DISCUSSION PAPER
0530

STRONG UNIFORM CONSISTENCY RESULTS OF THE WEIGHTED AVERAGE OF CONDITIONAL ARTIFICIAL DATA POINTS
C. HEUCHENNE

http://www.stat.ucl.ac.be
Strong uniform consistency results of the weighted average of conditional artificial data points

Cédric Heuchenne¹

Institut de Statistique

Université catholique de Louvain

August 17, 2005

Abstract

In this paper, we study strong uniform consistency of a weighted average of artificial data points. This is especially useful when information is incomplete (censored data, missing data . . .). In this case, reconstruction of the information is often achieved nonparametrically by using a local preservation of mean criterion for which the corresponding mean is estimated by a weighted average of new data points. This way of doing enlarges beyond incomplete data context and applies to the estimation of the conditional mean of specific functions of complete data points. As a consequence, we establish the strong uniform consistency of the Nadaraya-Watson (1964) estimator for general transformations of the data points. This result generalizes the one of Härdle, Janssen and Serfling (1988). In addition, the strong uniform consistency of a modulus of continuity will be obtained for this estimator. Applications of those two results are detailed for some popular estimators.

KEY WORDS: Kernel estimation; Nonparametric regression; Right censoring.

¹ This research was supported by ‘Projet d’Actions de Recherche Concertées’, No. 98/03–217, and by the IAP research network nr. P5/24 of the Belgian government.
1 Introduction

In many regression contexts where the data are incomplete, one has to reconstruct missing information by using other data points. In particular, if \( Z \) denotes a data point, \( X \) the covariate and \( \Delta \) is a binary variable equal to 1 if the data point \( Z \) is complete (in this case \( Z = Y \); the true data point) and 0 if it is incomplete, a natural way to reconstruct a function \( t(Y|x) \) at \( X = x \) and for \( t \in I \) is to take
\[
\varphi_t(Z|x) = E[\varphi_t(Y|x)|Y > Z, x](1-\Delta).
\]
In the case of missing data, \( E[\varphi_t(Y|x)|Y > Z, x] = E[\varphi_t(Y|x)|x] \). In censored regression, this scheme with \( \varphi_t(Y|x) = Y \) has been used by Buckley and James (1979), Koul, Susarla and Van Ryzin (1981), Leurgans (1987), Fan and Gijbels (1994), Heuchenne and Van Keilegom (2004) among others. In estimation with missing data, this kind of new data points has been proposed by e.g. Cheng (1994), Chu and Cheng (1995) and Cheng and Chu (1996).

As explained in Heuchenne and Van Keilegom (2005) for nonparametric estimation with censored data, \( \varphi_t(Y|x) \) can be any function of \( x, t \) and \( Y \), e.g., \( Y, Y^2 \) or \( I(Y > Z) \); for fixed \( t \in I \), if the objective is to estimate \( E[Y|x], E[Y^2|x] \) or \( E[I(Y \leq t)|x] = P(Y \leq t|x) \) respectively. Therefore, there is a need to construct a general asymptotic theory for a nonparametric estimator of \( E[\varphi_t(Y|x)|x] \) in the complete (incomplete) data case.

More precisely, let \( \{\Gamma_t, t \in I\} \) be a family of real valued measurable functions on \( R \) and suppose we want to estimate
\[
E[\Gamma_t(Z, \Delta|x)] = \sum_{\delta=0,1} \int \Gamma_t(z, \delta|x)dH_\delta(z|x),
\]
where \( I \) is a possibly infinite or degenerate interval in \( R \), \( x \in R_X \), a compact interval in \( R \) and \( H_\delta(y|x) = P(Z \leq y, \Delta = \delta|x) \) (\( \delta = 0, 1 \)). A natural nonparametric estimator for this conditional mean is given by
\[
\frac{\sum_{i=1}^n K(x-Z_i)\Gamma_t(Z_i, \Delta_i|x)}{\sum_{i=1}^n K(x-Z_i)}.
\]
In the case \( \Gamma_t(Z, \Delta|x) = Z \), this estimator reduces to the usual Nadaraya-Watson (1964) estimator and in the case \( \Gamma_t(Z, \Delta|x) = I(Z \leq t) \), we obtain the Stone (1977) estimator.

The objective of Section 3 is to provide the almost sure convergence of (1.2) uniformly in \( x, t \) with the rate \( (na_n)^{-1/2}(\log n)^{1/2} \). Now, suppose \( s, t \in I \) with \( |t-s| \leq d_n \). In Section 4, we aim to obtain the almost sure convergence of the modulus of continuity based on
uniformly in \(x, s, t, |t - s| \leq d_n\), with the rate \((na_n)^{-1/2}(\log n)^{1/2}d_n^{1/2}\). The useful purpose of these results is illustrated for some typical examples in Section 2.

2 Examples and Assumptions

Example 2.1 (Nonparametric estimation of conditional location and scale functions for complete data)

Suppose \(Y_1, \ldots, Y_n\) are \(n\) i.i.d. random variables corresponding to \(X_1, \ldots, X_n\), \(n\) i.i.d. covariates with distribution \(F_X(x) = P(X_1 \leq x)\). Let \(F(t|x) = P(Y_1 \leq t|X_1 = x)\) be the conditional distribution of the response given the covariate. Usual location and scale estimators are given by

\[
\hat{m}_{ST}(x) = \int_0^1 \hat{F}^{-1}(s|x)L(s) \, ds,
\hat{\sigma}_{ST}^2(x) = \int_0^1 \hat{F}^{-1}(s|x)^2L(s) \, ds - \hat{m}_{ST}^2(x),
\]

where \(\hat{F}(\cdot|x)\) is the Stone (1977) estimator and \(L(s)\) is a given score function satisfying \(\int_0^1 L(s)ds = 1\). If the objective is to estimate

\[
\int_0^1 F^{-1}(s|x)L(s) \, ds
\]

and

\[
\int_0^1 F^{-1}(s|x)^2L(s) \, ds,
\]

it is clear that \(\Gamma_1(Y|x) = YL(F(Y|x))\) for (2.2) and \(\Gamma_2(Y|x) = Y^2L(F(Y|x))\) for (2.3). Since the data points \(\Gamma_i(Y|x)\) depend themselves on \(F(Y|x)\), they are estimated by \(YL(\hat{F}(Y|x))\) and \(Y^2L(\hat{F}(Y|x))\) so that the classical Nadaraya-Watson estimator based on those data points corresponds to (2.1).

Note that when \(L(s) = I(0 \leq s \leq 1)\), \(\hat{m}_{ST}(x)\) and \(\hat{\sigma}_{ST}^2(x)\) reduce to estimators of the conditional mean and variance. Theorem 3.3 of the next section thus enables to prove at the same time the strong uniform consistency of estimators of any location and scale functions defined by the score function \(L\). This is achieved in two steps: first, an application of Theorem 3.3 for data points \(I(Y_i \leq t)\) \((i = 1, \ldots, n)\) in order to delete
the Stone estimators in the expressions $YL(\hat{F}(Y|x))$ and $Y^2L(\hat{F}(Y|x))$, and second, an application of the same theorem on the functions $\Gamma_1(Y|x)$ and $\Gamma_2(Y|x)$.

**Example 2.2 (Nonparametric estimation of conditional location and scale functions for censored data)**

Now, suppose $Y_1, \ldots, Y_n$ are possibly right censored by $C_1, \ldots, C_n$ i.i.d. random variables with distribution function $G(t|x) = P(C_1 \leq t|X = x)$. The observed random variable for the covariate $X_i$ is therefore the pair $(Z_i, \Delta_i), i = 1, \ldots, n$, with $Z_i = Y_i \wedge C_i$ and $\Delta_i = I(Y_i \leq C_i)$. We will now assume independence of $Y_i$ and $C_i$ conditionally on $X_i$. Usual location and scale estimators are given by

$$\hat{m}_B(x) = \int_{-\infty}^{\hat{T}} yL(\hat{F}(y|x)) \, d\hat{F}(y|x)$$

and

$$\hat{\sigma}_B^2(x) = \int_{-\infty}^{\hat{T}} y^2L(\hat{F}(y|x)) \, d\hat{F}(y|x) - \hat{m}_B^2(x),$$

where $\hat{F}(\cdot|x)$ is the Beran (1981) estimator given by

$$\hat{F}(t|x) = 1 - \prod_{Z_i \leq t, \Delta_i = 1} \left\{ 1 - \frac{W_i(x, a_n)}{\sum_{j=1}^{n} I(Z_j \geq Z_i)W_j(x, a_n)} \right\} I(t < Z_{(n)}),$$

with $W_i(x, a_n) = \frac{K\left(\frac{x-X_i}{a_n}\right)}{\sum_{j=1}^{n} K\left(\frac{x-X_j}{a_n}\right)}$.

are the Nadaraya-Watson weights and $L(s)$ is a given score function satisfying $\int_{0}^{1} L(s) \, ds = 1$. In order to avoid consistency problems in the right tails of the Beran estimators, $\hat{T}$ is chosen smaller than $\inf_{x} \tau_{H(\cdot|x)}$, where $H(y|x) = P(Z \leq y|x)$ and $\tau_{F(\cdot)} = \inf\{t : F(t) = 1\}$ for some $F$. Seeing that the objective is to estimate $E[YI(Y \leq \hat{T})L(F(Y|x))|x]$ and $E[Y^2I(Y \leq \hat{T})L(F(Y|x))|x]$ with an estimator of the Nadaraya-Watson type, we rewrite (2.4) and (2.5) as

$$\hat{m}_B(x) = \sum_{i=1}^{n} W_i(x, a_n)\hat{\Gamma}_{13}(Z_i, \Delta_i|x),$$

and

$$\hat{\sigma}_B^2(x) = \sum_{i=1}^{n} W_i(x, a_n)\hat{\Gamma}_{14}(Z_i, \Delta_i|x) - \hat{m}_B^2(x),$$

where $\Gamma_{13}(Z_i, \Delta_i|x)$ and $\Gamma_{14}(Z_i, \Delta_i|x)$.
where

\[ \hat{\Gamma}_{i3}(Z_i, \Delta_i|x) = Z_i I(Z_i \leq \tilde{T}) L(\hat{F}(Z_i|x)) \Delta_i + \frac{\int_{Z_i < \tilde{T}} y L(\hat{F}(y|x)) d\hat{F}(y|x)}{1 - \hat{F}(Z_i \wedge \tilde{T}|x)}(1 - \Delta_i), \]

and

\[ \hat{\Gamma}_{i4}(Z_i, \Delta_i|x) = Z_i^2 I(Z_i \leq \tilde{T}) L(\hat{F}(Z_i|x)) \Delta_i + \frac{\int_{Z_i < \tilde{T}} y^2 L(\hat{F}(y|x)) d\hat{F}(y|x)}{1 - \hat{F}(Z_i \wedge \tilde{T}|x)}(1 - \Delta_i). \]

