Statistica Sinica **12**(2002), 1085-1092

2^n BEHAVIORAL EXPERIMENTS USING PARETO OPTIMAL CHOICE SETS

Damaraju Raghavarao and Daozhi Zhang

Temple University and Purdue Pharma L.P.

Abstract: In behavioral experiments the respondents make their choices based on several factors called attributes. When the attributes are related to benefits and costs and respondents choose one attribute at a time, the highest level for benefit attributes and the lowest level for cost attributes will be selected. The relative importance given by the respondents to the attributes is not determined. If a set of profiles (factorial treatment combinations) of the attributes is given to the respondents to choose, it is possible to determine the relative importance of the attributes under some conditions. If the choice set has a dominating profile (or dominated profile) almost surely that profile will be selected (or not selected). Thus the set of profiles given to the respondent known as the choice set should not have dominating or dominated profiles. Choice sets with no dominating or dominated profiles are called Pareto optimal subsets. In this paper, we consider the connectedness and optimality of designs with Pareto optimal choice sets for n attributes each at 2 levels. Balanced incomplete block designs are helpful to reduce the choice set sizes without sacrificing optimality and this paper shows some interesting application of them.

Key words and phrases: Balanced incomplete block design, choice set, connected main effect plan, information per profile, Pareto optimal design.

1. Introduction

Consider a behavioral experiment in which an investigator wishes to study the relative importance new graduates put on factors (called attributes) like salary, vacation and pension plan in selecting a first job. To fix the ideas, consider two salary levels, say \$30k and \$35k per annum; vacation levels of two weeks for the first five years and three weeks thereafter, or three weeks from the beginning; two levels of pension plans, one starting immediately and the other starting after five years. If a respondent is asked to choose the levels of each factor, naturally the best level for each factor will be selected. In reality, some trade-offs are needed and cannot be determined by asking the respondent about each factor. For this purpose, we create job profiles of the characteristics of the jobs and ask the respondent to make a choice. For this, we need a choice set of profiles of the jobs. Profile are denoted by abc where a, b and c are 0 or 1 according as the salary, the vacation time and the pension plan are at their low or high levels, respectively. If we present a choice set with three profiles, $\{111, 011, 110\}$, the first dominates the other two and the respondent's choice is trivially made. However, if the choice set has profiles $\{101, 011, 110\}$, different respondents will select different profiles and one gains an insight into the importance of salary, vacation levels, and pension plans. The second choice set has no dominating or dominated profile. Different types of responses can be obtained from respondents on the choice set provided to them, such as choose the best profile, rank the profiles, rate the best profile on a point-scale, etc. In the analysis we make the assumption that the selection is solely due to the characteristics of the levels of the factors in the selected profile, and the role of the choice set is to create a decision making environment. We avoid giving large choice sets or sets of size one to create a genuine decision making environment.

We now formulate the problem and discuss the main results. Consider a 2^n experiment with $n (n \geq 2)$ attributes $A_1, \ldots, A_n$, each at two levels denoted by 0 and 1. A combination of the levels of the n attributes will be called a profile and denoted by $(u_1, \ldots, u_n)$. Let S be the set of all the profiles. A Pareto Optimal (PO) subset of S is defined as follows.

Definition 1.1. A subset T of S is said to be a PO subset if for every two distinct profiles $(u_1, \ldots, u_n)$, $(w_1, \ldots, w_n) \in T$, there exist subscripts i and j $(i \neq j)$ such that $u_i < w_i$ and $u_j > w_j$.

By giving a PO choice set, one can find respondent trade-offs between attributes.

Let $y_{u_1,\ldots,u_n}$ be the response to the profile $(u_1, \ldots, u_n)$, the average score of the profile or the logit of the proportion of times the profile is selected, and assume the main effects model

$$y_{u_1,\ldots,u_n} = \mu + \sum_{i=1}^{n} \alpha^i_{u_i} + e_{u_1,\ldots,u_n}, \tag{1.1}$$

where μ is the general mean, $\alpha^i_{u_i}$ is the effect of the u_i level of the *ith* attribute A_i. Different assumptions can be made regarding the distribution of random error terms $e_{u_1,\ldots,u_n}$. However, for the purpose of developing optimal designs, we assume the $e_{u_1,\ldots,u_n}$ are independently and identically distributed with a mean 0 and constant variance σ^2. The available profiles in the choice set may have competing effect (or cross effect) on the selected profile, but to keep the number of parameters under control we ignore such cross effects. This is a limitation of our study.

