Conditional Inference and Information in Response-Adaptive Allocation Designs

Adam Lane

December 2, 2020

Outline

SETTING

Phase II and III clinical trials

① Study efficacy of a set of treatments.

② An experimental objective is to precisely estimate the treatment success probabilities.

③ An ethical objective is to minimize the number of observed treatment failures.

Balancing Efficiency and Ethics

Response-adaptive allocation designs are a popular schema for balancing experimental and ethical objectives.

1. Alter the probability an observation is assigned to the available treatments based on the accrued responses.
2. More weight can be added to the treatments that have been observed to be more successful.

EXAMPLE: RANDOMIZED PLAY THE WINNER RULE

The Randomized Play the Winner Rule [RPW(α,β)] is an adaptive urn design

1. It is defined for two treatments 1 and 2 with success probabilities p_1 and p_2, respectively.
2. Initially an urn is comprised of α balls of type 1 and 2, representing the corresponding treatments.
3. If a success is observed on treatment 1 or a failure is observed on treatment 2 then β type 1 balls are added to the urn.
4. Otherwise β type 2 balls are added.
5. Can significantly reduce the total number of expected treatment failures.

EXAMPLE: SUCCESS DRIVEN DESIGN

A success driven design [SDD(α,β)] is also an adaptive urn design

1. Initializes in the same manner as the RPW rule.
2. If a success is observed on treatment $k(=1,2)$ then β type k balls are added to the urn.
3. Otherwise the urn remains unchanged.
4. This design is also aimed at reduce the total number of expected treatment failures.

NOTATION

In general we are tasked with assigning n subjects to K treatments.

① Responses are observed sequentially.

② The response of the ith the subject is denoted $Y_i (= 0, 1)$, $i = 1, \ldots, n$.

③ The treatment received by the ith subject is described by a vector $\boldsymbol{T}(i) = (T_{i1}, \ldots, T_{iK})$, where $T_{ik} = 1$ if treatment k is received by subject i and 0 otherwise. Note $\sum_{k=1}^{K} T_{ik} = 1$.

④ The total number of success on treatment k from the first i subjects is $S_k(i) = \sum_{j=1}^{i} Y_j T_{jk}$, let $\boldsymbol{S}(i) = \{S_1(i), \ldots, S_K(i)\}^T$.

⑤ The corresponding number of the first i subjects receiving treatment k is denoted $N_k(i) = \sum_{j=1}^{i} T_{jk}$ and $\boldsymbol{N}(i) = \{N_1(i), \ldots, N_K(i)\}^T$.

RESPONSE-ADAPTIVE ALLOCATION DESIGN

A response adaptive allocation design is defined by a rule for assigning subjects to treatment given the observed data.

① Let $\boldsymbol{X}(i) = \{\boldsymbol{S}(i), \boldsymbol{N}(i)\}$ and $\mathscr{F}(i) = \sigma\{\boldsymbol{X}(1), \ldots, \boldsymbol{X}(i)\}$.

② Define the allocation rule as

$$\xi_{ik} = P\{T_{ik} = 1 | \mathscr{F}(i-1)\}$$

for $i = 1, \ldots, n$ and $k = 1, \ldots, K$.

③ This probability can have support [0,1]

④ The case of 0,1 is a deterministic design.

⑤ If it is (0,1) it is a randomized design.

REQUIRED CONDITIONS

In this work we require the following conditions throughout.

① The total sample size, n, is fixed.

② The number of treatments, K, is finite.

③ The conditional success probabilities satisfy $P\{Y_i = 1 | T_{i-1,k} = 1\} = p_k \in (0,1)$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, K$.

④ Conditional on the observed data from the first $i-1$ subjects the distribution of $P\{T_{ik} = 1 | \mathscr{F}(i-1)\}$ does not depend on $\boldsymbol{p} = (p_1, \ldots, p_K)$.

Joint Likelihood

It is well documented that the joint likelihood from an adaptive experiment is the same as if the experiment contained no adaptation, i.e.,

$$L_f\{\boldsymbol{p}; \boldsymbol{s}(n), \boldsymbol{n}(n)\} = \prod_{k=1}^{K} p_k^{s_k(n)} (1-p_k)^{n_k(n)-s_k(n)},$$

where $\boldsymbol{s}(n) = (s_1, \ldots, s_K)^T$ and $\boldsymbol{n}(n) = (n_1, \ldots, n_K)^T$ are the observed realizations of the random variable $\boldsymbol{S}(n)$, $\boldsymbol{N}(n)$.

