テクニカルノート

ウェブ付録：『因果推論の計量経済学』(日本評論社、2024 年：ISBN 978-4-535-54043-9)

川口康平・澤田真行

2024 年 9 月 13 日

はじめに

このテクニカルノートでは、『因果推論の計量経済学』の各所で提示した命題の数学的な証明や、解説した内容のテクニカルな部分の解説を提供しています。

書籍内とテクニカルノートの対応は、T2.1 [*1] のマークで示しています。本書内の該当箇所には、テクニカルノートへの言及と、このマークが付いています。

テクニカルノートでの解説は、以下の目次の各章で提供されています。本書を読み進めながら、適宜こちらのノートも参照してください。

目次

[*1] 「T」はテクニカルノートの T、「2」は章番号に対応しています。

第 2 章　無作為化実験

T2.1 命題 2.1 の証明

平均処置効果

$$\tau = \frac{1}{n}\sum_{i=1}^{n}(Y_i^*(1) - Y_i^*(0)) \tag{1}$$

τ の推定量は、

$$\begin{aligned}\hat{\tau} = \bar{Y_1} - \bar{Y_0} &= \frac{1}{n_1}\sum_{i=1}^{n} Z_iY_i - \frac{1}{n_0}\sum_{i=1}^{n}(1-Z_i)Y_i \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Z_iY_i^*(1)}{n_1/n} - \frac{(1-Z_i)Y_i^*(0)}{n_0/n}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{n}{n_1}Z_iY_i^*(1) - \frac{n}{n_0}(1-Z_i)Y_i^*(0)\right).\end{aligned} \tag{2}$$

中心化した処置ベクトルを以下のように定義する。

$$D_i = Z_i - \frac{n_1}{n} = \begin{cases}\frac{n_0}{n} & \text{if } Z_i = 1 \\ -\frac{n_1}{n} & \text{if } Z_i = 0\end{cases} \tag{3}$$

すると、

$$\mathbb{E}\left[D_i\right] = 0 \tag{4}$$

$$\mathbb{V}\left[D_i\right] = \mathbb{E}\left[D_i^2\right] = \frac{n_0n_1}{n^2}, \tag{5}$$

また $i \neq j$ のとき

$$\mathbb{P}\left[D_iD_j = d\right] = \begin{cases}\frac{n_1(n_1-1)}{n(n-1)} & \text{if } d = \frac{n_0^2}{n^2} \\ 2\frac{n_1n_0}{n(n-1)} & \text{if } d = -\frac{n_1n_0}{n^2} \\ \frac{n_0(n_0-1)}{n(n-1)} & \text{if } d = \frac{n_1^2}{n^2} \\ 0 & \text{otherwise.}\end{cases} \tag{6}$$

これらより以下が導ける

$$\mathbb{E}\left[D_iD_j\right] = \begin{cases}\frac{n_0n_1}{n^2} & \text{if i = j} \\ -\frac{n_1n_0}{n^2(n-1)} & \text{if } i \neq j\end{cases} \tag{7}$$

τ の推定量は D_i を用いて以下のように書き直せる。

$$
\begin{aligned}
\bar{Y_1} - \bar{Y_0} &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{n}{n_1}\left(D_i + \frac{n_1}{n}\right)Y_i^*(1) - \frac{n}{n_0}\left(\frac{n_0}{n} - D_i\right)Y_i^*(0)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(Y_i^*(1) - Y_i^*(0)\right) + \frac{1}{n}\sum_{i=1}^{n} D_i\left(\frac{n}{n_1}Y_i^*(1) + \frac{n}{n_0}Y_i^*(0)\right) \\
&= \tau + \frac{1}{n}\sum_{i=1}^{n} D_i\left(\frac{n}{n_1}Y_i^*(1) + \frac{n}{n_0}Y_i^*(0)\right)
\end{aligned} \tag{8}
$$

$Y_i^+ = \frac{n}{n_1}Y_i^*(1) + \frac{n}{n_0}Y_i^*(0)$ と定義すると,

$$
\mathbb{V}[\hat{\tau}] = \mathbb{V}\left[\frac{1}{n}\sum_{i=1}^{n} D_i Y_i^+\right] = \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^{n} D_i Y_i^+\right)^2\right]. \tag{9}
$$

式 (9) を展開すると,

$$
\begin{aligned}
\mathbb{V}[\hat{\tau}] &= \mathbb{E}\left[\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} D_i D_j Y_i^+ Y_j^+\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}(Y_i^+)^2\mathbb{E}\left[D_i^2\right] + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}\mathbb{E}\left[D_i D_j\right]Y_i^+ Y_j^+ \\
&= \frac{n_0 n_1}{n^4}\sum_{i=1}^{n}(Y_i^+)^2 - \frac{n_0 n_1}{n^4(n-1)}\sum_{i=1}^{n}\sum_{j\neq i} Y_i^+ Y_j^+ \\
&= \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}(Y_i^+)^2 - \frac{n_0 n_1}{n^4(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n} Y_i^+ Y_j^+ \\
&= \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n} Y_i^+(Y_i^+ - \overline{Y^+}) \\
&= \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}(Y_i^+ - \overline{Y^+})^2 + \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}\overline{Y^+}(Y_i^+ - \overline{Y^+}) \\
&= \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}(Y_i^+ - \overline{Y^+})^2.
\end{aligned} \tag{10}
$$

上の 4 式目の導出は、$\frac{n_0 n_1}{n^4(n-1)}\sum_{i=1}^{n}(Y_i^+)^2$ を 1 項目に足して 2 項目から引いている。ここで Y_i^+ の定義を代入すると、

$$
\begin{aligned}
\mathbb{V}[\hat{\tau}] &= \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}\left(\frac{n}{n_1}Y_i^*(1) + \frac{n}{n_0}Y_i^*(0) - \left(\frac{n}{n_1}\bar{Y}^*(1) + \frac{n}{n_0}\bar{Y}^*(0)\right)\right)^2 \\
&= \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}\left(\frac{n}{n_1}Y_i^*(1) - \frac{n}{n_1}\bar{Y}^*(1)\right)^2 \\
&\quad + \frac{n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}\left(\frac{n}{n_0}Y_i^*(0) - \frac{n}{n_0}\bar{Y}_i^*(0)\right)^2 \\
&\quad + \frac{2 n_0 n_1}{n^3(n-1)}\sum_{i=1}^{n}\left(\frac{n}{n_1}Y_i^*(1) - \frac{n}{n_1}\bar{Y}^*(1)\right)\left(\frac{n}{n_0}Y_i^*(0) - \frac{n}{n_0}\bar{Y}_0\right) \\
&= \frac{n_0}{n n_1(n-1)}\sum_{i=1}^{n}\left(Y_i^*(1) - \bar{Y}_i^*(1)\right)^2 + \frac{n_1}{n n_0(n-1)}\sum_{i=1}^{n}\left(Y_i^*(0) - \bar{Y}_i^*(0)\right)^2 \\
&\quad + \frac{2}{n(n-1)}\sum_{i=1}^{n}\left(Y_i^*(1) - \bar{Y}_i^*(1)\right)\left(Y_i^*(0) - \bar{Y}_i^*(0)\right).
\end{aligned} \tag{11}
$$

S^2_{01} の定義を考えると

$$
\begin{aligned}
S^2_{01} &= \frac{1}{n-1}\sum_{i=1}^{n}\left(Y^*_i(1)-\bar{Y}_1-\left(Y^*_i(0)-\bar{Y}^*(0)\right)\right)^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(Y^*_i(1)-\bar{Y}^*(1)\right)^2+\frac{1}{n-1}\sum_{i=1}^{n}\left(Y^*_i(0)-\bar{Y}^*(0)\right)^2 \\
&\quad -\frac{2}{n-1}\sum_{i=1}^{n}\left(Y^*_i(1)-\bar{Y}^*(1)\right)\left(Y^*_i(0)-\bar{Y}^*(0)\right) \\
&= S^2_1+S^2_0-\frac{2}{n-1}\sum_{i=1}^{n}(Y^*_i(1)-\bar{Y}^*(1))\left(Y^*_i(0)-\bar{Y}^*(0)\right)
\end{aligned} \tag{12}
$$

となり、分散が導出される。

$$
\begin{aligned}
\mathbb{V}[\hat{\tau}] &= \frac{n_0}{nn_1}S^2_1+\frac{n_1}{nn_0}S^2_0+\frac{1}{n}\left(S^2_1+S^2_0-S^2_{01}\right) \\
&= \frac{S^2_0}{n_0}+\frac{S^2_1}{n_1}-\frac{S^2_{01}}{n}.
\end{aligned} \tag{13}
$$

