Biometrika, *pp*. 1–7

Supplementary material for "Optimal stratification in outcome prediction using baseline information"

BY FLORENCE H. YONG
Department of Biostatistics, Harvard University, Boston, Massachusetts 02115, U.S.A.
florenceyong@mail.harvard.edu

LU TIAN
Department of Biomedical Data Science, Stanford University, Stanford, California 94305, U.S.A.
lutian@stanford.edu

SHENG YU
Center for Statistical Science, Tsinghua University, Beijing, China
syu@tsinghua.edu.cn

TIANXI CAI AND L.J. WEI
Department of Biostatistics, Harvard University, Boston, Massachusetts 02115, U.S.A.
tcai.hsph@gmail.com, wei@hsph.harvard.edu

APPENDIX 1

The dynamic programming algorithm for optimal stratification

We will describe the dynamic programming algorithm for identifying the optimal grouping in this section. Below we first provide a brief introduction to the dynamic programming algorithm. To this end, assume that our objective is to find the minimum total cost of n stages:

$$\min_{(a_1,\ldots,a_n)} \sum_{t=1}^{n} c_t(s_t, a_t)|_{s_1},$$

where s_t is the state of stage t, a_t is the action we take at stage t, and $c_t(s_t, a_t)$ is the cost associated with state s_t and action a_t. The state of the next stage s_{t+1} is determined by both s_t and a_t: $s_{t+1} = f_t(s_t, a_t)$, $t = 1, \ldots, n-1$, and initial state s_1 is given. If we know that the minimum total cost from stage $m+1$ through stage n starting at state $s_{m+1} = s$ is

$$C_{m+1}(s) = \min_{(a_{m+1},\ldots,a_n)} \sum_{t=m+1}^{n} c_t(s_t, a_t)|_{s_{m+1}=s},$$

© 2016 Biometrika Trust

then the optimal cost from stage m at a given state s is simply

$$\begin{aligned} C_m(s) &= \min_{(a_m, a_{m+1}, \ldots, a_n)} \sum_{t=m}^{n} c_t(s_t, a_t)|_{s_m = s} \\ &= \min_{a_m} \left\{ c_m(s, a_m) + \min_{(a_{m+1}, \ldots, a_n)} \sum_{t=m+1}^{n} c_t(s_t, a_t)|_{s_{m+1} = f_m(s, a_m)} \right\} \\ &= \min_{a_m} \left[c_m(s, a_m) + C_{m+1}\{f_m(s, a_m)\} \right]. \end{aligned}$$

Thus we can start from the minimum cost $C_n(s) = \min_{a_n} c_n(s, a_n)$ at stage n to consecutively find $C_{n-1}(s), C_{n-2}(s), \ldots, C_2(s)$ and $C_1(s)$ and the corresponding optimal solutions at stages $n-1, n-2, \ldots, 2$ and 1.

Our problem is more complicated than the formulation above due to the presence of constraints, but the basic principle remains the same. Without loss of generality, we assume that the data consists of $\{Y_i, w_i, \hat{\mu}(V_i)\}, i = 1, \ldots, n$, with $\hat{\mu}(V_1) < \hat{\mu}(V_2) < \ldots < \hat{\mu}(V_n)$. Here Y_i and w_i are response and associated nonnegative weight for the ith observation. The objective is to group n observations into K strata: $S_k, k = 1, \ldots, K$, such that

$$\sum_{k=1}^{K} \sum_{i \in S_k} |Y_i - \bar{Y}(S_k)| w_i,$$

is minimized under the constraints that

$$n_k \geq np_0 \quad \text{and} \quad \bar{Y}(S_k) - \bar{Y}(S_{k-1}) \geq d_0,$$

where p_0 is the minimum stratum fraction, S_i denotes the set of observations in the ith stratum and

$$\bar{Y}(S) = \frac{\sum_{i \in S} w_i Y_i}{\sum_{i \in S} w_i}, \text{ for } S \subset \{1, \ldots, n\}.$$

