



An Improved Method for Generating Factorial Effects and their Treatment Combinations in a Full Symmetric Factorial Designs

Dauran N. S.^a and Danbaba A.^b

^{a,b}Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria
Email: sani.nasiru1@udusok.edu.ng

ARTICLE INFORMATION

Article history:

Received 30 April 2019

Revised 08 May 2019

Accepted 15 May 2019

Available online 06 June 2019

Keywords:

Factorial design, treatment combinations, main effects, factorial interactions, pascal triangle, lexicographical order, Yates standard order and symmetric.

ABSTRACT

In factorial experiments the generation of treatment combination is the beginning of the statistical analysis. Plans presented in standard texts are not necessarily suitable in most cases. A systematic procedure for generating treatment combination has been presented by [1] and latter was extended by [2]. In this research we show an alternative method using Pascal triangle and lexicographical order to obtain number of factorial effects and treatment combinations respectively. The new procedures save space and time.

1. Introduction

Design of experiments has played a fundamental role in the statistical curriculum, practice, and research ever since R.A. Fisher founded the modern discipline [3]. It has been successfully applied in many fields of scientific investigation. These include agriculture, medicine, and behavioral research as well as chemical, manufacturing and high-technical industries. Concepts like randomization, confounding, and aliasing, which originated in factorial experiments, have found applications beyond their initial motivation. An important application of statistical methods to industrial research is in the factorial experiments in connection with the improvement of manufacturing processes. This applies to chemical and biological processes, formulation of pharmaceutical preparations and, in fact, to most types of industrial research. For example, in chemical research it is frequently required to determine the effect of certain changes in reaction conditions or methods of manufacture on the yield and quality of chemical products. Typical reaction conditions which might be varied are temperature of reaction, time of reaction, rates and methods of agitation, concentration and amounts of reactants, different catalysts, etc. Similar considerations arise in chemicals manufactured by fermentation processes, e.g. penicillin, streptomycin, industrial alcohol, lactic acid, etc., for which we may wish to examine the effect of changes in the conditions of fermentation on the yield and quality of the products. The objective of all such investigations is usually to improve the yield or the quality of the product, or to produce the product more economically.

The factorial experiment in which each of the possible combinations of the levels of several factors are allocated to one or more experimental units is an experimental technique [4]. An alternative formulation for such treatment combination was also given using Kronecker product

[5]. Other standard text books that also reported the method of Yate order include [6, 7, 8, 9, 10], 11 and 12].

Factorial design is a strategy in which factors are simultaneously varied, instead of one at time experiment [13]. The common factorial design is 2^n which is applicable when many factors are to be investigated, with the aim of finding out which factorial effects and interaction between factors are the most influential on the response of the experiment where n is the number of factors. Factorial experiments have come to play a prominent role in such investigations because of their advantages in economy, precision and accuracy. Besides having applicability to the screening of factors and the study of the ranges of factors, i.e., the exploratory stages of scientific investigation, factorial designs are of use in studying detail qualitative factors such as varietal differences.

Factorial designs have been used for many years in agricultural and biological research, where the practice has been to produce a large number of observations relative to the number of factors studied (e.g., by replication).

An early stage in the construction of two-level factorial experiments is the generating of treatments combinations and besides the Yates standard order as proposed by [14], most researchers and standard texts give no guidance as to the generation of these treatment combinations other than by trial and error, hence there is need to come up with a systematic procedure to cater for this problem.

1.1 Symmetric Factorial Designs

Although the 2^n factorial represents only a special case of the general S^n factorial experiment, that is, n factors with S levels each, it deserves special consideration because of its practical importance and its special algebraic and combinatorial representation. The list of treatment combination is called the design matrix and is denoted by **D**. For a 2^n factorial, the design matrix contains n columns and $N = 2^n$ rows. There is a column for each of the n variables, and each row gives a treatment combinations. Normally the runs are listed in standard order due to [1]. An experiment in which each of the 2^n possible combinations of factor levels occurs as a treatment is called a complete 2^n design. The two levels for each factor are generally called *low* and *high*. These terms have clear meanings if the factors are quantitative, but they are often used as labels even when the factors are not quantitative. Note that “off” and “on” would work just as well, but low and high are the usual terms.

