# ON OPTIMIZING MULTI-LEVEL DESIGNS: POWER UNDER BUDGET CONSTRAINTS

TODD C. HEADRICK[1*] AND BRUNO D. ZUMBO[2]

*Southern Illinois University-Carbondale  and  University of British Columbia*

## Summary

This paper derives a procedure for efficiently allocating the number of units in multi-level designs given prespecified power levels. The derivation of the procedure is based on a constrained optimization problem that maximizes a general form of a ratio of expected mean squares subject to a budget constraint. The procedure makes use of variance component estimates to optimize designs during the budget formulating stages. The method provides more general closed form solutions than other currently available formulae. As such, the proposed procedure allows for the determination of the optimal numbers of units for studies that involve more complex designs. A method is also described for optimizing designs when variance component estimates are not available. Case studies are provided to demonstrate the method.

*Key words:* budget constraint; effect size; Lagrange multiplier; level of randomization; multi-level design; optimization; power; variance components

## 1. Introduction

The efficient allocation of economic resources in experimental designs has long been a topic of discussion (e.g. Deming, 1953; Brooks, 1955; Overall & Dalal, 1965; Cochran, 1977; Hsieh, 1988; Donner, Brown & Brasher, 1990; Muller *et al.*, 1992; Marcoulides, 1993; Snijders & Bosker, 1993; Raudenbush, 1997; Moerbeek, van Breukelen & Berger, 2000, 2001; Headrick & Zumbo, 2001). For example, Brooks (1955) derived a procedure for determining the optimal subsampling number, subject to a budget constraint, for a design that involved two-stage sampling. This procedure was extended to three-stage sampling (Cochran, 1977 p. 285) and also used to determine optimal sampling numbers for some multi-level designs (e.g. Moerbeek *et al.*, 2000; Headrick & Zumbo, 2001).

Consider an $r$-level design that consists of $n_i$ units at the $i$th level where $i = 1, \ldots, r$. For any particular design, there may be a number of different null hypotheses formulated and tested. However, in the test of any specific hypothesis, the ratio of expected mean squares has the general form

$$R = \frac{\sigma_0^2 + m\sigma_{r+1}^2}{\sigma_0^2} \, . \tag{1}$$

The term $\sigma_0^2$ is defined as the total component of variance present in both the numerator and denominator of (1) and is expressed as a linear combination of at most $r$ variance components.

The component $\sigma_{r+1}^2$ is the additional term in the numerator that is associated with some treatment effect. The term $m$ is a general coefficient that denotes the total magnitude of all other coefficients associated with $\sigma_{r+1}^2$.

Most modern statistical software packages (e.g. MINITAB, 2000) estimate a non-centrality parameter $\delta$ using estimates of the variance components from a set of data. In the context of (1), the non-centrality parameter can be expressed as

$$\delta = \frac{km\sigma_{r+1}^2}{\sigma_0^2}\,, \tag{2}$$

where $k$ is the number of means (from $k$ populations) in the null hypothesis.

Consider, for example, (1) and (2) in terms of a two-level design that consists of taking two measurements at $n_1$ time periods on each of $n_2$ randomly selected subjects from $k$ different treatment populations. Given appropriate assumptions, (1) and (2) can be expressed by setting $r = 2$ as

$$R = \frac{\sigma_1^2 + 2n_1\sigma_2^2 + 2n_1 n_2\sigma_3^2}{\sigma_1^2 + 2n_1\sigma_2^2}\,, \tag{3}$$

$$\delta = \frac{k2n_1 n_2\sigma_3^2}{\sigma_1^2 + 2n_1\sigma_2^2}\,, \tag{4}$$

where $\sigma_0^2 = \sigma_1^2 + 2n_1\sigma_2^2$, $m = 2n_1 n_2$, $\sigma_3^2 = \sum \tau_j^2/k$, and where $\tau_j$ is the hypothesized deviation of the $j$th treatment mean. The denominator of (3) consists of $r = 2$ independent sources of random variation which have expected values of zero and variances $\sigma_1^2$ and $\sigma_2^2$. Thus, $R$ in (3) would represent the ratio of expected mean squares associated with the omnibus test of significant treatment differences.

When economic resources are scarce, what are the optimal sizes of $n_1$ and $n_2$ for the design described above? For general situations where power is a concern, this question can be formulated in terms of $r$-level designs as: what are the optimal integer values of $n_1, \ldots, n_r$ to use in an $r$-level design such that the selected units of $n_i$ yield a targeted level of power for a specific hypothesis test at minimum cost?