T2.2 命題 2.2 の証明

母集団のサイズを n^*、$\mathbf{Y}^*(1), \mathbf{Y}^*(0)$ を母集団の潜在結果ベクトルとする。R_i を、個体 i が標本に含まれるときに 1 をとり、標本に含まれないときに 0 をとる変数とする。標本のサイズが n であれば、$\sum_{i=1}^{n^*} R_i = n$ となる。R_i は平均 $\frac{n}{n^*}$、分散 $\frac{n}{n^*}(1-\frac{n}{n^*})$ のベルヌーイ分布であり、i と j が異なるときの R_i と R_j の間の共分散は、

$$
\frac{n(n-1)}{n^*(n^*-1)}-\frac{n^2}{n^{*2}} \tag{14}
$$

となる。

以下のように母集団平均処置効果、標本平均処置効果などを定める。

$$
\tau = \frac{1}{n^*}\sum_{i=1}^{n^*}(Y_i(1)-Y_i(0)) \tag{15}
$$

$$
\sigma^2_{01} = \frac{1}{n^*}\sum_{i=1}^{n^*}(Y_i(1)-Y_i(0)-\tau)^2 \tag{16}
$$

$$
\tau_{fs} = \frac{1}{n}\sum_{i=1}^{n^*}R_i(Y_i(1)-Y_i(0)) \tag{17}
$$

母集団の潜在結果を固定して標本平均処置効果の無作為抽出による R_i に関する期待値を計算すると、以下の通り、母集団平均処置効果に一致する。

$$
\begin{aligned}
\mathbb{E}\left[\tau_{fs}|\mathbf{Y}^*(1),\mathbf{Y}^*(0)\right] &= \frac{1}{n^*}\sum_{i=1}^{n}\mathbb{E}[R_i](Y_i(1)-Y_i(0)) \\
&= \frac{1}{n}\sum_{i=1}^{n^*}\frac{n}{n_*}(Yi(1)-Y_i(0)) = \tau
\end{aligned} \tag{18}
$$

一方、母集団の潜在結果を固定して標本平均処置効果の無作為抽出による R_i に関する分散を計算すると、以下のようになる。

$$
\begin{aligned}
\mathbb{V}[\tau_{fs}|\mathbf{Y}^*(0),\mathbf{Y}^*(1))] &= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n^*} R_i\left(Y_i(1)-Y_i(0)\right)-\tau\right)^2 \middle| \mathbf{Y}^*(0),\mathbf{Y}^*(1)\right] \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n^*}\left(R_i-\frac{n}{n^*}\right)\left(Y_i(1)-Y_i(0)-\tau\right)\right)^2 \middle| \mathbf{Y}^*(0),\mathbf{Y}^*(1)\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n^*}\sum_{j=1}^{n^*}\mathbb{E}\left[\left(R_i-\frac{n}{n^*}\right)\left(R_j-\frac{n}{n^*}\right)\right. \\
&\quad \left.\times\left(Y_i(1)-Y_i(0)-\tau\right)\left(Y_j(1)-Y_j(0)-\tau\right)|\mathbf{Y}^*(0),\mathbf{Y}^*(1)\right] \\
&= \frac{1-n/n^*}{nn^*}\sum_{i=1}^{n^*}\left(Y_i(1)-Y_i(0)-\tau\right)^2 \\
&\quad -\frac{1}{n^2}\left(\frac{n(n-1)}{n^*(n^*-1)}-\frac{n^2}{n^{*2}}\right)\sum_{i=1}^{n^*}\sum_{j\neq i}\left(Y_i(1)-Y_i(0)-\tau\right)\left(Y_j(1)-Y_j(0)-\tau\right) \\
&= \frac{\sigma_{01}^2}{n}-\frac{\sigma_{01}^2}{n^*}-\frac{1}{n^2}\left(\frac{n(n-1)}{n^*(n^*-1)}-\frac{n^2}{n^{*2}}\right) \\
&\quad \times\sum_{i=1}^{n^*}\sum_{j\neq i}\left(Y_i(1)-Y_i(0)-\tau\right)\left(Y_j(1)-Y_j(0)-\tau\right)
\end{aligned} \tag{19}
$$

ここで、$n^*\to\infty, \frac{n}{n^*}\to 0$ のとき、

$$
\mathbb{V}[\tau_{fs}|Y(0),Y(1)]\to\frac{\sigma_{01}^2}{n} \tag{20}
$$

となる。すなわち、母集団のサイズが標本のサイズと比較して十分大きいとき、標本平均処置効果の無作為抽出による R_i に関する分散は σ_{01}^2 になる。

次に、平均値の差の推定量

$$
\hat{\tau}=\bar{Y}_1-\bar{Y}_0 \tag{21}
$$

を考える。無作為抽出 R_i と無作為割当 Z_i を用いて表現すると、以下のようになる。

$$
\hat{\tau}=\frac{1}{n_1}\sum_{i=1}^{n^*}R_iZ_iY_i-\frac{1}{n_0}\sum_{i=1}^{n^*}R_i(1-Z_i)Y_i \tag{22}
$$

母集団における潜在結果および標本を固定して、無作為割当による Z_i に関する期待値を計算すると、

$$
\begin{aligned}
\mathbb{E}\left[\hat{\tau}|R,\mathbf{Y}^*(1),\mathbf{Y}^*(0)\right] &= \frac{1}{n_1}\sum_{i=1}^{n^*}R_i\mathbb{E}[Z_i]Y_i-\frac{1}{n_0}\sum_{i=1}^{n^*}R_i\mathbb{E}[1-Z_i]Y_i \\
&= \frac{1}{n}\sum_{i=1}^{n^*}R_i(Y_i(1)-Y_i(0))=\tau_{fs}
\end{aligned} \tag{23}
$$

となり、標本平均処置効果に一致する。

最後に、母集団における潜在結果を固定して、R_i、Z_i の両方に関する $\hat{\tau}$ の期待値を計算すると、

$$
\begin{aligned}
\mathbb{E}\left[\hat{\tau}|\mathbf{Y}^*(1),\mathbf{Y}^*(0))\right] &= \mathbb{E}\left[\mathbb{E}_{\lesssim}\left[\hat{\tau}|R,\mathbf{Y}^*(\mathbf{1}),\mathbf{Y}^*(\mathbf{0})\right]|\mathbf{Y}^*(1),\mathbf{Y}^*(0)\right] \\
&= \mathbb{E}\left[\hat{\tau}|\mathbf{Y}^*(1),\mathbf{Y}^*(0)\right]=\tau
\end{aligned} \tag{24}
$$

となり、母集団平均処置効果に一致する。

このときの $\hat{\tau}$ の分散は、

$$\begin{aligned}
\mathbb{V} &= \mathbb{V}[\hat{\tau}|\mathbf{Y}^*(1), \mathbf{Y}^*(0)] \\
&= \mathbb{E}[\hat{\tau}^2|\mathbf{Y}^*(1), \mathbf{Y}^*(0))] - \{\mathbb{E}[\hat{\tau}|\mathbf{Y}^*(1), \mathbf{Y}^*(0))]\}^2 \\
&= \mathbb{E}\left[\mathbb{E}[\hat{\tau}^2|\mathbf{R}, \mathbf{Y}^*(1), \mathbf{Y}^*(0)]|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right] - \left(\mathbb{E}\left\{\mathbb{E}[\hat{\tau}|\mathbf{R}, \mathbf{Y}^*(1), \mathbf{Y}^*(0)]|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right]\right)^2 \\
&= \mathbb{E}\left[\mathbb{E}[\hat{\tau}^2|\mathbf{R}, \mathbf{Y}^*(1), \mathbf{Y}^*(0)] - \{\mathbb{E}[\hat{\tau}|\mathbf{R}, \mathbf{Y}^*(1), \mathbf{Y}^*(0)]\}^2|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right] \\
&+ \mathbb{E}\left[\{\mathbb{E}[\hat{\tau}|\mathbf{R}, \mathbf{Y}^*(1), \mathbf{Y}^*(0)]\}^2|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right] - \left[\mathbb{E}\left\{\mathbb{E}[\hat{\tau}|\mathbf{R}, \mathbf{Y}^*(1), \mathbf{Y}^*(0)]|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right\}\right]^2 \\
&= \mathbb{E}[\mathbb{V}[\hat{\tau}|R, \mathbf{Y}^*(1), \mathbf{Y}^*(0)]|\mathbf{Y}^*(1), \mathbf{Y}^*(0)] + \mathbb{V}[\mathbb{E}\left[\hat{\tau}|R, \mathbf{Y}^*(1), \mathbf{Y}^*(0)|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right]] \qquad (25) \\
&= \mathbb{E}\left[\frac{s_0^2}{n_0} + \frac{s_1^2}{n_1} - \frac{s_{01}^2}{n}\middle|\mathbf{Y}^*(1), \mathbf{Y}^*(0)\right] + \mathbb{V}[\tau_{fs}|\mathbf{Y}^*(1), \mathbf{Y}^*(0)] \\
&= \frac{\sigma_0^2}{n_0} + \frac{\sigma_1^2}{n_1} - \frac{\sigma_{01}^2}{n} + \frac{\sigma_{01}^2}{n} - \frac{\sigma_{01}^2}{n^*} - \frac{1}{n^2}\left(\frac{n(n-1)}{n^*(n^*-1)} - \frac{n^2}{n^{*2}}\right)\sum_{i=1}^{n^*}\sum_{j\neq i}(Y_i(1) - Yi(0) - \tau) \\
&(Y_j(1) - Y_j(0) - \tau)
\end{aligned}$$

