UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Cambridge Working Papers in Economics

Cambridge Working Papers in Economics: 1760

UNIFORM CONVERGENCE OF SMOOTHED DISTRIBUTION FUNCTIONS WITH AN APPLICATION TO DELTA METHOD FOR THE LORENZ CURVE

Shin Kanaya Debopam Bhattacharya

11 December 2017

In this note, we present some theoretical results useful for inference on a population Lorenz curve for income and expenditure distributions, when the population density of the distribution is not (uniformly) bounded away from zero, and potentially has thick tails. Our approach is to define Hadamard differentiability in a slightly nonstandard way, and using it to establish a functional delta method for the Lorenz map. Our differentiability concept is nonstandard in that the perturbation functions, which are used to compute the functional derivative, are assumed to satisfy certain limit conditions. These perturbation functions correspond to a (nonparametric) distribution function estimator. Therefore, as long as the employed estimator satis.es the same limit conditions, which we verify in this paper, the delta method and corresponding asymptotic distribution results can be established.

Uniform Convergence of Smoothed Distribution Functions with an Application to Delta Method for the Lorenz Curve*

Shin Kanaya†
University of Aarhus and CREATES

Debopam Bhattacharya‡
University of Cambridge

First draft: July, 2017.
This draft: December 11, 2017

1 Introduction

In this note, we present some theoretical results useful for inference on a population Lorenz curve for income and expenditure distributions, when the population density of the distribution is not (uniformly) bounded away from zero, and potentially has thick tails. Our approach is to define Hadamard differentiability in a slightly nonstandard way, and using it to establish a functional delta method for the Lorenz map. Our differentiability concept is nonstandard in that the perturbation functions, which are used to compute the functional derivative, are assumed to satisfy certain limit conditions. These perturbation functions correspond to a (nonparametric) distribution function estimator. Therefore, as long as the employed estimator satisfies the same limit conditions, which we verify in this paper, the delta method and corresponding asymptotic distribution results can be established.

For verifying these limit conditions, we establish novel (simultaneous) uniform convergence results for a (smoothed) distribution function estimator and its corresponding density estimator using a so-called boundary kernel. We allow the left end point of the support to be bounded, which is a natural feature of income and consumption distributions, e.g. due

*Bhattacharya acknowledges financial support from the European Research Council. Kanaya acknowledges financial support from the Danish Council for Independent Research, Social Sciences (Grant No. DFF - 4182-00279).

†Department of Economics, University of Aarhus, Fuglesangs Alle 4, Aarhus V, 8210, Denmark. E-mail: skanaya@econ.au.dk.

‡Faculty of Economics, University of Cambridge. E-mail: db692@cam.ac.uk.

to presence of minimum wages in the case of income or due to minimum expenditure required for sustenance in the case of consumption. If the support has a finite boundary point, nonparametric estimators with a usual symmetric kernel incur the so-called boundary bias, resulting in inconsistency. The use of a boundary kernel allows us to overcome the boundary bias problem and establish uniform convergence of our nonparametric estimators.

The above uniform convergence results are used here as intermediate steps to establish the delta method for the Lorenz curve. But they are interesting in their own right and can be used in other contexts. For instance, consider semi/nonparametric problems that involve multiple estimation steps as in the case of partially-linear regression, nonparametric regression with generated regressors, and matching estimators, c.f. Andrews (1994), and Hahn and Ridder (2012). In these problems, it is often the case that the entire functional form of preliminary kernel based estimators determine final semi/nonparametric estimators, and thus their uniform convergence is required to establish the asymptotic distributional results of the final estimators. When the support of relevant variables has a finite end point, one has to employ some boundary correction method; otherwise, the uniform convergence of the (preliminary) kernel estimators fails, resulting in inconsistency of the final estimators. Indeed, these issues seem to have been ignored in some previous papers on semiparametric problems. For the case with no boundary bias problem, various useful uniform results of kernel estimators have been developed in the literature (e.g., Hansen, 2008), but the corresponding results for boundary corrected kernel estimators have not been formally investigated; our new uniform results fill this gap. Our application of the above results to the Lorenz curve also extends and modifies a previous claim in Bhattacharya, (2007 Claim 1), by resolving some technical issues that arise in establishing differentiability when the density of the income distribution declines to zero with thick tails, e.g. a Pareto distribution. Finally, Fang and Santos (2015) have generalized the Delta method under directional Hadamard differentiability, which is significantly weaker than standard Hadamard differentiability, in that it does not require linearity of the functional derivative. While the notion of functional differentiability we consider here is also weaker than standard Hadamard differentiability, it is different from their generalization; the functional derivative of the Lorenz curve in our case does indeed satisfy linearity. In independent work, Kaji (2017) has derived Hadamard differentiability results for general L-statistics, which provide an alternative way to establish an asymptotic distribution theory for Lorenz curves.

The rest of this paper is organized as follows. In the next section, we discuss our weaker notion of the functional differentiability for the Lorenz curve and verify its differentiability under a set of high-level conditions. Section 3 introduces a new kernel based nonparametric estimator, and investigates its uniform convergence properties, which are also of independent interest. Finally, we verify that our new estimator indeed satisfies the high level conditions for the differentiability and discuss the functional delta method for the Lorenz curve. All proofs can be found in the Appendix.

2 Functional differentiability of the Lorenz curve

In this section, we introduce our basic setup and establish differentiability of the Lorenz curve. To this end, let $F(\cdot)$ be the cumulative distribution function (CDF) of a random variable X with support $[l, \infty)$, where the left end point l is supposed to be finite and known to researchers, e.g. the minimum wage in case of income. To fix ideas, we will suppose from now on that X is individual income. We also set $l = 0$ for notational economy.

Let α be a mapping from a $[0,1]$-valued, non-decreasing, continuous function $F(\cdot)$ on $[0, \infty)$ to a continuous function $\alpha(F)(\cdot)$ on $[0,1]$ defined as

$$\alpha(F)(p) := \int_0^p F^{-1}(u)\, du, \tag{1}$$

where $F^{-1}(u) := \inf_{x \in [0,\infty)} \{F(x) \geq u\}$. We show that this mapping α is Hadamard differentiable in a restricted sense that $[\alpha(F_t) - \alpha(F)]/t$ has a well defined limit when a class of perturbed functions, $\{F_t(\cdot)\}_{t \in (0,1]}$, satisfies certain limit conditions as $t \to 0$ (presented as **C.2** below), where $F_t(\cdot)$ is defined as $F_t(\cdot) := F(\cdot) + th_t(\cdot)$ through a class of functions, $\{h_t(\cdot)\}_{t \in (0,1]}$, and $F(\cdot)$ is the limit of $F_t(\cdot)$ at which the derivative of α is computed.

The usual definition of the Hadarmard differentiability (e.g., Section 20.2 of van der Vaart, 1998) requires that the convergence of $[\alpha(F_t) - \alpha(F)]/t$ take place for *every* $\{h_t(\cdot)\}$ converging to $h(\cdot)$. In contrast, our strategy to show the differentiability of α is to impose some restrictions on limit behavior of $\{h_t(\cdot)\}$ or equivalently that of $\{F_t(\cdot)\}$. This is innocuous for our eventual purpose of applying the functional Delta method. We consider a particular nonparametric estimator of the cumulative distribution function (CDF) $F(\cdot)$ in the next section, which corresponds to the class of $\{F_t(\cdot)\}$, where the index t corresponds to the sample size n in that $t \to 0$ is interpreted as $1/\sqrt{n} \to 0$. It turns out that our estimator satisfies such limit behavior (**C.2**), as shown in the next section. Therefore, it is sufficient to establish Hadamard differentiability in the restricted sense.

We impose the following conditions on the CDF $F(\cdot)$ at which the derivative of α is defined and on the set of functions $\{F_t(\cdot)\} = \{F_t(\cdot)\}_{t \in (0,1]}$:

C.1 Let $F(\cdot)$ be a cumulative distribution function (CDF): $[0, \infty) \to [0, 1)$, and it satisfies the following properties: i) $F(\cdot)$ is continuously differentiable with its probability density function, $f(x) = (d/dx) F(x)$, satisfying

$$\sup_{x \in [0,\infty)} f(x) < \infty, \text{ and } f(x) > 0 \text{ for each } x \in [0,\infty).$$

ii) There exists some non-increasing function $g(\cdot) : [0,\infty) \to (0,\infty)$ such that $f(x) \geq g(x) (> 0)$ and

$$\int_0^\infty [1 - F(x)] [f(x) / g(x)] dx < \infty.$$

C.2 i) For each $t \in (0,1]$, $F_t(\cdot)$ is a CDF on $[0,\infty)$ that is continuously differentiable and has the derivative $f_t(x)$ at each $x \in [0,\infty)$ ($f(0)$, the derivative of $F(x)$ at $x = 0$, is interpreted as a one-side, right derivative).

ii) For the function $g(\cdot)$ introduced in **C.1**, let

$$Q_t(u) := \frac{1-u}{g\left(F_t^{-1}(u)\right)},$$

where $F_t^{-1}(u) := \inf_{x \in [0,\infty)} \{F_t(x) \geq u\}$. For some (sufficiently small) $\eta > 0$, $\{Q_t(\cdot)\}_{t \in (0,\eta]}$ is uniform integrable (with respect to the Lebesgue measure on $[0,1]$).

The monotonicity condition on $g(\cdot)$, i.e. (ii) of **C.1** does not allow for $f(0) = 0$. While this case might also be accommodated by suitably restricting behavior of $f(x)$ at and near $x = 0$, the condition of $f(0) > 0$ does not appear restrictive for income distributions and is maintained throughout the paper, which makes our proof arguments simple and transparent. The integrability condition (ii) of **C.1** implies the existence of the first moment of X since $f(x)/g(x) \geq 1$ and $E[X] = \int_0^\infty [1 - F(x)] dx$.

As for the condition ii) of **C.2**, we say that $\{Q_t(\cdot)\}_{t \in (0,\eta]}$ uniformly integrable if

$$\lim_{\rho \to \infty} \sup_{t \in (0,\eta]} \int_0^1 \mathbf{1}\{Q_t(u) \geq \rho\} Q_t(u) du = 0,$$

which is the standard definition (for the finite-measure case) in the literature. For a fixed $\eta > 0$, **C.2** allows us to define a space/class of functions, say, denoted by $\bar{D}^\eta_{[0,\infty)}$. In view of this, the differentiability in our restricted sense may be regarded as the standard Hadamard differentiability of a mapping from $\bar{D}^\eta_{[0,\infty)}$ to the set of continuous function on $[0,1]$. However, since we may take any arbitrarily small η, we may interpret ii) as a condition on the limit behavior of $Q_t(\cdot)$ as $t \to 0$.

