Session 2
Quantile, M-quantile & Expectile Regression for Discrete Outcomes

Nikos Tzavidis (University of Southampton)
Ray Chambers (University of Wollongong)
James Dawber(University of Wollongong)
Nicola Salvati (University of Pisa)

Pisa, July, 2016

Outline

- Robust Estimation for Generalized Linear Models
- Quantile, M-quantile, Expectile regression for discrete outcomes
- Count Response
  - Quantile regression (Machado & Santos Silva, 2005, JASA)
  - Asymmetric Maximum Likelihood (Efron, 1992, JASA)
  - M-Quantile regression (Tzavidis et al., 2014, SMMR)
- Examples

Quantile & M-Quantile Regression for a Continuous Response: A Review

- Regression: model for the **mean** of y given x
  $\rightarrow E(y|x) = x^T\beta$
- Quantile regression: model for the **quantiles** of y given x
  $\rightarrow \mu_y(q|x) = x^T\beta_q$.
  (Koenker and Bassett, 1978, Econometrica)
- M-Quantile regression: $\mu_y(q|x;\psi) = x^T\beta_q$.
  (Breckling and Chambers, 1988, Biometrika)

Estimating Equations

For fixed q compute $\hat{\beta}(q)$ by

$$min \sum_{i=1}^{n} \rho_q \left(\frac{y_i - x_i^T \beta_q}{\sigma_q} \right)$$

- ρ_{LS} leads to linear expectile regression
- ρ_{LAV} leads to linear quantile regression
- ρ_{Huber} leads to linear M-quantile regression

Quantile Regression - A Likelihood Perspective (Yu & Moyeed, 2001, Statistics and Probability Letters)

Quantile Regression - A Likelihood Perspective

- Minimization of the ρ_{LAV} is equivalent to the maximization of a likelihood function formed by combining independently distributed Asymmetric Laplace densities

$$y \sim \ ALD(\mu, \sigma, q),$$

with pdf

$$f(y|\mu,\sigma,q) = \frac{q(1-q)}{\sigma} exp\Bigg(- \rho_q \Bigg(\frac{y-\mu}{\sigma} \Bigg) \Bigg)$$

- For fixed q the location parameter μ is modelled as a function of x.
- $\mu = x^T \beta_q$

Robust Estimation for Generalized Linear Models

Cantoni & Ronchetti (JASA, 2001)

- y_i from Exponential Family
- $E(y_i) = \mu_i$; $V(y_i) = V(\mu_i)$; $g(\mu_i) = x_i^T \beta$
- $\sum_{i=1}^{n} \frac{(y_i - \mu_i)}{V(\mu_i)} \frac{\partial}{\partial \beta} \mu_i = 0$, (McCullagh & Nelder, 1989)
- Large deviations of y_i from μ_i or leverage points$->$ influence
- $\sum_{i=1}^{n} \frac{\psi(r_i) w(x_i)}{V^{1/2}(\mu_i)} \frac{\partial}{\partial \beta} \mu_i - \alpha(\beta) = 0$ (Huber quasi-likelihood)
- r_i Pearson residuals; $w(x_i)$ controls leverage points
- Robust estimation for the Binomial and Poisson GLMs

Quantile, M-Quantile & Expectile Regression for Counts

1. Machado & Santos Silva (JASA, 2005; Quantile regression; Frequentist paradigm)
2. Lee & Neocleous (JRSS C, 2010; Quantile regression; Bayesian paradigm)
3. Efron (JASA,1992; Expectile regression via Asymmetric Maximum Likelihood)
4. Tzavidis et al.(Statistical Methods in Medical Research, 2014; M-Quantile regression). Extending Cantoni & Ronchetti (JASA,2001)

Quantile Regression for Counts

Machado & Santos Silva (JASA, 2005)

Quantiles of y are not continuous.

$$\mu_y(q) = \inf\left\{\delta \in \underbrace{supp(y)}_{=\mathbb{N}_0} | P(y \leq \delta) \geq q\right\}$$

