Balanced Incomplete Block Designs for v Treatments in b=v(v-1) Incomplete Blocks of Size k=2

BU-91-M

W. T. Federer*

May, 1958

Homogeneous experimental material often is available only in pairs: e.g., left and right readings or exposures, twins, opposite halves of a leaf, opposite leaves on a plant, left and right members of a subject, etc. If more than two treatments are being compared in an experiment, it is necessary to use incomplete blocks of size two if the variation among pairs is to be eliminated from the residual error mean square in the experiment.

In other situations, lack of experimental material necessitates the use of incomplete blocks of size two. For example, certain analyses or observations on eggs require one-half of an egg; in other cases, each treatment requires an entire greenhouse, nutrient tank, counter in a store, and incubator of which only two are available; etc.

Similar incomplete block designs may be constructed for both situations, but regardless of the situation it is often necessary to introduce a new feature in the design and/or analysis because of the peculiarity of the experimental situation. Such was the case in an experiment involving methods of cooling in a greenhouse. Enough equipment was available for only two greenhouses. Eight or nine different treatments were to be used. The experimental unit was one day in one greenhouse. It was desired to balance the treatment effects with greenhouse effects since the experiment would be conducted for 56 days if 8 treatments were used and 72 days if 9 treatments were used. In the absence of any knowledge concerning treatment differences or experimental error, it was decided to have 14 or 16 replicates per treatment.

---

*Biometrics Unit, Department of Plant Breeding, Cornell University

Construction of the design

A simple procedure of constructing designs of the desired type is available. Examples for v=3 and 4 are used to illustrate the procedure. The designs for v=3 and 4 follow:

v=3

Incomplete block no.	Greenhouse I	Greenhouse II
1	A	B
2	B	C
3	C	A
4	A	C
5	B	A
6	C	B

v=4

Incomplete block no.	Greenhouse I	Greenhouse II
1	A	B
2	B	C
3	C	D
4	D	A
5	A	C
6	B	D
7	C	A
8	D	B
9	A	D
10	B	A
11	C	B
12	D	C

All possible combinations of the v treatments plus the mirror images comprise the v(v-1) incomplete blocks. Thus, for v=3 the possible combinations are AB, AC, and BC and the mirror images are BA, CA, and CB to make up the v(v-1)=3(2)=6 arrangements for the incomplete blocks. Likewise for v=4 the possible combinations of two treatments are AB, AC, AD, BC, BD, and CD and the mirror images are BA, CA, DA, CB, DB, and DC; these make up the 12 arrangements.

The extension for v= any integer follows immediately from the above. For v treatments in b=v(v-1) incomplete blocks of size k=2, there are r=2(v-1) replicates on each treatment since one of the conditions for a balanced lattice design is that bk=vr. Hence, r=2v(v-1)/v=2(v-1).

Intrablock analysis

The yield of the ith treatment in the jth incomplete block in the hth greenhouse is assumed to be

$$Y_{ijh}=n_{ij}(\mu+\alpha_h+\tau_i+\beta_j+\epsilon_{hij}) \qquad (1)$$

where n_{ij}=1 if the ith treatment falls in the jth incomplete block and =0 otherwise, μ = an effect common to every observation, α_h = a fixed effect common to the hth greenhouse, τ_i = a fixed effect common to the ith treatment, β_j = a fixed effect common to the jth incomplete block, ϵ_{hij} = a random effect normally and independently distributed with mean zero and variance σ_ϵ^2, h=1,2, v=1,2,•••,v, and j=1,2,•••, b=v(v-1).

