5. Other Examples of Semiparametric Models

野間 久史
統計数理研究所
2013年5月28日
E-mail: noma@ism.ac.jp
URL: http://www.ism.ac.jp/~noma/

Contexts

2

Contexts

3

Location-Shift回帰モデル

$$Y_i = \mu(X_i, \beta) + \varepsilon_i, \qquad i = 1, \dots, n$$

- Y_i: 結果変数（Continuous, Univariate）
- X_i: q次元共変量（Not fixed, Random）
- $\mu(X_i, \beta)$: 系統的な成分
- ε_i: ランダムな成分（※ 誤差項ではない）
  - $\varepsilon_i \coprod X_i$: 独立性の仮定
  - $E[\varepsilon_i] = 0$ という制約は置かない

Bickel et al. (1993), Manski (1984)

4

Semiparametricモデル

- Y_iの分布型は、ε_iの分布によって決まる
- 一方、Y_iの分布のLocationは、共変量 X_iの値によって決まり、どの程度Shiftするのかは、β によって決まる
- β : 関心のあるパラメータ
- $p_\varepsilon(\varepsilon), p_X(x)$ [ε_i, X_iの分布型] : 局外要因（本質的な関心はない）

5

制約モーメントトモデルとの違い

- 線形制約モーメントトモデル

$$Y_i = \alpha + \beta_1 X_{i1} + \cdots + \beta_q X_{iq} + \varepsilon_i$$
$$E[\varepsilon_i | X_i] = 0$$

$\varepsilon_i \coprod X_i$の仮定は必ずしも必要でない

- Location-shiftモデルのほうが強い制約をもったセミパラメトリックモデルといえる
- Efficiency Boundも小さくなるはず

6

局外接空間と直交補空間

- 目標：βのセミパラメトリックRAL推定量を求めて、有効推定量を与える
  - RAL推定量の影響関数の空間を求める必要がある
- まずは、局外接空間とその直交補空間を求めよう

7

復習1：定理 4.3（p.67）

- Semiparametric RAL推定量が存在するならば、影響関数は、

$$\{\phi(Z) + \mathfrak{I}^{\perp}\}$$

という空間に含まれる
  - $\phi(Z)$: 任意のSemiparametric RAL推定量の影響関数
  - $\mathfrak{I}$: セミパラメトリック接空間

8

復習2：定理 4.3（p.67）

- Semiparametric有効な推定量が存在するならば、影響関数は一意であり、

$$\phi_{\text{eff}}(Z) = \phi(Z) - \Pi\{\phi(Z)|\mathfrak{I}^{\perp}\}$$
$$= \Pi\{\phi(Z)|\mathfrak{I}\}$$

と書ける
  - $\phi(Z)$: 任意のSemiparametric RAL推定量の影響関数
  - $\mathfrak{I}$: セミパラメトリック接空間

9

局外接空間（定理 5.1）

- Location-shirtモデルの局外接空間は、以下のように与えられる

$$\Lambda = \Lambda_{1s} \oplus \Lambda_{2s}$$

$$\Lambda_{1s} = \left[a_1{}^{q\times 1}(\varepsilon): E\{a_1{}^{q\times 1}(\varepsilon)\} = 0^{q\times 1}\right]$$

$$\Lambda_{2s} = \left[a_2{}^{q\times 1}(X): E\{a_2{}^{q\times 1}(X)\} = 0^{q\times 1}\right]$$

また、$\Lambda_{1s} \perp \Lambda_{2s}$である。

10

射影の分解（定理 5.2）

- Λ の直交補空間は、

$$\Lambda^{\perp} = [\{h - \prod(h|\Lambda)\} \text{ for all } h \in \mathcal{H}]$$

と書ける。
- $\Lambda_{1s} \perp \Lambda_{2s}$であるので、

$$\prod(h|\Lambda) = \prod(h|\Lambda_{1s}) + \prod(h|\Lambda_{2s})$$

となる（定理 5.2）

11

局外接空間の直交補空間

- 任意の $h(\varepsilon, X) \in \mathcal{H}$において、$\Lambda_{2s}$への射影は、$\prod(h|\Lambda_{2s}) = E\{h(\varepsilon, X)|X\}$で与えられる
- 同様に、$\prod(h|\Lambda_{1s}) = E\{h(\varepsilon, X)|\varepsilon\}$である
- したがって、局外接空間の直交補空間は

$$\Lambda^{\perp} = ([h(\varepsilon, X) - E\{h(\varepsilon, X)|\varepsilon\} - E\{h(\varepsilon, X)|X\}] \text{ for all } h \in \mathcal{H})$$

と書ける

12

セミパラ推定量の影響関数

- $h(\varepsilon, X) = g(\varepsilon, X) - E\{g(\varepsilon, X)\}$
  - $g(\varepsilon, X)$: (ε, X)のq次元関数
  - 平均ゼロになるように基準化
- $\Lambda^{\perp}$の元の表現:

  $g(\varepsilon, X) - G_\varepsilon(X) - G_X(\varepsilon) + E\{g(\varepsilon, X)\}$
  - $G_\varepsilon(X) = E\{g(\varepsilon, X)|X\}$
  - $G_X(\varepsilon) = E\{g(\varepsilon, X)|\varepsilon\}$

13

下準備

- $\varepsilon \coprod X$ なので、$X = x$とFixすると、
  - $G_\varepsilon(x) = E\{g(\varepsilon, x)\}$
- $\varepsilon \coprod X$ なので、$\varepsilon = \varepsilon^*$ とFixすると、
  - $G_X(\varepsilon^*) = E\{g(\varepsilon^*, X)\}$
- それぞれの一致推定量は

$$\hat{G}_\varepsilon(x) = \frac{1}{n}\sum_{i=1}^{n} g(\varepsilon_i, x), \hat{G}_X(\varepsilon) = \frac{1}{n}\sum_{i=1}^{n} g(\varepsilon^*, X_i)$$

