An Empirical Likelihood Approach of The Estimation of Linear Functionals with Side Information

Shan Wang

Department of Mathematical Sciences
Purdue School of Science
Indiana University-Purdue University at Indianapolis

July 30, 2013

Outline

Problem we are interested in

- Let

$$Z_1, \ldots, Z_n \sim i.i.d. \quad Z$$

Problem we are interested in

- Let

$$Z_1, \ldots, Z_n \sim i.i.d. \quad Z$$

- We are intersted in estimating the expectation

$$\theta = E(\psi(Z)) = \int \psi dP$$

for some known square-integrable function ψ from $\mathscr{Z}$ into $\mathbb{R}^d$ with side information expressed by

Problem we are interested in

- Let

$$Z_1, \ldots, Z_n \sim i.i.d. \quad Z$$

- We are intersted in estimating the expectation

$$\theta = E(\psi(Z)) = \int \psi dP$$

for some known square-integrable function ψ from $\mathscr{Z}$ into $\mathbb{R}^d$ with side information expressed by

$$E[u_k(Z)] = 0, \quad k = 1, \ldots, m$$

Problem we are interested in

- Let

$$Z_1, \ldots, Z_n \sim i.i.d. \quad Z$$

- We are intersted in estimating the expectation

$$\theta = E(\psi(Z)) = \int \psi dP$$

for some known square-integrable function ψ from $\mathscr{Z}$ into $\mathbb{R}^d$ with side information expressed by

$$E[u_k(Z)] = 0, \quad k = 1, ..., m$$

where $u_1, ...u_m$ are measurable functions from $\mathscr{Z}$ to $\mathbb{R}$ such that $Var(u_k(Z)) < \infty$.

Examples

- Example 1.1
  Estimation of Distribution function: Let

$$\psi_{x,y}(X,Y) = \mathbf{1}[X \leq x, Y \leq y]$$

for known x, y, with side infromation e.g. known marginal distribution of X. Then $\theta = E(\psi(X,Y) = F(x,y)$ is the distribution function.

Examples

- Example 1.1
  Estimation of Distribution function: Let

$$\psi_{x,y}(X,Y) = \mathbf{1}[X \le x, Y \le y]$$

for known x, y, with side infromation e.g. known marginal distribution of X. Then $\theta = E(\psi(X,Y) = F(x,y)$ is the distribution function.

Example 1.3

Example 1.3

- Estimation of depth function:

Example 1.3

- Estimation of depth function:
- Let $\psi_x(X) = \frac{x-X}{\|x-X\|}$, then

$$D(x, F) = 1 - \|E\psi(x - X)\|$$

is called the depth function, and $D(x, F)$ is the depth value of x w.r.t distribution F.

Example 1.3

- Estimation of depth function:
- Let $\psi_x(X) = \frac{x-X}{\|x-X\|}$, then

$$D(x, F) = 1 - \|E\psi(x - X)\|$$

is called the depth function, and $D(x, F)$ is the depth value of x w.r.t distribution F.

- With side infromation, for instance, X is spherical symmetric about some center μ, we can estimate the spatial outlyingness function $O(x, F) = E(\psi(x - X))$, then get the plug-in estimator of $D(x, F)$.

Method

- The approach we are using is Empirical Likelihood introduced by Owen(1988, 1990, 2001).

Method

- The approach we are using is Empirical Likelihood introduced by Owen(1988, 1990, 2001).
- Incoporrating side information:
  The usual empirical estimation

$$\theta = \frac{1}{n}\sum_{j=1}^{n}\psi(Z_i),$$

does not use the side information.

Method

- The approach we are using is Empirical Likelihood introduced by Owen(1988, 1990, 2001).
- Incoporrating side information:
  The usual empirical estimation

$$\theta = \frac{1}{n}\sum_{j=1}^{n}\psi(Z_i),$$

  does not use the side information.
- Improved estimation is in the sense of least dispersed regular estimates.

Nonparametric (Empirical) Likelihood Approach

- Let $X_1, ..., X_n \sim$ iid distribution function (DF) F_0. Let F be an arbitrary DF.

Nonparametric (Empirical) Likelihood Approach

- Let $X_1, ..., X_n \sim$ iid distribution function (DF) F_0. Let F be an arbitrary DF.
- Nonparametric likelihood of F:

$$\mathscr{L}_n(F) = \prod_{i=1}^{n} \{F(X_i) - F(X_i-)\}.$$

Nonparametric (Empirical) Likelihood Approach

- Let $X_1, ..., X_n \sim$ iid distribution function (DF) F_0. Let F be an arbitrary DF.
- Nonparametric likelihood of F:
$$\mathscr{L}_n(F) = \prod_{i=1}^n \{F(X_i) - F(X_i-)\}.$$
- Nonparametric Likelihood Ratio Function:
$$\mathscr{R}_n(\mu) = \frac{\max\{\mathscr{L}_n(F) : E_F(X) = \mu\}}{\max\{\mathscr{L}_n(F) : F\}}$$

