

WOJCIECH NIEMIRO and JACEK WESOŁOWSKI (Warszawa)

FIXED PRECISION OPTIMAL ALLOCATION IN TWO-STAGE SAMPLING

Abstract. Two-stage sampling schemes arise in survey sampling, especially in situations when the complete update of the frame is difficult. In this paper we solve the problem of fixed precision optimal allocation in two special two-stage sampling schemes. The solution is based on reducing the original question to an eigenvalue problem and then using the Perron–Frobenius theorem.

1. Introduction. We consider a finite population consisting of some number of subpopulations. Subpopulations themselves are split into clusters. We want to estimate total values of a variable in the subpopulations. Assume that the (expected) length of the whole sample is fixed. Our aim is to allocate the sample in subpopulations and clusters in such a way that relative errors of the estimates in all subpopulations are the same and as low as possible. Such a problem falls in the following general scheme. Let z_i be a vector describing the allocation inside the i th subpopulation and $V_i(z_i)$ be the respective relative error. We want to minimize the common value, say $F = V_i(z_i)$ (for all i 's), under the constraints $\sum_i g_i(z_i) = c$, where the g_i 's are given functions and c is a constant. Such a problem can be solved in two steps: 1) For each i separately, minimize $V_i(z_i)$ under the constraints $g_i(z_i) = c_i$. Let $V_i^*(c_i)$ be the solution. 2) Minimize $F = V_i^*(c_i)$ under the constraints $\sum_i c_i = c$. Section 2 describes the first phase of the solution while the second phase is examined in Section 3. In Section 4 we present a numerical example. It is related to a survey of small enterprises, led regularly in Poland by Główny Urząd Statystyczny (Central Statistical Office). Indeed, our work was motivated by the need to improve the sampling strategy in this survey.

2000 *Mathematics Subject Classification*: Primary 62D05.

Key words and phrases: survey sampling, estimator of total variance, subpopulation, minimization with constraints.

2. Optimal allocation in two-stage sampling. In this section we restrict our considerations to a fixed, finite subpopulation $U = \{1, \dots, N\}$ partitioned into clusters: $U = U_1 \cup \dots \cup U_M$. The objective of a survey is the total value t of a variable \mathcal{Y} , i.e. $t = \sum_{u \in U} y_u$, where y_u denotes the value of \mathcal{Y} for unit $u \in U$. Suppose we are to use a two-stage sampling scheme. At the first stage we choose a random sample $V = \{U_1, \dots, U_M\}$ from the population of clusters. The clusters are therefore *primary sampling units* (PSUs). At the second stage we draw random samples from the previously chosen clusters. *Secondary sampling units* (SSUs) are elements of U .

Throughout this paper, simple random sampling (SRS) means sampling without replacement, with equal probabilities. We consider the following two sampling schemes:

1. stratified SRS (SSRS) at the first stage followed by SRS at the second stage;
2. SRS at the first stage followed by SSRS at the second stage.

In both cases, we seek an optimal allocation subject to constraints on the (expected) size of the final sample and (expected) size of the primary sample. The latter can be expressed either in terms of the number of PSUs or the number of SSUs.

2.1. SSRS–SRS scheme. The population of PSUs is partitioned into strata: $V = W_1 \cup \dots \cup W_H$. In what follows, the subscript “ h, j, u ” will refer to the u th SSU in the j th PSU in the h th stratum (we just rearrange objects to make notation readable). Let $\#(W_h) = M_h$ be the number of PSUs in the h th stratum and let m_h be the number of PSUs sampled from this stratum ($h = 1, \dots, H$). Write $U_{h,j}$ for the j th PSU in the h th stratum and let $\#(U_{h,j}) = N_{h,j}$ ($j = 1, \dots, M_h$). If a PSU $U_{h,j}$ belongs to the primary sample, then we draw a sample of $n_{h,j}$ SSUs from it at the second stage. Write also $N_h = \sum_{j \in W_h} N_{h,j}$.

