SAMPLING VARIATION OF AGE-ADJUSTED RATES 1/

Regina Loewenstein and Jack Elinson
Columbia University School of Public Health and Administrative Medicine

Introduction

Analysts of data from health surveys frequently encounter indices of medical care that vary widely by age. It is often desired to compare these indices among classes, e.g. income classes, which have different age compositions. Reports of health surveys describe many methods of analyzing this type of data, including rates for specific age groups within classes 2/, multiple regression or analysis of variance with age as one variable 3/, and comparison of age-adjusted rates 4/.

This paper will discuss methods of studying the sampling variation of age-adjusted means and proportions. The procedures include estimation of confidence intervals of age-adjusted means and tests of hypotheses about their homogeneity.

This research was done as part of a collaborative health interview survey covering 5344 persons in the Washington Heights Health District of New York City, conducted by the Columbia University School of Public Health and Administrative Medicine, 1960-1961. Questions about medical care in the year prior to the interview were included for the Patterns of Medical Care Study of the New York City Department of Health 5/.

The index of medical care used for illustrative purposes here is the reported number of physician visits per person per year. The distribution of this variable is not normal for any age group. For example, about half of all persons covered in these interviews had no reported visits, one-fourth had one through three visits, one-fifth had four through 14 visits, and the remaining 5 percent reported 15 through 100 visits.

Estimating Variances for Specific Age Groups

Variances of age-adjusted means are functions of variances of specific age groups which are estimated, of course, with consideration of the sample design.

In this survey, a self-weighting sample of housing units with families to be interviewed was selected by a two-stage stratified cluster sampling plan with varying first-stage sampling rates among strata, and with uniform first and second stage sampling rates within each stratum 6/. The ten strata were defined by geographical location, racial composition in 1957 and rent in 1950.

From this sample design an equal number of observations was expected from each cluster within the same stratum. When this sampling plan is used, the expression 7/ frequently suggested for the variance of a mean per sampling unit includes terms that are variances of cluster totals within each stratum. These variances of cluster totals would be meaningful only if there was an approximately equal number of observations from each cluster in the same stratum.

But, the number of interviews within clusters from the same stratum varied widely for several reasons. First, clusters within each stratum had been estimated to have approximately the same number of housing units, based on Block Statistics from the 1950 Census and from maps showing the number of floors, but not the number of apartments, in each building. Second, even if the number of housing units in clusters had been equal, response rates among clusters varied. Third, there was more than one family in almost 10 percent of the housing units in the sample.

In addition, the number of persons per family varied among clusters, and response rates were lower among smaller families 8/. Consequently, the number of persons covered in interviews from clusters within the same stratum varied; for example, from 6 to 58 in one stratum.

The number of persons of specific age groups covered in interviews from clusters in each stratum varied even more, because clusters had different age distributions and because there were higher refusal rates among older persons 8/. In one stratum, for example, 20 of 75 clusters had no persons 65 years and older, and four clusters had 6 to 16 persons in this age group. Thus, variances of age-specific means could not be determined by a method that assumed an approximately equal number of observations per cluster.

Random Group Method

Therefore, the random group method was used to estimate variances of means for specific age groups 9/. It was decided to use thirty random groups for computations.

The 5344 persons were randomly assigned to 31 groups of equal size, each group with approximately the same distribution of persons among the ten strata. Cases in the 31st group were randomly distributed to other groups when necessary 10/. All persons in the same family were assigned to the same random group.

For computation of the variance of a specific age group, the number of persons from that age group in each of thirty random groups was to be equal. The machine first distributed all persons in an age group who had been assigned to random groups 1 through 30 to 30 different locations. Persons in the 31st group were randomly distributed to locations with the smallest number of persons.

For example, there were 1531 persons 35 through 54 years old covered in the interviews. After the 31st group had been distributed among the smallest groups, there were five groups with 46 persons and 25 groups with 47 through 63 persons. The program instructed the machine to find the sum of visits reported for all 46 persons in each of the five groups, and the sum for 46 randomly selected persons in each of the other 25 locations. Thus, in the random group method, computation of the variance for this age group was based on 30 times 46, that is, 1380 randomly selected persons, or 90 percent of all persons in this age group.

