Analysis of proportional odds models with censoring and errors-in-covariates

Samiran Sinha

Texas A&M University
sinha@stat.tamu.edu

Latent Variables 2016 Conference
University of South Carolina, South Carolina

This is a joint work with **Yanyuan Ma**

October 12, 2016

Overview

- Preliminaries
- Problem statement
- Method without errors in covariates
- Method with errors in covariates
- Simulation study
- Application to a real data set
- Summary

Basics of the proportional odds model

- T: Time-to-event, X: a scalar continuous covariate, $\mathbf{Z}$: p-vector of covariates
- Under the PO model: $\mathrm{pr}(T \leq t|X, Z) = \frac{\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z}+\beta_2 X)}{1+\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z}+\beta_2 X)}$
- The hazard function:

$$\lambda(t|X, \mathbf{Z}) = \frac{\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z} + \beta_2 X)}{1 + \Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z} + \beta_2 X)} \times \frac{\partial\Lambda(t)}{\partial t}$$

- Important point that unlike the proportional hazard model, here the ratio of two hazards corresponding to two sets of covariates at time t is not free from t
- Right censored data: Murphy et al. (1997); Current status data: Rossini & Tsiatis (1996);

Quick comparison between two semiparametric models

	Proportional hazard	Proportional odds
Dist. Func.	$1 - \exp\{-\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z} + \beta_2 X)\}$	$\frac{\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z}+\beta_2 X)\}}{1+\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z}+\beta_2 X)\}}$
Hazard Func.	$\frac{\partial\Lambda(t)}{\partial t}\exp(\boldsymbol{\beta}_1^T\mathbf{Z} + \beta_2 X)$	$\frac{\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z}+\beta_2 X)}{1+\Lambda(t)\exp(\boldsymbol{\beta}_1^T\mathbf{Z}+\beta_2 X)} \times \frac{\partial\Lambda(t)}{\partial t}$
Interpretation of $\Lambda(t)$	Cumulative hazard when $X = 0$, $\mathbf{Z} = \mathbf{0}$	Odds of the event when $X = 0$, $\mathbf{Z} = \mathbf{0}$

Problem statement

- T is subject to right censoring
- Assumption: censoring time C is independent of T conditional on X and $\mathbf{Z}$
- Here we do not observe X, rather $W_1^*, \ldots, W_m^*$ are observed
- Assume that $W_j^* = X + U_j^*$ (additive measurement errors), $U_j^* \sim$ a symmetric distribution
- Goal is consistent estimation of $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \beta_2)^T$, and Λ while
  - no distributional assumption will be made on X
  - except symmetry, no other assumption will be made on the distribution of U^*

Literature review

- Errors in covariates, proportional hazard model: Prentice (1982), Nakamura (1992), Zhou and Wang (2000), Huang and Wang (2000), Hu and Lin (2002), Zhuker (2005), and others
- Some important points about Huang and Wang (2000)
  - no distributional assumption on X and U^* (not even symmetry)
  - made a clever use of the partial likelihood function that allowed them to estimate the finite dimensional parameters and infinite dimensional parameters separately

Literature review

- Cheng and Wang (2001) considered errors in covariate in the linear transformation model (it includes the proportional odds model as a special case)
  - parametrically modeled $U_i^* - U_{i'}^*$ by a symmetric distribution (such as normal)
  - parametrically modeled $X_i - X_{i'}$ by a symmetric distribution (such as normal)
  - generally produces biased results if the support of C is significantly shorter than that of T

Literature review

- Sinha and Ma (2014) considered errors in covariate in the linear transformation model (it includes the proportional odds model as a special case)
  - assumed the distribution of U^* to be symmetric, but did not model it parametrically
  - modeled the distribution of X parametrically

