Bayesian Quantile Regression for Longitudinal Studies with Nonignorable Missing Data

Original by
Ying Yuan and Guosheng Yin

Review by
Nayeon Kwon and Hyunwoo Im

December 7, 2021

Overview

Introduction

- Longitudinal studies often suffer from attrition, which may lead to biased estimates of the model parameters if the missing data are non-ignorable or informative.
- Modeling longitudinal data with non-ignorable missing data has drawn substantial attention (Wu and Carroll, 1988; Wu and Bailey, 1989; Little, 1993; Diggle and Kenward, 1994; among others).
- The majority of these methods focuses on mean regression. In contrast, very limited research has been conducted on quantile regression (QR) for longitudinal studies with non-ignorable missing data.
- In this article, we propose a shared-parameter QR model to deal with non-ignorable missing data, in which the quantile regression model for the longitudinal process is linked to the missing data model via sharing common random effects.

Pediatric AIDS Study

- A double-blinded randomized pediatric AIDS trial (AIDS Clinical Trials Group 128, Brady et al., 1996)
- This is designed to compare the efficacy of a lower dosage(90 mg/m2/dose) of zidovudine with a higher dosage(180 mg/m2/dose) to treat HIV-infected children with mild to moderate symptoms.
- A total of 424 subjects were enrolled with 216 subjects randomized to a low-dose group and 208 to a high-dose group.
- The CD4 cell count was collected for the participants at the study entry and every 12 weeks up to 200 weeks.
- A slower decline of the CD4 cell count represents a better treatment effect.
- As these two dosages of zidovudine may have different efficacy effects on patients with different CD4 cell counts, the QR method can be naturally used to examine the treatment effects at various quantiles of the conditional distribution of the CD4 cell counts over time.

Pediatric AIDS Study (Cont'd)

- A problem is that there was a substantial amount of intermittent missing data and a large number of dropouts.
- Only 52% of the subjects completed 3 years of follow-up for the low-dose and 45% for the high-dose group.
- In addition, approximately 50% of the subjects experienced one or more intermittent missing CD4 cell counts.
- In AIDS studies, missing data are often associated with the level or underlying rate of the change in the CD4 cell counts. These types of missing data are typically non-ignorable (Wu and Carroll, 1988; De Gruttola and Tu, 1994; and Hogan and Laird, 1997).

Preliminary Analysis: Median Regression

- In a preliminary analysis, we applied the median regression model separately for each subject and plotted the estimated individual slopes against the dropout time.
- As shown in Figure 1, subjects with lower slopes tended to drop out earlier, especially in the low-dose arm, indicating that missing data might be informative.

Preliminary Analysis: Median Regression (Cont'd)

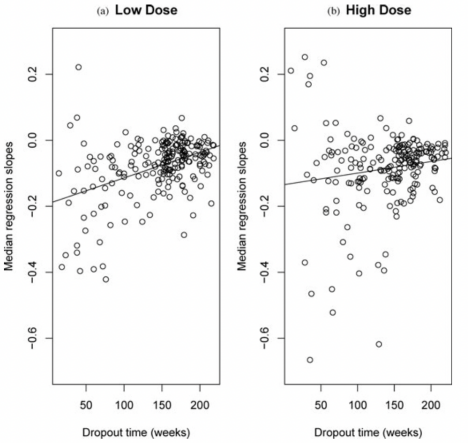

Figura: Individual median regression slopes for the square root of the CD4 cell count versus dropout times.

Proposal model

- To deal with the non-ignorable missing data, we propose a shared-parameter QR model, in which individual-level QR parameters are shrunk toward a population value by penalizing the standard check function of QR.
- We transform the penalized check function to a random-effects model in the likelihood framework.
- Assume that the missing data process is associated with the longitudinal outcome process via the shared latent subject-specific random effects.
- In the Bayesian paradigm, the estimation and inference based on the proposed model can be easily implemented using the Markov chain Monte Carlo (MCMC) procedure.

Methods: QR

- Let y_i denote the outcome of interest, and let x_i denote the corresponding covariate vector of subject i, for i = 1,...,n.
- The τth QR(Quantile Regression) model takes the form of

$$Q_{y_i}(\tau|x_i) = x_i^T\beta, \text{ where } 0 < \tau < 1.$$

- The regression coefficient β is estimated by minimizing check function :

$$\sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta), \text{ where } \rho_\tau(u) = u\{\tau - I(u < 0)\}.$$

- The check function is closely related to the ALD.

