Sankhyā : The Indian Journal of Statistics
Special Issue on Quantile Regression and Related Methods
2005, Volume 67, Part 2, pp 187-199

Isotonic Quantile Regression: Asymptotics and Bootstrap

Jason Abrevaya
Purdue University, W. Lafayette, IN, USA

Abstract

This paper considers a nonparametric model in which (i) the conditional quantile function is assumed to be nondecreasing and (ii) the distribution of the error disturbance may depend upon the covariate. For this general model, the (pointwise) limiting distribution of the isotonic quantile regression estimator (Casady and Cryer (1976)) is derived. Since the bootstrap is not consistent for this estimator, an adjusted version of the bootstrap is discussed as an alternative method for inference. An empirical application to birthweights is considered.

AMS (2000) *subject classification.* 62G08, 62G09, 62G20.
Keywords and phrases. Isotonic quantile regression, cube-root asymptotics, bootstrap.

1 Introduction

For a given $\lambda \in (0, 1)$, consider the following model:

$$Q_\lambda(y|x) = f_\lambda(x), \tag{1}$$

where Q_λ is the λ-quantile function (e.g., the median function for $\lambda = 0.5$) and f_λ is a nondecreasing function to be estimated. Equivalently, equation (1) could be written as a model in terms of an error disturbance as follows:

$$y = f_\lambda(x) + \epsilon, \quad \text{where } Q_\lambda(\epsilon|x) = 0. \tag{2}$$

To be precise, letting G denote the (conditional) cdf of $\epsilon|x$ and using the standard definition of a quantile, the quantile restriction on ϵ in (2) is equivalent to $\inf\{e : G(e|x) \geq \lambda\} = 0$ or, if G is continuous with positive density everywhere, $G(0|x) = \lambda$. Casady and Cryer (1976) proposed the *isotonic quantile regression estimator* of f_λ, generalizing the isotonic median regression estimator previously considered by Cryer et. al. (1972).[1] Other nonparametric

[1] Casady and Cryer (1976) and Casady et. al. (1972) use the terms *monotone percentile regression estimator* and *monotone median regression estimator*, respectively.

estimators, such as kernel-based estimators or local polynomial estimators, could also be used to estimate f_λ (without imposing monotonicity); such estimators are not the focus of this paper.

Suppose that the observed sample is $\{(y_i, x_i)\}_{i=1}^n$, where $x_1, \ldots, x_n$ are design points in $[0, 1]$. The isotonic quantile regression estimator, which we will denote as $\hat{f}_{\lambda,n}$, has the following min-max representation given by Casady and Cryer (1976):

$$\hat{f}_{\lambda,n}(x_0) \equiv \min_{u \leq x_0} \max_{v \geq x_0} Q_\lambda(y_i : u \leq x_i \leq v). \tag{3}$$

The estimator also minimizes the objective function

$$\sum_{i=1}^{n} (y_i - f(x_i))(\lambda - 1(y_i \leq f(x_i))) \tag{4}$$

over all nondecreasing functions f.[2] Due to its min-max representation in (3), it is well-known that the estimator can be computed using a pool adjacent violators (PAV) algorithm, similar to the case of isotonic mean regression (Brunk (1970) and Hanson et. al. (1973)) and isotonic median regression (Robertson and Wright (1973)).

Casady and Cryer (1976) proved pointwise consistency of $\hat{f}_{\lambda,n}$. The (pointwise) limiting distribution of the isotonic mean regression estimator has been known for some time (see, e.g., Prakasa Rao (1969) and Brunk (1970)). For the quantile case, the only analogous result is given by Wang and Huang (2002), who provide the limiting distribution for the median case ($\lambda = 0.5$) when the distribution of ϵ does not depend upon x. This paper generalizes Wang and Huang (2002) in two important directions. First, the model in (2) allows for heteroscedastic error disturbances; the theoretical implication is that, even in the median case, the limiting distribution depends upon the conditional distribution of $\epsilon | x$ (rather than the marginal distribution of ϵ). Second, the theoretical results generalize the median case by considering all $\lambda \in (0, 1)$ values. This latter generalization is particularly important for empirical applications where conditional quantiles other than the median are of interest.

