DECONVOLUTION PROBLEMS IN TIME SERIES

Young K. Truong
Department of Biostatistics
University of North Carolina
Chapel Hill, N.C. 27599

Jianqing Fan
Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27599

November 7, 1990

Abstract

The problem of estimating the density function and the regression function involving errors-in-variables in time series is considered. Under appropriate conditions, it is shown that the rates obtained in Fan (1991), Fan and Truong (1990) are also achievable in the context of dependent observations. Consequently, the results presented here extend our previous results for cross-sectional data to the longitudinal ones.

[0] *Abbreviated title.* Measurement Errors in Time Series

AMS 1980 subject classification. Primary 62G20. Secondary 62G05, 62J99.

Key words and phrases. Nonparametric regression; kernel estimator; errors in variables; optimal rates of convergence; deconvolution; stationary processes.

1 Introduction

Let (U_i, Y_i) denote a bivariate stationary time series and let $m(\cdot)$ denote the regression function so that $m(u) = E(Y_1|U_1 = u)$. Suppose now that the series $\{U_i\}$ is not observable, instead $X_i = U_i + \varepsilon_i$ is available. Based on the realization $(X_1, Y_1), \ldots, (X_n, Y_n)$, is it possible to estimate the function $m(\cdot)$? The problem just stated is not a well defined one, since there is an identifiability problem about ε_i. This can be resolved by assuming that the error ε_i is independent of U_i and has a known distribution. To simplify our discussion, we further assume that $\{\varepsilon_i\}$ forms an iid sequence of random variables.

There has been a great deal of interest in this so called errors-in-variable regression problem. For a comprehensive approach of parametric $m(\cdot)$, see Stefanski and Carroll (1985, 1987) and Fuller (1987). Recently, Fan and Truong (1990) proposed a nonparametric procedure based on the method of deconvoluting kernel and it is shown, under the iid assumption, this class of estimators possesses various optimal properties. In a subsequent paper, Fan, Truong and Wang (1990) addresses issues on numerical examples and confidence intervals. An important feature of deconvoluting kernel method is that it provides useful diagnostic tools in working with regression problems involving errors-in-variables. Another is its flexibility in dealing with different type of error distributions. Both of these features are regrettably missing in most of the current parametric or semiparametric approaches (Whittemore, 1989).

Fan and Truong (1990), Fan, Truong and Wang (1990) were dealing with the so called cross-sectional data, the current approach continues the same line of research by extending them to time series or longitudinal data. Under appropriate regularity and mixing conditions, the rates of convergence of the deconvoluting kernel density estimators are shown to be compatible to results of Carroll and Hall (1988), Fan (1991) and Zhang (1990) in the iid case. They are also compatible with the result obtained by Masry (1991) in the stationary case. The rates of convergence of the deconvoluting kernel regression estimators are identical to the results of Fan and Truong (1990) for the iid case. Since iid is a special case of the stationary sequence of random variables, the rates obtained here are therefore optimal by virtue of Fan (1991) and Fan and Truong (1990).

2 Deconvoluting Kernel Estimators

2.1 Density Function Estimation

Let $\{X_i\}$ and $\{U_i\}$ denote a stationary time series so that $X_i = U_i + \varepsilon_i$, where ε_i are iid with mean zero and variance σ^2. Denote the density function of U_1 by $f(\cdot)$. Given a sequence of observations $X_1, \ldots, X_n$, consider the following deconvoluting kernel estimator of $f(\cdot)$:

$$\hat{f}_n(u) = (nh_n)^{-1} \sum K_n\left(\frac{u - X_i}{h_n}\right), \tag{2.1}$$

where $\{h_n\}$ is a sequence of positive numbers converging to zero, and

$$K_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx) \frac{\phi_K(t)}{\phi_\varepsilon(t/h_n)}\, dt \tag{2.2}$$

with

$$\phi_K(t) = \int e^{itx} K(x)\, dx, \qquad \phi_\varepsilon(t) = \int e^{itx} f_\varepsilon(x)\, dx. \tag{2.3}$$

That is, $\phi_K(\cdot)$ is the Fourier transform of the kernel function $K(\cdot)$ and $\phi_\varepsilon(\cdot)$ is the characteristic function of the error variable ε. The above estimator has been considered extensively in the literature. See, for example, Carroll and Hall (1988), Fan (1991), Masry (1991), Stefanski and Carroll (1990) and Zhang (1990).