14

Semiparametric推定量

- RAL推定量の影響関数は、局外接空間に直交する
- $G_\varepsilon(x), G_X(\varepsilon^*), E\{g(\varepsilon, x)\}$が既知であれば、自然な推定量は以下の推定方程式から得られる

$$\sum_{i=1}^{n} [g\{\varepsilon_i(\beta), X_i\} - G_\varepsilon(X_i) - G_X\{\varepsilon_i(\beta)\} + E\{g(\varepsilon, X)\}] = 0^{q\times 1}$$

$$\text{where } \varepsilon_i(\beta) = Y_i - \mu(X_i, \beta).$$

15

Semiparametric推定量

- でも実際は、$G_\varepsilon(x), G_X(\varepsilon^*), E\{g(\varepsilon, x)\}$は既知ではない
- 未知量を推定量に置き換え

$$\sum_{i=1}^{n} \left(g\{\varepsilon_i(\beta), X_i\} - \hat{G}_{\varepsilon(\beta)}(X_i) - \hat{G}_X\{\varepsilon_i(\beta)\} + \hat{E}[g\{\varepsilon(\beta), X\}]\right) = 0^{q\times 1}$$

  - ただし、$\hat{E}[g\{\varepsilon(\beta), X\}] = n^{-1}\sum_{i=1}^{n} g\{\varepsilon_i(\beta), X_i\}$
- 一致性・漸近正規性を持つセミパラメトリック推定量が得られる

16

スコア関数（定理 5.3）

- (Y, X) の同時分布の分解

$$p_{Y,X}(y, x, \beta) = p_\varepsilon\{y - \mu(x, \beta)\}p_X(x).$$

- β についてのスコア関数

$$S_\beta^{q\times 1}(y, x) = \left.\frac{\partial \log p_{Y,X}(y, x, \beta)}{\partial \beta}\right|_{\beta=\beta_0} = -D^{T q\times 1}(x, \beta_0)S_\varepsilon(\varepsilon),$$

$$D(x, \beta_0) = \frac{\partial \mu(x, \beta_0)}{\partial \beta^T}, \quad S_\varepsilon(\varepsilon) = \frac{\partial \log p_\varepsilon(\varepsilon)}{d\varepsilon}.$$

- $E\{S_\varepsilon(\varepsilon)\} = 0$（定理 5.3）

17

有効スコア関数

- $E\{S_\varepsilon(\varepsilon)\} = 0$であるので、

$$S_\beta(\varepsilon, X) - \Pi[S_\beta|\Lambda] = -D^T(X, \beta_0)S_\varepsilon(\varepsilon) + E\{D^T(X, \beta_0)\}S_\varepsilon(\varepsilon) = -\{D^T(X, \beta_0) - E\{D^T(X, \beta_0)\}\}S_\varepsilon(\varepsilon).$$

  - Slide 16の影響関数の表現から、第2, 4項がキャンセル

18

局所有効推定量

- $p_\varepsilon(\varepsilon), p_X(x)$ が既知ならば、$E\{D^T(X,\beta)\}$, $S_\varepsilon(\varepsilon)$も既知なので、局所有効な推定量は以下の推定方程式から得られる

$$\sum_{i=1}^{n}[D^T(X_i,\beta) - E\{D^T(X,\beta)\}]S_\varepsilon\{Y_i - \mu(X_i,\beta)\} = 0.$$

19

局所有効推定量

- でも実際は、$E\{D^T(X,\beta)\}$は未知なので、Empiricalな推定量に置き換える

$$\sum_{i=1}^{n}\{D^T(X_i,\beta) - \bar{D}^T(\beta)\}S_\varepsilon\{Y_i - \mu(X_i,\beta)\} = 0,$$

$$\text{where } \bar{D}^T(\beta) = n^{-1}\sum_{i=1}^{n} D^T(X_i,\beta).$$

20

$p_\varepsilon(\varepsilon)$のWorkingモデル

- $S_\varepsilon(\varepsilon)$は未知の局外要因 $p_\varepsilon(\varepsilon)$で決まる
- $p_\varepsilon(\varepsilon)$に対して、適当なWorkingモデルを仮定
  - 必ずしも正しいモデルでなくてよい
- モデル誤特定の場合、先の推定方程式から得られる推定量 $\hat{\beta}_n$の漸近的な性質は？
  - $S_\varepsilon(\varepsilon)$を任意の関数 $\kappa(.)$ に置き換え

21

定理 5.4

- 任意の $\kappa(.)$ に対して、推定方程式

$$\sum_{i=1}^{n}\{D^T(X_i,\beta) - \bar{D}^T(\beta)\}\kappa\{Y_i - \mu(X_i,\beta)\} = 0,$$

の解 $\hat{\beta}_n$は漸近正規性を持ち、

$$n^{1/2}(\hat{\beta}_n - \beta_0) \xrightarrow{\mathcal{D}} N(0,\Sigma),$$

$$\Sigma = E\left\{\frac{\partial\kappa(\varepsilon)}{\partial\varepsilon}\right\}^{-2}\mathrm{var}\{\kappa(\varepsilon)\}\left[\mathrm{var}\left\{D^T(X,\beta_0)\right\}\right]^{-1}.$$

である

22

略 証（定理 5.4）

- （頑張って）漸近展開をして、影響関数を求めると、

$$\left[-E\left\{\frac{d\kappa(\varepsilon)}{\partial\varepsilon}\right\}\mathrm{var}\left\{D^T(X,\beta_0)\right\}\right]^{-1}\underbrace{\left\{D^T(X_i,\beta_0) - \mu_D\right\}}_{O-E}\underbrace{\{\kappa(\varepsilon_i) - \mu_\kappa\}}_{O-E}.$$

となる

- $n^{\frac{1}{2}}(\hat{\beta}_n - \beta_0)$ は、平均ゼロの正規分布に収束し，漸近共分散行列は、影響関数の共分散行列に一致する

