Journal of Econometrics 191 (2016) 196–216

Contents lists available at ScienceDirect

Journal of Econometrics

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

A direct approach to inference in nonparametric and semiparametric quantile models

Yanqin Fan[a,*], Ruixuan Liu[b]

[a] Department of Economics, University of Washington, Seattle, WA 98195, United States

[b] Department of Economics, Emory University, Atlanta, GA 30322, United States

ARTICLE INFO

Article history:
Received 3 October 2013
Received in revised form
8 August 2014
Accepted 15 January 2015
Available online 30 October 2015

JEL classification:
C12
C14
C21

Keywords:
Generic confidence interval
Generic confidence band
Partially linear quantile regression
Single-index quantile regression
Rearranged quantile curve

ABSTRACT

We construct a generic confidence interval for a conditional quantile via the direct approach. It avoids estimating the conditional density function of the dependent variable given the covariate and is asymptotically valid for any conditional quantile, any conditional quantile estimator, and any data structure, provided that certain weak convergence of the conditional quantile process holds for the original quantile estimator. We also construct a generic confidence band for the conditional quantile function across a range of covariate values. By using Yang–Stute estimator and two semiparametric quantile functions, we demonstrate the flexibility and simplicity of the direct approach. The advantages of our new confidence intervals are borne out in a simulation study.

Published by Elsevier B.V.

1. Introduction

In their seminal paper, Koenker and Bassett (1978) propose to use linear quantile regression to examine effects of an observable covariate on the distribution of a dependent variable other than the mean. Since then, linear quantile regression has become a dominant approach in empirical work in economics, see e.g., Buchinsky (1994) and Koenker (2005). Linearity adopted in Koenker and Bassett (1978) has been relaxed to accommodate possibly nonlinear effects of the covariates on the conditional quantile of the dependent variable in nonparametric and semiparametric quantile regression models. The 'check function' approach of Koenker and Bassett (1978) has been extended to estimating these models as well, see e.g., Chaudhuri (1991), Yu and Jones (1998) and Guerre and Sabbah (2012) for local polynomial estimation of regression quantiles; Lee (2003) and Song et al. (2012) for partial linear quantile regression models; Ichimura and Lee (2010), and Kong and Xia (2012) for single index quantile regression models.[1]

For nonparametric conditional quantiles, an alternative estimation approach to the 'check function' approach is taken in Stute (1986), Bhattacharya and Gangopadhyay (1990), Fan and Liu (2015), and Li and Racine (2008), among others. In this approach, the conditional distribution function of the dependent variable Y given the covariate X is estimated first and the generalized inverse of this estimator at a given quantile level $p \in (0, 1)$ is taken as an estimator of the pth conditional quantile. Stute (1986) and Bhattacharya and Gangopadhyay (1990) focus on univariate covariate and estimate the conditional distribution function by k-NN method, while Fan and Liu (2015) and Li and Racine (2008) allow for multivariate covariate and adopt respectively k-NN and kernel estimators of the conditional distribution function. Donald and Hsu (2014) use this approach to estimate conditional quantiles of potential outcomes in a treatment effect model.

* Correspondence to: Department of Economics, University of Washington, Box 353330, Seattle, WA 98195, United States. Tel.: +1 206 543 8172; fax: +1 206 685 7477.

E-mail address: fany88@uw.edu (Y. Fan).

[1] Conditional quantile function also plays an important role in the non-separable structural econometrics literature, see e.g., Chesher (2003), and in the study of quantile treatment effects, see e.g., Fan and Park (2011).

http://dx.doi.org/10.1016/j.jeconom.2015.01.009
0304-4076/Published by Elsevier B.V.

Under regularity conditions, existing work establish asymptotic normality of the conditional quantile estimators which is the basis for the Wald-type inference, i.e., using the t statistic to test hypotheses or form confidence intervals (CI) for the true conditional quantiles. Regardless of the approach used to estimate the conditional quantile in parametric, semiparametric, or nonparametric quantile regression models, one common feature of the asymptotic distributions of the conditional quantile estimators is that their asymptotic variances depend on the conditional (quantile) density function of Y given $X = x$ and some even depend on the density function of X, see e.g., Horowitz (1998), Li and Racine (2008), Hardle and Song (2010), and Song et al. (2012), among others. As a result, inference procedures for the conditional quantiles based on the asymptotic distributions of these estimators require consistent estimators of the conditional (quantile) density function of Y given $X = x$ and/or the density of X both involving bandwidth choices. Numerical evidence presented in De Angelis et al. (1993), Horowitz (1998), and Kocherginsky et al. (2005) shows that although asymptotically valid, these inference procedures are sensitive in finite samples to the choice of the smoothing parameter used to estimate the conditional (quantile) density function.