Methods: QR (Cont'd)

- The density function of an ALD with a scale parameter of 1 is

$$f(y|\mu,\tau) = \tau(1-\tau)exp\{-\rho_\tau(y-\mu)\}$$

  where τ determines the quantile level, and μ is the location parameter.

- Minimizing check function is equivalent to maximizing the likelihood function of y_i, by assuming y_i from an ALD with $\mu = x_i^T\beta$.

- The relationship between the check function and ALD can be used to reformulate the QR method in the likelihood framework.

Modeling Longitudinal Data

- Considering a longitudinal study in which n subjects are repeatedly measured at J prespecified time points, let y_{ij} denote the outcome for the ith subject measured at the jth time point, for i = 1,...,n, j = 1,...,J.
- For the τth regression quantile, we propose the following l_2-penalized check function,

$$\sum_{i=1}^{n} \sum_{j \in J_{i,obs}} \rho_\tau(y_{ij} - x_{ij}^T \beta - z_{ij}^T b_i) + \frac{1}{2} \sum_{i=1}^{n} b_i^T \Lambda^{-1} b_i, \text{ where}$$

  x_{ij} and z_{ij}: vectors of covariates,
  b_i: vector of unknown subject-specific effects,
  Λ: symmetric nonsingular matrix.
- By introducing the penalty term, we shirink the individual effects b_i toward 0.
- The amount of shrinkage is controlled by the tuning parameter Λ.

Modeling Longitudinal Data (Cont'd)

- The penalized check function can be cast into the likelihood frame work of a random-effects model as follows:

$$y_{ij}|b_i \sim ALD(\tau, x_{ij}^T\beta + z_{ij}^T b_i)$$
$$b_i \sim N(0, \Lambda)$$

- These distributional assumptions are used solely to ensure that, conditional on the tuning parameter Λ, the likelihood of model matches the penalized check function so that minimizing the penalized check function can be achieved by maximizing the likelihood of the random effects model.

Modeling Non-Ignorable Missing Data

- To account for the non-ignorable missing data, we model the intermittent missing data and dropout processes, and connect them with the longitudinal outcome process.
- Define the indicator for the missing data status s_{ij} :
  - $s_{ij} = \mathcal{O}$ if measurement j of subject i is observed.
  - $s_{ij} = \mathcal{I}$ if measurement j of subject is intermittent missing.
  - $s_{ij} = \mathcal{D}$ if subject i drops out at measurement j.
- Assume that the outcome measurement at j=1 is observed for all subjects, and dropout is an absorbing state.

Modeling Non-Ignorable Missing Data (Cont'd)

- We link the missing data process with the longitudinal outcome process by assuming that they share the same random effects b_i.
- Particularly, we model the missing data process using the transition probabilities as follows:

$$\pi_{ij}^{O} = Pr(s_{ij} = O | s_{i(j-1)} \neq D, b_i) = \frac{1}{1 + \sum_{k \in (I,D)} exp(w_{ij}^T \alpha^{(k)} + b_i^T \gamma^{(K)})}$$

$$\pi_{ij}^{I} = Pr(s_{ij} = I | s_{i(j-1)} \neq D, b_i) = \frac{exp(w_{ij}^T \alpha^{(I)} + b_i^T \gamma^{(I)})}{1 + \sum_{k \in (I,D)} exp(w_{ij}^T \alpha^{(k)} + b_i^T \gamma^{(K)})} \quad (1)$$

$$\pi_{ij}^{D} = Pr(s_{ij} = D | s_{i(j-1)} O, b_i) = \frac{exp(w_{ij}^T \alpha^{(D)} + b_i^T \gamma^{(D)})}{1 + \sum_{k \in (I,D)} exp(w_{ij}^T \alpha^{(k)} + b_i^T \gamma^{(K)})}$$

where w_{ij} is a vector of covariates, $\alpha^{(K)}$ is its associated regression parameter, and $\gamma^{(K)}$ governs the relationship between the random effects b_i and the missing data process.

