SMOOTHED VERSIONS OF STATISTICAL FUNCTIONALS FROM A FINITE POPULATION *

Hitoshi Motoyama** and Hajime Takahashi***

We will consider the central limit theorem for the smoothed version of statistical functionals in a finite population. For the infinite population, Reeds (1976) and Fernholz (1983) discuss the problem under the conditions of Hadamard differentiability of the statistical functionals and derive Taylor type expansions. Lindeberg-Feller's central limit theorem is applied to the leading term, and controlling the remainder terms, the central limit theorem for the statistical functionals are proved. We will modify Fernholz's method and apply it to the finite population with smoothed empirical distribution functions, and we will also obtain Taylor type expansions. We then apply the Erdös-Rényi central limit theorem to the leading linear term to obtain the central limit theorem. We will also obtain sufficient conditions for the central limit theorem, both for the smoothed influence function, and the original non-smoothed versions. Some Monte Carlo simulation results are also included.

Key words and phrases: asymptotic normality, central limit theorem, differentiable functional, empirical distribution function, finite population, functional Taylor series expansions, Hadamard differentiable, influence function, kernel smoothing, official statistics, opinion poll, simple random sampling, statistical functional, survey sampling, uniform topology.

1. Introduction

We will consider the central limit theorem for the statistical functionals in a finite population. A reason why we assume an infinite population is to simplify both the theory and the computation. However, in the area of sampling surveys (official statistics, opinion poll, etc.), the sample size is fairly large compared to the population size, and thus it may be inappropriate to apply classical statistical theory, especially large sample theory, directly to these problems. On the other hand, the progress of modern computer technology makes it possible to work directly with finite population problems in many

This version 10/10/2008. This paper has been subsequently published in the *Journal of Japan Statistical Society*, December 2008, Vol. 38, No.3, pp.475-504.

*This article is a part of the first author's doctoral thesis submitted to Hitotsubashi University.

**The Graduate School of Economics, Hitotsubashi University, 2-1 Naka Kunitachi-shi Tokyo Japan and Statistical Information Institute for Consulting and Analysis, 6-3-9 Minami-Aoyama Minato-ku Tokyo Japan. This research is supported in part by the Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology 15653014, 17330046 and in part a grant from the 21st Century COE Program "Research Unit for Statistical Analysis in Social Sciences" at Hitotsubashi University.

***The Graduate School of Economics, Hitotsubashi University, 2-1 Naka Kunitachi-shi Tokyo Japan. This research is supported in part by the Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology 18203013.

areas of statistical applications, and methods based on resampling, such as the bootstrap method, help us to measure the accuracy of the estimators. Here, the justification of resampling methods may depend on the central limit theorem for the estimators and we will prove the asymptotic normality of the estimators by applying the theory of statistical functionals developed by von Mises (1947), Reeds (1976), Fernholz (1983), and Takahashi (1988), among others.

We will also consider smoothed versions of the empirical distribution functions. Although it may sound strange in the finite population problem, in some practical situations it is reasonable to assume that the underlying distribution function converges to the continuous distribution function as the population size goes to infinity. Here we have implicitly assumed the probability space where continuous distribution functions are also defined. The finite probability spaces are embedded in the space which makes it possible to consider the limits operation. In this case, the smoothed bootstrap may be applied, and it is worth considering the smoothed version of the empirical distribution function and the statistical functionals defined on it. For these reasons, we will derive asymptotic normality of smoothed statistical functionals for a simple random sample from a finite population. The non-smoothed version will be obtained as a simple corollary to our results.

To fix the idea, we let $\Omega^{(N)} = \{x_1, \ldots, x_N\}$ be a mutually distinct finite population of size N, and a simple random sample $(X_1, X_2, \ldots, X_n)$ is taken from $\Omega^{(N)}$ without replacement. More precisely, let $(\pi_1, \ldots, \pi_N)$ take all possible permutations of $(1, \ldots, N)$ with common probability $(N!)^{-1}$, and $X_i = x_{\pi_i}, 1 \le i \le n$.

Define a population distribution function (d.f.)

$$F_N(x) = \frac{1}{N} \sum_{i=1}^{N} I_{(-\infty, x]}(x_i),$$

and an empirical distribution function $(X_1, X_2, \ldots, X_n)$ by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty, x]}(X_i).$$

Let T be a statistical functional defined on the set of distribution functions, including both the population distribution function and all empirical distribution functions (see von Mises (1947), Fernholz (1983), Reeds (1976), Takahashi (1988)), then the parameter of interest is expressed by $T(F_N)$ and its naive estimate may be given by $T(F_n)$. However, as a finite population distribution function tends to become a smooth function as N gets larger, it may be more appealing to replace F_n by its smoothed version $\widetilde{F}_n$, the kernel distribution function estimator, which is to be defined below. This type of statistic $T(\widetilde{F}_n)$ is used in the context of the smoothed bootstrap (Silverman and Young (1987), Young (1990)) and smoothed quantiles (Falk (1985)). Fernholz (1993) derives asymptotic normality of smoothed statistical functionals in I.I.D. settings. We consider the finite population counterpart and obtain its asymptotic distribution. In order to ensure that the population distribution converges to the sufficiently continuous distribution function, we need the following Assumption A. This is the standing assumption to be used throughout this paper;

ASSUMPTION A. *There exist some sequences $B_N > 0$ and constant $M > 0$ satisfying the condition*

$$0 < \inf_N B_N \leq \sup_N B_N \leq M < \infty \tag{1.1}$$

such that for any bounded set $\mathrm{B} \subset \mathbf{R}$

$$|F_N(x) - F_N(y)| \leq B_N|x - y| + O(1/N), \quad \textit{as} \quad N \to \infty \tag{1.2}$$

holds uniformly in $x, y \in \mathrm{B}$.

Assumption A represents the situation where the population distribution goes to the Lipschitz continuous function uniformly in any bounded set of $\mathbf{R}$ as $N \to \infty$. Also we note that Assumption A assures that the amount of jumps of F_N are at most $O(1/N)$.

Now, we have used and will often use Landau's notation and its probabilistic version; for $\{p_n, n \geq 1\}$ *and* $\{q_n, n \geq 1\}$, $p_n = o(q_n)$, as $n \to \infty$, if $\frac{p_n}{q_n} \to 0$, as $n \to \infty$, ($p_n = o_p(q_n)$, as $n \to \infty$, if $\frac{p_n}{q_n} \to 0$ in probability, as $n \to \infty$) and $p_n = O(q_n)$, as $n \to \infty$, if $|\frac{p_n}{q_n}| < M$ for all $n \geq 1$ and some $0 < M < \infty$. ($p_n = O_p(q_n)$, as $n \to \infty$, if $\mathrm{P}(|\frac{p_n}{q_n}| < M) > 1 - \epsilon$ for all $n \geq 1$, $0 < \epsilon < 1$, and some $0 < M < \infty$).

We are ready to define a smoothed empirical distribution function. For each $n \geq 1$, a kernel d.f. estimator $\widetilde{F}_n$ is defined by taking the convolution of F_n with some density k_n; $\widetilde{F}_n = F_n * k_n$. In our case,

$$\begin{aligned} \widetilde{F}_n(x) &= F_n * k_n(x) = \int F_n(x-t)k_n(t)dt \\ &= \int F_n(x-t)dK_n(t) = \int K_n(x-t)dF_n(t) \\ &= \frac{1}{n}\sum_{i=1}^{n} K_n(x - X_i), \end{aligned} \tag{1.3}$$

where $K_n(x) = \int_{-\infty}^{x} k_n(t)dt$.

We will next define a *regular kernel sequence* $\{k_n, n \geq 1\}$, which will be used to define our smoothed empirical distribution function. Let k be a symmetric kernel function (not necessarily nonnegative) satisfying $\int k(x)dx = 1$ and let $\{a_n, n \geq 1\}$ be a sequence of positive real numbers. The sequence of kernels $\{k_n, n \geq 1\}$ defined by

$$k_n(x) = \frac{1}{a_n}k\Big(\frac{x}{a_n}\Big), \quad n \geq 1, \tag{1.4}$$

will be called a *kernel sequence* if $a_n = o(1)$, as $n \to \infty$. Note that if $\{k_n, n \geq 1\}$ is a kernel sequence, then the sequence of d.f. $K_n(x) = \int_{-\infty}^{x} k_n(t)dt$ converges weakly to the d.f. Δ, where

$$\Delta(x) = \begin{cases} 0, \ x < 0 \\ 1, \ x \geq 0 \end{cases} . \tag{1.5}$$

We will consider a restricted class of the kernel sequences, which will be called a regular sequence. A regular sequence will be needed when we prove the $o_p(n^{-\frac{1}{2}})$ convergence of the remainder terms of the Taylor series expansion of the statistical functionals.

Definition 1. [Fernholz (1991, 1993)] *A kernel sequence $\{k_n\}$ is regular if there exists a sequence $\{b_n\}$ of positive real numbers such that $b_n = o(n^{-1/2})$ and*

$$\int_{|t|>b_n} |k_n(t)|dt = o(n^{-1/2}). \tag{1.6}$$

Csörgö and Horváth (1995) considers the similar but more restrictive regularity conditions in investigating the asymptotic properties of smoothed empirical and quantile processes in I.I.D. settings. Imposing these types of regularity conditions on kernel might be unavoidable to some extent without imposing more smoothness conditions on population distributions (Yukich (1992) and van der Vaart (1994)).

We close this section with some comments. Campbell (1980) proposes the use of statistical functionals in the finite population and gives a sketch of the proof for the asymptotic normality in various sampling schemes. She, however, uses essentially the I.I.D. result in proving the asymptotic normalities. We will fill in the incompleteness of her arguments and give legitimate proofs for these results. Some of the related results on the finite population problem are obtained by Motoyama and Takahashi (2003), where the rate of convergence to a normal distribution of statistical functionals in simple random sampling is obtained. For L-statistics in survey sampling problems, we refer readers to Shao (1994).

