第39卷第5期 新疆大学学报(自然科学版)(中英文) Vol.39, No.5
2022年9月 Journal of Xinjiang University(Natural Science Edition in Chinese and English) Sept., 2022

A New Quantile Regression Model Based on Uncertainty Theory*

LI Tianyong, HU Xijian†

(*School of Mathematics and System Sciences, Xinjiang University, Urumqi Xinjiang 830017, China*)

Abstract: In this paper, a new uncertain quantile regression model based on uncertainty theory is proposed. Under the framework of uncertainty theory, the definitions of uncertain quantile and uncertain loss function are given and the related theorems are proved. Then, parameter estimates, forecast values and confidence intervals are obtained for the uncertain quantile regression model by using the optimization method of mathematical programming. Finally, two numerical simulations are shown to illustrate the effectiveness and robustness of the model.

Key words: uncertainty theory; quantile regression; parameter estimate; optimization

DOI: 10.13568/j.cnki.651094.651316.2022.03.16.0001

CLC number: O21 **Document Code**: A **Article ID**: 2096-7675(2022)05-0530-012

引文格式: 李天用, 胡锡健. 一类基于不确定理论的分位数回归模型[J]. 新疆大学学报(自然科学版)(中英文), 2022, 39(5): 530-541.

英文引文格式: LI Tianyong, HU Xijian. A new quantile regression model based on uncertainty theory[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English), 2022, 39(5): 530-541.

一类基于不确定理论的分位数回归模型

李天用, 胡锡健

(新疆大学 数学与系统科学学院, 新疆 乌鲁木齐 830017)

摘 要: 基于不确定理论, 提出了一类新的不确定分位数回归模型. 在不确定理论的框架下, 给出了分位点、损失函数等相关概念的定义以及相关定理的证明, 并利用优化方法, 给出了不确定分位数回归模型的参数估计、预测值及预测区间. 最后, 通过数值模拟及案例分析验证了不确定分位数回归模型的鲁棒性和有效性.

关键词: 不确定理论; 分位数回归; 参数估计; 最优化

0 Introduction

Quantile regression is a method to obtain parameter estimates by minimizing the sum of the absolute values of the weighted residuals. Since 1978, when Koenker and Bassett[1] first introduced the concept of quantile regression, it has attracted the attention of many scholars and has been widely used in various fields, such as data processing, systems engineering and probability statistics. Eide and Showalter[2] used quantile regression to study the effect of education on intergenerational income, this method is less restrictive on the random disturbance term in the model and reduces the cost of the experiment. Beirlant and Goegebeur[3] modeled in data processing by Pareto index, this model can estimate extreme quantile problems and effectively analysis the tail weights on the conditional distribution of response variables. Somers and Whittaker[4] applied the quantile regression into retail credit risk assessment practices, solutions at different quartiles can reasonably explain the different distributions in the financial services industry. Nikolaou[5] analyzed the effect of different magnitude shocks on the real exchange rate based on semi-parametric and non-parametric quantile regression models, which clearly reflect the pattern of exchange rate changes at extreme quantile. Yu and Moyeed[6] introduced the idea of Bayesian quantile regression and proved the robustness of quantile regression to heteroskedasticity. Tang and Kong[7] applied the idea of quantile regression to

---

* **Received Date:** 2022-03-16

Foundation Item: The research is supported by National Natural Science Foundation of People's Republic of China (11961065).

Biography: LI Tianyong(1996-), female, Phd candidate, research fields: uncertainty theory.

† **Corresponding author:** HU Xijian(1964-), male, doctor, professer, E-mail: xijianhu@xju.edu.cn.

the linear semi-parametric model of functional data. Sun[8] further introduced the idea of quantile regression into functional data and established a functional data quantile regression model. He, Yan and Xu[9] combined the support vector integral quantile regression method with fuzzy information granulation to construct a quantile regression model based on fuzzy information granulation and support vector.

To the best of our knowledge, the quantile regression has been proposed usually in the framework of probability theory. Because the traditional quantile regression methods all assume that we have a large number of precise observations, and some observations are imprecise within a certain range, e.g. height around 1.8 m, distance from the school building about 500 m, weight of the dog between 10 and 20 kg, etc. So, the results of the regression analysis method under probability theory will have a large bias, we need to use experts' estimated imprecise data. Based on the above research background and problems, Liu[10–11] proposed uncertainty theory to fill these gaps and widely applied into various fields such as uncertain programming[12], uncertain calculus[13], uncertain statistics[14] and uncertain differential[15]. At the same time, Liu and his students put forward the principle of uncertain least squares method, the parameters were also estimated using experimental data and the sum of squares of distances from uncertain distributions[14]. They also introduced point estimates of the unknown parameters in the uncertain multiple regression model[16]. Shortly after, Lio and Liu[17] made interval predictions of the response uncertain variables to determine the range of values of the estimates. Ye and Liu[18] derived uncertain least squares estimates of the unknown parameters, in which the disturbance terms are assumed to have normal uncertainty distributions. Liu and Yang[19] applied the least absolute principle into the estimates of the unknown parameters in an uncertain multiple regression model. To further improve the uncertainty theory, this paper develops an uncertain quantile regression (UQR) model based on previous studies combining quantile regression model and uncertainty theory. From the data perspective, using uncertain data can make the estimation results more accurate. From the model perspective, UQR can not only measure the effect of regression variables at the center of the distribution, but also fit the curves corresponding to different quantiles, which can analyze the characteristics of the data more comprehensively.

The rest of the paper is organized as follows. Section 1 reviews the basics of uncertainty theory. Section 2 defines the concepts of uncertain quantile and uncertain loss function. In Section 3, the UQR model is introduced. The uncertain parameter estimates, residual distributions and confidence intervals of the response variables are also obtained. In Section 4, numerical simulations are performed to compare the mean square errors of the original data and the uncertain data, and the robustness of the UQR model is demonstrated. The conclusions are given in Section 5.

1 Preliminaries

In this section, we introduce some fundamental concepts and theorems based on uncertainty theory, including uncertain measure, uncertain variable, uncertainty distribution, uncertain expected value and uncertain variance.

Definition 1[10] Let $\mathcal{L}$ be a σ-algebra on a nonempty set Γ. A set function $\mathcal{M} : \mathcal{L} \to [0,1]$ is called an uncertain measure, if it satisfies the following axioms:

Axiom 1 (Normality Axiom) $\mathcal{M}\{\Gamma\} = 1$ for the universal set Γ.

Axiom 2 (Duality Axiom) $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ for any event Λ.