Nonparametric (Empirical) Likelihood Approach

- Let $X_1, ..., X_n \sim$ iid distribution function (DF) F_0. Let F be an arbitrary DF.
- Nonparametric likelihood of F:

$$\mathscr{L}_n(F) = \prod_{i=1}^n \{F(X_i) - F(X_i-)\}.$$

- Nonparametric Likelihood Ratio Function:

$$\mathscr{R}_n(\mu) = \frac{\max\{\mathscr{L}_n(F) : E_F(X) = \mu\}}{\max\{\mathscr{L}_n(F) : F\}}$$

- Note $\mathscr{L}_n(F) = 0$ if F is continuous. Choose F to be supported on $\{X_1, \ldots, X_n\}$ and place point mass π_i at X_i. Then

$$\pi \in \mathscr{P}_n, \quad E_F(X) = \pi_1 X_1 + ... + \pi_n X_n.$$

where $\mathscr{P}_n = \{\pi = (\pi_1, \ldots, \pi_n) : \pi_i \geq 0, \pi_1 + ... + \pi_n = 1\}$.

Nonparametric (Empirical) Likelihood Approach

- Let $X_1, ..., X_n \sim$ iid distribution function (DF) F_0. Let F be an arbitrary DF.
- Nonparametric likelihood of F:
$$\mathscr{L}_n(F) = \prod_{i=1}^n \{F(X_i) - F(X_i-)\}.$$
- Nonparametric Likelihood Ratio Function:
$$\mathscr{R}_n(\mu) = \frac{\max\{\mathscr{L}_n(F) : E_F(X) = \mu\}}{\max\{\mathscr{L}_n(F) : F\}}$$
- Note $\mathscr{L}_n(F) = 0$ if F is continuous. Choose F to be supported on $\{X_1, \ldots, X_n\}$ and place point mass π_i at X_i. Then
$$\pi \in \mathscr{P}_n, \quad E_F(X) = \pi_1 X_1 + ... + \pi_n X_n.$$
where $\mathscr{P}_n = \{\pi = (\pi_1, \ldots, \pi_n) : \pi_i \geq 0, \pi_1 + ... + \pi_n = 1\}$.

The Wilks Theorem of EL for Univariate Mean

- Since $\max\{\mathscr{L}_n(F) : F\} = (1/n)^n$, it follows

$$\begin{aligned}\mathscr{R}_n(\mu) &= \frac{\sup\{\prod_{i=1}^n \pi_i : \pi \in \mathscr{P}_n, \sum_{i=1}^n \pi_i X_i = \mu\}}{(1/n)^n} \\ &= \sup\left\{\prod_{i=1}^n n\pi_i : \pi \in \mathscr{P}_n, \sum_{i=1}^n \pi_i X_i = \mu\right\}\end{aligned}$$

The Wilks Theorem of EL for Univariate Mean

- Since $\max\{\mathscr{L}_n(F) : F\} = (1/n)^n$, it follows

$$\begin{aligned}\mathscr{R}_n(\mu) &= \frac{\sup\{\prod_{i=1}^n \pi_i : \pi \in \mathscr{P}_n, \sum_{i=1}^n \pi_i X_i = \mu\}}{(1/n)^n} \\ &= \sup\left\{\prod_{i=1}^n n\pi_i : \pi \in \mathscr{P}_n, \sum_{i=1}^n \pi_i X_i = \mu\right\}\end{aligned}$$

- Suppose $X_1, ..., X_n$ are iid with $0 < Var(X_1) < \infty$. Then

$$-2\log \mathscr{R}_n(\mu_0) \Rightarrow \chi_1^2$$

The Wilks Theorem of EL for Univariate Mean

- Since $\max\{\mathscr{L}_n(F) : F\} = (1/n)^n$, it follows

$$\mathscr{R}_n(\mu) = \frac{\sup\{\prod_{i=1}^n \pi_i : \pi \in \mathscr{P}_n, \sum_{i=1}^n \pi_i X_i = \mu\}}{(1/n)^n} = \sup\left\{\prod_{i=1}^n n\pi_i : \pi \in \mathscr{P}_n, \sum_{i=1}^n \pi_i X_i = \mu\right\}$$

- Suppose $X_1, ..., X_n$ are iid with $0 < Var(X_1) < \infty$. Then

$$-2\log \mathscr{R}_n(\mu_0) \Rightarrow \chi_1^2$$

Construction

Construction

- Available side information expressed via

$$E(u_k(Z)) = 0, \quad k = 1, ..., m$$

Construction

- Available side information expressed via

$$E(u_k(Z)) = 0, \quad k = 1, ..., m$$

- Following Owen's method,

$$\mathscr{R}_n = \sup\left\{\prod_{j=1}^{n} n\pi_j : \pi \in \mathscr{P}_n, \sum_{j=1}^{n} \pi_j u_k(Z_j) = 0, k = 1, 2, ..., m\right\},$$

where

$$\mathscr{P}_n = \{\pi = (\pi_1, \ldots, \pi_n)^\top \in [0,1]^n : \sum_{i=1}^{n} \pi_i = 1\}.$$