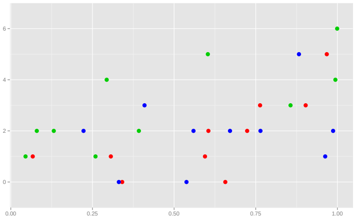
- Jittering: Impose artificial smoothness by adding to y random noise from a distribution with support in $(0, 1)$
- Here we use Uniform$(0, 1)$
- $1 - 1$ relationship between the conditional quantiles of the count and those of the jittered outcome
- Linear model (via ALD) for quantiles of jittered outcome
- Use monotonicity properties of quantiles
- Obtain estimates of the quantiles of y

1. Jittering the Response

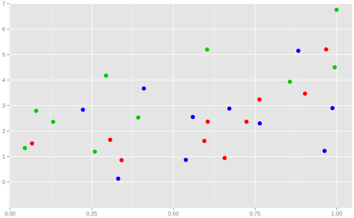
- Jittering: transform a discrete outcome into a continuous one,

$$z_i := y_i + u_i \qquad \text{with } u_i \sim U(0,1)$$
$$\text{independent from } y_i, x_i$$

y_i



$z_i := y_i + u_i$



2. Estimation of Quantiles for Jittered Outcome

- The quantile regression we use with the jittered outcome is

$$\mu_z(q|x_i) = q + exp(x_i^T \beta_q)$$

- This is because the quantiles of z are bounded below by q due to the addition of the uniform noise
- This can be implemented as an equivalent linear model on the following transformed outcome:

$$T(z_i, q) := \begin{cases} \log(\zeta) & , z_i \leq q \\ \log(z_i - q) & , z_i > q \end{cases}$$

with ζ a small positive constant

2. Estimation of Quantiles for Jittered Outcome

- A Linear quantile model for $T(z)$ is,

$$\mu_{T(z)}(q|x) = x_i^T \beta_q$$

- Estimation by using MLE via the ALD likelihood

2. Estimation of Quantiles for Jittered Outcome

- Jittering and estimating $\hat{\beta}_q$ is repeated M times leading to averaged jittering estimators,

$$\hat{\beta}_q = \frac{1}{M}\sum_{m=1}^{M} \hat{\beta}_q^{(m)}$$

- Estimated quantile function of z_i is

$$\widehat{\mu}_{z_i}(q|x_i) = \exp(x_i^T \hat{\beta}_q) + q$$

3. Quantiles Function of Discrete Outcome

Result

The q-quantile estimator for (the discrete) y_i is given by

$$\begin{aligned}\widehat{\mu}_y(q|x_i) &= \lceil \widehat{\mu}_z(q|x_i) - 1 \rceil \\ &= \lceil \exp(x_i^T \hat{\beta}_q) + q - 1 \rceil\end{aligned}$$

where $\lceil \cdot \rceil$ denotes the ceiling function

- See Theorem 2 (Machado & Santos Silva, JASA, 2005)

An Alternative View: Asymmetric Maximum Likelihood

Asymmetric Maximum Likelihood Estimation (Efron, JASA,1992)

Can be seen as the result of smoothing the objective function used to define the quantile regression estimator. Denote by D the deviance under a population model, e.g. Poisson and by D_w the weighted deviance

$$\begin{aligned} D_w(y,\mu) &= D(y,\mu), && y \leq \mu \\ D_w(y,\mu) &= wD(y,\mu), && y > \mu \end{aligned}$$

AML Details

Solution: Minimize $\sum_{i=1}^{n} D_w(y,\mu)$

- Leads to a version of quantile regression for counts

AML in More Detail

- Model for counts: Set $\mu_i = exp(x_i^T \beta_w)$

AML Details

Minimize

$$D_w(\beta_w) = 2\sum_{i=1}^{n}[y_i \log\left(\frac{y_i}{\mu_i(x_i^T \beta_w)}\right) - (y_i - \mu_i(x_i^T \beta_w))]w^{I(y_i > \mu_i(x_i^T \beta_w))}$$

$$\frac{\partial D_w(\beta_w)}{\partial \beta_w} = \sum_{i=1}^{n}\left[(y_i - \mu_i(x_i^T \beta_w))x_i\right]w^{I(y_i > \mu_i(x_i^T \beta_w))} = 0$$

