**Comparing the optimal two-phase designs when the first phase experiment is either CRD or RBD**

This write-up compares the designs found from the simulated annealing algorithm where the first phase experiments are either completely randomised design (CRD) or randomised block design (RBD).

The objective functions used when the first phase experiment is CRD is

Note that there is more weight on maximising the degrees of freedom (DF) associated with the treatment effects in the Between Animals Within Runs stratum, because it is important to conduct the test on the treatment effects where every treatment contrast is present in the Between Animals Within Runs stratum. As the number of the treatment increases, more weights should be on maximising the DF associated with the treatment effects in the Between Animals Within Runs stratum, because there is more chance that some DF can lost to the other stratum. The objective function can be simplified as

The current objective functions used when the first phase experiment is RBD is

Currently, I had another thought to simply this objective function by performing two checks initially. This first check is to make sure the average efficiency factors associated with the animal (Within Cages) effect in the Within Runs stratum is 100%. From all the designs have found so far, the canonical efficiency factors associated with the animal (Within Cages) effect in the Within Runs stratum is always 100%. Note that the information matrix can be expressed as

hence, the trace of the information matrix becomes the sum of the eigenvalues which can be shown as follows

The numbers of the eigenvalue equals to the degrees of the freedoms (DF) associated with the Between Animals Within Cages Within Runs stratum which can be computed form the trace of the projection matrix. Therefore, if the canonical efficiency factors are all ones, the sum of the canonical efficiency factors should be the same as the DF associated with the Between Animals Within Cages Within Runs stratum. This method avoids the eigenvalue decomposition and in theory should speed up the algorithm.

The second check is to make sure the treatment allocation to runs is connected, i.e. there should be v – 1 DF associated with treatment effects in the Between Animals (Within Cages) Within Runs stratum. Since we do need to compute the average efficiency factor associated with the treatment effects in the Between Animals (Within Cages) Within Runs stratum, hence the eigenvalue decomposition is required and the DF is computed from the number of the non-zero eigenvalues present.

If either of these checks failed, the objective function will give a very low value, i.e. zero, so there is still a very low chance of the design been accepted for the SA search. If the both checks are passed, then the following an objective function is used

I have used for the first weight and for the second weight, this is because the change in can be very subtle.

The design parameters that are studied here consists of v = 6, r\_b = 3, r\_t = 2, n\_\gamma = 4 and n\_R = 15. If the first phase experiment is RBD, then number of cage, n\_C = 3.

**The first phase is CRD**

The first phase theoretical ANOVA is expressed as follows

$ANOVA

DF e Ani

Between Ani

Trt 5 1 2

Residual 12 1 2

Within 18 1 0

$EF

Trt eff.Trt

Between Ani

Trt 6 1

Within

Based on the random effects table, all the treatment information is in the Between Animal stratum.

The following allocation of animals to runs and tags were found and expressed as matrix,

[,1] [,2] [,3] [,4]

[1,] "AK" "AD" "AG" "AN"

[2,] "AD" "AK" "AN" "AG"

[3,] "AA" "AP" "AQ" "AC"

[4,] "AP" "AA" "AC" "AQ"

[5,] "AI" "AH" "AJ" "AL"

[6,] "AH" "AI" "AL" "AJ"

[7,] "AM" "AR" "AO" "AB"

[8,] "AR" "AM" "AB" "AO"

[9,] "AF" "AF" "AE" "AE"

The following allocation of treatments to runs and tags were found and expressed as matrix,

[,1] [,2] [,3] [,4]

[1,] "e" "d" "a" "b"

[2,] "d" "e" "b" "a"

[3,] "a" "d" "e" "c"

[4,] "d" "a" "c" "e"

[5,] "c" "b" "d" "f"

[6,] "b" "c" "f" "d"

[7,] "a" "f" "c" "b"

[8,] "f" "a" "b" "c"

[9,] "f" "f" "e" "e"

The canonical efficiency factors and average efficiency factor associated with the treatment effects in the Between Animals Within Runs stratum for this design can be expressed as follows

$can.eff

[1] 0.9166667 0.9166667 0.8888889 0.7500000 0.7500000

$ave.eff

[1] 0.8370323

The second phase theoretical ANOVA is expressed as follows

DF e Ani Run

Between Run

Between Ani

Trt 4 1 2 4

Residual 4 1 0 4

Within

Between Ani

Tag 1 1 2 0

Trt 5 1 2 0

Residual 7 1 2 0

Residual

Tag 2 1 0 0

Residual 12 1 0 0

$EF

Tag Trt eff.Tag eff.Trt

Between Run

Between Ani

Trt 3/4 1/8

Residual

Within

Between Ani

Tag 9 2/3 1 1/9

Trt 7920/1577 1320/1577

Residual

Tag 9 1

Compared this ANOVA table of the first phase experiment, 5 DF associated with the animal effects are lost from the first phase experiment. The 4 DF of 5DF for the animals are now in the Between Runs stratum and 1 DF associated with tag effects is now in the Between Animals Within Runs stratum. In addition, the amount of treatment information remains for conducting the test is now 0.8370 from 100% of the first phase experiment.

