Classes of block designs and theirs application in the analysis of agricultural experiment organization

Dragan Randjelović, Ćemal Doličanin, Milan Randjelović
Agricultural faculty
University of Priština
38219 Lešak, Kosovo and Metohia
SERBIA

Abstract: - With the design $B(v,r_{\_1},...,r_{\_v},b,k_{\_1},...,k_{\_b},\lambda_{\{12\}},...,\lambda_{\{v-1,v\}})$,like one combinatorial configuration, can be presented organization of one experiment in which participate finite number v different elements some basic set V, which should organise in the b designs each with defined number $k_{\_j}$ (j=1,…,v) elements from this basic set, but so that every of this elements are exactly in r_i (i=1,…,b) designs and every pair of this different elements is in $\lambda_{\{v-1,v\}}$ designs. For the case $\boldsymbol{b}=\begin{pmatrix}\boldsymbol{v}\\ \boldsymbol{k}\end{pmatrix}$ we have so called balanced complete block-designs(BCBDs) and in other cases we have so called balanced incomplete block designs(BIBDs) which are more interesting for their application in determining the possibility of some experiment organization in many science discipline and also in agriculture and also lBIBDs are one of most effective one and more factorial experiment planes. In this paper are considered the conditions of presentation, construction and existence of some classes of BIBDs Because of very complicated prosedure for BIBDs statistical processing authors propose application of data more criterias analysis for anlalysis of experiment results. Three examples of their application in agricultural experiment organization is also given in the end of paper.

Key Words: BIBDs, Steiner system, symmetrical design, latin square, agricultural experiment organization, statistical analysis, data envelopment analysis

1. Blosk-design presenentation

Definition 1.1 Suppose that M is determinate or infinite set. Every set of subsets, which consists of elements of set M, is called configuration over the set M, it is marked with J and it is represented in form $J=\{S_1,S_2,\ldots,S_m\}$(every subset S_1 contains random number of elements).

Configuration can be graphically represented so that we assign a points of plane to the elements of the set $M=\{x_1, x_2,\ldots, x_n\}$ and every point, which belongs to set S_1, can be circled with a curve line. The configuration can be given with a graph where the points of plane (like elements of set M) are connected with appropriate circle in planes that represent subsets S_i.

Because of elaborated matrix-calculation the configuration is represented by so called: incident matrix. Suppose that $J=\{S_1,S_2,\ldots,S_m\}$ is configuration over set $M=\{x_1, x_2,\ldots, x_n\}$. For element x_j, j=1,2,…,n, we say that he is in incident with subset S_i i=1,2,…,m, if $x_j \in S_i$. Rectangular (0,1)-matrix $A=\{a_{ij}\}$, in form [m x n], which elements are defined with

$$a_{ij} = \begin{cases} 1 & \text{if } x_j \in S_i \\ 0 & \text{if } x_j \notin S_i \end{cases}$$

$\forall\, i = 1,2,\ldots,m \;\wedge\; \forall\, j =1,2,\ldots,n.$

It is called incident matrix of configuration J over set M. Because the configuration over set is given by subsets and because elements in those subsets are not organized and because schedule of this subsets in the configuration is not important, we can say that different notices of elements as in basic set, and also subsets in configuration are possible.It leads us to the conclusion that one configuration can be corresponded by more incident matrix. The question is: in which form we should give the configuration. The answer is very clear, we should change the configuration so that we get trivial form of incident matrix.

Definition 1.2 By design we mean any configuration $B=\{B_1,B_2,\ldots,B_b\}$ over finite set $V=\{a_1,a_2, \ldots,a_v\}$, where b i v are natural numbers.

Design can be defined like definite pair (V,B) where $V=\{a_1,a_2, \ldots,a_v\}$ is finite set of elements, and $B=\{B_1,B_2,\ldots,B_b\}$ set of subsets of different elements from V, or we can say set of block (we mean that $B_i \neq B_j$ for $i \neq j$).

Let we have design (V,B). For element a_j, $a_j \in V$ $j=1,2,\ldots,v$ we can say that it is incident to block B_i

, $B_i \in B$ $i=1,2,\ldots,b$ if $a_j \in B_i$. With k_j, $j=1,2,\ldots,b$ we will mark total number of elements a_i, $a_i \in V$ $i=1,2,\ldots,v$ which is incident to block B_{-j}.