The sum of the number of physician visits reported for an equal number of persons in each random group from each age group was used to compute variances as follows 9/:

If $\quad$ = number of reported physician visits for a person in one year,

K = number of persons in each random group used for computations,

T = number of random groups = 30

and X_g = total number of visits for K persons in group g,

then, the variance of X was estimated to be:

$$\frac{\sum_{g=1}^{T} X_g^2 - \left(\sum_{g=1}^{T} X_g\right)^2 \Big/ T}{K(T-1)} \qquad (1)$$

The variance of mean number of visits for each age group was estimated to be $\mathrm{Var}\,(\bar{X}) = \frac{\mathrm{Var}\,(X)}{30K}$, since $30K$ was the number of persons in each age group used to estimate $\mathrm{Var}\,(X)$. The standard error of the mean was estimated to be

$$\mathrm{S.E.}\ (\bar{X}) = \sqrt{\mathrm{Var}(\bar{X})}\ .$$

Table 1 summarizes estimates by the random group method for all persons and for persons in each of six age groups.

Comparisons of Variances and Means of Age Groups

Because the frequency distribution of visits was not normal, usual tests of homogeneity of variances might indicate differences that do not really exist 11/. Since the estimated variances for these six age groups ranged from 16 to 108 (Table 1), it seemed reasonable to reject the hypothesis that variances of the six age groups were homogeneous.

Therefore, tests of means that do not assume equal variances were needed. An approximate test of homogeneity of means of large samples allowing for unequal variances is a chi square test of homogeneity 12/. The least squares estimate of the mean is the weighted sum of observed means with the weights for each mean equal to the reciprocal of its estimated variance. The weighted sum of the squares of the deviations of the observed means from this estimate, using the same weights, has a chi square distribution. This test is expressed mathematically below.

If $\quad \bar{X}_i$ = observed mean number of physician visits of persons in age group i,

and $\mathrm{Var}(\bar{X}_i)$ = estimated variance of mean of age group i using the random group method (equals square of S.E. $(\bar{X}_i)$ in Table 1),

then $\quad \bar{X}_s = \sum_i \left[\bar{X}_i \,/\, \mathrm{Var}(\bar{X}_i)\right] \Big/ \sum_i \left[1 \,/\, \mathrm{Var}(\bar{X}_i)\right]$

= least squares estimate of the mean

and the statistic $\sum_i \left[(\bar{X}_i - \bar{X}_s)^2 \,/\, \mathrm{Var}(\bar{X}_i)\right]$ is approximately distributed as chi square with the number of degrees of freedom equal to one less than the number of groups.

Applying this test to mean visits of six age groups in Table 1, one rejects the hypothesis that the six means are equal. ($P<.001$)

One may also want to compare pairs of these means. Since these means were based on sufficiently large independent samples, and since the variances were estimated with sufficient degrees of freedom, i.e. 29, one may assume that the following ratio is approximately distributed as a normal deviate 13/:

$$\begin{aligned} Z &\doteq \left|\bar{X}_1 - \bar{X}_2\right| \Big/ \sqrt{\mathrm{Var}(\bar{X}_1 - \bar{X}_2)} \\ &\doteq \left|\bar{X}_1 - \bar{X}_2\right| \Big/ \sqrt{\mathrm{Var}(\bar{X}_1) + \mathrm{Var}(\bar{X}_2)} \\ &\doteq \left|\bar{X}_1 - \bar{X}_2\right| \Bigg/ \sqrt{\frac{\mathrm{Var}(X_1)}{N_1} + \frac{\mathrm{Var}(X_2)}{N_2}} \end{aligned}$$

where N_1 and N_2 are sizes of two independent samples from distributions with unequal variances and $\mathrm{Var}(X_i)$ are from Table 1.

Applying this test, one infers that the mean number of visits for persons 65 years and older is significantly higher than the mean for persons in each of the four youngest age groups, but not significantly different from the mean for persons 55 to 64 years old. (Tests done on 5 percent level).

Because of this wide variation of mean visits with age, comparison of means of classes with different age compositions might obscure variations due to factors other than age.

Age-Adjusted Means and Their Variances

Comparisons of age-adjusted means is one method to study differences of means of classes with different age compositions. Means of each of six age groups within an income class of size n, for example, were first computed. It was assumed that these six means were observed for a sample of size n with the same proportions in the six age groups as the total 5344 persons covered in interviews.