There are two methods for denoting a factor-level combination in a factorial design. The first uses letters and is probably the most common. For example we denote a factor level combination by a string of lower-case letters such as bcd to represent low level of factor A, high level of factor B, high level of factor C and high level of D. The second method uses numbers and generalizes to three-level and higher-order factorials as well. A factor-level combination is denoted by n binary digits, with one digit giving the level of each factor: a zero denotes a factor at its low level, and a one denotes a factor at its high level. Thus 000 is all factors at low level, the same as (1), and 011 is factors B and C at high level, the same as bc.

A 2^n factorial experiment was considered and denoted the treatments by (x_1, x_2, \dots, x_n) , where $x_r = 0$ or $1, r = 1, 2, \dots, n$. and the “ true effect” of the treatment (x_1, x_2, \dots, x_n) is denoted by $\emptyset(x_1, x_2, \dots, x_n)$ [15]. Then obtained the following model;

$$\begin{aligned} \phi(x_1, x_2, \dots, x_n) &= \mu + \sum_{i=1}^n \alpha_i F_i \\ &+ \sum_{i_1, i_2=1; i_1 < i_2}^k \alpha_{i_1} \alpha_{i_2} F_{i_1 i_2} + \sum_{i_1, i_2, i_3=1; i_1 < i_2 < i_3}^n \alpha_{i_1} \alpha_{i_2} \alpha_{i_3} F_{i_1 i_2 i_3} + \dots + \alpha_1 \dots \alpha_k F_{12\dots n} \end{aligned} \quad (1)$$

Where μ is the general mean, F_i the mean effect of the i^{th} factor, $F_{i_1 i_2}$ the two factor interaction between the i_1 and i_2 , and so on, $\alpha_i = 1$ or -1 according as $x_i = 1$ or 0 , $i = 1, \dots, n$. Let $y(x_1, x_2, \dots, x_n)$ be the observed response corresponding to (x_1, x_2, \dots, x_n) . The set of all level combinations can be represented by a $2^n \times n$ Matrix of -1 's and $+1$'s, where ± 1 's represent the two levels of each factor, respectively see [16].

As earlier mentioned a convenient way of obtaining set of treatment combinations is to apply the Yates algorithm. Yates order state as follows: Write down the observed treatment means in standard order. With $n = 3$ and $s = 2$, following this order, the eight treatment combinations are (1), a, b, ab, c, ac, bc and abc , where (1) denotes the control, a the application of treatment A only, ab the application of treatments A and B and so on. Two-level full factorial and fractional factorial designs are widely used in industrial experiments [17] and also in agricultural experiments [18] and [19] to assess the impact of factorial effects on a process.

The estimates of the factorial effects is obtained on the assumptions that the observations are uncorrelated and have equal variance, then in the case of 2^n factorial designs provide independent minimum variance estimates of the grand average and of the 2^{n-1} factorial effects.

Consider a factorial design with 3 two-level factors. The set of all level combinations for the 3 independent factors can be written as:

$$D = \begin{pmatrix} A & B & C \\ -1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \quad (2)$$

In general, for n independent factors, the matrix D obtained in a similar fashion is called the 2^n full factorial design matrix. The set of columns corresponding to all the main effects and interactions is called the 2^n full factorial model matrix, denoted by [16].

where $X = 2^n \times (2^n - 1)$ matrix defined by

$H_n = \{A, B, AB, C, AC, BC, ABC, \dots, ABCD \dots k\}_{2^n}$ and $n = 1, 2, \dots, k$ with entries -1 and 1 and

columns arranged in Yates order, where $A = (-1, \dots, -1, 1, \dots, 1)'$,

$B = (-1, \dots, -1, 1, \dots, 1, -1, \dots, -1, 1, \dots, 1)'$, ..., and $n = (-1, 1, -1, 1, \dots, -1)'$ stand for its n independent

columns, and the other columns are expressed as products of the n independent columns.

