النموذج الهرمي والتوزيع الاولي للقوى في انحدار التقسيم البيزي

أ. م. د. طاهر ريسان دخيل
قسم الاحصاء/ كلية الادارة والاقتصاد
جامعة القادسية
tahir_reisan@yahoo.com

المستخلص

تم التحري في هذا البحث عن الصلة بين النماذج الهرمية والتوزيع الاولي للقوى في الانحدار التقسيمي. بمعلومية مستوى التقسيم، تم تطوير صيغة محددة لمعلمة القوة a_0 لموازنة التوزيع الاولي للقوى للانحدار التقسيمي مع النموذج الهرمي المماثل. اضافة الى ذلك فقد قدرنا العلاقة بين a_0 ومستوى التقسيم من خلال النموذج الهرمي. وقد تم توضيح المنهجية المقترحة من خلال استخدام بيانات حقيقية.

المصطلحات الرئيسية للبحث/ النموذج الهرمي- التوزيع اللاحق- التوزيع الاولي للقوى- الانحدار التقسيمي .

مجلة العلوم
الاقتصادية والإدارية
المجلد 20
العدد 76
لسنة ٢٠١٤
الصفحات ١-١٦

1 Introduction

Bayesian parameter estimation in quantile regression (QReg) is often a difficult issue because of a standered conjugate prior distribution is not available. To solve this problem, Alhamzawi and Yu 2011 extended the power prior distribution of Ibrahim and Chen (2000) for Bayesian quantile regression. This prior is a conjugate prior distribution for Bayesian quantile regression. In this paper, we examine the relation between the power prior and the hierarchical model in (QReg). We investigate the relation between the power parameter and the qauntile level via the hierarchical model.

(QReg) models have received considerable attention over the years (see, Koenker 2005; Yu et al. 2003; Cade et al 2003). Since Yu and Moyeed (2001) Bayesian inference quantile regression (BQReg) has attracted a lot of attention in literature (see, Hanson and Johnson 2002; Geraci and Bottai 2007; Yu and Stander 2007; Reed and Yu 2009; Lancaster and Jun 2010; Yuan and Yin 2010; Alhamzawi et al. 2011, Kozumi and Kobayashi 2011, Alhamzawi and Yu 2012, Alhamzawi and Yu 2013). However, the prior distribution plays the most important role in Bayesian quantile regression (BQReg). Since being introduced in Ibrahim and Chen (2000), the power prior distribution has become a popular technique to incorporate the historical data into the current data. This power prior distribution has been widely used for a variety of applications. The relation between the power prior distribution and hierarchical models in generalized linear models has been discussed by Chen and Ibrahim (2006). The authors found expressions for the power parameter to calibrate the power prior distribution to a corresponding hierarchical model.
Alhamzawi and Yu 2011 extended the power prior distribution of Ibrahim and Chen (2000) for Bayesian quantile regression.

The rest of this paper is organized as follows. In Section 2, we introduce the hierarchical model in (QReg) based on the mixture representation of the asymmetric Laplace distribution. In Section 3, we define the power prior distribution in (QReg) and we define the power prior distribution based on the mixture representation. In Section 4, we explain the behavior of the posterior under the power prior distribution in (QReg), In Section 5, we present the propriety of the power prior distribution in (QReg), the relation between the power prior distribution and the hierarchical model, and the relation between the power parameter and the quantile. In Section 6, we demonstrate the proposed methodology for obtaining the guide value for the power parameter with real data.

2- Hierarchical Model

Consider the regression model,

$$y_i = x_i'\beta_p + \varepsilon_i \qquad (1)$$

where y_i is the outcome variable, $x' = (x_{i1}, x_{i2}, \dots, x_{ik})$ represent the k independent variables, β_p is a k×1 vector of regression coefficients and ε_i, i=1,...,n represent error terms which are identical and independent distributions. The distribution of the error is assumed unknown and is restricted to have the p^{th} quantile equal to zero and 0<p<1.

Following Yu and Moyeed (2001), we consider εi has asymmetric Laplace distribution (ALD) with density

$$f(y_i|\beta_p) = p(1-p)\exp\{-\rho_p(y_i - x_i'\beta_p)\} \qquad (2)$$

Where p determines the quantile level and $\rho_p(u) = (p - I(u<0))u$.

