Comparing Treatment with a Control Using Recovery of Inter-Block Information in Augmented Balance Incomplete Block Design

Badreldin A. Mohamed

Deparment of Community Health Sciences, College of Applied Medical Sciences, King Saud University

ABSTRACT

Many times, the experimenter wishes to compare treatments with a control, in incomplete block designs. To obtain greater precision for such comparisons, the control occurs more often in each block. The analysis of such designs using recovery of inter-block information is provided in this paper. Estimates of the weights for combining intra and inter-block estimates are also provided.

Key words: Control, incomplete block designs, recovery of inter-block information, estimation of weights.

Key words: Recovery Inter-Block Information, Balance Incomplete Design

INTRODUCTION

In many biomedical or agricultural experiments, the control treatment is in logically a different position from the rest of the treatments. The control plays a pivotal role and all other test treatments are compared to it. Many standard incomplete block designs are reinforced, by including the control treatment more often (Bateman *et al*., 1994; Creaser *et al*., 2009; Trietsch *et al*., 2009). Bechhofer and Tamane (1981), and John and Edwards (1986) have developed a general theory for optimal incomplete block designs intended for this purpose of comparing treatments with a control, with great precision than for other comparisons. We propose to give detail analysis of incomplete block designs augmented with a control with intra-block and with recovery of inter-block information, when blocks themselves are random. This will provide a wider basis of inference in many biomedical and agricultural problems. Also, the estimates of weights for combining inter and intra block estimates are also different in such reinforced designs as the expected value of the adjusted block sum of squares and other sums of squares also change.

Augmented incomplete block designs

Consider $v + 1$ treatments $t_0, t_1, \ldots, t_v$ with t_0 as the control, which are tested in b blocks of k plots each, in such a way that each pair of test treatments t_i, t_u ($i \neq u$; $i, u = 1, 2, \ldots, v$) occurs in λ_1 blocks and each test treatment t_i ($i = 1, \ldots, v$) occurs with the control t_0 in λ_0 blocks. If n_{ij} ($i=0, 1, \ldots, ; j = 1, \ldots, b$) denotes the number of times the i-th treatment occurs in the j-th block, $(v + 1) \times b$ matrix N of elements n_{ij} is the incidence matrix of the design. We assume that

$$\sum_{j}^{b} n_{oj} = r_0 \qquad \sum_{j=1}^{b} n_{ij} = r, \quad (i = 1,2, \ldots\ldots, v) \tag{2.1}$$

$$\sum_{j}^{b} n_{oj}^2 = s_0 \qquad \sum_{j=1}^{b} n_{ij}^2 = s, \quad (i = 1, 2, \ldots\ldots, v) \tag{2.2}$$

$$\sum_{i=0}^{v} n_{ij} = k, \tag{2.3}$$

The foregoing analysis can be worked out even if $\sum_{j}^{b} n_{ij}$ and $\sum_{j}^{b} n_{ij}^2$ change with i, but then the algebra becomes a little messy.

From (2.1) and (2.2)

$$NN' = \left[\begin{array}{c|c} s_0 & \lambda_0 E_{1v} \\ \hline \lambda_0 E_{v1} & (s - \lambda_1) I_v + \lambda_1 E_{vv} \end{array}\right] \tag{2.4}$$

$$NN' E_{v+1,\,1} = k\, NE_{b1} = k\, \underline{r} \tag{2.5}$$

Where

$E_{pq} = a$ is $p \times q$ matrix of unit elements

$\underline{r}$ = the column vector of elements $r_0, r_1, \ldots, r_v$

From (2.5),

$$K\, r_0 = s_0 + v\, \lambda_0 \tag{2.6}$$

$$K\, r = s + (v-1)\, \lambda_1 + \lambda_0\, ; \quad V\, r + r_0 = b\, k \tag{2.7}$$

Assuming the usual model,

$$y_{ij} = \mu + t_i + \beta_j + \varepsilon_{ij} \tag{2.8}$$

where y_{ij} is the yield of the i-th treatment in the j-th block (if $n_{ij} = 1$), β_j is the effect of the j-th block, t_i is the effect of the i-th treatment, μ is the general mean and ε_{ij} are the errors assumed to be independent normal with zero means and variance σ^2. In recovery of inter-block information, we further assume that the β_j are independent normal with means and variance σ_b^2. These are further independent of the ε_{ij}'s.