A sufficient condition of the uniform integrability (UI) condition ii) is that for some $\phi > 0$,

$$\limsup_{t \to 0} \int_0^\infty \{[1 - F_t(x)] / g(x)\}^{1+\phi} f_t(x) dx < \infty, \tag{2}$$

which is verified for our specific estimators in Theorem 3 below. The sufficiency follows from the Markov inequality and the fact that

$$\int_0^1 Q_t^{1+\phi}(u)du = \int_0^\infty \{[1 - F_t(x)]/g(x)\}^{1+\phi} f_t(x)\, dx, \tag{3}$$

which follows from change of variables with $u = F_t(x)$ (we refer to, e.g., Equation (1) and Footnote 3 of Falkner and Teschl, 2012 for change-of-variables formulae when $F_t(\cdot)$ is nondecreasing but may not be strictly increasing).[1] The condition of (2), though nonstandard as a condition for differentiability, is satisfied for a kernel based estimator of $F(\cdot)$ under a set of reasonable conditions, as we will see in the next section.

Given these, we can verify the differentiability of α:

Theorem 1 *Let α be the mapping (see 1), $F(\in \bar{D}_{[0,\infty)}) \mapsto \alpha(F)(\in C_{[0,1]})$, where $\bar{D}_{[0,\infty)}$ is the space of cumulative distribution functions on $[0,\infty)$ (equipped with the sup norm) each element of which satisfies the conditions in **C.1**, and $C_{[0,1]}$ is the set of continuous functions on $[0,1]$ (equipped with the $\mathbf{L}_1$ norm). Let D_α be the space of all realized sample paths of a F-Brownian bridge (i.e., $h(F^{-1}(\cdot)) \in D_\alpha$ is a standard Brownian bridge on $[0,1]$).*
Then, α is Hadamard differentiable at F tangentially to D_α in the following sense: there exists a linear functional $h \mapsto \alpha'_F(h) := -\int_0^p \frac{h(z(u))}{f(z(u))}du$ such that

$$\left\| \frac{\alpha(F + th_t) - \alpha(F)}{t} - \alpha'_F(h) \right\|_{\mathbf{L}_1} \to 0 \quad \text{as } t \downarrow 0,$$

*for any $\{h_t(\cdot)\}_{t\in(0,1]}$ with $h_t(\cdot)$ converging uniformly on $[0,\infty)$ to $h(\cdot) \in D_\alpha$ and $F_t(\cdot) = F(\cdot) + th_t(\cdot)$ being an element of the functional class $\{F_t(\cdot)\}_{t\in(0,1]}$ satisfying the conditions in **C.2**.*

3 Construction of a smooth distribution function Estimator

In this section, we construct an estimator that satisfies all the conditions of Theorem 1. In practical inference procedures, the set of the perturbation functions $\{F_t(\cdot)\}$ of the theorem

[1] It in general holds that $F_t^{-1}(F_t(x)) \leq x$, with equality holding when $F_t(x)$ is strictly increasing. However, on the region when $F_t(x)$ is not strictly increasing (i.e., flat), we have $F_t^{-1}(F_t(x)) < x$ but at the same time $f_t(x) = 0$. That is, only the set of x with $F_t^{-1}(F_t(x)) = x$ contribute to the calculation of the integral:

$$\int_0^1 Q_t^{1+\phi}(u)du = \int_0^\infty \left\{ \frac{[1 - F_t(x)]}{g(F_t^{-1}(F_t(x)))} \right\}^{1+\phi} f_t(x)\, dx,$$

leasing to (3).

corresponds to a sequence of estimators of the CDF $F(\cdot)$ (with its index being sample size n). Given a set of i.i.d. observations from the CDF F, $\{X_i\} = \{X_i\}_{i=1}^n$, we present a feasible nonparametric kernel based estimator $\hat{F}(\cdot)$ of $F(\cdot)$. Note that the index $t \in (0,1]$ of $\{F_t(\cdot)\}$ may correspond to the sample size n; in particular, $t \to 0$ is interpreted as $1/\sqrt{n} \to 0$. For consistency of notation, we let $t = 1/\sqrt{n}$ and subsequently interpret

$$F_t(\cdot) = F_{1/\sqrt{n}}(\cdot) = \hat{F}(\cdot). \tag{4}$$

As is evident from Theorem 1 and its conditions (**C.1-C.2**), our verification of the functional differentiability of α requires the existence of the densities of the true F and perturbed F_t (equivalently, the estimator $\hat{F}$). This requires the estimator $\hat{F}(x)$ to be smooth and possess derivative $\hat{f}(x) = (d/dx)\hat{F}(x)$. Therefore, we cannot use the empirical distribution function, $(1/n)\sum_{i=1}^n \mathbf{1}\{X_i \leq x\}$, which is discontinuous by construction. Instead, we consider a kernel-based smooth estimator $\hat{f}(x)$ and define our estimator of the CDF as the following smoothed empirical distribution function (SDF):

$$\hat{F}(x) := \int_0^x \hat{f}(y)\, dy. \tag{5}$$

For the construction of $\hat{f}(\cdot)$, we need to take into account the boundedness of the left-end point of the support of $F(\cdot)$, which is zero. When the support has a finite end point, it is known that the standard (Parzen-Rosenblatt type) kernel density estimator is not uniformly consistent over the entire support. The bias of the usual/standard kernel estimators with a symmetric kernel does not vanish (as $n \to \infty$) at and near the boundary. This is because the support of the kernel function exceeds that of the function to be estimated – the so called boundary bias problem. While the convergence of the density functions $\{f_t(\cdot)\}$ to $f(\cdot)$ is itself not necessarily required in Theorem 1, the inconsistency of the density estimator carries over to that of the CDF estimator, which is defined through integration as in (5), resulting in the violation of the conditions of Theorem 1. In particular, in our proof strategy, it does not appear possible to verify **C.2** without the uniform convergnece property of the density.

To overcome this problem, we introduce a so-called boundary kernel, i.e., a kernel function whose shape changes/adapts according to the location x. This shape adapting property guarantees the uniform convergence of the estimator over the entire support $[0, \infty)$ as shown in Theorem 2. Specifically, for a kernel function $K(u)$ with $\int_{-\infty}^{\infty} K(u)\, du = 1$, whose conditions are provided below, we let $a_0(p) := \int_{-\infty}^{p} K(u)\, du$ and define the following boundary kernel (indexed by x/b_n):

$$K_{x/b_n}(u) := \frac{1}{a_0(x/b_n)} K(u), \tag{6}$$

where $b = b_n > 0$ is a bandwidth/smoothing parameter tending to zero as $n \to \infty$ (the bandwidth is set as a parameter depending on n, selected by a researcher but its dependence on n is suppressed subseqently for notational simplicity). Given this, our first boundary corrected density estimator is defined as

$$\hat{f}_{\mathrm{B}}(x) := \frac{1}{nb_n} \sum_{i=1}^{n} K_{x/b_n}\left(\frac{x - X_i}{b_n}\right), \tag{7}$$

which depends on x and b_n through two routes, $a_0(\cdot)$ and $K(\cdot)$. While this estimator possesses the desirable consistency property uniformly over $x \in [0, \infty)$, it does not integrate to one. Thus, if we define a CDF estimator as

$$\hat{F}_{\mathrm{B}}(x) := \int_0^x \hat{f}_{\mathrm{B}}(y)\, dy, \tag{8}$$

$[1 - \hat{F}_{\mathrm{B}}(x)]$ may not approach zero as $x \to \infty$; thus it is uncertain if we can indeed verify the condition (2) in **C.2** (if we set $F_t(x) = \hat{F}_{\mathrm{B}}(x)$).[2] We therefore consider normalizing $\hat{f}_{\mathrm{B}}(x)$ by $\hat{F}_{\mathrm{B}}(\infty) := \int_0^\infty \hat{f}_{\mathrm{B}}(y)\, dy$. That is, we define our further modified density estimator as

$$\hat{f}(x) := \frac{\hat{f}_{\mathrm{B}}(x)}{\hat{F}_{\mathrm{B}}(\infty)}, \tag{9}$$

Given this (9) and the definition of (5), our CDF estimator can be written as

$$\hat{F}(x) = \frac{1}{\hat{F}_{\mathrm{B}}(\infty)\, nb_n} \sum_{i=1}^{n} \int_0^x K_{y/b_n}\left(\frac{y - X_i}{b_n}\right) dy. \tag{10}$$

Another boundary corrected CDF estimator has also been considered in Tenneiro (2013). In our setup/notation, his estimator may be interpreted as the one corresponding to $\hat{F}_{\mathrm{T}}(x) := \frac{1}{nb_n} \sum_{i=1}^{n} \int_0^x K_{x/b_n}\left(\frac{y-X_i}{b_n}\right) dy$ (note that his estimator is based on a more general boundary kernel). This $\hat{F}_{\mathrm{T}}(x)$ and our $\hat{F}(x)$ are similar but differ in two respects: Firstly, $\hat{F}_{\mathrm{T}}(x)$ may not satisfy $\hat{F}_{\mathrm{T}}(\infty) = 1$ as it does not have a normalization factor (while this can be easily modified); secondly and more importantly, the integration of the kernel function for $\hat{F}_{\mathrm{T}}(x)$ over $[0, x]$ does not concern the 'index' variable x/b_n, while that for our estimator does so. This point may be understood by comparing the summands in $\hat{F}_{\mathrm{T}}(x)$ and ours, say, $\int_0^x K_{x/b_n}\left(\frac{y-X_i}{b_n}\right) dy$ and $\int_0^x K_{y/b_n}\left(\frac{y-X_i}{b_n}\right) dy$. This difference leads to two consequences: 1) $\hat{F}_{\mathrm{T}}(x)$ is easier to compute, which usually has a closed-form expression, but our $\hat{F}(x)$ may not, often requiring numerical integration (even for a usual/simple underlying kernel

[2] We can see this point through the following: $\hat{F}_{\mathrm{B}}(\infty) = \frac{1}{n} \sum_{i=1}^{n} \int_{-X_i/h}^{\infty} \frac{1}{a_0(w + X_i/h)} K(w)\, dw$ (by changing variables); and if if $K(\cdot)$ is symmetric, it holds that $\int_{-X_i/h}^{\infty} \frac{1}{a_0(X_i/h)} K(w)\, dw = 1$ but not $\int_{-X_i/h}^{\infty} \frac{1}{a_0(w+X_i/h)} K(w)\, dw = \int_{-X_i/h}^{\infty} \frac{1}{a_0(X_i/h)} K(w)\, dw$ in general.

function $K(\cdot)$); 2) The derivative of $\hat{F}_{\mathrm{T}}(x)$, $\hat{f}_{\mathrm{T}}(x) = (d/dx)\hat{F}_{\mathrm{T}}(x)$, may not be consistent for $f(x)$, which may hamper the verification of the conditions of Theorem 1. In general, boundary correction requires different boundary kernels for estimating an original function and its derivative (c.f. the local linear method produces an estimator for a target function and one for its derivative but the latter is not defined as the derivative of the former; see also discussions in Section 8 of Jones, 1993). In contrast, our CDF estimator $\hat{F}(x)$ and its derivative $\hat{f}(x) = (d/dx)\hat{F}(x)$ are uniformly consistent, as shown in Theorem 2, which is a natural consequence of our construction of $\hat{F}(x)$ as the integral of the consistent estimator $\hat{f}(x)$. This simultaneous consistency property may be conveniently used in verifying the conditions of Theorem 1 (this task is undertaken in Theorem 3 below).

