CIRCULAR DESIGNS

By M. N. Das

Institute of Agricultural Research Statistics, New Delhi

(Received: February, 1959)

1. Introduction

Since the introduction of quasi-factorial designs by Yates (1936 *a*), a large number of incomplete block designs has been evolved to suit different situations. The additions to the list are balanced incomplete block designs by Yates (1936 *b*), partially balanced incomplete block designs by Bose and Nair (1939), rectangular lattice designs by Harshbarger (1947), linked block designs by Youden (1951), chain block designs by Youden and Conner (1953), a class of designs in two plot blocks by Kempthorne (1953),* generalised balanced designs by Das (1957), reinforced designs by Giri (1957) and Das (1958), etc.

All these designs suffer from one common limitation that they are not available for every number of treatments having any desired level of replications. The Reinforced designs introduced by the author (Das, 1958) are free from these limitations if there is no objection to using two unequal numbers of replications for the treatments.

Through further investigation on the problem it has been possible to evolve a type of incomplete block designs which is available for every number of treatments for any block size. These designs have been discussed in the present paper.

2. Definition of the Design

Let there be n equal arcs on the circumference of a circle denoted in order by $a_1, a_2, \ldots a_n$. If we form bigger arcs of size, A_{ki} such that it is the sum of k consecutive small arcs starting with a_i, we shall have in all n such arcs for n different values of i. Now, if we identify a_i with a set (s_i) of m treatments such that the different sets (s_i) are mutually exclusive, then the contents of the n arcs, A_{ki}, will form an incomplete block design with n blocks, mn treatments, mk block size and k replications.

* The author is grateful to Dr. K. R. Nair for calling his attention to Kempthorne's paper and to a discussion of it in Nair and Kishen's (1958) paper.

If the individual treatments be denoted by t_{ij} ($i = 1, 2, \ldots n_i$, $j = 1, 2, \ldots m$) and the set of all the m treatments on the arc, a_i, be denoted by (a_{ij}), the block contents of the design come out as below.

Block Nos.				
1	(a_{1j})	(a_{2j})	$\ldots$	(a_{kj})
2	(a_{2j})	(a_{3j})	$\ldots$	$(a_{k+1,\, j})$
3	(a_{3j})	(a_{4j})	$\ldots$	$(a_{k+2,\, j})$
.	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .
n	(a_{nj})	(a_{1j})	$\ldots$	$(a_{k-1,\, j})$

Such designs can be obtained for any number of treatments and block sizes.

As the placement of the treatments on the arcs of a circle simplifies very much the definition and analysis of the designs, these have been called circular designs.

3. Method of Intrablock Analysis

On the model $y_{ijm} = h + t_{ij} + b_m + \epsilon_{ijm}$, where t_{ij} is the effect of the j-th treatment on the i-th arc a_i, b_m, the m-th block effect and ϵ_{ijm}, a random variable with variance, σ^2, the normal equations for estimating the treatments after eliminating the block constants come out as below:

$$kt_{ij} - \frac{1}{m}\sum_j t_{ij} - \frac{k-1}{mk}\sum_j (t_{i+1,\, j} + t_{i-1,\, j})$$

$$- \frac{k-2}{mk}\sum_j (t_{i+2,\, j} + t_{i-2,\, j})$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$- \frac{k-s}{mk}\sum_j (t_{i+s,\, j} + t_{i-s,\, j})$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$- \frac{1}{mk}\sum_j (t_{i+k-1,\, j} + t_{i-k+1,\, j}) = Q_{ij} \qquad (1)$$

where $\sum_j t_{i+s,\,j}$ denotes the sum of those m treatments which are on the small arc a_{i+s}, *i.e.*, on the s-th arc after the i-th arc, similarly, $\sum_j t_{i-s,\,j}$ is the sum of the m treatments which are on the s-th arc before the i-th arc, and Q_{ij} is the adjusted total of the t_{ij} treatment, *i.e.*, sum of yields or of other character under observation for the k plots in which the treatment t_{ij} is replicated *minus* the sum of the k mean values per plot for that character in the k blocks in which that treatment has occurred.

