Nonparametric Estimation

Let $x_1, x_2, \cdots, x_n, \cdots$ be iid random variables with common df $F(x), x \in \mathbb{R}^p$ $(p \geq 1)$.

Here F is unknown, we write

$\mathcal{F} = \{F\}$: Class of all possible distributions having some property(property)

The class $\mathcal{F}$ is known.

$\theta(F)$ (real or vector valued): Functional defined on $\mathcal{F}$.

[$\mathcal{F}$: Abstract space.
Function defined on an abstract space is called functional.
In parametric Theory $\theta(F)$ is simply parameter]

Our problem is to estimate $\theta(F)$ on the basis of $x_1, x_2, \cdots, x_n, \cdots$

Estimability of $\theta(F)$:

WLG assume that $\theta(F)$ is real-valued. That is $\{\theta(F), F \in \mathcal{F}\} \subseteq \mathbb{R}^1$.
$\theta(F)$ is said to be estimable (or regular) if there exists a function $\phi(x_1, x_2, \cdots, x_m)$ such that $\theta(F) = E_F\, \phi(x_1, \cdots, x_m) \;\forall\, F \in \mathcal{F}$ --- (1)
The function $\phi(\cdot)$ is called a kernel of $\theta(F)$. The minimum sample size for which (1) holds is called the degree of the kernel.

U-Statistics

$x_1, x_2, \cdots, x_n$ are iid $F(x)$; $x \in \mathbb{R}^p$.

$\theta(F)$: Estimable functional of degree m. That is, $\exists$ a function $\phi(x_1, \cdots, x_m) \ni$

$E_F\, \phi(x_1, \cdots, x_m) = \theta(F) \quad \forall\, F \in \mathcal{F}$.

Comparing to $\phi(\cdot)$, define the following statistic,

$$U = U(x_1, \cdots, x_n) = \frac{1}{n(n-1)\cdots(n-m+1)} \sum_{P} \phi(x_{i_1}, \cdots, x_{i_m}) \quad \text{--- (2)}$$

where

$P = \{(i_1, \cdots, i_m) : 1 \leq i_1 \neq i_2 \neq \cdots \neq i_m \leq n\}$

$\equiv$ Set of all possible ${}^nP_m$ permutations.

The statistic defined by (2) is called U-statistic corresponding to the kernel $\phi(\cdot)$.

The kernel $\phi(\cdot)$ is said to be symmetric if

$$\phi(x_{i_1}, x_{i_2}, \cdots, x_{i_m}) = \phi(x_1, \cdots, x_m) \quad \forall (i_1, \cdots, i_m) \text{ of } (1, 2, \cdots, m).$$

Then, we may re-write (2) as

$$U = \frac{m!}{n(n-1)\cdots(n-m+1)} \sum_C \phi(x_{j_1}, \cdots, x_{j_m})$$

$$= \frac{1}{\binom{n}{m}} \sum_C \phi(x_{j_1}, \cdots, x_{j_m})$$

where $C = \{(j_1, \cdots, j_m) : 1 \le j_1 < j_2 < \cdots < j_m \le n\}$
$\equiv$ Set of all $\binom{n}{m}$ combinations.

Suppose $\phi(\cdot)$ is asymmetric. Then we define

$$\phi^0(x_1, \cdots, x_m) = \frac{1}{m!} \sum_{(i_1, \cdots, i_m) \text{ of } (1,2,\cdots,m)} \phi(x_{i_1}, \cdots, x_{i_m})$$

[Example : $m = 3$.

$$\phi^0(x_1, x_2, x_3) = \frac{1}{6}\{\phi(x_1, x_2, x_3) + \phi(x_1, x_3, x_2) + \phi(x_3, x_1, x_2) + \phi(x_2, x_1, x_3) + \phi(x_2, x_3, x_1) + \phi(x_3, x_2, x_1)\}$$

$$= \frac{1}{3!} \sum_{(i_1, i_2, i_3) \text{ of } (1,2,3)} \phi(x_{i_1}, x_{i_2}, x_{i_3}) \quad]$$

Then (2) may be written as

$$U = \frac{1}{\binom{n}{m}} \sum_{1 \le \alpha_1 < \cdots < \alpha_m \le n} \phi^0(x_{\alpha_1}, \cdots, x_{\alpha_m})$$

Here $\phi^0(\cdot)$ is a symmetric kernel.

