Are we witnessing the end of random sampling in surveys?

Dan Hedlin

Department of Statistics,

Stockholm University

– and would that be a problem?

Surveys have strong emphasis on representation/generality

(Kish 1987 page 7, Baker et al. 2013 sec 8.2, O'Muircheartaigh 1999)

Reasons for random sampling

Impartiality/transparency

- Transparently impartial (cf referee tossing a coin at start of game)
- Easy to communicate
- As distant from human mind as it can be (random selection as the most stupid selection method conceivable)
- A judgement sample opens the Pandora’s box of ”improvements” of the realised sample
- Easy to accept ”bad luck” in a random sample. In a judgement sample, there is no such a thing as ‘bad luck’.

Further reasons:
Statistical culture /Knowledgebase

- Coordinated samples, PRNs
- Incorporated in university courses
- Lots of literature, including textbooks at various levels
- Lots of experience
- Knowledge among practitioners
- Public acceptance

Disadvantages with random sampling

- **Persuading reluctant respondents in a longitudinal survey may create attrition**
- **Or may create nonresponse in other surveys** (Bergman and Brage 2008)
- **Is it cost-effective? Strong emphasis on representation/generality means that expensive units need to be sampled and approached.**
- **Long period of field work**
- **Increases general response burden**

Now focus on three sampling methods

1. Random sample from a perfect frame
2. Non-random sample from a perfect frame
3. Non-random sample from a frame that suffers from missing objects

- Nonresponse in all three

Inference in surveys

- Sample s
- Specify $f(\mathbf{Y}|\mathbf{X};\boldsymbol{\beta})$ for all variables $\mathbf{Y}$ for all *N* units. $\mathbf{X}$ to be used for sampling design and estimation.

1. Analytic aim: inference about $\boldsymbol{\beta}$
2. Descriptive aim: inference about $\mathbf{Y}_{\bar{s}}$ (complement)

- We need $f(\mathbf{Y}|\mathbf{X};\boldsymbol{\beta})$ for inference about $\mathbf{Y}_{\bar{s}}$
- Further, for 3, $f(\mathbf{Y}|\mathbf{X},\mathbf{Z};\boldsymbol{\beta})$, $\mathbf{Z}$ indicates availability on the frame (binary: missing/not missing)

1. **Sample selection ignorability criterion:**

$$f(\mathbf{I}_s|\mathbf{Y}_s, \mathbf{X}) = f(\mathbf{I}_s|\mathbf{X})$$

- **True also for many nonrandom sampling designs**

  (Little 1982, Smith 1983)

2. **To be able to ignore nonresponse:**

$$f(\boldsymbol{J}_r|\mathbf{I}_s, \mathbf{Y}_s, \mathbf{X}) = f(\boldsymbol{J}_r|\mathbf{I}_s, \mathbf{X})$$

3. **To be able to ignore frame availability:**

$$f(\mathbf{Y}|\mathbf{Z}, \mathbf{X}; \boldsymbol{\beta}) = f(\mathbf{Y}|\mathbf{X}; \boldsymbol{\beta})$$

   (Little 1982, Smith 1983, Valliant et al. 2003)

- **We restrict attention to ignorable sampling designs; nonrandom samples must be ignorable**
- **I will argue: it is sampling method 3 that is different.**
- **Does criterion 3 hold in practice? Sometimes it does, sometimes it does not. See e.g. Baker et al. (2013), Callegaro et al. (2014), Craig et al. (2013), Gotway Crawford (2013), Erens et al. (2014), Martinsson (2013), Sjöström (2012), Yeager et al. (2011), etc.**

On balanced samples

- What is balance?
- $\bar{\boldsymbol{x}}_s = \bar{\boldsymbol{x}}_r$ “Response set balance”

A balanced sample is good to have

- $\bar{y}_s - \bar{y}_r = (\bar{\boldsymbol{x}}_s - \bar{\boldsymbol{x}}_r)'\hat{\boldsymbol{\beta}}_r + \left(\hat{\boldsymbol{\beta}}_s - \hat{\boldsymbol{\beta}}_r\right)'\bar{\boldsymbol{x}}_s$

