An Evaluation of Inferential Procedures for Adaptive Clinical Trial Designs with Pre-specified Rules for Modifying the Sample Size

Gregory P. Levin, Sarah C. Emerson, & Scott S. Emerson

Cesar Torres

June 3, 2014

Overview

- Review
- Distribution of sampling density
- Simulation results
- Concerns (Criticisms)
- More simulation results

Review: Motivating Example

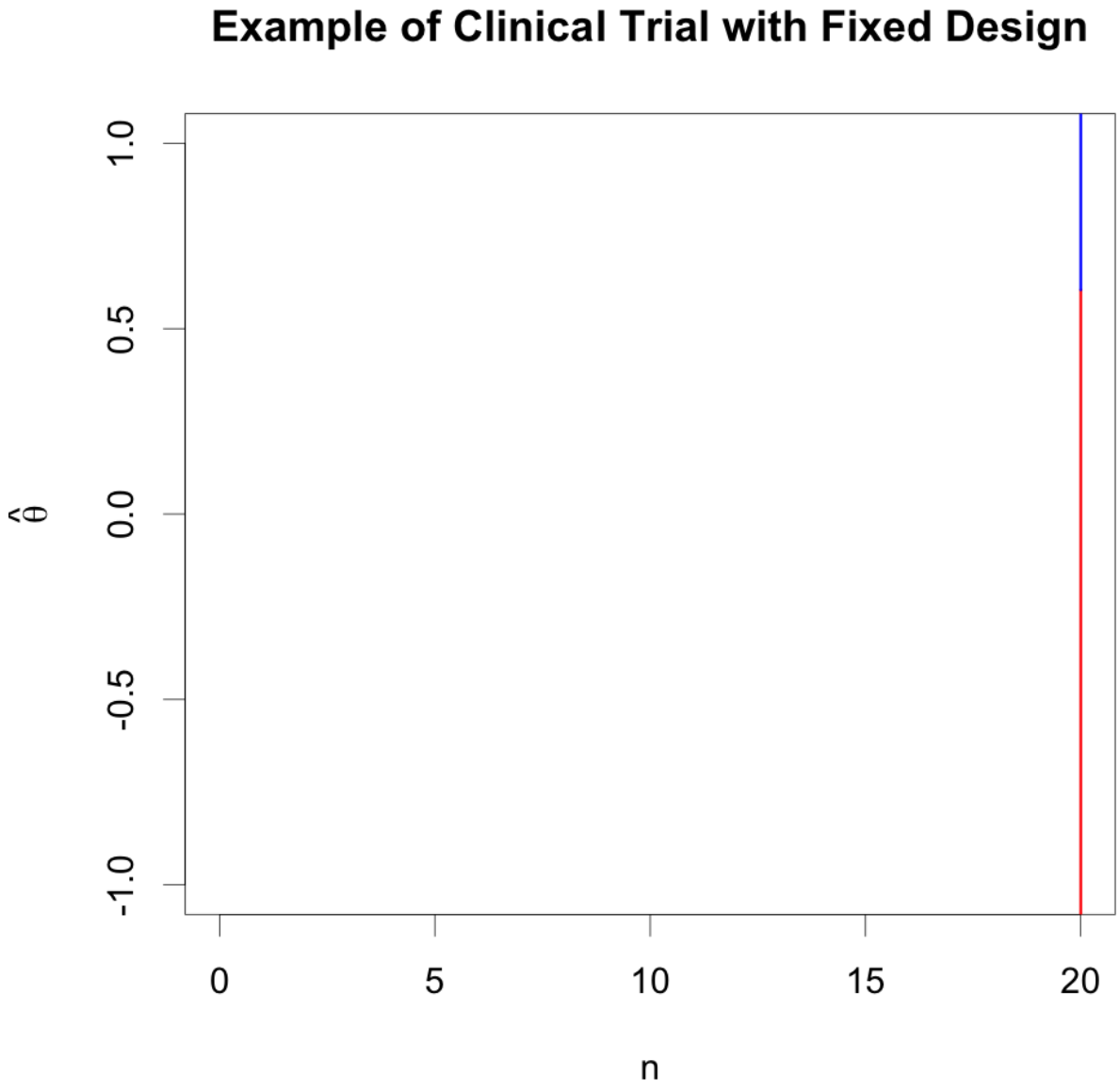
Suppose researchers want to cure [insert type of cancer here], because said cancer is bad.

- Two-arm clinical trial
- Can observe X_{plac_i}'s and X_{treat_i}'s
  - $X_{plac_i} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_{plac}, \sigma^2)$
  - $X_{treat_i} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_{treat}, \sigma^2)$
  - $\sigma^2 > 0$ known
- Defining $\theta := \mu_{treat} - \mu_{plac}$, interested in testing $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$
- Issue of concern: lots of treatments to evaluate

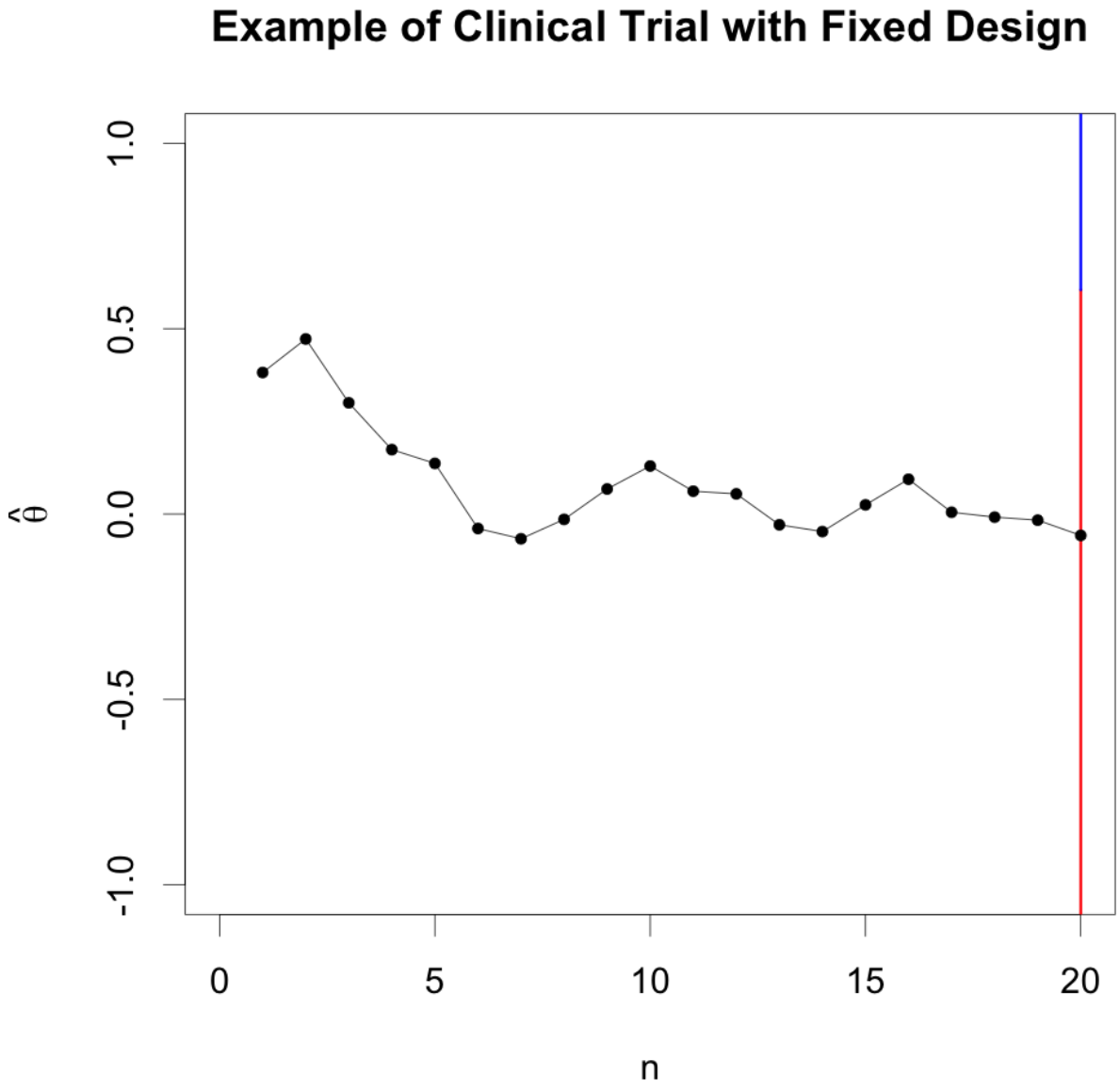
Review: Clinical Trial Designs

- "Well-understood" designs
  - Fixed design
  - Group sequential design
- "Less well-understood" designs
  - Adaptive design
    - The focus of this paper

