ON A LEAST SQUARES ADJUSTMENT OF A SAMPLED FREQUENCY TABLE WHEN THE EXPECTED MARGINAL TOTALS ARE KNOWN

By W. Edwards Deming and Frederick F. Stephan

1. Introduction. There are situations in sampling wherein the data furnished by the sample must be adjusted for consistency with data obtained from other sources or with deductions from established theory. For example, in the 1940 census of population a problem of adjustment arises from the fact that although there will be a complete count of certain characters for the individuals in the population, considerations of efficiency will limit to a sample many of the cross-tabulations (joint distributions) of these characters. The tabulations of the sample will be used to estimate the results that would have been obtained from cross-tabulations of the entire population.[1] The situation is shown in Fig. 1 in parallel tables for the universe and for the sample. For the universe the marginal totals $N_{i.}$ and $N_{.j}$ are known, but not the cell frequencies N_{ij}; for the sample, however, tabulation gives both the cell frequencies n_{ij} and the marginal totals $n_{i.}$ and $n_{.j}$.

In estimating any cell frequency of the universe, such as N_{ij}, three possibilities present themselves; from the sample one may make an estimate from the ith row alone, another from the jth column alone, and still another from the over-all ratio n_{ij}/n: specifically, the three estimates would be $n_{ij}N_{i.}/n_{i.}$, $n_{ij}N_{.j}/n_{.j}$, and $n_{ij}N/n$. As a result of sampling errors these will not be identical except by accident, and though any of them by itself may be considered accurate enough, still, if the whole $r \times s$ table of universe cell frequencies were so estimated, the marginal totals would not come out right. In this paper we present a rapid method of adjustment, which in effect combines all three of the estimates just mentioned, and at the same time enforces agreement with the marginal totals. The method is extended to varying degrees of cross-tabulation in three dimensions.

In any problem of adjustment where the conditions are intricate it is necessary to have a method that is straight-forward and self-checking; this becomes imperative when we realize that in the three-dimensional Case VII of the problem now at hand (*vide infra*), any adjustment in one cell must be balanced by adjustments in at least seven others. The method of least squares is one possible procedure for effecting an adjustment and at the same time enforcing certain conditions among the marginal totals. It is essentially a scheme for

[1] Examples will occur in the 1940 census publications. Further discussion of this problem and of the sampling procedure is given by the authors in "The sampling procedure of the 1940 population census," *Jour. Am. Stat. Assn.*, Vol. 35 (1940), pp. 615-630.

arriving at a set of calculated or adjusted observations that will satisfy the conditions of the problem, and at the same time minimize the sum of the weighted squares of the residuals, symbolized as

$$S = \Sigma w(n_c - n_0)^2 \tag{1}$$

n_c and n_0 being the calculated and observed numbers in a cell, and $n_c - n_0$ the corresponding residual. It is the nature of the conditions imposed on the adjusted values that distinguishes one type of problem from another. Least squares has the practical advantage of uniqueness, once the weights of the observations have been assigned, and it possesses the theoretical dignity of giving one kind of "best" estimates under ideal conditions of sampling. For our present purpose we shall minimize sums of the form

$$S = \Sigma(m_i - n_i)^2/n_i \tag{2}$$

n_i being the observed frequency in the ith cell, and m_i the calculated or adjusted frequency therein. The conditions among the m_i will arise from the fact that the marginal totals, after adjustment, must agree with their expected values, namely, the deflated marginal totals of the universe (for example, $m_{i.}$ and $m_{.j}$ as defined in eqs. (6) and (7)).

By definition, weight and variance are inversely proportional, hence the principle of least squares is identical with the minimizing of chi-square. Here the variance in the ith cell is $\nu_i(1 - \nu_i/n)$, where ν_i is the expected number in that cell, and n is the total number in the sample. Now if ν_i is sufficiently well approximated by n_i, it follows that if no cell contains an appreciable fraction of the whole sample (a circumstance requiring a fair sized number of cells—perhaps 100), the variance may be taken as ν_i for every i, and the minimized S can be used as chi-square. But regardless of the number of cells, if the n_i be not too much different from one another, so that the factor $1 - \nu_i/n$ may be treated as a constant, we still get the least squares solution by minimizing S as defined in eq. (2).

2. The two dimensional problem. Suppose that the data on two characteristics (e.g. age and highest grade of school completed) are obtained for each member of a universe of N individuals, and that tabulations of the data provide either (a) one set of marginal totals $N_{1.}, N_{2.}, \cdots, N_{r.}$; or ($b$) in addition, the marginal totals $N_{.1}, N_{.2}, \cdots, N_{.s}$. The nature of the tabulations is presumed such that it is not feasible to count the numbers N_{ij} in the cells, as would be done if one character were crossed with the other. Suppose, however, that for a sample of n individuals selected in a random manner from the universe, the two characters are crossed with each other, so that we know not only all the $s + r$ marginal totals $n_{.1}, \cdots, n_{r.}$ of the sample but also the numbers n_{ij} ($i = 1, 2, \cdots, r; j = 1, 2, \cdots, s$). The problem is to estimate the unknown frequencies N_{ij} in the cells of the universe. This will be done by finding the calculated or adjusted sample frequencies m_{ij} and then inflating them by the inverse sampling ratio N/n.