Minimization of the following sum of squares with respect to μ, α_h, τ_i, and β_j

$$\sum_{hij}\sum\sum n_{ij}(Y_{hij}-\mu-\alpha_h-\tau_i-\beta_j)^2 \qquad (2)$$

results in the following normal equations:

$$2b\mu+b(\alpha_1+\alpha_2)+r\sum_{i=1}^{v}\tau_i+2\sum_{i=1}^{b}\beta_j=Y_{\cdot\cdot\cdot}\ ; \qquad (3)$$

$$b(\mu+\alpha_h)+\frac{r}{2}\Sigma\tau_i+\Sigma\beta_j=Y_{h\cdot\cdot}\ ; \qquad (4)$$

$$r(\mu+\tau_i)+\frac{r}{2}(\alpha_1+\alpha_2)+\sum_j n_{ij}\beta_j=Y_{\cdot i\cdot}\ ; \qquad (5)$$

$$2(\mu+\beta_j)+(\alpha_1+\alpha_2)+\sum_i n_{ij}\tau_i=Y_{\cdot\cdot j}\ . \qquad (6)$$

Using the equations

$$\Sigma\tau_i=\Sigma\beta_j=\Sigma\alpha_i=0 \qquad (7)$$

results in unique solutions for μ, α_h, τ_i, and β_j. The estimates of μ and α_h satisfying equations (3), (4), and (7) are:

$$\hat{\mu}=\bar{y} ; \tag{8}$$

$$\hat{\alpha}_h=\bar{y}_{h\cdot\cdot}-\bar{y} . \tag{9}$$

The v equations involving only the τ_i are the usual ones for a non-orthogonal two-way classification, i.e., the fth equation is:

$$n_{f\cdot}\hat{\tau}_f-\sum_j \frac{n_{fj}}{n_{\cdot j}} \sum_i n_{ij}\hat{\tau}_i$$

$$=r\hat{\tau}_f-\frac{1}{2}\sum_i \hat{\tau}_i \sum_j n_{fj}n_{ij}$$

$$=Y_{\cdot f\cdot}-\sum_j n_{fj}\bar{y}_{\cdot\cdot j}=Q_{f\cdot} . \tag{10}$$

Since this is a balanced lattice, the $\sum_{j=1}^{b} n_{fj}n_{ij}=\lambda$, a constant for $f\neq i$. In this case $\lambda=2$. Now $\sum_{j=1}^{b} n_{ij}^2=2(v-1)=r$. Making use of these facts in (10) we find

$$(r-\frac{r}{2}+1)\hat{\tau}_i=Q_{i\cdot} , \tag{11}$$

or
$$\hat{\tau}_i=Q_{i\cdot}/v . \tag{12}$$

The gth equation of the b equations involving the $\hat{\beta}_j$ is

$$n_{\cdot g}\hat{\beta}_g-\sum_i \frac{n_{ig}}{n_{i\cdot}} \sum_j n_{ij}\hat{\beta}_j$$

$$=2\hat{\beta}_g-\frac{1}{r}\sum_j \hat{\beta}_j \sum_i n_{ig}n_{ij}$$

$$=Y_{\cdot\cdot g}-\sum_i n_{ig}\bar{y}_{\cdot i\cdot}=Q_{\cdot g} . \tag{13}$$

The b equations plus the equation $\Sigma\hat{\beta}_j=0$ reduces to v(v-1)/2 pairs of equations. Each pair of equations involves the block effect for a given treatment combination and the block effect for the mirror image of that treatment combination. Thus, let g = the treatment combination AB and let g' = the treatment combination BA; the resulting pair of equations is:

$$\hat{\beta}_g(2-\frac{1}{r})-\frac{1}{r}\hat{\beta}_{g'}=Y_{..g}-\bar{y}_{.A.}-\bar{y}_{.B.}=Q_{.g} \tag{14}$$

and

$$-\frac{1}{r}\hat{\beta}_g+(2-\frac{1}{r})\hat{\beta}_{g'}=Y_{..g'}-\bar{y}_{.B.}-\bar{y}_{.A.}=Q_{.g'} \quad . \tag{15}$$

Therefore,

$$\hat{\beta}_g=\frac{Q_{.g}(2r-1)+Q_{.g'}}{4(r-1)} \tag{16}$$

and

$$\hat{\beta}_{g'}=\frac{Q_{.g}+Q_{.g'}(2r-1)}{4(r-1)} \quad . \tag{17}$$

The adjusted treatment mean using only intrablock information is

$$\bar{y}_{.i.}'=\hat{\mu}+\hat{\tau}_i=\bar{y}+Q_{i.}/v \quad . \tag{18}$$

The analysis of variance may be set up in various forms, but the following form is considered to be useful here:

Source of variation	df	Sum of squares	m.s.
Total (uncorrected)	vr	$U=\Sigma\Sigma\Sigma n_{ij}Y^2_{hij}$	-
Correction for mean	1	$CT=Y^2_{...}/vr$	-
Total (corrected for mean)	vr-1	U-CT	-
Greenhouse	1	$(Y^2_{1..}+Y^2_{2..})/b-CT=G$	-
Incomplete block or day (ign. treatment effects)	b-1	$\sum_{j=1}^{b} Y^2_{..j}/2-CT=B$	-
Day x Greenhouse	(b-1)	U-CT-G-B=BG	-
Treatments (elim. day)	(v-1)	T'	-
Intrablock error	(b-1)	R=BG-T'	E_e
Day (elim. treatment)	(b-1)	B'	E_b
Treatment (ign. day effects)	(v-1)	$T=\Sigma Y^2_{.i.}/r-CT$	-

The treatment (eliminating day or incomplete block effects) sum of squares is computed as:

$$T'=\sum_{i=1}^{v}\hat{\tau}_iQ_{i.}=v\Sigma\hat{\tau}_i^2 \quad . \tag{19}$$

The block or day (eliminating treatment effects) sum of squares is computed as:

$$B'=\sum_{j=1}^{b}\hat{\beta}_jQ_{.j} \quad . \tag{20}$$

The variance of a difference between any two adjusted means utilizing only intrablock information is:

$$V(\hat{\mu}+\hat{\tau}_i-\hat{\mu}-\hat{\tau}_{i'})=V(\hat{\tau}_i-\hat{\tau}_{i'})=2\sigma_\epsilon^2/v \tag{21}$$

and is estimated by $2E_e/v$.

Likewise, the intrablock F test is

$$F\left[(v-1),(b-v)df\right]=\frac{T'/(v-1)}{E_e} \quad . \tag{22}$$

Interblock analysis

Minimization of the following sum of squares with respect to μ, α_h, τ_i, and β_j,

$$\omega \sum_{hij}\sum\sum n_{ij}(Y_{hij}-\mu-\alpha_h-\tau_i-\beta_j)^2$$

$$+\frac{\omega'}{2}\sum_j(Y_{\cdot\cdot j}-2\mu-\alpha_1-\alpha_2-\sum_i n_{ij}\tau_i)^2 \; , \tag{23}$$

results in the following normal equations:

$$(\omega+\omega')(2b\mu+b\Sigma\alpha_h+r\Sigma\tau_i)+2\omega\Sigma\beta_j=(\omega+\omega')Y_{\cdot\cdot\cdot} \quad ; \tag{24}$$

$$b(\omega+\omega')\mu+b\alpha_h+\frac{\omega'}{2}\Sigma\alpha_h+\frac{r}{2}(\omega+\omega\frac{\omega'}{2})\Sigma\tau_i=\omega Y_{\cdot\cdot h}+\frac{\omega'}{2}Y_{\cdot\cdot\cdot} \quad ; \tag{25}$$

$$r(\omega+\omega')\mu+\frac{r}{2}(\omega+\omega')\Sigma\alpha_h+r\omega\tau_f+\frac{\omega'}{2}\sum_i\tau_i\sum_j n_{fj}n_{ij}$$

$$+2\omega\Sigma n_{fj}\beta_j=\omega Y_{\cdot f\cdot}+\frac{\omega'}{2}\sum_j n_{fj}Y_{\cdot\cdot j} \quad ; \tag{26}$$

$$\omega\left[2(\mu+\beta_j)+\Sigma\alpha_h+\sum_i n_{ij}\tau_i\right]=\omega Y_{\cdot\cdot j} \quad . \tag{27}$$

