Published in final edited form as:

J Am Stat Assoc. 2016 ; 111(516): 1526–1530. doi:10.1080/01621459.2016.1242427.

Comment

Alexander R. Luedtke^a, Mark J. van der Laan^b

^aFred Hutchinson Cancer Research Center, Vaccine and Infectious Disease Division, Seattle, WA, USA

^bDivision of Biostatistics, UC Berkeley, Berkeley, CA, USA

1. Introduction

We commend Chen, Zeng, and Kosorok (CZK) for their impressive extension of outcome weighted learning (OWL) to the continuous dose setting. Their work not only shows that regret-consistent estimation is possible in this setting, but provides a robust and computationally feasible estimator to accomplish this task.

The authors prove a rate of convergence on the regret of their dose rule estimation procedure. In particular, they show that

$$\mathcal{R}(\hat{f}_n) \equiv \sup_f \mathcal{V}(f) - \mathcal{V}(\hat{f}_n) = O_P\left(n^{-\frac{1}{4 + 3d/\alpha}}\right),$$

where d is the dimension of X and α measures the underlying smoothness of the problem. If α is very large, then the above recovers a nearly $n^{-1/4}$ rate of convergence. At the end of Section 4, CZK argue that the rate of their method cannot be close to n^{-1} , the rate proved by Zhao et al. (2012) for OWL in the binary case under a margin assumption. In the binary treatment case, the fast rate n^{-1} can, under a margin assumption, be improved upon using a different class of estimators that directly estimates $E[R|A = 1, X] - E[R|A = 0, X]$ and estimates the optimal treatment rule at X by checking its sign. This result, presented in Theorem 8 of Luedtke and van der Laan (2016a), is an extension of the fast classification rate presented in Audibert and Tsybakov (2007).

In this commentary, we discuss the rates we believe to be (nearly) attainable by OWL-type methods and plug-in methods under smoothness assumptions. We make the following points:

-
Section 2:	We believe that OWL and similar methods that aim to directly optimize the value of an estimated rule can attain nearly $n^{-1/2}$ rates.
Section 3:	Plug-in methods can also attain nearly $n^{-1/2}$ rates.
Section 4:	Under interior maxima and concavity assumptions, plug-in methods can attain nearly n^{-1} rates.
-

When we started writing this commentary, we expected that the fast rate in Section 4 would not be attainable by direct value-based methods. We are now not clear on the matter, and look forward to reading future works by others (perhaps by CZK) to clarify this point.

Like CZK, we assume that A has support $\mathcal{A} \equiv [0, 1]$ and the outcome Y is positive and bounded, where without loss of generality we suppose that $Y \in [0, 1]$. Let $m_0(A, X) \equiv \mathbb{E}_{P_0}[R|A, X]$. For a dose rule $f: \mathcal{X} \rightarrow \mathbb{R}$, we let $m_0(X; f) \equiv m_0(f(X), X)$. We also let $m_0^*(x) \equiv \sup_a m_0(a, x)$. An optimal dose rule is defined as a dose rule f_{opt} satisfying $\mathcal{V}(f_{\text{opt}}) = \sup_f \mathcal{V}(f)$. As the supremum may not be achievable, such a rule f_{opt} is not guaranteed to exist, though for simplicity we assume that it does. An implication of this assumption is that each set $\mathcal{F}_{\text{opt}}(x) \equiv \{a: m_0(a, x) = \sup_{a'} m_0(a', x)\}$ is nonempty.

2. Faster Rate for OWL-Type Methods Under General Conditions

Using arguments from empirical risk minimization, we now discuss why we believe nearly $n^{-1/2}$ rates of convergence are attainable by OWL-type methods. Noting the close similarities between regularized estimation and function class restricted estimation, suppose one sets the penalization term in the OWL optimization problem to zero and restricts the class $\mathcal{F}_n$ over which $\hat{f}_n$ can be estimated. While likely not necessary, this procedural modification simplifies our brief analysis. Assume that there exists at least one $f_{\text{opt}} \in \mathcal{F}_n$ (this assumption is not crucial to the proceeding argument, but avoids the need to control the approximation error resulting from the restriction to $\mathcal{F}_n$). By the choice of $\hat{f}_n$ as the maximizer of $\hat{\mathcal{V}}_\phi$, $\hat{\mathcal{V}}_\phi(\hat{f}_n) \geq \hat{\mathcal{V}}_\phi(f_{\text{opt}})$. Hence,