The corresponding model matrix X for the 2^3 design is given by

$$X = \begin{pmatrix} A & B & C & AB & AC & BC & ABC \\ -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (3)$$

This representation of the factor level combinations is convenient since the columns of X denote the linear contrasts that estimate the main effects and interactions in a normal linear regression model by $\frac{XY}{2^n}$, where Y is the vector of observations corresponding to the factor level settings of each row of D .

1.2 Lexicographical Order

Lexicographical method is similar to dictionary (alphabetical) order [20]. If Word A is shorter than Word B, and every letter of Word A occurs in the same place in Word B, Word A comes before Word B ("compute" and "computer") If the first letter that differs in Word A comes before the corresponding letter in Word B, then Word A comes before Word B ("math" and "matter").

2. Methodology

2.1 Some Proposed Methods for Generating Factorial Effects and their Treatment Combinations

The classical method of generating factorial effects is inefficient of time and space. Hence we consider a new method that leads to a more formal and simpler in order to obtain main effects, interactions and treatment combinations.

Some alternative methods for generating factorial effects and their treatment combinations are proposed. These include using the general concept of Pascal triangle and lexicographical order.

2.1.1 A New Method of Generating 2^n Factorial Effects from Pascal Triangle

Under this subheading a proposition was made to obtain both the number of factorial effects and their treatment combinations as follows:

Proposition 2.1 The generation of orthogonal contrasts in 2^n factorial design follows Pascal triangle.

Proof.

In 2^n where $n = 2, 3, \dots$, be the real numbers. Then prove by induction.

For $k = 2$ the result is trivially true.

Now it is also true for $h = k$ since we have:

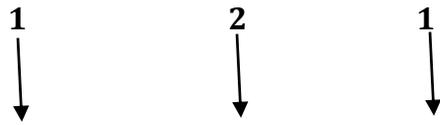
$n - \text{main effects}$,

$$\frac{n(n-1)}{2!} \text{ two-factor interaction, } \frac{n(n-1)(n-2)}{2 \times 3} \text{ three}$$

factor interaction effects, ..., $\frac{n(n-1)(n-2)\dots(n-k+1)}{k!} - (n-1)$ factor interactions and finally

1 $n - \text{factor interaction}$.

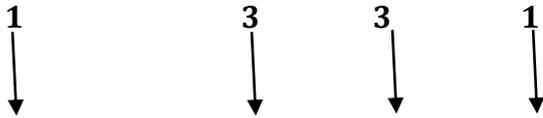
For $n = 2$



1 – identity 2 – me 1 – 2fi

This shows that in 2^2 factorial design the triangle generates 1–mean (*I*) the identity, 2–main effects (me), and 1–2–factor interactions (fi).

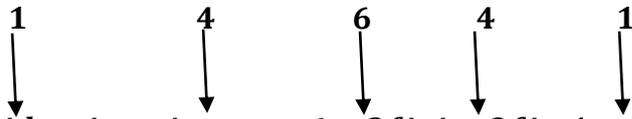
For $n = 3$



1 – identity 3 – me 3 – 2fi 1 – 3fi

For 2^3 factorial design the triangle will generates 1–mean (*I*) the identity, 3–main effects (me), 3–2–factor interactions (fi), and 1–3–factor interactions (fi),

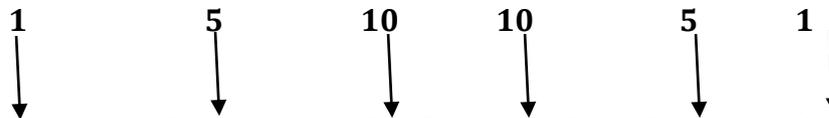
For $n = 4$



1 – identity 4 – me 6 – 2fi 4 – 3fi 1 – 4fi

For 2^4 factorial design the triangle will generates 1–mean (*I*) the identity, 4–main effects (me), 6–2–factor interactions (fi), 6–2–factor interactions (fi), and 1–4–factor interactions (fi),

