Smoothed Versions of Statistical Functionals from a Finite Population

Hitoshi Motoyama* and Hajime Takahashi †

This version: October, 2005

Abstract

We consider smoothed version of the empirical distribution functions from the finite population and the asymptotic behavior of the statistical functionals defined on the class of smoothed empirical distribution functions. Main parts of our results correspond to those of Fernholz (1991, 1993) in I.I.D. case.

---

*College of Economics, Aoyama Gakuin University, 4-4-25 Shibuya, Shibuya-ku, Tokyo 150-8366, Japan.
†Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi-shi, Tokyo 186-8601, Japan.
1 Introduction

In this article, we derive asymptotic normality of smoothed statistical functionals for a simple random sample from a finite population. For a statistical functional, naive empirical distribution function $F_n$ is commonly used (See Fernholz (1983), Reeds (1976), Takahashi (1988)). However, when a finite population distribution function tends to a smooth function, it may be more appealing to replace $F_n$ by a smoothed version $\tilde{F}_n$ of the empirical distribution function. This type of statistics is used in the context of smoothed bootstrap (Silverman and Young (1987), Young (1990)) and smoothed quantiles (Falk (1985)). Fernholz (1993) derive asymptotic normality of smoothed statistical functionals in IID settings. We consider a finite population case.

For finite population cases, Campbell (1980) proposes the use of statistical functional and gives the sketch of the proof for the asymptotic normality in various sampling cases. The authors investigate the rate of convergence to a normal distribution of statistical functionals in simple random sampling (Motoyama and Takahashi (2003)). For $L$-statistics in survey problems, we refer readers to Shao (1994).

First, we give some notations used in this article. Let $x_1, \ldots, x_N$ be a finite population of size $N$. A simple random sample $X_1, X_2, \ldots, X_n$ is taken, without replacement from the finite population. More precisely, let $(\pi_1, \ldots, \pi_N)$ take all possible permutations of $(1, \ldots, N)$ with common probability $(N!)^{-1}$, and $X_i = x_{\pi_i}, 1 \leq i \leq n$.

Define a distribution function (d.f.)

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

and an empirical distribution function

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

A kernel d.f. estimator $\tilde{F}_n$ is obtained by taking the convolution of $F_n$ with some density $k_n$, $\tilde{F}_n = F_n * k_n$. In our case,

$$\tilde{F}_n = F_n * k_n(x) = \int F_n(x-t)k_n(t)dt$$
$$= \int F_n(x-t)dK_n(t) = \int K_n(x-t)dF_n(t)$$
$$= \frac{1}{n} \sum_{i=1}^{n} K_n(x-X_i),$$

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

2
2 Regular Kernel Sequence

Let \( k \) be a symmetric kernel function (not necessary nonnegative) satisfying \( \int k(x)dx = 1 \) and let \( \{a_n\} \) be a sequence of positive real numbers. The sequence of densities \( \{k_n\} \) defined by

\[
k_n(x) = \frac{1}{a_n} k\left(\frac{x}{a_n}\right), \quad n \geq 1,
\]

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

**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 \( b_n = o\left(n^{-1/2}\right) \) and

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

3 Statistical Functionals

Let \( T_n = T_n(X_1, \ldots, X_n) \) be a statistics.

**Definition 2 (Fernholz (1983))**. 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 definition of \( T \) is assumed to contain the empirical d.f.’s for all \( n \geq 1 \), as well as d.f. \( F_N \). Unless otherwise specified, the range of \( T \) will be the set of real numbers.

**Example 1.** 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

\[
T(G) = G^{-1}(q), \quad 0 < q < 1, \quad G^{-1}(q) = \inf\{x : q < G(x)\}
\]

which allows

\[
X_{\lceil nq \rceil + 1} = T(F_n).
\]
Example 3. Linear combinations of order statistics, or L-estimator, statistics of the form

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

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

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

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

Then $F_N(X_1), \ldots, F_N(X_n)$ be a simple random sample without replacement from a finite population $1/N, 2/N, \ldots, (N - 1)/N, 1$. Let $U_n$ be the empirical distribution function corresponding to $F_N(X_1), \ldots, F_N(X_n)$. Then

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

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

and a statistical functional $T$ induces a functional $\tau$ on the d.f.s $U_n$ by

$$\tau(U_n) = T(F_n) = T(U_n \circ F_N).$$

We define for any d.f. $G$ on $[0, 1]$

$$\tau(G) = T(G \circ F_N)$$

when the right hand side is defined. So, for fixed $F_N$, the functional $T$ induces a functional $\tau$ on the space of d.f.s on $[0, 1]$. Therefore, 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 $\tau$ is the functional induced on $D[0, 1]$, the asymptotic properties of $T(\tilde{F}_n)$ can be determined by the differentiability of $\tau$.