Section 2 gives the limiting-distribution result for the isotonic quantile regression estimator. Section 3 discusses the use of the bootstrap as an alternative for asymptotic inference. Although Abrevaya and Huang (2005) have shown that the bootstrap is inconsistent for a class of cube-root estimators

[2]This fact follows from a straightforward extension of Robertson and Wright (1973, Theorem 2.7).

that includes isotonic estimators, their results suggest a simple method of asymptotic inference based upon the bootstrap. Section 4 considers an application to birthweights, in particular examining the quantiles of second-child birthweight conditional on first-child birthweight.

2 Results

Let $x_0 \in (0,1)$ denote a point and $\lambda \in (0,1)$ denote a quantile at which the conditional quantile function $f_\lambda(x_0)$ is to be estimated. As above, G is used to denote the distribution function of $\epsilon|x$. Then, whenever its derivative exists, let $g(z|x) \equiv G'(z|x)$ denote the density function of $\epsilon|x$.

Before stating the main result, the following assumptions are made:

ASSUMPTION 1. (Differentiability)

(i) For all x in a neighbourhood of x_0, $f_\lambda(x)$ is twice continuously differentiable.
(ii) For all x in a neighbourhood of x_0 and z in a neighbourhood of zero, $G(z|x)$ is twice continuously differentiable with respect to z with $g(z|x)$ strictly positive.
(iii) For all x in a neighbourhood of x_0 and z in a neighbourhood of zero, $G(z|x)$ and $g(z|x)$ are twice continuously differentiable with respect to x.

ASSUMPTION 2. (Boundedness)

The following functions are bounded by a constant for x in a neighbourhood of x_0 and z in a neighbourhood of zero: $f''_\lambda(x)$, $g'(z|x)$, $\frac{\partial^2 G(z|x)}{\partial x^2}$, and $\frac{\partial^2 g(z|x)}{\partial x^2}$.

The existence and boundedness of the various derivatives in these assumptions are used for the Taylor-series arguments in the derivation of the limiting distribution. These conditions could be weakened somewhat, but have been stated in the present form in order to simplify matters.

The limiting distribution of the isotonic quantile regression estimator is given by:

THEOREM 2.1 *Let Assumptions 1-2 hold. Let $H_n(x) \equiv n^{-1}\sum_{i=1}^{n} 1(x_i \leq x)$. Suppose that there is a distribution function H, which is continuously differentiable in a neighbourhood of x_0, with $h(x_0) > 0$ (where $h(x) \equiv H'(x)$) and*

$$max_x\,|H_n(x) - H(x)| = o(n^{-1/3}).$$

Then, if $f'_\lambda(x_0) > 0$,

$$\left\{\frac{nh(x_0)g(0|x_0)^2}{4\lambda(1-\lambda)f'_\lambda(x_0)}\right\}^{1/3}\left(\hat{f}_{\lambda,n}(x_0) - f_\lambda(x_0)\right) \xrightarrow{d} \mathcal{Z}, \tag{5}$$

where $\mathcal{Z}$ is the distribution of the maximum of $B(t)-t^2$ (where B is two-sided Brownian motion originating at zero).

The proof of Theorem 2.1 is provided in the Appendix. Note that the scaling constant contains $\lambda(1-\lambda)$, which is familiar from marginal quantile estimation and linear-regression quantile estimation (Koenker and Bassett (1978)).

The limiting distribution $\mathcal{Z}$ is common across a class of cube-root estimators of scalar parameters, including the isotonic mean regression estimator and the least median of squares estimator of the shift parameter. Groeneboom and Wellner (2001) give several additional examples of estimators having the limiting distribution $\mathcal{Z}$ (having different scaling constants than Theorem 2.1). Groeneboom (1989) derives an analytical expression for the density of $\mathcal{Z}$ in terms of Airy functions. Using this representation, Groeneboom and Wellner (2001) develop an algorithm to accurately compute all features of the distribution. As a result, asymptotic inference based upon Theorem 2.1 is feasible. Asymptotic confidence intervals can be constructed using the quantiles of $\mathcal{Z}$ and consistent estimators of the quantities $h(x_0)$, $g(0|x_0)$, and $f'_\lambda(x_0)$.

