Analysis of Critical Values of Response Rate in Single-Arm Two-Stage Phase II Clinical Trials

BIOSTAT 907 Term Project (Fall 2021)

Yi Liu

November 16, 2021

Overview

Single-arm phase II clinical trial

- A historical therapy vs. An experimental therapy
- Phase II: requires small sample size, test for safety and efficacy
- $H_0 : p = p_0$ vs. $H_1 : p = p_1$, test the response rate (RR) of different therapies ($p_1 = p_0 + \delta$, a clinically meaningful increasing)
  - Under H_0: $X \sim Bin(n, p_0)$
  - Single-stage design: Given $(\alpha^*, 1 - \beta^*, p_0, p_1)$, find a pair (n, a) to specify the sample size and rejection boundary
  - Start from $n \geq n_0$, find the smallest integer a such that based on level α^*, we reject H_0 if $X > a$ and otherwise fail to reject. Then if

    $$1 - \beta = P_1(X > a) \geq 1 - \beta^*$$

    the pair (n, a) is the best design. If not, continue to the next n
  - It is to *minimize the sample size*

Large Sample Approximation

Note that in a single-arm *single-stage* trial, given an (large) n, we can find the approximate value of a using large sample approximation

Denote $\bar{X} = X/n$

Under H_0:

$$\sqrt{n}(\bar{X} - p_0) \xrightarrow{d} N(0, p_0(1 - p_0))$$

Under H_1:

$$\sqrt{n}(\bar{X} - p_1) \xrightarrow{d} N(0, p_1(1 - p_1))$$

Large Sample Approximation

So, $P_0(X > a) \le \alpha^*$ is approximately equivalent to

$$\frac{\frac{a}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \ge z_{1-\alpha^*}$$

and $P_1(X > a) \ge 1 - \beta^*$ is approximately equivalent to

$$\frac{\frac{a}{n} - p_1}{\sqrt{\frac{p_1(1-p_1)}{n}}} \le z_{\beta^*}$$

where z_γ is the γ-quantile of $N(0,1)$ distribution

Large Sample Approximation

As such,

$$\sqrt{\frac{p_0(1-p_0)}{n}}z_{1-\alpha^*} + p_0 \leq \frac{a}{n} \leq \sqrt{\frac{p_1(1-p_1)}{n}}z_{\beta^*} + p_1$$

Then we approximate a/n by the average of the LHS and RHS

$$\frac{a}{n} \approx \frac{1}{2}(\underbrace{\sqrt{\frac{p_0(1-p_0)}{n}}z_{1-\alpha^*} + \sqrt{\frac{p_1(1-p_1)}{n}}z_{\beta^*}}_{\approx\ 0 \text{ as } n \text{ becomes large}} + p_0 + p_1)$$

Therefore

$$a \approx \frac{n}{2}(p_0 + p_1)$$

(Note a is related to n)

Two-Stage Design

Now, consider the single-arm *two-stage* design for phase II trials

Why two-stage?

- Economical and ethical consideration
- Stop early if no efficacy tested
- Data collected can be more informative than single-stage if we can proceed to phase III

Best Two-Stage Design

$H_0 : p = p_0$ vs. $H_1 : p = p_1$. $n = n_1 + n_2$ for two stages respectively

- $X_1 \sim Bin(n_1, p_i)$ is independent to $X_2 \sim Bin(n_2, p_i)$ under H_i, $i = 0, 1$
- $\alpha = P_0(X_1 > a_1, X_1 + X_2 > a)$
- $1 - \beta = P_1(X_1 > a_1, X_1 + X_2 > a)$
- Probability of early termination under H_0: PET = $P_0(X_1 \leq a_1)$
- Expected sample size under H_0: EN = $n_1 \times$ PET + $n \times (1 -$ PET$)$

Among designs with $\alpha \leq \alpha^*$ and $1 - \beta \geq 1 - \beta^*$

- **Optimal** design: (a_1, n_1, a, n) minimizes EN
- **Minimax** design: (a_1, n_1, a, n) minmizes $n = n_1 + n_2$

This is the case of without considering *superiority (upper) stopping* of stage 1 (or, only *futility (lower) stopping*)

Best Two-Stage Design

When we consider *both futility and superiority stopping*, there will be one more boundary value (b_1) to let us stop the trial immediately after stage 1:

