VARIANCE IN THE NATIONAL HEALTH INTERVIEW SURVEY DATA

Jai W. Choi and Robert J. Casady
National Center for Health Statistics

1. INTRODUCTION

This paper discusses how to present the sample variance for the data from the National Health Interview Survey (NHIS) and includes some details of the polynomial used to represent the NHIS sample variance points estimated by balanced half samples (BHS) method.

The current method has been used since the inception of the NHIS in 1958. The Current Population Survey (Edelman, 1967) and National Survey of Family Growth have also been using the same technique to present the sample variance points derived by Keyfitz procedures and BHS, respectively.

The NHIS procedure of variance presentation has never been organized in an understandable manner. Perhaps this is partly because there are no satisfactory answers to some of the assumptions made for the adoption of the method. The main purpose of this paper is to clarify crucial steps in presenting the variances of the NHIS data. No attempt is made to address the validity of the current methods.

Section 2 discusses how the NHIS variables with similar characteristics are grouped. Section 3 includes the estimation of sample variances for the items in each group by BHS method. Section 4 describes how to fit a curve to the variance points estimated in section 3. The estimation of the coefficients for the curve is presented in section 5. Finally, section 6 includes the comments on this method and other potential alternatives to the BHS method.

2. GROUPING OF NHIS VARIABLES

The variables in the NHIS publications are based on the complex sample survey and weighted four times to estimate the U. S. civilian noninstitutionalized population. The sample design and weighting procedures are not discussed in this paper.

When a sample is taken by cluster sampling and persons in the cluster are often correlated, it would be desirable from the previous experience to isolate certain classes of variables which have the similar score of intracluster correlation.

Rather, NHIS variables are classified according to four main characteristics, e.g., the range of variables, recall period, the length of data collection period, and the type of variables. These are briefly described below:

The three ranges of variables

Narrow range **(N)** -the statistics which estimate population attributes, for instance, the number of persons in a particular income group, and statistics for which the measure for a single individual for a specified period of reference period is usually either 0 or 1, on occasion may take on the value 2, and rarely 3.

Medium Range **(M)** -the statistics for which the measure for a single individual for a period of reference will rarely lie outside the range 0 to 5.

Wide range **(W)** -the statistics for which the measure for a single individual for a period of reference will range from 0 to a number in excess of 5, for instance, the number of days of bed disability experienced during the year.

Four reference periods

Type **A**- statistics on prevalence or incidence for which the period of reference in the questionnaire is 12 months.

Type **B**- incidence type statistics for which the period of reference in the question is 2 weeks

Type **C**- statistics for which the reference period in the questionnaire is 6 months.

Type **D**- statistics for which the refernce period in the questionnaire is 3 months.

Two lengths of data collection periods

The period of data collection is usually 1 or 4 quarters. These are represented with the number **1** or **4.** The data from 8 or 12 quarters are often combined when the size of the sample from four quarters is too small to have meaningful data analysis.

Two classes of variables

A set of curves for relative standard errors (RSE) of aggregates **(A)** and percentages **(P)** are commonly presented in the NHIS publication. RSE's of other statistics, e.g., ratios, and difference between two means, ratios, or percents, can be approximated from two or more of such curves usually assuming that the covariance terms can be ignored.

Thus, the data from the NHIS could be classified into at least one of 48 possible types of variables according to 2 types of estimates, 2 different durations of data collection period, 4 recall periods, and 3 ranges of variables.

Table A illustrates the symbols that are most commonly used. For instance, **A4BN** is used to represent the aggregate of 4 quarters data **(A4)**, based on a two week recall **(B)** and of the narrow range **(N)**. Similarly we can write symbols of percentages. For example, **P1AW** means the percentage from one quarter data **(P1)**, based on a 12 month recall **(A)**, and of the wide range **(W)**. Occasionally, one symbol represents more than one group of variables when these variables vary widely, e.g., **A4BW** is used to represent two different curves. One is for restricted activity and bed disability days, and the other is for work or school loss days as shown in Table B.