Although the case of fixed-design x has been considered here, the results of Theorem 2.1 easily extend to the case of random x. In that case, the functions H and h above would correspond to the marginal cdf and pdf, respectively, of the random variable from which the x_i's are drawn.

One drawback of the isotonic quantile regression estimator is its step-function nature. In particular, the estimator $\hat{f}_{\lambda,n}$ does not satisfy the twice differentiability assumption (Assumption 1(i)) of the underlying function f_λ. The issue of "regularizing" the isotonic estimator has been studied in the literature in the context of estimating the conditional *mean* function. Mammen (1991) provides a theoretical analysis of such an estimator, which is obtained by applying a kernel smoother to the isotonic point estimates. This estimator satisfies twice continuous differentiability as long as it is satisfied by the underlying kernel function used in the second-stage smoothing.[3] A similar estimation strategy should apply to quantile estimation as well.

[3] Note that the estimator obtained by smoothing and then isotonizing, which is also considered by Mammen (1991), would not satisfy such a differentiability condition.

An estimator analogous to the isotonize-then-smooth estimator of Mammen (1991) in the context of quantile estimation would be:

$$\hat{f}^{IS}_{\lambda,n}(x) \equiv (nh_n)^{-1} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h_n}\right) \hat{f}_{\lambda,n}(x_i).$$

The estimator $\hat{f}^{IS}$ would share the differentiability properties of the kernel function K. A theoretical analysis of this estimator is an interesting topic for future research. In particular, it would be interesting to determine whether $\hat{f}^{IS}$ provides an improved (pointwise) convergence rate (as is the case for mean estimation, where the improvement is from $n^{-1/3}$ to $n^{-2/5}$).

3 The Bootstrap

Since asymptotic inference based upon Theorem 2.1 requires three additional nonparametric estimators (to estimate $h(x_0)$, $g(0|x_0)$, and $f'_\lambda(x_0)$), the bootstrap is an appealing alternative as it does not require the choice of any smoothing parameters. However, the isotonic quantile estimator $\hat{f}_{\lambda,n}(x_0)$ belongs to a class of cube-root estimators for which Abrevaya and Huang (2005) have shown that the nonparametric bootstrap is inconsistent. Rather than having a limiting distribution related to the maximum of Brownian motion minus quadratic drift (see $\mathcal{Z}$ above), the bootstrap estimator has a limiting distribution related to the difference between the maxima of two such processes. Despite the inconsistency of the bootstrap, asymptotic inference using the bootstrap is still possible. Since the limiting distributions of the estimator and the bootstrap estimator are known, the bootstrap confidence intervals can be appropriately scaled to make them valid.

Consider the nonparametric bootstrap, where a bootstrap sample $\{(y_i^*, x_i^*)\}_{i=1}^n$ is drawn with replacement from the observed sample $\{(y_i, x_i)\}_{i=1}^n$. Let $\hat{f}^*_{\lambda,n}(x_0)$ denote the bootstrap estimator — i.e., the isotonic quantile regression estimator applied to the bootstrap sample. The limiting distribution of the bootstrap estimator, conditioning on the data, follows from Abrevaya and Huang (2005, Theorem 2):

$$\left\{\frac{nh(x_0)g(0|x_0)^2}{4\lambda(1-\lambda)f'_\lambda(x_0)}\right\}^{1/3} \left(\hat{f}^*_{\lambda,n}(x_0) - \hat{f}_{\lambda,n}(x_0)\right) \xrightarrow{d} \mathcal{D}, \tag{6}$$

with

$$\mathcal{D} \equiv \arg\max(-t^2 + B_1(t)) - \arg\max(-t^2 + B_1(t) + B_2(t)),$$

where B_1 and B_2 are independent two-sided Brownian motions originating at zero.

The distribution $\mathcal{D}$ in (6), unlike $\mathcal{Z}$, does not have a known analytical form for its density. As a result, Abrevaya and Huang (2005) use simulations in order to approximate the distribution. Table 1 reports selected quantiles for $\mathcal{Z}$ (from Groeneboom and Wellner (2001)) and $\mathcal{D}$ (from Abrevaya and Huang (2005)).[4] Since both $\mathcal{Z}$ and $\mathcal{D}$ are symmetric distributions, only upper quantiles are reported in the table.