となる。これは $n^* \to \infty, \frac{n}{n^*} \to 0$ のとき、

$$\frac{\sigma_0^2}{n_0} + \frac{\sigma_1^2}{n_1} \qquad (26)$$

に収束する。

第 6 章　回帰非連続デザインの基礎

T6.1〜6.5

仮定 1　カーネル関数 $K(\cdot)$ は、ある $r > 0$ について以下の条件を満たす

- (非負有界) ある $\bar{K}$ が存在し、$0 \leq K(\cdot) \leq \bar{K} < \infty$、
- (対称) $K(-u) = K(u)$、
- (積分して 1) $\int_{\mathcal{K}} K(u)du = 1$、
- (モーメントの存在) $\int_{\mathcal{K}} |u|^r K(u)du$。

仮定 2　$\{Y_i, S_i\}_{i=1}^n$ を i.i.d. 標本とする。S の密度関数を $f(s)$ とする。このとき、S のサポート $\mathcal{S}$ に含まれる区間 $[0, b], b < \infty$, 上の点 $\bar{s}$ について、以下の条件を満たす：
$f(\bar{s}) > 0$ かつ、点 $s = 0$ で $f(\cdot)$ が 1 階連続微分可能、$m(\cdot)$ が右側 3 階連続微分可能、$\mathbb{V}[\epsilon|S = \cdot]$ が右側連続、かつ、$f(\cdot), f^{'}(\cdot), m(\cdot), m^{'}(\cdot), m^{''}(\cdot)$ および $\mathbb{V}[\epsilon|S = \cdot]$ が $\mathcal{S}$ 上で有界。

定理 1（Hansen, 2020, Theorem 19.2, より）　仮定 1 および 2 のもとで、$nh \to \infty$ かつ $h \to 0$ であるときに、カーネル推定量の分散は

$$\sigma^2_{NW}(0_+) = \frac{(K^2 \mathbb{V}[\epsilon|S = 0_+])}{f(0)nh} + o_P(1/nh)$$

である。ただし、$K^2 \equiv \int_{\mathcal{K}} K(u)^2 du$ であり、$o_P(1/nh)$ は $1/nh$ で割っても (すなわち nh を掛けても) 0 に確率収束する項である。

■**証明**　ある $s \in \mathcal{S}$ について、$\sigma^2_{NW}(s) = \mathbb{V}[\hat{m}_{NW}(s)|S]$ であるが、

$$\hat{m}_{NW}(s) - \mathbb{E}[\hat{m}_{NW}(s)|S] = \frac{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)(Y_i - m(S_i))}{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)} = \frac{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\epsilon_i}{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)}$$

である。したがって、

$$\sigma^2_{NW}(s) = \frac{\mathbb{V}[\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\epsilon_i|S]}{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)^2} = (nh)^{-1}\frac{(nh)^{-1}\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)^2 \mathbb{V}[\epsilon|S_i]}{\left((nh)^{-1}\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\right)^2}$$

である。ここで、分母

$$\hat{f}(s) \equiv (nh)^{-1}\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)$$

はカーネル密度推定量であり、$f(s)$ に収束する。実際、変数変換 $v = s + hu$ を行うと、$f(s)$ の連続性より

$$\mathbb{E}[\hat{f}(s)] = \int_{-\infty}^{\infty} h^{-1} K\left(\frac{v - s}{h}\right) f(v)dv$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} K\left(u\right) f(s+hu)du = \\
&= \int_{-\infty}^{\infty} K(u)f(s)du + o(1) = f(s)
\end{aligned}
$$

ただし、$o(1)$ は $h \to 0$ で 0 に収束する項である。同様に、

$$
\begin{aligned}
\mathbb{V}[\hat{f}(s)] =& (nh)^{-2} n\mathbb{V}\left[K\left(\frac{S-s}{h}\right)\right] \\
\leq& (nh)^{-1}h^{-1}\int_{-\infty}^{\infty} K\left(\frac{v-s}{h}\right)^2 f(v)dv \\
=& (nh)^{-1}\int_{-\infty}^{\infty} K(u)f(s+hu)du = (nh)^{-1}f(s) + o(1)
\end{aligned}
$$

であるから、$\mathbb{V}[\hat{f}(s)] \to 0$ である。ここから、チェビシェフの不等式: 期待値 μ と分散 σ^2 を持つ有限な確率変数 X について、任意の $c > 0$ に対して

$$
\mathbb{P}[|X-\mu| \geq c] \leq \frac{\sigma^2}{c}
$$

を適用すれば、$\hat{f}(s) \to^{\mathbb{P}} f(s)$ である。

一方で、同様の手順により、分子

$$
\hat{v}(s) \equiv (nh)^{-1}\sum_{i=1}^{n} K\left(\frac{S_i - s}{h}\right)^2 \mathbb{V}[]\epsilon|S_i]
$$

の期待値は、

$$
\begin{aligned}
\mathbb{E}[\hat{v}(s)] &= \int_{-\infty}^{\infty} h^{-1}K\left(\frac{v-s}{h}\right)^2 \sigma^2(v)f(v)dv \\
&= \int_{-\infty}^{\infty} K(u)^2\sigma^2(s+hu)f(s+hu)du \\
&= \int_{-\infty}^{\infty} K(u)^2\sigma^2(s)f(s)du + o(1) = \int_{-\infty}^{\infty} K(u)^2 du\sigma^2(s)f(s) + o(1)
\end{aligned}
$$

であり、分散は

$$
\begin{aligned}
\mathbb{V}[\hat{v}(s)] =& (nh)^{-1}h^{-1}\mathbb{V}\left[K\left(\frac{S-s}{h}\right)^2\sigma^2(S)\right] \\
\leq& (nh)^{-1}h^{-1}\int_{-\infty}^{\infty} K\left(\frac{v-s}{h}\right)^4 \sigma^4(v)f(v)dv \\
=& (nh)^{-1}\int_{-\infty}^{\infty} K(u)^4\sigma^4(s+hu)f(s+hu)du \\
\leq& (nh)^{-1}\bar{K}^2\int_{-\infty}^{\infty} K(u)^2 du\sigma^4(s)f(s) + o(1)
\end{aligned}
$$

である。したがって、$\hat{v}(s) \to^{\mathbb{P}} K^2\sigma^2(s)f(s)$ である。したがって $f(s) > 0$ から、連続写像定理を適用すれば、

$$
\frac{\hat{v}(s)}{\hat{f}(s)^2} \to^{\mathbb{P}} \frac{K^2\sigma^2(s)f(s)}{f(s)^2}
$$

である。よって、

$$
\sigma^2_{NH}(s) = (nh)^{-1}\frac{\hat{v}(s)}{\hat{f}(s)^2} = \frac{K^2\sigma^2(s)}{nhf(s)} + o_P(1/nh)
$$

を得る。

定理 2（Hansen, 2020, Theorem 19.1, 19.5, より）　仮定 1 および 2 のもとで、$nh \to \infty$ かつ $h \to 0$ であるときに、カーネル推定量のバイアスは

$$Bias^{NW}(0_+) = hABias^{NW}_{boundary}(0_+) + h^2 ABias^{NW}_{interior}(0_+) + o_P(h^2) + O_p(\sqrt{h/n})$$

である。ただし、

$$ABias^{NW}_{boundary}(0_+) \equiv 2K_1 m'(0_+),$$

かつ

$$ABias^{NW}_{interior}(0_+) \equiv 0.5K_2 m''(0_+) + f(0)^{-1} K_2 f'(0_+) m'(0_+)$$

である。ここで定数 K_1 および K_2 は、$K_1 \equiv \int_{\mathcal{K}} K(u)udu$ および $K_2 \equiv \int_{\mathcal{K}} K(u)u^2 du$ であり、$O_p(\sqrt{h/n})$ は $\sqrt{h/n}$ で割っても確率有界な項である。

■証明　ある $s \in \mathcal{S}$ について、

$$\mathbb{E}[\hat{m}_{NW}(s)|S] = \frac{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\mathbb{E}[Y_i|S_i]}{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)} = \frac{(1/nh)\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)m(S_i)}{(1/nh)\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)}$$

である。ここで、分母は定理 1の証明にも現れたカーネル密度推定量 $\hat{f}(s)$ であり、分子についても

$$\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)m(S_i) = \sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\} + \sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)m(s)$$

とすれば、

$$\begin{aligned}
\mathbb{E}[\hat{m}_{NW}(s)|S] =& m(s)\frac{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)}{\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)} + \frac{(1/nh)\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\}}{(1/nh)\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)} \\
=& m(s) + \frac{(1/nh)\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\}}{\hat{f}(s)}
\end{aligned}$$