Construction

- Using Lagrange multipliers one derives that

$$\tilde{\pi}_j = \frac{1}{n}\frac{1}{1+\tilde{\zeta}^\top \mathbf{u}(Z_j)}, \quad j = 1, \ldots, n,$$

Construction

- Using Lagrange multipliers one derives that

$$\tilde{\pi}_j = \frac{1}{n} \frac{1}{1 + \tilde{\zeta}^\top \mathbf{u}(Z_j)}, \quad j = 1, \dots, n,$$

- where $\mathbf{u} = (u_1, \dots, u_m)^\top$ and $\tilde{\zeta}$ solves the equation

$$\sum_{j=1}^{n} \frac{\mathbf{u}(Z_j)}{1 + \zeta^\top \mathbf{u}(Z_j)} = 0.$$

under certain conditions, and the solution is unique.

- A natural estimate $\tilde{\theta}$ of θ is given by

$$\tilde{\theta} = \sum_{j=1}^{n} \tilde{\pi}_j \psi(Z_j).$$

- A natural estimate $\tilde{\theta}$ of θ is given by

$$\tilde{\theta} = \sum_{j=1}^{n} \tilde{\pi}_j \psi(Z_j).$$

- This estimator of θ improves the efficiency of the empirical estimator $\bar{\psi} = \frac{1}{n} \sum_{j=1}^{n} \psi(Z_j)$ as is shown later. As a matter of fact, it is the maximum empirical likelihood estimator (MELE) of θ.

Efficiency

Efficiency

- Finitely many known $u_1, ..., u_m$

Efficiency

- Finitely many known $u_1, ..., u_m$
  Example 4.1: Estimate distribution function when mean is zero

Efficiency

- Finitely many known $u_1, ..., u_m$
  Example 4.1: Estimate distribution function when mean is zero
- Let $X_1, ..., X_n \sim i.i.d. \quad X$ with side information $E(X) = 0$

Efficiency

- Finitely many known $u_1, ..., u_m$
  Example 4.1: Estimate distribution function when mean is zero
- Let $X_1, ..., X_n \sim i.i.d. \quad X$ with side information $E(X) = 0$
- Let $\psi_x(X) = \mathbf{1}[X \leq x]$, we're interested in estimating

$$\theta(x) = E(\psi(X)) = F(x)$$

where x is a known number.

Efficiency

- Finitely many known $u_1, ..., u_m$
  Example 4.1: Estimate distribution function when mean is zero
- Let $X_1, ..., X_n \sim i.i.d. \quad X$ with side information $E(X) = 0$
- Let $\psi_x(X) = \mathbf{1}[X \leq x]$, we're interested in estimating

$$\theta(x) = E(\psi(X)) = F(x)$$

  where x is a known number.
- By our approach,

$$\mathscr{R}_n = \sup\left\{\prod_{j=1}^{n} n\pi_j : \pi \in \mathscr{P}_n, \sum_{j=1}^{n} \pi_j(X_j) = 0\right\}$$

Efficiency

- Finitely many known $u_1, ..., u_m$
  Example 4.1: Estimate distribution function when mean is zero
- Let $X_1, ..., X_n \sim i.i.d. \quad X$ with side information $E(X) = 0$
- Let $\psi_x(X) = \mathbf{1}[X \leq x]$, we're interested in estimating

$$\theta(x) = E(\psi(X)) = F(x)$$

where x is a known number.
- By our approach,

$$\mathscr{R}_n = \sup\left\{ \prod_{j=1}^{n} n\pi_j : \pi \in \mathscr{P}_n, \ \sum_{j=1}^{n} \pi_j(X_j) = 0 \right\}$$

- A nature estimator is given by

$$\tilde{F}(x) = \tilde{\theta}(x) = \frac{1}{n} \sum_{j=1}^{n} \frac{\mathbf{1}[X_j \leq x]}{1 + \tilde{\zeta} X_j}.$$

THEOREM 1.1

- Suppose $u_1, \dots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If $W = E(\mathbf{u}(Z)\mathbf{u}(Z)^\top)$ is positive definite, then $\tilde{\theta}$ satisfies the stochastic expansion,

$$\tilde{\theta} = \bar{\psi} - \bar{\phi}_0 + o_p(n^{-1/2}),$$

where $\phi_0 = \Pi(\psi|[\mathbf{u}]) = E(\psi \otimes \mathbf{u}^\top)W^{-1}\mathbf{u}$.

