Published in final edited form as:

Stat Med. 2016 April 15; 35(8): 1245–1256. doi:10.1002/sim.6783.

Using pilot data to size a two-arm randomized trial to find a nearly optimal personalized treatment strategy

Eric B. Laber*, Ying-Qi Zhao, Todd Regh, Marie Davidian, Anastasios Tsiatis, Joseph B. Stanford, Donglin Zeng, Rui Song, and Michael R. Kosorok

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Abstract

A personalized treatment strategy formalizes evidence-based treatment selection by mapping patient information to a recommended treatment. Personalized treatment strategies can produce better patient outcomes while reducing cost and treatment burden. Thus, among clinical and intervention scientists, there is a growing interest in conducting randomized clinical trials when one of the primary aims is estimation of a personalized treatment strategy. However, at present, there are no appropriate sample size formulae to assist in the design of such a trial. Furthermore, because the sampling distribution of the estimated outcome under an estimated optimal treatment strategy can be highly sensitive to small perturbations in the underlying generative model, sample size calculations based on standard (uncorrected) asymptotic approximations or computer simulations may not be reliable. We offer a simple and robust method for powering a single stage, two-armed randomized clinical trial when the primary aim is estimating the optimal single stage personalized treatment strategy. The proposed method is based on inverting a plugin projection confidence interval and is thereby regular and robust to small perturbations of the underlying generative model. The proposed method requires elicitation of two clinically-meaningful parameters from clinical scientists and uses data from a small pilot study to estimate nuisance parameters which are not easily elicited. The method performs well in simulated experiments and is illustrated using data from a pilot study of time to conception and fertility awareness.

Keywords

Personalized medicine; sample size calculation; treatment regimes; nonregular asymptotics; projection confidence region

1. Introduction

A personalized treatment strategy formalizes clinical decision making as a function that maps individual patient information to a recommended treatment. By selecting the right treatment for the right patients, personalized treatment strategies can lead to better patient outcomes while reducing overmedication, cost, and patient burden [1]. There is a growing interest in conducting randomized clinical trials when one of the primary aims is the estimation of an optimal personalized treatment strategy [2]. However, at present, there are

*Correspondence to: North Carolina State University, Department of Statistics, SAS Hall, Raleigh, NC, 27518, USA.

no appropriate sample size formulae to inform the design of such a trial as existing methods treat estimation of a personalized treatment strategy as a secondary aim [3, 4, 5]. Furthermore, the expected outcome of the optimal personalized treatment strategy is nonregular (in the sense described in [6]); consequently, sample size formulae cannot be based on conventional asymptotic approximations of the sampling distribution of the estimated expected value of the optimal personalized treatment strategy [7, 8]. We propose a method for sizing a trial when the primary aim is the estimation of an optimal personalized treatment strategy that is based on inverting a plugin projection confidence interval [9, 10]. Because this interval is regular and robust to small perturbations of the underlying generative model so too is our sample size procedure. The method relies on auxiliary data in the form of a small pilot study or observational data and two simple clinician-elicited quantities. Simulations suggest that a pilot study with as few as 20 subjects may be sufficient to obtain reliable sample size estimators.

There is a large literature on estimating optimal treatment regimes using randomized or observational studies (see [11] and references therein). We consider a single treatment decision wherein baseline patient information is used to inform treatment choice; baseline information can include measures of disease stage severity, medical history, genetic and genomic information, as well as demographics, patient preferences and cost. The single decision problem is a special case of the more general problem of estimating a sequence of treatment rules, one per stage of clinical information, that map accrued patient information to a recommended treatment. Many existing estimation algorithms include single- and multi-stage variants. Regression-based methods, e.g., Q -learning [12, 13, 14, 15, 16], are a common and convenient tool for estimating interpretable treatment strategies from observational or randomized study data. Furthermore, extensions of these methods exist to accommodate a number of practical issues including: missing data [17], censored data [18], multiple outcomes [19, 20], binary outcomes [11], and high-dimensional data [21, 22]. Thus, we use Q -learning as the basis for constructing our sample size procedure. However, the proposed methodology could be extended to deal with other estimation procedures including direct value maximization [23, 24, 25, 26, 27], and A -learning [28, 29, 10, 16].

This work is partially motivated by our involvement with a clinical trial of fertility awareness and time to pregnancy (NICHD, K23 HD0147901, PI Joseph B. Stanford) and the planning of a follow-up trial. The original trial enrolled $n = 123$ subjects and demonstrated a promising but statistically insignificant benefit of a personalized treatment strategy. Furthermore, existing theory suggests that treatment response varies across subjects according to number of past successful pregnancies, length of sexual relationship, and other factors (see Section 5); thus, there is also scientific justification for a personalized treatment strategy in this context. Clinical investigators are therefore interested in running a follow-up trial with more subjects to estimate a high-quality personalized treatment strategy. The proposed methodology offers a principled method for sizing such a follow-up trial.

2. Q -learning

Q -learning is a simple regression-based method for estimating an optimal personalized treatment strategy. This is the assumed estimation method in the sample size calculations

given in Section 3. However, we do not assume that the posited regression model used in the Q -learning algorithm is correctly specified. Let capital letters like X and Y denote random variables and lowercase letters like x and y denote instances of these random variables. To estimate a personalized treatment strategy we will collect n independent identically distributed triples $\mathcal{D}(n) = \{(X_i, A_i, Y_i)\}_{i=1}^n$, one for each subject, where: $X \in \mathbb{R}^p$ denotes baseline, i.e., pre-treatment, subject information; $A \in \{-1, 1\}$ denotes the treatment received; and $Y \in \mathbb{R}$ denotes the observed outcome for which higher values correspond to better clinical outcomes. We assume the simplest setting in which A is independent of X and assigned randomly with $P(A = 1) = P(A = -1) = 1/2$. The more general setting, in which the distribution of A depends on X , can be handled at the expense of are more complex expression for the marginal mean outcome under a pre-specified policy, e.g., see [6, 24]; the extension of asymptotic arguments to this setting are straightforward (see Remark 2).

A personalized treatment strategy is a map $\pi: \mathbb{R}^p \rightarrow \{-1, 1\}$ from the space of patient characteristics into the space of available treatments. Under π , a patient with $X = x$ is recommended treatment $\pi(x)$. To define the optimal treatment strategy we use the framework of potential outcomes [30, 31]. Let $Y^*(a)$ denote the potential outcome under treatment $a \in \{-1, 1\}$. For any treatment strategy, π , define $Y^*(\pi) = \sum_{a \in \{-1, 1\}} Y^*(a) 1_{\pi(X)=a}$ to be the potential outcome under π . Define $V(\pi) \triangleq \mathbb{E}Y^*(\pi)$, so that $V(\pi)$ denotes the expected outcome if all subjects in the population are assigned treatment under π . An optimal treatment strategy, say π^{opt} , satisfies $V(\pi^{\text{opt}}) \geq V(\pi)$ for all π . To connect the potential outcomes with the observed data we assume: (i) $Y = Y^*(A)$ (consistency); (ii) there exists $c > 0$ so that $c < P(A = 1|X) < 1 - c$ with probability one (positivity); and (iii) $\{Y^*(-1), Y^*(1)\} \perp\!\!\!\perp A | X$ (no unmeasured confounders). Note that (ii) and (iii) are automatically satisfied by the randomized design considered here. However, if the pilot data (described below) are observational, then the validity of these assumptions must be considered carefully. Define $Q(x, a) \triangleq \mathbb{E}\{Y^*(a)|X = x, A = a\}$, then, under the foregoing assumptions, it can be shown that $Q(x, a) = \mathbb{E}Y|X = x, A = a$ and $V(\pi) = \mathbb{E}Q\{X, \pi(X)\}$. Thus, it can be seen that $\pi^{\text{opt}}(x) = \arg \max_{a \in \{-1, 1\}} Q(x, a)$.