Overall & Dalal (1965) proposed a procedure for obtaining the optimal numbers of units ($n_i$) for some multi-level designs based on (1) such that power could be maximized for a fixed budget, but their procedure lacks generality and is laborious. Specifically, it entails writing out all possible experimental situations given fixed prices and a total budget. For each of the experimental possibilities, the subsequent tasks are: (a) refer to previous research to estimate the variance components, (b) determine the non-centrality parameter, and (c) refer to power tables to determine the allocation that maximizes power subject to the total budget. Further, Overall & Dalal (1965) made no attempt to target a particular level of power, so that the optimal solutions may yield an unacceptably low level of power because of an inadequate initial budget.

The methods of Brooks (1955), Cochran (1977 p. 285) and Moerbeek *et al.* (2000) are based on minimizing error variance, i.e. the denominator in (1), to obtain the optimal allocation. The problem with these methods is that effect sizes, degrees of freedom, and power are not generally considered. Specifically, while these models do indeed minimize error variance they may in general yield solutions with an unacceptably low level of power due to an insufficient budget. The procedure suggested by Moerbeek *et al.* (2000 p. 281) allows for an effect

size to be used to determine the minimum funds needed to achieve a certain level of power. However, their procedure is limited to two groups.

Moerbeek *et al.* (2000 Table 2) offer convenient closed-form solutions to obtain the optimal numbers of units in the context of two- and three-level designs. These equations are only applicable for some designs. For example, the equation for the solution to $n_1$ given in Moerbeek *et al.* (2000 Table 2, for two levels, and randomization at the class level) is not general enough to provide an optimum for $n_1$ in (3) above.

Because of the work of researchers such as Glass (1976), Hedges (1981) and Cohen (1988), some methodologists (e.g. Kirk, 1996; Rosenthal, Rosnow & Rubin, 2000) are more concerned with the study and reporting of effect sizes — such as those related to $\sigma_{r+1}^2$ in (1). Further, effect sizes are available to the investigator (or can be estimated) at the initial stages of formulating a budget and selecting an appropriate experimental design (see e.g. Howell, 2002 p. 228).

## 2. Purpose of the study

In view of the above, what is needed is a general procedure that estimates the amount of adjustment to a budget necessary to bring power to a targeted level. More specifically, the purposes of the study are to (a) derive a procedure using Lagrange multipliers to determine the optimal numbers of units for multi-level designs with desired power targets, (b) provide more general closed form formulae than previous researchers (Moerbeek *et al.*, 2000) to enable the determination of optimal numbers of units for studies that involve complex designs, and (c) provide a method that optimizes multi-level designs when the estimates of variance components are unavailable. Case studies demonstrate the proposed procedure.

## 3. Mathematical development

### 3.1. When estimates of the variance components are available

We assume a balanced $r$-level design with fixed treatment conditions and where random allocation is carried out at one level. It is also assumed that there is one hypothesis of interest, the test of significance associated with the variance component $\sigma_{r+1}^2$ in (1). As such, if randomization is performed at level $i$ then $\sigma_0^2$ consists of a linear combination of $i$ variance components.

Given these assumptions, let (1) be expressed as

$$R = \frac{\sigma_1^2 + \sum_{i=1}^{r} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2}{\sigma_1^2 + \sum_{i=1}^{r-1} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2}, \tag{5}$$

where $p_1, \dots, p_{r-1}$ are non-negative integers, and $p_r$ is a positive integer. Similar to the role of $m$ in (1), the $p_i$ are symbols denoting the total magnitude associated with their respective variance component $\sigma_{i+1}^2$. The terms $n_1, \dots, n_r$ are the variables of concern to be selected in such a manner that $R$ is maximized subject to a budget constraint.

It can be shown that (3) is a special case of (5) when randomization is performed at the second level. With $r = 2$, (5) would appear as

$$R = \frac{\sigma_1^2 + p_1 n_1 \sigma_2^2 + p_2 n_1 n_2 \sigma_3^2}{\sigma_1^2 + p_1 n_1 \sigma_2^2}. \tag{6}$$

TABLE 1

*Solutions for $n_i^*$ and $\lambda^*$ in the context of two-level designs*

| Level randomized | $n_i^*$ | $\lambda^*$ |
|---|---|---|
| 1 | $\dfrac{B - q_2 c_2 n_2}{q_1 c_1 n_2}$ | $\dfrac{p_2 \sigma_3^2}{(\sigma_1 \sqrt{q_1 c_1})^2}$ |
| 2 | $\dfrac{B \sigma_2 \sqrt{p_1}}{\sigma_1 \sqrt{q_1 q_2 c_1 c_2} + \sigma_2 q_2 c_2 \sqrt{p_1}}$ | $\dfrac{p_2 \sigma_3^2}{(\sigma_1 \sqrt{q_1 c_1} + \sigma_2 \sqrt{p_1 q_2 c_2})^2}$ |

Setting $p_1 = p_2 = 2$ gives (3). If randomization were performed at the first level then setting $p_1 = 0$ and $p_2 = 2$ in (6) would give the appropriate expression for $R$ as

$$R = \frac{\sigma_1^2 + 2 n_1 n_2 \sigma_3^2}{\sigma_1^2} \, .$$

We suppose the total finite budget $B$ associated with (5) is

$$B = \sum_{i=1}^{r} c_i q_i \prod_{j=i}^{r} n_j \, , \tag{7}$$

where $c_i$ represents the price per unit of $n_i$ and the $q_i$ are fixed positive integers such that the term $q_i \prod_{j=i}^{r} n_j$ gives the total number of units in the design for level $i$. Thus, the $i$th product term in (7) $c_i q_i \prod_{j=i}^{r} n_j$ represents the total cost of level $i$.