$$\begin{aligned} \mathcal{R}(\hat{f}_n) &= \mathcal{V}(f_{\text{opt}}) - \mathcal{V}(\hat{f}_n) \\ &\leq \mathcal{V}(f_{\text{opt}}) - \mathcal{V}(\hat{f}_n) + \hat{\mathcal{V}}_\phi(\hat{f}_n) - \hat{\mathcal{V}}_\phi(f_{\text{opt}}) \\ &\leq 2 \sup_f |\hat{\mathcal{V}}_\phi(f) - \mathcal{V}(f)|. \end{aligned} \quad (1)$$

If $\frac{\partial^2}{\partial a^2} m_0(a, x)$ is continuous in a for each x and its magnitude is uniformly upper bounded by a constant $C' > 0$ that does not rely on a or x , then standard techniques from kernel density estimation with mean-zero kernels show that Theorem 1 of CZK strengthens to $|\mathcal{V}_\phi(f) - \mathcal{V}(f)| \leq C'' \phi^2$ for some C'' that does not depend on f . If $a \mapsto m_0(a, x)$ has three or more derivatives for each x , then this bound can be improved further by replacing the triangular kernel used by CZK with higher-order kernels. Noting that $\mathcal{V}_\phi(f)$ is the expectation of the randomness, in $\hat{\mathcal{V}}_\phi(f)$, this shows that we have bounded the bias of $\hat{\mathcal{V}}_\phi(f)$ by $C'' \phi^2$. As the upper bound does not rely on f , it only remains to study $\sup_f |\hat{\mathcal{V}}_\phi(f) - \mathcal{V}_\phi(f)|$. By studying the variance of $\hat{\mathcal{V}}_\phi(f) - \mathcal{V}_\phi(f)$, one can show that, for each f , $|\hat{\mathcal{V}}_\phi(f) - \mathcal{V}_\phi(f)| = O_P((n\phi)^{-1/2})$. It would seem that one could impose empirical process conditions ensuring that the class over which f is selected is not too large, thereby

establishing an up to $n^{-2/5}$ rate for $2 \sup_f |\hat{\mathcal{V}}_\phi(f) - \mathcal{V}(f)|$, therefore also for $\mathcal{R}(\hat{f}_n)$, provided $\mathcal{F}_n$ is not too large and contains at least one $f_{\text{opt}} \in \mathcal{F}_{\text{opt}}$.

3. Plug-in Dose Rules

We first provide a general description of plug-in dose rules. We then argue that, like OWL-type methods, plug-in methods also appear to be able to achieve an $n^{-1/2}$ rate of convergence under fairly general conditions.

Overview of plug-in dose rules.

Consider the estimator that first estimates m_0 with m_n and then chooses $\hat{f}_n(x) = \operatorname{argmax}_a m_n(a, x)$. We assume that each x -specific argmax is well-defined. In our discussion of estimators we will focus on kernel regression estimators, largely because their theoretical properties are well established. These estimators have been shown to perform well on a strict compact subset of the support, in the sense that one can prove uniform rates approaching $n^{-1/2}$ for both the regression function and its derivative in this set (see, e.g., Newey 1994; Hansen 2008). Local polynomial estimators tend to have better behavior than kernel estimators on the boundary of their support, and more generally one could consider a much richer class of nonparametric estimators.

A nearly $n^{-1/2}$ rate for plug-in methods under general conditions.

We start by making a simple observation, analogous to that made in Section 2, that allows us to obtain nearly $n^{-1/2}$ rate convergence rates on regret for our estimator even in the absence of an interior maxima and concavity assumption. Note that $m_n(x; \hat{f}_n) - m_n(x; f_{\text{opt}}) \geq 0$ for any x and $f_{\text{opt}} \in \mathcal{F}_{\text{opt}}$. Hence, similarly to (1),

$$m_0^*(x) - m_0(x; \hat{f}_n) \leq 2 \sup_{a \in \mathcal{A}'(x)} |m_n(a, \cdot) - m_0(a, \cdot)|, \quad (2)$$

where $\mathcal{A}'(x)$ is any set containing $\hat{f}_n(x)$ and $f(x)$. While one could always take $\mathcal{A}'(x) = \mathcal{A}$, it may be beneficial to look at some $\mathcal{A}'(x) \subset \mathcal{A}$ if m_n behaves poorly on the boundaries of the support. Here we let for simplicity, but the generality is useful in the proof of the results in the next section. For a function $g: \mathcal{A} \times \mathcal{X} \rightarrow \mathbb{R}$ and a $1 \leq p < \infty$, let