In practice, the Q -function, $Q(x, a)$, is not known. In Q -learning, an estimator of π^{opt} is formed by estimating $Q(x, a)$ from the observed data and then using a plug-in estimator. We will assume a linear working model of the form $Q(x, a; \theta) = x_1^\top \alpha + a x_2^\top \beta$, where $\theta = (\alpha^\top, \beta^\top)^\top$ and x_1, x_2 are known features derived from x . Note that x_1 and x_2 may contain polynomial terms or other nonlinear basis expansions. Let $\mathbb{P}_n$ denote the empirical measure, i.e., given *i.i.d.* random variables $Z_1, \dots, Z_n$ and function f defined on the support of Z , $\mathbb{P}_n f(Z) \triangleq n^{-1} \sum_{i=1}^n f(Z_i)$. Define least squares estimator $\hat{\theta}_n \triangleq \arg \min_{\theta} \mathbb{P}_n \{Y - Q(X, A; \theta)\}^2$. Subsequently, $\hat{\pi}_n(x) = \arg \max_{a \in \{-1, 1\}} Q(x, a; \hat{\theta}_n) = \text{sign}(x_2^\top \hat{\beta}_n)$, is the estimated optimal treatment strategy.

Define $\theta^* \triangleq \arg \min_{\theta} \mathbb{E}\{Y - Q(X, A; \theta)\}^2$ and subsequently

$\pi^*(x) \triangleq \arg \max_a Q(x, a; \theta^*) = \text{sign}(x_2^\top \beta^*)$. Thus, θ^* is the population analog of $\hat{\theta}_n$. However, π^* need not equal π^{opt} if the working model for the Q -function is not correctly specified (see, for example, [14, 16, 32]). In the context of a sample size calculation $\mathbb{E}Y(\pi^*)$

is more relevant than $\mathbb{E}Y(\pi^{\text{opt}})$ because the (possibly misspecified) postulated model for Q -function will be used in the analysis of the collected data. Our sample size procedure is robust in that provides sample size estimates that are correctly powered when this postulated model is incorrect.

2.1. Estimation of $V(\pi^*)$ and nonregularity

For any $\gamma \in \mathbb{R}^{\dim(x_2)}$ let $\pi_\gamma(x) \triangleq \text{sign}(x_2^\top \gamma)$, and write $V(\gamma)$ to denote $V(\pi_\gamma)$. Applying a change of measure it follows that $V(\gamma) = 2\mathbb{E}Y 1_{AX^\top \gamma > 0}$ [26, 24, 25]. Thus, $V(\beta^*)$ is a nonsmooth (nondifferentiable) functional of the underlying generative distribution [33, 34]. Consequently, there exists neither a regular nor an asymptotically unbiased estimator of $V(\beta^*)$ [35, 8]. To see how this impacts sample size calculation, define

$$\hat{V}_{\mathcal{D}(n)}(\gamma) \triangleq 2\mathbb{P}_n \left\{ Y 1_{AX^\top \gamma \geq 0} - \left(1_{AX^\top \gamma \geq 0} - \frac{1}{2} \right) (X_1^\top \hat{\alpha}_n + |X_2^\top \gamma|) \right\}, \quad (1)$$

then $V_{\mathcal{D}(n)}(\beta^*)$ is unbiased for $V(\beta^*)$ and $V_{\mathcal{D}(n)}(\hat{\beta}_n)$ is an augmented inverse probability weighted estimator of $V(\beta^*)$ [6]. Note that the first term in (1), $2\mathbb{P}_n Y 1_{AX^\top \gamma > 0}$, is a consistent estimator of $V(\gamma)$ and is nonparametric in the sense that it does not rely on a model for the

Q -function. The second term in (1), $-2\mathbb{P}_n (1_{AX^\top \gamma \geq 0} - 1/2)(X_1^\top \hat{\alpha}_n + |X_2^\top \gamma|)$ has mean zero and thus does not affect the consistency of $V_{\mathcal{D}(n)}(\gamma)$, instead this term is included to reduce variability of the estimator. See [24] for a detailed derivation of this estimator in a more general setting. A standard approach to sample size calculation would depend on the

limiting distribution of $\mathbb{G}_n \triangleq \sqrt{n} \{ \hat{V}_{\mathcal{D}(n)}(\hat{\beta}_n) - V(\beta^*) \}$. However, as stated above, because $V(\beta^*)$ is nonsmooth, $\mathbb{G}_n$ is nonregular and its sampling distribution is difficult to approximate in finite samples (see, for example, [36] and references therein). In fact, $V_{\mathcal{D}(n)}(\hat{\beta}_n)$ need not even be consistent for $V(\beta^*)$ (see [33]). Thus, sample size calculations should not be based directly on asymptotic approximations to the sampling distribution of $\mathbb{G}_n$.

3. Projection sample size estimator

Let V_0 denote an elicited expected outcome under standard care, and let $\delta > 0$ be an elicited value so that δV_0 denotes a clinically meaningful improvement in the expected outcome. For simplicity we treat V_0 as fixed, however, this is not essential, an alternative would be to replace V_0 with the mean outcome under a fixed treatment regime $\pi(x) \equiv a$ for all x .

Suppose that we observe data $\mathcal{D}_p = \{(X_i, A_i, Y_i)\}_{i=1}^{n_p}$ from a preliminary study; this preliminary study could be a standalone pilot study, an internal pilot, or an observational study. For simplicity, we assume that $\mathcal{D}_p$ is collected in a randomized pilot study of exactly the same form as the follow-up study we are trying to size, but with far fewer patients (see Remark 3 below). Our goal is to determine a sample size $n^* = n^*(\mathcal{D}_p)$ so that:

- (P1) The power to reject the hypothesis $V(\beta^*) = V_0$ in favor of $V(\beta^*) > V_0$ is at least $1 - \rho$ when $V(\beta^*) = (1 + \delta)V_0$;
- (P2) $P\{|V_{\mathcal{D}(n^*)}(\hat{\beta}_{n^*}) - V(\beta^*)| \leq \varepsilon\} = 1 - \rho$;

where $\mathcal{D}(n^*)$ is a follow-up study of size n^* , and $\rho, \varepsilon > 0$ are user chosen values. Conceptually, it is possible to have different values of ρ in (P1) and (P2), say ρ_1 and ρ_2 ; however, as we will show, because our method relies on inverting a confidence interval for $V(\beta^*)$, it would be necessary to choose the level of this confidence interval to be $\{1 - \min(\rho_1, \rho_2)\} \times 100$ to ensure that (P1) and (P2) hold simultaneously so there is no loss in generality by setting $\rho = \rho_1 = \rho_2$. The condition (P1) ensures that, with high probability, the follow-up study will produce significant evidence that the optimal personalized treatment strategy leads to better patient outcomes on average than standard care when indeed the optimal personalized treatment strategy yields clinically significantly better outcomes. Thus, (P1) regards the viability of a data-driven personalized treatment strategy. (P2) ensures that the estimated optimal treatment strategy using data from the follow-up study will have an empirical performance near that of the optimal strategy. Thus, (P2) regards the quality of the estimated optimal treatment strategy. An alternative to (P2) is to try to control $P(|V(\hat{\beta}_n) - V(\beta^*)| > \varepsilon)$, however, there is no guarantee that a confidence interval for $V(\hat{\beta}_n)$ actually contains $V(\beta^*)$. Because $V(\hat{\beta}_n)$ is a data-dependent parameter, constructing a confidence interval for this quantity is not well-studied (see [33, 34]).

Our sample size procedure works by constructing an estimator of the diameter of a confidence region for $V(\beta^*)$ as a function of the sample size and then solving for a sample size sufficiently large to ensure (P1) and (P2) hold with high probability. Constructing a sample size based on confidence region diameter is sometimes categorized as precision-based or accuracy-based sample size calculation [37, 38]. Suppose that based on n observations we can construct a $(1 - \rho) \times 100\%$ confidence interval for $V(\beta^*)$, say $\zeta_{n,1-\rho}$. Let $d(n)$ denote the diameter of $\zeta_{n,1-\rho}$. If $d(n^*) \leq \delta V_0$, then $V(\beta^*)$ and V_0 cannot simultaneously belong to $\zeta_{n^*,1-\rho}$ provided $V(\beta^*) \leq (1 + \delta)V_0$. Because $P\{V(\beta^*) \in \zeta_{n^*,1-\rho}\} = 1 - \rho$, we can distinguish $V(\beta^*)$ from V_0 with probability $1 - \rho$ when $V(\beta^*) \leq (1 + \delta)V_0$ by rejecting the hypothesis $V(\beta) = V_0$ when V_0 falls below the lower endpoint of $\zeta_{n^*,1-\rho}$. Thus, (P1) is satisfied. Furthermore, if $d(n^*) \leq \varepsilon$ and $V_{\mathcal{D}(n^*)}(\hat{\beta}_{n^*}) \in \zeta_{n^*,1-\rho}$ then $P\{|V_{\mathcal{D}(n^*)}(\hat{\beta}_{n^*}) - V(\beta^*)| \leq \varepsilon\} = 1 - \rho$, and (P2) is satisfied. Figure (1) shows a schematic illustrating the foregoing relationships between a confidence interval for $V(\beta^*)$ and the criteria (P1) and (P2). To estimate n^* we first construct a valid (asymptotic) confidence interval for $V(\beta^*)$ and an estimator $\hat{d}_{\mathcal{D}_p}(n)$ of the diameter function $d(n)$. Our estimator of n^* is given by $n^* \triangleq \inf\{n : \hat{d}_{\mathcal{D}_p}(n) \leq \min(\delta V_0, \varepsilon)\}$.

To construct a confidence interval for $V(\beta^*)$ we treat β^* as a nuisance parameter and construct a projection interval [9, 10]. For any fixed $\gamma \in \mathbb{R}^{\dim(x_2)}$ define

$$\sigma^2(\gamma) \triangleq 4\mathbb{E}(Y1_{AX^\top \gamma \geq 0} - \mathbb{E}Y1_{AX^\top \gamma \geq 0})^2 - 4\mathbb{E}\{Y1_{AX^\top \gamma \geq 0}(X_1^\top \alpha^* + |X_2^\top \gamma|)\} + 4\mathbb{E}\{(1_{AX^\top \gamma \geq 0} - 1/2)(X_1^\top \alpha^* + |X_2^\top \gamma|)\}^2,$$

and subsequently define the plug-in estimator

$$\hat{\sigma}_{\mathcal{D}(n)}^2(\gamma) \triangleq 4\mathbb{P}_n(Y1_{AX^\top \gamma \geq 0} - \mathbb{P}_n Y 1_{AX^\top \gamma \geq 0})^2 - 4\mathbb{P}_n\{Y 1_{AX^\top \gamma \geq 0}(X_1^\top \hat{\alpha}_n + |X_2^\top \gamma|)\} \\ + 4\mathbb{P}_n\{(1_{AX^\top \gamma \geq 0} - 1/2)(X_1^\top \hat{\alpha}_n + |X^\top \gamma|)\}^2.$$

For any fixed $\gamma \in \mathbb{R}^{dim(x_2)}$, under standard regularity conditions, it follows from the central limit theorem that $\sqrt{n}(\hat{V}_{\mathcal{D}(n)}(\gamma) - V(\gamma))$ converges to a normal random variable with mean zero and variance $\sigma^2(\gamma)$. Therefore, an asymptotic $(1 - \mu) \times 100\%$ confidence interval for $V(\gamma)$ based on pilot data $\mathcal{D}_p$ is

$$\mathbb{I}_\gamma(\mathcal{D}_p, n_p) \triangleq \left(\hat{V}_{\mathcal{D}_p}(\gamma) - z_{1-\mu/2} \hat{\sigma}_{\mathcal{D}_p}(\gamma) / \sqrt{n_p}, \hat{V}_{\mathcal{D}_p}(\gamma) + z_{1-\mu/2} \hat{\sigma}_{\mathcal{D}_p}(\gamma) / \sqrt{n_p} \right), \quad (2)$$

where $z_{1-\mu/2}$ is the $(1 - \mu/2) \times 100$ percentile of a standard normal random variable. Let $B \triangleq (X_1^\top, AX_2^\top)^\top$, define $\Omega = (\mathbb{E}BB^\top)^{-1} \mathbb{E}BB^\top \{Y - Q(X, A)\}^2 (\mathbb{E}BB^\top)^{-1}$ to be the asymptotic covariance of $\sqrt{n}(\hat{\theta}_n - \theta^*)$, and define Σ to be the submatrix of Ω corresponding to the asymptotic covariance of $\sqrt{n}(\hat{\beta}_n - \beta^*)$. Let $\hat{\Sigma}_{\mathcal{D}_p}$ denote the plug-in estimator of Σ based on the pilot data $\mathcal{D}_p$. Define

