Kernel estimation of multivariate cumulative distribution function

Rong Liu*, Lijian Yang

October, 2006

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824

Abstract
A smooth kernel estimator is proposed for multivariate cumulative distribution function, extending the work on Yamato (1973) on univariate distribution function estimation. Under assumptions of strict stationarity and geometrically strong mixing, we establish that the proposed estimator follows the same pointwise asymptotically normal limiting distribution of the empirical cdf, while the new estimator is a smooth instead of a step function as the empirical cdf. We also show that under stronger assumptions the smooth kernel estimator converges to the true cdf uniformly almost surely at a rate of \( (n^{-1/2} \log n) \). Simulated examples are provided to illustrate the theoretical properties. Using the smooth estimator, survival curves for US GDP growth are estimated conditional on the unemployment growth rate to examine how GDP growth rate depends on the unemployment policy. Another example of gold and silver price returns is given.

Keywords: bandwidth; Berry-Esseen bound; GDP; gold price return; kernel; rate of convergence; silver price return; strongly mixing; survival function; unemployment rate

Short Running Title. Kernel Distribution Estimation

1. Introduction
Let \( \{X_i = (X_{i1}, ..., X_{id})^T\}^n_{i=1} \) be a geometrically \( \alpha \)-mixing and strictly stationary sequence of \( d \)-dimensional variables, with a common probability density function \( f \in C^{(p+1)}(R^d) \) and cumulative distribution function \( F \in C^{(p+d+1)}(R^d) \), in which \( p \) is an odd integer. Traditionally, \( F \) is estimated by the empirical cumulative distribution function \( \hat{F}(x) = n^{-1} \sum_{i=1}^n I\{X_i \leq x\} \), whose theoretical properties have been well known. One drawback of \( \hat{F} \) is that it is a step function even when the true cdf \( F \) is smooth.

*Corresponding author.
E-mail address: liurong@stt.msu.edu (R. Liu).
Yamato (1973) proposed to estimate \( F \) by integrating a kernel density estimator of the density \( f \). To be precise, define the following kernel estimator of \( F \)

\[
\hat{F}(x) = \hat{F}_n(x) = \int_{-\infty}^{x} \hat{f}(u) \, du = n^{-1} \sum_{i=1}^{n} \int_{-\infty}^{x} K_h(X_i - u) \, du
\]

(1.1)

where \( \hat{f}(u) \) is the standard \( d \)-dimensional kernel density estimator (kde) of \( f(u) \) (see Bickel and Rosenblatt, 1973)

\[
\hat{f}(u) = n^{-1} \sum_{i=1}^{n} K_h(X_i - u), \quad K_h(u) = \prod_{\alpha=1}^{d} \frac{1}{h_\alpha} K\left( \frac{u_\alpha}{h_\alpha} \right), \quad u = (u_1, \ldots, u_d)^T
\]

in which \( h = (h_1, \ldots, h_d)^T \) are positive numbers depending on the sample size \( n \), called bandwidths. Throughout this paper, we denote

\[
h_{\text{max}} = \max (h_1, \ldots, h_d), \quad h_{\text{prod}} = h_1 \times \cdots \times h_d
\]

Theoretical properties of \( \hat{F}(x) \) as an estimator of the unknown true distribution function \( F(x) \) have been investigated by several authors for the case of \( d = 1 \) and under i.i.d assumptions, see for example Yamato (1973), Reiss (1981), Falk (1983) and more recently Cheng and Peng (2002). In this paper, we examine the pointwise asymptotic distribution of \( \hat{F}(x) \) and its global convergence rate for arbitrary dimension \( d \). Under some mild conditions (see Assumptions A1 to A4 in Section 2), we have proved that the smooth estimator \( \hat{F}(x) \) behaves asymptotically the same as the step function estimator \( \hat{F}(x) \). As our approach makes use of probability tools for strongly mixing sequences, our results are valid under weak dependence rather than independence.

The rest of the paper is organized as the following. In Section 2, we give Theorems 1, 2, the main theoretical results. In Section 3, we present Monte Carlo evidence that corroborates with the theory and two real data examples. The first real data example illustrates the stochastic dependence of GDP growth rate on unemployment growth rate in the US economy. Second example shows that gold and silver are substitute goods and their prices are strongly associated. All technical proofs are in the Appendix.

2. Asymptotic results

We list below some assumptions necessary for proving Theorems 1 and 2. Throughout this appendix, we denote by the same letters \( c, C \) etc., any positive constants, without distinction in each case.

(A1) The cumulative distribution function \( F \in C^{(p+d+1)}(R^d) \), in which \( p \) is an odd integer, while all \( (p+d+1) \)-th partial derivatives of \( F \) belong to \( L_1(R^d) \) and \( \max_{x \in R^d} |f(x)| \leq C \).

(A2) There exist positive constants \( K_0 \) and \( \lambda_0 \) such that \( \alpha(k) \leq K_0 \exp(-\lambda_0 k) \) holds for all \( k \), where the \( k \)-th order strong mixing coefficient of the strictly stationary process \( \{X_s\}_{s=-\infty}^{\infty} \) is defined as

\[
\alpha(k) = \sup_{B \in \sigma(X_{s,s \leq t}), C \in \sigma(X_{s,s \geq t+k})} |P(B \cap C) - P(B)P(C)|, \quad k \geq 1.
\]
(A3) As $n \to \infty$, $nh_{\text{prod}} \to \infty$, $n^{1/2}h_{\text{prod}}/(\log n)^{1/2} + n^{1/2}h_{\text{max}}^{p+1} \to 0$.

(A4) The univariate kernel function $K(\cdot)$ is $(p+1)$-th order, supported on $[-1,1]$ and Lipschitz continuous.

Assumptions (A1) to (A4) are all typical conditions in time series smoothing literature, see Bosq (1998) Chapter 2 for similar or even stronger assumptions.

**Theorem 1** Under Assumptions (A1)-(A4), as $n \to \infty$

$\sqrt{nV^{-1}(x)} \left( \hat{F}(x) - F(x) \right) \to N(0,1)$

where

$V(x) = \sum_{l=-\infty}^{\infty} \gamma(l), \gamma(l) = EI\{X_i \leq x\} I\{X_{i+l} \leq x\} - EI\{X_i \leq x\}^2.$

Theorem 1 shows that the smooth estimator $\hat{F}(x)$ has the same asymptotically the same as the step function estimator $\hat{F}(x)$. In particular, for iid process $\{X_s\}_{s=-\infty}^{\infty}$, the asymptotic variance function $V(x)$ reduces to the more familiar form of $\gamma(0) = F(x)\{1 - F(x)\}$.

