{ij} based on the kth half sample and its complement are then defined, respectively, by

$$\hat{P}_{ij}^{(k)} = \sum_{h=1}^{L} W_h[\alpha_h^{(k)}\hat{P}_{ijh1} + (1-\alpha_h^{(k)})\hat{P}_{ijh2}]$$

and

$$\tilde{P}_{ij}^{(k)} = \sum_{h=1}^{L} W_h[(1-\alpha_h^{(k)})\hat{P}_{ijh1} + \alpha_h^{(k)}\hat{P}_{ijh2}]$$

$$(i=1, \ldots, r;\ j=1, \ldots, c;\ k=1, \ldots, L). \tag{2.3}$$

Let $\hat{P}_{i\cdot}^{(k)}$, $\hat{P}_{\cdot j}^{(k)}$ and $\tilde{P}_{i\cdot}^{(k)}$, $\tilde{P}_{\cdot j}^{(k)}$ be the corresponding unbiased estimates of the marginal probabilities from the kth half sample and its complement, respectively. Then $\hat{P}_{ij}^{(k)}$, $\hat{P}_{i\cdot}^{(k)}$, and $\hat{P}_{\cdot j}^{(k)}$ are independent of $\tilde{P}_{ij}^{(k)}$, $\tilde{P}_{i\cdot}^{(k)}$, and $\tilde{P}_{\cdot j}^{(k)}$.

Thus the random variables

$$X_{ij}^{(k)} = \hat{P}_{ij}^{(k)} + \tilde{P}_{ij}^{(k)} - \hat{P}_{i\cdot}^{(k)}\tilde{P}_{\cdot j}^{(k)} - \tilde{P}_{i\cdot}^{(k)}\hat{P}_{\cdot j}^{(k)}$$

$$(i=1, \ldots, r-1;\ j=1, \ldots, c-1;\ k=1, \ldots, K) \tag{2.4}$$

have expectation zero under the null hypothesis (2.1).

Alternative statistics based on differences between cross products (rather than on differences between cell probability estimates and products of marginal probabilities) can be used, as proposed, e.g., by Bhapkar and Koch.[7] If we set

$$U_{ij}^{(k)} = \hat{P}_{ij}^{(k)}\tilde{P}_{rc}^{(k)} - \hat{P}_{ic}^{(k)}\tilde{P}_{rj}^{(k)}$$

$$(i=1, \ldots, r-1;\ j=1, \ldots, c-1;\ k=1, \ldots K), \tag{2.5}$$

then the random variables $U_{ij}^{(k)}$ also have expectation zero under the null hypothesis.

The multivariate random vectors

$$\mathbf{X}^{(k)} = (X_{11}^{(k)}, \ldots, U_{r-1\,c-1}^{(k)}). \tag{2.6}$$

and

$$\mathbf{U}^{(k)} = (U_{11}^{(k)}, \ldots, U_{r-1\,c-1}^{(k)}). \tag{2.7}$$

are each distributed asymptotically normal, with mean vector $\mathbf{0}$ under H_o for each $k=1, \ldots, K$. Neither the vectors $\mathbf{X}^{(k)}$ nor the vectors $\mathbf{U}^{(k)}$ are, however, independent, so their covariances must be evaluated in order to use them in test statistics.

3. Approximations of the Covariances

In the appendix it is shown that

$$\begin{aligned}
\operatorname{cov}(X_{ij}^{(k)}, X_{fg}^{(l)}) = 2\Big\{ & \sum_{h=1}^{L} W_h^2 [S_{(ij)(fg)h} \\
& - P_{i\cdot}S_{(\cdot j)(fg)h} - P_{\cdot j}S_{(i\cdot)(fg)h} \\
& - P_{f\cdot}S_{(i,j)(\cdot g)h} - P_{\cdot g}S_{(ij)(f\cdot)h} \\
& + P_{i\cdot}P_{f\cdot}S_{(\cdot j)(\cdot g)h} + P_{i\cdot}P_{\cdot g}S_{(\cdot j)(f\cdot)h} \\
& + P_{\cdot j}P_{f\cdot}S_{(i\cdot)(\cdot g)h} + P_{\cdot j}P_{\cdot g}S_{(i\cdot)(f\cdot)h}] \\
& + \Big(\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\Big)\Big(\sum_{h\in M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)h}\Big) \\
& + \Big(\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}\Big)\Big(\sum_{h\in M_{k,l}} W_h^2 S_{(\cdot j)(f\cdot)h}\Big) \\
& + \Big(\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\Big)\Big(\sum_{h\notin M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)h}\Big) \\
& + \Big(\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}\Big)\Big(\sum_{h\notin M_{k,l}} W_h^2 S_{(\cdot j)(f\cdot)h}\Big)\Big\}
\end{aligned} \tag{3.1}$$

where the parameters

$$S_{(ij)(fg)h} = \operatorname{cov}(\hat{P}_{ijha}, \hat{P}_{fgha})$$

$$S_{(i\cdot)(fg)h} = \operatorname{cov}(\hat{P}_{i\cdot ha}, \hat{P}_{fgha}) = \sum_{j=1}^{c} S_{(ij)(fg)h}$$

and similarly,

$S_{(\cdot j)(fg)h},\ S_{(ij)(f\cdot)h},\ S_{(ij)(\cdot g)h}$

$$S_{(i\cdot)(f\cdot)h} = \text{cov}\ (\hat{P}_{i\cdot ha}, \hat{P}_{f\cdot ha}) = \sum_{j=1}^{c} \sum_{g=1}^{c} S_{(ij)(fg)h}$$

and similarly,

$$S_{(i\cdot)(\cdot g)h},\ S_{(\cdot j)(f\cdot)h},\ S_{(\cdot j)(\cdot g)h} \tag{3.2}$$

are the covariances between the estimates of cell probabilities and marginal probabilities from the same PSU within the h^{th} stratum and the set $M_{k,l}$ is defined by

$$M_{k,l} = \{h : \alpha_h^{(k)} = \alpha_h^{(l)}\} \subset \{1, \ldots, L\}, \tag{3.3}$$

i.e., the set of strata in which the same PSU's are selected for the k^{th} and the l^{th} half sample. Similarly, it is easy to see that

$$\begin{aligned} \text{cov}\ (U_{ij}^{(k)}, U_{fg}^{(l)}) = {} & \text{cov}\ (\hat{P}_{ij}^{(k)} \tilde{P}_{rc}^{(k)}, \hat{P}_{fg}^{(l)} \tilde{P}_{rc}^{(l)}) \\ & - \text{cov}\ (\hat{P}_{ij}^{(k)} \tilde{P}_{rc}^{(k)}, \hat{P}_{fc}^{(l)} \tilde{P}_{rg}^{(l)}) \\ & - \text{cov}\ (\hat{P}_{ic}^{(k)} \tilde{P}_{rj}^{(k)}, \hat{P}_{fg}^{(l)} \tilde{P}_{rc}^{(l)}) \\ & + \text{cov}\ (\hat{P}_{ic}^{(k)} \tilde{P}_{rj}^{(k)}, \hat{P}_{fc}^{(l)} \tilde{P}_{rg}^{(l)}) \end{aligned} \tag{3.4}$$

where, as is shown in the appendix,

$$\begin{aligned} \text{cov}\ (\hat{P}_{iu}^{(k)} \tilde{P}_{ru'}^{(k)}, \hat{P}_{fv}^{(l)} \tilde{P}_{rv'}^{(l)}) = {} & \Big(\sum_{h \in M_{k,l}} W_h^2 S_{(iu)(fv)h} \Big) \Big(\sum_{h \in M_{k,l}} W_h^2 S_{(ru')(rv')h} \Big) + \Big(\sum_{h \notin M_{k,l}} W_h^2 S_{(iu)(rv')h} \Big) \\ & \Big(\sum_{h \notin M_{k,l}} W_h^2 S_{(ru')(fv')h} \Big) + P_{iu} P_{fv} \sum_{h \notin M_{k,l}} W_h^2 S_{(ru')(rv')h} \\ & + P_{ru'} P_{rv'} \sum_{h \in M_{k,l}} W_h^2 S_{(iu)(fv)h} + P_{iu} P_{rv'} \sum_{h \in M_{k,l}} \\ & W_h^2 S_{(ru')(fv)h} + P_{ru'} P_{fv} \sum_{h \notin M_{k,l}} W_h^2 S_{(iu)(rv')h} \end{aligned}$$

for $(u, u') = (j, c),\ (c, j);\ (v, v') = (g, c),\ (c, g).$ (3.5)

