European Journal of Operational Research 201 (2010) 500–504

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Unequal probability sampling from a finite population: A multicriteria approach

Emilio Carrizosa *

Facultad de Matemáticas, Universidad de Sevilla, Avda Reina Mercedes s/n, 41012 Sevilla, Spain

ARTICLE INFO

Article history:
Received 6 October 2007
Accepted 22 February 2009
Available online 21 March 2009

Keywords:
Multiple objective programming
Auditing
Hansen Hurwitz estimator

ABSTRACT

In this note we address the problem of determining selection probabilities for multipurpose surveys, when the aim is the simultaneous minimization of variances for each variable under study. A characterization of the set of Pareto-optimal designs is given for designs with replacement and also for a class of designs without replacement, namely, Poisson designs.

As an application, we describe a problem encountered in Auditing, where both the fraction of misstatements, and the average amount of such misstatements are of interest.

1. The problem

Let $\mathcal{U} = \{u_1, \dots, u_N\}$ be a finite population. Associated with each u_i we have an r-valued vector $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{ir})$. The vector ϑ of means,

$$\vartheta = (\vartheta_1, \dots, \vartheta_r) = \left(\frac{1}{N} \sum_{i=1}^{N} Y_{i1}, \frac{1}{N} \sum_{i=1}^{N} Y_{i2}, \dots, \frac{1}{N} \sum_{i=1}^{N} Y_{ir} \right) \tag{1}$$

is estimated by drawing from $\mathcal{U}$ a sample with replacement of size n, and considering as estimator $\widehat{\vartheta}$ the r-dimensional Hansen–Hurwitz estimator, [9], $\widehat{\vartheta} = (\widehat{\vartheta_1}, \dots, \widehat{\vartheta}_r)$ with

$$\widehat{\vartheta}_j = \frac{1}{N} \sum_{i=1}^{N} \frac{Y_{ij} f_i}{n \alpha_i}, \quad j = 1, 2, \dots, r. \tag{2}$$

Here f_i denotes the frequency of u_i in the sample, and α_i is the decision variable denoting the probability of selection of u_i at each draw. Since $(\alpha_1, \dots, \alpha_N)$ represents a probability vector, it is constrained to belong to the unit simplex Δ_N,

$$\Delta_N = \left\{ (\beta_1, \dots, \beta_N) : \sum_{j=1}^{N} \beta_j = 1, 0 \leqslant \beta_j \leqslant 1 \quad \forall j = 1, 2, \dots, N \right\}. \tag{3}$$

In this paper the Y_{ij} are assumed to be mutually independent random variables, with expected value $E(Y_{ij}) = \mu_{ij} < +\infty$ and finite variance $var(Y_{ij}) = E(Y_{ij}^2) - \mu_{ij}^2 = \sigma_{ij}^2 \geqslant 0$. For simplicity we assume in what follows:

$$\mu_{ij}^2 + \sigma_{ij}^2 > 0 \quad \forall i, j, \tag{4}$$

which holds when each Y_{ij} is not degenerate to zero. Observe also that the case of a fixed positive value for Y_{ij} is obtained by assuming that $\sigma_{ij} = 0$.

When the coefficients Y_{ij} are given, then each $\widehat{\vartheta}_j$ is unbiased for ϑ_j,

$$E\left(\frac{1}{N} \sum_{i=1}^{N} \frac{Y_{ij} f_i}{n \alpha_i} \middle| (Y_{1j}, \dots, Y_{Nj}) \right) = \frac{1}{N} \sum_{i=1}^{N} Y_{ij} \tag{5}$$

and the (design) variance of $\widehat{\vartheta}_j$ is given by

$$E\left(\left(\frac{1}{N} \sum_{i=1}^{N} Y_{ij} - \frac{1}{N} \sum_{i=1}^{N} \frac{Y_{ij} f_i}{n \alpha_i} \right)^2 \middle| (Y_{1j}, \dots, Y_{Nj}) \right)$$
$$= \sum_{i=1}^{N} \frac{Y_{ij}^2}{N^2 n \alpha_i} - \frac{1}{N^2 n} \left(\sum_{i=1}^{N} Y_{ij} \right)^2, \tag{6}$$

see p. 52 of [19].