**Theorem 2** Under Assumptions (A1)-(A4), as $n \to \infty$

$\sup_{x \in \mathbb{R}^d} \left| \hat{F}(x) - F(x) \right| = O(n^{-1/2} \log n)$ a.s.

while, for i.i.d. $X_1, \ldots, X_n$

$\sup_{x \in \mathbb{R}^d} \left| \hat{F}(x) - F(x) \right| = O\left(n^{-1/2}(\log n)^{1/2}\right)$ a.s.

We are unaware of any published results on the strong uniform rate of convergence for smooth estimation of multivariate distribution function based on strongly mixing data, as in Theorem 2. Since Assumptions (A1) to (A4) are mild, this strong theoretical result holds for most multiple time series data with continuous distributions. In the next section, we will present Monte Carlo evidence for Theorem 2, and illustrate the use of the smooth estimator $\hat{F}(x)$ with two real data examples.

### 3. Examples

#### 3.1 A simulated example

In the section, we examine the asymptotic result of Theorem 2 via simulation. The data are generated from the following distribution

$X = (X_1, X_2) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}\right)$
We use i.i.d. samples of size \( n = 50, 100, 200, 500 \). The number of replications in the simulation is set to be 100. For our bandwidth vector \( h \), we have taken the optimal bandwidths \( h_{\text{opt},n} \) according to Fan and Yao (2003), p. 200 and rescaled by a factor of \( n^{-1/3+1/5} \), i.e.

\[
h = n^{-1/3+1/5}h_{\text{opt},n} = \left(8\sqrt{\pi}/3\right)^{1/5} \times 2.0362 \times \hat{\sigma} n^{-1/3}
\]

where \( \hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2)^T \), \( \hat{\sigma}_i = \text{sd} (X_i) \) for \( i = 1, 2 \). Of interest in relation to Theorem 2 is the global maximal deviation

\[
D_n = \sup_{x \in \mathbb{R}^d} \left| \hat{F}(x) - F(x) \right|
\]

which is computed for each sample.

In Figure 1, the probability density function of \( D_n = \sup_{x \in \mathbb{R}^d} \left| \hat{F}(x) - F(x) \right| \) is estimated by kernel smoothing based on the 100 replications and plotted for \( n = 50, 100, 200, 500 \). Clearly the empirical distribution of \( D_n \) quickly collapses to 0, as sample size increases, conforming to Theorem 2.

(Insert Figure 1 about here)

We have put in Table 1 the sample mean \( \bar{D}_n \) and sample standard deviation \( \text{sd}(D_n) \) of \( D_n \) as well as \( (\log n/n)^{1/2} \bar{D}_n \), all computed over the 100 replications for each fixed sample size. Since in this example, \( X_1, \cdots, X_n \) are i.i.d., so \( (\log n/n)^{1/2} \bar{D}_n \) should be bounded according to Theorem 2. As one examines Table 1, all the \( (\log n/n)^{1/2} \bar{D}_n \) values change very little, which further corroborates with Theorem 2.

(Insert Table 1 about here)

Lastly we have plotted in Figure 2 the 3-D graphs of \( F \) and \( \hat{F} \) for sample size \( n = 50, n = 200, n = 500 \). It is easy to see that the estimation is quite satisfactory for all sample sizes while it becomes much better when sample size becomes larger. Based on the above observations, we believe our kernel estimator of multivariate cdf is a convenient and reliable tool.

(Insert Figure 2 about here)

### 3.2 GDP growth and unemployment

In this section, we discuss in detail the dependence of US GDP quarterly growth rate on unemployment rate. There are three types of unemployment: frictional, structural, and cyclical. Economists regard frictional and structural unemployment as essentially unavoidable in dynamic economy, so full employment is something less than 100% employment. The full-employment rate of unemployment is also referred to as the natural rate of unemployment. It does not mean the economy will always operate at the natural rate. The economy sometimes operates at an unemployment rate higher than the natural rate due to cyclical unemployment. In contrast, the economy may on some occasions achieve an unemployment rate below the natural rate. For example, during World War II, when the natural rate was about 4% and actual rate below 2% during 1943-1945. It is caused by the pressure of wartime production resulted in an almost unlimited demand for labor. The natural rate is not forever fixed. It was about 4% in the 1960s, and economists generally agreed that the natural rate was about 6%. Today, the consensus is that the rate is about 5.5%.
GDP gap denotes the amount by which actual GDP falls short of the theoretical GDP under the natural rate. Okun’s law, based on recent estimate, indicates that for every 1% which the actual unemployment rate exceeds the natural rate, a GDP gap of about 2% occurs. See Samuelson (1995), p.559 or McConnell and Brue (1999), p.214 for more details. In other words, if unemployment rate falls, then GDP growth rate increases. But unemployment rate can not keep falling because it moves around the natural rate. So it is useful to find the relationship between the GDP growth rate and unemployment growth rate.

Let \( X_{1t} = \) the seasonally adjusted quarterly unemployment growth rate in quarter \( t \), \( X_{2t} = \) the quarterly GDP growth rate in quarter \( t \), all data taken from the 1-st quarter of 1948 (\( t = 1 \)) to the 2-nd quarter of 2006 (\( t = 234 \)). Since both data have been seasonally adjusted, it is reasonable to treat \( X_{t} = (X_{1t}, X_{2t})^{T} \), \( t = 1, \ldots, 234 \) as a strictly stationary time series, which is shown in the timeplots. ACF plots also indicate that the assumption of \( \alpha \)-mixing is satisfied. The plots are shown in Figure 3.

For any fixed interval \( I = [a, b] \), the survival function of \( X_{2t} \) conditional on \( X_{1t} \in I \) is defined as

\[
S_I(x_1) = P(X_{2t} > x_1 | X_{1t} \in I) = \frac{P(X_{2t} > x_1, X_{1t} \in I)}{P(X_{1t} \in I)} = 1 - \frac{F(b, x_1) - F(a, x_1)}{F(b, +\infty) - F(a, +\infty)} \quad (3.2)
\]

in which \( F \) is the joint distribution function of \( X_{1t} \) and \( X_{2t} \).

The function \( S_I(x_1) \) can be approximated by the following plug-in estimator

\[
\hat{S}_I(x_1) = 1 - \frac{\hat{F}(b, x_1) - \hat{F}(a, x_1)}{\hat{F}(b, +\infty) - \hat{F}(a, +\infty)} \quad (3.3)
\]

in which \( \hat{F} \) is the kernel estimator of \( F \) defined in (1.1). According to Theorems 1 and 2, for any fixed \( x_1 \)

\[
|\hat{S}_I(x_1) - S_I(x_1)| = O_p(n^{-1/2})
\]

while

\[
\sup_{x_1 \in \mathbb{R}}|\hat{S}_I(x_1) - S_I(x_1)| = O(n^{-1/2} \log n) \quad a.s.
\]

so the estimator \( \hat{S}_I(x_1) \) is theoretically very reliable. We therefore draw probabilistic conclusions based on the smooth estimate \( \hat{S}_I(x_1) \) instead of the true \( S_I(x_1) \).