3. VARIANCE ESTIMATION BY BALANCED HALF SAMPLES

Two steps are applied to NHIS data in order to obtain a set of variance curves. First step is the derivation of relative variance by the balanced half sample (BHS) procedure. The second step is to fit a curve to a set of points representing such relative variances. This section discusses the first step of BHS method. The second step will be presented in the next section.

Table A
Classification of NHIS Variables for aggregates*

Recall period	Range of variable N	M	W
A (12 months)	A4AN	A4AM	A4AW
B (2 weeks)	A4BN	A4BM	A4BW
C (6 months)	A4CN	A4CM	A4CW
D (3 months)	A4DN	A4DM	A4DW

N: Narrow range (0-1); M: Medium range (0-5); W: Wide range (0-5 or over); * Similarly for percentage and one quarter data by replacing the first letter and second number with **P** and **1.**

All the NHIS sample PSU's are classified into one of the two types, self representing PSU's (SR PSU) and nonself representing PSU's (NSR PSU).

Two or more SR PSU's are regrouped into a new SR PSU, and then each new SR PSU is divided into two pseudo PSU's by systematically assigning odd numbered segments (subunits of PSU) into one pseudo-PSU and even numbered segments into the other pseudo-PSU. These two pseudo PSU's become a pair in a newly formed stratum.

The NSR PSU's are paired so that the two are similar in size and characteristics.

The main reason for pairing is that the pair thus formed reflects the stratum and the actual sampling procedures so that one can indirectly estimate an actaul sampling error through BHS procedures (McCarthy, 1966). The pairing of the pseudo-PSU's is accomplished independently for each of the four geographic regions (North-East, North-Central, South, and West). 149 pseudo strata each including two PSU's are summarized in another document (Schnack, 1974).

149 pairs give 152 replicates based on the 149 x 152 orthogonal matrix (Plackettand Burman, 1946). For each replicate, the original weights of individuals are newly ratio adjusted by the 60 cell age-race-sex table of the Current Population Survey (Jones, 1976), so that the weight of an individual could be changed over replication. This might give better estimates for the population and reduce sample variance.

Denote a sample estimate of X by x_i from the ith balanced half sample, i=1,...,152
Population X is estimated by

$$\bar{x} = \frac{1}{152} \sum_{i}^{152} x_i \qquad (3.1)$$

and the sample variance of $\bar{x}$ is given by

$$\text{var}(\bar{x}) = \frac{1}{152} \sum_{i}^{152} (x_i - \bar{x})^2 \qquad (3.2)$$

The size of the var($\bar{x}$) given in (3.2) depends on several factors in a practical situation. The most important factor is the difference between the two elements or two pseudo PSU's in a stratum. var($\bar{x}$) is directly related to this difference especially for a linear estimate.

When one stratum included many elements of different sizes and only two elements are randomly taken, it may happen that the two elements do not represent the stratum well and hence might distort the magnitute of the stratum variance. On the other hand if the stratum included only a few elements of similar size and character, the two elements so selected would correctly reflect the variance of that stratum.

4 CURVE FITTING

The basic assumption made on the curve fitting is that the sample variance is a function of estimates and that this is the only factor producing the variance and not the differences in the estimates.

Two main reasons for the use of this method are that fitting a curve to the variances provides the stability to these estimates and that there is not enough time nor money to compute variances for all items. In fact, if that could be done, it would be too messy and unstable to produce variances for users of data.

Let $x_1, x_2, \ldots, x_k$ be a set of k sample estimates for k items from $n_1, n_2, \ldots, n_k$ sample counts. The estimates are left unspecified.

The NHIS variances of these estimates are derived by the BHS method.

Let $V(x_1), V(x_2), \ldots, V(x_k)$ be the relative sample variances of $x_1, x_2, \ldots, x_k$. The relative sample variance is defined by

$$V(x_i) = \frac{\sigma^2(x_i)}{E^2(x_i)} \quad i=1,\ldots,k \qquad (4.1)$$

Now one may try to draw a curve to represent these k relative variances. Often this is impossible when these k points are widely scattered around and hence no one curve may represent them all. If some evidences show the existence of fitting curve, we could find such polynomial.

