Tatra Mt. Math. Publ. **39** (2008), 145–153

TATRA MOUNTAINS
Mathematical Publications

ON OPTIMAL DESIGNS IN RANDOM INTERCEPT MODELS

Rainer Schwabe — Thomas Schmelter

ABSTRACT. In the one-sample case standard optimal designs retain their optimality if a random intercept is present. In a multi-sample situation the variability of the intercept may have substantial influence on the choice of the optimal design.

1. Introduction

Mixed models have attracted growing interest in the biosciences, when replicated measurements are available from different individuals. While the corresponding statistical analysis is well-developed, only few results are available on optimal designs. For a recent survey we refer to Entholzner *et al.* (2005).

In the present note we focus on the widely used assumption of a random intercept, while the treatment effects are supposed to be fixed. In analysis of variance settings this situation is properly described by random block effects. For this random intercept setting there is a considerable number of different attempts for deriving optimal designs. Atkins and Cheng (1999) provide optimal designs for polynomial regression, Debusho and Haines (2006) derive optimal designs for linear regression under constraints on the replications and Fedorov and Hackl (1997, p. 75) present an equivalence theorem for this situation (see also the references therein).

In Section 2 we introduce the general random intercept model. For the one-sample situation it is shown in Section 3 that standard optimal designs from fixed effects models retain their optimality also in the presence of random intercepts. In Section 4 we consider a simple extension to the situation of the comparison of two treatments, when measurements at baseline are available. Already there the variability of the intercept has substantial influence on the choice of the optimal design. In the final Section 5 some problems of exact designs are discussed.

2000 Mathematics Subject Classification: 62K05, 62J05.
Keywords: optimal design, random intercept, random block effect, treatment comparison.

2. The random intercept model

By a linear model with random intercept we describe the situation that we have n individuals, $i = 1, \ldots, n$, with m_i observations each and the individual effects have only influence on the overall level of the response. Hence, the jth observation Y_{ij} at individual i can be written as

$$Y_{ij} = a_i + \mathbf{f}(x_{ij})^\top \, \boldsymbol{\beta} + \varepsilon_{ij} \,,$$

where x_{ij} is the experimental setting, $j = 1, \ldots, m_i$, $\mathbf{f} = (f_1, \ldots, f_p)^\top$ is a set of (known) regression functions and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a p-dimensional vector of parameters for the effects of the experimental settings. Here, $\mathbf{f}$ may be the identity for straight line regression, a set of dummy variables for analysis of variance models, or it may be even of a more complicated structure (see Section 4).

The term a_i denotes the individual random intercept with $E(a_i) = \mu$ and $\mathrm{var}(a_i) = \sigma_a^2$. The random observational errors ε_{ij} are assumed to be homoscedastic, $E(\varepsilon_{ij}) = 0\,,\ \mathrm{var}(\varepsilon_{ij}) = \sigma^2$, and all a_i and ε_{ij} are uncorrelated. Further analysis will depend on the variance ratio $d = \sigma_a^2/\sigma^2$. We will focus on the population parameters $\boldsymbol{\theta} = (\mu, \boldsymbol{\beta}^\top)^\top$ for the mean location and the effects of the experimental settings. For simplicity we assume that the number of observations per individual is constant, i.e., $m_i = m$.

Denote by $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{im})^\top$ the vector of observations for individual i. The corresponding covariance matrix $\mathrm{cov}(\mathbf{Y}_i) = \sigma^2\mathbf{V}$ is completely symmetric, $\mathbf{V} = \mathbf{I}_m + d\mathbf{1}_m\,\mathbf{1}_m^\top$, where $\mathbf{I}_m$ denotes the $m \times m$ identity matrix and $\mathbf{1}_m$ is a vector of length m with all entries equal to one. The individual fixed effect design matrix $\mathbf{X}_i = (\mathbf{1}_m \,|\, \mathbf{F}_i)$ can be decomposed into the first column of ones corresponding to the mean intercept μ and the effects design matrix $\mathbf{F}_i = \big(\mathbf{f}(x_{i1}) \,|\, \ldots \,|\, \mathbf{f}(x_{im})\big)^\top$ for the parameter vector $\boldsymbol{\beta}$.