As provided in Reed and Yu (2009) and Kozumi and Kobayashi (2009) that any variable has asymmetric Laplace distribution (ALD) with density (2) can be viewed as a mixture of an exponential and a scaled normal distribution given by

$$\varepsilon =^d (1-2p)v + \sqrt{2v}\xi, \qquad (3)$$

Where $v = [p(1-p)]^{-1}z$, z is a standard exponential variable, then it follows that each v_i has exponential distribution, $\exp(p(1-p)$, and ξ_i is a standard normal distribution. Now, the conditional distribution of each y_i given v_i is normal with mean $x_i'\beta_p + (1-2p)v_i$ and variance $2v_i$. Thus, the posterior density of β_p is given by.

$$f(\beta_p|y_i, v_i) \propto (v_i)^{-\frac{1}{2}} \exp\left\{-\frac{(y_i-(1-2p)v_i-x_i'\beta_p)^2}{4v_i}\right\}, \qquad (4)$$

and the complete data density of (y_i, v_i) is then given by

$$f(y_i, v_i|\beta_p) \propto (v_i)^{-\frac{1}{2}} \exp\left\{-\frac{(y_i-(1-2p)v_i-x_i'\beta_p)^2}{4v_i}\right\}\exp\{-p(1-p)v_i\}. \qquad (5)$$

Let $y = (y_1, \dots, y_n)$ and $v = (v_1, \dots, v_n)$, then the joint density of (y, v) is given by

$$f(y, v|\beta_p) = f(y|\beta_p, v)\pi(v).$$

If we integrating out the exponential variable, this leads to the likelihood

$$f(y|\beta_p) = \int f(y\,|\,\beta_p, v)\pi(v)dv. \qquad (6)$$

The model (1) with one historical dataset exist can be written as

$y_i = x_i'\beta_p + \varepsilon_i \,, i = 1, \dots, n,$ **and** $y_{0i} = x_{0i}'\beta_{0p} + \varepsilon_{0i} \,, i = 1, \dots, n_0,$ **(7)**

where β_p and β_{0p} denote the k regression coefficients for the current and historical study, respectively, $xi, and\ x_{0i}$ represent the k known covariates for the current and historical data, respectively, ε_i and ε_{0i} denote the error term associated with the subject i for the current and historical study, respectively.

Then we have the following hierarchical model

$$y_i = x_i'\beta_p + (1-2p)v_i + \sqrt{2v_i}\xi_i \quad \textbf{and} \quad y_{0i} = x_{0i}'\beta_{0p} + (1-2p)v_{0i} + \sqrt{2v_{0i}}\xi_{0i}$$

$$\beta_p|\mu_0, B_0 \sim N_k(\mu_0, B_0), \qquad \beta_{0p}|\mu_0, B_0 \sim N_k(\mu_0, B_0),$$

$$v_i \sim p(1-p)\exp\{-p(1-p)v_i\}, \quad v_{0i} \sim p(1-p)\exp\{-p(1-p)v_{0i}\},$$

$$\xi_i \sim \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\xi_i^2\right), \qquad \xi_{0i} \sim \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\xi_{0i}^2\right), \quad \textbf{(8)}$$

where μ_0 and B_0 are known fixed parameters, $v_{0i} = [p(1-p)]^{-1} z_{0i}$, z_{0i} is a standard exponential latent variable for the historical data, ξ_{0i} is a standard normal variable, and z_{0i} and ξ_{0i} are mutually independent.

3 - Power prior

Alhamzawi and Yu 2012 follow Ibrahim and Chen (2000) and define the power prior distribution for β_p in (QReg) for the current study as (Alhamzawi and Yu, 2012)

$$\pi(\beta_p|D_0, a_0) \equiv \frac{L(\beta_p|D_0)^{a_0}\pi_0(\beta_p)}{\int L(\beta_p|D_0)^{a_0}\pi_0(\beta_p)d\beta_p}$$

$$= \frac{L(\beta_p|D_0)^{a_0}\pi_0(\beta_p)}{g(a_0)}$$

$$\propto L(\beta_p|D_0)^{a_0}\pi_0(\beta_p)$$

$$= [\prod_{i=1}^{n_{0i}} p^{a_0}(1-p)^{a_0}\exp\{-a_0\rho_p(y_{0i} - x_{0i}'\beta_p)\}]\pi_0(\beta_p) \quad \textbf{(9)}$$

where D_0 represents the historical data, $0 \le a_0 \le 1$, a_0 determines by expert opinion, $L(\beta_p|D_0)$ denotes the likelihood function, $\pi_0(\beta_p)$ denotes the initial prior for β_p. Under the mixture representation (3), the joint power prior distribution for β_p and v_0 is given by.

$$\pi(\beta_p, \mathrm{v}_0|D_0, a_0) \propto \left(\prod_{i=1}^{n_{0i}} [f(\beta_p|y_{0i}, v_{0i})]^{a_0}\right)\pi_0(\beta_p), \quad (10)$$

where $f(\beta_p|y_{0i}, v_{0i})$ is (4) with (y_{0i}, v_{0i}) in place of (y_i, v_i), $\mathrm{v}_0 = (v_{01}, \dots, v_{0n_0})'$ and v_{0i} has exponential distribution, Exp(p(1− p)). The power priors (9) and (10) have several attractive properties. First, the power prior (9) can be obtained from the power prior (10) by integrating out the exponential variable. Further, the power priors (9) and (10) are always proper and have lower and upper bounds even if π_0 (β_p) is improper. In addition, the priors (9) and (10) depend on the quantile. The prior specification is completed by specifying a prior distribution for β_p . Let D denote to the current data and $\mathrm{v} = (v_1, \dots, v_n)'$, then the joint posterior distribution of β_p , v and v_0 is given by (Alhamzawi and Yu, 2012)

$$\pi(\beta_p, \mathrm{v}, \mathrm{v}_0|D, D_0, a_0) \propto \left(\prod_{i=1}^{n} [f(\beta_p|y_i, v_i)]\pi(v_i)\right)\left(\prod_{i=1}^{n_0} [f(\beta_p|y_{0i}, v_{0i})]^{a_0}\pi(v_{0i})\right)\pi_0(\beta_p), \quad (11)$$