Note that our simple boundary correction method (i.e., dividing the original kernel $K(u)$ by $a_0(x/b_n)$) may recover the consistency but does not allow for higher order bias correction. That is, the bias of our $\hat{f}(x)$ is at most $O(b_n)$ and inferior to $O\left(b_n^2\right)$, where the latter bias order is attained by kernel estimators with a usual symmetric (second order) kernel function when there is no boundary problem. Several papers have proposed how to construct second or higher order boundary kernels, including Müller (1991), Jones (1993), Müller and Wang (1994), and Zhang and Karunamuni (1998, 2000). While the use of such a sophisticated boundary kernel allows for the bias rate of $O\left(b_n^2\right)$ or faster, it may produce negative estimates of the density $f(x)$ for some x since the second or higher order kernels may take negative values. Negative estimates of the density $f(x)$ produce an estimate of $F(\cdot)$ that is not a proper CDF (i.e., $\neq 1$ at $x = \infty$ or decreasing). While some sort of regularization or normalization may be applied to correct these undesirable features, it may result in a complicated form of the corrected estimator. On the other hand, our estimator $\hat{f}(x)$ is non negative for any $x \in [0, \infty)$ by construction, and $\hat{F}(x)$ is shown to possess all the prerequisite of Theorem 1.

As an alternative to our $\hat{F}(\cdot)$ based on the boundary kernel in (6), we might be able to use a so-called asymmetric kernel and construct a CDF estimator that overcomes the boundary bias problem with a better bias rate (such as $O(b_n^2)$) but without losing the positivity of $\hat{f}(\cdot)$ (we refer to Hirukawa and Sakudo, 2014, and Igarashi and Kakizawa, 2017, for various forms of asymmetric kernels). Regardless of potentially better performances of asymmetric-kernel-based estimators, it may not be straightforward to establish uniform convergence results of such estimators, which do not appear to have been well investigated in the literature. In addition, asymmetric kernel based estimators do not involve a convolution operation, which is unlike boundary kernel based estimators (note that the standard kernel density estimator may be viewed as the convolution of the kernel function and the empirical distribution function).

For establishing the weak convergence result of smoothed CDF estimators, their convolution form appears to play an important role (see our discussions on the weak convergence of (10) in Section 3.2); it is uncertain if the asymmetric kernel based CDF estimator may satisfy the weak convergence property.

3.1 Uniform convergence of the new boundary corrected kernel estimators

In this subsection, we derive uniform convergence rates of our new nonparametric estimator $\hat{F}(\cdot)$ and its derivative $\hat{f}(\cdot)$. The rate results may be effectively used to show that the estimators satisfy all the conditions of Theorem 1, where **C.2** and other conditions of Theorem 1 are interpreted as the ones for $\hat{F}(\cdot)$ through (4). To this end, we introduce some additional conditions on $F(\cdot)$:

Assumption 1 *Let $\{X_i\}_{i=1}^n$ be an i.i.d. sample from the CDF $F(\cdot)$. i) $F(\cdot)$ satisfies the condition i) in **C.1**. and the density $f(\cdot)$ of $F(\cdot)$ is differentiable on $[0,\infty)$ with $\sup_{x\in[0,\infty)}|f'(x)| < \infty$.*
ii) There exist some (sufficiently large) constant $M > 0$ and some positive constant $\delta > 1$ such that

$$\left|f'(x)\right| \leq M_0 [1+x]^{-\delta} \quad \text{for any } x \geq 0,$$

for some $\delta > 1$.

The condition i) of Assumption 1 is fairly standard. The polynomial decaying condition ii) on the derivative $f'(\cdot)$ appears not to be restrictive, which may be effectively used to derive the uniform convergence of $\hat{F}_{\mathrm{B}}(\cdot)$.

We also set out the conditions on the kernel function used to compute $\hat{F}(x)$ and $\hat{f}(x)$:

Assumption 2 *The kernel function $K(\cdot): \mathbb{R} \to \mathbb{R}$ satisfies the following conditions: i) it is of bounded variation with $\sup_{u\in\mathbb{R}} K(u) < \infty$, $\int_{-\infty}^{\infty} K(u)\,du = 1$, and $\int_{-\infty}^{\infty} |uK(u)|\,du < \infty$.*
ii) There exist some constants $M_K > 0$ and $\nu > 2$ such that

$$|K(u)| \leq M_K [1+|u|]^{-\nu}.$$

The condition i) of Assumption 2 is fairly weak, including almost all forms of kernel functions found in the literature. It is sufficient (as a condition on $K(\cdot)$) for establishing the uniform convergence of the first step density estimator $\hat{f}_{\mathrm{B}}(\cdot)$ by using the covering number technique from empirical process theory (developed by Kanaya, 2017). The condition ii) is conveniently used to establish the uniform convergence of the (first step) CDF estimator $\hat{F}_{\mathrm{B}}(\cdot)$. For example, it is trivially satisfied the boundedness of the support of $K(\cdot)$.

We also note that i) - ii) permit $K(\cdot)$ to be discontinuous and/or non-symmetric; this weaker condition is maintained for the sake of generality. While the continuity of kernel functions has been imposed to establish the uniform convergence of kernel based estimators in Hansen (2008) and other work, it is not a necessary condition. The uniform convergence of kernel estimators may be established for discontinuous kernel functions by using the so-called covering number technique (from empirical process theory) developed by Kanaya (2017) (see the proof of Theorem 2 below).

Given Assumptions 1-2l, we are ready to state the uniform convergence result of our nonparametric estimators:

Theorem 2 *Suppose that i) of Assumption 1 and Assumption 2 hold. Then, the density estimator $\hat{f}(x)$ defined in (9) satisfies*

$$\hat{f}(x) - f(x) = O_p(\sqrt{(\log n)/nb_n}) + O_p(b_n), \tag{11}$$

uniformly over $x \in [0, \infty)$, as $n \to \infty$ and $b_n \to 0$ with $(\log n)/nb_n \to 0$. Suppose further that ii) of Assumption 1 holds. Then, the CDF estimator defined in (10) satisfies

$$\hat{F}(x) - F(x) = O_p(\sqrt{(\log n)/n}) + O_p(b_n), \tag{12}$$

uniformly over $x \in [0, \infty)$, as $n \to \infty$ and $b_n \to 0$.

The first part of Theorem 2 establishes the uniform convergence of our boundary corrected kernel density estimator $\hat{f}(\cdot)$. While similar sorts of conjectures/results has been presented in Müller (1991, p. 524), we here provide a formal proof. The second part for the bias corrected CDF estimator $\hat{F}(\cdot)$ appears to be new and is potentially useful in other applications. This theorem effectively establishes the simultaneous uniform convergence of $\hat{F}(\cdot)$ and its derivative $\hat{f}(\cdot)$.

Note that Tenreiro (2013) has also considered smoothed CDF estimation and investigated the uniform convergence of his estimator (Theorem 3.2). However, as discussed above, Tenreiro's result is not applicable to our case as his estimator is different, which does not appear to admit the *simultaneous* convergence.

Assumption 2 on $K(\cdot)$ and our construction of the boundary kernel in 6 do not allow for the second or higher order boundary kernel, resulting in the bias order of $O_p(b_n)$, which however is sharp in our setting. We again emphasize that the use of higher-order kernel is possible but it may produce negative estimates of the density, leading to an estimate of $F(\cdot)$ that can be decreasing over part of the support.

3.2 Verification of Condition C.2

Here, we verify that our boundary corrected CDF estimator $\hat{F}(\cdot)$ and its derivative $\hat{f}(\cdot)$ satisfy Condition C.2, which are interpreted as $F_t(\cdot)$ and $f_t(\cdot)$ through (4). To this end, we impose some additional conditions on the (true) $F(\cdot)$ and the kernel function $K(\cdot)$:

Assumption 3 *Let $F(\cdot)$ be the CDF introduced in **C.1** and $f(\cdot)$ be its probability density. Either of the following conditions is satisfied: a) There exists some (sufficiently large) constants $L, M \geq 1$ such that*

$$(1/M)\, x^{-\tau} \leq f(x) \leq M x^{-\tau} \quad \text{for any } x \geq L, \tag{13}$$

$$(1/M)\, L^{-\tau} \leq \inf_{z \in [0,L]} f(z), \tag{14}$$

for some $\tau > 2$; OR b) There exist some positive constant $M \geq 1$ and some strictly increasing function $m(\cdot) : [0, \infty) \to (0, \infty)$ such that

$$(1/M) \exp\{-m(x)\} \leq f(x) \leq M \exp\{-m(x)\} \quad \text{for any } x \geq 0, \tag{15}$$

and satisfies

$$\lim_{x\to\infty} \frac{\log x}{m(x)} \to 0 \quad \text{and} \quad \liminf_{x\to\infty} \frac{m(x)}{m(x-\bar{c})} = 1, \tag{16}$$

where some constant $\bar{c} > 0$, which may be arbitrarily small.

The bound condition (13) in Part a) of Assumption 3 specifies that the tail of $f(\cdot)$ declines, to some extent, smoothly and monotonically to zero. It appears to be reasonable for income distributions. In particular, Pareto distributions, which are often used to model an income distribution, would satisfy (13). We can also allow for the case when the convergence rates of the lower and upper bounds are not the same, say, $(1/M)\, x^{-\tau_1} \leq f(x) \leq M x^{-\tau_2}$ for $\tau_1 > \tau_2 > 2$, and the result of Theorem 3 may be established when the difference of τ_1 and τ_2 are not too large. For economy, we do not consider this case. An implication of (13) is that

$$1 - F(x) \leq M \int_x^{\infty} z^{-\tau} dz = \frac{M}{\tau - 1} x^{-\tau+1}, \tag{17}$$

which shall be used repeatedly in our subsequent proofs. The condition (14) is not at all restrictive as long as $f(x) > 0$ for any $x \geq 0$; we can always find some sufficiently large L and M satisfying (14).

Given (13)-(14), we can also find $g(\cdot)$, the lower bound function of $f(\cdot)$ introduced in **C.1**, as

$$g(x) := \begin{cases} \inf_{z\in[0,L]} f(z) & \text{for } x \in [0, L), \\ (1/M)\, x^{-\tau} & \text{for } x \geq L, \end{cases} \tag{18}$$

which is nonincreasing on $[0, \infty)$. We subsequently suppose that $g(x)$ takes this form when Part a) of Assumption 3 is satisfied. Using the upper bound of $f(x)$ in (13), that of $1-F(x)$ in (17), and this expression of $g(x)$, we can easily check that the integrability condition ii) of **C.1** is satisfied under $\tau > 2$.

The conditions in Part b) covers the cases when the tail decay speed of $f(\cdot)$ is faster than any polynomial functions, including tails of the log normal, exponential, and normal distributions. The bound condition (15) has implications similar to (13) in Part a). Note that for the polynomial decaying case, we can write $x^{-\tau} = \exp\{-\tau \log x\}$ but this is excluded by the first limit condition in (16). Given this expression of $x^{-\tau}$, we could write Parts a) and b) in a unifying manner. However, for the proof of Theorem 3, we use lower and upper bounds for $1-F(x)$. For Part a), we need to use tight bounds (since the tail decaying speed of $f(\cdot)$ is not fast enough; see the proof of Theorem 3), which can be easily derived under the condition (13) as in (17). In contrast, under the conditions in Part b), we have the faster decaying rate of $f(\cdot)$ and some less tight bounds of $1-F(x)$ are sufficient, which are to be derived below. This is the reason why we consider Parts a) and b) separately.