Proposed method

- Observed data on the ith subject, $(V_i, \Delta_i, \mathbf{Z}_i, W_{i1}, \ldots, W_{im})$, $V_i = \min(T_i, C_i)$, $\Delta_i = I(T_i \le C_i)$,
- Define $N_i(u) = I(V_i \le u, \Delta_i = 1)$, $Y_i(u) = I(V_i \ge u)$, $\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta}) = \exp(\boldsymbol{\beta}_1^T \mathbf{Z}_i + \beta_2 X_i)$
- Then,

$$M(t) = N(t) - \int_0^t Y(u) \frac{\lambda(u)\eta(X, \mathbf{Z}, \boldsymbol{\beta})}{1 + \Lambda(u)\eta(X, \mathbf{Z}, \boldsymbol{\beta})} du$$

  is a martingale with respect to filtration $\{\mathcal{F}_t : t \ge 0\}$, where $\mathcal{F}_t = \sigma\{Y(u), N(u), X, \mathbf{Z}, u \le t\}$
- Think $M(t)$ as a mean zero random variable conditional on X and $\mathbf{Z}$

Formation of estimating equations when X is observed

$$
\begin{aligned}
S_{\beta_1} &= \sum_{i=1}^{n} \int_0^{\tau} \underbrace{\mathbf{Z}_i \{1 + \Lambda(u)\eta(X_i, \mathbf{Z}_i, \beta)\} f\{\Lambda(u), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\}}_{\text{predicatble}} \\
&\quad \times \underbrace{\left\{ dN_i(u) - \frac{Y_i(u)\lambda(u)\eta(X_i, \mathbf{Z}_i, \beta)du}{1 + \Lambda(u)\eta(X_i, \mathbf{Z}_i, \beta)} \right\}}_{dM_i(u)} \\
&= \sum_{i=1}^{n} (\mathbf{Z}_i \Delta_i \{1 + \Lambda(V_i)\eta(X_i, \mathbf{Z}_i, \beta)\} f\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} \\
&\quad - \mathbf{Z}_i \eta(X_i, \mathbf{Z}_i, \beta) \left[F\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} - F(0, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]),
\end{aligned}
$$

$$
\begin{aligned}
S_{\beta_2} &= \sum_{i=1}^{n} \left(X_i \Delta_i \{1 + \Lambda(V_i)\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})\} f\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} \right. \\
&\quad \left. - X_i \eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta}) \left[F\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} - F(0, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}) \right] \right),
\end{aligned}
$$

- Here $F(\Lambda, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha})$ satisfies $\partial F(\Lambda, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha})/\partial \Lambda = f(\Lambda, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha})$
- The resulting estimating equations do not have X in the denominator that will allow us to do easy moment calculations

Estimation of Λ

$$
\begin{aligned}
S_\Lambda(u) &= \sum_{i=1}^{n} \{1+\Lambda(u)\eta(X_i, Z_i, \boldsymbol{\beta})\}\left\{dN_i(u) - Y_i(u)\frac{\lambda(u)\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})du}{1+\Lambda(u)\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})}\right\} \\
&= \sum_{i=1}^{n} \left[\{1+\Lambda(u)\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})\}dN_i(u) - Y_i(u)\lambda(u)\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})du\right], \text{ for all } u > 0.
\end{aligned}
$$

- To simplify the computation we did not include $f\{\Lambda(u), \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha}\}$ in $S_\Lambda(u)$

- Let the observed failure times be $0 < t_{n_1} < \cdots < t_{n_k}$
- Then

$$\widehat{\Lambda}(t_{n_1}) = \frac{\sum_{i=1}^{n} dN_i(t_{n_1})}{\sum_{i=1}^{n} \eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})\{Y_i(t_{n_1}) - dN_i(t_{n_1})\}}$$

- Other $\Lambda(t_{n_j})$'s can be estimated recursively as

$$\widehat{\Lambda}(t_{n_j}) = \frac{\sum_{i=1}^{n} dN_i(t_{n_j}) + \widehat{\Lambda}(t_{n_{(j-1)}}) \sum_{i=1}^{n} Y_i(t_{n_j})\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})}{\sum_{i=1}^{n}\{Y_i(t_{n_j}) - dN_i(t_{n_j})\}\eta(X_i, \mathbf{Z}_i, \boldsymbol{\beta})}, \text{ for } j = 1, \cdots, k.$$

