The Korean Communications
in Statistics Vol. 5, No. 2, 1998
pp. 353-359

A Simple Algorithm for Factorial Experiments in p^N

Dongwoo Kim[1)]

Abstract

Factorial designs with two-level factors represent the smallest factorial experiments. The system of notation and confounding and fractional factorial schemes developed for the 2^N system are found in standard textbooks of experimental designs. Just as in the 2^N system, the general confounding and fractional factorial schemes are possible in 3^N, 5^N, ... , and p^N where p is a prime number. Hence, we have the p^N system. In this article, the author proposes a new algorithm for constructing fractional factorial plans in the p^N system.

1. Introduction

The notion of factorial experiments was formally introduced into the scientific literature by R.A. Fisher in 1926. The basic idea was to study a set of factors affecting some phenomenon by varying combinations of factors in prescribed patterns. Prior to this ground-breaking paper, the standard approach in the scientific world had always been to study factors one-at-a-time. The scientist would very carefully hold conditions constant and then systematically study the factors one-at-a-time. This innovative idea, factorial experiments, has been taken up, developed, and extended by many statisticians and scientists. Some of the other early contributors to the statistical theory of factorial experiments were Yates (1937), Bose and Kishen (1940), and Bose (1947).

In its simplest form, a factorial experiment includes experimental runs with all possible combinations of levels of all factors considered. The total number of experimental runs is the product of the number of levels for each of the individual factors. When the experiment involves many factors, or factors at many levels, this leads to the obvious problem that a full replicate of the experiment becomes very large. The experiment becomes too large in two senses. The experimenter may have difficulty finding enough experimental material that is sufficiently uniform for the experiment, or the experimenter cannot afford the large number of experimental runs required. Fisher (1935) proposed blocking to avoid the first problem. Heterogeneity among experimental units increases experimental error. By blocking, experimenters can remove unwanted sources of variability among experimental units, and

1) Senior Statistical Consultant, Software Business Team, Samsung SDS Co., LTD.
707-19, Yoksam-Dong, Kangnam-Gu, Seoul, Korea, 135-080 dwkimsds@samsung.co.kr

thereby reduce the error.

The fundamental idea in blocking is to group experimental units into batches, called blocks in a manner that makes units within blocks as uniform as possible. Treatment combinations are then assigned within blocks to take advantage of this increased uniformity.

Finney (1945) devised the concept of fractional factorial experiments to get around the second problem. The fundamental idea in fractional factorial experiments is to examine a subset or fraction of the set of all possible treatment combinations. The challenge is to select the subset so that the responses to the resulting subset of treatment combinations can be interpreted easily. The usual strategy is to pick the fraction so that main effects and low-order interactions are retained and high-order interactions sacrificed.

2. The Fractionation Concept

To illustrate the fractionation concept involved without too much complexity, let us consider a small study, an experiment with selecting four observations from a 2^3 factorial plan, where the factors are A, B, and C, that is a 2^{3-1}, or in other words a half replicate of the 2^3 treatment combinations. For example, to select a half replicate, one chooses an effect, or more commonly an interaction to define the fraction, say ABC and then retains either the set of treatment combinations that enter the definition of the interaction with a positive sign or the alternative set of combinations that enter with a negative sign. Common practice is to formally represent these two alternate fractions by the symbols $I=+ABC$ and $I=-ABC$.

The first fraction is obtained by selecting treatment combinations for which $x_A+x_B+x_C=1$ (mod 2) and the second by selecting treatment combinations for which $x_A+x_B+x_C=0$ (mod 2). Now consider more complex example, an experiment with selecting eight observations from a 2^5 factorial plan, where the factors are A, B, C, D, and E, that is a 2^{5-2}, or in other words one fourth replicate of the 2^5 treatment combinations. To select a half replicate, one chooses an effect, or more commonly an interaction to define the fraction. For example, $I=ABC$ specifies a half replicate. To reduce down to one fourth replicate we must select another defining contrast. Choose CDE as another defining contrast.