The second limit condition (16) is satisfied by various functions that grows slower than exponential functions, say, $m(\cdot)$ can be any power of the logarithm, $(\log x)^{p_1}$ with $p_1 > 1$, any polynomial function, x^{p_2} with any $p_2 > 0$, or any product of them. It can be effectively used to establish an upper bound of $f(x-\bar{c})$ in terms of $m(x)$. We can also accommodate the case when $m(x)$ has some exponential or faster growth rate by considering the bounds of $f(x)$ as $(1/M)\exp\{-\exp\{m(x)\}\} \leq f(x) \leq M\exp\{-\exp\{m(x)\}\}$, which allows us to derive some reasonable upper bound of $f(x-\bar{c})$ and lower and upper bounds of $1-F(x)$ under some conditions analogous to (16). However, this case is omitted for brevity.

To establish Theorem 3, we also need to impose some additional conditions on the kernel $K(\cdot)$:

Assumption 4 *It holds that $K(u) \geq 0$ for any $u \in \mathbb{R}$ and $K(u) = 0$ for $|u| > L_K$ with some (sufficiently large) constant $L_K > 0$.*

The positivity of $K(\cdot)$ in Assumption 4 guarantees that $\hat{F}(\cdot)$ is a CDF for any (finite) realization of $\{X_i\}_{i=1}^n$. The boundedness of the support of $K(\cdot)$ is imposed for convenience.[3]

[3] We can also employ some kernel function whose support is the whole real line, say the normal kernel. However, in order to establish Theorem 3 under such a choice of $K(\cdot)$, we must restrict the tail decaying rate of $f(\cdot)$. That is, roughly speaking, the tail decaying rate of $K(\cdot)$ must be faster than that of $f(\cdot)$, meaning that a researcher, to some extent, needs to know the unknown density function's tail decay property. In contrast, if K with bounded support is used, it can allow for any fast tail decaying rate of $f(\cdot)$.

A specific choice of $K(\cdot)$ that satisfies all the conditions in Assumptions 2-4, is the Epanechnikov kernel, $K(u) = (3/4\sqrt{5})(1-u^2/5)$ for $|u| \leq \sqrt{5}$; $= 0$ otherwise.

Now we are ready present a result that our new CDF estimator $\hat{F}(\cdot)$ satisfies the integrability condition corresponding to the one in **C.2**:

Theorem 3 *Suppose that Assumptions 1-4 hold. If it also holds that*

$$b_n = o\left(n^{-1/2}\right) \quad \text{and} \quad 1/b_n = O\left(n^q\right) \quad \text{for some } q \in (0,1), \tag{19}$$

then, there exist some (sufficiently small) constant $\phi > 0$ and some constant $C_\phi \in (0,\infty)$ such that

$$\int_0^\infty \{[1-\hat{F}(x)]/g(x)\}^{1+\phi} \hat{f}(x)\, dx < C_\phi,$$

with probability approaching 1 (as $n \to \infty$), where the expression of C_ϕ is given in the proof.

The first bandwidth condition in (19) requires the bandwidth b_n to be undersmoothing (as the one intended for any of CDF and density estimators). As a result of this, the bias components in (11) and (12) of Theorem 2, corresponding to $O_p(b_n)$ are negligible relatively to the variance-effect components of the order $\sqrt{(\log n)/nb_n}$ and $\sqrt{(\log n)/n}$. The second condition in (19) requires b_n not to be too small, which for example, excludes $b_n = (\log n)^2/n$.

The condition of $b_n = o\left(n^{-1/2}\right)$ has also been used in Yukich (1992), where the weak convergence of smoothed CDF estimators under fairly weak conditions when there is no boundary problem. Yukich's result (Theorem 2.1) also holds in our case, i.e., we have

$$\sqrt{n}[\hat{F} - F] \Rightarrow G, \tag{20}$$

in the space of $l^\infty[0,\infty)$, where $l^\infty[0,\infty)$ is the set of all bounded functions on $[0,\infty)$ and $G := \{G(\cdot)\}_{x\in[0,\infty)}$ is a tight Brownian bridge process whose covariance is given by $F(x)[1-F(x)]$. We can formally prove this result under the conditions of Theorem 3; note in particular that the measure induced by the density $\frac{1}{a_0(x/b_n)b_n}K(u/b_n)$ weakly converges to the Dirac measure at zero as $b_n \to 0$ for any $x \in (0,\infty)$; we also refer to van der Vaart, 1994). This weak convergence result also corresponds to the condition in Theorem 1 that h_t converges to a path of a Brownian bridge.

To conclude this section, we point out that our boundary corrected smoothed CDF estimator $\hat{F}(\cdot)$ in (10) satisfies all the conditions imposed for the functional differentiability result of the mapping α, Theorem 1; thus, given the weak convergence result of (20), we can apply the functional Delta method, which allows us to establish the asymptotic distribution of the estimated α.

References

[1] Andrews, D. (1994) Asymptotics for semiparametric econometrics models via stochastic equicontinuity, *Econometrica* 62, 43-72.

[2] Bhattacharya, D. (2007) Inference on Inequality from Household Survey Data. *Journal of Econometrics* 137, 674-707.

[3] Falkner, N. and Teschl, G. (2012) On the substitution rule for Lebesgue-Stieltjes integrals. Expositiones Mathematicae 30, 412-418.

[4] Fang, Z. and Santos, A. (2015) Inference on directionally differentiabile functions, Working paper, U.C. San Diego.

[5] Hansen, B.E. (2008) Uniform convergence rates for kernel estimation with dependent data, *Econometric Theory* 24, 726-748.

[6] Jones, M.C. (1993) Simple boundary correction for kernel density estimation. *Statistics and Computing* 3, 135-146.

[7] Kaji, T (2017) Switching to the New Norm: From Heuristics to Formal Tests using Integrable Empirical Processes, mimeo. MIT.

[8] Kanaya, S. (2017) Uniform convergence rates of kernel-based nonparametric estimators for continuous time diffusion processes: a damping function approach, *Econometric Theory* 33, 874-914.

[9] Gasser, T., Müller, H.-G. and Mamitzsch, V. (1985) Kernels for nonparametric curve estimation. *Journal of Royal Statistical Society*, Series B 47, 238-252.

[10] Hahn, J. and G. Ridder (2012) The asymptotic variance of semi-parametric estimators with generated regressors. *Econometrica* 81, 315-340.

[11] Hansen, B.E. (2008) Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* 24, 726-748.

[12] Hirukawa, M. and Sakudo, M. (2014) Nonnegative bias reduction methods for density estimation using asymmetric kernels. *Computational Statistics and Data Analysis* 75, 112-123.

[13] Holland, G.B. (1999) *Real analysis*, 2nd ed. John Wiley & Sons Inc.

[14] Igarashi, G. and Kakizawa, Y. (2017) Bias corrections for some asymmetric kernel estimators. *Journal of Statistical Planning and Inference* 159, 37-63.

[15] Messer, K. and Goldstein, L. (1993) A new class of kernels for nonparametric curve estimation. *The Annals of Statistics* 21, 179-195.

[16] Müller, H.-G. (1991) Smooth optimum kernel estimators near endpoints. *Biometrika* 78, 521-530.

[17] Müller, H.-G. and Wang, J.-L. (1994) Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrika* 50, 61-76.

[18] Tenreiro, C. (2013) Boundary kernels for distribution function estimation. *REVSTAT - Statistical Journal* 11, 169-190.

[19] van der Vaart, A.W. (1994) Weak convergence of smoothed empirical processes. *Scandinavian Journal of Statistics* 21, 501-504.

[20] van der Vaart, A.W. and Wellner, J. (1996) *Weak Convergence and Empirical Processes with Applications to Statistics*. Springer.

[21] Yukich, J.E. (1992) Weak convergence of smoothed empirical processes. *Scandinavian Journal of Statistics* 19, 271-279.

[22] Zhang, S. and Karunamuni, R.J. (2000) On nonparametric density estimation at the boundary. *Journal of Nonparametric Statistics* 12, 197-221.

A Appendix

A.1 Proof of Theorem 1

In this section, we prove the functional differentiability of the mapping α:

Proof of Theorem 1. Consider a sequence of differentiable functions $\{h_t(\cdot)\}_{t\geq 0}$ that converges uniformly on $[0,\infty)$ to $h(\cdot)$ as $t\to 0$ with $F_t(\cdot)=F(\cdot)+th_t(\cdot)\in \bar{D}_{[0,\infty)}$ for any $t\in(0,1]$. Define the following two functions $[0,1)\to(0,\infty)$ as

$$z_t(u) := \inf_{x\in[0,\infty)}\{F_t(x)\geq u\} \quad \text{and} \quad z(u) := \inf_{x\in[0,\infty)}\{F(x)\geq u\}.$$

We also write $F_t^{-1}(u)=z_t(u)$ and $F^{-1}(u)=z(u)$, where $F(\cdot)$ is strictly increasing and the latter is the inverse function of $F(\cdot)$ in the usual sense. We shall show that as $t\to 0$,

$$\int_0^1 \left| \frac{\int_0^p z_t(u)\,du - \int_0^p z(u)\,du}{t} - \left(-\int_0^p \frac{h(z(u))}{f(z(u))}du\right)\right| dp \to 0.$$

Now, since $F(\cdot)$ is differentiable, for each $u\in(0,1)$, we can write

$$\begin{aligned} u &= F(z(u)) = F(z_t(u)) + [z(u)-z_t(u)]\,f(\tilde{z}_t(u)) \\ &= F_t(z_t(u)) + [z(u)-z_t(u)]\,f(\tilde{z}_t(u)) - th_t(z_t(u)), \end{aligned}$$

for some $\tilde{z}_t(u)$ lying between $z_t(u)$ and $z(u)$ (by the mean-value theorem), where $f_t(x)$ is the derivative of $F_t(x)$ with respect to x. Since any element in $\bar{D}_{[0,\infty)}$ is continuous and thus $F_t(z_t(u))=u$, we have

$$\frac{z_t(u)-z(u)}{t} = -\frac{h_t(z_t(u))}{f(\tilde{z}_t(u))}. \tag{21}$$

Therefore,

$$\begin{aligned} &\int_0^1 \left| \frac{\int_0^p z_t(u)\,du - \int_0^p z(u)\,du}{t} - \left(-\int_0^p \frac{h(z(u))}{f(z(u))}du\right)\right| dp \\ &\leq \int_0^1 \int_0^p \left| \frac{h_t(z_t(u))}{f(\tilde{z}_t(u))} - \frac{h(z(u))}{f(z(u))}\right| du dp \\ &= \int_0^1 (1-u) \left| \frac{h_t(z_t(u))}{f(\tilde{z}_t(u))} - \frac{h(z(u))}{f(z(u))}\right| du \\ &\leq \int_0^1 (1-u) \frac{|h_t(z_t(u)) - h(z_t(u))|}{f(\tilde{z}_t(u))}du + \int_0^1 (1-u) \frac{|h(z_t(u)) - h(z(u))|}{f(\tilde{z}_t(u))}du \\ &+ \int_0^1 (1-u) \left| \frac{h(z(u))}{f(\tilde{z}_t(u))} - \frac{h(z(u))}{f(z(u))}\right| du \\ &=: N_1(t) + N_2(t) + N_3(t), \end{aligned} \tag{22}$$

where the equality on the third line holds by changing the order of integration. We below show that the thee terms on the majorant side converge to zero.