Definition 3 (Fernholz (1993)). Let $\tau$ be a functional defined on an open subset $A$ of a normed vector space $V$ and let $G \in A$.

1. The functional $\tau$ is Gateaux differentiable at $G$ if there exists a continuous linear functional $\tau'_G$ defined on $V$ such that

$$\lim_{t \to 0} \frac{\tau(G + tH) - \tau(G) - \tau'_G(tH)}{t} = 0$$

when the right hand side is defined. So, for fixed $F_N$, the functional $T$ induces a functional $\tau$ on the space of d.f.s on $[0, 1]$. Therefore, 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 $\tau$ is the functional induced on $D[0, 1]$, the asymptotic properties of $T(\tilde{F}_n)$ can be determined by the differentiability of $\tau$.
for each $H \in \mathbf{V}$. In this case $\tau'_G$ will be called Gateaux derivative of $\tau$ at $G$.

2. The functional $\tau$ is Hadamard differentiable at $G$ if for any compact subset $K \subset \mathbf{V}$, (2) holds uniformly for $H \in K$. The linear functional will be called the Hadamard derivative of $\tau$ at $G$.

3. The functional $\tau$ is Fréchet differentiable at $G$ if for any bounded subset $K \subset \mathbf{V}$, (2) holds uniformly for $H \in K$. The linear functional will be called the Fréchet derivative of $\tau$ at $G$.

Clearly, 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 $\text{IF}_{T,F_N}$ defined by

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

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

$$\Delta_x(s) = \Delta(s - x).$$

If $T$ and $\tau$ are defined as above, then the Gateaux derivative $\tau$ at $U \equiv F_N \circ F_N^{-1}$, $\tau'_U$ and the influence function of $T$ at $F_N$ are related by

$$\text{IF}_{T,F_N}(x) = \tau'_U((\Delta_x - F_N) \circ F_N^{-1}),$$

since

$$\text{IF}_{T,F_N}(x) = \lim_{t \to 0} \frac{T(F_N + t(\Delta_x - F_N)) - T(F_N)}{t} = \lim_{t \to 0} \frac{\tau(U + t(\Delta_x - F_N) \circ F_N^{-1}) - \tau(U)}{t} = \tau'_U((\Delta_x - F_N) \circ F_N^{-1})$$

under the modification used in Fernholz (1983, pp.54-64) under which $F_N^{-1}(F_N(x)) = x$ holds. Note under the modification we have used, we also have $F_N(F_N^{-1}(u)) = u$.

In what follows, we assume the following assumption holds:

(A1)There exists $N_0$, such that for some sequences $K_N > 0$ satisfying the condition

$$0 < \inf_N K_N \leq \sup_N K_N < \infty$$

and for all $x, y \in \mathbf{R}$ such that $|x - y| \sim \alpha N^\gamma (\alpha > 0, \gamma > -1)$

$$|F_N(x) - F_N(y)| \leq K_N|x - y| \quad \forall N \geq N_0. \quad (3)$$

This assumption reflects the situation that the population distribution function tends to a smooth one.

Under the (A1), we shall prove the following proposition which corresponds to Theorem 2.3 of Fernholz (1991) for 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 $\tilde{F}_n$ be the smoothed empirical distribution function defined by (2) with a regular kernel sequence $\{k_n\}$. Then, under the assumption (A1), we have

$$\sqrt{n} \sup_{-\infty < x < \infty} |\tilde{F}_n(x) - F_n(x)| \rightarrow 0 \quad \text{a.s.}$$

The proof of this proposition will be done as Fernholz (1991). To prove this proposition, we will show the following lemma.

Lemma 1. 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 $\mathbb{R}$ such that the number of intervals in each $C_n$ is $O(n^\lambda)$ for some constant $\lambda$. 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 with respect to $F_N$. If $T_n$ be the maximum number of $X$'s with values in any $I \in C_n$, then

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

as $n \rightarrow \infty$.

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

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

where $\pi^* = P_F(I)$.