- When the last observation happens to be an event, we replace $\widehat{\Lambda}(t_{n_k})$ with a large value, larger than $\widehat{\Lambda}(t_{n_{k-1}})$, to facilitate further analysis

Choice of f

- When there is no measurement error, the score functions for the maximum likelihood estimator (Murphy et al., 1997) are obtained if we replace $f\{\Lambda(u), \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha}\}$ by $1/\{1+\Lambda(u)\eta(X, \mathbf{Z}, \boldsymbol{\beta})\}^2$ and multiply each summand of S_Λ by $1/\{1+\Lambda(u)\eta(X, \mathbf{Z}, \boldsymbol{\beta})\}$
- However, the presence of X in the expression $1/\{1+\Lambda(u)\eta(X, \mathbf{Z}, \boldsymbol{\beta})\}^2$ will cause difficulties as soon as X becomes unobservable (keeping in mind that our goal is to find corrected estimating equations)
- To circumvent this issue we shall take f free-from X

Estimating equations when X is unobserved

$$
\begin{aligned}
S^{\mathrm{me}}_{\beta_1} &= \sum_{i=1}^{n} (\Delta_i \mathbf{Z}_i \{1 + \Lambda(V_i) g_1(W_i, \mathbf{Z}_i, \boldsymbol{\beta})\} f\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} \\
&\quad - \mathbf{Z}_i g_1(W_i, \mathbf{Z}_i, \boldsymbol{\beta}) \left[F\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} - F(0, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha})\right]) = \mathbf{0}, \\
S^{\mathrm{me}}_{\beta_2} &= \sum_{i=1}^{n} (\Delta_i \{W_i + \Lambda(V_i) g_2(W_i, \mathbf{Z}_i, \boldsymbol{\beta})\} f\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} \\
&\quad - g_2(W_i, \mathbf{Z}_i, \boldsymbol{\beta}) \left[F\{\Lambda(V_i), \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}\} - F(0, \mathbf{Z}_i, \boldsymbol{\beta}, \boldsymbol{\alpha})\right]) = 0, \\
S^{\mathrm{me}}_{\Lambda} &= \sum_{i=1}^{n} \left[\{1 + \Lambda(u) g_1(W_i, \mathbf{Z}_i, \boldsymbol{\beta})\} dN_i(u) - Y_i(u) \lambda(u) g_1(W_i, \mathbf{Z}_i, \boldsymbol{\beta}) du\right] = 0,
\end{aligned}
$$

where

$$
g_1(W_i, \mathbf{Z}_i, \boldsymbol{\beta}) = \frac{\eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})}{\gamma_1}, \quad g_2(W_i, \mathbf{Z}_i, \boldsymbol{\beta}) = \frac{\eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})}{\gamma_1^2} (\gamma_1 W - \gamma_2),
$$

$\gamma_1 = E\{\exp(\beta_2 U_i)\}$, $\gamma_2 = E\{U_i \exp(\beta_2 U_i)\}$, and $U_i = \sum_{j=1}^{m} U^*_{ij} / m$.

- Good thing is that all three equations are free of unobserved X

Notion of corrected score

- It is important that $E(S_{\beta_1}^{\mathrm{me}}|V,\Delta,X,\mathbf{Z}) = S_{\beta_1}$, $E(S_{\beta_2}^{\mathrm{me}}|V,\Delta,X,\mathbf{Z}) = S_{\beta_2}$, and $E(S_{\Lambda}^{\mathrm{me}}|V,\Delta,X,\mathbf{Z}) = S_{\Lambda}$
- These are the "corrected scores": the effect of the measurement error is corrected because the original "scores" are recovered via the intermediate conditional expectation step
- As a result, as long as the original "scores" have mean zero, the "corrected" ones will also yield a consistent estimator