Table 1. Selected Quantiles of $\mathcal{Z}$ and $\mathcal{D}$

Quantile	$\mathcal{Z}$	$\mathcal{D}$
0.900	0.6642	0.7528
0.925	0.7434	0.9021
0.950	0.8451	1.0932
0.975	0.9982	1.3822
0.990	1.1715	1.7068
0.995	1.2867	1.9214

To illustrate how Table 1 can be used to "correct" a bootstrap confidence interval, suppose that a 90% confidence interval for $f_\lambda(x_0)$ is constructed using a series of bootstrap estimators. If the 90% confidence interval is written as

$$\left(\hat{f}_{\lambda,n}(x_0) - a, \hat{f}_{\lambda,n}(x_0) + b\right)$$

(where $a = b$ corresponds to a symmetric bootstrap interval), an asymptotically valid 90% confidence interval would be

$$\begin{aligned}&\left(\hat{f}_{\lambda,n}(x_0) - \frac{0.8451}{1.0932}a, \hat{f}_{\lambda,n}(x_0) + \frac{0.8451}{1.0932}b\right)\\ &\quad\approx \left(\hat{f}_{\lambda,n}(x_0) - 0.773a, \hat{f}_{\lambda,n}(x_0) + 0.773b\right).\end{aligned}$$

The values 0.8451 and 1.0932 correspond to the 0.95 quantiles of $\mathcal{Z}$ and $\mathcal{D}$, respectively, from Table 1. This method is illustrated in the empirical application of Section 4.

[4]The quantiles for $\mathcal{D}$ are based upon 10 million simulated draws. See Abrevaya and Huang (2005) for additional quantiles as well as 99% confidence intervals for the simulated values. Due to the large number of simulations, the confidence intervals are quite narrow. The widest 99% confidence interval, for the 0.995 quantile, is (1.9189, 1.9237).

4 Empirical Application

In this section, the isotonic quantile regression estimator is applied to a dataset on birthweights from the state of Arizona during the period 1993–2002. In particular, we estimate the conditional quantiles of second-child birthweight given first-child birthweight (for which the monotonicity assumption on f_λ is very plausible). The sample is restricted to (i) mothers who had their first two children during the period 1993–2002, (ii) mothers who had singleton births (no twins or higher plurality), and (iii) mothers who were married. Since the original data does not contain any uniquely identifying information for mothers (such as social security number), the birthdates of both parents are used in order to link the first two children born to a set of parents. The resulting sample consists of 47,684 observations. Table 2 reports descriptive statistics for the marginal distributions of first-child birthweight and second-child birthweight. The sample correlation between the two birthweights is 0.4326.

Table 2. Birthweight Sample Descriptive Statistics (in grams)

	First child	Second child
Mean	3351	3437
Stdev	523	504
10% quantile	2750	2863
25% quantile	3061	3146
50% quantile	3373	3445
75% quantile	3685	3742
90% quantile	3968	4040

Figure 1 shows the isotonic quantile regression estimates at five different quantiles (10%, 25%, 50%, 75%, 90%).

In Figure 2, the bootstrap method of the previous section is illustrated for the isotonic median estimator (the middle curve from Figure 1). The figure shows 90% symmetric confidence intervals based upon 1000 bootstrap iterations. Despite the cube-root convergence rate of the estimator, the intervals are rather tight due to the large sample size. Finally, we stress the importance of choosing a large number of bootstrap iterations in this context since the confidence intervals are based upon fairly extreme quantiles of the bootstrap distribution.

