

限界介入効果に関する諸考察

津 田 俊 樹*

I. はじめに

1990 年代以降，ミクロ計量経済学分野のなかで，データのみから識別できる測定対象を中心とした分析が主流となってきた（市村 [2010]）。その代表例が Local Average Treatment Effect (LATE) や Matching による推定などである。

LATE は Imbens and Angrist [1994] で提案された介入効果である。従来の操作変数の識別に課される仮定である外生性および関連性に加えて操作変数の変化によって人々が受ける処置の変化の方向が同じであることを仮定した単調性 (Monotonicity) を課すことによって，操作変数の変化によって処置が変わるような“従順な人々” (Complier) の平均的介入効果を識別することを可能にし，異質性のある介入効果の識別へと門戸を開いた (Cornelissen et al [2016])。また，Matching による推定とは共変量や傾向スコアの類似性でもって，ある結果変数に対応する反実仮想的な結果変数を抽出し，平均介入効果や処置を受けた人たちの平均介入効果といった伝統的な介入効果の推定を可能にしている。

近年では，個人の異質性に基づいた因果効果の分析が盛んであり，LATE の拡張として Marginal Treatment Effect (MTE) と呼ばれる介入効果が開発されている。MTE とは Generalized Roy Model によって処置行動を定式化した際の処置の決定に関する観測不可能な変数に基づく平均的介入効果である。この推定量は処置を決定する観測可能な変数 Z で条件づけた処置を受ける確率，傾向スコア $\Pr(D=1|Z)$ (以下， $P(Z)$ とする) の限界的な変動を利用した，Local Instrumental Variable (LIV) という手法で識別することが可能となる (Heckman and Vytlacil [1999])。

先に述べた LATE では操作変数の値によって処置が変動するような Complier の平均的な介入効果に着目しているのに対し，MTE では観測不可能な変数に基づいて平均的な介入効果を得ることができるという有用性がある。例えば，Brinch, Mogstad and Wiswall [2017] では家庭における子供の数が子供の教育年数に与える影響について，性別が同じ兄弟の存在の有無や双子出産であったかどうかなどのダミー変数を操作変数として用いた LATE による分析に一定の評価をしながらも，LATE の限界としてその効果の個人間での変動が大きいのではないかと懸念が生じることを主張し，MTE を使った分析を用いることで個人間により介入効果が大きく異なり，負にもなりうる結果を示している。また LATE では先ほど述べた通り，平均的介入効果といった経済政策の決定において示唆のある伝統的な介入効果を推定できない場合が存在するのに対し，傾向スコア $P(Z)$ が 0 から 1 の範囲で存在するという仮定が満たされる場合，MTE を使用することでそれらの

* 東京大学大学院経済学研究科修士課程

伝統的な介入効果を MTE の加重平均として推定することが可能である。また、ある一定の条件下では、まだ実行されていない政策との平均的な介入効果である Policy Relevant Treatment Effect (PRTE) を推定することまでもが可能となる。(Heckman [2010])

MTE の推定手法としては、パラメトリックな仮定を用いて推定をした Heckman, Tobias and Vytlacil [2003] や Akavik, Heckman and Vytlacil [2005] などがあり、またセミパラメトリックな推定手法は Heckman, Urzúa and Vytlacil [2006] の Appendix B や Carneilo and Lee [2009] などで提案されている。とりわけ、Heckman, Urzúa and Vytlacil [2006] の推定手法は、Carneilo, Heckman and Vytlacil [2011] や Brinch, Mogstad and Wiswall [2017] などの実証研究で扱われている。本論文では、Heckman, Urzúa and Vytlacil [2006] と Carneilo and Lee [2009] の手法を参考としたセミパラメトリック推定手法を構築し、その漸近正規性を証明した。

本論文の構成は以下の通りである。2 節で MTE の定式化および LIV で MTE が識別できることを説明し、3 節では具体的に自身が考案した推定方法を紹介する。4 節で漸近正規性を満たすために必要な仮定を挙げ、漸近正規性が成り立つことを示す。5 節ではモンテカルロシミュレーションを用いて自身の提案した手法、Carneilo and Lee [2009] の手法を比較した結果を紹介する。6 節はこの論文をまとめている。なお、4 節の漸近正規性を示した定理の証明は付録に詳述する。

II. MTE の定式化

II-1. MTE の定義

まず、各変数を以下のように定義する。

Y_i : 結果変数, D_i : 介入を決定する二値の内生変数, $D_i=1$ は個人 i が介入を受けたことを意味する。 X_i : 共変量, Z_i : 操作変数で共変量 X_i を全て含む。ただし、 $D \in \{0, 1\}$, $Y \in \mathbb{R}$, $X \in \mathcal{X} \subset \mathbb{R}^{dimX}$, $Z \in \mathcal{Z} \subset \mathbb{R}^{dimZ}$ また、サンプルサイズは n で、これらの変数はランダムに抽出されたものであるとする。

結果変数 Y の式を、反実仮想的な結果変数 $Y(0)$, $Y(1)$ を用いて以下のように定義する。

$$Y \equiv DY(1) + (1-D)Y(0) \quad (1)$$

ただし、 $Y(j)$ は、それぞれ $D=j$ の時の結果変数に対応する ($j=0$ または 1)。また、それぞれの反実仮想的な結果変数を以下のような式で定義する。

$$Y(j) \equiv \mu_j(X) + U_j \quad (2)$$

ただし、ここで $\mu_j(X) = E(Y(j)|X)$ で $\mu_j(X)$ は未知の関数とする。

介入ルールを Generalized Roy model を用いて定式化する。すなわち、

$$D \equiv 1\{\mu_D(Z) \geq U_D\}$$

とする。ただし、 $1\{\cdot\}$ は指示関数, U_D は連続確率変数, $\mu_D(\cdot)$ は Z に依存する未知の関数とする。この時、 U_D は各個人に依存する観測不可能な変数として解釈される。つまり、 U_D の値が低い個人は処置を受けやすい傾向があり、 U_D の値が高い人は処置を受けにくくなる。

ここで、 $F_{U_D|Z}(\cdot)$ を U_D の Z に関する条件付き累積確率分布とし、狭義単調増加関数であると仮定すると、

$$\mu_D(Z) \geq U_D \Leftrightarrow F_{U_D|Z}(\mu_D(Z)) \geq F_{U_D|Z}(U_D)$$

が成立する。新たに $P(Z) \equiv F_{U_D|Z}(\mu_D(Z))$, $V \equiv F_{U_D|Z}(U_D)$ と定義する。このとき、 Z で条件づけると $F_{U_D|Z}(U_D)$ は狭義単調増加関数であるから、逆関数が存在し微分可能、よって確率積分変換を用い

ると,

$$f_{V|Z}(v) = f_{U_D|Z}(F_{U_D}^{-1}(V)) \frac{1}{f_{U_D|Z}(F_{U_D}^{-1}(V))} = 1$$

となる (久保川 [2017])。

したがって、 Z で条件づけた V は区間 $(0, 1)$ 上の一様分布に従う。この結果を用いると、 Z と V が独立であれば,

$$E[D=1|Z] = \Pr(D=1|Z) = \Pr(P(Z) \geq V|Z) = P(Z)$$

より、 $P(Z)$ がデータから識別されることがわかる。

ここで、MTE の識別のために以下の仮定を課す。

仮定 1

(1-1) X に条件づけられたもとで $\mu_D(Z)$ は、退化分布を持たない。

(1-2) X に条件づけられたもとで、 Z は (U_D, U_0, U_1) と独立である。

(1-3) 確率変数 U_D が連続確率変数である。また X が条件づけられたもとで $\mu_D(Z)$ は連続確率変数となる。

(1-4) $E[Y(0)]$, $E[Y(1)]$ が有限、つまり存在する。

(1-5) $0 < \Pr(D=1|Z) < 1$

(1-6) $E[Y(1)|X=x, P(Z)=p, D=1]$ および $E[Y(0)|X=x, P(Z)=p, D=0]$ は p に関して連続微分可能。

(1-1) の仮定は、共変量 X として含まれない変数が少なくとも一つ Z の中に存在していることを示唆し、通常の操作変数の仮定である操作変数の関連性に対応する。また (1-2) は操作変数の外生性に対応する。

これらの仮定の下で MTE は次のように定義される。(Heckman and Vytlacil [2005])

$$\Delta_{MTE}(x, p) \equiv E[Y(1) - Y(0) | X=x, V=p]$$

ただし、 $x \in \mathcal{X}$, $p \in (0, 1)$ となる。これからわかるように MTE とは、観測されない異質性にに基づいた平均的な介入効果のことである。また、 $P(Z)$ が X の条件下の下 V と独立であることから,

$$\Delta_{MTE}(x, p) = E[Y(1) - Y(0) | X=x, V=p, P(Z)=p]$$

と書ける。直感的には、 X を条件づけたもとで $P(Z)$ があたかもランダムであるかのように振舞い、介入を受けるかどうかの決定を変化させることで、参加することと参加しないことが無差別であるような人々に対する限界的な介入効果をとらえることができる。