① This is identical to the likelihood of an experiment with n independent bernoulli trials.

② Only the likelihoods are the same; in response-adaptive allocation designs the minimal sufficient statistic for $\boldsymbol{p}$ is $\boldsymbol{X}(n)$.

③ The distribution of $\boldsymbol{X}(n)$ differs significantly from a design with n independent distributed bernoulli trials.

The Unconditional Maximum Likelihood Estimate (MLE)

The score function, the derivative of the log-likelihood is

$$\boldsymbol{\Psi}(\hat{\boldsymbol{p}}, \boldsymbol{n}) = \frac{\partial}{\partial \boldsymbol{p}} \log L_f(\boldsymbol{p}; \boldsymbol{s}, \boldsymbol{n}) = \Sigma_{\boldsymbol{n}}^{-1}(\hat{\boldsymbol{p}} - \boldsymbol{p}),$$

① $\hat{\boldsymbol{p}} = (\hat{p}_1, \ldots, \hat{p}_K)^T$, where $p_k = S_k(n)/N_k(n)$.

② $\Sigma_{\boldsymbol{n}}$ is a $K \times K$ diagonal matrix with kth diagonal entry $p_k(1 - p_k)/n_k(n)$.

③ setting $\boldsymbol{\Psi} = 0$ we are able to see that $\hat{\boldsymbol{p}}$ is the unconditional MLE of $\boldsymbol{p}$.

④ $\hat{\boldsymbol{p}}$ is not 1:1 with the minimal sufficient statistic $\boldsymbol{X}(n)$ and is not sufficient.

Outline

CONDITIONAL VERSUS UNCONDITIONAL

Here conditional implies conditional on the vector of treatment sample sizes $\boldsymbol{N}(n)$.

1. Wei et. al. (1990) heuristically found exact unconditional inference to be superior to conditional inference for the RPW design.
2. Intuitively, as Wei et. al. (1990) point out, the treatment sample sizes contain evidence regarding the effectiveness of the treatments; therefore, conditioning represents a significant loss of information.
3. The predominant opinion seems to me to be conditional information represents a loss of information.

Conditional versus Unconditional (cont.)

Begg (1990) presents some negative aspects of unconditional inference for the RPW design.

1. It is difficult to order the unobserved points in the sample space.
2. It is not necessarily sufficient to compare the observed $\hat{p}_k$ to the the corresponding $\hat{p}_k$ from an unobserved point in the sample space to reach a conclusion regarding the relative evidence in favor or against of a specific hypothesis.
3. This is a consequence of the MLE not being a sufficient statistic.
4. This ordering problem does not occur on the conditional space.

Unconditional Information

There are two common information quantities considered; expected and observed.

① The total expected Fisher information of the joint distribution is by definition the variance of the score function, i.e.,

$$I_{S,N}(\boldsymbol{p}) = Var[\boldsymbol{\Psi}].$$

② This will be referred to as the *unconditional expected information*.

UNCONDITIONAL INFORMATION

① Observed Fisher information is the the second derivative of the log-likelihood evaluated at the MLE

$$J_{S,N}(\hat{\boldsymbol{p}}) = \frac{\partial}{\partial \boldsymbol{p}} \boldsymbol{\Psi}(\hat{\boldsymbol{p}}, \boldsymbol{n}) \Big|_{\boldsymbol{p}=\hat{\boldsymbol{p}}} = \hat{\Sigma}_{\boldsymbol{n}}^{-1},$$

where $\hat{\Sigma}_{\boldsymbol{n}}$ is a $K \times K$ diagonal matrix with kth diagonal entry $\hat{p}_k(1-\hat{p}_k)/n_k(n)$.

② We will refer to this as *unconditional observed information*.