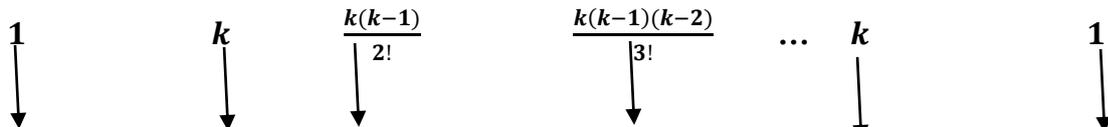
For $n = 5$



1 – identity 5 – me 10 – 2fi 10 – 3fi 5 – 4fi 1 – 5fi

For 2^5 factorial design the triangle will generate 1–mean (*I*) the identity, 5–main effects (me), 10–2–factor interactions (fi), 10–3–factor interactions (fi), 5–4–factor interactions (fi), and 1–5–factor interactions (fi),

For $n = k$



1 – identity k – me $\frac{k(k-1)}{2!}$ – 2fi $\frac{k(k-1)(k-2)}{3!}$ – 3fi k – (k – 1)fi 1 – kfi

For 2^k factorial design the triangle will generates 1–mean (*I*) the identity, k–main effects (fi), $\frac{k(k-1)}{2!}$ –2–factor interactions (fi), $\frac{k(k-1)(k-2)}{3!}$ –3–factor interactions (fi), ... , k–(k–1)–factor interactions (fi), and 1–k–factor interactions (fi),

3.2 Application of Lexicographical Method in Generating the Treatment Combinations

The above rules I and II to are being applied to obtain the following table for cases of $(s = 2, k = 2), (s = 2, k = 3), (s = 3, k = 3)$ and $(s = 3, k = 4)$.

Table 1. The List of S^K Treatment Combinations with S prime number and K number of factors