II-2. MTE の LIV による識別

LIV は局所操作変数と呼ばれるパラメーターで以下のように定義される。

$$\Delta_{LIV}(x, p) \equiv \frac{\partial E[Y | X=x, P(Z)=p]}{\partial p}$$

これは,

$$\begin{aligned} & \frac{\partial E[Y | X=x, P(Z)=p]}{\partial p} \\ &= \lim_{p' \rightarrow p} \frac{E[Y | X=x, P(Z)=p'] - E[Y | X=x, P(Z)=p]}{p' - p} \end{aligned}$$

であるので、 $E[Y|X=x, P(Z)=p]$ が微分可能であれば、LIV は $P(Z)$ を操作変数とした LATE の局所的な値として定義される。(Heckman and Vytlacil [1999])

以下で MTE が LIV によって識別されることを示す。

まず、

$$\begin{aligned} E[Y|X=x, P(Z)=p] &= E[DY(1) + (1-D)Y(0)|X=x, P(Z)=p] \\ &= E[DY(1)|X=x, P(Z)=p] + E[(1-D)Y(0)|X=x, P(Z)=p] \end{aligned}$$

となる。

$E[DY(1)|X=x, P(Z)=p]$ を考えると、繰り返し期待値の法則より

$$\begin{aligned} E[DY(1)|X=x, P(Z)=p] &= E[E[DY(1)|X=x, P(Z)=p, D]|X=x, P(Z)=p] \\ &= \Pr(D=1|X=x, P(Z)=p)E[Y(1)|X=x, P(Z)=p, D=1] \end{aligned}$$

となる。

$$\begin{aligned} &\Pr(D=1|X=x, P(Z)=p) \\ &= \Pr(V \leq P(Z)|X=x, P(Z)=p) \\ &= \Pr(V \leq p|X=x) = p \end{aligned}$$

より、

$$\begin{aligned} E[DY(1)|X=x, P(Z)=p] &= E[E[DY(1)|X=x, P(Z)=p, D]|X=x, P(Z)=p] \\ &= pE[Y(1)|X=x, P(Z)=p, D=1] \end{aligned}$$

となる。

同様にして

$$\begin{aligned} &E[(1-D)Y(0)|X=x, P(Z)=p] \\ &= E[E[(1-D)Y(0)|X=x, P(Z)=p, D]|X=x, P(Z)=p] \\ &= (1-p)E[Y(0)|X=x, P(Z)=p, D=0] \end{aligned}$$

となる。

したがって、

$$\begin{aligned} E[Y|X=x, P(Z)=p] &= \\ &pE[Y(1)|X=x, P(Z)=p, D=1] + (1-p)E[Y(0)|X=x, P(Z)=p, D=0] \end{aligned}$$

$Y(1)$, $Y(0)$ の定義や処置変数の定義を代入して整理すると

$$\begin{aligned} &pE[Y(1)|X=x, P(Z)=p, D=1] + (1-p)E[Y(0)|X=x, P(Z)=p, D=0] \\ &= \mu_0(x) + p(\mu_1(x) - \mu_0(x)) + pE[U_1|X=x, p \geq V] + (1-p)E[U_0|X=x, p < V] \end{aligned}$$

となる。

また、 $E[U_0|X=x]$ は定義より 0 なので繰り返し期待値の法則より、

$$\begin{aligned} 0 &= E[U_0|X=x, P(Z)=p] = E[E[U_0|X=x, D, P(Z)=p]|X=x, P(Z)=p] = \\ &pE[U_0|X=x, p \geq V] + (1-p)E[U_0|X=x, p < V] \end{aligned}$$

となる。よって

$$-pE[U_0|X=x, p \geq V] = (1-p)E[U_0|X=x, p < V]$$

であるから、

$$\begin{aligned} E[Y|X=x, P(Z)=p] &= \\ &\mu_0(x) + p(\mu_1(x) - \mu_0(x)) + pE[U_1|X=x, p \geq V] + (1-p)E[U_0|X=x, p < V] \\ &= \mu_0(x) + p(\mu_1(x) - \mu_0(x)) + p(E[U_1 - U_0|X=x, p \geq V]) \end{aligned}$$

となる。

この時,

$$K(P(Z)=v, X=x) \equiv v(E[U_1 - U_0 | X=x, v \geq V])$$

と定義する。

条件付き期待値の定義より,

$$E[U_1 | X=x, p \geq V] = \int_{\mathbb{R}} u_1 f(u_1 | X=x, p \geq V) du_1$$

また, 定義より

$$\begin{aligned} & f(u_1 | X=x, p \geq V) \\ &= \frac{\int_0^p f(u_1, v | X=x) dv}{\int_0^p f(v | X=x) dv} = \frac{\int_0^p f(u_1, v | X=x) dv}{\int_0^p f(v | Z=z, X=x) dv} \end{aligned}$$

二個目の等式は (1-2) による。

また, V は Z で条件づけると, 一様分布に従うので,

$$f(u_1 | X=x, p \geq V) = \frac{\int_0^p f(u_1, v | X=x) dv}{p}$$

したがって,

$$\begin{aligned} & E[U_1 | X=x, p \geq V] \\ &= \int_{\mathbb{R}} u_1 \frac{\int_0^p f(u_1, v | X=x) dv}{p} du_1 \\ &= \frac{1}{p} \int_0^p \int_{\mathbb{R}} u_1 f(u_1 | V=v, X=x) \cdot f(v | X=x) du_1 dv \\ &= \frac{1}{p} \int_0^p E[U_1 | V=v, X=x] dv \end{aligned}$$

が成り立つ。なお, 2 個目の等式では, $f(u_1, v | X=x) = f(u_1 | V=v, X=x) f(v | X=x)$ と Fubini の定理を用いた。また最後の等式では (1-2) を用いた。

同様にして,

$$E[U_0 | X=x, p \geq V] = \frac{1}{p} \int_0^p E[U_0 | V=v, X=x] dv$$

となるので,

$$K(p, x) = \int_0^p E[U_1 - U_0 | V=v, X=x] dv$$

となる。

したがって,

$$E[Y | X=x, P(Z)=p] = \mu_0(x) + p(\mu_1(x) - \mu_0(x)) + K(p, x) \quad (3)$$

(1-6) から $E[Y | X=x, P(Z)=p]$ が p で微分可能であるので,

$$\frac{\partial E[Y | X=x, P(Z)=p]}{\partial p} = \mu_1(x) - \mu_0(x) + \left. \frac{\partial K(v, x)}{\partial v} \right|_{v=p}$$

$$\begin{aligned}
 &= E[(\mu_1(X) + U_1) | V=p, X=x] - E[(\mu_0(X) + U_0) | V=p, X=x] \\
 &= E[Y(1) - Y(0) | X=x, V=p]
 \end{aligned} \tag{4}$$

となるので LIV によって MTE が識別できる。

Ⅲ. MTE の推定

まず, $\mu_j(X)$ が線形であるとする。すなわち

$$\mu_j(X) = X' \beta_j \quad (j=0, 1)$$

とおく。

また, 以下を仮定する。

仮定 2

(2-1) $\{(Y_i, X_i, Z_i, D_i)\}$ は独立同一分布に従う。

(2-2) $E[U_j | X, V] = E[U_j | V]$ となる。 ($j=0, 1$)

したがって, (2-2) のもとで

$$K(v, x) = \int_0^v E[U_1 - U_0 | V=t, X=x] dt = \int_0^v E[U_1 - U_0 | V=t] dt$$

となり, $K(v, x)$ は v にのみ依存する関数となる。以下, $K(v)$ とする。

以上の仮定を (4) に課すと MTE は

$$x'(\beta_1 - \beta_0) + \left. \frac{\partial K(v)}{\partial v} \right|_{v=p} = E[Y(1) - Y(0) | X=x, V=p]$$

となるので, MTE を推定するためにはそれぞれ $\beta_1, \beta_0, \left. \frac{\partial K(v)}{\partial v} \right|_{v=p}$ を推定する必要がある。