③ Under certain regularity conditions the unconditional expected information can be written

$$I_{S,N}(\boldsymbol{p}) = E[\Sigma_{\boldsymbol{n}}^{-1}].$$

Conditional Information

Define the conditional likelihood as

$$L_c(\boldsymbol{p}; \boldsymbol{s}, \boldsymbol{n}) = P\{\boldsymbol{S}(n) = \boldsymbol{s} | \boldsymbol{N}(n) = \boldsymbol{n}\}.$$

The conditional score function is

$$\boldsymbol{\Gamma}(\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{n}) = \frac{\partial}{\partial \boldsymbol{p}} \log L_c(\boldsymbol{p}; \boldsymbol{s}, \boldsymbol{n}).$$

Conditional expected and observed information are

$$I_{S|N}(\boldsymbol{p}) = Var[\boldsymbol{\Gamma} | \boldsymbol{N}(n)]$$

and

$$J_{S|N}(\hat{\boldsymbol{p}}) = \frac{\partial}{\partial \boldsymbol{p}} \boldsymbol{\Gamma}(\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{n}) \Big|_{\boldsymbol{p}=\hat{\boldsymbol{p}}}$$

respectively.

Conditional Information

Under the stated conditions and additionally if $1 \leq n_k(n)$ or all $k = 1, \ldots, K$ then

① the *conditional expected information* is

$$I_{S|N}(\boldsymbol{p}) = \Sigma_{\boldsymbol{n}}^{-1} Var[\hat{\boldsymbol{p}}|\boldsymbol{N}(n)]\Sigma_{\boldsymbol{n}}^{-1}$$

② and the *conditional observed information* is

$$J_{S|N}(\hat{\boldsymbol{p}}) = I_{S|N}(\hat{\boldsymbol{p}}) - D(\hat{\boldsymbol{p}})E[\hat{\boldsymbol{p}} - \boldsymbol{p}|\boldsymbol{N}(n)]J_{S,N}(\hat{\boldsymbol{p}}),$$

where $D(\boldsymbol{p}) = \mathrm{E}_{\boldsymbol{p}}\left[(\hat{\boldsymbol{p}} - \boldsymbol{p})|\boldsymbol{N}(n) = \boldsymbol{n}\right]\left[(\partial/\partial\boldsymbol{p})\Sigma_{\boldsymbol{n}}^{-1}\right]$.

RELATIVE EFFICIENCY

① The *observed relative efficiency*, defined as

$$J_{S|N}(\hat{\boldsymbol{p}})[J_{S,N}(\hat{\boldsymbol{p}})]^{-1}$$

which can be shown to be symmetric.

② To interpret Wei et. all (1990) statements regarding information more rigourously we might say that it is implied that observed relative efficiency must be less than $\boldsymbol{I}_K$.

③ Under fairly reasonable conditions, satisfied by both the SDD and RPW designs, the asymptotic relative efficiency

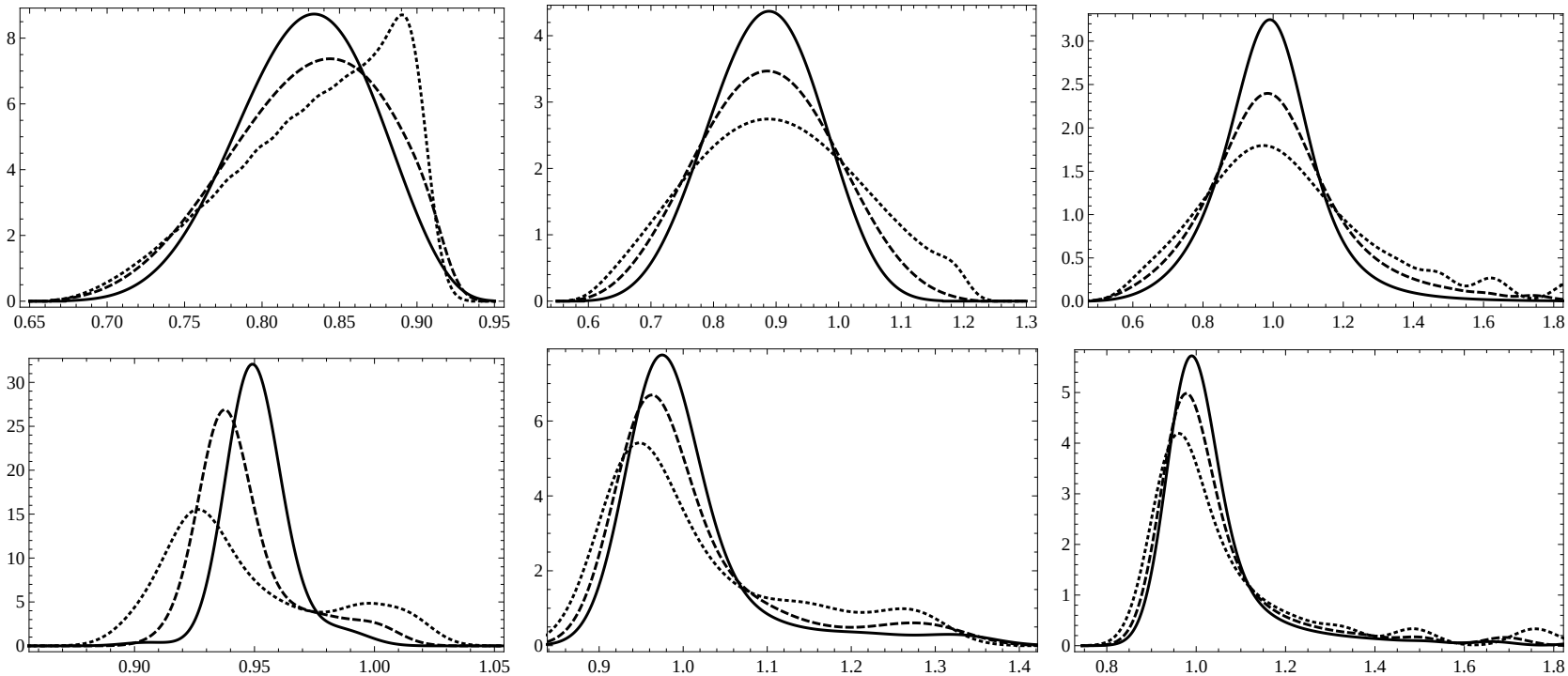
$$J_{S|N}(\hat{\boldsymbol{p}})[J_{S,N}(\hat{\boldsymbol{p}})]^{-1} - Var[\boldsymbol{Z}_n|\boldsymbol{N}(n)] \to 0$$

