UMYU Scientifica, Vol. 2 NO. 2, June 2023, Pp 001 – 006
https://doi.org/10.56919/usci.2223.001
ISSN: 2955 – 1145 (print); 2955 – 1153 (online)
A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina

ORIGINAL RESEARCH ARTICLE

A Statistical Model for the 2-Part Balanced Incomplete Block Design.

Muhammad Naziru Isah [1*], Garba Jamilu[2] and Abdulkarim Muhammad[3]
[1,2]Department of Statistics, Ahmadu Bello University, Zaria
[3]Department of Computer Science, Ahmadu Bello University, Zaria

ARTICLE HISTORY
Received April 06, 2023
Accepted May 21, 2023
Published June 20, 2023

KEYWORDS
Balanced Incomplete Block Design, Additive Model, ANOVA table, Model Parameters, Orthogonal Design, Adjusted Sum of Squares

© The authors. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/)

ABSTRACT
Multi-part Balanced Incomplete Block Designs combine many orthogonal balanced incomplete block designs in the same block. The design has been used to study cancer trials with medical centres as blocks. Various design layouts have been constructed for a restricted number of cancer types and drugs. However, parameter estimation, hypothesis testing, and model building for the constructed design layouts have not been considered. Based on the preceding, we proposed an additive model for the design (2-part Balanced Incomplete Block Designs) and estimated its parameters using the least square method. The parameter estimation precipitated the formation of the ANOVA table which paved way for the hypothesis testing procedure.

INTRODUCTION

A balanced incomplete block design (BIBD) is an incomplete block design in which c blocks have the same number of s plots; each treatment is replicated r times in the design. More so, each treatment occurs at most once in a block, and every pair of treatments occurs together λ times in the c blocks. The design is used to improve the efficiency and precision of the estimates when some fraction of treatments are allocated into blocks of size $s < t$ (Alawode *et al.*, 2012). Balance is achieved if every treatment occurs an equal number of times in the design and each pair of treatment occurs together an equal number of times across all the blocks in the design (Mason *et al.*, 2003).

Multi-part Balanced Incomplete Block Design combines many orthogonal balanced incomplete block designs in the same block (Bailey and Cameron, 2019). The design has been used in cancer trial experiments, allowing each medical center to treat only a restricted number of cancer types and a restricted number of drugs. Under this setup, various design layouts has been constructed Bailey and Cameron, (2019). Luna et al., (2022) utilized 2-part design to compared the efficiencies of Orthogonal array composite designs relative to the central composite design when there are a few missing observations from one factorial point and one additional point. Karmakar et al. (2023) used two-part designs in selecting the best possible components in integrated farming system that involve two groups of treatment arranged in incomplete blocks with respect to both groups, and the concurrence of treatment pairs within and between groups is constant. However, no standard statistical procedure has been proposed for the constructed layouts, such as parameter estimation, hypothesis testing and model building. This work therefore, considered 2-part Balanced Incomplete Block Design (2pBIBD), a combination of two orthogonal BIBDs in the same block and proposed a suitable statistical model for the design, estimated the model parameters and constructed the ANOVA table for inference-making.

DESIGN LAYOUT

Suppose we want to perform an experiment with two factors say factor 1 and factor 2, each at t_1 and t_2 levels respectively (i.e t_1 treatments of factor 1 and t_2 treatments of factor 2). Suppose also that the experiment involved blocks that cannot accommodate each factor's levels. This experiment can be arranged in an incomplete block design manner, and can be called 2-part balanced incomplete block design (2pBIBD) if the following conditions are satisfied; (Bailey and Cameron, 2019).

1. Each block can only accommodate s_1 of the t_1 treatments of factor 1 and s_2 of the t_2 treatments of factor 2. Where $s_1 < t_1$, $s_2 < t_2$ and $t_1 \geq t_2$

Correspondence: Muhammad N. I. Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.
✉ nimuhammad@abu.edu.ng Phone Number; +234 703 279 9771
How to cite: Muhammad N. I., Garba J., and Abdulkarim M. (2023). A Statistical Model for the 2-Part Balanced Incomplete Block Design. UMYU Scientifica, 1(2), 173 – 179. https://doi.org/10.56919/usci.2223.001

2. Each pair of treatment of factor 1 appear together in the same block λ_1 times. Also, each pair of treatment of factor 2 appear together in the same block λ_2 times
3. Each treatment from factor 1 appear with every treatment of factor 2 λ_{12} times, that is, the treatment of factor 1 and 2 are orthogonal

The layout of such design is given in Table 1, with $t_1 = 6$; $t_2 = 5$; $s_1 = 3$ and $s_2 = 2$.

