Optimization of nested step designs

Célia Fernandes[1], Paulo Ramos[2], João Tiago Mexia[3]

[1]Instituto Superior de Engenharia de Lisboa, Área de Matemática, Secção do Departamento de Engenharia de Electrónica e Telecomunicações e de Computadores, cfernandes@deetc.isel.ipl.pt
[2]Instituto Superior de Engenharia de Lisboa, Área de Matemática, Secção do Departamento de Engenharia de Electrónica e Telecomunicações e de Computadores, pramos@deetc.isel.ipl.pt
[3]Universidade Nova de Lisboa, Faculdade de Ciẹncias e Tecnologia, Departamento de Matemática, jtm@fct.unl.pt

SUMMARY

The number of treatments in a balanced nested design is the product of the number of levels in each factor. This number may be too large. As an alternative, in nested step designs there are as many sub-designs as factors plus one last sub-model used for error estimation. In each sub-design the branching is done only for the corresponding factor. The numbers of treatments is now the sum of the factor levels plus one. Moreover the amount of information for the different factors is more evenly distributed. We present a method to minimize the sum of the estimated variances of the estimators of the variance components.

Key words: Random effects, variance components, nested factors, nested step designs, optimization.

1. Introduction

In nested step designs (see Cox, Solomon, 2003) with u factors we have $u+1$ sub-designs. The usual balanced design has equal replication of all the levels that occur, so there are fewer degrees of freedom for estimating σ_i^2 for low i. The proposal here is use a_1 levels of factor 1, combined with a single level of all other factors, to give vector $\underline{y}_1$; then a new single level of factor 1, combined with a_2 new levels of factor 2, combined with a single level of all other factors, to give vector $\underline{y}_2$; and so on. In the last sub-model there is only one level for each factor nested inside the level of the preceding factor. Only for this model we consider replicates.

For $u=3$ we would have the design

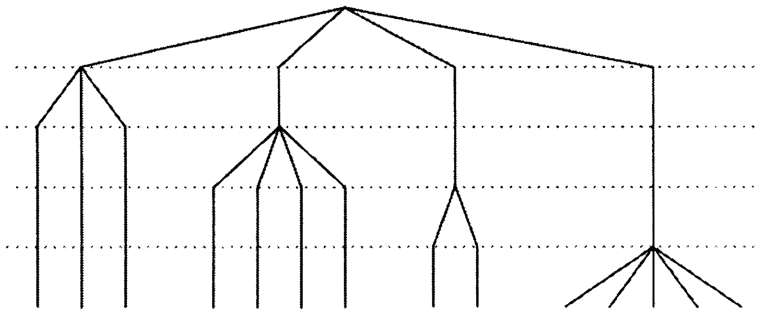

Figure 1. Nested step designs

The number of level combinations is $1+\sum_{j=1}^{u} a_j$ instead of, as in a balanced nested design, $\prod_{j=1}^{u} a_j$. Moreover, the number of observations will be $a_{u+1}+\sum_{j=1}^{u} a_j$ instead of $a_{u+1}\times\prod_{j=1}^{u} a_j$.

Thus the number of factor levels can be much higher then in the balanced case. This is important mainly for the first factors. Besides this the information, as we will see is distributed more evenly for all factors since we didn't have a concentration on the last factors as happen in the balanced case.

After presenting the model for step designs, we consider estimation and hypothesis testing.

The parameters in the models for step designs are the number of factors, the number of factor levels $a_1,\dots,a_u$ and the number of a_{u+1} replicates. If we have a variance component model for a step design we will be interested in minimizing the sum of the variance of our estimates of the variance components.

In what follows I_m will be the identity matrix of order m, and $\underline{1}_m$ the vector with m components equal to 1. We represent by $E(\underline{U})$ and $\Sigma(\underline{U})$ the mean vector and covariance matrix of $\underline{U}$. We write $\underline{y}\sim N(\underline{\mu},\sigma^2 M)$, when $\underline{y}$ is normal with $E(\underline{y})=\underline{\mu}$ and $\Sigma(\underline{y})=\sigma^2 M$, and $S\sim\gamma\chi_m^2$ if S is the product by γ of a central chi-square with m degrees of freedom. Moreover we take $\sum_{i=1}^{0} l_i=0$, whatever the l_i.

2. Model

In each sub-designs the branching is done only for the corresponding factor. The observations vector is

$$\underline{y} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \cdots \\ \underline{y}_u \\ \underline{y}_{u+1} \end{bmatrix}$$

with $\underline{y}_1$, $\underline{y}_2$, ..., $\underline{y}_u$, $\underline{y}_{u+1}$, independent vectors for sub-designs.