$\|g\|_{p, \mathcal{A}} \equiv (\int \sup_{a \in \mathcal{A}} |g(x, a)|^p dP_0(x))^{1/p}$ and $\|g\|_{\infty, \mathcal{A}} \equiv \operatorname{ess.} \sup_x \sup_{a \in \mathcal{A}} |g(x, a)|$, where the essential supremum is with respect to P_0 . The above yields that

$$\mathcal{R}(\hat{f}_n) \leq 2\|m_n - m_0\|_{1, \mathcal{A}} \leq 2\|m_n - m_0\|_{\infty, \mathcal{A}}.$$

If m_n does not behave erratically, then the supremum over a will be the same as the essential supremum over a , and so the righthand side behaves like $\sup_{a, x} |m_n(a, x) - m_0(a, x)|$. Under smoothness conditions, kernel regression estimators have been shown to obtain slightly slower than $n^{-1/2}$ rates in this setting.

4. Fast Rates for Plug-in Dose Rules

In this section, we give results suggesting that plug-in estimators can achieve fast rates provided (i) all of the x -specific dose maxima fall in the interior of A and (ii) $\mathcal{A}$ and (ii) $a \mapsto m_0(a, x)$ satisfies a concavity condition in neighborhoods of the optima. While the plug-in estimator that we analyze in this section is the same as that analyzed in last section, the assumptions on the underlying data-generating distribution are more strict. In particular, the previous section assumed neither interior maxima nor concavity. This in part explains the availability of a faster rate.

Intuition on the attainability of a fast rate.

A parametric example provides some insight into the potential for a fast rate for regret. This example is only a sketch: some extra work would be needed to make it precise. Suppose there are no covariates and the outcome regression $a \mapsto m_0(a) \equiv \mathbb{E}[R|A=a]$ is known to belong to the family $\{a \mapsto m_\eta(a) : \eta \in \mathbb{R}^k\}$, where we assume that this family is sufficiently smooth for all of the proceeding Taylor expansions to make sense. We let η_0 represent the value of η in the data-generating distribution. Under mild conditions, one can obtain an estimate η_n satisfying $\|\eta_n - \eta_0\|_2 = O_P(1/\sqrt{n})$, for example, via least-square estimation. Let $\hat{a}_n \equiv \operatorname{argmax}_a m_{\eta_n}(a)$ and $a_{\text{opt}} \equiv \operatorname{argmax}_a m_{\eta_0}(a)$, where we assume that a_{opt} is unique and well defined. Under mild conditions, the argmax continuous mapping theorem yields that $\hat{a}_n \rightarrow a_{\text{opt}}$ in probability. Expanding $a \mapsto \dot{m}_{\eta_0}(a) \equiv \frac{\partial}{\partial a'} m_{\eta_0}(a')|_{a'=a}$ about $a = a_{\text{opt}}$ yields that $\dot{m}_{\eta_0}(\hat{a}_n) - \dot{m}_{\eta_0}(a_{\text{opt}}) \approx (\hat{a}_n - a_{\text{opt}}) \ddot{m}_{\eta_0}(a_{\text{opt}})$. If a_{opt} and $\hat{a}_n$ are interior maxima of m_{η_0} and m_{η_n} , respectively, then $\dot{m}_{\eta_0}(a_{\text{opt}}) = \dot{m}_{\eta_n}(\hat{a}_n) = 0$. From the preceding Taylor expansion, it follows that

$$\hat{a}_n - a_{\text{opt}} \approx \frac{\dot{m}_{\eta_0}(\hat{a}_n) - \dot{m}_{\eta_n}(\hat{a}_n)}{\ddot{m}_{\eta_0}(a_{\text{opt}})} = O_P(\dot{m}_{\eta_p}(\hat{a}_n) - \dot{m}_{\eta_n}(\hat{a}_n))$$

provided $\ddot{m}_{\eta_0}(a_{\text{opt}}) < 0$. Linearizing $\eta \mapsto \dot{m}_\eta(\hat{a}_n)$ yields that

$$\begin{aligned} \dot{m}_{\eta_n}(\hat{a}_n) - \dot{m}_{\eta_0}(\hat{a}_n) &\approx \langle \eta_n - \eta_0, \nabla_\eta \dot{m}_\eta(\hat{a}_n)|_{\eta=\eta_0} \rangle \\ &= O_P(\|\eta_n - \eta_0\|_2 \|\nabla_\eta \dot{m}_\eta(\hat{a}_n)|_{\eta=\eta_0}\|_2). \end{aligned}$$