Axiom 3 (Subadditivity Axiom) For every countable sequence of events Λ_1, Λ_2, $\cdots$, we have

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty}\Lambda_i\right\} \le \sum_{i=1}^{\infty}\mathcal{M}\{\Lambda_i\}.$$

Axiom 4 (Product Axiom) Let $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ be uncertainty spaces for $k = 1,2,\cdots$. Then the product uncertain measure $\mathcal{M}$ is an uncertain measure satisfying

$$\mathcal{M}\Big\{\prod_{k=1}^{\infty}\Lambda_k\Big\} = \bigwedge_{k=1}^{\infty}\mathcal{M}_k\{\Lambda_k\},$$

where Λ_k are arbitrarily chosen events from $\mathcal{L}_k$ for $k = 1,2,\cdots$, respectively.

Definition 2[10] An uncertain variable ξ is a measurable function from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set of real numbers, i.e., for any Borel set B of real numbers, the set

$$\{\xi \in B\} = \{\gamma \in \Gamma \mid \xi(\gamma) \in B\}$$

is an event.

Definition 3[11] The uncertain variables $\xi_1, \xi_2, \cdots, \xi_m$ are said to be independent if

$$\mathcal{M}\Big\{\bigcap_{i=1}^{m}\{\xi_i \in B_i\}\Big\} = \bigwedge_{i=1}^{m} \mathcal{M}\{\xi_i \in B_i\},$$

for any Borel sets $B_1, B_2, \cdots, B_m$ of real number.

Theorem 1[14] Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If the function $f(x_1, x_2, \cdots, x_n)$ is strictly increasing with respect to $x_1, x_2, \cdots, x_m$ and strictly decreasing with $x_{m+1}, x_{m+2}, \cdots, x_n$, then the uncertain variable

$$\xi = f(\xi_1, \cdots, \xi_m, \xi_{m+1}, \cdots, \xi_n)$$

is an uncertain variable with inverse uncertainty distribution

$$\Psi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha), \cdots, \Phi_n^{-1}(1-\alpha)).$$

Definition 4[10] Let ξ be an uncertain variable. Then the expected value of ξ is defined as

$$E[\xi] = \int_0^{+\infty} \mathcal{M}\{\xi \ge x\}\,\mathrm{d}x - \int_{-\infty}^{0} \mathcal{M}\{\xi \le x\}\,\mathrm{d}x,$$

provided that at least one of the two integrals is finite.

For an uncertain variable ξ with regular uncertainty distribution $\Phi(x)$, its expected value can be obtained by

$$E[\xi] = \int_0^1 \Phi^{-1}(\alpha)\,\mathrm{d}\alpha.$$

Theorem 2[14] Let ξ and η be independent uncertain variables with finite expected values. Then for any real number a and b, we have

$$E[a\xi + b\eta] = aE[\xi] + bE[\eta].$$

Definition 5[10] Let ξ be an uncertain variable with finite expected value e. Then the variance of ξ is defined by

$$V[\xi] = E[(\xi - e)^2].$$

For an uncertain variable ξ with regular uncertainty distribution $\Phi(x)$ and finite expected value e, its variance can be obtained by

$$V[\xi] = \int_0^1 \left(\Phi^{-1}(\alpha) - e\right)^2 \mathrm{d}\alpha.$$

2 Uncertain Quantile

The core of the quantile regression model is the quantile in probability theory. In order to build the UQR model, we define the concepts of uncertain quantile and uncertain loss function as follows.

Definition 6 Let ξ be an uncertain variable with regular uncertainty distribution $\Phi(x)$. Then for any real number x and $0 \le \tau \le 1$, the τ uncertain quantile of ξ is defined as

$$\Phi^{-1}(\tau) = \inf\{x : \Phi(x) \ge \tau\}.$$

Definition 7 For a real number $\tau \in (0,1)$, we define an uncertain loss function $\rho_\tau(y)$ as follows,

$$\rho_\tau(y) = (\tau - I_{\{y<0\}})y = \begin{cases} \tau y, & \text{if } y \ge 0 \\ (\tau - 1)y, & \text{otherwise} \end{cases} \tag{1}$$

where the indicator function $I_{\{A\}}$ equals 1 for all elements of A and 0 otherwise.

According to the calculation of expectation in uncertainty theory, the expectation of uncertainty loss function in Definition 7 is proved as follows.

Theorem 3 Let ξ be an uncertain variable with an uncertainty distribution $\Phi(x)$. If the uncertain loss function is $\rho_\tau(y)$, then for any real number $\hat{x}$ and $\tau \in (0,1)$, the expected value of the uncertain loss function is defined by

$$E\rho_\tau(\xi-\hat{x}) = (\tau-1)\int_0^{\Phi(\hat{x})}\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha + \tau\int_{\Phi(\hat{x})}^1\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha \tag{2}$$

Proof According to formula (1) in Definition 7,

$$E[\rho_\tau(y)] = E[\rho_\tau(\xi-\hat{x})] \tag{3}$$

Denote the uncertain measure as $\mathcal{M}$ which reflects the personal confidence degree of an uncertain event that may happen, an uncertain variable as ξ and its uncertainty distribution as $\Phi(x)$. Then it follows from the definition of expected value[10] and subadditivity of uncertain measure[10] that

$$\begin{aligned}
E[\rho_\tau(\xi-\hat{x})] &= \int_0^{+\infty}\mathcal{M}\left\{\rho_\tau(\xi-\hat{x})\geq x\right\}\mathrm{d}x - \int_{-\infty}^0\mathcal{M}\left\{\rho_\tau(\xi-\hat{x})\leq x\right\}\mathrm{d}x\\
&= \int_0^{\infty}\mathcal{M}\left\{\left[\tau(\xi-\hat{x})\geq x\right]\cup\left[(\tau-1)(\xi-\hat{x})\leq x\right]\right\}\mathrm{d}x\\
&\leq \int_0^{\infty}\mathcal{M}\left\{\tau(\xi-\hat{x})\geq x\right\}+\mathcal{M}\left\{(\tau-1)(\xi-\hat{x})\leq x\right\}\mathrm{d}x\\
&= \int_0^{\infty}\left(1-\Phi\left(\frac{x}{\tau}+\hat{x}\right)+\Phi\left(\frac{x}{\tau-1}+\hat{x}\right)\right)\mathrm{d}x\\
&= \tau\left[\int_{\hat{x}}^{+\infty}(x-\hat{x})\,\mathrm{d}\Phi(x)\right]+(\tau-1)\left[\int_{-\infty}^{\hat{x}}(x-\hat{x})\,\mathrm{d}\Phi(x)\right]
\end{aligned} \tag{4}$$

Note that ξ is uncertain variable and

$$\xi-\hat{x} \tag{5}$$

is increasing with respect to ξ. The inverse distribution of the uncertain variable ξ is $\Phi^{-1}(\alpha)$. According to the operation law of the uncertain variable[10], and the inverse uncertainty distribution of uncertain variable (5) is

$$\Phi^{-1}(\alpha)-\hat{x},$$

therefore,

$$(\tau-1)\int_0^{\Phi(\hat{x})}\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha + \tau\int_{\Phi(\hat{x})}^1\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha.$$

Since Equation (4) can be showed as

$$E\rho_\tau(\xi-\hat{x}) = (\tau-1)\int_0^{\Phi(\hat{x})}\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha + \tau\int_{\Phi(\hat{x})}^1\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha,$$

for any $\tau\in(0,1)$.

