Statistics and Probability Letters 113 (2016) 30–37

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Estimation of extreme conditional quantiles through an extrapolation of intermediate regression quantiles

Fengyang He [a], Yebin Cheng [b,*], Tiejun Tong [c]

[a] *School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China*
[b] *Glorious Sun School of Business and Management, Donghua University, Shanghai, China*
[c] *Department of Mathematics, Hong Kong Baptist University, Hong Kong*

ARTICLE INFO

Article history:
Received 4 June 2015
Received in revised form 3 November 2015
Accepted 30 January 2016
Available online 10 February 2016

Keywords:
Extreme conditional quantile
Intermediate regression quantile
Extreme value

ABSTRACT

The paper proposes a new estimation method for the extreme conditional quantiles through an extrapolation of the intermediate regression quantiles for the linear quantile regression model, and obtains the asymptotic properties of the proposed estimator in the general framework.

1. Introduction

Quantile regression is a useful empirical tool in modern science for analyzing how a vector of regressor variables influences the features of the conditional distribution of a response variable. In many applications, the features of interest are the extremal (extreme order and intermediate order) quantiles of the conditional distribution. Chernozhukov (2005) developed a theory of quantiles regression in the tails, by integrating the linear quantile regression model and the extreme value theory (EVT). They obtained the large sample properties of extremal quantile regression estimators with the tails restricted to the domain of minimum attraction. Specifically, in large samples, they demonstrated intermediate order regression quantiles and their functionals converge to normal vectors with variance matrices dependent on the tail parameters and the regressor design, and extreme order regression quantiles converge weakly to arg min functionals of stochastic integrals of Poisson processes that depended on regressors. However, direct quantile estimation by means of conventional methods in tails is often unstable owing to data sparseness. In view of this, Chernozhukov (2005) noted that one interesting direction for further research is to examine the estimation of the extreme conditional quantiles defined through an extrapolation of the intermediate regression quantiles, so as to provide more practical and reliable inference procedures.

Based on the EVT, Wang et al. (2012) developed two methods for high conditional quantiles by first estimating the intermediate conditional quantiles and then extrapolating these estimators to the high tails. However, their estimation methods are only suitable for the case of heavy-tailed distributions with tail index $\gamma > 0$. Parametric models are also considered in, for example, Smith (1989) and Davison and Smith (1990), where some techniques for extreme values are extended to the point process view of high-level exceedance. Other papers on conditional extremes have focused on extremal regression quantiles in the nonlinear quantile regression models. Davison and Ramesh (2000) introduced a semi-parametric approach

* Corresponding author.
E-mail address: chengyebin@hotmail.com (Y. Cheng).

http://dx.doi.org/10.1016/j.spl.2016.01.024
0167-7152/

to model trends in sample extremes. Hall and Tajvidi (2000) provided a nonparametric estimation of the temporal trend when fitting parametric models to extreme values. In the circumstances of kernel smoothing, Daouia et al. (2011) extended the large sample theory into the tails in the particular setting of a heavy-tailed conditional distribution. Daouia et al. (2013) further developed a unified asymptotic theory for the kernel-smoothed conditional extremes in a general setting.

In this paper, we propose a new estimation method for the extreme conditional quantiles through an extrapolation of the intermediate regression quantiles. In particular, we will derive the asymptotic properties of the proposed estimator in a general framework with the error distribution ranging from short, light to heavy-tailed, corresponding to $\gamma < 0$, $\gamma = 0$ to $\gamma > 0$, respectively. Hence, from another point of view, our paper can also be treated as a useful supplement to the influential papers of Chernozhukov (2005) and Wang et al. (2012). The rest of the paper is organized as follows. In Section 2, we introduce the notations, assumptions, and the extreme value restrictions on the linear quantile regression model. In Section 3, we present our main results and eventually obtain the asymptotic normality for the extreme order-φ_T regression quantiles with $\varphi_T/\tau_T \to 0$ and $\tau_T \to 0$, $\tau_T T \to \infty$, where T stands for the sample size. In Section 4, we conduct a simulation study to assess the finite sample performance of the proposed estimator. We then illustrate an empirical example in Section 5 and provide the technical proofs in the supplementary materials (see Appendix A).