23

Remark 1

- 任意の $\kappa(.)$ に対して $g(\varepsilon,X) = D^T(X,\beta_0)\kappa(\varepsilon)$ についてのセミパラ推定量の影響関数は

$$D^T(X,\beta_0)\kappa(\varepsilon) - D^T(X,\beta_0)\mu_\kappa - \mu_D\kappa(\varepsilon) + \mu_D\mu_\kappa$$
$$= \left\{D^T(X,\beta_0) - \mu_D\right\}\{\kappa(\varepsilon) - \mu_\kappa\}$$

- 前頁の影響関数とProportional
- $p_\varepsilon(\varepsilon)$の正しいモデルを選んでいれば、有効推定量が得られる
  - $\hat{\beta}_n$は局所有効なセミパラ推定量といえる

24

Globally Efficientな推定量

- スコア関数 $\partial \log p_\varepsilon(\varepsilon)/\partial\varepsilon$ をノンパラメトリックに推定し、$\kappa(\varepsilon)$ を置き換え

  Bickel et al. (1993)
- が、このノンパラ推定量は、当然ながらサンプルサイズがかなり大きくないと、不安定である

25

Another Strategy

- $p_\varepsilon(\varepsilon)$に適当なパラメトリックモデルを仮定
  - $p_\varepsilon(\varepsilon, \xi)$; e.g. 正規分布などで考えるとよい
- ξ を擬似尤度法で推定

$$\prod_{i=1}^{n} p_\varepsilon\{Y_i - \mu(X_i, \hat{\beta}_n^I), \xi\},$$

- 得られた推定値 $\hat{\xi}_n$ をPlug-inしたEstimated Score $S_\varepsilon(\varepsilon, \hat{\xi}_n) = \partial \log p_\varepsilon(\varepsilon, \hat{\xi}_n)/\partial\varepsilon$ を推定方程式(Slide 23)に代入

26

Adaptive Local Efficiency

- このAdaptiveな推定量は、$p_\varepsilon(\varepsilon)$のモデルが真の分布を含む場合、局所有効となる
- 真の分布を含まない場合にも、一致性・漸近正規性を持つセミパラメトリック推定量になる
- 裏を返せば、真のモデルをうまく近似できるような柔軟なパラメトリックモデルを用いれば、$p_\varepsilon(\varepsilon)$を正しく推定はできなくても、$\beta$ については良好なEfficiencyの性質を持つ一致推定量が得られる

27

Example 1: 線形モデル

$$Y_i = X_i^T \beta + \varepsilon_i,$$

- $D^T(X_i, \beta) = X_i^{q\times 1}$ なので、推定方程式は

$$\sum_{i=1}^{n} (X_i - \bar{X}) S_\varepsilon(Y_i - X_i^T \beta) = 0.$$

となる

28

$p_\varepsilon(\varepsilon)$のWorkingモデル

- Workingモデルとして、$\varepsilon_i \sim N(\mu_\varepsilon, \sigma^2)$ とすると

$$S_\varepsilon(\varepsilon) = -\left\{\frac{\varepsilon - \mu_\varepsilon}{\sigma^2}\right\}.$$

- 先の推定方程式にPlug-inすると、

$$\sum_{i=1}^{n} (X_i - \bar{X}) \frac{(Y_i - X_i^T \beta - \mu_\varepsilon)}{\sigma^2} = 0.$$

- 整理すると、

$$\sum_{i=1}^{n} (X_1 - \bar{X})(Y_i - X_i^T \beta) = 0.$$

最小二乗法の推定関数に一致！

29

最小二乗推定量の解釈

$$\hat{\beta}_n = \left\{\sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})^T\right\}^{-1} \sum_{i=1}^{n} (X_i - \bar{X}) Y_i.$$

- $p_\varepsilon(\varepsilon)$ の真のモデルが正規分布である場合、Adaptive locally efficientなSemiparametric推定量である
- 真のモデルが正規分布でない場合も、$\hat{\beta}_n$は一致性を持つSemiparametric推定量に
  - 別な $p_\varepsilon(\varepsilon)$ を用いると、Efficiencyは改善する可能性あり

30

制約モーメントモデル

$$Y_i = \alpha + \beta_1 X_{i1} + \cdots + \beta_q X_{iq} + \varepsilon_i,$$

- 分散関数：

$$V(X_i) = \text{var}\,(Y_i|X_i) = \sigma^2 \text{ (a constant independent of } X_i).$$

- βの有効推定量の推定方程式

$$\sum_{i=1}^{n} D^T(X_i)V^{-1}(X_i)\{Y_i - \mu(X_i, \beta)\} = 0.$$

- 整理すると

$$\sigma^{-2}\sum_{i=1}^{n}(1, X_i^T)^T(Y_i - \alpha - X_i^T\beta) = 0.$$

31

セミパラ有効な推定量

- 少し頑張って、計算すると

$$\hat{\beta}_n = \left\{\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})^T\right\}^{-1}\sum_{i=1}^{n}(X_i - \bar{X})Y_i,$$

- やっぱり最小二乗推定量に
  - Location-shiftモデルの局所有効な推定量と一致
- 分散関数が定数である制約モーメントモデルの β の推定量の中で、漸近有効な推定量である

32

Contexts

- 5.1 Location-Shift Regression Model
- 5.2 Proportional Hazards Regression Model with Censored Data
- 5.3 Estimating the Mean in a Nonparametric Model
- 5.4 Estimating Treatment Difference in a Randomized Pretest-Posttest Study or with Covariate Adjustment
- 5.5 Remarks about Auxiliary Variables

33

Time to Event Data

- がん，循環器疾患の臨床試験など
  - Primary Endpointは、死亡や再発，心血管イベントまでの時間など
- ほとんどすべての臨床試験（観察研究も）で、全員にイベントを観測することはできない
  - Right Censoring（右側打ち切り）
  - Censoringを考慮した評価方法が必要