For any $\epsilon > 0$, we have

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

Let $k = \lfloor \epsilon \sqrt{n} \rfloor$, then we have $P(Y_I > \epsilon \sqrt{n}) = P(Y_I \geq k)$ and $k \geq n\pi^*$ for sufficiently large $n$ from $\pi^* = o(n^{-1/2})$.

Using well-known identity of hypergeometric distribution

$$\frac{p(y+1)}{p(y)} = \frac{(N\pi^* - y)(n-y)}{(y+1)(N - N\pi^* - n + y + 1)},$$

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

$$P(Y_I > k) \leq P(Y_I = k) \frac{(k+1)((N-n)(1-\pi^*) + (n-k) + 1)}{(k+1)((N-n)(1-\pi^*) + (n-k) + 1) - (N\pi^* - k)(n-k)} = P(Y_I = k)O(1),$$
It follows from Stirling’s formula that
\[
\Pr(Y_I = k) \sim \left( \frac{(N\pi^*)(N-N\pi^*)n(N-n)}{2\pi k(N\pi^*-k)N(n-k)(N-n-(N\pi^*-k))} \right)^{1/2} \frac{N}{N-n-(N\pi^*-k)}^N \frac{N\pi^*}{k(N(1-\pi^*)-(n-k))}^k \frac{n(N-n-(N\pi^*-k))}{(n-k)(N-n)}^n
\]
Here we have
\[
\left( \frac{(N\pi^*)(N-N\pi^*)n(N-n)}{2\pi k(N\pi^*-k)N(n-k)(N-n-(N\pi^*-k))} \right)^{1/2} = O(n^{-1/4})
\]
\[
\frac{N}{N-n-(N\pi^*-k)}^N = O(1)
\]
\[
\frac{N\pi^*}{k(N(1-\pi^*)-(n-k))}^k = o(n^{-1/2}).
\]
So for sufficiently large \(n\),
\[
\Pr(Y_I > k) \leq \left( \frac{n\pi^*}{\epsilon\sqrt{n}} \right)^{\epsilon\sqrt{n}} \left( \frac{n}{n - \epsilon\sqrt{n}} \right)^n.
\]
Recalling the fact
\[
\left( \frac{n}{n - \epsilon\sqrt{n}} \right)^n = e^{\epsilon\sqrt{n} + O(1)},
\]
we have
\[
\Pr(Y_I > k) \leq C \left( \frac{n\pi^*}{\epsilon\sqrt{n}} \right)^{\epsilon\sqrt{n}} e^{\epsilon\sqrt{n}} \leq C' e^{-\epsilon\sqrt{n}}
\]
for some constants \(C\) and \(C'\) and sufficiently large \(n\) since \(\pi^* = o(n^{-1/2})\) and \(k = [\epsilon\sqrt{n}] \leq n\pi^*\) for large \(n\).
Therefore
\[
\Pr(T_n > \epsilon\sqrt{n}) \leq C' \sum_{I \in C_n} e^{-\epsilon\sqrt{n}} = O(n^\lambda) e^{-\epsilon\sqrt{n}}
\]
since the number of intervals in \(C_n\) is \(O(n^\lambda)\). Hence
\[
\sum_{n=1}^{\infty} \Pr(T_n > \epsilon\sqrt{n}) < \infty,
\]
and since $\epsilon > 0$ was arbitrary, the first Borel-Cantelli Lemma implies that

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

Q.E.D.

Now we shall prove the proposition.

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

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

Immediately, we have $Q(0) = 0$,

$$\lim_{t \to \infty} Q(t) = 1,$$

and

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

$$= \sup_x (F_N(x + s + t) - F_N(x + s) + F_N(x + s) - F_N(x))$$

$$\leq \sup_x (F_N(x + s + t) - F_N(x + s)) + \sup_x (F_N(x + s) - F_N(x))$$

$$= Q(s) + Q(t). \tag{5}$$

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.

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 $\mathbb{R}$.