THEOREM 1.1

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If $W = E(\mathbf{u}(Z)\mathbf{u}(Z)^\top)$ is positive definite, then $\tilde{\theta}$ satisfies the stochastic expansion,

$$\tilde{\theta} = \bar{\psi} - \bar{\phi}_0 + o_p(n^{-1/2}),$$

where $\phi_0 = \Pi(\psi|[\mathbf{u}]) = E(\psi \otimes \mathbf{u}^\top)W^{-1}\mathbf{u}$.
Thus,

$$\sqrt{n}(\tilde{\theta} - \theta) \implies \mathcal{N}(0, V_1)$$

THEOREM 1.1

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If $W = E(\mathbf{u}(Z)\mathbf{u}(Z)^\top)$ is positive definite, then $\tilde{\theta}$ satisfies the stochastic expansion,

$$\tilde{\theta} = \bar{\psi} - \bar{\phi}_0 + o_p(n^{-1/2}),$$

where $\phi_0 = \Pi(\psi|[\mathbf{u}]) = E(\psi \otimes \mathbf{u}^\top)W^{-1}\mathbf{u}$.
Thus,

$$\sqrt{n}(\tilde{\theta} - \theta) \implies \mathcal{N}(0, V_1)$$

where $V_1 = Var(\psi(Z)) - Var(\phi_0(Z))$.

THEOREM 1.1

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If $W = E(\mathbf{u}(Z)\mathbf{u}(Z)^\top)$ is positive definite, then $\tilde{\theta}$ satisfies the stochastic expansion,

$$\tilde{\theta} = \bar{\psi} - \bar{\phi}_0 + o_p(n^{-1/2}),$$

where $\phi_0 = \Pi(\psi|[\mathbf{u}]) = E(\psi \otimes \mathbf{u}^\top)W^{-1}\mathbf{u}$.
Thus,

$$\sqrt{n}(\tilde{\theta} - \theta) \implies \mathcal{N}(0, V_1)$$

where $V_1 = Var(\psi(Z)) - Var(\phi_0(Z))$.

- By the above theorem, the estimator $\tilde{\theta}$ has smaller variance than the variance of the empirical estimator $\bar{\psi}$.

Example 4.1 continued:

Example 4.1 continued:

- In this case, $u(X) = X$, $W = E(X_1^2) = Var(X_j) = \sigma^2 > 0$

Example 4.1 continued:

- In this case, $u(X) = X$, $W = E(X_1^2) = Var(X_j) = \sigma^2 > 0$
- Recall that

$$\tilde{F}(x) = \tilde{\theta}(x) = \frac{1}{n}\sum_{j=1}^{n} \frac{\mathbf{1}[X_j \le x]}{1 + \tilde{\zeta} X_j} = F_n(x) - \bar{\phi}_0 + o_p(n^{-1/2}),$$

Example 4.1 continued:

- In this case, $u(X) = X$, $W = E(X_1^2) = Var(X_j) = \sigma^2 > 0$
- Recall that

$$\tilde{F}(x) = \tilde{\theta}(x) = \frac{1}{n}\sum_{j=1}^{n} \frac{\mathbf{1}[X_j \le x]}{1 + \tilde{\zeta}X_j} = F_n(x) - \bar{\phi}_0 + o_p(n^{-1/2}),$$

where

$$\phi_0(X) = E(\psi \otimes \mathbf{u}^\top)W^{-1}\mathbf{u} = \frac{E[\mathbf{1}[X \le x]X]X}{\sigma^2}.$$

Then

$$V_1 = F(x)(1 - F(x)) - E[\mathbf{1}[X \le x]X]^2$$

THEOREM 1.2

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \ldots, \hat{u}_m$ be their estimators respectively such that

THEOREM 1.2

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \ldots, \hat{u}_m$ be their estimators respectively such that

$$\max_{1 \le j \le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(n^{1/2}),$$

THEOREM 1.2

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \ldots, \hat{u}_m$ be their estimators respectively such that

$$\max_{1 \le j \le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(n^{1/2}),$$

$$|\hat{W}_n - W|_o = o_p(1),$$

THEOREM 1.2

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \ldots, \hat{u}_m$ be their estimators respectively such that

$$\max_{1 \leq j \leq n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(n^{1/2}),$$

$$|\hat{W}_n - W|_o = o_p(1),$$

where $\hat{W}_n = \frac{1}{n} \sum_{j=1}^{n} \hat{\mathbf{u}}_n(Z_j) \hat{\mathbf{u}}_n(Z_j)^\top$, and

THEOREM 1.2

- Suppose $u_1, \ldots, u_m$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \ldots, \hat{u}_m$ be their estimators respectively such that

$$\max_{1\le j\le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(n^{1/2}),$$

$$|\hat{W}_n - W|_o = o_p(1),$$

where $\hat{W}_n = \frac{1}{n}\sum_{j=1}^n \hat{\mathbf{u}}_n(Z_j)\hat{\mathbf{u}}_n(Z_j)^\top$, and

$$\frac{1}{n}\sum_{j=1}^n (\psi(Z_j) \otimes \hat{\mathbf{u}}(Z_j)^\top - E(\psi(Z_j) \otimes \hat{\mathbf{u}}(Z_j)^\top)) = o_p(1),$$