Here d_0 and p_0 are given a priori but K is unknown. To solve this optimization problem, we first consider the optimal grouping for observations $m, \ldots, n$ with the first stratum S_{mj1} comprised of j observations. That is, the optimal grouping $S_{mj1}, S_{mj2}, \ldots, S_{mjK_{mj}}$ for observations $m, \ldots, n$ minimizes the loss function

$$\sum_{k=1}^{K} \sum_{i \in S_k, i \geq m} |Y_i - \bar{Y}(S_k)| w_i,$$

under the constraints

$$n_k \geq np_0, \ \bar{Y}(S_k) - \bar{Y}(S_{(k-1)}) \geq d_0 \ \text{ and } \ S_1 = \{m, \ldots, m+j-1\}.$$

Here $j = 1, \ldots, n-m+1$. Let L_{mj} be the minimum L_1 loss for grouping observations $m, \ldots, n$ with j observations in the first stratum under the constraint above. Let the corresponding optimal grouping $S_{mj1}, \ldots, S_{mjK_{mj}}$ be denoted by $\mathcal{G}_{mj}$. If there is no stratification satisfying the constraints, e.g., when $j < np_0$, then $L_{mj} = +\infty$. In such a case, we let $\mathcal{G}_{mj} = \phi$ for convenience in notations. Also, denote $\bar{Y}(S_{mj1})$ by $\bar{Y}_{mj}$.

Like the standard dynamic programming algorithm, we start from the last observation and $(\mathcal{G}_{n1}, L_{n1}, \bar{Y}_{n1})$ can be obtained easily since $(\mathcal{G}_{n1}, L_{n1}, Y_{n1}) = (\{n\}, 0, Y_n)$ if $1 \geq np_0$ and $(\phi, +\infty, Y_n)$ otherwise. Assume that for $1 < m \leq n$ we have obtained

$$\begin{array}{ccc} (\mathcal{G}_{n1}, L_{n1}, \bar{Y}_{n1}) & & \\ (\mathcal{G}_{(n-1)1}, L_{(n-1)1}, \bar{Y}_{(n-1)1}) & (\mathcal{G}_{(n-1)2}, L_{(n-1)2}, \bar{Y}_{(n-1)2}) & \\ \cdots & & \\ (\mathcal{G}_{(m+1)1}, L_{(m+1)1}, \bar{Y}_{(m+1)1}) & \cdots & (\mathcal{G}_{(m+1)(n-m)}, L_{(m+1)(n-m)}, \bar{Y}_{(m+1)(n-m)}). \end{array}$$

We can construct $\{\mathcal{G}_{mj}, L_{mj}\}$ from the previous set of optimal solutions based on the fact that

$$L_{mj} = \min_{\{i|\bar{Y}_{mj}-\bar{Y}_{(m+j)i}\geq d_0, j\geq np_0\}} \left\{ \sum_{i=m}^{m+j-1} |Y_i - \bar{Y}_{mj}|w_i + L_{(m+j)i} \right\}$$

as in the general dynamic programming algorithm given above. Specifically, if $j < np_0$, then

$$(\mathcal{G}_{mj}, L_{mj}) = (\phi, +\infty).$$

For $np_0 \leq j \leq n-m+1$, since the first stratum of size j is fixed, we only need to choose the optimal grouping strategy for the remaining $n-m-j+1$ observations, while satisfying the necessary constraints. To examine the minimum incremental constraint between consecutive groups, we only need to consider the first two groups since the constraint is satisfied for all existing groupings $\mathcal{G}_{\tilde{m}\tilde{j}}, 1 \leq \tilde{j} \leq n+1-\tilde{m}, m+1 \leq \tilde{m}$. To this end, we may let

$$i^* = \text{argmin}_i l_{mj}(i),\ (i = 1, \ldots, n-m-j+1),$$

where

$$l_{mj}(i) = \begin{cases} \sum_{i=m}^{m+j-1} |Y_i - \bar{Y}_{mj}|w_i + L_{(m+j)i} & \text{if } \bar{Y}_{mj} - \bar{Y}_{(m+j)i} \geq d_0 \text{ and } L_{(m+j)i} < \infty \\ \infty & \text{if } \bar{Y}_{mj} - \bar{Y}_{(m+j)i} < d_0 \text{ or } L_{(m+j)i} = \infty \end{cases}.$$