In order to obtain simpler approximate expressions for the covariances, the following assumptions are made:

(a) For each stratum, h, a value, n_h, which depends only on the number of final units per PSU in stratum h, can be determined so that the first two moments of the variables $(n_h \hat{P}_{ijha})$ are approximately those of the multinomial distribution with parameters $(n_h, \{P_{ijh}\})$. This holds, for instance, when the same number of final units are selected in both PSU's of the same stratum (if sampling within PSU's is simple random) or when the same effective sample sizes are attained within both PSU's of the same stratum (if sampling within PSU's is clustered and intraclass correlations within strata are independent of (i,j)).

Under this assumption we obtain

$$S_{(ij)(fg)h} \simeq \frac{1}{n_h} [\delta_i^f \delta_j^g P_{ijh} - P_{ijh} P_{fgh}] \tag{3.6}$$

where δ_i^f is a Kronecker delta (equals 1 if $i=f$ and 0 otherwise). The values of $S_{(i.)(f.)h}$, $S_{(.j)(.g)h}$, $S_{(i.)(.g)h}$, $S_{(.j)(f.)h}$, $S_{(ij)(f.)h}$, $S_{(ij)(.g)h}$, $S_{(i.)(fg)h}$, and $S_{(.j)(fg)h}$ are obtained by summing (3.6) over the relevant indexes.

(b)

$$n_h \simeq W_h / f_0 (h = 1, \ldots, L) \tag{3.7}$$

where f_0 is some constant. This implies that the number of final sample units per PSU in a stratum (for the case of simple random sampling within PSU's) or the effective sample size (for the case of clustered sampling) is proportional to the weight of the stratum.

(c)

$$\sum_{h \in M_{k,l}} W_h P_{ijh} P_{fgh} \simeq w_{k,l} P_{ij} P_{fg} \tag{3.8}$$

where

$$w_{k,l} = \sum_{h \in M_{k,l}} W_h.$$

This holds exactly if the cell probabilities are independent of the stratum. In particular, (3.6) implies

$$\sum_{h=1}^{L} W_h P_{ijh} P_{fgh} \simeq P_{ij} P_{fg} \tag{3.9}$$

since $w_{k,k} = 1$.

Substituting the approximations (3.6), (3.7), and (3.8) and the hypothesis (2.1) in (3.1), we obtain under H_o:

$$\begin{aligned}\mathrm{cov}\,(X_{ij}^{(k)}, X_{fg}^{(l)}) &\simeq 2f_o\{\delta_i^f\delta_j^g P_{ij} - \delta_i^f P_{ij}P_{\cdot g} \\ &\quad - \delta_j^g P_{ij}P_{f\cdot} + P_{ij}P_{fg} \\ &\quad + f_o[w_{k,l}^2 + (1-w_{k,l})^2] \\ &\quad (\delta_i^f P_{i\cdot} - P_{i\cdot}P_{f\cdot}) \\ &\quad (\delta_j^g P_{\cdot j} - P_{\cdot j}P_{\cdot g})\} \\ &= 2f_o\{1 + f_o[w_{k,l}^2 + (1-w_{k,l})^2]\} \\ &\quad (\delta_i^f P_{i\cdot} - P_{i\cdot}P_{f\cdot}) \\ &\quad (\delta_j^g P_{\cdot j} - P_{\cdot j}P_{\cdot g}). \qquad (3.10)\end{aligned}$$

Thus, for the X statistic, the ratio of the covariance between cell estimates from different half samples to that of estimates from the same half sample is independent of the specific cells. This ratio is defined by

$$\rho_X(k,l) = \frac{\mathrm{cov}(X_{ij}^{(k)}, X_{fg}^{(l)})}{\mathrm{cov}(X_{ij}^{(k)}, X_{fg}^{(k)})} \simeq \frac{1 + f_o[w_{k,l}^2 + (1-w_{k,l})^2]}{1+f_o}. \qquad (3.11)$$

The selection of balanced half samples ensures that the number of strata with PSU's common to two different half samples is approximately constant over all possible pairs of different half samples. Thus the number of terms in the set $M_{k,l}$ for $k \neq l$ is approximately independent of k and l. The further assumption will be made that the sum of the weights of the strata with PSU's common to two half samples is approximately constant, i.e.,

$$\sum_{h \in M_{k,l}} W_h = w_{k,l} = \begin{cases} w = \dfrac{1}{K(K-1)} \sum\limits_{k \neq l} w_{k,l};\ k \neq l \\ 1;\ k = l \end{cases} \qquad (3.12)$$

It follows from (3.11) and (3.12) that the covariances for different half samples relate to those for the same half sample in a fixed ratio of approximately

$$\frac{\mathrm{cov}\,(X_{ij}^{(k)}, X_{fg}^{(l)})}{\mathrm{cov}\,(X_{ij}^{(k)}, X_{fg}^{(k)})} \simeq \rho X = \frac{1 + f_o[w^2 + (1-w)^2]}{1+f_o}\ (k \neq l). \qquad (3.13)$$

Substituting the approximations (3.6), (3.7), and (3.8) and the hypothesis (2.1) in (3.4) and (3.5), we obtain for the covariances between the U statistics

$$\mathrm{cov}\,(U_{ij}^{(k)}, U_{fg}^{(l)}) = \begin{cases} 2f_o^2 w^2 \delta_i^f \delta_j^g P_{ij} P_{rc} + f_o w(1-f_o w) P_{rc} \\ (\delta_i^f P_{rj} + P_{fj})\ (\delta_j^g P_{ic} + P_{ig});\ k \neq l \\ 2f_o^2 \delta_i^f \delta_j^g P_{ij} P_{rc} + f_o(1-f_o) P_{rc} \\ (\delta_i^f P_{rj} + P_{fj})\ (\delta_j^g P_{ic} + P_{ig});\ k = l. \end{cases} \qquad (3.14)$$

Thus the ratio of the covariance for different subsamples to that for the same subsamples is fixed for $(i,j) \neq (f,g)$

$$\rho_U = \frac{\mathrm{cov}\,(U_{ij}^{(k)}, U_{fg}^{(l)})}{\mathrm{cov}\,(U_{ij}^{(k)}, U_{fg}^{(k)})} = \frac{w(1-f_o w)}{(1-f_o)}$$

$$\text{for}\quad k \neq l;\ (i,j) \neq (f,g). \qquad (3.15)$$

If f_o is small, (3.15) will also hold approximately for $(i,j) = (f,g)$.

4. Tests of the Hypothesis

Chapman[4] has derived a test of the null hypothesis (2.1) based on the statistics

$$Z_{ij}^{(k)} = \hat{P}_{ij}^{(k)} - \tilde{P}_{i\cdot}^{(k)}\tilde{P}_{\cdot j}^{(k)}. \qquad (4.1)$$

The test relies on the following assumptions under H_0:

(a) $$E(Z_{ij}^{(k)}) = 0; \qquad (4.2)$$

(b) $$\mathrm{cov}(Z_{ij}^{(k)}, Z_{fg}^{(k)}) = \mathrm{cov}(Z_{i'j'}^{(k)}, Z_{f'g'}^{(k)}) \qquad (4.3)$$

for all $i \neq f, j \neq g, i' \neq g', j' \neq g'; k = 1, \ldots, K$;

and

(c) $$\mathrm{cov}(Z_{ij}^{(k)}, Z_{fg}^{(l)}) = 0 \qquad (4.4)$$

for all $k \neq l$ and all i, j, f, g.