Let \hat{t} be the usual π -estimator (Horvitz–Thompson estimator) of the total t . According to the well-known general formula [cf. Särndal, Svensson and Wretman (1992), Ch. 2], the variance of this estimator is given by

$$(1) \quad D^2(\hat{t}) = \sum_{h=1}^H \left(\frac{1}{m_h} - \frac{1}{M_h} \right) M_h^2 D_h^2 + \sum_{h=1}^H \frac{M_h}{m_h} \sum_{j=1}^{M_h} \left(\frac{1}{n_{h,j}} - \frac{1}{N_{h,j}} \right) N_{h,j}^2 S_{h,j}^2,$$

where $t_{h,j} = \sum_{u \in U_{h,j}} y_{h,j,u}$, $\bar{y}_{h,j} = t_{h,j}/N_{h,j}$ and $\bar{t}_h = \sum_{j=1}^{M_h} t_{h,j}/M_h$,

$$S_{h,j}^2 = \frac{1}{N_{h,j} - 1} \sum_{u \in U_{h,j}} (y_{h,j,u} - \bar{y}_{h,j})^2, \quad D_h^2 = \frac{1}{M_h - 1} \sum_{j=1}^{M_h} (t_{h,j} - \bar{t}_h)^2.$$

The problem is to minimize $D^2(\hat{t})$ as a function of the variables m_h and $n_{h,j}$, given that the expected number of SSUs in the primary sample is equal to n_I and the expected number of SSUs in the final sample is equal to n . Therefore, the constraints are the following:

$$(2) \quad \sum_{h=1}^H \frac{m_h}{M_h} \sum_{j=1}^{M_h} N_{h,j} = n_I, \quad \sum_{h=1}^H \frac{m_h}{M_h} \sum_{j=1}^{M_h} n_{h,j} = n.$$

Let **P1** denote the problem described above: minimize (1) subject to (2). Using the standard method of Lagrange multipliers, we easily obtain the following result.

PROPOSITION 1. *Provided that $\gamma_h = M_h D_h^2 - \sum_{j=1}^{M_h} N_{h,j} S_{h,j}^2 > 0$ for all h , the solution of **P1** has the following form: the optimal allocation is*

$$m_h = \frac{M_h}{N_h} \cdot \frac{\sqrt{N_h \gamma_h}}{\sum_{k=1}^H \sqrt{N_k \gamma_k}} n_I,$$

and

$$n_{h,j} = n \frac{M_h}{m_h} \cdot \frac{N_{h,j} S_{h,j}}{\sum_{k=1}^H \sum_{l=1}^{M_k} N_{k,l} S_{k,l}}.$$

The minimum value of the variance (1) is

$$D_{\text{opt}}^2(\hat{t}) = \frac{1}{n_I} \left(\sum_{h=1}^H \sqrt{N_h \gamma_h} \right)^2 + \frac{1}{n} \left(\sum_{h=1}^H \sum_{j=1}^{M_h} N_{h,j} S_{h,j} \right)^2 - \sum_{h=1}^H M_h D_h^2.$$

A slightly different problem is obtained if we replace the first of the constraints in (2) by a constraint on the expected size of the primary sample expressed in terms of PSUs. Put differently, instead of (2), we consider the following system of constraints:

$$(3) \quad \sum_{h=1}^H m_h = m, \quad \sum_{h=1}^H \frac{m_h}{M_h} \sum_{j=1}^{M_h} n_{h,j} = n.$$

Let **P2** denote the problem: minimize (1) subject to (3).

PROPOSITION 2. *Provided that $\gamma_h > 0$ for all h , the solution of **P2** has the following form: the optimal allocation of PSUs is*

$$m_h = \frac{\sqrt{M_h \gamma_h}}{\sum_{k=1}^H \sqrt{M_k \gamma_k}} m.$$

The formula for optimal $n_{h,j}$ is the same as in Proposition 1. The minimum value of the variance (1) is

$$D_{\text{opt}}^2(\hat{t}) = \frac{1}{m} \left(\sum_{h=1}^H \sqrt{M_h \gamma_h} \right)^2 + \frac{1}{n} \left(\sum_{h=1}^H \sum_{j=1}^{M_h} N_{h,j} S_{h,j} \right)^2 - \sum_{h=1}^H M_h D_h^2.$$

