Some Goals and Methods of Sensitivity Analysis

John L. Eltinge

Discussion of Lohr (2018, 2019)

FCSM/WSS Workshop on Sensitivity Analysis in the Integration of Multiple Data Sources

June 10, 2019

Acknowledgements and Disclaimer

The author thanks Sharon Lohr for the opportunity to review her materials in preparation for this workshop discussion; and thanks Chris Chapman, Moon Jung Cho, Wendy Martinez, Joe Schafer and participants in previous FCSM-WSS data quality workshops for helpful discussions of numerous questions on sensitivity analysis.

The views expressed here are those of the speaker and do not represent the policies of the United States Census Bureau.

Hearty Thanks to Sharon Lohr: As Always, a Very Insightful Presentation

Especially liked Lohr (2019) quote from Mallows (1997, 1998):

"Statistical arguments often fail because the basis for their assumptions is not spelled out"

Discussion: Spelling Out Multiple Dimensions of Sensitivity Analysis

I. Sensitivity OF What?

II. Sensitivity TO What?

III. What Would We DO?

I. Sensitivity OF What? - 1

A. Sensitivity of Estimation Results (realized random variable)

- Estimated model parameters θ (means, quantiles, regression coefficients, generalized linear models, hierarchical)

- Predictive distribution of substantive variable Y

I. Sensitivity OF What? - 2

B. Performance Profiles for Estimation of θ

Quality: Accuracy (MSE-TSE, interval properties),
Relevance, Timeliness, Comparability,
Coherence, Accessibility, Granularity
(Brackstone, 1999; CNSTAT, 2017; others)

Also: risk and cost (often dominate operations)

I. Sensitivity OF What? - 3

Operating Space Defined by

$Z =$ Environment (observed, uncontrolled)

$$X = \left(X_{Source}, X_{Method}, X_{System}, X_{Admin}\right)$$

= Design vector (resource decisions)

I. Sensitivity OF What? - 4

Schematic model: "Performance profile" vector

$$P = \ (Quality, Risk, Cost) = f_{\theta}(X, Z;\ \gamma) + e$$

e = residual effects (uncontrolled, unobserved)

γ = parameters of performance profile, dispersion

Spell out dominant layers of conditioning

II. Sensitivity TO What? - 1

A. Sensitivity (& Adjustment?) of Estimation:

- Extreme values of outcome variable, predictors, weights ("influential units")

- Model misspecification

- Wrong "plug in" values (e.g., imperfect calibration variables, per Dever and Valliant, 2010; outdated GVF for small domain estimation)

II. Sensitivity TO What? - 2

B. Per Lohr (2019) on "System Problems"

Sensitivity of Performance (Quality, Risk, Cost):

Inadequate Approximations to the True Design and Production Process, or Wrong γ

Ex: Level shift in P: Performance not as advertised

Ex: Rough surface – instability (high sensitivity)

II. Sensitivity TO What? - 3

C. Changes in Design Specifications X

1. Methodological design features:

   a. Data capture, record linkage, supplementary surveys, estimation

   b. "Added noise" for disclosure protection (e.g., Abowd and Schumtte, 2019)

II. Sensitivity TO What? - 4

2. Managerial: quality negotiated with data sources; IT standards; financial; training and other HR processes

3. Sensitivity to (ill-defined? unpredictable?) constraints on design settings X

II. Sensitivity TO What? - 5

D. Slippage from Nominal Design Settings X
"Operational Error"
(cf. "fault tolerant design" in engineering)

Ex: Fieldwork not as specified

Ex: Administrative source characteristics
differ from negotiated agreement
(definitions, incomplete data patterns)

II. Sensitivity TO What? - 6

E. Changes in Specific Environmental Conditions Z or Distribution of Z

Ex: Decline of public trust: “Consent to link”

Ex: Willingness to report crime through survey interviews, police reports

II. Sensitivity TO What? - 7

F. Related Puzzles:

- Observe Substantial Difference in Reported Results; Attribution to Specific *X*, *Z* Unclear

- Lohr (2019): Smoking, Crime Examples

- Longstanding "house effects" in surveys

II. Sensitivity TO What? - 8

G. Developing Numerical Results on Sensitivity:

1. Extend sample survey analysis methods to assessment of population coverage, linkage errors & entity resolution, definitional errors, incomplete data; estimation errors

