Randomization methods that depend on the covariates

Work by

Alessandro Baldi Antognini
Maroussa Zagoraiou
Alessandra Giovagnoli (*)

Department of Statistical Sciences
University of Bologna, Italy

PHASE-III CLINICAL TRIALS FOR COMPARING TWO TREATMENTS

PROBLEM:

Clinical Trials must

- maximize experimental information
- be randomized, to avoid bias
- minimize potential harm to the patients in the trial

In order to achieve a good trade-off between ethics and inference, allocation of patients to treatments must take into account also covariates

Covariates:

- are usually random (not under the experimenter's control)
- usually categorical or discrete; in continuous case may be discretized
- prestratification wrt the set of covariates of interest: in general not all of the **strata** with the same importance (a-priori information)

The Model

Treatments A and B

$$\delta_i = \begin{cases} 1, \text{ if } i-\text{th patient allocated to } A \\ 0, \text{ if } i-\text{th patient allocated to } B \end{cases}$$

$\mathbf{Z}_i$ = vector of covariates of i-th patient

Linear homoscedastic model

without covariates

$$E(Y_i) = \delta_i \mu_A + (1- \delta_i)\, \mu_B \qquad V(Y_i) = \sigma^2$$

without treatment/covariate interaction:

$$E(Y_i) = \delta_i \mu_A + (1- \delta_i)\, \mu_B + f(\mathbf{z}_i)^t \beta$$

with treatment/covariate interaction:

$$E(Y_i) = \delta_i \mu_A + (1 - \delta_i)\, \mu_B + f(\mathbf{z}_i)^t\, (\delta_i \beta_A + (1- \delta_i)\, \beta_B)$$

μ_A , μ_B baseline treatment effects, $f(\,)$ known vector function

Y_i **conditionally independent**

PARAMETERS OF INTEREST

1. No covariates: μ_A , μ_B
2. NO treat-covar interaction μ_A , μ_B , (β is nuisance)
3. WITH treat-covar interaction (μ_A , μ_B , β_A , β_B)

Variance-covariance matrix of OLS estimators of parameters:

- depends on allocation proportions of A , B within covariate strata
- is random , since covariates are random

parameters of interest $=$ γ

SEQUENTIAL TREATMENT ALLOCATION

IN THE ABSENCE OF COVARIATES

For the k-th patient

At the end of the trial with n assignments

- $n_A = \sum_{k=1}^{n} \delta_k$ number of subjects assigned to A
- $n_B = n - n_A$ number of subjects assigned to B
- $D_n = n_A - n_B$ global imbalance between the 2 groups

Balanced Design (i.e. $D_n = 0$) is **Universally Optimal**:

- $\min \det Var\,(\hat{\mu}_A, \hat{\mu}_B)$ $(D - optimal)$
- $\min Var\,(\hat{\mu}_A - \hat{\mu}_B)$ $(A - optimal)$

and also

- max Power of the test $H_0 : \mu_A = \mu_B$ vs $H_1 : \mu_A > \mu_B$

SEQUENTIAL TREATMENT ALLOCATION

WITH COVARIATES:
MARGINAL & JOINT BALANCE

- without interaction among covariates: **optimality** ⇒ **marginal balance**
  (for each covariate balancing the allocations of A and B within each level)

	Male		Female	
Young	$1A$	$2B$	$2A$	$1B$
Old	$2A$	$1B$	$0A$	$1B$

- with interaction among covariates: **optimality** ⇒ **joint balance**
  (balancing the allocations of A and B within each stratum)

	Male		Female	
Young	$2A$	$2B$	$3A$	$3B$
Old	$1A$	$1B$	$2A$	$2B$

Clearly, balance within strata ⇒ marginal balance (**but not viceversa**)

Minimization by Pocock & Simon (1975): marginal balance

Atkinson's BCD (1982): $\min\{Var\,(\hat{\mu}_A - \hat{\mu}_B)\}$

MEASURES OF IMBALANCE AND PREDICTABILITY

Because of randomization:

- **Inferential Loss** L_n: measures the increase in the variance of parameter estimates due to the imbalance $|D_n|$ caused by randomization

$$L_n = \frac{D_n^2}{n}$$

- **Predictability** SB_n: lack of randomness measured by the expected proportion of correct guesses in n assignments under the optimal strategy (*i.e. always to pick the under-represented treatment*)

$$J_i = \begin{cases} 1 & i\text{-th guess is correct} \\ 0 & \text{otherwise} \end{cases} \quad \Rightarrow \quad G_n = \frac{1}{n}\sum_{i=1}^{n} J_i$$

Selection Bias Indicator: $SB_n = E(G_n) \in \left[\frac{1}{2}\,;\,1\right]$

EFRON'S BCD (Biometrika, 1971)

Efron's BCD(p):

$$\Pr(\delta_{n+1} = 1 \mid n, D_n) = \begin{cases} p & D_n < 0 \\ \frac{1}{2} & D_n = 0 \\ 1-p & D_n > 0 \end{cases}, \qquad p \in \left[\tfrac{1}{2}; 1\right]$$

- favors the under-represented treatment
- the allocation depends only on the presence of imbalance (sign of D_n)
- does not depend on the degree of imbalance (value of D_n)

$p = 1 \Rightarrow$ Deterministic

$p = \frac{1}{2} \Rightarrow$ Simple Random Sampling

Efron suggests taking $p = \frac{2}{3}$

BALDI ANTOGNINI AND GIOVAGNOLI'S BCD (ABCD) (JRSSC, 2004)

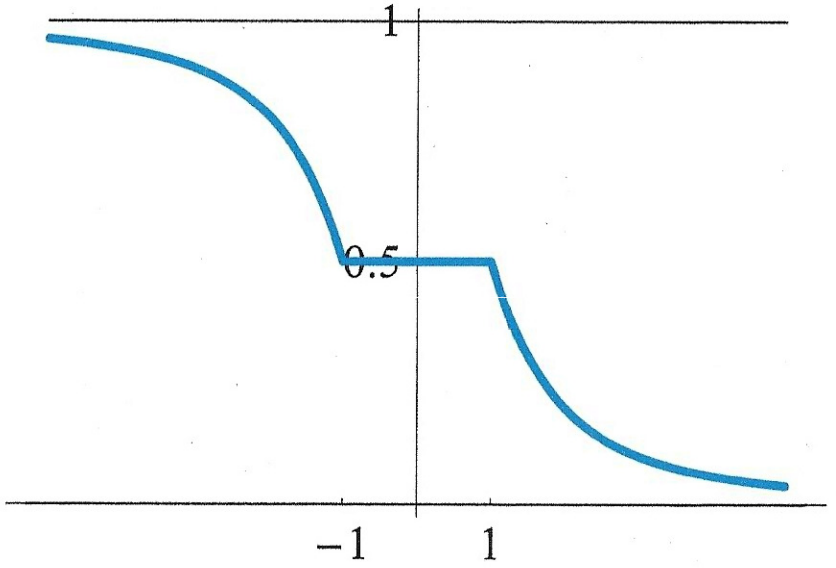
Let $F : \mathbf{Z} \to [0, 1]$ decreasing and such that $F(-x) = 1 - F(x)$

ABCD: $\Pr(\delta_{n+1} = 1 \mid n, D_n) = F(D_n)$

$\Rightarrow$ it favors the under-represented treatment increasingly as $|D_n|$ grows

---

E.g. F can be chosen s.t.



the rationale is to treat the case $D_n = \pm 1$ (i.e. n odd) as if the design was balanced, and to redress the balance in all the other cases

CATEGORICAL COVARIATES

Wlog, assume there are just 2 covariates, U and W, with levels

$$\{(u,w) \quad u = 0,1,\ldots,J \quad \text{and } w = 0,1,\ldots,L\}$$

Let the covariates' distribution in the population be

$$p_{jl} = \Pr(U = j,\ W = l)$$

This is assumed unknown, to be estimated from the experiment

The sample is divided into **strata** $j \times l$ of patients.