The individual information $\mathbf{X}_i^\top\mathbf{V}^{-1}\mathbf{X}_i = \mathbf{X}_i^\top\mathbf{X}_i - \frac{d}{1+md}\mathbf{X}_i^\top\mathbf{1}_m\,\mathbf{1}_m^\top\mathbf{X}_i$ is proportional to the inverse of the variance-covariance matrix $\mathrm{cov}(\hat{\boldsymbol{\theta}}_i)$ if $\mathbf{X}_i$ is of full column rank and $\boldsymbol{\theta}$ is estimated on an individual basis by the best linear unbiased estimator $\hat{\boldsymbol{\theta}}_i = (\mathbf{X}_i^\top\,\mathbf{V}^{-1}\,\mathbf{X}_i)^{-1}\,\mathbf{X}_i^\top\mathbf{V}^{-1}\mathbf{Y}_i = (\mathbf{X}_i^\top\mathbf{X}_i)^{-1}\mathbf{X}_i^\top\mathbf{Y}_i$.

On the population basis the best linear unbiased estimator can be computed as a matrix mean $\hat{\boldsymbol{\theta}} = \left(\sum_{i=1}^{n} \mathbf{X}_i^\top\mathbf{V}^{-1}\mathbf{X}_i\right)^{-1} \sum_{i=1}^{n} \mathbf{X}_i^\top\mathbf{V}^{-1}\mathbf{X}_i\hat{\boldsymbol{\theta}}_i$ of the individual $\hat{\boldsymbol{\theta}}_i$ if d is known. Then $\mathrm{cov}(\hat{\boldsymbol{\theta}}) = \sigma^2\,\mathbf{M}_d^{-1}$, where $\mathbf{M}_d = \sum_{i=1}^{n} \mathbf{X}_i^\top\,\mathbf{V}^{-1}\,\mathbf{X}_i$ is the information matrix on population basis. The subscript d indicates the dependence on the variance ratio d. As $\mathbf{M}_d = \sum_{i=1}^{n} \mathbf{X}_i^\top\,\mathbf{X}_i - \frac{d}{1+md}\sum_{i=1}^{n} \mathbf{X}_i^\top\,\mathbf{1}_m\,\mathbf{1}_m^\top\,\mathbf{X}_i$, the information is strictly decreasing in the positive semidefinite sense if the variance ratio increases, i.e., $\mathbf{M}_d \geq \mathbf{M}_{d'}$ for $d' > d$.

Partitioning the information matrix according to μ and β yields

$$\mathbf{M}_d = \frac{1}{1+md}\left(\begin{array}{c|c} nm & \sum_{i=1}^n \mathbf{1}_m^\top \mathbf{F}_i \\ \hline \sum_{i=1}^n \mathbf{F}_i^\top \mathbf{1}_m & (1+md)\sum_{i=1}^n \mathbf{F}_i^\top \mathbf{F}_i - d\sum_{i=1}^n \mathbf{F}_i^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{F}_i \end{array}\right).$$

If interest is in the fixed effects $\boldsymbol{\beta}$ only, then by the rules for inverting partitioned matrices the corresponding partial information $\mathbf{M}_{\boldsymbol{\beta},d} = \mathrm{cov}(\hat{\boldsymbol{\beta}})/\sigma^2$ equals

$$\mathbf{M}_{\boldsymbol{\beta},d} = \sum_{i=1}^n \mathbf{F}_i^\top \mathbf{F}_i - \frac{d}{1+md}\sum_{i=1}^n \mathbf{F}_i^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{F}_i - \frac{1}{nm}\frac{1}{1+md}\sum_{i=1}^n \mathbf{F}_i^\top \mathbf{1}_m \sum_{j=1}^n \mathbf{1}_m^\top \mathbf{F}_j\,.$$

We also consider the limiting models for $d = 0$ and $d \to \infty$, respectively:

For $d = 0$ we obtain the fixed effects model without individual intercepts

$$Y_{ij} = \mu + \mathbf{f}(x_{ij})^\top \boldsymbol{\beta} + \varepsilon_{ij}\,.$$