Convergence of $N_1(t)$ in (22). Since $\sup_{x\in[0,\infty)}|h_t(x)-h(x)|\to 0$ (as $t\to 0$) by the definition and

$$N_1(t)\le \sup_{x\in[0,\infty)}|h_t(x)-h(x)|\int_0^1\frac{1-u}{f(\tilde z_t(u))}du,$$

the convergence of $N_1(t)$ follows if it holds that

$$\limsup_{t\to 0}\int_0^1\frac{1-u}{f(\tilde z_t(u))}du<\infty, \tag{23}$$

which is to be shown below.

By the construction of $\tilde z_t(u)$ through the mean-value theorem, which is on the line segment connecting $z_t(u)$ to $z(u)$ for each u, and by the non-increasing property of $g(\cdot)$ in **C.1**, we must always have either

$$g(z(u))\le g(\tilde z_t(u))\le g(z_t(u)) \quad\text{or}\quad g(z(u))\ge g(\tilde z_t(u))\ge g(z_t(u)).$$

Thus, letting

$$\begin{aligned}A_t&:=\{u\in[0,\infty)\mid g(z(u))\le g(\tilde z_t(u))\le g(z_t(u))\}\quad\text{and}\\ B_t&:=\{u\in[0,\infty)\mid g(z(u))\ge g(\tilde z_t(u))\ge g(z_t(u))\},\end{aligned}$$

we have

$$\begin{aligned}\int_0^1\frac{1-u}{g(\tilde z_t(u))}du&=\int_{u\in A_t}\frac{1-u}{g(\tilde z_t(u))}du+\int_{u\in B_t}\frac{1-u}{g(\tilde z_t(u))}du\\ &\le\int_{u\in A_t}\frac{1-u}{g(z(u))}du+\int_{u\in B_t}\frac{1-u}{g(z_t(u))}du\\ &\le\int_0^1\frac{1-u}{g(z(u))}du+\int_0^1\frac{1-u}{g(z_t(u))}du. \end{aligned}\tag{24}$$

By changing variables, we can check the boundedness of the two terms on the RHS:

$$\int_0^1\frac{1-u}{g(z(u))}du=\int_0^\infty[1-F(x)][f(x)/g(x)]\,dx<\infty\quad\text{and} \tag{25}$$

$$\limsup_{t\to 0}\int_0^1\frac{1-u}{g(z_t(u))}du<\infty,$$

where the former is imposed in ii) of **C.1** and and the latter holds since the UI condition ii) of **C.2** imlies that

$$\begin{aligned}&\limsup_{t\to 0}\int_0^1\frac{1-u}{g(z_t(u))}du\\ &\le\sup_{t\in(0,\eta]}\int_0^1\mathbf{1}\{Q_t(u)\ge\rho\}\,Q_t(u)du+\sup_{t\in(0,\eta]}\int_0^1\mathbf{1}\{Q_t(u)<\rho\}\,Q_t(u)du\\ &\le 1+\rho,\end{aligned}$$

for some sufficiently large constant ρ (independent of t). Therefore, we have obtained the desired result (23), implying the convergence of $N_1(t)$.

Convergence of $N_2(t)$ in (22). To obtain the desired result, we shall show I) the pointwise convergence of the integrand variable of the second term in (**22**):

$$\tilde{Y}_{2,t}(u) := (1-u)\frac{|h(z_t(u)) - h(z(u))|}{f(\tilde{z}_t(u))} \to 0 \text{ for each } u, \text{ as } t \to 0, \tag{26}$$

and II) the uniform integrability (UI) of $\tilde{Y}_t$. Given these I) and II), the Vitali convergence theorem (see, e.g., p. 187 of Holland, 1999) implies that $N_2(t) \to 0$ as $t \to 0$.

To show I), recall the definition of $F_t(x) = F(x) + th_t(x) \in \bar{D}_{[0,\infty)}$ (with $h_t(\cdot)$ uniformly converging to $h(\cdot)$). Noting the uniform convergence of $F_t(\cdot)$ (to $F(\cdot)$ on $[0,\infty)$) and the following inequality:

$$\begin{aligned}\sup_{u\in[0,1)}|F(z(u)) - F(z_t(u))| &= \sup_{u\in[0,1)}|u - F(z_t(u))| \\ &= \sup_{u\in[0,1)}|F_t(z_t(u)) - F(z_t(u))| = \sup_{x\in[0,\infty)}|F_t(x) - F(x)|,\end{aligned}$$

wherw the second equality uses the fact that $F_t(z_t(u)) = u$ (by the definition of $z_t = \inf_{x\in[0,\infty)}\{F_t(x) \geq u\}$), we also have the uniform convergence of $F(z_t(u))$ to $F(z(u))$ on $[0,1)$ (as $t \to 0$). Then, since $F^{-1}(u) = z(u)$ is continuous, it holds that

$$|z(u) - z_t(u)| = \left|F^{-1}(F(z(u))) - F^{-1}(F(z_t(u)))\right| \to 0 \text{ for each } u \in [0,1),$$

which implies that

$$|z(u) - \tilde{z}_t(u)| \to 0 \text{ for each } u \in [0,1), \tag{27}$$

where we recall the definition of $\tilde{z}_t(u)$ through the mean-value theorem. Now, by the continuity of $h(\cdot)$, which is a realized path of the Brownian bridge, as well as that of $f(\cdot)$, we have $|h(\tilde{z}_t(u)) - h(z(u))| \to 0$ and $f(\tilde{z}_t(u)) \to f(z(u))$ for each $u \in [0,1)$ and obtain the pointwise convergence (26).

Next, we verify II), the UI of $\tilde{Y}_{2,t}$. To this end, recall that the limit $h(\cdot)$ of $h_t(\cdot)$ is defined as the standard Brownian bridge's path, and the value space of $z(u)$ is exactly the whole $[0,\infty)$. Given these facts, we can see the uniform blondeness of $h(z(\cdot))$ over $u \in [0,1]$, that is,

$$\sup_{x\in[0,\infty)}|h(x)| = \sup_{u\in[0,1)}|h(z(u))| =: \bar{C}_h < \infty.$$

This also implies that $\sup_{u\in[0,1)} h(z_t(u)) \leq \sup_{x\in[0,\infty)}|h(x)| = \bar{C}_h$. Using the monotonicity of $g(\cdot)$ and the same argument as that for (24),

$$\tilde{Y}_{2,t}(u) \leq (1-u)\left|h(z_t(u)) - h(z(u))\right| \left[\frac{1}{g(z_t(u))} + \frac{1}{g(z(u))}\right]$$
$$\leq 2\bar{C}_h \left[\frac{1-u}{g(z(u))} + \frac{1-u}{g(z_t(u))}\right]. \tag{28}$$

Here, $\frac{1-u}{g(z(u))}$ is independent of t, whose integrability has been checked in (25), and the UI of $Q_t(u) = \frac{1-u}{g(z_t(u))}$ is supposed in ii) of **C.2**. That is, the upper bound of $\tilde{Y}_{2,t}(u)$ is shown to be uniformly integrable, and thus $\tilde{Y}_{2,t}(u)$ itself is also uniformly integrable. Therefore, given the results I) and II), we have established the convergence of $N_2(t)$.

Convergence of $N_3(t)$ in (22). We can show the desired result in the same way as for $N_2(t)$. To this end, look at

$$N_3(t) \leq \sup_{x\in[0,\infty)} |h(x)| \times \int_0^1 \left|\frac{1-u}{f(\tilde{z}_t(u))} - \frac{1-u}{f(z(u))}\right| du.$$

The convergence of $N_3(t)$ follows if the integrand on the RHS satisfies the pointwise convergence (to zero for each u as $t \to 0$) and UI. The former holds by (27) and the continuity of $f(\cdot)$, and the latter holds if $\frac{1-u}{f(\tilde{z}_t(u))}$ is UI. By the same argument as before, we can see that

$$\frac{1-u}{f(\tilde{z}_t(u))} \leq \frac{1-u}{g(z(u))} + \frac{1-u}{g(z_t(u))}$$

and the UI of the majorant side has been already verified for $\tilde{Y}_{2,t}(u)$ through discussions after (28). Now, the proof is completed. ■

A.2 Proof of Theorem 2

In this section, we provide the proof of Theorem 2, the uniform convergence results for our kernel estimators. To this end, we first derive the convergence rate of the normalization component $\hat{F}_{\mathrm{B}}(\infty)$ in Lemma 1, which is required to investigate the properties of $\hat{f}(\cdot)$ and $\hat{F}(\cdot)$ that are based on $\hat{F}_{\mathrm{B}}(\infty)$:

Lemma 1 *Suppose that i) of Assumption 1 and i) - ii) of Assumption 2 hold. Then, the CDF estimator at $x = \infty$, $\hat{F}_{\mathrm{B}}(\infty) := \int_0^\infty \hat{f}_{\mathrm{B}}(y)\,dy$, defined through (8) satisfies*

$$E[|\hat{F}_{\mathrm{B}}(\infty) - 1|] = O_p(b_n) \quad \textit{as } b_n \to 0.$$

Given this lemma, whose proof is provided below, we are ready to derive the uniform convergence rates of our boundary corrected kernel estimators $\hat{f}(\cdot)$ and $\hat{F}(\cdot)$:

Proof of Theorem 2. We first derive the uniform convergence rate of the density estimator $\hat{f}(x)$ in (11). Since $\hat{f}(x) = \hat{f}_{\mathrm{B}}(x)/\hat{F}_{\mathrm{B}}(\infty)$ and $\hat{F}_{\mathrm{B}}(\infty) = 1 + O_p(b_n)$ (by Lemma 1), we can write

$$\begin{aligned}\hat{f}(x) - f(x) &= \frac{\hat{f}_{\mathrm{B}}(x) - f(x)}{\hat{F}_{\mathrm{B}}(\infty)} + f(x)\frac{1 - \hat{F}_{\mathrm{B}}(\infty)}{\hat{F}_{\mathrm{B}}(\infty)} \\ &= [\hat{f}_{\mathrm{B}}(x) - f(x)][1 + o_p(1)] + f(x)\frac{O_p(b_n)}{1 + o_p(1)}, \end{aligned} \tag{29}$$

uniformly over $x \in [0,\infty)$. The second term on the RHS is $O_p(b_n)$ uniformly over $x \in [0,\infty)$ given that $\sup_{x\in[0,\infty)} f(x) < \infty$. To analyze the first term, consider the following decomposition:

$$\begin{aligned}\left|\hat{f}_{\mathrm{B}}(x) - f(x)\right| &\leq \left|[\hat{f}_{\mathrm{B}}(x) - E[\hat{f}_{\mathrm{B}}(x)]\right| + \left|E[\hat{f}_{\mathrm{B}}(x)] - f(x)\right| \\ &=: \Pi_{1,n}(x) + \Pi_{2,n}(x). \end{aligned} \tag{30}$$

