Towards Generalized Error-in-Variables Modelling:

Extended Corrected Score Functions and Partially Identified Measurement Error Correction

T. Augustin

Department of Statistics
Ludwig-Maimilians-Universität München (LMU Munich)

Sketch of the Argumentation and Overview

- Omnipresence of measurement error, severe bias in statistical analysis when neglecting it
- Powerful correction methods based on the “classical model of testing theory”

  construct unbiased estimating functions → zero expectation → consistency and asymptotic normality
  in particular Nakamura's corrected score functions

Manski's Law of Decreasing Credibility

Reliability !? Credibility ?
"The credibility of inference decreases with the strength of the assumptions maintained." (Manski (2003, p. 1))

Charles Manski[1]

[1] `http://faculty.wcas.northwestern.edu/~cfm754/`; June 19, 2019

Sketch of the Argumentation and Overview

- Omnipresence of measurement error, severe bias in statistical analysis when neglecting it
- Powerful correction methods based on the "classical model of testing theory"

  construct unbiased estimating functions → zero expectation → consistency and asymptotic normality
  in particular Nakamura's corrected score functions

Sketch of the Argumentation and Overview

- Omnipresence of measurement error, severe bias in statistical analysis when neglecting it
- Powerful correction methods based on the "classical model of testing theory"

  construct unbiased estimating functions → zero expectation → consistency and asymptotic normality
  in particular Nakamura's corrected score functions

- The underlying assumptions are very restrictive, and rarely satisfied in social surveys
- Law of Decreasing Credibility
- Relax assumptions:
  1) What can be done within the classical framework?
  2) What can be done beyond the classical framework?

The Basic Model of Classical Testing Theory (CTT)

$$\text{Measurement} = \text{True Value} + \text{Error}$$

$$X_i^*[j] = X_i[j] + U_i[j]\,, \quad i = 1, \ldots, n,\ j = 1, \ldots, p$$

Assumptions on the distribution

$$\mathbb{E}(U_i[j]) = 0 \qquad [A1.1]$$
$$Var(U_i[j]) = \sigma_j^2 \qquad [A1.2]$$
$$U_i[j] \sim N(0, \sigma_j^2) \qquad [A1.3]$$

Independence Assumptions "⊥" (Uncorrelatedness)

$$U_i[j] \perp X_i[j] \qquad [A2.1]$$
$$U_{i_1}[j] \perp U_{i_2}[j] \quad i_1 \neq i_2 \qquad [A2.2]$$
$$U_i[j_1] \perp U_i[j_2] \quad j_1 \neq j_2 \qquad [A2.3]$$
$$U_{i_1}[j_1] \perp X_{i_2}[j_2] \quad i_1 \neq i_2;\ j_1 \neq j_2 \qquad [A2.4]$$

The Basic Model of Classical Testing Theory (CTT)

The meaningfullness of many well-known measures strongly depends on the Basic Model of CTT.

- reliability $Rel(X, X^*)$
- relationship between $Rel(X, X^*)$ and $Corr(X, X^*)$
- Split-Half-Reliability
- Spearman-Brown formula
- Cronbach's alpha

Corrected Score Functions

- Measurement error correction: Find an estimating function $\psi^{X^*}(\mathbf{Y}, \mathbf{X}^*, \vartheta)$ in the error prone data with

$$\mathbb{E}_{p_\vartheta} \psi^{X^*}(\mathbf{Y}; \mathbf{X}^*; \vartheta) = \mathbf{0}.$$

- Nakamura (1990, Biometrika): proceed indirectly and find $g(\cdot)$ such that

$$\mathbb{E}(\mathbb{E}(g(\Psi_{naive})|\text{ideal data})) = \mathbb{E}(g(\Psi_{naive})) = 0 = \mathbb{E}(\Psi_{ideal})$$

  $g(\cdot)$ with

$$\mathbb{E}g(\Psi_{naive}|\text{ideal data}) = \Psi_{ideal}$$

  corrected score function

- Existence for regular GLMs, common survival models etc., given the Basic Model of CTT holds.