2.2. SRS-SSRS scheme. Now assume that at the first stage PSUs are sampled according to the SRS scheme. Let M be the number of PSUs in the population V and assume the number of PSUs in the primary sample is m . Each of the PSUs (clusters) is partitioned into H strata. Let $N_{j,h}$ stand for the number of SSUs in the h th stratum of the j th cluster ($h = 1, \dots, H_j$, $j = 1, \dots, M$; note that we need a different labeling of units than in the previous subsection). We apply stratified SRS at the second stage, sampling $n_{j,h}$ units from the h th stratum of the j th cluster. The general formula for the variance of the π -estimator \hat{t} of the total t now assumes the following form:

$$(4) \quad D^2(\hat{t}) = \left(\frac{1}{m} - \frac{1}{M} \right) M^2 D^2 + \frac{M}{m} \sum_{j=1}^M \sum_{h=1}^{H_j} \left(\frac{1}{n_{j,h}} - \frac{1}{N_{j,h}} \right) N_{j,h}^2 S_{j,h}^2,$$

where $S_{j,h}^2$ is the variance of the variable \mathcal{Y} in the h th stratum of the j th PSU and

$$D^2 = \frac{1}{M-1} \sum_{j=1}^M (t_j - \bar{t})^2$$

denotes the variance of totals of \mathcal{Y} in PSUs (it is a variable defined in the population V of PSUs). Consider the following constraint on the expected sample size:

$$(5) \quad \frac{m}{M} \sum_{j=1}^M \sum_{h=1}^{H_j} n_{j,h} = n.$$

Denote by **P3** the problem of minimizing (4) under constraint (5).

PROPOSITION 3. *The allocation problem P3 has the Neyman-type solution*

$$n_{j,h} = n \frac{M}{m} \cdot \frac{N_{j,h} S_{j,h}}{\sum_{l=1}^M \sum_{k=1}^{H_l} N_{l,k} S_{l,k}}.$$

The optimal variance is

$$D_{\text{opt}}^2(\hat{t}) = \frac{1}{m} M \gamma + \frac{1}{n} \left(\sum_{j=1}^M \sum_{h=1}^{H_j} N_{j,h} S_{j,h} \right)^2 - M D^2,$$

where

$$\gamma = MD^2 - \sum_{j=1}^M \sum_{h=1}^{H_j} N_{j,h} S_{j,h}^2.$$

If instead of fixing the sample size of PSUs, we fix the expected number of SSUs in the primary sample, i.e.

$$(6) \quad \frac{m}{M}N = n_I,$$

then the formulas for optimal allocation take slightly modified forms. Denote by **P4** the problem of minimizing (4) under constraints (5) and (6), given n_I and n .

PROPOSITION 4. *The allocation problem **P4** has the following solution:*

$$n_{j,h} = n \frac{n_I}{N} \cdot \frac{N_{j,h} S_{j,h}}{\sum_{l=1}^M \sum_{k=1}^{H_l} N_{l,k} S_{l,k}}.$$

The optimal variance is

$$D_{\text{opt}}^2(\hat{t}) = \frac{1}{n_I} N \gamma + \frac{1}{n} \left(\sum_{j=1}^M \sum_{h=1}^{H_j} N_{j,h} S_{j,h} \right)^2 - MD^2.$$

Let us sum up the considerations of this section. Observe that the formulas for the optimal variance for problems **P1**–**P4** (see Propositions 1–4) are of the same shape. In what follows we will be interested in the precision of the estimator \hat{t} , defined as $2\sqrt{D^2(\hat{t})}/t$. Of course, the optimal precision corresponds to the minimal variance. We can express the squared relative error in the following form:

$$(7) \quad D_{\text{opt}}^2(\hat{t})/t^2 = V(x, y) = \frac{A}{x} + \frac{B}{y} - C,$$

where x and y are allocation parameters ($x = n_I$ and $y = n$ for **P1**, **P4** and $x = m$ and $y = n$ for **P2**, **P3**) while A , B and C are population parameters.

3. Equating the precision in subpopulations. Now consider a population U partitioned into subpopulations (domains) D_i , $i = 1, \dots, L$. We turn to the second stage of our main problem, as described in the introduction. This means that we seek an allocation of the sample into subpopulations which gives equal and optimal precisions. Now, we apply formula (7) separately for each of the domains, arriving at

$$D_{\text{opt}}^2(\hat{t}_i)/t_i^2 = V_i(x_i, y_i) = \frac{A_i}{x_i} + \frac{B_i}{y_i} - C_i,$$

where the additional subscript i indicates the subpopulation D_i . In particular, t_i is the total value of \mathcal{Y} in the domain D_i and \hat{t}_i is its estimator. Let

us recall that x_i and y_i are *allocation* parameters, while A_i , B_i and C_i are now *subpopulation* parameters.

3.1. Main result. Our main problem is to find the smallest common value F of the squared relative error $V_i(x_i, y_i)$ in the domains D_i , $i = 1, \dots, L$. The solution is based on the following theorem.