Making use of the restrictions in (7), and the parameters of the design, the above equations reduce to:

$$\hat{\mu}=\bar{y} \quad ; \tag{28}$$

$$\hat{\alpha}_h=\bar{y}_{h\cdot\cdot}-\bar{y} \quad ; \tag{29}$$

$$[\,\omega(r+2)+\omega'(r-2)]\hat{\tau}_f/2$$

$$=\omega[Y_{.f.}-\sum_j n_{fj}\bar{y}_{..j}]+\frac{\omega'}{2}[\sum_j n_{fj}Y_{..j}$$

$$-\frac{r}{b}Y_{...}]=Q'_{f.} \quad . \tag{30}$$

The equations in $\hat{\beta}_j$ are not required in this analysis since the intrablock analysis of variance table yields all the required sums of squares. ω is estimated as

$$w=1/E_e \tag{31}$$

and

$$w'=\frac{v(r-1)}{2(b-1)E_b-(v-2)E_e} \quad . \tag{32}$$

The treatment mean adjusted for interblock information is simply $\hat{\mu}+\hat{\tau}_i$.

The estimated variance of a mean difference between any two adjusted means, utilizing interblock information, is:

$$\left\{\frac{4}{w(r+2)+w'(r-2)}\right\} \quad . \tag{33}$$

The results obtained above follow from the results in textbooks by Federer (1) and Kempthorne (3) and papers by Federer (2), Kempthorne (4), and Rao (5).

Some comments

It is possible to remove some types of interaction in designs of this type. For example, a greenhouse x treatment interaction could be separated out. Likewise, if one classified the b blocks into types or groups of blocks, then a type of day x treatment interaction could also be partitioned out. In the greenhouse experiment there is a possibility of a type of day x treatment interaction, but the intrablock error mean square was quite low. Therefore, a type of day x treatment interaction was not removed from the intrablock error mean square. Also, only intrablock information was recovered in the greenhouse experiment because the block eliminating treatment sum of squares contains a component for type of day.

An example. An illustrative example for the balanced lattice described above was obtained by letting $\hat{\mu} = 10$, $\hat{\alpha}_1 = -1$, $\hat{\alpha}_2 = 1$, $\hat{\tau}_A = -2$, $\hat{\tau}_B = 3$, $\hat{\tau}_C = -1$, $\hat{\beta}_1 = 0$, $\hat{\beta}_2 = 1$, $\hat{\beta}_3 = 4$, $\hat{\beta}_4 = -4$, $\hat{\beta}_5 = 2$, and $\hat{\beta}_6 = -3$, and then constructing the yields for each observation. The following design and yields were obtained:

Block No.	Greenhouse I	Greenhouse II	Total	Mean	Treatment Totals
1	A - 7	B - 14	21	10.5	$Y_{.A.}=26$
2	C - 9	A - 10	19	9.5	$Y_{.B.}=54$
3	B - 16	C - 14	30	15.0	$Y_{.C.}=40$
4	B - 8	A - 5	13	6.5	Treatment Means
5	C - 10	B - 16	26	13.0	$\bar{y}_{.A.}=6.5$
6	A - 4	C - 7	11	5.5	$\bar{y}_{.B.}=13.5$
Total	54	66	120	--	$\bar{y}_{.C.}=10.0$
Mean	9	11	--	10.0	

Since treatment and block effects are orthogonal to greenhouse effects, the greenhouse and mean effects are computed as:

$$\hat{\mu}=\bar{y}=\frac{120}{12}=10 \quad ;$$

$$\hat{\alpha}_1=\bar{y}_{1..}-\bar{y}=9-10=-1 \quad ;$$

$$\hat{\alpha}_2=\bar{y}_{2..}-\bar{y}=11-10=1 \quad .$$

These agree, as they should, with the values of $\hat{\mu}$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$ used to construct the example.

In obtaining solutions for the $\hat{\tau}_i$ and $\hat{\beta}_j$, it is helpful to set up a table of the n_{ij}'s. For the present example, the n_{ij}'s are:

$n_{11}=1$	$n_{12}=1$	$n_{13}=0$	$n_{14}=1$	$n_{15}=0$	$n_{16}=1$
$n_{21}=1$	$n_{22}=0$	$n_{23}=1$	$n_{24}=1$	$n_{25}=1$	$n_{26}=0$
$n_{31}=0$	$n_{32}=1$	$n_{33}=1$	$n_{34}=0$	$n_{35}=1$	$n_{36}=1$