2. Statistical Functionals

We will briefly review the theory of statistical functionals. We start with the definition of statistical functionals for the distribution functions on a finite population. We then define the three typical differentiations of the functionals, and the theory of Taylor series type expansions. The conditions under which the linear part of the expansions obeys the central limit theorem, and the conditions and the choice of the topology which guarantee the convergence of the remainder terms to zero as fast as $o(n^{-\frac{1}{2}})$ in probability, are the main issue of this section. We will modify the arguments of Reeds (1976) and Fernholz (1983) for the finite population problems with the usual empirical distribution functions, and then for the smoothed distribution functions.

Let θ be the parameter of interest. We suppose that θ is a functional of the underlying distribution F_N and we write $\theta = \theta_N(F_N)$. We also let $T_n = T_n(X_1, \ldots, X_n)$ be an estimator of θ. Then it is tempting to write $T_n = \theta_{N,n}(F_n)$. We will formalize this in a manner so that that we can conduct rigorous mathematical arguments in the following manner(cf. Fernholz (1983, 1993)).

DEFINITION 2. *When $T_n = T_n(X_1, \ldots, X_n)$ can be written as a functional T of the empirical distribution function F_n, $T_n = T(F_n)$, where T does not depend on n, then T is called a statistical functional. The domain of T is assumed to contain the empirical d.f.s for all $n \geq 1$, as well as the underlying true d.f. F_N. Unless otherwise specified, the range of T will be the set of real numbers. Moreover, we call $T(\widetilde{F}_n)$ a smoothed statistical functional where $\widetilde{F}_n$ is the smoothed empirical distribution function considered above.*

EXAMPLE 1. [SAMPLE MEAN] *The simplest statistic may be the sample mean;*

$$T_n(X_1, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Then for a general distribution function G, the functional defined by

$$T(G) = \int x dG(x)$$

satisfies $T_n(X_1, \ldots, X_n) = T(F_n)$.

EXAMPLE 2. [SAMPLE QUANTILE] *We define a statistical functional for a distribution function G by*

$$T(G) = G^{-1}(q), \quad 0 < q < 1, \quad G^{-1}(q) = \inf\{x : q \leq G(x)\}.$$

It follows that the q-th sample quantile is given by the statistical functional defined by

$$T(F_n) = X_{(\lceil nq \rceil)},$$

where $\lceil x \rceil$ is the smallest integer not less than x and $X_{(1)} \leq \ldots \leq X_{(n)}$ are the order statistics of $X_1, \ldots, X_n$.

EXAMPLE 3. [L-ESTIMATOR] *Statistics of the form*

$$T(F_n) = \sum_{i=1}^{s} \beta_i F_n^{-1}(q_i), \quad F_n^{-1}(q) = \inf\{x : q \leq F_n(x)\}$$

where $q_1, \ldots, q_s$ are numbers in $(0, 1)$.

EXAMPLE 4. [M-ESTIMATORS] *Let ψ be a real valued function of two variables and let T_n be defined implicitly by*

$$\sum_{i=1}^{n} \psi(X_i, T_n) = 0.$$

The corresponding functional is defined as a solution $T(G) = \theta$ of

$$\int \psi(x, \theta) dG(x) = 0.$$

Estimators of this form are called M-estimators.

Let $X_1, X_2, \ldots, X_n$ be a simple random sample without replacement from a finite population $x_1, \ldots, x_N$ with distribution function F_N. Then $F_N(X_1), \ldots, F_N(X_n)$ is a simple random sample without replacement from the finite population $\{1/N, \ldots, 1\}$. Let $U_n^\star(x) = \frac{1}{n}\sum_{i=1}^n I_{[0,x]}(F_N(X_i))$ be the empirical distribution function of $\{F_N(X_1), \ldots, F_N(X_n)\}$. Then

$$\begin{aligned} F_n(x) &= \frac{1}{n}\sum_{i=1}^n I_{(-\infty,x]}(X_i) \\ &= \frac{1}{n}\sum_{i=1}^n I_{[0,F_N(x)]}(F_N(X_i)) = U_n^\star \circ F_N(x). \end{aligned}$$

Now, the monotone increasing and continuous version of F_N with $F^{-1}(F_N(x)) = x$ and $F_N(F_N^{-1}(u)) = u$ will be given by

$$(2.1) \qquad F_{NS}(x) = \begin{cases} F_N(x_{(1)}) + \frac{2F_N(x_{(1)})}{\pi}\arctan(x - x_{(1)}) & x \le x_{(1)} \\ F_N(x_{(i)}) & x = x_{(i)}, \\ & i = 1, \ldots, N \\ F_N(x_{(i)}) + \frac{F_N(x_{(i+1)}) - F(x_{(i)})}{x_{(i+1)} - x_{(i)}}(x - x_{(i)}) & x \in [x_{(i)},\ x_{(i+1)}] \ , \\ & (x_{(i)} \ne x_{(i+1)}), \\ & i = 1, \ldots, N-1 \\ 1 + \frac{1}{N}\arctan(x - x_{(N)}) & x \ge x_{(N)} \end{cases}$$

where $x_{(1)} \le x_{(2)} \le \ldots \le x_{(N)}$ are the ordered characteristics of the population.

F_{NS} defined above is the strictly increasing continuous function (but not a distribution function!) satisfying $F_{NS}(x_i) = F_N(x_i)$ for all $i = 1, \ldots, N$ and $\|F_{NS} - F_N\| = O(1/N)$. Hence, defining $U_n(x) = \frac{1}{n}\sum_{i=1}^n I_{[0,x]}(F_{NS}(X_i))$, we also have

$$(2.2) \qquad \begin{aligned} F_n(x) &= \frac{1}{n}\sum_{i=1}^n I_{[0,F_N(x)]}(F_{NS}(X_i)) \\ &= \frac{1}{n}\sum_{i=1}^n I_{[0,F_{NS}(x)]}(F_{NS}(X_i)) = U_n \circ F_{NS}(x). \end{aligned}$$

A statistical functional T induces a functional τ of d.f.s U_n by

$$(2.3) \qquad \tau(U_n) = T(F_n) = T(U_n \circ F_{NS}).$$

Generally, for each F_N, we define a statistical functional τ of d.f. G on $[0, 1]$ by

$$(2.4) \qquad \tau(G) = T(G \circ F_{NS})$$

whenever the right hand side is defined. Hence, we can restrict our attention to d.f.s concentrated on $[0,1]$ and view them as elements of $D[0,1]$, the space of right continuous real valued functions on $[0,1]$ which have left limits.

If T is a statistical functional and τ is the functional induced in $D[0,1]$ by (2.3), the asymptotic properties of $T(F_n)$ and $T(\widetilde{F}_n)$ may be determined by the differentiability of τ. The asymptotic properties depend on the type of the differentiations. We will consider three different type of differentiations, and they are defined in the following(cf. Fernholz (1993));

DEFINITION 3. *Let τ be a functional defined on an open subset $\mathcal{A}$ of a normed vector space $\mathbf{V}$ and let $g \in \mathcal{A}$.*

1. The functional τ is Gateaux differentiable at g if there exists a continuous linear functional τ'_g defined on $\mathbf{V}$ such that

$$\lim_{t \to 0} \frac{\tau(g+th) - \tau(g) - \tau'_g(th)}{t} = 0 \tag{2.5}$$

for each $h \in \mathbf{V}$. In this case τ'_g will be called the Gateaux derivative of τ at g.
2. The functional τ is Hadamard differentiable at g if for any compact subset $K \subset \mathbf{V}$, (2.5) holds uniformly for $h \in K$. The linear functional τ'_g will be called the Hadamard derivative of τ at g.
3. The functional τ is Fréchet differentiable at g if for any bounded subset $B \subset \mathbf{V}$, (2.5) holds uniformly for $h \in B$. The linear functional τ'_g will be called the Fréchet derivative of τ at g.

Since singleton is compact, and the compact set is bounded, Fréchet differentiability implies Hadamard differentiability which in turn implies Gateaux differentiability.

For a statistical functional T and a d.f. F_N, the *influence function* of T at F_N is a real valued function IF_{T,F_N} defined by

$$\mathrm{IF}_{T,F_N}(x) = \frac{d}{dt} T(F_N + t(\Delta_x - F_N))|_{t=0} \tag{2.6}$$

where Δ_x is the d.f. of the point mass one at x, i.e. :

$$\Delta_x(s) = \Delta(s-x) = \begin{cases} 0, & s < x \\ 1, & s \geq x \end{cases}.$$

If T and τ are defined as above, then the Gateaux derivative τ'_U of τ at uniform distribution function U and the influence function of T at F_{NS} are related by

$$\mathrm{IF}_{T,F_{NS}}(x) = \tau'_U((\Delta_x - F_{NS}) \circ F_{NS}^{-1}), \tag{2.7}$$

since

$$\begin{aligned}
\mathrm{IF}_{T,F_{NS}}(x) &= \lim_{t\to 0} \frac{T(F_{NS}+t(\Delta_x - F_{NS})) - T(F_{NS})}{t} \\
&= \lim_{t\to 0} \frac{\tau(U + t(\Delta_x - F_{NS})\circ F_{NS}^{-1}) - \tau(U)}{t} \\
&= \tau_U'((\Delta_x - F_{NS})\circ F_{NS}^{-1}),
\end{aligned}$$

where we have used the fact that $F_{NS}^{-1}(F_{NS}(x)) = x$ and $F_{NS}(F_{NS}^{-1}(u)) = u$ hold from the monotone increasing property and the continuity of F_{NS}.

