AN EIGENPROBLEM APPROACH TO OPTIMAL EQUAL-PRECISION SAMPLE ALLOCATION IN SUBPOPULATIONS

Jacek Wesołowski

Central Statistical Office and Warsaw University of Technology, Warszawa, POLAND

BaNoCoSS IV
August, 24-28, 2015
Helsinki, FINLAND

with Robert Wieczorkowski, CSO Warsaw, POLAND

Plan

General problem

Let p be a sampling design in a population U, which is divided in disjoint subpopulations

$$U = \bigcup_{i=1}^{I} U_i.$$

The sample S drawn according to p is divided into

$$S = S_1 \cup \dots S_I,$$

where $S_i \subset U_i$, $i = 1, \dots, I$.
We want to estimate total τ_i of a variable $\mathcal{Y}$ in the subpopulation U_i using, say, an unbiased estimator $\hat{t}_i$ based on S_i for each $i = 1, \dots, I$.
The aim is to allocate the sample in such a way that e.g. standard errors (cv's) of $\hat{t}_i(\mathcal{Y})$, $i = 1, \dots, I$, are minimal and comparable (equal).

A toy example

Assume that

- p consists of independent *SRS* plans in U_is,
- the total size of the sample is n,
- $\hat{t}_i = N_i \bar{y}_{S_i}$, $i = 1, \ldots$.

Then cv^2 of $\hat{t}_i$ is

$$T_i = \left(\frac{1}{n_i} - \frac{1}{N_i}\right)\gamma_i^2,$$

where $N_i = \#(U_i)$, $n_i = \#(S_i)$, and γ_i is the cv of $\mathcal{Y}$ in the subpopulation U_i, $i = 1, \ldots, I$.
Thus

$$n_i = \frac{N_i \gamma_i^2}{\gamma_i^2 + T_i N_i}$$

A toy example, cont.

We postulate constancy of $\mathrm{cv}^2(\hat{t}_i)$:

$$T_i =: T, \qquad i = 1, \ldots, I,$$

together with the constraint

$$n = n_1 + \ldots + n_I$$

which give

$$n = \sum_{i=1}^{I} \frac{N_i \gamma_i^2}{\gamma_i^2 + T N_i}.$$

It is elementary to see that the equation has a unique solution T^*, which can be easily obtained numerically, nevertheless no analytic formula is available. Thus the allocation is

$$n_i = \frac{N_i \gamma_i^2}{\gamma_i^2 + T^* N_i}, \quad i = 1, \ldots, I.$$

A toy example - an alternative approach

Define $v_i = \frac{n_i}{\gamma_i}\lambda$, $i = 1, \ldots, I$. Then $\lambda = \frac{1}{n}\sum_{j=1}^{I} v_j\gamma_j$. Moreover,

$$\lambda\frac{\gamma_i}{v_i} - \frac{\gamma_i^2}{N_i} = T, \quad i = 1, \ldots, I.$$

That is

$$\frac{\gamma_i}{\sqrt{n}}\left(\sum_{j=1}^{I} \frac{\gamma_j}{\sqrt{n}} v_j\right) - \frac{\gamma_i^2}{N_i} v_i = T v_i, \quad i = 1, \ldots, I.$$

Finally, for the vectors $\underline{g} = \frac{1}{\sqrt{n}}\left(\frac{\gamma_i}{\sqrt{n}}, i = 1, \ldots, I\right)^T$ and $\underline{c} = \left(\frac{\gamma_i^2}{N_i}, i = 1, \ldots, I\right)^T$ we get

$$(\underline{g}\underline{g}^T - \mathrm{diag}\underline{c})\underline{v} = T\underline{v},$$

where $\underline{v} = (v_i, i = 1, \ldots, I)^T$.

A toy example - an alternative approach, cont.