in probability as $n \to \infty$, where $\boldsymbol{Z}_n = (Z_{n1} \ldots, Z_{nK})$ and

$$Z_{nk} = \sqrt{N_k(n)}\frac{\hat{p}_k - p_k}{\sqrt{p_k(1-p_k)}}.$$

DISTRIBUTION RESULTS

Smoothed histograms of the distribution of $\frac{1}{2}\mathrm{Tr}\{\mathrm{Var}[\boldsymbol{Z}_n|\boldsymbol{N}(n)]\}$ for the RPW (top) and SDD (bottom). In each row values of $p_1 = p_2 = 0.5$, 0.7 and 0.9 (left to right) were used. Within each figure each line represents a different sample size, $n = 25$ (dotted), $n = 50$ (dashed) and $n = 100$ (solid). Values larger than 1 correspond to situations where the conditional information is greater than unconditional information.



Outline

Conditional MLE

The conditional MLE (CMLE) defined as the solution to

$$\tilde{\boldsymbol{p}} = \arg\max_{\boldsymbol{p}} L_c(\boldsymbol{p}; \boldsymbol{s}, \boldsymbol{n}).$$

The above, with respect to $\boldsymbol{p}$, is equivalent to

$$\hat{\boldsymbol{p}} = \mathrm{E}_{\boldsymbol{p}} \left[\hat{\boldsymbol{p}} | \boldsymbol{N}(n) = \boldsymbol{n} \right].$$

BOOTSTRAP INTERVAL ESTIMATION

1. Rosenberger and Hu (2001) propose an unconditional bootstrap to analyze randomized allocation designs.
   1. Generate a bootstrap distribution by replicating the response adaptive allocation design using the observed success probabilities $\hat{\boldsymbol{p}}$ in place of the true success probabilities $\boldsymbol{p}$
   2. For each bootstrap replication denote $\boldsymbol{S}^*$ and $\boldsymbol{N}^*$ as the vectors of successes and sample sizes observed in the replication.
   3. Then the percentile bootstrap intervals are defined as the $\alpha/2$ and $(1-\alpha/2)$ quantiles of the bootstrap distribution of $\hat{\boldsymbol{p}}^*$.
2. An obvious analog to the unconditional bootstrap procedure is to repeat the procedure finding $\tilde{\boldsymbol{p}}^*$ in place of $\hat{\boldsymbol{p}}^*$ in each bootstrap replication and then base the intervals on the relevant conditional quantiles of $\tilde{\boldsymbol{p}}^*$.
3. The computation time required to compute $\tilde{\boldsymbol{p}}^*$ makes such a procedure impractical even for moderate sample sizes.
4. To reduce computation time a conditional bootstrap procedure is developed that exploits the monotonicity of $\tilde{p}_k$ with respect to $S_k(n)$ given $N_k(n)$.