Note: Since F_{NS} is a monotone increasing function, using the (right-)continuity of F_{NS} we have

$$\begin{aligned}
F_{NS}^{-1}(F_{NS}(x)) &= \inf\{x^* : F_{NS}(x^*) \geq F_{NS}(x)\} \\
&= \inf\{x^* : F_{NS}(x^*) = F_{NS}(x)\} = x.
\end{aligned}$$

And, from the continuity of F_{NS}, we have

$$F_{NS}(F_{NS}^{-1}(u)) = F_{NS}(\inf\{x : F_{NS}(x) \geq u\}) = u.$$

Here and in what follows, we assume

$$\int \mathrm{IF}_{T,F_{NS}}(x)dF_N(x) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{IF}_{T,F_{NS}}(x_i) = 0, \tag{2.8}$$

by appropriate choice of additive constant(Reeds (1976), p.38 and Serfling (1980), pp.222-223 Lemma A).

Under Assumption A, we will prove the following proposition which corresponds to Theorem 2.3 of Fernholz (1991) for the I.I.D. case.

Proposition 1. *Let $X_1, \ldots, X_n$ be a simple random sample without replacement from a finite population with distribution function F_N. Let F_n be the empirical distribution function and let $\widetilde{F}_n$ be the smoothed empirical distribution function defined by (1.3) with a regular kernel sequence $\{k_n\}$. Then, we have*

$$\sqrt{n} \sup_{-\infty<x<\infty} |\widetilde{F}_n(x) - F_n(x)| \to 0 \quad a.s. \quad n, N \to \infty. \tag{2.9}$$

Proof. The proof will be given in the Appendix of this article.

In order to evaluate the linear part of the smoothed statistical functional, we will show that the influence function of the smoothed statistical functional may be obtained by smoothing the influence function of the original functional. We will prove the next lemma under the slightly weaker conditions of Lemma 4 of Fernfolz (1993), which is suitable for our purpose.

Lemma 1. *Suppose τ is Gateau differentiable at the uniform distribution function U with derivative τ'_U. If the influence function* $\mathrm{IF} = \mathrm{IF}_{\mathrm{T,F_{NS}}}$ *is Lebesgue-Stieltjes integrable with respect to functions of bounded variation, we have*

$$\tau'_U(\widetilde{U}_n) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathrm{IF}}(X_i) + \tau'_U(U), \tag{2.10}$$

where $\widetilde{U}_n = \widetilde{F}_n \circ F_{NS}^{-1}$ *and* $\widetilde{\mathrm{IF}} = \mathrm{IF} * \mathrm{k_n}$.

Note that when a function g is right continuous having left limit, and a function h is a (right-)continuous monotone nondecreasing function having left limit, it is easily seen that the composition $g \circ h$ is a right continuous function having left limits.

Proof. Let t_i, $i = 1, \ldots, m$ be $-\infty < t_1 < \cdots < t_m < \infty$, and $\bar{t}_i$, $i = 1, \ldots, m-1$ are middle points $t_i \le \bar{t}_i \le t_{i+1}$. Then as $t_1 \to -\infty$, $t_m \to \infty$, $\max_i(t_{i+1} - t_i) \to 0$, $K_{n,x} = \Delta_x * k_n$ is approximated uniformly on $\mathbf{R}$ by the sum

$$S = \sum_{i=1}^{m} (K_n(t_{i+1}) - K_n(t_i))\Delta_{x-\bar{t}_i}.$$

Therefore, $K_{n,x} \circ F_{NS}^{-1}$ may be approximated by the function of the form $S \circ F_{NS}^{-1}$ in the space $D[0,1]$.

It follows from the linearity of τ'_U and (2.7) that

$$\begin{aligned} \tau'_U(S \circ F_{NS}^{-1}) &= \sum_{i=1}^{m} (K_n(t_{i+1}) - K_n(t_i))\tau'_U(\Delta_{x-\bar{t}_i} \circ F_{NS}^{-1}) \\ &= \sum_{i=1}^{m} (K_n(t_{i+1}) - K_n(t_i))(\mathrm{IF}(x - \bar{t}_i) + \tau'_U(U)). \end{aligned} \tag{2.11}$$

Since IF is Lebesgue-Stieltjes integrable with respect to K_n, the sum (2.11) converges to $\widetilde{\mathrm{IF}}(x) + \tau'_U(U)$ for each given x.

Hence from the continuity of τ'_U, we have

$$\tau'_U(K_{n,x} \circ F_{NS}^{-1}) = \widetilde{\mathrm{IF}}(x) + \tau'_U(U).$$

Since,

$$\widetilde{U}_n = \widetilde{F}_n \circ F_{NS}^{-1} = \frac{1}{n}\sum_{i=1}^{n} K_{n,X_i} \circ F_{NS}^{-1},$$

it follows that

$$\tau'_U(\widetilde{U}_n) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathrm{IF}}(X_i) + \tau'_U(U).$$

3. Remainder terms

In this section we will show the convergence in probability of the remainder terms from the linear approximations.

Suppose that $\tau : D[0,1] \to \mathbf{R}$ is differentiable at U. Also let the remainder term be

$$\mathrm{Rem}(tH) = \tau(U + tH) - \tau(U) - \tau'_U(tH).$$

It follows from the definition of the Hadamard differentiability that for any compact $K \subset D[0,1]$, we have

$$\lim_{t \to 0} \frac{\mathrm{Rem(tH)}}{t} = 0$$

uniformly in $H \in K$.

In evaluating the remainder term, choosing the norm or metric which give topologies plays an essential role. Usually, $D[0,1]$ is equipped with the Skorohod topologies (Billingsley (1999)). However, in Skorohod topologies, pointwise addition of functions is not a continuous operation and $D[0,1]$ is not a topological vector space (Billingsley (1999) p.137 Problem12.2). Hence in order to utilize von Mises's differentiation theory, we adopt the uniform norm topologies.

Nevertheless, problem still remains after adopting the uniform norm. It is well known that under the uniform norm the empirical distribution function may not be a random element of $D[0,1]$ under some circumstances (cf. Billingsley(1999) pp.157-158, Fernholz (1983) pp.34-35). We show the result along the line of Fernholz. Consider the situation that $n = 1$ and values of the population characteristic are distributed (independently of the sampling structures) uniformly on $[0,1]$. Define F_1 as the corresponding empirical distribution function, i.e.

$$F_1(x) = \Delta_{X_1}.$$

The random variable X_1 induces a probability measure μ on $[0,1]$ by $\mu(B) = \mathrm{P}(X_1 \in B)$ for any Borel set $B \subset [0,1]$. Since the value of X_1 is uniformly distributed, and from the fact $n = 1$, μ coincides with the Lebesgue measure on $[0,1]$.

Now, we define the open ball $\mathcal{O}_x$ in $D[0,1]$ with center $\Delta_x (x \in [0,1])$ and radius $1/2$ as $\mathcal{O}_x = \{G \in D[0,1] : \|G - \Delta_x\| < 1/2\}$. $\mathcal{O}_x$ is open in $D[0,1]$, so for any subset $B \subset [0,1]$, $O_B = \cup_{x \in B} \mathcal{O}_x$ is also open.

For any $x \in [0,1]$, $X_1 = x$ if and only if $\Delta_{X_1} \in \mathcal{O}_x$, so if F_1 is a measurable element then

$$\mathrm{P}(X_1 \in B) = \mathrm{P}(F_1 \in O_B)$$

for any set $B \subset [0,1]$. However, then all subsets of $[0,1]$ are Lebesgue measurable, which is false.

We will overcome this difficulty by using the method of Reeds (1976) and Fernholz (1983). (Note: Dudley (1992, 1994) and Dudley and Norvaiša (1999), and the references therein proposed the use of p-variation norm with high feasibility of Fréchet differentiability. However, we adopt Hadamard differentiation because the usefulness of Dudley's method to the finite population asymptotics is yet unknown to us.)

To start with, we will define the distance between a function $H \in D[0,1]$ and a set $K \subset D[0,1]$ by

$$\mathrm{dist}(H,K) = \inf_{G \in K} \|H - G\|$$

where

$$\|H - G\| = \sup_{0 \le x \le 1} |H(x) - G(x)|.$$

We include the next lemma for completeness.

LEMMA 2.[FERNHOLZ (1983), PP.35-36, LEMMA 4.3.1] *Let $Q : D[0,1] \times \mathbf{R} \to \mathbf{R}$ be a function and for any compact set $K \subset D[0,1]$, we suppose*

$$\lim_{t \to 0} Q(H,t) = 0$$

holds uniformly in $H \in K$. Let $\epsilon > 0$ and let δ_n be a sequence such that $\delta_n \downarrow 0$. Then for any compact set $K \subset D[0,1]$, there exists $n_0 \ge 1$ for which if for all $n > n_0$, $\mathrm{dist(H,K)} \le \delta_\mathrm{n}$ *implies*

$$|Q(H,\delta_n)| < \epsilon.$$

Since, $\sqrt{n}(U_n - U)$ is not a random element of $D[0,1]$, it may not be measurable and no probabilistic statement may be made on this term. In order to overcome the difficulty, Reeds uses the inner probability (Reeds (1976), pp.80-83, Fernholz (1983), p.37). Using his method, we will obtain the next lemma which is a modification of Fernholz (1983), p.37 Lemma 4.3.2 for the finite population.