[For $m = 2$

$$U = \frac{1}{n(n-1)} \sum\sum_{i_1 \ne i_2} \phi(x_{i_1}, x_{i_2})$$

$$= \frac{2}{n(n-1)} \left[\frac{\phi(x_1, x_2) + \phi(x_2, x_1)}{2} + \frac{\phi(x_1, x_3) + \phi(x_3, x_1)}{2} + \cdots + \frac{\phi(x_{n-1}, x_n) + \phi(x_n, x_{n-1})}{2} \right]$$

$$= \frac{1}{\binom{n}{2}} \sum\sum_{\alpha_1 < \alpha_2} \phi^0(x_{\alpha_1}, x_{\alpha_2}) \quad]$$

Hence, WLG, we take

$$U = \frac{1}{\binom{n}{m}} \sum_C \phi(x_{i_1}, \cdots, x_{i_m}),$$

where $\phi(\cdot)$ is a symmetric kernel of degree m.

Examples:

1. $\mathcal{F} = \{F : E_F|X| < \infty\}$.

Take $\theta(F) = E_F X = \mu(F)$

Obviously, $\mu(F) = E_F(X_1) \ \forall F \in \mathcal{F}$.

$\Rightarrow \mu(F)$ is an estimable functional with kernel $\phi(x_1) = x_1$ of degree $m=1$.

Corresponding U-statistic $U = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$.

2. $\mathcal{F} = \{F : E_F X^2 < \infty\}$

Take $\theta(F) = E_F(X^2) - \mu^2(F)$

$= V_F(X) = \sigma^2(F)$, say.

Observe that $E_F\{X_1 X_2\} = \mu^2(F)$.

$\Rightarrow \mu^2(F)$ is an estimable functional with kernel $\phi(x_1, x_2) = x_1 x_2$ of degree 2.

Unbiased estimator of $E_F(X^2)$ is X_1^2 or X_2^2.

Hence, an unbiased estimator of $\theta(F)$ is

$$\phi^{(1)}(x_1, x_2) = x_1^2 - x_1 x_2 \text{ or } \phi^{(2)}(x_1, x_2) = x_2^2 - x_1 x_2.$$

Define

$$\phi(x_1, x_2) = \frac{1}{2}\left[\phi^{(1)}(x_1, x_2) + \phi^{(2)}(x_1, x_2)\right] = \frac{1}{2}(x_1 - x_2)^2$$

Hence $E_F\,\phi(X_1, X_2) = \sigma^2(F) \ \forall F \in \mathcal{F}$.

$\Rightarrow \sigma^2(F)$ is an estimable functional with kernel $\phi(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$ of degree 2.

$\phi(\cdot)$ is also a symmetric kernel.

Corresponding U-statistic is

$$U = \frac{1}{\binom{n}{2}} \sum\sum_{1 \le i_1 < i_2 \le n} \left[\frac{1}{2}(x_{i_1} - x_{i_2})^2\right]$$

$$= \frac{1}{2\binom{n}{2}} \sum\sum_{1 \le i_1 < i_2 \le n} (x_{i_1} - \bar{x} - x_{i_2} + \bar{x})^2$$

$$= \frac{1}{4.\binom{n}{2}} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \{(x_{i_1} - \bar{x}) - (x_{i_2} - \bar{x})\}^2$$

$$= \frac{1}{4.\binom{n}{2}} \left[2n \sum_{i=1}^{n} (x_i - \bar{x})^2\right]$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = s^2.$$

It can be easily observed that

$$E_F(U) = \frac{1}{\binom{n}{2}} \sum\sum_{1\le i_1<i_2\le n} E_F\left\{\frac{1}{2}(x_{i_1}-x_{i_2})^2\right\}$$

$$= \frac{1}{\binom{n}{2}}\cdot\binom{n}{2}\sigma^2(F)$$

$$= \sigma^2(F) \quad \forall F \in \mathcal{F}.$$

$\Rightarrow$ U is an unbiased estimator of $\sigma^2(F)$.