V may be considered as a decreasing function of the x_i, e.g., for i=1,...,k,

$$V(x_i) = A + B/x_i \qquad B > 0 \qquad (4.2)$$

where A and B are regression coefficients.

Edelman(1967) gives various other alternative models and R^2 values for them. But only the model (4.2) enjoys some justifications for such use as illustrated in the two examples below.

Suppose that a binomial sample of "a" clusters is randomly taken from "A" clusters, each containing M elementary units and denote the average per element by p, e.g.,

$$p = \frac{x}{n}, \text{ Let } E(p) = P = \frac{X}{N} \quad \text{and}$$

$$\sigma^2(p) = \frac{(1-f)PQ}{N\ N}\,(1 + (M-1)\rho) \qquad (4.3)$$

(Cochran p.242), where Q = 1-P ,f = a/A and ρ is the intracluster correlation coefficient. The relative variance of p is, ignoring f,

$$V(p) = \frac{-(1+(M-1)\rho)}{N^2} + \frac{1}{X}\ \frac{(1+(M-1)\rho)}{N}$$

which is the form given in (4.2).

Suppose that "a" first stage clusters are randomly taken from "A" clusters, a sample of "b" second stage clusters is also randomly taken from each of "a" first stage clusters and that each of "b" second stage clusters included "m" elementay units. Suppose that the sample is taken with replacement. Denote the total number of elementary units by n , e.g., n = abm, and the average number by p = x/n. we can write (Choi, 1981)

$$\sigma^2(p) = \frac{PQ}{N}\left(1 + \frac{G}{N}\rho_2 + \frac{(G-H)}{N}\rho_1\right) \quad (4.5)$$

where G=abm(m -1), H= bm(bm - 1), ρ_1 and ρ_2 are the common intracluster correlation coefficients in the first and second stage clusters, respectively.

We can also express it in the form of A+B/x, e.g., the relative variance of p is

$$V(p) = -\frac{g}{N} + \frac{g}{X} \quad (4.6)$$

where $g = \left(1 + \frac{G}{n}\rho_2 + \frac{(G-H)}{n}\rho_1\right)$

which also takes the same form of (4.2).

This kind of arguments can be written in terms of design effects.

Define, using the equation (4.5), the design effects by $\text{Deff} = \frac{\sigma^2(p)}{PQ/N} = g \quad (4.7)$

(4.6) can be written as -Deff /N + Deff /X that also takes the same form of (4.2) if Deff is independent of X. The function (4.2) is the only form among the equations that Edelman (1967) presented, which enjoys this type of justification.

Consider the ith estimate p = x/y for the population P = X/Y where X is a subset of Y. Tomlin (1974, p5) shows that the relative variance of p can be written as

$$V(p) = V(X) - V(Y) \quad (4.8)$$

The variance of p can then, from the definition (4.2), be written by

$$\sigma^2(p) = \frac{B\ P(1-P)}{Y} \quad (4.9)$$

The form of above equation (4.9) is the same as the usual form multiplied by B. Equation (4.2) is the only form that produces such a result.

5 ESTIMATION OF THE PARAMETERS A AND B.

There are many different ways to estimate regression coefficients depending on the various assumptions made on the regression model. In this section, the weighted least square method is used in order to reflect the heteroscedasticity of error terms in the model.

Denote the relative variance by $V(x_i)$ of an estimate x_i, i=1,...,k. First they are plotted to observe whether they can be represented by a curve and a few far outlying points, if any, are deleted, which are often caused by some erronous processing.

Let $(x_i, V(x_i))$ for i=1,...,k be a set of pairs where x_i could be aggregates, percents, or other estimates. The x_i's can be considered as observable deterministic variables or random variables.

$V(x_i)$ might vary with the unobservable random error e_i which might include the errors of measurement and other unexplainable variables in (5.1) below. One may assume that the errors are uncorrelated but heteroscedastic.