LEMMA 3. *Let $\tilde{U}_n = \tilde{F}_n \circ F_{NS}^{-1}$. Then, for any $\epsilon > 0$, there is a compact set $K \subset D[0,1]$ and a positive sequence $\delta_{N,n} \downarrow 0$ such that*

$$\mathrm{P_*(dist(\sqrt{Nn/(N-n)}(\tilde{U}_n - U), K) \le \delta_{N,n}) > 1 - \epsilon}.$$

PROOF. Let $(U_n - U)^*$ be the continuous version of $U_n - U$ defined by Rosén (1964): it is a constant 0 function in the intervals $[0,\alpha)$ and $[\beta,1]$ where α and β are the smallest and the largest jump points respectively, and for all $t \in [\alpha,\beta]$, it is obtained by the linear interpolation between the left endpoints of constancy intervals for $U_n - U$. Under Assumption A, the difference between the original function and the continuous version is of the order $O(n^{-1})$, hence we have

$$\|(U_n - U)^* - (U_n - U)\| = O(n^{-1}) \quad \text{a.s.} \tag{3.1}$$

Next, by Proposition 1,

$$\begin{aligned} \|\tilde{U}_n - U_n\| &\le \sup_t |\tilde{F}_n(F_{NS}^{-1}(t)) - F_n(F_{NS}^{-1}(t))| \\ &= \sup_x |\tilde{F}_n(x) - F_n(x)| \\ &= o(n^{-1/2}) \quad \text{a.s.} \end{aligned}$$

Combining above inequalities, we have

$$\begin{aligned}\|(\tilde{U}_n - U) - (U_n - U)^*\| &\leq \|(\tilde{U}_n - U) - (U_n - U)\| + \|(U_n - U) - (U_n - U)^*\| \\ &= o(n^{-1/2}) \quad \text{a.s.}\end{aligned}$$

The random element $(Nn/(N-n))^{1/2}(U_n - U)^*$ converges weakly to Brownian bridge W° in $C[0,1]$ as $\min(n, N-n) \to \infty$ (Rosén (1964)). It follows that the set of probability measures $\mathcal{P} = \{\mathrm{P}_0, \mathrm{P}_n, n \geq 1\}$ is relatively compact, where P_n is a probability measure of $(Nn/(N-n))^{1/2}(U_n - U)^*$ and P_0 denotes the probability measure for Brownian bridge in the space $C[0,1]$. We also note that the space $C[0,1]$ is a completely separable metric space under the uniform norm. It follows from the converse part of the Prohorov's Theorem that $\mathcal{P}$ is tight and therefore for any $\epsilon > 0$ there exists a compact set $K \subset C[0,1]$ for which, for any integer $n \geq 1$,

$$\mathrm{P}_n(K) > 1 - \epsilon.$$

From the definition of P_n, the above equation is tantamount to

$$\mathrm{P}\Big((Nn/(N-n))^{1/2}(U_n - U)^* \in K\Big) > 1 - \epsilon.$$

We note here that remembering the sample $X_i = x_{\pi_i},\ 1 \leq i \leq n$ where $(\pi_1, \ldots, \pi_N)$ take all possible permutations of $(1, \ldots, N)$. Denoting "non"-sample $X_i^*,\ i = 1, \ldots, N-n$ as $X_i^* = x_{\pi_{n+i}}$ and defining the analogue of the empirical distribution function F_{N-n}^* as:

$$F_{N-n}^*(x) = \frac{1}{N-n} \sum_{i=1}^{N-n} I_{(-\infty, x]}(X_i^*).$$

Using the fact

$$n(F_n(x) - F_N(x)) = -(N-n)(F_{N-n}^*(x) - F_N(x)),$$

we have

$$F_n(x) - F_N(x) = -\frac{N-n}{n}(F_{N-n}^*(x) - F_N(x)),$$

or equivalently

$$F_{N-n}^*(x) - F_N(x) = -\frac{n}{N-n}(F_n(x) - F_N(x)).$$

So we may assume here and in the sequel that $n/N \leq 1/2$. (This is essentially the same argument of Erdös and Rényi (1959).)

Since, $C[0,1] \subset D[0,1]$, K is also compact in $D[0,1]$. Since $\|(U_n - U)^* - (\tilde{U}_n - U)\| = o(n^{-\frac{1}{2}})$ and by writing $\sqrt{Nn/(N-n)}(\tilde{U}_n - U) = \sqrt{Nn/(N-n)}\{((\tilde{U}_n - U) - (U_n - U)^*) + (U_n - U)^*\}$, it can be seen easily that whenever $\sqrt{Nn/(N-n)}(U_n - U)^* \in K$ holds we have (using $n/N \leq 1/2$)

$$\mathrm{dist}(\sqrt{Nn/(N-n)}(\tilde{U}_n - U), K) \leq \delta_{N,n}.$$

It follows that

$$\mathrm{P}_*(\mathrm{dist}(\sqrt{Nn/(N-n)}(\tilde{U}_n - U), K) \leq \delta_{N,n}) > 1 - \epsilon.$$

It is interesting to note that the empirical distribution function may not be a measurable function in $D[0,1]$ under the uniform norm, while the remainder term is a measurable function. This can be proved by showing that both the statistical functional and the influence function are measurable, then the remainder term, as a difference of these terms, becomes measurable.

It is now possible to evaluate the error under the probability measure.

LEMMA 4. *If $K \subset D[0,1]$ is an arbitrary compact set for which*

$$\frac{\mathrm{Rem(tH)}}{t} \to 0, \quad \text{as} \quad t \to 0$$

uniformly in $H \in K$, then

$$\sqrt{Nn/(N-n)}\mathrm{Rem}(\tilde{\mathrm{U}}_\mathrm{n} - \mathrm{U}) \xrightarrow{\mathrm{P}} 0.$$

PROOF. Let $\epsilon > 0$. Then by Lemma 3, there exists a compact set $K \subset D[0,1]$ and a positive sequence $\delta_{N,n} \downarrow 0$ for which

$$\mathrm{P}_*(\mathrm{dist}(\sqrt{Nn/(N-n)}(\tilde{U}_n - U), K) \leq \delta_{N,n}) > 1 - \epsilon/2.$$

It follows that there is a measurable set E_n for which

$$E_n \subset \{\mathrm{dist}(\sqrt{Nn/(N-n)}(\tilde{U}_n - U), K) \leq \delta_{N,n}\}$$

and

$$\mathrm{P}(E_n) > 1 - \epsilon$$

for all n.

We now apply Lemma 2 to $Q(H,t) = \mathrm{Rem}(tH)/t$. We find a constant $\delta_{N,\ n} \downarrow 0$ and a positive integer n_0, such that if $n,\ N-n > n_0$ and $\mathrm{dist}(H,K) \leq \delta_{N,\ n}$,

$$|\sqrt{Nn/(N-n)}\,\mathrm{Rem}(\sqrt{(N-n)/Nn}H)| < \epsilon$$

follows. Hence, for all $n,\ N-n > n_0$ and $H = \sqrt{Nn/(N-n)}(\tilde{U}_n - U)$, we have

$$\mathrm{P}(|\sqrt{Nn/(N-n)}\,\mathrm{Rem}(\tilde{U}_n - U)| < \epsilon) \geq \mathrm{P}(E_n) > 1 - \epsilon.$$

Hence the lemma follows.

REMARK 1. *Under the same conditions of Lemma 3 and Lemma 4, we can prove for the unsmoothed case*

$$\sqrt{Nn/(N-n)}\,\mathrm{Rem}(\mathrm{U}_\mathrm{n} - \mathrm{U}) \xrightarrow{\mathrm{P}} 0.$$

PROOF. The proof is similar to and simpler than those of Lemma 3 and Lemma 4.

4. Asymptotic Normality

We will present and prove our main results of this paper. We will give the asymptotic normality of the smoothed statistical functionals under the Hadamard differentiability in Theorem 1. We will also prove in Theorem 2 and Theorem 3 the asymptotic normality for smoothed and non-smoothed functionals respectively, under the condition which is given in terms of the original non-smoothed influence function. Combining Theorem 2 and Theorem 3, we claim that the asymptotic distributions of smoothed statistical functionals are the same as the those of non-smoothed functionals.

For small samples Fernholz (1997) proved that smoothed functionals are more efficient than non-smoothed ones when some regularity conditions are placed on the influence function in I.I.D. settings. For related results for the cases of smoothed bootstrap, we refer readers to Silverman and Young (1987), Hall, DiCiccio and Romano (1989), Polansky and Schucany (1997), and the references therein.

THEOREM 1. *Let $X_1, \ldots, X_n$ be the sequence of a random sample chosen without replacement from the distribution function with the Assumption A. And we also let $\{k_n\}$ be the sequence of regular kernels with finite first moment. Let T be a statistical functional and τ is the induced statistical functional on $D[0,1]$ by (2.3). Suppose τ is Hadamard differentiable at U with the influence function* $\mathrm{IF} = \mathrm{IF}_{\mathrm{T,F_N}}$, *which is Lebesgue-Stieltjes integrable with respect to functions of bounded variations. Suppose* $\overline{\widetilde{\mathrm{IF}}} = \mathrm{E}[\widetilde{\mathrm{IF}}] = \sum_{\mathrm{i}=1}^{\mathrm{N}} \widetilde{\mathrm{IF}}(\mathrm{x_i})/\mathrm{N}$ *and* $0 < \sigma_N^2 = \mathrm{Var}[\widetilde{\mathrm{IF}}] = \sum_{\mathrm{i}=1}^{\mathrm{N}} (\widetilde{\mathrm{IF}}(\mathrm{x_i}) - \overline{\widetilde{\mathrm{IF}}})^2/\mathrm{N} < \infty$.
If the IF *satisfies the Lindeberg condition of Erdös-Rényi;*

$$\lim_{n, N-n \to \infty} \frac{\sum_{P_\tau} (\widetilde{\mathrm{IF}}(x_i) - \overline{\widetilde{\mathrm{IF}}})^2}{\sum_{i=1}^{N} (\widetilde{\mathrm{IF}}(x_i) - \overline{\widetilde{\mathrm{IF}}})^2} = 0 \quad \text{for any} \quad \tau > 0,$$

where, P_τ is the subset of $P = \{1, \ldots, N\}$ such that

$$|\widetilde{\mathrm{IF}}(x_i) - \overline{\widetilde{\mathrm{IF}}}| > \tau \sqrt{n\Big(1 - \frac{n}{N}\Big)} \sigma_N.$$

Then , as n and $N - n \to \infty$,

$$(T(\widetilde{F}_n) - T(F_{NS}))/\sigma_{N,n} \xrightarrow{d} N(0,1),$$

where $\sigma_{N,n}^2 = N\sigma_N^2(1 - n/N)/n(N-1) = (N-n)\sigma_N^2/n(N-1)$.