**The first phase is RBD**

The first phase theoretical ANOVA is expressed as follows

$ANOVA

DF e Cag:Ani Cag

Between Cag 2 1 2 12

Between Cag:Ani

Trt 5 1 2 0

Residual 10 1 2 0

Within 18 1 0 0

$EF

Trt eff.Trt

Between Cag

Between Cag:Ani

Trt 6 1

Within

Now all the treatment information is in the Between Animals Within Cages stratum. Note that there are 2 DF associated with the Between Cages stratum, so the DF for the residual MS for the Between Animals Within Cages stratum is now 10 compared to the DF of the residual MS for the Between Animals stratum were 12 for the CRD.

The following allocation of cages and animals to runs and tags were found and expressed as matrix where the upper case letter denotes the Cage ID and numeric numbers denotes the animal ID within each cage,

[,1] [,2] [,3] [,4]

[1,] "B4" "C5" "C3" "B1"

[2,] "C5" "B4" "B1" "C3"

[3,] "A1" "A2" "A6" "A3"

[4,] "A2" "A1" "A3" "A6"

[5,] "C6" "B3" "C2" "B5"

[6,] "B3" "C6" "B5" "C2"

[7,] "B2" "C1" "C4" "B6"

[8,] "C1" "B2" "B6" "C4"

[9,] "A5" "A5" "A4" "A4"

One initial observation is that the Cages is confounded with Runs, because Run 1, 2, 5, 6, 7 and 8 contain Cages B and C whereas Run 3, 4 and 9 contain only Cage A.

The following allocation of treatments to runs and tags were found and expressed as matrix,

[,1] [,2] [,3] [,4]

[1,] "d" "e" "c" "a"

[2,] "e" "d" "a" "c"

[3,] "a" "b" "f" "c"

[4,] "b" "a" "c" "f"

[5,] "f" "c" "b" "e"

[6,] "c" "f" "e" "b"

[7,] "b" "a" "d" "f"

[8,] "a" "b" "f" "d"

[9,] "e" "e" "d" "d"

The canonical efficiency factors and average efficiency factor associated with the treatment effects in the Between Animals Within Cages Within Runs stratum for this design can be expressed as follows

$can.eff

[1] 1.0000000 0.8943376 0.8888889 0.8333333 0.6056624

$ave.eff

[1] 0.8204481

Note that this average efficiency factor is slightly smaller than the previous case of 0.8370

The second phase theoretical ANOVA is expressed as follows,

$ANOVA

DF e Cag:Ani Cag Run

Between Run

Between Cag 1 1 2 12 4

Between Cag:Ani

Trt 3 1 2 0 4

Residual 4 1 0 0 4

Within

Between Cag 1 1 2 12 0

Between Cag:Ani

Tag 1 1 2 0 0

Trt 5 1 2 0 0

Residual 6 1 2 0 0

Residual

Tag 2 1 0 0 0

Residual 12 1 0 0 0

$EF

Tag Trt eff.Tag eff.Trt

Between Run

Between Cag

Between Cag:Ani

Trt 1 1/6

Residual

Within

Between Cag

Between Cag:Ani

Tag 9 2/3 1 1/9

Trt 15600/3169 2600/3169

Residual

Tag 9 1

Compared this ANOVA table to the previous one, 4 DF associated with the animal effects are lost from the first phase experiment. The 3 DF of 4DF for the animals are now in the Between Runs stratum and 1 DF associated with tag effects is now in the Between Animals Within Runs stratum. In addition, amount of treatment information remain for conducting the test is now 0.8204 from 100% of the first phase experiment.