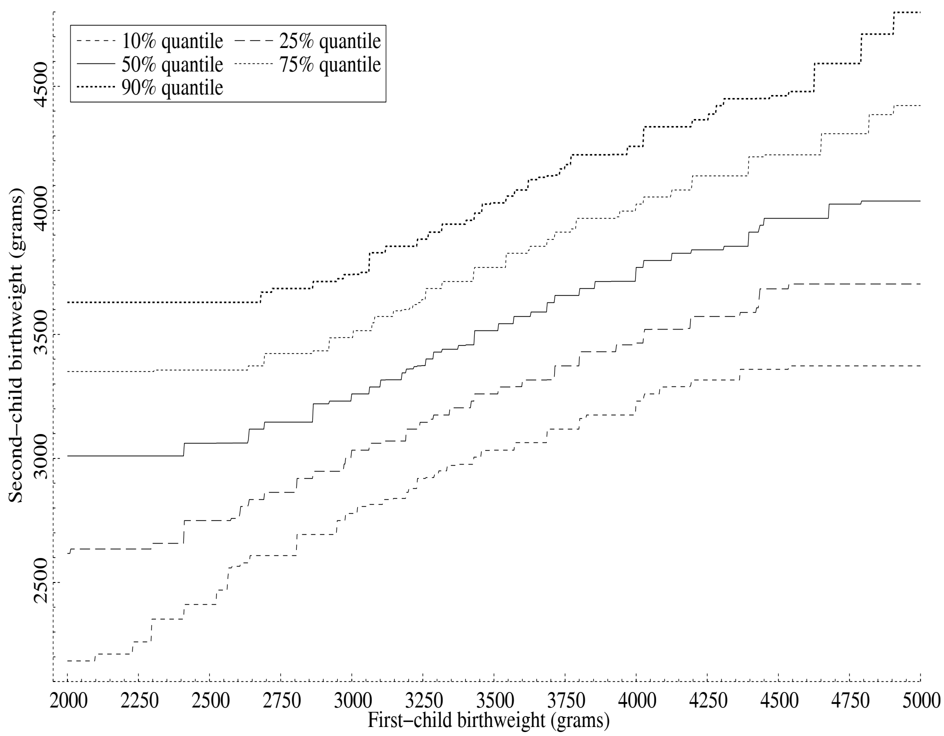

Figure 1. Isotonic Regression Results for Birthweight Data

Appendix

Proof of Theorem 2.1. To allow an easy comparison to Wang and Huang (2002), their proof and notation are followed as closely as possible. Additional notation is required to handle general λ and to deal with $G(\cdot|x)$ varying over x. The differences from their proof will be noted in the discussion. As in Wang and Huang (2002), assume for simplicity (and without loss of generality) that $x_i = i/n$ (so that $H(x) = x$ and $h(x) = 1$). Following standard arguments for the isotonic (mean) estimator (see, for example, Prakasa Rao (1969, Lemma 4.1)), the estimator $\hat{f}_\lambda(x_0)$ depends (asymptotically) only on those observations having $x_i \in [x_0 - cn^{-1/3}, x_0 + cn^{-1/3}]$, in the sense that

$$\lim_{c\to\infty} \liminf_{n\to\infty} P(\hat{f}_{\lambda,n}(x_0) = \hat{f}_{\lambda,n,c}(x_0)) = 1, \tag{7}$$