<i>s / no</i>	<i>The Algorithm</i>	<i>Factors and Levels</i>	<i>Re sult</i>
1	00,01,...,0 \bar{s}_2 ,10,11,...1 \bar{s}_2 ,..., \bar{s}_1 0, \bar{s}_1 1,... $\bar{s}_1\bar{s}_2$	$s = 2, k = 2$ $s = 3, k = 2$	00,01,10,11. 00,01,02,10,11,12,20,21,22.
2	000,010,...,0 \bar{s}_2 0,100,110,...1 \bar{s}_2 0,..., \bar{s}_1 00, \bar{s}_1 10,... $\bar{s}_1\bar{s}_2$ 0,001,011,...,0 \bar{s}_2 1,101,...,1 \bar{s}_2 1, ..., \bar{s}_1 01,..., $\bar{s}_1\bar{s}_2$ 1,...00 \bar{s}_3 ,01 \bar{s}_3 ,...,0 $\bar{s}_2\bar{s}_3$,10 \bar{s}_3 , 11 \bar{s}_3 ,...,1 $\bar{s}_2\bar{s}_3$,..., \bar{s}_1 0 \bar{s}_3 , \bar{s}_1 1 \bar{s}_3 ,..., \bar{s}_1 1 \bar{s}_3 ,..., $\bar{s}_1\bar{s}_2\bar{s}_3$	$s = 3, k = 3$	000,010,020,100,110,120,200, 210,220,001,011,101,110,121,201, 211,221,002,022,102,112,202,212, 222
3	0000,0100,...,0 \bar{s}_2 00,1000,...,1 \bar{s}_2 00,... \bar{s}_1 000,..., $\bar{s}_1\bar{s}_2$ 00,0010,0110,...0 \bar{s}_2 10,1010,...,1 \bar{s}_2 10, ..., \bar{s}_1 010, $\bar{s}_1\bar{s}_2$ 10,...,00 \bar{s}_3 0,01 \bar{s}_3 0,...,0 $\bar{s}_2\bar{s}_3$ 0, 10 \bar{s}_3 0,..., \bar{s}_1 0 \bar{s}_3 0,..., $\bar{s}_1\bar{s}_2\bar{s}_3$ 0,0001,0101,..., 0 \bar{s}_2 01,1001,...,1 \bar{s}_2 01,..., \bar{s}_1 001,..., $\bar{s}_1\bar{s}_2$ 01,0011, ...,0 \bar{s}_2 11,1011,...,1 \bar{s}_2 11,..., \bar{s}_1 011,..., $\bar{s}_1\bar{s}_2$ 11,00 \bar{s}_3 1, 01 \bar{s}_3 1,...,0 $\bar{s}_2\bar{s}_3$ 1,10 \bar{s}_3 1,...,1 $\bar{s}_2\bar{s}_3$ 1,..., \bar{s}_1 0 \bar{s}_3 1,...,1 $\bar{s}_2\bar{s}_3$ 1, ..., \bar{s}_1 0 \bar{s}_3 1, \bar{s}_1 1 \bar{s}_3 1, $\bar{s}_1\bar{s}_2\bar{s}_3$ 1,...,000 \bar{s}_4 ,010 \bar{s}_4 ,...,0 \bar{s}_2 0 \bar{s}_4 , ...100 \bar{s}_4 ,..., \bar{s}_1 00 \bar{s}_4 , \bar{s}_1 10 \bar{s}_4 ,..., $\bar{s}_1\bar{s}_2$ 0 \bar{s}_4 ,001 \bar{s}_4 ,011 \bar{s}_4 , ...,0 \bar{s}_2 1 \bar{s}_4 ,101 \bar{s}_4 ,111 \bar{s}_4 ,...,1 \bar{s}_2 1 \bar{s}_4 ,..., \bar{s}_1 01 \bar{s}_4 , \bar{s}_1 11 \bar{s}_4 ,..., $\bar{s}_1\bar{s}_2$ 1 \bar{s}_4 ,00 $\bar{s}_3\bar{s}_4$,01 $\bar{s}_3\bar{s}_4$,...,0 $\bar{s}_2\bar{s}_3\bar{s}_4$,10 $\bar{s}_3\bar{s}_4$ 11 $\bar{s}_3\bar{s}_4$,...,1 $\bar{s}_2\bar{s}_3\bar{s}_4$,..., \bar{s}_1 0 $\bar{s}_3\bar{s}_4$, \bar{s}_1 1 $\bar{s}_3\bar{s}_4$,..., $\bar{s}_1\bar{s}_2\bar{s}_3\bar{s}_4$	$s = 3, k = 4$	000,0100,0200,1000,1100,1200, 20002100,2200,0010,0110,0210, 1010,1110,1210,2010,2110,2210, 0020,0120,0220,1020,1120,1220, 2020,2120,2220,0001,0101,0201, 1001,1101,1201,2001,2101,2201, 0011,0111,1011,1111,1211,2011, 2111,2211,0021,0121,1021,1121, 1121,1221,2021,2121,2221,0002, 0102,0202,1002,1102,1202,2002, 2102,2202,0012,0112,0212,1012, 1112,1212,2012,2112,2212,0022, 0122,0222,1022,1122,1222,2022, 2122,2222.

Where $\bar{s}_1 = s_1 - 1$, $\bar{s}_2 = s_2 - 1$, $\bar{s}_3 = s_3 - 1$ and $\bar{s}_4 = s_4 - 1$

Table 1 display the treatment combinations of the S^K with the algorithm of how they are generated using $K = 2, 3, 4$ and $S=2,3$. The rigorous mathematical computation involved in kronecker product formulation of [2] in generating factorial effect and treatment combinations make it hard and cumbersome. The proposed method is a convenient one in achieving the same target. While [2] restrict the discussion on symmetric 2-level factorial designs this paper consider both symmetric and asymmetric s-level factorial design as can be seen in the application of the

algorithm for lexicographical order. Also the proposition under Pascal triangle can be used to calculate number of combinations of factorial effects and treatment combinations using the recursion formula.