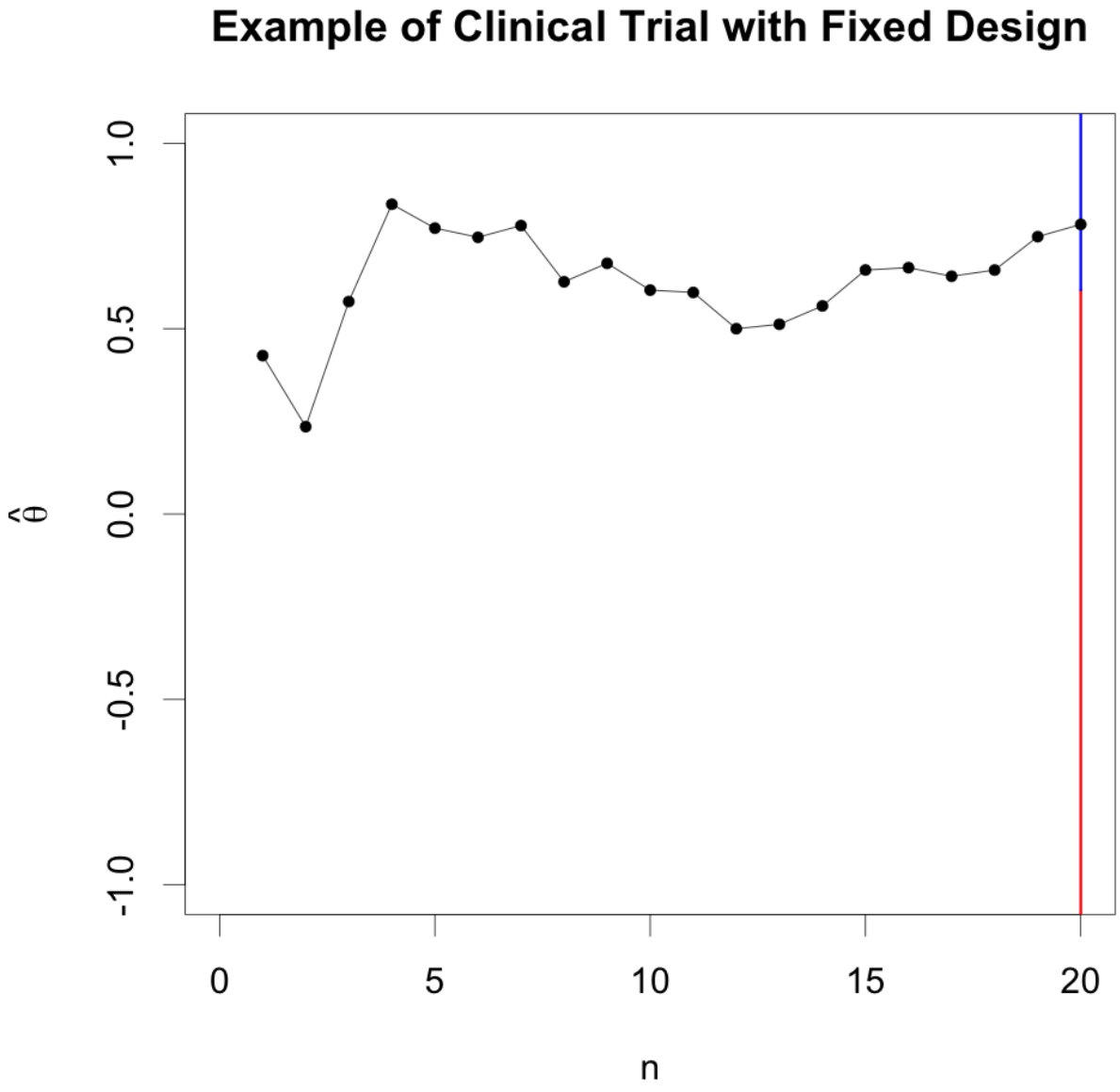
Review: Fixed Design



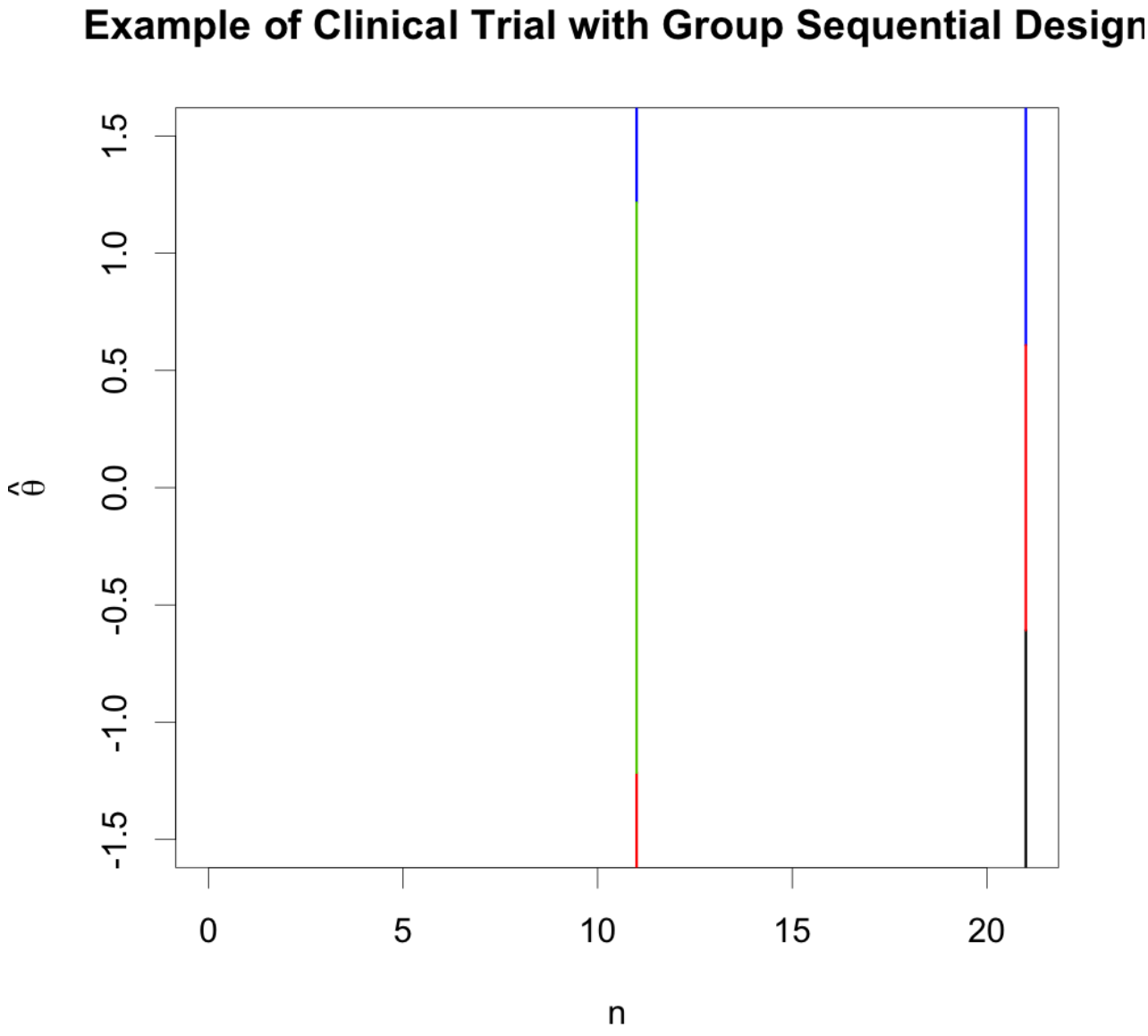
Review: Fixed Design



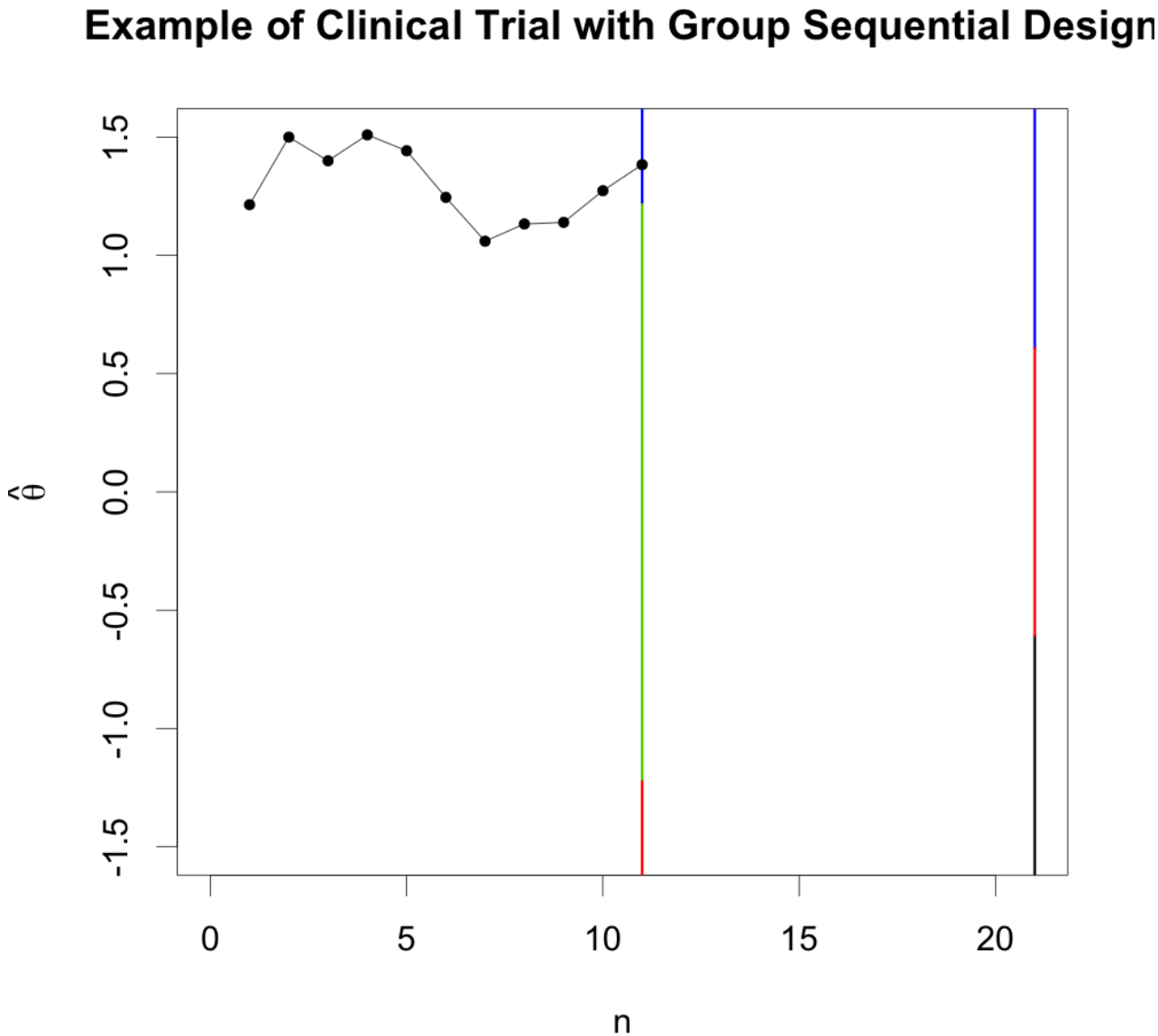
Review: Fixed Design



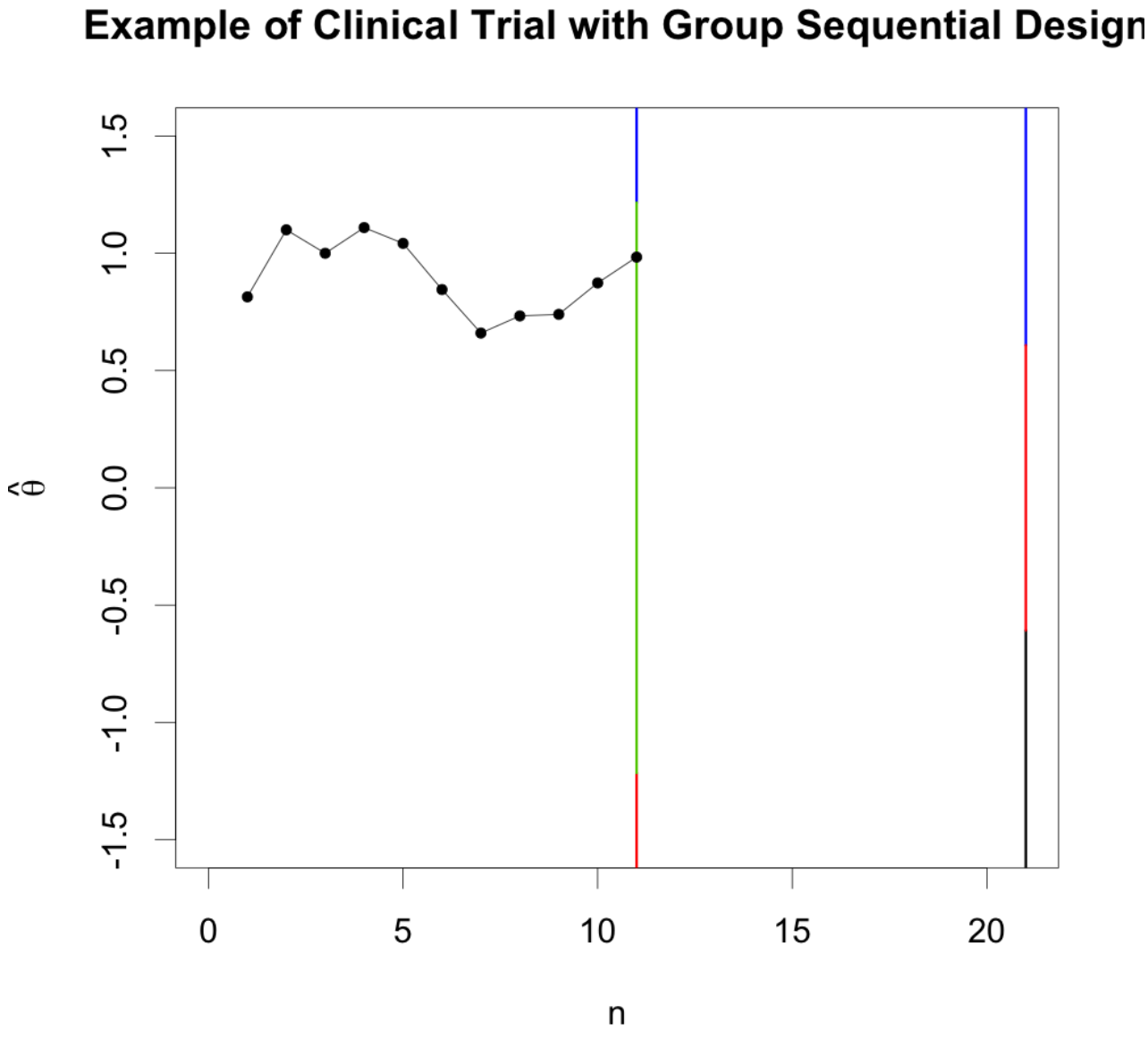
Review: Group Sequential Design



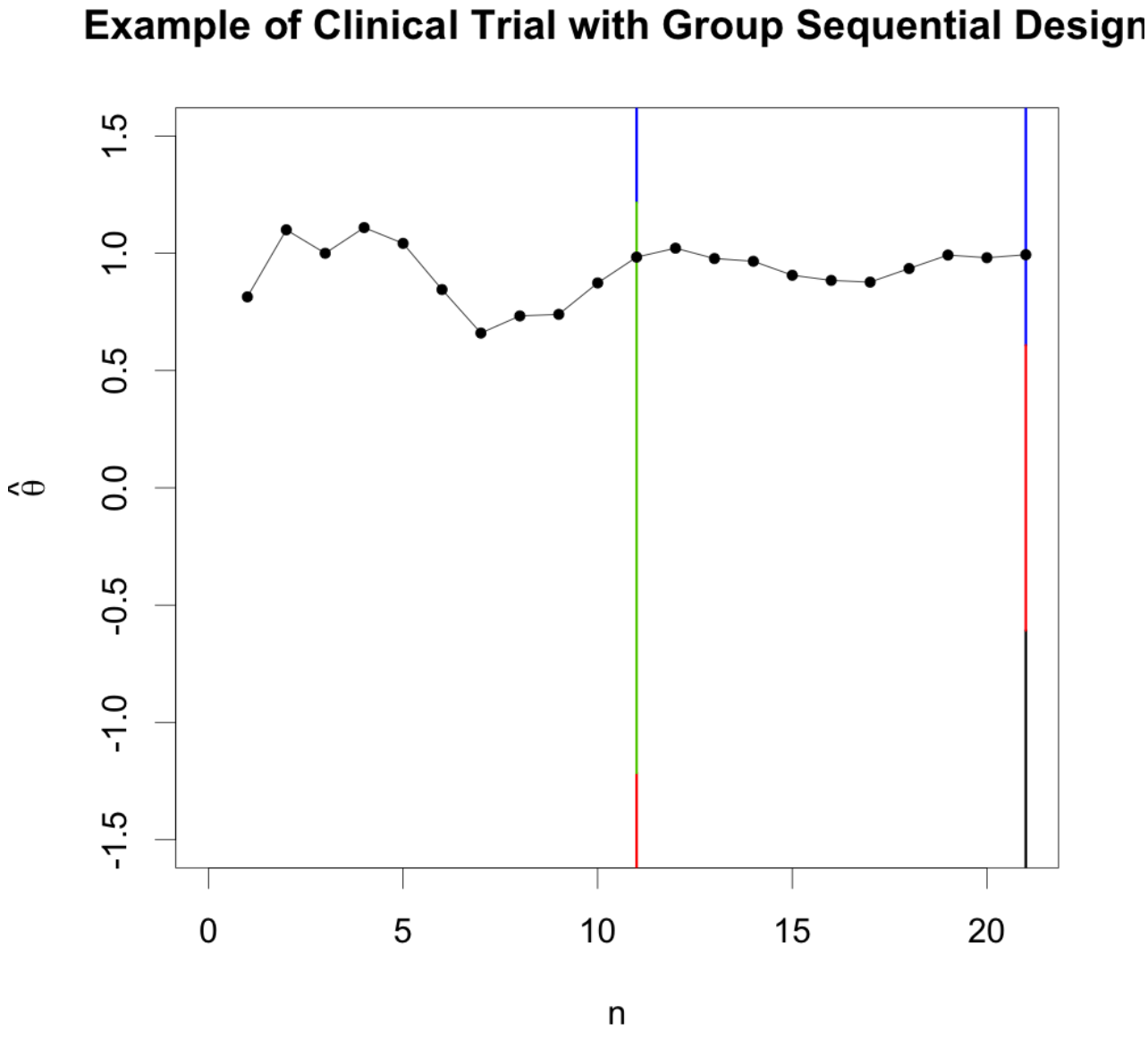
Review: Group Sequential Design



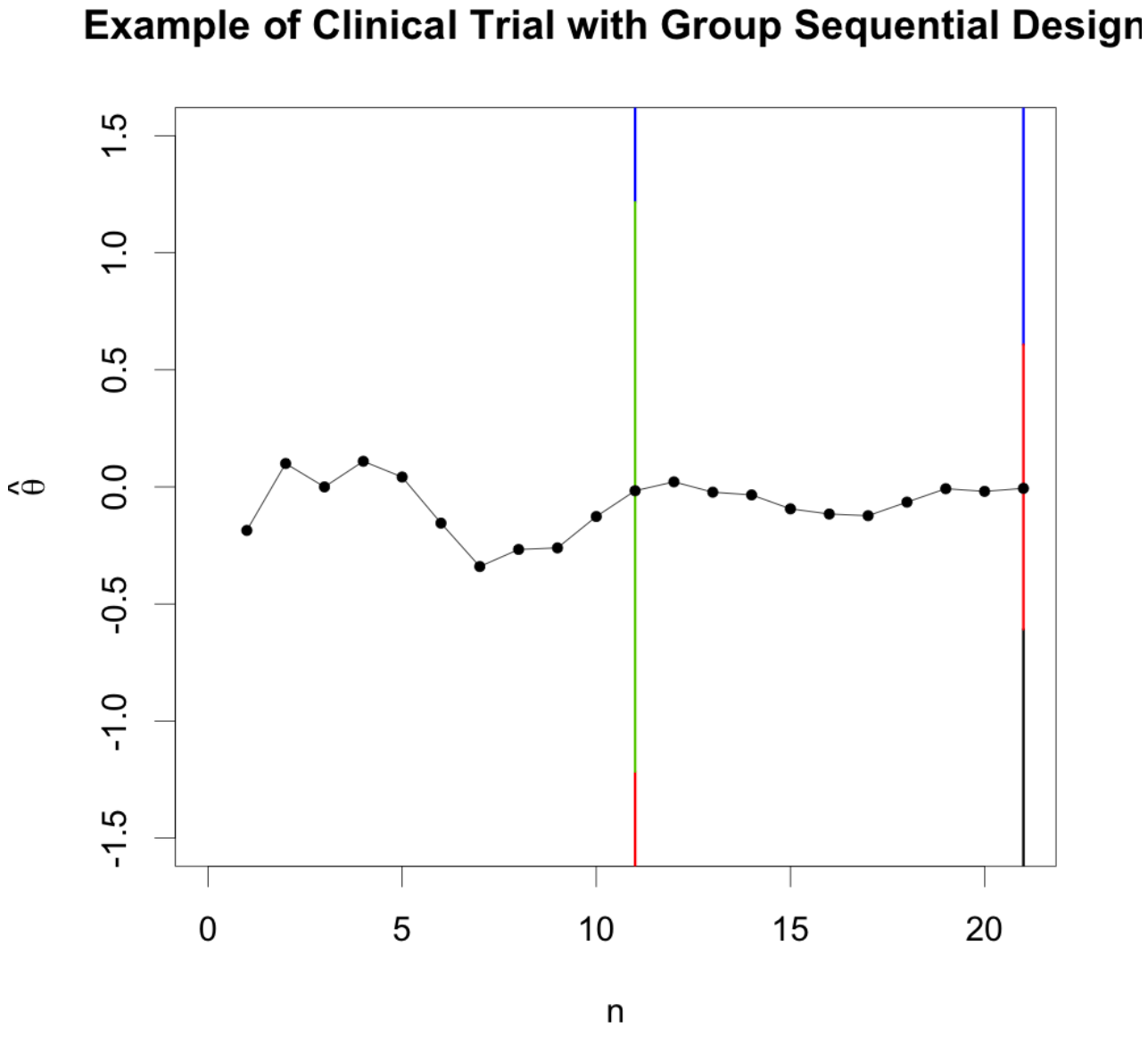
Review: Group Sequential Design



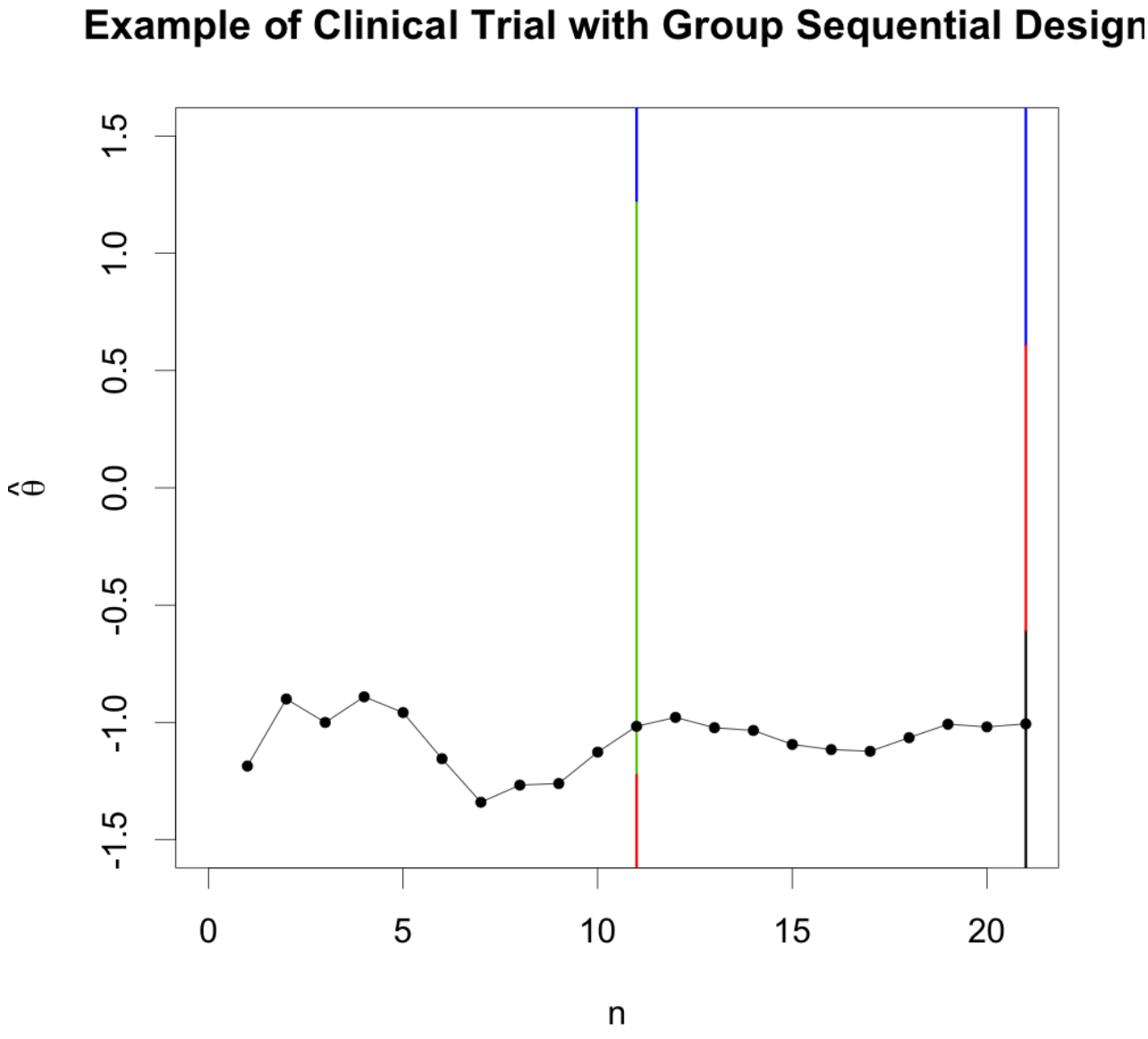