THEOREM 1.2 continued

Also,for some measurable function v_n from $\mathscr{X}$ into $\mathbb{R}^m$ such that $\int \mathbf{v}_n(Z)dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg, it satisfies

THEOREM 1.2 continued

Also,for some measurable function v_n from $\mathscr{Z}$ into $\mathbb{R}^m$ such that $\int \mathbf{v}_n(Z)dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg, it satisfies

$$\frac{1}{n}\sum_{j=1}^{n} E\left(|\hat{\mathbf{u}}_n(Z_j) - \mathbf{v}_n(Z_j)|^2\right) = o_p(1),$$

THEOREM 1.2 continued

Also,for some measurable function v_n from $\mathscr{Z}$ into $\mathbb{R}^m$ such that $\int \mathbf{v}_n(Z)dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg, it satisfies

$$\frac{1}{n}\sum_{j=1}^{n} E\left(|\hat{\mathbf{u}}_n(Z_j) - \mathbf{v}_n(Z_j)|^2\right) = o_p(1),$$

$$\frac{1}{n}\sum_{j=1}^{n} \hat{\mathbf{u}}_n(Z_j) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{v}_n(Z_j) + o_p(n^{-1/2}),$$

THEOREM 1.2 continued

- Then $\hat{\theta}$ satisfies the stochastic expansion,

$$\hat{\theta} = \bar{\psi} - \bar{\phi} + o_p(n^{-1/2}),$$

where $\phi = E(\psi \otimes \mathbf{u}^\top) W^{-1} \mathbf{u}$.

THEOREM 1.2 continued

- Then $\hat{\theta}$ satisfies the stochastic expansion,

$$\hat{\theta} = \bar{\psi} - \bar{\phi} + o_p(n^{-1/2}),$$

where $\phi = E(\psi \otimes \mathbf{u}^\top)W^{-1}\mathbf{u}$.
Thus

$$\sqrt{n}(\hat{\theta} - \theta) \implies \mathcal{N}(0, V_2),$$

where $V_2 = Var(\psi(Z)) - Var(\phi(Z))$

- Infinitely many constraints

- Infinitely many constraints

  Example 4.3 Estimate distribution function when the marginal distribution is known

- Infinitely many constraints

  Example 4.3 Estimate distribution function when the marginal distribution is known
- Suppose X, Y have joint distribution H. Suppose the marginal distribution function of X is known and equal to F_0.

- Infinitely many constraints

  Example 4.3 Estimate distribution function when the marginal distribution is known

- Suppose X, Y have joint distribution H. Suppose the marginal distribution function of X is known and equal to F_0.
- Suppose F_0 is continous, then the above information is equivalent to

$$\int a(x)dH(x,y) = \int a(x)dF_0(x) = 0, \quad \forall a \in L_{2,0}(F_0)$$

- Infinitely many constraints

  Example 4.3 Estimate distribution function when the marginal distribution is known
- Suppose X, Y have joint distribution H. Suppose the marginal distribution function of X is known and equal to F_0.
- Suppose F_0 is continous, then the above information is equivalent to

$$\int a(x)dH(x,y) = \int a(x)dF_0(x) = 0, \quad \forall a \in L_{2,0}(F_0)$$

- This can be reduced to countably many constriants

$$\int a_k(x)dH(x,y) = \int a_k(x)dF_0(x) = 0, \quad k = 1, 2, \ldots$$

where $a_k(x)$ is an othonormal basis of $L_{2,0}(F_0)$.

- Rewrite conditions to be $E(a_k(X)) = 0$

- Rewrite conditions to be $E(a_k(X)) = 0$
- Trigonometric basis $\phi_1, \phi_2, \ldots$ defined by

$$\phi_k(t) = \sqrt{2}\cos(k\pi t), \quad t \in [0,1], \quad k = 1, 2, \ldots.$$

- Rewrite conditions to be $E(a_k(X)) = 0$
- Trigonometric basis $\phi_1, \phi_2, ...$ defined by

$$\phi_k(t) = \sqrt{2}\cos(k\pi t), \quad t \in [0,1], \quad k = 1, 2, ...$$

- Then $u_k(X_j) = \phi_k(F_0(X_j)), \quad k = 1, 2, ..., m,$

- Rewrite conditions to be $E(a_k(X)) = 0$
- Trigonometric basis $\phi_1, \phi_2, ...$ defined by

$$\phi_k(t) = \sqrt{2}\cos(k\pi t), \quad t \in [0,1], \quad k = 1, 2, ...$$

- Then $u_k(X_j) = \phi_k(F_0(X_j)), \quad k = 1, 2, ..., m,$
- Estimate $\theta = F(x,y) = E(\psi(x,y))$, where