Carneilo and Lee [2009] の手法では, まず, 傾向スコア $\widehat{P(Z)}$ を三次 B-spline 法を用いて推定し, その情報をもとに, 部分線形推定量を

$$Y(1) = X' \beta_1 + U_1 \quad Y(0) = X' \beta_0 + U_0$$

に適用し, $\widehat{\beta}_1, \widehat{\beta}_0$ を得る。また,

$$E[U_1 | V=p] = E[U_1 | P(Z)=p, D=1] + p \left. \frac{\partial E[U_1 | P(Z)=v, D=1]}{\partial v} \right|_{v=p} \tag{5}$$

$$E[U_0 | V=p] = E[U_0 | P(Z)=p, D=0] - (1-p) \left. \frac{\partial E[U_0 | P(Z)=v, D=0]}{\partial v} \right|_{v=p} \tag{6}$$

であることを利用してそれぞれの項について局所線形多項式を使い, 先ほど推定した $\widehat{P(Z)}$ でもって推定し, $\widehat{E}[U_1 | V=p], \widehat{E}[U_0 | V=p]$ を得る。そして $\widehat{\beta}_1, \widehat{\beta}_0$ をつかって

$$\widehat{E}[Y(1) - Y(0) | X=x, V=p] = x'(\widehat{\beta}_1 - \widehat{\beta}_0) + \widehat{E}[U_1 | V=p] - \widehat{E}[U_0 | V=p]$$

で MTE を推定している。

一方, 本論文の手法では $\partial K(v) / \partial v|_{v=p}$ の推定に関して以下のようにする。

(3) と線形性の仮定により,

$$E[Y - (X' \beta_0 + P(Z) X' (\beta_1 - \beta_0)) | P(Z)=p] = K(p)$$

となる。よって, $Y - (X' \beta_0 + P(Z) X' (\beta_1 - \beta_0)) = \bar{Y}$ とし, 対応するデータを $\bar{Y}_i$ とすると

$$\tilde{Y}_i = K(P(Z_i)) + \varepsilon_i$$

となる。ただし $\varepsilon_i = \tilde{Y}_i - E[\tilde{Y}_i | P(Z) = P(Z_i)]$, $E[\varepsilon_i | P(Z)] = 0$ である。

このとき以下の最小化問題を考える

$$\min_{\alpha, \beta, \gamma} \sum_{i=1}^n (\tilde{Y}_i - \alpha - \beta(P(Z_i) - p) - \gamma(P(Z_i) - p)^2) K\left(\frac{P(Z_i) - p}{h}\right)$$

この時の最小解 $\hat{\beta}$ が $\partial K(v)/\partial v|_{v=p}$ の推定量である。

しかしながら, $\tilde{Y}_i$, $P(Z_i)$ は観測できない変数なので, 推定量 $\widehat{\tilde{Y}_i}$, $\widehat{P(Z_i)}$ に置き換え最適化問題

$$\min_{\alpha, \beta, \gamma} \sum_{i=1}^n (\widehat{\tilde{Y}_i} - \alpha - \beta(\widehat{P(Z_i)} - p) - \gamma(\widehat{P(Z_i)} - p)^2) K\left(\frac{\widehat{P(Z_i)} - p}{h}\right) \quad (7)$$

を解いて推定する。ただし $\widehat{\tilde{Y}_i} = Y_i - X_i \hat{\beta}_0 - \hat{P}_i(X_i(\hat{\beta}_1 - \hat{\beta}_0))$

よって MTE は (7) の最適解 $\partial \widehat{K(v)}/\partial v|_{v=p}$ と仮定 3 を満たす $\hat{\beta}_1, \hat{\beta}_0$ を用いて以下のように推定される。

$$\widehat{\text{MTE}}(p, x) = \widehat{E}[Y(1) - Y(0) | X = x, V = p] = x(\hat{\beta}_1 - \hat{\beta}_0) + \left. \frac{\partial \widehat{K(v)}}{\partial v} \right|_{v=p} \quad (8)$$

IV. MTE 推定量の漸近正規性

今までに課した仮定 1, 2 に加えて, いくつかの仮定を課することによって (8) の MTE 推定量が漸近正規性を持つことを示せる。

仮定 3

各 β_j の推定量, $\hat{\beta}_j$ に関して以下の関係式が成り立つ

$$\hat{\beta}_j - \beta_j = O_p(n^{-\frac{1}{2}})$$

Carneilo and Lee [2009] では, 自身らが提案した β_j の推定量に関していくつかの仮定の下で, 仮定 3 を満たすことが証明されている。

仮定 4

$K(v)$ が v に関して四回連続微分可能

仮定 5

カーネル関数が対称かつ二次のオーダーでコンパクトなサポートを持ち, リプシッツ連続。

仮定 6

$\widehat{P}_i \equiv \widehat{P(Z_i)}$ と置くと, 傾向スコアの推定量 $\widehat{P}_i$ は $\max_{i: 1 \leq i \leq n} |\widehat{P}_i - P_i| = o_p(h^2)$ を満たし, バンド幅 h は $h = O(n^{-\frac{1}{7}})$ を満たす。

仮定 7

局所多項式推定量 $\frac{\partial \widehat{E}(X_i | P(Z) = p)}{\partial P}$ が確率収束する。

仮定8

$\|\cdot\|$ をフロベニウスノルムとし、 $E[Y_i^4] < \infty$ 、 $E[\|X\|^4] < \infty$ を満たす。また、カーネル関数は以下を満たす。

$$\int_{-\infty}^{\infty} u^6 K(u) du < \infty, \quad \int_{-\infty}^{\infty} u^{12} (K(u))^2 du < \infty$$

仮定3-7は $\hat{\beta}_i$ や $\hat{P}_i$ などの推定値と真値の推定誤差がある収束スピードよりも速いことを仮定し、それらの推定量を使ってMTEを推定することがMTEの漸近正規性に影響を与えないようにしている。また仮定8は漸近正規性の保証やリアプノフの中心極限定理の適用に必要なモーメント条件である。

仮定1-8をすべて満たすとMTE推定量 $\hat{E}[Y(1)-Y(0)|X=x, V=p]$ に関して次の漸近正規性が得られる。

定理

仮定1-8が成り立つとき、MTE推定量に関して以下の漸近正規性が成り立つ。

$$\sqrt{nh^3} \left(\hat{E}[Y(1)-Y(0)|X=x, V=p] - E[Y(1)-Y(0)|X=x, V=p] - B(p)h^2 \right) \\ \xrightarrow{d} N \left(0, \frac{\kappa^{22}\sigma^2(p)}{f_{P(Z)}(p)\mu_2^2} \right)$$

ただし、 $f_{P(Z)}(p) > 0$ でそれぞれ

$$B(p) = \frac{1}{3!} \frac{\int u^4 K(u) du}{\int u^2 K(u) du} \frac{\partial^3 K(v, x)}{\partial v^3} \Big|_{v=p}, \quad \kappa^{22} = \int u^2 (K(u))^2 du, \quad \sigma^2(p) = E[\varepsilon_i^2 | P(Z)=p],$$

$$\mu_2 = \int u^2 K(u) du$$

この結果により自身が提案した推定量の分散およびバイアスが明示的となり、これらの値を推定することができれば正規分布の下でMTE推定量に関する仮説検定などを行うことができる。

V. シミュレーション

V-1. シミュレーションの方法

まず各変数は次のようにして生成した。

$$\begin{bmatrix} U_0 \\ U_1 \\ V \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3.6 & 0.4 & 0.16 \\ 0.4 & 4 & 2.8 \\ 0.16 & 2.8 & 2.2 \end{bmatrix} \right)$$

$$X_1 \sim N(-2, 4) \quad X_2 \sim N(2, 4) \quad Z_1 \sim N(-3, 9) \quad Z_2 \sim N(3, 4) \quad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

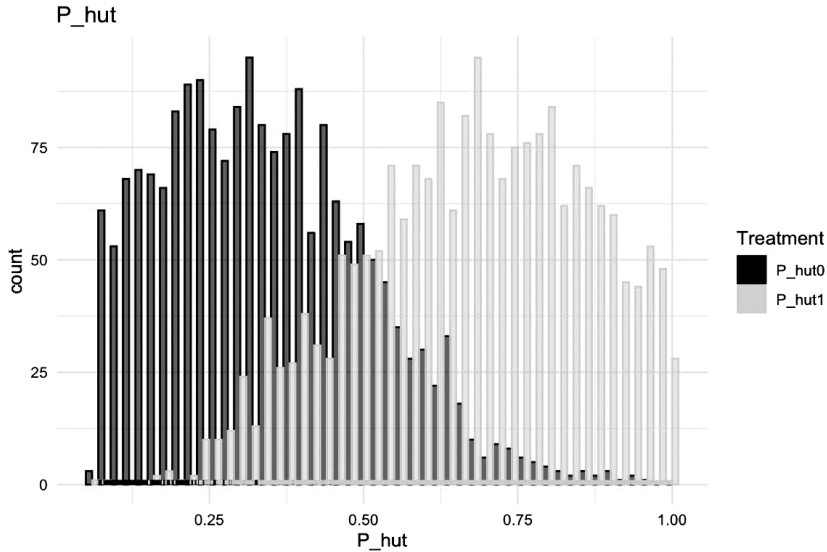


図 1. P_hut の処置別のヒストグラム

縦軸はデータの数, 横軸は $\hat{P}(Z)$ の値である。なおデータ数はそれぞれ P_hut0 が 1939, P_hut1 が 2059 である。

$$\beta_0 = \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix} \quad \beta_1 = \begin{bmatrix} 0.8 \\ 0.4 \end{bmatrix} \quad \gamma_0 = \begin{bmatrix} 0.7 \\ 0.8 \\ 0.6 \\ 0.5 \end{bmatrix}$$

$$D = 1[Z'\gamma_0 > V] \quad Y(1) = X'\beta_1 + U_1 \quad Y(0) = X'\beta_0 + U_0 \quad Y = DY(1) + (1-D)Y(0)$$