From (A1), for sufficiently large $n$ (therefore large $N$),

$$|Q(t) - Q(s)| \leq |\sup_x (F_N(x + t) - F_N(x)) - \sup_x (F_N(x + s) - F_N(x))| \leq K_N|t - s|.$$ 

For sufficiently large $n$, we have

$$P_{F_N}(I_j) = F_N(x_j) - F_N(x_{j-1}) \leq 2Q_n \leq Kb_n = o(n^{-1/2}).$$

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 $x \in \mathbb{R}$, $x \in I_j$ 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

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

8
where $T_n$ is defined as in Lemma 1. Therefore,

$$\sqrt{n}|\tilde{F}_n(x) - F_n(x)| \leq \sqrt{n} \int |F_n(x - t) - F_n(x)||k_n(t)|dt$$

$$\leq \sqrt{n} \int_{|t| \leq b_n} |F_n(x - t) - F_n(x)||k_n(t)|dt$$

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

$$\leq \frac{2T_n}{\sqrt{n}} \int_{|t| \leq b_n} |k_n(t)|dt + \sqrt{n} \int_{|t| > b_n} |k_n(t)|dt$$

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.

Q.E.D.

Applying the proposition we now proved, we can gain the following corollary:

**Corollary 1.** Under the assumption of the proposition 1,

(a) The smoothed and non-smoothed Kolmogorov-Smirnov statistics $\sqrt{n} \sup_x |\tilde{F}_n(x) - F_N(x)|$ 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 processes

$$\sqrt{n}(\tilde{F}_n(x) - F_N(x))/\sigma_{N,n} \overset{D}{\to} N(0,1), \quad n, N - n \to \infty,$$

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

**Proof.** By triangular inequality,

$$\sqrt{n} \sup_x |F_n(x) - F_N(x)| - \sqrt{n} \sup_x |\tilde{F}_n(x) - F_n(x)|$$

$$\leq \sqrt{n} \sup_x |\tilde{F}_n(x) - F_N(x)|$$

$$\leq \sqrt{n} \sup_x |F_n(x) - F_N(x)| + \sqrt{n} \sup_x |\tilde{F}_n(x) - F_n(x)|.$$
Lemma 2. Let $T$ be a statistical functional, $\tau$ be the functional indexed by $T$ on $D[0,1]$ by the relation (1), and $F_N$ satisfies (A1). Suppose $\tau$ is Gateaux differentiable at $U \equiv F_N \circ F_N^{-1}$ with derivative $\tau'_U$ and that the influence function $IF = IF_{T,F_N}$ is Stieltjes integrable with respect to bounded variation functions. If $X_1, \ldots, X_n$ is a simple random sample without replacement from $F_N$ and $\{k_n\}$ is a kernel sequence, then for $\tilde{U}_n = \tilde{F}_n \circ F_N^{-1}$ and $\tilde{IF} = IF_k$, 

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

Proof. Let $t_i$ and $\bar{t}_i$, $i = 1, \ldots, m$ be numbers such that $-\infty < t_1 < \cdots < t_m < \infty$, $t_i \leq \bar{t}_i \leq t_{i+1}$, $i = 1, \ldots, m - 1$. Then $K_{n,x}$ is approximated on $R$ by sums of the form

$$S = \sum_{i=1}^{m} (K_n(t_{i+1}) - K_n(t_i)) \Delta x_{t_{i+1} - t_i}$$

as $t_1 \to -\infty$, $t_m \to \infty$, and $\max_i(t_{i+1} - t_i)$. Therefore, $K_{n,x} \circ F_N^{-1}$ may be approximated as an element of $D[0,1]$ by functions of the form $S \circ F_N^{-1}$ for sufficiently large $n$. From the linearity of $\tau'_U$,

$$\tau'_U(S \circ F_N^{-1}) = \sum_{i=1}^{m} (K_n(t_{i+1}) - K_n(t_i)) \tau'_U(\Delta x_{t_{i+1} - t_i} \circ F_N^{-1})$$

$$= \sum_{i=1}^{m} (K_n(t_{i+1}) - K_n(t_i))(IF(x - \bar{t}_i) + \tau'_U(U)). \quad (6)$$

Since $IF$ is Stieltjes integrable with respect to $K_n$, for a given $x$ the sum (6) tends to $\tilde{IF}(x) + \tau'_U(U)$. Hence from the continuity of $\tau'_U$,

$$\tau'(K_{n,x} \circ F_N^{-1}) = \tilde{IF}(x) + \tau'_U(U).$$

Since

$$\tilde{U}_n = \tilde{F}_n \circ F_N^{-1} = \frac{1}{n} \sum_{i=1}^{n} K_{n,x_i} \circ F_N^{-1},$$

we have

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

Q.E.D.