A design is said to be a connected main effects plan, if we can estimate all the main effect contrasts from the model (1.1).

Wiley (1977) recognized the need of PO choice sets in choice experiments. Krieger and Green (1991) extended that work and constructed orthogonal and PO subsets (or close to PO subsets). Huber and Hansen (1986) gave some empirical results on comparison of PO designs and orthogonal designs, and reported that the PO designs predict better. Huber and Zwerina (1996) proposed the concept of utility balanced designs, which are similar to PO designs. They also gave methods for constructing them.

Raghavarao and Wiley (1998) considered a general setting with any number of levels for the attributes and obtained connected main effects plans. Their results can be strengthened and optimal designs can be obtained when each attribute has 2 levels. According to them, in case of a 2^n experiment, $S_l = \{(u_1, \ldots, u_n) | \sum_{i=1}^{n} u_i = l\}$ is a PO subset, and no single PO subset S_l is a connected main effects plan. They further showed that the design based on $S_{[\frac{n}{2}]}$ and $S_{[\frac{n}{2}]+1}$ is a connected main effects plan, where $[l]$ is the integral part of l.

In this paper, we consider the PO subsets S_l and show that any two of them, S_l and S_k $(0 < l < k < n,\ l \neq k)$, is a connected main effects plan (Section 2). We examine the optimality of such designs generally (Section 3) and specifically: S_l and S_{l+1} in Section 4; S_l and S_{n-1} in Section 5. It may be noted that these choice sets may have many profiles and may be difficult to administer. In such cases, we can divide S_l and S_k into choice sets, each with a reasonable number of profiles (two or more). Different partitionings of PO choice sets may affect the selected profile by the respondent; but, because of our earlier discussion, this may not be of great concern. Generally, we ask that each choice set be PO, and that two PO subsets S_l and S_k provide connected main effects plans. A substantial reduction in choice set sizes, without sacrificing optimality, can be achieved through balanced incomplete block designs. These results are discussed in Section 6.

2. Connected Main Effects PO Designs

Consider two PO subsets, S_l and S_k. Let s denote the total number of profiles in S_l and S_k, and number them from 1 to s. An even number of respondents is selected and randomly divided into two equal groups. Responses are recorded for each profile from each of the two choice sets. We rewrite the model (1.1) as

$$y_i = \mu + \sum_{j=1}^{n} x_{ij}\beta_j + e_i, \tag{2.1}$$

where y_i is the response to the *ith* profile, $i = 1, \ldots, s$, μ is the general mean, β_j is the main effect of the *jth* attribute, x_{ij} is the level of the *jth* attribute in the *ith* profile, and the e_i's are independently and identically distributed with mean zero and variance σ^2.

Denote the design matrix by $\boldsymbol{D} = [\mathbf{1_s}\ \boldsymbol{X}]$, where $\boldsymbol{X}$ is the $s \times n$ matrix whose $(i,j)th$ element is x_{ij} and $\mathbf{1_s}$ is a s-dimensional vector of ones. After straightforward algebra, the information matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \boldsymbol{D}'\boldsymbol{D} = \begin{bmatrix} a_0 & a_1\mathbf{1_n} \\ a_1\mathbf{1_n} & (a_0 - a_2)\boldsymbol{I_n} + a_2\boldsymbol{J_n} \end{bmatrix},$$

where $\boldsymbol{I_n}$ is the $n \times n$ identity matrix, $\boldsymbol{J_n}$ is the $n \times n$ matrix of all ones, $a_0 = s$, $a_1 = \sum_{i=1}^{s} x_{i\alpha}$, and $a_2 = \sum_{i=1}^{s} x_{i\alpha}x_{i\beta}$ for any $\alpha,\ \beta = 1,\ldots,n$, $\alpha \neq \beta$.