For the least squares solution we seek those values of m_{ij} that minimize[2]

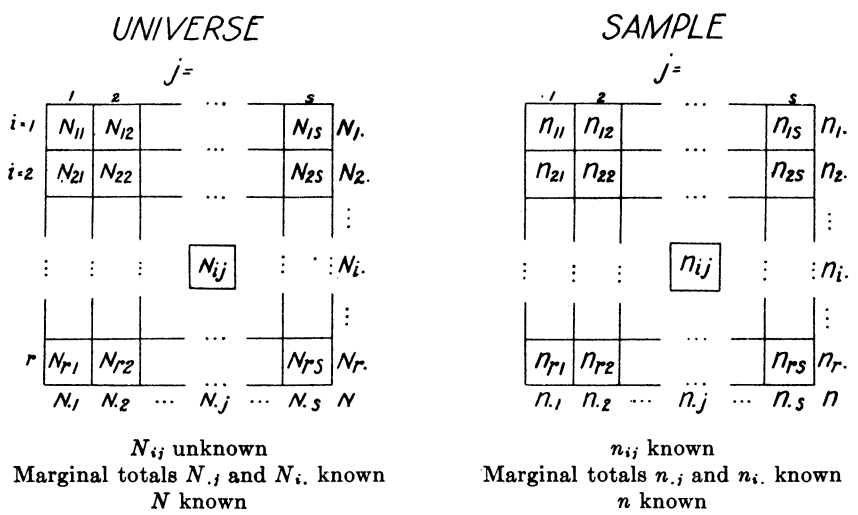
$$S = \Sigma(m_{ij} - n_{ij})^2/n_{ij} \tag{3}$$

wherein the m_{ij} are subjected to one of the following sets of conditions:

Case I: *One set of marginal totals known.* Assume $N_{1.}, N_{2.}, \cdots, N_{r.}$ to be known. Then we require

$$\sum_j m_{ij} = m_{i.}, \qquad i = 1, 2, \cdots, r. \tag{4}$$

These r equations constitute r conditions on the adjusted m_{ij}.



N_{ij} unknown
Marginal totals $N_{.j}$ and $N_{i.}$ known
N known

n_{ij} known
Marginal totals $n_{.j}$ and $n_{i.}$ known
n known

FIG. 1. SHOWING THE SYSTEM OF NOTATION FOR THE CELL FREQUENCIES AND MARGINAL TOTALS OF THE UNIVERSE AND THE SAMPLE IN THE TWO DIMENSIONAL PROBLEM

Case II: *Both sets of marginal totals known.* Here the adjusted cell frequencies must satisfy not only condition (4) but also

$$\sum_i m_{ij} = m_{.j} \qquad j = 1, 2, \cdots, s - 1 \tag{5}$$

there being now a total of $r + s - 1$ conditions. In both cases,

$$m_{i.} = N_{i.}n/N, \tag{6}$$

$$m_{.j} = N_{.j}n/N. \tag{7}$$

In other words, $m_{i.}$ and $m_{.j}$ are the deflated marginal totals, i.e., $N_{i.}$ and $N_{.j}$ divided by the actual sampling ratio N/n. The $m_{i.}$ and $m_{.j}$ are not independent, for

[2] The sign $\sum$ will denote summation over all possible cells, unless otherwise noted. $\sum_i$ will denote summation over all values of i, and similarly for an inferior j or k. The dot, as in $n_{.j}$, will signify the result of summing the n_{ij} over all values of i in the jth column.

$$N_{.1} + N_{.2} + \cdots + N_{.s} = N_{1.} + N_{2.} + \cdots + N_{r.} = N. \tag{8}$$

It is for this reason that if i runs through all r values in eq. (4), then j can run through only $s - 1$ in eq. (5). A similar equation also exists for the marginal totals of the sample, namely,

$$n_{.1} + n_{.2} + \cdots + n_{.s} = n_{1.} + n_{2.} + \cdots + n_{r.} = n. \tag{9}$$

Solution of the two dimensional Case I. Assuming that the adjusted values of the m_{ij} have been found, let each take on a small variation δm_{ij} ; then the differentials of eqs. (3) and (4) show that

$$\tfrac{1}{2}\delta S = \Sigma\{(m_{ij} - n_{ij})/n_{ij}\}\delta m_{ij} = 0 \qquad \text{(one equation)}, \tag{10}$$

$$\sum_j \delta m_{ij} = 0, \qquad i = 1, 2, \cdots, r \quad (r \text{ equations}). \tag{11}$$

Multiply now eq. (11i) by the arbitrary Lagrange multiplier $-\lambda_{i.}$, and add eqs. (10) and (11) to obtain

$$\Sigma\{(m_{ij} - n_{ij})/n_{ij} - \lambda_{i.}\}\delta m_{ij} = 0. \qquad \text{(one equation)}. \tag{12}$$

By the usual argument, one may now set each brace equal to zero, recognizing that the r Lagrange multipliers are then no longer arbitrary but must satisfy the relation

$$m_{ij} = n_{ij}(1 + \lambda_{i.}). \tag{13}$$

The adjusted frequencies m_{ij} can be computed at once as soon as the $\lambda_{i.}$ are found. To evaluate them one may rewrite the conditions (4) using the right-hand member of (13) for m_{ij} , obtaining

$$m_{i.} = n_{i.}(1 + \lambda_{i.}). \tag{14}$$

Another way to arrive at this same relation is to sum each member of eq. (13) in the ith row. However obtained $\lambda_{i.}$ is now known, since $m_{i.}$ and $n_{i.}$ are known, and in fact eq. (13) now gives

$$m_{ij} = n_{ij}(m_{i.}/n_{i.}). \tag{15}$$

The adjustment is thus a simple proportionate one by rows, the cells in any one row all being raised or lowered by the proportionate adjustment in the row total. Case I thus amounts to r independent one dimensional proportionate adjustments, one for each row, and any one or all may be carried out, as desired. This result can be obtained by a simpler approach but is presented in this way for consistency with later cases.

The minimized sum of squares may be computed directly, or from the row totals by seeing that

$$S = \sum_i (m_{i.} - n_{i.})^2/n_{i.}. \tag{16}$$

The term $(m_{i.} - n_{i.})^2/n_{i.}$ for the ith row may be considered separately, and

used as χ^2 with $s - 1$ degrees of freedom, or all rows may be combined into the minimized S as given in eq. (16), and used as χ^2 with $r(s - 1)$ degrees of freedom.