34

Notations

- T: Time to Event (Positive, Continuous)
- C: Time to Event or Censoring
- $p_{C|X}(c|x)$: X所与のもとでのCの条件付き分布

$$V = \min(T, C) : \text{(time on study)},$$
$$\Delta = I(T \le C) : \text{(failure indicator)},$$
$$\text{and } X = q\text{-dimensional covariate vector.}$$

$$Z_i = (V_i, \Delta_i, X_i),\ i = 1, \ldots, n. \quad \text{i.i.d.}$$

35

Coxの比例ハザードモデル

$$\lambda(u|X) = \lim_{h \to 0}\left\{\frac{P(u \le T < u + h | T \ge u, X)}{h}\right\}$$
$$= \lambda(u)\exp(\beta^T X),$$

- $\lambda(u)$: Baseline hazard ft. ⇐ 局外要因
  - 時点ごとのハザードの大きさを規定
- β: 関心のあるパラメータ（対数ハザード比）
  - 共変量のハザードへの影響の大きさを規定

Cox (1972)

36

Assumptions

- TとCは、X所与のもとで条件つき独立
  - Censoringがあるもとで、X所与のもとでのTの条件付き分布の識別可能性のために必要な仮定
- ただし、$p_{C|X}(c|x)$ については付加的な仮定は置かない（局外要因であるため）

37

ここでの目標

- Observed Data: $(V_i, \Delta_i, X_i),\ i = 1, \ldots, n.$
- 局外要因である $\lambda(u), p_{C|X}(c|x), p_X(x)$ に付加的な仮定を置くことなく、
- βについてのセミパラメトリックな推定量を求めること
  - 一致性，漸近正規性，有効性を持つ

38

(V, Δ, X)の密度関数

$$p_{V,\Delta,X}(v,\delta,x) = \left\{\lambda(v)\exp(\beta^T x)\right\}^{\delta} \exp\left\{-\Lambda(v)\exp(\beta^T x)\right\}$$
$$\left\{p_{C|X}(v|x)\right\}^{1-\delta} \left\{\int_v^{\infty} p_{C|X}(u|x)du\right\}^{\delta} p_X(x),$$

- where

$$\Lambda(v) = \int_0^v \lambda(u)du,$$

  - 累積ベースラインハザード関数

39

$C|X$の分布の表現

- $C|X$もハザード形式などで表現

$$\lambda_{C|X}(u|x) = \lim_{h\to 0}\left\{\frac{P(u \le C < u+h | C \ge u, X = x)}{h}\right\}$$

$$\Lambda_{C|X}(v|x) = \int_0^v \lambda_{C|X}(u|x)du.$$

$$p_{C|X}(v|x) = \lambda_{C|X}(v|x)\exp\{-\Lambda_{C|X}(v|x)\}$$

$$\int_v^{\infty} p_{C|X}(u|x)du = \exp\{-\Lambda_{C|X}(v|x)\},$$

40

対数尤度関数

- (V, Δ, X)の密度関数、ふたたび

$$p_{V,\Delta,X}(v,\delta,x) = \left\{\lambda(v)\exp(\beta^T x)\right\}^{\delta} \exp\left\{-\Lambda(v)\exp(\beta^T x)\right\}$$
$$\times \left\{\lambda_{C|X}(v|x)\right\}^{1-\delta} \exp\left\{-\Lambda_{C|X}(v|x)\right\} \times p_X(x).$$

- 対数尤度関数

$$\delta\left\{\log\lambda(v) + \beta^T x\right\} - \Lambda(v)\exp(\beta^T x)$$
$$+ (1-\delta)\log\lambda_{C|X}(v|x) - \Lambda_{C|X}(v|x) + \log p_X(x).$$

- 局外パラメータ

$$\eta = \{\lambda(v), \lambda_{C|X}(v|x), p_X(x)\},$$

41

局外接空間

- $\mathcal{H}$: (V, Δ, X)の平均ゼロのq-次元可測関数 $h(v, \delta, x)$ から構成されるヒルベルト空間
- $\mathcal{H}$の中に局外接空間を構成するために、パラメトリックサブモデルの局外接空間の平均二乗閉包を考える
- パラメトリックサブモデル：$\lambda(v, \gamma_1)$, $\lambda_{C|X}(v|x, \gamma_2)$, $p_X(x, \gamma_3)$
  - $\gamma_1, \gamma_2, \gamma_3$: r_1, r_2, r_3次元の局外パラメータ

42

局外接空間の分解

- $\gamma_1, \gamma_2, \gamma_3$は独立であり、対数尤度関数の中でも分離している
- パラメトリックサブモデルの局外接空間は、3つの直交する部分空間の直和となる

$$\Lambda = \Lambda_{1s} \oplus \Lambda_{2s} \oplus \Lambda_{3s},$$

$$\Lambda_{js} = \text{mean-square closure of } \{B^{q \times r_j} S_{\gamma_j}(V, \Delta, X) \text{ for all } B^{q \times r_j}\},$$

$$S_{\gamma_j}(v, \delta, x) = \partial \log p_{V,\Delta,X}(v, \delta, x, \gamma)/\partial \gamma_j$$

43

Λ_{3s} associated with $p_X(x)$

- Λ_{3s}は周辺分布 $p_X(x)$ についてのもの
- 4章の議論から、

$$\Lambda_{3s} = \left[\alpha^{q \times 1}(X) : E\left\{\alpha^{q \times 1}(X)\right\} = 0^{q \times 1}\right]$$

- $h(v, \delta, x) \in \mathcal{H}$の射影の表現

$$\Pi(h|\Lambda_{3s}) = E(h|X).$$

44

Notations

- $N_C(u) = I(V \leq u, \Delta = 0)$
  - Censoringに関するCounting Process
- $Y(u) = I(V \geq u)$
  - At risk process
- Martingale Increment

$$dM_C(u, x) = dN_C(u) - \lambda_{0C|X}(u|x)Y(u)du,$$

  - $\lambda_{0C|X}(u|x)$: 時点uにおけるCの条件付きハザード関数

45

Λ_{2s} associated with $\lambda_{C|X}(v|x)$

- Λ_{2s}は、以下の空間となる（補題 5.1）

$$\left\{\int \alpha^{q \times 1}(u, X) dM_C(u, X) \text{ for all functions } \alpha^{q \times 1}(u, x)\right\}.$$

