STATISTICAL DESIGNS FOR MIXTURES

by

W.T. Federer and D. Hall

BU-570-M*

Abstract

In nutritional, medical, educational, agronomic, and other types of experiments, it is necessary to use a mixture of k out of v treatments of interest and to obtain a single response to the mixture of k treatments. Utilizing balanced incomplete block and weighing design theory, it is possible to construct designs to estimate the general mixing effect of any given treatment and to estimate the specific mixing effect for any pair of treatments. Higher ordered specific mixing effects for triplets, quadruplets, etc. mixtures of treatments are estimable in certain cases. Minimal designs for obtaining solutions for general mixing effect parameters and specific mixing effect parameters for any pair of treatments, are discussed.

---

*In the Mimeo Series of the Biometrics Unit, 337 Warren Hall, Cornell University, Ithaca, N. Y. 14853.

STATISTICAL DESIGNS FOR MIXTURES

by

W.T. Federer and D. Hall

BU-570-M* August, 1975

1. Introduction

In many types of experimentation and survey investigations, it is possible to obtain only one response to k of v stimuli rather than being able to obtain responses to each of the k stimuli. For example, in studies on v nutritive elements given in diets of k elements to an individual, only one response to the k elements is obtainable. In studies on mixtures of cultivars, it may be impossible or undesirable to obtain individual yields and it may be that only the yield of the mixture of k cultivars is available. (See Federer [1975] and Federer et al. [1973].) In medical studies, in chemical research, (Free and Wilson [1964]), in educational studies, etc., similar situations arise. Also, for survey work alternatives to the randomized response technique have been proposed by Smith et al. [1974] and Raghavarao and Federer [1973]. Several additional illustrations of this nature could be added to this list but these should be sufficient to indicate that a statistical design and analysis problem for such situations does exist in the real world. Situations also exist wherein the response to individual stimuli are available even when used in combinations. A problem of statistical analyses exists here (see Federer [1975]).

---

*In the Mimeo Series of the Biometrics Unit, 337 Warren Hall, Cornell University, Ithaca, N.Y. 14853.

We shall consider statistical designs for investigations yielding one observation from a mixture of k stimuli out of v such stimuli, $k < v$. We shall follow the definitions for general mixing effects (gme), bispecific mixing effects (bisme) or interactions effects between pairs of stimuli, and trispecific mixing effects (trisme) or interactions peculiar to triplets of stimuli as given by Federer [1975] and shall indicate minimal and complete statistical designs for these cases.

It will be necessary to have at least as many groups (incomplete blocks) b of k stimuli each as there are parameters to estimate. For example, if only general mixing effect plus stimulus effect is to be estimated, then a minimal statistical design is obtained from a symmetrical balanced incomplete block design with parameters $v = b$, $r = k$, and λ. If gme plus stimulus effect plus bisme effects are to be estimated, then b must be greater than or equal to $v(v-1)/2$. If in addition, trisme's are to be estimated, then b must be greater than or equal to $v(v-1)(v-2)/6$, depending upon k and v.

2. Minimal Designs for gme

A symmetrical vkλ BIB design will suffice to obtain solutions for the mean effects from the v entries in an experiment. For example, let $v = 7 = b$, $r = 3 = k$, and $\lambda = 1$, and the groups of three out of seven (1,2,3,4,5,6,7) stimuli are:

Block			
1	1	2	4
2	2	3	5
3	3	4	6
4	4	5	7
5	5	6	1
6	6	7	2
7	7	1	3

.

The effect of stimuli i is obtained by taking one-third of the yields of blocks containing stimuli i minus one-sixth of the yields of the remaining blocks.

Symmetrical BIB designs exist for $v = 4t-1 = b$ in blocks of size $r = k = (v-1)/2$ and $(v+1)/2$ with $\lambda = (v-3)/4$ and $(v+1)/4$, respectively. Symmetrical BIB designs exist for every v when $k = v-1$ and may be obtained by dropping one row from a latin square design. For $v \neq 4t-1$ and $k \neq v-1$, it is sometimes possible to obtain partially balanced incomplete block designs and/or balanced incomplete block designs with $b \geq v$ from which solutions for the stimuli effects can be obtained.

The following balanced incomplete block designs may be found in Cochran and Cox [1957] for k = 3, 4, and 5:

k	v	r	b	λ	Plan # in C.&C.
3	6	5	10	2	11.4
3	6	10	20	4	11.5
3	7	3	7	1	11.7*
3	9	4	12	1	10.1
3	10	9	30	2	11.15
3	13	6	26	1	11.21
3	15	7	35	1	11.24
3	19	9	57	1	11.30
3	21	10	70	1	11.33
4	6	10	15	6	11.6
4	7	4	7	2	11.8*
4	8	7	14	3	11.10
4	9	8	18	3	11.11
4	10	6	15	2	11.16
4	13	4	13	1	11.22*
4	25	8	50	1	11.36
5	9	10	18	5	11.12
5	10	9	18	4	11.17
5	11	5	11	2	11.19*
5	21	5	21	1	11.34*
5	41	10	82	1	11.43

*Symmetrical balanced incomplete block design

Here we note that for most designs b is considerably larger than v and that the number is limited.