For the sub-vectors we have the models

$$\underline{y}_j = \underline{1}_{a_j}\left(\mu + \sum_{i=1}^{j-1} \beta_{j,i}\right) + \sum_{i=j}^{u+1} \underline{\beta}_{j,i}(a_j),\ j = 1,\ldots,u+1$$

where the $\beta_{j,i} \sim N\left(0, \sigma_i^2\right)$, $i = 1,\ldots,j-1$, $j = 1,\ldots,u+1$ and the $\underline{\beta}_{j,i}(a_j) \sim N\left(\underline{0}_{a_j}, \sigma_i^2 I_{a_j}\right)$, $i = j,\ldots,u+1$, $j = 1,\ldots,u+1$ are mutually independent. The error vector present in the last sub-model will be $\underline{e}_{u+1} = \underline{\beta}_{u+1,u+1}(a_{u+1})$. It is now straightforward to show that

$$E\left(\underline{y}_j\right) = 1_{a_j}\mu,\ j = 1,\ldots,u+1$$

and

$$\Sigma\left(\underline{y}_j\right) = \phi_j J_{a_j} + \psi_j I_{a_j},\ j = 1,\ldots,u+1$$

with $\phi_j = \sum_{i=1}^{j-1} \sigma_i^2$ and $\psi_j = \sum_{i=j}^{u+1} \sigma_i^2$, $j = 1,\ldots,u+1$. Thus

$$\Sigma\left[\left(I_{a_j} - \frac{1}{a_j} J_{a_j}\right)\underline{y}_j\right] = \psi_j\left(I_{a_j} - \frac{1}{a_j} J_{a_j}\right),\ j = 1,\ldots,u+1.$$

and so

$$S_j = \left\| \left(I_{a_j} - \frac{1}{a_j} J_{a_j}\right)\underline{y}_j \right\|^2$$

will be the product by ψ_j of a central chi-square with $a_j - 1$ degrees of freedom, $j = 1,\dots,u+1$, since $\left(I_{a_j} - \frac{1}{a_j} J_{a_j}\right)\underline{y}_j$ has null vector and $I_{a_j} - \frac{1}{a_j} J_{a_j}$ is an orthogonal projection matrix.

3. Inference

From

$$S_j \sim \gamma_{1,j}\chi^2_{a_j-1},\ j = 1,\dots,u+1$$

we get the unbiased estimators

$$\tilde{\gamma}_{1,j} = \frac{S_j}{a_j - 1},\ j = 1,\dots,u+1.$$

Since

$$\sigma_j^2 = \gamma_{1,j} - \gamma_{1,j+1},\ j = 1,\dots,u,$$

we obtain the unbiased estimators

$$\tilde{\sigma}_j^2 = \tilde{\gamma}_{1,j} - \tilde{\gamma}_{1,j+1},\ j = 1,\dots,u.$$

Moreover, we want to test

$$H_{0,j} : \sigma_j^2 = 0,\ j = 1,\dots,u$$

against

$$H_{1,j} : \sigma_j^2 > 0,\ j = 1,\dots,u.$$

Now, with

$$\theta_j = \frac{\gamma_{1,j}}{\gamma_{1,j+1}},\ j = 1,\dots,u$$

these hypothesis can be rewriten as

$$\begin{cases} H_{0,j} : \theta_j = 1 & j = 1,\dots,u \\ H_{1,j} : \theta_j > 1 & j = 1,\dots,u. \end{cases}$$

We point out that

$$\mathcal{F}_j = \frac{\frac{S_j}{a_j-1}}{\frac{S_{j+1}}{a_{j+1}-1}} = \frac{a_{j+1}-1}{a_j-1}\frac{S_j}{S_{j+1}} \sim \theta_j \frac{a_{j+1}-1}{a_j-1}\frac{\chi^2_{a_j-1}}{\chi^2_{a_{j+1}-1}},\ j=1,\dots,u$$

with $\theta_j = \frac{\gamma_{1,j}}{\gamma_{1,j+1}}$, $j=1,\dots,u$ is the product by θ_j of a variable with distribution F with a_j-1 and $a_{j+1}-1$ degrees of freedom. With $F(z|r,s)$ a central F distribution with r and s degrees of freedom and $f_{1-q,r,s}$ the corresponding quantile for the $1-q$ probability. The power functions of the tests with these statistics will be

$$Pot_j(\theta_j) = 1 - F\left(\frac{1}{\theta_j} f_{1-q,a_j-1,a_{j+1}-1} \middle| a_j-1, a_{j+1}-1\right),\ j=1,\dots,u.$$

These statistics increases with θ_j, $j=1,\dots,u$, so if we use them as test statistics, the corresponding F tests will be unbiased.

4. Optimization

Let $\dot{\gamma}_{1,1},\dots,\dot{\gamma}_{1,u+1}$ be positive estimators for the $\gamma_{1,1},\dots,\gamma_{1,u+1}$ obtained from pre-sampling. We now intend to, through a proper choice of the $a_1,\dots,a_u,a_{u+1}$, minimize

$$\sum_{j=1}^{u} \dot{V}ar(\tilde{\sigma}_j^2) = 2\sum_{j=1}^{u}\left(\frac{\dot{\gamma}_{1,j}^2}{a_j-1} + \frac{\dot{\gamma}_{1,j+1}^2}{a_{j+1}-1}\right)$$

where $\dot{V}ar(\tilde{\sigma}_j^2)$, $j=1,\dots,u$, is an estimate of $\tilde{\sigma}_j^2$ given the results of the pre-sampling.

