MODELING CONTINGENCY TABLES FROM COMPLEX SURVEYS

Peter B. Imrey
University of Illinois

Eugene Sobel
Mildred Francis
Independent Statistical Consultants

Over a period of decades, the ingenuity of sample survey design techniques has far surpassed the ability to specify and manipulate reasonable probability models for resultant data, and hence likelihood functions for parametric analysis. Controversies regarding the foundations of inference from survey data have inhibited the application of techniques from the mainstream of data analysis to statistics from complex surveys. Thus, sophisticated methods for point and interval estimation of simple functions such as means and ratios have been articulated for a variety of complex designs to which asymptotic normal distribution theory is felt to be applicable; practical devices, such as balanced repeated replication (BRR), have allowed survey analysts to bypass difficulties of variance derivations and computations. However, statistical inference based on variational or structural modeling (see Koch and Stokes (1980)) for survey data, e.g., multi-factor analysis of variance, has not been widely used by survey statisticians. Such inferences have appeared relatively frequently in the literatures of several disciplines which utilize surveys, but undertaken naively assuming probability distributions based upon simple random sampling (srs).

With the recent ascendancy of modeling techniques for the analysis of contingency tables (Grizzle, Starmer, and Koch (1969), Bishop, Fienberg, and Holland (1975), Gokhale and Kullback (1980)) and the attractiveness of their application to complex survey data, the various issues involved in modeling such data are receiving increasing attention within the domain of categorical data analysis. Thus, Koch, Freeman, and Freeman (1975), Freeman and Koch (1976), Freeman, Freeman, Brock, and Koch (1976), and Freeman, Freeman, and Brock (1977) essentially advocate use of weighted least-squares model-fitting and Wald (1943) statistics to model functions of counts from categorical data whenever, by randomization theory or superpopulation assumptions, these functions may be assumed Gaussian with consistently estimable covariance matrix. Their methods rely upon the practicality of full covariance estimation. Other authors have attempted to explore and adjust the usual srs-based test statistics to compensate for the effects of a complex survey design (e.g., Nathan (1975), Cowan and Binder (1978), Fellegi (1978), Rao and Scott (1979)), usually on the assumption that full design-based covariance estimation is impractical, but that variances or some limited set of design efficiencies (deffs) can be estimated and used in the adjustment process. Tomberlin (1979) has compared the distributions of a Wald statistic and Fellegi's (1978) statistic for the same hypothesis concerning the fit of a log-linear model. Fellegi's statistic utilizes less information about the survey design effects. Simplifying models for complex data collection processes have been proposed (e.g., Altham (1976), Cohen (1976), Brier (1978), Rao and Scott (1979), Tomberlin (1980)). Koch, Freeman, and Tolley (1975) have discussed a naive heuristic estimator for log-linear models, obtained by substituting data from a complex survey design into srs-based likelihood equations. Imrey, Francis, and Sobel (1979) have shown that certain efficiency results for estimating means apply equally well to estimating parameters in a modeling context.

The purpose of this note is to summarize certain aspects of our state of knowledge through remarks upon and straightforward extensions of previous work by several authors. The point of view will derive from weighted least-squares model fitting; the topics of 1) covariance structure, 2) estimation efficiency, and 3) null distributions of conventional test statistics will be discussed.

II. Weighted Least-Squares Model-Fitting

A population P has s subpopulations, $P_{\underset{\sim}{i}}$, corresponding to different values of the vector subscript $\underset{\sim}{i}$. Independent random samples are drawn from the $P_{\underset{\sim}{i}}$, and their elements classified into categories indexed by the vector subscript $\underset{\sim}{j}$, yielding an observed vector $\underset{\sim}{n}$ of subpopulation category counts $n_{\underset{\sim}{i}\underset{\sim}{j}}$ ordered lexicographically by subscript $\underset{\sim}{i}\underset{\sim}{j}$. The sample from $P_{\underset{\sim}{i}}$ may result from srs or a complex multi-stage sample design, possibly with stratification. The design may differ for different values of $\underset{\sim}{i}$. The proportion or probability vector characterizing the finite or infinite population $P_{\underset{\sim}{i}}$ is $\underset{\sim}{\pi}_{\underset{\sim}{i}}$ and the vector $\underset{\sim}{\pi}$, which strings-out the $\underset{\sim}{\pi}_{\underset{\sim}{i}}$ lexicographically by subscript, thus characterizes the joint subpopulation-response category structure of P. The vector $\underset{\sim}{n}$, in conjunction with knowledge of the sampling conditions C under which it was generated, is to be used to draw inferences about functions of $\underset{\sim}{\pi}$.