While (c) holds approximately for large L, it can easily be shown, on the basis of computations similar to those in the appendix, that (a) and (b) do not hold in general, even approximately.

The statistics $\mathbf{X}^{(k)}$ defined by (2.6) do, however, have expectation zero, and approximate tests of the hypothesis can be derived on the basis of the covariance approximations of the previous section. Set

$$\mathbf{Y}'_k=(Y_{k1},\ldots,Y_{kp})$$
$$=(X^{(k)}_{1,1},\ldots,X^{(k)}_{1,c-1},\ldots,X^{(k)}_{r-1,1},\ldots,$$
$$X^{(k)}_{r-1,c-1}),\ (k=1,\ldots,K) \quad (4.5)$$

where $p=(r-1)(c-1)$. Then asymptotically,

$$\mathbf{Y}_k \sim N(\boldsymbol{\mu},\Sigma)\ (k=1,\ldots,K) \quad (4.6)$$

with

$$\boldsymbol{\mu}'=(\mu_1,\ldots,\mu_p) \quad (4.7)$$

and

$$\Sigma=((\sigma_{uv}))=((\mathrm{cov}\,(Y_{ku},Y_{kv})))\,(u,v=1,\ldots,p), \quad (4.8)$$

is defined by the appropriate value of cov $(X^{(k)}_{ij}, X^{(k)}_{fg})$ independently of k according to (3.1) since $M_{k,k}=\{1,\ldots,L\}$.

Next, according to the approximation (3.13), we have

$$\rho\Sigma=((\rho\sigma_{uv}))$$
$$=((\mathrm{cov}\,(Y_{ku},Y_{lv})))\,(u,v=1,\ldots,p) \quad (4.9)$$

for any $k\neq l$ where $\rho=\rho X$ is defined by (3.13). Morrison[8] has shown that, under the conditions (4.6)–(4.9), if we define

$$\bar{\mathbf{Y}}=\frac{1}{K}\sum_{k=1}^{K}\mathbf{Y}_k, \quad (4.10)$$

then

(a) $$\bar{\mathbf{Y}}\sim N\left(\boldsymbol{\mu},\left[\frac{1+(K-1)\rho}{K}\right]\Sigma\right); \quad (4.11)$$

(b) $$\mathbf{A}=\sum_{k=1}^{K}(\mathbf{Y}_k-\bar{\mathbf{Y}})(\mathbf{Y}_k-\bar{\mathbf{Y}})'$$
$$\sim W\left(p,K-1,(1-\rho)\Sigma\right); \quad (4.12)$$

i.e.,
A is distributed p-Wishart with $K-1$ degrees of freedom and variance matrix $(1-\rho)\Sigma$; and
(c) $\mathbf{Y}$ and $\mathbf{A}$ are independent.

Two different test procedures can be suggested based on the above results.
(a) From (4.11), under H_0

$$\bar{\mathbf{Y}}\sim N\left(\mathbf{O},\left[\frac{1+(K-1)\rho}{K}\right]\Sigma\right). \quad (4.13)$$

Let $\tilde{\Sigma}$ be the estimate of Σ obtained by substituting the sample estimates of P_{ijh} in (3.1) or the sample estimates of P_{ij} in (3.10) for $k=l$. Set

$$\mathbf{C}=\frac{1+(K-1)\rho}{K}\tilde{\Sigma}. \quad (4.14)$$

Then

$$G=\bar{\mathbf{Y}}'\mathbf{C}^{-1}\bar{\mathbf{Y}} \quad (4.15)$$

is the approximate large sample test statistic, distributed asymptotically χ^2 with p degrees of freedom[9] under H_0.
(b) Define

$$\mathbf{B}=\frac{1+(K-1)\rho}{1-\rho}\mathbf{A}. \quad (4.16)$$

Then

$$\mathbf{B}\sim W(p,K-1,[1+(K-1)\rho]\Sigma). \quad (4.17)$$

Thus

$$T^2=K(K-1)\,\bar{\mathbf{Y}}'\mathbf{B}^{-1}\bar{\mathbf{Y}} \quad (4.18)$$

is distributed under H_0 as Hotelling's p-dimensional T^2 with K–1 degrees of freedom so that H_0 can be tested by comparing

$$F=\frac{K-p}{(K-1)p}T^2=\frac{K(K-p)}{p}\bar{\mathbf{Y}}'\mathbf{B}^{-1}\bar{\mathbf{Y}} \quad (4.19)$$

with the critical value of the F distribution with p and $K-p$ degrees of freedom.

The same tests can be performed with the U statistics (2.7). If we set

$$\mathbf{Y}'_k=(Y_{k1},\ldots,Y_{kp})=$$
$$(U^{(k)}_{1,1},\ldots,U^{(k)}_{1,c-1},\ldots,U^{(k)}_{r-1,1},\ldots,U^{(k)}_{r-1,c-1}),$$
$$(k=1,\ldots,K), \quad (4.20)$$

replace ρX by ρU (defined by (3.15)), and replace $\tilde{\Sigma}$ by the substitution of the sample estimates of P_{ij} in (3.4) and (3.5), the tests defined by (4.15) and (4.19) are valid under the same assumptions.

An alternative test using the cross product ratio could be based on the statistics

$$V_{ij}=\ln\left[\frac{\hat{P}_{ij}\hat{P}_{rc}}{\hat{P}_{ic}\hat{P}_{rj}}\right]$$

$$(i=1,\ldots,r-1;j=1,\ldots,c-1). \qquad (4.21)$$

The covariance matrix of these statistics can be approximated by the appropriate Taylor expansion as

$$\operatorname{cov}(V_{ij},V_{fg})=\sum_{\substack{u=1,r\\v=j,c}}\ \sum_{\substack{u'=f,r\\v'=g,c}}(-1)^{\alpha(u,v,u',v')}\frac{\operatorname{cov}(\hat{P}_{uv},\hat{P}_{u'v'})}{P_{uv}P_{u'v'}} \qquad (4.22)$$

where $\alpha(u,v,u',v')=\delta_u^r+\delta_v^c+\delta_{u'}^r+\delta_{v'}^c$.

The covariances, $\operatorname{cov}(\hat{P}_{uv},\hat{P}_{u'v'})$, can then be estimated by the balanced half-sample method:

$$\hat{\operatorname{cov}}(\hat{P}_{uv},\hat{P}_{u'v'})=\frac{1}{4K}\sum_{k=1}^{K}(\hat{P}_{uv}^{(k)}-\tilde{P}_{uv}^{(k)})(\hat{P}_{u'v'}^{(k)}-\tilde{P}_{u'v'}^{(k)}). \qquad (4.23)$$

Finally, set $\mathbf{V}'=(V_{11},\ldots,V_{r-1,c-1})$ and let $\mathbf{C}=((\hat{\operatorname{cov}}(V_{ij},V_{fg})))$ be defined by (4.22) with the covariances estimated by (4.23) and with $P_{uv},P_{u'v'}$ replaced by their sample estimates $\hat{P}_{uv}$ and $\hat{P}_{u'v'}$, respectively. Then the large sample Wald statistic $\mathbf{VC}^{-1}\mathbf{V}'$ can be used to test the null hypothesis with asymptotic distribution under the null hypothesis of $\chi^2(p)$.

5. Numerical Examples

The data used for the examples are from the noncertainty urban strata of the Israel Labour Force Survey for the period October–December 1968. The primary sampling units are towns and the stratification criteria are size, region, and type of population. Two PSU's are selected within each stratum with probability proportional to size (number of inhabitants), but, as size varies little within strata, the selection can be regarded for all practical purposes as equal probability sampling. Within PSU's, households (25–80 per PSU) are sampled random-systematically, so final selection probabilities are equal. For the first example, the characteristics cross-classified were labor force participation (2 classes) and age (5 classes). Simple sample estimates of the values of P_{ijh} were obtained from the two PSU's in each stratum and were used together with the average sample size, n_h, in each stratum, to obtain estimates of $cov(X_{ij}^{(k)},X_{fg}^{(l)})$ as defined by (3.1).