The first term $\Pi_{n,1}$ on the RHS is the so-called variance term. Recalling the definition of $\hat{f}_{\mathrm{B}}(x)$, observe that

$$\Pi_{1,n}(x) \leq \frac{1}{a_0(0)}\left|\frac{1}{nb_n}\sum\nolimits_{i=1}^{n}\left[K\left(\frac{x - X_i}{b_n}\right) - E[K\left(\frac{x - X_i}{b_n}\right)]\right]\right|.$$

We can apply the same arguments as in the proofs of Theorem 4 of Hansen (2008), under the uniform boundedness of the density and kernel functions, $f(\cdot)$ and $K(\cdot)$, and i.i.d. assumption of $\{X_i\}_{i\geq 1}$ to obtain $\frac{1}{nb_n}\sum_{i=1}^{n}\left[K\left(\frac{x-X_i}{b_n}\right) - E[K\left(\frac{x-X_i}{b_n}\right)]\right] = O_p(\sqrt{(\log n)/nb_n})$ uniformly over $x \in [0,\infty)$. Note that our kernel function $K(\cdot)$ may not satisfy Hansen's conditions (his Assumption 3 and/or equation (22)) (as we have not imposed its continuity in particular). However, by using a technique based on the covering numbers from empirical process theory, developed by Kanaya (2017), we can significantly relax Hansen's conditions on kernel functions. For more details, we refer to Theorem 2, Lemma A.3, and their proofs of Kanaya (2017).

To analyze the second term on the RHS of (30), which is the bias of $\hat{f}_{\mathrm{B}}(x)$, we look at

$$\begin{aligned}\Pi_{2,n}(x) &= \left|\frac{1}{a_0(x/b_n)}\int_0^{\infty}\frac{1}{b_n}K\left(\frac{x-p}{b_n}\right)f(p)\,dp - f(x)\right| \\ &\leq \frac{1}{a_0(x/b_n)}\left|\int_{x/b_n}^{-\infty}|K(q)|\,|f(x - qb_n) - f(x)|\,dq\right| \\ &\leq \frac{1}{a_0(x/b_n)}\int_{-\infty}^{x/b_n}|K(q)|\left|qb_n f'(\tilde{x})\right|dq \\ &\leq \frac{b_n}{a_0(0)}\int_{-\infty}^{\infty}|qK(q)|\,dq \times \sup_{z\in[0,\infty)}\left|f'(z)\right| = O(b_n), \end{aligned}$$

uniformly over $x \in [0, \infty)$, where we have used the mean-value theorem with some $\tilde{x}$ on the line segment connecting x and $x - qb_n$. From these, we can conclude that $\Pi_{1,n}(x) + \Pi_{2,n}(x) = O_p(\sqrt{(\log n)/nb_n}) + O_p(b_n)$ uniformly over $x \in [0, \infty)$, which, together with (29 and (30), establishes the desired result (11).

Next, we derive the result (12), the uniform convergence rate of $\hat{F}(x)$. Recalling the definition of $\hat{F}(x) = \hat{F}_{\mathrm{B}}(x) / \hat{F}_{\mathrm{B}}(\infty)$, we can write

$$\begin{aligned} \hat{F}(x) - F(x) &= \frac{\hat{F}_{\mathrm{B}}(x) - F(x)}{\hat{F}_{\mathrm{B}}(\infty)} + \frac{F(x)}{\hat{F}_{\mathrm{B}}(\infty)}[1 - \hat{F}_{\mathrm{B}}(\infty)] \\ &= [\hat{F}_{\mathrm{B}}(x) - F(x)][1 + o_p(1)] + O_p(b_n), \end{aligned} \tag{31}$$

uniformly over $x \in [0, \infty)$, where the second equality has used Lemma 1. To analyze the term $[\hat{F}_{\mathrm{B}}(x) - F(x)]$, we consider the following decomposition:

$$\begin{aligned} |\hat{F}_{\mathrm{B}}(x) - F(x)| &\leq |\hat{F}_{\mathrm{B}}(x) - E[\hat{F}_{\mathrm{B}}(x)]| + |E[\hat{F}_{\mathrm{B}}(x)] - F(x)| \\ &= \bar{\Pi}_{1,n}(x) + \bar{\Pi}_{n,2}(x), \end{aligned} \tag{32}$$

where we shall derive the convergence rates of the two terms on the RHS. To investigate the term $\Pi_{n,1}$ on the RHS, we let $\bar{K}(s) := \int_{-\infty}^{s} K(u)\,du$. Then, by the changing variables,

$$\begin{aligned} \bar{\Pi}_{n,1}(x) &\leq \frac{1}{a_0(0)} \left| \frac{1}{nb_n} \sum\nolimits_{i=1}^{n} \left[\int_0^x K\left(\frac{y - X_i}{b_n}\right) dy - E[\int_0^x \left(\frac{y - X_i}{b_n}\right) dy] \right] \right| \\ &= \frac{1}{a_0(0)} \left| \frac{1}{n} \sum\nolimits_{i=1}^{n} \left[\bar{K}\left(\frac{x - X_i}{b_n}\right) - E[\bar{K}\left(\frac{x - X_i}{b_n}\right)] \right] \right| \\ &+ \frac{1}{a_0(0)} \left| \frac{1}{n} \sum\nolimits_{i=1}^{n} [\bar{K}(-X_i/b_n) - E[\bar{K}(-X_i/b_n)] \right| \\ &=: \bar{\Pi}_{n,11}(x) + \bar{\Pi}_{n,12}, \end{aligned}$$

where we note that $\bar{\Pi}_{n,12}$ is independent of x. By changing variables and applying the bounded convergence theorem, we can compute

$$\mathrm{Var}[\bar{K}(-X_i/b_n) - E[\bar{K}(-X_i/b_n)]] = b_n f(0) \int_0^{\infty} \bar{K}^2(-u) du + o(b_n),$$

uniformly over $i \in \{1, \ldots, n\}$, where the boundedness of $\int_0^{\infty} \bar{K}^2(-u)du$ is guaranteed by Assumption 2. Thus, by the i.i.d. condition of $\{X_i\}$, $E\left[|\bar{\Pi}_{n,12}^2\right] = O(b_n/n)$ and $\bar{\Pi}_{n,12} = O_p(\sqrt{b_n/n})$. To analyze $\bar{\Pi}_{n,11}(x)$, we employ a technique based on the covering numbers as in Kanaya (2017): Define a set of functions, $p\,(\in [0, \infty)) \mapsto \bar{K}\left(\frac{x - p}{b_n}\right)\,(\in [0, 1])$, indexed by (x, b_n) as

$$\bar{\mathcal{K}} := \left\{ \bar{K}\left(\frac{x - p}{b_n}\right) \middle|\ x \geq 0 \ \text{ and } \ b_n > 0 \right\}.$$

By the same arguments as in the proof of Lemma A.3 of Kanaya (2017), this functional set $\bar{\mathcal{K}}$ is Euclidean, where we refer to pp. 905-906 of Kanaya (2017) for the definition of being Euclidean. By the uniform boundedness of $\bar{K}(\cdot)$, we can check that $\mathrm{Var}[\bar{K}\left(\frac{x-X_i}{b_n}\right) - E[\bar{K}\left(\frac{x-X_i}{b_n}\right)]]$ is uniformly bounded over x and i. Then, using the Euclidean property of $\bar{\mathcal{K}}$ and the Bernstein exponential inequality for i.i.d. random variables (see, e.g., p. 102 of van der Vaart and Wellner, 1996), we can show that $\sup_{x\in[0,\infty)} \bar{\Pi}_{n,11}(x) = O_p(\sqrt{(\log n)/n})$. Verification of this result is quite analogous to the proof Theorem 2 of Kanaya (2017), and its details are omitted for brevity. From these, we can see that

$$\begin{aligned}\bar{\Pi}_{1,n}(x) &\le \bar{\Pi}_{1,n}(x) + \bar{\Pi}_{n,12} \\ &= O_p(\sqrt{(\log n)/n}) + O_p(\sqrt{b_n/n}) = O_p(\sqrt{(\log n)/n}). \end{aligned} \tag{33}$$

We next investigate the term $\bar{\Pi}_{n,2}(x)$. By changing the order of integration, changing of variables, and using the Taylor expansion, we have

$$\begin{aligned} E[\hat{F}_{\mathrm{B}}(x)] &= \int_0^\infty \left\{ \int_0^x \frac{1}{a_0(y/b_n)\, b_n} K\left(\frac{y-p}{b_n}\right) dy \right\} f(p)\, dp \\ &= \int_0^x \frac{1}{a_0(y/b_n)} \left\{ \int_{-\infty}^{y/b_n} K(q)\, f(y - qb_n)\, dp \right\} dy \\ &= \int_0^x \frac{1}{a_0(y/b_n)} \left\{ \int_{-\infty}^{y/b_n} K(q) \left[f(y) - qb_n f'(\tilde{y}) \right] dq \right\} dy \\ &= \int_0^x \frac{1}{a_0(y/b_n)} \left\{ \int_{-\infty}^{y/b_n} K(q)\, dp \right\} f(y)\, dy + \int_0^x \frac{1}{a_0(y/b_n)} \left\{ -b_n \int_{-\infty}^{y/b_n} qK(q)\, f'(\tilde{y})\, dq \right\} dy, \end{aligned} \tag{34}$$

where $\tilde{y}$ is on the line segment connecting $y - qb_n$ to y. The first term on the RHS of (34) is equal to $F(x)$. To find a bound for the second term, note that we can write $\tilde{y} = y - \lambda q b_n$ for some $\lambda \in [0,1]$, which depends on y, q, and b_n. Thus, for $q \in (-\infty, y/2b_n]$, we have $\tilde{y} \ge y/2$ and

$$\left| f'(\tilde{y}) \right| \le M_0 \left[1 + y/2\right]^{-\delta},$$

by the condition ii) of Assumption 1. Therefore,

$$
\begin{aligned}
\bar{\Pi}_{2,n}(x) &= \left|E[\hat{F}_{\mathrm{B}}(x)] - F(x)\right| \leq \frac{b_n}{a_0(0)} \int_0^x \left\{ \int_{-\infty}^{y/b_n} |q|\, K(q) \left|f'(\tilde{y})\right| dq \right\} dy \\
&\leq \frac{b_n}{a_0(0)} \int_0^x \left\{ \int_{-\infty}^{y/2b_n} |q| K(q) \left|f'(\tilde{y})\right| dq + \int_{y/2b_n}^{y/b_n} qK(q) \left|f'(\tilde{y})\right| dq \right\} dy \\
&\leq \frac{b_n}{a_0(0)} \int_0^x \left\{ M_0 [1+y/2]^{-\delta} \int_{-\infty}^{\infty} |qK(q)|\, dq + \sup_{z\in(0,\infty]} \left|f'(z)\right| \int_{y/2b_n}^{\infty} qK(q)\, dq \right\} dy \\
&= O(b_n) \quad \text{uniformly over } x \in [0,\infty), \qquad (35)
\end{aligned}
$$

where the last equality follows since

$$
\int_0^x [1+y/2]^{-\delta}\, dy \leq \int_0^\infty [1+y/2]^{-\delta}\, dy < \infty \text{ and}
$$

$$
\int_0^x \left(\int_{y/2b_n}^{\infty} qK(q)\, dq \right) dy \leq \int_0^\infty \left(\int_y^\infty qK(q)\, dq \right) dy < \infty,
$$

which hold by the condition $\delta > 1$ in Assumption 1 and the exponential tail decay condition on $K(\cdot)$ in Assumption 2, respectively.