- $1 - \alpha = P_0(X_1 \leq a_1) + P_0(a_1 < X_1 < b_1, X_1 + X_2 \leq a)$
- $\beta = P_1(X_1 \leq a_1) + P_1(a_1 < X_1 < b_1, X_1 + X_2 \leq a)$
- Probability of early termination under H_i: $\mathrm{PET}(p_i) = P_i(X_1 \leq a_1) + P_i(X_1 \geq b_1), i = 0, 1$
- Expected sample size: $\mathrm{EN}(p_i) = n_1 \times \mathrm{PET}(p_i) + n \times (1 - \mathrm{PET}(p_i)), i = 0, 1$
  - $\mathrm{EN} = \dfrac{\mathrm{EN}(p_0) + \mathrm{EN}(p_1)}{2}$

Among designs with $\alpha \leq \alpha^*$ and $1 - \beta \geq 1 - \beta^*$

- **Optimal** design: (a_1, b_1, n_1, a, n) minimizes EN
- **Minimax** design: (a_1, b_1, n_1, a, n) minmizes $n = n_1 + n_2$

What are we looking for?

Recall that $a \approx \frac{n}{2}(p_0 + p_1)$ in single-stage trials

Is there any similar pattern(s) in the critical values of hypothesis test in two-stage trials?

- With only futility stopping, is there any pattern(s) in a_1/n_1 and a/n over different choices of $(\alpha, 1-\beta, p_0, p_1)$?
- With both futility and superiority stopping considered, is there any pattern(s) in $a_1/n_1, b_1/n_1$ and a/n over different choices of $(\alpha, 1-\beta, p_0, p_1)$?
- Our hunch
  - $a_1/n_1 \approx p_0$
  - $b_1/n_1 \approx p_1$
  - $a/n \approx (p_0 + p_1)/2$
- Anything else?

Experiment

- We did some numerical experiments to find a sense
- Consider the following choices of the parameters
  - $p_0 \in [0.05, 0.7]$ with increment of 0.005
  - $p_1 = p_0 + 0.2$ and $p_0 + 0.25$
  - level of test = $\alpha^* = 0.05, 0.1$
  - power = $1 - \beta^* = 0.8, 0.85, 0.9$

Experiment

- For only futility stopping, plot the figures of
  - $a_1 - n_1 p_0$
  - $a - n(p_1 + p_0)/2$
  - and corresponding ratio figures, such as $a_1/n_1 - p_0$
- For considering both two stoppings, plot the figures of
  - $a_1 - n_1 p_0$
  - $b_1 - n_1 p_1$
  - $a - n(p_1 + p_0)/2$
  - and corresponding ratio figures
- Investigate bias and variability over p_0

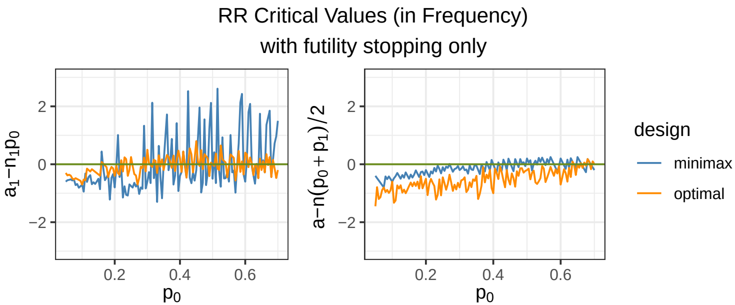
Result

Because of the similar results we get, we just present the case of frequency trends under

- $(\alpha^*, 1-\beta^*) = (0.1, 0.8),\ (0.05, 0.8),\ (0.1, 0.9),\ (0.05, 0.9)$
- $p_1 = p_0 + 0.2$
  - futility stopping only
  - both futility and superiority stopping

Result

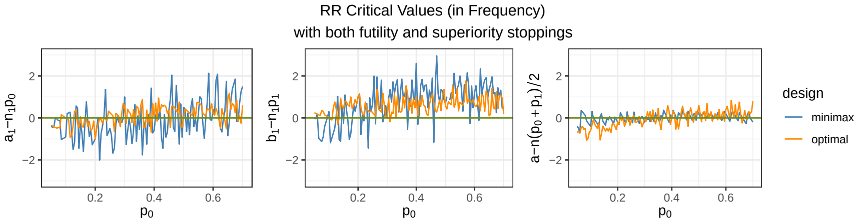
$(\alpha^*, 1-\beta^*) = (0.1, 0.8),\ p_1 - p_0 = 0.2$



Stage	Design	Median	Mean	IQR	SD
Stage 1	Optimal	-0.080	-0.067	0.473	0.310
	Minimax	-0.420	-0.051	0.807	0.931
Stage 2	Optimal	-0.555	-0.544	0.530	0.350
	Minimax	-0.125	-0.143	0.317	0.225