Modeling Non-Ignorable Missing Data (Cont'd)

- The logarithm of the conditional likelihood for the missing data process of subject i is :

$$\sum_{j=2}^{J} log f(s_{ij}|b_i) = \sum_{j=2}^{J} \{ I(s_{ij} = O) log \pi_{ij}^{(O)} + I(s_{ij} = I) log \pi_{ij}^{(I)} + I(s_{ij} = D) log \pi_{ij}^{(D)} \},$$

  where after subject i drops out, the rest of s_{ij}'s are undefined.

- By sharing the random effects b_i with the longitudinal outcome process, the missing data model accounts for a non-ignorable missing data mechanism (Little, 1995).

Posterior Estimation

- Let (y,s) denote the observed data.The likelihood of the observed data is given by :

$$L(y, s|\beta, \Lambda, \alpha^{(K)}, \gamma^{(k)} = \prod_{i=1}^{n} \int \prod_{j \in J_{i,obs}} f(y_{ij}|b_i) \prod_{j=2}^{J} f(s_{ij}|b_i) f(b_i) db_i, \text{ k=I,D}$$

- In the Bayesian Paradigm, let $P(\beta, \Lambda, \alpha^{(K)}, \gamma^{(K)})$ denote the prior distribution of the unknown parameters. The joint posterior distribution of these parameters is given by :

$$P(\beta, \Lambda, \alpha^{(K)}, \gamma^{(K)}|y, s) \propto L(y, s|\beta, \Lambda, \alpha^{(K)}, \gamma^{(K)}) P(\beta, \Lambda, \alpha^{(K)}, \gamma^{(K)}).$$

- We assign non-informative-prior distributions to the unknown parameters as follows,

$$\beta, \alpha^{(K)}, \gamma^{(K)} \propto 1, \text{ k=I,D}$$
$$\Lambda^{-1} \sim WI(q, cI)$$

where WI(q,cI) denotes a Wishart distribution with q degrees of freedom, and a scale matrix cI with c a small constant and I the identity matrix.

Posterior Estimation (Cont'd)

- We use Gibbs sampler to obtain posterior distributions of the unknown parameters.
- The full conditional distributions of the model parameters, except for Λ, do not have closed forms.
- The adaptive rejection Metropolis sampling algorithm is used to sample from these distributions.

Simulation Study

- This simulation study mimicked the setting of the pediatric AIDS clinical trial by taking the sample size n=200, and assuming that each subject had 14 scheduled longitudinal measurements.
- Simulated data from the model

$$
\begin{aligned}
y_{ij} \mid b_{0i}, b_{1i} &= \beta_0 + \beta_1 x_{ij} + b_{0i} + b_{1i} x_{ij} + \epsilon_{ij} \\
\pi_{ij}^{(o)} &= \frac{1}{1 + \sum_{k \in (I,D)} exp(\gamma_0^{(k)} + \gamma_1^{(k)} b_{1i})} \\
\pi_{ij}^{(I)} &= \frac{exp(\gamma_0^{(I)} + \gamma_1^{(I)} b_{1i})}{1 + \sum_{k \in (I,D)} exp(\gamma_0^{(k)} + \gamma_1^{(k)} b_{1i})} \\
\pi_{ij}^{(D)} &= \frac{exp(\gamma_0^{(D)} + \gamma_1^{(D)} b_{1i})}{1 + \sum_{k \in (I,D)} exp(\gamma_0^{(k)} + \gamma_1^{(k)} b_{1i})}
\end{aligned}
\tag{2}
$$

where x_{ij} was the standardized measurement time, $x_{ij} = (j - 7.5)/4.18$ for $j = 1, ..., 14$.

Simulation Setting I

- The error ϵ_{ij} from three different distribution:
  - The standard normal distribution $N(0, 1)$
  - A $t_{(3)}$ distribution with three degrees of freedom
  - A $\chi^2_{(3)}$ distribution with three degrees of freedom
- Generate the random intercept b_{0i} $N(0, 4)$
- The random slope b_{1i} from four different distributions
  - $N(0, 2)$
  - A $t_{(3)}$ distribution with a scale parameter of $\sqrt{2}$
  - A unimodal skewed mixture distribution $0.7N(1.3, 1.6^2) + 0.3N(-3.033, 1.6^2)$
  - A bimodal mixture distribution $0.45N(-2, 1.5^2) + 0.55N(1.636, 1.5^2)$