2.2 Regression Function Estimation

Given the bivariate series $(X_1, Y_1), \ldots, (X_n, Y_n)$, the regression function $m(\cdot)$ can be estimated by the following deconvoluting kernel estimator:

$$\hat{m}_n(u) = \sum_j K_n\left(\frac{u - X_j}{h_n}\right) Y_j \Big/ \sum_i K_n\left(\frac{u - X_i}{h_n}\right) \tag{2.4}$$

where $K_n(\cdot)$ is given in (2.2). See Fan and Truong (1990) and Fan, Truong and Wang (1990) for optimal results and numerical examples in the cross-sectional situation.

3 Performance of kernel estimators

The sampling behaviors of the kernel estimators (2.1) and (2.4) considered in the previous section will be treated here. The rates of convergence of these estimators depend on the smoothness of error distributions, which can be classified into:

- Super smooth of order β: if the characteristic function of the error distribution $\phi_\varepsilon(\cdot)$ satisfies

$$d_0|t|^{\beta_0}\exp(-|t|^\beta/\gamma) \leq |\phi_\varepsilon(t)| \leq d_1|t|^{\beta_1}\exp(-|t|^\beta/\gamma) \quad \text{as } t \to \infty, \tag{3.1}$$

  where d_0, d_1, β, γ are positive constants and β_0, β_1 are constants.

- Ordinary smooth of order β: if the characteristic function of the error distribution $\phi_\varepsilon(\cdot)$ satisfies

$$d_0|t|^{-\beta} \leq |\phi_\varepsilon(t)| \leq d_1|t|^{-\beta} \quad \text{as } t \to \infty, \tag{3.2}$$

  for positive constants d_0, d_1, β.

For example,

$$\text{Super smooth distributions :} \quad \begin{cases} N(0,1) & \text{with } \beta = 2, \\ \frac{1}{\pi}\frac{1}{1+x^2} \quad \text{Cauchy } (0,1) & \text{with } \beta = 1. \end{cases}$$

$$\text{Ordinary smooth distributions :} \quad \begin{cases} \frac{\alpha^p}{\Gamma(p)}x^{p-1}e^{-\alpha x} \quad \text{(Gamma)} & \text{with } \beta = p, \\ \frac{1}{2}e^{-|x|} \quad \text{(double exponential)} & \text{with } \beta = 2\ . \end{cases}$$

The rates of convergence depend on β, the order of smoothness of the error distribution. They also depend on the smoothness of the regression function $m(\cdot)$ and regularity conditions on the marginal distribution which are given as follows.

Condition 1. *Let $a < b$.*

a. *The marginal density $f(\cdot)$ has a bounded kth derivative on the interval (a, b).*

b. *The characteristic function of U_1 is absolutely integrable.*

c. *The characteristic function of the error distribution* $\phi_\varepsilon(\cdot)$ *does not vanish.*

d. *The marginal density* $f(\cdot)$ *of the unobserved* U_1 *is bounded away from zero on the interval* $[a, b]$.

e. *The regression function* $m(\cdot)$ *has a bounded* k*th derivative.*

f. *For some* $\theta > 2$, $\sup_{a \le u \le b} E(|Y_1|^\theta | U_1 = u) < M$.

The rates depend on the following condition of the kernel function:

Condition 2. *The kernel* $K(\cdot)$ *is a* k*th order kernel. Namely,*

$$\int_{-\infty}^{\infty} K(x)dx = 1, \int_{-\infty}^{\infty} x^k K(x)\,dx \neq 0,$$

$$\int_{-\infty}^{\infty} x^j K(x)\,dx = 0, \quad \text{for } j = 1, \ldots, k-1.$$

Let $\mathcal{F}_t$ and $\mathcal{F}^t$ denote the σ-fields generated respectively by (U_i, Y_i), $-\infty < i \le t$, and (U_i, Y_i), $t \le i < \infty$. Given a positive integer k, set [Rosenblatt (1956)]

$$\alpha(k) = \sup\{|\, P(A \cap B) - P(A)P(B)\,| : A \in \mathcal{F}_t \quad \text{and} \quad B \in \mathcal{F}^{t+k}\}.$$

The stationary sequence is said to be α-mixing or strongly mixing if $\alpha(k) \to 0$ as $k \to \infty$.

Condition 3. *The stationary sequence* (U_i, Y_i) *is* α *mixing with*

$$\sum_{j>N} \alpha^{1-2/\theta}(j) = O(N^{-1}) \quad \text{as} \quad N \to \infty.$$

It follows from the assumption on $\{\varepsilon_i\}$ that (X_i, Y_i) is also α-mixing.

