ELSEVIER

Statistics & Probability Letters 35 (1997) 141–143

STATISTICS & PROBABILITY LETTERS

Efficient rounding of sampling allocations

Friedrich Pukelsheim

Institut für Mathematik, Universität Augsburg, D-86135 Augsburg, Germany

Received 1 December 1996

Abstract

It is shown that for rounding fractional sampling allocations to integers, the rounding method of Jefferson maximizes an efficiency coefficient that is motivated by variance minimization. © 1997 Elsevier Science B.V.

AMS classifications: 62D05; 65G05

Keywords: Adams rounding method; Efficient design apportionment; Efficient sampling apportionment; Jefferson rounding method; Neyman–Tschuprow allocation; Proportional allocation

1. Introduction

In stratified sampling schemes, the proportion w_h of observations in stratum h is often chosen to be proportional to N_h (proportional allocation), or to $N_hS_h/\sqrt{c_h}$ (optimal allocation). As usual, N_h, S_h, c_h denote the population size, the standard deviation of the survey variable, and the per-observation cost in stratum $h = 1,\ldots,L$. Given a total number of observations n, the *fair quota* of observations in stratum h would be nw_h which, however, will generally fail to be an integer.

Since the variance function around the optimum is rather flat, see Raj (1972, p. 56), rounding the fair quotas nw_h to a nearby integer changes the variance only slightly, and is of no relevance for practical purposes.

Nevertheless, it seems worthwhile to point out that, of the many rounding methods that are available and that are aptly summarized and discussed by Balinski and Young (1982), the method of Jefferson appears to be best suited to provide an appropriate discretization for sampling purposes.

2. Efficient sampling apportionment

Under the optimal allocation, the estimated total has a standard deviation that is proportional to

$$\sum_h N_hS_h\sqrt{c_h} \propto \sum_h w_hc_h,$$

PII S0167-7152(97)00007-2

see Hedayat and Sinha (1991, p. 270). In order to compare the vector $\boldsymbol{w} = (w_1, \dots, w_L)$ with an integer vector $\boldsymbol{n} = (n_1, \dots, n_L)$ that has component sum $\sum n_h = n$, we introduce the minimum of the likelihood ratio of $\boldsymbol{w}$ relative to $\boldsymbol{n}/n$,

$$\varepsilon_{\boldsymbol{w}/\boldsymbol{n}}(n) = \min_h \frac{w_h}{n_h/n}.$$

Then the optimal standard deviation is bounded from below according to

$$\sum_h w_h c_h \geqslant \varepsilon_{\boldsymbol{w}/\boldsymbol{n}}(n) \sum_h \frac{n_h}{n} c_h.$$

Neglecting finite population corrections and higher-order terms that are due to the discretization of w_h to n_h/n, the sum $\sum_h n_h c_h/n$ determines the standard deviation of the estimated population total. Hence ε^2 serves as a variance efficiency bound,

$$1 \geqslant \frac{\left(\sum_h w_h c_h\right)^2}{\left(\sum_h n_h c_h/n\right)^2} \geqslant \left(\varepsilon_{\boldsymbol{w}/\boldsymbol{n}}(n)\right)^2.$$

For a given sample size n we may now choose the apportionment vector $\boldsymbol{n}$ so as to maximize the efficiency bound $\varepsilon_{\boldsymbol{w}/\boldsymbol{n}}(n)$.

Theorem. *The rounding method that maximizes the efficiency bound $\varepsilon_{\boldsymbol{w}/\boldsymbol{n}}(n)$ is the Jefferson rounding method.*

Proof. We need to find $\min_{\boldsymbol{n}} \max_h n_h/w_h$. Proposition 3.10 in Balinski and Young (1982, p. 105) states that the minimum is achieved by any apportionment $\boldsymbol{n}$ that is obtained using the Jefferson rounding method. For a proof one can adapt the arguments of Section 12.7 in Pukelsheim (1993). □

3. Discussion

We would like to point out that the efficient design apportionment of Pukelsheim and Rieder (1992) builds on the Adams rounding method. Both are divisor methods. Following Balinski and Young (1982) rounding methods other than divisor methods suffer from severe paradoxes and should not be utilized.

It is remarkable, though, that the Adams method and the Jefferson method are the extreme members of the list of traditional methods in Balinski and Young (1982, p. 99). Whereas the Adams method is the divisor method based on rounding fractional remainders up, the Jefferson method is the divisor method based on rounding them down.

The Jefferson method is biased to favor big weights over smaller ones. That is, in terms of the efficient sampling apportionment a stratum with a large weight $N_h S_h/\sqrt{c_h}$ will be more likely to be assigned the next observation, when passing from sample size n to $n+1$. In this sense, the efficient sampling apportionment follows the same rules of conduct that Cochran (1977, p. 98) emphasizes for the optimal allocation.

One of the few textbooks that explicitly addresses the rounding problem is Hedayat and Sinha (1991, p. 272). In their Example 9.5, they discuss 3 strata $N_1 = 60$, $N_2 = 90$, $N_3 = 50$, with standard deviations $S_1 = 2S_2 = 4S_3$. For the proportional allocation, the efficient sampling apportionment for sample size $n = 30$ is $n_1 = 9$, $n_2 = 14$, $n_3 = 7$, which coincides with the recommendation of Hedayat and Sinha. For optimum allocation, the efficient sampling apportionment offers the two allocation vectors (16, 11, 3) and (15, 12, 3), of which the first is the one recommended by Hedayat and Sinha.

References

Balinski, M.L., Young H.P., 1982. Fair Representation. Yale University Press, New Haven and London.
Cochran, W.G., 1977. Sampling Techniques, 3rd ed. Wiley, New York.
Hedayat, A.S., Sinha B.K. 1991. Design and Inference in Finite Population Sampling. Wiley, New York.
Pukelsheim, F., 1993. Optimal Design of Experiments. Wiley, New York.
Pukelsheim, F., Rieder S., 1992. Efficient rounding of approximate designs. Biometrika 79, 763–70.
Raj, D., 1972, The Design of Sample Surveys. McGraw-Hill, New York.