Adaptive Methods in Clinical Research

Practical 2: Group Sequential Designs

In this practical, we will compare the properties of a single-stage (non-adaptive) design with a variety of possible group sequential designs.

Suppose we are comparing a new experimental treatment for coronary heart disease with a control treatment, where the primary outcome is normally distributed. After discussion with the clinical investigators, we decide on the following design parameters for a two-arm trial:

- One-sided type I error rate of 5% ($\alpha = 0.05$)
- Power of 90% to detect a treatment effect of 0.5 ($\delta = 0.5, \beta = 0.1$)
- Standard deviation of the outcomes $\sigma = 1$
- Allocation ratio of 1:1 for the two arms

1. Calculate the sample size that would be required for a single-stage design with the above design parameters. [Hint: use `power.t.test` in R]

2. To investigate the different group sequential designs, we will use the OptGS Shiny App which is available at `https://newcastle-biostatistics.shinyapps.io/optgs/` [Note: please ask any questions you have on what all of the input parameters and outputs mean!]

   Go to the 'Design' tab and (using the default inputs), click on 'Update outputs' to get some example output. You should see an 'Operating characteristics summary', which has column headings defined as follows:

   - τ: the assumed value of the treatment effect θ
   - $P(\tau)$: the probability of rejecting H_0 when $\theta = \tau$
   - $ESS(\tau)$: the expected sample size when $\theta = \tau$
   - $SDSS(\tau)$: the standard deviation of the sample size when $\theta = \tau$
   - $MSS(\tau)$: two columns giving the median and modal sample size when $\theta = \tau$
   - $E_j(\tau)$: probability of stopping for efficacy at stage j when $\theta = \tau$
   - $F_j(\tau)$: probability of stopping for lack of benefit at stage j when $\theta = \tau$
   - $S_j(\tau)$: probability of stopping for efficacy or lack of benefit at stage j when $\theta = \tau$
   - cum$\{S_j(\tau)\}$: cumulative probability of stopping for efficacy or lack of benefit at stage j when $\theta = \tau$
   - max n: maximum sample size

3. Consider a two-stage group sequential design (again using the above design parameters) using O'Brien-Fleming stopping boundaries. Write down the following:

   a) Maximum sample size

   b) Expected sample under the null

   c) Expected sample size under the alternative

   d) Probability of stopping for lack of benefit at stage 1 under the null

   e) Probability of stopping for efficacy at stage 1 under the alternative

   How does this design compare with the single-stage design found earlier?

4. Try using Pocock stopping boundaries instead. What changes? Which of the two stopping boundaries do you prefer?

5. Now try near-optimal stopping boundaries with two different types of optimality criteria: Null-optimal and Alternative-optimal. Verify that these designs have a lower ESS under the null and alternative hypotheses, respectively, than either of the designs using O'Brien-Fleming or Pocock stopping boundaries. How large are these differences?

6. Find three-stage O'Brien-Fleming and Pocock stopping boundaries. What are the advantages and disadvantages of going from two to three stages? What do you think will happen if you consider designs with more than three stages?

7. Investigate what happens when you change the allocation ratio from 1:1 to 2:1 in favour of the experimental arm. When would such an approach be justified?