The information matrix of main effects after eliminating μ is

$$\boldsymbol{C_m} = (a_0 - a_2)\boldsymbol{I_n} + \Big(a_2 - \frac{a_1^2}{a_0}\Big)\boldsymbol{J_n}. \tag{2.2}$$

Theorem 2.1. *For a 2^n experiment, the design based on two PO subsets S_l and S_k is a connected main effects plan, where $l \neq k$, $0 < l,\ k < n$.*

Proof. Without loss of generality, take $l < k$. Clearly, $a_0 = \binom{n}{l} + \binom{n}{k}$,

$$a_1 = \frac{2l-n}{n}\binom{n}{l} + \frac{2k-n}{n}\binom{n}{k}, \tag{2.3}$$

$$a_2 = \frac{(n-2l)^2 - n}{n(n-1)}\binom{n}{l} + \frac{(n-2k)^2 - n}{n(n-1)}\binom{n}{k}. \tag{2.4}$$

Then $a_0 - a_2 = \frac{4l(n-l)}{n(n-1)}\binom{n}{l} + \frac{4k(n-k)}{n(n-1)}\binom{n}{k} > 0$, when $l \neq 0$ or $k \neq n$. Moreover $a_0[(a_0-a_2)+n(a_2-\frac{a_1^2}{a_0})] = \frac{4(k-l)^2}{n}\binom{n}{l}\binom{n}{k} > 0$ when $k \neq l$. When $l = 0$ and $k = n$, the eigenvalues of $\boldsymbol{C_m}$ are zero and $\boldsymbol{C_m}$ is singular. When $l \neq 0$ or $k \neq n$, the eigenvalues of $\boldsymbol{C_m}$ are non-zero and $\boldsymbol{C_m}$ is non-singular. Thus any two PO sets S_l and S_k (except $l = 0$, $k = n$) is a connected main effects plan.

3. Orthogonal and Optimal PO Designs

To compare designs with different number of profiles, we use the Information Per Profile (IPP) in the design as an optimality criterion. The Information Per Profile (θ) is

$$\theta = \frac{n}{a_0\ \mathrm{trace}(\boldsymbol{C_m^{-1}})}, \tag{3.1}$$

where $\boldsymbol{C_m^{-1}}$ is

$$\boldsymbol{C_m^{-1}} = \frac{1}{a_0 - a_2}\Big[\boldsymbol{I_n} - \frac{a_0a_2 - a_1^2}{a_0(a_0 - a_2) + n(a_0a_2 - a_1^2)}\boldsymbol{J_n}\Big] \tag{3.2}$$

for any connected main effects plan composed of choice sets S_l and S_k.

The next theorem whose proof is long and involved (see Zhang (2001) for details) provides optimal designs for 2^n when $n = m^2$ or $n = m(m+2)$. $\boldsymbol{A}$ sketch of it is given in the Appendix.

Theorem 3.1. (1) $\theta = 1$ *if and only if* $\boldsymbol{A}$ *is a diagonal matrix.* (2) *For* $n = m^2$, $\theta = 1$ *for the choice sets* $S_{\frac{n-m}{2}}$ *and* $S_{\frac{n+m}{2}}$; *for* $n = m(m+2)$, $\theta = 1$ *for the choice sets* $S_{\frac{n-m}{2}}$ *and* $S_{\frac{n+m+2}{2}}$ (*or* $S_{\frac{n-m-2}{2}}$ *and* $S_{\frac{n+m}{2}}$). (3) *In all other cases,* $\theta < 1$.

From Theorem 3.1, it follows that the optimal choice sets for a 2^4 experiment are $S_1 = \{0001, 0010, 0100, 1000\}$, and $S_3 = \{1110, 1101, 1011, 0111\}$.

From (2.2) and the proof of Theorem 3.1, we want a_2 to be close to a_1^2/a_0 to have orthogonality, and a_1^2 to be small to have small variances.