\( \hat{\Gamma}_{i3}(Z, \Delta|x) \) and \( \hat{\Gamma}_{i4}(Z, \Delta|x) \) actually estimate

\[ \Gamma_{i3}(Z, \Delta|x) = Z I(Z \leq \tilde{T}) L(F(Z|x)) \Delta + \frac{\int_{Z < \tilde{T}} y L(F(y|x)) dF(y|x)}{1 - F(Z \wedge \tilde{T}|x)}(1 - \Delta), \]

and

\[ \Gamma_{i4}(Z, \Delta|x) = Z^2 I(Z \leq \tilde{T}) L(F(Z|x)) \Delta + \frac{\int_{Z < \tilde{T}} y^2 L(F(y|x)) dF(y|x)}{1 - F(Z \wedge \tilde{T}|x)}(1 - \Delta), \]

respectively. It is easy to check that

\[ E[\Gamma_{i3}(Z, \Delta|x)|x] = E[Y I(Y \leq \tilde{T}) L(F(Y|x))|x], \]

and

\[ E[\Gamma_{i4}(Z, \Delta|x)|x] = E[Y^2 I(Y \leq \tilde{T}) L(F(Y|x))|x]. \]

As for the complete data case, Theorem 3.3 enables to prove the strong uniform consistency of estimators of any location and scale functions (truncated by \( \tilde{T} \)) defined by the score function \( L \). Note that in order to use Theorem 3.3 with the functions \( \hat{\Gamma}_{i3}(Z, \Delta|x) \) and \( \hat{\Gamma}_{i4}(Z, \Delta|x) \), we first need to delete the Beran estimators that appear in \( \hat{\Gamma}_{i3}(Z, \Delta|x) \) and \( \hat{\Gamma}_{i4}(Z, \Delta|x) \). This can be done by using Proposition 4.3 of Van Keilegom and Akritas (1999).

**Example 2.3 (Estimation of a conditional distribution function under the heteroscedastic model)**

Now, suppose in the previous example that we want to estimate the conditional distribution function of the response given the covariate when the relation \( Y = m^0(X) + \sigma^0(X) \varepsilon^0 \) with \( \varepsilon^0 \) independent of \( X \) is assumed. The corresponding preservation of means criterion is: construct new indicators for which the conditional mean equals the asked conditional
distribution function and which use the above heteroscedastic model. More precisely, this estimator is a weighted sum of data points $\hat{\Gamma}_{i\xi}(Z_i, \Delta_i|x) \sim 1, \ldots, n$, that approximate

$$\Gamma_{i\xi}(Z_i, \Delta_i|x) = I(Z_i \leq t)\Delta_i + \frac{F^0_{\xi}(T_i^x) - F^0_{\xi}(E^0_{i\xi})}{1 - F^0_{\xi}(E^0_{i\xi})} (1 - \Delta_i), \quad (2.10)$$

where $E^0_{i\xi} = \frac{Z_i \wedge T_i \wedge m^0(x)}{\sigma^0(x)}$, $E^0_{\xi} = \frac{Z_i - m^0(x)}{\sigma^0(x)}$, $E^0_{T} = E^0_{i\xi} \wedge T$, $T_i^x = \frac{T_i \wedge m^0(x)}{\sigma^0(x)}$. $T_x = T \sigma^0(x) + m^0(x)$ and $T < \tau_{H^0}(\cdot)$, where $H^0(y) = P(E^0 \leq y)$. We refer the reader to Heuchenne and Van Keilegom (2005) for complete description and explanation of this estimator. The same paper also provides proofs of the strong uniform consistency and the strong uniform consistency of the modulus of continuity for the estimator based on those new data points. Those proofs largely use Theorems 3.3 and 4.3.

**Example 2.4 (Nonparametric regression with missing data)**

Suppose in Example 2.1 that some $Y_i$, $i = 1, \ldots, n$, are possibly missing. In this case, $\Delta_i = 0$ if $Y_i$ is a missing data and $\Delta_i = 1$ otherwise. Moreover, the MAR (missing at random) assumption requires that

$$P(\Delta = 1|X, Y) = P(\Delta = 1|X) = p(X) \quad (2.11)$$

(see Little and Rubin, 1987, p.14). In this context, a simple idea (similar to the one developed by Chu and Cheng, 1995) to estimate a regression function is to construct a Nadaraya-Watson estimator with new data points given by

$$\hat{Y}_{i}^* = Y_i \Delta_i + \hat{m}_S(X_i)(1 - \Delta_i), \quad i = 1, \ldots, n,$$

where $\hat{m}_S(x)$ is the Nadaraya-Watson estimator based on the complete pairs:

$$\hat{m}_S(x) = \frac{\sum_{i=1}^{n} K\left(\frac{x - X_i}{a_n}\right)Y_i \Delta_i}{\sum_{i=1}^{n} K\left(\frac{x - X_i}{a_n}\right)\Delta_i}.$$

Therefore, two first applications of Theorem 3.3 with data points $\Gamma_{i\xi}(Z, \Delta|x) = Y \Delta$ and $\Gamma_{i\xi}(Z, \Delta|x) = \Delta$ allow with assumption (2.11) to proof the uniform strong consistency of $\hat{m}_S(x)$. Next, if $f_X(x) = \frac{df_X(x)}{dx}$ and $p(x)$ are uniformly Lipshitz continuous and $m_S(x)$ is two times continuously differentiable, the uniform strong consistency of (1.2) with data points $\Gamma_{i\xi}(Z, \Delta|x) = Y_i^* = Y_i \Delta_i + m_S(X_i)(1 - \Delta_i)$ is obtained in two steps. First, replace $m_S(X_i)$ by $m_S(x) + (X_i - x)m'_S(x) + O(a_n^2)$. Then, by similar developments as in Corollary 1 (ii) of Theorem 2 in Masry (1996), $m'_S(x) \sum_{i=1}^{n} W_i(x, a_n)(X_i - x)(1 - \Delta_i) = O(a_n^2)$ a.s. Second, a third application of Theorem 3.3 allows to obtain the result.
Strong uniform consistency and modulus of continuity proofs are achieved in three steps. First, we consider new data points $\gamma_i(Z_t, \Delta_i|x)$, $i = 1, \ldots, n$, $t \in I$, $x \in R_X$, and kernels that are defined by indicators. Second, we combine those data points to obtain the $\Gamma_t(Z_t, \Delta_i|x)$ used in this section and we sum indicators to construct kernels of step-function form. Third, by using a number of indicators that tends to infinity in the step-function kernel, we show the announced results for usual smooth kernels. The assumptions we need for the proofs of the results of Sections 3 and 4 are listed below.

(A1) $\gamma_i(\cdot, \cdot|\cdot)$ is Lipshitz on $R_X$ (compact) uniformly in $t \in I$:

$$
\sup_{|x-x_j| \leq d, x,x_j \in R_X} \sup_{i \in I} |\gamma_i(z, \delta|x) - \gamma_i(z, \delta|x_j)| \leq L_0(z, \delta|x_j)d
$$

where $L_0(\cdot, \cdot|\cdot)$ is a (positive) function independent of $t$ such that $E[L_0(Z, \Delta|x)^6] \leq L_6 < \infty$ for all $x \in R_X$.

(A2) $0 \leq \gamma_i(z, \delta|x) \leq \gamma_{t'}(z, \delta|x)$, $t < t' \in I$, for all $x, z$ and $\delta = 0, 1$.

(A3) $g(t|x) = E[\gamma_t(Z, \Delta|x)]$ is a continuous function of $t$ in $I$ for all $x$.

(A4) The limit functions $\gamma_{t'} = \lim_{t \to t'} \gamma_t$ and $\gamma_* = \lim_{t \to t_*} \gamma_t$ exist and are finite a.s. (w.r.t. $H(z) = P(Z \leq z)$) for all $x$, where $t_* = \inf \{t : t \in I\}$, $t_* \geq -\infty$ and $t^* = \sup \{t : t \in I\}$, $t^* \leq \infty$.

(A5) For all $x$, $E[\gamma_{t'}(Z, \Delta|x)^{6\lambda}] \leq M_{6\lambda} < \infty$ for some $\lambda$, $2 < \lambda < \infty$; in the case $\lambda = \infty$, $\sup_{x,z,\delta} |\gamma_{t'}(z, \delta|x)| < \infty$.

(A6) Let $\{c_n\}$ a nonnegative sequence satisfy (i) $0 \leq c_n \to 0$, (ii) $\Delta_n = nc_n/\log n \to \infty$, (iii) $c_n^{-1} \leq (n/\log n)^{1-2/\lambda}$, for $\lambda$ as in (A5).

(A7)(i) $F_X(x)$ is differentiable with respect to $x$ with derivative $f_X(x)$.

(ii) $H_\delta(x, y)$ is differentiable with respect to $(x, y)$.

(iii) $H_\delta(y)$ is differentiable with respect to $y$.