According to uncertainty theory, the uncertain quantile can be defined as Theorem 4.

Theorem 4 Let ξ be an uncertain variable with uncertainty distribution $\Phi(x)$. Then for any real number $\hat{x}$, we have

$$\Phi(\hat{x}) = \tau.$$

Proof According to the uncertain loss function $\rho_\tau(y)$, $\tau\in(0,1)$, Formula (2) and Theorem 3, we find the minimum value of the objective function:

$$E\rho_\tau(\xi-\hat{x}) = (\tau-1)\int_0^{\Phi(\hat{x})}\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha + \tau\int_{\Phi(\hat{x})}^1\left(\Phi^{-1}(\alpha)-\hat{x}\right)\mathrm{d}\alpha.$$

Taking the derivative of $\hat{x}$ in the Equation (2), we get

$$\begin{aligned}
0 &= (\tau-1)\int_0^{\Phi(\hat{x})}(-1)\,\mathrm{d}\alpha+\tau\int_{\Phi(\hat{x})}^1(-1)\,\mathrm{d}\alpha\\
&= (\tau-1)(-\alpha)\,|_0^{\Phi(\hat{x})}+\tau(-\alpha)\,|_{\Phi(\hat{x})}^1\\
&= -\Phi(\hat{x})(\tau-1)+(\Phi(\hat{x})-1)\tau\\
&= \Phi(\hat{x})-\tau,
\end{aligned}$$

which can be calculated as

$$\Phi(\hat{x}) = \tau.$$

Since $\Phi(x)$ is a regular uncertainty distribution, it has a unique solution, $\Phi(\hat{x}) = \tau$. Therefore, the minimum estimated expectation of the real number $\hat{x}$ is the uncertain quantile.

3 UQR Model

In this section, UQR model is presented, parameters are estimated for different quantiles, and the residual distributions, forecast values, confidence intervals for the corresponding quantiles are given.

3.1 Statistical Inference of UQR

The uncertain least absolute deviation estimator was derived by Liu and Yang[19]. According to the definition of uncertain quantile, the uncertain quantile is 0.5 in Reference [19], while other quartiles are not explained. However, Liu and Yang only analyzed the change pattern of the data in the center, other data are not analyzed. To make up for this deficiency, we define a unary uncertain quantile regression model as follows.

Definition 8 Given imprecisely observed data y_i, x_i $(i = 1,2,\cdots,n)$ characterized as independent uncertain variables with regular uncertainty distributions Φ_i, Ψ_i $(i = 1,2,\cdots,n)$. These data satisfy the unary linear regression model, i.e.,

$$y_i = \beta_0 + x_i\beta_1 + \varepsilon_i, \quad i = 1,2,\cdots,n.$$

When τ is uncertain quantile, the unary linear regression model is expressed as

$$y_i = \beta_{0\tau} + x_i\beta_{1\tau} + \varepsilon_{i\tau}, \quad i = 1,2,\cdots,n, \tau \in (0,1) \tag{6}$$

where $\beta_{0\tau}$, $\beta_{1\tau}$ are unknown parameters, $\varepsilon_{i\tau}$ are disturbance terms.

Definition 9 If the data satisfy the linear regression model (6) in which τ is uncertain quantile, then uncertain quantile regression estimators $\beta^*_{0\tau}$, $\beta^*_{1\tau}$ of $\beta_{0\tau}$, $\beta_{1\tau}$ respectively are the optimal solutions of the following minimization problem:

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \sum_{i=1}^{n} E\rho_\tau(|\, y_i - \beta_{0\tau} - x_i\beta_{1\tau}\,|) \tag{7}$$

further,

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \Big(\sum_{y_i < x_i\beta_{1\tau}+\beta_{0\tau}} E((1-\tau)\,|\, y_i - \beta_{0\tau} - x_i\beta_{1\tau}\,|) + \sum_{y_i \ge \beta_{0\tau}+x_i\beta_{1\tau}} E(\tau\,|\, y_i - \beta_{0\tau} - x_i\beta_{1\tau}\,|) \Big),$$

where $i = 1,2,\cdots,n$, $\tau \in (0,1)$.

To analyze the parameter estimation problem, the parameter estimation of the UQR model can be turned to a simple optimization solution problem.

Theorem 5 If the data satisfy the linear regression model (6) in which τ is uncertain quantile, then uncertain quantile regression estimators $\beta^*_{0\tau}$, $\beta^*_{1\tau}$ of $\beta_{0\tau}$, $\beta_{1\tau}$ respectively are the optimal solutions of the following minimization problem:

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \sum_{i=1}^{n} E\rho_\tau(|\, y_i - \beta_{0\tau} - x_i\beta_{1\tau}\,|),$$

further,

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \Big(\sum_{y_i < x_i\beta_{1\tau}+\beta_{0\tau}} E((1-\tau)\,|\, y_i - \beta_{0\tau} - x_i\beta_{1\tau}\,|) + \sum_{y_i \ge \beta_{0\tau}+x_i\beta_{1\tau}} E(\tau\,|\, y_i - \beta_{0\tau} - x_i\beta_{1\tau}\,|) \Big),$$

which can be calculated as

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \Big(\sum_{y_i < \beta_{0\tau}+x_i\beta_{1\tau}} (1-\tau) \int_0^{\upsilon_i} |\, \Phi_i^{-1}(\alpha) - \beta_{0\tau} - \Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}\,|\,\mathrm{d}\alpha + \sum_{y_i \ge \beta_{0\tau}+x_i\beta_{1\tau}} \tau \int_{\upsilon_i}^{1} |\, \Phi_i^{-1}(\alpha) - \beta_{0\tau} - \Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}\,|\,\mathrm{d}\alpha \Big),$$

where

$$\Psi_i^{-1}(\alpha,\beta_{1\tau}) = \begin{cases} \Psi_i^{-1}(1-\alpha), & \beta_{1\tau} \ge 0, \\ \Psi_i^{-1}(\alpha), & \beta_{1\tau} < 0, \end{cases}$$

and

$$\upsilon_i = \begin{cases} \frac{\beta_{0\tau}-(b_{1i}-\beta_{1\tau}b_{2i})}{a_{1i}-\beta_{1\tau}a_{2i}-b_{1i}+\beta_{1\tau}b_{1i}}, & \beta_{1\tau} \geq 0, \\ \frac{a_{1i}-\beta_{1\tau}a_{2i}-\beta_{0\tau}}{a_{1i}-\beta_{1\tau}a_{2i}+b_{1i}-\beta_{1\tau}b_{2i}}, & \beta_{1\tau} < 0, \end{cases}$$

for $i = 1,2,\cdots,n$, $\tau \in (0,1)$.