I was thinking that maybe I can use the allocation of the animals from the first case where the first phase is CRD to the second case where the first phase is RBD. Since I just need to add the cage component to the allocation, the new allocation of cages and animals to runs and tags becomes

[,1] [,2] [,3] [,4]

[1,] "B5" "A4" "B1" "C2"

[2,] "A4" "B5" "C2" "B1"

[3,] "A1" "C4" "C5" "A3"

[4,] "C4" "A1" "A3" "C5"

[5,] "B3" "B2" "B4" "B6"

[6,] "B2" "B3" "B6" "B4"

[7,] "C1" "C6" "C3" "A2"

[8,] "C6" "C1" "A2" "C3"

[9,] "A6" "A6" "A5" "A5"

The following allocation of treatments to runs and tags is still the same as before,

[,1] [,2] [,3] [,4]

[1,] "e" "d" "a" "b"

[2,] "d" "e" "b" "a"

[3,] "a" "d" "e" "c"

[4,] "d" "a" "c" "e"

[5,] "c" "b" "d" "f"

[6,] "b" "c" "f" "d"

[7,] "a" "f" "c" "b"

[8,] "f" "a" "b" "c"

[9,] "f" "f" "e" "e"

The canonical efficiency factors and average efficiency factor associated with the treatment effects in the Between Animals Within Cages Within Runs stratum for this design can be expressed as follows

$can.eff

[1] 0.9160823 0.8888889 0.8051929 0.7309597 0.4736910

$ave.eff

[1] 0.7207011

which is significantly smaller than the previous two cases.

The second phase theoretical ANOVA is expressed as follows,

$ANOVA

DF e Cag:Ani Cag Run

Between Run

Between Cag

Trt 2 1 2 6 4

Between Cag:Ani

Trt 2 1 2 0 4

Residual 4 1 0 0 4

Within

Between Cag

Trt 2 1 2 6 0

Between Cag:Ani

Tag 1 1 2 0 0

Trt 5 1 2 0 0

Residual 5 1 2 0 0

Residual

Tag 2 1 0 0 0

Residual 12 1 0 0 0

$EF

Tag Trt eff.Tag eff.Trt

Between Run

Between Cag

Trt 7/9 7/54

Between Cag:Ani

Trt 7/9 7/54

Residual

Within

Between Cag

Trt 7/11 7/66

Between Cag:Ani

Tag 9 2/3 1 1/9

Trt 79440/18371 289/401

Residual

Tag 9 1

Note that Between Cages stratum has some treatment information. This can explained from the designs where within each run there are always different set of cage; for example, the first run there are 1 Cage A, 2 Cage B and 1 Cage C. In addition, treatments are also different Within the Same Cage and Run; for example, the Cage B of the first run has treatment e and a. Hence, even though treatment effects are not confounded with Cages, but treatment effects are confounded with Cages in this Phase 2 block structure. For the previous case, each run contains either two set of Cage B and C or Cage A; so there is no any partial confounding between the cages and runs; hence, the treatment information stays intact in the Between Animals Within Cages stratum.

The partial confounding between Cages and Runs can be confirmed by fitting the Cages as fixed effects and Runs as block effects, the theoretical ANOVA table can be expressed as

$ANOVA

DF e Run

Between Run

Cag.c1 1 1 4

Cag.c2 1 1 4

Residual 6 1 4

Within

Cag.c1 1 1 0

Cag.c2 1 1 0

Residual 25 1 0

$EF

Cag.c1 Cag.c2 eff.Cag.c1 eff.Cag.c2

Between Run

Cag.c1 3 1/4

Cag.c2 9 3/4

Within

Cag.c1 9 3/4

Cag.c2 3 1/4

Note that one of two Cage contrast derived from the basic contrasts were fitted to the ANOVA table. In one contrast, there is 1/4 of Cage information in the Between Runs stratum and 3/4 of Cage information in the Within Runs stratum. The other contrast has 3/4 of Cage information in the Between Runs stratum and 1/4 of Cage information in the Within Runs stratum.

In conclusion, we cannot simply use the allocation of animals form the CRD of the Phase 1 experiment to be used for the RBD for the Phase 1 experiment. This is because when designing the two-phase experiment where the first phase is CRD, we only need to concern the amount of the treatment information that is maximised in the Between Animals Within Runs stratum. For designing the two-phase experiment where the first phase is RBD, we need to make sure the amount of the treatment information that is maximised in the Between Animals Within Cages Within Runs stratum.

**Case 2**

The following designs shows that when the first phase experiment is either CRD or RBD; the test for the treatment effects can be conducted for second phase experiment with the same levels of accuracy.

The design parameters that are studied here consists of v=6, r\_b = 2, r\_t = 2, n\_\gamma = 4 and n\_R = 15. If the first phase experiment is RBD, then number of cage, n\_C = 2.