3. Kendall's τ.

Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ be iid $F(\underline{x})$; $\underline{x} \in \mathbb{R}^2$

Write $\underset{\sim}{x_i} = \begin{pmatrix} x_{1i} \\ x_{2i} \end{pmatrix}$; $i=1(1)n$.

Assume F is continuous.

$\mathcal{F}$ = Class of all bivariate continuous distributions.

For any two random variables $\underline{x}_1, \underline{x}_2$, we say that there is a concordance if

$$[x_{11}-x_{12}>0,\ x_{21}-x_{22}>0] \cup [x_{11}-x_{12}<0,\ x_{21}-x_{22}<0]$$

$$\Leftrightarrow [(x_{11}-x_{12})(x_{21}-x_{22})>0]$$

This is a discordance if $[(x_{11}-x_{12})(x_{21}-x_{22})<0]$.

Write π_c = Prob.[Concordance]

$$= P_F\{(x_{11}-x_{12})(x_{21}-x_{22})>0\}$$

$$\pi_d = P_F\{(x_{11}-x_{12})(x_{21}-x_{22})<0\}$$

Obvious that

$$\pi_c + \pi_d = 1 \quad \forall F \in \mathcal{F}$$

Consider the functional

$$\theta(F) = \pi_c - \pi_d.$$

Obvious that

i) $|\theta(F)| \in [0,1]$

ii) $\theta(F) = +1$ for perfect association

$= -1$ " " disociation

$= 0$ for independence.

Here $\theta(F)$ can be used as a suitable measure of association between two variables. It is called Kendall's tau and is denoted by $\tau(=\tau(F))$.
Also note that

$$\tau(F) = 2\pi_c - 1$$
$$= 1 - 2\pi_d.$$

Now define,

$$\delta(z) = \left.\begin{array}{ll} +1 & \text{if } z > 0 \\ -1 & \text{if } z < 0 \\ 0 & \text{if } z = 0 \end{array}\right\}$$

If Z is continuous, we have

$$P\{\delta(Z) = 0\} = 0.$$

Write

$$\phi(\underline{x}_1, \underline{x}_2) = \delta(x_{11} - x_{12}) \cdot \delta(x_{21} - x_{22})$$

Note that $\phi(\underline{x}_1, \underline{x}_2) = \phi(\underline{x}_2, \underline{x}_1)$

Here

$$E_F\,\phi(\underline{x}_1, \underline{x}_2) = 1 \cdot P_F\{(x_{11} - x_{12})(x_{21} - x_{22}) > 0\}$$
$$- 1 \cdot P_F\{(x_{11} - x_{12})(x_{21} - x_{22}) < 0\}$$
$$= \pi_c - \pi_d$$
$$= \tau(F) \quad \forall F \in \mathcal{F}.$$

$\Rightarrow$ $\tau(F)$ is an estimable functional with kernel $\phi(\cdot)$ of degree 2.

Corresponding U-statistic is

$$U = \frac{1}{\binom{n}{2}} \sum\sum_{1 \le i_1 < i_2 \le n} \phi(x_{i_1}, x_{i_2})$$

Observe that

$$E_F(U) = \tau(F) \qquad \forall F \in \mathcal{F}.$$

Now it can be easily observed that

$$U = \frac{\sum\sum_{1 \le i_1 < i_2 \le n} \delta(x_{1i_1} - x_{1i_2})\,\delta(x_{2i_1} - x_{2i_2})}{\sqrt{\sum_{1 \le i_1 < i_2 \le n} \delta^2(x_{1i_1} - x_{1i_2}) \cdot \sum_{1 \le i_1 < i_2 \le n} \delta^2(x_{2i_1} - x_{2i_2})}}$$

= Product moment correlation coefficient.

= t, say.

Mean and variance of U:

$$U = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \cdots < i_m \le n} \phi(x_{i_1}, \ldots, x_{i_m})$$

Now, $E_F\, \phi(X_1, X_2, \ldots, X_m) = \theta(F)\ \forall F \in \mathcal{F}$

$$\Rightarrow E_F(U) = \binom{n}{m}^{-1}\binom{n}{m}\theta(F) \quad [\because \phi(x_{i_1}, \ldots x_{i_m}) = \phi(x_1, \ldots, x_m)\ \forall (i_1, \ldots, i_m) \text{ of } (1, 2, \ldots, n)]$$
$$= \theta(F) \quad \forall F \in \mathcal{F}$$

$\Rightarrow$ U is an unbiased estimator of $\theta(F)$.