Simulation Setting II

- In model (2), the parameter $\gamma_1^{(k)}$ determines the missing mechanism:
  - if $\gamma_1^{(k)} = 0$, the missing data are missing at random (MAR).
  - if $\gamma_1^{(k)} \neq 0$, the missing data are nonignorable.
- Considered $(\gamma_0^{(I)}, \gamma_1^{(I)}, \gamma_0^{(D)}, \gamma_1^{(D)}) = (-2.4, 0, -2.8, 0)$, and $(-6, 1, -8.9, 1.5)$ to simulate the cases of MAR and nonignorable missing data, respectively.
- The values of $\gamma_0^{(I)}$ and $\gamma_0^{(D)}$ are chosen such that, on average, 50% of the subjects experienced one or more incidences of intermittent missing data or dropout.
- Set $\beta_0 = 2$, $\beta_1 = 4$, and simulated 500 replicated datasets under each configuration.

The simulation results

ϵ_{ij}	b_{1i}	QR model	$\tau = 0.25$ Bias	SE	CP(%)
$N(0, 1)$	$N(0, 2)$	Before deletion	0.00	0.10	95.4
		Random effects	−0.47	0.11	0.8
		Shared parameter	0.01	0.12	96.8
	$t_{(3)}$	Before deletion	−0.01	0.17	93.8
		Random effects	−0.70	0.15	0.4
		Shared parameter	−0.06	0.16	94.2
	Skewed	Before deletion	−0.01	0.16	94.8
		Random effects	−0.86	0.16	0.2
		Shared parameter	0.03	0.18	94.2
	Bimodal	Before deletion	0.00	0.18	95.4
		Random effects	−0.97	0.18	0.0
		Shared parameter	0.03	0.21	95.0

Figura: Estimates of β_1 under $\epsilon_{ij} \sim N(0, 1)$ and different random distributions

- The random-effects QR yields biased estimators of β_1 and poor coverage probabilities.
- In contrast, the shared-parameter QR took the missing data into account so that the estimation bias is effectively corrected and the coverage probabilities are reasonably accurate.

The simulation results (Cont'd)

ϵ_{ij}	b_{1i}	QR model	$\tau = 0.25$ Bias	SE	CP(%)
$t_{(3)}$	$N(0, 2)$	Before deletion	0.00	0.11	94.6
		Random effects	−0.46	0.11	0.3
		Shared parameter	0.02	0.13	95.0
	$t_{(3)}$	Before deletion	0.00	0.17	95.0
		Random effects	−0.70	0.15	0.6
		Shared parameter	−0.05	0.16	95.6
	Skewed	Before deletion	−0.01	0.16	94.4
		Random effects	−0.87	0.16	0.2
		Shared parameter	0.04	0.18	93.4
	Bimodal	Before deletion	0.00	0.18	95.0
		Random effects	−0.97	0.19	0.0
		Shared parameter	0.06	0.21	95.0

Figura: Estimates of β_1 under $\epsilon_{ij} \sim t_{(3)}$ and different random distributions

- As expected, the estimates based on the shared-parameter QR are less efficient than those under the ideal before-deletion QR, but the efficiency loss is quite small.

The simulation results (Cont'd)

- When the missing data are MAR, both the estimates based on the random-effects QR and the shared-parameter QR have negligible biases and reasonable coverage probabilities, but the shared-parameter QR is slightly less efficient due to its complex model structure.
- The results for β_0 are similar except that the bias of β_0 under the random-effects QR is substantially smaller than that of β_1 because the dropout process only depends on the random slope b_{1i}.
  $\rightarrow$ The observed data are a biased sample of the complete data with respect to β_1, but not β_0.
- The simulation study shows that the estimates of the regression parameters based on the shared-parameter QR were robust to both the error and random-effects distributions.