Result

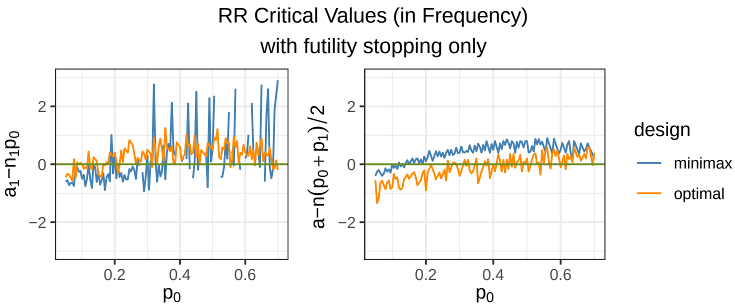
$$(\alpha^*, 1-\beta^*) = (0.1, 0.8),\ p_1 - p_0 = 0.2$$



Stage	Design	Median	Mean	IQR	SD
Stage 1 Lower	Optimal	0.135	0.102	0.680	0.438
	Minimax	-0.160	-0.117	1.068	0.877
Stage 1 Upper	Optimal	0.550	0.593	0.600	0.409
	Minimax	0.945	0.756	1.215	0.946
Stage 2 Lower	Optimal	-0.050	-0.064	0.528	0.386
	Minimax	0.020	0.012	0.305	0.202

Result

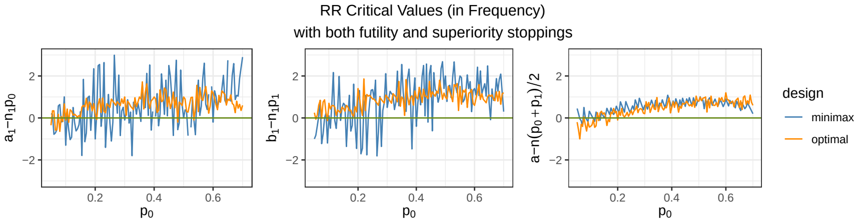
$(\alpha^*, 1-\beta^*) = (0.05, 0.8),\ p_1 - p_0 = 0.2$



Stage	Design	Median	Mean	IQR	SD
Stage 1	Optimal	0.345	0.315	0.530	0.375
	Minimax	-0.040	0.603	1.437	1.427
Stage 2	Optimal	-0.100	-0.124	0.545	0.393
	Minimax	0.440	0.404	0.437	0.316

Result

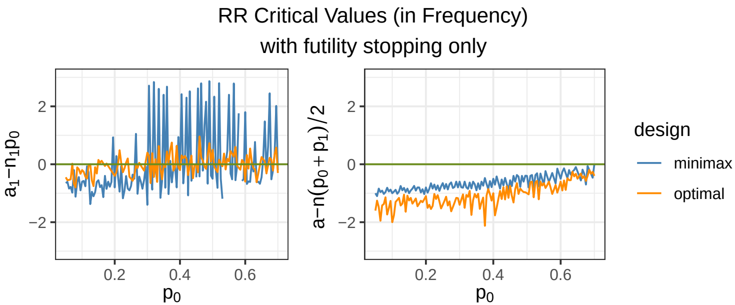
$$(\alpha^*, 1-\beta^*) = (0.05, 0.8),\ p_1 - p_0 = 0.2$$



Stage	Design	Median	Mean	IQR	SD
Stage 1 Lower	Optimal	0.650	0.628	0.565	0.446
	Minimax	0.500	0.664	1.467	1.188
Stage 1 Upper	Optimal	0.900	0.882	0.557	0.424
	Minimax	1.000	0.844	1.318	1.082
Stage 2 Lower	Optimal	0.500	0.440	0.527	0.422
	Minimax	0.600	0.545	0.367	0.294

Result

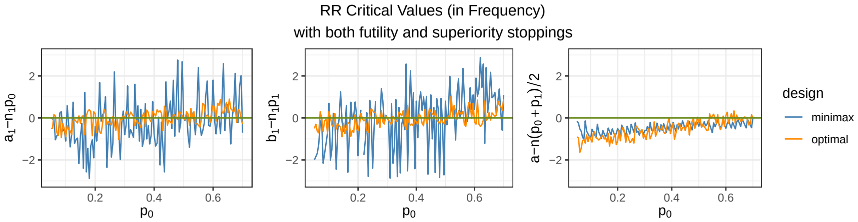
$(\alpha^*, 1-\beta^*) = (0.1, 0.9),\ p_1 - p_0 = 0.2$