For a given vector $(\alpha_1, \dots, \alpha_N) \in \Delta_N$ of selection probabilities, let $\varepsilon_j(\alpha_1, \dots, \alpha_N)$ denote the expected squared error in the j-th component if the sample is drawn with probabilities α_j at each stage, and we consider as random variables (under the above-mentioned assumptions) Y_{ij} and f_i as well, i.e., the expectation is computed with respect to both the sampling design and the distribution of the variables Y_{ij} :

$$\varepsilon_j(\alpha_1, \dots, \alpha_N) = E\left(\left(\frac{1}{N} \sum_{i=1}^{N} Y_{ij} - \frac{1}{N} \sum_{i=1}^{N} \frac{Y_{ij} f_i}{n \alpha_i} \right)^2 \right). \tag{7}$$

By (6), and, since, for fixed j, variables Y_{ij} are assumed to be mutually independent, one has that

* Tel.: +34 954557943; fax: +34 954622800.
E-mail address: ecarrizosa@us.es

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2009.02.036

$$\varepsilon_j(\alpha_1,\ldots,\alpha_N) = E\left(E\left(\left(\frac{1}{N}\sum_{i=1}^N Y_{ij} - \frac{1}{N}\sum_{i=1}^N \frac{Y_{ij}f_i}{n\alpha_i}\right)^2\right) \middle| (Y_{1j},\ldots,Y_{Nj})\right) = \sum_{i=1}^N \frac{E(Y_{ij}^2)}{N^2 n\alpha_i} - \frac{1}{N^2 n} E\left(\left(\sum_{i=1}^N Y_{ij}\right)^2\right) = \sum_{i=1}^N \frac{\sigma_{ij}^2+\mu_{ij}^2}{N^2 n\alpha_i} - \frac{1}{N^2 n}\left(\sum_{i=1}^N \sigma_{ij}^2 + \left(\sum_{i=1}^N \mu_{ij}\right)^2\right). \tag{8}$$

One seeks a probability vector $(\alpha_1,\ldots,\alpha_N) \in \Delta_N$ making small *all* errors ε_j given in (7). In other words, one faces the multiple-objective problem of simultaneous minimization of all errors, when the probabilities α_j are the decision variables. Observe that, by (8), those designs with some α_i equal to 0 yield an infinite value for each ε_j, and thus will be automatically discarded for optimality.

A first approach to address such multiple-objective problem would consist of reducing the problem to one single-objective optimization problem, as largely discussed in the literature of Multiple-Objective Optimization. In this sense, one error measure could be minimized imposing that the remaining errors are below specified threshold values, as suggested e.g. in [3,5,6,12,14]. This yields a problem of the form

$$\begin{array}{ll} \min & \varepsilon_l(\alpha_1,\ldots,\alpha_N) \\ \text{s.t} & \varepsilon_j(\alpha_1,\ldots,\alpha_N) \leqslant b_j \quad \forall j\neq l, \\ & (\alpha_1,\ldots,\alpha_N) \in \Delta_N \end{array} \tag{9}$$

for given bounds b_j, where Δ_N is defined in (3).

Alternatively, one can minimize a weighted sum of the errors. In other words, an optimization problem of the form

$$\begin{array}{ll} \min & \sum_{j=1}^r \omega_j \varepsilon_j(\alpha_1,\ldots,\alpha_N) \\ \text{s.t} & (\alpha_1,\ldots,\alpha_N) \in \Delta_N \end{array} \tag{10}$$

for given non-negative weights $\omega_1,\ldots,\omega_r$, is considered. See e.g. [5,6]. Another option, e.g. [11], might be to minimize the highest error, i.e., to solve the optimization problem

$$\begin{array}{ll} \min & \max_{j=1}^r \varepsilon_j(\alpha_1,\ldots,\alpha_N) \\ \text{s.t} & (\alpha_1,\ldots,\alpha_N) \in \Delta_N. \end{array} \tag{11}$$

Instead of using scalar problems as those mentioned above, we follow a multicriteria approach, and seek selection probabilities $\alpha_1,\ldots,\alpha_N$ minimizing simultaneously the r criteria. In other words, we consider the nonlinear multiple-objective optimization problem

$$\begin{array}{ll} \min & (\varepsilon_1(\alpha_1,\ldots,\alpha_N), \varepsilon_2(\alpha_1,\ldots,\alpha_N),\ldots,\varepsilon_r(\alpha_1,\ldots,\alpha_N)) \\ \text{s.t} & (\alpha_1,\ldots,\alpha_N) \in \Delta_N \end{array} \tag{12}$$

and we seek the set $\mathscr{P}$ of *Pareto-optimal* solutions to (12).