$\mathbb{T}(\mathcal{D}_p, n_p) \triangleq \left\{ \gamma : n_p(\gamma - \hat{\beta}_{\mathcal{D}_p}) \sum_{D_p}^{-1} (\gamma - \hat{\beta}_{\mathcal{D}_p}) \leq \chi_{dim(x_2), 1-\xi}^2 \right\}$, where $\chi_{dim(x_2), 1-\xi}^2$ is the $(1 - \xi) \times 100$ percentile of a χ^2 -distribution with $dim(x_2)$ degrees of freedom. Thus, $\mathbb{T}(\mathcal{D}_p, n_p)$ is an asymptotic $(1 - \xi) \times 100\%$ confidence interval for β^* . An asymptotic $(1 - \xi - \mu) \times 100\%$ projection confidence interval for $V(\beta^*)$ is therefore given by $\bigcup_{\gamma \in \mathbb{T}(\mathcal{D}_p, n_p)} \mathbb{I}_\gamma(\mathcal{D}_p, n_p)$. This interval has diameter $d(n_p) = \sup_{\gamma \in \mathbb{T}(\mathcal{D}_p, n_p)} \mathbb{I}_\gamma(\mathcal{D}_p, n_p) - \inf_{\gamma \in \mathbb{T}(\mathcal{D}_p, n_p)} \mathbb{I}_\gamma(\mathcal{D}_p, n_p)$. For an arbitrary $n > n_p$ an estimator of $d(n)$ is given by $\hat{d}_{\mathcal{D}_p}(n) \triangleq \sup_{\gamma \in \mathbb{T}(\mathcal{D}_p, n)} \mathbb{I}_\gamma(\mathcal{D}_p, n) - \inf_{\gamma \in \mathbb{T}(\mathcal{D}_p, n)} \mathbb{I}_\gamma(\mathcal{D}_p, n)$, and subsequently $n^* = \inf\{n : \hat{d}_{\mathcal{D}_p}(n) \leq \min(\delta V_0, \varepsilon)\}$.

Remark 1—Our sample size procedure is based on choosing a sufficient number of samples to ensure that the width of a confidence interval is sufficiently small. It may be helpful to note that the proposed procedure is analogous to the following familiar example. Suppose we desire a confidence interval for the mean of a random variable Z of width no more than ω . Let Z_n denote the sample mean of random sample $Z_1, \dots, Z_n$. The natural $(1 - \rho) \times 100\%$ confidence interval is $\bar{Z}_n \pm z_{1-\rho/2} \sigma / \sqrt{n}$ which has diameter $d(n) = 2z_{1-\rho/2} \sigma / \sqrt{n}$. Using pilot data $\mathcal{D}_p$ we construct estimator $\hat{\sigma}_{\mathcal{D}_p}$ of σ and subsequently the estimator $\hat{d}_{\mathcal{D}_p}(n) = 2z_{1-\rho/2} \hat{\sigma}_{\mathcal{D}_p} / \sqrt{n}$ of $d(n)$; solving for $n^* = \inf\{n : \hat{d}_{\mathcal{D}_p}(n) \leq \omega\}$ gives the familiar expression $\hat{n}^* = 4z_{1-\rho/2}^2 \hat{\sigma}_{\mathcal{D}_p}^2 / \omega^2$.

Remark 2—If the preliminary data are observational then the proposed sample size estimator can be applied with slight modification. To distinguish between the generative model for the pilot data and the data to be collected in the follow-up trial being sized, write $\mathcal{D}_p = \{(X_i^p, A_i^p, Y_i^p)\}_{i=1}^n$. We assume that for any y, x, a $P(Y^p = y | X^p = x, A^p = a) = P(Y = y | X = x, A = a)$, and $P(X^p = x) = P(X = x)$. Furthermore, we require that the pilot study satisfy the consistency, positivity, and no unmeasured confounders assumptions stated in Section 2.

Under these assumptions, the only difference between the population in the pilot study and the population in the randomized trial is the treatment assignment mechanism; note, weaker assumptions are possible. Let $p(a|x)$ denote the treatment assignment mechanism $P(A^p = a | X^p = x)$. Then,

$$\sigma^2(\gamma) = 8\mathbb{E}p(A^p | X^p) (Y^p 1_{A^p X^p \top \gamma \geq 0} - 2\mathbb{E}p(A^p | X^p) Y^p 1_{A^p X^p \top \gamma \geq 0})^2 - 8\mathbb{E}p(A^p | X^p) \{Y^p 1_{A^p X^p \top \gamma \geq 0} (X_1^{p \top} \alpha^* + |X_2^{p \top} \gamma|) + 8\mathbb{E}p(A^p | X^p) \{(1_{A^p X^p \top \gamma \geq 0} - 1/2)(X_1^{p \top} \alpha^* + |X^p \top \gamma|)\}\}^2,$$

and $\Omega = \{\mathbb{E}p(A^p | X^p) B^p B^{p \top}\}^{-1} \mathbb{E}p(A^p | X^p) B^p B^{p \top} \{Y^p - Q(X^p, A^p)\}^2 \{\mathbb{E}p(A^p | X^p) B^p B^{p \top}\}^{-1}$, and $(\alpha^{* \top}, \beta^{* \top})^\top = \{\mathbb{E}p(A^p | X^p) B^p B^{p \top}\}^{-1} \mathbb{E}p(A^p | X^p) B^p Y^p$. Thus, with knowledge or an estimator of $p(a|x)$ it is possible to construct plugin estimators of the quantities needed in the sample size formula with observational data.

Remark 3—The proposed sample size calculation does not require that the posited Q -function is correctly specified. However, if one is willing to assume that the Q -function is correctly specified then: (i) $V(\beta^*)$ is still non-regular; and (ii) the proposed sample size procedure can be applied with a slight modification. Assume that $Q(x, a) = x_1^\top \alpha^* + a x_2^\top \beta^*$ so that the posited linear model is correctly specified. We assume X_1 has been centered so that $\mathbb{E}X_1 = 0$. Then, $V(\beta) = \mathbb{E}Q\{X, \pi_\beta(X)\} = \mathbb{E}\{\text{sgn}(X_2^\top \beta) X_2^\top \beta^*\}$ and $V(\beta^*) = \mathbb{E}|X_2^\top \beta^*|$. Define $\tilde{V}_{\mathcal{D}(n)}(\gamma) = \mathbb{P}_n |X_2^\top \gamma|$ then $\sqrt{n}(\tilde{V}_{\mathcal{D}(n)}(\gamma) - \mathbb{E}|X_2^\top \gamma|)$ is asymptotically normal with mean zero and variance $\tau^2(\gamma) = \mathbb{E}\{|X_2^\top \gamma| - \mathbb{E}|X_2^\top \gamma|\}^2$. Define

$$\tilde{\tau}_{\mathcal{D}(n)}^2(\gamma) \triangleq \mathbb{P}_n \{|X_2^\top \gamma| - \mathbb{P}_n |X_2^\top \gamma|\}^2 \text{ and}$$

$$\tilde{\mathbb{I}}_\gamma(\mathcal{D}_p, n_p) \triangleq \left(\tilde{V}_{\mathcal{D}_p}(\gamma) - z_{1-\mu/2} \tilde{\tau}_{\mathcal{D}_p}(\gamma) / \sqrt{n_p}, \tilde{V}_{\mathcal{D}_p}(\gamma) + z_{1-\mu/2} \tilde{\tau}_{\mathcal{D}_p}(\gamma) / \sqrt{n_p} \right).$$

Letting $\mathbb{I}(\mathcal{D}_p, n_p)$ be defined as above, then $\bigcup_{\gamma \in \mathbb{I}(\mathcal{D}_p, n_p)} \tilde{\mathbb{I}}_\gamma(\mathcal{D}_p, n_p)$ is a valid $(1 - \xi - \mu) \times 100\%$ valid (asymptotic) confidence interval for $V(\beta^*)$. This interval can be inverted to construct a sample size. A potential drawback of this approach is that it depends heavily on the correctness of the postulated analysis model.