Denoting

$$\sum_j t_{i\pm s,\,j} \quad \text{by} \quad a_{i\pm s}, \quad a_{i+s} + a_{i-s} \quad \text{by} \quad x_{is}$$

and

$$\sum_j Q_{i+s,\,j} \quad \text{by} \quad Q_{i+s},$$

the equations at (1) can be written as

$$kt_{ij} - \frac{a_i}{m} - \frac{1}{mk}\sum_{s=1}^{k-1}(k-s)\,x_{is} = Q_{ij}. \tag{2}$$

Adding these equations over the m treatments on the i-th arc, we get

$$k\,(k-1)\,a_i - \sum_{s=1}^{k-1}(k-s)\,x_{is} = kQ_i. \tag{3}$$

Adding two such equations corresponding to the two arcs which are at a distance of p on either side of the i-th arc, we get

$$k\,(k-1)\,x_{ip} - \sum_{s=1}^{k-1}(k-s)\,\{x_{i\,(s+p)} + x_{i(s-p)}\} = kQ_i^p \tag{4}$$

where $Q_i^p = (Q_{i+p} + Q_{i-p})$ and p goes from 1 to d where $d = (n/2)$ or $(n-1)/2$ according as n is even or odd. When n is even, the last equation becomes the one corresponding to the $n/2$-th arc from the i-th arc.

A convenient way of writing these equations (4) is to write on a circle the different x_{ij}'s where j goes from 0 to $n-1$, x_{i0} being equal to $2a_i$ and $x_{i\,(n/2)}$ to $2a_{i+(n/2)}$ when n is even and further the x_{ij} with j greater than d being equated to x_{il} where $l = n - j$, whatever n may be. Now, the equations (4) can be obtained in terms of these x_{ij}'s by writing the equations exactly in the same manner in which equations (1) were obtained in terms of the t_{ij}'s written on the circle provided m be taken as unity. The equations corresponding to x_{i0} as also $x_{i(n/2)}$ when n is even are to be halved as they represent twice the equations of their treatments.

Corresponding to $d+1$, x_{ij}'s we shall thus get $d+1$ equations involving a_i and x_{ij}'s $(j = 1, 2, \ldots .\ d)$.

In each of these equations the sum of the coefficients of the different x_{ij}'s will be zero. In the equation of a_i the coefficient of $x_{i(k-1)}$ is -1, while in the equation of x_{i1} the coefficient of x_{ik} is -1 and it will not contain any x_{ij} where $j > k$. In this way the successive equations will have the highest value of j in x_{ij} as $k+1$, $k+2$, ... All the equations up to the one which contains x_{id} as the highest x_{ij}, as judged by the value of j, the property that the highest x_{ij} has the coefficient -1 will hold. This accounts for $d+2-k$ out of $d+1$ equations leaving $k-1$ more in which the highest x_{ij} may not have -1 as its coefficient. The set of equations (4) with p varying from 0 to $d+1-k$ will be called set A equations, while the other $(k-1)$ equations with p varying from $d+2-k$ to d will be called set B equations.

From the equations corresponding to $p = 1$, $x_{i(k-1)}$ can be eliminated with the help of the equation at (3) which may be called the equation with $p = 0$. Similarly, from the equation, $p = 2$, $x_{i(k-1)}$ and x_{ik} can be eliminated with the help of the equations for $p = 0$ and 1. In general, from the p-th equation all x_{ij}'s where j varies from $(k-1)$ to $(k+p-2)$ can be eliminated with the help of the equations corresponding to all the smaller values of p.

Thus, we shall get another set of $d+2-k$ equations from the set A such that each of them contains all the x_{ij}, j varying from 0 to $k-2$ with only one more x_{ij} with the coefficient -1 where j is greater than $k-2$ and takes different values in the different equations.