PROOF. It follows from Lemma 1 that

$$\begin{aligned} & (T(\widetilde{F}_n) - T(F_{NS}))/\sigma_{N,n} \\ = & (\tau(\widetilde{U}_n) - \tau(U))/\sigma_{N,n} \\ = & \tau'_U(\widetilde{U}_n - U)/\sigma_{N,n} + \mathrm{Rem}(\widetilde{U}_n - U)/\sigma_{N,n} \\ = & \Big\{\frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathrm{IF}}(X_i)\Big\}/\sigma_{N,n} + \mathrm{Rem}(\widetilde{U}_n - U)/\sigma_{N,n}. \end{aligned}$$

Since both the influence function and the functional are measurable, the above equation tells us that $\mathrm{Rem}(\widetilde{U}_n - U)$ is an random element in $D[0,1]$. Also, by the central limit theorem of Erdös and Rényi (1959) and Hájek (1960), the first term in the right most side of the above equation converges in distribution to $N(0,1)$ as n and $N-n \to \infty$. Also, we note that by Lemma 4 the second term $\mathrm{Rem}(\widetilde{U}_n - U)/\sigma_{N,n}(= \sqrt{nN/(N-n)}\sqrt{(N-1)/N}\mathrm{Rem}(\widetilde{U}_n - U)/\sigma_N = \sqrt{n(N-1)/(N-n)}\mathrm{Rem}(\widetilde{U}_n - U)/\sigma_N)$ converges to 0 in probability. The theorem follows from Slutzky's lemma.

The next theorem is the central limit theorem for smoothed functionals under the Erdös-Rényi condition for the original non-smoothed influence function, which is the extension of Theorem 1 of Fernholz (1993). In order to prove the theorem, we need the following lemma which is a modification of Lemma 2 of Fernholz (1993).

LEMMA 5. *Let $\{k_n\}$ be the sequence of regular kernels with finite first moment, and $X_1, \ldots, X_n$ be the sequence of a random sample chosen without replacement from the distribution function with Assumption A. Suppose that F_{NS} defined before have bounded derivatives $f_{NS} = F'_{NS}$. If the function ϕ is bounded function and Lebesgue-Stieltjes integrable with respect to functions of bounded variation.*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\tilde{\phi}(X_i) - \phi(X_i)) \xrightarrow{p} 0.$$

PROOF. First we write

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\tilde{\phi}(X_i) - \phi(X_i)) &= \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}(\tilde{\phi}(X_i) - \phi(X_i)) \\ &= \sqrt{n}\int(\tilde{\phi}(x) - \phi(x))dF_n(x) \\ &= \sqrt{n}\iint(\phi(x-t) - \phi(x))k_n(t)dt\, dF_n(x).\end{aligned}$$

We divide the range of integration of the inner integral,

$$\begin{aligned}\sqrt{n}\int(\phi(x-t) - \phi(x))k_n(t)dt &= \sqrt{n}\int_{|t|\le b_n}(\phi(x-t) - \phi(x))k_n(t)dt \\ &\quad + \sqrt{n}\int_{|t|>b_n}(\phi(x-t) - \phi(x))k_n(t)dt.\end{aligned}$$

The second term of the above equation is dominated by

$$\sqrt{n}\int_{|t|>b_n}|\phi(x-t) - \phi(x)||k_n(t)|dt,$$

which converges to 0 from the definition of the regular kernels and the fact that ϕ is a bounded function.

As for the first term, we integrate the term with respect to F_N

$$\begin{aligned}&\sqrt{n}\iint_{|t|\le b_n}(\phi(x-t) - \phi(x))k_n(t)dtdF_N(x) \\ &= \sqrt{n}\int_{|t|\le b_n}\int(\phi(x-t) - \phi(x))dF_N(x)k_n(t)dt.\end{aligned}$$

From the assumption, a sequence of functions F_{NS} defined before have bounded derivatives f_{NS} for which

$$\|F_N - F_{NS}\| = O(\frac{1}{N}).$$

From the assumption that function ϕ is bounded, it follows that there is a positive constant $\mathrm{C} > 0$, such that

$$\begin{aligned}
|\int(\phi(x-t)-\phi(x))dF_{NS}(x)| &= |\int \phi(x-t)f_{NS}(x)dx - \int \phi(x)f_{NS}(x)dx| \\
&= |\int \phi(x)(f_{NS}(x) - f_{NS}(x+t))dx| \\
&\leq \mathrm{C}|t|.
\end{aligned}$$

Hence,

$$\begin{aligned}
&|\int(\phi(x-t)-\phi(x))dF_N(x)| \\
&= |\int(\phi(x-t)-\phi(x))d(F_N(x) - F_{NS}(x) + F_{NS}(x))| \\
&\leq |\int(\phi(x-t)-\phi(x))d(F_N(x) - F_{NS}(x))| \\
&\quad +|\int(\phi(x-t)-\phi(x))dF_{NS}(x)| \\
&\leq O\Big(\frac{1}{N}\Big) + \mathrm{C}|t|
\end{aligned}$$

from the assumption that ϕ is bounded.

By substituting the above into the original equation, we have

$$\begin{aligned}
&\leq \mathrm{C}\sqrt{n}\int_{|t|\leq b_n} |t||k_n(t)|dt + O\Big(\frac{\sqrt{n}b_n}{N}\Big) \\
&\leq 2\mathrm{C}\sqrt{n}b_n + o\Big(\frac{1}{N}\Big) \to 0.
\end{aligned}$$

THEOREM 2. *In addition to the conditions of Theorem 1, we assume that* $\mathrm{IF} = \mathrm{IF}_{\mathrm{T,F_{NS}}}$ *is a bounded function and* F_{NS} *has bounded derivatives.*
Also suppose $\overline{\mathrm{IF}} = \mathrm{E[IF]} = \sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{IF(x_i)/N}$ *and* $0 < \sigma_N^2 = \mathrm{Var[IF]} = \sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{IF(x_i)} - \overline{\mathrm{IF}})^2/\mathrm{N} < \infty$.
Replacing the Erdös-Rényi condition of Theorem 1, if the IF *satisfies the Lindeberg condition of Erdös-Rényi;*

$$\lim_{n,N-n\to\infty} \frac{\sum_{P_\tau}(\mathrm{IF(x_i)} - \overline{\mathrm{IF}})^2}{\sum_{i=1}^{N}(\mathrm{IF(x_i)} - \overline{\mathrm{IF}})^2} = 0 \quad \textit{for any} \quad \tau > 0,$$

where, P_τ *is the subset of* $P = \{1, \ldots, N\}$ *such that*

$$|\mathrm{IF(x_i)} - \overline{\mathrm{IF}}| > \tau\sqrt{\mathrm{n}\Big(1 - \frac{\mathrm{n}}{\mathrm{N}}\Big)}\sigma_{\mathrm{N}}.$$

Then, as n and $N-n \to \infty$,

$$(T(\widetilde{F}_n) - T(F_{NS}))/\sigma_{N,n} \xrightarrow{d} N(0,1),$$

where $\sigma^2_{N,n} = N\sigma^2_N(1-n/N)/n(N-1) = (N-n)\sigma^2_N/n(N-1)$.

PROOF. It follows from Lemma 1 that

$$\begin{aligned}(T(\widetilde{F}_n) - T(F_{NS}))/\sigma_{N,n} &= (\tau(\widetilde{U}_n) - \tau(U))/\sigma_{N,n} \\ &= \tau'_U(\widetilde{U}_n - U)/\sigma_{N,n} + \mathrm{Rem}(\widetilde{U}_n - U)/\sigma_{N,n} \\ &= \Big\{\frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathrm{IF}}(X_i)\Big\}/\sigma_{N,n} + \mathrm{Rem}(\widetilde{U}_n - U)/\sigma_{N,n}.\end{aligned}$$

Using the argument of Erdös and Rényi (1959) (the same argument in the proof of Lemma 3). Remembering the sample $X_i = x_{\pi_i},\ 1 \le i \le n$ where $(\pi_1, \ldots, \pi_N)$ take all possible permutations of $(1, \ldots, N)$. Denote "non"-sample $X^*_i,\ i = 1, \ldots N-n$ as $X^*_i = x_{\pi_{n+i}}$.

Using the fact

$$\sum_{i=1}^{n} \mathrm{IF}(X_i) + \sum_{i=1}^{N-n} \mathrm{IF}(X^*_i) = \sum_{i=1}^{N} \mathrm{IF}(x_i),$$

we have

$$\frac{1}{n}(\sum_{i=1}^{n} \mathrm{IF}(X_i) - \sum_{i=1}^{N} \mathrm{IF}(x_i)) = -\frac{1}{N-n}(\sum_{i=1}^{N-n} \mathrm{IF}(X^*_i) - \sum_{i=1}^{N} \mathrm{IF}(x_i)).$$

So we may assume that $n/N \le 1/2$.

By Lemma 5, we have (using $n/N \le 1/2$)

$$\sqrt{N/n(N-n)}\sum_{i=1}^{n}(\widetilde{\mathrm{IF}}(X_i) - \mathrm{IF}(X_i)) \xrightarrow{p} 0$$

The theorem follows from the central limit theorem for the finite population (Erdös and Rényi (1959) and Hájek (1960)), and Lemma 4.

The last theorem, when compared to the previous Theorem 2, displays that both the smoothed and non-smoothed statistical functional have the same distribution in the limit.

THEOREM 3. *Under the same conditions of Theorem 2, we have, as n and $N-n$ $\to \infty$,*

$$(T(F_n) - T(F_{NS}))/\sigma_{N,n} \xrightarrow{d} N(0,1),$$

where $\sigma^2_{N,n} = N\sigma^2_N(1-n/N)/n(N-1) = (N-n)\sigma^2_N/n(N-1)$, $0 < \sigma^2_N = \mathrm{Var[IF]} = \sum_{i=1}^{N}(\mathrm{IF(x_i)} - \overline{\mathrm{IF}})^2/\mathrm{N} < \infty$, *and* $\overline{\mathrm{IF}} = \mathrm{E[IF]} = \sum_{i=1}^{N} \mathrm{IF(x_i)/N}$.