Equations (5) and (7) can be combined to give the Lagrangean

$$L(n_1, \ldots, n_r, \lambda) = f(n_1, \ldots, n_r) + \lambda \big( B - g(n_1, \ldots, n_r) \big) , \tag{8}$$

where $\lambda$ is the Lagrange multiplier and $R = f(n_1, \ldots, n_r)$ is the objective function from (5) that is maximized with respect to $n_1, \ldots, n_r$ subject to $B = g(n_1, \ldots, n_r)$ in (7) for fixed values of $\sigma_1^2$, $\sigma_{i+1}^2$, $c_i$, $p_i$ and $q_i$ for $i = 1, \ldots, r$.

If randomization is performed at level $r$ in the design, then the optimal solutions for $n_1^*, \ldots, n_r^*$ and $\lambda^*$ are expressed as follows (see the Appendix):

$$n_1^* = \frac{\sigma_1}{\sigma_2} \sqrt{\frac{q_2 c_2}{p_1 q_1 c_1}} \, , \tag{9}$$

$$n_i^* = \frac{\sigma_i}{\sigma_{i+1}} \sqrt{\frac{p_{i-1} q_{i+1} c_{i+1}}{p_i q_i c_i}} \quad \text{for } i = 2, \ldots, r-1 \, , \tag{10}$$

$$n_r^* = \frac{B \sigma_r \sqrt{p_{r-1}}}{\sigma_1 \sqrt{q_1 q_r c_1 c_r} + \sum_{i=2}^{r-1} \sigma_i \sqrt{p_{i-1} q_i q_r c_i c_r} + \sigma_r q_r c_r \sqrt{p_{r-1}}} \, , \tag{11}$$

$$\lambda^* = \frac{p_r \sigma_{r+1}^2}{\big( \sigma_1 \sqrt{q_1 c_1} + \sum_{i=2}^{r} \sigma_i \sqrt{p_{i-1} q_i c_i} \big)^2} \, , \tag{12}$$

where $\lambda^*$ represents the increase in $f(n_1^*, \ldots, n_r^*)$ given a one-unit increase in $B$.

TABLE 2

*Solutions for $n_i^*$ and $\lambda^*$ in the context of three-level designs*

| Level randomized | $n_i^*$ | $\lambda^*$ |
|---|---|---|
| 1 | $\dfrac{B - q_2 c_2 n_2 n_3 - q_3 c_3 n_3}{q_1 c_1 n_2 n_3}$ | $\dfrac{p_3 \sigma_4^2}{(\sigma_1 \sqrt{q_1 c_1})^2}$ |
| 2 | $\dfrac{\sigma_2 \sqrt{p_1}\,(B - q_3 c_3 n_3)}{\sigma_1 n_3 \sqrt{q_1 q_2 c_1 c_2} + \sigma_2 q_2 c_2 n_3 \sqrt{p_1}}$ | $\dfrac{p_3 \sigma_4^2}{(\sigma_1 \sqrt{q_1 c_1} + \sigma_2 \sqrt{p_1 q_2 c_2})^2}$ |
| 3 | $\dfrac{B \sigma_3 \sqrt{p_2}}{\sigma_1 \sqrt{q_1 q_3 c_1 c_3} + \sigma_2 \sqrt{p_1 q_2 q_3 c_2 c_3} + \sigma_3 q_3 c_3 \sqrt{p_2}}$ | $\dfrac{p_3 \sigma_4^2}{(\sigma_1 \sqrt{q_1 c_1} + \sigma_2 \sqrt{p_1 q_2 c_2} + \sigma_3 \sqrt{p_2 q_3 c_3})^2}$ |

Suppose random allocation is carried out at the $i$th level where $1 \le i \le r - 1$. In this situation, $n_i^*$ is obtained by substituting fixed integers for $n_{i+1}, \ldots, n_r$ and $i - 1$ equations from (9) and (10) into (7) and then solving for $n_i^*$. Tables 1 and 2 give the formulae for the optimal solutions of $n_i^*$ and $\lambda^*$ for two- and three-level designs. The extension of these results to larger designs is shown by the structure of the formulae.