If the latter norm on the right is O_P then the left-hand side is $O_P(\|\eta_n - \eta_0\|_2) = O_P(n^{-1/2})$, and so by the preceding display $\hat{a}_n - a_{\text{opt}} = O_P(n^{-1/2})$. Finally, a second-order Taylor expansion of $a \mapsto m_{\eta_0}(a)$ about a_{opt} shows that

$$\begin{aligned} \mathcal{R}(\hat{a}_n) &\equiv m_{\eta_0}(a_{\text{opt}}) - m_{\eta_0}(\hat{a}_n) \\ &\approx \frac{1}{2}(\hat{a}_n - a_{\text{opt}})^2 \ddot{m}_{\eta_0}(a_{\text{opt}}) = O_P(n^{-1}) \end{aligned}$$

where the final equality relies on $\ddot{m}_0(a_{\text{opt}}) > -\infty$. It is straightforward to extend this result to parametric families that include covariates provided the x -specific optima are all interior optima and the second derivatives at the optima are uniformly bounded away from zero and negative infinity. In practice, we do not expect the outcome regression to belong to a parametric family, and so in the next section we outline an extension of this result to general models.

Beyond the parametric case.

We now give results for a general estimator of $(a, x) \mapsto m_0(a, x)$. The main results are purely analytic and so, when presenting these results, we will replace m_n and $\hat{f}_n$ by the nonrandom quantities m and $\hat{f}$. We will then use m_n and $\hat{f}_n$ to discuss the implications of our analytic results for the plug-in estimator. While we are thorough in proving our analytic result, we only provide high level sketches of the implication for estimators due to space constraints.

For a given $\delta: \mathcal{X} \rightarrow (0, 1]$, define the x -specific neighborhood $B_\delta(x)$ by the set of all doses a such that $\inf_{a_{\text{opt}} \in \mathcal{F}_{\text{opt}}(x)} |a - a_{\text{opt}}| < \delta(x)$, that is, the set of all doses whose distance from the set $\mathcal{F}_{\text{opt}}(x)$ is less than $\delta(x)$. For $g: \mathcal{A} \times \mathcal{X} \rightarrow \mathbb{R}$, define the neighborhood-specific norm

$$\|g\|_{p, \delta} \equiv \left(\int \sup_{a \in B_\delta(x)} |g(a, x)|^p dP_0(x) \right)^{1/p}.$$

We now define several events that will serve as regularity conditions in the upcoming theorem. In particular, we assume that these events occur with either high probability or with probability 1. Each of these events relies on the neighborhood B_δ described in the preceding paragraph, but this dependence is omitted in the notation. For a function $\tilde{m}: \mathcal{A} \times \mathcal{X} \rightarrow [0, 1]$, define the event that $\tilde{m}$ is differentiable in $B_\delta(x)$ and has only interior maxima. In particular,

$$\mathcal{E}_1(\tilde{m}) = \left\{ x: a \mapsto |\tilde{m}(a, x)| \text{ is differentiable in } B_\delta(x) \text{ and } \max_{a \in [0, 1]} \tilde{m}(a, x) < \max_{a \in [0, 1]} \tilde{m}(a, x) \right\}.$$

Note that the neighborhood $B_\delta(x)$ in $\mathcal{E}_1(\tilde{m})$ depends on m_0 and not on $\tilde{m}$. For any $\gamma > 0$, define the Hölder-type event

$$\begin{aligned} \mathcal{E}_2(\gamma) \\ = \{x: |\ddot{m}_0(a, x) - \ddot{m}_0(a', x)| \leq \gamma |a - a'| \text{ for all } a, a' \in B_\delta(x)\}. \end{aligned}$$

The next event states that m_0 is only large in a neighborhood of $\mathcal{F}_{\text{opt}}$. For any $\epsilon > 0$, define

$$\mathcal{E}_3(\epsilon) \equiv \{x: m_0^\star(x) \geq m_0(a, x) + \epsilon \text{ for all } a \notin B_\delta(x)\}.$$