Proof According to Definition 9, then UQR estimators $\beta^*_{0\tau}$, $\beta^*_{1\tau}$ of $\beta_{0\tau}$, $\beta_{1\tau}$ respectively are the optimal solutions of the following minimization problem:

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \Big(\sum_{y_i<x_i\beta_{1\tau}+\beta_{0\tau}} E((1-\tau)\,|\,y_i-\beta_{0\tau}-x_i\beta_{1\tau}\,|) + \sum_{y_i\geq\beta_{0\tau}+x_i\beta_{1\tau}} E(\tau\,|\,y_i-\beta_{0\tau}-x_i\beta_{1\tau}\,|) \Big) \tag{8}$$

Noting imprecise observations y_i, x_i $(i = 1,2,\cdots,n)$ have uncertainty distributions Φ_i, Ψ_i $(i = 1,2,\cdots,n)$, so the inverse uncertainty distributions of y_i, x_i $(i = 1,2,\cdots,n)$ are $\Phi_i^{-1}(\alpha)$, $\Psi_i^{-1}(\alpha)$ $(i = 1,2,\cdots,n)$. The uncertain variables

$$y_i-\beta_{0\tau}-x_i\beta_{1\tau} \tag{9}$$

increase with the increase of y_i, increase with the increase of x_i when $\beta_{1\tau} < 0$; and increase with the decrease of x_i when $\beta_{1\tau} \geq 0$, $i = 1,2,\cdots,n$, $\tau \in (0,1)$. According to the operation law of uncertain variables[10], the inverse uncertainty distributions of uncertain variables (9) are

$$\Upsilon_i^{-1}(\alpha) = \Phi_i^{-1}(\alpha) - \beta_{0\tau} - \Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau},$$

where

$$\Psi_i^{-1}(\alpha,\beta_{1\tau}) = \begin{cases} \Psi_i^{-1}(1-\alpha), & \beta_{1\tau} \geq 0, \\ \Psi_i^{-1}(\alpha), & \beta_{1\tau} < 0, \end{cases}$$

for $i = 1,2,\cdots,n$, $\tau \in (0,1)$. Therefore, we take $\Phi_i^{-1}(\alpha)-\beta_{0\tau}-\Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau} = 0$, for $i = 1,2,\cdots,n$, $\tau \in (0,1)$. Suppose $\Phi_i(\alpha)$, $\Psi_i(\alpha)$ are linear uncertainty distributions, $\Phi_i(\alpha)$ have inverse uncertainty distributions $\Phi_i^{-1}(\alpha) = (1-\alpha)a_{1i} + \alpha b_{1i}$, $\Psi_i(\alpha)$ have inverse uncertainty distributions $\Psi_i^{-1}(\alpha) = (1-\alpha)a_{2i} + \alpha b_{2i}$, when $\beta_{1\tau} < 0$,

$$\Psi_i^{-1}(\alpha,\beta_{1\tau}) = \Psi^{-1}(\alpha) = (1-\alpha)a_{2i} + \alpha b_{2i},$$

and

$$\Phi_i^{-1}(\alpha) - \Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau} = \beta_{0\tau},$$

then

$$(1-\alpha)a_{1i} + \alpha b_{1i} - [(1-\alpha)a_{2i} + \alpha b_{2i}]\beta_{1\tau} = \beta_{0\tau},$$

so that

$$\alpha = \frac{a_{1i}-\beta_{1\tau}a_{2i}-\beta_{0\tau}}{a_{1i}-\beta_{1\tau}a_{2i}+b_{1i}-\beta_{1\tau}b_{2i}},$$

for $i = 1,2,\cdots,n$, $\tau \in (0,1)$. When $\beta_{1\tau} \geq 0$, $\Psi_i^{-1}(1-\alpha,\beta_{1\tau}) = \Psi_i^{-1}(1-\alpha) = \alpha a_{2i} + (1-\alpha)b_{2i}$,

$$\Phi_i^{-1}(\alpha) - \Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau} = \beta_{0\tau},$$

then

$$\alpha a_{1i} + (1-\alpha)b_{1i} - [\alpha a_{2i} + (1-\alpha)b_{2i}]\beta_{1\tau} = \beta_{0\tau},$$

so that

$$\alpha = \frac{\beta_{0\tau}-(b_{1i}-\beta_{1\tau}b_{2i})}{a_{1i}-\beta_{1\tau}a_{2i}-b_{1i}+\beta_{1\tau}b_{1i}},$$

for $i = 1,2,\cdots,n$, $\tau \in (0,1)$. According to Theorem 3, Formula (8) can be calculated as follows

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)} \Big(\sum_{y_i<\beta_{0\tau}+x_i\beta_{1\tau}} (1-\tau) \int_0^{\upsilon_i} |\,\Phi_i^{-1}(\alpha)-\beta_{0\tau}-\Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}\,|\,\mathrm{d}\alpha + \sum_{y_i\geq\beta_{0\tau}+x_i\beta_{1\tau}} \tau \int_{\upsilon_i}^1 |\,\Phi_i^{-1}(\alpha)-\beta_{0\tau}-\Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}\,|\,\mathrm{d}\alpha \Big),$$

where

$$\Psi_i^{-1}(\alpha,\beta_{1\tau}) = \begin{cases} \Psi_i^{-1}(1-\alpha), & \beta_{1\tau} \geq 0, \\ \Psi_i^{-1}(\alpha), & \beta_{1\tau} < 0, \end{cases}$$

and

$$\upsilon_i=\begin{cases}\frac{\beta_{0\tau}-(b_{1i}-\beta_{1\tau}b_{2i})}{a_{1i}-\beta_{1\tau}a_{2i}-b_{1i}+\beta_{1\tau}b_{1i}}, & \beta_{1\tau}\geq 0,\\ \frac{a_{1i}-\beta_{1\tau}a_{2i}-\beta_{0\tau}}{a_{1i}-\beta_{1\tau}a_{2i}+b_{1i}-\beta_{1\tau}b_{2i}}, & \beta_{1\tau}<0,\end{cases}$$

for $i=1,2,\cdots,n$, $\tau\in(0,1)$.

3.2 Stability Testing

Normally, the establishment of the statistic models should include stability testing. However, the data are imprecise in some cases, and it is difficult to test imprecise data in probility theory. So we give Definition 10 to introduce the specific method of UQR model testing using imprecise data.