We recall that $(\alpha_1^*,\ldots,\alpha_N^*) \in \Delta_N$ is said to be Pareto-optimal if there exists no $(\alpha_1,\ldots,\alpha_N) \in \Delta_N$ satisfying

$$\varepsilon_j(\alpha_1,\ldots,\alpha_N) \leqslant \varepsilon_j(\alpha_1^*,\ldots,\alpha_N^*), \quad j = 1,2,\ldots,r, \tag{13}$$

with at least one inequality strict. See e.g. [3,7,11,13,20,21] for other multiple-objective design problems and [4] for an introduction.

2. Results

2.1. Sampling with replacement

Define, for each $j = 1,\ldots,N$, the scalars λ_j, γ_j as

$$\lambda_j = \frac{1}{N^2 n}, \quad \gamma_j = -\frac{1}{N^2 n}\left(\sum_{i=1}^N \sigma_{ij}^2 + \left(\sum_{i=1}^N \mu_{ij}\right)^2\right). \tag{14}$$

By construction, each λ_j is strictly positive. Moreover, (8) implies that

$$\varepsilon_j(\alpha_1,\ldots,\alpha_N) = \lambda_j \sum_{i=1}^N \frac{\sigma_{ij}^2+\mu_{ij}^2}{\alpha_i} + \gamma_j. \tag{15}$$

This simple form of the objectives involved enables us to characterize $\mathscr{P}$, as shown below.

Proposition 1. *Optimal solutions to (9)–(11) are elements of* $\mathscr{P}$.

Proof. By (14), each ε_j is a strictly convex function, and thus the optimization problems (9)–(11), with objectives monotonic in the errors ε_j, have a unique optimal solution, which is then Pareto-optimal, [4]. □

We now give a characterization of Pareto-optimality.

Proposition 2. *Given* $(\alpha_1^*,\ldots,\alpha_N^*) \in \Delta_N$, *the following statements are equivalent.*

1. $(\alpha_1^*,\ldots,\alpha_N^*) \in \mathscr{P}$.
2. *There exists* $(\omega_1,\omega_2,\ldots,\omega_r)$, $\sum_{j=1}^r \omega_j = 1,\ \omega_j \geqslant 0\ \forall j$, *such that*

$$\alpha_i^* = \frac{\sqrt{\sum_{j=1}^r \omega_j\left(\sigma_{ij}^2+\mu_{ij}^2\right)}}{\sum_{l=1}^N \sqrt{\sum_{j=1}^r \omega_j\left(\sigma_{lj}^2+\mu_{lj}^2\right)}}, \quad i = 1,2,\ldots,N. \tag{16}$$

Proof. Consider a scalarized version of (12) in the form (10) for some non-negative $\omega_1,\ldots,\omega_r$, $\sum_{j=1}^r \omega_j = 1$. Problem (10) can be analytically solved. Indeed, dropping the nonnegativity constraint, we have a convex problem with one linear constraint. Necessary and sufficient optimality conditions are given by taking Lagrange multipliers, yielding the $(\alpha_1^*,\ldots,\alpha_N^*)$ in (16) as unique optimal solution, which also satisfies the non-negativity constraints. In other words, such $(\alpha_1^*,\ldots,\alpha_N^*)$ is the unique optimal solution to (10). See [16] for a similar result on a related problem. This implies in particular that, for each $(\omega_1,\ldots,\omega_r)$, the optimal solution to (10) is Pareto-optimal.

Conversely, the objectives ε_j are strictly convex, and thus any Pareto-optimal solution to (12) is optimal solution to some problem of type (10), e.g. [4].

Hence, the set $\mathscr{P}$ coincides with the set of solutions of the form (16), as asserted. □

This result implies that, if the unit simplex is discretized by generating a dense enough grid of points ω, and for each such ω the corresponding α from (16) is constructed, then one obtains a finite set of Pareto-optimal solutions which approximates accurately $\mathscr{P}$. In the bivariate case, one can plot the trade-off curve for $\varepsilon_1, \varepsilon_2$, as in Fig. 1.

2.2. Application in auditing

As an application, consider an Auditing problem, in which one has a population $\mathscr{U}$ of N items to be audited. For each item u_i, the reported book value $X_i > 0$ is known, and two parameters are of main interest, [8]:

- Fraction of misstated items;
- Average amount of misstatement per item.