3. Minimal Designs for gme and bisme Effects

In order to illustrate the nature of statistical designs resulting in solutions for gme and sme effects, consider the following example. Vet $v = 7$ and $k = 3$, then the total number (35) of combinations is given by:

block		block		block		block		block	
1	1 2 4	8	3 5 6	15	1 2 3	22	1 4 7	29	2 5 7
2	2 3 5	9	4 6 7	16	1 2 5	23	1 6 7	30	3 4 5
3	3 4 6	10	5 7 1	17	1 2 7	24	2 3 4	31	3 4 7
4	4 5 7	11	6 1 2	18	1 3 5	25	2 3 6	32	3 5 7
5	5 6 1	12	7 2 3	19	1 3 6	26	2 4 6	33	3 6 7
6	6 7 2	13	1 3 4	20	1 4 5	27	2 4 7	34	4 5 6
7	7 1 3	14	2 4 5	21	1 4 6	28	2 5 6	35	5 6 7

Design $D_1(7,7,3,3,1)$ | Design $D_2(7,7,3,3,1)$ | Design $D_3(7,21,3,9,3)$

The complete balanced incomplete block design D with parameters $v = 7$, $b = 35$, $k = 3$, $r = 15$, and $\lambda = 5$ is a union of three BIB designs, that is, $D = \bigcup_{i=1}^{3} D_i = D_1 \cup D_2 \cup D_3$. The situation in the general case appears to be unresolved and would appear to revolve around the solution for the following mathematical problem: How and when does one decompose a difference set into subsets which are themselves difference sets?

Hall [1975] has noted that design D_3 results in solutions for gme and bisme effects when the following yield equation from a randomized complete block design holds

$$Y_{hi} = \mu + \rho_h + \sum_{j=1}^{v} n_{ij}\left(\tau_j + \delta_j + \sum_{g=2}^{v} n_{ig}\gamma_{jg}\right) + \epsilon_{hi}\ , \tag{3.1}$$

where μ is a general mean effect, ρ_h is the effect of the h'th block, τ_j is the effect of the j'th stimuli in the presence of itself, δ_j is the general mixing effect of the j'th stimuli, γ_{jg} is the bispecific mixing effect of stimuli

j and g, $j < g$, the ϵ_{hi} are IID$(0,\sigma_\epsilon^2)$, and the n_{ij} and n_{ig} are zeros or ones denoting whether or not g and j are in the mixture. Solutions for $(\tau_j+\delta_j)$ are obtained. Federer [1975] has shown how to design experiments allowing solutions for the τ_j and δ_j separately.

The number of blocks b must be equal to or greater than $v(v-1)/2$ and designs of the form of D_3 are desired. Hence, if $b = v(v-1)/2$, the $bk = kv(v-1)/2 = vr$, $r = k(v-1)/2$, and $\lambda = r(k-1)(v-1) = k(k-1)/2$. In order for this design to exist k and/or $v-1$ must be even, otherwise r would not be an integer. Some BIB designs with these parameters would be:

k	v	b	r	λ	Plan # in C.&C.
3	5	10	6	3	
3	7	21	9	3	
3	9	36	12	3	
3	11	55	15	3	
3	13	78	18	3	
⋮					
4	5	10	8	6	
4	6	15	10	6	11.6
4	7	21	12	6	
4	8	28	14	6	
4	9	36	16	6	
4	10	45	18	6	
4	11	55	20	6	
4	12	66	22	6	
⋮					
5	7	21	15	10	
5	9	36	20	10	
5	11	55	25	10	
5	13	78	30	10	
⋮					

Only one of these designs is listed in Cochran and Cox [1957]. It would appear that these designs would need to be constructed perhaps by methods given in Raghavarao [1971].

4. Minimal Designs for gme, bisme, and trisme Effects

Suppose the yield equation for a mixture of k of v stimuli in a randomized complete block design is

$$Y_{hi} = \mu + \rho_h + \sum_{j=1}^{v-2} n_{ij}\left(\tau_j + \delta_j + \sum_{g=2}^{v-1} n_{ig}\left(\gamma_{jg} + \sum_{f=3}^{v} n_{if}\Pi_{hgf}\right)\right) + \epsilon_{hi} , \tag{4.1}$$

where Π_{hjg} represents an effect due to the mixture of stimuli h, j, and g, $j < h < g$, and the remaining effects are as defined for equation (3.1).