(iv) For the density $f_{X|Z,\Delta}(x|z, \delta)$ of $X$ given $(Z, \Delta)$, $\sup_{x,z} |f_{X|Z,\Delta}(x|z, \delta)| < \infty$, $\sup_{x,z} |f_{X|Z,\Delta}(x|z, \delta)| < \infty$ and $\sup_{x,z} |\tilde{f}_{X|Z,\Delta}(x|z, \delta)| < \infty$ ($\delta = 0, 1$), where $f_{X|Z,\Delta}(x|z, \delta)$ ($\tilde{f}_{X|Z,\Delta}(x|z, \delta)$) denotes the first (second) derivative of $f_{X|Z,\Delta}(x|z, \delta)$ with respect to $x$.

(A8) Define new data points as $\Gamma_t(z, \delta|x) = \sum_{i=1}^{i_0} q_i \gamma_{ti}(z, \delta|x)$, $z \in R$, $t \in I$, $x \in R_X$, $\delta = 0, 1$, with fixed and finite $i_0, q_1, \ldots, q_{i_0}$ and with families $\{\gamma_{ti}, t \in I\}$, $1 \leq i \leq i_0$, satisfying assumptions (A1)-(A5), with common $\lambda$ in (A5).
(A9)(i) Consider kernel sequences of step-function form, $K_n(u) = \sum_{j=1}^{j_n} m_{nj}I(-b_{nj} \leq u \leq b_{nj})$, $u \in R$, with $\{j_n\}, \{m_{nj}\}, \{b_{nj}\}$ nonnegative constants such that $|2 \sum_{j=1}^{j_n} m_{nj}b_{nj} - 1| = O(\max(\Delta_n^{-1/2}, a_n^2))$, with $j_n = O(n^s)$, $s > 0$ and $\Delta_n = na_n/\log n$.

(ii) $\sup_n \sum_{j=1}^{j_n} m_{nj}b_{nj}^{1/2} < \infty.$

(iii) $\sup_n \sum_{j=1}^{j_n} m_{nj}b_{nj}^3 < \infty.$

(A10) Let $\{c_n\}$ and $\{d_n\}$ two nonnegative sequences satisfy (i) $0 \leq c_n, d_n \to 0$, (ii) $\Delta_n = nc_n/\log n \to \infty$, (iii) $c_n^{-1} \leq d_n(n/\log n)^{1-2/\lambda}$ for $\lambda$ as in (A5).

(A11) The data points $\gamma_t(Z_i, \Delta_i|x), t \in I$, $x \in R_X$, $i = 1, \ldots, n$, have the following mean-Lipshitz properties: (i) $\sup_{x \in R_X; |t-s| \leq d_n, s \in I, d_n \to 0} |E[\gamma_t(Z, \Delta|x) - \gamma_s(Z, \Delta|x)]| \leq C_Ld_n,$

(ii) $\sup_{x \in R_X; |t-s| \leq d_n, s \in I, d_n \to 0} E[(\gamma_t(Z, \Delta|x) - \gamma_s(Z, \Delta|x))^2] \leq C_L^{2}d_n$, for $n$ sufficiently large.

(A12)(i) – (iii) Consider kernel sequences of the same form and with the same assumptions as in (A9) except that $|2 \sum_{j=1}^{j_n} m_{nj}b_{nj} - 1| = O(\max(\Delta_n^{-1/2}d_n^{-1/2}, a_n^2))$ in (A9) (i).

3 Strong uniform consistency of the weighted average of artificial data points

We start by showing two preliminary results which will lead to the strong uniform consistency of the estimator (1.2).

**Proposition 3.1** Assume (A6), (A7). Then,

$$P(M_{0n}(c_n) > C_0\Delta_n^{-1/2} + C_1c_n^2) = O(n^{-2}),$$

where

$$M_{0n}(c_n) = \sup_{x \in R_X} \sup_{t \in I} \frac{1}{2nc_n} \sum_{i=1}^{n} \gamma_t(Z_i, \Delta_i|x)I(x - c_n < X_i \leq x + c_n)$$

$$- \sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_t(z, \delta|x)h_\delta(x,z)dz,$$

$\gamma_t(z, \delta|x), t \in I, x \in R_X, z \in R, \delta = 0, 1,$ satisfy assumptions (A1)-(A5).
Proposition 3.2 Assume (A7)-(A9). $a_n$ satisfies (i) $a_n B_n \to 0$, (ii) $n a_n b_n / \log n \to \infty$ and (iii) $a_n^{-1} \leq b_n (n / \log n)^{1-2/\lambda}$, where $b_n = \min_j b_{nj}$, $B_n = \max_j b_{nj}$ and $\lambda$ is given as in (A5). Then,

$$
\sup_{x \in R_x} \sup_{t \in I} |d_{tn}(x) - d_t(x)| = O(\max(\Delta_n^{-1/2}, a_n^2)) \text{ a.s.,}
$$

where

$$
d_{tn}(x) = \frac{1}{na_n} \sum_{i=1}^{n} \Gamma_t(Z_i, \Delta_i|x) K_n \left( \frac{x - X_i}{a_n} \right),
$$

and

$$
d_t(x) = \sum_{\delta = 0, 1} \int_{-\infty}^{\infty} \Gamma_t(z, \delta|x) h_\delta(x, z) dz.
$$

Theorem 3.3 Assume (A7), (A8). For the sequence $a_n$, we suppose (i) $a_n \to 0$, (ii) $na_n^{5/2} / \log n \to \infty$, (iii) $a_n^{-5/2} \leq (n / \log n)^{-2/\lambda}$, where $\lambda$ is given as in (A5), and (iv) $na_n^2 \to 0$. $K$ is a symmetric kernel with bounded support, bounded first derivative and $\int K(u) du = 1$. Then,

$$
\sup_{x \in R_x} \sup_{t \in I} |d_{tn}(x) - d_t(x)| = O(\Delta_n^{-1/2}) \text{ a.s.,}
$$

where $d_{tn}(x)$ and $d_t(x)$ are defined with kernel $K$ and $\Delta_n = na_n / \log n$. Moreover, if \( \inf_{x \in R_x} |f_X(x)| > 0 \),

$$
\sup_{x \in R_x} \sup_{t \in I} \left| \frac{\sum_{i=1}^{n} K \left( \frac{X_i - X}{a_n} \right) \Gamma_t(Z_i, \Delta_i|x)}{\sum_{i=1}^{n} K \left( \frac{X_i - X}{a_n} \right)} - \frac{d_t(x)}{f_X(x)} \right| = O(\Delta_n^{-1/2}) \text{ a.s.}
$$

Remark 3.4 (density estimator) If we denote $f_{nX}(x) = (1/na_n) \sum_{i=1}^{n} K \left( \frac{X_i - X}{a_n} \right)$, the classical density estimator, we have using Proposition 3.3 with $\Gamma(Z_i, \Delta_i|x) = 1$ that $\sup_{x \in R_X} |f_{nX}(x) - f_X(x)| = O(\Delta_n^{-1/2})$ a.s., since $\sup_{x \in R_X} |f_X(x)| < \infty$.

Remark 3.5 (moment conditions) For a number of artificial data points, the moment conditions in (A1) and (A5) are not used. Indeed, those data points can often be of the form $\gamma_\nu(Z_i, \Delta_i|x) \leq \gamma_\nu^*(Z_i, \Delta_i)$ and such that $L_0(Z_i, \Delta_i|x) \leq L_0^*(Z_i, \Delta_i)$. In this case, strong law of large numbers can be immediately used with $(1/n) \sum_{i=1}^{n} \gamma_\nu^*(Z_i, \Delta_i) - E[\gamma_\nu^*(Z, \Delta)]$ and $(1/n) \sum_{i=1}^{n} L_0^*(Z_i, \Delta_i) - E[L_0^*(Z, \Delta)]$ in the appendix. The terms $V_n \Delta_n^{-1/2}$ and $2W_n \Delta_n^{-1/2}$ can then be treated outside Proposition 3.1 and be directly introduced in (A.13) in the proof of Proposition 3.2 (see the appendix) such that the final result of
Theorem 3.3 is preserved.

**Remark 3.6 (boundary effects)** The degree of smoothing of \( f_{X|Z}(x|z, \delta) \) allows via (A7) (iv) to obtain the artificial order \( O(\Delta_n^{-1/2}) \) near the boundaries of \( R_X \). If we suppose for instance the weaker condition
\[
\sup_{|x-x'| \leq d, \, x, x' \in R_X} \sup_{t \in I} | \sum_{\delta = 0, 1} \int \gamma_t(z, \delta|x)(h_\delta(x', z) - h_\delta(x, z))dz | \leq Cd,
\]
instead of (A7) (iv), then the more realistic rate \( O(a_n) \) can be obtained near the boundaries.

**Remark 3.7 (bandwidth assumptions)** The bandwidth parameter \( a_n \) could tend to zero more slowly. Indeed, the condition \( na_n^4 \to 0 \) of Theorem 3.3 can be written with another power on \( a_n \). By example, if \( na_n^5 \,(\log n)^{-1} = O(1) \), Theorem 3.3 also holds if \( na_n^3 / \log n \to \infty \) and \( a_n^{-3} \leq (n / \log n)^{1-2/\lambda} \).

**Remark 3.8 (artificial data representation)** The representation
\[
\Gamma_t(z, \delta|x) = \sum_{i=1}^{i_0} q_i \gamma_{ti}(z, \delta|x),
\]
needed in the above proofs, requires nonnegative \( \gamma_{ti}(z, \delta|x) \), \( i = 1, \ldots, i_0 \). This assumption is not restrictive since any random variable \( X \) with real values can be represented by \( X = \max(X, 0) - (-\min(X, 0)) \), where the two terms of this difference are nonnegative.

**Remark 3.9 (Extension to local linear estimator with conditional new data points)** The extension of Theorem 3.3 to local linear estimator is easily obtained by similar developments as in Corollary 1 (ii) of Theorem 2 in Masry (1996) and if \( f_{X}(x) \) is uniformly Lipshitz continuous. Indeed, using those arguments, the local linear estimator reduces to the classical weighted sum of conditional new data points dicussed above.