4. Simulation experiments

In this section we examine the finite sample performance of the proposed sample size procedure using simulated experiments. We use data from the following generative model:

$$\begin{aligned} U_i &\sim \text{Bernoulli}(\nu), & X_i &\sim N_p(0, I_p), \\ Z_i &= (I - U_i P_{\beta^*}) X_i, & A_i &\sim \text{Uniform}\{-1, 1\}, \\ e_i &\sim N(0, 1), & Y_i &= X_i^\top \alpha^* + A_i Z_i^\top \beta^* + e_i, \end{aligned}$$

where P_{β^*} denotes the projection matrix onto the linear span of β^* . Thus, under this class of generative models $\nu \times 100\%$ of the subjects will have the same expected outcome regardless

of which treatment they receive, i.e., the treatment effect for these patients is zero. We use a working model of the form $Q(x, a, \theta) = x^\top a + ax^\top \beta$ so that the working model is only correctly specified when $\nu = 0$. In our experiments we fix $p = 5$ and $a^* \equiv 1$. We include intercept terms in our working models so the dimension of θ is $2p + 2 = 12$. Furthermore, for any fixed value of $\nu \in [0, 1]$ we choose $\beta^* \equiv c$ where c is chosen to be the smallest positive constant so that $V(\beta^*)$ equals a pre-specified value. In our experiments we consider $\nu = 0, 0.05, 0.10, 0.25, 0.50, 0.75$ and $V(\beta^*) = 2, 2.25$. We fix $\delta = \varepsilon = V_0 = 1$. We set $\rho = 0.80$ and, following recommended best practice in [9], obtain an 80% confidence interval for $V(\beta^*)$ by setting $\xi = 0.01$ thereby constructing a 99% confidence interval for β , and setting $\mu = 0.19$ thereby constructing an 81% confidence interval for $V(\gamma)$ for each fixed γ in the confidence set for β^* . For illustration, we also implement the sample-size formula

$$\tilde{n}^*(\mathcal{D}_p) = 4z_{1-(\alpha+\eta)/2}^2 \hat{\sigma}_{\mathcal{D}_p}^2(\hat{\beta}_{n_p}) / (\delta V_0)^2 \text{ which is a plugin estimator derived from the}$$

asymptotic normality of $\sqrt{n_p} \{ \hat{V}_{\mathcal{D}_p}(\beta^*) - V(\beta^*) \}$ discussed in Section 3. However, as discussed previously, normal approximations are not valid in this context without strong (unverifiable) assumptions (e.g., [35, 8]) and even when these conditions hold the small sample performance of these approximations may be poor [36, 7, 39, 33, 34]. A confidence interval using the normal approximation with data $\mathcal{D}$ is $\hat{V}_{\mathcal{D}}(\hat{\beta}_n) \pm z_{1-(\alpha+\eta)/2} \hat{\sigma}_n(\hat{\beta}_n) / \sqrt{n}$.

We repeat the following procedure 1000 times. First, we generate pilot data $\mathcal{D}_p$ of size $n_p = 20$. Second, we use the sample size procedure described in the preceding section to obtain $n^*(\mathcal{D}_p)$; because n_p is small, least squares estimation of all $2p + 2$ parameters in our working model is unstable, hence in these experiments we use ridge-regression to estimate $\theta = (a^\top, \beta^\top)^\top$ with the amount of regularization chosen by BIC. The confidence intervals $\mathbb{T}_{\mathcal{D}_p, n_p}$ are adjusted accordingly (details given in the appendix). Finally, we generate a follow-up study, say $\mathcal{D}^* = \mathcal{D} \{n^*(\mathcal{D}_p)\}$ and subsequently the projection interval $\bigcup_{\gamma \in \mathbb{T}_{\mathcal{D}^*, n^*(\mathcal{D}_p)}} \mathbb{I}_{\gamma} \{ \mathcal{D}^*, n^*(\mathcal{D}_p) \}$ for $V(\beta^*)$. To approximate the union over all $\gamma \in \mathbb{T}_{\mathcal{D}^*, n^*(\mathcal{D}_p)}$ we randomly sample

100,000 points from $\mathbb{T}_{\mathcal{D}^*, n^*(\mathcal{D}_p)}$ using rejection sampling with a $\text{Normal}(\hat{\beta}_{n_p}, n_p^{-1} \sum_{\mathcal{D}_p})$ proposal distribution. Table 1 shows the estimated coverage, mean, and standard deviation of the proposed sample size procedure, as well as that of the normal-approximation-based estimator. The estimated coverage of $V(\beta^*)$ is not significantly different at the 0.01 level from the target 0.80 in all but one case for the proposed estimator. Thus, empirically objective (P2) was attained. The resultant interval covered V_0 significantly less than the $1 - 0.80 = 0.20$ target and thus the power to detect a difference between the adaptive strategy and standard care was significantly higher than the 0.80 target. Thus, empirically objective (P1) was attained. As expected, generally, both the mean and standard deviation of $n^*(\mathcal{D}_p)$ increase as the proportion of individuals with zero treatment effect (ν) increases. Not surprisingly, the normal based confidence interval generally predicted smaller sample sizes than the proposed projection based procedure. However, the normal approximation sample size covered $V(\beta^*)$ significantly less than the 0.80 target in 7 of the 12 generative models; furthermore, the normal based interval also included V_0 more frequently indicating less power.

4.1. Midstream sample size adjustment

A common strategy to increase precision when sample size estimates are based on estimated nuisance parameters is to use an interim sample size adjustment based on updated estimates of the nuisance parameters [40]. We illustrate a simple interim update for our sample size procedure. First, we generate pilot data $\mathcal{D}_p$ of size n_p . Second, we use our sample size procedure to estimate $n^*(\mathcal{D}_p)$. Third, we generate data $\mathcal{D}_1$ in a follow-up study with $\hat{n}_1^*(\mathcal{D}_p) \triangleq [\hat{n}^*(\mathcal{D}_p)/2]$ subjects. Fourth, treating $\mathcal{D}_1$ as a new pilot study, we collect data, say $\mathcal{D}_2$, on an additional $\hat{n}_2^*(\mathcal{D}_1) \triangleq \max\{0, \hat{n}^*(\mathcal{D}_1) - \hat{n}_1^*(\mathcal{D}_p)\}$ subjects. Let $\mathcal{D}^* \triangleq \mathcal{D}_1 \cup \mathcal{D}_2$. Finally, we construct a confidence interval for $V(\beta^*)$ based on the $\hat{n}^*(\mathcal{D}_p, \mathcal{D}_1) = \hat{n}_1^*(\mathcal{D}_p) + \hat{n}_2^*(\mathcal{D}_1)$ subjects in $\mathcal{D}^*$. Because our approach uses a projection interval, uncertainty about parameter values translates into larger diameters and hence larger estimated sample sizes. Consequently, the midstream update typically results in smaller sample sizes.