**The first phase is CRD**

The first phase theoretical ANOVA is expressed as follows

$ANOVA

DF e Ani

Between Ani

Trt 5 1 2

Residual 6 1 2

Within 12 1 0

$EF

Trt eff.Trt

Between Ani

Trt 4 1

Within

The following allocation of animals to runs and tags were found and expressed as matrix,

[,1] [,2] [,3] [,4]

[1,] "AA" "AB" "AK" "AD"

[2,] "AB" "AA" "AD" "AK"

[3,] "AE" "AF" "AG" "AC"

[4,] "AF" "AE" "AC" "AG"

[5,] "AJ" "AI" "AH" "AL"

[6,] "AI" "AJ" "AL" "AH"

The following allocation of treatments to runs and tags were found and expressed as matrix,

[,1] [,2] [,3] [,4]

[1,] "a" "b" "e" "d"

[2,] "b" "a" "d" "e"

[3,] "e" "f" "a" "c"

[4,] "f" "e" "c" "a"

[5,] "d" "c" "b" "f"

[6,] "c" "d" "f" "b"

The canonical efficiency factors and average efficiency factor associated with the treatment effects in the Between Animals Within Runs stratum for this design can be expressed as follows

$can.eff

[1] 1.00 1.00 1.00 0.75 0.75

$ave.eff

[1] 0.8823529

The second phase theoretical ANOVA is expressed as follows

$ANOVA

DF e Ani Run

Between Run

Between Ani

Trt 2 1 2 4

Residual 3 1 0 4

Within

Between Ani

Tag 1 1 2 0

Trt 5 1 2 0

Residual 3 1 2 0

Residual

Tag 2 1 0 0

Residual 7 1 0 0

$EF

Tag Trt eff.Tag eff.Trt

Between Run

Between Ani

Trt 1 1/4

Residual

Within

Between Ani

Tag 6 1

Trt 60/17 15/17

Residual

Tag 6 1

**The first phase is RBD**

The first phase theoretical ANOVA is expressed as follows

$ANOVA

DF e Cag:Ani Cag

Between Cag 1 1 2 12

Between Cag:Ani

Trt 5 1 2 0

Residual 5 1 2 0

Within 12 1 0 0

$EF

Trt eff.Trt

Between Cag

Between Cag:Ani

Trt 4 1

Within

The following allocation of cages and animals to runs and tags were found and expressed as matrix where the upper case letter denotes the Cage ID and numeric numbers denotes the animal ID within each cage,

[,1] [,2] [,3] [,4]

[1,] "A1" "A3" "B4" "B6"

[2,] "A3" "A1" "B6" "B4"

[3,] "A4" "A2" "B3" "B5"

[4,] "A2" "A4" "B5" "B3"

[5,] "A5" "A6" "B1" "B2"

[6,] "A6" "A5" "B2" "B1"

The Cage effects are confounded with tags as Tag 1, 2 contain only Cage A and Tag 3, 4 contain Cage B.

The following allocation of treatments to runs and tags were found and expressed as matrix,

[,1] [,2] [,3] [,4]

[1,] "a" "c" "d" "f"

[2,] "c" "a" "f" "d"

[3,] "d" "b" "c" "e"

[4,] "b" "d" "e" "c"

[5,] "e" "f" "a" "b"

[6,] "f" "e" "b" "a"

The canonical efficiency factors and average efficiency factor associated with the treatment effects in the Between Animals Within Cages Within Runs stratum for this design can be expressed as follows

$can.eff

[1] 1.00 1.00 1.00 0.75 0.75

$ave.eff

[1] 0.8823529

Note that the canonical efficiency factors and average efficiency factor are the same as before.

The second phase theoretical ANOVA is expressed as follows,

$ANOVA

DF e Cag:Ani Cag Run

Between Run

Between Cag:Ani

Trt 2 1 2 0 4

Residual 3 1 0 0 4

Within

Between Cag

Tag 1 1 2 12 0

Between Cag:Ani

Trt 5 1 2 0 0

Residual 3 1 2 0 0

Residual

Tag 2 1 0 0 0

Residual 7 1 0 0 0

$EF

Tag Trt eff.Tag eff.Trt

Between Run

Between Cag:Ani

Trt 1 1/4

Residual

Within

Between Cag

Tag 6 1

Between Cag:Ani

Trt 60/17 15/17

Residual

Tag 6 1

Due to the confounding of the Cages and Tags, 1DF associated with the tag effects that were in the Between Animals stratum before is now in the Between Cages stratum. The amount of treatment information for this designs is still 0.8823.