Letting $\underset{\sim}{n}_{\underset{\sim}{i}}$ be the segment of $\underset{\sim}{n}$ corresponding to $\underset{\sim}{\pi}_{\underset{\sim}{i}}$, define $p_{\underset{\sim}{i}\underset{\sim}{j}}$, $\underset{\sim}{p}_{\underset{\sim}{i}}$, and $\underset{\sim}{p}$ by $\underset{\sim}{p}_{\underset{\sim}{i}} = \underset{\sim}{n}_{\underset{\sim}{i}}/\underset{\sim}{1}'\underset{\sim}{n}_{\underset{\sim}{i}}$, where $\underset{\sim}{1}$ is a vector of units. Let $\underset{\sim}{V}^*_{\underset{\sim}{i}}$ be the true covariance matrix of $\underset{\sim}{p}_{\underset{\sim}{i}}$ and $\hat{\underset{\sim}{V}}^*_{\underset{\sim}{i}}$ be a consistent estimator of $\underset{\sim}{V}^*_{\underset{\sim}{i}}$ based on the data.

[To arrive at $\hat{V}_i^*$, more information than that contained in n_i is typically needed--for instance, counts within sampled clusters, prior to aggregation into n_i.] Let V_p^* and $\hat{V}_p^*$ be block-diagonal matrices formed from the V_i^* and $\hat{V}_i^*$, respectively.

When asymptotic conditions apply, weighted least-squares procedures may be used to fit linear models to functions of survey data. Let $F(\pi)$ be a u-vector of continuous, linearly independent functions with partial derivatives through second-order, and let $H(z) = [\frac{(dF_1)}{d\pi}, \frac{(dF_2)}{d\pi}, \ldots, \frac{(dF_u)}{d\pi}]'$, evaluated at z; write $H = H(\pi)$, $\hat{H} = H(p)$. When $F(p) = F$ is asymptotically normally distributed, its asymptotic covariance matrix is $V_F^* = (HV_p^*H')$, consistently estimable by $\hat{V}_F^* = (\hat{H}\hat{V}_p^*\hat{H}')$. Under a general linear hypothesis $F(\pi) = X\beta$, $\hat{\beta}^* = (X'\hat{V}_F^{*-1}X)^{-1} X'\hat{V}_F^{*-1}F$ is an asymptotically efficient (BAN) estimator of the parameter vector β, with asymptotic covariance matrix $V_\beta^* = (X'V_F^{*-1}X)^{-1}$, estimable by $\hat{V}_\beta^* = (X'\hat{V}_F^{*-1}X)^{-1}$. The residual quadratic form statistic $(F(p)-X\hat{\beta}^*)'\hat{V}_F^{*-1}(F(p)-X\hat{\beta}^*) = Q_{Fit}^*$ is a Wald (1943) statistic, asymptotically $\chi^2_{\text{rank } X}$ for the fit of X, while the form $Q_C^* = (C\hat{\beta}^*)' [C\hat{V}_\beta^* C']^{-1} (C\hat{\beta}^*)$ is also a Wald statistic, asymptotically $\chi^2_{\text{rank } C}$ for $H_o : C\beta = 0$ against the general alternative. This formulation for categorical data analysis includes linear and log-linear models for cell probabilities as well as models of complete, partial, and marginal symmetry and a variety of other approaches of interest (Grizzle, Starmer, and Koch (1969), Koch and Reinfurt (1971), Grizzle and Williams (1972), Forthofer and Koch (1973)).

III. Covariance Matrix Estimation

Implementation of the above approach requires a consistent estimator $\hat{V}_i^*$ from each subsample. However, the estimation of covariances from survey data has not been common practice, and in most cases is claimed to added a substantial increment to the overall computational burden of survey analysis. As a result, there has been a tendency to disregard the possibility of direct plug-in estimation of covariance matrices V_i^* in favor of either: i) general procedures such as BRR (Kish and Frankel (1974)) for estimation of V_F^* without formulation and estimation of the V_i^*; ii) sampling models which simplify the structure and estimation of the V_i^*; or iii) adjustments of conventional srs-based methods to approximate the results of covariance-based methods.