Denoting this set of equations by set A_1, they can be written as:

$$k(k-1)\,a_i - \sum_{s=1}^{k-1} (k-s)\,x_{is} = kQ_i$$

$$b_{10}x_{i0} + b_{11}x_{i1} + \ldots + b_{1(k-2)}\,x_{i(k-2)} - x_{ik} = P_i^{\,1}$$

$$b_{20}x_{i0} + b_{21}x_{i1} + \ldots + b_{2(k-2)}\,x_{i(k-2)} - x_{i,\,k+1} = P_i^{\,2}$$

$$\ldots\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$b_{p0}x_{i0} + b_{p1}x_{i1} + \ldots + b_{p(k-2)}\,x_{i(k-2)} - x_{i(k+p-1)} = P_i^{\,p}$$

$$(p = 0, 1, \ldots\ldots\ d+1-k) \qquad (5)$$

In each of these equations $\sum_{j=0}^{k-2} b_{pj} = 1$ so that the sum of the coefficients of x_{ij}'s is zero. The different b_{pj} for any given j can be obtained through the recurrence relations:

$$b_{pj} = -2b_{(p-1)j} - 3b_{(p-2)j} \ldots\ldots - (k-1)\,b_{(p-k+2)j} + k(k-1)\,b_{(p-k+1)j} - (k-1)\,b_{(p-k)j} - \ldots\ldots - 2b_{(p-2k+3)j} - b_{(p-2k+2)j}. \qquad (6)$$

Thus, once we know the first $2(k-1)$ b_{pj}'s, the others can be obtained from the recurrence relation. Similarly, $p^p{}_i$ for p greater than $2(k-1)$ can be obtained from the recurrence relation having the same set of coefficients excepting that one more term, *viz.*, kQ_i^p is to be added to it.

With the help of the equations in set A_1, it is possible to eliminate all the x_{ij}'s where $j > k-2$ from each of the $(k-1)$ equations in set B. Only $(k-2)$ of these equations are independent, each containing $(k-1)$ unknowns. Thus, one more equation is needed to solve for x_{i0} in which alone we are interested. This last equation can be obtained by adding (i) all the equations in set A_1, when n is odd or (ii) all the equations excepting the last one and then adding half times the last equation. Evidently, the total when n is odd, is

$$x_{i0} \sum_{p=0}^{d+1-k} b_{p0} + \sum_{j=1}^{k-2} x_{ij} \sum_{p=0}^{d-k+1} b_{pj} - \sum_{j=k-1}^{d} x_{ij} = \sum_{p=0}^{d+1-k} P_i^p \qquad (7)$$

where $b_{00} = k(k-1)/2$ and the other b_{pj}'s can be obtained from equation (3).

From this equation $\sum_{j=k-1}^{d} x_{ij}$ can be eliminated when n is odd with the restriction $\sum_{i,j} t_{ij} = 0$ which becomes $a_i + \sum_{j=1}^{d} x_{ij} = 0$. When n is even, $a_i + \sum_{j=1}^{d-1} x_{ij} + \frac{1}{2} x_{id} = 0$.

Hence, the summation of the equations as in (ii) above is necessary.

Once the solution of a_i's are obtained those of t_{ij}'s can be obtained from the relation:

$$t_{ij} = \frac{a_i}{m} + \frac{Q_{ij}}{k} - \frac{\sum_j Q_{ij}}{km}. \qquad (8)$$

The adjusted S.S. can now be obtained from

$$\Sigma\, t_{ij} Q_{ij}.$$

The adjusted S.S. can also be obtained as the total of the S.S. due to a_i's and the S.S. due to t_{ij}'s within the different a_i's. The first part can be obtained from $\Sigma a_i Q_i$ while the second can be obtained in the usual manner from the treatment total t_{ij}'s separately for each i and then adding the S.S. for the different i's.

4. Analysis with Recovery of Interblock Information

The normal equation in this case can be written, following Rao (1947), as

$$kwt_{ij} - \sum_j t_{ij} \left\{\frac{wk - w'}{mk}\right\} - \frac{w - w'}{mk} \sum_{s=1}^{k-1} (k - s)\, x_{is} = p_{ij}' \quad (9)$$

where

$$p_{ij}' = wQ_{ij} + \frac{w'}{km}\left(B_{ij} - \frac{kG}{n}\right),$$

B_{ij} is the sum of those block totals in which t_{ij} occurs and G, w and w' have the usual meaning.

