On Two Methods of Analysing Balanced Incomplete Block Designs

Acha, Chigozie Kelechi
Department of Statistics
College of Natural and Applied Sciences
Michael Okpara University of Agriculture
Umudike Abia State, Nigeria
specialgozie@yahoo.com

Abstract

The paper briefly discusses the balanced incomplete–block design (BIBD's) and further compares two methods of analyzing them-the classical and vector space analysis of variance (ANOVA) methods. These methods are applied differently to the data arising from the balanced incomplete block designs (BIBD's). The basic interest is to compare the performance of the two methods of analysis on the available data from National Root Crops Research Institute (NRCRI) Umudike, Abia State, Nigeria. To achieve this, we shall consider treatment (adjusted), block (adjusted) treatment (not adjusted) in the classical ANOVA method and the vector space ANOVA method. Block is adjusted to know if the experiment is symmetric balanced incomplete block design (SBIBD). Though both methods are statistically significant, the classical method gives a minimum variance when compared with the variance of the vector space. Also the classical ANOVA method was easier to compute and more convenient to handle than the vector ANOVA method. The classical ANOVA method is found to be preferable to the vector space ANOVA method in this research work.

Keywords: Treatments, Blocks, Experiment, Homogeneity, Replication.

1.0 Introduction

The basic concepts of the statistical design of experiments and data analysis were discovered in the early part of the 20th century as a cost effective research design tool to help improve yields in farming. Since then, many types of designs experiments and analysis techniques have been developed to meet the diverse needs of researchers and engineers. One of such experimental designs is called the block design. A block design is a set together with a family of subsets (repeated subsets are allowed at times) whose members are chosen to satisfy some set of properties that are deemed useful for a particular application. These applications come from many areas, including experimental design, finite geometry, software testing, cryptography, and algebraic geometry. Many variations have been examined, but the most intensely studied are the balanced incomplete block designs (BIBDs) which historically were related to statistical issues in the design of experiments. There exist several studies carried out in the area of the balanced incomplete block design. Yates (1936), Bose (1939), Fisher (1940) and Bose (1949) have extensively worked and concluded that an incomplete block design with treatment, each replicated r times in b blocks of size k, is said to be group divisible (GD) if the treatment can be divided into m groupseach with t treatments, so that the treatment belonging to the same group occurs together in λ_1 blocks and treatment belonging to different group occurs together in λ_2 blocks.

Though, complete block design is widely used by researchers more than the balanced incomplete block designs, because the missing data are computed before analysis. In this work, two methods of analyzing balanced incomplete-blocks designs are reviewed for comparison: Classical and the Vector space analysis-of-variance methods. Although, classical ANOVA method can be used to analyze BIBD without computing the missing data (see, Gomez, 1983; Montgomery, 1976; Cochran and Cox, 1967); none has tried to compare the vector space ANOVA method with the classical method with respect to the minimum variance of the methods and also the ease of their computation. The objective of this study therefore is to examine which of them is better method of analyzing SBIBD with reference to minimum variance and ease of computation.

Hypothesis

Ho: MSE (C) = MSE (V)

Where;

MSE (C) = Mean square error for the approach

MSE (V) = Mean square error for vector space approach

2.0 The Philosophy of the Paper

Since the basic concepts of the statistical design of experiments and data analysis were discovered in the early part of the 20th century as a cost effective research design tool to help improve yields in farming, many types of design experiments and analysis techniques have been developed to meet the diverse needs of researchers and engineers. This relevance of SBIBDs underscores the interest it has generated among researches in recent times, which could be explained by its application in various areas especially in information theory (Lee, Yoo, Park and Chun, 2006). In pursuit of this interest, a new generalized expression on parameters of SBIB designs has been obtained and all the symmetric balanced incomplete block (SBIB) designs have been characterized (see Mohan, Kageyama & Nair, 2004, Ionin, 1998 & 1999 and Mohan, 1999).

A complete-block design is one of the most widely used designs. It is used when it is possible to run all the treatment combinations in each book. In situations where it is not possible to run all the treatment combinations in each block due to shortage of experimental apparatus or facilities or the physical size of the block, an incomplete-block design is used: see, for example, Acha (2010), Montgomery (1976). An incomplete-block design is balanced if there exists, a constant λ such that $\lambda_{ij} = \lambda$ for all i and j where λ is the number of times each pair of treatments appear together in the treatment (see Street and Street 1987).