Definition 10 If the data satisfy the linear regression model (6) in which τ is uncertain quantile, the residuals of τ uncertain quantile regression are expressed as

$$\hat{\varepsilon}_{i\tau}=y_i-(\beta^*_{0\tau}+x_i\beta^*_{1\tau}),\quad i=1,\cdots,n,\tau\in(0,1).$$

$\beta^*_{0\tau}$, $\beta^*_{1\tau}$ are UQR estimators of $\beta_{0\tau}$, $\beta_{1\tau}$, respectively. Furthermore, with the assumption that $E[\hat{\varepsilon}_{i\tau}]=e_\tau$, $V[\hat{\varepsilon}_{i\tau}]=\sigma^2_\tau$, uncertain mean square errors of $\varepsilon_{i\tau}$ are mse_τ, $(i=1,\cdots,n)$, $\tau\in(0,1)$. We can estimate expected values, variances, mean square errors of τ uncertain quantile regression as $\hat{e}_\tau$, $\hat{\sigma}^2_\tau$, mse_τ, respectively.

$$\hat{e}_\tau=\frac{1}{n}\sum_{i=1}^{n}E[\hat{\varepsilon}_{i\tau}],$$

$$\hat{\sigma}^2_\tau=\frac{1}{n}\sum_{i=1}^{n}E[(\hat{\varepsilon}_{i\tau}-\hat{e}_\tau)^2],$$

$$mse_\tau=\frac{1}{n}\sum_{i=1}^{n}E[(\hat{\varepsilon}_{i\tau})^2],$$

where $i=1,2,\cdots,n$, $\tau\in(0,1)$.

For convenience of calculation, the following formula is given based on uncertainty theory and the concept of Definition 10.

Theorem 6 If the data satisfy the linear regression model (6) in which τ is uncertain quantile, then the expected values $\hat{e}_\tau$, variances $\hat{\sigma}^2_\tau$ and uncertain mean square errors mse_τ of $\varepsilon_{i\tau}$ can be respectively calculated as follows

$$\hat{e}_\tau=\frac{1}{n}\sum_{i=1}^{n}\int_0^1(\Phi_i^{-1}(\alpha)-(\beta^*_{0\tau}+\Psi_i^{-1}(\alpha,\beta^*_{1\tau})\beta^*_{1\tau}))\mathrm{d}\alpha \tag{10}$$

$$\hat{\sigma}^2_\tau=\frac{1}{n}\sum_{i=1}^{n}\int_0^1(\Phi_i^{-1}(\alpha)-(\beta^*_{0\tau}+\Psi_i^{-1}(\alpha,\beta^*_{1\tau})\beta^*_{1\tau}-\hat{e}_\tau))^2\mathrm{d}\alpha \tag{11}$$

$$mse_\tau=\frac{1}{n}\sum_{i=1}^{n}\int_0^1(\Phi_i^{-1}(\alpha)-(\beta^*_{0\tau}+\Psi_i^{-1}(\alpha,\beta^*_{1\tau})\beta^*_{1\tau}))^2\mathrm{d}\alpha \tag{12}$$

where

$$\Psi_i^{-1}(\alpha,\beta_{1\tau})=\begin{cases}\Psi_i^{-1}(1-\alpha), & \beta_{1\tau}\geq 0,\\ \Psi_i^{-1}(\alpha), & \beta_{1\tau}<0,\end{cases}$$

for $i=1,2,\cdots,n$, $\tau\in(0,1)$ and $\beta^*_{0\tau}$, $\beta^*_{1\tau}$ are UQR estimators of $\beta_{0\tau}$, $\beta_{1\tau}$, respectively.

Proof According to Definition 10, the residuals of τ UQR are expressed as

$$\hat{\varepsilon}_{i\tau}=y_i-(\beta^*_{0\tau}+x_i\beta^*_{1\tau}),$$

for $i=1,\cdots,n$, $\tau\in(0,1)$ and the valuation of $\beta_{0\tau}$, $\beta_{1\tau}$ are $\beta^*_{0\tau}$, $\beta^*_{1\tau}$, respectively. According to the operation law of uncertain variables[10], inverse uncertainty distributions $F^{-1}_{i\tau}(x)$ of $\hat{\varepsilon}_{i\tau}$ are

$$F^{-1}_{i\tau}=\Phi_i^{-1}(\alpha)-\beta^*_{0\tau}-\Psi_i^{-1}(\alpha,\beta^*_{1\tau})\beta^*_{1\tau},$$

where $i = 1,\cdots,n$, $\tau \in (0,1)$. Then we have

$$E[\hat{\varepsilon}_{i\tau}] = \int_0^1 F_{i\tau}^{-1}(\alpha)\mathrm{d}\alpha \tag{13}$$

$$E[(\hat{\varepsilon}_{i\tau} - e_\tau)^2] = \int_0^1 (F_{i\tau}^{-1}(\alpha) - e_\tau)^2\mathrm{d}\alpha \tag{14}$$

$$E[\hat{\varepsilon}_{i\tau}^2] = \int_0^1 (\varPhi_i^{-1}(\alpha) - (\beta_{0\tau}^* + \varPsi_i^{-1}(\alpha,\beta_{1\tau}^*)\beta_{1\tau}^*))^2\mathrm{d}\alpha \tag{15}$$

where

$$\varPsi_i^{-1}(\alpha,\beta_{1\tau}^*) = \begin{cases} \varPsi_i^{-1}(1-\alpha), & \beta_{1\tau}^* \geq 0, \\ \varPsi_i^{-1}(\alpha), & \beta_{1\tau}^* < 0, \end{cases}$$

for $i = 1,\cdots,n$, $\tau \in (0,1)$. Theorem 6 can be directly obtained from formulas (13)~(15).

In the following paragraphs, Definition 11 and Definition 12 give the calculations of the forecast values and confidence intervals.

Definition 11 Given a new uncertain variable $\tilde{x}$ with a regular uncertainty distribution $\tilde{\varPsi}$ in τ uncertain quantile regression model (6), the forecast values $\hat{y}_\tau$ are predicted to be

$$\mu_\tau = E(\hat{y}_\tau) = E(\beta_{0\tau}^* + \tilde{x}\beta_{1\tau}^*) + \hat{e}_\tau, \quad \tau \in (0,1).$$