In Figure 4, the estimated conditional survival curve \( \hat{S}_I(x_1) \) is plotted for intervals \( I = [-0.08, -0.04], I = [-0.02, 0.02], I = [0.04, 0.08] \). Clearly, when the unemployment growth rate is between \(-0.08\) and \(-0.04\), the chance to have the GDP growth rate higher than 1.5% is the greatest, which is about 0.2. This is in accordance with the Okun’s law that the growth in GDP is the associated with the unemployment rate. So if policymakers want to achieve high GDP growth rate, they should better find ways to lower the unemployment rate. One can even estimate the probabilities of GDP growth rates given the policy of unemployment, which is the interval \( I \). If current unemployment rate is close to the natural rate, then the \( I \) is an interval close to 0, such as \([-0.02, 0.02]\); if the current unemployment...
rate is much higher than the natural rate, then the \( I \) is an negative interval, i.e., trying to lower the unemployment rate.

(Insert Figure 4 about here)

On the other hand, the survival function of \( X_{1t} \) conditional on \( X_{2t} \) can be computed similarly. If certain level of GDP growth rate is planned to be achieved, one can estimate the conditional probabilities of different unemployment growth rates.

### 3.3 Gold and silver price returns

In this section, we discuss in detail the dependence of price returns of gold on silver. Let \( X_{1t} = \) the monthly gold price return in quarter \( t \), \( X_{2t} = \) the monthly silver price return in quarter \( t \), all data taken from the February of 1996 (\( t = 1 \)) to the August of 2006 (\( t = 127 \)). Since both data have been seasonally adjusted, it is reasonable to treat \( X_t = (X_{1t}, X_{2t})^T \), \( t = 1, ..., 127 \) as a strictly stationary time series, which is shown in the timeplots. ACF plots also indicate that the assumption of \( \alpha \)-mixing is satisfied. The plots are shown in Figure 5.

(Insert Figure 5 about here)

Similar to the previous example, for any fixed interval \( I = [a, b] \), survival function \( S_I(x_1) \) and its estimate \( \hat{S}_I(x_1) \) are defined as in (3.2) and (3.3) respectively. We again base our inference on the estimated function \( \hat{S}_I(x_1) \).

In Figure 6, the estimated conditional survival curve \( \hat{S}_I(x_1) \) is plotted for intervals \( I = [-0.10, -0.06] \), \( I = [-0.02, 0.02] \), \( I = [0.06, 0.10] \). Clearly, when the silver price return is higher, the gold price increases faster. This is in accordance with the economic theory of substitute goods, i.e., increase in the price of one good causes increases of demand of other substitutes, hence the increases of the prices of substitutes. So gold and silver are clearly substitutes to each other. See Samuelson (1995), p81 for more details.

(Insert Figure 6 about here)

On the other hand, the survival function of \( X_{1t} \) conditional on \( X_{2t} \) can be computed similarly. That is the conditional probability of silver price return based on gold price return.

### Acknowledgements

This research has been supported in part by NSF award DMS 0405330.

### Appendix

#### A.1 Preliminaries

In the proofs that follow, we use \( U \) and \( u \) to denote sequences of random variables that are uniformly \( O \) and \( o \) of certain order.

**Lemma A.1** [Sunklodas (1984), Theorem 1] Let \( \{\xi_i\}_{i=1}^n \) be an \( \alpha \)-mixing sequence with \( E\xi_n = 0 \). Denote \( d_\delta := \max_{1 \leq i \leq n} \left\{ E|\xi_i|^{2+\delta} \right\}, 0 < \delta \leq 1 \), \( S_n = \sum_{i=1}^n \xi_i \), \( \sigma_n^2 := ES_n^2 \geq c_0 n \) for some
Lemma A.3

If \( \alpha(n) \leq K_0 \exp(-\lambda_0 n) \), \( \lambda_0 > 0 \), \( K_0 > 0 \), then there exist \( c_1 = c_1(K_0, \delta) \), \( c_2 = c_2(K_0, \delta) \), such that

\[
\Delta_n = \sup_{z} \left| P\left\{ \sigma_n^{-1} S_n < z \right\} - \Phi(z) \right| \leq c_1 \frac{d_4}{c_0 \sigma_n^{\delta}} \left\{ \log \left( \frac{\sigma_n / c_0^{1/2}}{\lambda} \right) \right\}^{1 + \delta}
\]

(A.1)

for any \( \lambda \) with \( \lambda_1 \leq \lambda \leq \lambda_2 \), where

\[ \lambda_1 = c_2 \left\{ \log \left( \frac{\sigma_n / c_0^{1/2}}{\lambda} \right) \right\}^b / n, b > 2 (1 + \delta) / \delta; \lambda_2 = 4 (2 + \delta) \delta^{-1} \log \left( \frac{\sigma_n / c_0^{1/2}}{\lambda} \right). \]

Lemma A.2 (Bernstein’s inequality, Bosq, 1998, Theorem 1.4). Let \( \{\xi_i\} \) be a zero mean real valued process, \( S_n = \sum_{i=1}^{n} \xi_i \). Suppose that there exists \( c > 0 \) such that for \( i = 1, \ldots, n \), \( k \geq 3, E |\xi_i|^k \leq c k^2 k! E \xi_i^2 < + \infty, m_r = \max_{1 \leq i \leq N} \| \xi_i \|_r, r \geq 2 \). Then for each \( n > 1 \), integer \( q \in [1, n/2] \), each \( \varepsilon > 0 \) and \( k \geq 3 \)

\[
P\left\{ \sum_{i=1}^{n} \xi_i > n \varepsilon n \right\} \leq a_1 \exp \left( -\frac{q \varepsilon_n^2}{25 m_r^2 + 5 c \varepsilon_n} \right) + a_2 (k) \alpha \left( \left[ \frac{n}{q + 1} \right] \right) \frac{2k}{2k+1}
\]

where

\[ a_1 = \frac{2 n}{q} + 2 \left( 1 + \frac{\varepsilon_n^2}{25 m_r^2 + 5 c \varepsilon_n} \right), a_2 (k) = 11 n \left( 1 + \frac{5 m_k^{2k/(2k+1)}}{\varepsilon_n} \right) \]

A.2 Proof of Theorem 1

Lemma A.3

\[ E \left\{ \hat{F}(x) \right\} = F(x) + U(h_{\max}^p) = F(x) + u(n^{-1/2}) \]