  (Särndal & Lundquist 2014)
- **Does small $(\bar{\boldsymbol{x}}_s - \bar{\boldsymbol{x}}_r)$ imply small $\left(\hat{\boldsymbol{\beta}}_s - \hat{\boldsymbol{\beta}}_r\right)$?**
- **The answer is “in most cases, probably yes”**

  (Särndal & Lundquist 2014, Sec. 6)

- **So what is the causal mechanism small $(\bar{\boldsymbol{x}}_s - \bar{\boldsymbol{x}}_r)$ -> small $(\hat{\boldsymbol{\beta}}_s - \hat{\boldsymbol{\beta}}_r)$?**

Loosely speaking, it is: Small variance of response propensities (in groups defined by x) (Särndal & Lundquist 2014)

- **Note that we know $\bar{\boldsymbol{x}}_s, \bar{\boldsymbol{x}}_r$ and that we can manipulate $\bar{\boldsymbol{x}}_r$ by adaptive sampling**

  (Schouten et al. 2013, Särndal & Lundquist 2014)

- **They start with a random sample. Necessary?**

Paralleling the reasoning in Särndal & Lundquist (2014)

Let

- $\boldsymbol{U}$ be the population, with $\boldsymbol{x}$ known for all units
- $\boldsymbol{r}$ be a non-random sample
- $\boldsymbol{s}$ a random sample *that you never drew*

Define

- $t = (\overline{\boldsymbol{x}}_r - \overline{\boldsymbol{x}}_U)'\boldsymbol{b}_r$ (known)
- $v = (\boldsymbol{b}_r - \boldsymbol{b}_s)'\overline{\boldsymbol{x}}_U$ (unknown; it is $\boldsymbol{b}_s$ that is unkown)

where…

- $J_k = \begin{cases} 1 & \textbf{if } k \in r \\ 0 & \textbf{if } k \notin r \end{cases}$

- $\boldsymbol{b}_r = \left(\sum_s J_k \boldsymbol{x}_k \boldsymbol{x}'_k\right)^{-1} \left(\sum_s J_k \boldsymbol{x}_k y_k\right)$

- $\boldsymbol{b}_s = \left(\sum_s \boldsymbol{x}_k \boldsymbol{x}'_k\right)^{-1} \left(\sum_s \boldsymbol{x}_k y_k\right)$

- **Aim: reduce** $\bar{y}_s - \bar{y}_r = t + v$

- Often feasible (?) to reduce t by reducing $\bar{\boldsymbol{x}}_r - \bar{\boldsymbol{x}}_U$ through some form of adaptive sampling
- If the J_k are constant, then $\boldsymbol{b}_r = \boldsymbol{b}_s$
- If the variance among the J_k is small, $\boldsymbol{b}_r$ should be fairly close to $\boldsymbol{b}_s$
- Either reduce $\bar{\boldsymbol{x}}_r - \bar{\boldsymbol{x}}_U$ or make the variance among the J_k small.

Representativeness

- Meaning of ’a response set is representative (of some target population)’?
- Roughly: zero variance among the J_k (Schouten et al. 2012)
- Barry Schouten and co-workers proposed an ‘R-indicator’ to follow variance among the J_k while collecting data

Conclusions and further issues

- **Suppose you are successful in balancing the set of responses. Does it matter whether you have started from a random sample or a nonrandom, ignorable sample? It would seem that it does not.**
- **A more practical issue: If you strive for balancing the response set, is it easier to start from a random sample?**
- **What is best, balancing response set or adjusting through estimation? Some evidence that balancing is slightly better** (Schouten et al. 2014) **You can do both** (Särndal 2011)

- **Is it practicably doable to achieve balance?**
- **Of course, there is a broader picture** (Schouten et al. 2012)