To find the variance of U, we define the following random variables:

$$E_F\, \phi(x_1, \ldots, x_m) = \theta(F)$$

$$\phi_c(x_1, \ldots, x_c) = E_F\, \phi(x_1, \ldots, x_c, X_{c+1}, \ldots, X_m)$$
$$= E_F\{\phi(X_1, \ldots, X_c, X_{c+1}, \ldots, X_m) / X_1 = x_1, \ldots, X_c = x_c\}, \quad 1 \le c \le m.$$

Also write $\phi_0 = \theta(F)$

Here note that $\phi_m(x_1, \ldots, x_m) = \phi(x_1, \ldots, x_m)$.

Further write,

$$\xi_c(F) = V_F[\phi_c(X_1, \ldots, X_c)]$$
$$= E_F[\phi_c^2(X_1, \ldots, X_c)] - \theta^2(F),$$

Since

$$E_F\, \phi_c(X_1, \ldots, X_c) = E_F E_F\{\phi(X_1, \ldots, X_c, X_{c+1}, \ldots, X_m) / X_1, \ldots, X_c\}$$
$$= E_F[\phi(X_1, \ldots, X_m)]$$
$$= \theta(F) \quad \forall\ 0 \le c \le m.$$

Next we want to verify that

$$\xi_c(F) \le \xi_{c+1}(F), \quad \xi_0(F) = 0;\ c = 1, 2, \ldots, m. \quad \text{--- (1)}$$

Here $\xi_{c+1}(F) = E_F[\phi_{c+1}^2(X_1, \ldots, X_{c+1})] - \theta^2(F)$

$$= \underset{\{X_1, \ldots, X_c\}}{E_F}\ \underset{X_{c+1} / \{X_1, X_2, \ldots, X_c\}}{E_F}[\phi_{c+1}^2(X_1, \ldots, X_{c+1})] - \theta^2(F)$$

[Jenson Inequality $\Rightarrow$ for any convex function $\psi(x)$, $E(\psi(x)) \geq \psi(E(x))$.
If $\psi(x) = x^2$, we have
$E(x^2) \geq [E(x)]^2$]

$$\geq E_{F\,\{x_1,\dots,x_c\}} \left[E_{F\, x_{c+1}/(x_1,\dots,x_c)} \left(\phi_{c+1}(x_1,\dots,x_c,x_{c+1}) \right) \right]^2 - \theta^2(F)$$

[By applying Jensen's Inequality on conditional expectation]

$$= E_{F\,\{x_1,\dots,x_c\}} \left[E_{F\, x_{c+1}/(x_1,\dots,x_c)} \left(E_{F\,(x_{c+2},\dots,x_m)/(x_1,\dots,x_{c+1})} \phi(x_1,\dots,x_{c+1},x_{c+2},\dots,x_m) \right] ^2 - \theta^2(F)$$

$$= E_{F\,\{x_1,\dots,x_c\}} \left[E_{F\,(x_{c+1},\dots,x_m)/\{x_1,\dots,x_c\}} \phi(x_1,x_2,\dots,x_c,x_{c+1},x_{c+2},\dots,x_m) \right]^2 - \theta^2(F)$$

$$= E_F \{ \phi_c(x_1,x_2,\dots,x_c) \}^2 - \theta^2(F)$$

$= \zeta_c(F)$ which implies (1) for all c.