3.1 Rates of Convergence in the Density Case

The deconvoluting kernel density estimator will be considered in this section. We start with super smooth error distributions:

Theorem 1. *Suppose Conditions* 1(*a*)–(*c*), 2 *and* 3 *hold and that* [*see* (3.1)]

$$|\phi_\varepsilon(t)| \geq d_0 |t|^{\beta_0} \exp(-|t|^\beta/\gamma).$$

Assume that $\phi_K(t)$ *has a bounded support on* $|t| \leq M_0$. *Then, for bandwidth* $h_n = c(\log n)^{-1/\beta}$ *with* $c > M_0(2/\gamma)^{1/\beta}$,

$$|\hat{f}_n(u) - f(u)| = O_p((\log n)^{-k/\beta}), \quad u \in [a, b].$$

To compute the variance of the deconvoluting density estimator for the ordinary smooth error distributions, we need the following condition on the tail of $\phi_\varepsilon(t)$:

$$t^\beta \phi_\varepsilon(t) \to c_0, \quad t^{\beta+1}\phi'_\varepsilon(t) \to c_1 \quad \text{and} \quad |t^{\beta+2}\phi''_\varepsilon(t)| = o(1) \quad \text{as} \quad t \to \infty \tag{3.3}$$

for some constants $c_0, c_1 \neq 0$. Note that (3.3) is a special case of (3.2).

Theorem 2. *Suppose Conditions* 1(*a*)–(*c*), 2 *and* 3 *hold and that*

$$\int_{-\infty}^{\infty} \left|t^{\beta+1}\right| (|\phi_K(t)| + |\phi'_K(t)|)\, dt < \infty, \int_{-\infty}^{\infty} \left|t^{\beta+1}\phi_K(t)\right|^2 dt < \infty.$$

Then, under the ordinary smooth error distribution (3.3) *and* $h_n = dn^{-1/(2k+2\beta+1)}$ *with* $d > 0$,

$$|\hat{f}_n(u) - f(u)| = O_p\left(n^{-k/(2k+2\beta+1)}\right), \quad u \in [a, b].$$

Remarks.

- Under the iid assumption, Fan (1991) showed that the above rates are optimal in the minimax sense. Hence, they are also optimal in the stationary setup, since iid random variables are special cases of strong mixing stationary process. For more details on rate optimality in the deconvoluting density estimation for the iid case, see Fan (1991).
- In Theorem 1, only the first half of (3.1) is needed for establishing the so-called achievable rates. The other half is used for showing the lower bound (rates optimality).

3.2 Rates of Convergence in the Regression Case

The rates of convergence for the deconvoluting regression function estimator will be discussed in this section. We begin with the super smooth error distributions:

Theorem 3. *Suppose Conditions* 1–3 *hold and that*

$$|\phi_\varepsilon(t)| \geq d_0|t|^{\beta_0}\exp(-|t|^\beta/\gamma).$$

Assume that $\phi_K(t)$ *has a bounded support on* $|t| \leq M_0$. *Then, for bandwidth* $h_n = c(\log n)^{-1/\beta}$ *with* $c > M_0(2/\gamma)^{1/\beta}$,

$$|\hat{m}_n(u) - m(u)| = O_p((\log n)^{-k/\beta}), \quad u \in [a,b].$$

To compute the variance of the deconvoluting regression estimator for ordinary smooth error distributions, the condition similar to density estimation given in Section 3.1 is required. See (3.3).

Theorem 4. *Suppose Conditions* 1–3 *hold and that*

$$\int_{-\infty}^{\infty}\left|t^{\beta+1}\right|(|\phi_K(t)| + |\phi_K'(t)|)\,dt < \infty, \int_{-\infty}^{\infty}\left|t^{\beta+1}\phi_K(t)\right|^2 dt < \infty.$$

Then, under the ordinary smooth error distribution (3.3) *and* $h_n = dn^{-1/(2k+2\beta+1)}$ *with* $d > 0$,

$$|\hat{m}_n(u) - m(u)| = O_p\left(n^{-k/(2k+2\beta+1)}\right), \quad u \in [a,b].$$

Remark. Under the iid assumption, Fan and Truong (1990) showed that the rates given in Theorems 3 and 4 can not be improved. Hence, these rates are optimal for the stationary sequences. More detailed discussion on rate optimality of the deconvoluting regression estimators for the iid case can be found in Fan and Truong (1990).