4. Posterior Behavior under the power prior

To demonstrate the behavior of the marginal posterior distribution of β_p under the power prior with respect to different values for the power parameter. We simulate two data sets for the current and historical study. For the current study, 350 observation was generate from the model $y_i = 10 - x_i + \varepsilon_i$, where x_i was simulated from a uniform distribution on the interval (0, 10) and $\varepsilon_i \sim N(0,1)$.

For the historical data, 200 observation was generate from the model $y_{0i} = 9 - 1.5x_{0i} + \varepsilon_{0i}$, where x_{0i} was simulated from a uniform distribution on the interval (0, 10) and $\varepsilon_i \sim N(0,1)$. We take a multivariate normal distribution with mean zero and variance covariance matrix $B_0 = 100I$ as initial prior for β_p. Figure 1 and 2 compared the marginal posterior densities for $\beta_{(0)_p}$ and $\beta_{(1)_p}$ for p=95% and 75% respectively m for improper prior with the posterior densities of $\beta_{(0)_p}$ and $\beta_{(1)_p}$ for the power prior. Clearly, the power prior is more informative than improper prior, due to the small range of posterior densities.

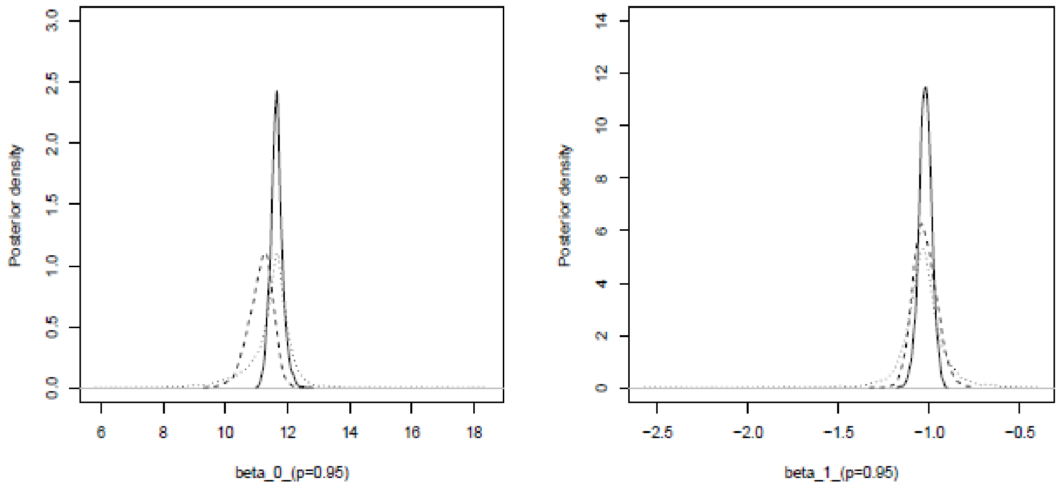

Figure 1: plots of posterior densities for $\beta_{0.95}$ where the dotted curve is for improper uniform prior (a_0= 0), the dashed and solid curves are for power priors with power parameter a_0= 0.50 and 0.90 respectively.

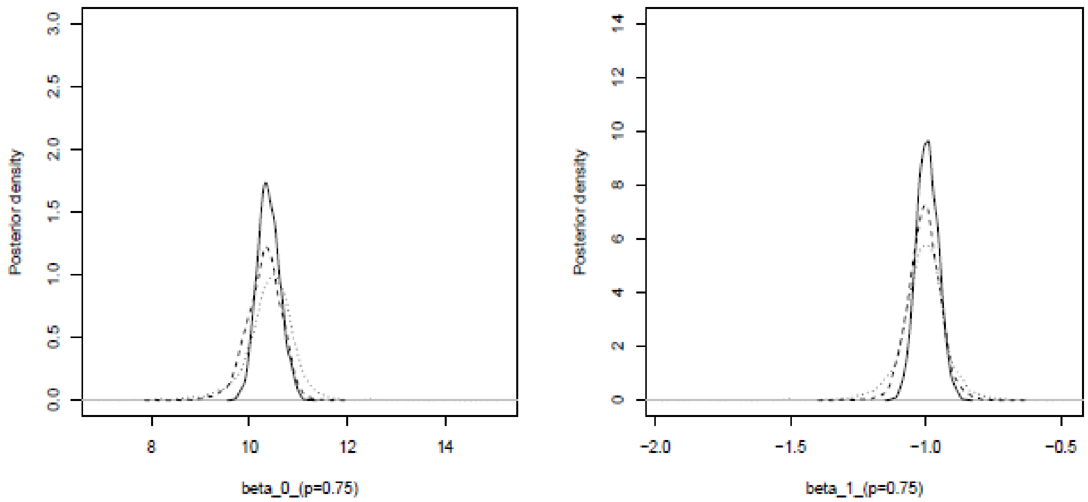

Figure 2: plots of posterior densities for $\beta_{0.75}$ where the dotted curve is for improper uniform prior (a_0= 0), the dashed and solid curves are for power priors with power parameter a_0= 0.50 and 0.90 respectively.