を得る。すでに $\hat{f}(s) \to^{\mathbb{P}} f(s)$ であることは示されているから、以後は分子に注目する。ここで、分子の期待値は

$$\begin{aligned}
\mathbb{E}\left[(1/nh)\sum_{i=1}^n K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\}\right] =& (1/h)\mathbb{E}\left[K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\}\right] \\
=& (1/h)\int_{\mathcal{K}} K\left(\frac{v - s}{h}\right)(m(v) - m(s))f(v)dv \\
=& \int_{\mathcal{K}} K(u)(m(s + hu) - m(s))f(s + hu)du
\end{aligned}$$

である。条件付き期待値関数と密度関数の連続性より、以下のテイラー展開

$$\begin{aligned}
m(s + hu) - m(s) =& m'(s)hu + (1/2)m''(s)h^2u^2 + o(h^2) \\
f(s + hu) =& f(s) + f'(s)hu + o(h)
\end{aligned}$$

を用いれば、

$$\begin{aligned}
&\int_{\mathcal{K}} K(u)(m(s + hu) - m(s))f(s + hu)du \\
&= \int_{\mathcal{K}} K(u)(m'(s)hu + (1/2)m''(s)h^2u^2 + o(h^2))(f(s) + f'(s)hu + o(h))du \\
&= h\left(\int_{\mathcal{K}} uK(u)du\right)m'(s)(f(s) + o(h))
\end{aligned}$$

$$
\begin{aligned}
&+ h^2 \left(\int_{\mathcal{K}} u^2 K(u)du\right)(m'(s)f'(s) + (1/2)m''(s)f(s) \\
&+ h^3 \left(\int_{\mathcal{K}} u^3 K(u)du\right)(m''(s)f'(s)) + o(h^2).
\end{aligned}
$$

である。

ここで、s がサポートの内点であるならば、カーネル関数の対称性より、正の整数 r が奇数であるならば、

$$
\begin{aligned}
\int_{-\infty}^{\infty} u^r K(u)du &= \int_0^{\infty} u^r K(u)du + \int_{-\infty}^{0} u^r K(u)du \\
&= \int_0^{\infty} u^r K(u)du + \int_0^{\infty} (-u)^r K(-u)du \\
&= \int_0^{\infty} u^r K(u)du + (-1)^r \int_0^{\infty} u^r K(u)du = 0
\end{aligned}
$$

が成り立つ。しかしながら、s がサポートの端点であるならば、奇数モーメントは 0 とはならない。したがって、

$$
\begin{aligned}
&\mathbb{E}\left[(1/nh)\sum_{i=1}^{n} K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\}\right] \\
&= hK_1 m'(s)f(s) + h^2 K_2(m'(s)f'(s) + (1/2)m''(s)f(s)) + o(h^2)
\end{aligned}
$$

である。ここで、内点であれば $K_1 = 0$ であるため、h に比例する項が消えることになる。

同様に、分子の分散を求めると

$$
\begin{aligned}
&\mathbb{V}\left[(1/nh)\sum_{i=1}^{n} K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\}\right] \\
=&n^{-1}h^{-2}\mathbb{V}\left[K\left(\frac{S_i - s}{h}\right)\{m(S_i) - m(s)\})\right] \\
\leq&n^{-1}h^{-2}\mathbb{E}\left[K\left(\frac{S_i - s}{h}\right)^2\{m(S_i) - m(s)\}^2)\right] \\
=&(1/nh)(1/h)\int_{\mathcal{K}} K\left(\frac{v - s}{h}\right)^2 (m(v) - m(s))^2 f(v)dv \\
=&(1/nh)\int_{\mathcal{K}} K(u)^2 (m(s + hu) - m(s))^2 f(s + hu)du
\end{aligned}
$$

である。ここで、先のテイラー展開を用いれば

$$
\begin{aligned}
(m(s + hu) - m(s))^2 f(s + hu) =&((m'(s))^2 u^2 h^2 + o(h^2))(f(s) + f'(s)hu + o(h)) \\
=&m'(s)^2 u^2 f(s)h^2 + o(h^2)
\end{aligned}
$$

であるから、分子の分散は

$$
\begin{aligned}
&(1/nh)\int_{\mathcal{K}} K(u)^2 (m'(s)^2 u^2 f(s)h^2 + o(h^2))du \\
&\leq(1/nh)\bar{K}\int_{\mathcal{K}} K(u)(m'(s)^2 u^2 f(s)h^2 + o(h^2))du \\
&=(1/nh)\left(\bar{K}\int_{\mathcal{K}} u^2 K(u)\,du m'(s)^2 f(s)h^2 + o(h^2)\right) = \quad (h/n)\bar{K}K_2(m'(s))^2 f(s) + o(h/n) \leq O(h/n)
\end{aligned}
$$

である。したがって、

$$
\begin{aligned}
\mathbb{E}[\hat{m}_{NW}(s)|S] - m(s) =&hK_1 m'(s) \\
&+ h^2 K_2(m'(s)f'(s)/f(s) + (1/2)m''(s)) + o_p(h^2) + O_p(\sqrt{h/n})
\end{aligned}
$$

である。

定理 3（Hansen, 2020, Theorem 19.6, より） 仮定 1 および 2 のもとで、$nh \to \infty$ かつ $h \to 0$ であるときに、局所線形推定量のバイアスは

$$Bias^{LL}(0_+) = h^2 ABias^{LL}(0_+) + o_P(h^2) + O_p(\sqrt{h/n})$$

であり、分散は

$$\sigma^2_{LL}(0_+) = \frac{K^2_{LL}\mathbb{V}[\epsilon|S=0_+]}{f(0_+)nh} + o_P(1/nh)$$

である。ただし、

$$ABias^{LL}(0_+) \equiv 0.5m''(0_+)\frac{(K_2^2 - K_1K_3)}{K_2K_0 - K_1^2}$$

である。ここで定数 K_0 および K_3 はそれぞれ、$K_0 \equiv \int_{\mathcal{K}} K(u)du$ および $K_3 \equiv \int_{\mathcal{K}} K(u)u^3du$ である、$K^2_{LL} \equiv \frac{\int(K_2-uK_1)^2K^2(u)du}{(K_2K_0-K_1^2)^2}$ である。

■証明 (Imbens and Kalyanaraman, 2012, Lemma A1 より)

以下においては、$s=0$ とする。$K_h(S_i) = K\left(\frac{S_i-s}{h}\right)/h$ とし、$W \equiv diag(K_h(S_1),\ldots,K_h(S_n))$ とし、Y を Y_i のベクトルとする。ここで、R を $n\times 2$ の 1 と S_i のベクトルを並べた行列とすれば、局所線形推定量の端点推定は $e_1 = \begin{pmatrix}1 & 0\end{pmatrix}'$ を用いて

$$\hat{m}(0) = e_1'(R'WR)^{-1}R'WY$$

と書ける。ここから、バイアス

$$\mathbb{E}[\hat{m}(0)|S] - m(0) = e_1'(RWR)^{-1}R'WM - m(0)$$

、ただし、$M = \begin{pmatrix} m(S_1) \\ \vdots \\ m(S_n)\end{pmatrix}$ を考える。

ここでのの整数 r について、$F_r = (1/n)\sum_{i=1}^n K_h(S_i)S_i^r$ とすれば、

$$(1/n)R'WR = \begin{pmatrix} (1/n)\sum_{i=1}^n K_h(S_i) & (1/n)\sum_{i=1}^n K_h(S_i)S_i \\ (1/n)\sum_{i=1}^n K_h(S_i)S_i & (1/n)\sum_{i=1}^n K_h(S_i)S_i^2 \end{pmatrix} = \begin{pmatrix} F_0 & F_1 \\ F_1 & F_2\end{pmatrix}$$

と書ける。この期待値は、$v = uh$ の変数変換を行えば、密度関数のテイラー展開により、

$$\begin{aligned}
\mathbb{E}[F_r] &= (1/h)\int_{\mathcal{K}} K\left(\frac{v}{h}\right)v^r f(v)dv \\
&= \int_{\mathcal{K}} K(u)u^r h^r f(uh)du \\
&= h^r \int_{\mathcal{K}} K(u)u^r(f(0) + f'(0)hu + o(h))du \\
&= h^r \int_{\mathcal{K}} K(u)u^r du f(0) + o(h^{r+1}).
\end{aligned}$$

同様に分散は、

$$\begin{aligned}
\mathbb{V}[F_r] \leq & (1/n)\mathbb{E}[K_h(S_i)^2 S_i^{2r}] \\
=& (1/nh)(1/h)\int_{\mathcal{K}} K(v/h)^2 v^{2r} f(v)dv \\
=& (1/nh)\int_{\mathcal{K}} K(u)^2u^{2r}h^{2r}f(uh)du \leq O\left(\frac{h^{2r}}{nh}\right) = o((h^r)^2)
\end{aligned}$$

である。したがって、$F_r = h^r f(0)K^r + o_p(h^r)$、ただし、$K^r = \int_{\mathcal{K}} K(u)u^r du$ である。

