Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

A Study of the Half-Cauchy-Exponential Prior in Quantile Regression

Mohammed Obayes Kadhim[a], Ahmed Alhamzawi[b]*

[a]*University of AL-Qadisiyah, College of Computer Science and Information Technology, Iraq. Email: m.m.96moh@gmail.com.*

[b]*University of AL-Qadisiyah, College of Science, Department of Mathematics, Iraq. Email: ahmed.alhamzawi@qu.edu.iq.*

ARTICLE INFO

Article history:
Received: 30 /10/2023
Rrevised form: 2 /12/2023
Accepted : 15 /12/2023
Available online: 30 /12/2023

Keywords:

Quantile Regression,

Half-Cauchy Distribution,

Exponential distribution,

Gibbs Sampler.

ABSTRACT

A modification of the familiar half-Cauchy prior is considered. The modification consists of writing the half-Cauchy prior as the product of the gamma and inverse gamma distributions plus adding an exponential distribution on the scale parameter. Furthermore, we consider this model in the setting of the quantile regression structure. Additionally, The Gibbs sampler is calculated for this model. Finally, the properties of this model are demonstrated using simulated data and it is shown that this method performs very well compared to other distributions.

MSC..

https:// 10.29304/jqcsm.2023.15.41352

1. Introduction

Three are many advantages of using quantile regression over mean regression [1,7]. Thus, it is objective to use this type of regression with a new type prior. Let $y = (y_1, \cdots, y_n)^T$ be a vector of dependent variables, with a $n \times p$ design matrix of independent variables $X = (x_1, \cdots, x_p)$, a $p \times 1$ of vector of unknown regression coefficient $\beta = (\beta_1, \cdots, \beta_p)^T$ and $\epsilon = (\epsilon_1, \cdots, \epsilon_n)^T$ where $\epsilon_i \sim N(0, \sigma^2)$, then we can write our linear model as

$$y = X\beta + \epsilon, \qquad \textbf{(1)}$$

The basic framework of frame work of quantile regression is to express the wth quantile regression model by defining the inverse cumulative distribution function $Q_{y_i}(w|x_i)$ of y_i given x_i as

*Corresponding author: Mohammed Obayes Kadhim

Email addresses: : m.m.96moh@gmil.com

Communicated by 'sub etitor'

$$Q_{y_i}(w|x_i) = \beta_0 + x_i'\beta. \tag{2}$$

where we have chosen a Jefferies prior for the variance σ^2. To find the values of the coefficient β, we use the same minimization methods in mean regression but with small modifications by the form

$$\sum_{i=1}^{n} \rho_w\,(y_i - x_i^T\beta), \tag{3}$$

where we define the quantile check loss function of ρ_w as

$$\rho_w(x) = xw - xI(x \leq 0) \tag{4}$$

or written in another way

$$\rho_w(s) = \{ws, if\;\; s \geq 0, (w-1)s, if\;\; s < 0. \tag{5}$$

Instead of using the minimization of (3), we could use the Bayesian way to maximize the likelihood function where our error distribution is given by the asymmetric Laplace distribution

$$f(\epsilon_n|\tau) = w(1-w)\tau\{-\tau\rho_w(\epsilon_n)\}, \tag{6}$$

with the probability of the dependent variable y is obtained by

$$f(y|X,\beta,w) = \prod_{i=1}^{n} w(1-w)\tau\{-\tau\rho_w(y_i - x_i^T\beta)\} \tag{7}$$

To maximize (7), we will use the work of [9] by writing the asymmetric Laplace distribution as a mixture of an exponential distribution with a scaled type of the normal distribution and therefore we get

$$y_i = x_i^T\beta + \eta_1 s_i + \eta_2\tau^{-1/2}\sqrt{s_i}q_i \tag{8}$$

$$s|\tau \sim \prod_{i=1}^{n} \tau(-\tau s_i) \tag{9}$$

$$q \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\left(-\frac{1}{2}q_i^2\right) \tag{10}$$

such that $s = (s_1, \dots, s_n)$, $q = (q_1, \dots, q_n)$

$$\eta_1 = \frac{1-2w}{w(1-w)} \qquad and \qquad \boldsymbol{\eta}_2 = \sqrt{\frac{2}{w(1-w)}}. \tag{11}$$

Next, we will discuss the properties of this model and derive its relevant algorithm, then compare its result with other known models.