By (31)-(33) as well as (35), we can obtain the conclusion of the theorem, completing the proof. ■

Proof of Lemma 1. By the definition of $\hat{F}_{\mathrm{B}}(\infty)$ and change of variables, we can write

$$
\hat{F}_{\mathrm{B}}(\infty) = \frac{1}{n} \sum_{i=1}^n \int_{-X_i/b_n}^{\infty} \frac{1}{a_0(w + X_i/b_n)} K(w)\, dw =: \frac{1}{n} \sum_{i=1}^n \eta_i. \qquad (36)
$$

Given the definition of $a_0(p) = \int_{-\infty}^{p} K(u)\, du$, we have $a_0(\infty) = \int_{-\infty}^{\infty} K(u)\, du = 1$ and $a_0(w + X_i/b_n) \geq a_0(0)$ for any $w \geq -X_i/b_n$. Then,

$$
\begin{aligned}
|\eta_i - 1| &\leq \int_{-X_i/b_n}^{\infty} \left| \frac{1}{a_0(w + X_i/b_n)} - \frac{1}{a_0(\infty)} \right| K(w)\, dw \\
&\quad + \left| \int_{-X_i/b_n}^{\infty} \frac{1}{a_0(\infty)} K(w)\, dw - \int_{-\infty}^{\infty} \frac{1}{a_0(\infty)} K(w)\, dw \right| \\
&\leq \frac{1}{a_0^2(0)} \int_{-X_i/b_n}^{\infty} [a_0(\infty) - a_0(w + X_i/b_n)]\, |K(w)|\, dw + \frac{1}{a_0(0)} \int_{-\infty}^{-X_i/b_n} |K(w)|\, dw \\
&=: \Delta_{1,i} + \Delta_{2,i}.
\end{aligned}
$$

By ii) of Assumption 2, we can find some constant $\tilde{M}_K > 0$ such that $\int_{-\infty}^{-s} |K(w)| \, dw \leq \tilde{M}_K [1+s]^{-\nu+1}$ for $s \geq 0$ since since $\nu > 2$. Thus,

$$\begin{aligned}
E[\Delta_{2,i}] &= \frac{1}{a_0(0)} \int_0^\infty \left(\int_{-\infty}^{-p/b_n} |K(w)| \, dw \right) f(p) \, dp \\
&\leq \frac{\tilde{M}_K}{a_0(0)} \int_0^\infty [1+p/b_n]^{-\nu+1} f(p) \, dp \\
&= \frac{\tilde{M}_K}{a_0(0)} b_n \int_0^\infty [1+q]^{-\nu+1} f(qb_n) \, dq \\
&\leq \frac{\tilde{M}_K}{a_0(0)} b_n \int_0^\infty [1+q]^{-\nu+1} \, dq \times \sup_{z\in[0,\infty)} f(z) = O(b_n),
\end{aligned}$$

uniformly over i, where the last equality holds since $\nu > 2$ and $\int_0^\infty [1+q]^{-\nu+1} \, dq < \infty$. To find a bound for $\Delta_{1,i}$, noting that $w \in [-X_i/b_n, \infty)$, we have

$$a_0(\infty) - a_0(w + X_i/b_n) = \int_{w+X_i/b_n}^\infty K(u) \, du \leq \tilde{M}_K [1 + |w + X_i/b_n|]^{-\nu+1}$$

and

$$\begin{aligned}
E[\Delta_{1,i}] &\leq \frac{\tilde{M}_K}{a_0^2(0)} \int_0^\infty \left\{ \int_{-p/b_n}^\infty [1 + |w + p/b_n|]^{-\nu+1} |K(w)| \, dw \right\} f(p) \, dp \\
&= \frac{\tilde{M}_K}{a_0^2(0)} b_n \int_0^\infty \left\{ \int_{-q}^\infty [1 + |w+q|]^{-\nu+1} |K(w)| \, dw \right\} f(qb_n) \, dq \\
&\leq \frac{\tilde{M}_K}{a_0^2(0)} b_n \int_0^\infty \left\{ \int_{-q}^\infty [1 + |w+q|]^{-\nu+1} |K(w)| \, dw \right\} dq \times \sup_{z\in[0,\infty)} f(z) = O(b_n),
\end{aligned}$$

uniformly over i, where the last equality holds since

$$\begin{aligned}
\int_0^\infty \left\{ \int_{-q}^\infty [1 + |w+q|]^{-\nu+1} |K(w)| \, dw \right\} dq &= \int_0^\infty \left\{ \int_0^\infty [1+|u|]^{-\nu+1} |K(u-q)| \, du \right\} dq \\
&= \int_0^\infty [1+|u|]^{-\nu+1} \left\{ \int_0^\infty |K(u-q)| \, dq \right\} du \\
&\leq \int_0^\infty [1+|u|]^{-\nu+1} du \times \int_{-\infty}^\infty |K(-q)| \, dq < \infty.
\end{aligned}$$

Therefore, we have shown that $E[|\eta_i - 1|] = O(b_n)$ uniformly over i, which, together with (36), implies the conclusion of the lemma. The proof is completed. ■

A.3 Proof of Theorem 3

In this section, we provide the proof of Theorem 3:

Proof of Theorem 3. We first consider Part a). The proof for Part b) can be done analogously. Our proof proceeds in two steps. In the first step, we derive uniform upper bounds of the kernel estimators $\hat{f}(x)$ and $1-\hat{F}(x)$ as well as that of $K\left(\frac{x-X_i}{b_n}\right)$, and, in the second step, we show the boundedness of the integral.

Step 1: Given the uniform convergence result in Theorem 2, we can write

$$\hat{f}(x) = f(x) + o(1) \leq 2f(x) \quad \text{for any } x \in [0, L]. \tag{37}$$

with probability approaching one as $n \to \infty$ (w.p.a. 1). To find another bound of $\hat{f}(x)$ for $x \geq L$, let

$$c_n := [nb_n/(\log n)^2]^{1/2\tau}. \tag{38}$$

Since the bandwidth b_n is selected as $o(1/\sqrt{n})$, we can also write

$$\hat{f}(x) = f(x) + O_p(\sqrt{(\log n)/nb_n}),$$

by Theorem 2. Given the definition of c_n and Assumption 1, we can check that $f(x)$ is larger than $\sqrt{(\log n)/nb_n}$ for any $x \in [L, c_n]$ in that

$$\sup_{x\in[L,c_n]} \frac{\sqrt{(\log n)/nb_n}}{f(x)} \leq \frac{\sqrt{(\log n)/nb_n}}{(1/M)\, c_n^{-\tau_1}} = M/\sqrt{\log n} = o(1),$$

and thus

$$\hat{f}(x) \leq 2f(x), \quad \text{for any } x \in [L, c_n], \tag{39}$$

w.p.a. 1.

We now derive some bounds for $1-\hat{F}(x)$. Given the uniform convergence of $\hat{F}(x)$ in Theorem 2, for any $x \in (0, L)$,

$$0 \leq 1-\hat{F}(x) = 1-F(x) + o(1) \leq 2[1-F(x)] \quad \text{for any } x \in [0, L]. \tag{40}$$

To find another bound $1-\hat{F}(x)$ for $x \geq L$, observe that the following lower bound of $1-F(x)$ holds:

$$1-F(x) = \int_x^\infty f(z)\, dz \geq \frac{1}{M}\int_x^\infty z^{-\tau} dz = \frac{1}{M(\tau_1 - 1)} x^{-\tau+1},$$

by Assumption 1. Given the uniform convergence rate of $\hat{F}(x)$ in Theorem 2 and the specified choice of b_n, we can write

$$\hat{F}(x) = F(x) + O_p(\sqrt{(\log n)/n}) \text{ uniformly over } x \in [0, \infty).$$

Given these, we can compute

$$\sup_{x\in[L,c_n]} \frac{\sqrt{(\log n)/n}}{1-F(x)} \leq \frac{\sqrt{(\log n)/n}}{[1/M(\tau_1-1)]c_n^{-\tau+1}}$$
$$= O(\sqrt{n^{-1/\tau}(\log n)^{-(\tau-2)/2}\, b_n^{(\tau-1)/\tau}}) = o(1),$$

implying that, w.p.a. 1,

$$1 - \hat{F}(x) \leq 2[1 - F(x)] \text{ for any } x \in [L, c_n]. \tag{41}$$

To derive a useful bound for the kernel function $K\left(\frac{x-X_i}{b_n}\right)$, let

$$d_n := [n \log n]^{1/(\tau-1)}. \tag{42}$$

Then, observe that

$$\Pr\left[\bigcup_{i=1}^{n} \{X_i \geq d_n/2\}\right] \leq n \frac{M}{\tau_2 - 1} (d_n/2)^{-\tau+1} = O((\log n)^{-1}) \to 0.$$

This implies that, w.p.a. 1, $\max_{1 \leq i \leq n} X_i \leq d_n/2$ and thus,

$$\left|\frac{x - X_i}{b_n}\right| \geq \left|\frac{x - d_n/2}{b_n}\right| > L_K \text{ for any } x \geq d_n \text{ and } i \in \{1, \ldots, n\},$$

implying that, w.p.a. 1,

$$K\left(\frac{x - X_i}{b_n}\right) = 0 \text{ for any } x \geq d_n \text{ and } i \in \{1, \ldots, n\}, \tag{43}$$

by the compactness of the support of $K(\cdot)$ imposed in iii) of Assumption 2.

Step 2: We here investigate the boundedness of the integral. To this end, with sight abuse of integral notation, we write

$$\begin{aligned} \int_0^\infty \{[1 - \hat{F}(x)]/g(x)\}^{1+\phi} \hat{f}(x)\, dx &= \left(\int_0^{c_n} + \int_{c_n}^{d_n} + \int_{d_n}^\infty\right) \{[1 - \hat{F}(x)]/g(x)\}^{1+\phi} \hat{f}(x)\, dx \\ &=: I_{1,n} + I_{2,n} + I_{3,n}, \end{aligned}$$

where c_n and d_n are defined in (38) and (42), respectively. We below analyze the three terms on the RHS separately and show that

$$I_{2,n} = o_p(1), \quad I_{3,n} = o_p(1),$$

and $I_{1,n}$ is bounded, whose upper bound in given in (44)-(45). Therefore, the constant C_ϕ is, for example, given by the twice of the upper bound of $I_{1,n}$.