ここから、

$$
\begin{aligned}
((1/n)R'WR)^{-1} =& \begin{pmatrix} F_0 & F_1 \\ F_1 & F_2 \end{pmatrix}^{-1} \\
=& \frac{1}{F_0F_2 - F_1^2} \begin{pmatrix} F_2 & -F_1 \\ -F_1 & F_0 \end{pmatrix} \\
=& \begin{pmatrix} \frac{f(0)K_2}{f(0)f(0)K_2 - f(0)^2K_1^2} + o_p(1) & -(1/h)\left(\frac{f(0)K_1}{f(0)f(0)K_2 - f(0)^2K_1^2} + o_p(1)\right) \\ -(1/h)\left(\frac{f(0)K_1}{f(0)f(0)K_2 - f(0)^2K_1^2} + o_p(1)\right) & (1/h^2)\left(\frac{f(0)}{f(0)f(0)K_2 - f(0)^2K_1^2} + o_p(1)\right) \end{pmatrix} \\
=& \begin{pmatrix} \frac{K_2}{(K_2 - K_1^2)f(0)} + o_p(1) & -(1/h)\left(\frac{K_1}{(K_2 - K_1^2)f(0)} + o_p(1)\right) \\ -(1/h)\left(\frac{K_1}{(K_2 - K_1^2)f(0)} + o_p(1)\right) & (1/h^2)\left(\frac{1}{(K_2 - K_1^2)f(0)} + o_p(1)\right) \end{pmatrix} \\
=& \begin{pmatrix} O_p(1) & O_p(1/h) \\ O_p(1/h) & O_p(1/h^2) \end{pmatrix}
\end{aligned}
$$

である。

このとき、テイラー展開より、

$$
R'WM = R'W \begin{pmatrix} m(0) + m'(0)S_1 + (1/2)m''(0)S_1^2 + r_1 \\ \vdots \\ m(0) + m'(0)S_n + (1/2)m''(0)S_n^2 + r_n \end{pmatrix}
$$

ただし剰余項 r_i は $|r_i| \leq \sup_s |m'''(s)||S_i^3|$ を満たす。したがって、

$$
\begin{aligned}
R'WM \leq& R'WR \begin{pmatrix} m(0) \\ m'(0) \end{pmatrix} + R'W(1/2)m''(0) \begin{pmatrix} F_2 \\ F_3 \end{pmatrix} + o_p(h^2) \\
=& R'WR \begin{pmatrix} m(0) \\ m'(0) \end{pmatrix} + R'W(1/2)m''(0) \begin{pmatrix} h^2K_2f(0) + o_p(h^2) \\ h^3K_3f(0) + o_p(h^3) \end{pmatrix} + o_p(h^2)
\end{aligned}
$$

である。以上より、

$$
\begin{aligned}
e_1'(R'WR)^{-1}R'WM - m(0) =& e_1'(R'WR)^{-1}R'W(1/2)\left(m''(0)\begin{pmatrix} h^2K_2f(0) + o_p(h^2) \\ h^3K_3f(0) + o_p(h^3) \end{pmatrix}\right) + o_p(h^2) \\
=& (1/2)m''(0)\frac{K_2^2 - K_1K_3}{K_2 - K_1^2}h^2 + o_p(h^2).
\end{aligned}
$$

を得る。

同様に、分散を考えると、

$$
\mathbb{V}[\hat{m}(0)|S] = e_1'(R'WR)^{-1}R'W\Sigma WR(R'WR)^{-1}e_1
$$

ただし、$\Sigma = diag(\sigma^2(S_1), \ldots, \sigma^2(S_n))$ である。このとき、

$$
(1/n)R'W\Sigma WR = \begin{pmatrix} (1/n)\sum_{i=1}^n K_h^2(S_i)\sigma^2(S_i) & (1/n)\sum_{i=1}^n K_h^2(S_i)\sigma^2(S_i)S_i \\ (1/n)\sum_{i=1}^n K_h^2(S_i)\sigma^2(S_i)S_i & (1/n)\sum_{i=1}^n K_h^2(S_i)\sigma^2(S_i)S_i^2 \end{pmatrix}
$$

であり、$G_r = (1/n)\sum_{i=1}^n K_h^2(S_i)\sigma^2(S_i)S_i^r$ とすれば、F_r と同様に

$$
G_r = h^{r-1}f(0)K_r + o_p(h^{r-1}),
$$

ただし $K_r \equiv \int_{\mathcal{K}} u^rK^2(u)du$ であるから、

$$
\begin{aligned}
nV =& \frac{1}{(F_0F_2 - F_1^2)^2}e_1' \begin{pmatrix} F_2 & -F_1 \\ -F_1 & F_0 \end{pmatrix} \begin{pmatrix} G_0 & G_1 \\ G_1 & G_2 \end{pmatrix} \begin{pmatrix} F_2 & -F_1 \\ -F_1 & F_0 \end{pmatrix} e_1 \\
=& \frac{F_2^2G_0 - 2F_1F_2G_1 + F_1^2G_2}{(F_0F_2 - F_1^2)^2}
\end{aligned}
$$

を得る。したがって、

$$
V = \frac{\sigma^2(0)}{f(0)nh}\left(\frac{K_2^2K_0 - 2K_1K_2K_1 + K_1^2K_2}{(K_0K_2 - K_1^2)^2}\right) + o_p(1/nh)
$$

を得る。

定理 4（Imbens and Kalyanaraman, 2012, Lemma 3.1）　仮定 1, 2 のもとで、局所線形推定による切片項推定量の漸近 2 乗誤差は

$$2^{-2}\left(\frac{\sigma^2(0)}{f(0)n}\right)^{4/5} m''(0)^{2/5}\left(T(K)\right)^{2/5}$$

で最小化される。このとき、

$$T(K) = (K_{LL}^2)^2\left|\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right|$$

はカーネルのみに依存する項である。

■証明　定理 3より、局所線形推定量の漸近 2 乗誤差は

$$(1/4)h^4(m''(0))^2\left(\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right)^2 + \frac{K_{LL}^2\sigma^2(0)}{f(0)nh}$$

である。1 階条件は

$$h^3(m''(0))^2\left(\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right)^2 - \frac{K_{LL}^2\sigma^2(0)}{f(0)nh^2} = 0$$

であり、

$$h = \left(m''(0)^{-2}\left(\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right)^{-2} K_{LL}^2\frac{\sigma^2(0)}{f(0)n}\right)^{1/5}$$

である。これを漸近 2 乗誤差に代入すると

$$\left(2^{-2}\left(K_{LL}^2\frac{\sigma^2(0)}{f(0)n}\right)^{4/5} m''(0)^{2/5}\left(\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right)^{2/5}\right) + \frac{(K_{LL}^2)^{4/5}m''(0)^{2/5}\left(\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right)^{2/5}\sigma^2(0)^{4/5}}{f(0)^{4/5}n^{4/5}}$$

であるから、漸近 2 乗誤差は

$$2^{-2}\left(\frac{\sigma^2(0)}{f(0)n}\right)^{4/5} m''(0)^{2/5}\left((K_{LL}^2)^2\frac{K_2^2 - K_1K_3}{K_2K_0 - K_1^2}\right)^{2/5}$$

である。

このとき、

仮定 3　カーネル関数 $K(\cdot)$ は有界なサポートをもち、リプシッツ連続である。すなわち、ある定数 $\bar{K}$ が存在し、$K(\cdot)$ のサポート上の任意の点 s, s' について

$$|K(s) - K(s')| \leq \bar{K}|s - s'|$$

である。

仮定 3 のもとで以下の定理が示されている。

定理 5（Cheng, Fan, and Marron, 1997, Theorem 2）　仮定 1, 2, 3 を考える。仮定 1と 3を満たす任意のカーネル関数の中で、$T(K)$ を最小化するカーネルは三角カーネル $K(u) = (1 - u)\mathbf{1}\{u \in [0, 1]\}$ である。

第 9 章　差の差法の基礎

T9.1

命題 1　(**Abadie et al., 2010, Appendix B**) 潜在結果関数について、以下のモデルが成り立つとする：

$$Y^*_{it}(0) = \delta_t + \theta_t Z_i + \boldsymbol{\lambda}'_t \boldsymbol{\mu}_i + \epsilon_{it}$$

ただし、$\boldsymbol{\lambda}_t = (\lambda_{t1}, \ldots, \lambda_{tF})$ は $(1 \times F)$ 次元の共通ファクターベクトル、$\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{iF})'$ は $(F \times 1)$ 次元のファクターローディングベクトルである。このとき、ϵ_{it} が時間と個体について i.i.d であり、$\mathbb{E}[\epsilon_{it}|\{Z_i, \boldsymbol{\mu}_i\}_{i=1}^J] = 0$ であるとする。さらに、ある $\bar{\xi}$ が存在し、任意の整数 $M, 1 \leq M \leq T_0$ について $M^{-1} \sum_{t=T_0-M+1}^{T_0} \boldsymbol{\lambda}'_t \boldsymbol{\lambda}_t$ の最小固有値 $\xi(M)$ が $\xi(M) > \underline{\xi} > 0$ を満たし、ある定数 $\bar{\lambda}$ について $|\lambda_{tf}| \leq \bar{\lambda}$ を満たすとする。さらに $\sum_{t=1}^{T_0} \boldsymbol{\lambda}'_t \boldsymbol{\lambda}_t$ が非特異であるとする。