The solutions of $n_i^*$ require that they be integers. However, (9), (10) and (11) in general do not yield such numbers. Therefore, the following rule for rounding (Cameron, 1951) is used with respect to (9) and (10): if $n_i$ are positive integers such that $n_i < n_i^* < n_i + 1$, round up if $n_i^{*2} > n_i(n_i + 1)$; otherwise round down. If randomization is performed at level $i$, then $n_i^*$ should be rounded in such a manner that the realized budget does not exceed the initial budget estimate.

### 3.2. Targeting levels of power

Assume that random allocation is carried out at the $i$th level in an $r$-level design. It is possible that $n_i^*$ may yield an undesirable level of power because of an inappropriate value of $B$ in (7). As a result, we define $\tilde{n}_i^*$ as the integer number of units yielding a targeted level of power $\pi_*$ that is at least as large as a prespecified power threshold point $\pi_0$. More specifically, $\pi_*$ is defined to be an element on the interval

$$0 < \pi_a < \pi_0 \le \pi_* < \pi_b < 1, \tag{13}$$

where $\pi_a$ and $\pi_b$ are associated with $\tilde{n}_i^* - 1$ and $\tilde{n}_i^* + 1$ units, respectively. Both $\pi_a$ and $\pi_b$ are considered undesirable levels of power because $\pi_a$ falls below $\pi_0$ and $\pi_b$ is at the point where the cost of an additional unit beyond $\tilde{n}_i^*$ exceeds the gain in power.

To determine the amount of change to the initial realized budget $\Delta B$ such that power is at $\pi_*$, an estimate of $R$ in (5) is first obtained. This estimate is computed as

$$\hat{R} = \frac{\hat{\sigma}_1^2 + \sum_{i=1}^{r} p_i \left( \prod_{j=1}^{i} n_j^* \right) \hat{\sigma}_{i+1}^2}{\hat{\sigma}_1^2 + \sum_{i=1}^{r-1} p_i \left( \prod_{j=1}^{i} n_j^* \right) \hat{\sigma}_{i+1}^2}. \tag{14}$$

Second, the point on a non-central F distribution that yields $\pi_0$ in (13) is obtained as

$$F_0 = 1 + \frac{\delta_0}{k}, \tag{15}$$

where the non-centrality parameter $\delta_0$ is a function of the required power $\pi_0$ and is determined from (A6) in the Appendix. Using (14) and (15), $\Delta B$ is then determined as

$$\Delta B = \frac{F_0 - \hat{R}}{\lambda^*} \,. \tag{16}$$

To obtain an integer solution for $\tilde{n}_i^*$ we relate the following inequality to (13) as

$$B_0 = B + \Delta B < B_* = B + (\tilde{n}_i^* - n_i^*)\frac{\partial B}{\partial n_i^*} < B_b = B + \big((\tilde{n}_i^* + 1) - n_i^*\big)\frac{\partial B}{\partial n_i^*} \,, \tag{17}$$

where $B_*$ is the budget selected and is associated with power $\pi_*$.

### 3.3. When estimates of the variance components are unavailable

In the absence of estimates of the variance components an approach the experimenter can take is to ask what would be a *minimum* effect size worth detecting. Using Cohen's (1988 p. 274) definition of a standardized effect size $f$ and a minimum value of $\sigma_{r+1}^2$, the component $\sigma_1^2$ in (5) can be determined as

$$\sigma_1^2 = \frac{\sigma_{r+1}^2}{f^2} = \frac{\sum \tau_j^2 / k}{f^2} \,, \tag{18}$$

where $\sum \tau_j^2 / k$ is defined as in (3) and (4).

A useful contribution from Brooks (1955 Table 1) is the pilot study that demonstrated the wide range of values that $n_1$ may take and still maintain at least 90% precision of the true optimum $n_1^*$ in (9). Commenting on Brooks (1955), Cochran (1977 p. 282) noted:

> Because of the flatness of the optimum, these [variance component] ratios need not be obtained with high accuracy ... the wide interval between the lower and upper limits [that maintain 90% precision] is striking in nearly all cases

Similar points were also made by Moerbeek *et al.* (2000 p. 278) on the flatness of the optimal solutions.

In view of this, it is convenient to estimate variance components in the manner suggested by Cochran (1977 p. 282) as

$$\frac{\sigma_i^2}{\sigma_{i+1}^2} = \frac{1 - \rho_{i+1}}{\rho_{i+1}} \quad \text{for } i = 1, \dots, r-1 \,, \tag{19}$$

where $\rho_{i+1}$ is the intra-class correlation between the elements of level $i+1$. Estimates of $\rho_{i+1}$ can be obtained from studies that report reliability estimates. For example, reliability coefficients such as KR-21 (Allen & Yen, 1979 p. 84) are often reported on instruments that take duplicate or repeated measures. Otherwise, if other intra-class correlations are unknown, the interval considered for estimating the variance components is $\rho_{i+1} \in [0.05, 0.50]$. Values of $\rho_{i+1}$ outside this interval are considered unusually low or high intra-class correlations (Cochran, 1977 p. 282). Values of $\rho_{i+1}$ are often related to the size of the unit being considered, e.g. larger units are usually associated with smaller values of $\rho_{i+1}$. Substituting the estimates for $\rho_{i+1}$ and $\sigma_i^2$ into (19) and solving for $\hat{\sigma}_{i+1}^2$ gives

$$\hat{\sigma}_{i+1}^2 = \frac{\hat{\sigma}_i^2 \hat{\rho}_{i+1}}{1 - \hat{\rho}_{i+1}} \quad \text{for } i = 1, \dots, r-1 \,. \tag{20}$$