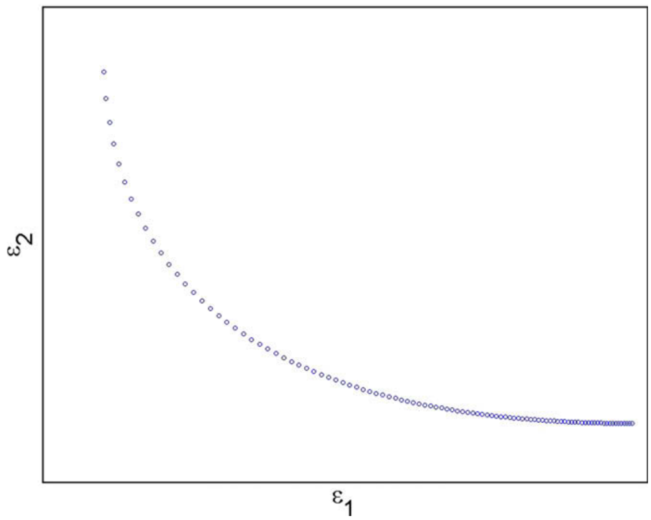

Fig. 1. Trade-off curve.

PPS sampling, frequently referred in this context as *Dollar-Unit Sampling*, amounts to choosing individuals with probabilities proportional to their book value X_i, and seems to be specially suited for the second objective, whereas a simple random sampling (SRS) might be more convenient for the first criterion, unless the likelihood of misstatement can be related with the book value. Whether a PPS design, a SRS, or a different design "in between" these two extreme designs is implemented, may be left to the analyst. For instance, in an real Auditing problem the author has recently been involved on European Regional Development Funds, the Commission Regulation (EC) 1828/2006, [17], states in its Article 17 that

> The method used to select the sample and to draw conclusions from the results shall take account of internationally accepted audit standards and be documented. Having regard to the amount of expenditure, the number and type of operations and other relevant factors, the audit authority shall determine the appropriate statistical sampling method to apply.

Hence, ample room exists to decide which sampling design should be used. We consider here the problem of describing the Pareto-optimal designs when two criteria are considered: minimization of the variance of the Hansen–Hurwitz estimators of the fraction of misstated items, and the average amount of misstatement per item.

Misstatement of item u_i is modeled as a Bernoulli variable Y_{i1}, with common success probability $\mu_{i1} = \mu$, and thus common variance $\sigma^2_{i1} = \mu(1-\mu)$.

Misstatement amount in a *misstated unit* with book value x is assumed to be a random variate, with second moment $\varphi(x)$. For instance, one may impose the second moment $\varphi(x)$ to be proportional to x^2,

$$\varphi(x) = \tau x^2 \tag{17}$$

for a given τ which, in practice, should be estimated from a test sample m_0 of misstated units, via, for instance, a ratio estimator,

$$\hat{\tau} = \frac{\sum_{k \in m_0} y_k^2}{\sum_{k \in m_0} x_k^2}, \tag{18}$$

where the pairs (x_k, y_k) represent the book value and misstatement in each unit in m_0.

Define Y_{i2} as the amount of misstatement of unit u_i. We have that

$$E(Y_{i2}^2 | Y_{i1} = 1) = \varphi(X_i), \tag{19}$$

thus

$$E(Y_{i2}^2) = \mu E(Y_{i2}^2 | Y_{i1} = 1) = \mu\varphi(X_i). \tag{20}$$

By Proposition 2, a vector $(\alpha_1^*, \ldots, \alpha_N^*)$ is Pareto-optimal iff it is of the form (16), which in this setting amounts to saying that there exists ω, $0 \leqslant \omega \leqslant 1$, such that, for each $i = 1, 2, \ldots, N$, α_i^* has the form

$$\begin{aligned}\alpha_i^* &= \frac{\sqrt{(1-\omega)\mu + \omega\mu\varphi(X_i)}}{\sum_{l=1}^N \sqrt{(1-\omega)\mu + \omega\mu\varphi(X_l)}} \\ &= \frac{\sqrt{1-\omega+\omega\varphi(X_i)}}{\sum_{l=1}^N \sqrt{1-\omega+\omega\varphi(X_l)}}.\end{aligned} \tag{21}$$

The two extreme cases for the parameter ω, namely $\omega = 0$ and $\omega = 1$, model the situations in which absolute priority is given, respectively, to the analysis of the fraction of misstated units and the average amount of misstatement per unit, and they yield, for each i, $\alpha_i^* = \frac{1}{N}$ and $\alpha_i^* \propto \sqrt{\varphi(X_i)}$, i.e., $\alpha_i^* \propto \sqrt{\tau X_i^2} \propto X_i$ under (17). In other words, the extreme cases of the trade-off parameter ω yield two well-known sampling schemes, namely, SRS and Dollar-Unit Sampling, which are shown to be Pareto-optimal by Proposition 2. Depending on the importance given to the first criterion against the second, one value of ω should be taken, and thus one specific sampling plan would be obtained by applying formula (21).