After n steps

$N_n(j,l)$ = # of patients with covariates $U = j,\ W = l$ (stratum $j \times l$).

$\pi_n(j,l)$ = proportion of allocations to ***A*** in stratum $j \times l$

Global Imbalance	D_n
Marginal Imbalance	$D_n(j),\ D_n(l)$
Joint Imbalance	$D_n(j,l)$

COVARIATE ADAPTIVE BCD

For each stratum (j,l) define $F_{jl} : \mathbf{Z} \to [0,1]$ decreasing and symmetric with

$$F_{jl}(-x) = 1 - F_{jl}(x)$$

At step k, the k+1- subject with profile (j,l) is assigned to A with probability

$$F_{jl}(D_k(j,l))$$

where $D_k(j,l)$ is the current imbalance within stratum $j \times l$

***Loss** measured by*

$$L_n = \sum_{j=1}^{J} \sum_{l=1}^{L} \frac{D_n^2(j,l)}{N_n(j,l)}$$

if model contains all interaction effects

$\Rightarrow$ *Strong departures from balance may not have a great impact on the loss if* observed for covariate profiles that occur frequently in the population

PROPERTIES OF CA-BCD

- If $F_{jl}(\cdot)$'s are the same $\forall(j,l)$, balance is forced in the same way in each stratum, for each level of the covariates and also globally.

  Different $F_{jl}(\cdot)$'s reflect relative importance of strata due for intance to *a-priori* information, e.g.

$$F_{jl}(x) = \begin{cases} \frac{1}{2} & x \in [0;1] \\ (1+x^{g_{jl}})^{-1} & x \geq 1, \end{cases} \qquad \text{where} \qquad g_{jl} \propto p_{jl}^{-1}$$

- Global, marginal e joint relative imblance converge to 0 a.s. Faster convergence to balance of the C-ABCD compared to competing designs

- Loss of efficiency L_n vanishes very rapidly in probability

FINITE SAMPLE PROPERTIES

Expectation and variance (within brackets) of L_n and SB_n, for $n = 150$ and 500, under the full model with $p_{00} = 0.2$, $p_{01} = 0.4$, $p_{10} = 0.3$, $p_{11} = 0.1$

(1000 SIMULATED RUNS)

	L_{150}	SB_{150}	L_{500}	SB_{500}
Pocock& Simon	1.09 (1.7392)	0.70 (0.0011)	1.02 (1.8610)	0.71 (0.0003)
Atkinson's BCD	0.82 (0.3458)	0.55 (0.0010)	0.81 (0.3308)	0.53 (0.0003)
C-A BCD(F_{jl}^g)	0.24 (0.0213)	0.61 (0.0013)	0.07 (0.0018)	0.61 (0.0004)

COMBINING ETHICS AND INFERENCE: A TWO-STEP APPROACH

WITHOUT COVARIATES

1. Find *target* treatment allocation which is optimal w.r.t. a compound criterion that combines inferential precision and number of patients receiving the better treatment

→ Compound Optimal Target Allocation

2. Allocate patients sequentially according to a randomized rule such that the treatment allocation converges to the target

(see Baldi Antognini & Giovagnoli, 2010)

WITH COVARIATES

Same procedure, possibly with different degrees of randomness according to each patient's covariates