Obviously, $\mathbf{M}_d$ tends to $\mathbf{M}_0 = \sum_{i=1}^n \mathbf{X}_i^\top \mathbf{X}_i$ for $d \to 0$. Similarly, $\mathbf{M}_{\boldsymbol{\beta},d}$ tends to $\mathbf{M}_{\boldsymbol{\beta},0} = \sum_{i=1}^n \mathbf{F}_i^\top \mathbf{F}_i - \frac{1}{nm}\sum_{i=1}^n \mathbf{F}_i^\top \mathbf{1}_m \sum_{j=1}^n \mathbf{1}_m^\top \mathbf{F}_j$.

For $d \to \infty$ we introduce the fixed effects model with fixed individual effects

$$Y_{ij} = \mu_i + \mathbf{f}(x_{ij})^\top \boldsymbol{\beta} + \varepsilon_{ij}\,.$$

Here, the full parameter vector $(\mu_1, \ldots, \mu_n, \beta_1, \ldots, \beta_p)^\top$ has dimension $n+p$ and the corresponding information matrix has the form

$$\mathbf{M}_\infty = \left(\begin{array}{c|c} m\,\mathbf{I}_n & \begin{array}{c} \mathbf{1}_m^\top \mathbf{F}_1 \\ \vdots \\ \mathbf{1}_m^\top \mathbf{F}_n \end{array} \\ \hline \mathbf{F}_1^\top \mathbf{1}_m \cdots \mathbf{F}_n^\top \mathbf{1}_m & \sum_{i=1}^n \mathbf{F}_i^\top \mathbf{F}_i \end{array}\right).$$

For $\boldsymbol{\beta}$ the corresponding partial information matrix can be calculated by

$$\mathbf{M}_{\boldsymbol{\beta},\infty} = \sum_{i=1}^n \mathbf{F}_i^\top \mathbf{F}_i - \frac{1}{m}\sum_{i=1}^n \mathbf{F}_i^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{F}_i\,.$$

Hence, we obtain the following result, which establishes the partial information matrix $\mathbf{M}_{\boldsymbol{\beta},d}$ as a convex combination of its limiting counterparts.

Lemma 1.

$$\mathbf{M}_{\boldsymbol{\beta},d} = \frac{1}{1+md}\,\mathbf{M}_{\boldsymbol{\beta},0} + \frac{md}{1+md}\,\mathbf{M}_{\boldsymbol{\beta},\infty}\,.$$

Note that the partial information matrix $\mathbf{M}_{\boldsymbol{\beta},d}$ tends to $\mathbf{M}_{\boldsymbol{\beta},\infty}$ as d tends to ∞.

3. Design issues

The quality of the estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$ depends on the experimental settings x_{ij} , $i = 1, \ldots, n$, $j = 1, \ldots, m$, through the information matrices $\mathbf{M}_d$ and $\mathbf{M}_{\boldsymbol{\beta},d}$, respectively. The aim in experimental design is to choose those settings from a design region $\mathcal{X}$ in order to minimize the covariance $\operatorname{cov}(\hat{\boldsymbol{\theta}})$ or $\operatorname{cov}(\hat{\boldsymbol{\beta}})$ or parts of it, which is equivalent to maximize the corresponding information matrices $\mathbf{M}_d$ or $\mathbf{M}_{\boldsymbol{\beta},d}$, respectively. As those matrices are not completely ordered, a uniform optimization is not possible, in general. Therefore, some real valued functionals that lay emphasis on particular properties of the estimators will be optimized. The most popular design criterion is the D-criterion, which aims at maximizing the determinant of the information matrix $\mathbf{M}_d$. This is equivalent to minimizing the volume of a confidence ellipsoid for $\boldsymbol{\theta}$ under the assumption of normality.