CONFIDENCE INTERVALS FOR THE CMLE

The following modified conditional bootstrap procedure can be used to generate conditional confidence intervals for $p_k, k = 1, \ldots, K$.

① Generate B replicates of the response adaptive allocation design using the observed success probabilities $\hat{\boldsymbol{p}}$ in place of the true success probabilities $\boldsymbol{p}$.

② Compute S_k^* and $\boldsymbol{N}^*$ for each replication. Delete bootstrap replicates such that $\boldsymbol{N}^*$ is not equal to $\boldsymbol{N}(n)$. This creates a conditional bootstrap distribution. Let B_C denote the number of bootstrap samples satisfying the conditional requirement.

③ Let $S_k^{*(1)}, \ldots, S_k^{*(B_C)}$ be the ordered conditional bootstrap samples. Let $\boldsymbol{S}^{[k]}(n)$ be the vector of observed successes with $S_k(n)$ removed.

④ Let $\tilde{p}_k^z$ denote the CMLE calculated using $[\boldsymbol{S}^{[k]}(n), S_k^{*(B_C z)}, \boldsymbol{N}(n)]$, where $z \in (0, 1)$. The interval $(\tilde{p}_k^{\alpha/2}, \tilde{p}_k^{1-\alpha/2})$ is the $1 - \alpha$ two-sided conditional confidence interval for p_k.

The above bootstrap procedure creates intervals for each $p_k, k = 1, \ldots, K$ separately.

Outline

Comparative Methods

1. Conditional Bootstrap.
2. Wald confidence intervals with continuity correction.
3. Rosenberger and Hu (2001) unconditional bootstrap.

Reported Characteristics

① Relative length defined as $\text{RL}_k = L_{\hat{p}_k}/L_{\tilde{p}_k}$, where $L_{\hat{p}_k}$ and $L_{\tilde{p}_k}$ are the the 95% Bonferonni simultaneous unconditional and conditional intervals for $p_k, k = 1, 2$, respectively.

② The interpretation of the RL is that values greater than 1 indicate that conditional inference is better for the respective measure.

③ Coverage of the of the simultaneous intervals.

④ Values of $\alpha = \beta = 1$, $n = 25, 50$, $p_1 = 0.1, 0.3, 0.5, 0.7, 0.9$, $p_2 = 0.1, 0.3, 0.5, 0.7, 0.9$ and $\alpha = 0.05$

⑤ Prior to initializing the adaptive designs one sample was place on each of the available treatments. This is to avoid cases where the MLE is undefined $N_1(n), N_2(n) = 0$.

Results: SDD $n = 25$

The coverage of the conditional bootstrap, Wald and unconditional intervals are denoted $C_{\tilde{p}}$, $C^W_{\hat{p}}$ and $C^B_{\hat{p}}$. The length of the Wald and unconditional bootstrap intervals relative to the conditional bootstrap are denoted RL^W_k and RL^B_k.

p_1	p_2	RL^B_1	RL^W_1	RL^B_2	RL^W_2	$C_{\tilde{p}}$	$C^B_{\hat{p}}$	$C^W_{\hat{p}}$
0.1	0.1	0.84	1.29	0.84	1.29	0.9952	0.9973	0.9988
0.3	0.1	0.98	1.18	0.85	1.33	0.9816	0.9337	0.9802
0.3	0.3	0.98	1.22	0.98	1.22	0.9760	0.9466	0.9749
0.5	0.1	1.11	1.13	0.85	1.38	0.9598	0.9378	0.9620
0.5	0.3	1.15	1.17	0.97	1.27	0.9600	0.9419	0.9603
0.5	0.5	1.16	1.21	1.16	1.21	0.9484	0.9289	0.9479
0.7	0.1	1.14	1.15	0.87	1.43	0.9541	0.9516	0.9667
0.7	0.3	1.24	1.17	0.97	1.32	0.9495	0.9412	0.9618
0.7	0.5	1.33	1.20	1.15	1.26	0.9394	0.9289	0.9495
0.7	0.7	1.37	1.23	1.37	1.23	0.9348	0.8995	0.9512
0.9	0.1	1.08	1.26	0.87	1.48	0.9891	0.9885	0.9975
0.9	0.3	1.16	1.27	0.95	1.38	0.9850	0.9806	0.9925
0.9	0.5	1.27	1.29	1.11	1.31	0.9698	0.9733	0.9807
0.9	0.7	1.38	1.30	1.32	1.28	0.9643	0.9154	0.9801
0.9	0.9	1.42	1.33	1.42	1.33	0.9687	0.9248	0.9968