Proof. Using the integral form of Taylor expansion

\[
f(u + hv) = f(u) + \sum_{r=1}^{p} \frac{1}{r!} \left( h_1 v_1 \frac{\partial}{\partial u_1} + \cdots + h_d v_d \frac{\partial}{\partial u_d} \right)^r f(u) + R_{p+1}
\]

in which we denote \( hv = (h_1 v_1, \ldots, h_d v_d)^T \), and where

\[ R_{p+1} = R_{p+1}(u, hv) = \int_0^1 \left\{ \frac{1}{p!} \left( h_1 v_1 \frac{\partial}{\partial u_1} + \cdots + h_d v_d \frac{\partial}{\partial u_d} \right)^{p+1} f(u + thv) \right\} dt. \]

Hence

\[
E \left\{ \hat{F}(x) \right\} = E \int_{-\infty}^{x} K_h (X_i - u) du = \int_{\mathbb{R}^d} \int_{-\infty}^{x} K_h (z - u) du f(z) dz
\]

\[ = \int_{-\infty}^{x} du \int_{[-1,1]^d} f(u + hv) K(v) dv \]
\begin{align*}
= & \int_{-\infty}^{x} du \int_{[-1,1]^d} \left[ f(u) + \sum_{r=1}^{p} \frac{1}{r!} \left( h_1 v_1 \frac{\partial}{\partial u_1} + \cdots + h_d v_d \frac{\partial}{\partial u_d} \right)^r f(u) + R_{p+1} \right] K(v) dv \\
= & \int_{-\infty}^{x} f(u) du + \int_{-\infty}^{x} du \int_{[-1,1]^d} \left[ \sum_{r=1}^{p} \frac{1}{r!} \left( h_1 v_1 \frac{\partial}{\partial u_1} + \cdots + h_d v_d \frac{\partial}{\partial u_d} \right)^r f(u) + R_{p+1} \right] K(v) dv \\
= & \int_{-\infty}^{x} f(u) du + \int_{-\infty}^{x} du \int_{[-1,1]^d} R_{p+1} K(v) dv \\
= & F(x) + \int_{-\infty}^{x} du \int_{[-1,1]^d} \left[ \int_{0}^{1} \left\{ \frac{1}{r!} \left( h_1 v_1 \frac{\partial}{\partial u_1} + \cdots + h_d v_d \frac{\partial}{\partial u_d} \right)^{p+1} f(u+tv) \right\} dt \right] K(v) dv \\
= & F(x) + U \left( h_{\text{max}}^{p+1} \right) = F(x) + u \left( n^{-1/2} \right),
\end{align*}

where the last two steps make use of Assumptions (A1) and (A3) sequentially. \( \square \)

Lemma A.4

\[ E \left\{ \int_{-\infty}^{x} K_h (X_i - u) du \int_{-\infty}^{x} K_h (X_j - u) du \right\} = EI \{ X_i \leq x \} I \{ X_j \leq x \} + U \left( h_{\text{max}} \right) \]

Proof.

\[ E \left\{ \int_{-\infty}^{x} K_h (X_i - u) du \int_{-\infty}^{x} K_h (X_j - u) du \right\} = E \left( \int_{-\infty}^{x} K_h (u_i - X_i) du \int_{-\infty}^{x} K_h (u_j - X_j) dv \right) \]

\[ = \int_{-\infty}^{x} f(v_i) f(v_j | v_i) dv_i dv_j \int_{-\infty}^{x} K_h (u_i - v_i) du_i \int_{-\infty}^{x} K_h (u_j - v_j) du_j \]

\[ = \int_{-\infty}^{x} f_{i,j}(v_i, v_j) dv_i dv_j \int_{-\infty}^{\frac{x-v_i}{h}} K(w_i) dw_i \int_{-\infty}^{\frac{x-v_j}{h}} K(w_j) dw_j \]

\[ = \int_{-\infty}^{x} f_{i,j}(v_i, v_j) dv_i dv_j \hat{K} \left( \frac{x-v_i}{h} \right) \hat{K} \left( \frac{x-v_j}{h} \right), \text{ here } \hat{K} (x) = \int_{-1}^{x} K(u) du, x \geq -1 \]

\[ = \int_{-\infty}^{x} f_{i,j}(v_i, v_j) dv_i dv_j \hat{K} \left( \frac{x-v_i}{h} \right) \hat{K} \left( \frac{x-v_j}{h} \right) \]

\[ = \int_{-1}^{x} f_{i,j}(x-hw_i, x-hw_j) \hat{K} \left( \frac{x-v_i}{h} \right) \hat{K} \left( \frac{x-v_j}{h} \right) (h_1 \cdots h_d)^2 dw_i \cdots dw_d dw_j \cdots dw_{jd} \]
\[ \leq \sum_{i=1}^{d} \sum_{k=1}^{d} \int_{w_i \in [-1,1], w_j \in [-1,1]} f_{i,j}(x - hw_i, x - hw_j) \tilde{K}(w_i) \tilde{K}(w_j) (h_1 \cdots h_d)^2 \, dw_i \, dw_j \]

\[ + \int_{w_{i1}>1, w_{j1}>1} f_{i,j}(x - hw_i, x - hw_j) (h_1 \cdots h_d)^2 \, dw_i \, dw_j \]

\[ = \sum_{i=1}^{d} \sum_{k=1}^{d} \int_{-1}^{1} f(x_{i1} - h_1 w_{i1}) \tilde{K}(w_{i1}) h_1 \, dw_{i1} \]

\[ \int_{-1}^{\infty} \cdots \int_{-1}^{\infty} f_{i,j}(x - hw_i, x - hw_j|x_{i1} - h_1 w_{i1}) h_2 \cdots h_d h_1 \cdots dw_{i1} \cdots dw_{jd} \]

\[ + \int_{-\infty}^{x-h} \int_{-\infty}^{x-h} f_{i,j}(u_i, u_j) \, du_i \, du_j \]

\[ = \sum_{i=1}^{d} \sum_{k=1}^{d} \left\{ C \left( \int_{-1}^{1} f(x_i) - h_i w_i f'(x_i) + o(h_i) \right) \tilde{K}(w_i) h_i \, dw_{i1} \right\} + EI \{ X_i \leq x \} I \{ X_j \leq x \} + U(h_{\text{max}}) \]

\[ = EI \{ X_i \leq x \} I \{ X_j \leq x \} + U(h_{\text{max}}) \]

and

\[ E \left\{ \int_{-\infty}^{x} K_h(x_i - u) \, du \int_{-\infty}^{x} K_h(x_j - u) \, du \right\} \geq EI \{ X_i \leq x - h \} I \{ X_j \leq x - h \} \]

So

\[ E \left\{ \int_{-\infty}^{x} K_h(x_i - u) \, du \int_{-\infty}^{x} K_h(x_j - u) \, du \right\} = EI \{ X_i \leq x \} I \{ X_j \leq x \} + U(h_{\text{max}}) \]

\[ \square \]