Total number of blocks B_j, $j=1,2,\ldots,b$, incident to element a_i, $i=1,2,\ldots,v$, we will mark with r_i. With λ_{it} we will mark the total number of elements of the set

(that kind of a_i ,$a_t \in B_j$ for every $i=1,2,\ldots,v$ and $t=1,2,\ldots,v$, $i \neq j$). Because of undefinity of the elements of the blocks we can say that $\lambda_{it}=\lambda_{ti}$, so it is valid only to observe cases i< t. The numbers v, b, r_i, kj, λ_{it}, are called argument of given design.

The use of this kind of configuration in solving the combinatory problems is very complicated.

Because of that we will only observe balanced incompleted block-designs BIBDs, marked with (v,r,b,k, λ) - configurations over finite set V, where the set V consists of v mutual different elements and configuration B is made of b blocks, every block is made of exactly k< v elements from V, and every element from V appears in exactly r<b blocks and every pair of different elements from V appears in exactly λ blocks(see [1]-[3]).

2.Conditions for existence and construction of block-designs

2.1 Existence of block-designs

It can be proved that if one (v,r,b,k, λ)-configuration over finite set of elements exists, then two arguments are in the roundly servitude of permanent three, what will be explained in the following theorem which will be given without proof.

Theorem 2.1 If exists balanced design with arguments v,r,b,k,λ, over finite set of elements V i.e exists (v,r,b,k, λ) - configuration over finite set V, then next equalities are correct

(2.1) $bk = vr$

and

(2.2) $r(k – 1) = \lambda(v – 1)$.

Theorem 2.1 gives necessary but not enough conditions for existence of designs. Namely, if some of arguments v,r,b,k, λ satisfy relations from theorem 2.1, we are not sure that suitable configuration really exists: also, since the arguments are natural numbers, when we give three of them, sometimes it is not possible to define the other two only by using relations from theorem 2.1.

The large number of testing in existence of designs has been performed, thanks to the evolution of the computers, depending on some arguments that could satisfy conditions from theorem 2.1 it is decided :

- for large values of argument v, balanced incompleted block designs always exists and for small values it never does.
- for k = 3 and k = 4 theorem gives enough conditions for appropriate configuration, but not for k = 5.

Let us have some (v,r,b,k, λ)-configuration over finite set of elements V and we know its incident matrix $A=\{a_{ij}\}$, which is rectangular, in form [b x v].

Every its column contains r units, and every row contains k units. Scalar product of two mutual different vector-columns is equal to the number of appearance of the pair of different elements from V in configuration, namely it equals λ. The scalar product of any vector-column with its ownself equals argument r, namely equals the number of appearance of any elements V in design. Those obvious attributes of incident matrix of (v,r,b,k, λ)-configuration, enable us to get necessary and enough conditions of its existents.

Theorem 2.2 Let the rectangular (o,1) matrix $A=\{a_{ij}\}$, in form [b x v] is incident matrix of some (v,r,b,k, λ)-configuration over finite set of elements V. Then the next equality are correct

$$A^T A = (r - \lambda) I_v + \lambda J_v ,$$

(2.3) and

$$A J_{v \times 1} = k J_{b \times 1}.$$

Vice versa is also correct.

Unfortunately, for determining of the arguments using the theorem 2.2, it is necessary to solve suitable matrix equality and that is not possible without corresponding mathematics method which is not developed until now.

Therefore we must take interest in other criterion, if they are not too complicated, for determination necessary and enough conditions of existence of some design and also we can catch sight of some design category attribute, for which it is easier to ascertain conditions of existence. From the classes of this special BIBDS the most famous one is Steiner system that consists of BIBDs with $\lambda = 1$. Most interesting class known as Steiner triple system consists of BIBDs with k=3.

Theorem 2.3 Necessary and enough conditions for the existence of Steiner triple system are that v can be given in one of two forms:

v = 6t+1 or v = 6t+3 for t = 0,1,…

Theorem 2.4 Steiner triple system exists if and only if for the argument v to be valid next equality

(2.4) $v \equiv 1 (mod 6)$ or $v \equiv 3 (mod 6)$.

Definition 2.1 BIBDs with b=v (automatic follow k = r - see (2.1)) is called symmetrical design.

Theorem 2.5 If exists one (v,k,λ) design then :

a) if v is even then $k-\lambda$ is full square of some natural number;

b) if v is odd then equation $z^2=(k-\lambda)x^2 + (-1)^{\{(v-1)/2\}}\lambda y^2$ has in the set of natural numbers untrivial solutions for x,y,z.

Let us have some (v,r,b,k,λ)-configuration and b is divisible with r ($b \equiv 0 (mod\ r)$). Suppose that the blocks B_i of this configuration, $B=\{B_1,B_2,\ldots,B_b\}$, can be classified in r family $\beta 1, \beta_{2,\ldots,} \beta_r$, and in spite of the fact that two condition are satisfied :

(2.6)The blocks from each family β_i , i=1,2,…,r , in the union consist of all elements of V and that exactly one time.