That is $\zeta_c(F)$ is increasing function of c.

$$\Rightarrow 0 \leq \zeta_1(F) \leq \zeta_2(F) \leq \dots \leq \zeta_m(F)$$

$\theta(F)$ is said to be stationary of order c if

$$\zeta_1(F) = \dots = \zeta_c(F) = 0, \quad \zeta_{c+1}(F) > 0.$$

Expression of variance

$$U = \frac{1}{\binom{n}{m}} \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} \phi(x_{i_1},\dots,x_{i_m})$$

$$\Rightarrow V_F(U) = \binom{n}{m}^{-2} \sum_{C} \sum_{C'} Cov_F \{ \phi(x_{i_1},\dots,x_{i_m}), \phi(x_{j_1},\dots,x_{j_m}) \},$$

where $C = \{ (i_1,\dots,i_m) : 1 \leq i_1 < i_2 < \dots < i_m \leq n \}$
$C' = \{ (j_1,\dots,j_m) : 1 \leq j_1 < j_2 < \dots < j_m \leq n \}$

Now, as $x_1, x_2, \dots, x_n$ are independent, we have,

$$Cov_F \{ \phi(x_{i_1},\dots,x_{i_m}), \phi(x_{j_1},\dots,x_{j_m}) \} \neq 0 \quad \text{if } \{i_1,\dots,i_m\} \cap \{j_1,\dots,j_m\} \neq \phi$$
$$= 0 \quad \text{if } \quad " \quad = \phi.$$

Let 'c' be the number of letters in common between $\{i_1,\dots,i_m\}$ and $\{j_1,\dots,j_m\}$.

Then, as $x_1,\dots,x_n$ are iid, we have

$$Cov_F \{ \phi(x_{i_1},\dots,x_{i_m}), \phi(x_{j_1},\dots,x_{j_m}) \}$$
$$= Cov_F \{ \phi(x_1,\dots,x_c,x_{c+1},\dots,x_m), \phi(x_1,\dots,x_c,x_{m+1},\dots,x_{2m-c}) \}$$
$$= E_F \{ \phi(x_1,\dots,x_c,x_{c+1},\dots,x_m) . \phi(x_1,\dots,x_c,x_{m+1},\dots,x_{2m-c}) \} - \theta^2(F)$$
$$= E_F\, E_F \{ \phi(x_1,\dots,x_c,x_{c+1},\dots,x_m) . \phi(x_1,\dots,x_c,x_{m+1},\dots,x_{2m-c}) / (x_1,\dots,x_c) \} - \theta^2(F)$$
$$= E_F \left[E_F \{ \phi(x_1,\dots,x_c,x_{c+1},\dots,x_m) / (x_1,\dots,x_c) \} . E_F \{ \phi(x_1,\dots,x_c,x_{m+1},\dots,x_{2m-c}) / (x_1,\dots,x_c) \} \right] - \theta^2(F)$$

(For given $(x_1,\dots,x_c)$, the above two functions are independent)

$$= E_F [\phi_c^2(x_1,\dots,x_c)] - \theta^2(F)$$
$$= V_F [\phi_c(x_1,\dots,x_c)]$$
$$= \zeta_c(F)$$

Hence $V_F(U) = \binom{n}{m}^{-2} \sum_{c=1}^{m} \binom{n}{m}\binom{m}{c}\binom{n-m}{m-c} \cdot \zeta_c(F)$.

$$= \sum_{c=1}^{m} \frac{\binom{m}{c}\binom{n-m}{m-c}}{\binom{n}{m}} \cdot \zeta_c(F)$$

$$= \sum_{c=1}^{m} H(c|m,n) \cdot \zeta_c(F),$$

where $H(x|m,n)$ = pmf of Hypergeometric (m,n).

Expansion of $V_F(U)$:

$$V_F(U) = \sum_{c=1}^{m} \frac{\binom{m}{c}\binom{n-m}{m-c}}{\binom{n}{m}} \zeta_c(F)$$

$$= \frac{\binom{m}{1}\binom{n-m}{m-1}}{\binom{n}{m}} \zeta_1(F) + \frac{\binom{m}{2}\binom{n-m}{m-2}}{\binom{n}{m}} \zeta_2(F) + \frac{\binom{m}{3}\binom{n-m}{m-3}}{\binom{n}{m}} \zeta_3(F) + \cdots + \frac{1}{\binom{n}{m}} \zeta_m(F)$$

$$= \frac{m^2}{n} \zeta_1(F) + \left\{ \frac{\binom{m}{1}\binom{n-m}{m-1}}{\binom{n}{m}} - \frac{m^2}{n} \right\} \zeta_1(F) + \frac{\binom{m}{2}\binom{n-m}{m-2}}{\binom{n}{m}} \zeta_2(F) + \cdots + \frac{1}{\binom{n}{m}} \zeta_m(F)$$