PROOF. The proof is performed in a similar manner as those of Theorem 1 and Theorem 2.

We first note

$$\begin{aligned}&(T(F_n)-T(F_{NS}))/\sigma_{N,n}\\&=(\tau(U_n)-\tau(U))/\sigma_{N,n}\\&=\tau'_U(U_n-U)/\sigma_{N,n}+\mathrm{Rem}(U_n-U)/\sigma_{N,n}\\&=\Big\{\frac{1}{n}\sum_{i=1}^{n}\mathrm{IF}(X_i)\Big\}/\sigma_{N,n}+\mathrm{Rem}(U_n-U)/\sigma_{N,n}.\end{aligned}$$

Since both the influence function and the functional are measurable, the above equation tells us that $\mathrm{Rem}(U_n-U)$ is a random element in $D[0,1]$. Also, by the central limit theorem of Erdös and Rényi (1959) and Hájek (1960), the first term in the right most side of the above equation converges in distribution to $N(0,1)$ as $n, N-n\to\infty$. Also, we note that by Remark 1, the second term $\mathrm{Rem}(U_n-U)/\sigma_{N,n}(=\sqrt{nN/(N-n)}\sqrt{(N-1)/N}\mathrm{Rem}(U_n-U)/\sigma_N=\sqrt{n(N-1)/(N-n)}\mathrm{Rem}(U_n-U)/\sigma_N)$ converges to 0 in probability. The theorem follows from Slutzky's lemma.

5. Monte Carlo Simulation

In this section, we present Monte Carlo simulation results for sample median

$$F_n^{-1}(0.5) \tag{5.1}$$

and inter-quartile range

$$F_n^{-1}(0.75)-F_n^{-1}(0.25). \tag{5.2}$$

These statistics are often used (ex. Deaton (1997) p.22 Table 1.2. Consumption and income for panel households, Côte d'Ivoire, 1985-86) because they are less affected by outliers which may appear in wide class of data.

5.1. Quasi-populations

We use the following simulated log-normal quasi-populations of different sizes whose mean and standard deviation of the distribution on the log scale is 3 and 0.4 respectively.

(i) generated values of log normal random number of size $1,000$

(ii) generated values of log normal random number of size $5,000$

(iii) generated values of log normal random number of size $10,000$

We use these populations because they approximate many economic variables such as household income and savings.

5.2. Construction of the Confidence Intervals

To construct the confidence intervals, we calculate the influence functions of the median and the inter-quartile range.

As is well-known(cf. Huber (1981)), the influence function of a non-smoothed median is

$$\mathrm{IF}_{T,F}(x) = \frac{\Delta_{x-F^{-1}(0.5)} - 0.5}{f(F^{-1}(0.5))} = \begin{cases} \frac{-1}{2f(F^{-1}(0.5))} & x \leq F^{-1}(0.5) \\ \frac{1}{2f(F^{-1}(0.5))} & x > F^{-1}(0.5) \end{cases}$$

and the influence function of a non-smoothed inter-quartile range is

$$\mathrm{IF}_{T,F}(x) = \frac{\Delta_{x-F^{-1}(0.75)} - 1 + 0.75}{f(F^{-1}(0.75))} - \frac{\Delta_{x-F^{-1}(0.25)} - 1 + 0.25}{f(F^{-1}(0.25))},$$

where F is an underlying distribution function and f is its density.

In order to obtain the influence functions of the smoothed functionals, we utilize the following Lemma of Fernholz (1993).

LEMMA 6.[FERNHOLZ (1993), PROPOSITION 2] *Let k be fixed, and define the smoothed functional $\widetilde{T}$ for general distribution function G as;*

$$\widetilde{T}(G) = T(G * k). \tag{5.3}$$

If T is Gateaux differentiable in a neighborhood of F_{NS} including $\widetilde{F} = F * k$, then

$$\mathrm{IF}_{\widetilde{\mathrm{T}},\mathrm{F}} = \mathrm{IF}_{\mathrm{T},\widetilde{\mathrm{F}}} * \mathrm{k}. \tag{5.4}$$

Utilizing this lemma, we calculate the influence functions of smoothed functionals.

Let $K(x) = \int_{-\infty}^{x} k(t)dt$ be the distribution function with respect to kernel k, and let $\tilde{f}$ be the density function with respect to the smoothed (cumulative) distribution $\widetilde{F} = F * k$. Then, the influence function of the smoothed median is

$$\mathrm{IF}_{\widetilde{T},F}(x) = \frac{K(x - F^{-1}(0.5)) - 0.5}{\tilde{f}(F^{-1}(0.5))}$$

and the influence function of the smoothed inter-quartile range is

$$\mathrm{IF}_{\widetilde{T},F}(x) = \frac{K(x - F^{-1}(0.75)) - 1 + 0.75}{\tilde{f}(F^{-1}(0.75))} - \frac{K(x - F^{-1}(0.25)) - 1 + 0.25}{\tilde{f}(F^{-1}(0.25))}.$$

In our simulations, we smooth two functionals with uniform distribution $U[-0.1, 0.1]$, which satisfies the conditions of regular kernel and it is easy to calculate the convolutions.

Using these influence functions, we construct the confidence intervals as follows(we describe the non-smoothed case, smoothed case is similar):

(i) Calculate the variance of the influence function σ_N^2 as

$$\sigma_N^2 = \mathrm{Var}[\mathrm{IF}] = \sum_{i=1}^{N}(\mathrm{IF}(x_i) - \overline{\mathrm{IF}})^2/N, \tag{5.5}$$

$$\overline{\mathrm{IF}} = \mathrm{E}[\mathrm{IF}] = \sum_{i=1}^{N} \mathrm{IF}(x_i)/N. \tag{5.6}$$

(ii) Approximate

$$\frac{T(F_n) - T(F_N)}{\sigma_{N,n}} \tag{5.7}$$

with standard normal distribution ($\sigma_{N,n}^2 = N\sigma_N^2(1 - n/N)/n(N-1) = (N - n)\sigma_N^2/n(N-1)$).

(iii) Construct a one-sided $100(1-\alpha)$ confidence interval as

$$\big(-\infty, T(F_n) + \sigma_{N,n} z_\alpha\big] \tag{5.8}$$

and a two-sided $100(1-\alpha)$ confidence interval as

$$\big[T(F_n) - \sigma_{N,n} z_{\alpha/2}, T(F_n) + \sigma_{N,n} z_{\alpha/2}\big] \tag{5.9}$$

where z_α is the upper $\alpha\%$ point of the standard normal distribution.

5.3. Results of Monte Carlo Simulations

In order to evaluate the fruits of theoretical facts, we calculate the empirical coverage ratio of confidence intervals constructed by our normal approximations. The simulated samples of sampling fractions 10% and 30% are chosen 100,000 times repeatedly, then the relative frequencies that the intervals contain the true value of parameter are evaluated. We can judge that the intervals are precise when the empirical coverage probability is close to the nominal confidence coefficient. Although readers may claim that these sampling fractions are extraordinarily high in reality, in a stratified population or in a selected unit of population, fractions of these types are not exceptional.

In what follows, the numbers in parentheses in tables for two-sided intervals are the lengths of the intervals.

As a whole, except for the slightest difference between the smoothed and the unsmoothed cases, we can see the following features:

(i) All the intervals, especially the cases with population sizes larger than 5000, display better features for both the one-sided and two-sided situations.

(ii) The case when the sampling fraction is 10% is better than the case of a sampling fraction 30% in small sample situations.

(iii) The intervals of higher confidence coefficients perform better than those of lower confidence coefficients.

Compared to the case of a sampling fraction of 10%, the case of 30% seems to be not as good when population and sample sizes are relatively small. However, our normal approximations show good performance for large size samples, so they are very useful for application to large scale sample surveys.

Table 1. One-sided confidence intervals for sample medians(sampling fraction10%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 100 population size 1000	0.95278	0.98441	0.99839
sample size 500 population size 5000	0.91295	0.96173	0.99483
sample size 1000 population size 10000	0.89981	0.9559	0.99309

Table 2. Two-sided confidence intervals for sample medians(sampling fraction10%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 100 population size 1000	0.93512 (3.142083)	0.96572 (3.744022)	0.99494 (4.920478)
sample size 500 population size 5000	0.90636 (1.405430)	0.95292 (1.674672)	0.99251 (2.200892)
sample size 1000 population size 10000	0.89818 (0.9937106)	0.94831 (1.184079)	0.99091 (1.556143)

Table 3. One-sided confidence intervals for sample medians(sampling fraction30%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 300 population size 1000	0.93353	0.98207	0.99866
sample size 1500 population size 5000	0.90336	0.95278	0.99256
sample size 3000 population size 10000	0.93079	0.95512	0.99066

Table 4. Two-sided confidence intervals for sample medians(sampling fraction30%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 300 population size 1000	0.91039 (1.599870)	0.96548 (1.906363)	0.99424 (2.505385)
sample size 1500 population size 5000	0.88282 (0.7156098)	0.93482 (0.8527015)	0.99091 (1.120640)
sample size 3000 population size 10000	0.90292 (0.5059727)	0.94974 (0.6029035)	0.98661 (0.7923494)

Table 5. One-sided confidence intervals for smoothed sample medians(sampling fraction10%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 100 population size 1000	0.95445	0.98435	0.99834
sample size 500 population size 5000	0.91447	0.96367	0.99552
sample size 1000 population size 10000	0.90483	0.95786	0.993

Table 6. Two-sided confidence intervals for smoothed sample medians(sampling fraction10%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 100	0.93529	0.96752	0.99521
population size 1000	(3.165274)	(3.771655)	(4.956795)
sample size 500	0.90714	0.95581	0.99304
population size 5000	(1.415939)	(1.687195)	(2.217350)
sample size 1000	0.89895	0.95001	0.99115
population size 10000	(1.000993)	(1.192757)	(1.567548)