We assume that $\sum_{j=1}^{u+1}(a_j-1)=n$, with n known, so we are led to use Lagrange multipliers. Thus we will have the auxiliary function

$$\begin{aligned} L(a_1,\dots,a_{u+1},\lambda) &= 2\sum_{j=1}^{u}\left(\frac{\dot{\gamma}_{1,j}^2}{a_j-1} + \frac{\dot{\gamma}_{1,j+1}^2}{a_{j+1}-1}\right) + \lambda\left[\sum_{j=1}^{u+1}(a_j-1)-n\right] = \\ &= 2\frac{\dot{\gamma}_{1,1}^2}{a_1-1} + 4\sum_{j=2}^{u}\frac{\dot{\gamma}_{1,j}^2}{a_j-1} + 2\frac{\dot{\gamma}_{1,u+1}^2}{a_{u+1}-1} + \lambda\left[\sum_{j=1}^{u+1}(a_j-1)-n\right] \end{aligned}$$

To lighten the writing we take $c_j = a_j - 1$, $j = 1,\dots,u+1$, getting

$$\begin{cases} \dfrac{\partial L}{\partial c_1} = 0 \\ \\ \dfrac{\partial L}{\partial c_j} = 0 \\ \\ \dfrac{\partial L}{\partial c_{u+1}} = 0 \\ \\ \dfrac{\partial L}{\partial \lambda} = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda = 2\dfrac{\dot{\gamma}_{1,1}^2}{c_1^2} \\ \\ \lambda = 4\dfrac{\dot{\gamma}_{1,j}^2}{c_j^2} \qquad j = 2,\dots,u \\ \\ \lambda = 2\dfrac{\dot{\gamma}_{1,u+1}^2}{c_{u+1}^2} \\ \\ \sum_{j=1}^{u+1} c_j - n = 0 \end{cases}$$

From the first three conditions we obtain

$$b_j^2 = 2\frac{\dot{\gamma}_{1,j}^2}{\dot{\gamma}_{1,1}^2} \Rightarrow b_j = \sqrt{2\frac{\dot{\gamma}_{1,j}^2}{\dot{\gamma}_{1,1}^2}} \text{ for } b_j = \frac{c_j}{c_1}, j = 2,\dots,u$$

$$b_{u+1}^2 = \frac{\dot{\gamma}_{1,u+1}^2}{\dot{\gamma}_{1,1}^2} \Rightarrow b_{u+1} = \sqrt{\frac{\dot{\gamma}_{1,u+1}^2}{\dot{\gamma}_{1,1}^2}} \text{ for } b_{u+1} = \frac{c_{u+1}}{c_1}.$$

From the last condition we have

$$c_1 = \frac{n}{1 + \sum_{j=2}^{u+1} b_j}.$$

We know that $c_1 = \frac{c_j}{b_j}$, $j = 2,\dots,u+1$, then using the previous results we get

$$c_j = \frac{nb_j}{1 + \sum_{l=2}^{u+1} b_l}, \quad j = 2,\dots,u+1.$$

Thus the a_j, $j = 1,\dots,u+1$, will be approximately proportional to n.

In what follows we will show one application of this theory:

Application:

$$\tilde{\gamma}^2_{1,1}=5,33;\ \tilde{\gamma}^2_{1,2}=3,45;\ \tilde{\gamma}^2_{1,3}=2,11;\ \tilde{\gamma}^2_{1,4}=1,09;\ \tilde{\gamma}^2_{1,5}=0,56$$

We obtain

Table 1. Results

n	$Var(\hat{\sigma}_1^2)$	$Var(\hat{\sigma}_2^2)$	$Var(\hat{\sigma}_3^2)$	$Var(\hat{\sigma}_4^2)$	a_1	a_2	a_3	a_4	a_5
100	0,6643	0,4297	0,3281	0,2763	25	29	22	16	8

During the development of this applications we use the fact of

$$\gamma_{1,j} > \gamma_{1,j+1},\ j=1,\ldots,u.$$

5. Conclusions

We conclude with the following remarks:

1. We point out that, $\overline{J}_s = I_s - \frac{1}{s}J_s$ and $\frac{1}{s}J_s$ are mutually orthogonal projection matrices. These matrices constitute a basis for a commutative Jordan algebra (see Seely, 1971), the variance-covariance matrices of the sub-models belong to such algebras.
2. The possibility of negative estimators has been considered by many authors, for instance see Nelder (see Nelder, 1954). The main inference to be had when we get $\tilde{\sigma}_i^2 < 0$ is that σ_i^2 must be null.
3. Further work is intended to recover the degrees of freedom, one per sub-model, that are wasted.
4. The optimization considered in *section 4* is a way of using the information in the preliminary estimators when all are positive.
5. The nested step designs turned out to be an valid alternative for the balanced nested designs. Thus, we can work with less observations and the amount of information for the different factors is more evenly distributed.

REFERENCES

Cox D., Solomon P. (2003). *Components of Variance*. Chapman and Hall.

Nelder J. (1954). The interpretation of negative components of variance. *Biometrika*, Vol. 41, p. 544-548.

Seely J. (1971). Quadratic Subspaces and Completeness. *Ann. Math. Stat.*, Vol. 43, p. 710-721.