This procedure is expressed mathematically.

If n_i = number of persons of income class in age group i,

$\sum_i n_i = n$ = total number of persons in income class,

X_{ij} = number of physician visits for j'th person in i'th age group,

and $Var(X_i)$ = estimated variance of visits for persons in age group i of income class,

then

$$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij} / n_i = \text{mean number of visits of persons in age group i of income class}$$

and

$$Var(\bar{X}_i) = Var(X_i) / n_i = \text{variance of mean of age group i of income class.}$$

If W_i = proportion of 5344 persons in age group i,

then

$$\bar{X}_a = \sum_i (W_i \bar{X}_i) = \text{age-adjusted mean of income class.}$$

The variance of the age-adjusted mean is

$$Var(\bar{X}_a) = Var\ (\sum_i W_i \bar{X}_i)$$

$$= \sum_i W_i^2 Var(\bar{X}_i) \text{ 14/}$$

$$= \sum_i W_i^2\ (Var(X_i) / n_i) \qquad (2)$$

The estimated standard error of the age-adjusted mean is S.E. $(\bar{X}_a) = \sqrt{Var(\bar{X}_a)}$,

and the 95 percent confidence interval is estimated to be $\bar{X}_a \pm 2$ S.E. $(\bar{X}_a)$.

The example to be discussed compares age-adjusted mean number of visits per person in each of four income classes with different age distributions. For example, 32 percent of persons in the lowest income class was 65 years or older in contrast with 7 to 10 percent in the three other income classes.

Since estimated variances of age-adjusted means of each income class are functions of estimated variances for each age-income class, methods of estimating these variances will be described.

In the compromise method; the variance of X found by the random group method for all persons in each age group was used as the estimate of $Var(X_i)$ for persons of the corresponding age group in each income class. More specifically, the $Var(X_i)$ in Table 1 were substituted in equation (2) to estimate variances of age-adjusted means of each income class.

In the optimum method; more precise estimates of $Var(X_i)$ could be obtained by applying the random group method to each of 24 age-income classes. Because of the small number of persons in some of these age-income classes, the allocation to random groups by the procedure described above resulted in some random groups with no or very few persons. Therefore, modifications of this procedure would be needed, including the use of fewer than 30 groups and/or random groups without homogeneous distributions among the strata used for sample selection.

In the direct method; $Var(X_i)$ would be estimated by $\sum_j (X_{ij} - \bar{X}_i)^2 / (n_i - 1)$ for each age-income class. This method would be expected to underestimate variances because it ignores within cluster correlation.

For example, the variance estimated for all persons 35 to 54 years by the random group method was 108. (Table 1) Variances estimated by the direct method for persons of this age group in each of four income classes were considerably smaller; ranging from 33 to 74.

It can be shown mathematically that the variance for each age group estimated by the compromise method is the sum of variances within four income classes plus a weighted sum of differences of income class means from the overall mean 15/.

Thus, the first method is a compromise between the optimum method that gives more precise estimates but requires a great deal of time and money, and the direct method that underestimates variances.

The results of computations by the compromise method are shown in Table 2. Estimated 95 percent confidence intervals of age-adjusted means for four income classes were very similar. A chi square test of homogeneity of these means indicated that the age-adjusted mean number of reported physician visits per person per year did not differ significantly among four income classes. (It should be noted here that these reported visits included paid, prepaid and free visits in homes, offices and clinics.)

This finding of homogeneity of age-adjusted mean visits among income classes suggests further study. Since the average family size was smallest in the lowest income class (Table 2), future analyses might preferably use an index of family income per person. Other analyses might include comparisons of age-adjusted means of income-ethnic classes, for example.

Age-Adjusted Proportions

Age-adjusted proportions of persons with characteristics that vary with age can also be used to compare classes with different age compositions.

One method of studying sampling variation of age-adjusted proportions will be discussed.

Let n_i = number of persons of income class in age group i,

$\sum_i n_i = n$ = total number of persons in income class,

W_i = weight = proportion of 5344 persons in age group i,

p_{ir} = proportion of n_i persons in category r, for example, persons with at least one visit,

and $q_{ir} = 1 - p_{ir}$.