TABLE 3

*Summary of a two-level design when the estimates of the variance components*
*are available from a dataset. Randomization is at the second (animal) level.*

| | |
|---|---|
| 1. | Cost for each repeated measure: $c_1 = \$5$. |
| 2. | Cost for each animal to enter the experiment: $c_2 = \$100$. |
| 3. | Initial budget estimate: $B = \$5250$. |
| 4. | Variance components: $\hat{\sigma}_1^2 = 0.01908$, $\hat{\sigma}_2^2 = 0.00698$, $\hat{\sigma}_3^2 = 0.00244$. |
| 5. | Values of $k$, $p_i$ and $q_i$: $k = 3$, $p_1 = 4$, $p_2 = 4$, $q_1 = 12$, $q_2 = 3$. |
| 6. | Optimal integer solutions: $n_1^* = 2$, $n_2^* = 12$. |
| 7. | Degrees of freedom: $\upsilon_1 = k - 1 = 2$, $\upsilon_2 = k(n_2^* - 1) = 33$. |
| 8. | Realized budget: $B = \$5040$. |
| 9. | Minimum power threshold point $\pi_0$ in (13): $\pi_0 = 0.80$. |
| 10. | Critical $Z$ from the $N(0, 1)$ distribution associated with $\pi_0$: $Z = -0.84$. |
| 11. | Critical value from the central $F_{\upsilon_1, \upsilon_2}$ distribution: $F_{0.05, 2, 33} = 3.28$. |
| 12. | Non-centrality parameter from (A6) and (A7): $\delta_0 = 10.41$. |
| 13. | Estimate of (14): $\hat{R} = 4.13$. |
| 14. | Point on the non-central $F_{\upsilon_1, \upsilon_2, \delta_0}$ distribution that yields $\pi_0$ in (15) $F_0 = 4.47$. |
| 15. | Lagrange multiplier (Table 1): $\lambda^* = 0.000622$. |
| 16. | Required change to the realized budget in (16): $\Delta B = (F_0 - \hat{R})/\lambda^* = \$547$. |
| 17. | Budget associated with $\pi_0$ in (17): $B_0 = B + \Delta B = \$5587$. |
| 18. | Value of $\partial B/\partial n_2^*$ in (17): $c_1 q_1 n_1^* + c_2 q_2 = \$420$. |
| 19. | Selected budget in (17) and associated with power of $\pi_*$ in (13): $B_* = B + (\tilde{n}_2^* - n_2^*)(\partial B/\partial n_2^*) = \$5040 + (14 - 12)(\$420) = \$5880$. |

## 4. Case studies

### 4.1. When estimates of the variance components are available

Consider a two-level design where a researcher is formulating a budget to study the effect of conditioned suppression on animal behaviour. Of interest to the researcher is the $F$ test for differences between groups. The experimenter has data from a repeated measures design with three factors as reported by Howell (2002 pp. 502–508) that was used to study the effects of conditioned suppression on three groups of rats.

The researcher wants to give $n_1$ repeated measures across four different cycles to each of $n_2$ animals in three groups. Suppose the experimenter estimates an initial budget of $B = \$5250$ and wants a power level of at least $\pi_0 = 0.80$. Table 3 gives a summary of the steps for determining the optimal solutions for this experiment. The variance component estimates were obtained from the dataset using MINITAB (2000). As indicated in Table 3, the optimal solutions are $n_1^* = 2$ repeated measures and $\tilde{n}_2^* = 14$ animals for each group. This requires the initial budget to be increased to $B_* = \$5880$ to achieve the targeted power level $\pi_*$ in (13).

### 4.2. When estimates of the variance components are unavailable

Consider a three-level design where subjects are randomly assigned to either an experimental or a control group. The design consists of taking $n_1$ duplicate measures for each of $n_2$ repeated measures on $n_3$ subjects from each group. Suppose it is considered meaningful to detect a difference of 20 units between a new treatment intervention and the standard treatment, with power of at least $\pi_0 = 0.90$. Using (18) and Cohen's (1988 p. 274) estimate of a medium effect size $f = 0.50$ gives $\hat{\sigma}_1^2 = 400 = (\frac{1}{2}(100 + 100))/0.25$.