However, the philosophy of this paper briefly discusses the balanced incomplete-block designs (BIBD's) and further compares two methods of analyzing them, here called the classical and vector space analysis-of-variance methods. The Classical and Vector space analysis-of-variance methods involve a statistical method for making simultaneous comparisons between two or more means. These methods were differently applied to data collected and were represented in various analysis-of-variance tables. The case study data employed here were obtained from the National Root Crops Research Institute (NRCRI) Umudike, Abia State, Nigeria. The data are on varieties of cassava with different ratios of Nitrogen, Phosphorous and Potassium (NPK). The data are secondary data, whose entries are ab initio arranged as BIBD by the experimenter.

3.0 Symmetric Balanced Incomplete-Block Design

Here, we assume that there are t treatments and b blocks. In addition, we assume each block contains k treatment each of which occurs r times in the design (or is replicated r times), and that there are N = tr = bk total observations; see, for example Montgomery (1976). Furthermore, the number of times each pair of treatments appears in the same block is

$$\lambda = \frac{r(k-1)}{t-1}$$

If t = b, the design is said to be symmetric. The parameter λ must be an integer. A design is said to be symmetric when the row and the column effects of the design are interchanged and still the error sum of squares so calculated in each remains the same, see, Montgomery (1976).

- **The Classical ANOVA:** In statistics, analysis of variance is a collection of statistical models and their associated procedures which compare means by splitting the overall observed variance into different parts (see Hill and Williams 2005). The analysis of variance for classical ANOVA is summarized in Table 1, Tables 2 and Table 3.

Table 1: Analysis of Variance for the Balanced Incomplete Block Design when Treatment is Adjusted

Sources of Variance	Sum of square	Degree of Freedom	Mean Square	F_0
Universal Mean	$\frac{(Y)^2}{N}$	1	-	
Treatment (adjusted)	$\frac{k\sum Q_i^2}{\lambda t}$	t- 1	$\frac{SS_{Treatment\ (adjusted)}}{t-1}$	$F_o = \frac{MS_{Treatment(adjustment)}}{MS_E}$
Blocks	$\frac{\sum Y_j^2}{k} - \frac{Y^2}{N}$	b-1	$\frac{SS_{Block}}{b-1}$	
Error	SS_E (by subtraction)	N – t – b – 1	$\frac{SS_E}{N-t-b-1}$	
Total	$\sum_i \sum_j Y_{ij}^2$	N		

Source: Montgometry (1976), Chapter 6.

Table 2: Analysis of Variance for the Balanced Incomplete Block Design when Treatment is not Adjusted

Sources of variation	Sum of squares	Degree of freedom	Mean Square	F_0
Universal Mean	$\frac{(Y)^2}{N}$	1	-	
Blocks	$\frac{\sum Y_j^2}{k} - \frac{Y^2}{N}$	b- 1	$\frac{SS_{Block}}{b-1}$	$F_0 = \frac{MS_{Block}}{MS_E}$
Treatment	$\frac{\sum Y_i^2}{k} - \frac{Y^2..}{N}$	t – 1	$\frac{SS_{Treatment}}{t-1}$	
Error	SS_E (by subtraction)	N – t – b - 1	$\frac{SS_E}{N-t-b+1}$	
Total	$\sum_i \sum_j Y_{ij}^2$	N		

Source: Montgometry (1976), Chapter 6

Table 3: Analysis of Variance for the Balanced Incomplete Block Design when Block is Adjusted

Sources of variation	Sum of squares	Degree of freedom	Mean Square	F_0
Universal Mean	$\frac{(Y)^2}{N}$	1	-	
Blocks(adjusted)	$\frac{r\sum_{j=1}^{b}(Q_j^i)^2}{\lambda b}$	b- 1	$\frac{SS_{Block(adjusted)}}{b-1}$	$F_o = \frac{MS_{Block(adjusted)}}{MS_E}$
Treatment	$\frac{\sum Y_i^2}{k} - \frac{Y^2..}{N}$	t – 1	$\frac{SS_{Treatment}}{t-1}$	
Error	SS_E	N – t – b - 1	$\frac{SS_E}{N-t-b+1}$	
Total	$\sum_i \sum_j Y_{ij}^2$	N		

Source: Montgometry (1976) Chapter 6

➢ **The Vector space Method**: A set, V is a vector space over a field, F (for example, the field of real or of complex numbers) if given.