　このとき、

$$\begin{aligned} \sum_{i=1}^J w_i^* Y^*_{it}(0) &= Y^*_{0t}(0), \forall t \in \{1, \ldots, T_0\} \\ \sum_{i=1}^J w_i^* Z_i &= Z_0 \end{aligned}$$

を満たすような w^* が存在するならば、

$$Y^*_{0t}(0) - \sum_{i=1}^J w_i^* Y^*_{it}(0) = R_{1t} + R_{2t} + R_{3t}$$

ただし、

$$\begin{aligned} R_{1t'} &= \boldsymbol{\lambda}'_t \left(\sum_t \boldsymbol{\lambda}'_t \boldsymbol{\lambda}_t \right)^{-1} \sum_t \boldsymbol{\lambda}_t \sum_{i=1}^J w_i^* \epsilon_{it} \\ R_{2t'} &= -\boldsymbol{\lambda}'_t \left(\sum_t \boldsymbol{\lambda}'_t \boldsymbol{\lambda}_t \right)^{-1} \sum_t \boldsymbol{\lambda}_t \epsilon_{0t} \\ R_{3t'} &= \sum_{i=1}^J w_i^* (\epsilon_{it'} - \epsilon_{0t'}) \end{aligned}$$

について、$\mathbb{E}[R_{2t}] = \mathbb{E}[R_{3t}] = 0$ かつ

$$\mathbb{E}[|R_{1t}|] \leq C(p)^{1/p} \left(\frac{\bar{\lambda}^2 F}{\underline{\xi}} \right) J^{1/p} \max \left\{ \frac{\bar{m}_p^{1/p}}{T_0^{1-1/p}}, \frac{\bar{\sigma}}{T_0^{1/2}} \right\}$$

である。ここで、

$$\bar{m} = \max_{i \in \{1, \ldots, J\}} (1/T_0) \sum_{t=1}^{T_0} \mathbb{E}[|\epsilon_{it}|^p]$$

$$\bar{\sigma} = \max_{i \in \{1,\ldots,J\}} T_0^{-1} \sum_t \sqrt{\mathbb{E}[|\epsilon_{jt}|^2]}$$

である。

■**証明**　$t > T_0$ における予測誤差は

$$\begin{aligned}
&Y_{0t}^*(0) - \sum_{i=1}^J w_i^* Y_{it}^*(0) \\
&= \theta_t \left(Z_0 - \sum_{i=1}^J w_i^* Z_j \right) \\
&\quad + \boldsymbol{\lambda}_t \left(\boldsymbol{\mu}_0 - \sum_{i=1}^J w_i^* \boldsymbol{\mu}_i \right) + \sum_{i=1}^J w_i^* (\epsilon_{0t} - \epsilon_{it})
\end{aligned}$$

と分解できるが、同様に個体 i の統制期間の顕在結果ベクトル $\mathbf{Y}_i = [Y_{i1}, \ldots, Y_{iT_0}]'$ について、

$$\begin{aligned}
&\mathbf{Y}_0^*(0) - \sum_{i=1}^J w_i^* \mathbf{Y}_i^*(0) \\
&= \boldsymbol{\theta} \left(Z_0 - \sum_{i=1}^J w_i^* Z_j \right) \\
&\quad + \boldsymbol{\lambda} \left(\boldsymbol{\mu}_0 - \sum_{i=1}^J w_i^* \boldsymbol{\mu}_i \right) + \sum_{i=1}^J w_i^* (\boldsymbol{\epsilon}_0 - \boldsymbol{\epsilon}_i)
\end{aligned}$$

を得る。

$$\boldsymbol{\lambda}'\boldsymbol{\lambda} = \sum_t \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t$$

は非特異であると仮定したので、上記の統制期間における関係は非観測ベクトル $\boldsymbol{\mu}_0 - \sum_{i=1}^J w_i^* \boldsymbol{\mu}_i$ について解くことができ、

$$\begin{aligned}
&Y_{0t}^*(0) - \sum_{i=1}^J w_i^* Y_{it}^*(0) \\
&= (\boldsymbol{\lambda}_t (\boldsymbol{\lambda}'\boldsymbol{\lambda})^{-1} \boldsymbol{\lambda}')(\mathbf{Y}_0^*(0) - \sum_{i=1}^J w_i^* \mathbf{Y}_i^*(0)) \\
&\quad + (\theta_t - \boldsymbol{\lambda}(\boldsymbol{\lambda}'\boldsymbol{\lambda})^{-1}\boldsymbol{\lambda}'\boldsymbol{\theta}) \left(Z_0 - \sum_{i=1}^J w_i^* Z_j \right) \\
&\quad - \boldsymbol{\lambda}_t (\boldsymbol{\lambda}'\boldsymbol{\lambda})^{-1} \boldsymbol{\lambda}' \sum_{i=1}^J w_i^* (\boldsymbol{\epsilon}_0 - \boldsymbol{\epsilon}_i) \\
&\quad + \sum_{i=1}^J w_i^* (\epsilon_{0t} - \epsilon_{it})
\end{aligned}$$

とできる。ここで、w^* が

$$\begin{aligned}
\sum_{i=1}^J w_i^* Y_{it}^*(0) &= Y_{0t}^*(0), \forall t \in \{1, \ldots, T_0\} \\
\sum_{i=1}^J w_i^* Z_i &= Z_0
\end{aligned}$$

を満たしていることを用いれば、予測誤差は R_{1t}, R_{2t}, R_{3t} に分解できる。このとき、$t > T_0$ において、w_i^* を含んでいない R_{2t} は期待値 0 であり、R_{3t} は w_i^* に依存しているものの $t > T_0$ の ϵ_{it} には依存していないた

め、同様に期待値 0 である。一方で、R_{1t} は w_i^* およびその荷重が決定される統制期間の ϵ_{it} に依存しており、必ずしも期待値 0 ではない。

一方で、

$$\boldsymbol{\lambda}_t \left(\sum_t \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t \right)^{-1} \boldsymbol{\lambda}_t'$$

が対称かつ正定値であるから、Cauchy-Schwartz 不等式より、

$$\begin{aligned}
&\left(\boldsymbol{\lambda}_s \left(\sum_t \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t \right)^{-1} \boldsymbol{\lambda}_t' \right)^2 \\
&\leq \left(\boldsymbol{\lambda}_s \left(\sum_t \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t \right)^{-1} \boldsymbol{\lambda}_s' \right) \left(\boldsymbol{\lambda}_t \left(\sum_t \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t \right)^{-1} \boldsymbol{\lambda}_t' \right) \\
&\leq \left(\frac{\bar{\lambda}^2 F}{T_0 \underline{\xi}} \right)
\end{aligned}$$

である。ここで、

$$\bar{\boldsymbol{\epsilon}}_i^L = \sum_s \left(\sum_t \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t \right)^{-1} \boldsymbol{\lambda}_s' \epsilon_{is}$$

とすれば、ある偶数 p について $|\epsilon_{it}|^p$ が $i \in \{1, \ldots, J\}, t \in \{1, \ldots, T_0\}$ について存在する時、Hölder の不等式より

$$\mathbb{E}\left[\sum_{i=1}^J w_i^* |\bar{\boldsymbol{\epsilon}}_i^L| \right] \leq \left(\mathbb{E}\left[\sum_{i=1}^J |\bar{\boldsymbol{\epsilon}}_i^L|^p \right] \right)^{1/p}$$

を得る。さらに、Rosenthal の不等式を用いれば、Ibragimov and Sharakhmetov (2002) に示される定数 $C(p)$ について、

$$\mathbb{E}[|\bar{\boldsymbol{\epsilon}}_i^L|^p] \leq C(p) \left(\frac{\bar{\lambda}^2 F}{T_0 \underline{\xi}} \right)^p \max \left\{ \frac{1}{T_0^p} \sum_{t=1}^{T_0} \mathbb{E}[|\epsilon_{it}|^p], \left(\frac{1}{T_0^2} \sum_{t=1}^{T_0} \mathbb{E}[|\epsilon_{it}|^2] \right)^{p/2} \right\}$$

ここで、

$$\bar{m} = \max_{i \in \{1, \ldots, J\}} (1/T_0) \sum_{t=1}^{T_0} \mathbb{E}[|\epsilon_{it}|^p], \bar{\sigma} = \sqrt{\max_{i \in \{1, \ldots, J\}} (1/T_0) \sum_{t=1}^{T_0} \mathbb{E}[|\epsilon_{it}|^2]}$$

とすれば、

$$\mathbb{E}[|R_{1t}|] \leq C(p)^{1/p} \left(\frac{\bar{\lambda}^2 F}{T_0 \underline{\xi}} \right) J^{1/p} \max \left\{ \frac{\bar{m}^{1/p}}{T_0^{1-1/p}}, \frac{\bar{\sigma}}{T_0^{1/2}} \right\}$$