These events define the regularity condition used in the theorems to ensure that m_0 and m are well-behaved:

(RC) There exists some $\delta: \mathcal{X} \rightarrow (0, 1]$, $\gamma > 0$, $\epsilon > 0$ such that the event $\mathcal{E}_1(m_0) \cap \mathcal{E}_1(m) \cap \mathcal{E}_2(\gamma) \cap \mathcal{E}_3(\epsilon)$ occurs with P_0 probability 1, where we remind the reader that we have omitted the dependence of $\mathcal{E}_1(m_0)$, $\mathcal{E}_1(m)$, $\mathcal{E}_2(\gamma)$ and $\mathcal{E}_3(\epsilon)$, on δ in the notation.

The theorems differ in the concavity condition they specify. For any $t > 0$, define the event

$$\mathcal{E}_4(t) = \{x: \ddot{m}_0(a, x) \leq -t \text{ for all } a \in B_\delta(x)\}.$$

The occurrence of this event ensures that m_0 is well-peaked in a neighborhood of $\mathcal{F}_{\text{opt}}$.

Implicit in $\mathcal{E}_4(t)$ is that $a \mapsto m(a, x)$ is twice differentiable for the x contained in the event. We let $\overline{\mathcal{E}}_4(t)$ denote the complement of $\mathcal{E}_4(t)$.

Theorem 1 [Second derivative bounded away from zero]. Suppose that (RC) holds and that, for the same function $\delta: \mathcal{X} \rightarrow (0, 1]$, there exists some $\tau > 0$ such that $P_0\{\mathcal{E}_4(t)\} = 1$ for all $t \leq \tau$. For any $p \geq 1$, we have that

$$\begin{aligned} \mathcal{R}(\hat{f}) &\leq 1/(2\tau)\|\dot{m} - \dot{m}_0\|_{2,\delta}^2 + \gamma/\tau^2\|\dot{m} - \dot{m}_0\|_{3,\delta}^3 \\ &\quad + (2/\epsilon)^p\|m - m_0\|_{p,\mathcal{A}}^p. \end{aligned}$$

If $\|m - m_0\|_{\infty,\mathcal{A}} < \epsilon/2$, then the final term above can be replaced by zero.

Given a sufficiently differentiable kernel and sufficient smoothness conditions on m_0 , kernel regression estimators can obtain rates of nearly $n^{-1/2}$ for the convergence of $\dot{m}_n$ to $\dot{m}_0$ in supremum norm (Newey 1994; Hansen 2008). Thus, provided (RC) is satisfied by m_n with probability approaching 1, it will be possible to obtain rates slightly slower than n^{-1} given sufficient smoothness.

Theorem 2 [Absolute second derivative small with low probability]. Suppose that (RC) holds and that, for the same function $\delta: \mathcal{X} \rightarrow (0, 1]$, there exists some $\beta > 0$, $K_0 > 0$ such that $P_0\{\overline{\mathcal{E}}_4(t)\} \leq K_0 t^\beta$. Further suppose that $\|\dot{m} - \dot{m}_0\|_{2,\delta} > 0$. For any $p \geq 1$, there exist constants K_1 and K_2 relying on K_0 , β , γ such that

$$\begin{aligned} \mathcal{R}(\hat{f}) &\leq K_1\|\dot{m} - \dot{m}_0\|_{3,\delta}^{2r(\beta)}\|m - m_0\|_{3,\delta}^{3z(\beta)} \\ &\quad + K_2\|\dot{m} - \dot{m}_0\|_{3,\delta}^{3[r(\beta) - z(\beta)]}\|m - m_0\|_{3,\delta}^{6z(\beta)} \\ &\quad + (2/\epsilon)^p\|m - m_0\|_{p,\mathcal{A}}^p, \end{aligned}$$

where above $z(\beta) \equiv 1/(2\beta + 3)$ and $r(\beta) \equiv 2\beta z(\beta)$. If $\|m - m_0\|_{\infty,\mathcal{A}} < \epsilon/2$, then the final term above can be replaced by zero.