Five different approximations of (3.1) were compared as follows:

(a) $\hat{cov}_0(X_{ij}^{(k)},X_{fg}^{(l)})$ –
obtained from (3.1) by substitution of the approximations (3.6) for the covariances defined by (3.2)—assumption (a) of section 3.

(b) $\hat{cov}_1(X_{ij}^{(k)},X_{fg}^{(l)})$ –
obtained from the previous approximation by substitution of the approximation (3.7)—assumption (b) of section 3.

(c) $\hat{cov}_2(X_{ij}^{(k)},X_{fg}^{(l)})$ –
obtained from the previous approximation by substitution of the approximation (3.8)—assumption (c) of section 3.

(d) $\hat{\mathrm{cov}}_3(X_{ij}^{(k)},X_{fg}^{(l)})$—
obtained from the previous approximation by substitution of the approximation (3.12).

(e) $\hat{\mathrm{cov}}_4(X_{ij}^{(k)},X_{fg}^{(l)})$—
obtained from the previous approximation by substitution of the null hypothesis (2.1), i.e., substitution of (3.10) with the approximation (3.12).

Table A gives the values of these approximations for $k=l$ (independent of the value of k) and their average, maximal, and minimal values over all pairs $k\neq l$. It should be noted that by definition $\hat{\mathrm{cov}}_2(X_{ij}^{(k)},X_{fg}^{(l)})=\hat{\mathrm{cov}}_3(X_{ij}^{(k)},X_{fg}^{(l)})$ for $k=l$ and that $\hat{cov}_3(X_{ij}^{(k)},X_{fg}^{(l)})$ and $\hat{\mathrm{cov}}_4(X_{ij}^{(k)},X_{fg}^{(l)})$ are independent of the values of k and l for all $k\neq l$ and for all $k=l$.

While the differences between the last four approximations are slight, it can be seen from table A that the first approximation differs from them considerably in some cases. This is due to the fact that in this example there are some serious departures from (3.7) (assumption (b) of section 3), as can be seen from table B.

The difference between n_h and W_h/f is due in this case to large PSU size variations within strata. It should, however, be pointed out that even the large departures from (3.7) do not affect the covariance approximation very seriously. The other assumptions made for the remaining approximations have virtually no effect.

Table A. *Approximations of cov* $(X_{ij}^{(k)}, X_{fg}^{(l)})$ $(\times 10^5)$

	Approximation	$i=1$, $j=1$: $k \neq l$ Mean	$k \neq l$ Maximum	$k \neq l$ Minimum	$k=l$	$i=1$, $j=2$: $k \neq l$ Mean	$k \neq l$ Maximum	$k \neq l$ Minimum	$k=l$	$i=1$, $j=3$: $k \neq l$ Mean	$k \neq l$ Maximum	$k \neq l$ Minimum	$k=l$	$i=1$, $j=4$: $k \neq l$ Mean	$k \neq l$ Maximum	$k \neq l$ Minimum	$k=l$
$f=1$, $g=1$	(a)	14.005	14.010	14.004	14.021												
	(b)	13.039	13.041	13.039	13.053												
	(c)	13.288	13.291	13.288	13.302												
	(d)		13.288		13.302												
	(e)		13.123		13.137												
$f=1$, $g=2$	(a)	−2.412	−2.412	−2.413	−2.415	14.943	14.949	14.942	14.960								
	(b)	−2.379	−2.379	−2.379	−2.381	13.885	13.887	13.885	13.900								
	(c)	−2.341	−2.341	−2.342	−2.344	14.017	14.020	14.017	14.032								
	(d)		−2.341		−2.344		14.017		14.032								
	(e)		−2.309		−2.312		14.085		14.100								
$f=1$, $g=3$	(a)	−4.356	−4.356	−4.358	−4.363	−3.831	−3.831	−3.833	−3.836	20.703	20.712	20.701	20.726				
	(b)	−3.063	−3.063	−3.064	−3.067	−3.295	−3.295	−3.294	−3.298	17.245	17.249	17.244	17.263				
	(c)	−3.006	−3.006	−3.006	−3.009	−3.349	−3.349	−3.349	−3.352	17.476	17.479	17.475	17.495				
	(d)		−3.006		−3.009		−3.349		−3.352		17.475		17.495				
	(e)		−3.123		−3.126		−3.401		−3.405		17.849		17.868				
$f=1$, $g=4$	(a)	−3.660	−3.660	−3.662	−3.663	−5.182	−5.181	−5.184	−5.188	−7.983	−7.982	−7.987	−7.992	24.458	24.469	24.456	24.488
	(b)	−4.147	−4.147	−4.148	−4.152	−4.953	−4.953	−4.954	−4.959	−6.587	−6.586	−6.588	−6.593	21.941	21.945	21.940	21.966
	(c)	−4.392	−4.392	−4.393	−4.397	−4.993	−4.993	−4.993	−4.998	−6.844	−6.844	−6.845	−6.851	22.475	22.480	22.475	22.501
	(d)		−4.392		−4.397		−4.993		−4.998		−6.844		−6.851		22.475		22.501
	(e)		−4.671		−4.676		−5.086		−5.092		−6.878		−6.886		23.286		23.312

Table B. *Values of* n_h *and* W_h/f

h	1	2	3	4	5	6	7	8	9	10
n_h	71.0	65.5	30.0	41.0	54.0	39.0	34.5	39.5	41.0	40.0
W_h/f	60.5	56.8	19.1	45.6	61.8	59.4	61.4	44.6	29.2	17.3

The values of the Chapman statistics [4] obtained for this example are

$$T_1 = 45.0, \text{ based on } (4.1),$$

and

$T_2 = 51.0$, based on the dual of (4.1) with $\hat{P}_{ij}$, $\tilde{P}_{i\cdot}$, and $\tilde{P}_{\cdot j}$ replaced by $\tilde{P}_{ij}$, $\hat{P}_{i\cdot}$, and $\hat{P}_{\cdot j}$, respectively.

Four different values of the G statistic defined by (4.15) for the X statistic were calculated:

$$G_0 = 49.2$$
$$G_1 = 56.2$$
$$G_2 = G_3 = 55.7$$
$$G_4 = 53.3$$

where G_a $(a = 0, 1, 2, 3, 4)$ is based on the approximation cov_a $(X_{ij}^{(k)}, X_{fg}^{(l)})$.

It can be seen that the differences between the G statistics, due to the various simplifying assumptions, are small.

The G_o value (4.15) for the U statistic (2.7) was 39.0 in this example. While this is considerably lower than the value obtained for the X statistic, it together with the remaining values obtained, still far exceeds the critical chi-square value at any practical level of significance.

A further comparison of values of G_0 for the X and U statistics was made on three 3 x 2 contingency tables from the same survey which indicated much smaller departures from the null hypothesis. The values obtained were as follows:

Data set:	I	II	III
G_0 for X statistic:	.953	3.93	12.93
G_0 for U statistic:	.928	3.90	12.16

These values are close enough for all practical purposes. The other statistics, however, performed poorly for these examples, showing large divergences.

Simulations of 350 sets of sample frequencies for the same 10 strata 3 x 2 table were obtained from three sets of cell probabilities, one of which satisfied the null hypothesis while the other two represented increasing departures from the null hypothesis (see details in Nathan's paper).[10] From each set of sample frequencies, the Chapman statistics (T_1, T_2), three approximations of Wilks' statistics (G_1, G_2, G_3), and Hotelling's F were computed for the X statistic. The relative frequencies of the number of times each of the statistics exceeded the critical chi-square values for nominal levels of significance of .01, .05, and .10 are given in table C. These relative frequencies estimate the powers of the statistic and again indicate small differences between the two variants of Chapman's statistic and between the various approximations of Wilks' statistics. In general, higher estimated powers are achieved for statistics with higher estimated levels of significance, but Hotelling's statistic indicates smaller power than Wilks even though it has a higher actual level of significance.