Under this situation, we may write

$$V(x_i) = f(x_i) + e_i \quad i=1,\dots,k \quad (5.1)$$

where $f(x_i) = A + B/x_i$ given in (4.2).

Ordinary Least Square (OLS) method is to minimize,

$$L_1 = \sum_i^k (V(x_i) - f(x_i))^2 \quad (5.2)$$

Let $\hat{V}_i = \hat{V}(x_i)$ be the estimates of $V(x_i)$ from BHS method. To a certain extent, it may be possible that the heteroscedasticity of the model can be reflected by weighting L_1 inversly to $\hat{V}_i$, e.g.,

$$L_2 = \sum_i^k \left(\frac{\hat{V}_i - f(x_i)}{\hat{V}_i}\right)^2 \quad (5.3)$$

L_2 is the sum of squares of the observed relative variance minus predicted relative variance divided by observed relative variance.

The estimation can also be improved by iteration. The minimization of L_2 will give the initial estimates $\hat{A}_0$ and $\hat{B}_0$, that will be used as initial values to the following iteration.

Let L_3 be the system to be repeatedly used to estimate A_j and B_j. The subscript j is for the jth time estimation. Start with $\hat{A}_0$ and $\hat{B}_0$ in $\hat{f}_0(x_i)$ to find the next estimates of A_1 and B_1 in $f_1(x_i)$ by minimizing L_3,

$$L_3 = \sum_i^k \left(\frac{\hat{V}_i - f_j(x_i)}{\hat{f}_{j-1}(x_i)}\right)^2 \quad (5.4)$$

where $\hat{f}_{j-1}(x_i) = \hat{A}_{j-1} + \frac{\hat{B}_{j-1}}{x_i}$, $f_j(x_i) = A_j + \frac{B_j}{x_i}$.

Only difference between (5.3) and (5.4) is that $\hat{V}_i$ in the denominator of (5.3) is replaced by $\hat{f}_{j-1}(x_i)$.

For the initial values A_0 and B_0, we solve

$$\frac{\partial L_2}{\partial A} = \frac{\partial L_2}{\partial B} = 0 \qquad (5.5)$$

which gives $\hat{A}_0$ and $\hat{B}_0$. Similarly we may find $\hat{A}_j$ and $\hat{B}_j$ from

$$\frac{\partial L_3}{\partial A_j} = \frac{\partial L_3}{\partial B_j} = 0 \qquad (5.6)$$

$$\hat{A}_j = \left(\sum_i^k \frac{1}{\hat{f}^2_{j-1}}\right)^{-1} \left(\sum_i^k \frac{\hat{V}_i}{\hat{f}^2_{j-1}} - \hat{B}_j \sum_i^k \frac{1}{x_i \hat{f}^2_{j-1}}\right) \qquad (5.7)$$

$$\hat{B}_j = \frac{\left(\sum_i^k \frac{\hat{V}_i}{x_i \hat{f}^2_{j-1}}\right)\left(\sum_i^k \frac{1}{\hat{f}^2_{j-1}}\right) - \left(\sum_i^k \frac{\hat{V}_i}{\hat{f}^2_{j-1}}\right)\left(\sum_i^k \frac{1}{x_i \hat{f}^2_{j-1}}\right)}{\left(\sum_i^k \frac{1}{\hat{f}^2_{j-1}}\right)\left(\sum_i^k \frac{1}{x_i^2 \hat{f}^2_{j-1}}\right) - \left(\sum_i^k \frac{1}{x_i \hat{f}^2_{j-1}}\right)^2} \qquad (5.8)$$

Observe that, if $\hat{f}_{j-1}$ is replaced by $\hat{V}_i$ in (5.7) and (5.8), these two systems give the initial values of $\hat{A}_0$ and $\hat{B}_0$.