Further design criteria include the A-criterion, which aims at minimizing the trace of the standardized covariance matrix $\mathbf{M}_d^{-1}$. If interest is in the effects $\boldsymbol{\beta}$ only, $D_{\boldsymbol{\beta}}$- and $A_{\boldsymbol{\beta}}$-optimality are defined in terms of the determinant and the trace of the inverse $\mathbf{M}_{\boldsymbol{\beta},d}^{-1}$ of the corresponding partial information matrix. As it is readily seen, $\det(\mathbf{M}_d) = \frac{nm}{1+md} \det(\mathbf{M}_{\boldsymbol{\beta},d})$ holds by the formula for the determinant of partitioned matrices. Hence, D- and $D_{\boldsymbol{\beta}}$- optimality coincide also in random intercept models, a well-known fact in the fixed effects setting.

Lemma 2. *A design (x_{ij}) is D-optimal if and only if it is $D_{\boldsymbol{\beta}}$-optimal.*

Another class of criteria is based on the standardized variance function

$$v_d(x) = \operatorname{var}\big(\hat{\mu} + \mathbf{f}(x)^\top \hat{\boldsymbol{\beta}}\big)/\sigma^2 = \big(1, \mathbf{f}(x)^\top\big)\mathbf{M}_d^{-1}\big(1, \mathbf{f}(x)^\top\big)^\top$$

for the prediction of the mean response on the design region, $x \in \mathcal{X}$. The integrated mean squared error (IMSE) criterion, for example, aims at minimizing $\int_{\mathcal{X}} v_d(x)\,\mathrm{d}x$, while the G-criterion aims at minimizing the maximal variance over the design region $\max_{x\in\mathcal{X}} v_d(x)$. Note that for fixed effects models the D- and G-optimality are equivalent in the approximate theory due to the celebrated Kiefer-Wolfowitz equivalence theorem. However, this coincidence does not hold in models with random effects in general, as will be seen later.

If we consider designs which are uniform across the individuals, i.e., for which the experimental settings are the same for each individual, $x_{ij} \equiv x_j$, then the situation dramatically simplifies. In this case the individual design matrices coincide, $\mathbf{F}_i = \mathbf{F}_1$ and $\mathbf{X}_i = \mathbf{X}_1$, respectively, and $\mathbf{X}_1$ has to be of full column rank to allow for estimability of $\boldsymbol{\theta}$. Moreover, $\hat{\boldsymbol{\theta}} = \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{\theta}}_i$ reduces to the average of the individually fitted values for the parameters.

The standardized covariance matrix $\mathbf{M}_d^{-1}$ decomposes additively into the corresponding matrix $\mathbf{M}_0^{-1}$ for the fixed effects model without individual intercepts

and the variability of the random intercept (see, e.g., Entholzner *et al.*, (2005)). Hence, for linear criteria like the A-, $A_{\boldsymbol{\beta}}$- and IMSE-criterion the optimal design in the fixed effects model without individual intercepts ($d = 0$) remains optimal for all values of the variability d. For the reduced information matrix we observe

$$\mathbf{M}_{\boldsymbol{\beta},0} = n \left(\mathbf{F}_1^\top \mathbf{F}_1 - \frac{1}{m} \mathbf{F}_1^\top \mathbf{1}_m \mathbf{1}_m^\top \mathbf{F}_1 \right) = \mathbf{M}_{\boldsymbol{\beta},\infty}$$

and, consequently, by Lemma 1 $\mathbf{M}_{\boldsymbol{\beta},d} = \mathbf{M}_{\boldsymbol{\beta},0}$ is independent of d. Thus, the D-optimal design for the fixed effects model without individual intercepts is D- and D_β-optimal for every $d \geq 0$ in view of Lemma 2.

Moreover, for the variance function we obtain $v_d(x) = \frac{d}{n} + v_0(x)$. As a consequence, also the G-optimal design is independent of the value of d.

4. Treatment comparison with measurements at baseline

If different treatments are to be compared, out of which only one can be administered to each individual, then it is evident that the experimental conditions cannot be chosen uniformly across the individuals.