Table 1: Concise form of the layout

Block	Cancer types	Drug
1	T1, T2, T3	V1, V5
2	T1, T5, T6	V1, V2
3	T1, T3, T4	V2, V3
4	T1, T2, T6	V3, V4
5	T1, T4, T5	V4, V5
6	T2, T4, T5	V1, V3
7	T2, T3, T5	V2, V4
8	T3, T5, T6	V3, V5
9	T3, T4, T6	V1, V4
10	T2, T4, T6	V2, V5

Source: (Bailey and Cameron, 2019)

This layout is presented in Bailey and Cameron (2019); the design is for six cancer types and five drugs; each block (medical centre) can only accommodate 3 cancer types and 2 drugs. The layout can be presented in a different form as in Table 2, see (Bailey and Cameron, 2019)

Details on how to come up with minimum number of blocks given different number of treatments and block sizes for the two factors is found in Bailey and Cameron (2019).

Table 2: Layout of 2-Part Balanced Incomplete Block Design

Block	Cancer					
	T1	T2	T3	T4	T5	T6
1	V1, V5	V1, V5	V1, V5			
2	V1, V2				V1, V2	V1, V2
3	V2, V3		V2, V3	V2, V3		
4	V3, V4	V3, V4				V3, V4
5	V4, V5			V4, V5	V4, V5	
6		V1, V3		V1, V3	V1, V3	
7		V2, V4	V2, V4		V2, V4	
8			V3, V5		V3, V5	V3, V5
9			V1, V4	V1, V4		V1, V4
10		V2, V5		V2, V5		V2, V5

Source: (Bailey and Cameroon, 2019)

THE PROPOSED MODEL

This research proposed an additive model with fixed treatments and blocks.

$$y_{ijk} = \mu + \alpha_{(i)} + \beta_{(j)} + \gamma_k + \varepsilon_{ijk}, \qquad (1)$$

Where:

y_{ijk} is the observation in block k that receive treatment i of factor 1 and treatment j of factor 2

μ is the grand mean

$a_{(i)}$ is the treatment i effect of factor 1

$\beta_{(j)}$ is the treatment j effect of factor 2

γ_k is the block k effect

ε_{ijk} is the random error associated with observation y_{ijk}

under the assumption that;

$$\sum_i^{t_1}\alpha_i = \sum_j^{t_2}\beta_j = \sum_k^{c}\gamma_k = 0, \quad \text{and } \varepsilon_{ijk} \sim NID(0,\sigma_\varepsilon^2)$$

where

$$i = 1,2\ldots,t_1, \quad j = 1,2\ldots,t_2, \quad \text{and} \quad k = 1,2\ldots,c$$

Note that the subscript i and j of treatments $a_{(i)}$ & $\beta_{(j)}$ where written in braces because the treatments depend on the block in the model (Mills, 2014)

The total number of observations N in the experiment is given by

$$\begin{aligned} N &= cs_1s_2 \\ &= t_1r_1s_2 \quad \text{or} \quad t_2r_2s_1 \end{aligned}$$

The design can be considered as a Balanced Incomplete Block (t_1, s_1, λ_1) Design replicated $s_2 - \text{times}$ or a Balanced Incomplete Block (t_2, s_2, λ_2) Design replicated $s_1 - \text{times}$.

Therefore all the relationships between the Balanced Incomplete Block Design parameters are satisfied.

ESTIMATION OF PARAMETERS

The model parameters were estimated using Least Square Estimation.

Given the model,

$$y_{ijk} = \mu + \alpha_{(i)} + \beta_{(j)} + \gamma_k + \varepsilon_{ijk},$$

We need to estimate the parameters to minimize the sum of squares error.

