Regression Methods for Longitudinal Data with Missing Observations and Mismeasured Measurements

Grace Y. Yi

Department of Statistics & Actuarial Science

University of Waterloo

Outline

- INCOMPLETE LONGITUDINAL DATA
- MOTIVATING EXAMPLES
- MODEL FORMULATION
- ESTIMATION PROCEDURES
- REGRESSION FOR BINARY DATA
- SIMULATION STUDY
- DISCUSSION

Incomplete Longitudinal Data

NOTATION

- n subjects are followed up longitudinally at m occasions

$$Y_{i1} \quad Y_{i2} \quad Y_{i3} \quad Y_{i4} \quad Y_{i5} \quad Y_{i6}$$

- Y_{ij}: continuous response; $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2}, ..., Y_{im})'$
- $R_{ij} = I(Y_{ij}$ is observed); $\boldsymbol{R}_i = (R_{i1}, R_{i2}, ..., R_{im})'$
- $\boldsymbol{x}_{ij}$: covariate vector

MODEL OF INTEREST

- $\boldsymbol{\mu}_i = E(\boldsymbol{Y}_i | \boldsymbol{x}_i)$: mean vector

SELECTION MODELS (Little & Rubin 1987)

$$f(\boldsymbol{Y}_i, \boldsymbol{R}_i | \boldsymbol{x}_i; \boldsymbol{\theta}, \boldsymbol{\alpha}) = f(\boldsymbol{Y}_i | \boldsymbol{x}_i; \boldsymbol{\theta}) f(\boldsymbol{R}_i | \boldsymbol{Y}_i, \boldsymbol{x}_i; \boldsymbol{\alpha})$$

MISSING DATA MECHANISMS (Little & Rubin 2002)

- Missing Completely At Random (MCAR)
  $$f(\boldsymbol{R}_i | \boldsymbol{Y}_i, \boldsymbol{x}_i; \boldsymbol{\alpha}) = f(\boldsymbol{R}_i | \boldsymbol{x}_i; \boldsymbol{\alpha})$$
- Missing At Random (MAR)
  $$f(\boldsymbol{R}_i | \boldsymbol{Y}_i, \boldsymbol{x}_i; \boldsymbol{\alpha}) = f(\boldsymbol{R}_i | \boldsymbol{Y}_i^{obs}, \boldsymbol{x}_i; \boldsymbol{\alpha})$$
- Not Missing At Random (NMAR)
  $$f(\boldsymbol{R}_i | \boldsymbol{Y}_i, \boldsymbol{x}_i; \boldsymbol{\alpha}) = f(\boldsymbol{R}_i | \boldsymbol{Y}_i^{obs}, \boldsymbol{Y}_i^{mis}, \boldsymbol{x}_i; \boldsymbol{\alpha})$$

Motivating Examples

- URINE DATA (Liu & Liang 1992)
  - 7 consecutive daily urine samples
  - 408 men participated in the study
    only 397 complete measurements
  - response: systolic blood pressure
  - covariates: age, body mass index
    daily urinary sodium chloride
- DIABETES TRIAL (Hu & Lachin 2001)
  - 9 repeated measurements of albumin excretion rate
  - incomplete response measurements (some just had 5 measurements)
  - covariates: HDL cholesterol level
    systolic blood pressure

FEATURES

- longitudinal data: a response Y with covariates $\boldsymbol{x}$ is recorded at each assessment
- missing observations: some response measurements are not available
- measurement error in covatiates:

$$\boldsymbol{x}_{ij} = (\omega_{ij}, \boldsymbol{z}'_{ij})'$$: $p \times 1$ covariate vector

ω_{ij}: error-prone

$\boldsymbol{z}_{ij}$: error-free

Model Formulation

RESPONSE MODEL

- Mean and Variance:
  - $\mu_{ij} = \mathrm{E}(Y_{ij}|\boldsymbol{x}_i)$
  - $v_{ij} = \mathrm{var}(Y_{ij}|\boldsymbol{x}_i)$

- Regression Model:

$$\mu_{ij} = g^{-1}(\boldsymbol{x}'_{ij}\boldsymbol{\beta})$$
$$v_{ij} = \phi h^{-1}(g^{-1}(\boldsymbol{x}'_{ij}\boldsymbol{\beta}))$$