$$\psi(x,y) = \mathbf{1}[X \leq x, Y \leq y],$$

- Rewrite conditions to be $E(a_k(X)) = 0$
- Trigonometric basis $\phi_1, \phi_2, ...$ defined by

$$\phi_k(t) = \sqrt{2}\cos(k\pi t), \quad t \in [0,1], \quad k = 1, 2, ...$$

- Then $u_k(X_j) = \phi_k(F_0(X_j)), \quad k = 1, 2, ..., m,$
- Estimate $\theta = F(x, y) = E(\psi(x, y))$, where

$$\psi(x, y) = \mathbf{1}[X \le x, Y \le y],$$

- then a nature estimator would be

$$\tilde{\theta} = \frac{1}{n}\sum_{j=1}^{n} \frac{\mathbf{1}[X \le x, Y \le y]}{1 + \tilde{\zeta}^{\top}\Phi(F_0(X_j))},$$

where $\Phi(t) = (\phi_1(t), ..., \phi_m(t))^{\top}$

THEOREM 1.3

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If

THEOREM 1.3

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If

$$\max_{1 \leq j \leq n} |\mathbf{u}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

THEOREM 1.3

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. If

$$\max_{1 \leq j \leq n} |\mathbf{u}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

the $m_n \mathbf{x} m_n$ disperson matrix W_n is regular and satisfies

$$|S_n - W_n|_o = o_p(m_n^{-1})$$

THEOREM 1.3

- Suppose that $u_1, u_2, \ldots$ are known functions *s.t.* $E[u_k(Z)] = 0$ and sencond moment finite. If

$$\max_{1\leq j\leq n} |\mathbf{u}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

the $m_n \mathbf{x} m_n$ disperson matrix W_n is regular and satisfies

$$|S_n - W_n|_o = o_p(m_n^{-1})$$

and

$$|\frac{1}{n}\sum_{j=1}^{n}\left(\psi(Z_j)\otimes \mathbf{u}(Z_j)^{\top} - E\left(\psi(Z_j)\otimes \mathbf{u}(Z_j)^{\top}\right)\right)|_o = o_p(m_n^{-1/2}).$$

THEOREM 1.3(continued)

- Then $\tilde{\theta}_n$ satisfies, as m_n tends to infinity, the stochastic expansion,

$$\tilde{\theta}_n = \bar{\psi} - \bar{\varphi}_0 + o_p(n^{-1/2}),$$

where $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$ is the projection of ψ onto the closed linear span $[\mathbf{u}_\infty]$.

THEOREM 1.3(continued)

- Then $\tilde{\theta}_n$ satisfies, as m_n tends to infinity, the stochastic expansion,

$$\tilde{\theta}_n = \bar{\psi} - \bar{\varphi}_0 + o_p(n^{-1/2}),$$

where $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$ is the projection of ψ onto the closed linear span $[\mathbf{u}_\infty]$.
Thus

$$\sqrt{n}(\tilde{\theta}_n - \theta) \implies \mathcal{N}(0, V_3).$$

where $V_3 = Var(\psi(Z)) - Var(\varphi_0(Z))$.

Example 4.3 continued

- Apply theorem 1.3, check conditions :
  For $u_k(Z_j) = \phi_k(F_0(X_j)) = \sqrt{2}\cos(k\pi F_0(X_j))$,

Example 4.3 continued

- Apply theorem 1.3, check conditions :
  For $u_k(Z_j) = \phi_k(F_0(X_j)) = \sqrt{2}\cos(k\pi F_0(X_j))$,
- (1) u_k has mean zero and finte variances

Example 4.3 continued

- Apply theorem 1.3, check conditions :
  For $u_k(Z_j) = \phi_k(F_0(X_j)) = \sqrt{2}\cos(k\pi F_0(X_j))$,
- (1) u_k has mean zero and finte variances
- (2) $\max_{1\leq j\leq n} |\mathbf{u}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}) \Longleftarrow m_n^3 = o_p(n)$

Example 4.3 continued

- Apply theorem 1.3, check conditions :
  For $u_k(Z_j) = \phi_k(F_0(X_j)) = \sqrt{2}\cos(k\pi F_0(X_j))$,
- (1) u_k has mean zero and finte variances
- (2) $\max_{1\le j\le n}|\mathbf{u}_n(Z_j)| = o_p(m_n^{-1}n^{1/2}) \Longleftarrow m_n^3 = o_p(n)$
- (3) W_n is regular and satisfies $|S_n - W_n|_o = o_p(m_n^{-1})$

$$\Longleftarrow W_n = I, m_n^4 = o_p(n)$$

Example 4.3 continued

- Apply theorem 1.3, check conditions :
  For $u_k(Z_j) = \phi_k(F_0(X_j)) = \sqrt{2}\cos(k\pi F_0(X_j))$,
- (1) u_k has mean zero and finte variances
- (2) $\max_{1\leq j\leq n} |\mathbf{u}_n(Z_j)| = o_p(m_n^{-1}n^{1/2}) \Longleftarrow m_n^3 = o_p(n)$
- (3) W_n is regular and satisfies $|S_n - W_n|_o = o_p(m_n^{-1})$