As $\beta \rightarrow \infty$, so that the concavity assumption in the above theorem approaches that of Theorem 1, $r(\beta) \rightarrow 1$ and $z(\beta) \rightarrow 0$. Often the term with leading constant K_2 in the above

theorem will be of smaller order than the term with leading constant K_1 . If both norms in the above converge to zero at rate slightly slower than $n^{-1/2}$, as we would expect given enough smoothness, then the above theorem gives a rate that is slightly slower than $n^{-\frac{4\beta+3}{4\beta+6}}$ if m_n is a kernel estimator. For large β this rate upper bound is nearly n^{-1} , but as β gets smaller (absolute second derivative of m_0 small with higher probability), the rate upper bound is only slightly faster than the $n^{-1/2}$ rate outlined in Section 3.

5. Discussion

We have outlined theoretical results for an alternative procedure for estimating the optimal dose rule. This method uses a plug-in estimate of the outcome regression and selects the dose that maximizes this function within each stratum of covariates. In parallel to results from the classification literature, we showed that one can achieve fast rates on this estimation procedure under a concavity assumption. These rates are much faster than the almost $n^{-1/4}$ rates presented for OWL. In contrast to the classification and the discrete treatment literatures, it is not clear to us whether or not these rates are specific to plug-in estimation strategies. As we outlined at the end of Section 3, we suspect that OWL-type methods can achieve a faster than $n^{-1/4}$ rate ($n^{-1/2}$) given restrictions on the entropy of the candidate function class for f . We would not be surprised to find that it can also achieve an n^{-1} rate under these concavity assumptions. We would be interested in hearing CZK's insights on this open question.

Despite the proven fast rate of plug-in estimation strategies, there are compelling reasons to prefer value-based estimation strategies to plug-in-based strategies. In particular, estimation strategies seeking to directly optimize the value should be more robust than plug-in methods. This robustness may be especially relevant in high-dimensional settings. Consider a class $\mathcal{F}$ of dose rules. In methods that directly optimize the value (of which OWL is an approximation), one typically expects that, even for a misspecified class $\mathcal{F}$, the method does the best job possible under misspecification in the sense that $\mathcal{V}(f_n) \rightarrow \sup_{f \in \mathcal{F}} \mathcal{V}(f)$. Plug-in

methods also yield a class $\mathcal{F}$ of possible dose rules, namely, by first specifying a class $\mathcal{M}$ of candidate estimates for the outcome regression m_0 . This class subsequently yields a class of dose rules $\mathcal{F}_{\mathcal{M}} \equiv \{x \mapsto \operatorname{argmax}_a m(a, x) : m \in \mathcal{M}\}$. This procedure typically yields that $\operatorname{Risk}(m_n) \rightarrow \inf_{m \in \mathcal{M}} \operatorname{Risk}(m)$, but this convergence does not typically imply that $\mathcal{V}(f_n) \rightarrow \sup_{f \in \mathcal{F}_{\mathcal{M}}} \mathcal{V}(f)$. Thus, as discussed by CZK, it would appear that OWL performs better than plug-in methods under misspecification.

Nonetheless, we do not expect that OWL methods are strictly superior to plug-in methods. In CZK's simulation, there was no strict ordering between OWL and plug-in methods. In small samples, when one cannot reasonably expect to have a strong estimate of the outcome regression, OWL outperformed its plug-in competitors. As sample size increased (and covariate dimension remained fixed), plug-in procedures become comparable or superior to OWL methods. As is standard in our work, in practice we suggest using a cross-validation-based selector such as super-learning to select between using OWL methods and a plug-in procedure (Luedtke and van der Laan 2016b). A reasonable follow-up question is what

criterion the super-learner should aim to optimize. We believe it best to use a direct value-based objective function (or an approximation thereof) because this makes it possible to maintain the robustness of CZK's approach. We would be interested in hearing the CZK's opinion on this matter.

We again congratulate CZK for making an important advance in extending optimal treatment rules to the continuous treatment setting. We look forward to seeing the many subsequent advances to come in this area.

6. Proofs

For any $t > 0$, define the event

$$\mathcal{E}_5(t) \equiv \mathcal{E}_1(m_0) \cap \mathcal{E}_1(m) \cap \mathcal{E}_2(\gamma) \cap \left\{ x: \hat{f}(x) \in B_\delta(x) \right\} \cap \mathcal{E}_4(t).$$

We give a lemma before proving our theorems.

Lemma 1.