When $n \neq m^2$ or $m(m+2)$, we find the best two choice sets among S_l and S_{l+1} (Section 4) and among S_l and S_{n-l} (Section 5). Because S_l, S_{l+1} and S_{l+2} provide a connected design to estimate main effects and 2 factor interactions (Zhang (2001)), one can augment the choice sets S_l and S_{l+1}, considered in Section 4, by S_{l+2} to estimate two factor interactions, if needed. While reducing the choice set sizes, we are going to use balanced incomplete block designs and their complements. With this in mind we develop best choice sets of the form S_l and S_{n-l} in Section 5.

4. Best PO Sets S_l and S_{l+1}

Theorem 4.1. *The design based on* $S_{[\frac{n}{2}]}$ *and* $S_{[\frac{n}{2}]+1}$ *has maximum IPP among designs based on* S_l *and* S_{l+1}.

Proof. Taking $k = l + 1$ in the expressions of a_0, a_1, and a_2, straightforward algebra yields

$$\theta(l) = \frac{n}{a_0 \ \mathrm{Trace}(\boldsymbol{C_m^{-1}})} = \frac{2(l+1)(n-l)}{n(n+1)}. \tag{4.1}$$

Now $\frac{\theta(l+1)}{\theta(l)} = \frac{(l+2)(n-l-1)}{(l+1)(n-l)}$ is larger, equal or less than 1, according as l is less, equal or larger than $\frac{n}{2} - 1$. Thus $l = [\frac{n}{2}]$ maximizes $\theta(l)$ and completes the proof of the Theorem.

The choice sets given in Theorem 4.1 were also given by Raghavarao and Wiley (1998).

5. Best PO Sets S_l and S_{n-l}

Consider the connected main effects plans based on two PO subsets S_l and S_{n-l} for $l = 1, \ldots, n-1$. Without loss of generality, take $l \leq [\frac{n}{2}]$.

Taking $k = n - l$ in the expressions of a_0, a_1, and a_2, straightforward algebra shows that

$$\theta = \frac{n}{a_0 \ \mathrm{Trace}(\boldsymbol{C_m^{-1}})} = \frac{(1-z)(1+(n-1)z)}{1+(n-2)z}, \tag{5.1}$$

where $z = \frac{a_2}{a_0} = \frac{(n-2l)^2-n}{n(n-1)}$, as a function of l, decreases with respect to l.

Lemma 5.1. *θ, as a function of z, decreases on $z \geq 0$ and increases on $z \leq 0$.*

Theorem 5.1. *The design based on $S_{[\frac{n-\sqrt{n}}{2}]}$ and $S_{n-[\frac{n-\sqrt{n}}{2}]}$ has maximum IPP among designs based on S_l and S_{n-l}.*

Proof. Note that $z \geq 0$ if and only if $l \leq \frac{n-\sqrt{n}}{2}$. From Lemma 5.1, we see that θ, as a function of l, decreases if $l \geq \frac{n-\sqrt{n}}{2}$ and increases if $l \leq \frac{n-\sqrt{n}}{2}$. This completes the proof.

The optimal designs given in Theorem 5.1 are those given in Theorem 3.1 when $n = m^2$.

6. Some Reduced Size Optimal PO Designs

As indicated in Section 1, the PO design based on subsets S_l and S_k may have too many profiles. For example, when $n = 9$ the optimal design based on S_3 and S_6 has 84 profiles in each choice set. This is impractical.

We provide designs using choice sets with fewer profiles through Balanced Incomplete Block (BIB) designs that have the same Information Per Profile (IPP). For the definition of BIB designs and their complements, see Raghavarao (1971).

Consider a BIB design with parameters $v = n,\ b,\ k,\ r,\ \lambda$ and the complementary $v^* = n$, $b^* = b$, $k^* = n-k$, $r^* = b-r$ and $\lambda^* = b-2r+\lambda$. From the first design we form a choice set S_k^* of b profiles, where the *ith* profile corresponds to the *ith* block, with the symbol present interpreted as the high level, and absent as the low level of that attribute. We can similarly form S_{n-k}^*.