## 4 Modulus of continuity for the weighted average of conditional synthetic data points

The development of this section is similar to Section 3. The strong uniform consistency of the modulus of continuity is established via two preliminary results.
Proposition 4.1 Assume (A7), (A10). Then,
\[ P(M_{n}(c_n) > C_{m0} \Delta_n^{-1/2} d_n^{1/2} + C_{m1} c_n^2 d_n) = O(n^{-2}), \]
where
\[ M_{n}(c_n) = \sup_{x \in R_x | t-s| \leq d_n} \sup_{s, t \in I} \left| \frac{1}{2n c_n} \sum_{i=1}^{n} (\gamma_i(Z_i, \Delta_i|x) - \gamma_s(Z_i, \Delta_i|x)) I(x - c_n < X_i \leq x + c_n) - \sum_{\delta=0.1}^{\infty} \int_{-\infty}^{\infty} (\gamma_i(z, \delta|x) - \gamma_s(z, \delta|x)) h_{\delta}(x, z) dz \right|, \]
\[ \gamma_i(z, \delta|x), t \in I, x \in R_X, z \in R, \delta = 0, 1, \text{satisfy assumptions (A1)-(A5) and (A11)}. \]

Proposition 4.2 Assume (A7), (A8), (A11), (A12). \( a_n \) and \( d_n \) satisfy (i) \( a_n B_n \rightarrow 0, d_n \rightarrow 0, (ii) na_n b_n/\log n \rightarrow \infty \) and (iii) \( a_n^{-1} \leq b_n d_n(n/\log n)^{1-2/\lambda} \), where \( b_n = \min_{j \leq j_n} b_{n_j}, B_n = \max_{j \leq j_n} b_{n_j} \), and \( \lambda \) is given as in (A5). Then,
\[ \sup_{x \in R_X | t-s| \leq d_n} \sup_{s, t \in I} \left| d_{stn}(x) - d_{st}(x) \right| = O(\max(\Delta_n^{-1/2} d_n^{1/2}, a_n d_n)) \text{ a.s.,} \]
where
\[ d_{stn}(x) = \frac{1}{na_n} \sum_{i=1}^{n} (\Gamma_i(Z_i, \Delta_i|x) - \Gamma_s(Z_i, \Delta_i|x)) K_n(x - X_i/a_n), \]
\[ d_{st}(x) = \sum_{\delta=0.1}^{\infty} \int_{-\infty}^{\infty} (\Gamma_i(z, \delta|x) - \Gamma_s(z, \delta|x)) h_{\delta}(x, z) dz. \]

Proof. The proof is along the same lines as the proof of Proposition 3.2.

Theorem 4.3 Assume (A7), (A8), (A11). \( a_n \) and \( d_n \) satisfy (i) \( a_n \rightarrow 0, d_n \rightarrow 0, (ii) na_n^{5/2} d_n^{-1/2} \log n \rightarrow \infty, (iii) a_n^{-5/2} \leq d_n^{1/2}(n/\log n)^{1-2/\lambda} \), where \( \lambda \) is given as in (A5), (iv) \( \log n/na_n d_n = O(1) \) and (v) \( na_n^{3}\rightarrow 0 \). \( K \) is a symmetric kernel with bounded support, bounded first derivative and \( \int K(u)du = 1 \). Then,
\[ \sup_{x \in R_X | t-s| \leq d_n} \sup_{s, t \in I} \left| d_{stn}(x) - d_{st}(x) \right| = O(\Delta_n^{-1/2} d_n^{1/2}) \text{ a.s.,} \]
where \( d_{stn}(x) \) and \( d_{st}(x) \) are defined with kernel \( K \) and \( \Delta_n = na_n/\log n \). Moreover, if \( na_n^{5/2} / \log n \rightarrow \infty \) and \( \inf_{x \in R_X} f_X(x) > 0 \),
\[ \sup_{x \in R_X | t-s| \leq d_n} \sup_{s, t \in I} \left| \frac{\sum_{i=1}^{n} K \left( \frac{x - X_i}{a_n} \right) \Gamma_{ts}(Z_i, \Delta_i|x)}{\sum_{i=1}^{n} K \left( \frac{x - X_i}{a_n} \right)} - \frac{d_{st}(x)}{f_X(x)} \right| = O(\Delta_n^{-1/2} d_n^{1/2}) \text{ a.s.,} \]
where \( \Gamma_{ts}(Z_i, \Delta_i|x) = \Gamma_t(Z_i, \Delta_i|x) - \Gamma_s(Z_i, \Delta_i|x) \).

Remark 4.4 (bandwidth assumptions) If \( na_n^{3}(\log n)^{-1} = O(1) \), Theorem 4.3 also holds if \( na_n^{3}/\log n \rightarrow \infty \) and \( a_n^{-3} \leq d_n^{1/2}(n/\log n)^{1-2/\lambda} \).
Appendix : Proofs of main results

Proof of Proposition 3.1. Let \( f_n = \Delta_n^{-1/2}c_n \). We have
\[
M_{2n}(c_n) = \sup_{x \in R_X} \sup_{t \in I} \left| \frac{1}{2nc_n} \sum_{i=1}^{n} \gamma_i(Z_i, \Delta_i|x)I(x - c_n < X_i \leq x + c_n) \right|
\]
\[
- \frac{1}{2cn} \int_{x-c_n}^{x+c_n} \sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_i(z, \delta|x) h_\delta(s, z) dz ds
+ \sup_{x \in R_X} \sup_{t \in I} \left| \frac{1}{2cn} \int_{x-c_n}^{x+c_n} \sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_i(z, \delta|x) h_\delta(s, z) dz ds \right|
- \sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_i(z, \delta|x) h_\delta(x, z) dz
\]
\[
= \sup_{x \in R_X} \sup_{t \in I} |M_{2n}(x)| + \sup_{x \in R_X} |M_{2n}(x)|.
\]

First, we treat the term \( M_{2n}(x) \). It is given by
\[
\sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_i(z, \delta|x) \left\{ \frac{1}{2cn} \int_{x-c_n}^{x+c_n} h_\delta(s, z) ds - h_\delta(x, z) \right\} dz.
\]

Using two Taylor developments of order three around \( x \), we get
\[
\frac{1}{2cn} \int_{x-c_n}^{x+c_n} h_\delta(s, z) ds - h_\delta(x, z) = (c_n^2/12)[\tilde{f}_X|Z, \Delta(\theta_1|z, \delta) + \tilde{f}_X|Z, \Delta(\theta_2|z, \delta)] h_\delta(z),
\]
where \( \theta_1 (\theta_2) \) is between \( x + c \) and \( x \) (and \( x - c \)). Since \( \sup_{x,z} |\tilde{f}_X|Z, \Delta(x, z, \delta)| < \infty \) (\( \delta = 0, 1 \)) and \( \sup_{x \in R_X, t \in I} E[\gamma_i(Z, \Delta|x)] < \infty \) with \( \gamma_i(Z, \Delta|x) \geq 0 \),
\[
\sup_{x \in R_X, t \in I} |M_{2n}(x)| \leq C_1c_n^2 \quad \text{(A.1)}
\]

Let \( L_X \) the length of \( R_X \) and divide \( R_X \) into \( \left[ \frac{2L_X}{f_n} \right] \) intervals of length smaller or equal to \( f_n \) ([x] denotes the integer part of x). Denote \( x_0 = \inf\{x : x \in R_X\} \), \( I_X \) the set of points \( \{x_k = x_0 + k[\frac{2L_X}{f_n}]^{-1}L_X, \ 1 \leq k \leq \left[ \frac{2L_X}{f_n} \right] - 1 = L_X^n\} \) and \( x_{L_X+1} = \sup\{x : x \in R_X\} \) which limit the intervals. Using the Lipshitz condition (A1), we can rewrite for \( 1 \leq j \leq L_X^n \),
\[
\sup_{x \in R_X} \sup_{t \in I} |M_{1n}(x)| \quad \text{(A.2)}
\]
\[
\leq \max_{x_j \in I_x} \sup_{x \in [x_{j-1}, x_{j+1}]} \sup_{t \in I} \left| \frac{1}{2nc_n} \sum_{i=1}^{n} \gamma_i(Z_i, \Delta_i|x_j)I(x - c_n < X_i \leq x + c_n) \right|
- \frac{1}{2cn} \int_{x-c_n}^{x+c_n} \sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_i(z, \delta|x_j) h_\delta(s, z) dz ds
+ \max_{x_j \in I_x} \left( E[\gamma_i(Z, \Delta|x_j)] + |V_n(x_j)| \right) \Delta^{-1/2}_n
= (2c_n)^{-1} \max_{x_j \in I_x} \sup_{x \in [x_{j-1}, x_{j+1}]} \sup_{t \in I} |M_{3n}(x_j, x)| + \max_{x_j \in I_x} \left( E[\gamma_i(Z, \Delta|x_j)] + |V_n(x_j)| \right) \Delta^{-1/2}_n,
\]

12
where \( V_n(x_j) = (1/2n) \sum_{i=1}^{n} L_0(Z_i, \Delta|x_j|) - (E[L_0(Z, \Delta|x_j|)]/2) \). We have

\[
P(\max_{x_j \in I_X} (E[L_0(Z, \Delta|x_j|)] + |V_n(x_j)|) > 2C_2) \\
\leq \sum_j \{ P(|2V_n(x_j)| > 2C_2) + P(E[L_0(Z, \Delta|x_j|)] > C_2) \},
\]

where the second term on the right hand side of the above expression is zero when \( C_2 > L_6^{1/6} \). For the first term, we use an extension of Chebyshev’s inequality:

\[
P(\left| n^{-1} \sum_{i=1}^{n} L_0(Z_i, \Delta_i|x_j|) - E[L_0(Z, \Delta|x_j|)] \right| > 2C_2) \\
\leq \frac{1}{(2nC_2)^6} E[\left\{ \sum_{i=1}^{n} (L_0(Z_i, \Delta_i|x_j|) - E[L_0(Z, \Delta|x_j|)]) \right\}^6] = O(n^{-3}),
\]

since \( E[L_0(Z, \Delta|x_j|)^6] \leq L_6 < \infty \). Then, with \( L_X^n O(n^{-3}) = o(n^{-2}) \),

\[
P(\max_{x_j \in I_X} (E[L_0(Z, \Delta|x_j|)] + |V_n(x_j)|) > 2C_2) = o(n^{-2}),
\]

for which, using Borel-Cantelli Lemma, we obtain

\[
V_n = \max_{x_j \in I_X} (E[L_0(Z, \Delta|x_j|)] + |V_n(x_j)|) = O(1) \text{ a.s.}
\]