4 Proofs

Lemma 4.1. *Under the conditions given in Theorems 1 or 2, then*

$$E\hat{f}_n(u) - f(u) = \frac{1}{h_n}\int_{-\infty}^{\infty}[f(v) - f(u)]K(\frac{u-v}{h_n})\,dv$$

$$= f(u)b_{1,k}(u)h_n^k(1+o(1)),$$

where

$$b_{1,k}(u) \equiv (-1)^{k+1} \cdot \frac{f^{(k)}(u)}{k!} f^{-1}(u) \int_{-\infty}^{\infty} v^k K(v)\, dv.$$

Proof. This follows from Lemma 6.2 of Fan and Truong (1990) and Taylor's expansion.

The following lemma computes the variances of $\hat{f}_n(u)$ for super smooth errors and ordinary smooth errors, respectively.

Lemma 4.2. *Under the conditions given in Theorem 1,*

$$\text{Var}\left(\hat{f}_n(u)\right) = o(h_n^{2k}).$$

Under the conditions given in Theorem 2,

$$\text{Var}\left(\hat{f}_n(u)\right) = O\left((nh_n^{2\beta+1})^{-1}\right).$$

Proof. Set $K_{n,j} = K_n((u - X_j)/h_n)$. Suppose first that the error is super smooth. Then

$$\begin{aligned}
&\text{Var}\left(\hat{f}_n(u)\right) \\
&= \frac{1}{n^2h_n^2}\text{Var}\left(\sum K_{n,j}\right) \\
&\leq \frac{1}{nh_n^2} E\left|K_n\left(\frac{u-X}{h_n}\right)\right|^2 + \frac{1}{nh_n^2}\sum_j \text{Cov}(K_{n,1}, K_{n,1+j}) \\
&\leq \frac{1}{nh_n^2} E\left|K_n\left(\frac{u-X}{h_n}\right)\right|^2 + \frac{1}{nh_n^2}\sum_j \|K_{n,1}\|_\infty^2 \alpha(|j|).
\end{aligned}$$

The last inequality follows from Corollary A.1 of Hall and Heyde (1980). Now according to Fan and Truong (1990),

$$\|K_{n,i}\|_\infty \leq \sup_x |K_n(x)| = O\left(\exp(|M_0/h_n|^\beta/\gamma)/h_n\right).$$

By Condition 3 and $h_n = c(\log n)^{-1/\beta}$,

$$\text{Var}(\hat{f}_n(u)) = O\left(\frac{1}{nh_n^4}\exp(2|M_0/h_n|^\beta/\gamma)\right) = o(h_n^{2k}).$$

Suppose now that the error distribution is ordinary smooth. Invoke the integration by parts twice and by an analogous argument in Lemma 6.4 of Fan and Truong (1990), there exists a constant C such that

$$|h_n^\beta K_n(v)| \le \frac{C}{1+|v|^2}. \tag{4.4}$$

Let $f_1(\cdot)$ and $f_{1,j}(\cdot,\cdot)$ denote the density functions of X_1 and (X_1, X_j), respectively. Then

$$EK_{n,1} = EK_n\left(\frac{u-X_1}{h_n}\right) = h_n \int K_n(v) f_1(u - vh_n)\, dv = O(h_n^{1-\beta}), \tag{4.5}$$

$$EK_{n,1}K_{n,j} = h_n^2 \int\int K_n(v_1)K_n(v_2) f_{1,j}(u - v_1 h_n, u - v_2 h_n)\, dv_1 dv_2 = O(h_n^{2-2\beta}). \tag{4.6}$$

Similarly,

$$E|K_{n,1}|^\theta = O(h_n^{1-\beta\theta}), \quad \theta > 1. \tag{4.7}$$

Set

$$\frac{1}{nh_n^2} \sum_j \mathrm{Cov}(K_{n,1}, K_{n,1+j}) = \Sigma_1 + \Sigma_2,$$

where

$$\Sigma_1 = \frac{1}{nh_n^2} \sum_{j \le N} \mathrm{Cov}(K_{n,1}, K_{n,1+j})$$

and

$$\Sigma_2 = \frac{1}{nh_n^2} \sum_{j > N} \mathrm{Cov}(K_{n,1}, K_{n,1+j}).$$

Choose $N \sim h_n^{-1}$. It follows from (4.6) that

$$\Sigma_1 = O\left(\frac{1}{nh_n^{2\beta+1}}\right).$$

By (4.7) and Corollary A.2 of Hall and Heyde (1980),

$$\begin{aligned}
\Sigma_2 &= \frac{1}{nh_n^2} \sum_{j>N} \left(E|K_{n,1}|^\theta\right)^{2/\theta} \alpha^{1-2/\theta}(j) \\
&= O\left(\frac{1}{nh_n^{2\beta+1}}\right) h_n^{2/\theta-1} \sum_{j>N} \alpha^{1-2/\theta}(j) \\
&= o\left(\frac{1}{nh_n^{2\beta+1}}\right).
\end{aligned}$$

This completes the proof of Lemma 2.