このモデルに従ってサンプル数 5000 のデータを生成し, 識別のための条件として 5000 個のデータの中から $D=1$ の中で最小の $\hat{P}(Z)$ と $D=0$ の中で最大の $\hat{P}(Z)$ との間に含まれるものだけを使用して推定を行なった。なお図 1 はそれぞれ $D=1$ の $\hat{P}(Z)$ を P_hut1, $D=0$ の $\hat{P}(Z)$ を P_hut0 としてヒストグラムを描いたものである。なお, B スプライン法ではそれぞれの説明変数に対してノット数を 10 個として推定し, (5), (6) および (7) の推定では, バンド幅 $h=0.2$ として推定を行った。

V-2. 結果

表 1 は, Carneilo and Lee [2009] の手法および自身が提案した手法の二つの推定量に関して, それらのバイアスと分散を比較したものである。これらは, サンプル数 5000 のブートストラップで推定した。また X_1, X_2 はそれぞれ -2, 2 として固定し, 0.1 から 0.9 までの九つの V の値でもって評価した。

表 1 を見ると, 自身の手法に関して 0.1 や 0.9 などの境界値付近では推定された分散の値が Carneilo らの手法と比べ大きくなっているが, 0.4 や 0.5 付近の値では分散は Carneilo らの手法と同じほどになっている。またバイアスに関しても同じような傾向を見せており, V の値が中央値付近であるときバイアスと分散に関して Carneilo らの手法と同じくらいのパフォーマンスをみせている。

表 2 は、データの生成方法より理論的に求められた MTE の値、Carneilo らの手法で求めた推定値、自身が提案した手法の推定値を先ほどと同様に X_1, X_2 はそれぞれ $-2, 2$ として固定し、 0.1 から 0.9 までの九つの V の値でもって比較したものであり、図 2 はそれらを MTE の理論値のグラフとともにプロットしたものである。また、図 3 はそれぞれ左から Carneilo らの手法のバイアスを取り除いた 95 パーセンタイル信頼区間、自身の手法のバイアスを取り除いた 95 パーセンタイル信頼区間をプロットしたものである。この結果からもやはり V が $0.2 \sim 0.8$ あたりでは、信頼区間に関して同じようなふるまいを見せてはいるが、境界値付近ではばらつきが大きくなっていることがわかる。

表 1. ブートストラップ法により推定した分散とバイアス

V_value	bias_C	bias_CH	var_C	var_CH
0.1	-0.39407	0.013475	0.49625	1.108122
0.2	-0.19154	0.02872	0.173013	0.347845
0.3	-0.25123	-0.1531	0.212377	0.241885
0.4	-0.03007	-0.00479	0.103903	0.080468
0.5	0.058357	0.014558	0.097585	0.080071
0.6	0.118807	0.044122	0.116315	0.120633
0.7	0.216839	0.128757	0.263856	0.242721
0.8	0.210513	0.318126	0.58014	0.761042
0.9	0.580583	0.817363	0.968915	1.735848

それぞれ C が Carneilo and Lee (2009) の手法、CH が自身の提案した手法を表す。

表 2 : 理論値と各推定値の比較

V_value	TMTE	theta_C	theta_CH	unbiased_C	unbiased_CH
0.1	-1.53786	-2.76905	-1.97376	-2.374981303	-1.987236106
0.2	-1.00995	-2.2124	-2.3302	-2.020860845	-2.358919551
0.3	-0.62928	-1.59773	-1.91498	-1.346499957	-1.761875514
0.4	-0.30402	-0.88244	-1.04827	-0.852365983	-1.043474632
0.5	0	-0.09843	-0.02149	-0.156789191	-0.03605181
0.6	0.304017	0.689073	0.933356	0.570266817	0.889233884
0.7	0.629281	1.438492	1.628057	1.221653331	1.499299883
0.8	1.009945	2.144606	1.891076	1.934092406	1.572949691
0.9	1.537862	2.827371	1.567815	2.246787939	0.750452679

それぞれ C が Carneilo and Lee (2009) の手法、CH が自身の提案した手法を表す。
なお二列目は理論的に求められた MTE の値である。

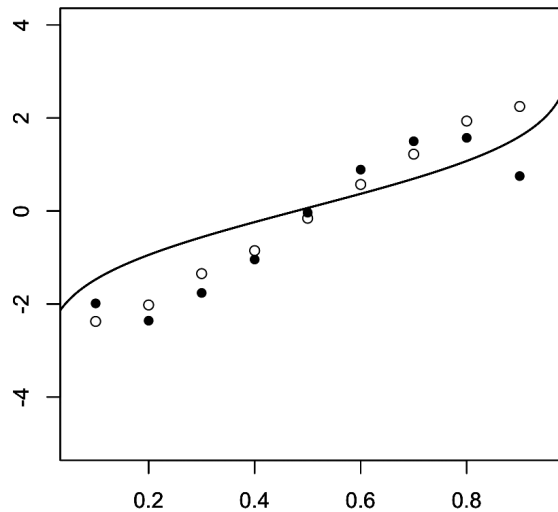


図 2：理論値と各推定値の比較

バイアスを取り除いた MTE 推定量をプロットしたもの。縦軸は MTE の値，横軸は V の値を表す。黒線は理論値を表し，点はそれぞれ○が Carneilo の手法，●が自身の手法で推定した推定値である。

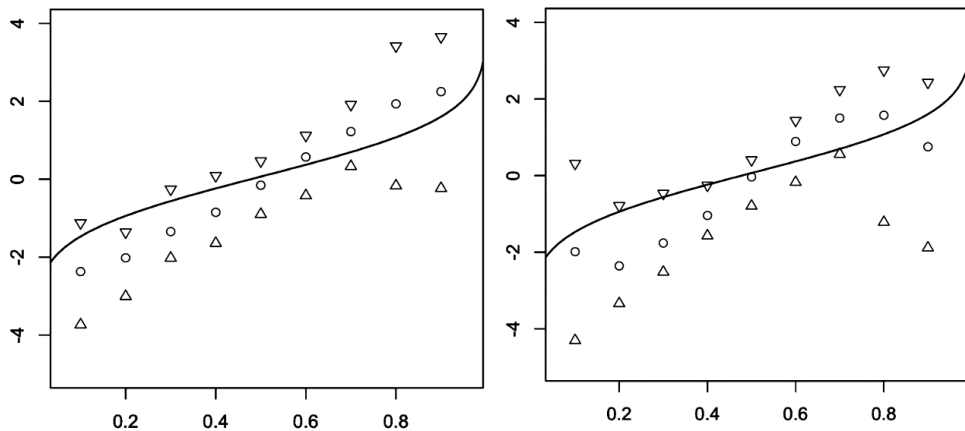


図 3：バイアスを取り除いて求めた 95 パーセントイル信頼区間

それぞれ左から Carneilo らの手法のバイアスを取り除いた 95 パーセントイル信頼区間，自身の手法のバイアスを取り除いた 95 パーセントイル信頼区間をプロットしたものである。なお，黒線は理論値を表し，点は▽・△がそれぞれ信頼区間の上端・下端に対応し，○はバイアスを取り除いた MTE の推定値である。

Ⅵ. 結論

本論文では Heckman らの手法と Carneilo らの手法の二つを基礎としながら、 β_l の推定に際しては仮定 3 を置き、 $\partial K(v)/\partial v|_{v=p}$ の推定に関する理論的分析を進め、いくつかの仮定の下で MTE 推定量が漸近正規性を持つことを証明した。また、数値シミュレーションによって、自身が提案した手法は、Carneilo らが開発したセミパラメトリック MTE 推定量と、特に V の値が中央値付近の値を取るときに、同程度のふるまいを見せることが明らかとなり¹⁾、より簡便な方法で Carneilo らの手法と同程度の精度を達成できることを示した。