Adding over j,

$$a_i \left\{(k - 1)\, w + \frac{w'}{k}\right\} - \frac{w - w'}{k} \sum_{s=1}^{k-1} (k - s)\, x_{is} = \sum_j p_{ij}'. \quad (10)$$

Writing

$$\alpha = (k - 1)\, w + \frac{w'}{k} \quad \text{and} \quad \beta = \frac{w - w'}{k},$$

these equations become

$$\alpha a_i - \beta \sum_{s=1}^{k-1} (k - s)\, x_{is} = \sum_j p_{ij}'. \quad (11)$$

Dividing by β and putting $\alpha/\beta = c$, we can write the equations as

$$ca_i - \sum_{s=1}^{k-1} (k - s)\, x_{is} = \frac{\sum_j p_{ij}'}{\beta} = P_i \text{ (say)}. \quad (12)$$

These equations are similar to those at (3) and the same techniques of solving them can be applied excepting that in these equations the sum of the coefficients of the x_{ij}'s will not be zero and in the recurrence relations $k(k - 1)$ is to be replaced by c.

5. Particular Cases

Case 1.— When $k = 2$

The general equation in set A_1, *viz.*, equations (5) reduces to

$$b_{p0}x_{i0} - x_{i(p+1)} = P_i^p. \tag{13}$$

As the sum of the coefficients of the equations is zero, $b_{p0} = 1$. The equation, thus, becomes

$$x_{i0} - x_{i(p+1)} = P_i^p. \tag{14}$$

Adding (14) over p from 0 to $d - 1$,

$$dx_{i0} - \sum_{p=0}^{d-1} x_{i(p+1)} = \sum_{p=0}^{d-1} P_i^p. \tag{15}$$

Taking the restriction,

$$\Sigma\, t_{ij} = 0 = \frac{x_{i0}}{2} + \sum_{p=0}^{d-1} x_{i(p+1)}.$$

When n is odd, we get

$$x_{i0}(d + \tfrac{1}{2}) = \sum_{p=0}^{d-1} P_i^p$$

i.e.,

$$a_i = \frac{1}{2d+1} \sum P_i^p.$$

In this case it is possible to express $\Sigma\, P_i^p$ as a function of Q_i^p and can be shown to be equal to

$$\sum_{m=1}^{d-1} (d - m)(d - m + 1)\, Q_i^m.$$

Hence,

$$a_i = \frac{1}{2d+1} \sum_{m=0}^{d-1} (d - m)(d - m + 1)\, Q_i^m. \tag{16}$$

When n is even

$$a_i = \frac{1}{2d} \sum_{m=0}^{d-1} (d - m)^2\, Q_i^m. \tag{16 a}$$

t_{ij} can now be obtained from:

$$t_{ij} = \frac{a_i}{m} + \frac{Q_{ij}}{2} - \frac{Q_i}{2m}.$$

Variance of $(t_{ij} - t_{ij'}) = \sigma^2$.

Variance of

$$(t_{ij} - t_{i\pm p,\, j}) = \sigma^2 \left\{1 - \frac{1}{m} + \frac{2p(2d - p + 1)}{m(2d + 1)}\right\} \tag{17}$$

where p varies from 1 to d when n is odd

and can be obtained by collecting the coefficients of Q_{ij} and $Q_{i\pm p,\, j}$ in the estimate of $(t_{ij} - t_{i\pm p,\, j})$ and subtracting them.

When n is even the variance at (17) becomes

$$\left\{\left(1 - \frac{1}{m}\right) + \frac{p(2d - p)}{md}\right\} \sigma^2 \text{ for } p \text{ from 1 to } d.$$

The average variance when n is odd comes to

$$\frac{6d(m - 1) + 4d(d + 1)}{3(mn - 1)} \sigma^2.$$

Hence the efficiency factor in this case is

$$\frac{3(mn - 1)}{6d(m - 1) + 4d(d + 1)}.$$

For analysis with recovery of interblock information, equation (12) becomes in this case

$$ca_i - x_{i1} = P_i$$

and the equations corresponding to (5) are

$$b_{p0} x_{i0} - x_{i(p+1)} = P_i^p,$$

where P_i^p is defined in terms of P_i

whence the solution of a_i is

$$a_i = \frac{\Sigma P_i^p}{\left(1 + 2 \sum_{p=0}^{d-1} b_{p0}\right)}.$$

when n is odd.