- An operation, vector addition defined in V, denoted v + w (where v, w$\in$ v) and
- An operation scalar multiplication in V, denoted a x v (where v$\in$V and a$\in$F)

The total variation of vector space summarized in the Table 4 and Table 5

Table 4: The v-spaces analysis of variance

SYMBOL	Name	Vector Space (v)	DIM	Fitted value = Projection of y onto v-space	//Fit/I2 = Crude SS
U	Mean or Universal factor	V_u	I	$Y_u^i (overallmean)$	$Total^2$
Ca	Mean effect of Cassava	V_{ca}	N_{ca}	$Y_{ca}^i\ (cassava\ mean)$	$\sum_{Cassava}^{N} \frac{(cassava\ Total)^2}{t}$
F	Mean effect of fertilizer	V_f	N_f	$Y_f^i (fertilizer\ mean)$	$\sum_{fertilizer\ level}^{Level} \frac{(fertilizer\ Total)^2}{b}$
E	F –by – Ca interaction	V	N	$Y_E^1 = Y$	$\sum Y_i^2$

Table 5: The w-spaces analysis of variance

SYMBOL	W-SPACE	DIM	EFFECT	//EFFECT//2=SS
U	$W_u = V_u$	1	$Y_u = Y_u^i$	$CSS_u = SS_u$
Ca	$W_{ca} = V_{ca} \cap V_u 1_{-}$	$N_{ca} - 1$	$Y_{ca} = Y_{ca}^i - Y_u$	$SS_{ca} = CSS_{ca} - SS_u$
F	$W_f = V_f \cap V_u 1_{-}$	$N_f - 1$	$Y_f = Y_f^i - Y_u$	$SS_f = CSS_f - SS_u$
E	$W_E = (V_{ca} \cap V_f) 1_{-}$	By subtraction	$Y_E = Y - Yca - Y_f - Y_u$	$SS_E = CSS_E - SS_{ca} - SS_f - SS_u$

- **Notations**

GD= group divisible, **b**=block, **r**=replication, **k**=block size, **t**=treatments, **m**=number of groups, λ_1=first group of blocks, λ_2=second group of blocks, **C**=classical, **V**=vector, **N**=total, **U**=universal mean, **F**=fertilizer, **Ca**=cassava, **SS=**sum of squares **ANOVA**=analysis of variance, **SBIBD**=symmetric balanced incomplete block design, **MSE**= mean square error.

4.0 Results On The Classical And The Vector Space Analysis-Of-Variance Methods

a) The classicalanalysis-of-variancemethods

Experiment A

A balanced incomplete-block design (BIBD) that is symmetric was used to study the yield on four varieties of cassava with four different rates of NPK. These rates were administered in addition to the natural manure. The data collected from the experiment are tabulated in the layout of table 6.

Table 6: Qualitative layout of the experiment

Block 1	B	C	D
Block 2	A	C	D
Block 3	A	B	C
Block 4	A	B	D

Table 7: Raw data for the analysis of variance for experiment

Block (cassava)					
Treatment (fertilizer)	1	2	3	4	Yi
1	2		20	7	29
2		32	14	3	49
3	4	13	31		48
4	0	23		11	34
Yi	6	68	65	21	Y.(160)

Source: NCRP under NRCRI Umudike (1992) entries sponsored by IFAD (World Bank)

The data in table 7 for the fertilizer experiment is a symmetric balanced incomplete block design and its analysis of variance is shown in tables 8, 9 and 10.

Table 8: Analysis of variance when treatment is adjusted for experiment

Sources of variation	Sum of square	Degrees freedom	Mean squares	F-ratio	F-test at 0.01	F-test at 0.05
Universal mean	2133.33	1				
Treatment (adjusted)	880.83	3	293.61	4.04	3.62	3.41
Block	100.67	3	18.33			
Error	363.17	5	72.63			
Total	3478	12				

Since $F_{cal} \geq F_{tab}$ at different levels of probability, we conclude that the fertilizer applied has a significant effect on the cassava.

Table 9: Analysis of variance when block is adjusted for experiment

Sources of variation	Sum of squares	Degrees of freedom	Mean squares
Universal mean	2133.33	1	
Treatment	975.34	3	
Blocks (adjusted)	6.17	3	
Error	363.17	5	72.63
Total	3478	12	

Table 10: Analysis of variance when treatment is not adjusted for experiment

Sources of Variation	Sum of squares	Degree of Freedom	Mean Square	F-ratio	F-test at 0.01	F-test At 0.05
Universal mean	2133.33	1				
Treatment	975.34	3	325.11	6.05	3.62	3.41
Blocks	100.67	3				
Error	268.66	5	53.732			
Total	3478	12				

The analysis of variance in Table10 shows that $F_{cal} \geq F_{tab}$ at different levels of probability, we conclude that the fertilizer applied has a significant effect on the cassava.

b) The vector space analysis-of-variance methods

Table 11: Original data of the experiment

TREATMENT	1	2	3	4	TOTAL
1	2		20	7	29
2		32	14	3	49
3	4	13	31		48
4	0	23		11	34
TOTAL	6	68	65	21	160

To give the V-space and the W-space, we calculate the crude sum of square for the V-space and the actual sum of squares for the W-space, such that we have in summary.