Using Hadamard derivative, Fernholz (1993) obtained the asymptotic normality of the smoothed statistic $T(\tilde{F}_n)$ for I.I.D. case. We show the analogous result for finite population case under Fréchet differentiability.
Theorem 1. Let \( X_1, \ldots, X_n \) be a simple random sample from a distribution function satisfies with (A1). Let \( \{k_n\} \) be a bounded regular kernel sequence. Let \( T \) be a statistical functional and \( \tau \) be the statistical functional induced by \( T \) on \( D[0,1] \) according to (1). If \( \tau \) is Fréchet differentiable at \( U \) with the sup norm and the influence function \( IF = IF_{T,F_N} \) is Stieltjes integrable with respect to bounded variation functions with \( \mu = E[\tilde{IF}] \) and \( 0 < \sigma^2 = \text{Var}[\tilde{IF}] < \infty \), Further suppose that \( \tilde{IF} \)'s satisfy the Lindeberg type condition such that:

\[
\lim_{n,N-n\to\infty} \frac{\sum_{i=1}^{N} P_x (IF(x_i) - \tilde{IF})^2}{\sum_{i=1}^{N} (IF(x_i) - \tilde{IF})^2} = 0 \quad \text{for any } \tau > 0,
\]

where \( P_x \) be the subset of elements of \( P = \{1, \ldots, N\} \) on which the inequality \( |IF(x_i) - \tilde{IF}| > \tau \sqrt{n} \sigma \), and \( \tilde{IF} \) is the arithmetic average of \( IF \)'s.

Then

\[
\sqrt{n}(T(\tilde{F}_n) - T(F_N) - \mu_N)/\sigma_N \xrightarrow{D} N(0,1),
\]

as \( n, N - n \to \infty \).

Proof. For sufficienctly large \( n \), we have

\[
\sqrt{n}(T(\tilde{F}_n) - T(F_N) - \mu_N)/\sigma_N = \sqrt{n}(\tau(\tilde{U}_n) - \tau(U) - \mu_N)/\sigma_N
\]

\[
= \sqrt{n}\{\tau(U_n - U) - \mu_N\}/\sigma_N + \sqrt{n}\text{Rem}(\tilde{U}_n - U)/\sigma_N
\]

\[
= \frac{1}{\sigma_N} \sqrt{n} \sum_{i=1}^{n} (IF(X_i) - \mu_N) + \sqrt{n}\text{Rem}(\tilde{U}_n - U)/\sigma_N
\]

by lemma 2. From the central limit theorem for finite population (Erdös and Rényi (1959) and Hajek (1960)), the first term of the last equation tends to \( N(0,1) \) as \( n, N - n \to \infty \).

As for the remainder term, \( \sqrt{n}\text{Rem}(\tilde{U}_n - U) \) goes to 0 in probability since from the definition of Fréchet differentiability \( \text{Rem}(\tilde{F}_n - F_N) = o_P(\text{sup}_x |\tilde{F}_N(x) - F_N(x)|) \) and from Corollary 1 \( \text{sup}_x |\tilde{F}_N(x) - F_N(x)| = O_p(n^{-1/2}) \) so \( \text{Rem}(\tilde{F}_n - F_N) = o_P(n^{-1/2}) \). Q.E.D.

We used Fréchet differentiability of \( \tau \) instead of Hadamard differentiability that Fernholz (1994) had used. This may be somewhat stronger assumption. We present asymptotic normality under the Hadamard differentiability in the next version of our research.

4 Appendix

We shall show the Glivenko-Cantelli type theorem for the non-smoothed empirical distribution function. Then we state the consistency of the smoothed empirical distribution
function. The results we show in the Appendix is not directly used in deriving our results. However, we think these are interesting results theirselves.

For every fixed \( x \), \( I_{(-\infty,x)}(x_i) \) consist of \( N \times F_N \) and \( N(1 - F_N) \) 0. By the finite population strong law of large numbers for \( U \)-statistics (Nandi and Sen (1963), Sen (1970)), there is a set \( A_x \) of asymptotically probability 0 such that

\[
\lim_{n,N \to \infty} |F_n(x, \omega) - F_N(x)| = 0, \quad \text{for } x \text{ fixed and } \omega \in A_x^c.
\]

Now we prove the following Glivenko-Cantelli type theorem.