Assume from prior research that the instrument used to take duplicate measures has good reliability, $\hat{\rho}_2 = 0.80$. The experimenter uses an estimate of $\hat{\rho}_3 = 0.25$ which is the

TABLE 4

*Summary of a three-level design when the estimates of the variance components are not available from a dataset. Randomization is at the third (subjects) level*

| | |
|---|---|
| 1. | Cost for duplicating each measure: $c_1 = \$10$. |
| 2. | Cost for each repeated measure: $c_2 = \$50$. |
| 3. | Cost for training each subject: $c_3 = \$100$. |
| 4. | Initial budget estimate: $B = \$12\,500$. |
| 5. | Variance components: $\hat{\sigma}_1^2 = 400$, $\hat{\sigma}_2^2 = 1600$, $\hat{\sigma}_3^2 = 533.33$, $\sigma_4^2 = 100$. |
| 6. | Values of $k$, $p_i$ and $q_i$: $k = 2$, $p_1 = p_2 = p_3 = 1$, $q_1 = q_2 = q_3 = 2$. |
| 7. | Optimal integer solutions: $n_1^* = 1$, $n_2^* = 3$, $n_3^* = 22$. |
| 8. | Degrees of freedom: $\upsilon_1 = k - 1 = 1$, $\upsilon_2 = k(n_3^* - 1) = 42$. |
| 9. | Realized budget: $B = \$12\,320$. |
| 10. | Minimum power threshold point $\pi_0$ in (13): $\pi_0 = 0.90$. |
| 11. | Critical $Z$ from the $N(0, 1)$ distribution associated with $\pi_0$: $Z = -1.28$. |
| 12. | Critical value from the central $F_{\upsilon_1, \upsilon_2}$ distribution: $F_{0.05,1,42} = 4.07$. |
| 13. | Non-centrality parameter from (A6): $\delta_0 = 10.78$. |
| 14. | Estimate of (14): $\hat{R} = 2.83$. |
| 15. | Point on the non-central $F_{\upsilon_1, \upsilon_2, \delta_0}$ distribution that yields $\pi_0$ in (15): $F_0 = 6.39$. |
| 16. | Lagrange multiplier (Table 2): $\lambda^* = 0.000150$. |
| 17. | Required change to the realized budget in (16): $\Delta B = (F_0 - \hat{R})/\lambda^* = \$23\,733$. |
| 18. | Budget associated with $\pi_0$ in (17): $B_0 = B + \Delta B = \$36\,053$. |
| 19. | Value of $\partial B/\partial n_3^*$ in (17): $c_1 q_1 n_1^* n_2^* + c_2 q_2 n_2^* + c_3 q_3 = \$560$. |
| 20. | Selected budget in (17) and associated with power of $\pi_*$ in (13): $B_* = B + (\tilde{n}_3^* - n_3^*)(\partial B/\partial n_3^*) = \$12\,320 + (65 - 22)(\$560) = \$36\,400$. |

approximate midpoint of the intra-class correlation interval considered above. It follows from (20) that the variance component estimates are $\hat{\sigma}_2^2 = 1600$ and $\hat{\sigma}_3^2 = 533.33$.

Suppose the researcher estimates an initial budget of $B = \$12\,500$. Table 4 summarizes the information for this experiment and the steps to follow for determining the optimal number of units for this design. As Table 4 shows, the required budget is $B_* = \$36\,400$ for $n_1^* = 1$ measure for each of $n_2^* = 3$ repeated measures on each of $\tilde{n}_3^* = 65$ subjects in each group. Thus, the initial budget estimate was approximately one-third of the budget necessary to achieve the targeted level of power $\pi_*$.

In practice, when the variance components are *a priori* unknown, it is prudent to take an approach that provides a conservative estimate of $\hat{R}$ in (14). The consequence of this approach is that a larger budget is needed to ensure a minimum targeted level of power is achieved.

## 5. Conclusion

The proposed procedure determines the efficient allocation of resources in the context of multi-level designs. Specifically, given a fixed budget, various prices, and estimates of the variance components for a design, the Lagrange multiplier method locates the point where the ratio of expected mean squares is at a maximum.

The procedure presented simplifies the Overall & Dalal (1965) procedure for determining optimal solutions to the extent that there is no need to list all possible combinations of potential solutions. The example and discussion in Overall & Dalal (1965) on the special 'zero-overhead' case for a simple repeated measures design is also subsumed under the proposed method. That is, if there is no cost associated with each subject (i.e. $c_2$ is arbitrarily close to zero in Table 1), then each subject is tested once and the design uses as many subjects as the budget allows.

The case studies examined show that the procedure enables the researcher to estimate the amount of funds needed to adjust power to a desired level. Using effect sizes and Lagrange multipliers the method provides the researcher with the appropriate information to formulate an accurate budget.