Table 12: Vector space analysis of variance result for experiment 1b

SYMBOL	W-SPACE	DIM		EFFECT			$\|EFFECT\|^2$	MEAN SQUARES	F-CAL	F-TEST 0.05	F-TEST 0.01
				72.5	72.5	72.5					
M	W_μ	1	72.5	72.5	72.5		2133.33				
			72.5	72.5		72.5					
			72.5		72.5	72.5					
F	Wr	3		72.67	72.67	72.67					
			53.5	53.5	53.5		975.34	325.11	1089	3.86	6.99
			72	72		72					
			74		74	74					
Ca	Wca	3		74.67	72.67	73.67					
			69	74.67	72.67		100.67	29.85			
			69	74.67		73.67					
			69		72.67	73.67					
				74	71	73					
E	W	9	67	75	72		14.33	268.66			
			68	75		73					
			72		75	75					

Since $F_{ca}\geq F_{tab}$ at different levels of probability, we conclude that the fertilizer applied has a significant effect on the cassava.

Summarized results of the symmetric balanced incomplete block design

	EXPERIMENT	
	0.05	0.01
TREATMENT (NOT ADJUSTED)	SIGNIFICANT	SIGNIFICANT
TREATMENT (ADJUSTED)	SIGNIFICANT	SIGNIFICANT
VECTOR SPACE	SIGNIFICANT	SIGNIFICANT

5.0 Summary and Conclusion

Considering the results obtained from the experiment A, symmetric balanced incomplete-block design (SBIBD) used in this research work, the treatments were all significant at 0.05 and 0.01 levels of probability. Though the conclusion of both methods are the same (that is, statistically significant), the classical method gives a minimum variance when compare with the variance of the vector space.

However, other supportive observations were made as follows, that:

(i) the classical ANOVA method seems easier and more convenient to handle than the vector ANOVA method,

(ii) the result of the block adjustment conclude that experiment A is SBIBD and

(iii) the two methods can be used to analyze a symmetric balanced incomplete-block design (SBIBD).

Finally, this study concludes that the classical ANOVA method is more efficient in the analysis of SBIBD than the vector space ANOVA method.

References

1. Acha, C. K. (2010). Impact of Housing and Age of Birds on Egg Production. *International Journal of Numerical Mathematics*, 5(2): 275 – 298.
2. Barnes J. and Adams R. (2011). Wikipedia, File://A:/ Vector Space Definition and Much More From Answer Com.Htm.
3. Bose, R. C. (1939). On the Construction of Balanced Incomplete-Block Designs, *Annals of Eugenics*, 9: 353-399.
4. Bose, R. C. (1949). A Note on Fishers Inequality for Balanced Incomplete-Block Designs, *Annala of Math. Stat.* 20: 619-620.
5. Cochran, W. G and Cox, G. M. (1957). *Experimental Designs*, J. Wiley, New York.
6. Fisher, R. A. (1940). An Examination of the Different Possible Solutions of a Problem in Incomplete-Blocks, *Annals of Eugenics*, 10:52-75.
7. Gomez. A. K. and Gomez, A. A. (1983). *Statistical Procedures for Agricultural Research*, John Wiley and Sons, Inc, New York.
8. Hill D. H. and Williams W. C. (2011). *The America Heritage Dictionary of the English Language,* File://A:/ Analysis of Variance: Definition and Much More from Answer Com.Htm.
9. Ionin, Y.J. (1998). A Technique for Constructing Symmetric Designs, *Designs, Codes and Cryptography* 14: 147-158.
10. Ionin, Y.J. (1999). Building Symmetric Designs with Building Sets, *Designs, Codes and Cryptography* 17: 159-175.
11. Lee, O., Yoo, S., Park, B. and Chun, I. (2006). The Design And Analysis of An Efficient Load Balancing Algorithm Employing The Symmetric Balanced Incomplete Block Design, *Information Sciences,* 176: 2148-2160.

12. Mohan, R. N. (1999). A New Series of Affine μ-Resolvable (d+1)-Associate Class PBIB Designs, *Indian Journal of Pure and Applied Mathematics.* 30: 106-111.
13. Mohan, R. N., Kageyama, S. and Nair, M. M. (2004).On a Characterization of Symmetric Balanced Incomplete Block Designs. *Discussiones Mathematicae Probability and Statistics*, 24: 41-58.
14. Montgomery, D. C. (1976). *Design and Analysis of Experiments*, John Wiley and Sons, Inc, New York.
15. Street, A. I. and Street, D. J. (1987). *Combinatorics of Experimental Design*, Oxford University Press, Oxford.
16. Yates, F. (1936). Incomplete Randomized Blocks, *Annals of Eugenics*, 7: 121-140.