Review: Group Sequential Design



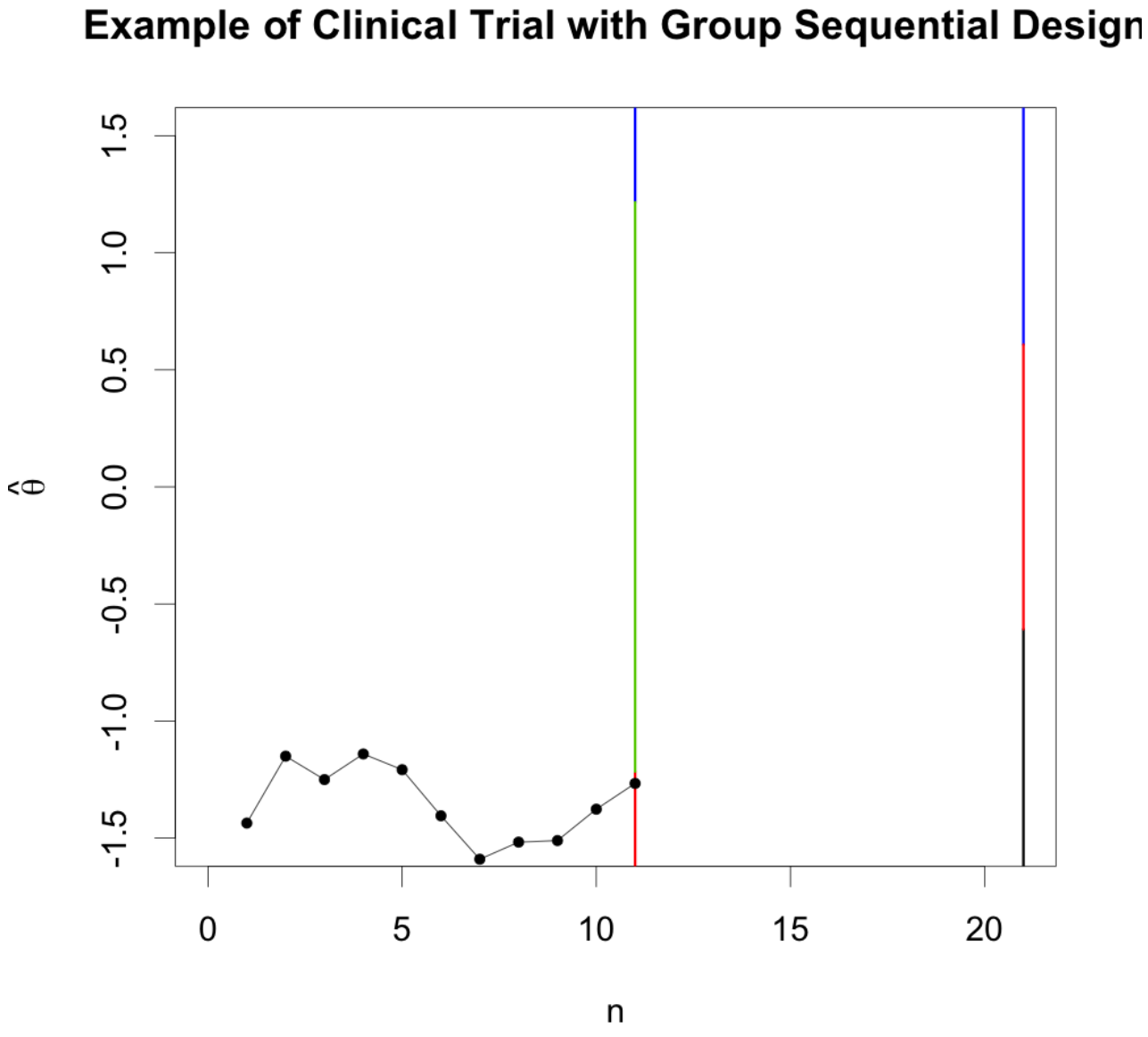
Review: Group Sequential Design



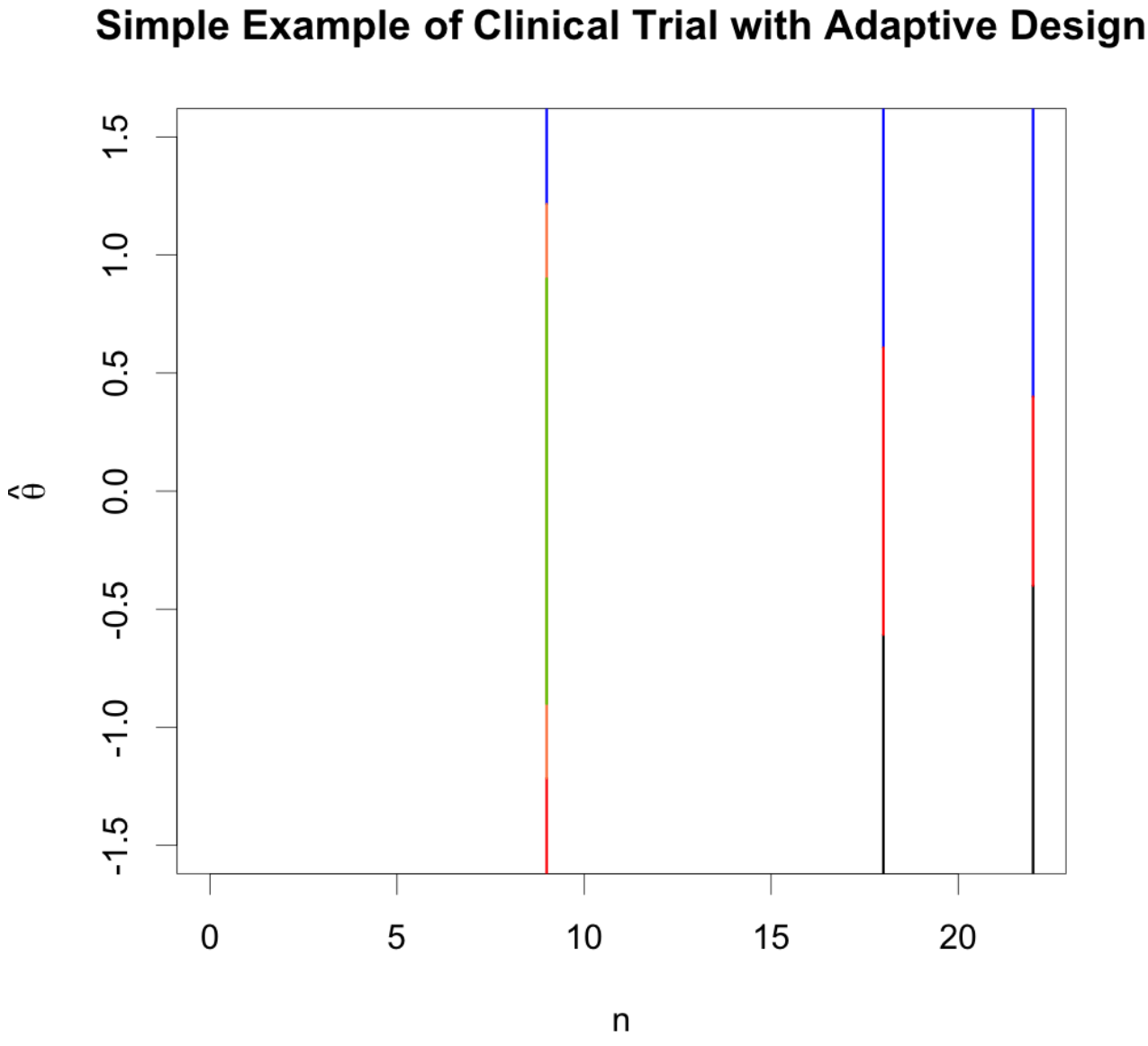
Review: Group Sequential Design



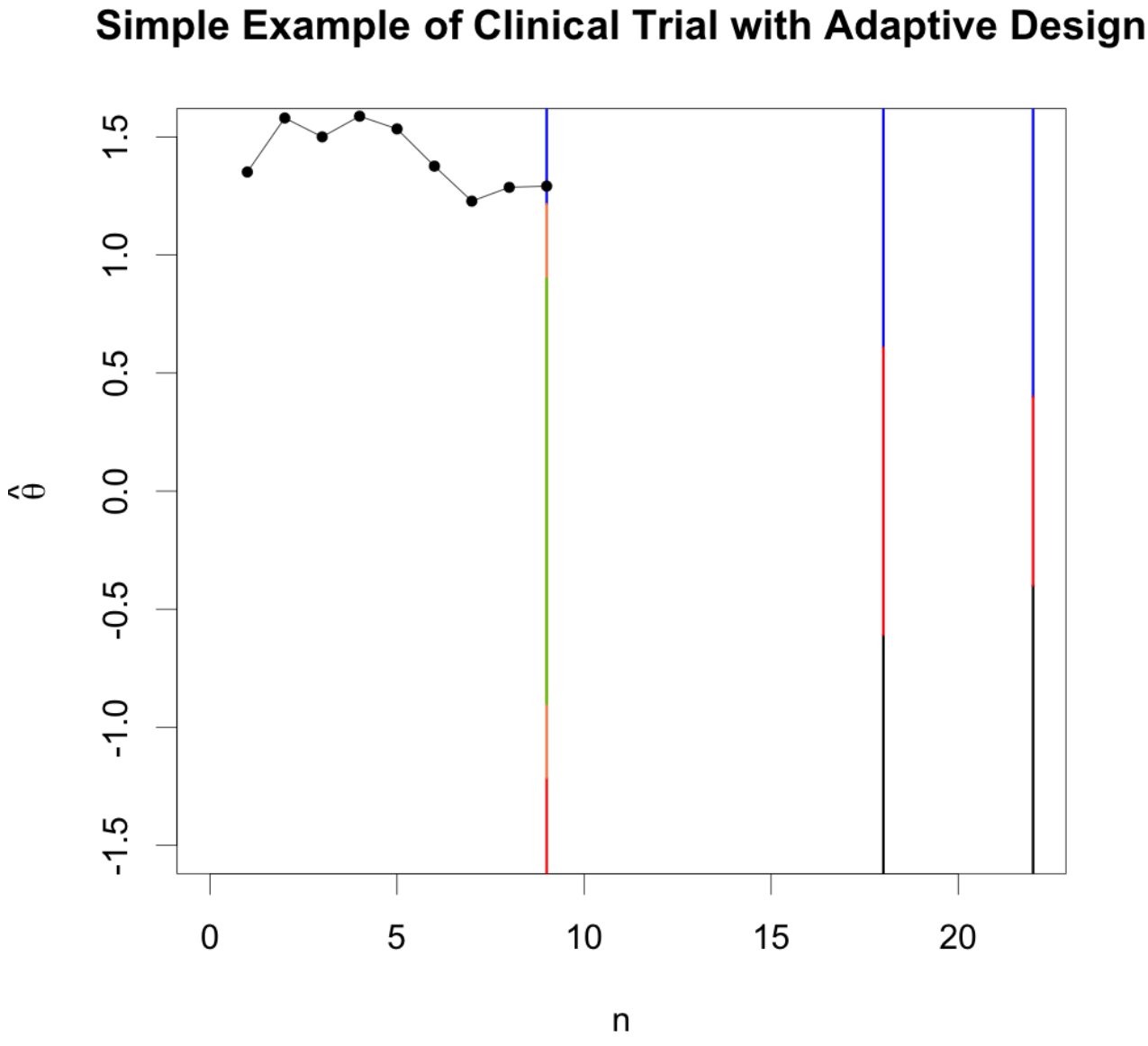
Review: Group Sequential Design



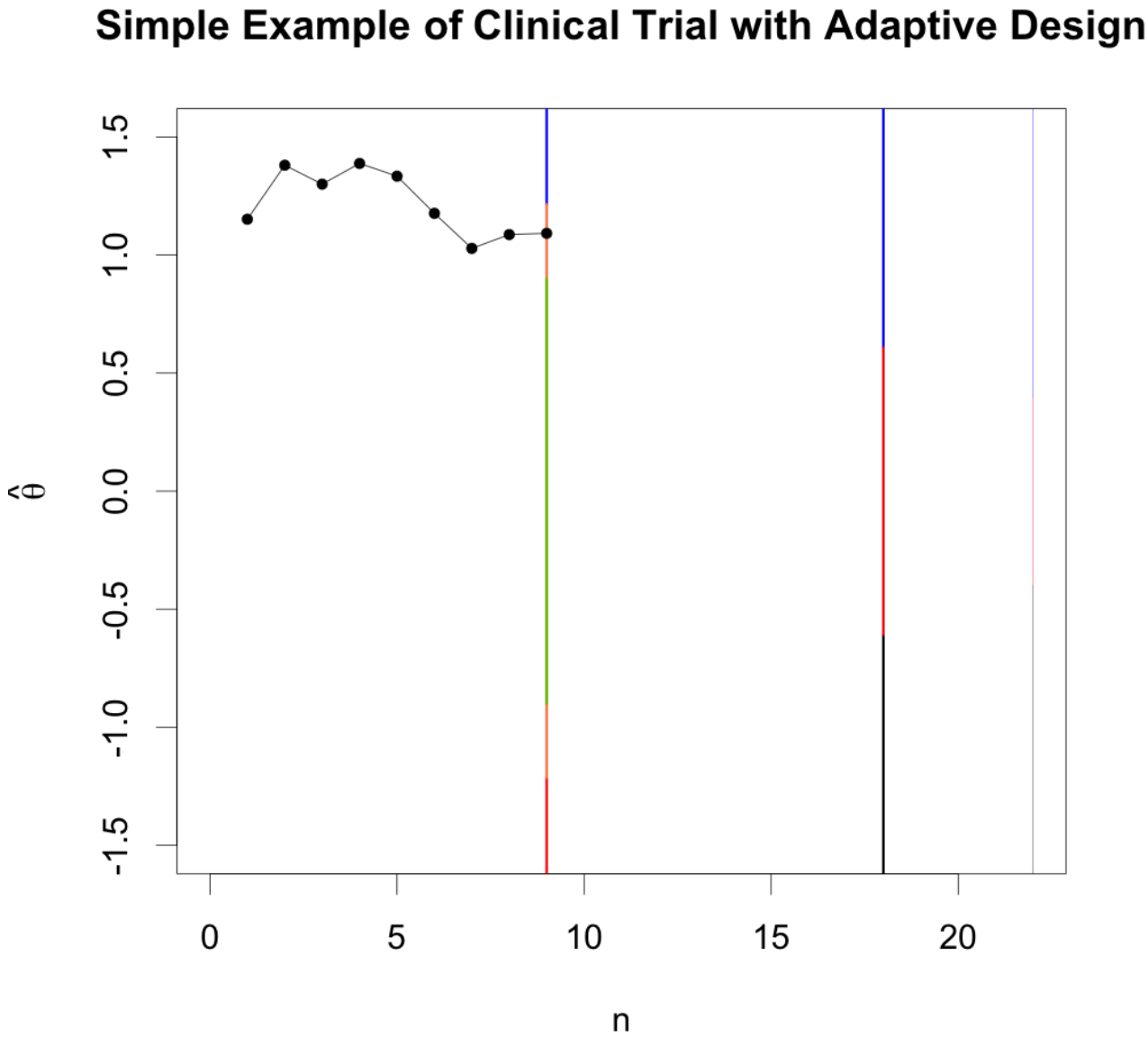
Review: Adaptive Design



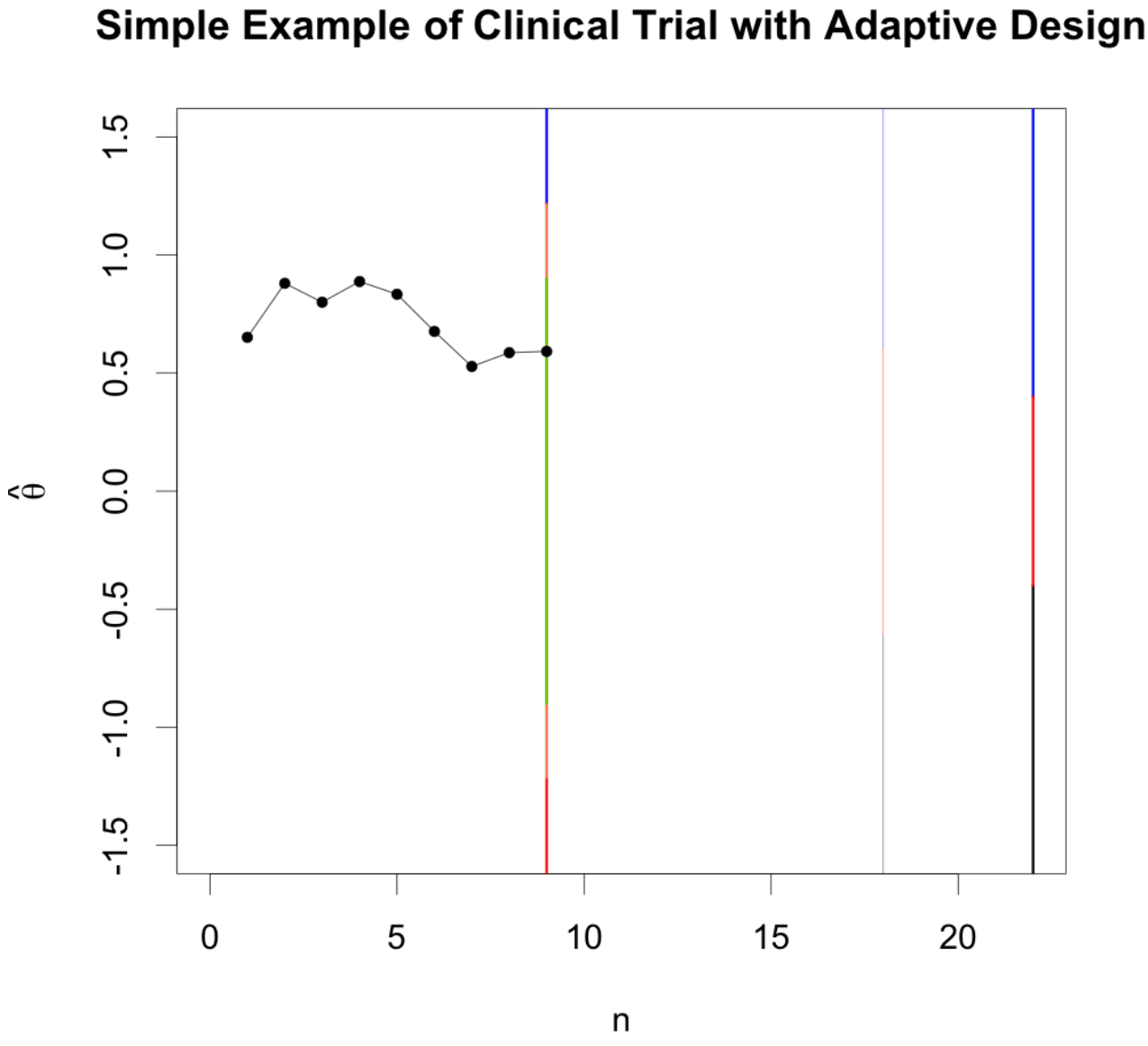
Review: Adaptive Design



Review: Adaptive Design



Review: Adaptive Design



Review: Inference when using GS or Adaptive Designs

- Neyman-Pearson lemma, Karlin-Rubin theorem not applicable
  - Likelihood ratio not monotone non-decreasing when using group-sequential-like designs
- Need some way (some ordering) to determine what are "extreme" observations under the null hypothesis

Review: Considered Orderings

- Sample mean
- Signed LR: If $\forall$ fixed θ^*,

$$sign\left(\hat{\theta}_{(1)} - \theta^*\right) \frac{f\left(\text{outcome 1}|\theta = \hat{\theta}_{(1)}\right)}{f(\text{outcome 1}|\theta = \theta^*)} > sign\left(\hat{\theta}_{(2)} - \theta^*\right) \frac{f\left(\text{outcome 2}|\theta = \hat{\theta}_{(2)}\right)}{f(\text{outcome 2}|\theta = \theta^*)},$$

  then outcome 1 ordered higher than outcome 2, with $\hat{\theta}_{(i)}$ the sample mean from outcome i
- Conditional Error Ordering: Outcomes ordered according to the stage-wise p-value of "backward image"