(Lohr & Raghunathan,2017; Elliott & Valliant, 2017; Steorts, 2015; Meng, 2018; Rao and Molina, 2015)

II. Sensitivity TO What? - 9

2. Extend tools from Total Survey Error (TSE) analyses (e.g., Biemer et al., 2017)?

3. Align customary model diagnostics with high-priority sensitivity-analysis issues?

II. Sensitivity TO What? - 10

4. Extend utility- and prior-elicitation methods from Bayesian framework? (e.g., O'Hagan et al., 2006; Garthwaite et al., 2005)

5. Align with literature on transparency, reproducibility and replicability (e.g., Stodden et al, 2014; NASEM, 2019)

III. What Would We DO? - 1

Lohr (2019): “Systems problems need systems solutions”

- Actions in response to sensitivity analysis results:

A. Communication with internal and external stakeholders – align with information base

- Reported measures of uncertainty to reflect (most?) dominant sources and sensitivity – TSE extensions

- Note implicit conditioning and limitations – polling case

III. What Would We DO? - 2

B. Remediation steps:

Change design (X) to reduce sensitivity

1. Analysis methods, e.g.:
   - Hierarchical models
   - Bayesian model averaging

2. Other steps to "smooth" the performance profile P?

3. Does sensitivity analysis provide traction for (1), (2)?

IV. Summary: Sensitivity Analysis

A. Sensitivity OF What?

B. Sensitivity TO What?

C. What Would We DO?

Thank You!

John L. Eltinge

Assistant Director for Research

and Methodology

U.S. Census Bureau

John.L.Eltinge@census.gov

References (1)

Abowd, John M., and Ian M. Schmutte (2019). “An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices.” American Economic Review.

Biemer, Paul P., Edith de Leeuw, Stephanie Eckman, Brad Edwards, Frauke Kreuter, Lars E. Lyberg, N. Clyde Tucker, Brady T. West (Editors) (2017). *Total Survey Error in Practice.* New York: Wiley.

Brackstone, Gordon (1999). Managing Data Quality in a Statistical Agency. *Survey Methodology* **25**, 139-149.

Dever, J.A. and R.L. Valliant (2010). A Comparison of Variance Estimators for Poststratification to Estimated Control Totals." *Survey Methodology*, **36**, 45-56.

Elliott, Michael R. and Richard Valliant (2017). Inference for Nonprobability Samples. *Statistical Science* **32**, 249-264

Eltinge, John L. (2013). Integration of matrix sampling and multiple-frame methodology. Proceedings of the 59th World Statistical Congress. https://www.statistics.gov.hk/wsc/IPS033-P4-S.pdf

References (2)

Garthwaite, Paul H., Kadane, Joseph B. and O'Hagan, Anthony (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, **100**, 680–701.

Lohr, Sharon L. (2018). Measuring Uncertainty with Multiple Sources of Data. *Proceedings of Statistics Canada Symposium 2018.*

Lohr, Sharon L and Trivellore E. Raghunathan (2017). Combining Survey Data with Other Data Sources. *Statistical Science* **32**, 293-312

Meng, Xiao-Li (2018). Statistical Paradises and Paradoxes in Big Data (I): Law of Large Populations, Big Data Paradox and the 2016 U.S. Presidential Election. *Annals of Applied Statistics* 1-42.

National Academies of Sciences, Engineering, and Medicine (2017). *Federal Statistics, Multiple Data Sources, and Privacy Protection: Next Steps*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24893.

References (3)

National Academies of Sciences, Engineering, and Medicine (2019). Reproducibility and Replicability in Science. Washington, DC: The National Academies Press. https://doi.org/10.17226/25303

O'Hagan, A., C.E. Buck, A. Daneshkhah, J.R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakley and T. Rakow (2006). *Uncertain Judgements: Eliciting Experts' Probabilities.* Chichester: Wiley.

Rao, J.N.K. and I. Molina (2015). *Small Area Estimation, Second Edition.* New York: Wiley.

Steorts, Rebecca (2015). Entity Resolution with Empirically Motivated Priors. *Bayesian Analysis* **10**, 849-875.

Stodden, V, F. Leisch and R.D. Peng (2014). *Implementing Reproducible Research*. London: CRC Press