5. Main results

The power prior proposed by Ibrahim and Chen (2000) has been constructed to be a useful class of informative prior in Bayesian analysis. This prior depend on the availability of the historical data, and in the context of Bayesian analysis when such data is available the prior would be better proper because it is well known that any informative Bayesian analysis requires a proper prior distribution, thus the propriety of the power prior is of critical importance. In this section we discuss the relation between power priors and hierarchical models. Thus, we present Lemma 1 and 2 to introduce the marginal posterior distribution for β_p under the power prior, Lemma 3 introduce the marginal posterior distribution for β_p under the hierarchical model and Lemma 4 discuss the relation between power priors and hierarchical models.

Lemma 1. The marginal posterior distribution of β_p under the power prior (10) with multivariate normal distribution as initial prior for β_p is given by

$$\beta_p|y,X,V,y_0,X_0,V_0,a_0 \sim N_k(A^{-1}B,A^{-1}),$$

Where

$$A=\frac{1}{2}X'VX+\frac{a_0}{2}X_0'V_0X_0+B_0^{-1},$$

$$B=\frac{1}{2}X'Vu+\frac{a_0}{2}X_0'V_0u_0+B_0^{-1}\mu_0, \qquad (12)$$

here, $y=(y_1,\dots,y_n)$, $X=(x_1,\dots,x_n)'$, $y_0=(y_{01},\dots,y_{0n_0})$, $X_0=(x_{01},\dots,x_{0n_0})'$, $V=diag(v_1,\dots,v_n)$, $V_0=diag(v_{01},\dots,v_{0n_0})$, $u=(u_1,\dots,u_n)$, $u_0=(u_{01},\dots,u_{0n_0})$, **with** $u_i=y_i-(1-2p)v_i$, **and** $u_{0i}=y_{0i}-(1-2p)v_{0i}$.

The proof of Lemma (1) with the details of the Gibbs sampler is given in Appendix.

Lemma 2. The marginal posterior distribution of β_p under the power prior (10) with uniform prior distribution as initial prior for β_p is given by

$$\beta_p|y,X,V,y_0,X_0,V_0,a_0 \sim N_k(A_1^{-1}B_1,A_1^{-1}),$$

Where

$$A_1=\frac{1}{2}X'VX+\frac{a_0}{2}X_0'V_0X_0,$$

$$B_1=\frac{1}{2}X'Vu+\frac{a_0}{2}X_0'V_0u_0, \qquad (13)$$

The proof of Lemma (2) is similar to proof of Lemma (1).

Lemma 3. The marginal posterior distribution of β_p for the hierarchical model (8) in (QReg) is given by

$\beta_p|y,X,V,y_0,X_0,V_0 \sim N_k(A_2^{-1}B_2,A_2^{-1}),$

Where

$A_2=\frac{1}{2}X'VX+B_0^{-1}-(2B_0-\left(B_0^{-1}+\frac{1}{2}X_0'V_0X_0\right)^{-1})^{-1},$

$$B_2=\frac{1}{2}X'V_u+\left(2B_0-\left(B_0^{-1}+\frac{1}{2}X_0'V_0X_0\right)^{-1}\right)^{-1}\left(B_0^{-1}+\frac{1}{2}X_0'V_0X_0\right)^{-1}\frac{1}{2}X_0'V_0\text{u}_0.$$

The proof of Lemma 3 in Appendix.

Lemma 4. The posterior distributions of the quantile coefficient β_p given in Lemma 2 and 3 are identical distributions if and only if

$$a_0(I+B_0X_0V_0X_0)=I \qquad (14)$$

Proof: Similar to Chen and Ibrahim (2006), If $A_1=A_2$ then we have

$\frac{1}{2}X'VX+\frac{a_0}{2}X_0'V_0X_0=\frac{1}{2}X'VX+B_0^{-1}-(2B_0-\left(B_0^{-1}+\frac{1}{2}X_0'V_0X_0\right)^{-1})^{-1},$

and this lead to $2^{-1}a_0B_0X_0'V_0X_0=I-(2I-(I+2^{-1}B_0X_0'V_0X_0)^{-1})^{-1}$. a little algebra shows $a_0B_0X_0'V_0X_0[B_0X_0'V_0X_0+I]=B_0X_0'V_0X_0$

$$a_0(I+B_0X_0'V_0X_0)=I \qquad (15)$$

Similarly, it can be shown that $B_1=B_2$ if and only if $a_0[B_0X_0'V_0X_0+I]=I$.