Solution of the two dimensional Case II. In addition to eqs. (11) we now have also

$$\sum_i \delta m_{ij} = 0 \qquad j = 1, 2, \cdots, s - 1 \tag{17}$$

which comes by differentiating eqs. (5). By addition of eqs. (10), (11), and (17), after multiplying eq. (11i) by $-\lambda_{i.}$ and eq. (17j) by $-\lambda_{.j}$, we obtain

$$\Sigma\{(m_{ij} - n_{ij})/n_{ij} - \lambda_{i.} - \lambda_{.j}\}\delta m_{ij} = 0. \tag{18}$$

Equating each brace to zero, as before, we find that

$$m_{ij} = n_{ij}(1 + \lambda_{i.} + \lambda_{.j}) \tag{19}$$

wherein $\lambda_{.s}$ is to be counted 0. The adjustment is now no longer proportionate by rows, but involves every cell.

To evaluate the Lagrange multipliers in eq. (19) we may sum the two members downward and across in Fig. 1 and obtain the $r + s - 1$ normal equations

$$\begin{aligned} n_{i.}\lambda_{i.} + \sum_j n_{ij}\lambda_{.j} &= m_{i.} - n_{i.}, \qquad i = 1, 2, \cdots, r \\ \sum_i n_{ij}\lambda_{i.} + n_{.j}\lambda_{.j} &= m_{.j} - n_{.j}, \qquad j = 1, 2, \cdots, s - 1. \end{aligned} \tag{20}$$

These can be reduced for numerical computation. The top row solved for $\lambda_{i.}$ gives

$$\lambda_{i.} = (1/n_{i.})\{m_{i.} - \sum_j n_{ij}\lambda_{.j}\} - 1 \tag{21}$$

whereupon by substitution into the bottom row of eqs. (20) we arrive at the $s - 1$ normal equations

$$\begin{array}{lllll} \lambda_{.1} & \lambda_{.2} & \cdots & \lambda_{.s-1} & = 1 \\ \hline n_{.1} - \sum_i \dfrac{n_{i1}n_{i1}}{n_{i.}} & - \sum_i \dfrac{n_{i1}n_{i2}}{n_{i.}} & \cdots & - \sum_i \dfrac{n_{i1}n_{i,s-1}}{n_{i.}} & = m_{.1} - \sum_i \dfrac{n_{i1}m_{i.}}{n_{i.}} \\ & n_{.2} - \sum_i \dfrac{n_{i2}n_{i2}}{n_{i.}} & \cdots & - \sum_i \dfrac{n_{i2}n_{i,s-1}}{n_{i.}} & = m_{.2} - \sum_i \dfrac{n_{i2}m_{i.}}{n_{i.}} \\ & & & \vdots & \vdots \\ & & & n_{.s-1} - \sum_i \dfrac{n_{i,s-1}n_{i,s-1}}{n_{i.}} & = m_{.s-1} - \sum_i \dfrac{n_{i,s-1}m_{i.}}{n_{i.}} \\ & & & & 0. \end{array} \tag{22}$$

Because of symmetry in the coefficients, those below the diagonal are not shown, indeed, in a systematic computation, they are not used. The 0 in the bottom

row is appended for the computation of the minimized S, if desired. The number of Lagrange multipliers to be solved for directly is $s - 1$, and the remaining ones come by substitution into eq. (21), $\lambda_{.s}$ being counted 0.

A simple procedure for calculating the coefficients in the normal equations (22) is to set up a preparatory table by dividing each n_{ij} in the ith row by $\sqrt{n_{i.}}$; also to write down $m_{i.}/\sqrt{n_{i.}}$ for that row, for use on the right-hand side of the normal equations (compare Tables I and II). In machine calculation the constant divisor $\sqrt{n_{i.}}$ would be left on the keyboard until the entire ith row is divided; or, if reciprocal multiplication is preferred, the multiplier $1/\sqrt{n_{i.}}$ would be left on. From this preparatory table, the cumulation of squares and cross-products in the vertical gives the required summations for the coefficients. The sum check would be applied in the usual manner.

3. A numerical example of the two dimensional Case II. The fact is that in practice one need not bother about forming and solving the normal equations because they will be displaced by a simplifying iterative procedure, to be explained in a later section. For illustration, however, we may do an example both ways, first using the normal equations and the adjustment (19), later on accomplishing the same results by the quicker method.

We may start with the unitalicized numbers in the 4×6 array of Table I, assuming these to be the sampling frequencies n_{ij} to be adjusted. Actually, they were obtained by deflating 1/20th (for a supposed 5 per cent sample) the New England age $\times$ state table on p. 1108 of vol. 2 of the *Fifteenth Census of the U. S.*, 1930, then varying the deflated values by chance with Tippett's numbers to get our sampling frequencies n_{ij}. The italicized entries in Table I represent the final (adjusted) m_{ij}, and it is these that we now set out to get. We start off with the sampling frequencies n_{ij} and the known marginal totals $m_{.1}$, $m_{.2}$, etc., where $m_{i.} = N_{i.}n/N$, $m_{.j} = N_{.j}n/N$, as in eqs. (6) and (7). The Lagrange multipliers shown along the left-hand and top borders arise in the calculations now to be undertaken.

Table II is the preparatory table, advised at the close of the last section. It is derived from Table I by dividing the ith row of sample frequencies by $\sqrt{n_{i.}}$. For example, the entry 8.64 in the cell $i = 3, j = 2$ comes by dividing 419 by $\sqrt{2352}$, 419 being the entry in the cell of the same indices in Table I, and 2352 being the sum of the third row. The sums at the bottom and right-hand side are for checking the formation of the normal equations. The cumulations of squares and cross-products along the vertical give the summations required for the normal eqs. (22), which now appear numerically as eqs. (23).