From the treatment equations (formula (10)),

$$\hat{\tau}_A = \frac{1}{3}\left\{Y_{.A.}-\bar{y}_{..1}-\bar{y}_{..2}-\bar{y}_{..4}-\bar{y}_{..6}\right\}$$

$$= \frac{1}{3}\left\{26-10.5-9.5-6.5-5.5\right\} = -2 \quad ;$$

$$\hat{\tau}_B = \frac{1}{3}\left\{Y_{.B.}-\bar{y}_{..1}-\bar{y}_{..3}-\bar{y}_{..4}-\bar{y}_{..5}\right\}$$

$$= \frac{1}{3}\left\{54-10.5-15.0-6.5-13.0\right\} = 3 \quad ;$$

$$\hat{\tau}_C = \frac{1}{3}\left\{40-9.5-15.0-13.0-5.5\right\} = -1 \quad .$$

From the block equations (formulae (16) and (17), we find:

$$\hat{\beta}_1 = \frac{1}{2}\left\{Y_{..1} - \frac{1}{3}\left(Y_{.A.}+Y_{.B.}-\bar{y}_{..1}-\bar{y}_{..4}\right)\right\}$$

$$= \frac{1}{2}\left\{21 - \frac{1}{3}\left(26+54-10.5-6.5\right)\right\} = 0 \quad ;$$

$$\hat{\beta}_2 = \frac{1}{2}\left\{Y_{..2} - \frac{1}{3}\left(Y_{.C.}+Y_{.A.}-\bar{y}_{..2}-\bar{y}_{..6}\right)\right\}$$

$$= \frac{1}{2}\left\{19 - \frac{1}{3}\left(40+26-9.5-5.5\right)\right\} = 1 \quad ;$$

$$\hat{\beta}_3 = \frac{1}{2}\left\{30 - \frac{1}{3}\left(54+40-15.0-13.0\right)\right\} = 4 \quad ;$$

$$\hat{\beta}_4 = \frac{1}{2}\left\{13 - \frac{1}{3}\left(26+54-6.5-10.5\right)\right\} = -4 \quad ;$$

$$\hat{\beta}_5 = \frac{1}{2}\left\{26 - \frac{1}{3}\left(40+54-13.0-15.0\right)\right\} = 2 \quad ;$$

$$\hat{\beta}_6 = \frac{1}{2}\left\{11 - \frac{1}{3}\left(40+26-5.5-9.5\right)\right\} = -3 \quad .$$

The values for the $\hat{\tau}_i$ and $\hat{\beta}_j$ agree, as they should, with the values used in constructing the example.

The adjusted treatment means using only intrablock information are $\bar{y}'_{.A.}=\hat{\mu}+\hat{\tau}_A=10-2=8$, $\bar{y}_{.B.}=10+3=13$, and $\bar{y}_{.C.}=10-1=9$.

The total sum of squares is obtained in the usual manner as 1388-1200=188; the greenhouse sum of squares is $(54-66)^2/12=12$; the treatment (ignoring block) sum of squares is

$$\frac{26^2+54^2+40^2}{4} - \frac{120^2}{12} = 1298-1200=98 \; ;$$

the block (ignoring treatment) sum of squares is

$$\frac{21^2+19^2+30^2+13^2+26^2+11^2}{2} - 1200 = 134 \quad ;$$

the treatment (eliminating block) sum of squares is

$$\Sigma\hat{\tau}_i Q_i=(-2)(-6)+3(9)+(-1)(-3)=42 \quad ;$$

the block (eliminating treatment) sum of squares is

$$\Sigma\hat{\beta}_j Q_{.j}=0(1)+1(2.5)+4(6.5)-4(-7.0)+2(2.5)-3(5.5)=78 \quad .$$

The $Q_{.j}$ are the block totals minus the sum of treatment means appearing in the block. The error sum of squares is obtained by subtraction as 188-12-98-78=188-12-42-134=0, as it should, since no allowance was made for error in this example.