TABLE C. *Relative frequencies of times nominal significance level exceeded using nonproportional sampling (350 simulations)*

Hypothesis	Significance level	Chapman		Wilks			Hotelling
		T_1	T_2	G_1	G_2	G_3	F
H_0	.01	.011	.017	.029	.026	.026	.083
	.05	.046	.037	.103	.097	.091	.180
	.10	.071	.169	.166	.166	.160	.286
H_1	.01	.060	.063	.151	.149	.149	.191
	.05	.160	.146	.326	.309	.306	.337
	.10	.226	.209	.437	.420	.417	.437
H_2	.01	.337	.311	.577	.563	.563	.420
	.05	.497	.506	.760	.754	.754	.654
	.10	.591	.594	.837	.834	.831	.754

In order to eliminate the effect of the different actual levels of significance, unbiased estimates of the Expected Significance Level (ESL) proposed by Dempster and Schatzoff[11] were computed for each alternative. The estimated ESL is the Mann-Whitney statistic, which is based on comparisons of the values of statistics obtained under the null hypothesis with those obtained under the alternative hypotheses and measures the relative efficiencies of the statistics independently of the actual significance levels attained.

The estimated ESL values based on 250 simulations for each alternative are given in table D. As before, the results indicate that the differences between variations of the statistics within groups are small as compared with the differences between groups and are, in fact, not significant, while differences between groups are significant (at the 1-percent level). The results thus indicate that for the given parameters of the two alternatives, Wilks' statistic, with any of the three approximations, is more efficient than Chapman's (either variation), while Hotelling's statistic is less efficient than Chapman's.

TABLE D. *Estimates and rank of Expected Significance Level (ESL) using nonproportional sampling* (250 *simulations*)

Statistic	H_1		H_2	
	ESL	Rank	ESL	Rank
Chapman $-T_1$	.6301	4	.8734	4
Chapman $-T_2$	.6249	5	.8735	5
Wilks $-G_1$	.6554	1	.9080	1
Wilks $-G_2$	.6547	3	.9078	3
Wilks $-G_3$	.6549	2	.9075	2
Hotelling $-F$	.5999	6	.8005	6

A further 250 simulations were carried out for each hypothesis, with sample sizes proportional to strata weights (i.e., $n_h = nW_h$). In this case, taking into account the previous results, only one statistic from each group was computed—Chapman's T_1, Wilks' G_1, and Hotelling's F. In addition, the log-likelihood ratio statistic based on the overall marginal table was computed as follows:

$$H = -2\left[n(\ln n) - \sum_i n_{i\cdot}(\ln n_{i\cdot}) - \sum_j n_{\cdot j}(\ln n_{\cdot j}) + \sum_{i,j}(\ln n_{ij})\right]$$

and compared with the critical values of $\chi^2(p)$.

Both from the relative frequencies of times the critical values were exceeded, given in table E, and from the estimated ESL's given in table F, it is seen that the naive test has greater power than the test based on Wilks' statistic although the difference in ESL is not significant. Thus this computationally simple test can be used in the case of proportional sampling without any loss of efficiency.

TABLE E. *Relative frequencies of times nominal significance level exceeded using proportional sampling (250 simulations)*

Hypothesis	Significance level	Chapman T_1	Wilks G_1	Hotelling F	Log-likelihood—H
H_0	.01	.004	.016	.064	.016
	.05	.016	.052	.156	.048
	.10	.040	.092	.196	.092
H_1	.01	.052	.116	.160	.100
	.05	.140	.272	.336	.256
	.10	.204	.360	.456	.356
H_2	.01	.404	.604	.448	.592
	.05	.584	.792	.684	.776
	.10	.660	.864	.780	.864

TABLE F. *Estimates and rank of Expected Significance Level (ESL) using proportional sampling (250 simulations)*

Statistic	H_1		H_2	
	ESL	Rank	ESL	Rank
Chapman—T_1	.7041	3	.9232	3
Wilks—G_1	.7336	2	.9463	2
Hotelling—F	.6848	4	.8703	4
Log-likelihood—H	.7351	1	.9466	1

It should be noted that the ranking of the ESL's of the statistics used in the nonproportional sampling case remains the same in the proportional case, thus strengthening the previous results.

6. The Case of a 2x2 Table

For the special case of a 2x2 table ($r=c=2$), some simplifications of the tests are possible. Thus the statistics (2.4) and (2.5) become

$$X^{(k)}=\hat{P}_{11}^{(k)}+\tilde{P}_{11}^{(k)}-\hat{P}_{1\cdot}^{(k)}\tilde{P}_{\cdot 1}^{(k)}-\tilde{P}_{1\cdot}^{(k)}\hat{P}_{\cdot 1}^{(k)}$$

$$=[\hat{P}_{11}^{(k)}\tilde{P}_{22}^{(k)}-\hat{P}_{12}^{(k)}\tilde{P}_{21}^{(k)}]+[\hat{P}_{22}^{(k)}\tilde{P}_{11}^{(k)}-\hat{P}_{21}^{(k)}\tilde{P}_{12}^{(k)}] \qquad (6.1)$$

and

$$U^{(k)}=\hat{P}_{11}^{(k)}\tilde{P}_{22}^{(k)}-\hat{P}_{12}^{(k)}\tilde{P}_{21}^{(k)} \qquad (6.2)$$

so that

$$X^{(k)}=U^{(k)}+U'^{(k)} \qquad (6.3)$$

where

$$U'^{(k)}=\tilde{P}_{11}^{(k)}\hat{P}_{22}^{(k)}-\tilde{P}_{12}^{(k)}\hat{P}_{21}^{(k)}. \qquad (6.4)$$

$U^{(k)}$ and $U'^{(k)}$ can be shown to be independent and, from (3.10) and (3.14) under the null hypothesis,

$$\mathrm{Var}\,(X^{(k)})=2f_o(1+f_o)P_{11}P_{22}=2\,\mathrm{Var}\,(U^{(k)}). \qquad (6.5)$$

The variates (4.5) are univariate, so that if

$$\bar{X}=\frac{1}{K}\sum_{k=1}^{K}X^{(k)};\ \bar{U}=\frac{1}{K}\sum_{k=1}^{K}U^{(k)}, \qquad (6.6)$$

the test statistics to be used instead of (4.15) are

$$\sqrt{G_x}=\sqrt{\frac{K}{1+(K-1)\rho X}}\,\frac{\bar{X}}{\sqrt{2f_0(1+f_0)\,\hat{P}_{11}\hat{P}_{22}}} \qquad (6.7)$$

and

$$\sqrt{G_u}=\sqrt{\frac{K}{1+(K-1)\rho U}}\,\frac{\bar{U}}{\sqrt{f_0(1+f_0)\hat{P}_{11}\hat{P}_{22}}}, \qquad (6.8)$$

both distributed asymptotically standard normal under H_0. Similarly, (4.18) and (4.19) can be replaced by

$$T_x=\sqrt{\frac{1-\rho X}{1+(K-1)\rho X}}\,\frac{\bar{X}}{\sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(X^{(k)}-\bar{X})^2/K}} \qquad (6.9)$$

and

$$T_u=\sqrt{\frac{1-\rho U}{1+(K-1)\rho U}}\,\frac{\bar{U}}{\sqrt{\frac{1}{K-1}\sum_{k=1}^{K}(U^{(k)}-\bar{U})^2/K}} \qquad (6.10)$$

and compared with the critical Student's t values (with $k-1$ degrees of freedom).

REFERENCES

[1] Bhapkar, V. P.: Some tests for categorical data. *The Annals of Mathematical Statistics.* 32(1):72–83, March 1961.