	No.	$\lambda_{.1}$	$\lambda_{.2}$	$\lambda_{.3}$	=	1
	1	7413	−3549	−2354	=	3197×10^{-2}
(23)	2		4441	−544	=	2356
	3			3129	=	−3222
	4					0

Performing the solution by any favorite procedure one will obtain

(24) $\lambda_{.1} = .01182 \quad \lambda_{.2} = .01490 \quad \lambda_{.3} = .00119$

TABLE I

A table of artificial sample frequencies, an artificial 5 percent sample of native white persons of native white parentage attending school, by age by state, New England, 1930. The adjusted frequency m_{ij} in each cell is shown italicized just below the corresponding sample frequency n_{ij}

Age			7 to 13	14 & 15	16 & 17	18 to 20	
		$j =$	1	2	3	4	$n_{i.}$
		$\lambda_{.j} =$	.0118	.0149	.0012	0	$m_{i.}$
State	i	$\lambda_{i.}$					
Maine	1	−.0146	3623	781	557	313	5274
			3613	*781*	*550*	*308*	*5252*
New Hampshire	2	−.0003	1570	395	251	155	2371
			1588	*401*	*251*	*155*	*2395*
Vermont	3	.0234	1553	419	264	116	2352
			1608	*435*	*270*	*119*	*2432*
Massachusetts	4	−.0162	10538	2455	1706	1160	15859
			10492	*2452*	*1680*	*1141*	*15766*
Rhode Island	5	−.0230	1681	353	171	154	2359
			1662	*350*	*167*	*150*	*2330*
Connecticut	6	−.0034	3882	857	544	339	5622
			3915	*867*	*543*	*338*	*5662*
		$n_{.j}$	22847	5260	3493	2237	33837
		$m_{.j}$	*22877*	*5285*	*3462*	*2213*	*33837*

The adjusted m_{ij} (italicized) are rounded off, hence when summed may occasionally disagree a unit or so with the expected marginal totals (also italicized), the latter arise by deflation from the universe rather than by direct addition of the m_{ij}.

whereupon by substitution into eq. (21) comes

(25)
$$\lambda_{1.} = -.0146 \quad \lambda_{4.} = -.0162$$
$$\lambda_{2.} = -.0003 \quad \lambda_{5.} = -.0230$$
$$\lambda_{3.} = +.0234 \quad \lambda_{6.} = -.0034.$$

The next step is to compute the m_{ij} by eq. (19). Table I is now bordered with the Lagrange multipliers for a convenient arrangement of the factors required, and the calculation is completed. It will be noted that, for example

(26) $$m_{32} = 419(1 + .0234 + .0149) = 435.$$

The m_{ij} thus calculated are shown italicized in Table I. The marginal totals, found by adding the m_{ij} just calculated, do not agree exactly everywhere with the expected totals, because of rounding off to integers: the errors of closure, however, are slight, and it is a simple matter to raise or lower some of the larger cells by a unit or two to force exact satisfaction of the conditions, if this is desired.

4. The three dimensional problem. Here the N cards of the universe are sorted and counted for one and perhaps a second and third characteristic, and possibly crossed by pairs in various combinations (Cases I–VII). The sample of n, however, is crossed by all three characteristics, which is to say that the

TABLE II

This comes by dividing each sample frequency in Table I by the corresponding $\sqrt{n_{i\cdot}}$. *(This operation would ordinarily be done a row at a time)*

	j =				$m_{i\cdot}/\sqrt{n_{i\cdot}}$	Sum
	1	2	3	4		
$i = 1$	49.89	10.75	7.67	4.31	72.32	144.94
2	32.24	8.11	5.15	3.18	49.19	97.87
3	32.02	8.64	5.44	2.39	50.15	98.64
4	83.68	19.49	13.55	9.21	125.19	251.12
5	34.61	7.27	3.52	3.17	47.97	96.54
6	51.77	11.43	7.26	4.52	75.51	150.49
Sum	284.21	65.69	42.59	26.78	420.33	839.60

cell frequencies n_{ijk} are all known (refer to Fig. 2). As before, the adjusted frequencies are required.

Case I: One set of slice totals known. Assume the slice totals $N_{1\cdot\cdot}$, $N_{2\cdot\cdot}$, $\cdots$, $N_{r\cdot\cdot}$ to be known; the conditions are then

(27) $$\sum_{jk} m_{ijk} = m_{i\cdot\cdot} = N_{i\cdot\cdot}n/N \qquad i = 1, 2, \cdots r$$

being r in number. The summation to be minimized is

(28) $$S = \Sigma(m_{ijk} - n_{ijk})^2/n_{ijk}$$

being similar to that in eq. (3), except that now there are three indices to be summed over instead of two. Following a procedure similar to that used before, we differentiate eqs. (27) and (28) and introduce the r Lagrange multipliers λ_i.

with eq. (27). The steps are identical with those of the two dimensional Case I, and the result is at once

$$m_{ijk} = n_{ijk}(1 + \lambda_{i..}) = n_{ijk}(m_{i..}/n_{i..}). \tag{29}$$

This adjustment, like that shown by eq. (15), is a simple proportionate one, but this time by slices rather than by columns. All cell frequencies having the same i index are raised or lowered in the same proportion.