(2.7)All the blocks which belong to one family β_i , i=1,2,…,r are mutual disjunctive.

Definition 2.2 Given (v,r,b,k,1)-configuration which blocks B_i , i=1,2,…,b, can be groped in r family, and satisfy condition (2.6) i (2.7) , is called solvable.

Remark 2.1 Condition $b \equiv 0 (mod\ r)$, which is valid for solvable block-schemas, is causing and condition $v \equiv 0 (mod\ k)$.

Theorem 2.6 Each $(n^2, n+1, n^2+n, n, 1)$-configuration , if n is a natural nummber, is solvable.

Rectangle matrix order rXs which consists of elements of the set $M_n=\{x_1, x_2,\ldots, x_n\}$ $r \le n$, $s \le n$, with attribute that each element from this set appears maximum one time in each its row and column, is called Latin rectangle order rXs over the set of elements M_n.

In the case r=s Latin rectangle is called Latin square. Because of the better clearness of notions that we consider, we will (without reduction of generality) mark elements of the set with arabian numbers 1,2,…,n. Let us have Latin rectangle L

It is easy to see that the rows of the Latin rectangle are different permutations of the elements from the set M_n . Because of those elements from its columns, like subsets $M_n=\{x_1, x_2,\ldots, x_n\}$ i=1,2,…,n, form one configuration $J=\{S_1,S_2,\ldots,S_n\}$ over the set M_n , with attribute that each S_i , i=1,2,…,n, consists of r mutually diferent elements from M_n and we have each element from the set M_n that appears in exactly r subsets S_i. For r=n $J=\{M_n\}$. Reversed is not valid i.e. each configuration we can not uniform assign with one Latin rectangle then each configuration correspond to unempty set of Latin rectangles.

Theorem 2.7 Let us have Latin rectangle L order rXs, $r \le s \le n$, over the set of elements $M_n=\{x_1, x_2,\ldots, x_n\}$. If we mark with N(i) number of appearences of elements on i-th place in given Latin rectangle, this rectangle could be expanded in Latin square order nXn over the set M_n , only and only if $N(i) \ge r+s-n$ for each i, i=1,2,…,n. Taking connection between Latin rectangle i.e. square with combinatorial configuration enable their examination especially in the domain of existence, presentation and construction(see [5]-[8]).

2.2 Construction of block-designs

The third problem in relation with BIBDs is their construction (presentation and existence are already considered).

Definition 2.2 Let us have matrix A like one incident matrix some (v,k,λ) - symmetrical configuration. The technique of receiving a new so called "dual design" using A^T is practicly based on the change of role the blocks and elements in starting matrix A.

Definition 2.3 Let us have some (v,k,λ) - symmetrical configuration over the set V, which is defined with the blocks $B_1,B_2,\ldots,B_v$. We choose arbitrary B_v and form the new blocks $B_i^{'}=B_i \setminus B_v$ for i=1,2,...,v-1.Now with the blocks $B_i^{'}$ for i=1,2,...,v-1 we obtain a new so called "residual design" in relation with starting (v,k,λ) - symmetrical configuration.It is over the set $V'=V \setminus B_v$.

with following parameters $v'=v-k, b'=v-1, k'=k-\lambda, \lambda'=\lambda$.

Theorem 2.8 Let natural numbers v,k,λ satisfy equality $k(k-1)=\lambda(v-1)$ for $\lambda=1,2$. Then if exists (v-k,k,v-1,k-λ , λ) configuration it also exists and (v,k, λ) configuration which is residual in relation with starting design(see [2]).

3. Main results

The scientists make experiments to affirm their hypotheses or to choose the best from available possibilties. It is necessary before the beginning of one experiment to make its most effective plan and then to check its organization possibility. In agricultural experiments organization it is especially important because they can be repeated often only next year. Mathematical instruments for that making of most effective plan are good to search on the base of balanced incomplete block designs (BIBDs) while they can most effectively represent one or more factorial agricultural experiments in which participate finite number of different elements which should be organised in some blocks of defined number so that every of this elements are in smaller number blocks and every pair of this elements is in necessary number of blocks. But it is very comlicated to make statistical analysis for experiment planes based on the BIBDs and therefor author propose like better possiblity application of data more criterias analysis(see [4]).