Once i^* is identified with a finite loss $l_{mj}(i^*) < \infty$, we let

$$L_{mj} = l_{mj}(i^*) \quad \text{and} \quad \mathcal{G}_{mj} = \{m, \ldots, m+j-1\} \cup \mathcal{G}_{(m+j)i^*}.$$

If all $l_{mj}(i) = \infty$, then we set $(\mathcal{G}_{mj}, L_{mj}) = (\phi, \infty)$ again. Therefore, one may construct $(\mathcal{G}_{mj}, L_{mj}), j = 1, \ldots, n-m+1$ by tracking $(\mathcal{G}_{\tilde{m}\tilde{j}}, L_{\tilde{m}\tilde{j}}), 1 \leq \tilde{j} \leq n+1-\tilde{m}, m+1 \leq \tilde{m}$. Repeating this process for $m = n-1, \ldots, 1$, $(\mathcal{G}_{1j}, L_{1j}), j = 1, \ldots, n$, the optimal grouping for n observations $\{1, \ldots, n\}$ with the first group having j observations, are obtained in the end. The optimal stratification is simply $\mathcal{G}_{1j^*}$, where

$$j^* = \text{argmin}_j L_{1j},\ (j = 1, \ldots, n).$$

The complexity of the algorithm is $O(n^3)$ and therefore the computation can be slow when n is big. In such a case, one may pre-group observations with similar $\hat{\mu}(V_i)$s together before applying the dynamic programming. One way to achieve this is to divide the interval containing all the estimated scores into subintervals and represent all the $\hat{\mu}(V_i)$s in the same subinterval by its center. In this way, we effectively reduce the choices of potential grouping while using the original Y_is and w_is to calculate the $\bar{Y}_k$ and prediction error. The computation speed can be substantially improved without sacrificing much precision in locating the optimal stratification scheme.

APPENDIX 2

The Proof of Lemma 1

Let

$$L_n(\mathbf{c}) = n^{-1} \sum_{i=1}^{n} |Y_i - f(V_i|\mathbf{c})|$$

and

$$L(\mathbf{c}) = \mathrm{E}|Y_i - f_0(V_i|\mathbf{c})|,$$

where $\mathbf{c} = (c_1, \cdots, c_K)^{\mathrm{T}}, -\infty = c_1 < c_2 < \ldots < c_K = \infty$, $\hat{\mu}(V_i) = g(\hat{\beta}^{\mathrm{T}} Z_i)$, $\mu(V_i) = g(\beta_0^{\mathrm{T}} Z_i)$,

$$f(V_i|\mathbf{c}) = \sum_{k=1}^{K} \hat{\mu}_Y(c_{k-1}, c_k) I\{\hat{\mu}(V_i) \in (c_{k-1}, c_k]\},$$

$$f_0(V_i|\mathbf{c}) = \sum_{k=1}^{K} \mu_Y(c_{k-1}, c_k) I\{\mu(V_i) \in (c_{k-1}, c_k]\}$$

$$\hat{\mu}_Y(a,b) = \frac{n^{-1}\sum_{i=1}^n Y_i I\{\hat{\mu}(V_i) \in (a,b]\}}{n^{-1}\sum_{i=1}^n I\{\hat{\mu}(V_i) \in (a,b]\}} \quad \text{and} \quad \mu_Y(a,b) = \mathrm{E}\{Y|\mu(V_i) \in (a,b]\}.$$