Let

$$\psi = \sum_i\sum_j\sum_k \varepsilon_{ijk}^2 = \sum_i\sum_j\sum_k (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k)^2. \quad (2)$$

To estimate the grand mean, we have

$$\frac{\partial\psi}{\partial\mu} = -2\sum_i\sum_j\sum_k (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k),$$

Setting the equation to zero gives

$$\sum_i\sum_j\sum_k (y_{ijk} - \hat{\mu} - \alpha_i - \beta_j - \gamma_k) = 0$$

$$\sum_i\sum_j\sum_k y_{ijk} - cs_1s_2\hat{\mu} - cs_2\sum_i\alpha_i - cs_1\sum_j\beta_j - s_1s_2\sum_k\gamma_k = 0$$

implying that

$$\hat{\mu} = \frac{\sum_i\sum_j\sum_k y_{ijk}}{cs_1s_2} = \bar{y}_{\ldots} \quad (3)$$

To estimate the block effect, we have

$$\frac{\partial\psi}{\partial\gamma_k} = -2\sum_i\sum_j (y_{ijk} - \mu - \alpha_i - \beta_j - \hat{\gamma}_k)$$

setting the equation to zero gives

$$\sum_i\sum_j (y_{ijk} - \mu - \alpha_i - \beta_j - \hat{\gamma}_k) = 0$$

$$\sum_i\sum_j y_{ijk} - s_1s_2\mu - s_2\sum_i\alpha_i - s_1\sum_j\beta_j - s_1s_2\gamma_k = 0$$

implying that

$$\gamma_k = \frac{\sum_i\sum_j y_{ijk}}{s_1s_2} - \hat{\mu} = \bar{y}_{..k} - \bar{y}_{\ldots} \quad (4)$$

Since $\alpha_{(i)}$ & it depend on the block in the model,

we need to rewrite the model in the form:

$$y_{ijk} = \mu + \gamma_k + \sum_{a=1}^{t_1}\delta_{ik}^a\alpha_a + \sum_{b=1}^{t_2}\delta_{jk}^b\beta_b + \varepsilon_{ijk}, \quad (5)$$

where δ_{ik} is the incidence plot ik, and

$$\delta_{ik}^{a} = \begin{cases} 1 & if \quad treatment \quad a \quad is \quad present \quad in \quad cell \quad ik \\ 0 & if \quad treatment \quad a \quad is \quad not \quad present \quad in \quad cell \quad ik \end{cases}$$

therefore

$$\sum_{i} \delta_{ik}^{a} = n_{ak} \ \& \ \sum_{i} \sum_{k} \delta_{ik}^{a} = r_1,$$

where n_{ak} is the incidence of occurrence of treatment *a* in block *k*.

Similarly, δ_{jk} is the incidence on plot *jk*, and

$$\delta_{jk}^{b} = \begin{cases} 1 & if \quad treatment \quad b \quad is \quad present \quad in \quad cell \quad jk \\ 0 & if \quad treatment \quad b \quad is \quad not \quad present \quad in \quad cell \quad jk \end{cases}$$

therefore

$$\sum_{j} \delta_{jk}^{b} = n_{bk}$$
$$\sum_{j} \sum_{k} \delta_{jk}^{b} = r_2.$$

where n_{bk} is the incidence of occurrence of treatment *b* in block *k*.

So we re-estimate the block effect as;

$$\frac{\partial \psi}{\partial \gamma_k} = -2 \sum_{i} \sum_{j} (y_{ijk} - \mu - \sum_{a=1}^{t_1} \delta_{ik}^{a} \alpha_a - \sum_{b=1}^{t_2} \delta_{jk}^{b} \beta_b - \gamma_k)$$