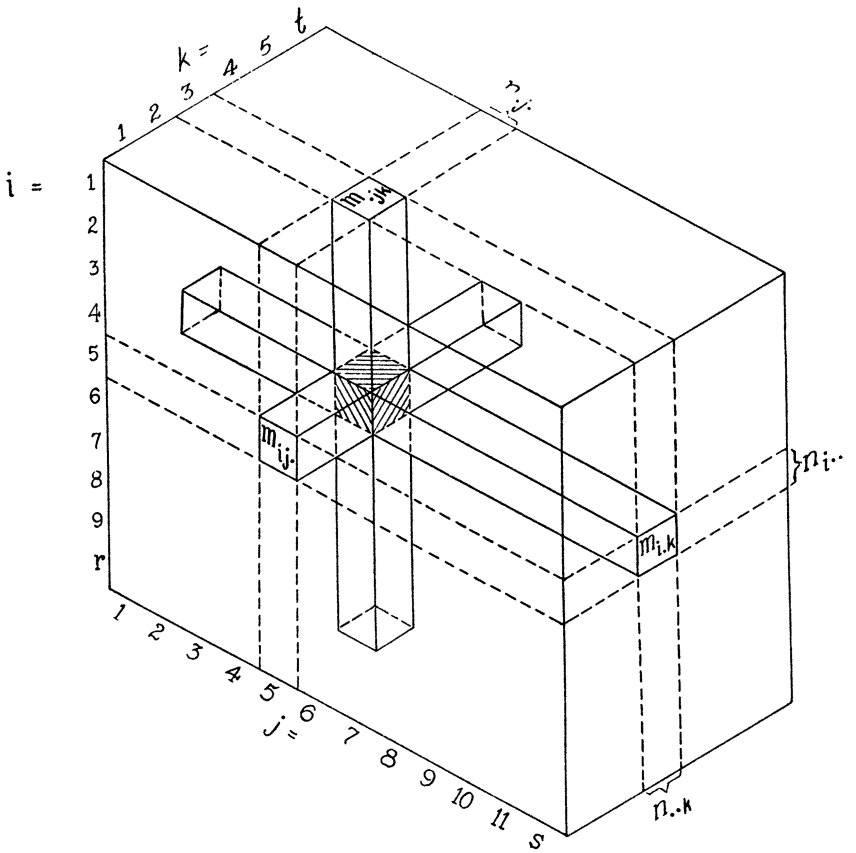


FIG. 2. SHOWING THE SYSTEM OF NOTATION FOR THE CELL FREQUENCIES AND MARGINAL TOTALS IN THE THREE DIMENSIONAL SAMPLE

Case II: *Two sets of slice totals known.* Here, in addition to the slice totals of Case I we know also

$$N_{.1.}, N_{.2.}, \cdots, N_{.s.}$$

whence arise the $s - 1$ additional conditions

$$\sum_{ik} m_{ijk} = m_{.j.} = N_{.j.}n/N, \qquad j = 1, 2, \cdots, s - 1. \tag{30}$$

Using the Lagrange multiplier $\lambda_{.j.}$ here, and $\lambda_{i..}$ with eq. (27) as before, we find that

$$m_{ijk} = n_{ijk}(1 + \lambda_{i..} + \lambda_{.j.}) \tag{31}$$

in which $\lambda_{.s.}$ is to be counted zero. This adjustment is proportionate by tubes, the ratio m_{ijk}/n_{ijk} being constant along the ijth tube and in fact equal to $m_{ij.}/n_{ij.}$, independent of k. Unfortunately we do not here know the face totals $m_{ij.}$ and are unable to make use of the proportionality as we shall in Case IV.

To solve for the $r + s - 1$ Lagrange multipliers we sum the members of eq. (31) over j and then over i and arrive at the normal equations

$$\begin{aligned} n_{i..}\lambda_{i..} + \sum_j n_{ij.}\lambda_{.j.} &= m_{i..} - n_{i..}, \quad i = 1, 2, \cdots, r, \\ \sum_i n_{ij.}\lambda_{i..} + n_{.j.}\lambda_{.j.} &= m_{.j.} - n_{.j.}, \quad j = 1, 2, \cdots, s-1. \end{aligned} \tag{32}$$

These can be reduced to $s - 1$ equations in precisely the same way that eqs. (20) were reduced, but because of the iterative process to come further on, we shall not pursue the reduction here.

Case III: All three sets of slice totals known. All slice totals

$$\begin{aligned} &N_{.1.}, N_{.2.}, \cdots, N_{.s.} \\ &N_{1..}, N_{2..}, \cdots, N_{r..} \\ &N_{..1}, N_{..2}, \cdots, N_{..t} \end{aligned}$$

now being known, in addition to conditions (27) and (30) we require here

$$\sum_{ij} m_{ijk} = m_{..k} = N_{..k}n/N, \qquad k = 1, 2, \cdots, t-1 \tag{33}$$

which makes a total of $r + (s - 1) + (t - 1)$ or $r + s + t - 2$ conditions. The same kind of manipulation as used heretofore gives

$$m_{ijk} = n_{ijk}(1 + \lambda_{i..} + \lambda_{.j.} + \lambda_{..k}) \tag{34}$$

with $\lambda_{.s.}$ and $\lambda_{..t}$ to be counted zero. The adjustment is no longer proportionate by slices or tubes, but involves every cell. In practice, once the normal equations are solved and the Lagrange multipliers worked out, one proceeds very much as in the two dimensional Case II: for each of the t slices, corresponding to the t values of k, there will be a two dimensional adjustment, the 1 in eq. (19) being replaced now by $1 + \lambda_{..k}$.

The normal equations for the Lagrange multipliers can be found by performing double summations on eq. (34). The result is

$$\begin{aligned} n_{i..}\lambda_{i..} + \sum_j n_{ij.}\lambda_{.j.} + \sum_k n_{i.k}\lambda_{..k} &= m_{i..} - n_{i..}, \quad i = 1, 2, \cdots, r, \\ \sum_i n_{ij.}\lambda_{i..} + n_{.j.}\lambda_{.j.} + \sum_k n_{.jk}\lambda_{..k} &= m_{.j.} - n_{.j.}, \quad j = 1, 2, \cdots, s-1, \\ \sum_i n_{i.k}\lambda_{i..} + \sum_j n_{.jk}\lambda_{.j.} + n_{..k}\lambda_{..k} &= m_{..k} - n_{..k}, \quad k = 1, 2, \cdots, t-1. \end{aligned} \tag{35}$$

If these calculations were to be carried out, one would simplify the computation by solving the top row for $\lambda_{i..}$, getting

$$\lambda_{i..} = (1/n_{i..})\{m_{i..} - \sum_j n_{ij.}\lambda_{.j.} - \sum_k n_{i.k}\lambda_{..k}\} - 1 \tag{36}$$

and then substituting this into the middle and last rows of eqs. (35) to get a reduced set of $s + t - 2$ normal equations for the Lagrange multipliers $\lambda_{.j.}$ and $\lambda_{..k}$, the numerical values of which when set back into eq. (36) give the $\lambda_{i..}$. In all the summations of eqs. (35) and (36), $\lambda_{.s.}$ and $\lambda_{..t}$ would be counted zero. But here again, the iterative process to be explained later will displace the use of normal equations, so actually we are not interested in reducing them.