ADDITIVE ERROR MODEL

$$W_{ij} = \omega_{ij} + e_{ij}$$

where e_{ij} has mean 0 and mgf $m(t)$

- Estimation of Parameters:
  - validation data sample
  - repeated measurements of ω_{ij}
  - if neither is available, then sensitivity analysis can be conducted

MISSING DATA PROCESS

- Notation:
  - monotone missing data patterns:
    $$R_{ij} = 0 \Rightarrow R_{ik} = 0 \text{ for } k > j$$
  - drop-out time:
    $$M_i = \sum_{j=1}^{m} R_{ij} + 1$$
  - conditional probability:
    $$\lambda_{ij} = P(R_{ij} = 1 | R_{i,j-1} = 1, \boldsymbol{y}_i, \boldsymbol{x}_i)$$
  - marginal probability:
    $$\pi_{ij} = P(R_{ij} = 1 | \boldsymbol{y}_i, \boldsymbol{x}_i)$$

- Conditional Method:
  - Model

$$\text{logit}\ \lambda_{ij} = \boldsymbol{u}'_{ij}\boldsymbol{\alpha}$$

$\boldsymbol{u}_{ij}$: consisting of $\boldsymbol{z}_{ij}$ and observed responses

- Conditional Method:
  - Model
$$\text{logit } \lambda_{ij} = \boldsymbol{u}_{ij}'\boldsymbol{\alpha}$$
$\boldsymbol{u}_{ij}$: consisting of $\boldsymbol{z}_{ij}$ and observed responses
  - Estimating $\boldsymbol{\alpha}$
    - Likelihood: $L_i(\boldsymbol{\alpha}) = \prod_{t=2}^{m_i-1} \lambda_{it} \cdot (1 - \lambda_{im_i})$
    - score: $\boldsymbol{S}_i(\boldsymbol{\alpha}) = \partial \ell_i(\boldsymbol{\alpha}) / \partial \boldsymbol{\alpha}$
    - $\sum_{i=1}^{n} \boldsymbol{S}_i(\boldsymbol{\alpha}) = \mathbf{0}$
    - $\pi_{ij} = P(R_{ij} = 1 | \boldsymbol{y}_i, \boldsymbol{x}_i) = \prod_{t=2}^{j} \lambda_{it}$
  - MAR mechanisms are accommodated

- Marginal Method:
  - Model

$$\text{logit}\ \pi_{ij} = \boldsymbol{u}'_{ij}\boldsymbol{\alpha}$$

$\boldsymbol{u}_{ij}$: consisting of $\boldsymbol{z}_{ij}$ and observed responses

- Marginal Method:
  - Model

$$\text{logit } \pi_{ij} = \boldsymbol{u}_{ij}'\boldsymbol{\alpha}$$

$\boldsymbol{u}_{ij}$: consisting of $\boldsymbol{z}_{ij}$ and observed responses

  - Estimating $\boldsymbol{\alpha}$
    - estimating functions for $\boldsymbol{\alpha}$:

$$\boldsymbol{S}(\boldsymbol{\alpha}) = \sum_{i=1}^{n} \boldsymbol{S}_i(\boldsymbol{\alpha})$$

    - $\boldsymbol{S}_i(\boldsymbol{\alpha}) = \frac{\partial \boldsymbol{\pi}_i'}{\partial \boldsymbol{\alpha}} \boldsymbol{W}_i^{-1}(\boldsymbol{R}_i - \boldsymbol{\pi}_i),$ for $i = 1, 2, ..., n$
    - $\boldsymbol{W}_i$ is the working matrix:
      e.g., $\boldsymbol{W}_i = \text{diag}(\pi_{ij}(1 - \pi_{ij}), j = 1, 2, ..., m)$

  - MAR and NMAR can be accommodated.