**Proof of Theorem 1.**

Let

\[ \xi_{i,n}(x) = \int_{-\infty}^{x} K_h(x_i - u) \, du - E \left\{ \int_{-\infty}^{x} K_h(x_i - u) \, du \right\} \]

\[ \xi_i(x) = EI \{ X_i \leq x \} - EI \{ X_i \leq x \} \]

\[ S_n = S_n(x) = \sum_{i=1}^{n} \xi_{i,n}(x) = \sum_{i=1}^{n} \left\{ \int_{-\infty}^{x} K_h(x_i - u) \, du - E \left\{ \int_{-\infty}^{x} K_h(x_i - u) \, du \right\} \right\} \]

We have \( ES_n = 0 \).

\[ S_n = n \sum_{i=1}^{n} \left[ \frac{1}{n} \int_{-\infty}^{x} K_h(x_i - u) \, du - E \left\{ \frac{1}{n} \int_{-\infty}^{x} K_h(x_i - u) \, du \right\} \right] \]
= n \left[ \hat{F}(x) - \left\{ F(x) + u(n^{-1/2}) \right\} \right]

Let

$$\gamma(l) = \gamma(l, x) = EI \{X_1 \leq x\} I \{X_{1+t} \leq x\} - EI \{X_1 \leq x\}^2$$

Then

$$\text{cov} \left( \xi_{i,n}, \xi_{j,n} \right) = E \xi_{i,n} \xi_{j,n} - E \xi_{i,n} E \xi_{j,n}$$

$$= \left[ E \left\{ \int_{-\infty}^{\infty} K_h(x - u) \, du \int_{-\infty}^{\infty} K_h(x - u) \, du \right\} - \left( E \int_{-\infty}^{\infty} K_h(x - u) \, du \right)^2 \right]$$

$$= \left[ EI \{X_i \leq x\} I \{X_j \leq x\} + U(h_{\text{max}}) - [EI \{X_i \leq x\} + U(h_{\text{max}})]^2 \right]$$

$$= \left[ EI \{X_i \leq x\} I \{X_j \leq x\} - [EI \{X_i \leq x\}]^2 \right] + U(h_{\text{max}})$$

$$= \left[ EI \{X_1 \leq x\} I \{X_{1+t} \leq x\} - [EI \{X_1 \leq x\}]^2 \right] + U(h_{\text{max}})$$

$$= \gamma(l) + U(h_{\text{max}})$$

according to Lemmas A.3 and A.4.

$$\sigma_n^2 = ES_n^2 = \text{var} \left( S_n \right) = \text{var} \left( \sum_{i=1}^{n} \xi_{i,n} \right) = \sum_{i,j=1}^{n} \text{cov} \left( \xi_{i,n}, \xi_{j,n} \right)$$

$$= n \sum_{|h| < n} \left( 1 - \frac{|h|}{n} \right) \left\{ \gamma(l) + U(h_{\text{max}}) \right\} + n \sum_{|h| < n} \text{cov} \left( \xi_{i,n}, \xi_{(1+l),n} \right)$$

$$= n A_n + n B_n$$

$$\gamma(l) = |EI \{X_1 \leq x\} I \{X_{1+h} \leq x\} - EI \{X_{1+h} \leq x\}|$$

$$= |P(\omega : X_1(\omega) \leq x) \cap \omega : X_{1+h}(\omega) \leq x) - P(\omega : X_1(\omega) \leq x) P(\omega : X_{1+h}(\omega) \leq x)|$$

$$\leq \alpha(h) \leq K_0 e^{-\lambda_0 h}$$

$$|\text{cov} \left( \xi_{1,n}, \xi_{(1+l),n} \right)| \leq 4 \|\xi_{1,n}\|_\infty \|\xi_{(1+l),n}\|_\infty \alpha(h) \leq 4 K_0 \exp(-\lambda_0 h)$$

$$A_n = \sum_{|h| \leq c \log n} \left( 1 - \frac{|h|}{n} \right) \left\{ \gamma(l) + U(h_{\text{max}}) \right\}$$

$$= \sum_{|h| \leq c \log n} \left( 1 - \frac{|h|}{n} \right) \gamma(l) + \sum_{|h| \leq c \log n} \left( 1 - \frac{|h|}{n} \right) U(h_{\text{max}})$$

$$\to \sum_{|h| \leq c \log n} \gamma(l) + U(h_{\text{max}} \log n) \to \sum_{l=-\infty}^{\infty} \gamma(l)$$
\[
\sum_{l=-\infty}^{\infty} |\gamma(l)| \leq \gamma(0) + 2 \sum_{l=1}^{\infty} K_0 \exp(-\lambda_0 l) < \infty
\]

\[
\lim_{n \to \infty} n \text{var}(\overline{\xi}_n) = \lim_{n \to \infty} \sum_{|l| \leq n} \left( 1 - \frac{|l|}{n} \right) \gamma(l) = \sum_{|l| \leq n} \gamma(l)
\]

So \[\sum_{|l| \leq n} \gamma(l) > 0\]

\[
B = \sum_{c \log n < |l| < n} \text{cov}(\xi_{1,n}, \xi_{(1+l),n}) \leq \sum_{|h| > c \log n} \left( 1 - \frac{|l|}{n} \right) 4K_0K_0 \exp(-\lambda_0 l)
\]

For \(c \geq 2/\lambda_0\),\]

\[
B \leq \frac{4K_0 e^{-\lambda_0 c \log n}}{1 - e^{-\lambda_0}} = \frac{K_0 n^{-c \lambda_0}}{1 - e^{-\lambda_0}} \leq C_1 n^{-2}
\]

So \[\sigma_n^2 = nA + nB \rightarrow n \sum_{l=-\infty}^{\infty} \gamma(l) \geq c_0 n \text{ when } n \text{ is large}\]

then by (A.1) in Lemma A.1,

\[
\Delta_n = \sup_{z} |P\{\sigma_n^{-1} S_n < z\} - \Phi(z)| \leq c_1 \frac{d}{c_0 \sigma_n^3} \left\{ \log \left( \frac{\sigma_n}{c_0^{1/2}} \right) / \lambda \right\}^{1+\delta}
\]

Let \[\delta = 1, \lambda = 4 (2 + \delta) \delta^{-1} \log \left( \frac{\sigma_n}{c_0^{1/2}} \right) = 12 \log \left( \frac{\sigma_n}{c_0^{1/2}} \right), d = 1\]

\[
\Delta_n \leq \frac{c_1}{c_0 \sigma_n} 12^{-2} = \frac{c}{\sigma_n} = O \left( n^{-1/2} \right)
\]

\[
\frac{S_n}{\sigma_n} \rightarrow N(0, 1)
\]

\[
n \left( \hat{F}(x) - F(x) \right) \sigma_n + u \left( n^{1/2} \right) \rightarrow N(0, 1)
\]

By Slutsky’s theorem,

\[
\frac{\sqrt{n} \left( \hat{F}(x) - F(x) \right)}{\sqrt{V(x)}} \rightarrow N(0, 1)
\]
A.3 Proof of Theorem 2

