Adaptive Methods in Clinical Research

Practical 3: Designing Multi-Arm Multi-Stage Studies

You are the lead statistician of a clinical trial seeking to evaluate the utility of Telmisartan, an anti-hypertensive drug, to reduce insulin resistance in HIV positive patients on anti-retroviral therapy. Four doses, 40, 80, 120 and 160 mg are to be compared to placebo in a double blind study. The primary endpoint of the study is change in insulin resistance from baseline to 24 weeks.

There are several requirements. Specifically, the probability to claim incorrectly that any of investigation arms has a treatment effect different from the control (when in fact all of the effects are equal to the effect of the control) should be no more than 5%. Moreover, the probability to indentify the superior arm having an interesting value $\delta = 2$ while the rest arms correspond to the effect of uninteresting value $\delta_0 = 0.5$ should be 90%. You also have information from previous studies that the standard deviation of the primary endpoint is $\sigma = 4.4$ for both investigation and control arms, and you are confident in this value. Currently, several design options are being considered:

a) Design 0: No interim analyses, to assign equal number of patients to each dose and control arm.

b) One interim analysis after half of all patients been enroled, to assign equal number of patients to each dose, and equal number of patients at both stages. Several options are considered:

- Design 1: No early stopping for efficacy at the interim analysis, and drop any treatment that have test statistic below 0.
- Design 2: Early stopping for efficacy at the interim if the test statistic for at least one of doses is above $\frac{4}{3}c$ (where c is the critical efficacy bound at the end of the second stage) but drop any doses at the interim for which the test statistic is below 0.

You need to make a decision which of these options is the most favorable. An accurate way to do it is to compare the operating characteristics of these designs.

1. To compute the operating characteristics of the designs, you can use `R`-package `MAMS`.

   ```
   install.packages (MAMS) # if you do not have it yet
   library(MAMS)
   ```

   The first step is to evalute the "null" option, which is a single-stage design. The `MAMS`-package can compute the final (and only) critical value using the Dunnett's test that will be ensure a control of the family-wise error rate (FWER). Specifically, the function `mams` can be used.

   Use the code below (also given in the file `P3-MAMS-template.R`) as a starting point and plug-in all missing values given the information above.

```
design.0<-mams(K= , # number of experimental arms
     J= , # number of stages
     alpha= , # FWER
     power= ,  # Desired Power
     r=1, r0=1,  # Allocation ratio (1 to 1)
     delta= ,    # interesting value of effect
     delta0= , # uninteresting value of effect
 p=NULL, p0=NULL, # interesting/uninterseting effects on probability scale
                  # (leave as is)
    sd=     # standard deviation)
```

Run the code. Check the design critical value and maximum (and fixed) sample size. What is the sample size required for this single-stage design?

Now, we need to ensure that the FWER is controlled and the design achieves the desired power. For this, one can use simulations using the computed bounds. Use the code below as a starting point and plug-in values obtained above and the information available in the trial.

```
design.0.sims.null<-mams.sim(nsim=20000, # number of simulations
nMat=matrix(c(  ), nrow=1, ncol=5), # the matrix with sample sizes to be used
                                    # at each stage
pv=NULL,
l=c(  ),   # lower critical values at each stage
u=c(  ),   # upper critical value at each stage
deltav=c(0,0,0,0), # the treatment effect at investigation arms
                   # under the null hypothesis
sd= )        # standard deviation

summary(design.0.sims.null) # display the operating characteristics
design.0.sims.null$typeI    # display FWER
```

Run the code and check FWER.

```
summary(design.0.sims.null) # display the operating characteristics
design.0.sims.null$typeI    # display FWER
```

Now, we need to check the characteristics (power) under the least favourable configuration. Similarly, one can use the simulations

```
design.0.sims.lfc<-mams.sim(nsim=20000, # number of simulations
nMat=matrix(c(  ), nrow=1, ncol=5), # the matrix with sample sizes to be used
                                    # at each stage
pv=NULL,
```

```
l=c(  ),   # lower critical values at each stage
u=c(  ),   # upper critical value at each stage
deltav=c(2,0.5,0.5,0.5), # the treatment effect at investigation arms
                         # under the least favourable configuration (LFC)
sd= )        # standard deviation

summary(design.0.sims.lfc) # display the operating characteristics
design.0.sims.lfc$power   # display Power
```

Why do you think it is important to start the analysis from this null option?

2. Now, we will investigate the alternative options that allow for interim decisions to be made. Use function `mams` to find the boundary values for the design and the maximum sample sizes. You can use the code above as a starting point. However, now you have different number of stages, and you need to define the customised bounds. For example, it can be done as below

```
design.1<-mams(...,

 lshape="fixed", # use fixed value for the lower bound
  lfix= , # choose the lower value for dropping arms
  ushape="fixed",  # use fixed value for the lower bound
   ufix= ) # choose the upper value for stopping early for efficacy
           # Hint: instead of Inf you can use large value (e.g. 50) to decrease
           # computational time

design.1   # display characteristics
```

For Design 2, you will need to use a customised efficacy bound. This can be achieve by supplying the function to `ushape` variable. The function should require exactly one argument (e.g. the critical value) and return a vector of the length equal to the number of stages.

This will provide you with the maximum sample size, the cumulative sample size at each stage, lower and upper bounds to be used in the adaptive designs.

Use these for your simulations in `mams.sim` when checking FWER and Power. Note that you now have two stages, so in the corresponding function you need to change (among other things)

```
mams.sim(..., nMat=matrix(c(  ,  ), nrow=2, ncol=5),...,
u=c(   ,   ),
l=c(   ,    ),...)
```

Evaluate both designs and complete the table below:

Design	Lower Bound	Upper Bound	Final Crit. Value	SS per stage	Max SS	Expected SS under NULL	Expected SS under LFC
Design 0							
Design 1							
Design 2							

Given these results, what recommendation will you make on the design for this trial? Why? Can you foreseen any disadvantages of the chosen design?