Then, $P_a = \sum_i (W_i p_{ir})$

= age-adjusted proportion of income class in category r,

and $Var(P_a) = \sum_i Var(W_i p_{ir}) = \sum_i W_i^2 \; Var\;(p_{ir})$.

The standard error of the age-adjusted rate is S.E. $(P_a) = \sqrt{Var(P_a)}$, and the 95 percent confidence interval is estimated to be

$$P_a \pm 2\; S.E.\; (P_a).$$

Estimates of Var (p_{ir}) that consider the sample design can be obtained by the random group method. Let the variable X_{ir} be defined as 1 for persons of age group i in category r and as 0 for persons not in category r. The mean of X_{ir} for age group i is an estimate of p_{ir}, and the $Var(X_i)$ estimated by expression (1) is an estimate of $Var(p_{ir})$. Comparisons of estimated confidence intervals and tests of homogeneity of $\bar{X}_a$ would be equivalent to analyses about P_a.

This method can be used, for example, to compare age-adjusted proportions of persons in each income class who had visits to an outpatient department.

Appropriate extensions would be needed to develop significant tests of age-adjusted distributions.

Summary

A health survey included questions about physician visits, which varied widely with age. Age-adjusted means were used to compare mean visits among classes with different age compositions.

Because of the complex sample design, the random group method was used to estimate variances for each age group.

Variance of the age-adjusted mean of each class was estimated as $\sum_i W_i^2\; (Var(X_i)\;/\;n_i)$, where $Var(X_i)$ was the variance estimated for all persons of the i'th age group by the random group method. Using these means and variances in a chi square test of homogeneity, one accepts the hypothesis that the age-adjusted mean visits of four income classes were equal.

Implications

This paper described the application of well-known statistical techniques to the study of sampling variation of age-adjusted means and proportions. Implications for other research work are:

1. The random group method is appropriate to estimate variances when one or more of the following circumstances obtain:

   a. When the method frequently suggested for estimating variances for the specific sample design cannot be applied because there are vastly different numbers of observations in sampling units expected to have approximately the same number of cases.

   b. When the frequently proposed method of estimating variances for that sample design would require a prohibitive amount of time and/or money if used for many variables and many types of classes.

2. The optimum method to estimate variances of age-adjusted means would use variances for each age group within each class, estimated by a procedure that considers the sample design, including the random group method.

   In a compromise method that requires less time and/or money, variances estimated for all persons of each age group by a method that considers the sample design can be used as estimates for variances of persons of corresponding ages within each class.

   Variances of each age group estimated by the compromise method would always be larger than corresponding variances estimated by the optimum method by an amount related to the difference between class means.

   Tests using variances estimated by the compromise method would be more conservative. The probability of accepting a false hypothesis would be greater with the compromise method than with the optimum method.

Table 1

Mean, Variance and Standard Error of Reported Physician Visits per Person per Year, by Age Group, Estimated by Random Group Method

Age Group (years) i	Number of persons covered in interviews N	Mean number of visits $\bar{X}$	Number of persons used to estimate variance $30K$	Estimated variance $Var(X)$	Estimated standard error of mean $S.E.(\bar{X})$[2]
Total	5344	3.1	5340[1]	110.7	.14
Under 5	386	2.2	270	16.1	.24
5-14	593	1.5	480	32.2	.26
15-34	1358	2.8	1290	49.2	.20
35-54	1531	3.0	1380	107.9	.28
55-64	743	3.9	570	87.0	.39
65 and older	733	4.9	540	103.5	.44

----

[1] The number of persons used to estimate variance for persons of all ages was greater than the total persons used to estimate variances for each age group.

[2] The standard error of the mean was estimated to be: $S.E.\ (\bar{X}) = \sqrt{\frac{Var(X)}{30K}}$.