2. The regression prior

In this paper, we consider a modification the horseshoe prior proposed in [8,9] presented by the following hierarchical form

$$\beta_i|\sigma^2, \omega_i \sim N(0, \sigma^2\omega_i),$$
$$\omega_i^{1/2} \sim C^+(0, \gamma), \quad \textbf{(12)}$$

by placing an exponential prior in the parameter γ. We will try to compare the result of this model with a family models. We can simplify this equation by writing the above model as

$$\beta_i|\sigma^2, \omega_i \sim N(0, \sigma^2\omega_i\Omega_i\gamma_i),$$
$$\omega_i \sim G\left(\frac{1}{2}, 1\right),$$
$$\Omega_i \sim IG\left(\frac{1}{2}, 1\right), \quad \textbf{(13)}$$
$$\gamma_i \sim Exp(\lambda),$$

where we have used the fact that the half-cauchy prior can be written as the product of the gamma and inverse gamma distributions. Now, we can write our full hierarchal model as

$$y_i = x_i^T\beta + \xi_1 s_i + \xi_2\tau^{-1/2}\sqrt{s_i}q_i \quad \textbf{(14)}$$

$$s|\tau \sim \prod_{i=1}^{n} \tau(-\tau s) \quad \textbf{(15)}$$

$$q \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\left(-\frac{1}{2}q_i^2\right) \quad \textbf{(16)}$$

$$\beta_i|\omega_i, \Omega_i, \tau \sim N(0, \tau^{-1}\prod_{i=1}^{N} z_i) \quad \textbf{(17)}$$

$$\tau \sim G(c_0, d_0) \quad \textbf{(18)}$$

Next, we will present a method for updating the hyperparameters that will be added to the Gibbs sampler.

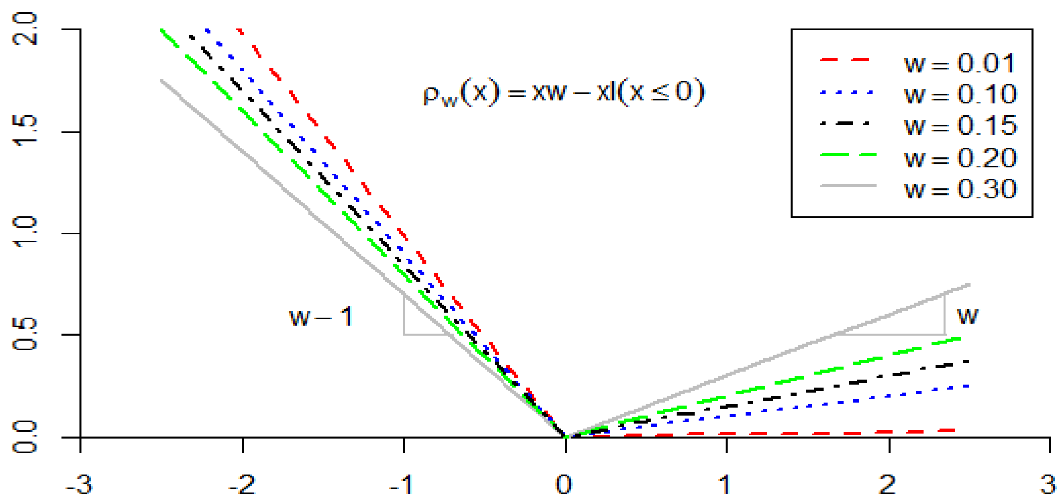

Figure 1. A plot showing check loss function ρ_w for different values of w.