The simulation results (Cont'd)

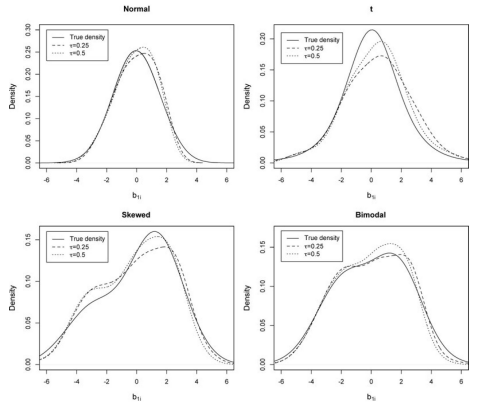

Figura: Posterior distribution of random effects b_{1i} under different true distributions for b_{1i}

- When the true distribution of b_{1i} is not normal, the posterior estimates of b_{1i} could be adaptively adjusted to the truth, providing a possible explanation for the robustness of estimating β_1 under various distributions for the random effects.

Application: Data Analysis

- This section illuminates the proposed method with the pediatric AIDS data.
- Let t_{ij} be the jth measurement time for the ith subject, y_{ij} be the square root of CD4 cell count measured at t_{ij}, and x_i be a binary treatment indicator with $x_i = 0$ denoting the high-dose arm.
- At the τth regression quantile, we considered

$$\begin{aligned} &Q_{y_{ij}}(\tau \mid x_i, t_{ij}, b_{0j}, b_{1i}) \\ &= \beta_0 + \beta_1 x_i + \beta_2 t_{ij} + \beta_3 x_i t_{ij} + b_{0j} + b_{1i} t_{ij}, \end{aligned} \tag{3}$$

where the β's characterized the population-level trajectory, $b_{0i} \sim N(0, \lambda_0)$ and $b_{1i} \sim N(0, \lambda_1)$

Application: Data Analysis (Cont'd)

- The missing data model was given by

$$
\begin{aligned}
\pi_{ij}^{(o)} &= \frac{1}{1+\sum_{k\in(I,D)} exp(\gamma_0^{(k)}+\gamma_1^{(k)}b_{0i}+\gamma_2^{(k)}b_{1i}+\gamma_3^{(k)}x_i+\gamma_4^{(k)}x_ib_{0i}+\gamma_5^{(k)}x_ib_{1i})} \\
\pi_{ij}^{(I)} &= \frac{exp(\gamma_0^{(I)}+\gamma_1^{(I)}b_{0i}+\gamma_2^{(I)}b_{1i}+\gamma_3^{(I)}x_i+\gamma_4^{(I)}x_ib_{0i}+\gamma_5^{(I)}x_ib_{1i})}{1+\sum_{k\in(I,D)} exp(\gamma_0^{(k)}+\gamma_1^{(k)}b_{0i}+\gamma_2^{(k)}b_{1i}+\gamma_3^{(k)}x_i+\gamma_4^{(k)}x_ib_{0i}+\gamma_5^{(k)}x_ib_{1i})} \\
\pi_{ij}^{(D)} &= \frac{exp(\gamma_0^{(D)}+\gamma_1^{(D)}b_{0i}+\gamma_2^{(D)}b_{1i}+\gamma_3^{(D)}x_i+\gamma_4^{(D)}x_ib_{0i}+\gamma_5^{(D)}x_ib_{1i})}{1+\sum_{k\in(I,D)} exp(\gamma_0^{(k)}+\gamma_1^{(k)}b_{0i}+\gamma_2^{(k)}b_{1i}+\gamma_3^{(k)}x_i+\gamma_4^{(k)}x_ib_{0i}+\gamma_5^{(k)}x_ib_{1i})}
\end{aligned}
\tag{4}
$$

- In the MCMC procedure, we recorded 10,000 draws after 1000 burn-in iterations.
- Checked the convergence of Markov chains with the Gelman–Rubin convergence statistic, the shrinkage factor
- After 1000 burn-in iterations, the values of the shrinkage factors became very close to 1, suggesting the convergence of these chains.