**Proposition 2.**

\[
D_n(\omega) \equiv \sup_{-\infty < x < \infty} |F_n(x, \omega) - F_N(x)| \to 0 \quad \text{a.s. as } n, N \to \infty.
\]

**Note:** By the right continuity of \( F_n \), the supremum above is unchanged if \( x \) is restricted to the rationals. So \( D_n \) are random variables.

**Proof.** The proof will be done as that of infinite population version of Glivenko-Cantelli theorem, see for example Billingsley (1995). As already seen, \( \lim_{n,N} |F_n(x, \omega) - F_N(x)| = 0 \) except on a set \( A_x \) of probability 0. Similary, applying the finite population strong law of large numbers with

\[
F_n(x-, \omega) = \frac{1}{n} \sum_{j=1}^n I_{(-\infty,x)}(X_i),
\]

we have \( \lim_{n,N} |F_n(x-, \omega) - F_N(x-)| = 0 \) except on a set \( B_x \) of asymptotically probability 0.

Let

\[
\varphi(u) = \inf[x : u \leq F_N(x)], \quad x_{m,k} = \varphi\left(\frac{k}{m}\right), \quad m \geq 1, \quad 1 \leq k < m,
\]

Then we have

\[
F_N(\varphi(u) -) \leq u \leq F_N(\varphi(u)).
\]

Hence

\[
F_N(x_{m,k-1}) - F_N(x_{m,k-1}) \leq \frac{1}{m},
\]

\[
F_N(x_{m,1}) \leq \frac{1}{m},
\]

\[
F_N(x_{m,m}) \geq 1 - \frac{1}{m}.
\]

Here we let

\[
D_{m,n}(\omega) = \max_k \{|F_n(x_{m,k}, \omega) - F_N(y)|, |F_n(x_{m,k}, \omega) - F_N(x-)|\}.
\]
So, when $x_{m,k-1} \leq x < x_{m,k}$,

\[ F_n(x, \omega) \leq F_n(x_{m,k-1}, \omega) \leq F_N(x_{m,k-1}) + D_{m,n}(\omega) \leq F_N(x) + m^{-1} + D_{m,n}(\omega), \]

\[ F_n(x, \omega) \geq F_n(x_{m,k}, \omega) \geq F_N(x_{m,k}) - D_{m,n}(\omega) \geq F_N(x) - m^{-1} - D_{m,n}(\omega). \]

In deriving the inequalities, we use

\[ \frac{k-1}{m} \leq F_N(x_{m,k-1}) \leq F_N(x) \leq F_N(x_{m,k}) \leq \frac{k}{m}. \]

Similarly, for $x < x_{m,1}$,

\[ F_n(x, \omega) \leq F_n(x_{m,1}, \omega) \leq F_N(x_{m,1}) + D_{m,n}(\omega) \leq F_N(x) + m^{-1} + D_{m,n}(\omega), \]

\[ F_n(x, \omega) \geq F_n(x_{m,0}, \omega) \geq F_N(x_{m,0}) - D_{m,n}(\omega) \geq F_N(x) - m^{-1} - D_{m,n}(\omega), \]

and for $x \geq x_{m,m}$,

\[ F_n(x, \omega) \leq F_n(\infty, \omega) \leq F_N(\infty) + D_{m,n}(\omega) \leq F_N(x) + m^{-1} + D_{m,n}(\omega), \]

\[ F_n(x, \omega) \geq F_n(x_{m,m-1}, \omega) \geq F_N(x_{m,m-1}) - D_{m,n}(\omega) \geq F_N(x) - m^{-1} - D_{m,n}(\omega). \]

Hence

\[ D_n(\omega) \leq D_{m,n}(\omega) + m^{-1}. \]

If $\omega$ lies outside the union $A$ of all the $A_{mk}$ and $B_{mk}$, then $\lim_{n,N} D_{m,n}(\omega) = 0$ so $\lim_{n,N} D_n(\omega) = 0$. However, $A$ has probability 0 for large $n$ and $N$. The proposition follows.

Q.E.D.

Therefore, combining the proposition 1 and 2, we have the following Glivenko-Cantelli theorem of the smoothed empirical distribution function from a finite population.

**Proposition 3.** Under the conditions of Lemma 1, we have for all $x \in \mathbb{R}$,

\[ \sup_{-\infty < x < \infty} |\tilde{F}_n(x, \omega) - F_N(x)| \to 0 \text{ a.s. as } n, N \to \infty. \]

**References**