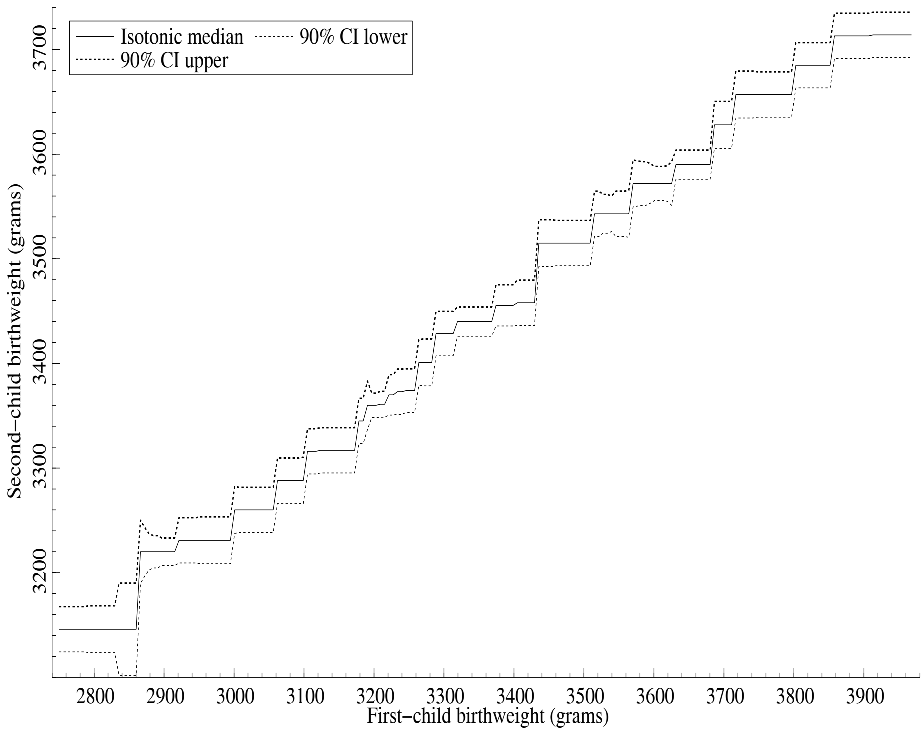

Figure 2. Bootstrap-Based 90% Confidence Intervals for Isotonic Median Regression

where

$$\hat{f}_{\lambda,n,c}(x_0) \equiv \min_{u\le x_0}\max_{v\ge x_0} Q_\lambda(y_i : u \le x_i \le v; u, v \in [x_0 - cn^{-1/3}, x_0 + cn^{-1/3}]). \quad (8)$$

Letting $t_i \equiv n^{1/3}(x_i - x_0)$, we can re-write $\hat{f}_{\lambda,n,c}(x_0)$ as follows:

$$\begin{aligned}\hat{f}_{\lambda,n,c}(x_0) &= \min_{s\le 0}\max_{t\ge 0} Q_\lambda(f_\lambda(x_i) + \epsilon_i : s \le t_i \le t; s, t \in [-c, c]) \\ &= \min_{s\le 0}\max_{t\ge 0} Q_\lambda(f_\lambda(x_0) + f'_\lambda(x_0)(x_i - x_0) \\ &\quad +\epsilon_i : s \le t_i \le t; s, t \in [-c, c]) + o(n^{-1/3}) \\ &= f_\lambda(x_0) + \min_{s\le 0}\max_{t\ge 0} Q_\lambda(f'_\lambda(x_0)n^{-1/3}t_i \\ &\quad +\epsilon_i : s \le t_i \le t; s, t \in [-c, c]) + o(n^{-1/3}).\end{aligned}$$

Boundedness of f''_λ in a neighbourhood of x_0 is used in the second equality above. Define the following quantities:

$$\begin{aligned} z_i &\equiv f'_\lambda(x_0)n^{-1/3}t_i + \epsilon_i \\ \delta_{\lambda,n,c}(s,t) &\equiv Q_\lambda(z_i : s \le t_i \le t; s,t \in [-c,c]) \\ \Delta_{\lambda,n,c} &\equiv \min_{s\le 0}\max_{t\ge 0} \delta_{\lambda,n,c}(s,t) \\ H_n(z,t) &\equiv \begin{cases} n^{-2/3}\sum_{tn^{-1/3}\le t_i\le 0} 1(z_i \le z) & \text{if } -c \le t \le 0 \\ n^{-2/3}\sum_{0\le t_i\le tn^{-1/3}} 1(z_i \le z) & \text{if } 0 \le t \le c. \end{cases} \end{aligned}$$

In this notation, $\hat{f}_{\lambda,n,c}(x_0) - f_\lambda(x_0) = \Delta_{\lambda,n,c} + o(n^{-1/3})$. Since $[H_n(z,t) - H_n(z,t)]/(t-s)$ is the empirical distribution function of z_i for $x_i \in [s,t]$, we have:[5]

$$\begin{aligned} \delta_{\lambda,n,c}(s,t) &= Q_\lambda(z_i : s \le t_i \le t; s,t \in [-c,c]) \\ &= \inf\left\{z : \frac{H_n(z,t) - H_n(z,s)}{t-s} = \lambda; s,t \in [-c,c]\right\} \\ &= \inf\left\{z : H_n(z,t) - H_n(z,s) = \lambda(t-s); s,t \in [-c,c]\right\}. \quad (9) \end{aligned}$$