Estimation Procedures

COVARIATES ARE ERROR-FREE [a]

$$\boldsymbol{U}_i(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \boldsymbol{D}_i'[\boldsymbol{V}_i^{-1/2}\boldsymbol{\Omega}_i^{-1}\boldsymbol{V}_i^{-1/2}] \cdot \boldsymbol{\Delta}_i(\boldsymbol{\alpha}) \cdot \boldsymbol{\epsilon}_i$$

- $\boldsymbol{D}_i = \partial\boldsymbol{\mu}_i'/\partial\boldsymbol{\beta}$
- $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2}, ..., \epsilon_{im})'$: $\epsilon_{ij} = Y_{ij} - \mu_{ij}$
- $\boldsymbol{V}_i = \text{var}(\boldsymbol{Y}_i)$
- $\boldsymbol{\Omega}_i =$ correlation matrix $[r_{ijk}]^{-1}$
- $\boldsymbol{\Delta}_i(\boldsymbol{\alpha}) = \text{diag}(I(R_{ij} = 1)/\pi_{ij}, 1 \leq j \leq m)$

[a] Horvitz & Thompson 1952; Robins et al. 1995; Yi & Cook 2002 a, b

COVARIATES ARE ERROR-PRONE [a]

find $\boldsymbol{U}^*_{i\beta_s}(\boldsymbol{\beta}, \boldsymbol{\alpha}; W, Y, z)$ of the observed data such that

$$\mathrm{E}_{W|X}[U^*_{i\beta_s}(\boldsymbol{\beta}, \boldsymbol{\alpha}; W, Y, z)] = U_{i\beta_s}(\boldsymbol{\beta}, \boldsymbol{\alpha}; \omega, Y, z)$$

then

$$U^*_{i\beta_s}(\boldsymbol{\beta}, \boldsymbol{\alpha}; W, Y, z) = 0$$

is an unbiased estimating equation for β_s

Denote

$$\boldsymbol{U}^*_i = (U^*_{i\beta_1}, \ldots, U^*_{i\beta_p})'$$

[a] Nakamura 1990

COMMENTS

$$U_{i\beta_s} = \sum_{j=1}^{m}\sum_{k=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \eta_{ij} \cdot r_{ikj} v_{ik}^{-1/2} v_{ij}^{-1/2} \cdot \frac{\partial \mu_{ij}}{\partial \beta_s} \cdot (Y_{ij} - \mu_{ij})$$

- $\eta_{ij} = 1$: optimal (Robins et al. 1995)
- $\eta_{ij} = 1$, $r_{ikj} = I(k = j)$: working indep. matrix
- $r_{ikj} = I(k = j)$:
  $U_{i\beta_s} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \eta_{ij} \cdot v_{ij}^{-1} \cdot \frac{\partial \mu_{ij}}{\partial \beta_s} \cdot (Y_{ij} - \mu_{ij})$
- $U_{i\beta_s} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \eta_{ij}^{*} \cdot \frac{\partial \mu_{ij}}{\partial \beta_s} \cdot (Y_{ij} - \mu_{ij})$

METHODS

- Moment Identities:
  - $\mathrm{E}_{W|\omega}\{W_{ij}\} = \omega_{ij}$
  - $\mathrm{E}_{W|\omega}\{[m(t)]^{-1} \cdot e^{tW_{ij}}\} = e^{t\omega_{ij}}$
  - $\mathrm{E}_{W|\omega}\{[m(t)]^{-1} \cdot [W_{ij} - (m(t))^{-1}m'(t)]e^{tW_{ij}}\} = \omega_{ij}e^{t\omega_{ij}}$

- Optimal Estimating Functions $U_{i\beta_s}$:

$$U_{i\beta_s} = \sum_{j=1}^{m}\sum_{k=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot r_{ikj} v_{ik}^{-1/2} v_{ij}^{-1/2} \cdot \frac{\partial \mu_{ij}}{\partial \beta_s} \cdot (Y_{ij} - \mu_{ij})$$

REGRESSION MODELS [a]

- Linear Regression
- Quadratic Regression
- Gamma Regression
- Inverse Gaussian Regression
- Poisson Regression