Ethical Aim

To maximize the expected percentage of patients who receive the better treatment

$\theta(\mathbf{z})$ = expected difference between treatment effects

$$|\theta(\mathbf{z})| = ethical\ gain$$

Without treatment/covariate interaction:

$$\theta(\mathbf{z}) = \mu_A - \mu_B$$

The superiority of *A* or *B* is uniformly constant over the covariates

With treatment/covariate interaction

The relative performance of the treatments depends on the patient's covariates

$$\theta(\mathbf{z}) = E(Y_i|\delta_i = 1, \mathbf{Z}_i = \mathbf{z}) - E(Y_i|\delta_i = 0, \mathbf{Z}_i = \mathbf{z})$$
$$= \mu_A - \mu_B + \mathbf{f}(\mathbf{z})^t(\boldsymbol{\beta}_A - \boldsymbol{\beta}_B)$$

Optimal allocation for ethical concern

% of subjects to the better treatment: $\frac{1}{n}\sum_{j=0}^{J}\sum_{l=0}^{L} N(j,l)\left\{\frac{1}{2}-\left(\frac{1}{2}-\pi(j,l)\right)sgn(\theta(j,l))\right\}$

Ethical criterion: $\tilde{\Phi}^{E}(\boldsymbol{\pi})=\sum_{j=0}^{J}\sum_{l=0}^{L}|\theta(j,l)|\,p_{jl}\left\{\frac{1}{2}-\left(\frac{1}{2}-\pi(j,l)\right)sgn(\theta(j,l))\right\}$

Optimal Ethical Target: to assign all patients to the better treatment (for each patient)

WITHOUT treatment/cov interaction $\pi_E^*(j,l)=\left(1_{\{\mu_A-\mu_B>0\}}\right)\ \forall(j,l)$

it does not depend on the covariates

WITH treatment/cov interaction $\pi_E^*(j,l)=\left(1_{\{\theta(j,l)>0\}}\right)\ \forall(j,l)$

Optimality for inference

Usual optimality criteria Φ^I (D-optimality, A-optimality…) applied to Var-Covar of OLS gives random quantity.

We take expected value with respect to the distribution of the covariates

$$\tilde{\Phi}^I(\boldsymbol{\pi}) = E_{\mathbf{Z}}\left(\Phi^I(\boldsymbol{\pi})\right) \propto \left(\prod_{j=0}^{J} \prod_{l=0}^{L} \pi(j,l)(1-\pi(j,l))\ n^2 p(j,l)^2 \right)^{-1}$$

Compound Optimal Designs

Compound criterion: combines ethical gain and inferential precision by flexible weights ω, 1-ω

Minimize

$$\Psi_{\omega}(\boldsymbol{\pi}) = \omega \cdot \left(\Psi^{E}(\boldsymbol{\pi})\right)^{-1} + (1-\omega) \cdot \left(\Psi^{I}(\boldsymbol{\pi})\right)^{-1}$$

Ψ^{E}, Ψ^{I} are the ethical and inferential criteria $\tilde{\Phi}^{I}(\boldsymbol{\pi})$ and $\tilde{\Phi}^{E}(\boldsymbol{\pi})$ standardized in [0,1]

The weight ω is chosen on the basis of the relative importance of the two criteria and can be fixed a priori or modelled as an increasing function of $E_{\boldsymbol{Z}}(|\theta(\boldsymbol{z})|)$

Adaptive Randomization With treatment-covariates interaction

Optimal target allocation depends on the stratum covariates, the inferential criterion and the unknown parameters

Modify the randomization probabilities as the experiment goes along in order to gradually approach the desired target

Randomization probabilities for the current patient will depend on

previous patients' treatment assignments;

previous responses;

previous covariates;

and the **covariates** of the **current patient**

New method

Reinforced Doubly –Adaptive Biased Coin Design (RD-BCD)
(Baldi Antognini & Zagoraiou, 2012)

THE REINFORCED DOUBLY-ADAPTIVE BIASED COIN DESIGN (RD-BCD)