Like Chen and Ibrahim (2006), we can use the connection between the power prior and the hierarchical model to specify a guide value for a_0 in (QReg). To achieve this we can write equation (15) as $a_0[B_0X_0'Z_0X_0+p(1-p)I]=p(1-p)I$, where $Z_0=diag\left(z_{01},\dots,z_{0n_0}\right)$.

Since Z_0 is random then the guide value for a_0 is the posterior expectation of

$\frac{kp(1-p)}{kp(1-p)+tr(B_0X_0'Z_0X_0)}$

That is,

$$\hat{a}_0=E\left(\frac{kp(1-p)}{kp(1-p)+tr(B_0X_0'Z_0X_0)}\right) \qquad (16)$$

The posterior expectation is taken with respect to Z_0, where B_0 is constant . Equation (16) reflects the relation between the power parameter and the quantile level.

6. Wage data

We consider data from the British Household Panel Survey. The data represent the wage distribution among British workers which was previously analyzed by Yu et al. (2005) and Alhamzawi and Yu (2012). Similar to Alhamzawi and Yu (2012), we use the data for the year 2000 as current data and for year 1994 as historical data. Our model is described as follows

$$\ln(Y_i) = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 D_i,$$

Where S_i is the number of years of schooling, E_i is the potential experience (approximated by the age minus years of schooling minus 6), and D_i is equal to 1 for public sector workers and 0 otherwise. We consider (QReg) model to fit the current and the historical data. We take B_0 to be a fixed diagonal matrix such that $B_0 = 100I$. We use Gibbs sampler to sample β_p and β_{0p} from their respective distribution. We specify $\hat{a}_0$ from equation (16). Table 1 summarizes the posterior mean for β_p and β_{0p} under the hierarchical model. The posterior distribution for β_p under the power prior are summarized in Table 2 for different value for the power parameter including $\hat{a}_0$. Clearly, the posterior distribution for the regression parameter under the power prior with $\hat{a}_0$ are fairly close to those obtained under the hierarchical model.

7. Appendixes

Proof Lemma 1. First, consider equation (4). In vector notation, the likelihood function for current and historical data are, respectively, given by

$$f(y|\beta_p,\sigma,\mathrm{v}) \propto \sigma^{-\frac{n}{2}}(\prod_{i=1}^{n} v_i^{-\frac{1}{2}})\exp(-\tfrac{1}{4}(u - X\beta_p)'V(u - X\beta_p) \quad (17)$$

$$f(y_0|\beta_p,\sigma,\mathrm{v}_0) \propto \sigma^{-\frac{n_0}{2}}(\prod_{i=1}^{n_0} v_{0i}^{-\frac{1}{2}})\exp(-\tfrac{1}{4}(u_0 - X_0\beta_p)'V_0(u_0 - X_0\beta_p) \quad (18)$$

here, $y = (y_1, \ldots, y_n)$, $X = (x_1, \ldots, x_n)'$, $y_0 = (y_{01}, \ldots, y_{0n_0})$, $X_0 = (x_{01}, \ldots, x_{0n_0})'$, $V = diag(v_1, \ldots, v_n)$, $V_0 = diag(v_{01}, \ldots, v_{0n_0})$, $\mathrm{u} = (u_1, \ldots, u_n)$, $\mathrm{u}_0 = (u_{01}, \ldots, u_{0n_0})$, $u_i = y_i - (1-2p)v_i$, **and** $u_{0i} = y_{0i} - (1-2p)v_{0i}$.

The posterior distribution of β_p can then be calculated using the joint posterior distribution in equation (11). We have

$$f(\beta_p, \mathrm{v}, \mathrm{v}_0|D, D_0, a_0) \propto (\prod_{i=1}^{n} v_i^{-\frac{1}{2}})\exp(-\frac{1}{4}(u - X\beta_p)'V(u - X\beta_p)$$
$$\times \prod_{i=1}^{n} \exp\{-p(1-p)v_i\}$$
$$\times (\prod_{i=1}^{n_0} v_{0i}^{-\frac{a_0}{2}})\exp(-\tfrac{a_0}{4}(u_0 - X_0\beta_p)'V_0(u_0 - X_0\beta_p))$$

$$\times \prod_{i=1}^{n_0} \exp\{-p(1-p)v_{0i}\}$$
$$\times \exp\{-\tfrac{1}{2}(\beta-\mu_0)' B_0^{-1}(\beta-\mu_0)\} \qquad \textbf{(19)}$$

From (19), the full conditional distribution of β_p is given by

$$f(\beta_p|\sigma, \mathrm{v}, \mathrm{v}_0, D, D_0, a_0)$$
$$\propto \exp\left\{-\frac{1}{4}(u-X\beta_p)'V(u-X\beta_p) - \frac{a_0}{4}(u_0-X_0\beta_p)'V_0(u_0-X_0\beta_p)\right\}$$
$$\times \exp\{-\frac{1}{2}(\beta_p-\mu_0)'B_0^{-1}(\beta_p-\mu_0)\}$$