Review: Point Estimates

Three point estimates considered

- Sample mean (MLE) $\hat{\theta}$
- Bias adjusted mean (BAM) $\hat{\eta}$: the value θ for which $\hat{\theta}$ is the mean
- Median-unbiased estimate (MUE) $\hat{\zeta}$: the value θ for which $\hat{\theta}$ is the median

Distribution of Sampling Density

Law of Total Probability:

$$\begin{aligned} F_{\hat{\theta}|\theta}(x) &= P_\theta\left(\hat{\theta} \leq x\right) \\ &= \sum_{i=0}^{n} P_\theta\left(\hat{\theta} \leq x | \mathcal{C}_i\right) P_\theta(\mathcal{C}_i) \\ &= \sum_{i=0}^{n} F_{\hat{\theta}|\theta,\mathcal{C}_i}(x) P_\theta(\mathcal{C}_i). \end{aligned}$$

Taking derivatives:

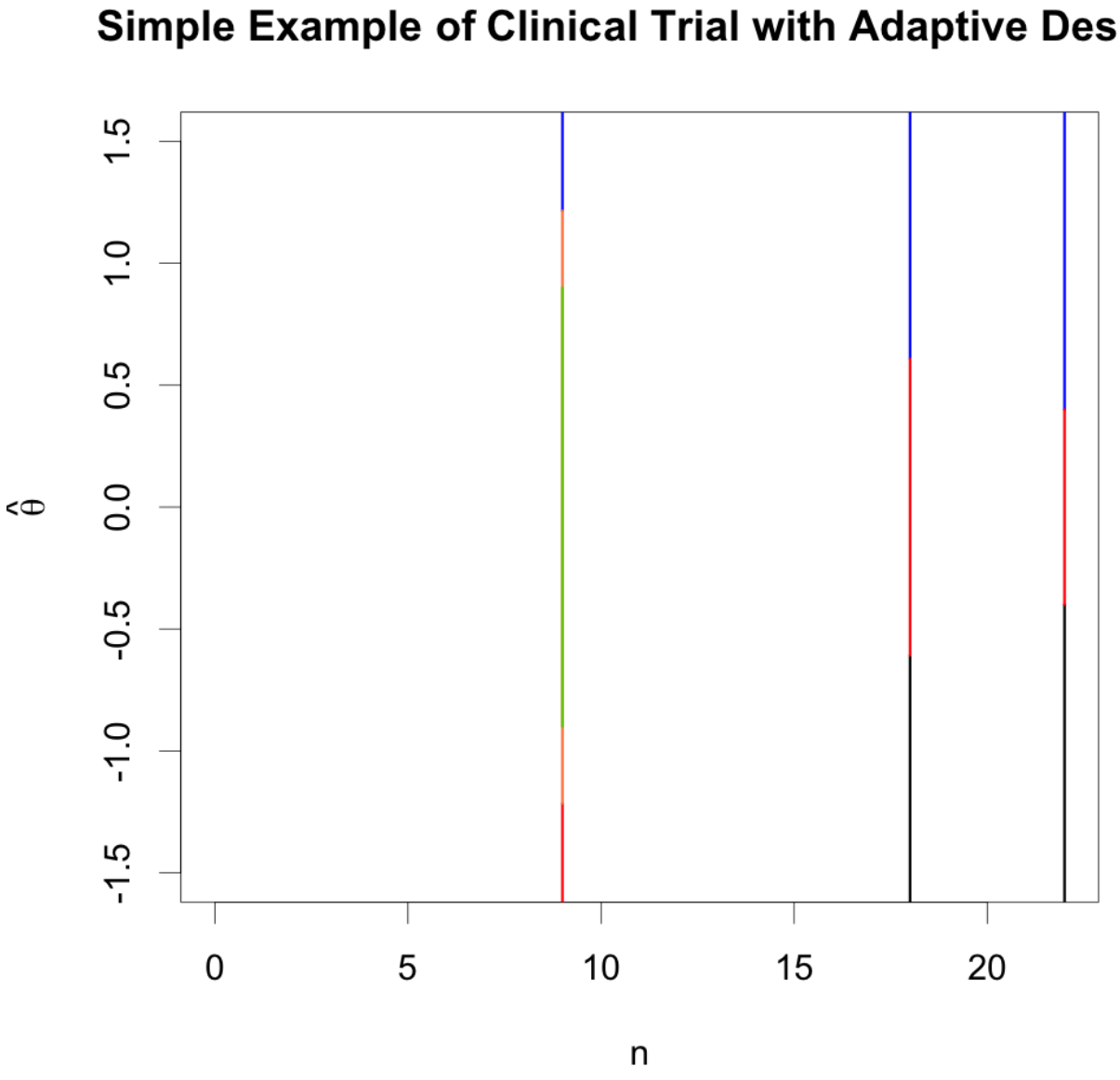
$$\begin{aligned} f_{\hat{\theta}|\theta}(x) &= \frac{d}{dx} F_{\hat{\theta}|\theta}(x) \\ &= \frac{d}{dx} \sum_{i=0}^{n} F_{\hat{\theta}|\theta,\mathcal{C}_i}(x) P_\theta(\mathcal{C}_i) \\ &= \sum_{i=0}^{n} \frac{d}{dx} F_{\hat{\theta}|\theta,\mathcal{C}_i}(x) P_\theta(\mathcal{C}_i) \\ &= \sum_{i=0}^{n} f_{\hat{\theta}|\theta,\mathcal{C}_i}(x) P_\theta(\mathcal{C}_i). \end{aligned}$$

Distribution of Sampling Density

$\mathcal{C}_0$: the stopping region.

$$
\begin{aligned}
F_{\hat{\theta}|\theta,\mathcal{C}_0}(x) &= P_\theta\Big(\hat{\theta} \le x | \mathcal{C}_0\Big) \\
&= P_\theta\Big(\hat{\theta} \le x | \hat{\theta}_1 \notin (a_1, d_1)\Big) \\
&= \frac{P_\theta\Big(\hat{\theta} \le x, \hat{\theta}_1 \notin (a_1, d_1)\Big)}{P_\theta\Big(\hat{\theta}_1 \notin (a_1, d_1)\Big)} \\
&= \frac{F_{\hat{\theta}_1|\theta}(x) \times 1_{\{\hat{\theta}_1 \notin (a_1, d_1)\}}}{P_\theta\Big(\hat{\theta}_1 \notin (a_1, d_1)\Big)}
\end{aligned}
$$

Review: Adaptive Design



Distribution of Sampling Density

Taking derivatives once more:

$$
\begin{aligned}
f_{\hat{\theta}|\theta,\mathcal{C}_0}(x) &= \frac{d}{dx} F_{\hat{\theta}|\theta,\mathcal{C}_0}(x) \\
&= \frac{f_{\hat{\theta}_1|\theta}(x) \times 1_{\left\{\hat{\theta}_1 \notin (a_1, d_1)\right\}}}{P_\theta\left(\hat{\theta}_1 \notin (a_1, d_1)\right)}.
\end{aligned}
$$

Distribution of Sampling Density

$\mathcal{C}_i, i \geq 1$: a continuation region.

- $m =$ sample size at interim analysis
- $N = m + n =$ sample size at final analysis
- $\hat{\theta} = \frac{m}{N} \times \hat{\theta}_1 + \frac{n}{N} \times \hat{\theta}_2$

Distribution of Sampling Density

$\mathcal{C}_i, i \geq 1$: a continuation region.

$$
\begin{aligned}
F_{\hat{\theta}|\hat{\theta}_1=x}(z) &= P_\theta\left(\hat{\theta} \leq z | \hat{\theta}_1 = x\right) \\
&= P_\theta\left(\frac{m}{N} \times \hat{\theta}_1 + \frac{n}{N} \times \hat{\theta}_2 \leq z | \hat{\theta}_1 = x\right) \\
&= \text{(some algebra)} \\
&= F_{\hat{\theta}_2|\theta}\left(\frac{N}{n}\left(z - \frac{mx}{N}\right)\right)
\end{aligned}
$$

Derivative:

$$
f_{\hat{\theta}|\hat{\theta}_1=x}(z) = f_{\hat{\theta}_2|\theta}\left(\frac{N}{n}\left(z - \frac{mx}{N}\right)\right)\frac{N}{n}
$$

Distribution of Sampling Density

$\mathcal{C}_i, i \geq 1$: a continuation region. Convolution:

$$f_{\hat{\theta}|\mathcal{C}_i}(z) = \int_{-\infty}^{\infty} f_{\hat{\theta}|\hat{\theta}_1}(z) f_{\hat{\theta}_1}(x) dx$$

- R can compute this numerically.

Simulation Results

Settings:

- Recalling that $\theta := \mu_{treat} - \mu_{plac}$, interested in testing $H_0 : \theta \leq 0$ vs. $H_1 : \theta > 0$
- Assumed: $\sigma^2 = 0.5$
- Desired: Level $\alpha = 0.025$ at $\theta = 0$, power of 0.9 at $\theta = 1$
- Continuation region from original GS design divided into 10 equally sized continuation regions
- Adaptive rule: Final sample size $N^*(t) = 2.02N - 1.627\,(t - 1.96)$, with t the midpoint of the new continuation region.
- Standard boundaries derived similarly to those in GS design