Results: SDD $n = 50$

The coverage of the conditional bootstrap, Wald and unconditional intervals are denoted $C_{\tilde{p}}$, $C^W_{\hat{p}}$ and $C^B_{\hat{p}}$. The length of the Wald and unconditional bootstrap intervals relative to the conditional bootstrap are denoted RL^W_k and RL^B_k.

p_1	p_2	RL^B_1	RL^W_1	RL^B_2	RL^W_2	$C_{\tilde{p}}$	$C^B_{\hat{p}}$	$C^W_{\hat{p}}$
0.1	0.1	0.84	1.17	0.84	1.17	0.9955	0.9984	0.9981
0.3	0.1	1.08	1.07	0.80	1.23	0.9649	0.9196	0.9605
0.3	0.3	1.05	1.12	1.06	1.12	0.9592	0.8547	0.9520
0.5	0.1	1.10	1.06	0.79	1.29	0.9668	0.9546	0.9704
0.5	0.3	1.22	1.09	0.99	1.17	0.9575	0.9012	0.9564
0.5	0.5	1.26	1.12	1.26	1.12	0.9472	0.8732	0.9456
0.7	0.1	1.06	1.08	0.80	1.35	0.9653	0.9638	0.9737
0.7	0.3	1.17	1.09	0.95	1.23	0.9589	0.9421	0.9635
0.7	0.5	1.32	1.12	1.22	1.17	0.9449	0.8890	0.9477
0.7	0.7	1.44	1.15	1.45	1.15	0.9339	0.8753	0.9440
0.9	0.1	1.03	1.15	0.81	1.39	0.9496	0.9446	0.9520
0.9	0.3	1.07	1.16	0.93	1.28	0.9542	0.9578	0.9529
0.9	0.5	1.16	1.18	1.15	1.21	0.9538	0.9445	0.9478
0.9	0.7	1.29	1.21	1.45	1.19	0.9511	0.9065	0.9517
0.9	0.9	1.39	1.25	1.40	1.25	0.9721	0.9295	0.9764

Results: RPW $n = 25$

The coverage of the conditional bootstrap, Wald and unconditional intervals are denoted $C_{\tilde{p}}$, $C^W_{\hat{p}}$ and $C^B_{\hat{p}}$. The length of the Wald and unconditional bootstrap intervals relative to the conditional bootstrap are denoted RL^W_k and RL^B_k.

p_1	p_2	RL^B_1	RL^W_1	RL^B_2	RL^W_2	$C_{\tilde{p}}$	$C^B_{\hat{p}}$	$C^W_{\hat{p}}$
0.1	0.1	0.84	1.23	0.84	1.23	0.9935	0.9968	0.9989
0.3	0.1	0.90	1.11	0.84	1.26	0.9779	0.9229	0.9762
0.3	0.3	0.90	1.14	0.90	1.14	0.9817	0.8959	0.9693
0.5	0.1	0.93	1.07	0.83	1.30	0.9747	0.9529	0.9632
0.5	0.3	0.95	1.10	0.91	1.19	0.9722	0.9384	0.9555
0.5	0.5	0.97	1.14	0.97	1.14	0.9551	0.9118	0.9338
0.7	0.1	0.96	1.07	0.83	1.35	0.9701	0.9581	0.9717
0.7	0.3	0.98	1.09	0.91	1.25	0.9674	0.9495	0.9634
0.7	0.5	1.01	1.12	1.00	1.19	0.9526	0.9096	0.9421
0.7	0.7	1.08	1.17	1.08	1.17	0.9468	0.8945	0.9444
0.9	0.1	0.99	1.16	0.83	1.43	0.9912	0.9872	0.9976
0.9	0.3	1.01	1.18	0.90	1.35	0.9879	0.9835	0.9930
0.9	0.5	1.04	1.20	1.02	1.28	0.9785	0.9732	0.9812
0.9	0.7	1.10	1.24	1.16	1.24	0.9730	0.9182	0.9796
0.9	0.9	1.20	1.28	1.20	1.28	0.9700	0.9201	0.9967