While all these approaches are valid and desirable, they do not fully substitute for explicit formulation of a design-based covariance structure and subsequent covariance estimation independent of restrictive assumptions about the nature of clustering. True covariance structure can be used for sensitivity analyses (with respect to sample design and nature of clustering) of linear model parametric inference, using both analytic and simulation methods. Inferences without restrictive assumptions about the nature of clustering are preferable in principle to conservative approximations based on models, unless the model is backed by a very strong rationale. The difficulty of covariance matrix specification, and magnitude of associated computational burden, are easily overestimated in an era of rapidly expanding computing capabilities vis a vis storage and array manipulation.

With regard to specification, every textbook of sampling theory contains formulae for variances of estimators from standard complex survey designs. These are easily extended to cover the categorized data setting, essentially multivariate, elaborated in Section II. For instance, Cochran (1977) discusses the cluster-size weighted mean per element $\bar{\bar{Y}}_R$ from a two-stage sample. A sample of n from N primary units is drawn by srs at the first stage; when cluster ℓ is drawn, m_ℓ elements are subsampled by srs. $\bar{\bar{Y}}$ and $\bar{M}$ are the true mean per element and average cluster size, while $\bar{Y}_\ell$ and $S^2_{2\ell}$ are the mean and variance of cluster ℓ. The first and second stage sampling fractions are denoted by $f_1 = n/N$ and $f_{2\ell} = m_\ell/M_\ell$. The asymptotic variance of $\bar{\bar{Y}}_R$ is

$$V(\bar{\bar{Y}}_R) = \frac{(1-f_1)}{n\bar{M}^2} \sum_\ell M_\ell^2(\bar{Y}_\ell - \bar{\bar{Y}})^2/(N-1) \qquad \text{(III. 1)}$$

$$+ (f_1/(n\bar{M})^2) \sum_\ell \frac{(1-f_{2\ell})}{m_\ell} M_\ell^2 S^2_{2\ell}.$$

In the categorical framework, let π_ℓ be the vector of relative proportions corresponding to cluster ℓ, and π be that for the population (for simplicity, we consider a single population and drop the subscript i for now). The analogue of $\bar{\bar{Y}}_R$ is $\tilde{p} = (\sum_{k=1}^{n}(M_k/m_k)\, n_k)/(\sum_{k=1}^{n} M_k)$, where n_k is the observed distribution into categories of the subsample from the kth sampled primary unit.

Then (III.1) is easily extended to

$$\tilde{V}_{\tilde{p}}^{*} = \frac{(1-f_1)}{n\bar{M}^2} \sum_{\ell} M_{\ell}^2 (\tilde{\pi}_{\ell}-\tilde{\pi})(\tilde{\pi}_{\ell}-\tilde{\pi})'/(N-1)$$

$$+ (f_1/(n\bar{M})^2) \sum_{\ell} \frac{(1-f_{2\ell})M_{\ell}^3}{m_{\ell}(M_{\ell}-1)} (D_{\tilde{\pi}_{\ell}} - \tilde{\pi}_{\ell}\tilde{\pi}_{\ell}') \qquad (III.2)$$

where $D_{\tilde{\pi}}$ is the diagonal matrix with $\tilde{\pi}_{\ell}$ on the diagonal. $V_{\tilde{p}}^{*}$ may be estimated (consistently)

by $$\hat{V}_{\tilde{p}}^{*} = \frac{(1-f_1)}{n\bar{M}^2} \sum_{k} M_k^2 (\tilde{p}_k-\tilde{p})(\tilde{p}_k-\tilde{p})'/(n-1) \qquad (III.3)$$

$$+ (f_1/(n\bar{M})^2) \sum_{k} \frac{(1-f_{2k})M_k^3}{m_k(M_k-1)} (D_{\tilde{p}_k} - \tilde{p}_k\tilde{p}_k')$$

where $\tilde{p}_k = \tilde{n}_k/\tilde{1}'\tilde{n}_k$. If $\bar{M}$ is unknown, it may be replaced by $\bar{M}^* = \sum_k M_k/n$.