Let

$\underline{T}$ = vector of the treatment totals $T_0, T_1, \ldots, T_v$

$\underline{B}$ = vector of the block totals $B_1, \ldots, B_b$

g = grand total of the yield.

We also define:

C = diag ($r_0, r, r, \ldots\ldots, r$) (2.9)

$Q = \underline{T} - k^{-1} N \underline{B}$ (2.10)

Where diag stands for a diagonal matrix and the elements Q_i of Q are the adjusted treatments totals. Further let

$W = 1/\sigma^2$; $W_1 = 1/(\sigma^2 + k\ \sigma_b^2)$ (2.11)

$\underline{Q}_1 = k^{-1} N \underline{B} - (g/(v+1)) E_{v+1,1}$; (with elements Q_{1i}) (2.12)

$C_1 = k^{-1} N N' - (b\ k)^{-1} \underline{r}\ \underline{r}'$ (2.13)

It is well known that the intra-block estimates of the treatment effects, when β_j are fixed, are obtained from the normal equations

$$Q = C\ \underline{\hat{t}} \qquad (2.14)$$

While the combined inter and intra-block estimates $\underline{t}^*$ are obtained from the combined normal equations

$$W \underline{Q} + W_1 \underline{Q}_1 = (WC + W_1 C_1)\ \underline{t}^* \qquad (2.15)$$

The matrix C and $WC + W_1 C_1$ are singular as their rows add up to zero and we take the additional equation $\underline{r}'\ \underline{\hat{t}} = 0$ for (2.14) and $\underline{r}'\ \underline{t}^* = 0$

for (2.15), which then becomes

$$W \underline{Q} + W_1 \underline{Q}_1 = (WC + k^{-1} W_1 N N')\ \underline{t}^* \qquad (2.16)$$

An element by element comparison of the matrix C of (2.14) with the matrix $WC + k^{-1} W_1 N N'$ of (2.16) shows that (2.16) can be obtained from (2.14) by changing:

r_0 to $W r_0$

s_0 to $(W - W_1)\, s_0$

r to $W\, r$

Q to $W\, \underline{Q} + W_1\, \underline{Q_1}$

s to $(W - W_1)\, s_0$

λ_0 to $((W - W_1)\, \lambda_0$

λ_1 to $((W - W_1)\, \lambda_1$ (2.17)

And if we make the changes in a solution of (2.14), we shall get a solution of (2.16). However, it must be remembered that the changes must be made, prior to any simplification using (2.6) or (2.7).

Intra and inter-block estimates

From (2.14), (2.9) and (2.4), a solution of (2.14) can be written as:

$$\hat{\underline{t}} = C^{-}\underline{Q} \quad (3.1)$$

Where C^{-} is a generalize inverse of C, with elements C^{iu} ($i,u = 0, 1, 2, ..., v$). In particular, a solution will be

$$\hat{t}_0 = r\, k\, Q_0\, (r\, r_0\, k - r\, s_0 + r_0\, \lambda_0)^{-1} \quad (3.2)$$

$$\hat{t}_i = (r\, k - s + \lambda_1)^{-1}\, k\, Q_i + (r\, r_0\, k - r\, s_0 + r_0\, \lambda_0)^{-1}\, (r\, k - s + \lambda_1)^{-1}\, (r\, \lambda_0 - r\, \lambda_1)\, k\, Q_0 \quad (i = 1 \ldots v) \quad (3.3)$$

Then

$$\hat{t}_i - \hat{t}_0 = k\, (r\, k - s + \lambda_1)^{-1}\, Q_i + (r\, \lambda_0 - r\, \lambda_1 - r^2\, k + r\, s - r\, \lambda_1)\, k\, Q_0 + (r\, r_0\, k - r\, s_0 + r_0\, \lambda_0)^{-1}\, (r\, k - s + \lambda_1)^{-1} \quad (3.4)$$