The results from (5.7) and (5,8) are repeatedly used for the next iteration until the predetermined conditions specified in (5.9) and (5.10) are satisfied, i.e.,

$$\left|\frac{A_j - A_{j-1}}{A_j}\right| \leq 0.01 \qquad (5.9)$$

$$\left|\frac{B_j - B_{j-1}}{B_j}\right| \leq 0.01 \qquad (5.10)$$

It has been proved that the roots always exist in NHIS situation. The iteration usually stops after about five times of repetition to meet these conditions. Let the final estimates be $\hat{A}$ and $\hat{B}$. The relative variance can now be estimated by the curve

$$\tilde{V}(x_i) = \hat{A} + \hat{B}/x_i \qquad (5.11)$$

for aggregate x_i without going back to the derivation.

Fortran programs for the BHS procedures and curve fitting method are available at the National Center for Health Statistics.

An estimate $\tilde{V}$ is sometimes negative especially when x is large. In order to avoid such cases, one may impose the condition

$$\tilde{V}(x_i) \geq 0 \quad \text{or} \quad \hat{A} \geq -\hat{B}/x_i \qquad (5.12)$$

Often $\hat{A}$ is very small in NHIS situation. The relative variance can be approximated by

$$\tilde{V}(x_i) = \hat{B}/x_i$$

Often R^2 is used to measure the fitness of the curve to the data points. R^2 is defined by

$$R^2 = \frac{\sum_i^k (\tilde{V}(x_i) - \bar{V})^2}{\sum_i^k (\hat{V}(x_i) - \bar{V})^2} \qquad (5.13)$$

where $$\bar{V} = \frac{1}{k} \sum_i^k \hat{V}(x_i) \qquad (5.14)$$

R^2 is close to one when $\tilde{V}(x_i)$ is a good predictor of $\hat{V}(x_i)$ and close to zero when $\tilde{V}(x_i)$ is poor predictor.

Table B

Estimates of A and B and cut-off (CP*) points for NHIS aggregate data (x).

Symbol and type of x	year	A	B	CP*
A4AN-Persons in any age-sex-color category#.	'73	-0.000023	3,258.9	36
	'79	-0.000022	3,377.5	38
	'80	-0.000031	3,830.3	43
A4BN-Acute conditions. Persons injured	'73	0.000208	52,761.1	588
	'79	0.000074	69,598.2	774
	'80	0.000191	74,000.8	824
A4BW-Restricted activity and bed disability days	'73	0.000085	482,754.0	5,369
	'79	0.000406	558,050.8	6,229
	'80	0.000452	630,292.1	7,039
A4BW-Work or school loss days	'73	0.000104	303,506.0	3,376
	'79	-0.000135	350,808.0	3,892
	'80	0.000225	337,261.3	3,757
A4BM-Physician visits(P) or Dental visits (D).	@'73	0.000029	85,139.6	946
	P'79	0.000011	117,187.2	1,302
	D'79	0.000095	91,761.8	1,021
	@'80	0.000070	106,516.3	1,185
A4CN-Short stay hospital discharges	'73	0.000014	5,532.0	62
	'79	-0.000007	6,204.6	69
	'80	0.000076	6,579.0	73
A4CW-Short stay hospital days for discharges	'73	0.000519	98,041.8	1,096
	'79	0.000168	130,704.6	1,455
	'80	0.004483	111,371.8	1,302

\# except for persons or total number in any age-sex-color category in population. This curve is also used for persons with activity limitation or with hospital episode.

@ The data for P and D are combined.

* $30 \approx \sqrt{A+B/CP} \times 100$. CP is the number of counts in thousands.

In each group of NHIS variables, there are many items due to the age-sex-race and socio-economic-health categories. A sample of about 100 items is randomly taken from 680 data items in a group. The RSE for each of these items is calculated by BHS. These 100 points of RSE's are used to draw a curve to represent the sample variance for the group.

Let S_i be square roots of (5.11) multiplied by 100, i.e., relative standard error in percent.

The estimates of A, B, and the cut-off points in thousands at 30% relative standard error are given in Table B for the aggregate data from 1973, 1979, and 1980 National Health Interview Surveys. Here only a sample of items is used to draw a curve for the group. The use of partial items rather than all items might have an impact on the curve and hence it may be necessary to have this impact reflected on the variance curve.