- $w = 1 \rightarrow$ Poisson GLM

M-Quantiles for Counts: An Estimating Equation Approach

- Extending Cantoni & Ronchetti (JASA, 2001)
- Proposed in Tzavidis et al. (2014, Statistical Methods in Medical Research)

Let $\mu_y(q|x_i;\psi) = \exp(x_i^T\beta_q) = \mu_{iq}$. Estimate β_q by using the following estimating equation ,

$$\sum_{i=1}^{n}\Big[\psi_q\Big\{\frac{(y_i-\mu_{iq})}{V^{1/2}(\mu_{iq})}\Big\}w(x_i)\frac{1}{V^{1/2}(\mu_{iq})}\big(\frac{\partial\mu_{iq}}{\partial\beta_q}\big)-a(\beta_q)\Big]=0$$

- Estimation: Fisher scoring algorithm
- $V(\mu_{iq}) = \mu_{iq}$
- $a(\beta_q)$ computed under Poisson

Linking AML with the M-Quantile Extension of Cantoni & Ronchetti

AML Details

Start with Efron (1992): Minimize

$$D_w(\beta_w) = 2\sum_{i=1}^{n}[y_i \log\left(\frac{y_i}{\mu_i}\right) - (y_i - \mu_i)]w^{I(y_i>\mu_i)}$$

$$\frac{\partial D_w(\beta_w)}{\partial \beta_w} = \sum_{i=1}^{n}\left[(y_i - \mu_i)x_i\right]w^{I(y_i>\mu_i)} = 0$$

Linking AML with M-Quantile Extension of Cantoni & Ronchetti

Extension of Cantoni & Ronchetti

Consider the extension of Cantoni & Ronchetti:

$$\sum_{i=1}^{n}\Big[\psi_q\Big\{\frac{(y_i-\mu_{iq})}{V^{1/2}(\mu_{iq})}\Big\}w(x_i)\frac{1}{V^{1/2}(\mu_{iq})}\Big(\frac{\partial\mu_{iq}}{\partial\beta_q}\Big)-a(\beta_q)\Big]$$

With large tuning constant in Huber influence function and $w(x_i)=1$

$$\sum_{i=1}^{n}\Big[(y_i-\mu_{iq})w_{qi}x_i\Big]=0,$$

with

$$w_{qi}=\Big[\Big(\frac{q}{1-q}\Big)I(y_i>\mu_{iq})+I(y_i\leqslant\mu_{iq})\Big]$$

Corresponds to Efron (1992) with $w=\frac{q}{1-q}$

Comparing Different Approaches to Quantile Regression for Counts: An Example

- Generate data under the Poisson model:
- $\eta = 0.8 + 0.1x$; $y \sim Poisson(exp(\eta))$
- Compare
    - M-Quantile regression
    - AML regression
    - Quantile regression (via ALD & jittering)

Comparing Different Approaches to Quantile Regression for Counts: An Example

Table: Parameter estimates: Comparing different approaches to quantile regression for counts

Est.	MQ $c = 0.8$	QR	MQ$c = 100$	AML
$\hat{\beta}_{0q25}$	0.51	0.43	0.57	0.57
$\hat{\beta}_{1q25}$	0.08	0.09	0.08	0.08
$\hat{\beta}_{0q50}$	0.77	0.62	0.83	0.83
$\hat{\beta}_{1q50}$	0.09	0.12	0.08	0.08
$\hat{\beta}_{0q75}$	1.10	1.08	1.10	1.10
$\hat{\beta}_{1q75}$	0.08	0.08	0.07	0.07

R Packages - Count outcomes

Table: Quantile, M-quantile & Expectile Regression

Model	*R* Package
Poisson	*glm, family=Poisson*
Robust GLM	*glmrob, family=Poisson*,(*robustbase*)
Quantile (ALD)	*lqm.counts*, (*lqmm*)
M-Quantile	CountMQ [1]
Expectile	*amlpoisson*,(*VGAM*)

[1] Tzavidis et al., SMMR,2014