We have $(u-X\beta_p)'V(u-X\beta_p)$ and $a_0(u_0-X_0\beta_p)'V_0(u_0-X_0\beta_p)$ into sum of squares

$$(u-X\beta_p)'V(u-X\beta_p) = (u-X\beta_p^*)'V(u-X\beta_p^*) + (\beta_p-\beta_p^*)'X'VX(\beta_p-\beta_p^*),$$
$$a_0(u_0-X_0\beta_p)'V_0(u_0-X_0\beta_p)$$
$$= a_0\Big((u_0-X_0\beta_p^*)'V_0(u_0-X_0\beta_p^*) + (\beta_p-\beta_p^*)'X_0'V_0X_0(\beta_p-\beta_p^*)\Big).$$

We set $X'VX\beta_p^* = X'V$, and $X_0'V_0X_0\beta_p^* = X_0'V_0\mathrm{u}_0$. Then the posterior distribution of β_p is given by

$$\beta_p|y, X, V, y_0, X_0, V_0, a_0 \sim N_k(A^{-1}B, A^{-1}),$$

Where

$$A = \frac{1}{2}X'VX + \frac{a_0}{2}X_0'V_0X_0 + B_0^{-1},$$
$$B = \tfrac{1}{2}X'V_u + \tfrac{a_0}{2}X_0'V_0u_0 + B_0^{-1}\mu_0,$$

To complete our MCMC-based computation technique, the full conditional distribution of v_{0i} is given by

$$\pi(v_{0i}|\beta_p, a_0, D_0)$$
$$\propto (v_{0i})^{-a_0/2}\exp\left\{-\tfrac{a_0}{4v_{0i}}\left(y_{0i} - x_{0i}'\beta_p - (1-2p)v_{0i}\right)^2 - p(1-p)v_{0i}\right\}$$
$$\propto (v_{0i})^{-a_0/2}\exp\left\{-\tfrac{a_0}{4v_{0i}}(y_{0i}-x_{0i}'\beta_p)^2 - \left(\tfrac{a_0(1-2p)^2}{4}\right) + p(1-p))v_{0i}\right\}$$
$$= v_{0i}^{-a_0/2}\exp\left\{-\tfrac{1}{2}[\tfrac{a_0(y_{0i}-x_{0i}'\beta_p)^2}{2}v_{0i}^{-1} + (\tfrac{a_0(1-2p)^2}{2} + 2p(1-p))v_{0i}]\right\}$$

Thus, the full conditional distribution of v_{0i} is a generalized inverse Gaussian (GIG) distribution. In the same way we can deduce that the full conditional distribution of v_{0i} is also GIG distribution; that is,

$$\pi(v_i|\beta_p) \propto v_i^{-1/2}\exp\left\{-\frac{1}{2}[\frac{(y_i-x_i'\beta_p)^2}{2}v_i^{-1} + (\frac{1}{2})v_i]\right\}$$

Proof of Lemma 3.

$$\pi(\beta_p, \beta_{0p}, \mu_0 | y, y_0, \mathrm{v}, \mathrm{v}_0) \propto$$

$$\exp\left\{-\frac{1}{4}(u - X\beta_p)' V(u - X\beta_p)\right\} \exp\left\{-\frac{1}{4}(u_0 - X_0\beta_p)' V_0(u_0 - X_0\beta_p)\right\}$$

$$\times \exp\left\{-\frac{1}{2}(\beta_p - \mu_0)' B_0^{-1}(\beta_p - \mu_0)\right\} \exp\left\{-\frac{1}{2}(\beta_p - \mu_0)' B_0^{-1}(\beta_p - \mu_0)\right\}.$$

After Integrating out β0p and rearrange the terms, we can get

$$\pi(\beta_p, \mu_0 | y, y_0, \mathrm{v}, \mathrm{v}_0) \propto \exp\left\{-\frac{1}{4}(u - X\beta_p)' V(u - X\beta_p) - \frac{1}{2}(\beta_p - \mu_0)' B_0^{-1}(\beta_p - \mu_0)\right\}$$

$$\times \exp\left\{-\frac{1}{2}[\mu_0'\left(B_0^{-1} - \left(B_0 + \frac{1}{2}B_0X_0'V_0X_0B_0\right)^{-1}\right)\mu_0 - \mu_0'\left(I + \frac{1}{2}X_0'V_0X_0B_0\right)^{-1} X_0'V_0u_0]\right\}.$$

then, integrating out μ_0 leads to

$$\pi(\beta_p | y, y_0, \mathrm{v}, \mathrm{v}_0) \propto \{ -\tfrac{1}{2}[\beta_p'(\tfrac{1}{2}X'VX + B_0^{-1} - \left(2B_0 - \left(B_0^{-1} + \tfrac{1}{2}X_0'V_0X_0\right)^{-1}\right)^{-1})\beta_p$$

$$-\beta_p'X'Vu + \left(2B_0 - \left(B_0^{-1} + \frac{1}{2}X_0'V_0X_0\right)^{-1}\right)^{-1}\left(B_0^{-1} + \frac{1}{2}X_0'V_0X_0\right)^{-1} X_0'V_0u_0)]\}.$$