3. The Sampler

We will derive the conditional posteriors of our prior as follows:

- **For** β

$$P(\beta|X, y, \dots) \propto P(y|\beta, \dots)\pi(\beta), \propto \left\{-\frac{(y - X\beta - \eta_1 v)^T S^{-1}(y - X\beta - \eta_1 s)}{2\tau^{-1}\eta_2^2}\right\} \times \left\{-\tau\frac{\beta^T D^{-1}\beta}{2}\right\} \\ \propto \left\{-\frac{\tau}{2}[-2\eta_2^{-2}(y - \eta_1 s)^T S^{-1}X\beta + \eta_2^{-2}\beta^T X^T S^{-1}X\beta + \beta^T D^{-1}\beta]\right\} \quad \textbf{(19)}$$

with the mean $\mu = \eta_2^{-2}\Sigma^{-1}X^T S^{-1}(y - \eta_1 s)$, $S = diag(s_1, \dots, s_n)$, $D = daig(\omega_1\Omega_1\gamma_1, \dots, \omega_p\Omega_p\gamma_p)$ and variance $\Sigma = \eta_2^{-2}X^T S^{-1}X + D^{-1}$. Therefore, we have the normal distribution

$$\beta|X, y, \dots \sim N(\mu, \Sigma^{-1}\sigma^2). \quad \textbf{(20)}$$

- **For** ω_i

$$P(\omega_i|X,y,\dots) \propto \pi(\beta_i|\omega_i,\Omega_i,\gamma_i,\tau)\pi(\omega_i) \propto \frac{1}{\sqrt{\omega_i}}\left\{-\frac{\tau\beta^T D^{-1}\beta}{2}\right\} \times (\omega_i)^{\frac{1}{2}-1}\{-\omega_i\},$$
$$\propto (\omega_i)^{-1}\left\{-\frac{1}{2}\left[\frac{\tau\beta^T\beta}{\Omega_i}\omega_i^{-1} + 2\omega_i\right]\right\}, \qquad \textbf{(21)}$$

herefore, the posterior conditional distribution of ω_i is given by generalized inverse-gaussian distribution

$$\omega_i|X,y,\dots \sim GIG\left(\frac{\tau\beta^T\beta}{\Omega_i},2,0\right). \qquad \textbf{(22)}$$

- **For Ω_i**

$$P(\Omega_i|X,y,\dots) \propto \pi(\beta_i|\omega_i,\Omega_i,\gamma_i,\tau)\pi(\Omega_i) \propto \frac{1}{\sqrt{\Omega_i}}\left\{-\frac{\tau\beta^T D^{-1}\beta}{2}\right\} \times (\Omega_i)^{-\frac{1}{2}-1}\left\{-\frac{1}{\Omega_i}\right\},$$
$$\propto (\Omega_i)^{-1-1}\left\{-\left[\frac{\tau\beta^T Z^{-1}\beta}{2}+1\right](\Omega_i)^{-1}\right\}, \qquad \textbf{(23)}$$

Thus, we have the inverse-gamma distribution

$$\Omega_i|X,y,\dots \sim IG\left(1,\frac{\tau\beta^T D^{-1}\beta}{2}+1\right). \qquad \textbf{(24)}$$

- **For γ_i**

$$P(\gamma_i|X,y,\dots) \propto \pi(\beta_i|\omega_i,\Omega_i,\gamma_i,\tau)\pi(\gamma_i) \propto \frac{1}{\sqrt{\gamma_i}}\left\{-\frac{\tau\beta^T D^{-1}\beta}{2}\right\} \times \{-\lambda\,\gamma_i\},$$
$$\propto (\gamma_i)^{-\frac{1}{2}-1}\left\{-\frac{1}{2}\left[\frac{\tau\beta^T\beta}{\omega_i\Omega_i}(\gamma_i)^{-1} + 2\lambda\,\gamma_i\right]\right\}, \qquad \textbf{(25)}$$

And hence, we have the generalized inverse-gaussian distribution

$$\gamma_i|X,y,\dots \sim GIG\left(\frac{\tau\beta^T\beta}{\omega_i\Omega_i},2\lambda,\frac{1}{2}\right). \qquad \textbf{(26)}$$