Another interesting consequence of (21) is the fact that the set of Pareto-optimal vectors is independent of the expected fraction μ of items with misstatements, since each α_i^* is proportional to $\sqrt{1-\omega+\omega\varphi(X_i)}$.

2.3. Sampling without replacement. Poisson sampling

Addressing multipurpose survey designs via Multiple-Objective methods is not restricted to designs with replacement, as analyzed in Sections 2.1, 2.2 above. Indeed, if a sampling scheme without replacement is used to estimate the r-dimensional parameter ϑ given in (1), instead of the Hansen–Hurwitz estimator, one can use the so-called Horvitz–Thompson estimator [10],

$$\hat{\vartheta}_j = \frac{1}{N}\sum_{i=1}^N \frac{Y_{ij} I_i}{\pi_i}, \quad j = 1, 2, \ldots, r, \tag{22}$$

where I_i is the random variate which takes the value 1 if u_i is selected and takes the value 0 otherwise, and π_i denotes the probability unit u_i belongs to the sample, i.e., the so-called first-order inclusion probability.

The multiple-objective problem to be addressed now has the same form than (12), namely

$$\begin{array}{ll} \min & (\varepsilon_1(d), \varepsilon_2(d), \ldots, \varepsilon_r(d)) \\ \text{s.t} & d \in \mathscr{D}, \end{array} \tag{23}$$

where $\mathscr{D}$ is a class of designs without replacement, and each ε_j represents the expected squared error under design $d \in \mathscr{D}$ in the jth component, if we consider as random variables Y_{ij} and I_i as well, i.e., if the expectation is computed jointly with respect to the sampling design and the distribution of the variables Y_{ij},

$$\begin{aligned}\varepsilon_j(d) &= E\left(\left(\frac{1}{N}\sum_{i=1}^N Y_{ij} - \frac{1}{N}\sum_{i=1}^N \frac{Y_{ij} I_i}{\pi_i}\right)^2\right) \\ &= \frac{1}{N^2}\sum_{i=1}^N \frac{1-\pi_i}{\pi_i}(\sigma_{ij}^2 + \mu_{ij}^2) + \frac{1}{N^2}\sum_{\substack{i,k=1 \\ i \neq k}}^N \frac{\pi_{ik} - \pi_i\pi_k}{\pi_i\pi_k}\mu_{ij}\mu_{kj}.\end{aligned} \tag{24}$$

Here π_{ik} denotes the second-order inclusion probability, i.e., the probability that units u_i, u_k are simultaneously in the sample. For simplicity, we write π_i and π_{ik} i.o. $\pi_i(d)$ and $\pi_{ik}(d)$, although it should be clear that both the first and the second order inclusion probabilities are design-dependent.

For a particular case of class of designs $\mathscr{D}$, a characterization similar to the one of Section 2.1 is easily obtained. Indeed, let us consider the class $\mathscr{D}$ of *Poisson designs*, e.g. [2]: Scalars $\alpha_i,\ 0 \leqslant \alpha_i \leqslant 1\ \forall i = 1,\ldots,N$ are given, and each unit u_i is selected with probability α_i, independently of the remaining units. Poisson designs have the advantage that samples are easily drawn: independent variables $X_1,\ldots,X_N$, uniformly distributed on $[0,1]$ are generated; the sample consists of those u_i with $X_i \leqslant \alpha_i$. Observe that the size of the samples generated is a random variate, with expected value $\sum_{i=1}^N \alpha_i$, and variance $\sum_{i=1}^N \alpha_i(1-\alpha_i)$.