  - $\alpha^{q \times 1}(u, x)$は、$(u, x)$についての任意の$q$-次元関数

46

Notations

- $N(u) = I(V \leq u, \Delta = 1)$
  - イベントに関する Counting Process
- $Y(u) = I(V \geq u)$
  - At risk process
- Martingale Increment

$$dM(u, X) = dN(u) - \lambda_0(u) \exp(\beta_0 X) Y(u) du$$

47

Λ_{1s} associated with $\lambda(v)$

- Λ_{1s}は、以下の空間となる（補題 5.2）

$$\Lambda_{1s} = \left\{\int \alpha^{q \times 1}(u) dM(u, X) \text{ for all } q\text{-dimensional functions } \alpha^{q \times 1}(u)\right\}$$

  - $\alpha^{q \times 1}(u)$は、u についての任意の q-次元関数

48

β のRAL推定量

$$\Lambda = \Lambda_{1s} \oplus \Lambda_{2s} \oplus \Lambda_{3s},$$

- Λ は3つの直交する部分空間の直和
- β のRAL推定量の影響関数は、Λ の直交補空間に存在する

49

直交補空間（定理 5.5）

- 局外接空間の直交補空間は、以下で与えられる

$$\Lambda^{\perp} = \left[\int \{\alpha(u, X) - a^*(u)\} \, dM(u, X) \text{ for all } \alpha^{q \times 1}(u, x) \right],$$

$$a^*(u) = \frac{E\left\{\alpha(u, X) \exp(\beta_0^T X) Y(u)\right\}}{E\left\{\exp(\beta_0^T X) Y(u)\right\}}.$$

50

$a^*(u)$の推定量

- $a^*(u)$の推定量（分子・分母の標本平均をPlug-in）

$$\hat{a}^*(u) = \frac{n^{-1} \sum_{i=1}^{n} \alpha(u, X_i) \exp(\beta_0^T X_i) Y_i(u)}{n^{-1} \sum_{i=1}^{n} \exp(\beta_0^T X_i) Y_i(u)}.$$

- Lenglartの不等式より、

$$n^{-1/2} \sum_{i=1}^{n} \int \{\alpha(u, X_i) - a^*(u)\} \, dM_i(u, X_i)$$
$$= n^{-1/2} \sum_{i=1}^{n} \int \{\alpha(u, X_i) - \hat{a}^*(u)\} \, dM_i(u, X_i) + o_p(1).$$

51

もうちょっと整理して

$$\sum_{i=1}^{n} \int \left\{ \alpha(u, X_i) - \frac{\sum_{j=1}^{n} \alpha(u, X_j) \exp(\beta_0^T X_j) Y_j(u)}{\sum_{j=1}^{n} \exp(\beta_0^T X_j) Y_j(u)} \right\} \underbrace{dM_i(u, X_i)}_{\substack{\| \\ dN_i(u) - \lambda_0(u) \exp(\beta_0^T X_i) Y_i(u)}}$$

$$= \sum_{i=1}^{n} \int \left\{ \alpha(u, X_i) - \frac{\sum_{j=1}^{n} \alpha(u, X_j) \exp(\beta_0^T X_j) Y_j(u)}{\sum_{j=1}^{n} \exp(\beta_0^T X_j) Y_j(u).} \right\} dN_i(u).$$

52

影響関数

- 真値の周りで推定方程式を展開することで、

$$\sum_{i=1}^{n} \int \left\{ \alpha(u, X_i) - \frac{\sum \alpha(u, X_j) \exp(\beta^T X_j) Y_j(u)}{\sum \exp(\beta^T X_j) Y_j(u)} \right\} dN_i(u) = 0,$$

の解となる推定量は、$\Lambda^{\perp}$の元である

$$\int \{\alpha(u, X_i) - a^*(u)\} \, dM_i(u, X_i).$$

にProportionalな影響関数を持つ

53

β のセミパラメトリック推定量

- 任意の $\alpha(u, x)$ に対して、推定方程式

$$\sum_{i=1}^{n} \Delta_i \left\{ \alpha(V_i, X_i) - \frac{\sum_{j=1}^{n} \alpha(V_i, X_j) \exp(\beta^T X_j) Y_j(V_i)}{\sum_{j=1}^{n} \exp(\beta^T X_j) Y_j(V_i)} \right\} = 0.$$

の解となる推定量のクラス

  - β のRAL推定量のすべての影響関数を含むセミパラメトリック推定量のクラスとなる

54

有効スコア

- 対数尤度を β について微分

$$S_\beta = \int X^{q\times 1} dM(u,X).$$

- Λ 上へ射影したあとの残差として、有効スコアを求めると

$$S_{\rm eff} = \int \left\{ \underline{X} - \frac{E\{\underline{X}\exp(\beta_0^T X)Y(u)\}}{E\{\exp(\beta_0^T X)Y(u)\}} \right\} dM(u,X)$$

$\alpha(u,X) = X$のとき！

55

β の有効推定量

- $\alpha(u,X) = X$のときの推定方程式を考えてやればよい

$$\sum_{i=1}^{n} \Delta_i \left[X_i - \left\{ \frac{\sum_{j=1}^{n} X_j \exp(\beta X_j) Y_j(V_i)}{\sum_{j=1}^{n} \exp(\beta X_j) Y_j(V_i)} \right\} \right] = 0.$$

- 結局、Cox (1972, 1975) の部分尤度に
- Coxの最大部分尤度推定量は、比例ハザードモデルのもとで、Globally Efficientなセミパラメトリック推定量である

56

Contexts

- 5.1 Location-Shift Regression Model
- 5.2 Proportional Hazards Regression Model with Censored Data
- 5.3 Estimating the Mean in a Nonparametric Model
- 5.4 Estimating Treatment Difference in a Randomized Pretest-Posttest Study or with Covariate Adjustment
- 5.5 Remarks about Auxiliary Variables