Conclusion: $\underline{v}$ is an eigenvector of the matrix $\mathbf{D} = \underline{g}\underline{g}^T - \mathrm{diag}\,\underline{c}$ and T is its respective eigenvalue.
Since

$$\sum_{i=1}^{I} \frac{g_i^2}{c_i} = \frac{N}{n} > 1$$

it follows that there is a unique simple positive eigenvalue $\lambda =: T^*$. Then it follows from the Perron-Frobenius Theorem that the respective eigenvector $\underline{v}^*$ has all components which are positive (up to a multiplication).
Finally the allocation is:

$$n_i = n \frac{v_i^* \gamma_i}{\sum_{j=1}^{I} v_j^* \gamma_j}, \quad i = 1, \ldots, I.$$

Remarks

- if T_i, $i = 1, \ldots, I$, are fixed, then the total sample size n is uniquely determined;
- if T_i, $i = 1, \ldots, I$, are just bounded above by some fixed numbers, minimization of n is a valid issue;
  - solved for *SSRS* in Choudhry, Rao and Hidiroglou (2012) through non-linear programming (Newton-Rawson procedure);
  - alternative numerical method for multivariate variables (together with strata construction) given in Lednicki and Wieczorkowski (2003) uses Nelder-Mead simplex algorithm;
- for two-stage ($SSRS_I$, SRS_{II}) sampling when no subpopulations are considered and a single constraint is given through expected total cost standard methods go back to Cochran's book;
- we extend the approach from Niemiro and Wesołowski (2001), where common precision allocation for either ($SSRS_I$, SRS_{II}) or (SRS_I, $SSRS_{II}$) were considered.

General minimization scheme

We consider such sampling plans (with stratification on the first stage) and estimators for which $\mathrm{cv}^2(\hat{t}_i)$ in U_i is of the form

$$T_i = \sum_{h=1}^{H_i} \frac{1}{x_{i,h}} \left(A_{i,h} + \sum_{j \in \mathcal{V}_{i,h}} \frac{B_{i,h,j}}{z_{i,h,j}} \right) - c_i \qquad (1)$$

where $\mathcal{V}_{i,h}$ is (a set of indices) related to $\mathcal{W}_{i,h}$ (the hth strata of PSUs in U_i) and the constraints are

$$\sum_{i=1}^{I} \sum_{h=1}^{H_i} x_{i,h} = x \qquad (2)$$

and

$$\sum_{i=1}^{I} \sum_{h=1}^{H_i} x_{i,h} \sum_{j \in \mathcal{V}_{i,h}} \alpha_{i,h,j} z_{i,h,j} = y \qquad (3)$$

where $A_{i,h} > 0$, $B_{i,h,j} \geq 0$, $c_i > 0$ and $\alpha_{i,h,j} > 0$ are some given population quantities and the constraints $x > 0$ and $y > 0$ are also known.

Objects of interest

Define vectors

$$\underline{a} = \frac{1}{\sqrt{x}} \left(\sum_{h=1}^{H_i} \sqrt{A_{j,h}},\ i = 1, \ldots, I \right),$$

$$\underline{b} = \frac{1}{\sqrt{z}} \left(\sum_{h=1}^{H_i} \sum_{j \in \mathcal{V}_{i,h}} \sqrt{\alpha_{i,h,j} B_{i,h,j}},\ i = 1, \ldots, I \right),$$

and

$$\underline{c} = (c_i,\ i = 1, \ldots, I)$$

and a matrix

$$\mathbf{D} = \underline{a}\underline{a}^T + \underline{b}\underline{b}^T - \mathrm{diag}(\underline{c}).$$

Minimization

Theorem. *Assume that* $\mathbf{D}$ *has a unique simple eigenvalue* $\lambda > 0$ *and* $\mathbf{D}\underline{v} = \lambda\,\underline{v}$.

Then the solution of the problem of minimization of T, *under* $I + 2$ *constraints:* $T = T_i$, $i = 1, \ldots, I$, *as given in* (1), (2) *and* (3) *is*

$$x_{i,h} = x \frac{v_i \sqrt{A_{i,h}}}{\sum_{k=1}^{I} v_k \sum_{g=1}^{H_k} \sqrt{A_{k,g}}}, \tag{4}$$

$$z_{i,h,j} = \frac{z}{x_{i,h}} \frac{v_i \sqrt{\frac{B_{i,h,j}}{\alpha_{i,h,j}}}}{\sum_{k=1}^{I} v_k \sum_{g=1}^{H_k} \sum_{r \in \mathcal{V}_{k,g}} \sqrt{\alpha_{k,g,r} B_{k,g,r}}}. \tag{5}$$

and

$$T = \lambda.$$

A sketch of proof

Consider the Lagrange function

$$F(T,\underline{x},\underline{z}) = T + \sum_{i=1}^{I} \rho_i \left(\sum_{h=1}^{H_i} \frac{1}{x_{i,h}} \left(A_{i,h} + \sum_{j\in\mathcal{V}_{i,h}} \frac{B_{i,h,j}}{z_{i,h,j}} \right) - c_i - T \right)$$

$$+ \sum_{i=1}^{I} \sum_{h=1}^{H_i} x_{i,h} \left(\mu + \nu \sum_{j\in\mathcal{V}_{i,h}} \alpha_{i,h,j} z_{i,h,j} \right).$$