Case IV: *One set of face totals known.* It may be that the rs face totals

$$N_{11.}, N_{12.}, \cdots, N_{ij.}, \cdots, N_{rs.}$$

are known from crossing the i and j characters in the universe. The conditions are then

$$\sum_k m_{ijk} = m_{ij.} = N_{ij.}n/N \qquad \begin{matrix} i = 1, 2, \cdots, r, \\ j = 1, 2, \cdots, s. \end{matrix} \tag{37}$$

The adjustment here turns out to be

$$m_{ijk} = n_{ijk}(1 + \lambda_{ij.}); \tag{38}$$

but by summing both sides over the index k to evaluate $\lambda_{ij.}$ it is seen that

$$m_{ij.} = n_{ij.}(1 + \lambda_{ij.}), \tag{39}$$

whence

$$m_{ijk} = n_{ijk}(m_{ij.}/n_{ij.}). \tag{40}$$

This adjustment is thus proportionate by tubes, like that in eq. (31), though here the factor $m_{ij.}/n_{ij.}$ is known and eq. (40) can be applied at once.

Case V: *One set of face totals, and one set of slice totals known.* Sometimes, in addition to the rs face totals of Case IV, the slice totals

$$N_{..1}, N_{..2}, \cdots, N_{..t}$$

will also be known, in which circumstances the conditions (37) are to be accompanied by

$$\sum_{ij} m_{ijk} = m_{..k} = N_{..k}n/N, \qquad k = 1, 2, \cdots, t - 1. \tag{41}$$

The same procedure as previously applied yields now

$$m_{ijk} = n_{ijk}(1 + \lambda_{ij.} + \lambda_{..k}) \tag{42}$$

with $\lambda_{..t}$ to be counted zero. Summations performed over k, and then over i and j together, give the normal equations

$$(43)\qquad \begin{aligned} n_{ij.}\lambda_{ij.} + \sum_k n_{ijk}\lambda_{..k} &= m_{ij.} - n_{ij.}, \\ \sum_{ij} n_{ijk}\lambda_{ij.} + n_{..k}\lambda_{..k} &= m_{..k} - n_{..k}. \end{aligned}$$

The number of equations is $rs + t - 1$, since $\lambda_{..t}$ does not exist. As before, a simplification can be effected by solving the top row for $\lambda_{ij.}$ and making a substitution into the lower one, but because of the great advantage of the iterative process to be seen further on, we shall not carry out the reduction.

Before going on it might be noted that although this case is three dimensional, it reduces to the two dimensional Case II if one considers that $ij.$ is one index running through the values 11, 12, $\cdots$, 21, 22, $\cdots$, rs, and that $..k$ is a second index running through the values 1, 2, $\cdots$, t. This can be seen by the similarity between eqs. (43) and (20).

Case VI: *Two sets of face totals known.* If in addition to the face totals of Case IV, the face totals

$$N_{.11}, N_{.12}, \cdots, N_{.st}$$

are also known from further crossing the j and k characters in the universe, we shall require

$$(44)\qquad \sum_i m_{ijk} = m_{.jk} = N_{.jk}n/N, \qquad \begin{matrix} j = 1, 2, \cdots, s, \\ k = 1, 2, \cdots, t-1 \end{matrix}$$

in addition to the conditions (37). In place of eq. (40) of Case IV we now find that

$$(45)\qquad m_{ijk} = n_{ijk}(1 + \lambda_{ij.} + \lambda_{.jk})$$

in which $\lambda_{.jt}$ is to be counted zero for all j. No simple relation such as eq. (40) is possible here, because the adjustment is not proportionate by tubes; the Lagrange multipliers must be evaluated. This can be accomplished by summing the members of eq. (45) over k and i in turn, resulting in the normal equations

$$(46)\qquad \begin{aligned} n_{ij.}\lambda_{ij.} + \sum_k n_{ijk}\lambda_{.jk} &= m_{ij.} - n_{ij.}, \\ \sum_i n_{ijk}\lambda_{ij.} + n_{.jk}\lambda_{.jk} &= m_{.jk} - n_{.jk}. \end{aligned}$$

Since $\lambda_{.jt}$ does not exist for any values of j, the number of equations is $rs + s(t-1) = s(r + t - 1)$. They break up at once into s sets each of $r + t - 1$ equations, one set for every j value. In fact, the problem can be considered as s sets of the two dimensional Case II. Any one value of j gives a slice, which can be looked upon as fulfilling the specifications of the two dimensional Case II. Each set of normal equations can be reduced in the same manner that eqs. (20) were reduced.

Case VII: *All three sets of face totals known.* All totals now being known, we require

$$\sum_k m_{ijk} = m_{ij.} = N_{ij.}n/N, \qquad \begin{matrix} i = 1, 2, \cdots, r, \\ j = 1, 2, \cdots, s, \end{matrix} \tag{37}$$

$$\sum_i m_{ijk} = m_{.jk} = N_{.jk}n/N, \qquad \begin{matrix} j = 1, 2, \cdots, s, \\ k = 1, 2, \cdots, t-1, \end{matrix} \tag{44}$$

$$\sum_j m_{ijk} = m_{i.k} = N_{i.k}n/N, \qquad \begin{matrix} i = 1, 2, \cdots, r-1, \\ k = 1, 2, \cdots, t-1. \end{matrix} \tag{47}$$

The adjusting relation is

$$m_{ijk} = n_{ijk}(1 + \lambda_{ij.} + \lambda_{.jk} + \lambda_{i.k}) \tag{48}$$

in which $\lambda_{.jt}$ is to be counted zero for any j, $\lambda_{r.k}$ for any k, and $\lambda_{i.t}$ for any i. The normal equations for the Lagrange multipliers are

$$\begin{aligned} n_{ij.}\lambda_{ij.} + \sum_k n_{ijk}\lambda_{.jk} + \sum_k n_{ijk}\lambda_{i.k} &= m_{ij.} - n_{ij.} \\ \sum_i n_{ijk}\lambda_{ij.} + n_{.jk}\lambda_{.jk} + \sum_i n_{ijk}\lambda_{i.k} &= m_{.jk} - n_{.jk} \\ \sum_j n_{ijk}\lambda_{ij.} + \sum_j n_{ijk}\lambda_{.jk} + n_{i.k}\lambda_{i.k} &= m_{i.k} - n_{i.k} \end{aligned} \tag{49}$$

being $rs + rt + st - r - s - t + 1$ in number. They can be reduced in the same way that previous normal equations have been reduced; but here again, the iterative process will render the use of normal equations unnecessary, except for theoretical purposes, e.g. justification of the iterative process.