Coefficient of $\zeta_1(F)$ from the 2nd term

$$= \frac{m^2}{n} \left[\frac{(n-m)_{(m-1)}}{(n-1)_{(m-1)}} - 1 \right]$$

$$= \frac{m^2}{n} \cdot \frac{\text{a polynomial in } n \text{ of degree } (m-2)?}{\text{a polynomial in } n \text{ of } (m-1)?}$$

$$= \frac{\lambda_{n_1}}{n}, \text{ where } \lambda_{n_1} = O(1), \text{ as } n \to \infty$$

Coefficient of $\zeta_2(F)$

$$= \frac{\binom{m}{2}\binom{n-m}{m-2}}{\binom{n}{m}} = \frac{\lambda_{n_2}}{n}, \text{ say, where } \lambda_{n_2} = o(1) = O\left(\tfrac{1}{n}\right), \text{ as } n \to \infty,$$

Coefficient of $\zeta_3(F)$ $= \dfrac{\binom{m}{3}\binom{n-m}{m-3}}{\binom{n}{m}} = \dfrac{\lambda_{n_3}}{n}$, say, where $\lambda_{n_3} = O\left(\frac{1}{n^2}\right)$, as $n \to \infty$

- - - - -

Coefficient of $\zeta_m(F) = \dfrac{\lambda_{n_m}}{n}$, where, $\lambda_{n_m} = O\left(\frac{1}{n^{m-1}}\right)$ as $n \to \infty$.

Hence the expression is

$$V_F(U) = \frac{m^2}{n} \zeta_1(F) + \frac{1}{n} \left[\lambda_{n_1} \zeta_1(F) + \lambda_{n_2} \zeta_2(F) + \cdots + \lambda_{n_m} \zeta_m(F) \right],$$

where $\lambda_{n_1} = O\left(\frac{1}{n}\right)$, $\lambda_{n_c} = O\left(\frac{1}{n^{c-1}}\right)$; $c \geq 2$.

Note 1: $V_F(U) < \infty$ iff $\zeta_m(F) < \infty$, $V_F(U) > 0$ iff $\zeta_1(F) > 0$ and $V_F(U) = 0$ iff $\zeta_m(F) = 0$.

Note 2: If $\zeta_m(F) < \infty$, then $V_F(U) \to 0$ as $n \to \infty$ and hence U is a consistent estimator of $\theta(F)$ [as $E_F(U) = \theta(F)$].

Result: Under appropriate condition(s), as $n \to \infty$,

$$\sqrt{n}\,[U - \theta(F)] \xrightarrow{\mathcal{D}} N(0, m^2 \xi_1(F))$$

Proof: Note that, as $X_1, \dots, X_n$ are iid, $\phi_1(X_1), \dots, \phi_1(X_n)$ are iid with

$$E_F[\phi_1(X_1)] = \theta(F)$$

$$V_F[\phi_1(X_1)] = \xi_1(F) < \infty$$

Hence, by CLT, as $n \to \infty$,

$$Y_n = \sum_{i=1}^{n} m \frac{[\phi_1(X_i) - \theta(F)]}{\sqrt{n}} \xrightarrow{\mathcal{D}} N(0, m^2 \xi_1(F)).$$

Writing $Z_n = \sqrt{n}\,(U - \theta(F))$.

We have to show that $Z_n - Y_n \xrightarrow{P} 0$, as $n \to \infty$ which holds if

$$E_F(Z_n - Y_n)^2 \to 0, \text{ as } n \to \infty. \quad \text{--- (1)}$$

$$E_F(Z_n - Y_n)^2 = n\, E_F(U - \theta(F))^2 + m^2 \cdot E_F[\phi_1(X_1) - \theta(F)]^2 - 2m\, Cov_F\{U, \sum_{i=1}^{n}\phi_1(X_i)\} \quad \text{--- (2)}$$