**Lemma A.5** Denote \( g_{m_1, \ldots, m_d} = (a_{1,m_1}, \ldots, a_{d,m_d}) \in \mathbb{R}^d \), \( 1 \leq m_i \leq M_i \) and

\[
A_n = \max_{1 \leq m_i \leq M_i} \left| \hat{F} (g_{m_1, \ldots, m_d}) - E \{ \hat{F} (g_{m_1, \ldots, m_d}) \} \right|
\]

\[
B_n = \max_{1 \leq m_i \leq M_i} \left| \hat{F} (g_{m_1, \ldots, m_d}) - F (g_{m_1, \ldots, m_d}) \right|
\]

If \( \max (M_1, \ldots, M_d) \leq Cn \), then

\[
A_n + B_n = O \left( n^{-1/2} \log n \right) \text{ a.s.}
\]

while, for i.i.d. \( X_1, \ldots, X_n \)

\[
A_n + B_n = O \left( n^{-1/2} (\log n)^{1/2} \right) \text{ a.s.}
\]

**Proof.** Let

\[
\xi_{in} = \xi_{in,m_1,\ldots,m_d} = \xi_{i,n} (g_{m_1,\ldots,m_d}) = \int_{-\infty}^{g_{m_1,\ldots,m_d}} K_h (X_i - u) \, du - E \left\{ \int_{-\infty}^{g_{m_1,\ldots,m_d}} K_h (X_i - u) \, du \right\}
\]

We have \( E\xi_{in} = 0 \).

\[
\hat{F} (g_{m_1,\ldots,m_d}) - E \hat{F} (g_{m_1,\ldots,m_d}) = n^{-1} \sum_{i=1}^{n} \xi_{in}
\]

\[
E(\xi_{in}^2) = E \left( \int_{-\infty}^{g_{m_1,\ldots,m_d}} K_h (X_i - u) \, du - E \left\{ \int_{-\infty}^{g_{m_1,\ldots,m_d}} K_h (X_i - u) \, du \right\} \right)^2 \leq 1
\]

\[
E \left( |\xi_{in}|^k \right) = E \left( |\xi_{in}|^{k-2} \xi_{in}^2 \right)
\]

\[
= E \left[ \int_{-\infty}^{g_{m_1,\ldots,m_d}} K_h (X_i - u) \, du - E \left\{ \int_{-\infty}^{g_{m_1,\ldots,m_d}} K_h (X_i - u) \, du \right\} \right]^{k-2} \xi_{in}^2
\]

\[
\leq 1^{k-2} E(\xi_{in}^2), \; k \geq 2
\]

By Lemma A.2 (Bernstein’s inequality),

\[
P \left\{ \left| \sum_{i=1}^{n} \xi_{in} \right| > n\varepsilon_n \right\} \leq a_1 \exp \left( - \frac{q\varepsilon_n^2}{25m_2^2 + 5q\varepsilon_n} \right) + a_2 (k) \alpha \left( \left[ \frac{n}{q + 1} \right] \right)^{\frac{2k}{2k+1}}
\]

\[
a_1 = \frac{2n}{q} + 2 \left( 1 + \frac{\varepsilon_n^2}{25m_2^2 + 5q\varepsilon_n} \right), a_2 (k) = 11n \left( 1 + \frac{5m_k^{2k/(2k+1)}}{\varepsilon_n} \right)
\]

12
Let \( k = 3, a_2 (3) = 11 \), \( m_2^2 = \frac{b_{11}}{\varepsilon_n} \leq 1, \varepsilon_n = \frac{\log n}{\sqrt{n}} \)

\[
P \left\{ \left| \sum_{i=1}^{n} \xi_{in} \right| > n \varepsilon_n \right\} \leq a_1 \exp \left( -\frac{q^2 \varepsilon_n^2}{25m_2^2 + 5c \varepsilon_n} \right) + a_2 (3) \alpha \left( \frac{n}{q+1} \right)^{6/7}
\]

\[
m_3 = \max_{1 \leq i \leq N} \| \xi_i \|_3 \leq \left\{ E (\xi_i^3) \right\}^{1/3} \leq 1
\]

take \( q \) such that \( \left[ \frac{n}{q+1} \right] \geq c_0 \log n \), \( q \geq \frac{c_1 n}{\log n} \)

\[
\frac{q^2 \varepsilon_n^2}{25m_2^2 + 5c \varepsilon_n} = \frac{q (\log n)^2}{25m_2^2 + 5c \varepsilon_n} = \frac{c_1 n}{\log n} \frac{a_2 (\log n)^2}{25} = \frac{c_2 a^2 \log n}{25}
\]

\[
a_1 = 2 \frac{n}{q} + 2 \left( 1 + \frac{\varepsilon_n^2}{25m_2^2 + 5c \varepsilon_n} \right) = O (\log n)
\]

\[
a_2 (3) = 11n \left( 1 + \frac{5}{\varepsilon_n} \right) \leq 11n \left( 1 + \frac{5}{an^{-1/2} \log n} \right) \leq 11n \left( 1 + \frac{5}{a \log n} \right) = O (n)
\]

\[
\alpha \left( \frac{n}{q+1} \right)^{6/7} \leq \left( K_0 e^{-\lambda_0} \left[ \frac{n}{q+1} \right] \right)^{6/7} \leq C n^{-6 \lambda_0 c_0 / 7}
\]

\[
P \left\{ \left| \sum_{i=1}^{n} \xi_{in} \right| > n \varepsilon_n \right\} \leq O (\log n) \exp \left( -c_2 a^2 \log n \right) + C n^{1 - 6 \lambda_0 c_0 / 7}
\]

\[
\leq C n^{-(d+2)}
\]

for \( c_0, c_2 \) large enough.

\[
P \left\{ \max_{1 \leq m_i \leq M_i} n^{-1} \left| \sum_{i=1}^{n} \xi_{in,m_1,\ldots,m_d} \right| > an^{-1/2} \log n \right\}
\]

\[
\leq \sum_{m_1=1}^{M_1} \cdots \sum_{m_d=1}^{M_d} P \left\{ n^{-1} \left| \sum_{i=1}^{n} \xi_{in,m_1,\ldots,m_d} \right| > an^{-1/2} \log n \right\}
\]

\[
\leq C n^{-(d+2)} M_1 \cdots M_d \leq C n^{-2}
\]