The variance of the comparison $\hat{t}_i - \hat{t}_0$ is (Klaus and Kempthorne, 2005)

$$\sigma^2\, (c^{ii} + c^{oo} - c^{io} - c^{oi}) \quad (3.5)$$

Which becomes

$$\frac{k\sigma^2}{rk-s+\lambda_1}+\frac{rk\sigma^2}{rr_0-rs_0+r_0\lambda_0}-\frac{(r\lambda_0-r_0\lambda_1)k\sigma^2}{(rr_0-rs_0+r_0\lambda)(rk-s+\lambda_1)} \quad (3.6)$$

on substituting for C^{iu} 's which are coefficients of Q_u in the solution $\hat{t}_i$ given by (3.2) – (3.7).

Further simplification can be done using (2.6), (2.7) but we postpone it for reasons explained earlier. By making the changes outlined in (2.17), the combined inter and intra-block estimates of $t_i - t_0$ is:

$$t_i^* - t_0^* = \frac{k(WQ_i + W_1Q_i)}{Wrk-(W-W_1)(s-\lambda_1)} + \frac{k(WQ_0+W_1Q_{10})[(W-W_1)(r\lambda_0-r_0\lambda_1+rs-r\lambda_1)-Wr_2k]}{[krr_0W-(W-W_1)][rkW-(W-W_1)(s-\lambda_1)]} \quad (3.7)$$

The covariance matrix of $\underline{Q}$ is σ^2 C but that of $W\,\underline{Q} + W_1\,\underline{Q}_1$ is:

$$WC + \frac{1}{k}W_1(NN' - b^{-1}\underline{r}\;\underline{r}') \quad (3.8)$$

Hence, the variance of the combined estimate $t_i^* - t_0^*$ is obtained from that $\hat{t}_i$ - $\hat{t}_0$

By making the changes (2.17) in (3.16) and replacing σ^2 by 1 in (3.6), we then get:

$$V(t_i^*-t_0^*) = \frac{k}{Wrk-(W-W_1)(s-\lambda_1)} + \frac{rk}{Wrr_0-(W-W_1)(rs_0-\lambda_0r_0)} - \frac{k(W-W_1)(r\lambda_0-r_0\lambda_1)}{[Wrr_0-(W-W_1)(rs_0-\lambda_0r_0)][Wrk-(W-W_1)(s-\lambda_1)]}$$

(3.9)

The inter-block estimate alone $\tilde{t}_i - \tilde{t}_0$ of $t_i - t_0$ can be obtained from (3.7) by putting $W = 0$, $W' = 1$, provided $b > v + 1$.

It is

$$\tilde{t}_i - \tilde{t}_0 = \frac{kQ_{1i}}{s-\lambda_1} + \frac{kQ_0(r_0 + r\lambda_1 - r\lambda_0 - rs)}{(rs_0-\lambda_0r_0)(s-\lambda_1)} \quad (3.10)$$

We may now simplify (3.9) and (3.10), if desired, by using (2.6), (2.7). The variance of $\tilde{t}_i - \tilde{t}_0$ is obtained by putting $W = 0$ in 3.9.

Estimation of the weights *W*, *W₁*

To obtain the estimates of σ^2, σ_b^2 and hence those of W, W_1, we find the expected values of the variance sums of squares in the analysis of variance table. Those sums of squares correspond to the fixed block effect model but now we assume them to be random. The sum of squares are given in table (1) below.

Table 1: Analysis of variance table

source	df	ss	ss	df	Source
Blocks (unadjusted)	$b - 1$	$k^{-1}\underline{BB} - (bb)^{-1}g^2$	$(b-1)\,E_b$ by subtraction	$b-1$	Blocks (adjusted)
1- treatments (adjusted)	$(v + 1) - 1$	$\hat{\underline{t}}'\underline{Q}$	$\frac{T_0^2}{r_0} + \frac{\sum_{i=1}^{v} T_i^2}{r} - \frac{g^2}{bk}$	v	Treatments (unadjusted)
Intra-block error	$f = bk - b - v$	$F\,E_I$ by subtraction	$F\,E_I$ Carried over	f	Error
Total (corrected)	$bk-1$			$bk-1$	Total