Firstly, we will show that

$$\sup_{\mathbf{c}} |L_n(\mathbf{c}) - L(\mathbf{c})| = o_p(1),$$

where the sup is over all $\mathbf{c}$ such that $\mathrm{pr}\{\mu(V_i) \in (c_{k-1}, c_k]\} \geq \delta_0 > 0$. Since K is bounded and takes only finite number of possible values, it is sufficient to show the above uniform convergence for any fixed K. To this end, we note that the coverage number $N_{[]}(\epsilon, \mathcal{F}, L_1) < \infty$ for the class of functions $\mathcal{F} = \{yI(g(\beta^{\mathrm{T}} z) \in (a,b]) \mid \max(|a|, |b|, \|\beta\|_1) < C_0\}$ or $\{I\{g(\beta^{\mathrm{T}} z) \in (a,b]\} \mid \max(|a|, |b|, \|\beta\|_1) < C_0\}$, where $C_0 < \infty$ is a constant. Thus it follows from the Glivenko-Cantelli theorem that

$$\sup_{\max\{|a|,|b|,\|\beta\|_1\}<C_0} \left| n^{-1}\sum_{i=1}^n Y_i I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\} - \mathrm{E}\left[Y_i I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\right] \right| = o_p(1)$$

and

$$\sup_{\max\{|a|,|b|,\|\beta\|_1\}<C_0} \left| n^{-1}\sum_{i=1}^n I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\} - \mathrm{pr}\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\} \right| = o_p(1),$$

which implies that

$$\sup_{(a,b,\beta)\in\Omega_0} \left| \frac{n^{-1}\sum_{i=1}^n Y_i I\{\hat{g}(\beta^{\mathrm{T}} Z_i) \in (a,b]\}}{n^{-1}\sum_{i=1}^n I\{\hat{g}(\beta^{\mathrm{T}} Z_i) \in (a,b]\}} - \mathrm{E}\{Y|g(\beta^{\mathrm{T}} Z_i) \in (a,b]\} \right| = o_p(1), \tag{A1}$$

where $\Omega_0 = \{a, b, \beta \mid \mathrm{pr}\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\} \geq \delta_0, \max\{|a|, |b|, \|\beta\|_1\} < C_0\}$. Next, consider

$$U_n(a,b,\beta) = n^{-1}\sum_{i=1}^n I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\} \left| Y_i - \frac{n^{-1}\sum_{i=1}^n Y_i I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}}{n^{-1}\sum_{i=1}^n I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}} \right|.$$

It follows from (A1) that

$$\sup_{(a,b,\beta)\in\Omega_0} \left| U_n(a,b,\beta) - n^{-1}\sum_{i=1}^n I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\big|Y_i - \mathrm{E}\{Y|g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\big| \right| = o_p(1).$$

Now, consider the class of functions $\mathcal{F} = \{I\{g(\beta^{\mathrm{T}} v) \in (a,b]\}|y - \tilde{\mu}(a,b,\beta)| \mid (a,b,\beta) \in \Omega_0\}$, where $\tilde{\mu}(a,b,\beta)$ has continuous partial derivatives with respect to a, b and β. The covering number of the class is finite as well, and it follows from the Glivenko-Cantelli theorem that

$$n^{-1}\sum_{i=1}^n I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\big|Y_i - \mathrm{E}\{Y|g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\big|$$

uniformly converges to $u(a,b,\beta) = \mathrm{E}\left\{I\{g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\big|Y_i - \mathrm{E}\{Y|g(\beta^{\mathrm{T}} Z_i) \in (a,b]\}\big|\right\}$ over the set Ω_0 and thus

$$\sup_{(a,b,\beta)\in\Omega_0} \left| U_n(a,b,\beta) - u(a,b,\beta) \right| = o_p(1).$$

Coupled with the fact that $u(a,b,\hat{\beta}) - u(a,b,\beta_0) = o_p(1)$ and triangular inequality, it implies that

$$\sup_{(a,b,\beta_0)\in\Omega_0} \left| U_n(a,b,\hat{\beta}) - u(a,b,\beta_0) \right| = o_p(1).$$