[2] Garza-Hernandez, T., An Approximate Test of Homogeneity on the Basis of a Stratified Random Sample. Unpublished master's thesis, Cornell University, 1961.

[3] Nathan, G.: Tests of independence in contingency tables from stratified samples, in Johnson, N. L., and Smith, H., Jr., eds., *New Developments in Survey Sampling.* New York. Wiley-Interscience, 1969. pp. 578–600.

[4] Chapman, D. W., An Approximate Test of Independence Based on Replications of a Complex Sample Survey Design. Unpublished master's thesis, Cornell University, 1966.

[5] McCarthy, P. J.: Pseudoreplication: Half samples. *Review of the International Statistical Institute.* 37(3) : 239–264, 1969.

[6] National Center for Health Statistics: Replication: An approach to the analysis of data from complex surveys. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 14. Public Health Service. Washington. U.S. Government Printing Office, Apr. 1966.

[7] Bhapkar, V. P., and Koch, G. G.: On the hypotheses of 'no interaction' in contingency tables. *Biometrics.* 24(3) : 567–594, Sept. 1968.

[8] Morrison, D. F.: On the distribution of squares and cross products of normal variates in the presence of intra-class correlation. *The Annals of Mathematical Statistics.* 33(4):1461–1463, Dec. 1962.

[9] Wilks, S. S.: *Mathematical Statistics.* New York. Wiley, 1962.

[10] Nathan, G.: A simulation comparison of tests for independence in stratified cluster sampling. *Bulletin of the International Statistical Institute* 14(2): 289–295, 1971.

[11] Dempster, A. P., and Schatzoff, M.: Expected significance level as a sensitivity index for test statistics. *Journal of the American Statistical Association.* 60(310):420–436, June 1965.

ooo

APPENDIX I

PROOFS OF (3.1) AND (3.5)

Denote for fixed values of k, l

$$M(1) = M_{k,l}$$

and

$$M(2) = \{h : h \notin M_{k,l}\}, \qquad (A.1)$$

and define

$$\hat{P}_{ij}^{(k)}(u) = \sum_{h \in M(u)} W_h [\alpha_h^{(k)} \hat{P}_{ijhl} + (1-\alpha_h^{(k)}) \hat{P}_{ijh2}],$$

$$\tilde{P}_{ij}^{(k)}(u) = \sum_{h \in M(u)} W_h [(1-\alpha_h^{(k)}) \hat{P}_{ijhl} + \alpha_h^{(k)} \hat{P}_{ijh2}]$$

$$(u=1,2; i=1,\ldots,r; j=1,\ldots,c; k=1,\ldots,K). \qquad (A.2)$$

Then

$$\hat{P}_{ij}^{(k)} = \hat{P}_{ij}^{(k)}(1) + \hat{P}_{ij}^{(k)}(2),$$

$$\tilde{P}_{ij}^{(k)} = \tilde{P}_{ij}^{(k)}(1) + \tilde{P}_{ij}^{(k)}(2). \qquad (A.3)$$

Also,

$$E[\hat{P}_{ij}^{(k)}(u)] = E[\tilde{P}_{ij}^{(k)}(u)] = P_{ij}(u), \qquad (A.4)$$

where

$$P_{ij}(u) = \sum_{h \in M(u)} W_h [\alpha_h^{(k)} P_{ijh1} + (1-\alpha_h^{(k)}) P_{ijh2}]. \qquad (A.5)$$

Thus

$$\operatorname{cov}(\hat{P}_{ij}^{(k)} \tilde{P}_{i'j'}^{(k)}, \hat{P}_{fg}^{(l)} \tilde{P}_{f'g'}^{(l)})$$

$$= \operatorname{cov}\left[\sum_{u,v=1}^{2} \hat{P}_{ij}^{(k)}(u) \tilde{P}_{i'j'}^{(k)}(v), \sum_{u,v=1}^{2} \hat{P}_{fg}^{(l)}(u) \tilde{P}_{f'g'}^{(l)}(v)\right] = \sum_{(u_1,u_2,u_3,u_4)}$$

$$\operatorname{cov}[\hat{P}_{ij}^{(k)}(u_1) \tilde{P}_{i'j'}^{(k)}(u_2), \hat{P}_{fg}^{(l)}(u_3) \tilde{P}_{f'g'}^{(l)}(u_4) \qquad (A.6)$$

where the summation is over all the 16 possible combinations.

It can easily be seen that $\hat{P}_{ij}^{(k)}(u)$ is independent of $\tilde{P}_{i'j'}^{(k)}(v)$, $\hat{P}_{fg}^{(l)}(v)$, and $\tilde{P}_{f'g'}^{(l)}(v)$ for $u \neq v$ (and similarly for all other pairs of estimates from mutually exclusive subsets of strata). Also, the following are pairs of independent variables:

$$(\hat{P}_{ij}^{(k)}(1), \tilde{P}_{f'g'}^{(l)}(1)), \quad \tilde{P}_{i'j'}^{(k)}(1), \hat{P}_{fg}^{(l)}(1)),$$

$$(\hat{P}_{ij}^{(k)}(2), \hat{P}_{fg}^{(l)}(2)), (\tilde{P}_{i'j'}^{(k)}(2), \tilde{P}_{f'g'}^{(l)}(2)),$$

as each of the two estimates in any pair is derived from different PSU's in each stratum.

Each of the covariance terms in the summation (A.6) is of the form $\operatorname{cov}(xy, x'y')$, where the pair of random variables (x, x') is independent of the pair of random variables (y, y'). Under these conditions it is easily verified that

$$\operatorname{cov}(xy, x'y') = \operatorname{cov}(x, x') \operatorname{cov}(y, y')$$

$$+ E(x) E(x') \operatorname{cov}(y, y') + E(y) E(y') \operatorname{cov}(x, x').$$

In particular, if, in addition, y and y' are independent, then

$$\operatorname{cov}(xy, x'y') = E(y)\,E(y')\operatorname{cov}(x, x'). \tag{A.7}$$

The evaluation of the components of (A.6) is then obtained by noting that

$$\operatorname{cov}[\hat{P}_{ij}^{(k)}(1), \hat{P}_{fg}^{(l)}(1)] = E\left[\left\{\sum_{h\in M(1)} W_h(\hat{P}_{ijh}^{(k)} - P_{ijh})\right\}\left\{\sum_{h\in M(u)} W_h(\hat{P}_{fgh}^{(l)} - P_{fgh})\right\}\right]$$
$$= \sum_{h,h'}\sum_{\in M(1)} W_h W_{h'} E[(\hat{P}_{ijh}^{(k)} - P_{ijh})(\hat{P}_{fgh'}^{(l)} - P_{fgh'})]. \tag{A.8}$$

But

$$E[(\hat{P}_{ijh}^{(k)} - P_{ijh})(\hat{P}_{fgh'}^{(l)} - P_{fgh'})] = \begin{cases} S_{(ij)(fg)h};\ h=h' \text{ and } \alpha_h^{(k)} = \alpha_h^{(l)} \\ 0;\ \text{otherwise,} \end{cases} \tag{A.9}$$

where $S_{(ij)(fg)}$ is defined by (3.2). Thus

$$\operatorname{cov}[\hat{P}_{ij}^{(k)}(1), \hat{P}_{fg}^{(l)}(1)] = \operatorname{cov}(\tilde{P}_{ij}^{(l)}(1), \tilde{P}_{fg}^{(k)}(1)) = \sum_{h\in M(1)} W_h^2 S_{(ij)(fg)h}. \tag{A.10}$$

Similarly,

$$\operatorname{cov}[\hat{P}_{ij}^{(l)}(2), \tilde{P}_{fg}^{(k)}(2)] = \operatorname{cov}(\tilde{P}_{ij}^{(l)}(2), \hat{P}_{fg}^{(k)}(2)) = \sum_{h\in M(2)} W_h^2 S_{(ij)(fg)h}. \tag{A.11}$$