を得た。

T10.1

定理 6　(Goodman-Bacon, 2021, **Theorem 1**) N 人 T 期の観測について、観測期間を通じて処置 Z を受け取らない統制群 U と、$k=1,\ldots,K$ について、$k\in(1,T]$ 期に処置 Z を受ける K 種の処置群に分割できるとする。ここで、t_i を i が処置を受ける期とし、i が統制群 U ならば $t_i=U$ とする。さらに、

$$\bar{y}_b^{POST(a)}=\frac{1}{T-(a-1)}\sum_{t=a}^{T}\left[\frac{\sum_i Y_{it}1\{t_i=b\}}{\sum_i 1\{t_i=b\}}\right],$$

は b 期に処置を受ける群の a 期以後の顕在結果平均、

$$\bar{y}_b^{PRE(a)}=\frac{1}{a-1}\sum_{t=1}^{a-1}\left[\frac{\sum_i Y_{it}1t_i=b}{\sum_i 1t_i=b}\right],$$

は b 期に処置を受ける群の a 期より前の顕在結果平均、

$$\bar{y}_b^{MID(a,b)}=\frac{1}{b-(a-1)}\sum_{t=a}^{b-1}\left[\frac{\sum_i Y_{it}1t_i=b}{\sum_i 1t_i=b}\right],$$

は b 期に処置を受ける群の a 期から b 期までの顕在結果平均、

$$\bar{D}_k=T^{-1}\sum_{t=1}^{T}1t\geq k,$$

$n_k=N^{-1}\sum_{i=1}^{N}1t_i=k$, $n_{ab}=\frac{n_a}{n_a+n_b}$ とする。

このとき、最小 2 乗法による以下の TWFE 回帰

$$Y_{it}=\alpha_{i\cdot}+\alpha_{\cdot t}+\beta Z_{it}+\epsilon_{it}$$

における β の推定量 $\hat{\beta}$ について、

$$\hat{\beta}=\sum_{k\neq U}s_{kU}\hat{\beta}_{kU}^{2\times2}+\sum_{k\neq U}\sum_{l>k}[s_{kl}^{k}\beta_{kl}^{2\times2,k}+s_{kl}^{l}\hat{\beta}_{kl}^{2\times2,l}]$$

を得る。ただし、

$$\begin{aligned}\hat{\beta}_{kU}^{2\times2}&=(\bar{y}_k^{POST}(k)-\bar{y}_k^{PRE}(k))-(\bar{y}_U^{POST}(k)-\bar{y}_U^{PRE}(k)),\\\hat{\beta}_{kl}^{2\times2,k}&=(\bar{y}_k^{MID}(k,l)-\bar{y}_k^{PRE}(k))-(\bar{y}_l^{MID}(k,l)-\bar{y}_l^{PRE}(k)),\\\hat{\beta}_{kl}^{2\times2,l}&=(\bar{y}_l^{POST}(l)-\bar{y}_l^{MID}(k,l))-(\bar{y}_k^{POST}(l)-\bar{y}_k^{MID}(k,l))\end{aligned}$$

であり、

$$s_{kU}=\frac{(n_k+n_U)^2n_{kU}(1-n_{kU})\bar{D}_k(1-\bar{D}_k)}{\hat{V}^D},$$

$$s_{kl}^k = \frac{((n_k+n_l)(1-\bar{D}_l^2)n_{kl}(1-n_{kl})\frac{\bar{D}_k-\bar{D}_l}{1-\bar{D}_l}frac1-\bar{D}_k1-\bar{D}_l}{\hat{V}^D},$$

$$s_{kl}^l = \frac{((n_k+n_l)\bar{D}_k)^2 n_{kl}(1-n_{kl})\frac{\bar{D}_l}{\bar{D}_k}\frac{\bar{D}_k-\bar{D}_l}{\bar{D}_k}}{\hat{V}^D}$$

である。 このとき、$\sum_{k\neq U} s_{kU} + \sum_{k\neq U}\sum_{l>k}[s_{kl}^k + s_{kl}^l] = 1$ を満たす。

■証明　Frisch–Waugh–Lovell 定理より、$\hat{\beta}$ は個体・時間について平均差分を取った $\tilde{D}_{it} = D_{it} - \bar{D}_i - (\bar{D}_t - \bar{\bar{D}})$ についての単回帰推定量

$$\frac{(1/nT)\sum_i\sum_t Y_{it}\tilde{D}_{it}}{(1/nT)\sum_i\sum_t \tilde{D}_{it}^2}$$

に等しい。ここで、$k(i)$ を i が属するグループの処置時点とすると、このときの分子は $(\bar{D}_{k(i)t} - \bar{D}_{k(i)})$ の項を足し引きすることで、

$$(1/nT)\sum_i\sum_t Y_{it}\tilde{D}_{it} = (1/nT)\sum_i\sum_t Y_{it}[(D_{it}-\bar{D}_i) - (\bar{D}_{k(i)t}-\bar{D}_{k(i)}) + (\bar{D}_{k(i)t}-\bar{D}_{k(i)}) - (\bar{D}_t - \bar{\bar{D}}]$$

となる。ここで、$(D_{it}-\bar{D}_i) - (\bar{D}_{k(i)t}-\bar{D}_{k(i)}) = 0$ であることに注意すれば、

$$(1/nT)\sum_i\sum_t Y_{it}\tilde{D}_{it} = (1/nT)\sum_i\sum_t Y_{it}[(\bar{D}_{k(i)t}-\bar{D}_{k(i)}) - (\bar{D}_t - \bar{\bar{D}}]$$

である。したがって、個体についての和は処置期間グループごとの和に置き換えられるため、以後 $k(i)$ を k と表す。このとき、

$$\sum_k n_k(1/T)\sum_t \bar{Y}_{kt}[(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_t - \bar{\bar{D}})]$$

を観察すると、

$$\begin{aligned}(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_t - \bar{\bar{D}}) =& (\bar{D}_{kt}-\bar{D}_k) - \sum_l(\bar{D}_{lt}-\bar{D}_l)\\ =& (1-n_k)(\bar{D}_{kt}-\bar{D}_k) - \sum_{l\neq k}(\bar{D}_{lt}-\bar{D}_l)\end{aligned}$$

である。ここで、$1-n_k = \sum_{l\neq k} n_l$ であるから、

$$(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_t - \bar{\bar{D}}) = \sum_{l\neq k} n_l[(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_{lt}-\bar{D}_l)]$$

である。したがって、

$$\begin{aligned}&\sum_k n_k(1/T)\sum_t \bar{Y}_{kt}[(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_t - \bar{\bar{D}})]\\ =& \sum_k\sum_{l\neq k} n_l n_k(1/T)\sum_t \bar{Y}_{kt}[(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_{lt}-\bar{D}_l)]\\ =& \sum_k\sum_{l>k} n_l n_k(1/T)\sum_t (\bar{Y}_{kt}-\bar{Y}_{lt})[(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_{lt}-\bar{D}_l)]\end{aligned}$$

である。ただし、最後の等式では、l についての和では平均差の表現の符号が反転することを利用している。

　ここでグループごとに上の表現 $\sum_k\sum_{l>k} n_l n_k(1/T)\sum_t(\bar{Y}_{kt}-\bar{Y}_{lt})[(\bar{D}_{kt}-\bar{D}_k) - (\bar{D}_{lt}-\bar{D}_l)]$ を分解する。統制群については $\bar{D}_{Ut} - \bar{D}_U = 0$ であり、

$$\begin{aligned}&-(1/T)\sum_{t<k}(\bar{Y}_{kt}-\bar{Y}_{Ut})\bar{D}_k + (1/T)\sum_{t\geq k}(\bar{Y}_{kt}-\bar{Y}_{Ut})(1-\bar{D}_k)\\ &= \left[(\bar{Y}_{kt}^{POST(k)} - \bar{Y}_{kt}^{PRE(k)}) - (\bar{Y}_{Ut}^{POST(k)} - \bar{Y}_{Ut}^{PRE(k)})\right]\bar{D}_k(1-\bar{D}_k)\end{aligned}$$