## Appendix

### A.1. Derivation of the optimal solutions

Substituting (5) and (7) into (8) yields

$$L = \frac{\sigma_1^2 + \sum_{i=1}^{r} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2}{\sigma_1^2 + \sum_{i=1}^{r-1} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2} + \lambda \left( B - \sum_{i=1}^{r} c_i q_i \prod_{j=i}^{r} n_j \right). \tag{A1}$$

The partial derivatives of (A1) with respect to $\lambda$ and $n_i$ are expressed as

$$\frac{\partial L}{\partial \lambda} = 0 = B - \sum_{i=1}^{r} c_i q_i \prod_{j=i}^{r} n_j, \tag{A2}$$

$$\begin{aligned}
\frac{\partial L}{\partial n_i} = f_{n_i} - \lambda g_{n_i} = 0 = {} & \frac{\sum_{i=1}^{r} \partial_{n_i} \left( p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2 \right)}{\sigma_1^2 + \sum_{i=1}^{r-1} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2} \\
& - \frac{\left( \sum_{i=1}^{r-1} \partial_{n_i} \left( p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2 \right) \right) \left( \sigma_1^2 + \sum_{i=1}^{r} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2 \right)}{\left( \sigma_1^2 + \sum_{i=1}^{r-1} p_i \left( \prod_{j=1}^{i} n_j \right) \sigma_{i+1}^2 \right)^2} \\
& - \lambda \sum_{i=1}^{r} \partial_{n_i} \left( c_i q_i \prod_{j=i}^{r} n_j \right),
\end{aligned} \tag{A3}$$

for all $i = 1, \ldots, r$. Solving the $r$ equations in (A3) for $\lambda$ gives

$$\lambda = \frac{f_{n_1}}{g_{n_1}} = \cdots = \frac{f_{n_r}}{g_{n_r}}. \tag{A4}$$

Sequentially solving $f_{n_i}/g_{n_i} - f_{n_{i+1}}/g_{n_{i+1}} = 0$ for $n_i$, where $i = 1, \ldots, r-1$, gives the closed-form formulae for $n_i^*$ in (9) and (10). The $r-1$ expressions on the right-hand sides of (9) and (10) are subsequently substituted into (A2) and (A4) to obtain the equations for $n_r^*$ and $\lambda^*$ in (11) and (12).

It is only necessary to satisfy the first-order conditions because it can be shown that the objective function $R$ in (A1) is explicitly quasiconcave. This implies that evaluating $R$ at $n_i^*$ for $i = 1, \ldots, r$ is an absolute constrained maximum. To show that $R$ is explicitly quasiconcave we must have (Chiang, 1984 p. 398)

$$R(v) > R(u) \Rightarrow R\big(\theta u + (1 - \theta)v\big) > R(u), \tag{A5}$$

where $0 < \theta < 1$, and where $v = (n_1^*, \ldots, n_r^*)$ and $u = (n_1^*, \ldots, n_r^*)$ are distinct arbitrary vectors from the convex domain of $R$. By inspection of (5), the conditions in (A5) must hold because $\sigma_1^2, \ldots, \sigma_{r+1}^2$, $n_1^*, \ldots, n_r^*$ and $p_r$ are all positive, $p_1, \ldots, p_{r-1}$ are non-negative, and the constraint set in (A2) is convex because it is a linear equation.

## A.2. Calculation of the non-centrality parameter

The value of $F_0$ in (15) is based on the following expression (Winer, Brown & Michels, 1991 p. 136):

$$z_0 = \frac{\sqrt{(2\upsilon_2 - 1)\dfrac{\upsilon_1 F_{\alpha,\upsilon_1,\upsilon_2}}{\upsilon_2}} - \sqrt{2(\upsilon_1 + \delta) - \dfrac{\upsilon_1 + 2\delta}{\upsilon_1 + \delta}}}{\sqrt{\dfrac{\upsilon_1 F_{\alpha,\upsilon_1,\upsilon_2}}{\upsilon_2} + \dfrac{\upsilon_1 + 2\delta}{\upsilon_1 + \delta}}}, \tag{A6}$$

where $z_0$ is a critical value from the $\mathrm{N}(0, 1)$ distribution, $\delta$ is the non-centrality parameter, and $F_{\alpha,\upsilon_1,\upsilon_2}$ denotes the $(1-\alpha)$-quantile of the central F distribution with $\upsilon_1$ and $\upsilon_2$ degrees of freedom.

If $F$ is distributed as the $\mathrm{F}_{\upsilon_1,\upsilon_2,\delta}$ distribution, then $\Pr(F > F_{\alpha,\upsilon_1,\upsilon_2}) \approx \Pr(z > z_0)$. Setting $z_0$ in (A6) to the value associated with $\pi_0$ in (13) and solving for $\delta$ gives $\delta_0$. Because $\delta_0$ is of the form in (2) it follows that $F_0 = 1 + \delta_0/k$ which appears in (15).