To treat the first term on the right hand side of (A.2), we introduce some notations. Let

\[
G_{tn}(x_j, x) = n^{-1} \sum_{i=1}^{n} \gamma_{t}(Z_i, \Delta_i|x_j|) I(X_i \leq x),
\]

and

\[
G_t(x_j, x) = E[G_{tn}(x_j, x)] = \int_{x_0}^{x} \sum_{\delta=0,1} \int_{-\infty}^{\infty} \gamma_{t}(z, \delta|x_j|) h_{\delta}(s, z) dz ds.
\]

Therefore,

\[
|M_{3tn}(x_j, x)| = |G_{tn}(x_j, x + c_n) - G_{tn}(x_j, x - c_n) - [G_t(x_j, x + c_n) - G_t(x_j, x - c_n)]| \\
\leq 2 \sup_{|t| \leq c_n} |G_{tn}(x_j, x + z) - G_{tn}(x_j, x) - [G_t(x_j, x + z) - G_t(x_j, x)]| \\
= 2M_{4tn}(x_j, x, c_n).
\]

By conditions (A2) – (A5), the functions \( g(t|x_j|) \), \( j = 1, \ldots, L_X^n \), are nondecreasing, continuous in \( t \) with finite limits \( g(t^*|x_j|) \) and \( g(t_*|x_j|) \) as \( t \rightarrow t^* \) and \( t_* \). Divide \( I \) in \( O(f_n^{-2}) \) intervals such that \( |g(t_1|x_j|) - g(t_*|x_j|)| \leq f_n, |g(t_{k+1}|x_j|) - g(t_k|x_j|)| \leq f_n \), for \( k = 1, \ldots N_{n_j} - 1, \)

13
\[ |g(t_1|x_1) - g(t_{N_{nj}}|x_1)| \leq f_n, \quad j = 1, \ldots, L^*_X. \]

Let \( I_{nj} \) denote the set \( \{t_1, \ldots, t_{N_{nj}}, t^*\} \) and \( I^*_j \) the set \( \{(t_1, t_1), (t_1, t_2), \ldots, (t_{N_{nj}}, t^*)\} \). Clearly, the cardinality \( N_{nj} \) of \( I^*_j \) is bounded by

\[
N_{nj} \leq \frac{2(g(t^*|x_j) - g(t_1|x_j))}{f_n}.
\]

Also, for fixed \( j, x, z, n \), the functions \( G_{tn}(x, x+z) - G_{tn}(x, x) \) and \( G_{t}(x, x+z) - G_{t}(x, x) \) are monotone in \( t \) and have finite limits as \( t \to t_1, t^* \). We therefore have

\[
|G_{tn}(x, x+z) - G_{tn}(x, x) - [G_{t}(x, x+z) - G_{t}(x, x)]| \\
\leq \max_{t \in I_{nj}} |G_{tn}(x, x+z) - G_{tn}(x, x) - [G_{t}(x, x+z) - G_{t}(x, x)]| \\
+ \max_{(s,t) \in I^*_j} |G_{t}(x, x+z) - G_{s}(x, x+z) - [G_{t}(x, x) - G_{s}(x, x)]|. \quad (A.4)
\]

It is easily shown that the right hand side of the above expression is bounded by

\[
2 \int_{R_X} \sum_{\delta=0,1} \int_{-\infty}^{\infty} (\gamma_t(z, \delta|x_j)) - \gamma_s(z, \delta|x_j)) h_\delta(s, z) dz ds \\
= 2 \sum_{\delta=0,1} \int_{-\infty}^{\infty} (\gamma_t(z, \delta|x_j)) - \gamma_s(z, \delta|x_j)) h_\delta(z) dz \\
\leq 2(g(t|x_j) - g(s|x_j)) \leq 2f_n,
\]

using monotonicity of \( \gamma_t \) with respect to \( t \). Therefore,

\[
\max_{x_j \in I_X} \sup_{x \in [x_{j-1}, x_{j+1}]} M_{4tn}(x_j, x, c_n) \leq \max_{x_j \in I_X} \max_{x \in [x_{j-1}, x_{j+1}]} M_{4tn}(x_j, x, c_n) + 2f_n. \quad (A.5)
\]

Let

\[
M_{5tn}(x_j, x, z) = |G_{tn}(x, x+z) - G_{tn}(x, x) - [G_{t}(x, x+z) - G_{t}(x, x)]|.
\]

We have

\[
\max_{x_j \in I_X} \sup_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{nj}} M_{5tn}(x_j, x, z) \quad (A.6) \\
\leq \max_{x_j \in I_X} \max_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{nj}} M_{5tn}(x_j, x, z), \\
\max_{x_j \in I_X \setminus \{x_n\}} \max_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{nj}} M_{5tn}(x_j, x, z), \\
\max_{x_j \in I_X \setminus \{x_n\}} \max_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{nj}} M_{5tn}(x_j, x, z), \\
\max_{x_j \in I_X \setminus \{x_n\}} \max_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{nj}} M_{5tn}(x_j, x, z), \\
\max_{x_j \in I_X \setminus \{x_n\}} \max_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{nj}} M_{5tn}(x_j, x, z). \]
Introducing $G_{tn}(x_j, x_{j+k})$ and $G_t(x_j, x_{j+k})$ for $k = -1, 0$ or $1$, it is easily shown that

\[
\max_{x_j \in I_x} \sup_{x \in [x_{j-1}, x_{j+1}]} \max_{t \in I_{tn}} M_{4tn}(x_j, x, c_n) \leq 2 \max_{x_j \in I_x} \max_{t \in I_{tn}} \sup_{k \in \{-1, 0, 1\}} \|x\| \leq c_n M_{5tn}(x_j, x_{j+k}, z).
\]

Now, put $Q_n = M_{\lambda}f_n^{-1/(\lambda-1)}$, where, for all $x$, $(E[(\gamma_t(Z, \Delta|x)^{\lambda})^{1/\lambda}] \leq M_{\lambda} < \infty$ for some $\lambda$, $2 < \lambda < \infty$. In the case $\lambda = \infty$, $M_{\infty}$ denotes then $\sup_{x, z, \delta} |\gamma_t(z, \delta|x)|$. Also, put

\[
H_{tn}(x_j, x) = n^{-1} \sum_{i=1}^{n} \gamma_i(Z_i, \Delta_i|x_j)\big(\gamma_i(Z_i, \Delta_i|x_j) \leq Q_n\big)I(X_i \leq x),
\]

and define $M_{6tn}(x_j, x, z)$ by substitution of $H_{tn}$ for $G_{tn}$ and $E[H_{tn}]$ for $G_t$ in $M_{5tn}(x_j, x, z)$. That yields

\[
\sup_{x \in R_x} \sup_{t \in I} |M_{1tn}(x)| \leq 2c_n^{-1} \max_{x_j \in I_x} \max_{t \in I_{tn}} \max_{k \in \{-1, 0, 1\}} \|x\| \leq c_n M_{6tn}(x_j, x_{j+k}, z) + 2f_n c_n^{-1}(1 + V_n/2 + W_n + \theta_n),
\]

where

\[
W_n = f_n^{-1} \max_{x_j \in I_x} \max_{t \in I_{tn}} \max_{k \in \{-1, 0, 1\}} \|x\| \leq c_n M_{7tn}(x_j, x_{j+k}, z),
\]

\[
M_{7tn}(x_j, x, z) = |G_{tn}(x_j, x + z) - G_{tn}(x_j, x) - [H_{tn}(x_j, x + z) - H_{tn}(x_j, x)]|,
\]

\[
\theta_n = f_n^{-1} \max_{x_j \in I_x} \max_{t \in I_{tn}} \max_{k \in \{-1, 0, 1\}} \|x\| \leq c_n M_{8tn}(x_j, x_{j+k}, z),
\]

and

\[
M_{8tn}(x_j, x, z) = |G_{t}(x_j, x + z) - G_{t}(x_j, x) - [E[H_{tn}(x_j, x + z)] - E[H_{tn}(x_j, x)]]|.
\]

Using (A2), (A4) and the fact that $f_n^{-1} = (Q_n/M_{\lambda})^{\lambda-1}$, we have

\[
M_{\lambda}^{-1}W_n \leq Q_n^{-1} \max_{x_j \in I_x} n^{-1} \sum_{i=1}^{n} \gamma_{t_i}(Z_i, \Delta_i|x_j)\big(\gamma_{t_i}(Z_i, \Delta_i|x_j) > Q_n\big) \leq \max_{x_j \in I_x} n^{-1} \sum_{i=1}^{n} \gamma_{t_i}(Z_i, \Delta_i|x_j)^\lambda,
\]

if $\lambda < \infty$ and $W_n = 0$ if $\lambda = \infty$. Next, for $\lambda < \infty$,

\[
P\left(\max_{x_j \in I_x} n^{-1} \sum_{i=1}^{n} \gamma_{t_i}(Z_i, \Delta_i|x_j)^\lambda > C_3\right) \leq \sum_{j} \{P(n^{-1} \sum_{i=1}^{n} \gamma_{t_i}(Z_i, \Delta_i|x_j)^\lambda - E[\gamma_{t_i}(Z, \Delta|x_j)^\lambda] > C_3/2\} + P\left(E[\gamma_{t_i}(Z, \Delta|x_j)^\lambda] > C_3/2\right)
\]
where the second term on the right hand side of the above expression is zero when $C_3/2 > M_\lambda^\lambda$. For the first term, we also use the extension of Chebyshev’s inequality:

\[
P(n^{-1} \sum_{i=1}^{n} \gamma_{t^i} (Z_i, \Delta_i|x_j) - E[\gamma_{t^i} (Z_i, \Delta_i|x_j)] > C_3/2)
\]

\[
\leq \frac{64}{(nC_3)^n} E[\{ \sum_{i=1}^{n} (\gamma_{t^i} (Z_i, \Delta_i|x_j) - E[\gamma_{t^i} (Z_i, \Delta_i|x_j)]) \}^6] = O(n^{-3}),
\]

since $E[\gamma_{t^i} (Z_i, \Delta_i|x_j)] \leq M_6 < \infty$. Then, with $L_X O(n^{-3}) = o(n^{-2})$,

\[
P(\max_{x_j \in I_X} n^{-1} \sum_{i=1}^{n} \gamma_{t^i} (Z_i, \Delta_i|x_j) > C_3) = o(n^{-2}),
\]

for which, using Borel-Cantelli Lemma, we obtain $W_n = O(1)$ a.s. We also see that

\[
\theta_n \leq \frac{\max_{x_j \in R_X} E[\gamma_{t^i} (Z_i, \Delta_i|x_j)]}{M_\lambda^{\lambda-1}} \leq M_\lambda,
\]

if $\lambda < \infty$ and $\theta_n = 0$ if $\lambda = \infty$.