57

Nonparametric Estimation

- $Z_1, Z_2, \dots, Z_n \sim p(z)$
  - パラメトリックな分布型は仮定しない
  - モーメントが存在する（可積分）であることだけは仮定
- $\mu = E(Z)$のセミパラメトリック（ノンパラメトリック）なRAL推定量のクラスを求めたい

58

影響関数の空間

- 定理 4.4（p.68）から、接空間 $\mathfrak{J}$ はヒルベルト空間 $\mathcal{H}$ そのもの
  - 直交補空間 $\mathfrak{J}^{\perp}$ は $\mathcal{H}$ の原点に対応する零点のみの部分空間
- 影響関数の空間 $\phi(Z) + \mathfrak{J}^{\perp}$ は、高々1つの影響関数によって構成される空間に
- μ のRAL推定量の影響関数が存在するならば、有効影響関数であり、対応する推定量がセミパラメトリック有効な推定量である

59

セミパラ有効な推定量

- TrivialなRAL推定量
  - $\bar{Z} = n^{-1}\sum Z_i$
- 影響関数：$(Z - \mu_0)$
  - これが、唯一かつ有効な影響関数
- つまり、標本平均がノンパラメトリックモデルのセミパラメトリック有効な推定量に

60

Contexts

- 5.1 Location-Shift Regression Model
- 5.2 Proportional Hazards Regression Model with Censored Data
- 5.3 Estimating the Mean in a Nonparametric Model
- 5.4 Estimating Treatment Difference in a Randomized Pretest-Posttest Study or with Covariate Adjustment
- 5.5 Remarks about Auxiliary Variables

Randomized Clinical Trials

- 2つの治療のランダム化臨床試験
  - 参加者は、2つの治療のうち、いずれかに確率(δ, 1 – δ)でランダムに割り付けられる
- 治療効果の評価
  - 2つの群の平均的な反応の差を推定
- 治療後の標本平均の差
- 治療前後の変化値の標本平均の差

共変量情報の有効活用

- ランダム割り付け
  - 2群の対象者背景は、平均的には均等
  - 共変量調整をしなくても、交絡による偏りのない治療効果の評価は保証されている
- 実際には、Baseline共変量などの情報が
  - 治療効果の推定に、これらの情報を有効活用することで、精度(検出力)を上げることができないか？

Pretest-Posttest design

- A = (0, 1); Treatment Indicator
- Y_1: Pretest measurement
- Y_2: Posttest(一定期間後の)measurement
  - Posttest measurementにおける治療効果の推定
  - 前後値の変化量における治療効果を推定

治療効果の指標

- Posttestの平均の差

$$\beta = E(Y_2|A=1) - E(Y_2|A=0),$$

- Baselineからの変化量の平均の差

$$\beta = E(Y_2 - Y_1|A=1) - E(Y_2 - Y_1|A=0),$$

- 理論的には、$E(Y_1|A=1) = E(Y_1|A=0)$ なので、この2つの指標は等しい

Naïve Estimators

- データセット：$Z_i = (Y_{1i}, A_i, Y_{2i}), i = 1, \dots, n$
- よく使われる推定量

$$\hat\beta_n = \frac{\sum_{i=1}^n A_i Y_{2i}}{\sum_{i=1}^n A_i} - \frac{\sum_{i=1}^n (1-A_i) Y_{2i}}{\sum_{i=1}^n (1-A_i)},$$

$$\hat\beta_n = \frac{\sum A_i (Y_{2i} - Y_{1i})}{\sum A_i} - \frac{\sum (1-A_i)(Y_{2i} - Y_{1i})}{\sum (1-A_i)}.$$

- 分布型の仮定は置かないことが一般的
  - セミパラメトリック推定量に

ANCOVA

- 共分散分析で共変量調整

$$Y_{2i} = \eta_0 + \beta A_i + \eta_1 Y_{1i} + \varepsilon_i$$

  - 推定は、昔ながらの最小二乗法
  - 他にも、強い予後因子があれば入れるのが一般的（分散でトクをする）
- 一般線形モデルとして考えれば、これもセミパラメトリック推定量に

67

$\hat{\beta}_n$の影響関数

- $\hat{\beta}_n$(two-sample difference) の影響関数

$$n^{1/2}(\hat{\beta}_n - \beta_0) = n^{1/2}\left(\frac{\sum A_i Y_{2i}}{\sum A_i} - \mu_2^{(1)}\right) - n^{1/2}\left(\frac{\sum(1-A_i)Y_{2i}}{\sum(1-A_i)} - \mu_2^{(0)}\right)$$

$$\text{where } \beta_0 = E(Y_2|A=1) - E(Y_2|A=0) = \mu_2^{(1)} - \mu_2^{(0)}.$$

68

$\hat{\beta}_n$の影響関数

$$n^{-1/2}\left\{\frac{\sum A_i(Y_{2i} - \mu_2^{(1)})}{\underset{\substack{\Downarrow p \\ \delta}}{(\sum A_i/n)}}\right\} - n^{-1/2}\left\{\frac{\sum(1-A_i)(Y_{2i} - \mu_2^{(0)})}{\underset{\substack{\Downarrow p \\ 1-\delta}}{(\sum(1-A_i)/n)}}\right\}$$

$$= n^{-1/2}\sum\left\{\frac{A_i}{\delta}(Y_{2i} - \mu_2^{(1)}) - \frac{(1-A_i)}{(1-\delta)}(Y_{2i} - \mu_2^{(0)})\right\} + o_p(1).$$

なので、i 番目の対象者の影響関数は

$$\varphi(Z_i) = \frac{A_i}{\delta}(Y_{2i} - \mu_2^{(1)}) - \frac{(1-A_i)}{(1-\delta)}(Y_{2i} - \mu_2^{(0)}).$$