Now, note the fact that

$$L_n(\mathbf{c}) = \sum_{k=1}^{K} U_n(c_{k-1}, c_k, \hat{\beta}) \quad \text{and} \quad L(\mathbf{c}) = \sum_{k=1}^{K} u(c_{k-1}, c_k, \beta_0),$$

we have

$$\sup_{\mathbf{c}} \left| L_n(\mathbf{c}) - L(\mathbf{c}) \right| = o_p(1).$$

Secondly, we will derive the upper bound of $L(\hat{\mathbf{c}})$ as $n \to \infty$. To this end, let the constraint be written as $\mathcal{S}_n(\mathbf{c}) \geq 0$, where

$$\mathcal{S}_n(\mathbf{c}) = \begin{pmatrix} \bar{Y}_2 - \bar{Y}_1 - d_0 \\ \cdots \\ \bar{Y}_K - \bar{Y}_{K-1} - d_0 \\ n^{-1}\sum_{i=1}^{n} I\{c_1 < \hat{\mu}(V_i) \leq c_2\} - p_0 \\ \cdots \\ n^{-1}\sum_{i=1}^{n} I\{c_{K-1} < \hat{\mu}(V_i) \leq c_K\} - p_0 \end{pmatrix}.$$

We also define the limiting constraint by $\mathcal{S}_0(\mathbf{c}) \geq 0$, where

$$\mathcal{S}_0(\mathbf{c}) = \begin{pmatrix} \bar{\mu}_2 - \bar{\mu}_1 - d_0 \\ \cdots \\ \bar{\mu}_K - \bar{\mu}_{K-1} - d_0 \\ \text{pr}\{c_1 < \tilde{\mu}(V_i) \leq c_2\} - p_0 \\ \cdots \\ \text{pr}\{c_{K-1} < \tilde{\mu}(V_i) \leq c_K\} - p_0 \end{pmatrix}.$$

Let $\mathbf{c}_0$ be the minimizer of $L(\mathbf{c})$ subject to the constraint $S_0(\mathbf{c}) \geq 0$ and $\hat{\mathbf{c}}$ be the minimizer of $L_n(\mathbf{c})$ subject to the constraint $\mathcal{S}_n(\mathbf{c}) \geq 0$. Furthermore, we let

$$\hat{\mathbf{c}}_\epsilon = \text{argmin}_{\mathbf{c}:\mathcal{S}_0(\mathbf{c})\geq\epsilon} L_n(\mathbf{c}) \quad \text{and} \quad \mathbf{c}_\epsilon = \text{argmin}_{\mathbf{c}:\mathcal{S}_0(\mathbf{c})\geq\epsilon} L(\mathbf{c}).$$

Under a rather mild condition that the numbers of strata of both stratification rules $\mathbf{c}_{\tilde{\epsilon}}$ and $\mathbf{c}_0$ are the same for some $\tilde{\epsilon} > 0$,

$$L(\mathbf{c}_\epsilon) \to L(\mathbf{c}_0) = L_0, \quad \text{as} \quad \epsilon \to 0.$$

A sufficient condition for the existence of such a $\tilde{\epsilon}$ is that the optimal grouping $\mathbf{c}_0$ is unique and the set $\{(c_1, \ldots, c_{K_0}) \mid \mathcal{S}_0(\mathbf{c}) \geq 0\}$ is not contained by a $K_0 - 1$ dimensional hyperplane in R^{K_0}, where $K_0 + 1$ is the dimension of the vector $\mathbf{c}_0$. Now, since $\mathcal{S}_n(\mathbf{c}) - S(\mathbf{c}) = o_p(1)$,

$$\text{pr}\Big[\{\mathbf{c} \mid \mathcal{S}_0(\mathbf{c}) \geq \varepsilon\} \subseteq \{\mathbf{c} \mid \mathcal{S}_n(\mathbf{c}) \geq 0\}\Big] \to 1, \quad \text{as } n \to \infty,$$

which implies that

$$\text{pr}\left\{L_n(\hat{\mathbf{c}}) \leq L_n(\hat{\mathbf{c}}_\varepsilon)\right\} \to 1 \quad \text{as } n \to \infty.$$