Using the above, the covariance terms of the summation (A.6) are evaluated as follows:

$$\operatorname{cov}(\hat{P}_{ij}^{(k)}(u)\tilde{P}_{i'j'}^{(k)}(u), \hat{P}_{fg}^{(l)}(u)\tilde{P}_{f'g'}^{(l)}(u)) = \left(\sum_{h\in M(u)} W_h^2 S_{(ij)(fg)h}\right)\left(\sum_{h\in M(u)} W_h^2 S_{(i'j')(f'g')h}\right)$$
$$+ P_{i'j'}(u)P_{f'g'}(u)\sum_{h\in M(u)} W_h^2 S_{(ij)(fg)h}$$
$$+ P_{ij}(u)P_{fg}(u)\sum_{h\in M(u)} W_h^2 S_{(i'j')(f'g')h} \tag{A.12}$$

for $u=1$, with the expression for $u=2$ obtained by interchanging (fg) and $(f'g')$. The other terms of (A.6) are obtained as follows:

$$\operatorname{cov}(\hat{P}_{ij}^{(k)}(u)\tilde{P}_{i'j'}^{(k)}(u), \hat{P}_{fg}^{(l)}(v)\tilde{P}_{f'g'}^{(l)}(v)) = 0 \quad \text{for } u \neq v. \tag{A.13}$$

$$\operatorname{cov}(\hat{P}_{ij}^{(k)}(1)\tilde{P}_{i'j'}^{(k)}(v_1), \hat{P}_{fg}^{(l)}(1)\tilde{P}_{f'g'}^{(l)}(v_2)) = P_{i'j'}(v_1)P_{f'g'}(v_2)\sum_{h\in M(1)} W_h^2 S_{(ij)(fg)h} \quad \text{for } (v_1, v_2) \neq (1,1). \tag{A.14}$$

$$\operatorname{cov}(\hat{P}_{ij}^{(k)}(v_1)\tilde{P}_{i'j'}^{(k)}(1), \hat{P}_{fg}^{(l)}(v_2)\tilde{P}_{f'g'}^{(l)}(1)) = P_{ij}(v_1)P_{fg}(v_2)\sum_{h\in M(1)} W_h^2 S_{(i'j')(f'g')h} \quad \text{for } (v_1, v_2) \neq (1,1). \tag{A.15}$$

$$\operatorname{cov}(\hat{P}_{ij}^{(k)}(2)\tilde{P}_{i'j'}^{(k)}(v_1), \hat{P}_{fg}^{(l)}(v_2)\tilde{P}_{f'g'}^{(l)}(2)) = P_{i'j'}(v_1)P_{fg}(v_2)\sum_{h\in M(2)} W_h^2 S_{(ij)(f'g')h} \quad \text{for } (v_1, v_2) \neq (2,2). \tag{A.16}$$

$$\operatorname{cov}(\hat{P}_{ij}^{(k)}(v_1)\tilde{P}_{i'j'}^{(k)}(2), \hat{P}_{fg}^{(l)}(2)\tilde{P}_{f'g'}^{(l)}(v_2)) = P_{ij}(v_1)P_{f'g'}(v_2)\sum_{h\in M(2)} W_h^2 S_{(i'j')(fg)h} \quad \text{for } (v_1, v_2) \neq (2,2). \tag{A.17}$$

Substituting (A.12)–(A.17) in (A.6), we obtain

$$
\begin{aligned}
&\operatorname{cov}\left(\hat{P}^{(k)}\tilde{P}^{(k)}_{i'j'},\ \hat{P}^{(l)}_{fg}\tilde{P}^{(l)}_{f'g'}\right)\\
&=\Bigl(\sum_{h\in M_{k,l}} W_h^2 S_{(ij)(fg)h}\Bigr)\Bigl(\sum_{h\in M_{k,l}} W_h^2 S_{(i'j')(f'g')h}\Bigr)\\
&+\Bigl(\sum_{h\notin M_{k,l}} W_h^2 S_{(ij)(f'g')h}\Bigr)\Bigl(\sum_{h\notin M_{k,l}} W_h^2 S_{(i'j')(fg)h}\Bigr)\\
&\qquad+P_{ij}P_{fg}\sum_{h\in M_{k,l}} W_h^2 S_{(i'j')(f'g')h}\\
&\qquad+P_{ij}P_{f'g'}\sum_{h\in M_{k,l}} W_h^2 S_{(ij)(fg)h}\\
&\qquad+P_{ij}P_{f'g'}\sum_{h\notin M_{k,l}} W_h^2 S_{(i'j')(fg)h}\\
&\qquad\qquad+P_{i'j'}P_{fg}\sum_{h\notin M_{k,l}} W_h^2 S_{(ij)(f'g')h}.
\end{aligned}\tag{A.18}
$$

This is the result used in (3.5).

To prove (3.1), note that

$$
\begin{aligned}
\operatorname{cov}\left(X^{(k)}_{ij}, X^{(l)}_{fg}\right) &= \operatorname{cov}\Bigl[\bigl(\hat{P}^{(k)}_{ij}+\tilde{P}^{(k)}_{ij}-\hat{P}^{(k)}_{i\cdot}\tilde{P}^{(k)}_{\cdot j}\\
&\quad-\tilde{P}^{(k)}_{i\cdot}\hat{P}^{(k)}_{ij}\bigr),\ \hat{P}^{(l)}_{fg}+\tilde{P}^{(l)}_{fg}-\hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}-\tilde{P}^{(l)}_{f\cdot}\hat{P}^{(l)}_{\cdot g}\bigr)\Bigr]\\
&= A-B+C,
\end{aligned}
$$

where

$$
\begin{aligned}
A &= \operatorname{cov}\left(\hat{P}^{(k)}_{ij}, \hat{P}^{(l)}_{fg}\right)+\operatorname{cov}\left(\hat{P}^{(k)}_{ij}, \tilde{P}^{(l)}_{fg}\right)\\
&\quad+\operatorname{cov}\left(\tilde{P}^{(k)}_{ij}, \hat{P}^{(l)}_{fg}\right)+\operatorname{cov}\left(\tilde{P}^{(k)}_{ij}, \tilde{P}^{(l)}_{fg}\right);\\
B &= \operatorname{cov}\left(\hat{P}^{(k)}_{ij}, \hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\right)+\operatorname{cov}\left(\hat{P}^{(k)}_{ij}, \tilde{P}^{(l)}_{f\cdot}\hat{P}^{(l)}_{\cdot g}\right)\\
&\quad+\operatorname{cov}\left(\tilde{P}^{(k)}_{ij}, \hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\right)+\operatorname{cov}\left(\tilde{P}^{(k)}_{ij}, \tilde{P}^{(l)}_{f\cdot}\hat{P}^{(l)}_{\cdot g}\right)\\
&\quad+\operatorname{cov}\left(\hat{P}^{(l)}_{fg}, \hat{P}^{(k)}_{i\cdot}\tilde{P}^{(k)}_{\cdot j}\right)+\operatorname{cov}\left(\hat{P}^{(l)}_{fg}, \tilde{P}^{(k)}_{i\cdot}\hat{P}^{(k)}_{\cdot j}\right)\\
&\quad+\operatorname{cov}\left(\tilde{P}^{(l)}_{fg}, \hat{P}^{(k)}_{i\cdot}\tilde{P}^{(k)}_{\cdot j}\right)+\operatorname{cov}\left(\tilde{P}^{(l)}_{fg}, \tilde{P}^{(k)}_{i\cdot}\hat{P}^{(k)}_{\cdot j}\right);\\
C &= \operatorname{cov}\left(\hat{P}^{(k)}_{i\cdot}\tilde{P}^{(k)}_{\cdot j}, \hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\right)+\operatorname{cov}\bigl(\tilde{P}^{(k)}_{i\cdot}\hat{P}^{(k)}_{\cdot j},\\
&\quad\hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\bigr)+\operatorname{cov}\left(\hat{P}^{(k)}_{i\cdot}\tilde{P}^{(k)}_{\cdot j}, \tilde{P}^{(l)}_{f\cdot}\hat{P}^{(l)}_{\cdot g}\right)\\
&\quad+\operatorname{cov}\left(\tilde{P}^{(k)}_{i\cdot}\hat{P}^{(k)}_{\cdot j}, \tilde{P}^{(l)}_{f\cdot}\hat{P}^{(l)}_{\cdot g}\right).
\end{aligned}\tag{A.19}
$$