Then the equations

$$\frac{\partial F}{\partial x_{i,h}} = 0 \quad \text{and} \qquad \frac{\partial F}{\partial z_{i,h,j}} = 0$$

together with the constraints after some algebra allow to write

$$\sqrt{\frac{\mu}{\rho_i}} \sum_{h=1}^{H_i} \sqrt{A_{i,h}} + \sqrt{\frac{\nu}{\rho_i}} \sum_{h=1}^{H_i} \sum_{j\in\mathcal{V}_{i,h}} \sqrt{\alpha_{i,h,j} B_{i,h,j}} - c_i = T.$$

The above can be reduced to $\mathbf{D}\underline{v} = T\underline{v}$ with $v_i = \sqrt{\rho_i}$, $i = 1, \ldots, I$. $\square$

Priority adjustments

In practice it is often important to give different (known) priority weights $\kappa_i > 0$ to $\mathrm{cv}^2(\hat{t}_i)$. Then, instead of $T_i = T$, we have $T_i = \kappa_i T$, and thus

$$\kappa_i T = \sum_{h=1}^{H_i} \frac{1}{x_{i,h}} \left(A_{i,h} + \sum_{j \in \mathcal{V}_{i,h}} \frac{B_{i,h,j}}{z_{i,h,j}} \right) - c_i.$$

Minimization problem in such a case is easily reduced to the previous one by proper adjustments: $A_{i,h} \to A_{i,h}/\kappa_i$, $B_{i,h,j} \to B_{i,h,j}/\kappa_i$ and $c_i = c_i/\kappa_i$. That is, the vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$ are modified and consequently also the matrix **D**.

The solution is read out from Theorem with modified **D** and modified $A_{i,h}$, $B_{i,h,j}$ and c_i.

$(SSRS_I, SSRS_{II})$ scheme

For any subpopulation $i = 1, \ldots, I$ the variance of $\hat{t}_i$ is

$$D^2(\hat{t}_i) = \sum_{h=1}^{H_i} \left(\frac{1}{m_{i,h}} - \frac{1}{M_{i,h}} \right) M_{i,h}^2 D_{i,h}^2$$

$$+ \sum_{h=1}^{H_i} \frac{M_{i,h}}{m_{i,h}} \sum_{j \in \mathcal{W}_{i,h}} \sum_{g=1}^{G_{i,h,j}} \left(\frac{1}{n_{i,h,j,g}} - \frac{1}{N_{i,h,j,g}} \right) N_{i,h,j,g}^2 S_{i,h,j,g}^2,$$

where $M_{i,h} = \#\mathcal{W}_{i,h}$ and

$$D_{i,h}^2 = \frac{1}{M_{i,h}-1} \sum_{j \in \mathcal{W}_{i,h}} (t_j - \bar{t}_{i,h})^2.$$

$(SSRS_I, SSRS_{II})$, *cont*. scheme

That is in terms of symbols used in Theorem we have:
$\mathcal{V}_{i,h} = \{(j,g) : j \in \mathcal{W}_{i,h}, \, g \in \{1, \ldots, G_{i,h,j}\}\}$,

$$A_{i,h} = \tfrac{M_{i,h}}{\tau_i^2} \left(M_{i,h} D_{i,h}^2 - \sum_{j \in \mathcal{W}_{i,h}} \sum_{g=1}^{G_{i,h,j}} N_{i,h,j,g} S_{i,h,j,g}^2 \right),$$

$$B_{i,h,(j,g)} = \frac{M_{i,h} N_{i,h,j,g}^2 S_{i,h,j,g}^2}{\tau_i^2}$$

and

$$c_i = \tfrac{1}{\tau_i^2} \sum_{h=1}^{H_i} M_{i,h} D_{i,h}^2.$$

($SSRS_I$, $SSRS_{II}$) scheme, cont.