Setting the equation to zero gives

$$\sum_{i} \sum_{j} (y_{ijk} - \mu - \sum_{a=1}^{t_1} \delta_{ik}^{a} \alpha_a - \sum_{b=1}^{t_2} \delta_{jk}^{b} \beta_b - \gamma_k) = 0$$

$$\sum_{i} \sum_{j} y_{ijk} - s_1 s_2 \mu - s_2 \sum_{a=1}^{t_1} n_{ak} \alpha_a - s_1 \sum_{b=1}^{t_2} n_{bk} \beta_b - s_1 s_2 \gamma_k = 0$$

$$\hat{\gamma}_k = \frac{y_{\bullet\bullet k}}{s_1 s_2} - \hat{\mu} - \frac{\sum_{a=1}^{t_1} n_{ak} \hat{\alpha}_a}{s_1} - \frac{\sum_{b=1}^{t_2} n_{bk} \hat{\beta}_b}{s_2} \quad (6)$$

To estimate the treatment effect of factor 1, we have

$$\frac{\partial \psi}{\partial \alpha_i} = -2 \sum_{j} \sum_{k} (y_{ijk} - \mu - \hat{\alpha}_i - \beta_j - \gamma_k)$$

setting the equation to zero gives

$$\sum_{j} \sum_{k} (y_{ijk} - \mu - \hat{\alpha}_i - \beta_j - \gamma_k) = 0$$

$$\sum_{j} \sum_{k} y_{ijk} - r_1 s_2 \mu - r_1 s_2 \hat{\alpha}_i - r_1 \sum_{j} \beta_j - s_2 \sum_{k} n_{ik} \gamma_k = 0$$

where $n_{ik} = \begin{cases} 1 & if \quad treatment \quad i \quad is \quad present \quad in \quad block \quad k \\ 0 & if \quad treatment \quad i \quad is \quad not \quad present \quad in \quad block \quad k \end{cases}$

implying that

$$\begin{aligned} \alpha_i &= \frac{\sum_{j} \sum_{k} y_{ijk}}{r_1 s_2} - \hat{\mu} - \frac{s_2 \sum_{k} n_{ik} \hat{\gamma}_k}{r_1 s_2} \\ &= \frac{y_{i\bullet\bullet}}{r_1 s_2} - \hat{\mu} - \frac{\sum_{k} n_{ik}}{r_1} \left[\frac{y_{\bullet\bullet k}}{s_1 s_2} - \hat{\mu} - \frac{\sum_{a=1}^{t_1} n_{ak} \hat{\alpha}_a}{s_1} - \frac{\sum_{b=1}^{t_2} n_{bk} \hat{\beta}_b}{s_2} \right] \end{aligned}$$

$$\begin{aligned} \therefore r_1 s_1 s_2 \alpha_i &= s_1 y_{i\bullet\bullet} - \sum_{k} n_{ik} y_{\bullet\bullet k} + s_2 \sum_{k} n_{ik} (\sum_{a=1}^{t_1} n_{ak} \hat{\alpha}_a) + s_1 \sum_{k} n_{ik} (\sum_{b=1}^{t_2} n_{bk} \hat{\beta}_b) \\ &= s_1 y_{i\bullet\bullet} - \sum_{k} n_{ik} y_{\bullet\bullet k} + s_2 \left[\sum_{k} n_{ik}^{2} \hat{\alpha}_i + \sum_{k} n_{ik} \sum_{a \neq i}^{t_1} n_{ak} \hat{\alpha}_a \right] + s_1 \lambda_{12} \sum_{b=1}^{t_2} \hat{\beta}_b \\ &= s_1 y_{i\bullet\bullet} - \sum_{k} n_{ik} y_{\bullet\bullet k} + s_2 \left[r_1 \hat{\alpha}_i + \lambda_1 \sum_{a \neq i}^{t_1} \hat{\alpha}_a \right] \end{aligned}$$

$$\therefore \left(r_1 s_1 s_2 - s_2 r_1 + s_2 \lambda_1 \right) \alpha_i = s_1 \left[y_{i\bullet\bullet} - \frac{\sum_{k} n_{ik} y_{\bullet\bullet k}}{s_1} \right] = s_1 A_i,$$

where

$$A_i = y_{i\bullet\bullet} - \frac{\sum_{k} n_{ik} y_{\bullet\bullet k}}{s_1}, \quad (7)$$