Choice of f when X unobserved

- We take $f\{\Lambda(u), \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha}\} = 1/\{1 + \Lambda(u)\eta(X^*, \mathbf{Z}, \boldsymbol{\beta})\}^2$, where we take $X^* = E(X|\mathbf{Z})$ calculated using a proposed model for X given $\mathbf{Z}$ (bearing similar spirit as the regression calibration)
- However, there is no harm for replacing X^* by $E^*(X|\mathbf{Z})$, a misspecified model for the conditional expectation of X given $\mathbf{Z}$
- Importantly, unlike in the classical regression calibration treatment, our estimator will remain consistent whether the proposed model is correct or incorrect

- Furthermore, one can simply bypass the specification of a model for the distribution of X given $\mathbf{Z}$, and directly assume a model $X^* = \mu(\mathbf{Z}, \boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$, the additional parameter can be obtained through solving

$$\sum_{i=1}^{n} \frac{\partial \mu(\mathbf{Z}_i, \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \{W_i - \mu(\mathbf{Z}_i, \boldsymbol{\alpha})\} = 0$$

Estimation of γ_1 and γ_2

- Note $\gamma_1 = E\{\exp(\beta_2 U_i)\} = \{\mathcal{M}(\beta_2/m)\}^m$, where $\mathcal{M}(\cdot)$ denotes the moment generating function of U_{ij}^*, $U_i = \sum_{j=1}^m U_{ij}^*/m$
- Making use of the symmetry assumption of the distribution of U_{ij}^*, we have $\mathcal{M}(\beta_2/m) = (2\sum_{j,k=1,j<k}^m E\left[\exp\{(W_{ij}^* - W_{ik}^*)\beta_2/m\}\right]/m(m-1))^{1/2}$.
- $$\widehat{\gamma}_1 = \left[\frac{2}{nm(m-1)}\sum_{j,k=1,j<k}^m \sum_{i=1}^n \exp\{(W_{ij}^* - W_{ik}^*)\beta_2/m\}\right]^{m/2}$$

- Observe that $\gamma_2 = E\{U_i \exp(\beta_2 U_i)\} = \partial E\{\exp(\beta_2 U_i)\}/\partial\beta_2$
- Then we can derive a consistent estimator of γ_2

$$\begin{aligned}\widehat{\gamma}_2 &= \left(\widehat{\gamma}_1\right)^{(m-2)/m} \\ &\times \frac{1}{nm(m-1)} \sum_{j,k=1,j<k}^{m} \sum_{i=1}^{n} (W_{ij}^{*} - W_{ik}^{*}) \exp\{(W_{ij}^{*} - W_{ik}^{*})\beta_2/m\}\end{aligned}$$

- Good thing is that both $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ are functions of observable random variables

Complete estimation procedure

Step 0. Form $W_i = m^{-1}\sum_{j=1}^m W_{ij}^*$ for $i = 1, \ldots, n$. Obtain $\widehat{\boldsymbol{\alpha}}$
Step 1. Form $\widehat{\gamma}_1(\boldsymbol{\beta})$ and $\widehat{\gamma}_2(\boldsymbol{\beta})$, both are functions of $\boldsymbol{\beta}$
Step 2. For fixed $\boldsymbol{\beta}$ and $\widehat{\gamma}_1(\boldsymbol{\beta})$, form

$$\widehat{\Lambda}\{t_{n_1}; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})\} = \frac{\sum_{i=1}^n \widehat{\gamma}_1(\boldsymbol{\beta}) dN_i(t_{n_1})}{\sum_{i=1}^n \eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})\left\{Y_i(t_{n_1}) - dN_i(t_{n_1})\right\}}$$

and

$$\widehat{\Lambda}\{t_{n_j}, \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})\} = \frac{\sum_{i=1}^n \{\widehat{\gamma}_1(\boldsymbol{\beta}) dN_i(t_{n_j}) + Y_i(t_{n_j})\widehat{\Lambda}\{t_{n_{j-1}}, \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})\}\eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})\}}{\sum_{i=1}^n \eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})\{Y_i(t_{n_j}) - dN_i(t_{n_j})\}}$$

for $u = t_{n_1}, \ldots, t_{n_k}$.