To illustrate the effect of this imbalance we consider a simplified model with linear response in time (or dosage) for several mutually exclusive treatments,

$$Y_{kij} = a_{ki} + \beta_k \, t_{kij} + \varepsilon_{kij} \, .$$

Here, Y_{kij} is the jth observation at individual i in treatment group k, where we have p treatments ($k = 1, \ldots, p$), n_k different individuals in each group and $m_{ki} = m$ observations at each individual. The term a_{ki} denotes the baseline of individual i in treatment group k, which are assumed to come from the same population, $E(a_{ki}) = \mu$, $\operatorname{var}(a_{ki}) = \sigma_a^2$. The corresponding experimental conditions $x_{kij} = (k, t_{kij})$ now consist of a discrete component k for the treatment and a continuous component t for time or dosage, thus, resulting in an analysis of covariance model. We assume that time points t can be chosen from the unit interval, $0 \leq t \leq 1$, maybe after standardization. The effects parameter vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is the vector of the slopes β_k related to the kth treatment, which we deliberately assume to be fixed. (Random slopes will cause significantly more severe problems, which will be treated elsewhere.) If we set $f_k(k', t) = t$ for $k = k'$, and $f_k(k', t) = 0$ otherwise, then the present model fits into the framework of Section 2.

For simplicity we assume that the number n of individuals is a multiple of p such that equal numbers of individuals, $n_k = n/p$, can be assigned to each

treatment. Furthermore, we confine ourselves to the situation where for all individuals within each treatment group measurements are taken at the same time points, $t_{kij} = t_{kj}$. Such designs can be shown to be optimal in a generalized setup (S c h m e l t e r , (2006)). Symmetry considerations suggest that those time points should be independent of the treatment, $t_{kj} = t_j$, and by majorization arguments observations should be restricted to the extreme time points $t = 0$ (at baseline) and $t = 1$ (at maximal dosage or elapsed time), when we allow for multiple observations at single time points. The corresponding designs are, then, completely characterized by the numbers $m(0)$ and $m(1) = m - m(0)$ of observations at $t = 0$ or 1, respectively. Thus, design optimization reduces to the search for the optimal proportion $w = \frac{m(0)}{m}$ of observations at baseline, $t = 0$. As all these simplifications are only valid in a generalized setup, we also allow for optimal weights w not being a multiple of $1/m$. For practical applications some rounding may be required, which will be discussed in the final Section 5.

For the comparison of $p = 2$ treatments optimal designs have been obtained in the fixed effects model without individual intercepts, $Y_{kij} = \mu + \beta_k\, t_j + \epsilon_{ijk}$, by S c h w a b e (1996, p. 109–110). Generalization of this result to $p > 2$ establishes that $w = \frac{1}{p+1}$ is D-optimal for $d = 0$. On the other hand side $w = \frac{1}{2}$ is seen to be $D_{\boldsymbol{\beta}}$-optimal in the fixed effects model with fixed individual intercepts, $Y_{kij} = \mu_{ki} + \beta_k\, t_j + \varepsilon_{ijk}$. Consequently, it can be suspected that the D-optimal proportion w varies continuously from $\frac{1}{p+1}$ to $\frac{1}{2}$ if the variance ratio d increases from zero to infinity.

For the present model the information matrix is given by

$$\mathbf{M}_d = \frac{m}{1+md}\left(\begin{array}{c|c} n & n_1(1-w)\mathbf{1}_p^\top \\ \hline n_1(1-w)\mathbf{1}_p & n_1(1-w)(1+wmd)\mathbf{I}_p \end{array}\right).$$

By the formula for the determinant of partitioned matrices we obtain $\det(\boldsymbol{M}_d) = cw(1-w)^p\,(1+wmd)^{p-1}$, where $c > 0$ is a generic constant. This is maximized by the D-optimal proportion

$$w = \frac{1}{4}\left(1 - \frac{p+1}{pmd} + \sqrt{\left(1 - \frac{p+1}{pmd}\right)^2 + \frac{8}{pmd}}\right),$$

which is increasing in d (see Figure 1) and tends to $\frac{1}{p+1}$ for $d \to 0$. Due to Lemma 2 this proportion is also $D_{\boldsymbol{\beta}}$-optimal for the treatment effects $\boldsymbol{\beta}$.