Table 2 shows the estimated coverage, mean, and standard deviation of the proposed sample size procedure based with a midstream update; to make the results comparable the table is constructed from the same 1000 Monte Carlo pilot data sets of size $n_p = 20$ used to construct Table 1. The mean and standard deviation of the total samples size $n^*(\mathcal{D}_p, \mathcal{D}_1)$ are both generally smaller when using the midstream update yet the coverages are comparable. This suggests that midstream update may be a viable strategy for reducing the number of subjects required to estimate an optimal personalized treatment strategy.

5. Time to pregnancy in normal fertility

A motivation for this work is our involvement with a randomized controlled trial of the effect of fertility awareness on time to pregnancy in normal fertility introduced in Section 1. We use data from this trial as pilot data to size a follow-up study. In this trial subjects were randomized to receive either standard care or follow the Creighton Model Fertility Care™ (CrM) System. The CrM involves patient education and daily patient self-monitoring to predict periods of fertility. Thus, there is an interest in identifying subjects for whom the burden of CrM is expected to be worthwhile. A personalized treatment strategy in this setting is map from baseline patient characteristics to a recommendation of CrM or standard care.

For simplicity, in our analysis we consider patients with complete information ($n = 109$ of an enrolled 123 subjects). Eligibility criteria required participants to be of proven fertility (full term pregnancy in past 8 years), and planning to conceive in the near future. For additional study details, see [41]. The outcome of interest is the minimum of eight cycles and the number of cycles before conception occurred. To be consistent with our development of finding strategies which maximize an expected outcome, we use eight minus the outcome of interest as our response. The two possible treatments are standard care (coded -1) and CrM (coded 1). Treatment was randomly assigned using a block-randomized design with blocking by age (blocks 18–29 and 30–35); thus, the no unmeasured confounders assumption (iii) of Section 2 is automatically satisfied. Variables thought to be predictive of benefit from CrM include number of previous successful pregnancies (Prg),

subject age (years, Age), length of sexual relationship (years, SexLen), and education level (indicator of post-secondary, Edu). Let X denote these predictors, A the binary treatment, and Y the outcome. We use a linear Q -function of the form $Q(x, a) = x^\top a + ax^\top \beta$. Residual diagnostics for this model suggested that SexLen and Age should be transformed to the log-scale and that Prg should be transformed to $1/\text{Prg}$. Standard regression diagnostics are provided in the appendix. Table 3 provides the estimated coefficients for the Q -function. There is a significant interaction between CrM and the reciprocal of Prg suggesting that the benefit of CrM is greatest for subjects with a single prior successful pregnancy. The estimated optimal decision rule is $\pi_n(\hat{x}) = \text{sign}(x^\top \hat{\beta}_n)$. This rule has an estimated value of 5.82 (which corresponds to an average of 2.18 months to conception); in contrast the estimated value of assigning standard care to all patients is 5.52 (2.48 months to conception) whereas the estimated value of assigning CrM to all patients is 4.92 (3.08 months to conception). Thus, while there is some indication that a personalized treatment strategy might be effective, the evidence is not significant. For example, an 80% confidence interval for the value of the optimal decision rule, $V(\beta^*)$, is [5.50, 6.13], which has a diameter of 0.63. It is of interest to know how this diameter would change in a larger, follow-up study. Figure 2 shows $\hat{n}^*$ as a function of the target diameter for an 80% interval. To see how this figure might be used, suppose that, based on the pilot study, one postulated that $V(\beta^*) = 5.90$, and $V_0 = 5.52$, the mean outcome treating all subjects. Then $\delta = V(\beta^*)/V_0 - 1 = 0.0688$ so that a diameter of $\delta V_0 = 0.38$ is required for 80% power to detect superiority of $V(\beta^*)$ to V_0 . According to Figure 2, the estimated required size for a follow-up study is $\hat{n}^* = 280$. If instead it was postulated that $V(\beta^*) = 6.12$, the required diameter is 0.5 and the required size for a follow-up study is $\hat{n}^* = 165$. In addition to informing the size of a follow-up study, the slope of the line in Figure 2 can be viewed as a rough estimate of the incremental gain in the precision of $V_{\mathcal{D}(n)}(\hat{\beta}_n)$ per subject.

6. Discussion

We proposed a procedure for estimating the required sample size for a two-arm randomized clinical trial when the primary aim is the estimation of a personalized treatment strategy. The proposed method is based on inverting a projection confidence interval for the expected outcome under the optimal treatment strategy. Empirically the proposed method appeared to make efficient use of limited pilot data and enjoyed increased precision when assisted by a midstream interim update. Our procedure relied on standard assumptions for causal inference for treatment strategies, i.e., consistency, positivity, and no unmeasured confounders. The question of what happens when there are unmeasured confounders in the pilot is an interesting question but one that is difficult to answer in generality. The reason is that even in simpler settings, e.g., estimating an optimal regime from observational data, the impact of unmeasured confounding depends on the nature and severity of the confounding. Thus, characterizing the impact of unmeasured confounding on the sample size calculation would require first cataloging the effects of unmeasured confounding on estimation of the nuisance parameters across the many possible ways in which confounding can occur. It would be interesting future work to investigate the effects of confounding on the sample size procedure.

We focused on Q -learning as our method of estimation, however, there is growing interest in direct value maximization methods [42, 23, 24, 25, 26, 27, 43]. Distribution theory for parameters estimated via direct value maximization methods is lacking at present; when this theory is developed the proposed sizing procedure should be directly applicable.

Furthermore, direct value maximization procedures often rely on a lower dimensional parameter than Q -learning, thus it is possible that the proposed sample size procedure will be less conservative when used with such methods.

Another important extension of this work is to sequences of treatment rules, also known as dynamic treatment regimes [28, 10]. Sample size procedures for sequential multiple assignment randomized trials [4] with the primary aim of constructing an optimal sequence of treatment rules are of increasing interest. The extension to multiple time points is complicated by the fact that in this setting the parameters indexing the Q -functions are nonregular [7, 32]. We are currently pursuing this extension.