Borel-Cantelli lemma implies that

\[
A_n = O \left( n^{-1/2} \log n \right) \text { a.s.}
\]

\[
B_n = \max_{1 \leq m_i \leq M_i} \left| \hat{F} (g_{m_1,\ldots,m_d}) - F (g_{m_1,\ldots,m_d}) \right|
\]
\[
\begin{align*}
&\leq \max_{1 \leq m_i \leq M_i} \left| \hat{F}(g_{m_1, \ldots, m_d}) - E \left\{ \hat{F}(g_{m_1, \ldots, m_d}) \right\} \right| + \max_{1 \leq m_i \leq M_i} \left| E \left\{ \hat{F}(g_{m_1, \ldots, m_d}) \right\} - F(g_{m_1, \ldots, m_d}) \right| \\
&= A_n + U \left( n^{-1/2} \right) \quad \text{by } A_n = O \left( n^{-\frac{1}{2}} \log n \right) \ a.s. \\
&= O \left( n^{-1/2} \log n \right) \ a.s.
\end{align*}
\]

If \( X_1, \ldots, X_n \) are i.i.d., we can get
\[
A_n + B_n = O \left( n^{-1/2} \left( \log n \right)^{1/2} \right) \ a.s.
\]
by using same steps above but applying Bernstein’s inequality of i.i.d. case. \( \square \)

**Lemma A.6** \( \forall A \subset \mathbb{R}^d, \)
\[
\int_A |K_h(v - u)| \, du \leq \int_{\mathbb{R}^d} |K_h(v - u)| \, du \leq \|K\|_{L^1}^d.
\]

**Proof.**
\[
\int_A |K_h(v - u)| \, du \leq \int_{\mathbb{R}^d} |K_h(v - u)| \, du \\
= \int_{\mathbb{R}^d} \prod_{\alpha=1}^d \frac{1}{h_\alpha} |K \left( \frac{v_\alpha - u_\alpha}{h_\alpha} \right)| \, du = \int_{\mathbb{R}^d} \prod_{\alpha=1}^d \frac{1}{h_\alpha} K \left( \frac{u_\alpha}{h_\alpha} \right) \, du \\
= \prod_{\alpha=1}^d \int_{-1}^1 |K(w_\alpha)| \, dw_\alpha \leq \|K\|_{L^1}^d
\]
\( \square \)

**Lemma A.7** Let \( -\infty = a_{1,1} < \cdots < a_{1,N_1} = \infty, -\infty = a_{2,1} < \cdots < a_{2,N_2} = \infty, -\infty = a_{d,1} < \cdots < a_{d,N_d} = \infty \) such that \( P(a_{i,k} \leq X_i \leq a_{i,k+1}) \leq 1/n \) for \( k = 1, \ldots, N_1 \) and \( i = 1, \ldots, d. \) Denote \( g_{n_1, \ldots, n_d} = (a_{1,n_1}, \ldots, a_{d,n_d}) \in \mathbb{R}^d. \) If \( \max(N_1, \ldots, N_d) \leq Cn, \) then
\[
n^{-1} \sum_{i=1}^n E \int_{g_{n_1, \ldots, n_d}} |K_h(X_i - u)| \, du = u \left( n^{-1/2} \left( \log n \right)^{1/2} \right) \ a.s.
\]

**Proof.**
\[
n^{-1} \sum_{i=1}^n \int_{g_{n_1, \ldots, n_d}} |K_h(X_i - u)| \, du \leq n^{-1} \sum_{i=1}^n \int_{g_{n_1, \ldots, n_d}} \int_{g_{n_1+1, \ldots, n_d+1}} |K_h(v - u)| \, du F(v) \\
= \int_{g_{n_1, \ldots, n_d}} \int_{g_{n_1+1, \ldots, n_d+1}} \int_{g_{n_1, \ldots, n_d}} |K_h(v - u)| \, du
\]

14
\[ \leq C \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) \]

according to Lemma A.6.

\[ \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) = \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) - \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) + \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) \]

\[ = \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) \]

So

\[ \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) = \int_{g_{n_1, \ldots, n_d - (h_1, \ldots, h_d)}}^{g_{n_1, \ldots, n_d + (h_1, \ldots, h_d)}} dF(v) + P(g_{n_1, \ldots, n_d} \leq X_i \leq g_{n_1, \ldots, n_d + 1}) \]

\[ \leq \left( \int_{a_{n_1} - h_1}^{a_{n_1} + h_1} \int_{a_{n_1} + 1}^{a_{n_1} + h_1} \int_{a_1}^{a_{n_1} + h_1} dF(v) \right) \]

Expanding

\[ \left( \int_{a_{n_1} - h_1}^{a_{n_1} + h_1} \int_{a_1}^{a_{n_1} + h_1} dF(v) \right) \]

For the \( 3^d - 2^d \) terms with \( \int_{a_{n_1} + h_1}^{a_{n_1} + h_1} \), is of order \( O(n^{-1}) \). For the other \( 2^d \) terms which do not have \( \int_{a_{n_1} + h_1}^{a_{n_1} + h_1} \), each one \( \leq h_{\text{prod max}} \left| f(x) \right| \).

So

\[ \frac{1}{n} \sum_{i=1}^{n} E \int_{g_{n_1, \ldots, n_d}}^{x} \left| K_h(X_i - u) \right| du \]

\[ \leq Ch_{\text{prod max}} \left| f(x) \right| + C \left( 3^d - 2^d \right) / n \]

\[ = u \left( n^{-1/2} \left( \log n \right)^{1/2} \right) a.s. \]

according to Assumptions (A1) and (A3). \( \square \)

**Lemma A.8** Under of same conditions of Lemma A.7,

\[ \frac{1}{n} \sum_{i=1}^{n} \int_{g_{n_1, \ldots, n_d}}^{x} \left| K_h(X_i - u) \right| du - E \int_{g_{n_1, \ldots, n_d}}^{x} \left| K_h(X_i - u) \right| du = O \left( n^{-1/2} \log n \right) a.s. \]

while, for i.i.d. \( X_1, \ldots, X_n \)

\[ \frac{1}{n} \sum_{i=1}^{n} \int_{g_{n_1, \ldots, n_d}}^{x} \left| K_h(X_i - u) \right| du - E \int_{g_{n_1, \ldots, n_d}}^{x} \left| K_h(X_i - u) \right| du = O \left( n^{-1/2} \left( \log n \right)^{1/2} \right) a.s. \]
Proof. Let

\[ \xi_{in} = \xi_{in,n_1,\ldots,n_d} = \xi_{i,n} (g_{n_1,\ldots,n_d}) \]

\[ = \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du - E \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \]

We have \( E \xi_{in} = 0 \).