Table 2

Estimated Standard Error and 95 Percent Confidence Interval of Age-adjusted Mean Reported Physician Visits per Person per Year, by Family Income Class

	Total	Family Income in 1960 or 1961			
		Under $3000	$3000-4999	$5000-7499	$7500 & over
Size of Family					
Number of persons of all ages	5344[1]	922	1221	1468	1125
Number of families interviewed	2216[1]	551	510	514	380
Number of persons per family	2.4	1.7	2.4	2.9	3.0
Mean visits per person	3.11	3.49	3.05	3.09	3.09
Age-adjusted mean visits per person					
Age-adjusted mean[2]		3.11	3.23	3.36	3.05
Estimated standard error[3]		.30	.26	.24	.26
Estimated 95 percent confidence interval[4]		2.5-3.7	2.7-3.8	2.9-3.8	2.5-3.6

---

1 Includes 608 persons in 261 families with family income not reported.

2 Weights (W_i) used to compute age-adjusted means were the proportions of 5344 persons in six age groups.

3 The estimated variances for each age group used to estimate the variance of age-adjusted means were $Var(X_i)$ from Table 1. The standard error of the age-adjusted mean was estimated to be

$$S.E. (\bar{X}_a) = \sqrt{\sum_i W_i^2 \; (Var(X_i) \; / \; n_i)} \; .$$

4 The 95 percent confidence interval was estimated to be $\bar{X}_a \pm 2$ S.E. (X_a).

FOOTNOTES

1/ This research was supported by the Health Research Council of New York City under Contract U-1053, and was conducted at the School of Public Health and Administrative Medicine, Columbia University.

2/ United States Department of Health, Education and Welfare, Public Health Service, National Health Survey, PHS Publication No. 584: Series B-19, Volume of Physician Visits, August 1960.

United States Department of Health, Education and Welfare, Public Health Service, National Center for Health Statistics, PHS Pub. No. 1000: Series 10, No. 9 - Medical Care, Health Status, and Family Income. May 1964.

3/ McNerney, W. J. and Study Staff. "Hospital and Medical Economics; A Study of Population, Services, Costs, Methods of Payment, and Controls," Vol. 1. Hospital Research and Educational Trust. Chicago, 1962.

4/ Kitagawa, Evelyn. "Standardized Comparisons in Population Research." Demography, Vol. 1, No. 1, 1964.

Levene, Howard. "Interrelations of Genetic and Environmental Factors in Disease." Annals of the New York Academy of Sciences. Vol. 84, Art. 17, 1960.

5/ This research was supported by Public Health Service Grant No. CH 00010-05 from the Division of Community Health Services, under direction of Edward A. Suchman.

6/ Elinson, J. and Loewenstein, R. "Community Fact Book for Washington Heights, New York City, 1960-61." School of Public Health and Administrative Medicine, Columbia University, 1963.

Glasser, M. "Master Sampling Plan for Columbia-Washington Heights Community Mental Health Project." School of Public Health and Administrative Medicine, Columbia University. Memorandum, December 15, 1958.

7/ In the original description of this sampling plan, Marvin Glasser suggested that variances per housing unit be computed by using Equation 5.14, page 318 of Hansen, M.; Hurwitz, W.; and Madow, W.G. "Sample Survey Methods and Theory; Methods and Applications." Vol. 1. John Wiley and Sons, Inc., 1956.

8/ Loewenstein, R.; Colombotos, J. and Elinson, J. "Interviews Hardest-to-Obtain in an Urban Health Survey." Proceedings of the Social Statistics Section, American Statistical Association, 1962.

9/ Hansen, M.; Hurwitz, W.; and Madow, W.G. "Sample Survey Methods and Theory; Methods and Applications." Vol. 1, pp. 440-444. John Wiley and Sons, Inc., 1956.

10/ Dr. Norman Matlin conceived of the concept of an additional group, and also wrote the program for this operation.

11/ Eisenhart, C. "The Assumptions underlying the Analysis of Variance." Biometrics, Vol. 3, No. 1, March 1947.

Scheffé, H. "The Analysis of Variance." John Wiley and Sons, 1959.

12/ Fertig, J.W. Mimeographed notes for Biostatistics 204, School of Public Health and Administrative Medicine, Columbia University, 1963.

13/ Walker, H. and Lev, J. "Statistical Inference." Henry Holt and Co., 1953.

14/ Based on Theorem: $\mathrm{Var}(\sum_i a_i X_i) = \sum a_i^2 \mathrm{Var}(X_i)$

if a is a constant.

15/ We are grateful to Professor John W. Fertig, School of Public Health and Administrative Medicine, Columbia University, for his contribution to this part of the paper, as well as his review of the entire manuscript and many helpful suggestions.