[a] Yi 2005

ASYMPTOTIC DISTRIBUTION

Under regularity conditions, we have, as $n \to \infty$,

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{P}^{-1}\boldsymbol{\Sigma}[\boldsymbol{P}^{-1}]')$$

where

- $\boldsymbol{P} = E\left[\partial \boldsymbol{U}_i^*(\boldsymbol{\beta}, \boldsymbol{\alpha})/\partial \boldsymbol{\beta}'\right]$
- $\boldsymbol{\Sigma} = E\{\boldsymbol{Q}_i(\boldsymbol{\beta}, \boldsymbol{\alpha})\boldsymbol{Q}_i'(\boldsymbol{\beta}, \boldsymbol{\alpha})\}$
- $\boldsymbol{Q}_i(\boldsymbol{\beta}, \boldsymbol{\alpha}) =$
  $\boldsymbol{U}_i^*(\boldsymbol{\beta}, \boldsymbol{\alpha}) - E(\partial \boldsymbol{U}_i^*(\boldsymbol{\beta}, \boldsymbol{\alpha})/\partial \boldsymbol{\alpha}')[E(\partial \boldsymbol{S}_i(\boldsymbol{\alpha})/\partial \boldsymbol{\alpha}')]^{-1}\boldsymbol{S}_i(\boldsymbol{\alpha})$

Regression for Binary Data

LOGISTIC MODEL

$$\text{logit}\ \mu_{ij} = \omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\beta_z$$
$$v_{ij} = \mu_{ij}(1 - \mu_{ij})$$

Regression for Binary Data

LOGISTIC MODEL

$$\text{logit}\ \mu_{ij} = \omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\beta_z$$
$$v_{ij} = \mu_{ij}(1-\mu_{ij})$$

$$U_{i\beta_s} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \eta_{ij} \cdot \underline{v_{ij}^{-1}\frac{\partial \mu_{ij}}{\partial \beta_s}} \cdot (Y_{ij}-\mu_{ij})$$

take $\eta_{ij} = 1 + e^{\omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z}$, or $1 + e^{-\omega_{ij}\beta_1 - \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z}$,

- $U_{i\beta}^{(1)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} \omega_{ij} \\ \boldsymbol{z}_{ij} \end{pmatrix} \cdot \{Y_{ij} + (Y_{ij}-1)e^{\omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z})\}$
- $U_{i\beta}^{(2)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} \omega_{ij} \\ \boldsymbol{z}_{ij} \end{pmatrix} \cdot \{Y_{ij}(1 + e^{-\omega_{ij}\beta_1 - \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z}) - 1\}$

$$U_{i\beta_s} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \eta_{ij}^{*} \cdot \frac{\partial \mu_{ij}}{\partial \beta_s} \cdot (Y_{ij} - \mu_{ij})$$

- take $\eta_{ij}^{*} = (1 + e^{\omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z})^3$, then

$$U_{i\beta}^{(3)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} \omega_{ij} \\ \boldsymbol{z}_{ij} \end{pmatrix} \cdot e^{\omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z}$$
$$\cdot \{Y_{ij} + (Y_{ij} - 1) e^{\omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z})\}$$

- take $\eta_{ij}^{*} = (1 + e^{\omega_{ij}\beta_1 + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_{ij}})^2 (1 + e^{-\omega_{ij}\beta_1 - \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z})$, then

$$U_{i\beta}^{(4)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} \omega_{ij} \\ \boldsymbol{z}_{ij} \end{pmatrix} \cdot \{Y_{ij}(1 + e^{-\omega_{ij}\beta_1 - \boldsymbol{z}'_{ij}\boldsymbol{\beta}_z}) - 1\}$$

CORRECTION TERMS

$$C_1(W_{ij}, r, t) = e^{r(W_{ij}t + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_{ij})}/m(rt), \quad r = 1, 2$$

$$C_2(W_{ij}, r, t) = [W_{ij} - m'(rt)/m(rt)]e^{r(W_{ij}t + \boldsymbol{z}'_{ij}\boldsymbol{\beta}_{ij})}/m(rt), \quad r = 1, 2$$