Now

$$\begin{aligned} n \cdot E_F[U - \theta(F)]^2 &= n \cdot V_F(U) \\ &= m^2 \cdot \xi_1(F) + [\lambda_{n_1}\xi_1(F) + \lambda_{n_2}\xi_2(F) + \dots + \lambda_{n_m}\xi_m(F)] \\ &= m^2 \cdot \xi_1(F) + O(\tfrac{1}{n}), \text{ as } n \to \infty, \text{ if } \xi_m(F) < \infty. \end{aligned}$$

$$m^2 E_F[\phi_1(X_1) - \theta(F)]^2 = m^2 \cdot V_F(\phi_1(X_1)) = m^2 \cdot \xi_1(F)$$

$$2m\, Cov_F(U, \sum_{i=1}^{n} \phi_1(X_i))$$

$$= 2m \cdot \binom{n}{m}^{-1} \sum_{1 \le i_1 < \dots < i_m \le n} \sum_{i=1}^{n} Cov_F\{\phi(X_{i_1}, \dots, X_{i_m}), \phi_1(X_i)\}$$

$$= 2m \cdot \binom{n}{m}^{-1} \binom{n}{m} \binom{m}{1} Cov_F\{\phi(X_1, \dots, X_m), \phi_1(X_1)\} \quad \text{[As } X_i\text{'s are iid]}$$

$$= 2m^2 [E_F\{\phi(X_1, \dots, X_m)\, \phi_1(X_1)\} - \theta^2(F)]$$

$$= 2m^2 [E_F E_F\{\phi_1(X_1) \cdot \phi(X_1, \dots, X_m) / X_1\} - \theta^2(F)]$$

$$= 2m^2 [E_F \phi_1(X_1) \cdot E_F\{\phi(X_1, \dots, X_m)/X_1\} - \theta^2(F)]$$

$$= 2m^2 [E_F \phi_1^2(X_1) - \theta^2(F)]$$

$$= 2m^2 V_F[\phi_1(X_1)]$$

$$= 2m^2 \xi_1(F)$$

Hence, the RHS of (2) is

$$m^2 \zeta_1(F) + O\left(\tfrac{1}{n}\right) + m^2 \zeta_1(F) - 2m^2 \zeta_1(F)$$
$$= O\left(\tfrac{1}{n}\right), \text{ as } n \to \infty$$

Hence the required result follows.

Note 1: Suppose $\zeta_1(F) > 0$. Then we get

$$\frac{\sqrt{n}\,(U - \theta(F))}{\sqrt{m^2 \zeta_1(F)}} \xrightarrow{\mathcal{D}} N(0,1), \text{ as } n \to \infty$$

Note 2: Suppose $0 < \zeta_1(F) < \cdots < \zeta_m(F) < \infty$

Then
$$\frac{U - \theta(F)}{\sqrt{V_F(U)}} = \frac{\sqrt{n}\,(U - \theta(F))}{\sqrt{m^2 \zeta_1(F)}} \cdot \sqrt{\frac{\frac{m^2}{n}\zeta_1(F)}{V_F(U)}}$$
$$= \frac{\sqrt{n}\,(U - \theta(F))}{\sqrt{m^2 \zeta_1(F)}} \cdot \sqrt{\frac{\frac{m^2}{n}\zeta_1(F)}{\frac{m^2}{n}\zeta_1(F) + O\left(\frac{1}{n^2}\right)}}$$
$$= \frac{\sqrt{n}\,(U - \theta(F))}{\sqrt{m^2 \zeta_1(F)}} \cdot \sqrt{\frac{1}{1 + O\left(\frac{1}{n}\right)}}$$
$$\xrightarrow{\mathcal{D}} N(0,1), \text{ as } n \to \infty \text{ [by note 1]}$$

Note 3: For every $\epsilon > 0$, we have

$$P_F\{|U - \theta(F)| < \epsilon\} = P_F\{\underset{\downarrow\,-\infty}{-\epsilon\sqrt{n}} < \underset{\downarrow\,\mathcal{D}\; N(0,\, m^2\zeta_1(F))}{\sqrt{n}\,(U - \theta(F))} < \underset{\downarrow\,+\infty}{\epsilon\sqrt{n}}\} \quad \text{as } n \to \infty$$

$$\to 1, \text{ as } n \to \infty$$

$$\Rightarrow U \xrightarrow{P} \theta(F)$$

i.e. U is a consistent estimator of $\theta(F)$.