References

1. Coalition PM. The case for personalized medicine. *Technical Report*, Coalition for personalized medicine. 2011
2. Hamburg MA, Collins FS. The path to personalized medicine. *New England Journal of Medicine*. 2010; 363(4):301–304. [PubMed: 20551152]
3. Oetting, A.; Levy, J.; Weiss, R.; Murphy, S. Statistical methodology for a smart design in the development of adaptive treatment strategies. In: Shrouf, P.; Keys, K.; Ornstein, K., editors. *Causality and Psychopathology*. Oxford University Press; New York: 2010. p. 179-205.
4. Murphy SA. An experimental design for the development of adaptive treatment strategies. *Statistics in medicine*. 2005; 24(10):1455–1481. [PubMed: 15586395]
5. Li Z, Murphy SA. Sample size formulae for two-stage randomized trials with survival outcomes. *Biometrika*. 2011; 98(3):503–518. [PubMed: 22363091]
6. Tsiatis, A. *Semiparametric theory and missing data*. Springer Verlag; 2006.
7. Laber EB, Lizotte DJ, Qian M, Pelham WE, Murphy SA, et al. Dynamic treatment regimes: Technical challenges and applications. *Electronic Journal of Statistics*. 2014; 8:1225–1272. [PubMed: 25356091]
8. Hirano K, Porter JR. Impossibility results for nondifferentiable functionals. *Econometrica*. 2012; 80(4):1769–1790.
9. Berger RL, Boos DD. P values maximized over a confidence set for the nuisance parameter. *Journal of the American Statistical Association*. 1994; 89(427):1012–1016.
10. Robins, JM. *Proceedings of the Second Seattle Symposium in Biostatistics*. Springer; 2004. Optimal structural nested models for optimal sequential decisions; p. 189-326.
11. Chakraborty, B.; Moodie, E. *Statistical Methods for Dynamic Treatment Regimes*. Springer; 2013.
12. Watkins C, Dayan P. Q -learning. *Machine learning*. 1992; 8(3):279–292.
13. Murphy SA. A generalization error for q -learning. *Journal of Machine Learning Research*. Jul.2005 6:1073–1097. [PubMed: 16763665]
14. Qian M, Murphy SA. Performance guarantees for individualized treatment rules. *Annals of Statistics*. 2011; 39(2):1180. [PubMed: 21666835]
15. Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano GA, Waxmonsky JG, Yu J, Murphy SA. Q -learning: A data analysis method for constructing adaptive interventions. *Psychological methods*. 2012; 17(4):478. [PubMed: 23025434]
16. Schulte P, Tsiatis A, Laber E, Davidian M. q - and a -learning methods for estimating optimal dynamic treatment regimes. *Statistical Science*. 2014; 29(4):640–661. [PubMed: 25620840]

17. Shortreed SM, Laber E, Scott Stroup T, Pineau J, Murphy SA. A multiple imputation strategy for sequential multiple assignment randomized trials. *Statistics in Medicine*. 2014; 33(24):4202–4214. [PubMed: 24919867]
18. Goldberg Y, Kosorok MR. Q-learning with censored data. *Annals of Statistics*. 2012; 40(1):529. [PubMed: 22754029]
19. Laber EB, Lizotte DJ, Ferguson B. Set-valued dynamic treatment regimes for competing outcomes. *Biometrics*. 2014; 70(1):53–61. [PubMed: 24400912]
20. Linn, K.; Laber, E.; Stefanski, L. Estimation of dynamic treatment regimes for complex outcomes: Balancing benefits and risks. In: Kosorok, M.; Moodie, E., editors. *Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine*. CRC Press; New York: 2015.
21. Lu W, Zhang HH, Zeng D. Variable selection for optimal treatment decision. *Statistical methods in medical research*. 2011 0962280211428 383.
22. Geng Y, Zhang HH, Lu W. On optimal treatment regimes selection for mean survival time. *Statistics in Medicine*. 2014
23. Orellana L, Rotnitzky A, Robins J. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part i: Main content. *Int Jn of Biostatistics*. 2010; 6(2)
24. Zhang B, Tsiatis AA, Laber EB, Davidian M. A robust method for estimating optimal treatment regimes. *Biometrics*. 2012; 68(4):1010–1018. [PubMed: 22550953]
25. Zhang B, Tsiatis AA, Davidian M, Zhang M, Laber E. Estimating optimal treatment regimes from a classification perspective. *Stat*. 2012; 1(1):103–114. [PubMed: 23645940]
26. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*. 2012; 107(499):1106–1118. [PubMed: 23630406]
27. Zhao Y, Zeng D, Laber E, Kosorok MR. New statistical learning methods for estimating optimal dynamic treatment regimes. *Journal of the American Statistical Association*. 2015; 110(510):583–598. [PubMed: 26236062]
28. Murphy SA. Optimal dynamic treatment regimes (with discussion). *Journal of the Royal Statistical Society: Series B*. 2003; 58:331–366.
29. Blatt, D.; Murphy, SA.; Zhu, J. Technical Report 04-63. The Methodology Center, Penn. State University; 2004. A-learning for approximate planning.
30. Rubin D. Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*. 1978:34–58.
31. Splawa-Neyman J, Dabrowska D, Speed T, et al. On the application of probability theory to agricultural experiments. essay on principles section 9. *Statistical Science*. 1990; 5(4):465–472.
32. Laber E, Linn L, Stefanski L. Interactive model-building for q -learning. *Biometrika*. 2014; 101(4): 831–847. [PubMed: 25541562]
33. Laber EB, Murphy SA. Adaptive confidence intervals for the test error in classification. *Journal of the American Statistical Association*. 2011; 106(495):904–913. [PubMed: 22053123]
34. Chakraborty B, Laber E, Zhao Y. Inference about the expected performance of a data-driven treatment regime. *Clinical Trials*. 2014; 0:1–10.
35. van der Vaart A. On differentiable functionals. *Annals of Statistics*. 1991; 19(1):178–204.
36. Leeb H, Pötscher BM. The finite-sample distribution of post-model-selection estimators and uniform versus nonuniform approximations. *Econometric Theory*. 2003; 19(1):100–142.
37. Kelley K, Maxwell SE, Rausch JR. Obtaining power or obtaining precision delineating methods of sample-size planning. *Evaluation & the Health Professions*. 2003; 26(3):258–287. [PubMed: 12971200]
38. Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. *Annu Rev Psychol*. 2008; 59:537–563. [PubMed: 17937603]
39. Chakraborty B, Murphy S, Strecher V. Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research*. 2010; 19(3):317–343. [PubMed: 19608604]

40. Wittes J, Brittain E. The role of internal pilot studies in increasing the efficiency of clinical trials. *Statistics in Medicine*. 1990; 9(1-2):65-72. [PubMed: 2345839]
41. Stanford J, Smith KMV. Impact of instruction in the creighton model fertilitycare system on time to pregnancy in couples of proven fecundity: Results of a randomised trial. *Paediatric and Perinatal Epidemiology*. 2014; 29(5)
42. Robins J, Orellana L, Rotnitzky A. Estimation and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine*. 2008:4678-4721. [PubMed: 18646286]
43. Zhao Y, Zeng D, Laber EB, Song R, Ming Y, Kosorok MR. Doubly robust learning for estimating individualized treatment with censored data. *Biometrika*. 2014; 102(1):1-16.

