

Sampling with minimal strata size requirements

Stas Kolenikov, Abt SRBI

Igor Griva, George Mason University

JSM 2016



Problem statement

- ▶ Finite population $\mathcal{U} = \bigcup_h \mathcal{U}_h$
- ▶ H strata of sizes N_h , $\sum_{h=1}^H N_h = N$
- ▶ Create a sampling design \mathcal{S} :
 1. total sample size $\sum_{h=1}^H n_h = n$
 2. minimal strata sizes m_h , $n_h \geq m_h \geq 0$

Additional $n - \sum_h m_h \geq 0$ units need to be distributed across the strata.

A simple generalization: allow for stratum-specific costs c_h ; the study size is then determined by the total budget $\sum_h n_h c_h \leq C$.

- ▶ Neyman (1934) considered sampling with different costs
- ▶ Choudhry, Rao & Hidioglou (2012) consider a similar allocation problem, resort to black box general numeric optimization
- ▶ Wright (2012) proposed an interesting connection to elections, presented a unit-by-unit algorithm

- ▶ Full consideration of a constrained optimization problem
- ▶ Explicit algorithm
- ▶ Extensions to more complicated designs (dual-frame RDD)

Example




	Total pop	Incidence	S_h^2	$m_h = 0$	$m_h = 20$	$m_h = 100$
CT	3,592,053	14.28%	0.122	69	67	100
ME	1,328,535	1.40%	0.014	9	20	100
MA	6,657,291	10.24%	0.092	111	107	100
NH	1,321,069	3.05%	0.030	13	20	100
NJ	8,874,374	18.59%	0.151	190	183	100
NY	19,594,330	18.17%	0.149	415	400	200
PA	12,758,729	6.15%	0.058	168	163	100
RI	1,053,252	13.28%	0.115	20	20	100
VT	626,358	1.63%	0.016	5	20	100

At the interactive session. . .

- ▶ . . . the slides will be rolling
- ▶ . . . I will walk you through an example in Excel
- ▶ . . . copies of the paper can be obtained



References

-  Goudhry, G. H., Rao, J. N. K. & Hidirolou, M. A. (2012), 'On sample allocation for efficient domain estimation', *Survey Methodology* **38**(1), 23–29.
-  Neyman, J. (1934), 'On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection', *Journal of the Royal Statistical Society* **109**, 558–606.
-  Wright, T. (2012), 'The equivalence of Neyman optimum allocation for sampling and equal proportions for apportioning the U.S. House of Representatives', *The American Statistician* **66**(4), 217–224.