$$\Longleftarrow W_n = I, m_n^4 = o_p(n)$$

- (4)
  $\frac{1}{n}\sum_{j=1}^{n}\left(\psi(Z_j)\otimes\mathbf{u}(Z_j)^\top - E\left(\psi(Z_j)\otimes\mathbf{u}(Z_j)^\top\right)\right) = o_p(m_n^{-1/2})$

$$\Longleftarrow m_n^2 = o_p(n)$$

Example 4.3 continued

- After satisfying all the conditions, we can apply theorem 1.3 that the estimator

Example 4.3 continued

- After satisfying all the conditions, we can apply theorem 1.3 that the estimator
$$\tilde{\theta}_n = \bar{\psi} - \bar{\varphi}_0 + o_p(n^{-1/2}),$$

Example 4.3 continued

- After satisfying all the conditions, we can apply theorem 1.3 that the estimator

$$\tilde{\theta}_n = \bar{\psi} - \bar{\varphi}_0 + o_p(n^{-1/2}),$$

and when $m_n^4 = o_p(n)$,

$$\sqrt{n}(\tilde{\theta}_n - \theta) \implies \mathcal{N}(0, V_3)$$

Example 4.3 continued

- Calculation of $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$:

Example 4.3 continued

- Calculation of $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$:

$$\varphi_0 = \Pi(\psi|L_{2,0}(F_0))$$

Example 4.3 continued

- Calculation of $\varphi_0 = \Pi(\psi | [\mathbf{u}_\infty])$:

$$\varphi_0 = \Pi(\psi | L_{2,0}(F_0))$$

$$\Leftrightarrow (\psi - \varphi_0) \perp L_{2,0}(F_0)$$

Example 4.3 continued

- Calculation of $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$:

$$\varphi_0 = \Pi(\psi|L_{2,0}(F_0))$$

$$\Leftrightarrow (\psi - \varphi_0) \perp L_{2,0}(F_0)$$

$$\Leftrightarrow E((\psi - \varphi_0)a) = 0, \forall a \in L_{2,0}(F_0)$$

Example 4.3 continued

- Calculation of $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$:

$$\varphi_0 = \Pi(\psi|L_{2,0}(F_0))$$

$$\Leftrightarrow (\psi - \varphi_0) \perp L_{2,0}(F_0)$$

$$\Leftrightarrow E((\psi - \varphi_0)a) = 0, \forall a \in L_{2,0}(F_0)$$

$$\Leftrightarrow E((\psi(X,Y) - \varphi_0(X))a(X)) = 0, \forall a \in L_{2,0}(F_0)$$

Example 4.3 continued

- Calculation of $\varphi_0 = \Pi(\psi|[\mathbf{u}_\infty])$:

$$\varphi_0 = \Pi(\psi|L_{2,0}(F_0))$$

$$\Leftrightarrow (\psi - \varphi_0) \perp L_{2,0}(F_0)$$

$$\Leftrightarrow E((\psi - \varphi_0)a) = 0, \forall a \in L_{2,0}(F_0)$$

$$\Leftrightarrow E((\psi(X,Y) - \varphi_0(X))a(X)) = 0, \forall a \in L_{2,0}(F_0)$$

Change it to the conditional expectation given $X = x$, we can solve

$$\varphi_0(x) = E(\psi(X,Y)|X = x)$$

Example 4.3 continued

- **Theorem:**
  Let $(X_1, Y_1), ..., (X_n, Y_n)$ be i.i.d. from H. Suppose X_1 has known distribution F_0, which is continous. Then $\tilde{\theta}$ is an efficient estimator of θ such that

$$\tilde{\theta} = \frac{1}{n}\sum_{j=1}^{n}\psi(X_j, Y_j) - \frac{1}{n}\sum_{j=1}^{n}\varphi_0(X_j) + o_p(n^{-1/2})$$

where φ_0 is the projector of ψ onto $L_{2,0}(F_0)$, with formular

$$\varphi_0(x) = E(\psi(X, Y)|X = x),$$

Thus

$$\sqrt{n}(\tilde{\theta}_n - \theta) \implies \mathcal{N}(0, \sigma^2).$$

where $\sigma^2 = Var(\psi) - Var(\varphi_0)$.