Using (A.10) and (A.11), we obtain

$$
A = 2\sum_{h=1}^{L} W_h^2 S_{(ij)(fg)h}. \tag{A.20}
$$

A typical term of B is

$$
\operatorname{cov}\left(\hat{P}^{(k)}_{ij}, \hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\right) = \sum_{i',j'}\sum_{f',g'}\operatorname{cov}\left(\hat{P}^{(k)}_{ij}\tilde{P}^{(k)}_{i'j'}, \hat{P}^{(l)}_{fg'}\tilde{P}^{(l)}_{f'g}\right), \tag{A.21}
$$

and by summing over (A.18) we obtain

$$
\operatorname{cov}\left(\hat{P}^{(k)}_{ij}, \hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\right) = P_{\cdot g}\sum_{h=1}^{L} W_h^2 S_{(ij)(f\cdot)h}. \tag{A.22}
$$

Evaluating the remaining terms of B similarly, we obtain

$$
\begin{aligned}
B = 2\sum_{h=1}^{L} W_h^2\bigl[&P_{i\cdot}S_{(\cdot j)(fg)h}+P_{\cdot j}S_{(i\cdot)(fg)h}\\
&+P_{f\cdot}S_{(ij)(\cdot g)h}+P_{\cdot g}S_{(ij)(f\cdot)h}\bigr].
\end{aligned}\tag{A.23}
$$

A typical term of C is again obtained by summing over (A.18) as follows:

$$
\begin{aligned}
&\operatorname{cov}\left(\hat{P}^{(k)}_{i\cdot}\tilde{P}^{(k)}_{\cdot j}, \hat{P}^{(l)}_{f\cdot}\tilde{P}^{(l)}_{\cdot g}\right)\\
&\quad=\sum_{i',j'}\sum_{f',g'}\operatorname{cov}\left(\hat{P}^{(k)}_{ij'}\tilde{P}^{(k)}_{i'j}, \hat{P}^{(l)}_{fg'}\tilde{P}^{(l)}_{f'g}\right)\\
&\quad=\Bigl(\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\Bigr)\Bigl(\sum_{h\in M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)h}\Bigr)\\
&\quad+\Bigl(\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}\Bigr)\Bigl(\sum_{h\notin M_{k,l}} W_h^2 S_{(f\cdot)(\cdot j)h}\Bigr)\\
&\quad+P_{i\cdot}P_{f\cdot}\sum_{h\in M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)h}\\
&\quad+P_{\cdot j}P_{\cdot g}\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\\
&\quad+P_{i\cdot}P_{\cdot g}\sum_{h\notin M_{k,l}} W_h^2 S_{(f\cdot)(\cdot j)h}\\
&\qquad\quad+P_{f\cdot}P_{\cdot j}\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}.
\end{aligned}\tag{A.24}
$$

Similarly,

$$\operatorname{cov}(\tilde{P}_{i\cdot}^{(k)}\hat{P}_{\cdot j}^{(k)}, \hat{P}_{f\cdot}^{(l)}\tilde{P}_{\cdot g}^{(l)}) = \operatorname{cov}(\hat{P}_{i\cdot}^{(k)}\tilde{P}_{\cdot j}^{(k)}, \tilde{P}_{f\cdot}^{(l)}\hat{P}_{\cdot g}^{(l)})$$

$$= \left(\sum_{h\in M_{k,l}} W_h^2 S_{(f\cdot)(\cdot j)h}\right)\left(\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}\right)$$

$$+ \left(\sum_{h\notin M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)}\right)\left(\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\right)$$

$$+ P_{f\cdot}P_{\cdot j}\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}$$

$$+ P_{i\cdot}P_{\cdot g}\sum_{h\in M_{k,l}} W_h^2 S_{(f\cdot)(\cdot j)h}$$

$$+ P_{\cdot j}P_{\cdot g}\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}$$

$$+ P_{i\cdot}P_{f\cdot}\sum_{h\notin M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)}. \qquad \text{(A.25)}$$

Thus

$$C = 2\left[\left(\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\right)\left(\sum_{h\in M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)h}\right)\right.$$

$$+ \left(\sum_{h\in M_{k,l}} W_h^2 S_{(f\cdot)(\cdot j)h}\right)\left(\sum_{h\in M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}\right)$$

$$+ \left(\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(f\cdot)h}\right)\left(\sum_{h\notin M_{k,l}} W_h^2 S_{(\cdot j)(\cdot g)h}\right)$$

$$+ \left(\sum_{h\notin M_{k,l}} W_h^2 S_{(f\cdot)(\cdot j)h}\right)\left(\sum_{h\notin M_{k,l}} W_h^2 S_{(i\cdot)(\cdot g)h}\right)$$

$$+ P_{i\cdot}P_{f\cdot}\sum_{h=1}^{L} W_h^2 S_{(\cdot j)(\cdot g)h}$$

$$+ P_{\cdot j}P_{\cdot g}\sum_{h=1}^{L} W_h^2 S_{(i\cdot)(f\cdot)h}$$

$$+ P_{i\cdot}P_{\cdot g}\sum_{h=1}^{L} W_h^2 S_{(f\cdot)(\cdot j)h}$$

$$\left. + P_{f\cdot}P_{\cdot j}\sum_{h=1}^{L} W_h^2 S_{(i\cdot)(\cdot g)h}\right] \qquad \text{(A.26)}$$

Substituting (A.20), (A.23), and (A.26) in (A.19), we obtain (3.1).

———— ○○○ ————

☆ U. S. GOVERNMENT PRINTING OFFICE : 1973 515-212/45

VITAL AND HEALTH STATISTICS PUBLICATION SERIES

Originally Public Health Service Publication No. 1000

Series 1. Programs and collection procedures.—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions, data collection methods used, definitions, and other material necessary for understanding the data.

Series 2. Data evaluation and methods research.—Studies of new statistical methodology including: experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, contributions to statistical theory.

Series 3. Analytical studies.—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Series 4. Documents and committee reports.—Final reports of major committees concerned with vital and health statistics, and documents such as recommended model vital registration laws and revised birth and death certificates.

Series 10. Data from the Health Interview Survey.—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, based on data collected in a continuing national household interview survey.

Series 11. Data from the Health Examination Survey.—Data from direct examination, testing, and measurement of national samples of the civilian, noninstitutional population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics; and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

Series 12. Data from the Institutional Population Surveys.—Statistics relating to the health characteristics of persons in institutions, and their medical, nursing, and personal care received, based on national samples of establishments providing these services and samples of the residents or patients.

Series 13. Data from the Hospital Discharge Survey.—Statistics relating to discharged patients in short-stay hospitals, based on a sample of patient records in a national sample of hospitals.

Series 14. Data on health resources: manpower and facilities.—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

Series 20. Data on mortality.—Various statistics on mortality other than as included in regular annual or monthly reports—special analyses by cause of death, age, and other demographic variables, also geographic and time series analyses.

Series 21. Data on natality, marriage, and divorce.—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports—special analyses by demographic variables, also geographic and time series analyses, studies of fertility.

Series 22. Data from the National Natality and Mortality Surveys.—Statistics on characteristics of births and deaths not available from the vital records, based on sample surveys stemming from these records, including such topics as mortality by socioeconomic class, hospital experience in the last year of life, medical care during pregnancy, health insurance coverage, etc.

For a list of titles of reports published in these series, write to:

Office of Information
National Center for Health Statistics
Public Health Service, HSMHA
Rockville, Md. 20852