$$=\hat{\beta}^{2\times 2}_{kU}\bar{D}_k(1-\bar{D}_k)$$

である。処置群について、$k<l<T$ として考えると、

$$\begin{aligned}
&-(1/T)\sum_{t<k}(\bar{Y}_{kt}-\bar{Y}_{lt})(\bar{D}_k-\bar{D}_l)+(1/T)\sum_{t\in[k,l)}(\bar{Y}_{kt}-\bar{Y}_{lt})(1-\bar{D}_k+\bar{D}_l)\\
&\quad-(1/T)\sum_{t\geq l}(\bar{Y}_{kt}-\bar{Y}_{lt})(\bar{D}_k-\bar{D}_l)\\
&\quad=-(\bar{Y}^{PRE(k)}_{kt}-\bar{Y}^{PRE(k)}_{lt})(\bar{D}_k-\bar{D}_l)(1-\bar{D}_l)+(\bar{Y}^{MID(k,l)}_{kt}-\bar{Y}^{MID(k,l)}_{lt})(\bar{D}_k-\bar{D}_l)(1-\bar{D}_k+\bar{D}_l)\\
&\quad-(\bar{Y}^{POST(l)}_{kt}-\bar{Y}^{POST(l)}_{lt}(\bar{D}_l(\bar{D}_k-\bar{D}_l))\\
&\quad=\left[(\bar{Y}^{MID(k,l)}_{kt}-\bar{Y}^{PRE(k)}_{kt})-(\bar{Y}^{MID(k,l)}_{kt}-\bar{Y}^{PRE(k)}_{lt})\right](\bar{D}_k-\bar{D}_l)(1-\bar{D}_k)\\
&\quad+\left[(\bar{Y}^{POST(l)}_{lt}-\bar{Y}^{MID(k,l)}_{lt})-(\bar{Y}^{POST(l)}_{kt}-\bar{Y}^{MID(k,l)}_{kt})\right]\bar{D}_l(\bar{D}_k-\bar{D}_l)\\
&\quad=(1-\bar{D}_l)^2\hat{\beta}^{2\times 2,k}_{kl}\left(\frac{\bar{D}_k-\bar{D}_l}{1-\bar{D}_l}\right)\left(\frac{1-\bar{D}_k}{1-\bar{D}_l}\right)+\bar{D}_k^2\hat{\beta}^{2\times 2,l}_{kl}\left(\frac{\bar{D}_l}{\bar{D}_k}\right)\left(\frac{\bar{D}_k-\bar{D}_l}{\bar{D}_k}\right)
\end{aligned}$$

となる。結果として、

$$\begin{aligned}
&\sum_k\sum_{l>k}n_ln_k(1/T)\sum_t(\bar{Y}_{kt}-\bar{Y}_{lt})[(\bar{D}_{kt}-\bar{D}_k)-(\bar{D}_{lt}-\bar{D}_l)]\\
&=\sum_{k\neq U}(n_k+n_U)^2n_{kU}(1-n_{kU})\bar{D}_k(1-\bar{D}_k)\hat{\beta}^{2\times 2}_{kU}\\
&\quad+\sum_{k\neq U}\sum_{l>k}((n_k+n_l)(1-\bar{D}_l))^2n_{kl}(1-n_{kl})\left(\frac{\bar{D}_k-\bar{D}_l}{1-\bar{D}_l}\right)\left(\frac{1-\bar{D}_k}{1-\bar{D}_l}\right)\hat{\beta}^{2\times 2,k}_{kl}\\
&\quad+\sum_{k\neq U}\sum_{l>k}((n_k+n_l)\bar{D}_k)^2n_{kl}(1-n_{kl})\left(\frac{\bar{D}_l}{\bar{D}_k}\right)\left(\frac{\bar{D}_k-\bar{D}_l}{\bar{D}_k}\right)\hat{\beta}^{2\times 2,l}_{kl}
\end{aligned}$$

を得る。同様の操作を分母についても行えば、2×2 係数の荷重の和に等しくなることがわかり、したがって係数の荷重和は 1 となる。

T10.2

命題 2 (Sant'Anna and Zhao, 2020, Theorem 1 に基づく) $m_{k,t}(W_i)=E[Y_{it}-Y_{it_k-1}-1\mid W_i,G_i^0=1]$ とする。以下の 2 重頑健表現

$$E[((G_i^k)/E[G_i^k]-((p_k(W_i)G_i^0)/(1-p_k(W_i)))/E[(p_k(W_i)G_i^0)/(1-p_k(W_i))])(Y_{it}-Y_{it_k-1}-m_{k,t}(W_i))]$$

を $m_{k,t}(W_i)$ と $p_k(W_i)$ の推定量、$\hat{m}_{k,t}(W_i)$ および $\hat{p}_k(W_i)$ で置き換えた 2 重頑健推定量は、(i) $\hat{m}_{k,t}(W_i)\to^{\mathbb{P}} m_{k,t}(W_i)$ であるが $\hat{p}_k(W_i)\to^{\mathbb{P}} p^*(W_i)\neq p_k(W_i)$ である、(ii) $\hat{p}_k(W_i)\to^{\mathbb{P}} p_k(W_i)$ であるが $\hat{m}_{k,t}(W_i)\to^{\mathbb{P}} m^*(W_i)\neq m_{k,t}(W_i)$ であるとき、のいずれにおいても

$$ATT_t^k=\mathbb{E}[\mathbb{E}[Y_{it}-Y_{it_k-1}\mid W_i,G_i^k=1]-\mathbb{E}[Y_{it}-Y_{it_k-1}\mid W_i,G_i^0=1]\mid G_i^k=1]$$

に収束する。

■証明 (i) のケース：

傾向スコア $p_k(W_i)$ の推定量が一致性を持つとき、$\hat{m}_{k,t}(W_i)$ の項は、推定量のある収束先 $m^*(W_i)$ について、

$$\mathbb{E}\left[\left(\frac{(G_i^k)}{\mathbb{E}[G_i^k]}-\frac{\frac{(p_k(W_i)G_i^0)}{(1-p_k(W_i))}}{\mathbb{E}[(p_k(W_i)G_i^0)/(1-p_k(W_i))]}\right)m^*(W_i)\right]=\mathbb{E}\left[\left(\mathbb{E}\frac{(G_i^k)}{E[G_i^k]}-\frac{\mathbb{E}[G_i^k|W_i]}{\mathbb{E}[\mathbb{E}[G_i^k\mid W_i]]}\right)m^*(W_i)\right]=0$$

となる。したがって、2 重頑健のために追加された項が 0 に収束するため、推定量は一致性を持つ。

(ii) のケース:

一方で、$\hat{m}_{k,t}(W_i)$ が $m_{k,t}(W_i)$ の一致推定量であるが、$\hat{p}_k(W_i)$ がある $p^*(W_i) \neq p_k(W_i)$ に収束する時、推定量の収束先は

$$\mathbb{E}\left[\left(\frac{(G_i^k)}{\mathbb{E}[G_i^k]} - \frac{\frac{(p^*(W_i)G_i^0)}{(1-p^*(W_i))}}{\mathbb{E}[(p^*(W_i)G_i^0)/(1-p^*(W_i))}]\right)(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))\right]$$

である。ここで

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{(G_i^k)}{\mathbb{E}[G_i^k]}\right)(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))\right] \\
&= \mathbb{E}\left[\left(\frac{(G_i^k)}{\mathbb{E}[G_i^k]}\right)(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))|G_i^k = 1\right]\mathbb{E}[G_i^k] \\
&= \mathbb{E}\left[\mathbb{E}[((G_i^k)/E[G_i^k])(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i)) \mid G_i^k = 1, W_i]G_i^k = 1\right]\mathbb{E}[G_i^k]
\end{aligned}$$

であり、これは ATT_t^k である。

一方で、

$$\mathbb{E}\left[\left(\frac{((p^*(W_i)G_i^0)/(1-p^*(W_i)))}{\mathbb{E}[(p^*(W_i)G_i^0)/(1-p^*(W_i))]}\right)(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))\right]$$

に着目すると、期待値の中身は $Y_{it} - Y_{it_k-1}$ を除いて $G_i^0 = 1$ と W_i で条件付ければ定数となる表現である。したがって、

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{((p^*(W_i)G_i^0)/(1-p^*(W_i)))}{\mathbb{E}[(p^*(W_i)G_i^0)/(1-p^*(W_i))]}\right)(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))\right] \\
&= \mathbb{E}\left[\left(\frac{((p^*(W_i))/(1-p^*(W_i)))}{\mathbb{E}[(p^*(W_i)G_i^0)/(1-p^*(W_i))]}\right)G_i^0(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))|G_i^0 = 1\right]P(G_i^0 = 1) \\
&= \mathbb{E}\left[\left(\frac{((p^*(W_i))/(1-p^*(W_i)))}{\mathbb{E}[(p^*(W_i)G_i^0)/(1-p^*(W_i))]}\right)(\mathbb{E}[Y_{it} - Y_{it_k-1}|W_i, G_i^0 = 1] - m_{k,t}(W_i))|G_i^0 = 1\right]P(G_i^0 = 1)
\end{aligned}$$

を得る。ここで、$\mathbb{E}[Y_{it} - Y_{it_k-1} \mid W_i, G_i^0 = 1] = m_{k,t}(W_i)$ であったから、

$$\mathbb{E}\left[\left(\frac{((p^*(W_i)G_i^0)/(1-p^*(W_i)))}{\mathbb{E}[(p^*(W_i)G_i^0)/(1-p^*(W_i))]}\right)(Y_{it} - Y_{it_k-1} - m_{k,t}(W_i))\right] = 0$$

が得られた。

参考文献

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105 (490), 493–505.

Goodman-Bacon, Andrew (2021) "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225 (2), 254–277.

Sant'Anna, Pedro H. C. and Jun Zhao (2020) "Doubly Robust Difference-in-Differences Estimators," *Journal of Econometrics*, 219 (1), 101–122.