Start with n_0 initial observation and estimates,

At each step $k > n_0$

- $\pi_k(j,l)$ is proportion of allocations to A in stratum j x l
- $\hat{\boldsymbol{\gamma}}_k$ is estimate of unknown parameters
- $\hat{\mathbf{p}}_k$ is estimate of stratum probabilities

the optimal target $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(\boldsymbol{\gamma}, \mathbf{p})$ can be estimated by $\hat{\boldsymbol{\pi}}_k^* = \boldsymbol{\pi}^*(\hat{\boldsymbol{\gamma}}_k, \hat{\mathbf{p}}_k)$

For suitable choice of function $\varphi(\bullet,\bullet,\bullet)$ current patient at step k+1 with covariates (j,l) will be allocated to A with probability

$$\varphi\left(\pi_k(j,l);\ \hat{\pi}_k^*(j,l);\ \hat{\mathrm{p}}_k(j,l)\right)$$

The Reinforced Doubly-Adaptive BCD

Definition of $\varphi(x,y,z)$

$$\varphi: (0,1)^3 \rightarrow [0,1]$$

1. $\varphi \searrow$ in x
2. $\varphi \nearrow$ in y
3. $\varphi(x, x, z) = x$ for all z
4. $\varphi(x, y, z)$ $\begin{cases} \searrow \text{ in } z \text{ if } x < y \\ \nearrow \text{ in } z \text{ if } x > y \end{cases}$
5. $\varphi(x, y, z) = 1 - \varphi(1-x, 1-y, z)$ for all z

NB A **different** φ *for each stratum can also be defined* $\varphi_{ij}(x,y,z)$

NB **The RD-BCD can be extended to several treatments**

PROPERTIES OF THE RD-BCD

- The allocation proportion converges to the target almost surely
- The estimated target converges to the desired target almost surely
- The estimators of the parameters are strongly consistent and asymptotically normal

SPECIAL CASE OF THE RD-BCD

Zhang *et al.*, Annals 2007 Covariate Adaptive Sequential Maximum Likelihood Design

$$\text{Target } \boldsymbol{\pi}^* = \boldsymbol{\pi}^*(\boldsymbol{\gamma}) \text{ continuous function of } \boldsymbol{\gamma}$$

at each step k, estimate $\boldsymbol{\gamma}$ with all the collected data up to that step by $\hat{\boldsymbol{\gamma}}_k$
$\Rightarrow$ the optimal target can be estimated by $\hat{\boldsymbol{\pi}}_k^* = \boldsymbol{\pi}^*(\hat{\boldsymbol{\gamma}}_k)$

$(k+1)$st patient with covariate $Z_{k+1} = (j,l)$
will be allocated to A with probability $\hat{\pi}_k^*(j,l)$

Drawbacks:

slower convergence, strong variability for small sample sizes

References

- A. C. Atkinson. 1982 *Optimum biased coin designs for sequential clinical trials with prognostic factors.* **Biometrika**, 69:61–67,
- Baldi Antognini A. and Giovagnoli A. (2010). *Compound optimal allocation for individual and collective ethics in binary clinical trials,* **Biometrika**, 97: 935-946.
- Baldi Antognini A. and Zagoraiou M. (2011). *The covariate-adaptive biased coin design for balancing clinical trials in the presence of prognostic factors,* **Biometrika**, 98: 519-535.
- Baldi Antognini A. and Zagoraiou M. (2012). *Multi-objective optimal designs in comparative clinical trials with covariates: the reinforced doubly-adaptive biased coin design,* **The Annals of Statistics**, 40: 1315–1345
- Hu F., Zhang L.X. and He, X. (2009). *Efficient randomized-adaptive designs*, **The Annals of Statistics** 37: 2543-2560.
- J. Pocock and R. Simon. 1975.*Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial*. **Biometrics**, 31:103–115
- Zhang L.X., Hu F., Cheung S.H. and Chan W.S. (2007). *Asymptotic properties of covariate-adjusted response-adaptive designs,* **The Annals of Statistics** 35: 1166-1182.