Calculation of $\hat{V}_{\tilde{p}}^{*}$ does not entail great storage or computational demands, particularly if the data are sorted by cluster or if an indicator function approach such as that taken in GENCAT (Landis, Stanish, and Koch (1976)) or MISCAT (Stanish, Koch, and Landis (1978)), for example, is applied. Various conditions on cluster size and sampling fractions lead to simplifications, further easing the computational problem. When clusters are large and of equal size,

$$\hat{V}_{\tilde{p}}^{*} = \frac{(1-f_1)}{n(n-1)} \sum_{k} (\tilde{p}_k-\tilde{p})(\tilde{p}_k-\tilde{p})' \qquad (III.\ 4)$$

$$+ (f_1/n^2) \sum_{k} \frac{(1-f_{2k})}{m_k} (D_{\tilde{p}_k} - \tilde{p}_k\tilde{p}_k')$$

is consistent; when the sampling fraction within clusters is held constant,

$$\hat{V}_{\tilde{p}}^{*} = \frac{(1-f_1)}{n\bar{M}^2} \sum_{k} M_k^2 (\tilde{p}_k-\tilde{p})(\tilde{p}_k-\tilde{p})'/(n-1) \qquad (III.\ 5)$$

$$+ (f_1(f_2^{-1}-1)((n\bar{M})^2) \sum_{k} M_k^2 (D_{\tilde{p}_k} - \tilde{p}_k\tilde{p}_k')/(M_k-1).$$

When both sampling fractions are small,

$$\hat{V}_{\tilde{p}}^{*} = (1/n(n-1)\bar{M}^2) \sum M_k^2 (\tilde{p}_k-\tilde{p})(\tilde{p}_k-\tilde{p})'$$

$$+ (1/nN\bar{M}^2) \sum_{k} M_k^2 (D_{\tilde{p}_k} - \tilde{p}_k\tilde{p}_k')/m_k \qquad (III.\ 6)$$

and, when all these conditions apply,

$$\hat{V}_{\tilde{p}}^{*} = (1/n(n-1)) \sum_{k} (\tilde{p}_k-\tilde{p})(\tilde{p}_k-\tilde{p})' \qquad (III.\ 7)$$

$$+ (1/mnN) \sum_{k} (D_{\tilde{p}_k} - \tilde{p}_k\tilde{p}_k').$$

Analogous expressions for more complex designs may be developed in a similar straightfoward fashion. Sampling models for the nature of clustering, such as Altham (1976), lead to simpler formulas as in Rao and Scott (1979). However, if one assumes that covariance estimation will generally become accepted as necessary in survey analysis, and that continued expansion of computing capabilities will occur, use of the design-based formulae such as (III. 3) - (III. 7) does not seem at all out of the question. Here, distinctions must be made between very large scale federal surveys with highly complex sampling strategies, on the one hand, and the many surveys performed by state and local governments, contract research groups and universities which incorporate elements of cluster sampling into the design without a highly elaborate architecture. For such relatively simple surveys, direct formula-based estimation may be at least as attractive as BRR, say, inasmuch as necessary software should not be hard to develop for tests and intervals of common interest. For the most complex designs, however, the various arguments that have been made in favor of some type of split-sample covariance estimation have greater validity.

IV. Estimation Efficiency

The efficiencies of two survey designs may be considered in the context of an intended linear model analysis. Let $\tilde{X}$ be a specified linear model for $\tilde{F}$, $V_{\tilde{p}}^{*}$ and $V_{\tilde{p}}^{**}$ be the covariance matrices under two competing survey designs, $\tilde{S}^* = (\tilde{H}\tilde{V}_{\tilde{p}}^{*}\tilde{H}')$, $\tilde{S}^{**} = (\tilde{H}\tilde{V}_{\tilde{p}}^{**}\tilde{H}')$, $V_{\tilde{\beta}}^{*} = (\tilde{X}'\tilde{S}^{*-1}\tilde{X})^{-1}$, $V_{\tilde{\beta}}^{**} = (\tilde{X}'\tilde{S}^{**-1}\tilde{X})^{-1}$. Then the design effect of plan 2 (**) relative to plan 1 (*) pertaining to a specified parametric function $\tilde{c}'\tilde{\beta}$ may be written as

$$\text{deff}\ (\tilde{c}'\tilde{\beta}) = ((\Sigma a_v^2\lambda_v/\Sigma a_v^2)+1) \qquad (IV.\ 1)$$