For the following discussions, set $A_n(u) = (nh_n)^{-1} \sum K_n((u - X_j)/h_n)(Y_j - m(u))$.

Lemma 4.3. *Given the conditions in Theorem 3 or 4, then*

$$\begin{aligned} EA_n(u) &= \frac{1}{h_n}\int_{-\infty}^{\infty}[m(v)-m(u)]K(\frac{u-v}{h_n})f(v)\,dv \\ &= f(u)b_{2,k}(u)h_n^k(1+o(1)), \end{aligned}$$

where

$$b_{2,k}(u) \equiv (-1)^{k+1}\left[\frac{(m(u)f(u))^{(k)}}{k!} - m(u)\frac{f^{(k)}(u)}{k!}\right] f^{-1}(u)\int_{-\infty}^{\infty} v^k K(v)\,dv.$$

Proof. This follows from Lemma 6.2 of Fan and Truong (1990) and Taylor's expansion.

Lemma 4.4. *Under the conditions given in Theorem 3,*

$$\text{Var}(A_n(u)) = o(h_n^{2k}).$$

While under the conditions given in Theorems 4,

$$\text{Var}(A_n(u)) = O\left((nh_n^{2\beta+1})^{-1}\right).$$

Proof. Set $K_{n,j} = K_n((u-X_j)/h_n)$. For the super smooth error distribution,

$$\begin{aligned} \text{Var}(A_n(u)) &= \frac{1}{n^2h_n^2}\text{Var}\left(\sum K_{n,j}(Y_j - m(u))\right) \\ &\leq \frac{1}{nh_n^2}E\left|K_n\left(\frac{u-X}{h_n}\right)(Y-m(u))\right|^2 \\ &\quad + \frac{1}{nh_n^2}\sum_j \text{Cov}(K_{n,1}(Y_1-m(u)), K_{n,1+j}(Y_{1+j}-m(u))). \end{aligned}$$

By Corollary A.2 of Hall and Heyde (1980),

$$\begin{aligned} &\text{Cov}(K_{n,1}(Y_1-m(u)), K_{n,1+j}(Y_{1+j}-m(u))) \\ &\leq 8\left(E|K_{n,1}|^\theta|Y_1-m(u)|^\theta\right)^{2/\theta}\alpha^{1-2\theta}(|j|) \\ &\leq 8\|K_{n,1}\|_\infty^2\left(E|Y_1-m(u)|^\theta\right)^{2/\theta}\alpha^{1-2\theta}(|j|). \end{aligned}$$

It follows from Condition 1(f),

$$\|K_{n,1}\|_\infty = \sup_x|K_n(x)| = O\left(\exp(|M_0/h_n|^\beta/\gamma)/h_n\right) \quad \text{and} \quad h_n = c(\log n)^{-1/\beta}$$

that

$$\operatorname{Var}(A_n(u)) = O\left(\frac{1}{nh_n^3}\exp(2|M_0/h_n|^\beta/\gamma)\right) = o(h_n^{2k}).$$

For the ordinary smooth error distributions, set $Z_i = Y_i - m(u)$ and

$$\frac{1}{nh_n^2}\sum_j \operatorname{Cov}(K_{n,1}Z_1, K_{n,1+j}Z_{1+j}) = \Sigma_1 + \Sigma_2,$$

where

$$\Sigma_1 = \frac{1}{nh_n^2}\sum_{j\le N} \operatorname{Cov}(K_{n,1}Z_1, K_{n,1+j}Z_{1+j})$$

and

$$\Sigma_2 = \frac{1}{nh_n^2}\sum_{j> N} \operatorname{Cov}(K_{n,1}Z_1, K_{n,1+j}Z_{1+j}).$$

