MEASURES OF BLOCK DESIGN EFFICIENCY RECOVERING INTERBLOCK INFORMATION

Walter T. Federer
337 Warren Hall, Biometrics Unit
Cornell University
Ithaca, NY 14853
and
Terry P. Speed
Division of Mathematics & Statistics, CSIRO
G.P.O. Box 1965
Canberra, ACT 2601, Australia

ABSTRACT. In evaluating goodness of a class of designs, researchers have used a measure of design efficiency proposed by F. Yates in the thirties. This measure considers only intrablock information and does not make use of the information contained in the interblock variance. The measures of efficiency proposed here are dependent upon the ratio of the interblock and intrablock components of variance, i.e., $\sigma_\beta^2/\sigma_\varepsilon^2 = \gamma$. The efficiency of one block design to a second may not remain invariant with respect to this ratio. Incomplete block designs which were inefficient under the intrablock measure, now become quite efficient for some ratios of γ. Likewise, the indications are that interblock information should always be recovered when analyzing data from experiments arranged in an incomplete block design.

I. INTRODUCTION. In the mid-thirties Yates (e.g. in 1937) introduced an efficiency factor for partially confounded factorials and for incomplete block designs. The factor is computed as the ratio of the average variance of a difference between two adjusted means (or for factorial effects) to the variance of a difference of two means from an orthogonal design such as a completely randomized or randomized complete block design assuming no change in the error variance σ_ε^2 for the two designs. It is common practice in statistical literature to present this efficiency factor for designs and to discuss optimality of classes of designs in terms of the Yates efficiency factor, which only makes use of the intrablock error variance. No use is made of the information contained in the interblock variance obtained from the incomplete blocks (eliminating treatment effects) sum of squares. A more proper efficiency factor should make use of the information contained in both the intrablock and the interblock variances. Some measures accomplishing this are presented in this paper.

BU-851-M in the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, NY 14853.

II. BALANCED INCOMPLETE BLOCK DESIGN. A classical balanced incomplete block design (BIBD) consists of v treatments arranged in b incomplete blocks of size k, k<v, with r repetitions of each treatment, and with each and every pair of treatments occurring together in an incomplete block λ times. The standard relations are $bk = rv$, $\lambda = r(k-1)/(v-1)$, and $e = (1-k^{1})/(1-v^{1}) = v(k-1)/k(v-1)$. The factor e is the Yates efficiency factor. The usual response model assumed for a BIBD is

$$Y_{hij} = n_{hij} (\mu + \rho_h + \beta_{hi} + \tau_j + \varepsilon_{hij}), \qquad (1)$$

where Y_{hij} is the response for the jth treatment in the ith incomplete block in the hth complete block, $h=1,\dots,r$; $i=1,\dots b/r$; $j=1,\dots,v$; n_{hij} is one if the jth treatment occurs in the hith incomplete block and zero otherwise; μ is a general mean effect; ρ_h is the hth complete block effect; β_{hi} is the hith random incomplete block effect distributed with mean zero and common variance σ_β^2, τ_j is the jth treatment effect, and ε_{hij} are random error effects which were distributed with mean zero and variance σ_ε^2. An analysis of variance is given in Table 1.

TABLE 1. Analysis of variance for a resolvable BIBD.

Source of variation	Degree of freedom	Expected value of mean square
Total	$bk = vr$	-
Correction for mean	1	-
Treatment (ignoring incomplete block effects)	$v-1$	-
Within treatments	$v(r-1)$	-
Blocks (eliminating treatment effects)	$b-1$	$\sigma_\varepsilon^2 + \frac{bk-v}{b-1}\sigma_\beta^2$ [1]
Complete blocks	$r-1$	
Incomplete blocks (elim. tr.)	$b-r$	$\sigma_\varepsilon^2 + \frac{r-1}{r}k\sigma_\beta^2$
Intrablock error	$vr-v-b+1$	σ_ε^2

[1] Expected mean square for $\rho_h = 0$, i.e., no complete block effects.

Intrablock information $\omega = 1/\sigma_\varepsilon^2$

Interblock information $\omega' = 1/(\sigma_\varepsilon^2 + k\sigma_\beta^2)$

For intrablock contrasts the variance of a difference between two estimated treatment effects is $2\sigma_\varepsilon^2/re$, when $e=(1-1/k)/(1-1/v)$. For interblock contrasts the variance of a difference between two treatments effects is

$$\frac{2(\sigma_\varepsilon^2 + k\sigma_\beta^2)}{r(1-e)} \quad . \tag{2}$$

For the combined estimator of a difference between two treatment effects the variance is

$$\frac{2}{r}\left\{\frac{e}{\sigma_\varepsilon^2} + \frac{1-e}{\sigma_\varepsilon^2 + k\sigma_\beta^2}\right\}^{-1}$$

$$= \frac{2\sigma^2_\varepsilon}{r}\left\{\frac{1+k\gamma}{1+ke\gamma}\right\}, \tag{3}$$

where $\gamma = \sigma_\beta^2/\sigma_\varepsilon^2$.