Table 7. One-sided confidence intervals for smoothed sample medians(sampling fraction30%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 300	0.93533	0.98113	0.9987
population size 1000			
sample size 1500	0.91001	0.95471	0.99343
population size 5000			
sample size 3000	0.93079	0.95512	0.99066
population size 10000			

Table 8. Two-sided confidence intervals for smoothed sample medians(sampling fraction30%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 300	0.90517	0.96556	0.99476
population size 1000	(1.611678)	(1.920433)	(2.523876)
sample size 1500	0.88967	0.94752	0.99132
population size 5000	(0.720961)	(0.8590779)	(1.129020)
sample size 3000	0.90915	0.95047	0.98905
population size 10000	(0.5096809)	(0.6073221)	(0.7981564)

Table 9. One-sided confidence intervals for sample inter-quartile ranges(sampling fraction10%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 100 population size 1000	0.87794	0.94572	0.99329
sample size 500 population size 5000	0.90063	0.95342	0.99092
sample size 1000 population size 10000	0.89934	0.95181	0.99202

Table 10. Two-sided confidence intervals for sample inter-quartile ranges(sampling fraction10%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 100	0.92177	0.96616	0.99485
population size 1000	(4.249439)	(5.063518)	(6.65459)
sample size 500	0.91433	0.95735	0.99087
population size 5000	(1.943841)	(2.316229)	(3.044041)
sample size 1000	0.919	0.96182	0.99323
population size 10000	(1.373149)	(1.636208)	(2.150341)

Table 11. One-sided confidence intervals for sample inter-quartile ranges(sampling fraction30%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 300 population size 1000	0.85486	0.9187	0.98698
sample size 1500 population size 5000	0.91475	0.95681	0.99182
sample size 3000 population size 10000	0.91218	0.95477	0.99003

Table 12. Two-sided confidence intervals for sample inter-quartile ranges(sampling fraction30%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 300	0.89541	0.9546	0.99358
population size 1000	(2.163709)	(2.578217)	(3.388351)
sample size 1500	0.91531	0.96037	0.99353
population size 5000	(0.9897555)	(1.179366)	(1.549950)
sample size 3000	0.9172	0.96021	0.99273
population size 10000	(0.6991734)	(0.8331163)	(1.094900)

Table 13. One-sided confidence intervals for smoothed sample inter-quartile ranges(sampling fraction10%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 100	0.879	0.94585	0.99331
population size 1000			
sample size 500	0.9024	0.95407	0.99073
population size 5000			
sample size 1000	0.89808	0.95104	0.99214
population size 10000			

Table 14. Two-sided confidence intervals for smoothed sample inter-quartile ranges(sampling fraction10%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 100	0.92194	0.96566	0.99518
population size 1000	(4.258238)	(5.074003)	(6.668369)
sample size 500	0.91574	0.95783	0.99081
population size 5000	(1.950190)	(2.323794)	(3.053983)
sample size 1000	0.91763	0.96177	0.99357
population size 10000	(1.376929)	(1.640712)	(2.156260)

Table 15. One-sided confidence intervals for smoothed sample inter-quartile ranges(sampling fraction30%).

	90%(one-sided)	95%(one-sided)	99%(one-sided)
sample size 300 population size 1000	0.8555	0.92003	0.98739
sample size 1500 population size 5000	0.91652	0.9582	0.99218
sample size 3000 population size 10000	0.91179	0.95488	0.99038

Table 16. Two-sided confidence intervals for smoothed sample inter-quartile ranges(sampling fraction30%, length of intervals in parentheses).

	90%(two-sided)	95%(two-sided)	99%(two-sided)
sample size 300 population size 1000	0.89756 (2.168189)	0.95435 (2.583556)	0.99312 (3.395367)
sample size 1500 population size 5000	0.91756 (0.9929883)	0.96083 (1.183218)	0.99348 (1.555012)
sample size 3000 population size 10000	0.91782 (0.7010979)	0.96035 (0.8354095)	0.99291 (1.097914)

Appendix

The proof of Proposition 1 will be shown along the lines of Fernholz (1991), and for this purpose we first show the following lemma. A covering C of $\mathbf{R}$ is a collection of intervals whose union is $\mathbf{R}$. The intervals need not have finite length.

LEMMA 6. *Let $X_1, \ldots, X_n$ be a simple random sample without replacement from a finite population with distribution function F_N. Let $\{C_n\}$ be a sequence of coverings of $\mathbf{R}$ such that the number of intervals in each C_n is $O(n^\lambda)$ for some constant λ. Suppose that $\max_{I \in C_n} P_{F_N}(I) = \max_{I \in C_n} P_{F_N}(X \in I) = o(n^{-1/2})$ where P_{F_N} stands for the probability generated by F_N. If T_n is the maximum number of X's with values in any $I \in C_n$, then*

$$\frac{T_n}{\sqrt{n}} \to 0,$$

as $n, N \to \infty$ a.s..

PROOF. For each $n \geq 1$, define Y_I as the number of X_i's with values in $I \in C_n$. Then Y_I is a random variable from a hypergeometric distribution with probability function

$$p(y) = \frac{\binom{N\pi^*}{y}\binom{N(1-\pi^*)}{n-y}}{\binom{N}{n}}, \quad y = 0, 1, \ldots, \min(n, N\pi^*),$$

where $\pi^* = \mathrm{P}_{F_N}(I)$. Note that $\pi^* = o(n^{-1/2})$ as $n, N \to \infty$.

For any $\epsilon > 0$, and $k = \lfloor \epsilon\sqrt{n} \rfloor$, we have

$$\mathrm{P}(T_n > \epsilon\sqrt{n}) \leq \sum_{I \in C_n} \mathrm{P}(Y_I > \epsilon\sqrt{n}) = \sum_{I \in C_n} \mathrm{P}(Y_I \geq k).$$

Using the well known identity of hypergeometric distribution,

$$\frac{p(y+1)}{p(y)} = \frac{(N\pi^* - y)(n-y)}{(y+1)(N - N\pi^* - n + y + 1)}, \quad \max(0, n + N\pi^* - N) \leq y \leq \min(n, N\pi^*) - 1,$$

we have the following inequalities for $k \geq n\pi^*$ (See Feller (1968) for binomial case).

$$\begin{aligned}\mathrm{P}(Y_I > k) &\leq \mathrm{P}(Y_I = k)\frac{(k+1)((N-n)(1-\pi^*) + (n-k) + 1)}{(k+1)((N-n)(1-\pi^*) + (n-k) + 1) - (N\pi^* - k)(n-k)} \\ &= \mathrm{P}(Y_I = k)O(1),\end{aligned}$$

from $\pi^* = o(n^{-1/2})$.

It follows from Stirling's formula that

$$\begin{aligned}\mathrm{P}(Y_I = k) \sim &\left(\frac{(N\pi^*)(N - N\pi^*)n(N-n)}{2\pi k(N\pi^* - k)N(n-k)(N - n - (N\pi^* - k))}\right)^{1/2} \\ &\left(\frac{N(1-\pi^*)(N-n)}{N(N - n - (N\pi^* - k))}\right)^N \left(\frac{N\pi^*}{N(1-\pi^*)(N(1-\pi^*) - (n-k))}\right)^{N\pi^*} \\ &\left(\frac{(N\pi^* - k)(n-k)}{k(N(1-\pi^*) - (n-k))}\right)^k \left(\frac{n(N - n - (N\pi^* - k))}{(n-k)(N-n)}\right)^n.\end{aligned}$$

Here we have

$$\left(\frac{(N\pi^*)(N-N\pi^*)n(N-n)}{2\pi k(N\pi^*-k)N(n-k)(N-n-(N\pi^*-k))}\right)^{1/2} = O(n^{-1/4})$$
$$\left(\frac{N(1-\pi^*)(N-n)}{N(N-n-(N\pi^*-k))}\right)^{N} = O(1)$$
$$\left(\frac{N\pi^*}{N(1-\pi^*)(N(1-\pi^*)-(n-k))}\right)^{N\pi^*} = O(1).$$

So for sufficiently large n, N and a constant $\mathrm{C} < \infty$,

$$\mathrm{P}(Y_I > k) \le \mathrm{C}\Big(\frac{n\pi^*}{\epsilon\sqrt{n}}\Big)^{\epsilon\sqrt{n}}\Big(\frac{n}{n-\epsilon\sqrt{n}}\Big)^{n}.$$

Recalling the fact

$$\Big(\frac{n}{n-\epsilon\sqrt{n}}\Big)^{n} = e^{\epsilon\sqrt{n}+O(1)},$$

we have

$$\begin{aligned}\mathrm{P}(Y_I > k) &\le \mathrm{C}\Big(\frac{n\pi^*}{\epsilon\sqrt{n}}\Big)^{\epsilon\sqrt{n}} e^{\epsilon\sqrt{n}} \\ &\le \mathrm{C}' e^{-\epsilon\sqrt{n}}\end{aligned}$$

for some constants C and C$'$ and sufficiently large n, as $k = \lfloor \epsilon\sqrt{n} \rfloor \ge n\pi^*$ for large n.

Therefore

$$\mathrm{P}(T_n > \epsilon\sqrt{n}) \le \mathrm{C}' \sum_{I \in C_n} e^{-\epsilon\sqrt{n}} = O(n^{\lambda})e^{-\epsilon\sqrt{n}}$$

since the number of intervals in C_n is $O(n^{\lambda})$. Hence

$$\sum_{n=1}^{\infty} \mathrm{P}(T_n > \epsilon\sqrt{n}) < \infty,$$

and since $\epsilon > 0$ was arbitrary, the first Borel-Cantelli Lemma implies that

$$\frac{T_n}{\sqrt{n}} \to 0 \quad \text{a.s.}$$

Now we will prove the proposition.