Theorem 6.1. *The IPP θ for the design based on S_k and S_{n-k} is the IPP θ^* for the design based on S_k^* and S_{n-k}^*.*

Proof. For the design based on S_k and S_{n-k}, taking $l = n-k$ in the expressions of a_0, a_1, and a_2 and simplifying, we get

$$\theta = \frac{n}{a_0\ \mathrm{trace}(\boldsymbol{C_m^{-1}})} = \frac{4k(n-k)(n-2k)^2}{n[(n-2k)^2(n-2)+n]}. \tag{6.1}$$

For the design based on S_k^* and S_{n-k}^*, the information matrix, $\boldsymbol{A}^*$, can be written as

$$\boldsymbol{A}^* = \begin{bmatrix} a_0^* & a_1^*\boldsymbol{1_n} \\ a_1^*\boldsymbol{1_n} & (a_0^*-a_2^*)\boldsymbol{I_n} + a_2^*\boldsymbol{J_n} \end{bmatrix},$$

where $a_0^* = 2b$, $a_l^* = 0$, and $a_2^* = 2(b-4r+4\lambda)$.
The IPP θ^* of the design based on S_k^* and S_{n-k}^* is

$$\theta^* = \frac{a_0^*-a_2^*}{a_0^*} \times \frac{a_0^*(a_0^*-a_2^*)+n(a_0^*a_2^*-a_1^{*2})}{a_0^*(a_0^*-a_2^*)+(n-1)(a_0^*a_2^*-a_1^{*2})} = \frac{4k(n-k)(n-2k)^2}{n[(n-2k)^2(n-2)+n]}. \tag{6.2}$$

Thus $\theta = \theta^*$, completing the proof of the theorem.

The following corollary is an immediate consequence.

Corollary 6.1.1. (1) *The design generated from the BIB design with parameters* $n = 4t+1$, $b = 8t+2$, $k = 2t$, $r = 4t$, $\lambda = 2t-1$ *and its complement has the same IPP as the design based on* $S_{[\frac{n}{2}]}$ *and* $S_{[\frac{n}{2}]+1}$ *provided* n *is a prime or a prime power.* (2) *The design generated from the BIB design with the parameters* $n = 4t+3$, $b = 4t+3$, $k = 2t+1$, $r = 2t+1$, $\lambda = t$ *and its complement has the same IPP as the design based on* $S_{[\frac{n}{2}]}$ *and* $S_{[\frac{n}{2}]+1}$, *for infinitely many* n. (3) *The design generated from the BIB design with the parameters* $n = m^2$, $b = m(m+1)$, $k = \frac{m(m-1)}{2}$, $r = \frac{m^2-1}{2}$, $\lambda = \frac{(m+1)(m-2)}{4}$ *and its complement has the same IPP, namely* 1, *as the design based on* $S_{\frac{n-m}{2}}$ *and* $S_{\frac{n+m}{2}}$, *provided* $m = 4t+3$ *is a prime or a prime power.* (4) *The design generated from the BIB design with the parameters* $n = m^2$, $b = 2m(m+1)$, $k = \frac{m(m-1)}{2}$, $r = m^2-1$, $\lambda = \frac{(m+1)(m-2)}{2}$ *and its complement has the same IPP, namely* 1, *as the design based on* $S_{\frac{n-m}{2}}$ *and* $S_{\frac{n+m}{2}}$, *provided* $m = 4t+1$ *is a prime or a prime power.*

Corollary 6.1.1 indicates that for $n = 9$, we can construct a connected main effects plan based on the choice set S_3^* and S_6^* with 12 profiles in each choice set.

Remark. It may be noted that the two choice sets S_l and S_k together do not form a PO subset. However, each respondent receives a PO subset S_l or S_k and in that sense the choice sets are PO subsets.

Acknowledgements

The authors are very thankful to the two referees and an associate editor for their helpful comments in improving the presentation of this paper.