参考文献

- 市村英彦「ミクロ実証分析の進展と今後の展望」(日本経済学会編『日本 経済学会 75 年史：回顧と展望』, 有斐閣, 2010 年), 289-361 ページ。
- 久保川達也『現代数理統計学の基礎』(共立講座 数学の魅力 11) 共立出版, 2017 年, 23-25 ページ。
- Aakvik, A., J. J. Heckman, and E. J. Vytlacil, "Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to norwegian vocational rehabilitation programs," *Journal of Econometrics*, 125, (1-2), 2005, pp. 15-51.
- Bates, D. M. and W. N. Venables, Package 'splines', 2016 ([https:// prsr.ism.ac.jp/~nakama/Rjp/splines-manual.pdf](https://prsr.ism.ac.jp/~nakama/Rjp/splines-manual.pdf), 2022 年 1 月 23 日アクセス)。
- Brinch, C.N., M. Mogstad and M. Wiswall, "Beyond LATE with a Discrete Instrument," *Journal of Political Economy*, 125, (4), 2017, pp. 985-1039.
- Carneiro, P., J. J. Heckman and E. Vytlacil, "Evaluating marginal policy changes and the average effect of treatment for individuals at the marginal," *Econometrica*, 78, (1), 2010, pp. 377-394.
- Carneiro, P. and S. Lee, "Estimating distributions of potential outcomes using local instrumental variables with application to changes in college enrollment and wage inequality," *Journal of Econometrics*, 49, (2), 2009, pp. 191-208.
- Cornelissen, T., C. Dustman, A. Raute and U. Schönberg, "From LATE to MTE: Alternative methods for the evaluation of policy interventions," *Labour Economics*, 41, 2016, pp. 47-60.
- Fan, J. and I. Gijbels, *Local Polynomial Modelling and Its Applications*, London, Chapman & Hall, 1996.
- Heckman, J. J., "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic literature*, 48, (2), 2010, pp. 356-398.
- Heckman, J. J., J. L. Tobias and E. Vytlacil, "Simple estimators for treatment parameters in a latent-variable framework," *Review of Economics and Statistics*, 85, (3), 2003, pp. 748-755.
- Heckman, J. J., S. Urzua, and E. Vytlacil, "Understanding instrumental variables in models with essential heterogeneity," *Review of Economics and Statistics*, 88, (3), 2006, pp. 389-432.
- Heckman, J. J. and E. Vytlacil, "Local Instrument Variable and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences*, 96, (8), 1999, pp. 4730-34.
- "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, (3),

1) なお、 $\mu_{-1}(X) - \mu_{-0}(X)$ の項は、仮定 3 に漸近正規性に必要な条件を集約させた。本論文では議論しなかったが Heckman, Urzua and Vytlacil (2006) の提案している推定法が仮定 3 を満たしているのであれば、本論文の結果を適用できるが、彼らの手法は $\widehat{P(Z)}$ を用いた二段階セミパラメトリック推定となるのでその分析はやや複雑である。詳しくは Carneilo and Lee [2009] や Ichimura and Lee [2010] などを参照されたい。

2005, pp. 669–738.

Ichimura, H. and S. Lee, “Characterization of the asymptotic distribution of semiparametric M-estimators,” *Journal of Econometrics*, 159, (2), 2010, pp. 252–266.

Li, Q. and J. Racine, *Nonparametric Econometrics: Theory and Practice*, Princeton University Press., 2007.

Racine, J., *An Introduction to the Advanced Theory and Practice of Nonparametric Econometrics: A Replicable Approach Using R*, Cambridge University Press, 2019.

付録：定理の証明

Step1

$$\left. \frac{\partial \widehat{E}[\widehat{Y} | \widehat{P(Z)=v}]}{\partial v} \right|_{v=p} = (e_2' (\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q)^{-1} (\widehat{X}'_Q \widehat{W}_h \widehat{Y})) + o_p((nh^3)^{-\frac{1}{2}}) \quad (A1)$$

であることを示す。

証明

(7) の推定量 $\widehat{\beta}$ は、次のように書ける。

$$\begin{aligned} \widehat{\beta} &= \left. \frac{\partial \widehat{E}[\widehat{Y} | \widehat{P(Z)=v}]}{\partial v} \right|_{v=p} = e_2' (\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q)^{-1} (\widehat{X}'_Q \widehat{W}_h \widehat{Y}) \\ \text{where } \widehat{X}_Q &= \begin{bmatrix} 1 & (\widehat{P}_1 - p) & (\widehat{P}_1 - p)^2 \\ \vdots & \vdots & \vdots \\ 1 & (\widehat{P}_n - p) & (\widehat{P}_n - p)^2 \end{bmatrix}, \\ \widehat{W}_h &= \text{diag} \left[\frac{1}{h} K \left(\frac{\widehat{P}_1 - p}{h} \right), \dots, \frac{1}{h} K \left(\frac{\widehat{P}_n - p}{h} \right) \right] \\ \widehat{Y} &= \begin{bmatrix} \widehat{Y}_1 \\ \vdots \\ \widehat{Y}_n \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

となる。

ただし、 $\widehat{Y}_i = Y_i - X_i \widehat{\beta}_0 - \widehat{P}_i (X_i (\widehat{\beta}_1 - \widehat{\beta}_0))$ とする。

この時、

$$\begin{aligned} \widehat{Y}_i &= Y_i - X_i \widehat{\beta}_0 - \widehat{P}_i (X_i (\widehat{\beta}_1 - \widehat{\beta}_0)) \\ &= Y_i - X_i (\beta_0 + O_p(n^{-\frac{1}{2}})) - (P_i + o_p(h^2)) (X_i ((\beta_1 - \beta_0) + O_p(n^{-\frac{1}{2}}))) \\ &= (Y_i - (X_i \beta_0 + P_i X_i (\beta_1 - \beta_0))) \\ &\quad - X_i O_p(n^{-\frac{1}{2}}) - P_i X_i O_p(n^{-\frac{1}{2}}) - X_i o_p(h^2) ((\beta_1 - \beta_0) + O_p(n^{-\frac{1}{2}})) \\ &= \bar{Y}_i - X_i o_p((nh^3)^{-\frac{1}{2}}) \end{aligned}$$

となる。なお、二個目の等式は仮定 3 および 5 より成り立ち、最後の等式は仮定 6 より成り立つ。

よって

$$\left. \frac{\partial \widehat{E}[\widehat{Y} | \widehat{P(Z)=v}]}{\partial v} \right|_{v=p} = e_2' (\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q)^{-1} (\widehat{X}'_Q \widehat{W}_h \widehat{Y}) - R_n(p) o_p((nh^3)^{-\frac{1}{2}})$$

となる。ただし、

$$R_n(p) = e_2' \left(\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q \right)^{-1} \left(\widehat{X}'_Q \widehat{W}_h X \right) \text{ where } X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

とする。

仮定 5, 仮定 6 および 仮定 7 より、

$$R_n(p) = e_2' (X'_Q W_h X_Q)^{-1} (X'_Q W_h X) (1 + o_p(1)) = O_p(1)$$

よって

$$\left. \frac{\partial E \left[\widehat{Y} \mid \widehat{P}(Z) = v \right]}{\partial v} \right|_{v=p} = e_2' \left(\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q \right)^{-1} \left(\widehat{X}'_Q \widehat{W}_h \bar{Y} \right) + o_p \left((nh^3)^{-\frac{1}{2}} \right)$$

Step2

$$\begin{aligned} & e_2' \left(\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q \right)^{-1} \left(\widehat{X}'_Q \widehat{W}_h \bar{Y} \right) - \beta \\ &= e_2' \left(\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q \right)^{-1} \left(\widehat{X}'_Q \widehat{W}_h \varepsilon \right) + e_2' \left(\widehat{X}'_Q \widehat{W}_h \widehat{X}_Q \right)^{-1} \left(\widehat{X}'_Q \widehat{W}_h r \right) + o_p(h^2) \\ &= e_2' (X'_Q W_h X_Q)^{-1} (X'_Q W_h \varepsilon) (1 + o_p(1)) + e_2' (X'_Q W_h X_Q)^{-1} (X'_Q W_h r) (1 + o_p(1)) + o_p \left(\frac{1}{\sqrt{nh^3}} \right) \end{aligned} \quad (\text{A2})$$

$$\text{となることを示す。ただし、} \quad r = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} \quad \beta = \begin{bmatrix} E[\bar{Y}_i \mid P(Z) = p] \\ \left. \frac{\partial E[\bar{Y}_i \mid P(Z) = v]}{\partial v} \right|_{v=p} \\ \left. \frac{1}{2} \frac{\partial^2 E[\bar{Y}_i \mid P(Z) = v]}{\partial v^2} \right|_{v=p} \end{bmatrix}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \text{ であり、各 } r_i \text{ は剰余項、} \varepsilon_i = \bar{Y}_i - E[\bar{Y}_i \mid P(Z) = P_i] \text{ とする。}$$