Stage	Design	Median	Mean	IQR	SD
Stage 1	Optimal	-0.040	-0.037	0.440	0.343
	Minimax	-0.450	0.078	0.930	1.254
Stage 2	Optimal	-1.120	-1.065	0.525	0.398
	Minimax	-0.690	-0.645	0.398	0.253

Result

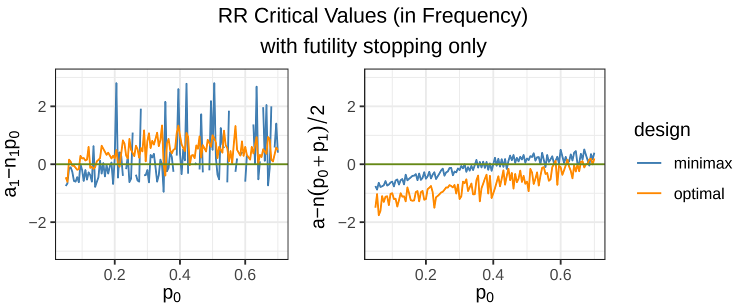
$(\alpha^*, 1-\beta^*) = (0.1, 0.9)$, $p_1 - p_0 = 0.2$



Stage	Design	Median	Mean	IQR	SD
Stage 1 Lower	Optimal	0.055	-0.012	0.560	0.401
	Minimax	-0.310	-0.221	1.215	1.134
Stage 1 Upper	Optimal	-0.075	-0.044	0.525	0.393
	Minimax	0.140	-0.083	1.700	1.394
Stage 2 Lower	Optimal	-0.550	-0.546	0.652	0.429
	Minimax	-0.465	-0.469	0.333	0.235

Result

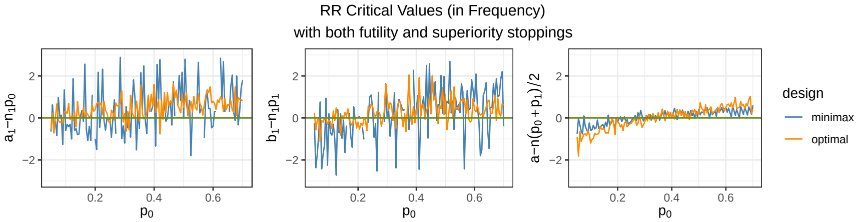
$$(\alpha^*, 1-\beta^*) = (0.05, 0.9),\ p_1 - p_0 = 0.2$$



Stage	Design	Median	Mean	IQR	SD
Stage 1	Optimal	0.445	0.450	0.465	0.374
	Minimax	0.060	0.745	1.757	1.611
Stage 2	Optimal	-0.710	-0.677	0.697	0.451
	Minimax	-0.060	-0.107	0.535	0.341

Result

$$(\alpha^*, 1-\beta^*) = (0.05, 0.9),\ p_1 - p_0 = 0.2$$



Stage	Design	Median	Mean	IQR	SD
Stage 1 Lower	Optimal	0.480	0.478	0.658	0.499
	Minimax	0.300	0.633	1.730	1.378
Stage 1 Upper	Optimal	0.250	0.287	0.688	0.553
	Minimax	0.230	-0.023	1.630	1.435
Stage 2 Lower	Optimal	0.100	-0.001	0.728	0.519
	Minimax	0.090	0.053	0.395	0.311

Summary

Motivated by the relationship between (a, n) and (p_0, p_1) in single-stage phase II trials, we investigated the trend of several critical values of testing RR in two-stage phase II trials with two kinds of stopping assumption

- Indeed, for both optimal and minimax design
  - $a_1 \approx n_1 p_0$
  - $b_1 \approx n_1 p_1$
  - $a \approx n(p_0 + p_1)/2$
- In many cases, there is no too much difference, or no special pattern in biases (of $a_1 - n_1 p_0$, $b_1 - n_1 p_1$ and $a - n(p_0 + p_1)/2$) between optimal and minimax design
- Optimal design has smaller **variations** (by variance, IQR) in $a_1 - n_1 p_0$ (and $b_1 - n_1 p_1$), where minimax design has smaller **variations** in $a - n(p_0 + p_1)/2$

Acknowledgement

The credit of the idea of this term project should be given to the instructor of this course, Dr. Sin-Ho Jung, and thanks for his two CTD softwares to help me to confirm the programming results.

References

- S-H Jung (2013). Randomized Phase II Cancer Clinical Trials (1st ed.). *Chapman and Hall/CRC*. (Chapter 2 and 3)

Thank you!

yi.liu98@duke.edu