THEOREM 1. *Let x , y and A_i , B_i , C_i , $i = 1, \dots, L$, be given positive numbers. Assume that $A_i/x + B_i/y - C_i > 0$ for all i . Put $a_i = \sqrt{A_i}$ and $b_i = \sqrt{B_i}$. The smallest $F > 0$ such that*

$$(8) \quad \frac{A_i}{x_i} + \frac{B_i}{y_i} - C_i = F \quad (\forall i = 1, \dots, L),$$

under the constraints $x_i, y_i > 0$, $i = 1, \dots, L$, and

$$(9) \quad \sum_{i=1}^L x_i = x, \quad \sum_{i=1}^L y_i = y,$$

is the largest eigenvalue of the matrix

$$\frac{1}{x} \mathbf{a} \mathbf{a}^T + \frac{1}{y} \mathbf{b} \mathbf{b}^T - \text{diag}(\mathbf{C}),$$

where $\mathbf{a} = (a_1, \dots, a_L)^T$, $\mathbf{b} = (b_1, \dots, b_L)^T$, $\mathbf{C} = (C_1, \dots, C_L)^T$. This eigenvalue is simple and positive. A corresponding eigenvector $\mathbf{v} = (v_1, \dots, v_L)^T$ has all components positive. The values of x_i 's and y_i 's corresponding to the minimum F are given by

$$x_i = \frac{a_i v_i}{\sum_{r=1}^L a_r v_r} x, \quad y_i = \frac{b_i v_i}{\sum_{r=1}^L b_r v_r} y.$$

Proof. Consider the function $f : (0, \infty)^{2L+1} \rightarrow \mathbb{R}$ defined by

$$f(x_1, \dots, x_L, y_1, \dots, y_L, F) = F.$$

We seek the minimum of f on the set

$$\mathcal{R} = \{(x_1, \dots, x_L, y_1, \dots, y_L, F) \in (0, \infty)^{2L+1} : (8) \text{ and } (9) \text{ hold}\}.$$

It is clear that f is bounded from below on this set, because $A_i/x_i \geq A_i/x$ and $B_i/y_i \geq B_i/y$. Let f^* be the infimum of f over \mathcal{R} . Notice that $\{(x_1, \dots, x_L, y_1, \dots, y_L, F) \in \mathcal{R} : F \leq f^* + 1\}$ is a compact set. Indeed, it is a (relatively) closed and bounded subset of $(0, \infty)^{2L+1}$. Moreover, this set is bounded away from the boundary of $(0, \infty)^{2L+1}$, because $x_i \rightarrow 0$ or $y_i \rightarrow 0$ together with (8) implies $F \rightarrow \infty$. Therefore, there exists a point $(x_1^*, \dots, x_L^*, y_1^*, \dots, y_L^*, F^*) \in \mathcal{R}$ such that $F^* = f^*$. This is a stationary point of f , under constraints (8) and (9). Using the method of Lagrange multipliers, we will find it and then show it is unique.

Consider the Lagrange function

$$(10) \quad F + \sum_{i=1}^L \lambda_i \left(\frac{A_i}{x_i} + \frac{B_i}{y_i} - C_i - F \right) + \alpha \left(\sum_{i=1}^L x_i - x \right) + \beta \left(\sum_{i=1}^L y_i - y \right),$$

where α , β and λ_i are the Lagrange multipliers. Differentiating (10) with respect to x_i and y_i , we infer that $\alpha(x_i^*)^2 = \lambda_i A_i$ and similarly $\beta(y_i^*)^2 = \lambda_i B_i$. Hence all the Lagrange multipliers α , β and λ_i have the same sign. Without loss of generality we can assume that they are all positive. Let us write $v_i = \sqrt{\lambda_i}$ and recall that $A_i = a_i^2$, $B_i = b_i^2$. We arrive at

$$x_i^* = \frac{a_i v_i}{\sqrt{\alpha}}, \quad y_i^* = \frac{b_i v_i}{\sqrt{\beta}}.$$

Use constraints (9) to get $\sqrt{\alpha} = \sum_{r=1}^L a_r v_r / x$ and $\sqrt{\beta} = \sum_{r=1}^L b_r v_r / y$. Now, plug in the expressions for x_i^* and y_i^* into (8). We obtain

$$\frac{a_i}{x v_i} \sum_{r=1}^L a_r v_r + \frac{b_i}{y v_i} \sum_{r=1}^L b_r v_r - C_i = F^*$$

or, equivalently,

$$\sum_{r=1}^L \left(\frac{a_i a_r}{x} + \frac{b_i b_r}{y} \right) v_r - C_i v_i = F^* v_i.$$