where the λ_v are eigenvalues of $\tilde{U} = (V_{\tilde{\beta}}^{**} - V_{\tilde{\beta}}^{*})$ and $\tilde{c} = \Sigma a_v\tilde{\lambda}_v$ for corresponding eigenvectors $\tilde{\lambda}_v$. Various summary measures of design effect may be useful in particular situations within the linear model context, viz:

i) $(1 + \lambda_1)$ = maximum effect on parameter;

ii) $(1 + \lambda_v)$ = effect on vth parameter, e.g. on a main effect of a log-linear model;

iii) $1 + (\text{tr}\ \tilde{U}/\text{rank}\ \tilde{X})$ = arithmetic mean effect on parameters;

iv) $[\det V_{\tilde{\beta}}^{**}/\det V_{\tilde{\beta}}^{*}]^{1/\text{rank}\ \tilde{X}}$ = geometric mean effect on parameters;

v) $u^{-1}(\text{tr}\ \tilde{X}\tilde{U}\tilde{X}')$ = arithmetic mean effect on predicted values $\hat{\tilde{F}} = \tilde{X}\hat{\tilde{\beta}}$ for $\tilde{F}$; etc.

Rao and Scott (1979) have discussed this issue in connection with simple goodness-of-fit tests and tests of independence within two-way contingency tables. They use measures based on cell probability design effects, or on design effects applicable to $\tilde{F}$. Measures suggested above may, theoretically, be quite different than measures based upon $\tilde{p}$ or $\tilde{F}$ in the absence of an underlying model; differences arise due to the interrelationship between the sample design, population clustering structure, nature of $\tilde{F}$

and X. In certain circumstances, such relation- might be exploited in sample design.

General theories about global efficiency of certain designs may be extended from standard univariate results to the linear model context. When V_p^* corresponds to proportionate stratified random sampling and $V_i^{**}=(D_\pi-\pi_i\pi_i')/n$ (for V_p corresponding to srs), Imrey, Francis, and Sobel (1979) have shown that U is positive definite, so that proportionate stratification is efficient for estimation of any linear model by any criterion based on first-order efficiency. Their model applies to any comparison in which $(V_p^{**}-V_p^*)$ (strong condition) or $(V_F^{**}-V_F^*)$ (weak condition) is positive definite. Hence, in a cluster random sampling design with N clusters each of size M and one population, let

$$R=(1/NM(M-1))2\Sigma\Sigma(y_{\ell z}-\pi)(y_{\ell z}-\pi)'[D_\pi-\pi\pi']^{-1} \quad \text{(IV.2)}$$

where $y_{\ell z}$ is an indicator vector corresponding to the zth element of the ℓth cluster. The "intracluster correlation matrix" R is an analogue of the intracluster correlation coefficient ρ, and

$$V_p^* \doteq \frac{(1-f)}{M}[I+(M-1)R]V_p \quad \text{(IV. 3)}$$

with f the sampling fraction. Generalizing the well-known results based on ρ, we may say that cluster random sampling is generally inefficient when R is positive-definite and efficient if R is negative definite. Similar results apply to other sampling structures based upon models of clustering, e.g., Altham's (1976) model of two-stage sampling and its extension to three-stage sampling, both discussed from a similar point of view by Rao and Scott (1979).

In these situations, the strong condition on cell estimators holds. As an example of a circumstance where it would not, but where the weaker condition relating to F might apply, consider a panel study of voter preferences at various times in a long Presidential campaign, and let F be the vector of marginal candidate preference proportions. It is quite conceivable that design effects on counts corresponding to particular patterns of response over time might be lower than one, with yet $(V_F - V_F^*)$ negative definite, where V_F corresponds to srs, in which case any model parameters corresponding to F would be estimated with a loss in efficiency by the * design.