UNBIASED ESTIMATING FUNCTIONS

$$\boldsymbol{U}^{*}_{i\beta}{}^{(1)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} Y_{ij}W_{ij} + (Y_{ij}-1)C_2(W_{ij},1,\beta_1) \\ Y_{ij}\boldsymbol{z}_{ij} + (Y_{ij}-1)\boldsymbol{z}_{ij}C_1(W_{ij},1,\beta_1) \end{pmatrix}$$

$$\boldsymbol{U}^{*}_{i\beta}{}^{(2)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} (Y_{ij}-1)W_{ij} + Y_{ij}C_2(W_{ij},1,\beta_1) \\ (Y_{ij}-1)\boldsymbol{z}_{ij} + Y_{ij}\boldsymbol{z}_{ij}C_1(W_{ij},1,\beta_1) \end{pmatrix}$$

$$\boldsymbol{U}^{*}_{i\beta}{}^{(3)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} Y_{ij}C_2(W_{ij},1,\beta_1) + (Y_{ij}-1)C_2(W_{ij},2,\beta_1) \\ Y_{ij}C_1(W_{ij},1,\beta_1)\boldsymbol{z}_{ij} + (Y_{ij}-1)C_1(W_{ij},2,\beta_1)\boldsymbol{z}_{ij} \end{pmatrix}$$

$$\boldsymbol{U}^{*}_{i\beta}{}^{(4)} = \sum_{j=1}^{m} \frac{I(R_{ij}=1)}{\pi_{ij}} \cdot \begin{pmatrix} (Y_{ij}-1)C_2(W_{ij},1,\beta_1) + Y_{ij}W_{ij} \\ (Y_{ij}-1)C_1(W_{ij},1,\beta_1)\boldsymbol{z}_{ij} + Y_{ij}\boldsymbol{z}_{ij} \end{pmatrix}$$

EFFICIENT ESTIMATOR

Let

$$\boldsymbol{\Phi}^*_{i\beta} = (\boldsymbol{U}^{*(1)}_{i\beta}, \boldsymbol{U}^{*(2)}_{i\beta}, \boldsymbol{U}^{*(3)}_{i\beta}, \boldsymbol{U}^{*(4)}_{i\beta})'$$

$$\boldsymbol{\Phi}^*_{\beta} = \frac{1}{n}\sum_{i=1}^{n} \Phi^*_{i\beta}, \quad \boldsymbol{\Sigma}^* = \mathrm{var}(\Phi^*_{\beta})$$

$$\boldsymbol{Q}^*(\boldsymbol{\beta}) = \boldsymbol{\Phi}^{*\prime}_{\beta}\boldsymbol{\Sigma}^{*-1}\boldsymbol{\Phi}^*_{\beta}$$

then

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} Q^*(\boldsymbol{\beta})$$

is an efficient estimator of $\boldsymbol{\beta}$.

Hansen 1982; Qu et al. 2000; Huang & Wang 2001

COMMENTS

- In actual implementation, $\mathbf{\Sigma}^*$ is replaced by

$$\tilde{\mathbf{\Sigma}}^* = \frac{1}{n^2}\sum_{i=1}^{n} \mathbf{\Phi}^*_{i\beta}\mathbf{\Phi}^{*\prime}_{i\beta}$$

-

$$Q^*(\hat{\boldsymbol{\beta}}) \sim \chi^2_{df}$$

with

$$\begin{aligned} df &= \dim(\mathbf{\Phi}^*_{i\beta}) - \dim(\boldsymbol{\beta}) \\ &= 3p \end{aligned}$$

Simulation Study

- Response Models:
  - Continuous response: $Y_{ij} \sim N(\mu_{ij}, 1.0)$ with

$$\mu_{ij} = \beta_0 + \omega_{ij}\beta_1$$

  - Count data: $Y_{ij} \sim Poisson(\mu_{ij})$ with

$$\log \mu_{ij} = \beta_0 + \omega_{ij}\beta_1$$

Simulation Study

- Response Models:
  - Continuous response: $Y_{ij} \sim N(\mu_{ij}, 1.0)$ with
$$\mu_{ij} = \beta_0 + \omega_{ij}\beta_1$$
  - Count data: $Y_{ij} \sim Poisson(\mu_{ij})$ with
$$\log \mu_{ij} = \beta_0 + \omega_{ij}\beta_1$$
- Measurement Error: $W_{ij} \sim N(\omega_{ij}, \sigma_e^2)$
- True covariate: $\omega_{ij} \sim N(0.5, 1.0)$
- Missing Data Models: $\text{logit}\lambda_{ij} = \alpha_0 + \alpha_1 y_{i,j-1}$
- Setting: $n = 200, 500$; 1000 simulations