In matrix notation, this is just

$$\left(\frac{1}{x} \mathbf{a} \mathbf{a}^T + \frac{1}{y} \mathbf{b} \mathbf{b}^T - \text{diag}(\mathbf{C}) \right) \mathbf{v} = F^* \mathbf{v},$$

where $\mathbf{v} = (v_1, \dots, v_L)^T$. Under our assumptions, all elements of the symmetric matrix $\mathbf{M} = \mathbf{a} \mathbf{a}^T / x + \mathbf{b} \mathbf{b}^T / y - \text{diag}(\mathbf{C})$ are positive. By the Perron–Frobenius theorem, the largest eigenvalue of \mathbf{M} is simple and it is *the only* eigenvalue with a positive corresponding eigenvector [cf. Kato (1981), Th. 7.3 in Ch. 1]. Hence it must be F^* and the corresponding eigenvector is \mathbf{v} . The conclusions of the theorem follow. ■

3.2. Corollaries. Now, let us specify the conclusions of Theorem 1 to the four problems considered in Section 2. We only have to express the parameters a_i , b_i and C_i using the formulas for the optimum variance for **P1–P4**, respectively. The formulas are given in Propositions 1–4. Note that now the additional subscript i , referring to the subpopulation, has to appear in each of these formulas ($i = 1, \dots, L$).

For **P1** we have

$$a_i = \sum_{h=1}^{H_i} \sqrt{N_{i,h} \gamma_{i,h} / t_i}, \quad b_i = \sum_{h=1}^{H_i} \sum_{j=1}^{M_{i,h}} N_{i,h,j} S_{i,h,j} / t_i,$$

$$C_i = \sum_{h=1}^{H_i} M_{i,h} D_{i,h}^2 / t_i^2, \quad x = n_I, \quad y = n.$$

Consequently, the solution of our fixed precision optimal allocation problem is given by

$$(11) \quad \begin{aligned} m_{i,h} &= x_i \frac{M_{i,h}}{N_{i,h}} \cdot \frac{\sqrt{N_{i,h} \gamma_{i,h}}}{\sum_{k=1}^{H_i} \sqrt{N_{i,k} \gamma_{i,k}}}, \\ n_{i,h,j} &= y_i \frac{M_{i,h}}{m_{i,h}} \cdot \frac{N_{i,h,j} S_{i,h,j}}{\sum_{k=1}^{H_i} \sum_{l=1}^{M_{i,k}} N_{i,k,l} S_{i,k,l}}. \end{aligned}$$

For **P2** we have b_i , C_i and y the same as above,

$$a_i = \sum_{h=1}^{H_i} \sqrt{M_{i,h} \gamma_{i,h}} / t_i, \quad x = m.$$

The solution of the fixed precision optimal allocation problem is given by

$$m_{i,h} = x_i \frac{\sqrt{M_{i,h} \gamma_{i,h}}}{\sum_{k=1}^{H_i} \sqrt{M_{i,k} \gamma_{i,k}}}$$

and $n_{i,h,j}$ as in (11).

For **P3** we have

$$\begin{aligned} a_i &= \sqrt{M_i \gamma_i} / t_i, \quad b_i = \sum_{j=1}^{M_i} \sum_{h=1}^{H_{i,j}} N_{i,j,h} S_{i,j,h} / t_i, \\ C_i &= M_i D_i^2 / t_i^2, \quad x = m, \quad y = n. \end{aligned}$$

The solution of the fixed precision optimal allocation problem is given by

$$(12) \quad m_i = x_i, \quad n_{i,j,h} = y_i \frac{M_i}{m_i} \cdot \frac{N_{i,j,h} S_{i,j,h}}{\sum_{l=1}^{M_i} \sum_{k=1}^{H_{i,l}} N_{i,l,k} S_{i,l,k}}.$$

For **P4** we have b_i , C_i and y as above,

$$a_i = \sqrt{N_i \gamma_i}, \quad x = n_I.$$

The solution of the fixed precision optimal allocation problem is given by

$$m_i = x_i \frac{M_i}{N_i}$$

and $n_{i,j,h}$ as in (12).