Since the intrablock contrast variance is of the form $2\sigma_\varepsilon^2/re$, it would be logical to have the combined estimator variance in the same for, i.e., $2\sigma_\varepsilon^2/re_1^*$, where

$$e_1^* = \frac{1+ke\gamma}{1+k\gamma} = 1 - \frac{k(1-e)}{1+k\gamma} \quad . \tag{4}$$

A second measure not involving e is

$$e_2^* = \frac{1+k\gamma}{1+(k+1)\gamma} = 1 - \frac{1}{k+1+1/\gamma} \tag{5}$$

The latter measure of efficiency depends only upon $\gamma = \sigma_\beta^2/\sigma_\varepsilon^2$ and k; note that (4) is also a function of k and γ since $e=k/(k+1)$ for $v=k^2$ and $r=k+1$. A comparison of the two measures is given in Table 2 for k=3,7, and 11. There is little to choose between e_1^* and e_2^* and it is suggested that e_1^* be used as a measure of efficiency rather than e_2^*. Note that as γ approaches zero the efficiency for all k approaches unity. When γ approaches infinity, the Yates efficiency factor e is approached for all k. For small k, e is relatively low indicating an inefficient design. However, e_1^* indicates that designs with small k are quite efficient if γ is 1/4 to 1/16, say. From these results, it is suggested that interblock information should always be recovered and that inefficiency of incomplete block design is not a problem.

TABLE 2. Intrablock - interblock efficiencies for various values of γ for k = 3, 7, and 11.

	k					
	3		7		11	
γ	e_1^*	e_2^*	e_1^*	e_2^*	e_1^*	e_2^*
0	1	1	1	1	1	1
1/32	.98	.97	.98	.98	.98	.98
1/16	.96	.95	.96	.96	.97	.96
1/4	.89	.88	.92	.92	.94	.94
1/2	.85	.83	.90	.90	.93	.93
1	.81	.80	.89	.89	.92	.92
2	.79	.78	.88	.88	.92	.92
4	.77	.76	.88	.88	.92	.92
∞	.75	.75	.875	.875	.917	.917

$$e_1^* = \frac{1 + ke\,\gamma}{1 + k\,\gamma}\,. \qquad e_2^* = \frac{1 + k\,\gamma}{1 + (k+1)\gamma}\,.$$

For a randomized complete block design, the variance of a difference between two arithmetic means is

$$\frac{2\sigma_e^2}{r}\left\{1 + \frac{v - k}{v - 1}\gamma\right\}, \tag{6}$$

whereas the variance of a difference between two adjusted treatment means from a BIBD is

$$\frac{2\sigma_e^2}{re_1^*} \tag{7}$$

Now (6) ≥ (7), their difference being

$$1 + \frac{v - k}{v - 1} \gamma - \frac{1 + k\gamma}{1 + ke\gamma} = \frac{v(v-k)(k-1)\gamma^2}{(v-1)[v-1 + v(k-1)\gamma]} . \tag{8}$$

(8) is zero when $\gamma = 0$ and/or $v = k$. Equation (8) could be another measure of intrablock - interblock efficiency. Perhaps a more appropriate measure would be a ratio rather a difference to obtain

$$e_3^* = 1/e_1^* \left(1 + \frac{(v-k)\gamma}{v-1}\right). \tag{9}$$

The measure e_3^* would conform more to the definition of efficiency originally presented by Yates but would include both intrablock and interblock information.

III. OTHER BLOCK DESIGNS. A class of generalized N-ary designs were discussed by Shafiq and Federer (1979, 1983). For these designs the response model equation (1) is replaced by

$$Y_{ghij} = \mu + \rho_h + \beta_{hi} + \tau_j + \varepsilon_{ghij}, \tag{10}$$

where $g = 0,..,n_{hij}$ and when $n_{hij} = 0$ there is no response Y_{ghij}. The other symbols are as defined in (1). The above authors generalized the Yates efficiency factor for this class of designs and they also proved that the Fisher inequality $v \le b$ holds for this general class of balanced block designs. The efficiency factor e_1^* for the generalized balanced block design is

$$e_1^* = 1 - \frac{c - \lambda}{r(k + \gamma^{-1})} , \tag{11}$$

where $c = \sum_h \sum_i n_{hij}^2$, n_{hij} is the number of times treatment j occurs in block hi, and $r = \sum_h \sum_i n_{hij}$ is the number of replicates for treatment j.

For the class of resolvable incomplete block designs known as lattices, the average variance of a difference between two adjusted treatment means is

$$\frac{2}{k+1} \left\{\frac{r}{(r-1)\omega + \omega'} + \frac{k-r+1}{r\omega}\right\}$$

$$= \frac{2\sigma_e^2}{r} \left(\frac{r}{k+1}\right)\left(\frac{r}{r-1 + (1+k\gamma)^{-1}} + \frac{k-r+1}{r}\right) . \qquad (12)$$

As before, we may take

$$e_1^* = \frac{k+1}{r}\left(\frac{r}{r-1 + (1+k\gamma)^{-1}} + \frac{k-r+1}{r}\right)^{-1}. \qquad (13)$$

Note that r is the number of confounding arrangements and r=2 for the simple (double) lattice, r=3 for the triple lattice, etc.

An intrablock - interblock measure of efficiency like e_3^* would be

$$e_3^* = \frac{\left[1 + \frac{r}{k+1}\left(\frac{1 - (1+k\gamma)^{-1}}{r-1 + (1+k\gamma)^{-1}}\right)\right]}{\left[1 + \frac{k}{k+1}\gamma\right]} \qquad (14)$$

Note that $v=k^2$ and r is the number of geometrical factorial effects confounded.

Using the average variance of a difference between two adjusted treatment effects, we could construct e_1^* and e_3^* for any class of incomplete block designs. The ideas in this paper may be used to construct efficiency factors similar to e_1^* and e_3^* for cubic lattices, for lattice squares, and for other designs.

IV. LITERATURE CITED.

Shafiq, M. and W. T. Federer (1979). "Generalized N-ary balanced block designs." *Biometrika* 66, pp. 115-123.

Shafiq, M. and W. T. Federer (1983). "General binary partially balanced designs." *Indian J. Agric. Statist.*, XXXV(2).

Yates, F. (1937). "The design and analysis of factorial experiments." *Imperial Bur. Soil Sci., Tech. Comm.* 35, pp. 1-95.