When n is even the denominator should be reduced by unity.

Case 2.—$k = 3$

The normal equations corresponding to (5) each of which contains only 3 unknowns become

$$(1 - \alpha_p)\, x_{i0} + \alpha_p x_{i1} - x_{i(p+2)} = P_i^p. \tag{18}$$

where

(i) p varies from 0 to $d-2$.

(ii) $\alpha_p = -2\alpha_{p-1} + 6\alpha_{p-2} - 2\alpha_{p-3} - \alpha_{p-4}$ (19)

(iii) $P_i^p = 3Q_i^p - 2P_i^{p-1} + 6P_i^{p-2} - 2p_i^{p-3} - P_i^{p-4}$ (20)

From the equation for a_i, it will be seen that $\alpha_0 = -2$.

Eliminating x_{i2} from the equation, $p = 1$, α_1 comes to be 9. Similarly, from equation with p equal to 2 and 3 values of α_2 and α_3 come to -32 and 121 respectively. The other α_p's can now be obtained from the recurrence relation (19).

From these equations we have again

$$P_i^0 = 3Q_i, \quad P_i^1 = 3Q_i^1 - 2P_i^0$$

$$P_i^2 = 3Q_i^2 - 2P_i^1 + 6P_i^0$$

$$P_i^3 = 3Q_i^3 - 2P_i^2 + 6P_i^1 - 2P_i^0$$

and

$$P_i^4 = 3Q_i^4 - 2P_i^3 + 6P_i^2 - 2P^1 - P_i^0.$$

The other values of P_i^p can be now obtained from the recurrence relation (20).

Summing the equations (18) over p from 0 to $d-2$ the sum becomes

$$x_{i0} \sum_{p=0}^{d-2} (1 - \alpha_p) + x_{i1} \Sigma\, \alpha_p - \sum_{j=2}^{d} x_{ij} = \sum_{p} P_i^p. \tag{21}$$

Taking the restriction

$$\sum_{ij} t_{ij} = 0 = a_i + \sum_{j=1}^{d} x_{ij},$$

we get when n is odd

$$\left(d - \tfrac{1}{2} - \sum_{p=0}^{d-1} \alpha_p\right) x_{i0} + (1 + \Sigma\, \alpha_p)\, x_{i1} = \Sigma\, p_i^p. \tag{22}$$

Taking one more equation, *viz.*, the one corresponding to $p = d-1$ and eliminating all the x_{ij}'s from $j > 1$ with the help of the equation at (21) another equation of the form

$$a_{d-1}\,(x_{i0} - x_{i1}) = p_i^{s-1} \tag{23}$$

can be obtained.

When n is greater than 8, a_{d-1} can be obtained from

$$a_{d-1} = -3a_{d-2} + 6a_{d-3} - 2a_{d-4} - a_{d-5}.$$

No easy summation of $\Sigma\, a_p$ seems possible in this case. $\sum_p P_i^p$ can be expressed as a linear function of Q_i^p's as below:

$$\sum_{p=0}^{d-2} P_i^p = 3 \sum_{p=0}^{d-2} Q_i^p \left\{ \sum_{\gamma=1}^{d-1-p} b_\gamma \right\}$$

where b_γ's are the different terms of the series, 1, -2, 10, -34, 131, ..., the recurrence relation for getting the other terms of the series being

$$b_\gamma = -2b_{\gamma-1} + 6b_{\gamma-2} - 2b_{\gamma-3} - b_{\gamma-4}.$$

The solution for a_i can now be obtained from the two equations at (22) and (23).

When n is even, the results will change to some extent as indicated in the general case.

6. Concluding Remarks

The method of construction of such designs can be generalised in the case of 2 plot blocks with $m = 1$ by first arranging the treatments on the circumference of a circle and then including in a block any treatment together with one more which is at a distance of, say, q from it in a given direction. The next block will then contain the treatment considered last in the previous block and another which is at a distance of q from it in the same direction. This process is to be repeated to obtain different blocks till we arrive at the same treatment with which we started. If q happens to be a prime factor of n or any multiple of the factor, the process will terminate before all the treatments appear in the blocks. In such cases the process is to be repeated with another treatment which has not occurred in any block previously and so on till each of the treatments appear twice. In this case the design will be disconnected with different sections.