If $x \in \mathcal{E}_5(t)$, then

$$m_0(x; f_{\text{opt}}) - m_0(x; \hat{f}) \leq t^{-1} \left[\dot{m}(x; \hat{f}) - \dot{m}_0(x; \hat{f}) \right]^2 \times \left(\frac{1}{2} + \gamma t^{-1} |\dot{m}(x; \hat{f}) - \dot{m}_0(x; \hat{f})| \right).$$

Proof.

Let $f_{\text{opt}}(x) = \arg\min_a |\hat{f}(x) - a|$, where the argmin is well-defined because $x \in \mathcal{E}_1(m_0)$, and therefore $a \mapsto m_0(a, x)$ is continuous in $B_\delta(x)$.

Because $x \in \mathcal{E}_2(\gamma)$ and $\hat{f}(x), f_{\text{opt}}(x) \in B_\delta(x)$, $\ddot{m}_0(\cdot, x)$ is continuous in the closed interval between $\hat{f}(x)$ and $f_{\text{opt}}(x)$, and therefore there exists an $\tilde{a}_1$ falling between $\hat{f}(x)$ and $f_{\text{opt}}(x)$ such that $\dot{m}_0(x; f_{\text{opt}}) - \dot{m}_0(x; \hat{f}) = [f_{\text{opt}}(x) - \hat{f}(x)] \ddot{m}_0(\tilde{a}_1, x)$. The fact that $\dot{m}_0(x; f_{\text{opt}}) = \dot{m}(x; \hat{f}) = 0$ shows that

$$\dot{m}(x; \hat{f}) - \dot{m}_0(x; \hat{f}) = [f_{\text{opt}}(x) - \hat{f}(x)] \ddot{m}_0(\tilde{a}_1, x). \quad (3)$$

Expanding $a \mapsto m(a, x)$ about $f_{\text{opt}}(x)$ shows that there also exists an $\tilde{a}_2$ between $\hat{f}(x)$ and $f_{\text{opt}}(x)$ such that $m_0(x; f_{\text{opt}}) - m_0(x; \hat{f}) = -\frac{1}{2} (f_{\text{opt}}(x) - \hat{f}(x))^2 \ddot{m}_0(\tilde{a}_2, x)$. Note that $\tilde{a}_1, \tilde{a}_2 \in B_\delta(x)$. Using that $x \in \mathcal{E}_2(\gamma)$ and that $|\tilde{a}_1 - \tilde{a}_2| \leq |f_{\text{opt}}(x) - \hat{f}(x)|$,

$$m_0(x; f_{\text{opt}}) - m_0(x; \hat{f}) \leq \frac{1}{2} \left[f_{\text{opt}}(x) - \hat{f}(x) \right]^2 \ddot{m}_0(\tilde{a}_1, x) + \gamma |\tilde{a}_1 - \tilde{a}_2|^3.$$

Upper bounding the second term on the right by $\gamma |f_{\text{opt}}(x) - \hat{f}(x)|^3$, plugging in (3) to replace the two instances of $f_{\text{opt}}(x) - \hat{f}(x)$, and using that $x \in \mathcal{E}_4(t)$ yields the result. $\square$

Proof of Theorem 1 and 2. First note that

$$\begin{aligned} \mathcal{R}(\hat{f}) &= \mathbb{E} \left[(m_0(X; f_{\text{opt}}) - m_0(X; \hat{f})) I\{f(X) \in B_\delta(X)\} \right] \\ &+ \mathbb{E} \left[(m_0(X; f_{\text{opt}}) - m_0(X; \hat{f})) I\{f(X) \notin B_\delta(X)\} \right]. \end{aligned}$$

Denote the terms on the right by T_1 and T_2 , respectively. Using (RC), $Y \in [0, 1]$, and (2),

$$\begin{aligned} T_2 &\leq P_0 \left\{ \hat{f}(X) \notin B_\delta(X), X \in \mathcal{E}_3(\epsilon) \right\} \\ &\leq P_0 \left\{ m_0^\star(X) - m_0(X; \hat{f}) \geq \epsilon \right\} \\ &\leq P_0 \left\{ \sup_a |m(a, X) - m_0(a, X)| \geq \epsilon/2 \right\}. \end{aligned}$$

If $\|m - m_0\|_{\infty, \mathcal{A}} < \epsilon/2$, then the upper bound is equal to zero. Alternatively, Markov's inequality yields that, for any p for which $\|m - m_0\|_{p, \mathcal{A}} < \infty$, $T_2 \leq (2/\epsilon)^p \|m - m_0\|_{p, \mathcal{A}}^p$.