69

$\mathcal{H}$ の分割

- 密度関数の分解

$$p_{Y_1,A,Y_2}(y_1, a, y_2) = p_{Y_1}(y_1)p_{A|Y_1}(a|y_1)p_{Y_2|Y_1,A}(y_2|y_1, a).$$

- 定理 4.5（p.71）より、ヒルベルト空間 $\mathcal{H}$ は、以下のように分割できる

$$\mathcal{H} = \mathscr{T}_1 \oplus \mathscr{T}_2 \oplus \mathscr{T}_3,$$

$$\mathscr{T}_1 = \{\alpha_1(Y_1) : E\{\alpha_1(Y_1)\} = 0\},$$
$$\mathscr{T}_2 = \{\alpha_2(Y_1, A) : E\{\alpha_2(Y_1, A)|Y_1\} = 0\},$$
$$\mathscr{T}_3 = \{\alpha_2(Y_1, A, Y_2) : E\{\alpha_2(Y_1, A, Y_2)|Y_1, A\} = 0\}.$$

※ 互いに排反で、直交する

70

局外接空間の表現

- ランダム化しているので、Y_1とAは独立

$$p_{A|Y_1}(a|y_1) = P_A(a) = \delta^a(1-\delta)^{(1-a)};$$

  - 完全に既知かつ未知パラメータ含まない
- 一方で、$p_{Y_1}(y_1), p_{Y_2|Y_1,A}(y_2|y_1, a)$にはパラメトリックな仮定は置いていない
- つまり、

$$\mathscr{T} = \mathscr{T}_1 \oplus \mathscr{T}_3.$$

71

直交補空間

- $\mathcal{H}$の分割から、接空間の直交補空間は

$$\mathscr{T}^{\perp} = \mathscr{T}_2.$$

- 定理 4.5 から、$\mathfrak{I}_2$の元は以下のように表現できる

$$h_{*2}(Y_1, A) - E\{h_{*2}(Y_1, A)|Y_1\}$$

  - $h_{*2}(.)$: Y_1, Aの任意の関数

72

直交補空間

- Aは2値変数なので、以下の表現もできる

$$h_{*2}(Y_1, A) = Ah_1(Y_1) + h_2(Y_1)$$

  - h_1, h_2は、Y_1の任意の関数
- ということは、直交補空間の元は、

$$\begin{aligned} &h_*(Y_1, A) - E\{h_*(Y_1, A)|Y_1\} \\ &\quad = Ah_1(Y_1) + h_2(Y_1) - \{E(A|Y_1)h_1(Y_1) + h_2(Y_1)\} \\ &\quad = (A-\delta)h_1(Y_1). \end{aligned}$$

73

RAL推定量のクラス

- 影響関数の空間 $\{\varphi(Z) + \mathscr{T}^{\perp}\}$ の表現：

$$\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)}) - \frac{(1-A)}{(1-\delta)}(Y_2 - \mu_2^{(0)}) + (A-\delta)h_*(Y_1)\right\}$$

  - $h_*(Y_1)$: Y_1の任意の関数
- よって、セミパラメトリックなRAL推定量は、

$$\frac{\sum A_i Y_{2i}}{\sum A_i} - \frac{\sum (1-A_i) Y_{2i}}{\sum (1-A_i)} + n^{-1}\sum\left(A_i - n^{-1}\sum A_i\right)h_*(Y_{1i}).$$

74

有効な影響関数

- 有効な影響関数は

$$\varphi(Z) - \Pi\{\varphi(Z)|\mathscr{T}^{\perp}\},$$

- 今回のケースでは、

$$\begin{aligned} &\frac{A}{\delta}(Y_2 - \mu_2^{(1)}) - \Pi\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|\mathscr{T}^{\perp}\right\} \\ &\quad - \left(\frac{1-A}{1-\delta}\right)(Y_2 - \mu_2^{(0)}) + \Pi\left\{\left(\frac{1-A}{1-\delta}\right)(Y_2 - \mu_2^{(0)})|\mathscr{T}^{\perp}\right\}. \end{aligned}$$

75

有効な影響関数

- $\mathscr{T}^{\perp} = \mathscr{T}_2$. なので、

$$\Pi\{\alpha(\cdot)|\mathscr{T}^{\perp}\} = \Pi\{\alpha(\cdot)|\mathscr{T}_2\} = E\{\alpha(\cdot)|Y_1, A\} - E\{\alpha(\cdot)|Y_1\}.$$

- したがって、

$$\Pi\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|\mathscr{T}^{\perp}\right\} = E\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|Y_1, A\right\} - E\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|Y_1\right\},$$

- また、

$$E\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|Y_1, A\right\} = \frac{A}{\delta}\left\{E(Y_2|Y_1, A=1) - \mu_2^{(1)}\right\},$$

$$E\left\{\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|Y_1\right\} = E(Y_2|Y_1, A=1) - \mu_2^{(1)}.$$

76

有効な影響関数

- ということで、

$$\Pi\left(\frac{A}{\delta}(Y_2 - \mu_2^{(1)})|\mathscr{T}^{\perp}\right) = \frac{A-\delta}{\delta}\left\{E(Y_2|Y_1, A=1) - \mu_2^{(1)}\right\}.$$

- 同様にして、

$$\Pi\left(\frac{1-A}{1-\delta}(Y_2 - \mu_2^{(0)})|\mathscr{T}^{\perp}\right) = -\frac{A-\delta}{1-\delta}\left\{E(Y_2|Y_1, A=0) - \mu_2^{(0)}\right\}.$$

77

有効な影響関数

- 以上のことから、

$$\begin{aligned} &\left\{\frac{A}{\delta}Y_2 - \frac{(A-\delta)}{\delta}E(Y_2|A=1, Y_1)\right\} \\ &\quad - \left\{\frac{(1-A)}{(1-\delta)}Y_2 + \frac{(A-\delta)}{(1-\delta)}E(Y_2|A=0, Y_1)\right\} \\ &\quad - \underbrace{\left(\mu_2^{(1)} - \mu_2^{(0)}\right)}_{\|\ \beta}. \end{aligned}$$