The sample size constraints have the form

$$\sum_{i=1}^{I}\sum_{h=1}^{H_i} m_{i,h} = m \tag{6}$$

and

$$\sum_{i=1}^{I}\sum_{h=1}^{H_i} \frac{m_{i,h}}{M_{i,h}} \sum_{j\in\mathcal{W}_{i,h}} \sum_{g=1}^{G_{i,h,j}} n_{i,h,j,g} = n, \tag{7}$$

that is $x = m$, $z = n$ and $\alpha_{i,h,(j,g)} = M_{i,h}^{-1}$.

$(SSRS_I, SSRS_{II})$ scheme, cont.

Assume that $A_{i,h} > 0$ and that **D** has a unique simple eigenvalue $\lambda > 0$ with eigenvector $\underline{v}$.

Then the allocation assuring optimal common-precision in subpopulations is

$$m_{i,h} = m \frac{v_i \sqrt{A_{i,h}}}{\sum_{k=1}^{I} v_k \sum_{r=1}^{H_k} \sqrt{A_{k,r}}}$$

and

$$n_{i,h,j,g} = n \frac{v_i \sqrt{B_{i,h,(j,g)} M_{i,h}}}{m_{i,h} \sum_{k=1}^{I} v_k \sum_{r=1}^{H_k} \sum_{l \in \mathcal{W}_{k,r}} \sum_{s=1}^{G_{k,r,l}} \sqrt{B_{k,r,(l,s)}/M_{k,r}}}.$$

$(SSRS_I, SRS_{II})$ scheme with $n_{i,h,j} = n_{i,h}$

Here $\#(\mathcal{V}_{i,h}) = 1$. We keep the notation $M_{i,h}$, $D^2_{i,h}$, c_i as above and above but with

$$A_{i,h} = \frac{M_{i,h}}{\tau_i^2}\left(M_{i,h}D^2_{i,h} - \sum_{j\in\mathcal{W}_{i,h}} N_{i,h,j}S^2_{i,h,j}\right)$$

and

$$B_{i,h} = \frac{M_{i,h}}{\tau_i^2}\sum_{j\in\mathcal{W}_{i,h}} N_{i,h,j}S^2_{i,h,j}.$$

The sample size constraint (6) is the same and (7) changes into

$$\sum_{i=1}^{I}\sum_{h=1}^{H_i} m_{i,h}n_{i,h} = n.$$

That is $x = m$, $z = n$ and $\alpha_{i,h} = 1$.

$(SSRS_I, SRS_{II})$ scheme with $n_{i,h,j} = n_{i,h}$, cont.

Assume that $\gamma_{i,h} > 0$ and that $\mathbf{D}$ has a unique simple eigenvalue $\lambda > 0$ with eigenvector $\underline{v}$.

Then the allocation assuring optimal common-precision in subpopulations is

$$m_{i,h} = m \frac{v_i \sqrt{A_{i,h}}}{\sum_{k=1}^{I} v_k \sum_{r=1}^{H_k} \sqrt{A_{k,h}}}$$

and

$$n_{i,h} = n \frac{v_i \sqrt{B_{i,h}}}{m_{i,h} \sum_{k=1}^{I} v_k \sum_{r=1}^{H_k} \sqrt{B_{k,r}}}.$$

$(SHR_I, SSRS_{II})$

Hartley-Rao (HR) scheme is πps (based on a size variable $\mathcal{Z}$) systematic sampling from randomly ordered list.

SHR_I means that HR scheme is used in each strata of PSUs.

The approximate variance of π-estimator $\hat{t}_i$ has the form

$$\sum_{h=1}^{H_i} \frac{1}{m_{i,h}} \sum_{j\in\mathcal{W}_{i,h}} \omega_{i,h,j}(1+\tilde{z}_{i,h,j}) - \sum_{h=1}^{H_i} \sum_{j\in\mathcal{W}_{i,h}} \omega_{i,h,j}\tilde{z}_{i,h,j}$$

$$+\sum_{h=1}^{H_i} \frac{1}{m_{i,h}} \sum_{j\in\mathcal{W}_{i,h}} \frac{1}{\tilde{z}_{i,h,j}} \sum_{g=1}^{G_{i,h,j}} \left(\frac{1}{n_{i,h,j,g}} - \frac{1}{N_{i,h,j,g}}\right) N^2_{i,h,j,g} S^2_{i,h,j,g}$$

where $\tilde{z}_{i,h,j} = \frac{z_j}{\sum_{k\in\mathcal{W}_{i,h}} z_k}$ and $\omega_{i,h,j} = \left(\frac{y_{i,h,j}}{\tilde{z}_{i,h,j}} - \sum_{k\in\mathcal{W}_{i,h}} y_k\right)^2 \tilde{z}_{i,h,j}$.