For example, consider the first case study with degrees of freedom $\upsilon_1 = 2$, $\upsilon_2 = 33$ and a critical point $F_{0.05,2,33} = 3.28$. Because $\pi_0 = 0.80$ in (13) we set $z_0 = -0.84$. Substituting these values into (A6) gives

$$-0.84 = \frac{\sqrt{(2(33) - 1)\dfrac{(2)(3.28)}{33}} - \sqrt{2(2 + \delta) - \dfrac{2 + 2\delta}{2 + \delta}}}{\sqrt{\dfrac{(2)(3.28)}{33} + \dfrac{2 + 2\delta}{2 + \delta}}}. \tag{A7}$$

Solving (A7) for $\delta$, using the equation solver FindRoot (Wolfram, 1999), gives $\delta_0 = 10.41$. From (15) and with $k = 3$ we have $F_0 = 1 + 10.41/3 = 4.47$.

### References

ALLEN, M.J. & YEN, W.M. (1979). *Introduction to Measurement Theory*. Monterey: Brooks-Cole.

BROOKS, S. (1955). The estimation of an optimum subsampling number. *J. Amer. Statist. Assoc.* **50**, 398–415.

CAMERON, J.M. (1951). The use of variance components in preparing schedules for the sampling of baled wool. *Biometrics* **7**, 83–96.

CHIANG, A.C. (1984). *Fundamental Methods of Mathematical Economics*, 3rd edn. New York: McGraw-Hill.

COCHRAN, W.J. (1977). *Sampling Techniques*, 3rd edn. New York: Wiley.

COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. New York: Academic Press.

DEMING, W.E. (1953). On a probability mechanism to attain an economic balance between the resultant error of non-response and the bias of non-response. *J. Amer. Statist. Assoc.* **51**, 24–53.

DONNER, A., BROWN, K.S. & BRASHER, P. (1990). A methodological review of non-therapeutic intervention trials employing cluster randomization 1979–1989. *Internat. J. Epidemiol.* **19**, 795–800.

GLASS, G.V. (1976). Primary, secondary, and meta-analysis of research. *Educ. Res.* **5**, 3–8.

HEADRICK, T.C. & ZUMBO, B.D. (2001). Optimizing experimental designs relative to variance components and price inputs. *Proceedings of the Biometrics Section*. American Statistical Association: Alexandria, VA.

HEDGES, L.V. (1981). Distributional theory for Glass's estimator of effect size and related estimators. *J. Educ. Statist.* **6**, 107–128.

HOWELL, D. (2002). *Statistical Methods for Psychology*, 5th edn. Belmont, CA: Duxbury-Wadsworth.

HSIEH, F.Y. (1988). Sample size formulae for intervention studies with the cluster as unit of randomization. *Statist. Med.* **8**, 1195–1201.

KIRK, R.E. (1996). Practical significance: a concept whose time has come. *Educ. Psychol. Measur.* **56**, 746–759.

MARCOULIDES, G.A. (1993). Maximizing power in generalizability studies under budget constraints. *J. Educ. Statist.* **18**, 197–206.

MINITAB (2000). *MINITAB for Windows, release* 13.3, MINITAB, Inc., State College, PA, USA.

MOERBEEK, M., VAN BREUKELEN, G.J.P. & BERGER, M.P.F. (2000). Design issues for experiments in multilevel populations. *J. Educ. Behav. Statist.* **25**, 271–284.

MOERBEEK, M., VAN BREUKELEN, G.J.P. & BERGER, M.P.F. (2001). Optimal experimental designs for multilevel logistic models. *J. Roy. Statist. Soc. Ser. D* **50**, 17–30.

MULLER, K., LAVANGE, L., RAMEY, S. & RAMEY, C.T. (1992). Power calculations for general linear models including repeated measures applications. *J. Amer. Statist. Assoc.* **87**, 1209–1226.

OVERALL, J.E. & DALAL, S.N. (1965). Design of experiments to maximize power relative to cost. *Psychol. Bull.* **64**, 339–350.

RAUDENBUSH, S.W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychol. Meth.* **2**, 173–185.

ROSENTHAL, R., ROSNOW, R.L. & RUBIN D.B. (2000). *Contrasts and Effect Sizes in Behavioral Research.* Cambridge, UK: Cambridge University Press.

SNIJDERS, T.A.B. & BOSKER, R.J. (1993). Standard errors and sample sizes for two-level research, *J. Educ. Statist.* **18**, 237–259.

WINER, B.J., BROWN, D.R. & MICHELS, K.M. (1991). *Statistical Principles in Experimental Designs*, 3rd edn. New York: McGraw-Hill.

WOLFRAM, S. (1999). *The Mathematica Book*, 4th edn. Cambridge, UK: Wolfram Media-Cambridge University Press.