Linear Regression ($n = 500$)

Missingness (α_0, α_1)	True	$\sigma_e = 0.25$ bias	s.e.	rate	$\sigma_e = 1.00$ bias	s.e.	rate
$(0.5, 0.5)$	$\beta_0 = 1$	0.0001	0.0328	95.7	0.0040	0.0530	94.5
	$\beta_1 = 1$	-0.0010	0.0310	95.1	-0.0071	0.0658	95.6
$(0.5, 0)$	$\beta_0 = 1$	0.0017	0.0385	96.0	0.0006	0.0609	95.3
	$\beta_1 = 1$	0.0007	0.0357	95.6	-0.0031	0.0751	94.7
$(0.5, -0.5)$	$\beta_0 = 1$	0.0005	0.0696	94.4	0.0046	0.1066	93.8
	$\beta_1 = 1$	-0.0021	0.0636	93.4	-0.0161	0.1370	94.2
$(0.5, 0.5)$	$\beta_0 = -1$	-0.0004	0.0725	93.0	-0.0127	0.1156	92.8
	$\beta_1 = -1$	0.0029	0.0637	93.8	0.0148	0.1458	93.8
$(0.5, 0)$	$\beta_0 = -1$	-0.0008	0.0392	95.8	-0.0046	0.0633	94.3
	$\beta_1 = -1$	0.0028	0.0361	94.6	0.0087	0.0785	94.2
$(0.5, -0.5)$	$\beta_0 = -1$	0.0010	0.0331	95.4	-0.0038	0.0531	94.9
	$\beta_1 = -1$	0.0000	0.0292	95.7	0.0067	0.0654	94.4

Poisson Regression ($n = 500$)

Missingness	True	$\sigma_e = 0.25$			$\sigma_e = 1.00$		
(α_0, α_1)		bias	s.e.	rate	bias	s.e.	rate
$(0.1, 0.1)$	$\beta_0 = 0.2$	-0.0012	0.0410	94.5	0.0017	0.0491	95.1
	$\beta_1 = 0.2$	0.0023	0.0343	93.4	-0.0018	0.0509	93.7
$(0.1, 0)$	$\beta_0 = 0.2$	0.0020	0.0448	94.1	0.0032	0.0522	93.8
	$\beta_1 = 0.2$	-0.0005	0.0369	94.6	-0.0032	0.0545	94.6
$(0.1, -0.1)$	$\beta_0 = 0.2$	0.0019	0.0465	94.6	0.0042	0.0530	95.4
	$\beta_1 = 0.2$	0.0006	0.0401	93.1	-0.0018	0.0569	94.4
$(0.1, 0.1)$	$\beta_0 = 0.2$	0.0015	0.0412	94.7	0.0065	0.0583	95.7
	$\beta_1 = 0.4$	-0.0011	0.0326	93.7	-0.0053	0.0566	94.4
$(0.1, 0)$	$\beta_0 = 0.2$	0.0016	0.0447	94.7	0.0045	0.0610	95.0
	$\beta_1 = 0.4$	0.0008	0.0325	94.6	-0.0055	0.0605	94.2
$(0.1, -0.1)$	$\beta_0 = 0.2$	0.0016	0.0495	94.1	0.0070	0.0708	94.4
	$\beta_1 = 0.4$	-0.0007	0.0366	94.2	-0.0060	0.0689	94.2

SUMMARY

- Finite sample biases are reasonably small, suggesting that the estimates obtained from the proposed methods are consistent.
- Standard error tends to increase as the magnitude in measurement error increases.
- Coverage rates agree well with the nominal level 95%, which indicates that the resultant estimators are reliable.
- Increasing sample size can reduce the magnitude of the finite sample bias and standard error, and the effect on the latter seems more striking.

Discussion

- We proposed a semiparametric approach in the sense that the full distribution form of the response process is not needed. Instead, only the marginal mean and variance structures are assumed.
- We considered a functional method for the measurement error model, where the distribution of the true covariates is not required.
- The proposed methods can be extended to the case with multiple error-prone covariates.
- More flexible missing data process can be accommodated by incorporating error-prone covariates into the modeling.