Furthermore, by the definition of $\hat{\mathbf{c}}_\varepsilon$ which minimizes $L_n(\mathbf{c})$ under the constraint $\mathcal{S}_0(\mathbf{c}) \geq \epsilon$,

$$L_n(\hat{\mathbf{c}}_\varepsilon) \leq L_n(\mathbf{c}_\varepsilon).$$

From the uniform convergence, for any $\delta > 0$,

$$\text{pr}\left\{L_n(\mathbf{c}_\varepsilon) > L(\mathbf{c}_\varepsilon) + \delta/2\right\} \to 0 \quad \text{and} \quad \text{pr}\left\{L(\hat{\mathbf{c}}) - \delta/2 > L_n(\hat{\mathbf{c}})\right\} \to 0 \quad \text{as } n \to \infty.$$

Therefore, for any $\delta > 0$, there exists an ε_0 such that $L(\mathbf{c}_{\varepsilon_0}) \leq L_0 + \delta$ and

$$\begin{aligned}
&\text{pr}(L(\hat{\mathbf{c}}) \leq L_0 + 2\delta) \\
&\geq \text{pr}\left\{L(\hat{\mathbf{c}}) \leq L_n(\hat{\mathbf{c}}) + \delta/2 \leq L_n(\hat{\mathbf{c}}_{\varepsilon_0}) + \delta/2 \leq L_n(\mathbf{c}_{\varepsilon_0}) + \delta/2 \leq L(\mathbf{c}_{\varepsilon_0}) + \delta\right\} \\
&\geq 1 - \text{pr}\left\{L(\hat{\mathbf{c}}) > L_n(\hat{\mathbf{c}}) + \delta/2\right\} - \text{pr}\left\{L_n(\hat{\mathbf{c}}) > L_n(\hat{\mathbf{c}}_{\varepsilon_0})\right\} - \text{pr}\left\{L_n(\mathbf{c}_{\varepsilon_0}) > L(\mathbf{c}_{\varepsilon_0}) + \delta/2\right\} \to 1,
\end{aligned}$$

as $n \to \infty$. It follows that the finite sample optimal stratification scheme minimizes the limit of the total of intra-stratum predicted error. The estimated stratification scheme approaches that of the optimal stratification scheme as the sample size goes to infinity.

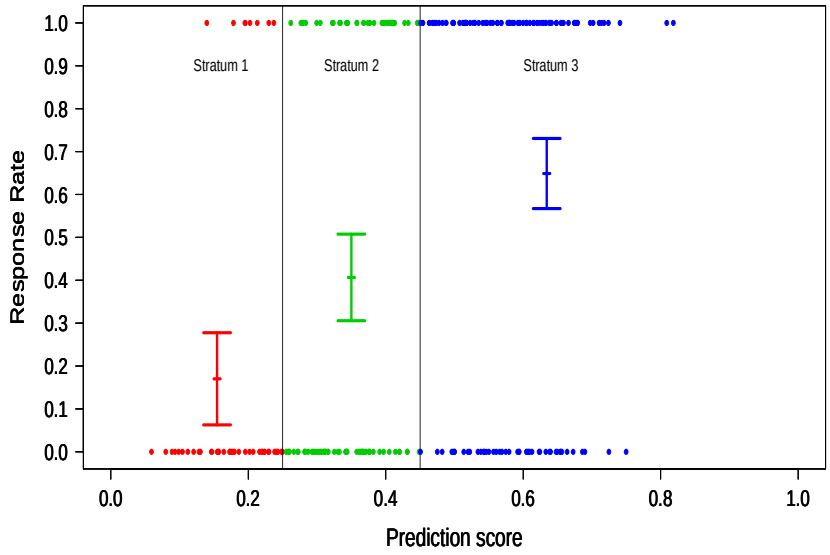

Fig. 1: Stratum-specific point and 95% confidence intervals for the response rates with the Part II data (denoted by dots) of study 320 by AIDS Clinical Trials Group with cutoff points $\hat{c}_1 = 0.25$ and $\hat{c}_2 = 0.45$.