Complete estimation procedure

Step 3. We obtain $\widehat{\beta}$ through solving

$$\sum_{i=1}^{n} \begin{pmatrix} \phi_{1,i} \\ \phi_{2,i} \end{pmatrix} = 0,$$

where

$$\begin{aligned}
\phi_{1,i} &= \mathbf{Z}_i \Delta_i \widehat{\gamma}_1(\boldsymbol{\beta}) \widehat{\Lambda}(V_i; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta}))\eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta}) f\{\widehat{\Lambda}(V_i; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})), \mathbf{Z}_i, \boldsymbol{\beta}, \widehat{\boldsymbol{\alpha}}\} \\
&\quad - \mathbf{Z}_i \eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})[F\{\widehat{\Lambda}(V_i; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})), \mathbf{Z}_i, \boldsymbol{\beta}, \widehat{\boldsymbol{\alpha}}\} - F(0, \mathbf{Z}_i, \boldsymbol{\beta}, \widehat{\boldsymbol{\alpha}})], \\
\phi_{2,i} &= \Delta_i [W_i \widehat{\gamma}_1^2(\boldsymbol{\beta}) + \widehat{\Lambda}(V_i; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta}))\{\widehat{\gamma}_1(\boldsymbol{\beta}) W_i - \widehat{\gamma}_2(\boldsymbol{\beta})\}\eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})] \\
&\quad \times f\{\widehat{\Lambda}(V_i; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})), \mathbf{Z}_i, \boldsymbol{\beta}, \widehat{\boldsymbol{\alpha}}\} \\
&\quad - \{\widehat{\gamma}_1(\boldsymbol{\beta}) W_i - \widehat{\gamma}_2(\boldsymbol{\beta})\}\eta(W_i, \mathbf{Z}_i, \boldsymbol{\beta})[F\{\widehat{\Lambda}(V_i; \boldsymbol{\beta}, \widehat{\gamma}_1(\boldsymbol{\beta})), \mathbf{Z}_i, \boldsymbol{\beta}, \widehat{\boldsymbol{\alpha}}\} - F(0, \mathbf{Z}_i, \boldsymbol{\beta}, \widehat{\boldsymbol{\alpha}})]
\end{aligned}$$

Complete estimation procedure

Step 4. Go to Steps 1 and 2 to obtain $\widehat{\gamma}_1(\widehat{\boldsymbol{\beta}})$ and $\widehat{\Lambda}\{u, \widehat{\boldsymbol{\beta}}, \widehat{\gamma}_1(\widehat{\boldsymbol{\beta}})\}$ respectively.

...

- In Step 3, we used a standard Newton-Raphson procedure
- In both the simulation and the data example, we used the classical regression calibration estimates as the initial value
- We also experimented with using the naive estimator as the initial value and the results are identical.

Asymptotic properties

Theorem. *Under some regularity conditions, when* $n \to \infty$,

i) there exists an estimator $\widehat{\boldsymbol{\beta}}$ *from the procedure described earlier so that* $|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}| \to 0$ *in probability and* $\sup_{u \in [0,\tau]} |\widehat{\Lambda}\{u, \widehat{\boldsymbol{\beta}}, \widehat{\gamma}_1(\widehat{\boldsymbol{\beta}})\} - \Lambda(u)| \to 0$ *in probability,*

ii) $\sqrt{n}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to \mathrm{Normal}(0, \Sigma_H^{-1}\Sigma_M\Sigma_H^{-T})$ *in distribution,*

iii) $\sqrt{n}[\widehat{\Lambda}\{t, \widehat{\boldsymbol{\beta}}, \widehat{\gamma}_1(\widehat{\boldsymbol{\beta}})\} - \Lambda(t)]$ *follows a zero-mean Gaussian process with a covariance kernel*

The good news is that Σ_M, Σ_H, and the above referenced covariance kernel are all consistently estimable