Application: Results

QR model		$\tau = 0.25$ Est.	$\tau = 0.25$ 95% CI	$\tau = 0.5$ Est.	$\tau = 0.5$ 95% CI	$\tau = 0.75$ Est.	$\tau = 0.75$ 95% CI
Random effects	β_0	21.40	(20.40, 22.41)	23.62	(22.56, 24.68)	25.97	(24.86, 27.08)
	β_1	0.14	(−1.84, 2.11)	0.19	(−1.94, 2.31)	0.24	(−2.01, 2.47)
	β_2	−3.19	(−3.54, −2.86)	−3.54	(−3.89, −3.20)	−3.82	(−4.19, −3.46)
	β_3	0.78	(0.11, 1.45)	0.82	(0.13, 1.51)	0.83	(0.127, 1.55)
	λ_0	104.49	(90.94, 120.45)	117.98	(102.61, 135.76)	129.61	(112.56, 149.17)
	λ_1	9.49	(7.83, 11.38)	10.67	(8.86, 12.75)	11.09	(9.14, 13.36)
Shared parameter	β_0	21.09	(20.07, 22.08)	23.33	(22.24, 24.39)	25.58	(24.45, 26.70)
	β_1	0.17	(−1.86, 2.14)	0.21	(−1.92, 2.28)	0.27	(−1.98, 2.47)
	β_2	−3.40	(−3.76, −3.06)	−3.75	(−4.11, −3.40)	−4.09	(−4.47, −3.72)
	β_3	0.77	(0.06, 1.44)	0.81	(0.08, 1.52)	0.84	(0.11, 1.57)
	λ_0	105.87	(92.09, 121.59)	119.77	(104.40, 137.65)	132.02	(114.93, 151.51)
	λ_1	9.55	(7.92, 11.41)	10.95	(9.10, 13.12)	11.69	(9.68, 14.04)

Figura: Estimates and 95% CIs of the model parameters for the pediatric AIDS data

- In both models, the estimates of β_1 are not significantly different from 0 at all three quantiles, indicating that the baseline CD4 count was well balanced between the two treatment arms by randomization.

Application: Results (Cont'd)

QR model		$\tau = 0.25$ Est.	$\tau = 0.25$ 95% CI	$\tau = 0.5$ Est.	$\tau = 0.5$ 95% CI	$\tau = 0.75$ Est.	$\tau = 0.75$ 95% CI
Random effects	β_0	21.40	(20.40, 22.41)	23.62	(22.56, 24.68)	25.97	(24.86, 27.08)
	β_1	0.14	(−1.84, 2.11)	0.19	(−1.94, 2.31)	0.24	(−2.01, 2.47)
	β_2	−3.19	(−3.54, −2.86)	−3.54	(−3.89, −3.20)	−3.82	(−4.19, −3.46)
	β_3	0.78	(0.11, 1.45)	0.82	(0.13, 1.51)	0.83	(0.127, 1.55)
	λ_0	104.49	(90.94, 120.45)	117.98	(102.61, 135.76)	129.61	(112.56, 149.17)
	λ_1	9.49	(7.83, 11.38)	10.67	(8.86, 12.75)	11.09	(9.14, 13.36)
Shared parameter	β_0	21.09	(20.07, 22.08)	23.33	(22.24, 24.39)	25.58	(24.45, 26.70)
	β_1	0.17	(−1.86, 2.14)	0.21	(−1.92, 2.28)	0.27	(−1.98, 2.47)
	β_2	−3.40	(−3.76, −3.06)	−3.75	(−4.11, −3.40)	−4.09	(−4.47, −3.72)
	β_3	0.77	(0.06, 1.44)	0.81	(0.08, 1.52)	0.84	(0.11, 1.57)
	λ_0	105.87	(92.09, 121.59)	119.77	(104.40, 137.65)	132.02	(114.93, 151.51)
	λ_1	9.55	(7.92, 11.41)	10.95	(9.10, 13.12)	11.69	(9.68, 14.04)

Figura: Estimates and 95% CIs of the model parameters for the pediatric AIDS data

- The estimates of β_2 under the shared-parameter QR are smaller than those under the random-effects QR across all the three quantiles, as the shared-parameter QR takes into account the fact that early dropouts were associated with lower slopes.