Now, define

\[
w_n = \frac{[2Q_n c_n]}{f_n} + 1
\]

and

\[
\eta_{njk} = x_{j+k} + \frac{r c_n}{w_n}, \text{ for } r = -w_n, -w_n + 1, \ldots, w_n.
\]

Defining

\[
\xi_{ntjk} = |H_{tn}(x_j, \eta_{njk}) - H_{tn}(x_j, x_{j+k}) - [E[H_{tn}(x_j, \eta_{njk})] - E[H_{tn}(x_j, x_{j+k})]] |
\]

we have

\[
\sup_{|z| \leq c_n} M_{\beta n}(x_j, x_{j+k}, z) \leq \max_{-w_n \leq r \leq w_n} \xi_{ntjk}
\]

\[
+ \max_{-w_n \leq r \leq w_n - 1} |E[H_{tn}(x_j, \eta_{njk(r+1)})] - E[H_{tn}(x_j, \eta_{njk(r)})]] |
\]

The second term of the right hand side of the above expression is bounded by

\[
Q_n \max_{-w_n \leq r \leq w_n - 1} \int_{\eta_{njk}}^{\eta_{njk(r+1)}} \sum_{\delta = 0,1} \int_{-\infty}^{\infty} h_\delta(s, z) dz ds
\]

\[
\leq Q_n \max_{-w_n \leq r \leq w_n - 1} \int_{\eta_{njk}}^{\eta_{njk(r+1)}} f_X(s) ds \leq Q_n C_4 (\eta_{njk(r+1)} - \eta_{njk}) \leq C_4 f_n / 2,
\]

16
where \(C_4\) is the Lipshitz constant of \(F_X(\cdot)\). The goal is therefore to calculate
\[
P(\sup_{x \in R_X} |M_{1n}(x)| > C_0 \Delta_n^{-1/2})
\]
\[
\leq P(2 \max_{j,t,k,r} \xi_{ntjkr} + f_n(2W_n + V_n) > (C_0 - 2 - C_4 - 2M_\lambda)f_n)
\]
\[
\leq \sum_{j,t,k,r} P(\xi_{ntjkr} > (1/6)(C_0 - 2 - C_4 - 2M_\lambda)f_n)
\]
\[
+ P(W_n > (1/6)(C_0 - 2 - C_4 - 2M_\lambda))
\]
\[
+ P(V_n > (1/3)(C_0 - 2 - C_4 - 2M_\lambda)),
\]
where \(C_0, C_2\) and \(C_3\) have to satisfy \(2M_\lambda < (C_3/M_\lambda^{\lambda-1}) = (1/6)(C_0 - C_4 - 2M_\lambda - 2)\) and \(L_6^{1/6} < C_2 = (1/6)(C_0 - C_4 - 2M_\lambda - 2)\). Therefore, we only have to treat the first term on the right hand side of the above expression. Defining \(C'_0 = (1/6)(C_0 - C_4 - 2M_\lambda)\), we have by Bernstein’s inequality,
\[
P(\xi_{ntjkr} > C'_0 f_n) \leq 2 \exp(-\nu_{tjknr}),
\]
where
\[
\nu_{tjknr} = \frac{C'_0^2 n^2 f_n^2}{2n\sigma_{tjknr}^2 + \frac{2}{3}nC'_0 f_n Q_n},
\]
and \(\sigma_{tjknr} = Var[D_{tjknr}]\), where
\[
D_{tjknr} = \gamma_t(Z, \Delta|x_j)I(\gamma_t(Z, \Delta|x_j) \leq Q_n)(I(X \leq \eta_{tjkr}) - I(X \leq x_{j+k})).
\]
We have
\[
\sigma_{tjknr}^2 \leq E[D_{tjknr}^2] \leq \int_{x_{j+k}}^{\eta_{tjkr}} \int_{x_{j+k}}^{\eta_{tjkr}} \sum_{\delta=0,1} \int_{-\infty}^{\gamma^2} \gamma_t^2(z, \delta|x_j)I(\gamma_t(z, \delta|x_j) \leq Q_n)h_\delta(s, z)dzds
\]
\[
\leq C_5 M_\lambda^2 c_n,
\]
using (A5), condition (A7) (iv) and where \(C_5 = \max_s \sup_{x,z} |f_X(x, \Delta|x, \delta)|\). Using (A6) (iii),
\[
Q_n f_n = M_\lambda f_n^{(\lambda-2)/\lambda-1} = M_\lambda \left(\frac{c_n \log n}{n}\right)^{(\lambda-2)/2(\lambda-1)} \leq M_\lambda c_n.
\]
We thus have by (A7) and (A8),
\[
\nu_{tjknr} \geq C_0'' \log n,
\]

17
with
\[ C''_0 = \frac{C_0^2}{2M_\lambda(C_5M_\lambda + \frac{1}{3}C''_0)} > \max(\frac{6}{3C_5 + 2}, \frac{3L_0^{1/3}}{M_\lambda(6M_\lambda C_5 + 2L_0^{1/6})}). \]

Therefore,
\[ \sum_{j,t,k,r} P(\xi_{ntjkr} > C''_0f_n) \leq 6 \sum_{j=1}^{L_0^X}(N_{nj} + 2)(2w_n + 1)n^{-C''_0}, \quad (A.9) \]

where \( C''_0 \) has to be chosen large enough so that the right hand side of (A.9) tends to zero sufficiently fast. Thus, the highest order term on the right hand side of (A.9) is
\[ 96\left(\frac{L_0^X M_\lambda Q_n c_n}{f_n^3}\right) n^{-C''_0}, \quad (A.10) \]

where
\[ N_{nj} \leq \frac{2g(t^*|x_j)}{f_n} \leq 2 \frac{M_\lambda}{f_n} \]

and
\[ \frac{Q_n c_n}{f_n} = M_\lambda c_n\left(\frac{n}{c_n \log n}\right)^{\lambda}, \]

Using (A6) (iii), (A.10) is bounded by
\[ L_\lambda\left(\frac{n}{c_n \log n}\right)^{\lambda-2} c_n n^{-C''_0} \leq \frac{n}{\log n}^{2\lambda} n^{-C''_0}, \]

where \( L_\lambda = 96M_\lambda^2 L_X \). Therefore, choosing \( C''_0 \geq 4 \) allows to write
\[ P\left( \sup_{x \in \mathcal{R}_X} \sup_{t \in I} |M_{1tn}(x)| > C_0 \Delta_n^{-1/2} \right) = O(n^{-2}). \quad (A.11) \]

By (A.1) and (A.11), we finally obtain
\[ P\left( \sup_{x \in \mathcal{R}_X} \sup_{t \in I} |M_{1tn}(x) + M_{2tn}(x)| > C_0 \Delta_n^{-1/2} + C_1 c_n^2 \right) = O(n^{-2}). \quad (A.12) \]

**Proof of Proposition 3.2.** Write
\[ \sup_{x \in \mathcal{R}_X} \sup_{t \in I} |d_{tn}(x) - d_t(x)| \leq \sum_{i=1}^{i_0} |q_i|(S_{ni}^{(1)} + S_{ni}^{(2)}), \]

where
\[ S_{ni}^{(1)} = 2 \sum_{j=1}^{j_n} m_{nj} b_{nj} \sup_{x \in \mathcal{R}_X} \sup_{t \in I} |M_{1tn}^{(i,j)}(x) + M_{2tn}^{(i,j)}(x)|, \quad (A.13) \]
\[
M_{1tn}^{(i,j)}(x) + M_{2tn}^{(i,j)}(x) = \sup_{x \in R_X} \sup_{t \in I} |\frac{1}{2n \delta_{nj}} \sum_{k=1}^{n} \gamma_{t_k}(Z_{k},\Delta_k|x)I(x - c_{nj} < X_k \leq x + c_{nj})
- \sum_{\delta=0.1}^{\delta} \int_{-\infty}^{\infty} \gamma_{t_k}(z,\delta|x)h_\delta(x,z)dz|,
\]
c_{nj} = a_n b_{nj} and
\[
S_{ni}^{(2)} = \sup_{x \in R_X} \sup_{t \in I}|(2 \sum_{j=1}^{j_n} m_{nj} b_{nj} - 1) \sum_{\delta=0.1}^{\delta} \int_{-\infty}^{\infty} \gamma_{t_k}(z,\delta|x)h_\delta(x,z)dz|
\leq C_5 M_\lambda|(2 \sum_{j=1}^{j_n} m_{nj} b_{nj} - 1)|.
\]
Next, define
\[
\varepsilon_n = 2C_0 \Delta_n^{-1/2} \sum_{j=1}^{j_n} m_{nj} b_{nj}^{1/2} + 2C_1 a_n^2 \sum_{j=1}^{j_n} m_{nj} b_{nj}^3.
\]
Then,
\[
P(S_{ni}^{(1)} > \varepsilon_n) \leq \sum_{j=1}^{j_n} P(\sup_{x \in R_X} \sup_{t \in I} |M_{1tn}^{(i,j)}(x) + M_{2tn}^{(i,j)}(x)| > C_0 \Delta_n^{-1/2} + C_1 a_n^2),
\]
with \( \Delta_n = \Delta_n b_{nj} \) and \( c_{nj} = a_n b_{nj} \). By using (A.12), we thus obtain
\[
P(S_{ni}^{(1)} > \varepsilon_n) \leq O(j_n n^{-2}).
\]
For \( s < 1 \) and using the Borel-Cantelli Lemma, we obtain
\[
S_{ni}^{(1)} = O(\varepsilon_n) a.s.,
\]
for which \( \varepsilon_n = O(\max(\Delta_n^{-1/2}, a_n^2)) \).