証明

$$m(P) = \begin{bmatrix} E[\bar{Y}_i \mid P(Z) = P_1] \\ \vdots \\ E[\bar{Y}_i \mid P(Z) = P_n] \end{bmatrix} \text{ とおくと、} \bar{Y} = m(P) + \varepsilon \text{ とかける。}$$

$$\text{このとき } E[\varepsilon_i \mid P(Z)] = 0$$

仮定 4 より、

$$\begin{aligned} & E[\bar{Y}_i \mid P(Z) = P_i] \\ &= E[\bar{Y}_i \mid P(Z) = p] + \left. \frac{\partial E[\bar{Y}_i \mid P(Z) = v]}{\partial v} \right|_{v=p} (P_i - p) \end{aligned}$$

$$+\frac{1}{2}\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p}(P_i-p)^2+r_i$$

となる。

したがって、

$$m(P)=X_Q\beta+r$$

であるから

$$\tilde{Y}=X_Q\beta+r+\varepsilon$$

と書ける。

よって

$$\begin{aligned} & e_2'(\hat{X}_Q'\hat{W}_h\hat{X}_Q)^{-1}(\hat{X}_Q'\hat{W}_h\tilde{Y})-e_2'\beta=e_2'(\hat{X}_Q'\hat{W}_h\hat{X}_Q)^{-1}(\hat{X}_Q'\hat{W}_h(\tilde{Y}-m(P))) \\ & +e_2'(\hat{X}_Q'\hat{W}_h\hat{X}_Q)^{-1}(\hat{X}_Q'\hat{W}_hm(P))-\frac{\partial E[\tilde{Y}_i|P(Z)=v]}{\partial v}\bigg|_{v=p}=e_2'(\hat{X}_Q'\hat{W}_h\hat{X}_Q)^{-1}(\hat{X}_Q'\hat{W}_h\varepsilon) \\ & +e_2'(\hat{X}_Q'\hat{W}_h\hat{X}_Q)^{-1}(\hat{X}_Q'\hat{W}_hm(P))-\frac{\partial E[\tilde{Y}_i|P(Z)=v]}{\partial v}\bigg|_{v=p} \end{aligned}$$

各 $m(P)$ の i 項は、仮定 4 より

$$\begin{aligned} E[\tilde{Y}_i|P(Z)=P_i]&=E[\tilde{Y}_i|P(Z)=p]+\frac{\partial E[\tilde{Y}_i|P(Z)=v]}{\partial v}\bigg|_{v=p}(\hat{P}_i-p) \\ & -\frac{\partial E[\tilde{Y}_i|P(Z)=v]}{\partial v}\bigg|_{v=p}(\hat{P}_i-P_i)+\frac{1}{2}\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p}(\hat{P}_i-P_i)^2 \\ & +\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p}(P_i-\hat{P}_i)(\hat{P}_i-p)+\frac{1}{2}\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p}(\hat{P}_i-p)^2 \\ & +r_i \end{aligned}$$

とも書けるから、

$$\begin{aligned} & E[\tilde{Y}_i|P(Z)=P_i]=[1, (\hat{P}_i-p), (\hat{P}_i-p)^2] \cdot \\ & \left[\begin{aligned} & E[\tilde{Y}_i|P(Z)=p]-\frac{\partial E[\tilde{Y}_i|P(Z)=v]}{\partial v}\bigg|_{v=p}(\hat{P}_i-P_i)+\frac{1}{2}\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p}(\hat{P}_i-P_i)^2 \\ & \frac{\partial E[\tilde{Y}_i|P(Z)=v]}{\partial v}\bigg|_{v=p}+\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p}(P_i-\hat{P}_i) \\ & \frac{1}{2}\frac{\partial^2 E[\tilde{Y}_i|P(Z)=v]}{\partial v^2}\bigg|_{v=p} \end{aligned} \right] +r_i \end{aligned}$$

$$= \left[1, \left(\hat{P}_i - p \right), \left(\hat{P}_i - p \right)^2 \right] \cdot \left[\begin{array}{c} E[\tilde{Y}_i | P(Z) = p] + o_p(h^2) \\ \frac{\partial E[\tilde{Y}_i | P(Z) = v]}{\partial v} \Big|_{v=p} + o_p(h^2) \\ \frac{1}{2} \frac{\partial^2 E[\tilde{Y}_i | P(Z) = v]}{\partial v^2} \Big|_{v=p} \end{array} \right] + r_i$$

よって,

$$m(P) = \hat{X}'_Q \left[\begin{array}{c} E[\tilde{Y}_i | P(Z) = p] + o_p(h^2) \\ \frac{\partial E[\tilde{Y}_i | P(Z) = v]}{\partial v} \Big|_{v=p} + o_p(h^2) \\ \frac{1}{2} \frac{\partial^2 E[\tilde{Y}_i | P(Z) = v]}{\partial v^2} \Big|_{v=p} \end{array} \right] + r$$

したがって,

$$\begin{aligned} & e_2' \left(\hat{X}'_Q \hat{W}_h \hat{X}_Q \right)^{-1} \left(\hat{X}'_Q \hat{W}_h \tilde{Y} \right) - \beta \\ &= e_2' \left(\hat{X}'_Q \hat{W}_h \hat{X}_Q \right)^{-1} \left(\hat{X}'_Q \hat{W}_h \varepsilon \right) \\ &+ e_2' \left[\begin{array}{c} E[\tilde{Y}_i | P(Z) = p] + o_p(h^2) \\ \frac{\partial E[\tilde{Y}_i | P(Z) = v]}{\partial v} \Big|_{v=p} + o_p(h^2) \\ \frac{1}{2} \frac{\partial^2 E[\tilde{Y}_i | P(Z) = v]}{\partial v^2} \Big|_{v=p} \end{array} \right] - \frac{\partial E[\tilde{Y}_i | P(Z) = v]}{\partial v} \Big|_{v=p} + e_2' \left(\hat{X}'_Q \hat{W}_h \hat{X}_Q \right)^{-1} \left(\hat{X}'_Q \hat{W}_h r \right) \\ &= e_2' (X'_Q W_h X_Q)^{-1} (X'_Q W_h \varepsilon) (1 + o_p(1)) + e_2' (X'_Q W_h X_Q)^{-1} (X'_Q W_h r) (1 + o_p(1)) + o_p \left(\frac{1}{\sqrt{nh^3}} \right) \end{aligned}$$

なお最後の等式は、仮定 5 および仮定 6 より成り立つ。

Step3

$$e_2' (X'_Q W_h X_Q)^{-1} (X'_Q W_h \varepsilon) = \left(e_2' Q^{-1} \left(\frac{1}{n} X'_Q W_h \varepsilon \right) \right) + o_p \left(\frac{1}{\sqrt{nh^3}} \right) \quad (\text{A3})$$

であることを示す。

$$\text{ただし, } Q^{-1} = \begin{bmatrix} \frac{\mu_4}{f_{P(Z)}(p)(\mu_4 - \mu_2^2)} & 0 & \frac{-\mu_2}{f_{P(Z)}(p)h^2(\mu_4 - \mu_2^2)} \\ 0 & \frac{1}{f_{P(Z)}(p)h^2\mu_2} & 0 \\ \frac{-\mu_2}{f_{P(Z)}(p)h^2(\mu_4 - \mu_2^2)} & 0 & \frac{1}{f_{P(Z)}(p)h^4(\mu_4 - \mu_2^2)} \end{bmatrix}, \quad \mu_4 = \int u^4 K(u) du$$

証明

$$\frac{1}{n} X_Q' W_h X_Q = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p) & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^2 \\ \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p) & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^2 & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^3 \\ \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^2 & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^3 & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^4 \end{bmatrix}$$

の各項に対して考察する。

$$\begin{aligned} E\left[\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)\right] &= \int \frac{1}{h} K\left(\frac{P - p}{h}\right) f_{P(Z)}(P) dP = \int \frac{1}{h} K(p) f_{P(Z)}(p + hp) h dp \\ &= f_{P(Z)}(p) + o(1) \end{aligned}$$

$$\begin{aligned} Var\left[\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)\right] &= \frac{1}{n} E\left[\left(\frac{1}{h} K\left(\frac{P_i - p}{h}\right)\right)^2\right] + O\left(\frac{1}{n}\right) \\ &= \frac{1}{n} \int \frac{1}{h^2} K(p)^2 f_{P(Z)}(p + hp) h dp + O\left(\frac{1}{n}\right) = O\left(\frac{1}{nh}\right) \end{aligned}$$