4. Numerical example. We consider a population of small (up to 5 employees) enterprises in a southern part of Poland of size 79578, divided into 4 subpopulations, say A, B, C, and D, of respective sizes 15144, 11582, 26960, 25892. Each of the subpopulations is clustered into regions. There are 170 regions in A, 128 in B, 306 in C and 326 in D, thus the total number

of regions equals 930. In each of the subpopulations A, B, C and D, the regions are stratified into 8, 7, 8 and 8 strata, respectively. The variable of interest is the total employment in subpopulations. We seek a fixed precision optimal allocation of the sample using the double stage SSRS–SRS scheme where PSUs are regions and SSUs are enterprises. We impose the following constraints: the total size of the primary unit sample is 10000 SSUs and the total size of the final sample equals 2000 SSUs.

Thus we are in the framework of the problem **P1** discussed above with $n_I = 10000$, $n = 2000$ and $L = 4$. First it was checked that all γ_h 's are positive. Using the definitions of the vectors (4-variate in our setting) **a**, **b** and **C** it was found that

$$\begin{aligned}\mathbf{a} &= (2.75441, 2.50996, 2.59874, 2.57227)^T, \\ \mathbf{b} &= (0.62892, 0.60228, 0.61864, 0.60082)^T, \\ \mathbf{C} &= (0.00063, 0.00059, 0.00028, 0.00028)^T.\end{aligned}$$

The eigenvalue problem has the following solutions (λ_i denotes the eigenvalue and \mathbf{v}_i a respective eigenvector):

$$\begin{aligned}\lambda_1 &= 0.00304, & \mathbf{v}_1 &= (0.49649, 0.46302, 0.52303, 0.51531)^T, \\ \lambda_2 &= 0.00028, & \mathbf{v}_2 &= (0.00203, 0.00183, -0.70362, 0.71057)^T, \\ \lambda_3 &= 0.00045, & \mathbf{v}_3 &= (0.46661, 0.57090, -0.47862, -0.47674)^T, \\ \lambda_4 &= 0.00061, & \mathbf{v}_4 &= (-0.73197, 0.67800, 0.04772, 0.04759)^T.\end{aligned}$$

Now using the solution given in the preceding section we notice that $F^* = \lambda_1$, consequently the precision of the estimate equals $2\sqrt{F^*} = 0.1102$. The allocation of the primary sample was computed according to the first part of the formula (11), in terms of the eigenvector \mathbf{v}_1 (and suitably rounded). For clarity, we present it in the form of pairs $(M_{i,h}, m_{i,h})$:

- for the subpopulation A:

$$(14, 2), (22, 3), (2, 1), (29, 3), (11, 2), (71, 13), (3, 1), (18, 4),$$

- for the subpopulation B:

$$(14, 3), (27, 5), (31, 6), (3, 1), (41, 10), (2, 1), (10, 2),$$

- for the subpopulation C:

$$(22, 2), (46, 5), (3, 1), (62, 5), (13, 1), (121, 15), (5, 1), (34, 3),$$

- for the subpopulation D:

$$(25, 2), (38, 3), (2, 1), (70, 6), (6, 1), (154, 18), (6, 1), (25, 3).$$

Similarly the allocation of the final sample of SSUs was obtained using the second part of (11) (and again rounded). Since this is a lengthy table,

we present here the pairs $(N_{i,h,j}, n_{i,h,j})$ only for stratum number 8 of the subpopulation A:

(127, 21), (126, 13), (80, 2), (11, 17), (107, 31), (114, 17), (123, 18), (87, 6),
 (112, 18), (139, 17), (113, 12), (127, 10), (87, 9), (119, 13), (139, 25),
 (119, 9), (226, 33), (145, 17).

References

- C. E. Särndal, B. Swensson and J. Wretman (1992), *Model Assisted Survey Sampling*, Springer, Berlin.
 T. Kato (1981), *A Short Introduction to Perturbation Theory for Linear Operators*, Springer, New York.

Wojciech Niemirow
 Institute of Applied Mathematics
 and Mechanics
 Warsaw University
 Banacha 2
 02-097 Warszawa, Poland
 E-mail: wniem@mimuw.edu.pl

Central Statistical Office
 Al. Niepodległości 208
 00-925 Warszawa, Poland

Jacek Wesołowski
 Faculty of Mathematics and
 Information Science
 Warsaw University of Technology
 Pl. Politechniki 1
 00-661 Warszawa, Poland

Central Statistical Office
 Al. Niepodległości 208
 00-925 Warszawa, Poland

*Received on 24.2.2000;
 revised version on 29.6.2000*

(1529)