For a Poisson design, (24) simplifies to

$$\varepsilon_j(\alpha_1,\ldots,\alpha_N) = \frac{1}{N^2}\sum_{i=1}^N \frac{\sigma_{ij}^2+\mu_{ij}^2}{\alpha_i} - \frac{1}{N^2}\sum_{i=1}^N \left(\sigma_{ij}^2+\mu_{ij}^2\right). \tag{25}$$

Hence, ε_j has the form (14), and thus the arguments in Section 2.1 can be repeated to show the following:

Proposition 3. *Given* $(\alpha_1^*,\ldots,\alpha_N^*)$, $\sum_{i=1}^N \alpha_i^* = n,\ 0 \leqslant \alpha_i^* \leqslant 1\ \forall i$, *the following statements are equivalent.*

1. *A Poisson design with selection probabilities* $(\alpha_1^*,\ldots,\alpha_N^*)$ *is a Pareto-optimal solution within the class of Poisson designs with expected size n, i.e.* $(\alpha_1^*,\ldots,\alpha_N^*)$ *is Pareto-optimal for*

$$\begin{aligned} \min\quad & (\varepsilon_1(\alpha_1,\ldots,\alpha_N),\varepsilon_2(\alpha_1,\ldots,\alpha_N),\ldots,\varepsilon_r(\alpha_1,\ldots,\alpha_N)) \\ \text{s.t}\quad & \textstyle\sum_{j=1}^N \alpha_j = n, \\ & 0 \leqslant \alpha_j \leqslant 1, \quad j = 1,\ldots,N. \end{aligned} \tag{26}$$

2. *There exists* $(\omega_1,\omega_2,\ldots,\omega_r)$, $\sum_{j=1}^r \omega_j = 1,\ \omega_j \geqslant 0\ \forall j$, *such that*

$$\alpha_i^* = n\frac{\sqrt{\sum_{j=1}^r \omega_j\left(\sigma_{ij}^2+\mu_{ij}^2\right)}}{\sum_{l=1}^N \sqrt{\sum_{j=1}^r \omega_j\left(\sigma_{lj}^2+\mu_{lj}^2\right)}}, \quad i = 1,2,\ldots,N. \tag{27}$$

Having random sample sizes, as happens with Poisson designs, is seen as a serious drawback both from a managerial viewpoint (the sample size, and thus the sampling cost is unknown in advance) and statistical viewpoint as well: the errors ε_j and its estimates may be rather large. Different variants of Poisson designs have been proposed to mitigate or avoid this undesirable effect [1,2] the most popular being the *conditional* Poisson design, in which samples are successively drawn from a Poisson design with probabilities vector $(\alpha_1,\ldots,\alpha_N)$ until a sample of size n is obtained. First and second order inclusion probabilities are obtained from the vector $(\alpha_1,\ldots,\alpha_N)$ using a recursive procedure [1]. However, the error functions ε_j do not have the nice and tractable form (14). As a matter of fact, the functions ε_j are not necessarily (quasi)convex. This is shown in Fig. 2 for a population $\mathscr{U}$ of $N = 10$, $n = 2$, and a variable $Y_1 = 5, Y_2 = \cdots = Y_{N-1} = 20, Y_N = 5$. The section of the function for $\alpha_2 = \cdots = \alpha_{N-1} = (n-0.5)/(N-2)$ is depicted for varying α_1.

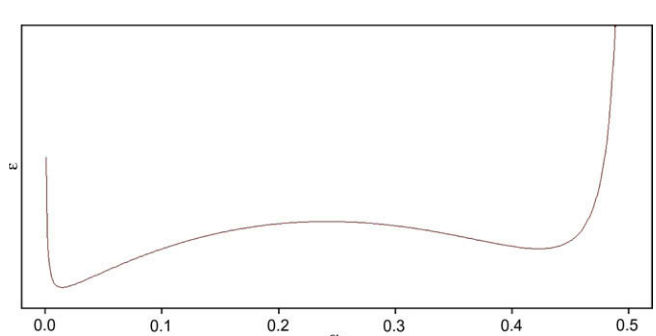

Fig. 2. Errors in conditional Poisson sampling.

Hence, one cannot guarantee that all Pareto-optimal solutions can be obtained by minimizing a weighted sum of the errors [4]. Moreover, problems of type (9)–(11) are not so tractable, since local search may not lead to global optima.

3. Discussion

In this note we have addressed the problem of finding multipurpose sampling designs using Multiple-Objective Programming ideas. The main result, with implications in Auditing, is a characterization of the set of Pareto-optimal designs when the designs under consideration are with replacement. The characterization obtained is given by a closed formula. It is then costless to get a fine discrete approximation to the Pareto-optimal set, to plot trade-off curves as in Fig. 1, and to analyze how the input parameters affect the Pareto set and the errors ε_j.