ただし、 $p' = \frac{P - p}{h}$ である。よって Theorem A.7 (Li and Racine 2007) より、

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) = f_{P(Z)}(p) + o_p\left(\frac{1}{\sqrt{nh}}\right) + o_p(1) = f_{P(Z)}(p) + o_p(1)$$

となる。

同様にして、

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p) &= o_p(h), & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^2 &= \mu_2 h^2 f_{P(Z)}(p) + o_p(h^2) \\ \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^3 &= o_p(h^3), & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right)(P_i - p)^4 &= \mu_4 h^4 f_{P(Z)}(p) + o_p(h^4) \end{aligned}$$

となる。

よって、

$$\frac{1}{n} X_Q' W_h X_Q = \begin{bmatrix} f_{P(Z)}(p) + o_p(1) & o_p(h) & \mu_2 h^2 f_{P(Z)}(p) + o_p(h^2) \\ o_p(h) & \mu_2 h^2 f_{P(Z)}(p) + o_p(h^2) & o_p(h^3) \\ \mu_2 h^2 f_{P(Z)}(p) + o_p(h^2) & o_p(h^3) & \mu_4 h^4 f_{P(Z)}(p) + o_p(h^4) \end{bmatrix}$$

となる。

$$A = \begin{bmatrix} a & b & c \\ b & c & d \\ c & d & e \end{bmatrix} \text{ のとき,}$$

$$A^{-1} = \frac{1}{ace + bdc + cbd - c^3 - b^2e - ad^2} \begin{bmatrix} ce - d^2 & cd - be & bd - c^2 \\ cd - be & ae - c^2 & bc - ad \\ bd - c^2 & bc - ad & ac - b^2 \end{bmatrix}$$

であるから、

$$\begin{aligned} \left(\frac{1}{n} X_Q' W_h X_Q\right)^{-1} &= \frac{1}{(f_{P(Z)}(p))^3 h^6 \mu_2 (\mu_4 - \mu_2^2) + o_p(1)} \cdot \\ &\begin{bmatrix} (f_{P(Z)}(p))^2 h^6 \mu_4 \mu_2 + o_p(h^6) & o_p(h^5) & \mu_2^2 h^4 (f_{P(Z)}(p))^2 + o_p(h^4) \\ o_p(h^5) & h^4 (f_{P(Z)}(p))^2 (\mu_4 - \mu_2^2) + o_p(h^4) & o_p(h^3) \\ \mu_2^2 h^4 (f_{P(Z)}(p))^2 + o_p(h^4) & o_p(h^3) & \mu_2 h^2 (f_{P(Z)}(p))^2 + o_p(h^2) \end{bmatrix} \quad (A4) \end{aligned}$$

となる。

$$\text{このとき } D(h) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & h^2 & 0 \\ 0 & 0 & h^4 \end{bmatrix} \text{ とおくと,}$$

$$\left(\left(\frac{1}{n} X_Q' W_h X_Q \right)^{-1} - Q^{-1} \right) D(h) = \begin{bmatrix} o_p(1) & o_p(h) & o_p(h^2) \\ \frac{1}{h} o_p(1) & o_p(1) & o_p(h) \\ \frac{1}{h^2} o_p(1) & \frac{1}{h} o_p(1) & o_p(1) \end{bmatrix} (1 + o_p(1))$$

となる。

$$(D(h))^{-1} \left(\frac{1}{n} X_Q' W_h \varepsilon \right) \text{ は定義より,}$$

$$(D(h))^{-1} \left(\frac{1}{n} X_Q' W_h \varepsilon \right) = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) \varepsilon_i \\ \frac{1}{nh^2} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) (P_i - p) \varepsilon_i \\ \frac{1}{nh^4} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) (P_i - p)^2 \varepsilon_i \end{bmatrix}$$

各項を考察する。

$$E \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) \varepsilon_i \right] = E \left[\frac{1}{h} K \left(\frac{P - p}{h} \right) \varepsilon_i \right] = E \left[\frac{1}{h} K \left(\frac{P - p}{h} \right) E[\varepsilon_i | P] \right] = 0$$

$$\begin{aligned} \text{Var} \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) \varepsilon_i \right] &= \frac{1}{n} E \left[\frac{1}{h^2} K \left(\frac{P - p}{h} \right)^2 \varepsilon_i^2 \right] = \frac{1}{n} \int \frac{1}{h^2} K(p)^2 E[\varepsilon_i^2 | P = p + hp] f_{P(Z)}(p + hp) h dp \\ &= \frac{1}{nh} \int K(p)^2 E[\varepsilon_i^2 | P = p + hp] f_{P(Z)}(p + hp) dp = O\left(\frac{1}{nh}\right) \end{aligned}$$

ただし, $p' = \frac{P-p}{h}$ である。よって,

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) \varepsilon_i = O_p\left(\frac{1}{\sqrt{nh}}\right)$$

同様にして,

$$\frac{1}{nh^2} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) (P_i - p) \varepsilon_i = O_p\left(\frac{1}{\sqrt{nh^3}}\right), \quad \frac{1}{nh^4} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) (P_i - p)^2 \varepsilon_i = O_p\left(\frac{1}{\sqrt{nh^5}}\right)$$

となる。

よって

$$(D(h))^{-1} \left(\frac{1}{n} X_Q' W_h \varepsilon \right) = \begin{bmatrix} O_p\left(\frac{1}{\sqrt{nh}}\right) \\ O_p\left(\frac{1}{\sqrt{nh^3}}\right) \\ O_p\left(\frac{1}{\sqrt{nh^5}}\right) \end{bmatrix}$$

であり, $e_2' \left(\left(\frac{1}{n} X_Q' W_h X_Q \right)^{-1} - Q^{-1} \right) \left(\frac{1}{n} X_Q' W_h \varepsilon \right) = o_p\left(\frac{1}{\sqrt{nh^3}}\right)$ となる。

$$\text{したがって, } e_2'(X_Q' W_h X_Q)^{-1} (X_Q' W_h \epsilon) = (e_2' Q^{-1} \left(\frac{1}{n} X_Q' W_h \epsilon \right) + o_p \left(\frac{1}{\sqrt{nh^3}} \right))$$

となる。

Step4

$$\sqrt{nh^3} e_2' Q^{-1} \left(\frac{1}{n} X_Q' W_h \epsilon \right) \xrightarrow{d} N \left(0, \frac{\kappa^{22} \sigma^2(p)}{f_{P(Z)}(p) \mu_2^2} \right) \quad (\text{A5})$$

となることを示す。

証明

$$\begin{aligned} \sqrt{nh^3} e_2' Q^{-1} \left(\frac{1}{n} X_Q' W_h \epsilon \right) &= \sqrt{nh^3} e_2' Q^{-1} D(h) (D(h))^{-1} \left(\frac{1}{n} X_Q' W_h \epsilon \right) \\ &= \sqrt{nh^3} \begin{bmatrix} 0 & \frac{1}{f_{P(Z)}(p) \mu_2} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) \epsilon_i \\ \frac{1}{nh^2} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) (P_i - p) \epsilon_i \\ \frac{1}{nh^4} \sum_{i=1}^n \frac{1}{h} K \left(\frac{P_i - p}{h} \right) (P_i - p)^2 \epsilon_i \end{bmatrix} \\ &= \sqrt{nh^3} \frac{1}{f_{P(Z)}(p) \mu_2} \frac{1}{nh^3} \sum_{i=1}^n K \left(\frac{P_i - p}{h} \right) (P_i - p) \epsilon_i \\ &= \frac{1}{f_{P(Z)}(p) \mu_2} \sum_{i=1}^n \frac{1}{\sqrt{nh^3}} K \left(\frac{P_i - p}{h} \right) (P_i - p) \epsilon_i \end{aligned}$$

となることが定義よりわかる。

$$Z_{n,i} = \frac{1}{\sqrt{nh^3}} K \left(\frac{P_i - p}{h} \right) (P_i - p) \epsilon_i$$

とすると,

$$\begin{aligned} \text{Var}(Z_{n,i}) &= \int \frac{1}{nh^3} (P - p)^2 K \left(\frac{P - p}{h} \right)^2 \sigma^2(P) f_{P(Z)}(P) dP \\ &= \int \frac{1}{nh^3} (hp)^2 K(p)^2 \sigma^2(p + hp) f_{P(Z)}(p + hp) h dp \\ &= \frac{1}{n} \kappa^{22} \sigma^2(p) f_{P(Z)}(p) + o\left(\frac{1}{n}\right) \end{aligned}$$