Application: Results (Cont'd)

QR model		$\tau = 0.25$ Est.	$\tau = 0.25$ 95% CI	$\tau = 0.5$ Est.	$\tau = 0.5$ 95% CI	$\tau = 0.75$ Est.	$\tau = 0.75$ 95% CI
Random effects	β_0	21.40	(20.40, 22.41)	23.62	(22.56, 24.68)	25.97	(24.86, 27.08)
	β_1	0.14	(−1.84, 2.11)	0.19	(−1.94, 2.31)	0.24	(−2.01, 2.47)
	β_2	−3.19	(−3.54, −2.86)	−3.54	(−3.89, −3.20)	−3.82	(−4.19, −3.46)
	β_3	0.78	(0.11, 1.45)	0.82	(0.13, 1.51)	0.83	(0.127, 1.55)
	λ_0	104.49	(90.94, 120.45)	117.98	(102.61, 135.76)	129.61	(112.56, 149.17)
	λ_1	9.49	(7.83, 11.38)	10.67	(8.86, 12.75)	11.09	(9.14, 13.36)
Shared parameter	β_0	21.09	(20.07, 22.08)	23.33	(22.24, 24.39)	25.58	(24.45, 26.70)
	β_1	0.17	(−1.86, 2.14)	0.21	(−1.92, 2.28)	0.27	(−1.98, 2.47)
	β_2	−3.40	(−3.76, −3.06)	−3.75	(−4.11, −3.40)	−4.09	(−4.47, −3.72)
	β_3	0.77	(0.06, 1.44)	0.81	(0.08, 1.52)	0.84	(0.11, 1.57)
	λ_0	105.87	(92.09, 121.59)	119.77	(104.40, 137.65)	132.02	(114.93, 151.51)
	λ_1	9.55	(7.92, 11.41)	10.95	(9.10, 13.12)	11.69	(9.68, 14.04)

Figura: Estimates and 95% CIs of the model parameters for the pediatric AIDS data

- Under the shared-parameter QR, both the high-dose and low-dose regimens were more effective for sicker patients with lower CD4 counts.

Application: Results (Cont'd)

QR model		$\tau = 0.25$ Est.	$\tau = 0.25$ 95% CI	$\tau = 0.5$ Est.	$\tau = 0.5$ 95% CI	$\tau = 0.75$ Est.	$\tau = 0.75$ 95% CI
Random effects	β_0	21.40	(20.40, 22.41)	23.62	(22.56, 24.68)	25.97	(24.86, 27.08)
	β_1	0.14	(−1.84, 2.11)	0.19	(−1.94, 2.31)	0.24	(−2.01, 2.47)
	β_2	−3.19	(−3.54, −2.86)	−3.54	(−3.89, −3.20)	−3.82	(−4.19, −3.46)
	β_3	0.78	(0.11, 1.45)	0.82	(0.13, 1.51)	0.83	(0.127, 1.55)
	λ_0	104.49	(90.94, 120.45)	117.98	(102.61, 135.76)	129.61	(112.56, 149.17)
	λ_1	9.49	(7.83, 11.38)	10.67	(8.86, 12.75)	11.09	(9.14, 13.36)
Shared parameter	β_0	21.09	(20.07, 22.08)	23.33	(22.24, 24.39)	25.58	(24.45, 26.70)
	β_1	0.17	(−1.86, 2.14)	0.21	(−1.92, 2.28)	0.27	(−1.98, 2.47)
	β_2	−3.40	(−3.76, −3.06)	−3.75	(−4.11, −3.40)	−4.09	(−4.47, −3.72)
	β_3	0.77	(0.06, 1.44)	0.81	(0.08, 1.52)	0.84	(0.11, 1.57)
	λ_0	105.87	(92.09, 121.59)	119.77	(104.40, 137.65)	132.02	(114.93, 151.51)
	λ_1	9.55	(7.92, 11.41)	10.95	(9.10, 13.12)	11.69	(9.68, 14.04)

Figura: Estimates and 95% CIs of the model parameters for the pediatric AIDS data

- $\hat{\beta}_3$ suggests that the low dose of zidovudine was superior to the high dose because the decline in the CD4 cell count was less steep in the low-dose arm.
- However, the superiority of the low-dose arm is slightly less for lower quantiles, as $\hat{\beta}_3 = 0.77$ versus 0.84 for $\tau = 0.25$ and 0.75, respectively.

Conclusion

- This approach does not pose any distributional assumption on the outcome variable, and is thus more robust than conventional mean regression.
- However, we do make a model assumption on the dropout process. Because the nonignorable missing data mechanism cannot be directly verified based on the observed data (Molenberghs et al., 2008).

References

- Yuan, Y., & Yin, G. (2009). Bayesian Quantile Regression for Longitudinal Studies with Nonignorable Missing Data. Biometrics, 66(1), 105–114.