The same multiple-objective problem for designs without replacement appears to be much harder, though, at the same time, very interesting from both a theoretical and a practical viewpoint. A characterization of Pareto-optimal designs within the class of Poisson designs is given. However, for fixed-size sampling variants of Poisson sampling, such as Conditional Poisson Sampling, a similar analysis does not seem to be possible, since the error functions are not even unimodal. Characterizing the Pareto-optimal designs for other classes $\mathscr{D}$, such as order sampling schemes, [18], deserves attention, but it does not seem either to be an easy task, since even the calculation of the probabilities is a hard problem [15].

Analysis of the problem using estimators different from the Hansen–Hurwitz and Horvitz–Thompson estimators, or under statistical assumptions different to those addressed in this paper, remains an open problem.

Acknowledgements

The author thanks the two referees for their valuable suggestions. The financial support from Spain via MICINN (MTM2008-03032), and Junta de Andalucía (FQM-329) is appreciated.

References

[1] N. Aires, Algorithms to find exact inclusion probabilities for conditional Poisson sampling and Pareto PPS sampling designs, Methodology and Computing in Applied Probability 1 (1999) 457–469.
[2] K. Brewer, L. Early, M. Hanif, Poisson, modified Poisson and collocated sampling, Journal of Statistical Planning and Inference 10 (1984) 15–30.
[3] E. Carrizosa, D. Romero Morales, A biobjective method for sample allocation in stratified sampling, European Journal of Operational Research 177 (2007) 1074–1089.
[4] V. Chankong, Y.Y. Haimes, Multiobjective Decision Making: Theory and Methodology, North-Holland, 1983.
[5] M. Clyde, K. Chaloner, The equivalence of constrained and weighted designs in multiple objective design problems, Journal of the American Statistical Association 91 (1996) 1236–1244.
[6] R.D. Cook, W.K. Wong, On the equivalence of constrained and compound optimal designs, Journal of the American Statistical Association 89 (1994) 687–692.
[7] J.E. Gentle, S.C. Narula, R.L. Valliant, Multicriteria optimization in sampling design, in: S. Ghosh, W. Schucany, T. Smith (Eds.), Statistics of Quality, Marcel-Dekker, New York, 1997, pp. 411–425.
[8] D.M. Guy, D.R. Carmichael, R. Whittington, Audit Sampling. An Introduction, Wiley, 2001.
[9] M.M. Hansen, W.N. Hurwitz, On the theory of sampling from finite populations, Annals of Mathematical Statistics 14 (1943) 333–362.
[10] D. Horvitz, D. Thompson, A generalization of sampling without replacement from a finite universe, Journal of the American Statistical Association 47 (1952) 663–685.
[11] L. Imhof, W.K. Wong, A graphical method for finding maximin efficiency designs, Biometrics 56 (2000) 113–117.
[12] C.M.S. Lee, Constrained optimal designs, Journal of Statistical Planning and Inference 18 (1988) 377–389.

[13] D. Malec, Selecting multiple-objective fixed-cost sample designs using an admissibility criterion, Journal of Statistical Planning and Inference 48 (1995) 229–240.
[14] S. Mandal, B. Torsney, K.C. Carriere, Constructing optimal designs with constraints, Journal of Statistical Planning and Inference 128 (2005) 609–621.
[15] A. Matei, Y. Tillé, Computational aspects of order PPS sampling schemes, Computational Statistics and Data Analysis 51 (2007) 3703–3717.
[16] J.A. Mayor, Optimal cluster selection probabilities to estimate the finite population distribution function under PPS cluster sampling, Test 11 (2002) 73–88.
[17] Official Journal of the European Union, 27 of December 2006, L 371/1–L 371/163.
[18] B. Rosén, Asymptotic theory for order sampling, Journal of Statistical Planning and Inference 62 (1997) 135–158.
[19] C.-E. Särndal, B. Swensson, J. Wretman, Model Assisted Survey Sampling, Springer, 1992.
[20] R. Valliant, J.E. Gentle, An application of mathematical programming to sample allocation, Computational Statistics and Data Analysis 25 (1997) 337–360.
[21] W.K. Wong, Recent advances in multiple-objective design strategies, Statistica Neerlandica 53 (1999) 257–276.