Then we have

$$\beta_p | y, X, V, y_0, X_0, V_0 \sim N_k(A_2^{-1}B_2, A_2^{-1}),$$

Where

$$A_2 = \frac{1}{2}X'VX + B_0^{-1} - (2B_0 - \left(B_0^{-1} + \frac{1}{2}X_0'V_0X_0\right)^{-1})^{-1},$$

$$B_2 = \tfrac{1}{2}X'V_u + \left(2B_0 - \left(B_0^{-1} + \tfrac{1}{2}X_0'V_0X_0\right)^{-1}\right)^{-1}\left(B_0^{-1} + \tfrac{1}{2}X_0'V_0X_0\right)^{-1}\tfrac{1}{2}X_0'V_0\mathrm{u}_0.$$

Table 1: Posterior estimates of $\beta_{(p)}$ and $\beta_{0(p)}$ under the hierarchical model.

p	parameter	posterior mean	parameter	posterior mean
0.95	$\beta_{0(p)}$	7.163	$\beta_{00(p)}$	7.188
	$\beta_{1(p)}$	0.027	$\beta_{01(p)}$	0.040
	$\beta_{2(p)}$	0.004	$\beta_{02(p)}$	0.019
	$\beta_{3(p)}$	-0.103	$\beta_{03(p)}$	-0.125
0.75	$\beta_{0(p)}$	6.840	$\beta_{00(p)}$	6.798
	$\beta_{1(p)}$	0.015	$\beta_{01(p)}$	0.021
	$\beta_{2(p)}$	0.009	$\beta_{02(p)}$	0.011
	$\beta_{3(p)}$	-0.031	$\beta_{03(p)}$	-0.028
0.50	$\beta_{0(p)}$	6.799	$\beta_{00(p)}$	6.806
	$\beta_{1(p)}$	0.020	$\beta_{01(p)}$	0.023
	$\beta_{2(p)}$	0.003	$\beta_{02(p)}$	0.008
	$\beta_{3(p)}$	0.012	$\beta_{03(p)}$	0.063
0.25	$\beta_{0(p)}$	6.572	$\beta_{00(p)}$	6.528
	$\beta_{1(p)}$	0.022	$\beta_{01(p)}$	0.019
	$\beta_{2(p)}$	0.006	$\beta_{02(p)}$	0.006
	$\beta_{3(p)}$	0.066	$\beta_{03(p)}$	0.097
0.05	$\beta_{0(p)}$	6.334	$\beta_{00(p)}$	6.313
	$\beta_{1(p)}$	0.019	$\beta_{01(p)}$	0.020
	$\beta_{2(p)}$	0.003	$\beta_{02(p)}$	0.004
	$\beta_{3(p)}$	0.098	$\beta_{03(p)}$	0.114

Table 2: Posterior estimates of $\beta_{(p)}$ under the power prior distribution.

p	a_0	$\beta_{0(p)}$ (95% CrI)	$\beta_{1(p)}$ (95% CrI)	$\beta_{2(p)}$ (95% CrI)	$\beta_{3(p)}$ (95% CrI)
0.95	0	6.032 (5.933, 6.147)	0.170 (0.112, 0.226)	0.013 (0.010, 0.033)	-0.347 (-0.834, 0.121)
	0.731	7.170 (7.138, 7.205)	0.027 (0.015, 0.043)	0.016 (0.002,0.027)	-0.119 (-0.138, -0.098)
	1	7.493 (7.434, 7.553)	0.031 (0.028, 0.035)	0.011 (0.009, 0.012)	-0.169 (-0.196, -0.147)
0.75	0	5.941 (5.853, 6.027)	0.071 (0.063, 0.125)	0.011 (0.009, 0.023)	0.027 (0.012, 0.048)
	0.49	6.976 (6.923, 7.028)	0.021 (0.019, 0.024)	0.010 (0.008, 0.011)	-0.004 (-0.035, 0.025)
	1	7.012 (6.953, 7.071)	0.028 (0.025, 0.031)	0.010 (0.008, 0.011)	-0.030 (-0.056, -0.006)
0.50	0	5.819 (5.665, 5.962)	0.030 (0.022, 0.054)	0.004 (0.001, 0.013)	0.037 (-0.019, 0.114)
	0.337	6.834 (6.702, 6.976)	0.021 (0.015, 0.026)	0.007 (0.001,0.011)	0.030 (-0.051, 0.116)
	1	6.851 (6.693, 6.991)	0.022 (0.016, 0.028)	0.007 (0.003, 0.011)	0.033 (-0.050, 0.113)
0.25	0	5.615 (5.535, 5.702)	0.013 (0.009, 0.019)	0.002 (0.001, 0.036)	0.041 (0.005, 0.117)
	0.51	6.557 (6.515, 6.600)	0.021 (0.019, 0.023)	0.006 (0.005, 0.007)	0.071 (0.042, 0.104)
	1	6.492 (6.457, 6.531)	0.018 (0.016, 0.019)	0.006 (0.005, 0.007)	0.061 (0.032, 0.092)
0.05	0	5.479 (5.395, 5.546)	0.006 (0.001, 0.013)	-0.005 (-0.003, 0.006)	0.091 (0.035, 0.153)
	0.57	6.337 (6.307, 6.365)	0.019 (0.017, 0.021)	0.004 (0.003, 0.005)	0.104 (0.081,0.127)
	1	5.895 (5.859, 5.927)	0.013 (0.012, 0.015)	0.004 (0.003, 0.006)	0.216 (0.181, 0.248)