The variance of the estimates $\hat{A}$, $\hat{B}$, and R^2 may be found by bootstrap method.

The series 10 publication for NHIS data includes the RSE curve of S_i for aggregates x_i.

Figure A shows the RSE of S_i for the aggregate x_i of short stay hospital days, **A4CW**, short stay hospital discharges, **A4CN**, and population characteritics, **A4AN** from 1980 NHIS.

Now the relative standard error of an estimate x_i is obtained from the prediction curve in Figure A without going through the direct calculation for individual items.

Relative variance of $p_i = x_i/y_i$ (x_i is subset of y_i) is,

$$V(p_i) = V(x_i) - V(y_i)$$

$$= \frac{B\ (1-P_i)}{Y_i P_i} \qquad (5.15)$$

Denote square root of (5.15) by S_i'. S_i' is RSE of p_i. Series 10 publication also included S_i' curve for given percent p_i. Figure B shows **P4AN** curve for RES's of percents for population characteristics from 1980 NHIS data.

The figure shows RSE's on the vertical scale and the percentages on the horizontal scale. In the figure, each curve is based on a different base number of y_i in millions.

The curve for four quarters data can be used for one quarter data when both are the same type of data based on the same recall period and when one quarter data is weighted up by the same weights as is the four quarter data.

A new curve for one quarter data should be drawn even for the same type of data when the weights for one quarter data are different from those of four quarter data. For instance, the number of acute conditions, e.g., common cold, is estimated quarterly to observe the seasonal changes of incidences. In this case, the weights used for one quarter data are different from those used for four quarter data. Thus, the curve for one quarter data should be different from that for four quarters data even for the same

Figure A **RSE of aggregate**

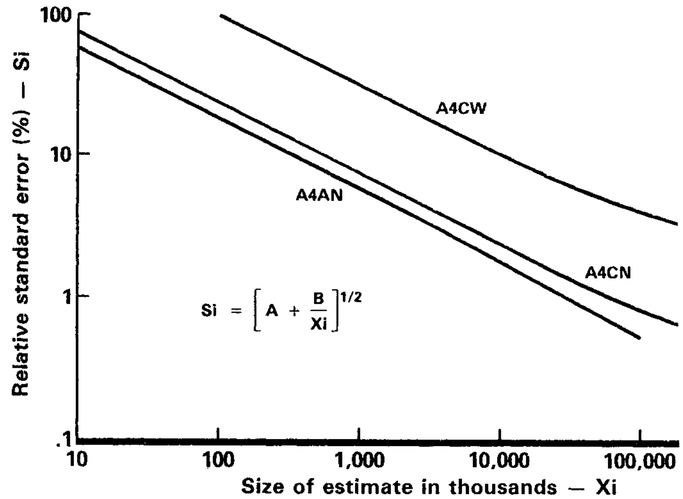
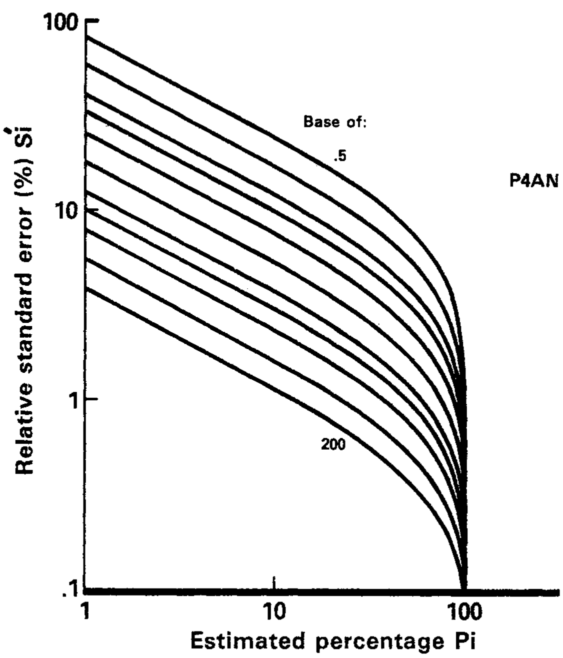


Figure B **RSE of percentages (Base of percentage shown on curves in millions)**



acute condition.