\[ E(\xi_{in}^2) = E \left( \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du - E \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \right)^2 \]

\[ = E \left( \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \right)^2 - \left( E \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \right)^2 \]

\[ \leq E \left( \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \right)^2 \]

\[ \leq C^2 \int_{g_{n_1,\ldots,n_d} + (h_{1,\ldots,h_d})}^{g_{n_1,\ldots,n_d}} dF (v) \leq C^2 / n \]

\[ E \left( |\xi_{in}|^k \right) = E \left( |\xi_{in}|^{k-2} \xi_{in}^2 \right) \]

\[ = E \left( \left| \int_{-\infty}^{g_{n_1,\ldots,n_d}} |K_h (X_i - u)| \, du - E \left( \int_{-\infty}^{g_{n_1,\ldots,n_d}} |K_h (X_i - u)| \, du \right) \right|^{k-2} \xi_{in}^2 \right) \]

\[ \leq (2C)^{k-2} / n E(\xi_{in}^2), \ k \geq 2 \]

\[ m_3 = \max_{1 \leq i \leq N} \| \xi_i \|_3 \leq \{ E(\xi_{in}^3) \}^{1/3} \leq C / n^{1/3} \]

Then

\[ \frac{1}{n} \sum_{i=1}^{n} \left| \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du - E \left( \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \right) \right| = O \left( n^{-1/2} \log n \right) \text{ a.s.} \]

can be proven similarly as the proof of Lemma A.5 by applying Lemma A.2 (Bernstein’s inequality). If \( X_1, \ldots, X_n \) are i.i.d., we can get

\[ \frac{1}{n} \sum_{i=1}^{n} \left| \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du - E \left( \int_{g_{n_1,\ldots,n_d}}^{x} |K_h (X_i - u)| \, du \right) \right| = O \left( n^{-1/2} (\log n)^{1/2} \right) \text{ a.s.} \]

by using same steps above but applying Bernstein’s inequality of i.i.d. case. □

Proof of Theorem 2.
Under the same conditions of Lemma A.7. We have

\[ \max_{1 \leq n \leq N} \left| \hat{F} (g_{n_1,\ldots,n_d}) - F (g_{n_1,\ldots,n_d}) \right| = O \left( n^{-1/2} \log n \right) \text{ a.s.} \]
by Lemma A.5.
For $\forall x = (x_1, \ldots, x_d) \neq g_{n_1, \ldots, n_d}$, then $\exists n_1, \ldots, n_d$ such that $F(g_{n_1, \ldots, n_d}) \leq F(x) \leq F(g_{n_1+1, \ldots, n_d+1})$.

$$
\left| \hat{F}(x) - \hat{F}(g_{n_1, \ldots, n_d}) \right| = \left| \frac{1}{n} \sum_{i=1}^{n} \int_{-\infty}^{x} K_h(x_i - u) \, du - \frac{1}{n} \sum_{i=1}^{n} \int_{g_{n_1, \ldots, n_d}}^{x} K_h(x_i - u) \, du \right|
= \left| \frac{1}{n} \sum_{i=1}^{n} \int_{g_{n_1, \ldots, n_d}}^{x} K_h(x_i - u) \, du \right| \leq \frac{1}{n} \sum_{i=1}^{n} \int_{g_{n_1, \ldots, n_d}}^{x} |K_h(x_i - u)| \, du
= \frac{1}{n} \sum_{i=1}^{n} \left\{ \int_{g_{n_1, \ldots, n_d}}^{x} |K_h(x_i - u)| \, du - E \int_{g_{n_1, \ldots, n_d}}^{x} |K_h(x_i - u)| \, du \right\}
+ \frac{1}{n} \sum_{i=1}^{n} E \int_{g_{n_1, \ldots, n_d}}^{x} |K_h(x_i - u)| \, du
= O\left(n^{-1/2} \log n\right) \text{ a.s.}
$$

according to Lemmas A.7 and A.8.

\[
\left| \hat{F}(x) - F(x) \right| = \left| \hat{F}(x) - \hat{F}(g_{n_1, \ldots, n_d}) + \hat{F}(g_{n_1, \ldots, n_d}) - F(g_{n_1, \ldots, n_d}) + F(g_{n_1, \ldots, n_d}) - F(x) \right|
\leq \left| \hat{F}(x) - \hat{F}(g_{n_1, \ldots, n_d}) \right| + \left| \hat{F}(g_{n_1, \ldots, n_d}) - F(g_{n_1, \ldots, n_d}) \right| + \left| F(g_{n_1, \ldots, n_d}) - F(x) \right|
= O\left(n^{-1/2} \log n\right) + O\left(n^{-1/2} \log n\right) + O\left(1/n\right) \text{ a.s.}
\]

according to Lemma A.5.
So
$$
\sup_{x \in \mathbb{R}^d} \left| \hat{F}(x) - F(x) \right| = O\left(n^{-1/2} \log n\right) \text{ a.s.}
$$
If $X_1, \ldots, X_n$ are i.i.d, we can prove

$$
\sup_{x \in \mathbb{R}^d} \left| \hat{F}(x) - F(x) \right| = O\left(n^{-1/2} (\log n)^{1/2}\right) \text{ a.s.}
$$

by using same steps as above. \(\square\)

References


Table 1: Mean and standard deviation of $D_n$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$D_n$</th>
<th>$(\log n/n)^{1/2} D_n$</th>
<th>sd($D_n$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.1018</td>
<td>0.3640</td>
<td>0.0405</td>
</tr>
<tr>
<td>100</td>
<td>0.0775</td>
<td>0.3614</td>
<td>0.0214</td>
</tr>
<tr>
<td>200</td>
<td>0.0595</td>
<td>0.3658</td>
<td>0.0184</td>
</tr>
<tr>
<td>500</td>
<td>0.0380</td>
<td>0.3411</td>
<td>0.0112</td>
</tr>
</tbody>
</table>
Figure 1: Estimation of maximal deviation densities: $n = 50$, solid line; $n = 100$, dot-dash line, $n = 200$, dashed line; $n = 500$, dotted line.
Figure 2: Plots of $F(x)$ and of $\hat{F}(x)$ in the simulated example: (a) Plot of $F(x)$; (b) Plot of $\hat{F}$ for $n = 50$; (c) Plot of $\hat{F}(x)$ for $n = 200$; (d) Plot of $\hat{F}(x)$ for $n = 500$. 

Figure 3: ACF plots and timeplots of GDP and unemployment quarterly growth rates: (a) ACF plot of GDP quarterly growth rate; (b) ACF plot of unemployment quarterly growth rate; (c) Timeplot of GDP quarterly growth rate; (d) Timeplot of unemployment quarterly growth rate.
Figure 4: Survival curves of GDP growth rate conditional on unemployment growth rate: $X_{1t} \in [-0.08, -0.04]$, crosses; $X_{1t} \in [-0.02, 0.02]$, solid; $X_{1t} \in [0.04, 0.08]$, dash.
Figure 5: ACF plots and timeplots of gold and silver price monthly returns: (a) ACF plot of gold price return; (b) ACF plot of silver price return; (c) Timeplot of gold price return; (d) Timeplot of silver price return.
Figure 6: Survival curves of gold price return conditional on silver price return: $X_{1t} \in [-0.10, -0.06]$, crosses; $X_{1t} \in [-0.02, 0.02]$, solid; $X_{1t} \in [0.06, 0.10]$, dash.