$(SHR_I, SSRS_{II})$, cont.

Here, $\mathcal{V}_{i,h} = \{(j,g) : j \in \mathcal{W}_{i,h},\ g \in \{1,\ldots,G_{i,h,j}\}\}$ and, with $D^2_{i,h} = \sum_{j\in\mathcal{W}_{i,h}} \omega_{i,h,j}(1+\tilde{z}_{i,h,j})$,

$$A_{i,h} = \frac{1}{\tau_i^2}\left(D^2_{i,h} - \sum_{j\in\mathcal{W}_{i,h}} \frac{1}{\tilde{z}_{i,h,j}} \sum_{g=1}^{G_{i,h,j}} N_{i,h,j,g} S^2_{i,h,j,g}\right),$$

$$B_{i,h,(j,g)} = \frac{N^2_{i,h,j,g} S^2_{i,h,j,g}}{\tau_i^2\, \tilde{z}_{i,h,j}}, \quad \text{and} \quad c_i = \frac{1}{\tau_i^2} \sum_{h=1}^{H_i} \sum_{i\in\mathcal{W}_{i,h}} \tilde{z}_{i,h,j}\omega_{i,h,j}.$$

The sample size constraint (6) is the same and (7) changes into

$$\sum_{i=1}^{I} \sum_{h=1}^{H_i} m_{i,h} \sum_{j\in\mathcal{W}_{i,h}} \tilde{z}_{i,h,j} \sum_{g=1}^{G_{i,h,j}} n_{i,h,j,g} = n.$$

That is $x = m$, $z = n$ and $\alpha_{i,h,(j,g)} = \tilde{z}_{i,h,j}$.

$(SHR_I, SSRS_{II})$, cont.

We assume $\mathbf{D} = \underline{a}\underline{a}^T + \underline{b}\underline{b}^T - \mathrm{diag}(\underline{c})$ has the unique simple positive eigenvalue λ and that $A_{i,h} > 0$. Then the (approximate) optimal equal-precision allocation is

$$m_{i,h} = m \frac{v_i \sqrt{A_{i,h}}}{\sum_{k=1}^{I} v_k \sum_{g=1}^{H_k} \sqrt{A_{k,g}}}$$

and

$$n_{i,h,j,g} = n \frac{v_i \sqrt{B_{i,h,(j,g)}/\tilde{z}_{i,h,j}}}{m_{i,h} \sum_{k=1}^{I} v_k \sum_{g=1}^{H_k} \sum_{r \in \mathcal{W}_{k,g}} \sum_{s=1}^{G_{k,g,r}} \sqrt{B_{k,g,(r,s)} \tilde{z}_{k,g,r}}},$$

where $\underline{v}$ is the unique (up to a scale) eigenvector (of eigenvalue λ) with all coordinates of the same sign.

(SHR_I, SRS_{II}) scheme with $n_{i,h,j} = n_{i,h}$

Here $\#(\mathcal{V}_{i,h}) = 1$. We keep the notation $D^2_{i,h}$, c_i as above. The constraint for the common precision reads

$$T = \sum_{h=1}^{H_i} \frac{1}{m_{i,h}} \left(A_{i,h} + \frac{B_{i,h}}{n_{i,h}} \right),$$

where

$$A_{i,h} = \frac{1}{\tau_i^2} \left(D^2_{i,h} - \sum_{j \in \mathcal{W}_{i,h}} \frac{N_{i,h,j} S^2_{i,h,j}}{\tilde{z}_{i,h,j}} \right)$$

and

$$B_{i,h} = \frac{1}{\tau_i^2} \sum_{j \in \mathcal{W}_{i,h}} \frac{N^2_{i,h,j} S^2_{i,h,j}}{\tilde{z}_{i,h,j}}.$$

(SHR_I, SRS_{II}) scheme with $n_{i,h,j} = n_{i,h}$, cont.