7. Appendix

7.1. Wald type confidence region for ridge regression

For fixed nonnegative ridge parameter λ , let $\hat{\theta}_n^\lambda = \arg \min_{\theta} \mathbb{P}_n(Y - Q(X, A; \theta))^2 + \lambda \theta^\top \theta$. Then, $\hat{\theta}_n^\lambda = (\mathbb{P}_n BB^\top + \lambda I)^{-1} \mathbb{P}_n BB^\top \hat{\theta}_n$. Recall $\theta = (a^\top, \beta^\top)^\top$, let U denote the constant matrix so that $U\theta = \beta$, and define $W_n^\lambda \triangleq U(\mathbb{P}_n BB^\top + \lambda I)^{-1} \mathbb{P}_n BB^\top$, so that $\hat{\beta}_n^\lambda = W_n^\lambda \hat{\beta}_n$. It can be seen that $\sqrt{n}(\hat{\beta}_n^\lambda - W_n^\lambda \beta^*)$ converges in distribution to a normal random variable with mean zero and variance-covariance matrix

$$\Sigma^\lambda = U(PBB^\top + \lambda I)^{-1} PBB^\top (Y - B^\top \theta^*) (PBB^\top + \lambda I)^{-1} U^\top,$$

which has plug-in estimator

$$\hat{\Sigma}^\lambda = U(\mathbb{P}_n BB^\top + \lambda I)^{-1} \mathbb{P}_n BB^\top (Y - B^\top \hat{\theta}_n^\lambda) (\mathbb{P}_n BB^\top + \lambda I)^{-1} U^\top.$$

A Wald type confidence region for β^* based on $\hat{\beta}_n^\lambda$ is given by

$$\left\{ \beta : n \left(\hat{\beta}_n^\lambda - W_n^\lambda \beta \right)^\top \left(\hat{\Sigma}^\lambda \right)^- \left(\hat{\beta}_n^\lambda - W_n^\lambda \beta \right) \leq \chi_{\dim(\beta^*), 1-\xi}^2 \right\},$$

where the u^- denotes the generalized inverse of u .

7.2. Regression diagnostics for Q-function in fertility study

Figure (3) displays standard regression diagnostic plots using data from the fertility study.

Figure 1.

Schematic showing how a confidence interval for $V(\beta^*)$ can be used to ensure (P1) and (P2) hold. If the diameter of the confidence interval is less than $(1 + \delta)V_0$, then the interval cannot simultaneously contain V_0 and $V(\beta^*)$. Furthermore, if the diameter of the confidence interval is less than ε and $\hat{V}_{\mathcal{D}(n^*)}(\hat{\beta}_n)$ belongs to the interval, then when the interval covers $V(\beta^*)$ it follows that $|\hat{V}_{\mathcal{D}(n^*)}(\hat{\beta}_n) - V(\beta^*)| \leq \varepsilon$.

Figure 2. Estimated sample size as a function of 80% confidence interval diameter. Reducing the interval width by a factor of two would require nearly a four-fold increase in sample size over the initial pilot study.

Figure 3. Regression diagnostics for fitted Q -function used in fertility study. The discreteness of the response manifests itself as parallel lines in the plot residuals vs. fitted values. Otherwise, there do not appear to be major departures from standard linear modeling assumptions.

Table 1

Estimated coverage, mean, and standard deviation for the proposed sample size procedure based on 1000 Monte Carlo replications. Target coverage of $V(\beta^*)$ is 0.80; bolded numbers indicate observed coverage is not significantly different from the target coverage at the 0.05 level.

Projection sample size					Normal sample size						
V^*	V_0	δ	ν	$\text{Cov } V(\beta^*)$	$\text{Cov } V_0$	$\mathbb{E}n^*(\mathcal{D}_p)$	$SD(n^*(\mathcal{D}_p))$	$\text{Cov } V(\beta^*)$	$\text{Cov } V_0$	$\mathbb{E}n^*(\mathcal{D}_p)$	$SD(n^*(\mathcal{D}_p))$
2	1	1	0	0.77	0.08	156	151	0.77	0.14	111	62
2	1	1	0.05	0.77	0.06	144	132	0.77	0.13	107	62
2	1	1	0.1	0.80	0.07	144	127	0.76	0.14	100	60
2	1	1	0.25	0.80	0.08	149	123	0.76	0.16	98	65
2	1	1	0.5	0.79	0.09	191	204	0.72	0.19	85	65
2	1	1	0.75	0.79	0.09	191	204	0.70	0.23	101	104
2.25	1	1	0	0.74	0.02	145	138	0.78	0.073	120	68
2.25	1	1	0.05	0.77	0.02	147	143	0.78	0.07	112	65
2.25	1	1	0.1	0.77	0.03	142	118	0.77	0.067	114	71
2.25	1	1	0.25	0.78	0.04	166	157	0.75	0.09	112	75
2.25	1	1	0.5	0.79	0.03	228	216	0.74	0.11	105	84
2.25	1	1	0.75	0.77	0.06	347	390	0.69	0.17	129	135

Table 2

Estimated coverage, mean, and standard deviation for the proposed sample size procedure with a midstream adjustment based on 1000 Monte Carlo replications. Target coverage of $V(\beta^*)$ is 0.80; bolded numbers indicate observed coverage is not significantly different from the target coverage at the 0.05 level.

V^*	V_0	δ	ν	Coverage $V(\beta^*)$	Coverage V_0	$\mathbb{E}\hat{n}_1^*(\mathcal{D}_p)$	$\mathbb{E}\pi(\mathcal{D}_p, \mathcal{D}_1)$	$SD(n(\mathcal{D}_p, \mathcal{D}_1))$
2	1	1	0	0.80	0.06	76	109	72
2	1	1	0.05	0.76	0.06	72	110	78
2	1	1	0.1	0.79	0.07	72	112	92
2	1	1	0.25	0.80	0.06	75	115	83
2	1	1	0.5	0.79	0.05	91	155	116
2	1	1	0.75	0.76	0.08	91	235	201
2.25	1	1	0	0.77	0.02	73	108	101
2.25	1	1	0.05	0.74	0.03	74	109	79
2.25	1	1	0.1	0.77	0.02	71	108	95
2.25	1	1	0.25	0.78	0.02	83	131	136
2.25	1	1	0.5	0.78	0.02	114	169	138
2.25	1	1	0.75	0.73	0.03	174	246	197

Table 3

Estimated regression coefficients for linear Q -function. The outcome is coded so that higher values correspond to shorter time to conception. There is a significant interaction between CrM and 1/Prg suggesting that the benefit of CrM is greatest for patients with only one successful prior pregnancy (which was the minimum required for study eligibility). Indeed, all subjects assigned CrM by the estimated optimal treatment strategy had one successful prior pregnancy.

	Estimate	Std. Error	t value	Pr(> t)
Intercept	22.10	6.11	3.61	0.00
CrM	-1.80	6.11	-0.294	0.770
log Age	-5.72	1.99	-2.88	0.005
1/Prg	-0.586	0.642	-0.906	0.367
Edu	0.316	0.194	1.63	0.106
log SexLen	0.844	0.550	1.535	0.128
CrM:log Age	-0.0840	1.99	-0.042	0.966
CrM:1/Prg	1.47	0.647	2.26	0.025
CrM:Edu	-0.0920	0.194	-0.474	0.636
CrM:log SexLen	0.699	0.550	1.27	0.206