- **For s_i**

$$P(s|X,y,\dots) \propto P(y|\beta,\dots)\pi(s|\tau) \propto \frac{1}{\sqrt{s}}\left\{-\frac{(y-X\beta-\eta_1 s)^T S^{-1}(y-X\beta-\eta_1 s)}{2\tau^{-1}\eta_2^2}\right\} \times \{-\tau s\}$$
$$\propto s^{\frac{1}{2}-1}\left\{-\frac{1}{2}\left[\tau\eta_2^{-2}(y-X\beta)^T S^{-1}(y-X\beta)+\tau\left(\frac{\eta_1^2}{\eta_2^2}+2\right)s\right]\right\} \quad \textbf{(27)}$$

again, we get the generalized gaussian distribution

$$\gamma_i|X,y,\dots \sim GIG\left(\frac{\tau(y_i-x_i^T\beta)^2}{\eta_2^2},\tau\left(\frac{\eta_1^2}{\eta_2^2}+2\right),\frac{1}{2}\right). \quad \textbf{(28)}$$

- **For τ**

$$P(\tau|X,y,\dots) \propto P(y|\beta,\dots)\pi(\beta_i|\omega_i,\Omega_i,\gamma_i,\tau)\pi(\tau)$$
$$\propto \left(\tau^{n/2}\left\{-\frac{(y-X\beta-\eta_1 s)^T S^{-1}(y-X\beta-\eta_1 s)}{2\tau^{-1}\eta_2^2}\right\}\right)$$
$$\times\left(\tau^{p/2}\left\{-\frac{\beta^T D^{-1}\beta}{2\tau^{-1}}\right\}\right)(\tau^n(-\tau s))\tau^{c_0-1}\{-d_0\tau\} \quad \textbf{(29)}$$
$$\propto \tau^{c_0+\left(\frac{3n+p}{2}\right)-1}\left\{-\tau\left[\sum_{i=1}^{n}\frac{(y_i-x_i^T\beta-\eta_1 s_i)^2}{2\eta_2^2 s_i}+\sum_{i=1}^{p}\frac{\beta_i^2}{2\omega_k\Omega_k\gamma_k}+d_0\right]\right\}$$

Therefore, the conditional distribution of τ is given by the gamma distribution

$$\gamma_i|X,y,\dots \sim G\left(c_0+\frac{3n+p}{2},\sum_{i=1}^{n}\frac{(y_i-x_i^T\beta-\xi_1 v_i)^2}{2\xi_2^2 v_i}+\frac{\beta^T\beta}{2\prod_{i=k}^{N} z_k}+d_0\right). \quad \textbf{(30)}$$

4. Simulation Study

In this section, we will show and demonstrate how our model, referred to as Half-Cauchy-Plus, difference in terms of prediction with respect to other models such as the horseshoe prior [13], the beta prime prior [10], the regular Bayesian quantile regression (Bqr) [12], Bayesian quantile regression with lasso penalty (qr.lasso) [11,15], and Bayesian quantile regression with the elastic net penalty (qr.enet) [14]. We will evaluate our methods with the mean squared error (MSE), the false positive rate (FPR) and the false negative rate (FNR).

Example 1 (almost all variables are active)

In this example, we consider the model where only some variables are not active by setting $\beta = (4,0,6,2,3,5,9,1,9,0,1,2)$, $w = 0.5$ and $\sigma^2 = 1$. The will simulate the covariates independently from $N(0,\Sigma)$ where Σ is defined such that it is elements are given by $0.5^{|i-j|}$ with i and j being then $(i,j)^{th}$ elements using 100 simulations. It is clear that the results show that the proposed works better than other distributions in Table 1.

Table 1. Results for Example 1.

	MSE (sd)	FPR (sd)	FNR (sd)
Half-Cauchy+	**1.0359 (0.4842)**	0.0000 (0.0000)	0.0400 (0.0749)
Horseshoe	1.0787 (0.9972)	0.0000 (0.0000)	0.0500 (0.1643)
Bqr	1.0561 (0.5818)	0.0000 (0.0000)	0.1600 (0.2351)
qr.lasso	1.1197 (0.7700)	0.0000 (0.0000)	0.1300 (0.5812)
qr.enet	1.0973 (0.5571)	0.0000 (0.0000)	0.2200 (0.5123)

Example 2 (Most variables are not active)

In the second example, we will study the opposite case where only few variables are active by setting $\beta = (7, 0, 0, 0, 0, 0, 1, 0)$, $w = 0.5$ and $\sigma^2 = 1$. Similarly, our proposed method gives better results and prediction accuracy than other distributions in Table 2 and Figures 1 and 2.