Corrected Score Functions

- Measurement error correction: Find an estimating function $\psi^{X^*}(\mathbf{Y}, \mathbf{X}^*, \vartheta)$ in the error prone data with

$$\mathbb{E}_{p_\vartheta} \psi^{X^*}(\mathbf{Y}; \mathbf{X}^*; \vartheta) = \mathbf{0}.$$

- Nakamura (1990, Biometrika): proceed indirectly and find $g(\cdot)$ such that

$$\mathbb{E}(\mathbb{E}(g(\Psi_{naive})|\text{ideal data})) = \mathbb{E}(g(\Psi_{naive})) = 0 = \mathbb{E}(\Psi_{ideal})$$

$g(\cdot)$ with

$$\mathbb{E}g(\Psi_{naive}|\text{ideal data}) = \Psi_{ideal}$$

corrected score function

- Existence for regular GLMs, common survival models etc., **given the Basic Model of CTT holds.**

- **But, remember the Law of Decreasing Credibility!**

What can be done within the classical framework?

What can be done within the classical framework?

- Measurement error with known dependency of components
- Nonnormal measurment error when distribution is known (regularity conditions on the moment generating functions)
- Heteroscedastic error when distribution is known
- Normal measurement error with a precise measurement model where mean or variance are depending on the true value seems to work

- Interesting results when replicates are available, mainly by work around CY Wang (e.g. Wang (2012, Biostatistics)) $\longrightarrow$ extend to longitudinal data with constant trend?

What can be done within the classical framework?

- Measurement error **with known dependency of components**
- Nonnormal measurment error **when distribution is known** (under regularity conditions on the moment generating function)
- Heteroscedastic error **when distribution is known**
- Normal measurement error **with a precise measurement model** where mean or variance are depending on the true value seems to work

Partial Identification

Change the underlying mindset:

Instead of

$$\text{data} + \text{strong assumptions} \longrightarrow \text{unique model}$$

Partial Identification

Change the underlying mindset:

Instead of

$$\text{data} + \text{strong assumptions} \longrightarrow \text{unique model}$$

now

$$\text{data} \longrightarrow \text{set of compatible model}$$

Reliable inference instead of overprecision!!

- Adding untenable assumptions to produce precise solution may destroy credibility of statistical analysis, and therefore its relevance for the subject matter questions.
- Make *realistic* assumptions and let the data speak for themselves!
- Extreme case: Consider the *set* of *all* models that are compatible with the data (and then add successively additional assumptions, if desirable)
- The results may be imprecise, but are more reliable.
- The extent of imprecision is related to the data quality!
- As a welcome by-product: clarification of the implication of certain assumptions
- Often still sufficient to answer subjective matter question

Successfully applied in different contexts

- Nonrandomly missing data (e.g. treatment evaluation): e.g. Manski (2002, Springer); Vansteelandt, Goetghebeur, Kenword, Molenberghs (2006, Stat. Sinica)
- Misclassification: Molinari (2008, J Econometrics); Küchenhoff, Augustin, Kunz (IntJApproxR2014)
- Interval data: e.g., Manski & Tamer (2002, Econometrica); Schollmeyer & Augustin (2015, IntJApproxR); reliable computing in engineering e.g., Ferson et al.; Kreinovich et al.
- Nonrandomly coarsened categorical/ordinal data Plass, Cattaneo, Augustin, Schollmeyer, Heumann: (2019, IntStatRev)
- Imprecise imputation for statistical matching: Endres, Fink & Augustin (2019, JOffStat)

Opportunities and Challenges in the Measurement Error Context

Opportunities and Challenges in the Measurement Error Context

- Use a set of 'plausible' measurement error models, for instance
  - Set of measurement error distributions
  - Set of dependencies
  - Generalized independence models
  - Structure models with varying parameters
- Set of corrected score functions
- (Convex hull of) parameter estimates

- Generalized confidence regions / tests

Concluding Remarks

- Omnipresence of measurement error, severe bias in statistical analysis when neglecting it
- Powerful correction methods based on the "classical model of testing theory"

  construct unbiased estimating functions → zero expectation → consistency and asymptotic normality
  in particular Nakamura's corrected score functions
- The underlying assumptions are very restrictive, and rarely satisfied in social surveys
- Law of Decreasing Credibility
- Relax assumptions:
  1) Something can be done within the classical framework!
  2) Great opportunities when going beyond the classical framework!