Applying Hölder inequality twice,

$$\begin{aligned}
&\operatorname{Cov}(K_{n,1}Z_1, K_{n,1+j}Z_{1+j}) \\
&= E\left(K_{n,1}Z_1^\theta\right)^{2/\theta}\cdot E\left(K_{n,1}K_{n,1+j}\right)^{1-2/\theta}\cdot O(h_n^{-2\beta/\theta}) + O(h_n^{2k})
\end{aligned}$$

By Condition 1(f) and (4.4)–(4.7),

$$\operatorname{Cov}(K_{n,1}Z_1, K_{n,1+j}Z_{1+j}) = O(h_n^{2-2/\theta-2\beta}).$$

Choose $N \sim h_n^{-1+2/\theta}$. Then

$$\Sigma_1 = O\left(\frac{1}{nh_n^{2\beta+1}}\right).$$

By Conditions 1(f), 3 and Corollary A.2 of Hall and Heyde (1980),

$$\begin{aligned}
\Sigma_2 &\le \frac{8}{nh_n^2}\sum_{j>N}\left(E|K_{n,1}Z_1|^\theta\right)^{2/\theta}\alpha^{1-2/\theta}(j) \\
&\le \frac{8M}{nh_n^2}\sum_{j>N}\left(E|K_{n,1}|^\theta\right)^{2/\theta}\alpha^{1-2/\theta}(j) \\
&= O\left(\frac{h_n^{2/\theta}}{nh_n^{2\beta+2}}\right)\sum_{j>N}\alpha^{1-2/\theta}(j) \\
&= o\left(\frac{1}{nh_n^{2\beta+1}}\right).
\end{aligned}$$

The last two equalities follow from (4.7) and Condition 3, respectively. This completes the proof of Lemma 4.

By Lemmas 4.1 and 4.2, it follows from the Chebyshev's inequality and the usual variance-bias decomposition that the conclusions of Theorems 1 and 2 hold.

To prove Theorem 3, note that

$$\hat{m}_n(u) - m(u) = \frac{A_n(u)}{\hat{f}_n(u)}.$$

Given $u \in [1, b]$, set $\epsilon = f(u)/2$. Then Condition 1(d) implies $\epsilon > 0$. By Theorem 1,

$$\lim_n P(\Omega_n) = 1, \tag{4.8}$$

where $\Omega_n = \{|\hat{f}_n(u) - f(u)| \leq \epsilon\}$. In particular, for n sufficiently large, $\hat{f}_n(u) > \epsilon > 0$ on Ω_n. Hence, by Chebyshev's inequality, Lemmas 4.3 and 4.4, on Ω_n,

$$|\hat{m}_n(u) - m(u)| \leq \frac{|A_n(u)|}{\epsilon} = O_p((\log n)^{-k/\beta}). \tag{4.9}$$

The conclusion of Theorem 3 follows from (4.8) and (4.9).

The argument for Theorem 4 follows similarly by virtue of Theorem 2, Lemma 4.3 and the second part of Lemma 4.4.

References

[1] Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvoluting a density. *J. Amer. Statist. Assoc.*, **83**, 1184–1186.

[2] Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist.* To appear.

[3] Fan, J. and Truong, Y. (1990). Nonparametric regression with errors-in-variables. *Institute of Statistics Mimeo Series #2023*, Univ. of North Carolina, Chapel Hill.

[4] Fan, J., Truong, Y. and Wang, Y. (1990). Measurement errors regression: A nonparametric approach. *Institute of Statistics Mimeo Series #2036*, Univ. of North Carolina, Chapel Hill.

[5] Fuller, W. A. (1987). *Measurement error models.* Wiley, New York.

[6] Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its applications.* Academic Press, New York.

[7] Masry, E. and Rice, J. A. (1991). Gaussian deconvolution via differentiation. To appear in *Can. J. Statist.*.

[8] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.* **42**, 43–47.

[9] Stefanski, L. A. and Carroll, R. J. (1985). Covariate measurement error in logistic regression. *Ann. Statist.*, **13**, 1335–1351.

[10] Stefanski, L. A. and Carroll, R. J. (1987). Conditional scores and optimal scores for generalized linear measurement-error models. *Biometrika*, **74**, 703–716.

[11] Stefanski, L. A. and Carroll, R. J. (1990). Deconvoluting kernel density estimators. *Statistics*, **21**, 169-184.

[12] Whittemore, A. S. (1989). Errors-in-variables regression using Stein estimates. *American Statistician.* **43**, 226–228.

[13] Zhang, C. H. (1990). Fourier methods for estimating mixing densities and distributions. *Ann. Statist.*, **18**, 806–830.