Appendix

Proof of Theorem 3.1. Clearly, $\theta = 1$ if A is a diagonal matrix. We prove the necessity of (1). If λ_i, $i = 1, \ldots, n$, are the eigenvalues of the matrix $\boldsymbol{C_m}$, Trace$(\boldsymbol{C_m}) = \sum_{i=1}^n \lambda_i = n(a_0 - \frac{a_1^2}{a_0})$. Since trace$(\boldsymbol{C_m^{-1}}) = \sum \lambda_i^{-1}$and the arithmetic mean is at least the harmonic mean of $\frac{\lambda_i}{a_0}$ for $i = 1, \ldots, n$, we have

$$\theta = \Big[\frac{1}{n}\sum \frac{1}{(\lambda_i/a_0)}\Big]^{-1} \le \frac{1}{n}\sum_{i=1}^{n}\frac{\lambda_i}{a_0} = 1 - \frac{a_1^2}{a_0^2} \le 1. \tag{A.1}$$

From (A.1), if $\theta = 1$ then $a_1 = 0$. In this case $\theta = (1 - \frac{a_2}{a_0}) \times (1 + (n-1)\frac{a_2}{a_0})(1 + (n-2)\frac{a_2}{a_0})^{-1}$, and is 1 if and only if $a_2 = 0$. Thus $\theta = 1$ if and only if $a_1 = 0$ and $a_2 = 0$. Then the design is orthogonal, and $\boldsymbol{A}$ is a diagonal matrix. This completes the proof of (1). When $a_1 = 0$ and $a_2 = 0$, we get

$-\frac{(n-2l)^2-n}{n(n-1)} \times \frac{2k-n}{2l-n}\binom{n}{k} + \frac{(n-2k)^2-n}{n(n-1)}\binom{n}{k} = 0$, which implies that

$$n = (n-2l)(2k-n). \tag{A.2}$$

Because the design based on S_l and S_k, and the design based on S_{n-l} and S_{n-k} have the same information matrix $\boldsymbol{C_m}$, we can take $k + l \geq n$. Rewrite (A.2) as

$$n = m(m+2i), \tag{A.3}$$

where $m = n-2l(>0)$ and $i = k+l-n(\geq 0)$. Now (A.3) is a necessary condition for $a_1 = 0$ and $a_2 = 0$. Sufficiency of this condition can be verified and is true for $i = 0$ and $i = 1$. This completes the proof of part (2) of the theorem.

If $i \geq 2$, we have $l = \frac{n-m}{2} = \frac{m(m+2i-1)}{2}$, $k = \frac{n+m+2i}{2}$, and

$$a_1 = \frac{(n-1)!}{l!(n-k)!(n-l)(n-l-1)\cdots(l+1)} \times \frac{\binom{k}{i}i!}{m+2i}[-f(i)+1],$$

$$f(i) = \frac{m\binom{l+m+i}{i}}{(m+2i)\binom{l}{i}}.$$

Then $f(2) = \frac{m^4+10m^3+31m^2+30m+8}{m^4+10m^3+31m^2+22m-24} > 1$, and $\frac{f(i+1)}{f(i)} = \frac{l(m+2i)+(m+i+1)(m+2i)}{l(m+2i)+2l-2i-i(m+2i)} > 1$. Thus $f(i) > 1$, if $i \geq 2$ and $a_1 < 0$, and in this case $\theta < 1$. This completes the proof of part (3) of the theorem.

References

Huber, J. and Hansen, D. (1986). Testing the impact of dimensional complexity and affective differences of paired concepts in adaptive conjoint analysis. *Adv. Consumer Res.* **14**, 159-163.

Huber, J. and Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *J. Marketing Res.* **23**, 307-317.

Krieger, A. M. and Green, P. E. (1991). Designing Pareto Optimal stimuli for multiattribute choice experiments. *Marketing Lett.* **2**, 337-348.

Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments.* John Wiley, New York.

Raghavarao, D. and Wiley, J. B. (1998). Estimating main effects with Pareto Optimal subsets. Australian and New Zealand, *J. Statist.* **40**, 425-432.

Wiley, J. B. (1978). Selecting Pareto Optimal subsets from multiattribute alternatives. *Adv. Consumer Res.* **5**, 171-174.

Zhang, D. (2001). Pareto optimal design in behavioral experiments. Unpublished Ph.D. dissertation, Temple University.

Department of Statistics, The Fox School of Business and Management, Temple University, Speakman Hall (006-00), Philadelphia, PA 19122.

E-mail: v5558e@vm.temple.edu

Purdue Biopharma L. P., 201 College Road East, Princeton, New Jersey 08540.

E-mail: zhang13@jeflin.tju.edu

(Received December 2000; accepted March 2002)