V. Behavior of SRS-Based Tests

Rao and Scott (1979) examine the asympotic distributions of conventional chi-square goodness-of-fit statistics for a single categorization, and for independence in two-way contingency tables. They suggest corrections for these statistics which account, in some circumstances, for design effects. The basic result is readily extended to the general linear model context. V_p^* is the true covariance matrix of p, V_p that based on srs. Without specifically accounting for it notationally, we include the multipopulation case. Thus, V_p and V_p^* may be block diagonal. Let $V_F = (HV_pH')$ (the asymptotic covariance matrix of $F(p)$ under srs), $V_\beta = (X'V_F^{-1}X)^{-1}$ (the asymptotic covariance matrix of $\hat{\beta}$ under srs), with $V_F^* = (HV_p^*H')$ (the true asymptotic covarance matrix of $F(p)$ under the actual survey design). $\hat{V}$ indicates a consistent estimate of the corresponding V, ordinarily determined by $\hat{V} = V(p)$ after writing $V = V(\pi)$. The following hold whether the $\hat{V}$ are unrestricted consistent estimates (e.g., unrestricted maximum likelihood estimates in which case minimum modified (Neyman) chi-square estimates, e.g., Wald statistics are involved) or consistent under hypothesis (e.g., restricted m.l.e.'s under the model, in which case Pearson chi-square statistics and minimum chi-square estimation are involved).

Lack-of-Fit Statistics: Under the model $F(\pi) = X\beta$, the lack-of-fit statistic $Q = (F-X\hat{\beta})'\hat{V}_F^{-1}(F-X\hat{\beta})$, where $\hat{\beta} = \hat{V}_\beta X'\hat{V}_F^{-1}F$, is asymptotically distributed as $\Sigma\lambda_i Z_i^2$, where the Z_i are independent N(0,1), and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{u-v}$ are the eigenvalues of $V_F^{-1}(M_X V_F^* M_X')$, with $M_X = [I-XV_\beta X'V_F^{-1}]$.

Tests of Linear Hypothesis = Under the model $F(\pi) = X\beta$, and $H_o: C\beta = 0$, the statistic $Q_C = (C\hat{\beta})'[C\hat{V}_\beta C']^{-1}(C\hat{\beta})$, is asymptotically distributed as $\Sigma\gamma_i Z_i^2$, where the Z_i are independent N(0,1) and $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_{\text{rank } C}$ are the eigenvalues of $V_{C\beta}^{-1}(M_C V_F^* M_C')$ where $V_{C\beta} = CV_\beta C'$ and $M_C = [CV_\beta X'V_F^{-1}]$.

Koch, Freeman, and Tolley (1975) have considered a similar situation, solution of srs-based likelihood equations from log-linear models, using data from complex sample surveys, and obtained a related result.

Since the asymptotic distributions of standard test statistics depend on V_p and V_p^* only through V_F and V_F^*, adjustments to these statistics may be made based upon knowledge of the design effects pertaining to F. Following Rao and Scott, note that when $V_F^* = \lambda V_F$, then $\lambda_1 = \lambda_2 = \ldots = \lambda$, and Q/λ and Q_C/λ have exact chi-square distributions. This occurs when $\tilde{V}_p^* = \lambda \tilde{V}_p$ and may occur otherwise. When such a relationship is not known, Rao and Scott suggest use of a function of the design efficiencies obtainable from the diagonals of $\tilde{V}_p^*$ and $\tilde{V}_p$ or $\tilde{V}_F^*$ and $\tilde{V}_F$ as divisor. For instance, estimates of λ_1 and γ_1, used as divisors for Q and Q_C produce conservative tests. Similarly, estimates of arithmetic mean eigenvalues from tr $\tilde{M}$, tr $\tilde{M}_C$ essentially standardize the mixture of chi-squares.

Extension of the above to characterize behavior of tests derived using any erroneous sampling model (not just srs) is obvious.

VI. Summary

Analysis of categorical data from complex survey designs has been considered from the veiwpoint of general linear model fitting and associated chi-square statistics. Many standard results on survey analysis generalize quite readily to this setting. In particular, in Section III covariance matrix expressions for complex surveys involving categorical data are easily generated from existing results for univariate continuous variables. In Section IV, elaborating on Imrey, Francis, and Sobel (1979), it was shown that basic efficiency theorems may also be generalized. Section V, elaborating on Rao and Scott (1979), clarifies the distributions of linear model test statistics obtained utilizing the wrong sampling model. These results should aid in comparing behavior of procedures which assume srs, now used commonly by social scientists and those advocated recently (e.g., by Koch, Freeman, and Freeman (1975)) which utilize estimated covariances based upon complex survey design structure.

REFERENCES

Altham, P.A.E. (1976), "Discrete variable analysis for individuals grouped into families," Biometrika 63, 263-269.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1974), Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA: MIT Press.