All the connected designs and each section of disconnected ones can be analysed as described in the paper if x_{is} is taken as the sum of the treatments which are at distances of qs on either side of the i-th treatment. The definition of Q_i^m's also will change correspondingly. The designs given by Kempthorne where $r = 2$ are obtainable through this method of construction by taking q equal to s, where s is as defined in Kempthorne's paper. While Kempthorne's design is available only for odd number of treatments when $r = 2$, the circular designs are available for all numbers of treatments and can be analysed through the same technique.

As for each value of q a separate design can be obtained, by combining two or more designs obtained from different values of q, designs with multiples of two replications can be obtained. All such designs which are obtainable by combining connected designs obtainable from the values 1 and 2 of q can also be analysed by the method given in the paper. In design with $m = 1$, $r = 2$ and $k = 2$, there will be only 1 *d.f.* for error. As such these designs are not suitable. In such cases the designs can be repeated any number of times, say, δ. The only adjustment needed in the analysis in case of repetition is to divide all the Q_i's by δ.

The circular designs are actually partially balanced incomplete block designs with associate classes of half or just less than half the number of treatments. When n is odd each $n_i = 2$ and when n is even one of the n_i's is 1 and the rest 2. Equations (4) evidently show that such designs are P.B.I.B. designs. All the designs become symmetrical when $m = 1$ and as such elimination of column effect is also nil.

For the analysis of designs with more than two replications a set of tables giving the coefficients in the different equations can be prepared for ready use with the help of the recurrence relations given in the paper. Such designs can also be adopted for both symmetrical and asymmetrical factorial designs. Thus, if n be 3, m be 4 and $k = 2$, then by taking n to be the levels of a factor and m, the number of combinations of two other factors each at 2 levels, we can get the asymmetrical factorial design 3×2^2 in blocks of 8 plots with 2 replications in which the main effect of the first factor only is affected with $\frac{1}{4}$ as the loss of information.

7. Summary

A new design called circular design has been defined and its method of both intra and inter-block analyses worked out. Such designs exist for all numbers of treatments and replications and do not involve any problem of construction. Expressions for analysis of particular cases with up to three replications, whatever the number of treatments, have been worked out.

8. Acknowledgement

The author is grateful to Dr. V. G. Panse, Statistical Adviser to the I.C.A.R., for his general guidance for research which produced this paper and for his encouragement for doing research work.

References

1. Bose, R. C. and Nair, K. R. "Partially balanced incomplete block designs," *Sankhya*, 1939, **4**, 337–73.
2. Das, M. N. .. "A generalised balanced design," *J. Ind. Soc. Agri. Stat.*, 1957, **9**, 18–30.
3. ——— .. "Reinforced incomplete block designs," *Ibid.*, 1958, **10**, 73–77.
4. Giri, N. C. .. "On reinforced partially balanced incomplete block designs," *Ibid.*, 1957, **9**, 41–51.
5. Harshbarger, B. .. "Rectangular lattices," *Virginia Agri. Expt. Sta. Memoir*, 1947, 1.
6. Kempthorne, O. .. "A class of experimental designs using blocks of two plots," *Ann. Math. Stat.*, 1953, **24**, 76–84.
7. Nair, K. R. and Kishen, K. "Recent developments in experimental design with special reference to the work in India," Presented at the 31*st session of the International Statistical Institute, Brussels*, September 1958.
8. Yates, F. .. "A new method of arranging variety trails involving a large number of varieties," *J. Agri. Sci.*, 1936 *a*, **26**, 424–55.
9. ——— .. "Incomplete randomised blocks," *Ann. Eug.*, 1936 *b*, **7**, 121–40.
10. Youden, W. J. .. "Linked blocks, a new class of incomplete block designs," *Biometrics*, 1951, **7**, 124.
11. ——— and Conner, W. S. "The chain block designs," *Ibid.*, 1953, 9 127–40.