**Proof of Theorem 3.3.** Let \( b_{nj} = j a_n^{3/2} \) and \( m_{nj} = K(j a_n^{3/2}) - K((j + 1) a_n^{3/2}) \) in (A9). (A9) (i) becomes
\[
|(2 \sum_{j=1}^{j_n} m_{nj} b_{nj} - 1)| \leq \int |K_n(u) - K(u)|du \leq C a_n^{3/2},
\]
for some \( C > 0 \). (A9) (ii) and (A9) (iii) become
\[
\sup_n a_n^{9/4} \sum_{j=1}^{j_n} j^{1/2}|K'(\theta_{nj})| < \infty
\]
and
\[ \sup_n a_n^6 \sum_{j=1}^{j_n} |K'(\theta_n)| < \infty, \]
where \( \theta_n \) is between \( ja_n^{3/2} \) and \((j + 1)a_n^{3/2}\). Therefore, we can choose \(0 < s < 1\) such that \(j_n = O(a_n^{-3/2})\). Next, if we denote \(d_{tn}(x, K), d_{tn}(x)\) using kernel \(K\), it is clear that
\[ \sup_{x \in \mathbb{R}^d} \sup_{t \in I} |d_{tn}(x, K) - d_{tn}(x, K_n)| \]
\[ \leq \sup_{x \in \mathbb{R}^d} \sup_{t \in I} \left( \frac{Da_n^{3/2}}{na_n} \sum_{i=1}^{n} |\Gamma_i(Z_i, \Delta_i)| I(x - a_n < X_i \leq x + a_n) \right) = O(a_n^{-3/2}) \text{ a.s.,} \]
for some constant \(D > 0\), a kernel support equal to \([-1, 1]\) and where Proposition 3.1 is used with \(c_n = a_n\) (with \(c_n = (L/2)a_n\) if \(L\) is the length of the support).
Finally, write
\[ \sup_{x \in \mathbb{R}^d} \sup_{t \in I} \left| \sum_{i=1}^{n} K\left( \frac{Z_i}{a_n} \right) \Gamma_i(Z_i, \Delta_i) - \frac{d_t(x)}{f_X(x)} \right| \leq \sup_{x \in \mathbb{R}^d} \sup_{t \in I} \left| \frac{d_t(x) - d_t(x)}{f_n(x)} \right| \]
\[ + \sup_{x \in \mathbb{R}^d} \sup_{t \in I} \left| \frac{d_t(x)(f_X(x) - f_n(x))}{f_X(x)f_n(x)} \right|. \]
If we use the fact that \(\inf_{x \in \mathbb{R}^d} |f_X(x)| > 0\) in addition to the obtained result for \(d_{tn}(x, K)\), the two terms on the right hand side of the above expression are \(O(\Delta_n^{-1/2}) \text{ a.s.} \)
\[ \text{since } \sup_{x \in \mathbb{R}^d} \sup_{t \in I} |d_t(x)| \text{ is bounded (using the definition (A8) of the points } \Gamma(\cdot, \cdot, \cdot)). \]

**Proof of Proposition 4.1.** Let
\[ M_{9n}(c_n) = \sup_{x \in \mathbb{R}^d} \sup_{[t-s] \leq d_n, s,t \in I} \left| \frac{1}{2nc_n} \sum_{i=1}^{n} (\gamma_t(Z_i, \Delta_i)|x) - \gamma_s(Z_i, \Delta_i)|x)I(x - c_n < X_i \leq x + c_n) \right| \]
\[ - \frac{1}{2c_n} \int_{x-c_n}^{x+c_n} \sum_{\delta=0,1} \int_{-\infty}^{\infty} (\gamma_t(z, \delta|x) - \gamma_s(z, \delta|x))h_\delta(s, z)dzds \]
\[ + \sup_{x \in \mathbb{R}^d} \sup_{[t-s] \leq d_n, s,t \in I} \left| \frac{1}{2c_n} \int_{x-c_n}^{x+c_n} \sum_{\delta=0,1} \int_{-\infty}^{\infty} (\gamma_t(z, \delta|x) - \gamma_s(z, \delta|x))h_\delta(s, z)dzds \right| \]
\[ - \sum_{\delta=0,1} \int_{-\infty}^{\infty} (\gamma_t(z, \delta|x) - \gamma_s(z, \delta|x))h_\delta(s, z)dz \]
\[ = \sup_{x \in \mathbb{R}^d} \sup_{[t-s] \leq d_n, s,t \in I} \left| M_{10stn}(x) \right| + \sup_{x \in \mathbb{R}^d} \sup_{[t-s] \leq d_n, s,t \in I} \left| M_{11stn}(x) \right|. \]
First, \(M_{11stn}(x)\) is treated as \(M_{2tn}(x)\) in Proposition 3.1 such that
\[ \sup_{x \in \mathbb{R}^d, [t-s] \leq d_n, s,t \in I} \left| M_{11stn}(x) \right| \leq C_m d_n c_n^2, \quad (A.14) \]
using (A11).

Divide $R_X$ into $\frac{2L_X}{f_{X,n}}$ intervals of length smaller or equal to $f_n d_n^{1/2}$. Denote $J_X$ the set of points \( \{ x_k = x_0 + k\frac{2L_X}{f_{X,n}} \} \), $1 \leq k \leq \frac{2L_X}{f_{X,n}} - 1 = L_X^m$ and $x_{L_X^m+1} = \sup \{ x : x \in R_X \}$ which limit the intervals. $M_{10\text{stn}}(x)$ is treated like (A.2) in Proposition 3.1, where $\gamma_t(\cdot, \cdot; \cdot)$ is replaced by $\gamma_t(\cdot; \cdot; \cdot)$, $V_n(x_j)$ by $V_n^d(x_j) = \frac{1}{n} \sum_{i=1}^{n} L_0(Z_i, \Delta_i|x_j) - E[L_0(Z, \Delta|x_j)]$ and $V_n$ by $V_n^d = \max_{x_j \in J_X} (2E[L_0(Z, \Delta|x_j)] + |V_n^d(x_j)|)$. Using Chebyshev’s inequality, $P(V_n^d > C_{m2}) = o(n^{-2})$ with $C_{m2}$ chosen larger than $4L_0^{d/6}$. A development similar to (A.3) and (A.6) is used to obtain

\[
\sup_{x \in R_X \cup \{ -\infty \}} \sup_{|t| \leq d_n} |M_{10\text{stn}}(x)| \leq 2c_n^{-1} \max_{x \in J_X} \max_{|s| \leq d_n} \max_{s \in I} \max_{k \in \{-1, 0, 1\}} \sup_{|z| \leq c_n} M_{12\text{stn}}(x_j, x_{j+k}, z) + V_n^d \Delta_n^{-1/2} d_n^{1/2},
\]  

(A.15)

where

\[
M_{12\text{stn}}(x_j, x, z) = |G_{stn}(x_j, x + z) - G_{stn}(x_j, x)| - |G_{st}(x_j, x + z) - G_{st}(x_j, x)|,
\]

\[
G_{stn}(x_j, x) = n^{-1} \sum_{i=1}^{n} (\gamma_t(Z_i, \Delta_i|x_j) - \gamma_s(Z_i, \Delta_i|x_j)) I(X_i \leq x),
\]

and

\[
G_{st}(x_j, x) = E[G_{stn}(x_j, x)] = \int_{x_0}^{x} \sum_{\delta=0,1} \int_{-\infty}^{\infty} (\gamma_t(z, \delta|x_j) - \gamma_s(z, \delta|x_j)) h_\delta(s, z) dz ds.
\]

Partition $I$ into $O(f_n d_n^{-3/2})$ intervals such that for each $x_j$, $j = 0, \ldots, L_X^m + 2$, $g(t|x_j) - g(t_0|x_j)$ is divided into $m_j = \left[ \frac{g(t|x_j) - g(t_0|x_j)}{C_{L,d_n}} \right]$ intervals of length $C_{m3}(x_j)C_Ld_n$, $1 \leq C_{m3}(x_j) \leq 2 (|g(t_0|x_j) - g(t_1|x_j)|) = C_{m3}(x_j)C_Ld_n$, $|g(t_{\alpha+1}|x_j) - g(t_0|x_j)| = C_{m3}(x_j)C_Ld_n$, $\alpha = 1, \ldots, m_j - 2$, $|g(t_0|x_j) - g(t_{m_j-1}|x_j)| = C_{m3}(x_j)C_Ld_n$, $t_0 = t_0$, $t_{m_j} = t^*$ for all $j$. Let $I_{ja} = [g(t_{\alpha-1}|x_j), g(t_{\alpha+1}|x_j)]$, $\alpha = 1, \ldots, m_j - 1$. For each $s, t \in I$ with $|t - s| \leq d_n$, there exists an interval $I_{ja}$ such that $g(s|x_j), g(t|x_j) \in I_{ja}$. Partition each $I_{ja}$ by a grid $g(t_{\alpha\beta}|x_j) = g(t_0|x_j) + \beta C_{m3}(x_j)C_{L,d_n}$, $\beta = -p_n, \ldots, p_n$, where $p_n = \lfloor \Delta_n^{1/2} d_n^{1/2} + 1 \rfloor$. Using (A7) (iv), (A4), (A5) and the monotonicity of $\gamma_t(Z, \Delta|x)$, (A.15) is majorized by

\[
2c_n^{-1} \max_{x_j \in J_X} \max_{k \in \{-1, 0, 1\}} \max_{1 \leq \alpha \leq m_j - 1} \max_{-p_n \leq \beta \leq p_n} \max_{|z| \leq c_n} M_{12\text{stn}}(x_j, x_{j+k}, z) + 4c_n^{-1} \max_{x_j \in J_X} \max_{1 \leq \alpha \leq m_j - 1} \max_{-p_n \leq \beta \leq p_n} C_5 c_n |g(t_{\alpha(\beta+1)}|x_j) - g(t_{\alpha\beta}|x_j)| + V_n^d \Delta_n^{-1/2} d_n^{1/2},
\]  

(A.16)

where $C_5$ is defined as in (A.7) and the second term of the above expression equals

\[
4C_5 \frac{C_{m3}(x_j)C_{L,d_n}}{p_n} \leq C_{m4} \Delta_n^{-1/2} d_n^{1/2},
\]

$C_{m4} = 8C_5C_L$. Now, put $T_n = M_\Delta(f_n^{-1} d_n^{-1/2})^{-1}$. 