Table 2. Results for Example 2.

	MSE (sd)	FPR (sd)	FNR (sd)
Half-Cauchy+	0.2996 (0.1989)	0.3000 (0.6749)	0.0000 (0.0000)
Horseshoe	0.4392 (0.3245)	0.4000 (0.6992)	0.0000 (0.0000)
Bqr	1.3473 (0.4414)	1.5000 (1.4337)	0.0000 (0.0000)
qr.lasso	1.0756 (0.3434)	4.3000 (1.4944)	0.0000 (0.0000)
qr.enet	1.1541 (0.3593)	4.1000 (1.2867)	0.0000 (0.0000)

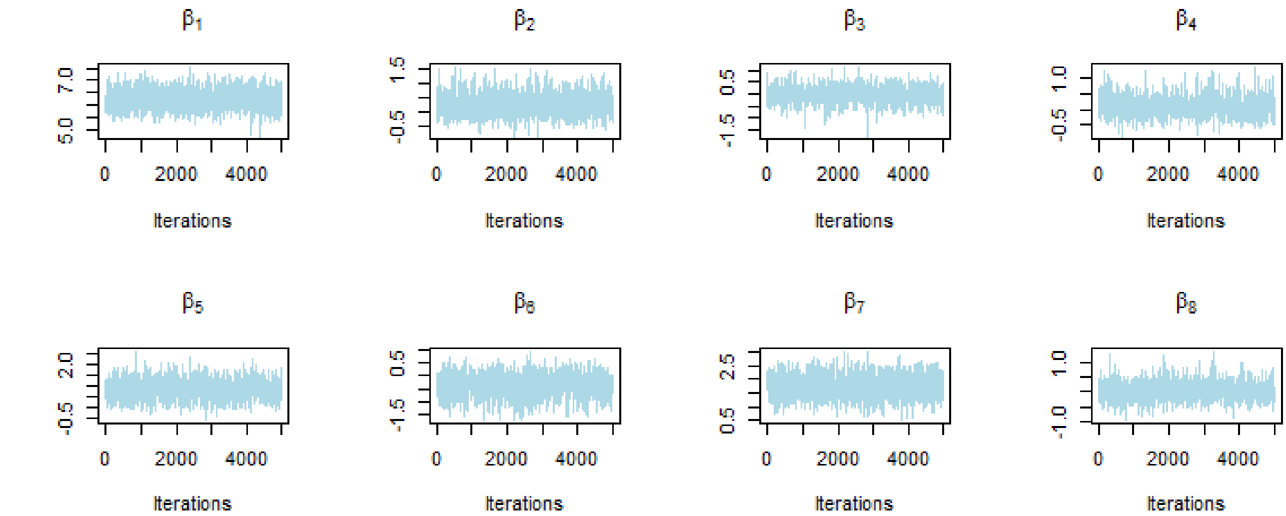

Figure 1. Trace plots of the data covariates.