For a given t, the expectation and variance of $H_n(z,t)$ can be expressed as follows:

$$E[H_n(z,t)] = tG(z|x_0) + \frac{t^2}{2}n^{1/3}\left(\frac{\partial G(z|x_0)}{\partial x} - g(z|x_0)f'_\lambda(x_0)\right) + o(n^{-1/3}) \quad (10)$$

and

$$Var[H_n(z,t)] = n^{-2/3}tG(z|x_0)(1 - G(z|x_0)) + o(n^{-2/3}). \quad (11)$$

The details for deriving (10) and (11) are given for the negative t case below. The positive t case is similar and omitted. For $t \le 0$, we have

$$\begin{aligned} E[H_n(z,t)] &= n^{-2/3}\sum_{tn^{-1/3}\le t_i\le 0} P(z_i \le z) \\ &= n^{-2/3}\sum_{tn^{-1/3}\le t_i\le 0} P(\epsilon_i \le z - f'_\lambda(x_0)n^{-1/3}t_i) \end{aligned}$$

[5]Due to a typographical error, $\delta_{n,c}$ appears without its arguments in the analogous equation of Wang and Huang (2002, p. 284).

$$\begin{aligned}
&= n^{-2/3} \sum_{tn^{-1/3} \le t_i \le 0} G(z - f'_\lambda(x_0) n^{-1/3} t_i | x_i) \\
&= n^{-2/3} \sum_{tn^{-1/3} \le t_i \le 0} \Big[G(z|x_i) - g(z|x_i) f'_\lambda(x_0) n^{-1/3} t_i + o(1) \Big] \\
&= n^{-2/3} \sum_{tn^{-1/3} \le t_i \le 0} \Big[G(z|x_0) + \frac{\partial G(z|x_0)}{\partial x} n^{-1/3} \\
&\quad t_i - g(z|x_0) f'_\lambda(x_0) n^{-1/3} t_i + o(1) \Big] \\
&= tG(z|x_0) + \frac{t^2}{2} n^{1/3} \left(\frac{\partial G(z|x_0)}{\partial x} - g(z|x_0) f'_\lambda(x_0) \right) + o(n^{-1/3}).
\end{aligned}$$

The boundedness (for x_i in a neighbourhood of x_0) of $g'(z|x_i)$, $\frac{\partial^2 G(z|x_i)}{\partial x^2}$, and $\frac{\partial^2 g(z|x_i)}{\partial x^2}$ is utilized in the fourth and fifth equalities of the expectation derivation. The main difference from Wang and Huang (2002, p. 285), aside from conditioning on x_0, is the term involving $\frac{\partial G(z|x_0)}{\partial x}$. Similarly, for $t \le 0$,

$$\begin{aligned}
Var[H_n(z,t)] &= n^{-2/3} \sum_{tn^{-1/3} \le t_i \le 0} P(z_i \le z) \\
&= n^{-2/3} t G(z|x_0)(1 - G(z|x_0)) + o(n^{-2/3}).
\end{aligned}$$

Finally, for $t_1, t_2 \in [-c, c]$ of the same sign,

$$\begin{aligned}
&Cov[H_n(z_1,t_1), H_n(z_2,t_2)] \\
&\quad = n^{-2/3}(t_1 \wedge t_2) G(z_1 \wedge z_2|x_0)(1 - G(z_1 \wedge z_2|x_0)) + o(n^{-2/3}), \quad (12)
\end{aligned}$$

and $Cov[H_n(z_1,t_1), H_n(z_2,t_2)] = 0$ otherwise. From (10), (11), and (12), it follows that as $n \to \infty$, the finite-dimensional distributions of $n^{1/3}[H_n(z,t) - tG(z|x_0)]$ converge to those of $B(z,t) + \left(\frac{\partial G(z|x_0)}{\partial x} - g(z|x_0) f'_\lambda(x_0) \right) t^2/2$ (for $t \in [-c, c]$), where $B(z,t)$ is a zero-mean Gaussian process with

$$Cov[B(z_1,t_1), B(z_2,t_2)] = (t_1 \wedge t_2) G(z_1 \wedge z_2|x_0)(1 - G(z_1 \wedge z_2|x_0)). \quad (13)$$