Simulation design

- Simulated 1,000 data sets, and each data set consists of $n = 500$ iid observations (the paper contains simulation studies for other n)
- $Z \sim \mathrm{Normal}(0, 1)$,
- $X \sim$a two-component mixture of normal distributions, $(1/3)\mathrm{Normal}(-0.6, 0.5^2) + (2/3)\mathrm{Normal}(1.25, 0.5^2)$ (for the purpose of showing that our method can handle any distribution for X)
- $T \sim$the proportional odds model with $\Lambda(t) = t^2$, and $\beta_1 = \beta_2 = 1$
- Censoring time
  - $C \sim \mathrm{Exp}(e^{2.25-X-Z})$ (20% censoring)
  - $C \sim \mathrm{Exp}(e^{0.75-X-Z})$ (50% censoring)
- $W^*_{ij} = X_i + U^*_{ij}$, $U^*_{ij} \sim \mathrm{Uniform}(-1.75, 1.75)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$

Methods used for comparison

- Naive (NV): Use MLE approach where X_i is replaced by $\overline{W}_i = (W^*_{i1} + W^*_{i2})/2$
- Regression calibration (RC): Use MLE approach where X_i is being replaced by

$$(1/\widehat{\sigma}^2 + 1/\widehat{\sigma}^2_U)\{\overline{W}_i/\widehat{\sigma}^2_U + (\widehat{\zeta}_0 + \widehat{\zeta}_1^{\mathrm{T}} Z_i)/\widehat{\sigma}^2\}$$

with $\widehat{\sigma}^2$, $\widehat{\sigma}^2_U$, $\widehat{\zeta}_0$ and $\widehat{\zeta}_1$ being the estimators of $\sigma^2 = \mathrm{var}(X|Z)$, $\sigma^2_U = \mathrm{var}(U)$, and ζ_0 and ζ_1 are the coefficients of the linear regression of X on Z

- Cheng and Wang (2001): took normal model for $X_i - X_{i'}$ and $U^*_{ij} - U^*_{i'j}$
- The proposed method: took $f\{\Lambda(t), Z, \boldsymbol{\beta}, \boldsymbol{\alpha}\} = \{1 + \Lambda(t)\exp(Z\beta_1 + X^*\beta_2)\}^{-2}$, where $X^* = \widehat{\alpha}_0 + \widehat{\boldsymbol{\alpha}}_1^T \mathbf{Z}$ with $\widehat{\alpha}_0$ and $\widehat{\boldsymbol{\alpha}}_1$ being the estimate of the coefficients of the linear model $\overline{W} = X + U = \alpha_0 + \boldsymbol{\alpha}_1^T \mathbf{Z} + \epsilon$,

Simulation results for $n = 500$

	$n=500$							
	NV		RC		CW		COR	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
	Censoring depends on X and Z							
	20% Censoring							
Bias	-0.78	-3.79	-0.79	-1.69	-1.46	-0.72	0.22	0.37
SD	0.90	0.73	0.90	0.98	0.99	1.40	1.33	2.34
MAD	0.89	0.72	0.91	1.00	1.01	1.38	1.30	2.26
ESE							1.21	2.40
CP							9.42	9.59
	50% Censoring							
Bias	-1.26	-3.94	-1.26	-1.89	-3.60	-2.32	0.35	0.54
SD	1.10	0.89	1.11	1.14	1.05	1.51	1.68	2.62
MAD	1.09	0.89	1.07	1.21	1.05	1.56	1.56	2.43
ESE							1.65	2.54
CP							9.68	9.43

[1]

[1] All entries are multiplied by 10, Bootstrap approach was used for calculating the SE of the CW method

Simulation results for $n = 1000$

	NV		RC		CW		COR	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
	Censoring depends on X and Z							
	20% Censoring							
Bias	−0.81	−3.79	−0.82	−1.70	−1.53	−0.77	0.08	0.27
SD	0.63	0.49	0.63	0.69	0.69	0.99	0.97	1.60
MAD	0.62	0.51	0.62	0.71	0.67	1.04	0.95	1.59
ESE							0.84	1.72
CP							9.43	9.69
	50% Censoring							
Bias	−1.29	−3.95	−1.29	−1.90	−3.63	−2.33	0.19	0.40
SD	0.75	0.57	0.76	0.78	0.72	1.08	1.18	1.76
MAD	0.72	0.58	0.71	0.80	0.73	1.12	1.12	1.67
ESE							1.11	1.76
CP							9.62	9.65