with the additional condition that

$$\sum_{i} A_i = 0$$

so

$$s_2\left(r_1 s_1 - r_1 + \lambda_1\right)\alpha_i = s_1 A_i$$
$$s_2\left(r_1(s_1 - 1) + \lambda_1\right)\alpha_i = s_1 A_i$$
$$s_2(\lambda_1(t_1 - 1) + \lambda_1)\alpha_i = s_1 A_i$$
$$s_2\lambda_1 t_1 \alpha_i = s_1 A_i$$

$$\therefore \alpha_i = \frac{s_1}{s_2\lambda_1 t_1} A_i \qquad (8)$$

To estimate the treatment effect of factor 2, we have

$$\frac{\partial\psi}{\partial\beta_j} = -2\sum_i\sum_k (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k) \qquad (9)$$

setting the equation to zero gives

$$\sum_i\sum_k (y_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k) = 0$$

$$\sum_i\sum_k y_{ijk} - r_2 s_1\mu - r_2\sum_i \alpha_i - r_2 s_1\beta_j - s_1\sum_k n_{jk}\gamma_k = 0$$

where $n_{jk} = \begin{cases} 1 & if \;\; treatment \;\; j \;\; is \;\; present \;\; in \;\; block \;\; k \\ 0 & if \;\; treatment \;\; j \;\; is \;\; not \;\; present \;\; in \;\; block \;\; k \end{cases}$

implying that

$$\beta_j = \frac{\sum_i\sum_k y_{ijk}}{r_2 s_1} - \hat{\mu} - \frac{s_1\sum_k n_{jk}\hat{\gamma}_k}{r_2 s_1}$$
$$= \frac{y_{\cdot j\cdot}}{r_2 s_1} - \hat{\mu} - \frac{\sum_k n_{jk}}{r_2}\left[\frac{y_{\cdot\cdot k}}{s_1 s_2} - \hat{\mu} - \frac{\sum_{a=1}^{t_1} n_{ak}\hat{\alpha}_a}{s_1} - \frac{\sum_{b=1}^{t_2} n_{bk}\hat{\beta}_b}{s_2}\right]$$

$$\therefore r_1 s_1 s_2 \beta_j = s_2 y_{\cdot j\cdot} - \sum_k n_{jk} y_{\cdot\cdot k} + s_2\sum_k n_{jk}\left(\sum_{a=1}^{t_1} n_{ak}\hat{\alpha}_a\right) + s_1\sum_k n_{jk}\left(\sum_{b=1}^{t_2} n_{bk}\hat{\beta}_b\right)$$
$$= s_2 y_{\cdot j\cdot} - \sum_k n_{jk} y_{\cdot\cdot k} + s_2\sum_{a=1}^{t_1}\hat{\alpha}_a + s_1\left[\sum_k n^2_{jk}\hat{\beta}_j + \sum_k n_{jk}\sum_{b\neq j}^{t_2} n_{bk}\hat{\beta}_b\right]$$
$$= s_2 y_{\cdot j\cdot} - \sum_k n_{jk} y_{\cdot\cdot k} + s_1\left[r_2\hat{\beta}_j + \lambda_2\sum_{b\neq j}^{t_2}\hat{\beta}_b\right]$$

$$\therefore \left(r_2 s_1 s_2 - s_1 r_2 + s_1\lambda_2\right)\beta_j = s_2\left[y_{\cdot j\cdot} - \frac{\sum_k n_{jk} y_{\cdot\cdot k}}{s_2}\right] = s_2 B_j,$$

where

$$B_j = y_{\cdot j\cdot} - \frac{\sum_k n_{jk} y_{\cdot\cdot k}}{s_2}, \qquad (10)$$

with the additional condition that

$$\sum_j B_j = 0$$

so

$$s_1\left(r_2 s_2 - r_2 + \lambda_2\right)\beta_j = s_2 B_j$$
$$s_1\left(r_2(s_2 - 1) + \lambda_2\right)\beta_j = s_2 B_j$$
$$s_1(\lambda_2(t_2 - 1) + \lambda_2)\beta_j = s_2 B_j$$
$$s_1\lambda_2 t_2\beta_j = s_2 B_j$$

$$\therefore \beta_j = \frac{s_2}{s_1\lambda_2 t_2} B_j \qquad (11)$$

ESTIMATING THE SUM OF SQUARES

The sum square total is given as;

$$SST = \sum_{ijk}(y_{ijk} - \bar{y}_{\dots})^2 = \sum_{ijk} y^2_{ijk} - \frac{Y^2_{\dots}}{N}. \qquad (12)$$