The boundedness of $I_{1,n}$. First, using the four upper bound derived in (37)-(41), we have

$$\begin{aligned} I_{1,n} &\leq \int_0^{c_n} \{2[1 - F(x)]/g(x)\}^{1+\phi} 2f(x)\, dx \\ &= 2^{2+\phi} \left\{\int_0^L \{[1 - F(x)]/g(x)\}^{1+\phi} f(x)\, dx + \int_L^\infty \{[1 - F(x)]/g(x)\}^{1+\phi} f(x)\, dx\right\}, \end{aligned} \tag{44}$$

where we can easily see the boundedness of the first term on the RHS. Using Assumption 1 and its implications derived in (17) and (18), the integral of the second term can be written as

$$\begin{aligned}\int_L^\infty \{[1-F(x)]/g(x)\}^{1+\phi} f(x)\,dx &\le \tilde{C}\int_L^\infty \left\{x^{-\tau+1}\times x^{\tau}\right\}^{1+\phi} x^{-\tau}dx \\ &= \tilde{C}\int_L^\infty x^{(1+\phi)-\tau}dx, \qquad (45)\end{aligned}$$

with some (sufficiently large) constant $\tilde{C}>0$. Recalling $\tau>1$, we can see that the integral $\int_L^\infty x^{(1+\phi)-\tau}dx$ is bounded if

$$(1+\phi)-\tau<-1 \;\Leftrightarrow\; \phi<\tau-2.$$

It is possible to pick some positive ϕ satisfying this inequality when $\tau>2$, which is maintained in Assumption 1.

The boundedness of $I_{2,n}$. For notational convenience, we let

$$\check{f}(x) := \frac{1}{nb_n}\sum_{i=1}^n K\left(\frac{x-X_i}{b_n}\right) \quad \text{and} \quad \check{F}^c(x) := \int_x^\infty \check{f}(y)\,dy.$$

Using these, we can write

$$\begin{aligned}\hat{f}(x) &= \frac{\hat{f}_{\mathrm{B}}(\infty)}{\hat{F}_{\mathrm{B}}(\infty)} = \frac{1}{\hat{F}_{\mathrm{B}}(\infty)}\frac{1}{nb_n}\sum_{i=1}^n \frac{1}{a_0(y/b_n)}K\left(\frac{x-X_i}{b_n}\right) \\ &\le \frac{1}{[1+o_p(1)]\,a_0(0)}\check{f}(x), \qquad (46) \\ 1-\hat{F}(x) &= \frac{1}{\hat{F}_{\mathrm{B}}(\infty)}[\hat{F}_{\mathrm{B}}(\infty)-\hat{F}_{\mathrm{B}}(x)] \\ &= \frac{1}{\hat{F}_{\mathrm{B}}(\infty)}\frac{1}{nb_n}\sum_{i=1}^n\int_x^\infty \frac{1}{a_0(y/b_n)}K\left(\frac{y-X_i}{b_n}\right)dy \\ &\le \frac{1}{[1+o_p(1)]\,a_0(0)}\check{F}^c(x), \qquad (47)\end{aligned}$$

uniformly over $x\in[0,\infty)$, where the ineuqlities have used $a_0(y/h)\ge a_0(0)$ and $\hat{F}_{\mathrm{B}}(\infty)=1+o_p(1)$ (by Lemma 1). Therefore, the boundedness of $I_{2,n}$ (in the probability sense) follows if that of

$$\check{I}_{2,n} := \int_{c_n}^{d_n}\{\check{F}^c(x)/g(x)\}^{1+\phi}\check{f}(x)\,dx$$

holds (since $I_{2,n}=O_p(1)\times\check{I}_{2,n}$). By the Hölder inequality,

$$\begin{aligned}\check{I}_{2,n} &\le \sqrt{\int_{c_n}^{d_n}|g(x)|^{-(1+\phi)}\,|\check{f}(x)|^2dx\times\int_{c_n}^{d_n}|g(x)|^{-(1+\phi)}\,|\check{F}^c(x)|^{2(1+\phi)}dx} \\ &=: \sqrt{\check{I}_{21,n}\times\check{I}_{22,n}}.\end{aligned}$$

To analyze $\check{I}_{21,n}$, we look at

$$\begin{aligned}
\check{I}_{21,n} &= \frac{1}{n^2 b_n}\sum_{i=1}^{n}\int_{c_n}^{d_n}|g(x)|^{-(1+\phi)}\frac{1}{b_n}K^2\left(\frac{x-X_i}{b_n}\right)dx \\
&+\frac{2}{n^2}\sum_{1\leq i<j\leq n}\int_{c_n}^{d_n}|g(x)|^{-(1+\phi)}\frac{1}{b_n^2}K\left(\frac{x-X_i}{b_n}\right)K\left(\frac{x-X_j}{b_n}\right)dx \\
&=: \frac{1}{n^2 b_n}\sum_{i=1}^{n}D_n(i)+\frac{2}{n^2}\sum_{1\leq i<j\leq n}\Gamma_n(i,j). \qquad (48)
\end{aligned}$$

By changing the order of integration and changing variables, we can compute the expectation of the summand of the first term on the RHS as follows:

$$\begin{aligned}
E[D_n(i)] &= \int_{c_n}^{d_n}|g(x)|^{-(1+\phi)}\left[\int_{-\infty}^{x/b_n}K^2(q)f(x-qb_n)\,dq\right]dx \\
&\leq M^{\tau(1+\phi)}\int_{c_n}^{d_n}x^{\tau(1+\phi)}\left[M|x/2|^{-\tau}\int_{-\infty}^{\infty}K^2(q)\,dq+\int_{x/2b_n}^{\infty}K^2(q)\,dq\sup_{z\in[0,\infty)}f(z)\right]dx \\
&\leq \check{C}_2\int_{c_n}^{d_n}x^{\tau\phi}dx,
\end{aligned}$$

uniformly over i, where the last inequality holds with a constant $\check{C}_2 := M^{\tau_1(1+\phi)+1}2^{\tau}\int_{-\infty}^{\infty}K^2(q)\,dq$ since $\int_{x/2b_n}^{\infty}K^2(q)\,dq=0$ for $x/2b_n>c_n/2b_n>L_K$ (by the boundedness of the support of $K(\cdot)$). For $\tau(1+\phi)-\tau>0$, we have $\int_{c_n}^{d_n}x^{\tau\phi}dx\leq d_n^{\tau\phi+1}$ and

$$E[D_n(i)]\leq \check{C}_2 d_n^{\tau\phi+1},\ \text{ uniformly over } i\in\{1,\ldots,n\}.$$

Since we have set $d_n=[n\log n]^{1/(\tau-1)}$ and $1/b_n=O(n^{-q_0})$ for some $q_0\in(0,1)$,

$$\begin{aligned}
\frac{1}{n^2 b_n}\sum\nolimits_{i=1}^{n}E[D_n(i)] &\leq \frac{1}{nb_n}\check{C}_2 d_n^{\tau\phi+1} \\
&= O(1)\times(\log n)^{\frac{\tau\phi}{\tau-1}}n^{\frac{\tau\phi}{\tau-1}-(1-q_0)}=o(1), \qquad (49)
\end{aligned}$$

where the last equality holds if

$$\frac{\tau\phi}{\tau-1}-(1-q_0)<0 \;\Leftrightarrow\; \phi<\frac{1}{\tau}(\tau-1)(1-q_0).$$

This inequality is satisfied for some (sufficiently small) $\phi>0$ since $\tau>2$ and $q_0\in(0,1)$.

Since X_i and X_j $(i<j)$ are independent, we can compute

$$E[\Gamma_n(i,j)]=\int_{c_n}^{d_n}|g(x)|^{-(1+\phi)}\left\{\int_{-\infty}^{x/b_n}K(q)f(x-qb_n)\,dq\right\}^2 dx.$$

The integral between the curly braces is bounded as follows:

$$\int_{-\infty}^{x/b_n} K(q)\, f(x-qb_n)\, dq = \int_{-\infty}^{x/2b_n} K(q)\, f(x-qb_n)\, dq + \int_{x/2b_n}^{x/b_n} K(q)\, f(x-qb_n)\, dq$$
$$\leq \int_{-\infty}^{\infty} K(q)\, dq \times M(x/2)^{-\tau} = O(1) \times x^{-\tau},$$

uniformly over $x \geq c_n\ (\geq L)$, and thus

$$\begin{aligned} E[\Gamma_n(i,j)] &= O(1) \times \int_{c_n}^{d_n} |g(x)|^{(1+\phi)} (x/b_n)^{-2\tau}\, dx \\ &\leq O(1) \times M^{1+\phi} \int_{c_n}^{\infty} x^{\tau(1+\phi)} (x/b_n)^{-2\tau}\, dx \\ &= O(1) \times b_n^{2\tau} c_n^{\tau(1+\phi)-2\tau+1}, \quad \text{uniformly over } i,j \in \{1,\ldots,n\}, \end{aligned}$$

where the last equality holds when

$$\tau(1+\phi) - 2\tau + 1 < -1 \;\Leftrightarrow\; \phi < (\tau-2)/2\tau.$$

which may be satisfied for any sufficiently small ϕ since $\tau > 2$. Therefore, given $c_n = [nb_n/(\log n)^2]^{1/2\tau}$ in (38) and $b_n = o\left(n^{-1/2}\right)$,

$$\frac{2}{n^2} \sum_{1\leq i<j\leq n} \Gamma_n(i,j) = o(1) \times (\log n)^{-1/2\tau} n^{-\frac{\tau}{2}-\frac{1}{4\tau}} = o(1),$$

which, together with (48) and (49), leads to $\check{I}_{21,n} = o_p(1)$.

To analyze the term $\check{I}_{22,n}$, note that

$$\check{I}_{22,n} \leq \int_{c_n}^{d_n} |g(x)|^{-(1+\phi)} |\check{F}^c(x)|^2 dx,$$

since $\check{F}^c(x) < 1$ for any $x \in [0,\infty)$ (by its definition). This upper bound of $\check{I}_{22,n}$ can be shown to be $o_p(1)$, which may be analyzed exactly in the same way as $\check{I}_{21,n}$. While details are omitted for brevity, we in particular note that both the kernel function and its integral have the same tail decay rate, i.e.,

$$K(x) \leq M_K \exp\{-c_K x\} \quad \text{and} \quad \int_x^{\infty} K(y)\, dy \leq (M_K/c_K) \exp\{-c_K x\},$$

This property leads to the same tail behavior of $\check{f}(x)$ and $\check{F}^c(x)$, and thus the same convergence speed of $\check{I}_{21,n}$ and $\check{I}_{22,n}$.

The boundedness of $I_{3,n}$. By (46), it is sufficient to show the boundedness of

$$I_{3,n} = O_p(1) \times \int_{d_n}^{\infty} \{[1-\hat{F}(x)]/g(x)\}^{1+\phi} \check{f}(x)\, dx.$$

By the bound derived in (43), we have

$$\check{f}(x) \leq \frac{1}{b_n} K\left(\frac{x - d_n/2}{b_n}\right) = 0 \quad \text{for } x \geq d_n,$$

w.p.a. 1. Therefore, $I_{3,n} = 0$ w.p.a. 1. Now, the proof is completed. ■