3.1 Analysis of variance in agricultural experiments

Considering results from the experiments in natural science in meaning analysis of variance i.e. when we consider three or more different characteristics-tretments in one basis set of experiment units (for primer case - two characteristics is used arithmetic mean for data analysis), we can observe two elementary forms of experiments(see [8]):
The first form of experiments is that where we have only one criteria by clasification units. By those experiments, called experiment with total random distribution, total variation is devided into two components: variation beetwen and inside groups. First variation results from the use of different treatments and the other is a consequence of accidental swinging inside each sample.
Mathematical model analysis of variance for each case can be given in two different forms: like additive which we will use and multiplicative.

Mathematical model analysis of variance in experiments with total random distribution is given with following relation:

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

where: X_{ij} is random variable j-th unit and i-th treatment (i,j=1,2,…,n), α_i presents effect i-th treatment and ε_{ij} is random variation inside units.

These experiments are suitable because they can include the big number of treatments without limitation of repetition and the statistics analysis of variance is very simple.
The second form of experiments is with two or more criteria by clasification units and we will consider those which have two criteria and that first criteria is treatment and second criteria is restriction of experiments error with the set apart from experiments error expected system variation which exists beside the tretments in experiment units. Total variation then have beside influnce of treatments and variation inside units and third part and that is beside treatments influence and expected influence one other system variation and therefore exists random block-design which is the plan of experiment in which units beside treatments are grouped and by known contolled system variation. They are realised like random block-design with two subforms:
* first subform of block-designs, which have plane of complete random distribution so called balanced complete block design (BCBDs).
Like first we set apart homogenous groups i.e. blocks towards criteria of classification which don't result from treatments and afterwards is one tretment applied at each units from group.
So, we have such a number of units how much tretments and the number of repetition of tretment is equal of the number of groups (it is possible that we have menyfold repetition of tretment in some group but it is obligatory to make random disposition of tretments individually by units inside group).
For this plane mathematical model analysis of variance we can given with following relation

$$X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where: X_{ij} is random variable j-th block and i-th treatment (i=1,2,…,t; j=1,2,…,b), α_i presents effect i-th treatment, β_j presents effect j-th block

and ε_{ij} is random variation of basis set which has for middle 0 and σ^2.
Mathematical model analysis of variance for latin square is given with following relation:

$$X_{ij} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$

where: X_{ijk} is random variable i-th rows, j-th columns and k-th treatments (i,j,k=1,2,…,n), α_i presents effect i-th rows, β_j presents effect j-th columns, γ_k presents effect k-th treatments and ε_{ijk} is random variable basic set which has for middle **0** and σ^2.
The main problem for wide use of BCBDs and the Latin squares like the special form of BIBDs is the number of treatments (each treatment in each block(group)) which increase error of experiment. Because of that Latin squares over the order 12X12 are rarely used.(the most frequently in use are Latin squares order 5X5 to 8X8).
***** second subform of block-dsigns which have plane of incomplete random distribution so called balanced incomplete block design i.e. BIBDs are most effective but it is with difficulty to make one useful mathematical model because this must take in consideration and influence which cause in the blocks unrepresented criterions.
In morefactorial experiments in which the number of considered units make total number of combination without repetition from t^f where t is the number of treatments and f is the number of factors each subjected to this number of tretments the situation is and more difficult to make on useful mathematical. model. Because of that authors, for this planes of experiments, propose application mathematical apparatus so called more criterias data analysis for anlalysis of experiment results.

3.2 More criterias data analysis

More criteias data analysis methods are grouped about two basis methods:
data envelopment analisys (DEA) which is without heviness coefificients and preference ranking organization method for enricment evaluations (PROMETHEE) with heviness coefificients for considered units.
It is easy to make the table of criterias values from executed experiments and with the application of method of linear programming, whic is in the basis both methods of more criteias data analysis, today we can produce and suitable information support in the form of suitable software package.
Example 3.1 Compare effect of different phosphorus fortilizers on the basis equal quantity of P_2O_5. They are applied at the meadow like latin square experiment with 4x4 following treatments: A = superphosphate 17%, B = superphosphate +nitrogen, C = superphosphate +nitrogen + potassium, D = control, unfertilized. Trsults of experiments (5X5) = $25m^2$/kg are given with the table 1.

rows	column 1	col. 2	col. 3	col. 4	total
1	17,00/A	15,00/B	19,50/D	26,00/C	77,50
2	14,30/D	21,20/A	22,40/C	17,50/B	75,40
3	25,00/C	13,20/D	16,70/B	23,30/A	78,20
4	13,80/B	26,00/C	24,00/A	17,40/D	81,20
total	70,10	75,40	82,60	84,20	312,30
	Treatment A	Trea. B	Trea. C	Trea. D	
	17,00	15,00	26,00	19,50	
	21,20	17,50	22,40	14,30	
	23,30	16,70	25,00	13,20	
	24,00	13,80	26,00	17,40	
total	85,50	63,00	99,40	64,40	
$\overline{X}$	21,37	15,75	24,85	16,10	