となる。ただし, $p' = \frac{P-p}{h}$ である。よって $S_n = \sum_{i=1}^n Z_{n,i}$ とすると,

$$\text{Var}(S_n) = \kappa^{22} \sigma^2(p) f_{P(Z)}(p) + o_p(1)$$

となる。

リアプノフの中心極限定理より

$$S_n \xrightarrow{d} N(0, \kappa^{22} \sigma^2(p) f_{P(Z)}(p))$$

であるから,

$$\sqrt{nh^3} e_2' Q^{-1} \left(\frac{1}{n} X_Q' W_h \epsilon \right) \xrightarrow{d} N \left(0, \frac{\kappa^{22} \sigma^2(p)}{f_{P(Z)}(p) \mu_2^2} \right)$$

が成り立つ。

Step5

$$(e_2'(X_Q'W_hX_Q)^{-1}(X_Q'W_hr))=B(p)h^2+o_p\left(\frac{1}{\sqrt{nh^3}}\right) \quad (A6)$$

となることを示す。

証明

$E[\tilde{Y}_i|P(Z)]$ は四回連続微分可能であるから、ある実数 $\lambda \in (0,1)$ が存在し

$$r_i = \frac{1}{3!} \frac{\partial^3 E[\tilde{Y}_i|P(Z)=v]}{\partial v^3} \Big|_{v=p} (P_i-p)^3 + \frac{1}{4!} \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^4$$

と書ける。

よって

$$\begin{aligned} & \left(\frac{1}{n} X_Q' W_h r \right) = \\ & \left[\begin{aligned} & \beta_3 \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) (P_i-p)^3 + \frac{1}{4!} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^4 \\ & \beta_3 \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) (P_i-p)^4 + \frac{1}{4!} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^5 \\ & \beta_3 \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) (P_i-p)^5 + \frac{1}{4!} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^6 \end{aligned} \right] \end{aligned}$$

となる。ただし $\beta_3 = \left(\frac{1}{3!} \frac{\partial^3 E[\tilde{Y}_i|P(Z)=v]}{\partial v^3} \Big|_{v=p} \right)$ とおいた。

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^4 \text{ に関して,}$$

$$\begin{aligned} & E\left[\frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i-p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^4 \right] \\ &= \int \frac{1}{h} K(u) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(hu)} h^4 u^4 f_{P(Z)}(v+hu) h du \\ &= \int K(u) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(hu)} h^4 u^4 f_{P(Z)}(v+hu) du = O(h^4) \\ & \text{Var} \left(\frac{1}{h} K\left(\frac{P_i-p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} (P_i-p)^4 \right) \\ &= \frac{1}{n} E \left[\frac{1}{h^2} \left(K\left(\frac{P_i-p}{h}\right) \right)^2 \left(\frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(P_i-p)} \right)^2 (P_i-p)^8 \right] + O\left(\frac{h^8}{n}\right) \\ &= \frac{1}{n} \int \frac{1}{h^2} (K(u))^2 \left(\frac{\partial^4 E[\tilde{Y}_i|P(Z)=v]}{\partial v^4} \Big|_{v=p+\lambda(hu)} \right)^2 h^8 u^8 f_{P(Z)}(v+hu) h du + O\left(\frac{h^8}{n}\right) \end{aligned}$$

$$= O\left(\frac{h^7}{n}\right) + O\left(\frac{h^8}{n}\right) = O\left(\frac{h^7}{n}\right)$$

よって

$$\frac{1}{4!} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i | P(Z) = v]}{\partial v^4} \bigg|_{v=p+\lambda(P_i-p)} (P_i - p)^4 = O_p\left(\sqrt{\frac{h^7}{n}}\right) + O_p(h^4) = o_p(h^3)$$

となる。

同様にして,

$$\begin{aligned} & \frac{1}{4!} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i | P(Z) = v]}{\partial v^4} \bigg|_{v=p+\lambda(P_i-p)} (P_i - p)^5 = o_p(h^4), \\ & \frac{1}{4!} \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) \frac{\partial^4 E[\tilde{Y}_i | P(Z) = v]}{\partial v^4} \bigg|_{v=p+\lambda(P_i-p)} (P_i - p)^6 = o_p(h^5) \end{aligned}$$

となる。

また,

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) (P_i - p)^3 = o_p(h^3), \quad \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) (P_i - p)^4 = \mu_4 h^4 f_{P(Z)}(p) + o_p(h^4) \\ & \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{P_i - p}{h}\right) (P_i - p)^5 = o_p(h^5) \end{aligned}$$

であるから,

$$\left(\frac{1}{n} X_Q' W_h r\right) = \begin{bmatrix} o_p(h^3) \\ \beta_3 \mu_4 h^4 f_{P(Z)}(p) + o_p(h^4) \\ o_p(h^5) \end{bmatrix}$$

したがって (A4) より

$$\begin{aligned} e_2'(X_Q' W_h X_Q)^{-1} (X_Q' W_h r) &= \frac{1}{(f_{P(Z)}(p))^3 h^6 \mu_2 (\mu_4 - \mu_2^2) + o_p(1)}. \\ & \begin{bmatrix} o_p(h^3) \\ \beta_3 \mu_4 h^4 f_{P(Z)}(p) + o_p(h^4) \\ o_p(h^5) \end{bmatrix} \\ &= \beta_3 \frac{\mu_4}{\mu_2} h^2 + o_p(h^2) = B(p) h^2 + o_p\left(\frac{1}{\sqrt{nh^3}}\right) \end{aligned}$$

となる。

Step6

$$\begin{aligned} & \sqrt{nh^3} \left(\widehat{E}[Y(1) - Y(0) | X = x, V = p] - E[Y(1) - Y(0) | X = x, V = p] - B(p) h^2 \right) \\ & \xrightarrow{d} N\left(0, \frac{\kappa^{22} \sigma^2(p)}{f_{P(Z)}(p) \mu_2^2}\right) \end{aligned}$$

を示す。

証明

$$\widehat{E}[Y(1) - Y(0) | X = x, V = p] - E[Y(1) - Y(0) | X = x, V = p]$$

$$= \left(x'(\hat{\beta}_1 - \hat{\beta}_0) + \frac{\partial K(\hat{v})}{\partial v} \Big|_{v=\hat{p}} \right) - \left(x'(\beta_1 - \beta_0) + \frac{\partial K(v)}{\partial v} \Big|_{v=p} \right)$$

仮定 3 より

$$\begin{aligned} & \widehat{E}[Y(1) - Y(0) | X=x, V=p] - E[Y(1) - Y(0) | X=x, V=p] \\ &= \left(\frac{\partial K(\hat{v})}{\partial v} \Big|_{v=\hat{p}} - \frac{\partial K(v)}{\partial v} \Big|_{v=p} \right) + O_p(n^{-\frac{1}{2}}) \end{aligned}$$

である。

(A1), (A2) より

$$\begin{aligned} & \widehat{E}[Y(1) - Y(0) | X=x, V=p] - E[Y(1) - Y(0) | X=x, V=p] = \\ & e_2'(\widehat{X}_Q \widehat{W}_h \widehat{X}_Q)^{-1}(\widehat{X}_Q \widehat{W}_h \bar{Y}) + o_p((nh^3)^{-\frac{1}{2}}) - \frac{\partial K(v)}{\partial v} \Big|_{v=p} + O_p(n^{-\frac{1}{2}}) \\ &= e_2'(X_Q W_h X_Q)^{-1}(X_Q W_h \varepsilon)(1 + o_p(1)) \\ &+ e_2'(X_Q W_h X_Q)^{-1}(X_Q W_h r)(1 + o_p(1)) + o_p\left(\frac{1}{\sqrt{nh^3}}\right) \end{aligned}$$

となる。また, (A3) (A5) (A6) より

$$\begin{aligned} & \widehat{E}[Y(1) - Y(0) | X=x, V=p] - E[Y(1) - Y(0) | X=x, V=p] \\ &= e_2'(X_Q W_h X_Q)^{-1}(X_Q W_h \varepsilon) + B(p)h^2 + o_p\left(\frac{1}{\sqrt{nh^3}}\right) \end{aligned}$$

(A3) (A5) より

$$\sqrt{nh^3} e_2'(X_Q W_h X_Q)^{-1}(X_Q W_h \varepsilon) \xrightarrow{d} N\left(0, \frac{\kappa^{22} \sigma^2(p)}{f_{P(Z)}(p) \mu_2^2}\right)$$

であるから

$$\begin{aligned} & \sqrt{nh^3} \left(\widehat{E}[Y(1) - Y(0) | X=x, V=p] - E[Y(1) - Y(0) | X=x, V=p] - B(p)h^2 \right) \\ & \xrightarrow{d} N\left(0, \frac{\kappa^{22} \sigma^2(p)}{f_{P(Z)}(p) \mu_2^2}\right) \end{aligned}$$

となる。