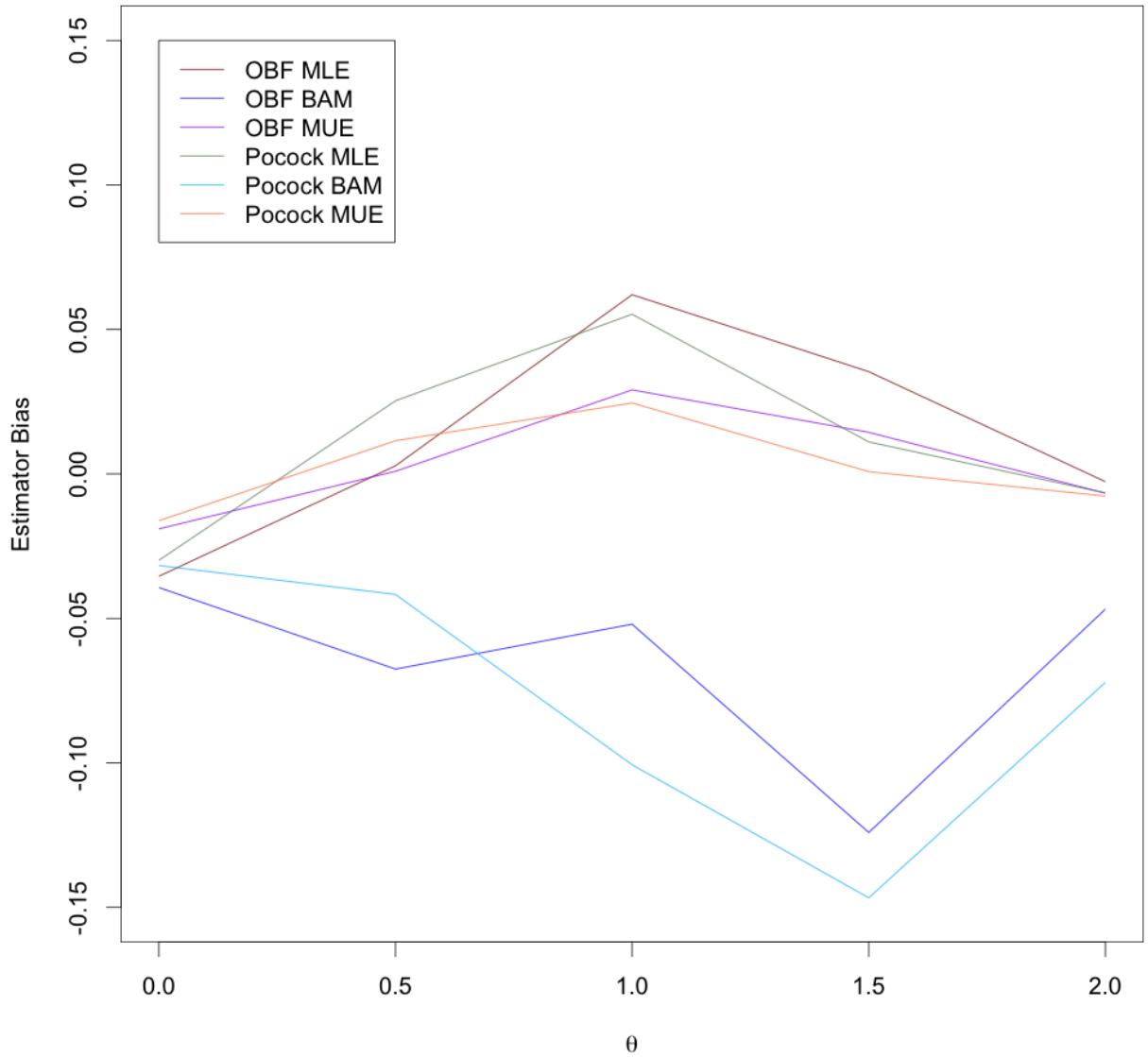
Simulation Results

Procedure:

- Through grid search, get boundaries and sample sizes needed to achieve desired size and power
  - Computationally demanding
- Run clinical trial (or simulate data)
  - Computationally easy
- Draw inference from observed data
  - Computationally **intense**