5. A simplified procedure—iterative proportions. It is well known in least squares that the number of Lagrange multipliers in any problem is equal to the number of conditions imposed on the adjustment. Here the conditions have appeared in sets, depending on which marginal totals are involved. By a comparison of eqs. (15) and (29) on the one hand, with eqs. (19), (31), (34), (42), (45), and (48) on the other, we see that wherever there was only one set of marginal totals involved we came out with a proportionate adjustment, but that in all other cases it was not so; the Lagrange multipliers involved were unfortunately related to one another through normal equations. We now make the observation, however, that as a first approximation the adjustments may all be considered proportionate, and we shall be able to write down an expression for the error in this approximation, and shall be able to eliminate it by a succession of proportionate adjustments.

Take the two dimensional Case II for an example. In eq. (21) one may recognize $(1/n_{i.}) \sum_j n_{ij}\lambda_{.j}$ as a weighted average of $\lambda_{.j}$ for the ith row. There will be a weighted average of $\lambda_{.j}$ for the first row, another for the second, etc., one for each value of i; consequently one may appropriately speak of the ith

average of $\lambda_{.j}$, writing it i-av. $\lambda_{.j}$. Substituting from eq. (21) into (19) one then sees the adjustment (19) appear as

$$m_{ij} = n_{ij}(m_{i.}/n_{i.} + \lambda_{.j} - i\text{-av. } \lambda_{.j}). \tag{50}$$

If, on the other hand, $\lambda_{.j}$ had been eliminated from eqs. (20), instead of $\lambda_{i.}$, the result would have been

$$m_{ij} = n_{ij}(m_{.j}/n_{.j} + \lambda_{i.} - j\text{-av. } \lambda_{i.}). \tag{51}$$

From either eq. (50) or (51) it is clear why the adjustment (19) is not proportionate by rows or columns, and why Case II does not break up into r or s sets of Case I: the reason is that $\lambda_{.j}$ in any cell is not necessarily equal to the average $\lambda_{.j}$ for that row, nor is $\lambda_{i.}$ in any cell necessarily equal to the average $\lambda_{i.}$ for that column. If nevertheless one were to make the simple proportionate adjustment

$$m'_{ij} = n_{ij}(m_{i.}/n_{i.}) \tag{52}$$

along the horizontal in the ith row, the horizontal conditions (4) will be enforced but not the vertical ones (5); i.e., it will be found that $m'_{i.} = m_{i.}$, but that usually not all $m'_{.j} = m_{.j}$. This is because eq. (52) effects only a partial adjustment, each m'_{ij} being in error through the disparity between the $\lambda_{.j}$ proper to the jth column, and the average of all the $\lambda_{.j}$ for the ith row, as seen in eq. (50). This error can then be diminished by turning the process around and subjecting these m'_{ij} to a proportionate adjustment in the vertical according to the equation

$$m''_{ij} = m'_{ij}(m_{.j}/m'_{.j}) \tag{53}$$

which may be considered an application of eq. (51) wherein the disparity between any $\lambda_{i.}$ and the average $\lambda_{i.}$ for the jth column has been neglected. It is the vertical conditions that will now be found satisfied, but perhaps not all of the horizontal ones, because some of the row totals may have been disturbed. The cycle initiated by eq. (52) is therefore repeated, and the process is continued until the table reproduces itself and becomes rigid with the satisfaction of all the conditions, both horizontal and vertical. The final results coincide with the least squares solution, which is thus accomplished without the use of normal equations.

Usually two cycles suffice. In practice the work proceeds rapidly, requiring only about one-seventh as much time as setting up the normal equations and solving them. The tables III–V show the various stages of the work when the method of iterative proportions is applied to the sample frequencies of Table I. It will be noticed that the results of the third approximation (Table V) are final, since if the process were continued, the table would only reproduce itself.

The same process can be extended to three or more dimensions with an even greater relative saving in time. To see how the method of iterative proportions

applies in one of the three dimensional cases, we may go back to Case III. By the substitution afforded through eq. (36) the adjusting eq. (34) may be put into the form

TABLE III

The method of iterative proportions applied to the data of Table I. First stage: A proportionate adjustment by rows by eq. (52). Note that $m'_{i.} = m_{i.}$, *but that* $m'_{.j} \neq m_{.j}$

	$j = 1$	2	3	4	m'	$m_{i.}$
$i = 1$	3608	778	555	312	5253	5252
2	1586	399	254	157	2396	2395
3	1606	433	273	120	2432	2432
4	10476	2441	1696	1153	15766	15766
5	1660	349	169	152	2330	2330
6	3910	863	548	341	5662	5662
$m'_{.j}$	22846	5263	3495	2235	33839	
$m_{.j}$	22877	5285	3462	2213		33837

TABLE IV

A continuation of the process initiated in Table III. The figures in Table III are now adjusted proportionately by columns according to eq. (53). The vertical totals $m''_{.j}$ *and* $m_{.j}$ *now are equal, but the agreement of the horizontal totals accomplished in Table III has been slightly disturbed*

	$j = 1$	2	3	4	$m''_{i.}$	$m_{i.}$
$i = 1$	3613	781	550	309	5253	5252
2	1588	401	252	155	2396	2395
3	1608	435	270	119	2432	2432
4	10490	2451	1680	1142	15763	15766
5	1662	350	167	151	2330	2330
6	3915	867	543	338	5663	5662
$m''_{.j}$	22876	5285	3462	2214	33837	
$m_{.j}$	22877	5285	3462	2213		33837

$$(54) \qquad m_{ijk} = n_{ijk}(m_{i..}/n_{i..} + \lambda_{.j.} + \lambda_{..k} - i\text{-av. } \lambda_{.j.} - i\text{-av. } \lambda_{..k}).$$