21
Define $M_{\lambda}$ for $2 < \lambda \leq \infty$ as in the proof of Proposition 3.1, $H_{\text{stn}}(x_j, x)$ by

$$
n^{-1} \sum_{i=1}^{n} (\gamma_i(Z_i, \Delta_i|x_j) - \gamma_s(Z_i, \Delta_i|x_j)) I(1(\gamma_i(Z_i, \Delta_i|x_j) - \gamma_s(Z_i, \Delta_i|x_j)) \geq T_n) I(X_i \leq x),
$$

and $M_{13\text{stn}}(x_j, x, z)$ by substitution of $H_{\text{stn}}$ for $G_{\text{stn}}$ and $E[H_{\text{stn}}]$ for $G_{st}$. (A.16) is then majorized by

$$
2c_n^{-1} \max_{x_j \in I_X} \max_{-p_n \leq \beta, \zeta \leq p_n} \max_{1 \leq \alpha \leq m_j-1} \sup_{|z| \leq c_n} M_{13\alpha\zeta\lambda_{\text{stn}}}(x_j, x_j + k, z) \\
+ 2c_n^{-1} f_n d_n^{1/2} (W_n^d + \theta_n^d) + (V_n + C_{m4}) \Delta_n^{-1/2} d_n^{1/2}, \quad (A.17)
$$

where $W_n^d$ and $\theta_n^d$ are defined similarly to $W_n$ and $\theta_n$ in the proof of Proposition 3.1. It is easy to check that $P(2W_n^d > C_{m5}) = o(n^{-2})$ and $2\theta_n^d < C_{m6}$, where $C_{m5}$ and $C_{m6}$ are chosen such that $C_{m5} > 2^{\lambda+2} M_{\lambda}$ and $C_{m6} = 2^{\lambda+1} M_{\lambda}$.

Next, consider

$$
\kappa_{njkr} = x_j + k + \frac{r c_n}{p_n}, \quad \text{for } r = -p_n, -p_n + 1, \ldots, p_n.
$$

Define $M_{ntn\alpha\zeta\lambda_{jkr}}$ by

$$
|H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, \kappa_{njkr}) - H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k) - [E[H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, \kappa_{njkr})] - E[H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k)]]|.
$$

For fixed $j, k, \alpha, \beta, \zeta, n$, $H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k + z) - H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k)$ and $E[H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k + z) - E[H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k)]$ are monotone with respect to $z$ and have finite limits in $x_j + k + c_n$ and $x_j + k - c_n$. Therefore,

$$
\sup_{|z| \leq c_n} M_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k, z) \\
\leq \max_{-p_n \leq \beta, \zeta \leq p_n} M_{ntn\alpha\zeta\lambda_{jkr}} \\
+ \max_{-p_n \leq \beta, \zeta \leq p_n} [E[H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k + \frac{r + 1}{p_n} c_n)] - E[H_{ntn\alpha\zeta\lambda_{jkr}}(x_j, x_j + k + \frac{r c_n}{p_n})]],
$$

where the second term on the right hand side of the above expression is bounded by

$$
4C_L C_5^2 c_n \Delta_n^{-1/2} d_n^{1/2}.
$$

Therefore, (A.17) is majorized by

$$
2c_n^{-1} \max_{x_j \in I_X} \max_{-p_n \leq \beta, \zeta \leq p_n} \max_{1 \leq \alpha \leq m_j-1} \max_{-p_n \leq \beta, \zeta, r \leq p_n} M_{ntn\alpha\zeta\lambda_{jkr}} \\
+(2W_n^d + V_n + C_{m4} + C_{m6} + C_{m7}) \Delta_n^{-1/2} d_n^{1/2}, \quad (A.18)
$$
where $C_{m7} = 8C_L C_5$.

Next,

$$P(\sup_{x \in \mathbb{R}^k} \sup_{t \leq t_n} |M_{10stn}(x)| > C_{m0} \Delta_n^{-1/2} d_n^{1/2})$$  \hfill (A.19)

$$\leq \sum_{j,k,\alpha,\beta,\xi,\eta} P(M_{nt_{\alpha}t_{\beta}jkr} > C_{m0} f_n d_n^{1/2}) + P(W_n^d > C_{m0}) + P(V_n^d > 2C_{m0}),$$

where $C_{m0} = (1/6)(C_{m0} - 16C_5 C_L - 2^{\lambda+1} M_\lambda)$ and $C_{m0}, C_{m2}$ and $C_{m5}$ have to satisfy

$$\max(2^{\lambda+1} M_\lambda, 2 L_n^{1/6}) < C_{m0} = C_{m2}/2 = C_{m5}/2.$$

By Bernstein’s inequality,

$$P(M_{nt_{\alpha}t_{\beta}jkr} > C_{m0} f_n d_n^{1/2}) \leq 2 \exp\left(-\frac{\varphi_{nt_{\alpha}t_{\beta}jkr}}{2nt_{\alpha}t_{\beta}jkr}\right),$$

where

$$\varphi_{nt_{\alpha}t_{\beta}jkr} = \frac{C_{m0}^2 \sigma_n^2 f_n^2 d_n}{2nt_{\alpha}t_{\beta}jkr + \frac{2}{3} n C_{m0} f_n d_n^{1/2} T_n},$$

$$\sigma_{nt_{\alpha}t_{\beta}jkr}^2 = \text{Var}[\Omega_{nt_{\alpha}t_{\beta}jkr}],$$

and

$$\Omega_{nt_{\alpha}t_{\beta}jkr} = (\gamma_{t_{\alpha}}(Z, \Delta|x_j) - \gamma_{t_{\beta}}(Z, \Delta|x_j))$$

$$\times I(\gamma_{t_{\alpha}}(Z, \Delta|x_j) - \gamma_{t_{\beta}}(Z, \Delta|x_j) \leq T_n)(I(X \leq \kappa_{jkr}) - I(X \leq x_{j+k})).$$

Using (A11) (ii), $\sigma_{nt_{\alpha}t_{\beta}jkr}^2 \leq C_{L2} C_5 c_n d_n$, and (A10) (iii), $T_n \Delta_n^{-1/2} \leq d_n^{1/2}$. Therefore,

$$\varphi_{nt_{\alpha}t_{\beta}jkr} \geq C_{m0}^\prime \log n,$$

with

$$C_{m0}^\prime = \frac{C_{m0}^2}{2(C_{L2} C_5 + \frac{1}{3} C_{m0})}.$$  

Finally,

$$\sum_{j,k,\alpha,\beta,\xi,\eta} P(M_{nt_{\alpha}t_{\beta}jkr} > C_{m0} f_n d_n^{1/2}) \leq 2 \sum_{j=1}^{L_n^\alpha} \sum_{k=-1}^{m_j-1} \sum_{\alpha=1}^{\sum_{-p_{n} \leq \beta, \xi, \eta \leq p_{n}}^{} \sum_{n^{-C_{m0}^\prime}}} \leq 2 \sum_{j=1}^{L_n^\alpha} \sum_{k=-1}^{m_j-1} \sum_{\alpha=1}^{\sum_{-p_{n} \leq \beta, \xi, \eta \leq p_{n}}^{} \sum_{n^{-C_{m0}^\prime}}^{}},$$

for which the highest order term on the right hand side is

$$96 \frac{L X_{M_\lambda} \Delta_n^2 c_n^{-1} n^{-C_{m0}^\prime}}{C_L} \leq 96 \frac{L X_{M_\lambda} n^2}{C_L (\log n)^2 n^{-C_{m0}^\prime}}.$$

23
Choosing $C_m$ sufficiently large finishes the proof.

**Proof of Theorem 4.3.** Let $b_{nj} = j n^{-3/2} d_n^{-1/2}$ and $m_{nj} = K(j n^{-3/2} d_n^{-1/2}) - K((j + 1) a_n^{-3/2} d_n^{-1/2})$ in (A12). (A12) (i) becomes

$$|(2 \sum_{j=1}^{j_n} m_{nj} b_{nj} - 1)| \leq \int |K_n(u) - K(u)| du \leq C n^{-3/2} d_n^{-1/2};$$

for some $C > 0$. (A12) (ii) and (A12) (iii) are easily satisfied using $j_n = O(a_n^{-3/2} d_n^{1/2})$ such that $c$ can then be chosen between 0 and 1. Next, let denote $d_{stn}(x, K)$, $d_{stn}(x)$ using kernel $K$. It is clear that

$$\sup_{x \in R_x} \sup_{|t-s| \leq a_n, s,t \in I} |d_{stn}(x, K) - d_{stn}(x, K_n)|$$

$$\leq \sup_{x \in R_x} \sup_{|t-s| \leq a_n, s,t \in I} \left( \frac{D a_n^{-3/2} d_n^{-1/2}}{n a_n} \sum_{i=1}^{n} \Gamma_{ts}(Z_i, \Delta_i |x)|I(x - a_n < X_i \leq x + a_n)\right)$$

$$= O(a_n^{-3/2} d_n^{1/2}) \text{ a.s.,}$$

for which we use Proposition 4.1 with $c_n = a_n$ (for a kernel support equal to $[-1, 1]$).

**References**