We now consider T_1 . For any $t > 0$,

$$\begin{aligned} T_1 &= \mathbb{E} \left[(m_0^\star(X) - m_0(X; \hat{f})) I_{\mathcal{E}_4(t)} I\{\hat{f}(X) \in B_\delta(X)\} \right] \\ &+ \mathbb{E} \left[(m_0^\star(X) - m_0(X; \hat{f})) I_{\overline{\mathcal{E}_4(t)}} I\{\hat{f}(X) \in B_\delta(X)\} \right]. \end{aligned}$$

Denote the terms on the right by S_1 and S_2 , respectively. We first bound S_1 . Applying (RC) shows that we can replace the indicator of the event $\mathcal{E}_4(t) \cap \{x: f(x) \in B_\delta(x)\}$ by $I_{\mathcal{E}_5(t)}$.

Applying Lemma 1 to bound $m_0^\star(x) - m_0(x; \hat{f})$, removing $\mathcal{E}_5(t)$ from the resulting expectation, and using that $|m(x; \hat{f}) - \dot{m}_0(x; \hat{f})| \leq \sup_{a \in B_\delta(x)} |\dot{m}(a, x) - \dot{m}_0(a, x)|$ when $\hat{f}(x) \in B_\delta(x)$,

$$S_1 \leq 2^{-1} t^{-1} \|\dot{m} - \dot{m}_0\|_{2, \delta}^2 + \gamma t^{-2} \|\dot{m} - \dot{m}_0\|_{3, \delta}^3. \quad (4)$$

We now turn to S_2 . For Theorem 1, the bound on Y and the bound on $P_0(\overline{\mathcal{E}_4(t)})$ yield that $S_2 \leq I(t > \tau)$. For Theorem 2, Hölder's inequality, (2), and the bound on $P_0(\overline{\mathcal{E}_4(t)})$ yield that $S_2 \leq (2K_0)^{2/3} \|m - m_0\|_{3, \delta} t^{2\beta/3}$. These inequalities hold for arbitrary $t > 0$. Choose $t = \tau$ for Theorem 1 and choose

$$t = \left(\frac{\|\dot{m} - \dot{m}_0\|_{3,\delta}^2}{2K_0^{1/2}\|m - m_0\|_{3,\delta}} \right)^{\frac{3}{2\beta+3}}$$

for Theorem 2. For Theorem 2, bound the appearance of $\|\dot{m} - \dot{m}_0\|_{2,\delta}$ in (4) by $\|\dot{m} - \dot{m}_0\|_{3,\delta}$. Recalling that $\hat{\mathcal{R}}(\hat{f}) = T_1 + T_2 = S_1 + S_2 + T_2$ and the established bounds on these quantities yield the result. $\square$

Acknowledgments

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The authors thank Marco Carone for helpful discussions.

Funding

Alex Luedtke and Mark vander Laan, respectively, acknowledge the support of NIAID at the NIH under award number UM1 AI068635 and NIH grant R01 AI074345-06.

References

- Audibert JY, and Tsybakov AB (2007), “Fast Learning Rates for Plug-in Classifiers,” *Annals of Statistics*, 35, 608–633. [1526]
- Hansen BE (2008), “Uniform Convergence Rates for Kernel Estimation With Dependent Data,” *Economic Theory*, 24, 726–748. [1527,1528]
- Luedtke AR, and van der Laan MJ (2016a), “Statistical Inference for the Mean Outcome Under a Possibly Non-Unique Optimal Treatment Strategy,” *Annals of Statistics*, 44, 713–742. [1526] [PubMed: 30662101]
- Luedtke (2016b), “Super-Learning of an Optimal Dynamic Treatment Rule,” *International Journal of Biostatistics*, 12, 305–332. [1529] [PubMed: 27227726]
- Newey WK (1994), “Kernel Estimation of Partial Means and a General Variance Estimator,” *Economic Theory*, 10, 1–21. [1527,1528]
- Zhao Y, Zeng D, Rush A, and Kosorok M (2012), “Estimating Individual Treatment Rules Using Outcome Weighted Learning,” *Journal of the American Statistical Association*, 107, 1106–1118. [1526] [PubMed: 23630406]