From the set up (2.8), it follows that, when β_j are normally distributed

$$V(Y_{ij}) = \sigma^2 + \sigma_b^2 \qquad (4.1)$$

$$Cov(Y_{ij}, Y_{i',j'}) = \begin{cases} \sigma_b^2 & \text{if } j = j' \\ 0 & otherwise \end{cases} \qquad (4.2)$$

$$V(\beta_j) = k(\sigma^2 + k\sigma_b^2) \qquad (4.3)$$

$$Cov(\beta_j, \beta_{j'}) = 0, \quad j \neq j', \qquad (4.4)$$

$$E(Y_{ij}) = \mu + t_i, \qquad (4.5)$$

$$V(\underline{T}) = \sigma^2 D + \sigma_b^2 NN' \qquad (4.6)$$

Where D = diagonal $(r_0, r, r\ldots r)$ (4.7)

From these after some algebra, we obtain

$$E((b-1)E_b) = (b-1)\sigma^2 + (bk - s_0/r_0 - vs/r)\,\sigma_b^2 \qquad (4.8)$$

$$E(E_I) = \sigma^2 \qquad (4.9)$$

Solving for σ^2 and σ_b^2 and substituting in W, W_1, the following estimates of W and W_1 are obtained.

$$\hat{W} = \frac{1}{E_I}, \hat{W}_1 = \frac{(bk - s_0/r_0 - vs/r)}{(bk - s_0/r_0 - vs/r)E_I + k(b-1)(E_b - E_I)} \qquad (4.10)$$

Example:

Consider an experiment in which 6 treatments of varying level of protein, riboflavin and total food intake were to be compared in a nutrition experiment on rats, where level 0 indicates the control. It was possible to obtain litters with four male rats each. A balanced incomplete block design of $(v + 1) = 6$ treatments in b = 7 blocks of size $k = 4$ rats was used. Each test treatment was replicated $r = 4$ times except the control which is replicated $r_0 = 8$ times. The variable measured was micrograms of riboflavin per 100 ml of blood serum. The coded data is given in table (2). Treatment numbers are given in the upper right hand corner of each cell. The analysis of variance table is given in table (3). The treatment estimates, elementary treatment contrasts and their variances are summarized in tables (4, 5, and 6).

Table 2: Data for the example

(0) 10.462	(1) 11.242	(2) 15.283	(3) 14.706
(0) 8.651	(1) 12.307	(2) 10.223	(5) 18.261
(0) 3.905	(2) 14.364	(3) 12.298	(4) 12.618

(0) 9.494	(2) 16.522	(4) 14.732	(5) 13.131
(0) 4.632	(0) 3.112	(1) 13.298	(4) 12.322
(0) 8.157	(0) 9.102	(3) 15.298	(5) 14.968
(1) 14.624	(3) 16.945	(4) 16.773	(5) 16.428

T_0 57.515	T_1 51.471	T_2 56.392	T_3 59.359	T_4 56.445	T_5 62.970
Q_0 -32.581	Q_1 1.654	Q_2 6.797	Q_3 7.538	Q_4 7.600	Q_5 8.992
Q_{10} -8.231	Q_{11} 0.652	Q_{12} 0.432	Q_{13} 2.656	Q_{14} -0.32	Q_{15} 4.813

Table 3: ANOVA table for testing treatment effects

Source of variation	degrees of freedom	Sum of squares	Mean square error
Blocks unadjusted	6	140.431	
1) treatment component	5	129.660	
2) inter-block error	1	10.770	10.770
Treatment adjusted	5	221.466	
Intra-block error	16	70.346	4.397
Total	27	432.243	

Table 4: ANOVA table for testing block effects

Source of variation	degrees of freedom	Sum of squares	Mean square error
Blocks adjusted	6	52.403	
Treatment unadjusted	5	309.494	
Intra-block error	16	70.346	4.397
Total	27	432.243	

From tables (5 and 6) we conclude that treatment five is more effective than the other treatments when compared to the control. However, the reduction in variance due to recovery of inter-block information is very small.