References

- Baker, R. et al. (2013). Report on the AAPOR task force on non-probability sampling. American Association for Public Opinion Research.
- Bergman, L. and Brage, R. (2008). Survey Experiences and Later Survey Attitudes, Intentions and Behaviour. Journal of Official Statistics, 24, 99-113.
- Callegaro, M., Villar, A., Yeager, D. S. and Krosnick, J. A. (2014). A critical review of studies investigating the quality of data obtained with online panels. In M. Callegaro, R. P. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, & P. J. Lavrakas (Eds.), Online panel research. A data quality perspective (pp. 23–53).Chichester, UK: Wiley.
- Craig, B.M., Hays, R.D., Pickard, A.S., Cella, D., Revicki, D.A. and Reeve, B.B. (2013). Comparison of US Panel Vendors for Online Surveys. Journal of Medical Internet Research, 15(11), e260. http://www.jmir.org/2013/11/e260 (retrieved 12/7/15).
- Erens, B., Burkill, S., Couper, M.P., Conrad, F., Clifton, S., Tanton, C., Phelps, A., Datta, J., Mercer, C.H., Sonnenberg, P., Prah, P., Mitchell, K.R., Wellings, K., Johnson, A.M. and Copas, A.J. (2014). Nonprobability Web Surveys to Measure Sexual Behaviors and Attitudes in the General Population: A Comparison With a Probability Sample Interview Survey. Journal of Medical Internet Research, 16(12), e276. http://www.jmir.org/2014/12/e276 (retrieved 12/7/15).

- Gotway Crawford, C.A. (2013). Comment. Journal of Survey Statistics and Methodology, 1, 118-124.
- Kish, L. (1987). Statistical Design for Research. New York: Wiley.
- Little, R. J.A. (1982). Models for Nonresponse in Sample Surveys. Journal of the American Statistical Association, 77, 237-250.
- Martinsson, J. (2013). Webbpaneler och slumpmässiga urval. En jämförelse av sju undersökningar. https://politologerna.wordpress.com/2013/03/14/webbpaneler-ochslumpmassiga-urval-en-jamforelse-av-sju-undersokningar/ (retrieved 12/7/15).
- O'Muircheartaigh, C. (1999). CASM: Successes, Failures, and Potential. In Cognition and Survey Research, eds M.G. Sirken, D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur, and R. Tourangeau. New York: Wiley, 39-62.
- Särndal, C.-E. (2011). The 2010 Morris Hansen Lecture Dealing with Survey Nonresponse in Data Collection, in Estimation. Journal of Official Statistics, 27, 1-21.
- Särndal, C.-E. and Lundquist, P. (2014). Accuracy in Estimation with Nonresponse: A Function of Degree of Imbalance and Degree of Explanation. Journal of Survey Statistics and Methodology, 1-27.
- Schouten, B., Bethlehem, J., Beullens, K., Kleven, Ø., Loosvelt, G., Luiten, A., Rutar, K., Shlomo, N. and Skinner, C. (2012). Evaluating, Comparing, Monitoring, and Improving Representativeness of Survey Response Through R-Indicators and Partial R-Indicators. International Statistical Review, 80, 382-399.

- Schouten, B., Calinescu, M. and Luiten, A. (2013). Optimizing quality of response through adaptive survey designs. Survey Methodology, 39, 29-58.
- Schouten, B., Cobben, F., Lundquist, P. and Wagner, J. (2014). Theoretical and Empirical Support for Adjustment of Nonresponse by Design. Discussion paper, 2014/15, Statistics Netherlands.
- Sjöström, T. (2012). Självrekryterade jämfört med slumpmässigt rekryterade paneler. Novus, Sweden.
- Smith, T.M.F. (1983). On the validity of inferences from non-random sample. Journal of the Royal Statistical Society, Series A, 146, 394-403.
- Valliant, R., Dorfman, A.H. and Royall, R.M. (2000). Finite Population Sampling and Inference: A Prediction Approach. New York: Wiley.
- Yeager, D.S., Krosnick, J.A., Chang, l., Javitz, H.S., Levendusky, M.S., Simpser, A., Wang, R. (2011). Comparing the accuracy of RRD telephone surveys and internet surveys conducted with probability and non-probability samples. Public Opinion Quarterly, 75, 709–747.