THEOREM 1.4

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \hat{u}_2, \ldots$ be their estimators respectively such that

THEOREM 1.4

- Suppose that $u_1, u_2, \dots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \hat{u}_2, \dots$ be their estimators respectively such that

$$\max_{1 \le j \le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

THEOREM 1.4

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \hat{u}_2, \ldots$ be their estimators respectively such that

$$\max_{1 \le j \le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

$$|\hat{W}_n - W_n|_o = o_p(m_n^{-1}),$$

THEOREM 1.4

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \hat{u}_2, \ldots$ be their estimators respectively such that

$$\max_{1 \le j \le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

$$|\hat{W}_n - W_n|_o = o_p(m_n^{-1}),$$

$$\frac{1}{n} \sum_{j=1}^{n} \Big(\psi(Z_j) \otimes \hat{\mathbf{u}}_n(Z_j)^\top - E\big(\psi(Z_j) \otimes \hat{\mathbf{u}}_n(Z_j)^\top\big) \Big) = o_p(m_n^{-1/2}),$$

THEOREM 1.4

- Suppose that $u_1, u_2, \ldots$ are known functions $s.t.$ $E[u_k(Z)] = 0$ and sencond moment finite. Let $\hat{u}_1, \hat{u}_2, \ldots$ be their estimators respectively such that

$$\max_{1\le j\le n} |\hat{\mathbf{u}}_n(Z_j)| = o_p(m_n^{-1} n^{1/2}),$$

$$|\hat{W}_n - W_n|_o = o_p(m_n^{-1}),$$

$$\frac{1}{n}\sum_{j=1}^{n}\left(\psi(Z_j)\otimes\hat{\mathbf{u}}_n(Z_j)^\top - E\left(\psi(Z_j)\otimes\hat{\mathbf{u}}_n(Z_j)^\top\right)\right) = o_p(m_n^{-1/2}),$$

$$\frac{1}{n}\sum_{j=1}^{n} E\left(|\hat{\mathbf{u}}_n(Z_j) - \mathbf{v}_n(Z_j)|^2\right) = o_p(1),$$

THEOREM 1.4 continued

- for some measurable function $\mathbf{v}_n$ from $\mathscr{L}$ into $\mathbb{R}^{m_n}$ such that $\int \mathbf{v}_n(Z)\,dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg.

$$\frac{1}{n}\sum_{j=1}^{n}\hat{\mathbf{u}}_n(Z_j) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{v}_n(Z_j) + o_p(n^{-1/2}),$$

THEOREM 1.4 continued

- for some measurable function $\mathbf{v}_n$ from $\mathscr{L}$ into $\mathbb{R}^{m_n}$ such that $\int \mathbf{v}_n(Z)\,dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg.

$$\frac{1}{n}\sum_{j=1}^{n} \hat{\mathbf{u}}_n(Z_j) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{v}_n(Z_j) + o_p(n^{-1/2}),$$

- Then $\hat{\theta}$ satisfies, as m_n tends to infinity, the stochastic expansion,

$$\hat{\theta} = \bar{\psi} - \bar{\varphi} + o_p(n^{-1/2}),$$

THEOREM 1.4 continued

- for some measurable function $\mathbf{v}_n$ from $\mathscr{L}$ into $\mathbb{R}^{m_n}$ such that $\int \mathbf{v}_n(Z)\, dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg.
$$\frac{1}{n}\sum_{j=1}^{n} \hat{\mathbf{u}}_n(Z_j) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{v}_n(Z_j) + o_p(n^{-1/2}),$$
- Then $\hat{\theta}$ satisfies, as m_n tends to infinity, the stochastic expansion,
$$\hat{\theta} = \bar{\psi} - \bar{\varphi} + o_p(n^{-1/2}),$$
where $\varphi = \Pi(\psi|[\mathbf{v}_\infty])$ is the projection of ψ onto the closed linear span $[\mathbf{v}_\infty]$.

THEOREM 1.4 continued

- for some measurable function $\mathbf{v}_n$ from $\mathscr{L}$ into $\mathbb{R}^{m_n}$ such that $\int \mathbf{v}_n(Z)\, dP = 0$ and $|\mathbf{v}_n|$ is Lindeberg.

$$\frac{1}{n}\sum_{j=1}^{n} \hat{\mathbf{u}}_n(Z_j) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{v}_n(Z_j) + o_p(n^{-1/2}),$$

- Then $\hat{\theta}$ satisfies, as m_n tends to infinity, the stochastic expansion,

$$\hat{\theta} = \bar{\psi} - \bar{\varphi} + o_p(n^{-1/2}),$$

where $\varphi = \Pi(\psi|[\mathbf{v}_\infty])$ is the projection of ψ onto the closed linear span $[\mathbf{v}_\infty]$.
Thus

$$\sqrt{n}(\hat{\theta}_n - \theta) \implies \mathcal{N}(0, V_4).$$

where

$$V_4 = Var(\psi(Z)) - Var(\varphi(Z)).$$

Conclusion and Further Topics

- Based on previous discussion, when we obtain the side information, we can give a more efficiency estimator to the linear functional we are interested in.

Conclusion and Further Topics

- Based on previous discussion, when we obtain the side information, we can give a more efficiency estimator to the linear functional we are interested in.
- The same method can apply to many topics, Especially interests in Missing Data.

Thank you!