Table 5: a- Intra-block model b- inter-block model

$\hat{t}_0$ = -6.156	$\hat{t}_1 - \hat{t}_0$ = 5.127	$\tilde{t}_0$ = 4.058	$\tilde{t}_1 - \tilde{t}_0$ = 9.539
$\hat{t}_1$ = -1.389	$\hat{t}_2 - \hat{t}_0$ = 6.596	$\tilde{t}_1$ = 13.597	$\tilde{t}_2 - \tilde{t}_0$ = 9.095
$\hat{t}_2$ = 0.080	$\hat{t}_3 - \hat{t}_0$ = 6.807	$\tilde{t}_2$ = 13.153	$\tilde{t}_3 - \tilde{t}_0$ = 13.547
$\hat{t}_3$ = 0.291	$\hat{t}_4 - \hat{t}_0$ = 6.827	$\tilde{t}_3$ = 17.605	$\tilde{t}_4 - \tilde{t}_0$ = 7.594
$\hat{t}_4$ = 0.311	$\hat{t}_5 - \hat{t}_0$ = 7.223	$\tilde{t}_4$ = 11.652	$\tilde{t}_5 - \tilde{t}_0$ = 17.859
$\hat{t}_5$ = 0.707		$\tilde{t}_5$ = 21.917	

a- var ($\hat{t}_i - \hat{t}_0$) = 1.884 b- var ($\tilde{t}_i - \tilde{t}_0$) = 20.074

Table 6: Combined inter and intra-block estimates

t_0^* = -1.079	$t_1^* - t_0^*$ = 5.385
t_1^* = 0.103	$t_2^* - t_0^*$ = 6.743
t_2^* = 0.394	$t_3^* - t_0^*$ = 7.203
t_3^* = 0.493	$t_4^* - t_0^*$ = 6.870
t_4^* = 0.421	$t_5^* - t_0^*$ = 7.840
t_5^* = 0.630	

var($t_i^* - t_0^*$) = 1.856

Using equation (4.10), we get:

$\hat{W} =$ 0.227 And $\hat{W}_1$ = 0.099

REFERENCES

Bateman G. l., Horn D., Payne W. R., Nicholls P.H.(1994); "Evaluation of fungicides applied to soil to control naturally occurring take-all using a balanced-incomplete-block design and very small plots"; *Annals of Applied Biology*; 124: 241-251.

Bechhofer R.F. and Tamhane A.C. (1981); "Incomplete block designs for Comp= aring treatment with a control: General theory"; *Technometrics*; 23: 45 -57.

John D. S. and Edwards D.(1986); "An asymptotically optimal subclass of balanced treatment incomplete block designs for comparisons with a control"; *Biometrika*; 73: 191 - 199.

Creaser J. I., Ward N. J., Rakauskas M. E., Shankwitz C., Boer E. R. (2009); "Effects of alcohol impairment on motorcycle riding skills"; *Accid Anal Prev*; 41: 906-13.

Klaus H., and Kempthorne O. (2005); "Design and analysis of experiments, advanced experimental design"; Wiley & Sons, NJ.

Trietsch J., van der Weijden T., Verstappen W., Janknegt R., Muijrers P., Winkens R., van Steenkiste B., Grol R., Metsemakers J.(2009); "A cluster randomized controlled trial aimed at implementation of local quality improvement collaborative to improve prescribing and test ordering performance of general practitioners: Study Protocol"; *Implement Sci*.; 17: 4-6.

مقارنة معالجة مع معالجة المراقبة باستخدام المعلومات المستخلصة من بين القطاعات في تصاميم القطاعات غير التامة المتزنة الممتدة

بدر الدين عبد الرحمن محمد

قسم علوم الرعاية الصحية، كلية العلوم الطبية التطبيقية ، جامعة الملك سعود

الخلاصة

في كثير من تجارب تصاميم القطاعات غير التامة يود الباحث مقارنة المعالجات مع معالجة المراقبة. وتكرر هذه المقارنات أكثر من مرة لكل قطاع للحصول علي مقارنات أكثر دقة. في هذه الورقة تم تناول مثل هذا النوع من تحليل التصاميم باستخدام المعلومات المستخلصة من بين القطاعات. وأيضاً تم تناول تقديرات الأوزان للنموذج الموحد بين القطاعات و داخل القطاعات