Equally well it could have been written

$$(55) \qquad m_{ijk} = n_{ijk}(m_{.j.}/n_{.j.} + \lambda_{i..} + \lambda_{..k} - j\text{-av. } \lambda_{i..} - j\text{-av. } \lambda_{..k}),$$

or

(56) $$m_{ijk} = n_{ijk}(m_{..k}/n_{..k} + \lambda_{i..} + \lambda_{.j.} - k\text{-av. } \lambda_{i..} - k\text{-av. } \lambda_{.j.}).$$

Any of these three equations shows why the adjustment (34) is not proportional by slices, and why this case does not break up into r or s or t sets of the three dimensional Case I. As a first approximation it does, as is now clear from these three equations, and by making successive proportionate adjustments we may thus arrive at the least squares values. To go about the work we could first calculate the values of

(57) $$m'_{ijk} = n_{ijk}(m_{i..}/n_{i..})$$

then

(58) $$m''_{ijk} = m'_{ijk}(m_{.j.}/m'_{.j.})$$

TABLE V

The cycle is commenced again. The figures of Table IV are subjected to a proportionate adjustment by rows, according to eq. (52). And since these results turn out to be almost a reproduction of Table IV but with both horizontal and vertical conditions satisfied, they are considered final. The agreement with the m_{ij} in Table I should be noted

	$j = 1$	2	3	4	$m'_{i.}$	$m_{i.}$
$i = 1$	3612	781	550	309	5252	5252
2	1587	401	252	155	2395	2395
3	1608	435	270	119	2432	2432
4	10492	2451	1680	1142	15765	15766
5	1662	350	167	151	2330	2330
6	3914	867	543	338	5662	5662
m'_{ij}	22875	5285	3462	2214	33836	
$m_{.j}$	22877	5285	3462	2213		33837

followed by

(59) $$m'''_{ijk} = m''_{ijk}(m_{..k}/m''_{..k}).$$

These three successive adjustments would constitute a cycle, which would then be repeated in whole or in part until the table becomes rigid with the satisfaction of all three sets of conditions.

6. Simplification when only one cell requires adjustment. On occasions it happens in sampling work that one is especially interested in one particular cell of the universe, and would like to have a result for it in advance before the other cells are adjusted. Sometimes it even happens that the others individually are of no particular concern. In such circumstances one merely places the cell

of interest in one corner of the table by an appropriate interchange of rows and columns, and then compresses the rest of the table into the cells adjacent to it. In the two dimensional Case II one would thus work with a 2×2 table, one corner cell being the one of special interest, the other three being the result of compression. The marginal totals of the row and column belonging to the cell of interest are unaffected. For illustration we may suppose that from the sample shown in Table I we require only m_{61}. We then start with the 2×2 Table VI, which is derived from Table I by compression. Commencing with Table VI, one might first adjust by rows according to eq. (52), then by columns by eq. (53). One cycle of iterative proportions is sufficient, as is seen in Table

TABLE VI

Derived from Table I by compression, the cell $i = 6, j = 1$, requiring adjustment

	$j = 1$	$j = 2 - 4$	$n_{i.}$	$m_{i.}$
$i = 1 - 5$	18965	9250	28215	28175
$i = 6$	3882	1740	5622	5662
$n_{.j}$	22847	10990	33837	
$m_{.j}$	22877	10960		33837

TABLE VII

A proportionate adjustment of Table VI

Rows adjusted by eq. (52)

18938	9237	28175
3910	1752	5662
22848	10989	33837

Columns adjusted by eq. (53)

18962	9213	28175
3915	1747	5662
22877	10960	33837

Conclusion: $m_{61} = 3915$

VII, and the value 3915 found for m_{61} is in good agreement with its value shown in Tables I and V. The scheme of compression provides a quick method of getting out an advance adjustment for a cell of special interest, and the result so obtained will ordinarily be in good agreement with what comes later when and if all the cells are adjusted.

In the three dimensional Cases II, III, V, VI, and VII, one compresses the original table to a $2 \times 2 \times 2$ table, and then uses the method of iterative proportions. (The other cases do not require consideration, since they are proportionate adjustments wherein one is already at liberty to adjust as few or as many cells as he likes without altering the equations or the routine.) The same procedure can be extended to the adjustment of two cells, the only modification

being that in two dimensions we shall compress to a 2×3 or a 3×3 table, depending on whether the two cells do or do not lie in the same row or column. In three dimensions we compress to a $2 \times 2 \times 3$, or a $2 \times 3 \times 3$, or a $3 \times 3 \times 3$ table; the first if the two cells lie in the same i, j, or k tube, the second if they lie in the same slice but not in the same tube, the third if they are in separate slices.

7. Some remarks on the accuracy of an adjustment. A least squares adjustment of sampling results must be regarded as a systematic procedure for obtaining satisfaction of the conditions imposed, and at the same time effecting an improvement of the data in the sense of obtaining results of smaller variance than the sample itself, under ideal conditions of sampling from a stable universe. It must not be supposed that any or all of the adjusted m_{ij} in any table are necessarily "closer to the truth" than the corresponding sampling frequencies n_{ij}, even under ideal conditions. As for the standard errors of the adjusted results, they can easily be estimated for the ideal case by making use of the calculated chi-square. For predictive purposes, however (which can be regarded as the only possible use of a census by any method, sample or complete), it is far preferable, in fact necessary, to get some idea of the errors of sampling by actual trial, such as by a comparison of the sampling results with the universe, as can often be arranged by means of controls. There is another aspect to the problem of error—even a 100 per cent count, even though strictly accurate, is not by itself useful for prediction, except so far as we can assert on other grounds what secular changes are taking place.

In conclusion it is a pleasure to record our appreciation of the assistance of Miss Irma D. Friedman and Mr. Wilson H. Grabill for putting the formulas and procedure into actual operation with census data, and thereby disclosing defects in earlier drafts of the manuscript.

BUREAU OF THE CENSUS,
WASHINGTON