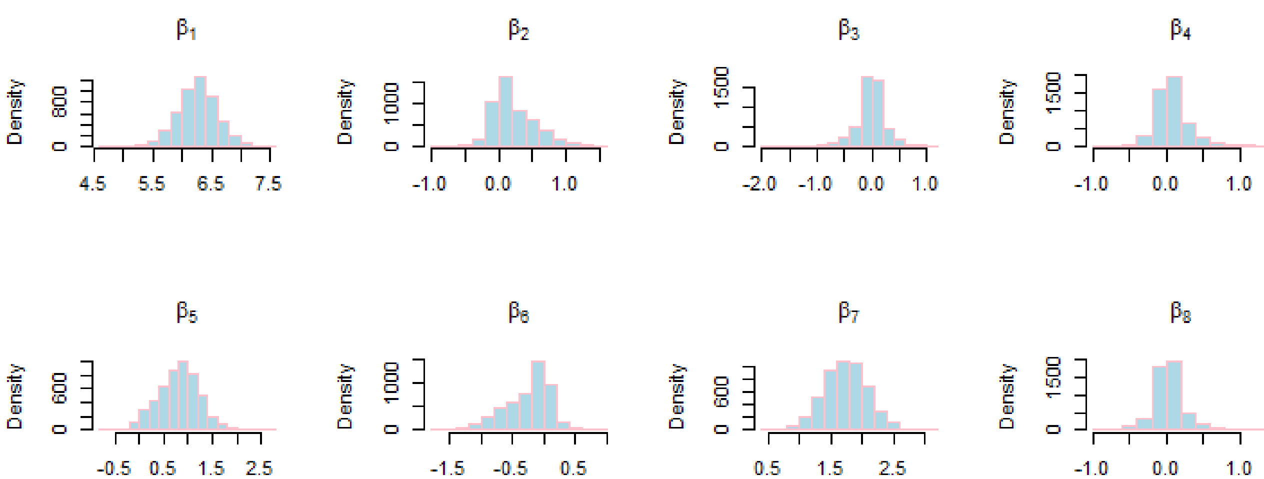

Figure 2. Histograms of the data covariates.

5. Concluding Remarks

In this paper, we have presented a modification of the Bayesian quantile regression using a half-Cauchy prior with an exponential distribution for the scale parameter. To study the properties of our proposition, we have compared our modification the regular Bayesian quantile regression [1,2], Bayesian quantile regression with lasso penalty [3], and Bayesian quantile regression with the elastic net penalty [4]. Through these simulations, we have showed how this models preforms comparatively better than other distributions.

References

[1] Alhamzawi, A. and G. S. Mohammad (2022). Statistical inference with normal-compound gamma priors in regression models. arXiv preprint arXiv:2212.13144 .

[2] Alhamzawi, R. and H. T. M. Ali (2020). Brq: An r package for Bayesian quantile regression. METRON 78 (3), 313–328.

[3] Alhamzawi, R. and H. Mallick (2020). Bayesian reciprocal lasso quantile regression. Communications in Statistics-Simulation and Computation, 1–16.

[4] Alhamzawi, R., K. Yu, V. Vinciotti, and A. Tucker (2011). Prior elicitation for mixed quantile regression with an allometric model. Environmetrics 22 (7), 911–920.

[5] Armagan, A., M. Clyde, and D. Dunson (2011). Generalized beta mixtures of gaussians. Advances in neural information processing systems 24.

[6] Bai, R. and M. Ghosh (2018). On the beta prime prior for scale parameters in high-dimensional bayesian regression models. arXiv preprint arXiv:1807.06539 .

[7] Carvalho, C. M., N. G. Polson, and J. G. Scott (2010). The horseshoe estimator for sparse signals. Biometrika 97 (2), 465–480.

[8] Koenker, R. and G. Bassett (1978). Regression quantiles. Econometrica: Journal of the Econometric Society, 33–50.

[9] Koenker, R. W. and V. d'Orey (1987). Algorithm as 229: Computing regression quantiles. Journal of the Royal Statistical Society. Series C (Applied Statistics) 36 (3), 383–393.

[10] Kozumi, H. and G. Kobayashi (2011). Gibbs sampling methods for Bayesian quantile regression. Journal of statistical computation and simulation 81 (11), 1565–1578.

[11] Li, Q., R. Xi, N. Lin, et al. (2010). Bayesian regularized quantile regression. Bayesian Analysis 5 (3), 533–556.

[12] Mallick, H. and N. Yi (2014). A new Bayesian lasso. Statistics and its interface 7 (4), 571.

[13] P´erez, M.-E., L. R. Pericchi, and I. C. Ram´ırez (2017). The scaled beta2 distribution as a robust prior for scales. Bayesian Analysis 12 (3), 615–637.

[14] Stamey, T. A., J. N. Kabalin, J. E. McNeal, I. M. Johnstone, F. Freiha, E. A. Redwine, and N. Yang (1989). Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. ii. radical prostatectomy treated patients. The Journal of urology 141 (5), 1076–1083.

[15] Yu, K. and R. A. Moyeed (2001). Bayesian quantile regression. Statistics & Probability Letters 54 (4), 437–447.