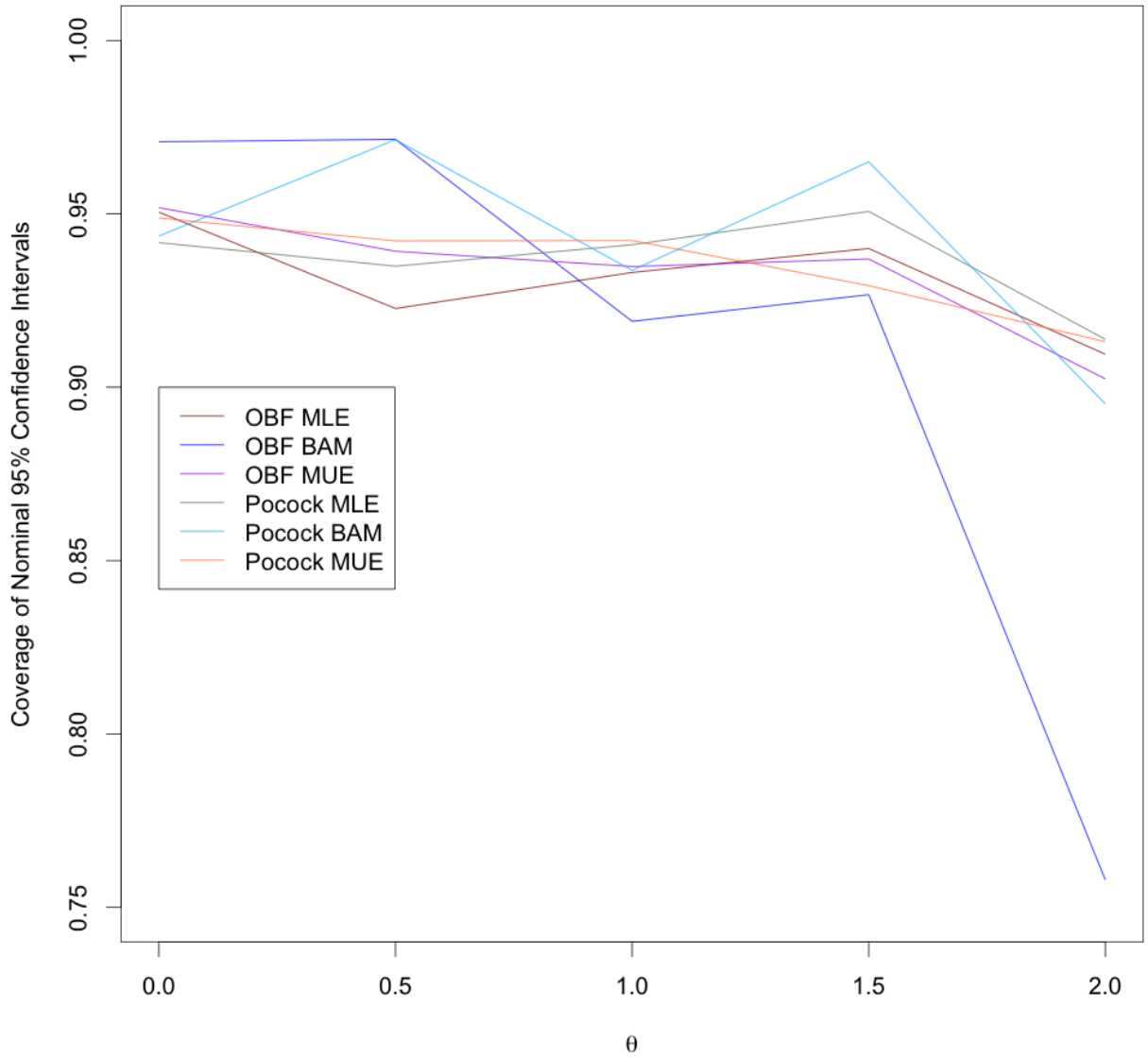
Simulation Results

Scenario 1: Distribution assumptions hold



Simulation Results

Scenario 1: Distribution assumptions hold



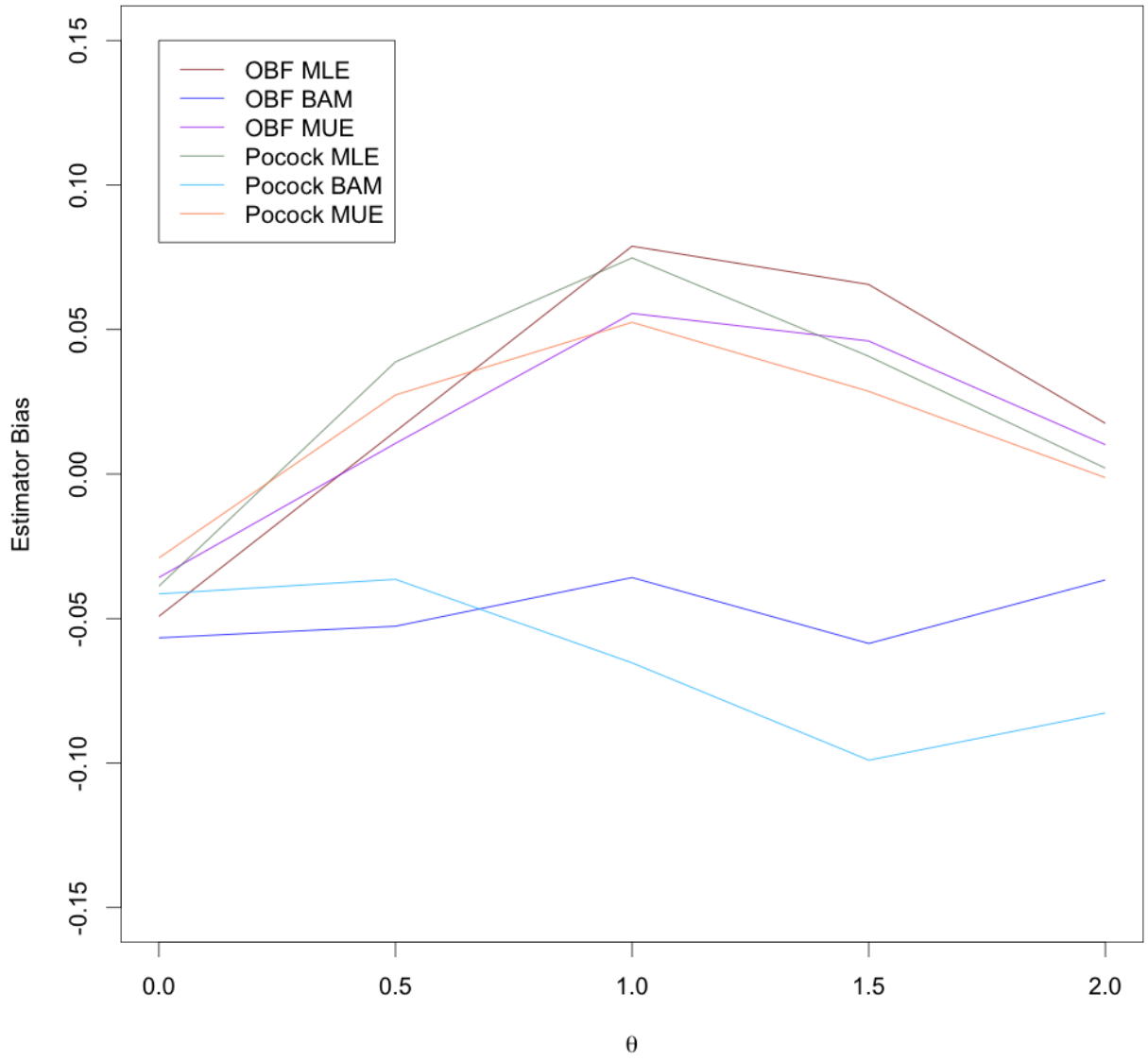
Concern

Distribution Assumptions

- Known variance
- Normality

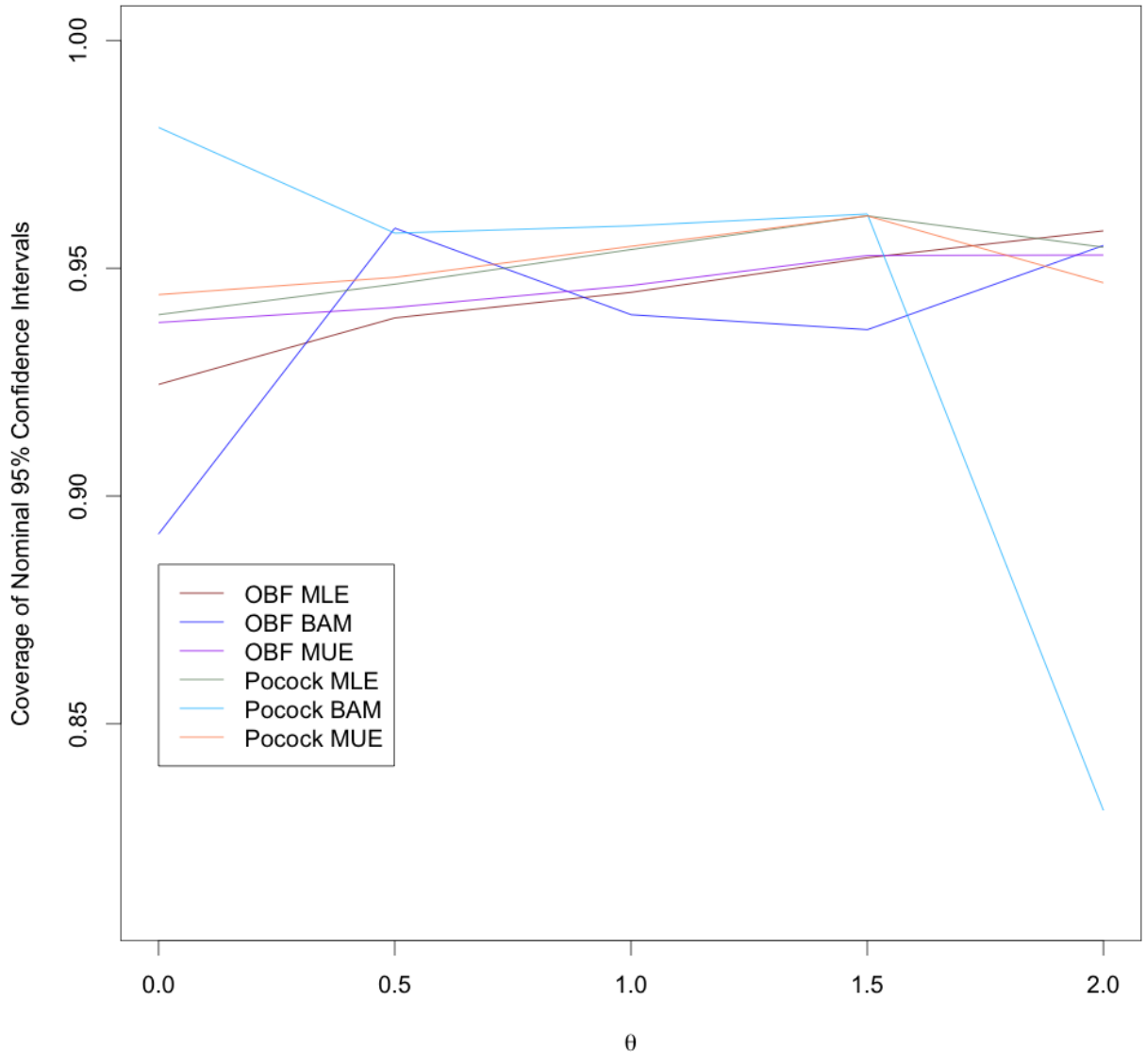
Simulation Results

Scenario 2: Normality holds, but true $\sigma^2 = 1$



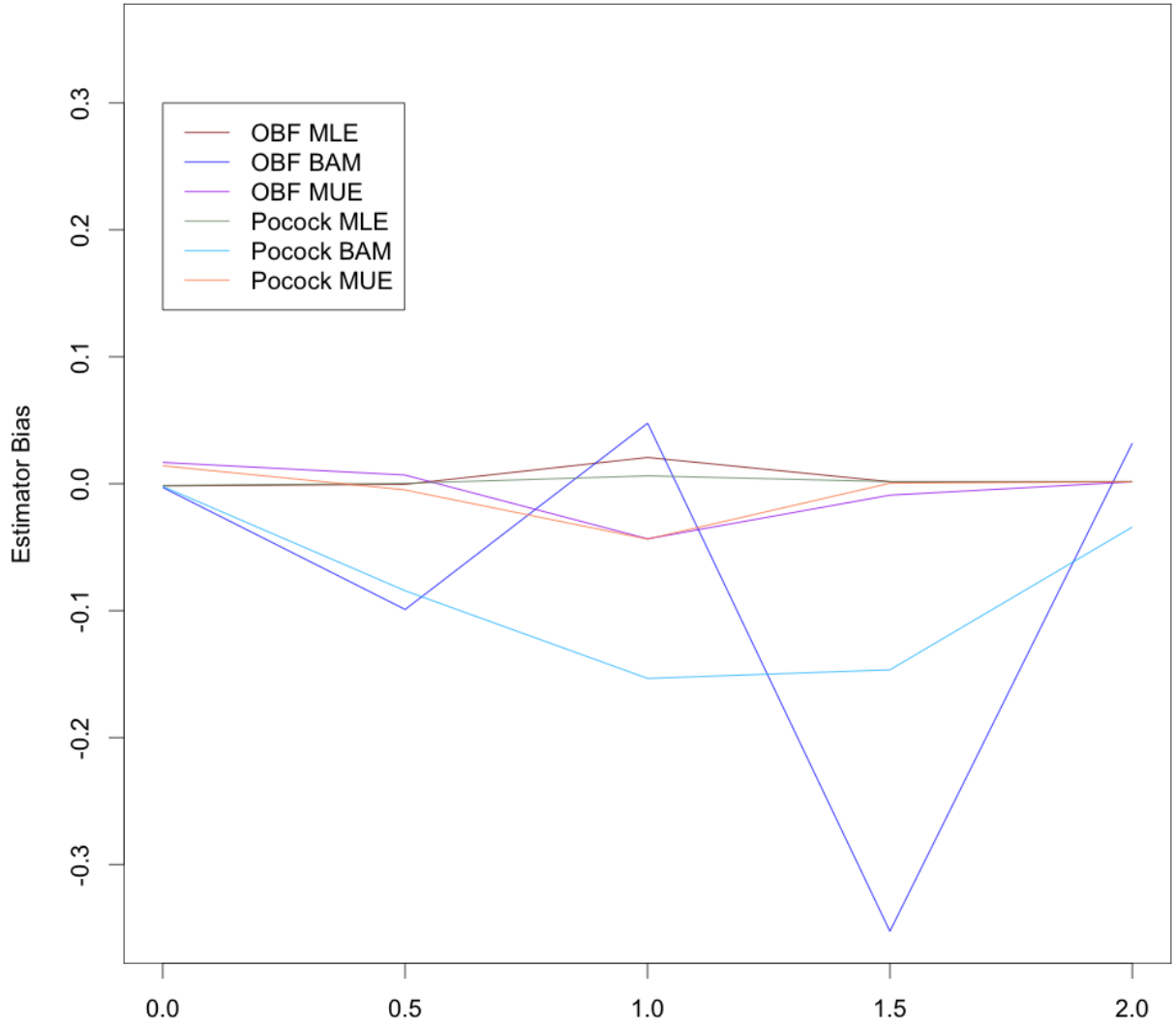
Simulation Results

Scenario 2: Normality holds, but true $\sigma^2 = 1$



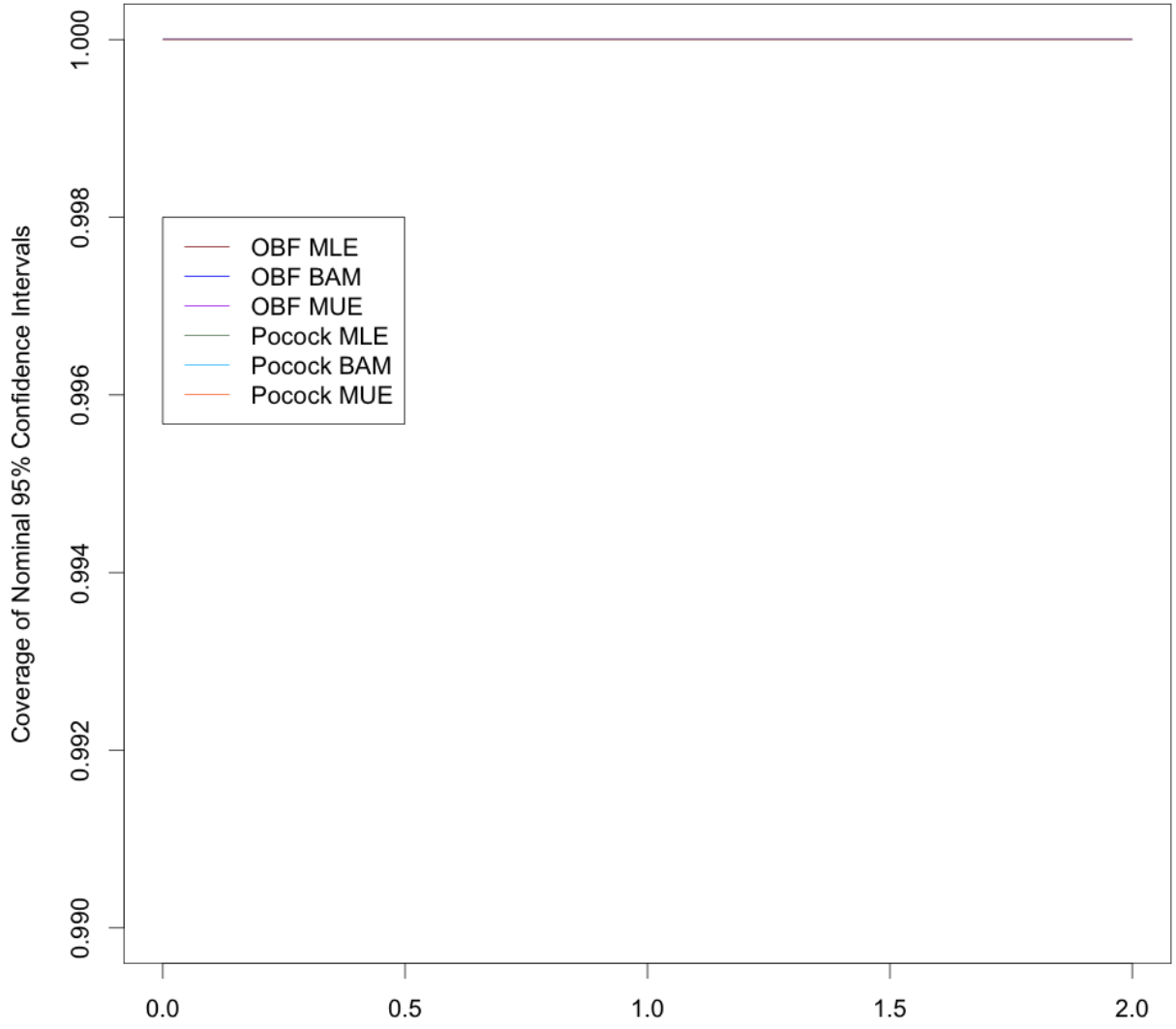
Simulation Results

Scenario 3: Data exponentially distributed, appropriately scaled and shifted so that $\sigma^2 = 0.5$ and $\theta \in (0, 2)$



Simulation Results

Scenario 3: Data exponentially distributed, appropriately scaled and shifted so that $\sigma^2 = 0.5$



Additional Concern

Knowledge of the final sample size is potentially unblinding.

- Same could be said of group sequential design, but group sequential design is widely accepted
  - Not a great answer, but it's something
- No clear way to quantify effects of such an unblinding

Summary

- Whether or not adaptive designs are a good idea, they are implemented to find cures for things such as [insert type of cancer here], so their properties need to be understood
- Under sample mean ordering and either type of boundary design, all 3 estimators do reasonably well, and confidence intervals do okay when θ is close to 0
- Inference not necessarily robust to violations of distribution assumptions

Questions?