The minimum number of blocks must be $b = v(v-1)(v-2)/6$, $v \geq 6$, and $k \geq 3$, in order to obtain solutions for the gme's, the bisme's, and the trisme's. In general for $k = 3$ in stimuli in a mixture, the following holds:

Source of variation	Number of parameters	Number of independent restraints	Degrees of freedom
Total	$\binom{v}{3}=v(v-1)(v-2)/6$	0	$\binom{v}{3}$
Mean	1	0	1
gme	v	1	v-1
bisme	v(v-1)/2	v	v(v-3)/2
trisme	v(v-1)(v-2)/6	$\Sigma_{i=1}^{v}(v-i) = v(v-1)/2$	v(v-1)(v-5)/6

The degrees of freedom for the various sources of variation for v = 6, 7, 8, 9, and 10 are:

Source of variation	Degrees of freedom v=6	v=7	v=8	v=9	v=10
Total	20	35	56	84	120
Mean	1	1	1	1	1
gme	5	6	7	8	9
bisme	9	14	20	27	35
trisme	5	14	28	48	75

Balanced incomplete block designs with parameters $v, b = v(v-1)(v-2)/6$, $k = 3$, $r = (v-1)(v-2)/2$, and $\lambda = v-2$ need to exist in order to obtain solutions for all the effects in equation (4.1).

For $k \geq 3$ and $v \geq 6$, b must be at least equal $v(v-1)(v-2)/6$. For example, when $k = 4$ and $v = 6$, $\binom{6}{4} < \binom{6}{3} = 20$ and therefore solutions cannot be obtained for the trisme's. The same is true when $k = 4$ and $v = 7$. However, for $v \geq 8$, $\binom{v}{4} > v(v-1)(v-2)/6$.

5. Discussion

The complete BIB design has $b = \binom{v}{k} = v!/k!(v-k)!$, v, k, $r = (v-1)!/(k-1)!(v-k)!$, and $\lambda = (v-2)!/(k-2)!(v-k)!$ and is denoted as CBIBD. Whenever $b > v(v-1)(v-2)/6$ in the CBIBD, solutions to all effects in equation (4.1) can be obtained. However, if a BIB is available for parameters v, k, $b = v(v-1)(v-2)/6$, $r = k(v-1)(v-2)/6$, and $\lambda = k(k-1)(v-2)/6$, this will be the minimal design and should be used to obtain the mixtures of k stimuli used as treatments in an experimental design.

Hall [1975] has considered the statistical analysis for equation (3.1) and has found the X'X matrix after adding restraints. It has a patterned form and is not too difficult to invert.

When $\binom{v}{k} = b < v(v-1)(v-2)/6$, Federer [1975] has shown how to design the combinations in order to obtain solutions for the various general and specific mixing effects.

As has been demonstrated many times in the past, a new scientific idea is often considered impractical or inconsequential. This was the case with the proposal by Jensen [1952] for growing mixtures of cultivars. The beneficial effects are just now being realized although experimental documentation has been available for many years (for example Jensen and Federer [1964, 1965]). The work of Free and Wilson [1964] appears on the way to being widely accepted in chemical research and the work of Federer et al. [1973] has been favorably received by agronomists. Thus it would appear that a statistician can aid in the design and analysis of investigations wherein mixtures of stimuli are used and that the scientific climate is such as to accept these ideas.

6. Literature Cited

1. Cochran, W.G. and Cox, G.M. [1957]. Experimental Designs, 2nd edition, John Wiley and Sons, Inc., New York.

2. Federer, W.T. [1975]. Statistical designs for mixtures of crops and other applications. BU-560-M in the Mimeo Series of the Biometrics Unit, Cornell University, June.

3. Federer, W.T., Hedayat, A., Lowe, C.C., and Raghavarao, D. [1973]. An application of spring balance designs to crop estimation with special reference to legumes and mixtures of crops. BU-472-M in the Mimeo Series of the Biometrics Unit, Cornell University. (To be published in the Agronomy Journal.)

4. Free, Jr., S.M. and Wilson, J.W. [1964]. A mathematical contribution to structure-activity studies. Journal of Medical Chemistry 7:395-399.

5. Hall, D.B. [1975]. Minimal designs to estimate bi-specific mixing ability. M.S. Thesis, Biometrics Unit, Cornell University, December.

6. Jensen, N.F. [1952]. Intra-varietal diversification in oat breeding. Agronomy Journal 44:30-34.

7. Jensen, N.F. and Federer, W.T. [1964]. Adjacent row competition in wheat. Crop Science 4:641-645.

8. Jensen, N.F. and Federer, W.T. [1965]. Competing ability in wheat. Crop Science 5:449-452.

9. Raghavarao, D. [1971]. Constructions and Combinatorial Problems in Design of Experiments. John Wiley and Sons, Inc., New York, London, Sydney, and Toronto.

10. Raghavarao, D. and Federer, W.T. [1973]. Application of BIB designs as an alternative to the randomized response method in survey sampling. BU-490-M in the Mimeo Series of the Biometrics Unit, Cornell University. (To be published in the Journal of the American Statistical Association.)

11. Smith, L.L., Federer, W.T., and Raghavarao, D. [1974]. A comparison of three techniques for eliciting truthful answers to sensitive questions. Social Statistics Section Proceedings of the American Statistical Association, pp. 447-452.