2. Model and assumptions

For the sake of consistency, we follow the similar notations as those in Chernozhukov (2005). Specifically, we let Y be the response variable in $\mathbb{R}$, and $X = (1, X'_{-1})'$ is a regressor vector of size d. The conditional distribution function of Y given $X = x$ is denoted by $F_Y(\cdot|x)$. The present focus is on $Q_Y(\tau|x) = F_Y^{-1}(\tau|x) = \inf\{y : F_Y(y|x) > \tau\}$, where $\tau \to 0$. Let $\{Y_t, X_t,\ t = 1, \dots, T\}$ be an independent sample of the random vector (Y, X) in $\mathbb{R} \times \mathbb{R}^d$. We investigate the following quantile regression model:

$$Q_Y(\tau|x) = x'\beta(\tau), \quad \forall \tau \in \mathcal{I},\ x \in \mathbf{X}, \tag{1}$$

where $\mathcal{I} = [0, \eta]$ for some $0 < \eta < 1$, and $\mathbf{X}$ is a compact subset of $\mathbb{R}^d$. The quantile slope $\beta(\tau)$ can be estimated by solving the least asymmetric absolute deviation problem:

$$\hat{\beta}(\tau) = \arg\min_{\beta\in\mathbb{R}^d} \sum_{t=1}^{T} \rho_\tau(Y_t - X_t'\beta) \quad \text{where } \rho_\tau(u) = u\{\tau - \mathbb{I}(u \le 0)\}. \tag{2}$$

For deriving the large sample properties of $\hat{\beta}(\tau)$ as $\tau \to 0$, Chernozhukov (2005) distinguished the following three cases of sample regression quantiles according to the classical theory of order statistics: (a) the extreme order case, where $\tau_T \downarrow 0$ and $\tau_T T \to k > 0$, (b) the intermediate order case, where $\tau_T \downarrow 0$ and $\tau_T T \to \infty$, and (c) the central order case, where $\tau \in (0, 1)$ is fixed and $T \to \infty$. Under case (c), one can apply the conventional methods to estimate $\beta(\tau)$.

Let v be a random variable with distribution function F_v and lower endpoint $x_* = 0$ or $x_* = -\infty$, where $x_* = \inf\{y : F_v(y) > 0\}$. According to Resnick (1987), F_v has tail of type 1, 2 or 3 if for

- type 1: as $t \downarrow x_* = 0$ or $-\infty$,

$$F_v(t + xa(t)) \sim F_v(t)e^x \quad \forall x \in \mathbb{R},\ \gamma \equiv 0,$$

- type 2: as $t \downarrow x_* = -\infty$,

$$F_v(tx) \sim x^{-\frac{1}{\gamma}} F_v(t) \quad \forall x > 0,\ \gamma > 0,$$

- type 3: as $t \downarrow x_* = 0$,

$$F_v(tx) \sim x^{-\frac{1}{\gamma}} F_v(t) \quad \forall x > 0,\ \gamma < 0,$$

where $a(t) = \int_{x_*}^{t} F_v(x)dx/F_v(t)$ is a positive function for $t > x_*$. The parameter γ is commonly called the extreme value index or tail index, and F_v with tail of type 1, 2 or 3 is said to be in the minimum domain of attraction (MDA). We also use $a(t) \sim b(t)$ to denote that $a(t)/b(t) \to 1$ as t tends to infinity.

To achieve the asymptotic results, we assume some regularity conditions as follows.

(C1) Suppose $F_v(\cdot)$ is twice differentiable and

$$\lim_{t\downarrow x_*} \frac{F_v(t)F_v''(t)}{\left(F_v'(t)\right)^2} = 1 + \gamma, \tag{3}$$

where $F_v'(\cdot)$ and $F_v''(\cdot)$ are the first and second derivatives of $F_v(\cdot)$, respectively.