Proof. Define the function Q on $\mathbf{R}_+$ by

$$Q(t) = \sup_x (F_N(x+t) - F_N(x)).$$

It is easily seen that $Q(0) = 0$, $\lim_{t\to\infty} Q(t) = 1$,and

$$\begin{aligned}Q(s+t) &= \sup_x (F_N(x+s+t) - F_N(x)) \\ &= \sup_x (F_N(x+s+t) - F_N(x+t) + F_N(x+t) - F_N(x)) \\ &\le \sup_x (F_N(x+s+t) - F_N(x+t)) + \sup_x (F_N(x+t) - F_N(x)) \\ &= Q(s) + Q(t).\end{aligned}$$

Since $\{k_n\}$ is regular, there exists a positive sequence $\{b_n\}$ satisfying the condition of Definition 1. We may assume $b_n^{-1} = o(n)$ without any restrictions, or else we may replace b_n by $\max\{b_n, n^{-3/4}\}$.

Let $Q_n = Q(b_n)$ and define $x_0 = -\infty, x_i = F_N^{-1}(iQ_n)$ for $iQ_n < 1$ and $x_k = \infty$ for $k = \inf\{i : iQ_n \geq 1\}$. The intervals $I_1 = (-\infty, x_1], I_2 = (x_1, x_2], \ldots, I_k = (x_{k-1}, \infty)$ define a covering C_n of $\mathbf{R}$.

From Assumption A,

$$|Q(t) - Q(s)| \leq |\sup_x(F_N(x+t) - F_N(x)) - \sup_x(F_N(x+s) - F_N(x))| \leq 2M|t-s| + O\Big(\frac{1}{N}\Big).$$

So, we have

$$\mathrm{P}_{F_N}(I_j) = F_N(x_j) - F_N(x_{j-1}) \leq 2Q_n \leq 4Mb_n + O\Big(\frac{1}{N}\Big) = o(n^{-1/2}).$$

The number of intervals in C_n satisfies the relationship $k \leq Q_n^{-1} + 1$. Since $b_n^{-1} = o(n)$ as $n \to \infty$ we have $nb_n \to \infty$ and hence $Q(nb_n) \to 1$, so for sufficiently large n, $Q(nb_n) > 1/2$. We see $Q_n^{-1} = O(n)$ from $Q(nb_n) \leq nQ(b_n) = nQ_n$. Therefore $\{C_n\}$ satisfies the assumption of Lemma 1.

For any $x \in \mathbf{R}$, $x \in I_j$ for some j, we have $(x - b_n) \in I_j \cup I_{j-1}$ and $(x + b_n) \in I_j \cup I_{j+1}$. Hence if $|t| \leq b_n$, then

$$\sqrt{n}|F_n(x-t) - F_n(x)| \leq \frac{2T_n}{\sqrt{n}},$$

where T_n is defined as in Lemma 6. Therefore,

$$\begin{aligned}
\sqrt{n}|\widetilde{F}_n(x) - F_n(x)| &\leq \sqrt{n}\int |F_n(x-t) - F_n(x)||k_n(t)|dt \\
&\leq \sqrt{n}\int_{|t|\leq b_n} |F_n(x-t) - F_n(x)||k_n(t)|dt \\
&\quad + \sqrt{n}\int_{|t|>b_n} |F_n(x-t) - F_n(x)||k_n(t)|dt \\
&\leq \frac{2T_n}{\sqrt{n}}\int_{|t|\leq b_n} |k_n(t)|dt + \sqrt{n}\int_{|t|>b_n} |k_n(t)|dt.
\end{aligned}$$

The first term of the last inequality converges almost sure to zero by Lemma 1 and the second term converges to zero since $\{k_n\}$ is a regular sequence. The proposition follows.

Corollary 1. *Under the assumption of Proposition 1,*
(a) The smoothed and non-smoothed Kolmogorov-Smirnov statistics $\sqrt{n}\sup_x|\widetilde{F}_n(x) - F_N(x)|$ *and* $\sqrt{n}\sup_x|F_n(x) - F_N(x)|$ *have the same asymptotic distribution.*
(b) For x *such that* $0 < f_1 \leq F_N(x) \leq f_2 < 1$ *in sufficiently large* N*, the normalized smoothed empirical process has a normal distribution as a limit*

$$\sqrt{n}(\widetilde{F}_n(x) - F_N(x))/\sigma_{N,n} \xrightarrow{d} N(0,1), \quad n, N-n \to \infty,$$

where $\sigma_{N,n} = ((N-n)/(N-1))F_N(x)(1-F_N(x))$.

PROOF. By triangular inequality,

$$\begin{aligned}
&\sqrt{n}\sup_x |F_n(x) - F_N(x)| - \sqrt{n}\sup_x |\widetilde{F}_n(x) - F_n(x)| \\
&\leq \sqrt{n}\sup_x |\widetilde{F}_n(x) - F_N(x)| \\
&\leq \sqrt{n}\sup_x |F_n(x) - F_N(x)| + \sqrt{n}\sup_x |\widetilde{F}_n(x) - F_n(x)|.
\end{aligned}$$

Hence, (a) follows from proposition 1.

Dividing

$$\sqrt{n}\{\widetilde{F}_n(x) - F_N(x)\} = \sqrt{n}\{\widetilde{F}_n(x) - F_n(x)\} + \sqrt{n}\{F_n(x) - F_N(x)\},$$

(b) follows from the proposition and the asymptotic normality of the hypergeometric distribution(Eeden and Runnenburg (1960)): $\sqrt{n}\{F_n(x) - F_N(x)\}/\sigma \xrightarrow{d} N(0,1)$ as n and $N - n \to \infty$.

Acknowledgements

We would like to thank an anonymous referee for valuable comments and constructive suggestions that helped to much improve the paper.

REFERENCES

[1] Billingsley, P. (1999). *Convergence of Probability Measures, 2nd. ed.*, John Wiley & Sons.

[2] Campbell, C. (1980). A Different View of Finite Population Estimation, *Proceedings of the Survey Research Methods Section, ASA*, 319–324.

[3] Csörgö, M. and L. Horváth (1995). On the Distance between Smoothed Empirical and Quantile Processes, *The Annals of Statistics*, **23**, 113–131.

[4] Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, The Johns Hopkins University Press.

[5] Dudley, R. (1992). Frechet Differentiability, p-Variation and Uniform Donsker Classes, *The Annals of Probability*, **20**, 1968–1982.

[6] Dudley, R. (1994). The Order of the Remainder in Derivatives of Composition and Inverse Operators for p-Variation Norms, *The Annals of Statistics*, **22**, 1–22.

[7] Dudley, R. and R. Norvaiša (1999). *Lecture Notes in Mathematics # 1703 Differentiability of Six Operators on Nonsmooth Functions and p-Variation*, Springer-Verlag.

[8] Erdös, P. and A. Rényi (1959). On the Central Limit Theorem for Samples from a Finite Population, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **4**, 49–61.

[9] Eeden, van C. and J.Th. Runnenburg (1960). Conditional Limit-distributions for the Entries in a 2×2-table, *Statistica Neerlandica*, **14**, 111–126.

[10] Falk, M. (1985). Asymptotic Normality of the Kernel Quantile Estimator, *The Annals of Statistics*, **13**, 428–433.

[11] Feller, W. (1968). *An Introduction to Probability Theory 3rd ed.*, John-Wiley & Sons, Inc.

[12] Fernholz, L.T. (1983). *Lecture Notes in Statistics # 19 von Mises Calculus for Statistical Functionals*, Springer-Verlag.

[13] Fernholz, L.T. (1991). Almost Sure Convergence of Smoothed Empirical Distribution Functions, *Scandinavian Journal of Statistics*, **18**, 255–262.

[14] Fernholz, L.T. (1993). Smoothed Versions of Statistical Functionals, in *New Directions in Statistical Data Analysis and Robustness*, Edited by Morgenthalter, S. Ronchetti, E. and W. A. Stahel., Birkhäuser Verlag.

[15] Fernholz, L.T. (1997). Reducing the Variance by Smoothing, *Journal of Statistical Planning and Inference*, **57**, 29–38.

[16] Hájek, J. (1960). Limiting Distributions in Simple Random Sampling from a Finite Population, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **5**, 361–374.

[17] Hall, P., DiCiccio, T.J. and J.P. Romano (1989). On Smoothing and the Bootstrap, *The Annals of Statistics*, **17**, 692–704.

[18] Huber, P. (1981). *Robust Statistics*, John Wiley & Sons.

[19] Motoyama, H. and H. Takahashi (2003). On Normal Approximation for Statistical Functional in Finite Population, Presented at Workshop in Hitotsubashi University, Tokyo, Japan.

[20] Polansky, A.M. and W.R. Schucany (1997). Kernel Smoothing to Improve Bootstrap Confidence Intervals, *Journal of the Royal Statistical Society, Series B(Statistical Methodology)*, **59**, 821–838.

[21] Reeds, J.A. III (1976). *On the Definition of von Mises Functionals*, Ph.D thesis, Harvard University, Cambridge, Massachusetts.

[22] Rosén, B. (1964). Limit Theorems for Sampling from Finite Populations, *Arkiv för Matematik*, **5**, 383–422.

[23] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, Inc.

[24] Shao, J. (1994). L-Statistics in Complex Survey Problem, *The Annals of Statistics*, **22**, 946–967.

[25] Silverman, B.W. and G.A. Young (1987). The bootstrap: To smooth or not to smooth?, *Biometrika*, **74**, 469–79.

[26] Takahashi, H. (1988). A Note on Edgeworth Expansions for the von Mises Functionals, *Journal of Multivariate Analysis*, **24**, 54–65.

[27] van der Vaart, A (1994). Weak Convergence of Smoothed Empirical Processes, *Scandinavian Journal of Statistics*, **21**, 501–504.

[28] von Mises, R. (1947). On the Asymptotic Distribution of Differentiable Statistical Functions, *The Annals of Mathematical Statistics*, **18**, 309-348.

[29] Young, G.A. (1990). Alternative Smoothed Bootstraps, *Journal of the Royal Statistical Society. Ser. B*, **52**, 477–484.

[30] Yukich, J.E. (1992). Weak Convergence of Smoothed Empirical Processes, *Scandinavian Journal of Statistics*, **19**, 271–279.