Table 1

Analysis of variance is given with table 2.

source of variation	level of freedom	sum of quadrate	middle of quadrate	expected middle of quadrate
rows	3	4,32	1,440	$\sigma^2 + \frac{4}{3}\sum_{i=1}^{4}\alpha_i^2$
columns	3	32,19	10,730	$\sigma^2 + \frac{4}{3}\sum_{j=1}^{4}\beta_j^2$
treat-ments	3	231,04	77,013	$\sigma^2 + \frac{4}{3}\sum_{k=1}^{4}\gamma_k^2$
error	6	35,16	5,860	σ^2
total	15	302,71		

Table 2

Example {3.2} Let us have 6 new sorts of corn and they must be planted on 10 fields each fields divided on 3 lots; each of sorts is planted on 5 different fields so that each pair of sorts is planted on different field exactly once (twice).

It is visible that the possibility of experiment organization depends of existence one balanced incomplete block design
with argument $v=6, r=5, b=10, k=3, \lambda=1(2)$ i.e. {6,5,10,3,1(2)}-configuration over set V={1,2,3,4,5,6}.
Let us mark the sorts of corn with the number 1,2,3,4,5,6,the fields with B1,B2,B3,B4,B5,B6,B7,B8,B9,B10.
We conclude that the existing of this configuration is possible only for $\lambda = 2$ with using the following known equality :
from (2.1)
$bk=vr \Rightarrow 10 \cdot 3=6 \cdot 5=30$
and from (2.2)
$r(k-1)= \lambda (v-1) \Rightarrow 5 \cdot (3-1)=2 \cdot (6-1)=10$.
So,for $\lambda =2$ we can plant the sorts of corn on the next schedule:
B1={1,2,3},B2={1,2,5},B3={1,3,4},B4={1,4,6}, B5={1,5,6},B6={2,3,6},B7={2,4,5},B8={2,4,6}, B9={3,4,5},B10={3,5,6}.
Example {3.3} Let us examine the existence of one (36,7,42,6,1) BIBD, using technique of residual BIBDs construction. (36,7,42,6,1) BIBD satisfy the necessary conditions for existence which are given with (2.1) and (2.2). But symmetrical configuration, if exists, must have next parameters: $v'=b'=b+1=43$, $r'=k'=k+1=7$, $\lambda'=\lambda =1$, consequently (43,7,1)-configuration. On the basis of theorem 2.5 equation $z^2=6x^2-y^2$ must have untrivial solution for x,y and z and that is unpossible. So (43,7,1) symmetrical configuration, for which is given starting "residual design",
can't exist and because of that and given starting (36,7,42,6,1) design can't exist.

4. Conclusion

Although uneffective BCBDs in relation with BIBDs we mustunderline advantages of latin square, which is also BCBDs, which application reduce the error of experiment but on other side repetition of each treatment in each block at least once by the planes of random block designs and latin square limit their real using because the big number of treatments increase the error of experiment.Because of that today more and more application have BIBDs which don’t include in each block each treatment than provide equal precision in comparison some pairs. Because the application classical statistic mathematical apparatus for result analysis of BIBDs is very difficult, the authors for this planes of experiments, propose application mathematical apparatus so called more criterias data analysis for anlalysis of experiment results..

References

[1]M.Aigner,*Combinatorial Theory*, Springer, 1997.
[2]T.Beth et. al., *Design Theory*, Cambridge University Press, 1993.
[3]I.Z. Milovanović, et. al., *Diskretna matematika*, Univerzitet Niš, Pelikan Niš, 2000.
[4]A. Charnes et al., *data envelopment analysis:Theory, Methodology and Applications*, Kluwer Academic Publisher, Boston, 1995.
[5]D.M.Randjelović et. al., Existence one class of Steiner block - schemas and their application in the agricultural experiment organization, *MASSEE2003*, Borovets Bulagaria,2003.
[6]D.M. Randjelović et. al., One class of design and their application in the experiment organization, *XXXVII IOC of Mining and Metallurgy*, Bor Sebia,2004.
[7]D.M.Randjelović et. al., Steiner systems and symmetrical design application in the agricultural experiment organization, *III Congress of mathematicians of Macedonia*, Struga Macedonia, 2005.
[8]S.Hadzivuković, *Statistički metodi,* Univeryitet u Novom sadu, 1991