4. Conclusion

The classical method of generating treatment combinations is only applicable to Yates' forward algorithm in estimating the effects of full factorial and the aliased subsets of fractional design. The Yates' order which is difficult to implement in fractional factorial design especially in the case of $n \geq 3$. This is mostly common in generating the blocks elements of principal block.

The existing method will pave way for estimating the factorial effects for not only the full and fractional factorial designs but for designs with confounding, partial confounding and split plot.

References

- [1] Yates, F.,(1937) . The Design and Analysis of Factorial Experiments. Imper. Bur. Soil Sci. Tech. Comm.135 144.
- [2] Kurkjian, B. and M. Zelen (1962). A calculus for factorial arrangements. Ann. Math. Statist. 33, 600-619
- [3] Rahul M. C.F. and Wu J. (2006). A Modern Theory of Factorial Designs. New York: Springer
- [4] Fisher R. A. (1926) The Arrangement of Field Experiments, Journal of Ministry of Agriculture, 23, 503-513.
- [5] Kurkjian, B. and M. Zelen (1963). Application of the calculus for factorial arrangements. Block and direct product designs. Biometrika 50, 63-73
- [6] Montgomery D. C. (2001). Design and Analysis of Experiment 5th edition"John Wiley and Son.
- [7] Garay W. D. (2010) A First Course In Design and Analysis of Experiments, John Wiley and Son.
- [8] Dean A. M.. and D. T. Voss (1999). Design and Analysis of Experiments, springer – Verlag new York, Inc.
- [9] Patrick D. and Andrew D. (2007) Analysis of Variance and Covariance, How to Choose and Construct Models for the Life Sciences" Cambridge University Press, UK.
- [10] Glenn G., Lawrence S. M., and Guarino, A.J (2008) Analysis of Variance Designs, A Conceptual and Computational approach with SPSS and SAS, Cambridge University , Press, UK
- [11] Klaus H. and Oscar K. (2007) Design and Analysis of Experiments A John Wiley and Sons, Inc., 2nd, (1)
- [12] Scott E. M. and Harold D. D. (2004) Designing Experiments and Analysing Data, A Model Comparison Perspective Lawrence Erlbaum Associates, Publishers Mahwah, New Jersey, London
- [13] Monica B. Carmen M. G., Martin M. A. and Antonio M. (2009). Application of the Factorial Design of Experiments to Biodiesel Production fro Lard. Fuel Processing Technology. 90(12) 1447 – 1451.
- [14] Yates, F. W. (1935). Complex experiments. Journal of the Royal Statistical Society, supplement 2, 181-247
- [15] Subir G. (1980) On Main Effect Plus One Plans for 2^m factorial, The Annals of Statistics, 8(4), 922-930
- [16] Shengli, Z., Dennis K. J. and Pengfei, L. (2016). A Note on the Construction of Blocked two-level Designs with general Minimum Lower Order Confounding. Journal of Statistical Planning and Inference. 172, 16-22.
- [17] Box G. E. P., Hunter J. S. and Hunter, W. G. (1978). Statistics For Experimenters. 2nd ed. New York:Wiley.
- [18] Kempthorne, O. (1952). The Design and Analysis of Experiments. Wiley Publications in Statistics. New York: John Wiley and Sons
- [19] Cochran, W.G. and Cox, G. M. (1957). Experimental Designs, 2nd edn. New York: Wiley.
- [20] Nathan W. (2008) Algorithms for Generating Permutations and Combinations www4.uwsp.edu/.../209-0809F-L10-Section06-03-Algorithms for Generating Permutations and Combinations