Results: RPW $n = 50$

The coverage of the conditional bootstrap, Wald and unconditional intervals are denoted $C_{\tilde{p}}$, $C^W_{\hat{p}}$ and $C^B_{\hat{p}}$. The length of the Wald and unconditional bootstrap intervals relative to the conditional bootstrap are denoted RL^W_k and RL^B_k.

p_1	p_2	RL^B_1	RL^W_1	RL^B_2	RL^W_2	$C_{\tilde{p}}$	$C^B_{\hat{p}}$	$C^W_{\hat{p}}$
0.1	0.1	0.85	1.13	0.87	1.13	0.9931	0.9940	0.9983
0.3	0.1	0.88	1.01	0.85	1.17	0.9754	0.9527	0.9652
0.3	0.3	0.88	1.05	0.89	1.05	0.9506	0.8872	0.9331
0.5	0.1	0.90	0.98	0.84	1.21	0.9826	0.9640	0.9750
0.5	0.3	0.91	1.01	0.88	1.08	0.9528	0.9028	0.9423
0.5	0.5	0.91	1.05	0.92	1.05	0.9578	0.9127	0.9461
0.7	0.1	0.90	0.97	0.84	1.26	0.9836	0.9667	0.9769
0.7	0.3	0.94	1.01	0.88	1.13	0.9583	0.9066	0.9492
0.7	0.5	0.96	1.04	0.92	1.08	0.9546	0.9093	0.9427
0.7	0.7	0.97	1.07	0.98	1.07	0.9489	0.8971	0.9430
0.9	0.1	0.93	1.04	0.81	1.33	0.9554	0.9510	0.9607
0.9	0.3	0.96	1.07	0.90	1.23	0.9542	0.9445	0.9502
0.9	0.5	1.00	1.10	1.01	1.16	0.9527	0.9260	0.9385
0.9	0.7	1.04	1.14	1.10	1.13	0.9531	0.9130	0.9467
0.9	0.9	1.14	1.19	1.15	1.19	0.9653	0.9377	0.9755

FLUOXETINE-PLACEBO CONTROLLED CLINICAL TRIAL

① Patients were stratified in to two groups - normal rapid eye movement latency (REML) and those with shortened REML.

② Trial initialized with the first six patients in each strata assigned to treatment according to a permuted block randomization.

③ After this initialization a a RPW(1,1) was used, once again separately for each strata.

④ The studies primary endpoint was the percent of patients with a $> 50\%$ reduction in the Hamilton Depression Scale ($HAMD_{17}$) between baseline and the final study visit.

⑤ The primary endpoint was not able to be measure quickly enough to update the urn and in its place the allocations rule was based on a surrogate endpoint.

⑥ A surrogate responder was defined as a patient exhibiting $> 50\%$ reduction from baseline, with respect to ($HAMD_{17}$), in two consecutive visits after at least three weeks of therapy.

Fluoxetine Trial Results

The below table presents the estimate and confidence intervals (broken out by strata) for the Placebo (p_1) and Fluoxetine (p_2) success probabilities for the surrogate endpoint.

REML	n	$\hat{p}_1$	$\hat{p}_2$	$\tilde{p}_1$	$\tilde{p}_2$	CI($\hat{p}_1$)	CI($\hat{p}_2$)	CI($\tilde{p}_1$)	CI($\tilde{p}_2$)
Shortened	29	0.18	0.58	0.10	0.71	(0.00,0.44)	(0.29,0.91)	(0.04,0.41)	(0.22,0.88)
Normal	32	0.56	0.57	0.62	0.53	(0.21,0.81)	(0.23,0.82)	(0.27,0.89)	(0.28,0.84)

Summary

① Using either conditional or unconditional inference there is little evidence of a fluoxetine benefit with respect to the surrogate marker.

② However, the conditional intervals were narrower than their unconditional counterpart for 3 of the 4 parameters of interest.

Outline

Conclusions

① For the SDD using conditional bootstrap provides is the best in terms of both coverage and efficiency for most of the parameter space.

② For the RPW using conditional bootstrap provides provides the best balance between coverage and efficiency.

③ Comparisons of asymptotic efficiency for the most part agree with the small sample results.