For the other criteria we have to determine the inverse

$$\mathbf{M}_d^{-1} = \frac{1}{nm}\left(\begin{array}{c|c} md + \frac{1}{w} & -\frac{1}{w}\mathbf{1}_p^\top \\ \hline -\frac{1}{w}\mathbf{1}_p & \frac{p(1+md)}{(1-w)(1+wmd)}\mathbf{I}_p + \frac{1}{w(1+wmd)}\mathbf{1}_p\mathbf{1}_p^\top \end{array}\right)$$

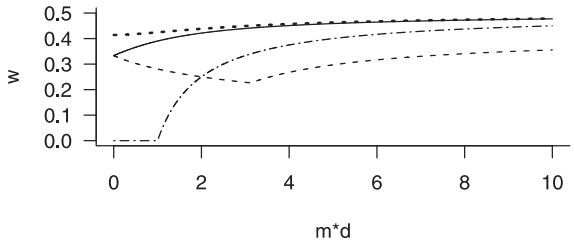

FIGURE 1. Optimal proportions w at $t = 0$: D- (—), G- (- - -), A_β-criterion ($\cdots$), and for minimizing $\max_k v_d(k,1)$ (-·-).

of the information matrix. Thus, we obtain for the variance function

$$v_d(k,t) = \frac{1}{nm}\left(md + \frac{1}{w} - \frac{2}{w}t + \left(\frac{1}{w} + \frac{p}{1-w} + \frac{(p-1)md}{1+wmd}\right)t^2\right),$$

which is constant in k. As the variance function is a polynomial of degree two in t with positive leading term, it attains its maximum at 0, where $v_d(k,0) = \frac{1}{nm}\left(md + \frac{1}{w}\right)$, or at 1, where

$$v_d(k,1) = \frac{1}{nm}\left(md + \frac{p}{1-w} + \frac{(p-1)md}{1+wmd}\right).$$

Before we determine the G-optimal proportion, first we consider the auxiliary criterion of minimizing $\max_k v_d(k,1)$ in order to illustrate the structural dependence on d. For the fixed effects model without individual intercepts ($d = 0$) all observations will be taken at 1, i.e., $w = 0$. This remains optimal for small $d > 0$. However, as d increases, some information at 0 may become valuable for the prediction at 1. If $md > \sqrt{p}/\sqrt{p-1}$, the variance $v_d(k,1)$ is minimized for

$$w' = \left(\sqrt{p} - \sqrt{p-1}\right)\left(\sqrt{p-1} - \frac{1}{md}\sqrt{p}\right).$$

Now, for the G-criterion also $v_d(k,0)$ has to be taken into account: $v_d(k,0)$ is monotonically decreasing in w and there is a unique point of intersection with $v_d(k,1)$ at

$$w'' = \frac{1}{4}\left(\sqrt{\left(p - 2 + \frac{p+1}{md}\right)^2 + \frac{8}{md}} - \left(p - 2 + \frac{p+1}{md}\right)\right).$$

As $v_d(k,1)$ is strictly convex in w, the G-optimal proportion is obtained as $w = \max(w', w'')$. For $p = 2$ this optimal proportion is plotted in Figure 1 together with the D-optimal proportions. From this it becomes evident that even for random intercept models the equivalence of D- and G-optimality is not retained, which for the fixed effects model without individual intercept holds due

to the Kiefer-Wolfowitz equivalence theorem. The non-monotonic dependence of the G-optimal proportion on the variability parameter d has to be highlighted. For $p = 2$ we have plotted the D- and G-efficiencies of the limiting optimal designs in dependence on the true variance ratio d in Figure 2. Apparently, the proportion $w_0 = \frac{1}{p+1}$, which is optimal for the fixed effects models without individual intercepts ($d = 0$), performs well over the whole region for d with a minimal D-efficiency of 0.92 for $d \to \infty$ and a minimal G-efficiency of 0.96 for $md \approx 1.10$. The D-optimal limiting proportion $w_{D,\infty} = 0.5$ shows a minimal D-efficiency of 0.94 at $d = 0$, whereas the G-optimal limiting proportion $w_{G,\infty} = \sqrt{2} - 1$ only results in a minimal G-efficiency of 0.87.