Based upon (9), we solve $H_n(z,t) - H_n(z,s) = \lambda(t - s)$ for z to derive the limiting distribution of $\Delta_{\lambda,n,c}$. Note that

$$\begin{aligned}
&H_n(z,t) - H_n(z,s) \\
&= G(z|x_0)(t-s) + n^{-1/3}\{[B(z,t) - B(z,s)] \\
&\quad + \frac{t^2-s^2}{2}\left(\frac{\partial G(z|x_0)}{\partial x} - g(z|x_0)f'_\lambda(x_0)\right)\right\}(1+o_p(1)) \\
&= G(0|x_0)(t-s) + g(0|x_0)z(t-s)(1+o(1)) + n^{-1/3} \\
&\quad \left\{[B(0,t) - B(0,s)] + \frac{t^2-s^2}{2}\left(\frac{\partial G(0|x_0)}{\partial x} - g(0|x_0)f'_\lambda(x_0)\right)\right\}(1+o_p(1)) \\
&= G(0|x_0)(t-s) + g(0|x_0)z(t-s)(1+o(1)) \\
&\quad + n^{-1/3}\left\{[B(0,t) - B(0,s)] - \frac{t^2-s^2}{2}g(0|x_0)f'_\lambda(x_0)\right\}(1+o_p(1)) \\
&= \lambda(t-s) + g(0|x_0)z(t-s)(1+o(1)) \\
&\quad + n^{-1/3}\left\{[B(0,t) - B(0,s)] - \frac{t^2-s^2}{2}g(0|x_0)f'_\lambda(x_0)\right\}(1+o_p(1)).
\end{aligned}$$

The third equality above follows from the fact that $G(0|x) = \lambda$ for all x (and therefore is not a function of x). Since $\frac{\partial G(0|x_0)}{\partial x}$ drops out, the resulting expression for $H_n(z,t) - H_n(z,s)$ is similar to that obtained in Wang and Huang (2002, p. 286) aside from the conditioning on x_0. The remainder of the proof, then, follows identically to Wang and Huang (2002), with the generalization

$$Cov[B(0,t), B(0,s)] = (s \wedge t)G(0|x_0)(1 - G(0|x_0)) = (s \wedge t)\lambda(1-\lambda). \quad (14)$$

Acknowledgements. This research was partially supported by a grant from the National Science Foundation (SES-0451660). The author thanks Probal Chaudhuri and an anonymous referee for their helpful suggestions and to Christopher Mrela of the Arizona Department of Health Services for providing the birthweight data used in Section 4.

References

Abrevaya, J. and Huang, J. (2005). On the bootstrap of the maximum score estimator, *Econometrica* **73**, 1175-1204.

Brunk, H.D. (1970). Estimation of isotonic regression, in *Nonparametric techniques in statistical inference*, Cambridge University Press, 177-195.

Casady, R.J. and J.D. Cryer (1976). *Monotone percentile regression*, *Ann. Statist.* **4**, 532-541.

Cryer, J.D., T. Robertson, F. T. Wright and Casady, R.J. (1972). *Monotone median regression*, *Ann. Math. Statist.* **43**, 1459-1469.

Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions, *Probability Theory and Related Fields* **81**, 79-109.

Groeneboom, P. and Wellner, J.A. (2001). Computing Chernoff's distribution, *J. Computational and Graphical Statistics* **10**, 388-400.

Hanson, D.L., G. Pledger and Wright, F.T. (1973). *On consistency in monotonic regression*, *Ann. Statist.* **1**, 401-421.

Koenker, R. and Bassett, G.W.(1978). Regression quantiles, *Econometrica* **46**, 33-50.

Mammen, E. (1991). Estimating a smooth monotone regression function, *Ann. Statist.* **19**, 724-740.

Prakasa Rao, B.L.S. (1969). Estimation of a unimodal density, *Sankyhā Ser. A* **31**, 23-36.

Robertson, T. and Wright, F.T. (1980). Algorithms in order restricted statistical inference and the Cauchy mean value property, *Ann. Statist.* **8**, 645-651.

Wang, Y. and Huang, J. (2002). Limiting distribution for monotone median regression, *J. Statist. Plann. Inference* **107**, 281-287.

Jason Abrevaya
Department of Economics
Purdue University
403 W. State Street
W. Lafayette, IN 47907-2056
E-mail: abrevaya@purdue.edu

Paper received: August 2004; revised March 2005.