The sum square block is given as;

$$SSB = s_1 s_2\sum_k \gamma^2_k = s_1 s_2\sum_k(\bar{y}_{\cdot\cdot k} - \bar{y}_{\dots})^2 = \frac{\sum_k y^2_{\cdot\cdot k}}{s_1 s_2} - \frac{Y^2_{\dots}}{N} \qquad (13)$$

The sum square treatment of factor 1 is given as;

$$SSF_1 = \frac{\lambda_1 t_1 s_2}{s_1}\sum_i \alpha_i^2 = \frac{s_1}{\lambda_1 t_1 s_2}\sum_i A_i^2 \qquad (14)$$

The sum square treatment of factor 2 is given as;

$$SSF_2 = \frac{\lambda_2 t_2 s_1}{s_2} \sum_i \beta_j^2 = \frac{s_2}{\lambda_2 t_2 s_1} \sum_j B_j^2 \quad (15)$$

Sum square error is obtained as;

$$SSE = SST - SSB - SSF_1 - SSF_2 \quad (16)$$

The ANOVA Table is given in Table 3

Table 3: ANOVA Table

Source of Variation	Degrees of Freedom	Sum of Square	Mean Square	F
Block	$c-1$	$\frac{\sum_k y_{\bullet\bullet k}^2}{s_1 s_2} - \frac{Y_{\bullet\bullet\bullet}^2}{N}$	$\frac{SSB}{c-1}$	
Factor 1	$t_1 - 1$	$\frac{s_1}{\lambda_1 t_1 s_2} \sum_i A_i^2$	$\frac{SSF_1}{t_1 - 1}$	$\frac{MSF_1}{MSE}$
Factor 2	$t_2 - 1$	$\frac{s_2}{\lambda_2 t_2 s_1} \sum_j B_j^2$	$\frac{SSF_2}{t_2 - 1}$	$\frac{MSF_2}{MSE}$
Error	$N - c - t_1 - t_2 + 2$	Difference	$\frac{SSE}{dfe}$	
Total	$N-1$	$\sum_{ijk} y_{ijk}^2 - \frac{Y_{\bullet\bullet\bullet}^2}{N}$		

CONCLUSSION

A new method of analyzing the 2-part Balanced Incomplete Block Design was developed by adopting the method in Mills (2014). An additive model with both treatments and block fixed was proposed; the unknown parameters in the model were estimated using the least square method. The sum of squares of the sources of variation in the design was also estimated, and finally, an ANOVA Table was presented for the given design.

REFERENCES

Alawode, O. A., Bamiduro, T. A., and Eludire, A. A. (2012). $(v+1, b+r+\lambda+1, r+\lambda+1, k, \lambda+1)$ Some balanced incomplete block designs (BIBDs) from lotto designs (LDs). *International journal of mathematical and computational sciences*, 6(8):1194–1197. **[Crossref]**

Bailey, R. A. and Cameron, P. J. (2019). Multi-part balanced incomplete-block designs. *Statistical papers*, 60(2):55–76. **[Crossref]**

Karmakar, S., Varghese, C., Harun, M., Vinaykumar, L. N., & Varghese, E. (2023). Experimental designs for the selection of integrated farming system components. *Current science*, 124(9), 1053-1057. **[Crossref]**

Luna, J., Jaynes, J., Xu, H., & Wong, W. K. (2022). Orthogonal array composite designs for drug combination experiments with applications for tuberculosis. *Statistics in medicine*, 41(17), 3380-3397. **[Crossref]**

Mason, R. L., Gunst, R. F., and Hess, J. L. (2003). *Statistical design and analysis of experiments: with applications to engineering and science*. John Wiley & Sons. ISBN: 9780471372165. Retrieved from online library

Mills, S. E. (2014). *STAT5505 Design of Experiments*. Citeseer, 5203 Herzberg Bldg., School of Mathematics and Statistics, Carleton. Retrieved from: mathstat