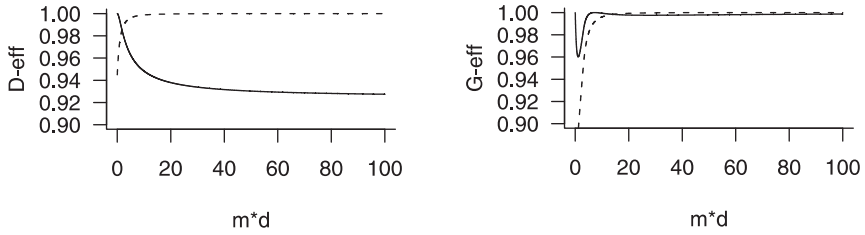

FIGURE 2. D- (left) and G-efficiencies (right): w_0 (—), $w_{\cdot,\infty}$ (- - -).

Similar to the G-criterion there is also a structural change if we are interested in contrasts of the treatment effects, e. g. in $\beta_1 - \beta_2$ for $p = 2$. More generally, if we are interested in a set of $p-1$ orthogonal contrasts in $\beta_1, \ldots, \beta_p$, then the corresponding covariance matrix is proportional to $\frac{p(1+md)}{(1-w)\,(1+wmd)}\,\mathbf{I}_{p-1}$, and the leading factor is minimized by the proportion $w = \frac{1}{2} - \frac{1}{2md}$, if $md > 1$, and $w = 0$ otherwise. Here, this also means, that, although contrasts are estimated efficiently at the highest level $t = 1$ in the fixed effects model without individual intercepts, additional information may come in from $t = 0$ if d becomes large.

Additionally, the $A_{\boldsymbol{\beta}}$-optimal proportions, which minimize belongs to formula

$$(\mathbf{M}^{-1}_{\boldsymbol{\beta},d}) = \frac{p}{nm}\left(\frac{1}{w} + \frac{p}{1-w} + \frac{(p-1)md}{1+wmd}\right),$$

are depicted in Figure 1. Even for this linear criterion there is a substantial dependence of the optimal design on d.

5. Discussion

In random intercept models optimal designs are not affected if all individuals can be treated under the same regime. In more complicated settings, where

different treatments are required, the optimal design may well depend on the intra-individual correlation, in particular, if the number of replications is large. However, the efficiencies of the locally optimal designs stay remarkably high, if the variability is misspecified. For a fixed number m of replication an optimal design can be realized by rounding the optimal proportion to one of the adjacent multiples of $1/m$. Further improvements may be achieved if not all of the individuals within one treatment group are measured at the same time points, in particular, if the number of replications is small.

REFERENCES

[AC99] ATKINS, J. E.—CHENG, C.-S.: *Optimal regression designs in the presence of random block effects*. J. Statist. Plann. Inference **77** (1999), 321–335.

[DH06] DEBUSHO, L. K.—HAINES, L. M.: *V- and D-optimal population designs for the simple linear regression model with a random intercept term*. J. Statist. Plann. Inference **138** (2008), 1116-1130.

[EBSS05] ENTHOLZNER, M.—BENDA, N.—SCHMELTER, T.—SCHWABE, R.: *A note on designs for estimating population parameters*. Listy Biometryczne – Biometrical Letters **42** (2005), 25–41.

[FH97] FEDOROV, V. V.—HACKL, P.: *Model-Oriented Design of Experiments*. Lecture Notes in Statist. 125, Springer-Verlag, New York, 1997.

[Sch06b] SCHMELTER, T.: *Considerations on group-wise identical designs for linear mixed models*. J. Statist. Plann. Inference **137** (2007), 4003–4010.

[Sch96] SCHWABE, R.: *Optimum Designs for Multi-Factor Models*. Lecture Notes in Statist. 113, Springer-Verlag, New York, 1996.

Received October 2, 2006

Rainer Schwabe
Institute for Mathematical Stochastics
Otto-von-Guericke-University, Magdeburg
D–39 016 Magdeburg
GERMANY
E-mail: rainer.schwabe@mathematik.uni-magdeburg.de

Thomas Schmelter
Bayer Schering Pharma AG
Global Biostatistics
D–13342 Berlin
GERMANY
E-mail: thomas.schmelter@bayerhealthcare.com