In order to have a sum of squares for error in this experiment, change the values in the example to read as follows:

Block No.	Greenhouse I	Greenhouse II	Total	Mean	Treatment Totals
1	A - 6	B - 15	21	10.5	$Y_{.A.}=26$
2	C - 8	A - 11	19	9.5	$Y_{.B.}=54$
3	B - 15	C - 15	30	15.0	$Y_{.C.}=40$
4	B - 9	A - 4	13	6.5	Treatment Means
5	C - 11	B - 15	26	13.0	$\bar{y}_{.A.}=6.5$
6	A - 5	C - 6	11	5.5	$\bar{y}_{.B.}=13.5$
Total	54	66	120	--	$\bar{y}_{.C.}=10.0$
Mean	9	11	--	10.0	

The totals in the above table are identical to those obtained for the first example. All effects will have the same values and all sums of squares, <u>except</u> the total and error, will remain the same. The total sum of squares will increase by 12 since 12 ones (either plus or minus) were used to change the values in order to introduce an error sum of squares different from zero and still retain the same parameters. The total sum of squares is 1400-1200=200.

The analysis of variance table for these sums of squares follows:

Source of variation	df	s.s.	m.s.
Total (uncorrected)	12	1400	--
Correction for mean	1	1200	--
Total (corrected)	11	200	--
Greenhouse	1	12	12
Treatment (ign. blocks)	2	98	--
Block (elim. treatment)	5	78	15.6
Error	3	12	4
Treatment (elim. blocks)	2	42	21
Blocks (ign. treatment)	5	134	--

The variance of a difference between two treatment means adjusted for intrablock information only is (formula (21)):

$$s_{\bar{d}}^2 = \frac{2(4)}{3} = \frac{8}{3} \quad .$$

Consider now the recovery of interblock information in the second example above. The first step is to compute the weights from formulae (31) and (32) as follows:

$$w = \frac{1}{4} = .2500$$

and

$$w' = \frac{3(4-1)}{2(6-1)(15.6)-(3-2)(4)} = \frac{9}{152} = .0592 \ .$$

The variance of a difference between two treatment means adjusted for interblock information is (formula (33)):

$$\frac{4}{.2500(4+2)+.0592(4-2)} = 2.4716 \quad .$$

The average effective error variance r/2 times the average variance. For the example it is equal to 4(2.4716)/2=4.9432. The efficiency of this design relative to the design obtained using each treatment in each greenhouse r/2=2 times is

$$\frac{(78+12)/(5+5)}{4.9432} \times 100=182\% \quad ,$$

or a gain in efficiency of 82 per cent. Only 5/9 as many replicates of the incomplete block design as of the complete block design was required to obtain the same precision.

The treatment means adjusted for interblock information are (formulae (28) and (30)):

$$\hat{\mu}+\hat{\tau}_A = \bar{y}+2Q'_{A.}/[w(r+2)+w'(r-2)]$$

$$=10+ \frac{2}{1.6184} \Big\{ .2500(26-10.5-9.5-6.5-5.5)$$

$$+.0592(10.5+9.5+6.5+5.5-4(10))\Big\}$$

$$=10+ \frac{2}{1.6184} \Big\{ -1.5000-.4736 \Big\}$$

$$=7.561 \quad ;$$

$$\hat{\mu}+\hat{\tau}_B = 10+ \frac{2}{1.6184} \Big\{ 2.2500+.2960 \Big\} =13.146 \quad ;$$

$$\hat{\mu}+\hat{\tau}_C = 10+ \frac{2}{1.6184} \Big\{ -.7500+.1776 \Big\} =9.293 \quad .$$

As a partial check 7.561+13.146+9.293=30.000=$v\bar{x}$ as it should.

Literature cited

1. Federer, W. T., Experimental design, Macmillan, New York, 1955.

2. Federer, W. T., Augmented designs, Paper presented at American Stat. Assoc. and Biometric Soc. Annual Meeting in Sept., 1957.

3. Kempthorne, O., Design and analysis of experiments, Wiley, New York, 1952.

4. Kempthorne, O., A class of experimental designs using blocks of two plots, Annals of Math. Stat. 24:76-84, 1953.

5. Rao, C. R., General methods of analysis for incomplete block designs, Jour. Amer. Stat. Assoc. 42:541-561, 1947.