Definition 12 Suppose in τ uncertain quantile regression, $\hat{\varepsilon}_\tau$ are normal uncertain variables, $\mathcal{N}(\hat{e}_\tau, \hat{\sigma}_\tau)$, $\tau \in (0,1)$. Using the linear regression model (6), the uncertainty distributions of $\hat{y}_\tau$ are $\varPhi_\tau$, and the inverse uncertainty distributions $\varPhi_\tau^{-1}(\alpha)$ of $\hat{y}_\tau$ are given as follows

$$\varPhi_\tau^{-1}(\alpha) = \beta_{0\tau}^* + \tilde{\varPsi}^{-1}(\alpha,\beta_{1\tau}^*)\beta_{1\tau}^* + \Theta_\tau^{-1}(\alpha),$$

where

$$\tilde{\varPsi}^{-1}(\alpha,\beta_{1\tau}^*) = \begin{cases} \tilde{\varPsi}^{-1}(1-\alpha), & \beta_{1\tau}^* \geq 0, \\ \tilde{\varPsi}^{-1}(\alpha), & \beta_{1\tau}^* < 0, \end{cases}$$

and

$$\Theta_\tau^{-1}(\alpha) = \hat{e}_\tau + \hat{\sigma}_\tau \frac{\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha},$$

for $\tau \in (0,1)$. $\Theta_\tau^{-1}(\alpha)$ are the inverse uncertainty distributions of the normal uncertain variables $\mathcal{N}(\hat{e}_\tau, \hat{\sigma}_\tau)$, $\tau \in (0,1)$. Therefore, according to the subadditivity of uncertain measures[10], we have

$$\mathcal{M}\{\mu_\tau - \beta_\tau \leq y_\tau \leq \mu_\tau + \beta_\tau\} \geq \varPhi_\tau(\mu_\tau + \beta_\tau) - \varPhi_\tau(\mu_\tau - \beta_\tau),$$

where $\tau \in (0,1)$. Thus, the prediction intervals of $\tilde{x}$ corresponding to y_τ are

$$[\mu_\tau - \beta_\tau, \mu_\tau + \beta_\tau],$$

and β_τ are the minimum values of Formula (16):

$$\varPhi_\tau(\mu_\tau + \beta_\tau) - \varPhi_\tau(\mu_\tau - \beta_\tau) \geq \alpha \tag{16}$$

where $\tau \in (0,1)$.

4 Numerical Simulation and Case Analysis

In this section, two experiments based on Model (6) are performed. Furthermore, we estimate the unknown parameters, expected values, variances, mean square errors, forecast values and confidence intervals for the UQR model. We have a linear uncertain variable $\mathcal{L}(a,b)$ with the distribution

$$\varPhi(x) = \begin{cases} 0, & x \leq a, \\ x - a/b - a, & a < x \leq b, \\ 1, & x > b, \end{cases}$$

where a and b are real numbers, and satisfying a<b.

4.1 Numerical Example

We describe the imprecise data expressed in interval forms as linear uncertain variables (Table 1).

Table 1 Imprecisely observed data

i	y_i	x_i	i	y_i	x_i
1	$\mathcal{L}(83,90)$	$\mathcal{L}(5,6)$	9	$\mathcal{L}(233,280)$	$\mathcal{L}(20,21)$
2	$\mathcal{L}(32,40)$	$\mathcal{L}(0,1)$	10	$\mathcal{L}(198,248)$	$\mathcal{L}(17,18)$
3	$\mathcal{L}(41,60)$	$\mathcal{L}(2,3)$	11	$\mathcal{L}(298,369)$	$\mathcal{L}(27,28)$
4	$\mathcal{L}(98,125)$	$\mathcal{L}(7,8)$	12	$\mathcal{L}(264,330)$	$\mathcal{L}(23,24)$
5	$\mathcal{L}(60,80)$	$\mathcal{L}(3,4)$	13	$\mathcal{L}(253,318)$	$\mathcal{L}(22,23)$
6	$\mathcal{L}(140,157)$	$\mathcal{L}(11,12)$	14	$\mathcal{L}(298,396)$	$\mathcal{L}(28,29)$
7	$\mathcal{L}(182,219)$	$\mathcal{L}(15,16)$	15	$\mathcal{L}(117,145)$	$\mathcal{L}(9,10)$
8	$\mathcal{L}(150,191)$	$\mathcal{L}(12,13)$			

By Theorem 5, if the data satisfy the linear regression model (6), the unknown parameters $\beta_{0\tau}^*$, $\beta_{1\tau}^*$ respectively are the optimal solutions to the following minimization problem:

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)}\Big(\sum_{y_i<\beta_{0\tau}+x_i\beta_{1\tau}}(1-\tau)\int_0^{\upsilon_i}|\varPhi_i^{-1}(\alpha)-\beta_{0\tau}-\varPsi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}|\,\mathrm{d}\alpha+\sum_{y_i\geq\beta_{0\tau}+x_i\beta_{1\tau}}\tau\int_{\upsilon_i}^1|\varPhi_i^{-1}(\alpha)-\beta_{0\tau}-\varPsi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}|\,\mathrm{d}\alpha\Big),$$

where

$$\varPsi_i^{-1}(\alpha,\beta_{1\tau})=\begin{cases}\varPsi_i^{-1}(1-\alpha), & \beta_{1\tau}\geq 0,\\ \varPsi_i^{-1}(\alpha), & \beta_{1\tau}<0,\end{cases}$$

and

$$\upsilon_i=\begin{cases}\frac{\beta_{0\tau}-(b_{1i}-\beta_{1\tau}b_{2i})}{a_{1i}-\beta_{1\tau}a_{2i}-b_{1i}+\beta_{1\tau}b_{1i}}, & \beta_{1\tau}\geq 0,\\ \frac{a_{1i}-\beta_{1\tau}a_{2i}-\beta_{0\tau}}{a_{1i}-\beta_{1\tau}a_{2i}+b_{1i}-\beta_{1\tau}b_{2i}}, & \beta_{1\tau}<0,\end{cases}$$

for $i=1,\cdots,n$, $\tau\in(0,1)$.

Through the formulations above, we can obtain estimates of the UQR in Table 2 when τ is equal to 0.1, 0.25, 0.5, 0.75 and 0.9, respectively.

Table 2 Estimated coefficients of each quantile

τ	0.1	0.25	0.5	0.75	0.9
β_0	29.03	29.30	27.22	10.14	10.16
β_1	10.41	10.47	11.17	12.64	12.64

According to Theorem 6, the estimated expected values and variances of ε_τ, $\tau=0.1,0.25,0.5,0.75,0.9$, respectively are shown in Table 3.

Table 3 Estimated expected values and variances

τ	0.1	0.25	0.5	0.75	0.9
e_τ	9.43	8.37	0.75	-2.65	-2.67
σ_τ^2	201.72	193.75	137.46	270.46	270.46

From Table 4, it can be seen that mean square errors with imprecise observations are as follows.

Table 4 Mean square errors under imprecise observations

τ	0.1	0.25	0.5	0.75	0.9
$mse_{1\tau}$	290.66	263.82	138.03	277.48	277.57

The mean square error in the UQR model can reflect the fit of different quartiles, and the whole pattern of the data can be judged based on it. As shown in Table 4, by calculating the mse_τ values at different quartiles, with the exception of the quantile with $\tau = 0.5$, all other mse_τ are about 300, from 0.1 to 0.5, the values of the quantiles increase and mse_τ decrease. From 0.5 to 0.9, mse_τ increase with the increase of quantiles. In other words, the data are mostly concentrated around the $\tau=0.5$ quantile regression line.