Application to an AIDS clinical trial data

- A randomized double-blinded study to investigate the effect of a single nucleoside or two nucleosides (different drugs) among HIV-1 infected adults (Hammer et al., 1996)
- Considered only $n = 1,036$ subjects who did not have antiretroviral treatment before this trial
- Treatment groups
  - 600 mg of zidovudine: $n_1 = 262$
  - 600 mg of zidovudine plus 400 mg of didanosine: $n_2 = 257$
  - 600 mg of zidovudine plus 2.25 mg of zalcitabine : $n_3 = 260$
  - 400 mg of didanosine: $n_4 = 257$

- T : the time to AIDs or death from the date the treatment started
- The average follow-up time was 32 months
- Only 85 subjects experienced the events during the follow-up time
- Two ($m = 2$) baseline CD4 measurements that were taken prior to the treatment started, were available
- CD4 cells help to fight infection; therefore, low CD4 counts indicates weak immune system and it is used as a marker of the stage of HIV disease

- Treatments were considered as **Z** with 600 mg of zidovudine being the reference category
- $W^*_{i1} W^*_{i2}$: logarithm of the two CD4 count at the baseline minus 5.89 for the $i^{\rm th}$ subject

Table for the data example

Covariates	NV		RC		CW		COR	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Z+D (Ref: Z)	−0.78	0.33	−0.76	0.33	−0.16	0.13	−0.80	0.34
Z+Z (Ref: Z)	−1.00	0.34	−0.99	0.34	−0.27	0.10	−0.99	0.36
D (Ref: Z)	−0.75	0.31	−0.75	0.31	−0.22	0.11	−0.81	0.34
log(CD4)	−2.19	0.40	−2.58	0.48	−0.85	0.19	−2.70	0.57

[2]

[2] Z:zidovudine, Z+D: zidovudine plus didanosine, Z+Z: zidovudine plus zalcitabine, and D:didanosine

Results with different choices of f

$$f\{\Lambda(u), \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha}\} = \left\{1 + \Lambda(u)\eta(X^*, \mathbf{Z}, \boldsymbol{\beta}, \boldsymbol{\alpha})\right\}^{-r}$$

Covariates		$r=0$	$r=1$	$r=2$	$r=5$	$r=10$	$r=15$
Z+D (Ref: Z)	Est.	−0.78	−0.79	−0.80	−0.83	−0.87	−0.90
	SE	0.36	0.36	0.36	0.35	0.34	0.34
Z+Z (Ref: Z)	Est.	−0.98	−0.99	−1.00	−1.03	−1.08	−1.12
	SE	0.37	0.36	0.36	0.35	0.35	0.35
D (Ref: Z)	Est.	−0.79	−0.81	−0.82	−0.84	−0.88	−0.91
	SE	0.34	0.34	0.33	0.33	0.33	0.33
log(CD4)	Est.	−2.69	−2.69	−2.70	−2.71	−2.70	−2.68
	SE	0.57	0.56	0.56	0.55	0.56	0.58

Summary

- We proposed a consistent *functional* method to analyze proportional odds models in the presence of errors in covariates
- We do not make any distributional assumption on the unobserved covariate X
- Other than symmetry, no assumption is made on the distribution of the measurement error

- Like other estimating equation based approaches, the proposed method is not guaranteed to produce unique solution in the small or large sample
- There is no fixed remedy to handle this situation in the errors in covariates context. If there are multiple solutions,
  - usually the solution close to the regression calibration approach can be reported as the estimate
  - alternatively, one can compute the approximate likelihood function after discretizing X, and then the solution that maximizes the likelihood can be taken be reported as the estimate

...

- No method is available to check goodness-of-fit in the errors-in-covariates case (not for any model, Cox, Proportional odds, AFT)
- We have developed an approximate graphical approach, but a theoretically sound goodness-of-fit test (or a diagnostic tool) is worth investigating

Thank **you all** !