Brier, S. S. (1978), Categorical Data Models for Complex Data Structures. Ph.D. Dissertation, University of Minnesota.

Cochran, W. G. (1977), Sampling Techniques (3rd ed.). New York: John Wiley and Sons.

Cohen, J. E. (1976), "The distribution of the chi-squared statistic under clustered sampling from contingency tables," J. Amer. Statist. Assoc. 71, 665-670.

Cowan, J. and Binder, D. (1978), "The effect of a two-stage sample design on tests of independence of a 2 x 2 table," Survey Methodology 4.

Fellegi, I. (1978), "Approximate tests of independence and goodness of fit based on stratified and multi-staged samples," Survey Methodology 4, 29-56.

Forthofer, R. N. and Koch, G. G. (1973), "An analysis for compounded functions of categorical data," Biometrics 29, 143-157.

Freeman, D. J., Freeman, J. L., and Brock, D. W. (1977), "Modularisation for the analysis of complex sample survey data," presented to International Association of Survey Statisticians at meeting of the International Statistics Institute, New Delhi, India.

Freeman, D. J., Freeman, J. L., Brock, D. W., and Koch, G. G. (1976), "Strategies in the multivariate analysis of data from complex surveys II," Int. Statist. Rev. 44, 317-330.

Freeman, D. J. and Koch, G. G. (1976), "An asymptotic covariance structure for testing hypotheses on raked contingency tables from complex surveys," 1976 Proceedings of the Social Statistics Section of the American Statistical Association,

Gokhale, D. V. and Kullback, S. (1978), The Information in Contingency Tables, Marcel Dekker, Inc.: New York.

Grizzle, J. E., Starmer, C. F., and Koch, G. G. (1969), "Analysis of categorical data by linear models," Biometrics 25, 489-504.

Grizzle, J. E. and Williams, O. D. (1972), "Log-linear models and tests of independence for contingency tables," Biometrics 28, 137-156.

Imrey, P. B., Francis, M. E., and Sobel, E. (1979), "Analysis of categorical data obtained by stratified random sampling I," Commun. Statist. A8(7), 653-70.

Kish, L. and Frankel, M. (1974), "Inferences from complex samples," J. R. Statist. Soc. B36, 1-37.

Koch, G. G., Freeman, D. J., and Freeman, J. L. (1975), "Strategies in the multivariate analysis of data from complex surveys," Int. Statist. Rev. 43, 59-78.

Koch, G. G., Freeman, D. J., and Tolley, H. D. (1975), "The asymptotic covariance structure of estimated parameters from contingency table log-linear models," University of North Carolina Institute of Statistics Mimeo Series No. 1046.

Koch, G. G. and Reinfurt, D. W. (1971), "The analysis of categorical data from mixed models," Biometrics 27, 157-173.

Koch, G. G., Gillings, D. B., and Stokes, M. E. (1980), "Biostatistical implications of design, sampling and measurement to the analysis of health science data," Annual Review of Public Health 1, 163-225.

Landis, J. R., Stanish, W. R., Freeman, J. L., and Koch, G. G. (1976), "A computer program for the generalized chi-square analysis of categorical data using weighted least squares (GENCAT)," Computer Programs in Biomedicine 6, 196-231.

Nathan, G. (1975), "Tests of independence in contingency tables from stratified proportional samples," Sankhya C37, 77-87.

Rao, J.N.K. and Scott, A. J. (1979), "Chi-squared tests for analysis of categorical data from complex surveys," 1979 Proceedings of the Section on Survey Research Methods of the American Statistical Association, 58-66.

Stanish, W. M., Koch, G. G., and Landis, J. R. (1978), "A computer program for multivariate ratio analysis (MISCAT)," Computer Programs in Biomedicine, 8, 197-207.

Tomberlin, T. J. (1979), "The analysis of contingency tables of data from complex samples," 1979 Proceedings of the Section on Survey Research Methods of the American Statistical Association, 152-157.
Tomberlin, T. J. (1980), "A model-based approach to the analysis of contingency tables of data from complex samples," 1980 Proceedings of the Section on Survey Research Methods of the American Statistical Association.
Wald, A. (1943), "Tests of statistical hypotheses concerning general parameters when the number of observations is large," Trans. Amer. Math. Soc. 54, 426-482.