Next, a new imprecise observation $\tilde{x} = \mathcal{L}(30,31)$ is given. From Table 5, the forecast values of the uncertain variables and the prediction intervals at the confidence level $\alpha = 95\%$ are predicted.

From the values of the coefficients for each quantile in Table 2, the corresponding lines are drawn as follows (Fig 1).

Table 5 Forecast values and confidence intervals

τ	μ	con$f$
0.10	355.98	[327.29,384.67]
0.25	356.92	[328.81,385.03]
0.50	368.51	[344.83,392.19]
0.75	392.96	[359.74,426.18]
0.90	392.96	[359.74,426.18]

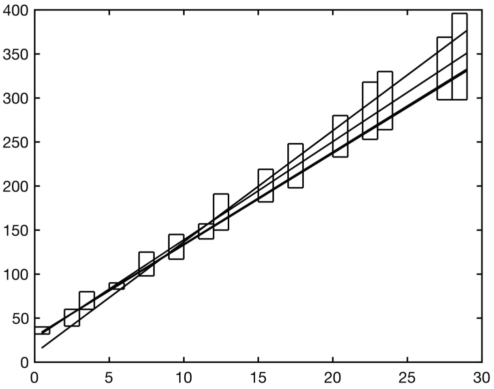

Fig 1 UQR diagram of imprecise data at five quantiles

4.2 The Case Analysis

The following case is taken from Reference [19]. Suppose an industrial engineer, employed by a company responsible for bottled soft drinks, is analyzing the working performance of a vending machine. A man randomly visits 25 retail stores equipped with vending machines to observe the delivery time (in minutes) and the quantity delivered (in cases) for each store in turn.

Due to the data recording mechanism, imprecision of the data is inevitable, so it is more reasonable to describe imprecise data as uncertain variables. In Table 6, data Y_i and X_i are from the original data in Reference [19], $i = 1, 2, \cdots, n$, respectively. The lower limit of each uncertain data y_i is the corresponding Y_i subtracted from itself by 20%, and the upper limit is Y_i itself, $i = 1, 2, \cdots, n$, respectively. The lower limit of each uncertain data x_i is the corresponding X_i itself, and the upper limit is the corresponding X_i of itself plus 1, $i = 1, 2, \cdots, n$, respectively.

Table 6 Data of delivery time and delivery volumes

i	Y_i	y_i	X_i	x_i	i	Y_i	y_i	X_i	x_i
1	16.18	$\mathcal{L}(12.944, 16.18)$	7	$\mathcal{L}(7,8)$	14	19.75	$\mathcal{L}(15.800, 19.75)$	6	$\mathcal{L}(6,7)$
2	11.50	$\mathcal{L}(9.200, 11.50)$	3	$\mathcal{L}(3,4)$	15	24.00	$\mathcal{L}(19.200, 24.00)$	9	$\mathcal{L}(9,10)$
3	12.03	$\mathcal{L}(9.624, 12.03)$	3	$\mathcal{L}(3,4)$	16	29.00	$\mathcal{L}(23.200, 29.00)$	10	$\mathcal{L}(10,11)$
4	14.88	$\mathcal{L}(11.904, 14.88)$	4	$\mathcal{L}(4,5)$	17	15.35	$\mathcal{L}(12.280, 15.35)$	6	$\mathcal{L}(6,7)$
5	13.75	$\mathcal{L}(11.000, 13.75)$	6	$\mathcal{L}(6,7)$	18	19.00	$\mathcal{L}(15.200, 19.00)$	7	$\mathcal{L}(7,8)$
6	18.11	$\mathcal{L}(14.488, 18.11)$	7	$\mathcal{L}(7,8)$	19	9.50	$\mathcal{L}(7.600, 9.50)$	3	$\mathcal{L}(3,4)$
7	8.00	$\mathcal{L}(6.400, 8.00)$	2	$\mathcal{L}(2,3)$	20	35.10	$\mathcal{L}(28.080, 35.10)$	17	$\mathcal{L}(17,18)$
8	17.83	$\mathcal{L}(14.264, 17.83)$	7	$\mathcal{L}(7,8)$	21	17.90	$\mathcal{L}(14.320, 17.90)$	10	$\mathcal{L}(10,11)$
9	79.24	$\mathcal{L}(63.392, 79.24)$	30	$\mathcal{L}(30,31)$	22	52.32	$\mathcal{L}(41.856, 52.32)$	26	$\mathcal{L}(26,27)$
10	21.50	$\mathcal{L}(17.200, 21.50)$	5	$\mathcal{L}(5,6)$	23	18.75	$\mathcal{L}(15.000, 18.75)$	9	$\mathcal{L}(9,10)$
11	40.33	$\mathcal{L}(32.264, 40.33)$	16	$\mathcal{L}(16,17)$	24	19.83	$\mathcal{L}(15.864, 19.83)$	8	$\mathcal{L}(8,9)$
12	21.00	$\mathcal{L}(16.800, 21.00)$	10	$\mathcal{L}(10,11)$	25	10.75	$\mathcal{L}(8.600, 10.75)$	4	$\mathcal{L}(4,5)$
13	13.50	$\mathcal{L}(10.800, 13.50)$	4	$\mathcal{L}(4,5)$	-	-	-	-	-

We choose a unary linear regression model

$$y_i = \beta_{0\tau} + x_i\beta_{1\tau} + \varepsilon_{i\tau}, \quad i = 1,2,\cdots,n, \tau \in (0,1)$$

to fit the uncertain observations in Table 6. According to Theorem 5, UQR estimators $\beta^*_{0\tau}$, $\beta^*_{1\tau}$ of $\beta_{0\tau}$, $\beta_{1\tau}$ respectively are the optimal solutions of the following formula minimization problem:

$$\min_{\beta_{0\tau},\beta_{1\tau},\tau\in(0,1)}\Big(\sum_{y_i<\beta_{0\tau}+x_i\beta_{1\tau}}(1-\tau)\int_0^{\upsilon_i}|\Phi_i^{-1}(\alpha)-\beta_{0\tau}-\Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}|\,\mathrm{d}\alpha+\sum_{y_i\geq\beta_{0\tau}+x_i\beta_{1\tau}}\tau\int_{\upsilon_i}^1|\Phi_i^{-1}(\alpha)-\beta_{0\tau}-\Psi_i^{-1}(\alpha,\beta_{1\tau})\beta_{1\tau}|\,\mathrm{d}\alpha\Big),$$

where

$$\Psi_i^{-1}(\alpha,\beta_{1\tau}) = \begin{cases} \Psi_i^{-1}(1-\alpha), & \beta_{1\tau}\geq 0, \\ \Psi_i^{-1}(\alpha), & \beta_{1\tau}<0, \end{cases}$$

$$\upsilon_i = \begin{cases} \frac{\beta_{0\tau}-(b_{1i}-\beta_{1\tau}b_{2i})}{a_{1i}-\beta_{1\tau}a_{2i}-b_{1i}+\beta_{1\tau}b_{1i}}, & \beta_{1\tau}\geq 0, \\ \frac{a_{1i}-\beta_{1\tau}a_{2i}-\beta_{0\tau}}{a_{1i}-\beta_{1\tau}a_{2i}+b_{1i}-\beta_{1\tau}b_{2i}}, & \beta_{1\tau}<0, \end{cases}$$

for $i = 1,\cdots,n$, $\tau \in (0,1)$.