Various alternative approaches have been proposed in the current literature to improve on the finite sample performance of Wald-type inferences. Most of these are developed for linear or parametric conditional quantile regression models. First, Goh and Knight (2009) propose a different scale statistic to standardize the estimator of the model parameter in linear quantile regression models resulting in a nonstandard inference procedure; second, Zhou and Portnoy (1996) construct CIs/bands directly from pairs of estimates of conditional quantiles in the location–scale forms of linear quantile regression models extending the direct or order statistics approach for sample quantiles in Thompson (1936), see also Serfling (1980), and van der Vaart (1998); third, Gutenbrunner and Jureckova (1992) employ rank scores to test a class of linear hypotheses; fourth, Whang (2006) and Otsu (2008) apply the empirical likelihood approach to parametric quantile regression models; lastly, MCMC related approaches have been proposed to improve standard resampling or simulation paradigms in parametric quantile regression models: He and Hu (2002) obtain their resampling estimators which solve one-dimensional estimating equation recursively along the generated Markov chain; and Chernozhukov et al. (2009) develop finite sample inference procedures based on conditional pivotal statistics. A nice survey of various inference procedures targeted at linear quantile regression models could be found in Kocherginsky et al. (2005).

Compared with parametric quantile regression models, inference in nonparametric and semiparametric quantile regression models is still in its infancy. The only alternative approach to the Wald-type and bootstrap inferences that is currently available is the empirical likelihood procedure in Xu (2013) for nonparametric quantile regression models.[2] In semiparametric quantile regression models including partial linear and single index models, only Wald-type and bootstrap inferences are available. Although the empirical likelihood approach in Xu (2013) avoids estimation of the conditional (quantile) density function and performs better than the Wald-type inference procedures, it is known to be computationally costly. Among existing approaches to inference in parametric quantile regression models, the direct approach and the Wald approach are the simplest to implement and least costly computationally. The direct approach only requires computing pairs of the quantile estimate. Moreover it does not rely on any estimate of the conditional (quantile) density function and exhibits superior finite sample performance to the Wald-type inference, see Zhou and Portnoy (1996). However, as discussed in Portnoy (2012), it appears that the direct approach in Zhou and Portnoy (1996) has theoretical justification only under location–scale forms of linear quantile regression models.

This paper aims at bridging this gap. Specifically, it makes two main contributions to inference on conditional quantiles. First, we construct a generic CI for a conditional quantile from any given estimator of the conditional quantile via the direct approach. Our generic CI makes use of two estimates of the conditional quantile function evaluated at two appropriately chosen quantile levels. If the original quantile estimator is monotone in the quantile level $p \in (0, 1)$, then the two estimates are computed from this estimator; else the two estimates are computed from the monotone rearranged version of the original quantile estimator as proposed in Chernozhukov et al. (2010). In contrast to the standard Wald type CI, ours circumvents the need to estimate the conditional density function of the dependent variable given the covariate. We show that our new CI is asymptotically valid for any conditional quantile (parametric, nonparametric, or semiparametric), any conditional quantile estimator (standard kernel, local polynomial or sieve estimates), and any data structure (random samples, time series, or censored data), provided that certain weak convergence of the conditional quantile process holds for the preliminary quantile estimator. In the same spirit, we also construct a generic confidence band (CB) for the conditional quantile function across a range of covariate values focusing on the nonparametric setting and a class of quantile estimators obtained from inverting proper estimators of the conditional distribution function of Y given X. Since members of this class of quantile estimators are monotone by construction, monotone rearrangement is avoided. Second, we use a specific estimator, the Yang–Stute (also known as the symmetrized k-NN) estimator for a nonparametric quantile function, and two popular semiparametric quantile functions to demonstrate that oftentimes by a judicious choice of the quantile estimator combined with the specific model structure, one may further take advantage of the flexibility and simplicity of the direct approach. For instance, by using the Yang–Stute estimator, we construct CIs and bands for a nonparametric and two semiparametric quantile functions that are free from additional bandwidth choices involved in estimating not only the conditional but also the marginal density functions and that are very easy to compute. The reason that we choose the Yang–Stute estimator is its simplicity and elegance; it inherits the so-called asymptotic distributional-free property (Stute, 1984b) and avoids estimating the covariate's density function (unlike standard kernel estimators), so we are able to eliminate all unnecessary tuning parameters in our CIs and CBs. Besides, as we directly invert conditional distribution functions, the resulting conditional quantile estimators are indeed monotone, so there is no need for monotone rearrangement. Of course, practitioners are free to choose their favorite preliminary quantile estimators and under the mild high level assumptions below, our generic CIs/CBs would apply.

Like the empirical likelihood CI for a nonparametric conditional quantile in Xu (2013), our CIs/bands for nonparametric conditional quantiles based on the Yang–Stute estimator internalize the conditional quantile density estimation of Y given X and the covariate density estimation and they are not necessarily symmetric. Compared with Xu (2013), our procedure is much easier to implement and does not require optimization. For conditional quantiles in partial linear and single index quantile regressions, direct applications of the generic CI and CB would require monotone rearrangement, but by making use of the model structures, we construct CIs and CBs that are easy to implement avoiding monotone rearrangement. A small scale simulation study demonstrates the advantages and feasibility of our CIs/bands over existing ones in practically relevant model set-ups.

[2] After finishing the first version of this paper, we came across Kaplan (2013), who proposed similar inference procedures for nonparametric conditional quantiles to Example 2.1 in Section 2 of this paper.

Download English Version:

https://daneshyari.com/en/article/5095742

Download Persian Version:

https://daneshyari.com/article/5095742

Daneshyari.com