78

未知量をどう扱う？

- 有効な影響関数を求める上で、以下の条件付き平均が未知である
  - $E(Y_2|A = 1, Y_1), E(Y_2|A = 0, Y_1)$
- 有限個のパラメータ ξ_1, ξ_0 でモデル化

$$E(Y_2|A = j, Y_1) = \zeta_j(Y_1, \xi_j), \;\; j = 0, 1.$$

- ξ_1, ξ_0 は、制約モーメントモデル（一般化線形モデル）として、推定してやればよい
  - $\hat{\xi}_1, \hat{\xi}_0$: セミパラメトリック推定量

79

有効なセミパラメトリック推定量

- Plug-inして得られる有効な推定量は、

$$\hat{\beta}_n = n^{-1} \sum_{i=1}^{n} \left[\left\{ \frac{A_i}{\delta} Y_{2i} - \frac{(A_i - \delta)}{\delta} \zeta_1(Y_{1i}, \hat{\xi}_{1n}) \right\} - \left\{ \frac{(1 - A_i)}{(1 - \delta)} Y_{2i} + \frac{(A_i - \delta)}{(1 - \delta)} \zeta_0(Y_{1i}, \hat{\xi}_{0n}) \right\} \right].$$

- $E(Y_2|A = 1, Y_1), E(Y_2|A = 0, Y_1)$のモデルが正しければ、Efficiency Boundを達成

80

Plug-inした影響関数

$$\frac{A}{\delta}(Y_2 - \mu_2^{(1)}) - \frac{(A - \delta)}{\delta}\{\zeta_1(Y_1, \xi_1^*) - \mu_2^{(1)}\} - \frac{(1 - A)}{(1 - \delta)}(Y_2 - \mu_2^{(0)}) - \frac{(A - \delta)}{(1 - \delta)}\{\zeta_0(Y_1, \xi_0^*) - \mu_2^{(0)}\}.$$

- $E(Y_2|A = 1, Y_1), E(Y_2|A = 0, Y_1)$のモデルは誤っていても、RAL推定量の影響関数のクラスに含まれる
  - 誤特定のもとでも、一致性・漸近正規性はOKなセミパラメトリック推定量に

81

まとめ

- セミパラメトリック推定量：$\hat{\beta}_n$
- $E(Y_2|A = 1, Y_1), E(Y_2|A = 0, Y_1)$のモデルを正しく与えられていれば、有効な推定量に
- $E(Y_2|A = 1, Y_1), E(Y_2|A = 0, Y_1)$のモデルを誤っていても、一致性・漸近正規性を持つセミパラメトリック推定量に（局所有効）

82

一般の共変量調整の問題

- Y_{1i}をBaseline共変量のベクトルに変更
- 割り付け変数 A とは独立なので、分布型に仮定を与えることなく、ここまでと同じ論理で、有効な影響関数を求めることができる
- βの局所有効なAdaptive semiparametric推定量は、まったく同じ型で与えられる

83

Contexts

- 5.1 Location-Shift Regression Model
- 5.2 Proportional Hazards Regression Model with Censored Data
- 5.3 Estimating the Mean in a Nonparametric Model
- 5.4 Estimating Treatment Difference in a Randomized Pretest-Posttest Study or with Covariate Adjustment
- 5.5 Remarks about Auxiliary Variables

84

Auxiliary Variablesの有用性

- Randomized Pretest-Posttestの事例では、補助変数 W の情報を利用することによって、β の推定精度を改善することができた
- これは、より一般的に成り立つ結果？
- どのような条件のもとで、W による精度の改善は図れるのか？

85

W の情報を使う？使わない？

- $\mathcal{H}^Z$: Zについての平均0の二乗可積分な q次元関数の空間
- $\mathcal{H}^{ZW}$: (Z, W)についての平均0の二乗可積分な q次元関数の空間
  - Z の周辺分布のみモデル化し（βのモデル）、(W, Z) については付加的な仮定は置かない

86

β のRAL推定量の影響関数

- Wの情報を使わない場合の影響関数の空間
  - $\phi(Z) + \mathfrak{I}^{Z\perp}$
  - $\mathfrak{I}^{Z\perp}$: $\mathcal{H}^Z$の接空間に直交する空間
- Wの情報を使う場合の影響関数の空間
  - $\phi(Z) + \mathfrak{I}^{WZ\perp}$
  - $\mathfrak{I}^{WZ\perp}$: $\mathcal{H}^{WZ}$の接空間に直交する空間

87

定理 5.6

- セミパラメトリックモデル $p_Z(z, \beta, \eta)$ から、Zの周辺分布モデルが仮定されているとする
- 条件付き分布 $p_{W|Z}(w|z)$に付加的な仮定を置かない場合、(W, Z) の同時分布におけるセミパラメトリックモデルの接空間 $\mathfrak{I}^{WZ}$の直交補空間 $\mathfrak{I}^{WZ\perp}$ は、Z の周辺分布のみをモデル化したときの接空間の直交補空間 $\mathfrak{I}^{Z\perp}$ に一致する

88

定理 5.6 の解釈

- $\mathfrak{I}^{WZ\perp} = \mathfrak{I}^{Z\perp}$ということは
  - β のRAL推定量の影響関数の空間はまったく同じ
  - 一般的には、W を考慮することによる推定精度の改善はまったくない
- W の情報を使いたくても、数学的なモデルの仮定を付与しないのであれば、β の推定精度で得することはない　⇔ 直感と整合

89

ランダム化比較試験では

- Randomized Pretest-Posttestの設定ではなぜ精度が改善された？
- ランダム化のため、治療の指示変数 A と Y_1 が独立だった
- デザインによって与えられた付加的な仮定によって、$\mathfrak{I}^{WZ\perp} \neq \mathfrak{I}^{Z\perp}$（前節参照）
  - ランダム化比較試験では、Wを考慮することによって、精度の改善を図ることができる

90