Through the formulations above, we can obtain estimates of the UQR in Table 7, when τ is equal to 0.1, 0.25, 0.5, 0.75, 0.9, respectively. According to Theorem 6, the estimated expected values and variances of ε_τ, τ=0.3,0.4,0.5,0.6,0.7, respectively are shown in Table 8.

Table 7 Estimated coefficients of each quantile

τ	0.3	0.4	0.5	0.6	0.7
β_0	1.04	1.51	3.21	0.06	-0.30
β_1	1.90	1.90	1.77	2.20	2.20

Table 8 Estimated expected values and variances

τ	0.3	0.4	0.5	0.6	0.7
e_τ	1.46	0.99	0.54	-0.30	0.06
σ^2_τ	14.56	14.56	16.14	16.93	16.93

Next, a new imprecise observation $\tilde{x} = \mathcal{L}(31,32)$ is given. From Table 9, the forecast values of the uncertain variables and the confidence intervals with the confidence level $\alpha = 95\%$ are predicted.

Table 9 Forecast values and confidence intervals

τ	μ	con$f$
0.3	62.47	[54.76,70.18]
0.4	62.47	[54.76,70.18]
0.5	59.45	[51.34,67.56]
0.6	69.05	[60.74,77.36]
0.7	69.05	[60.74,77.36]

By Table 7, the corresponding lines are drawn as follows (Fig 2).

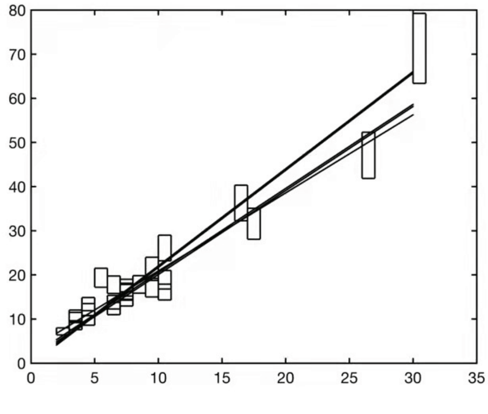

Fig 2 UQR diagram of imprecise data in different quantiles

The mean square errors are $mse_{2\tau}$ in Reference [19], and the mean square errors of UQR are $mse_{1\tau}$. The results are as follows in Table 10.

Table 10 Mean square errors

τ	0.3	0.4	0.5	0.6	0.7
$mse_{1\tau}$	16.69	15.35	16.43	17.03	16.93
$mse_{2\tau}$	25.27	24.67	23.81	17.89	19.49

In the UQR model, the mean square error is an important evaluation index. By comparing the mean square errors in Table 10, they show that the mean square errors of UQR are smaller than the mean square errors of the traditional quantile regression model in different quantiles. In other words, UQR estimation is more suitable.

5 Conclusion

This paper proposes a new uncertain quantile regression (UQR) model to compensate the uncertain least absolute deviations for uncertain multivariate regression model in Reference [19]. In terms of model applications, the UQR model can reasonably describe the variation of the response variables and predictor variables at each quantile. And potential different solutions have very useful interpretative significance in different quantiles. In terms of data, uncertain data can be applied to the UQR model. This new model combines leverages uncertainty theory and quantile regression to provide a more comprehensive explanation of the problem. So the UQR model is more reasonable and reliable.

References:

[1] KOENKER R, BASSETT J R G. Regression quantiles[J]. Journal of the Econometric Society, 1978, 46(1): 33-50.

[2] EIDE E R, SHOWALTER M H. Factors affecting the transmission of earning across generations: a quantile regression approach[J]. Journal of Human Resources, 1999, 34(2): 253-267.

[3] BEIRLANT J, GOEGEBEUR Y. Regression with response distributions of pareto-type[J]. Computational Statistics and Data analysis, 2003, 42(4): 595-619.

[4] SOMERS M, WHITTAKER J. Quantile regression for modelling distributions of profit and loss[J]. European Journal of Operational Research, 2007, 183(3): 1477-1487.

[5] NIKOLAOU K. The behaviour of the real exchange rate: evidence from regression quantiles[J]. Journal of Banking and Finance, 2008, 32(5): 664-679.

[6] YU K, MOYEED R A. Bayesian quantile regression[J]. Statistics and Probability Letters, 2001, 54(4): 437-447.

[7] TANG Q G, KONG L L. Quantle regression in functional linear semiparametric model[J]. Statistics, 2017, 51(6): 1342-1358.

[8] SUN M X. Functional data quantile regression model and its application (in Chinese)[D]. Wuhan: Huazhong University of Science and Technology, 2019.

[9] HE Y, YAN Y, XU Q. Wind and solar power probability density prediction via fuzzy information granulation and support vector quantile regression[J]. International Journal of Electrical Power and Energy Systems, 2019, 113: 515-527.

[10] LIU B. Uncertainty theory[M]. Berlin: Springer, 2007.

[11] LIU B. Some research problems in uncertainty theory[J]. Journal of Uncertain Systems, 2009, 3(1): 3-10.

[12] LIU B. Theory and practice of uncertain programming[M]. Berlin: Springer, 2009.

[13] LIU B. Fuzzy process, hybrid process and uncertain process[J]. Journal of Uncertain Systems, 2008, 2(1): 3-16.

[14] LIU B. Uncertainty theory: a branch of mathematics for modeling human uncertainty[M]. Berlin: Springer, 2010.

[15] MONTGOMERY D C, PECK E A, VINING G G. Introduction to linear regression analysis[M]. Hoboken: Wiley, 2021.

[16] YAO K, LIU B. Uncertain regression analysis: an approach for imprecise observations[J]. Soft Computing, 2018, 22(17): 5579-5582.

[17] LIO W, LIU B. Residual and confidence interval for uncertain regression model with imprecise observations[J]. Journal of Intelligent and Fuzzy Systems, 2018, 35(2): 2573-2583.

[18] YE T, LIU Y. Multivariate uncertain regression model with imprecise observations[J]. Journal of Ambient Intelligence and Humanized Computing, 2019, 11(11): 4941-4950.

[19] LIU Z, YANG Y. Least absolute deviations estimation for uncertain regression with imprecise observations[J]. Fuzzy Optimization and Decision Making, 2020, 19(1): 33-52.

责任编辑：艾合麦提·吾买尔