References

Alhamzawi, R. and Yu, K. (2011). "Power Prior Elicitation in Bayesian Quantile Regression". Journal of Probability and Statistics, vol. 2011, Article ID 874907, 16 pages, 2011.

Alhamzawi, R., K. Yu, and D. F. Benoit (2011). "Bayesian adaptive Lasso quantile regression". Statistical Modelling 12, 279–297.

Alhamzawi, R. and Yu, K. (2012). "Variable selection in quantile regression via Gibbs sampling". Journal of Applied Statistics, 39: 799-813.

Alhamzawi, R. and K. Yu (2013). "Conjugate priors and variable selection for Bayesian quantile regression". Computational Statistics & Data Analysis, in press.

Cade, Brian S., and Barry R. Noon. "A gentle introduction to quantile regression for ecologists." *Frontiers in Ecology and the Environment* 1.8 (2003): 412-420.

Chen, M. H. and J. G. Ibrahim (2006). "The relationship between the power prior and hierarchical models". Bayesian Analysis 1, 551–574.

Dunson, D. B. and J. A. Taylor (2005). "Approximate bayesian inference for quantiles". Nonparametric Statistics 17, 385–400.

Geraci, M. and M. Bottai (2007). "Quantile regression for longitudinal data using the asymmetric Laplace distribution". Biostatistics 8 (1), 140–154.

Hanson, T. and W. O. Johnson (2002). "Modeling regression error with a mixture of polya trees". Journal of the American Statistical Association 97, 1020–1033.

Ibrahim, J. G. and M. H. Chen (2000). "Power prior distributions for regression models". Statist. Sci 15 (1), 46–60.

Koenker, R. (2005). "Quantile Regression". New York: Cambridge University Press.

Kottas, A. and M. Krnja jic (2009). "Bayesian semiparametric modeling in quantile regression". Scandinavian Journal of Statistics 36, 297–319.

Kozumi, H. and G. Kobayashi (2009). "Gibbs sampling methods for Bayesian quantile regression". Technical report, Graduate school of Business Administration, Kobe University, 2-1 Rokko, Kobe 657-8501, Jaban.

Lancaster, T. and S. J. Jun (2010). "Bayesian quantile regression methods". Journal of Applied Econometrics 25, 287–307.

Li, Q., R. Xi, and N. Lin (2010). "Bayesian regularized quantile regression". Technical report, Department of Mathematics, Washington University.

Reed, C. and K. Yu (2009). "A partially collapsed Gibbs sampler for Bayesian quantile regression". Technical report, Department of Mathematical Sciences, Brunel University.

Scaccia, L. and P. J. Green (2003). "Bayesian growth curves using normal mixtures with nonparametric weights". Journal of Computational and Graphical Statistics 12, 208–331.

Schennach, S. (2005)."Bayesian exponentially tilted empirical likelihood". Biometrika 92, 31–46.

Tsionas, E. G. (2003). "Bayesian quantile inference". Journal of statistical computation and simulation 73, 659–674.

Yu, K., P. V. Kerm, and J. Zhang (2005). "Bayesian quantile regression: An application to the wage distribution in 1990s Britain". Indian Statistical Institute 67 (2), 359–377.

Yu, K., Z. Lu, and J. Stander (2003). "Quantile regression: applications and current research area". The Statistician 52, 331–350.

Yu, K. and R. A. Moyeed (2001). "Bayesian quantile regression". Statistics and Probability Letters. 54, 437–447.

Yu, K. and J. Stander (2007). "Bayesian analysis of a tobit quantile regression model". Journal of Econometrics 137, 260–276.

Yuan, Y. and G. Yin (2010). "Bayesian quantile regression for longitudinal studies with no ignorable missing data". Biometrics 66, 105–114.

A Note on the Hierarchical Model and Power Prior Distribution in Bayesian Quantile Regression

Abstract.

In this paper, we investigate the connection between the hierarchical models and the power prior distribution in quantile regression (QReg). Under specific quantile, we develop an expression for the power parameter (a_0) to calibrate the power prior distribution for quantile regression to a corresponding hierarchical model. In addition, we estimate the relation between the a_0 and the quantile level via hierarchical model. Our proposed methodology is illustrated with real data example.

Keywords: Hierarchical model; Posterior distribution; Power prior; Quantile Regression (QReg).