If all other conditions remain the same, a curve based on four quarters data may be used to deduce a curve, as a crude approximation, for the data from eight quarters or more assuming that total variance can be divided into 15% of between PSU's and 85% of within PSU's (Bryant, 1975).

6 COMMENTS

This method of curve fitting is a covenient procedure in smoothing the sample variances for a large number of items. Through this method, it is a simple matter for the users of data to find the sample variance of the data. But the accuracy of the smoothing model is a major concern. In order to increase this accuracy, we can further consider two problems, e.g., estimation of more accurate variance points and finding a curve which better represents these points.

When the data are based on a complex survey design, methods other than BHS procedure can also be considered to calculate the variance of data based on such complex sample surveys. Some of them are Linear approximation, Jackknife method, and Bootstrap method. BHS and these methods are useful especially when the sample elements reflect the population elements and actual survey design reasonably well.

Analysts frequently form pseudo elements, that differ from the original elements, and they often ignore the actual steps of sampling design. Then, they use one of these methods when it is too complex to find a closed form of variance. For instance, GSK method of analysis (Grizzle, Starmer, and Koch, 1969) utilizes an independent estimator of variance from Balanced Half Sample replication (BHS) or Linearization method for the National Health Examination Survey data and implant this result to derive the Wald statistic for a chi-square test statistic.

Another example is that Hidiroglou and Rao (1983) used Linearization (Williams, 1962) to estimate the sample variance of Canadian Health Survey data and used this result for chi-square testing.

A common problem to these procedures is that it is impossible to reflect the intracluster correlation correctly to variance estimation when a sample is taken by cluster sampling and the elements in the cluster are correlated. To avoid this problem, one may use a type of probability assumption defining the relationship of two members of any pair in the cluster (Choi, 1981). The sample variances thus derived will include such correlations and hence the curve based on these variances should be more efficient. Secondly, in order to obtain a more accurate curve, we need to reflect the actual characteristics of these variance points on the estimation of regression parameters. Such characteristics include the correlation between the estimates of variance points and these points are not usual variables but relative variances based on the weighted data from the complex sample surveys. Therefore, the improvements in these two problems will certainly increase the accuracy of curves in the presentation of such sample variances.

REFERENCES

Bryant, E. E.(1975): HIS Sampling Error. Unpublished memorandum to Director, HIS, NCHS, May 2,'75.

Choi, J. W. (1981). A further study on the analysis of categorical data from weighted cluster sample survey. ASA proceedings of the section on Social Statistics pp 217-222.

Cochran, W. G.(1977). Sampling techniques. 3rd. ed. Wiley.

Edelman, M. S.(1967). Curve fitting of keyfitz variances. Unpublished memorandum to R.H. Hanson.Bureau of the Census, August 16, 1967.

Grizzle, J. E.: Starmer, C. F.: and Koch, G. G. (1969). Analysis of categorical data by linear models. Biometrics 25, 389-504.

Hidiroglow, M. A. and Rao, J. N. K. (1983). Chi-square tests for the analysis of three-way contingency tables from the Canada Health Survey. Statistics Canada.

Jones, G.(1976). HES Variance-covariance matrix Unpublished document. ORM, NCHS.

McCarthy, P. (1966). Replication-an approach to the analysis of data from complex survey. Series 2, No. 14, NCHS.

Schnack, G. A. (1974). HIS variance-1973. Unpublished memorandum to E. L. Madigan. NCHS. April 3, 1974.

Tomlin, P. H. (1977). A Paradoxical situation regarding the b-parameter for a cross-tabulated estimate. Unpublished memorandum. Bureau of the Census. August 31, 1977.

Williams, W. H.(1962). The variance of an estimator with post-stratified weighting. JASA 57, 622-627.

ACKNOWLEDGEMENT

Authors are grateful to Joe Fred Gonzalez, Andrew White, and Randy Curtin, NCHS, for their comments that helped to improve this paper.