Since the constraints are

$$\sum_{i=1}^{I}\sum_{h=1}^{H_i} m_{i,h} = m \quad \text{and} \quad \sum_{i=1}^{I}\sum_{h=1}^{H_i} m_{i,h} n_{i,h} = n$$

we see that $\alpha_{i,h,j} = 1$ and by Theorem (if only $A_{i,h} > 0$ and $\mathbf{D}$ has the unique simple eigenvalue $\lambda > 0$) the (approximate) optimal equal-precision allocation is

$$m_{i,h} = m\,\frac{v_i\sqrt{A_{i,h}}}{\sum_{k=1}^{I} v_k \sum_{g=1}^{H_k} \sqrt{A_{k,g}}}$$

and

$$n_{i,h} = n\frac{v_i\sqrt{B_{i,h}}}{m_{i,h}\sum_{k=1}^{I} v_k \sum_{g=1}^{H_k} \sqrt{B_{k,g}}},$$

where $\mathbf{D}\underline{v} = \lambda\underline{v}$ and $\underline{v}$ has all coordinates of the same sign.

(SHR_I, SRS_{II}) for LFS

The variable of interest: **number of unemployed in a household**.

The subpopulations: **voivodships (NUTS2 level)**.

Pseudo-population was created on the basis of a survey (accompanying the last virtual census) of 20% dwellings in Poland. The records of the sample were cloned with the cloning multiplicity equal to rounded weights. Such a "population" had 13 243 000 dwellings.

As in the standard LFS the PSUs were the census areas and SSUs were the households. The total number of PSUs $m = 1872$ and of SSUs $n = 13676$.

Actually, only a quarter of the standard sample (one of four rotation groups) were considered.

(SHR_I, SRS_{II}) for LFS, cont.

The matrices $\underline{a}\underline{a}^T$, $\underline{b}\underline{b}^T$ and $\mathrm{diag}(\underline{c})$ were computed on the basis of the variables: "number of dwellings in a census area (PSU)", "number of unemployed in a household (SSU)" and "number of households in census areas".

The standard *eigen* function of *R* (see e.g. R Core Team, 2013) gives: (1) the largest eigenvalue $\lambda \approx 9.7\%$ which is the common cv of estimators; (2) respective eigenvector $\underline{v}$ which gives the sample allocation.

The R-code available at:

$$\mathrm{https://github.com/rwieczor/eigenproblem-sample-allocation}$$

(SHR_I, SRS_{II}) for LFS, cont.

The experiment with sample allocated according to $\underline{v}$ was repeated 100 times with estimators precision evaluated through a bootstrap method (described e.g. in McCarthy and Snowden (1985)): In each stratum a suitable bootstrap samples was taken 500 times. The bootstrap variance estimate was obtained by the usual Monte Carlo approximation based on independent bootstrap replicates.

The results were compared with other 100 independent experiments in which the sample was drawn from the "population" according to standard LFS procedure (see Popiński, 2006). The same variable was estimated and the precision was evaluated again through the bootstrap procedure.

The results in the next Table. Note an average of $\approx$ 14.5% gain in CV in subpopulations.

(SHR_I, SRS_{II}) for LFS, cont.

sub	s-SSU	s-PSU	e-SSU	e-SSU	s-CV	e-CV
PL11	884	130	888	127	10.6	9.4
PL12	1170	156	1033	153	10.8	9.5
PL21	884	130	861	130	10.9	9.4
PL22	1170	182	999	114	9.2	9.6
PL31	754	104	813	117	11.2	9.5
PL32	702	104	661	96	10.2	9.7
PL33	676	78	731	117	12.4	9.5
PL34	832	78	882	120	11.8	9.5
PL41	936	156	927	129	10.3	9.5
PL42	806	104	836	112	11.7	9.7
PL43	572	78	814	95	12.3	9.9
PL51	910	130	880	113	10.6	9.7
PL52	988	156	942	133	10.3	9.5
PL61	728	104	779	101	10.9	9.6
PL62	780	78	754	103	11.0	9.6
PL63	884	104	869	111	11.0	9.6
Sum	13676	1872	13669	1871		

Based on:

WESOŁOWSKI, J., WIECZORKOWSKI, R.
An eigenproblem approach to optimal equal-precision sample allocation in subpopulations.
Communications in Statistics. Theory and Methods
(2015) - in press

Available on arXiv - see: arXiv:1503.08686