The defining relation is $I=ABC=CDE=ABDE$. Note ABDE is the generalized interaction of ABC and CDE. Finally one must obtain the actual fraction, the actual set of treatment combinations, that are to be used by the scientist. There are numerous ways of approaching the job of finding a set of treatment combinations. One way to do this is to think of the set as the solutions to the equations $x_A+x_B+x_C=1$ (mod 2) and $x_C+x_D+x_E=1$ (mod 2). We must obtain solutions to the system of equations by modular arithmetic. Modular arithmetic is a simple computation by remainder. (mod 2) can have the value 0 or 1 since a number is divided by 2 and the possible remainder is 0 or 1. Similarly, (mod 3) generates 0,1, and 2. But the more generators, the harder solving equations. For example, the 2^{10-5} plan needs to solve the system of five equations. When the number of generators are increased, solutions may practically not be obtained without help of computers.

3. A Simple Algorithm for Factorial Designs

In order to solve the system of equations, a new algorithm is developed in this article. It can be easily implemented by using various computer languages. Especially, some sophisticated programs such as SAS, S-plus or some other statistical package programs require several lines of commands to implement this algorithm.

To illustrate the concept of the algorithm, consider 2^N factorial designs since factorial designs with two-level factors represent the smallest factorial experiments. As a simple example, consider a half fraction of a 2^3 factorial design which is defined by the interaction ABC. The actual plan for a complete replicate of a 2^3 factorial plan consists of 2^3 treatment combinations, all possible combinations of the two factors, each at two levels.

The algorithm we propose is based on modular and integer functions. The construction of a fractional factorial design by using this algorithm is straightforward. All we have to do is generating a full replicate of p^N factorial and selecting a subset or fraction of the set of all possible treatment combinations. For convenience, denote the modular function a mod b as $\mathrm{mod}(a, b)$. Since we have three factors, we need to establish three functional generators. For the first factor, $\mathrm{mod}(i, 2)$ is obvious choice since the solutions of i mod 2 are within a set F_1 of elements 0, 1 where i is an integer, $0 \leq i \leq 2^3 - 1$. Similarly, $\mathrm{mod}(i, 2^2)$ generates a set F_2 of elements 0, 1, 2, 3 for the second factor. By the same way, $\mathrm{mod}(i, 2^3)$ produces a set F_3 of elements 0, 1, ..., 7. All three sets should belong to the *Galois Field,* GF(2), since two-level factors are considered. However, the sets F_2 and F_3 are not in GF(2). To convert them to ones in GF(2), we select proper divisors, and apply integer functions to their results. Consequently, the full replicate of the 2^3 factorial design is constructed by using three modular functions $\mathrm{mod}(i, 2)$, $\mathit{integer}(\mathrm{mod}(i, 2^2)/2)$ and $\mathit{integer}(\mathrm{mod}(i, 2^3)/2^2)$. From this full replicate, the final fraction can be found by selecting treatment combinations that satisfies the equation $x_A + x_B + x_C = 1 \pmod 2$.

The idea illustrated in this example can be extended to the experimental situation with N factors. The system of the algorithm for confounding and fractional factorial schemes developed for the 2^N system extend directly to the p^N system, for any prime number p. The details of the general confounding and fractional factorial method for the p^N system can be found in standard textbooks such as P.W.M. John (1971), Anderson and McLean (1974), Raghavarao (1971), Raktoe, Hedayat, and Federer(1981), and Das and Giri (1986).

Algorithm

1. The number of factors $N > 0$, where N are integers, and p is a prime number.
2. Construct a full replicate by using modular and integer functions;

   Do i= 0 to $p^N - 1$

   Factor1=mod(i, p)

   Factor2=integer(mod(i, p^2*)/ p)*

   Factor3=integer(mod(i, p^3*)/* p^2*)*

   .

   .

   FactorN=integer(mod(i, p^N*)/* p^{N-1}*)*

   End
3. Choose a fraction from the above full replication that satisfies the system of equations.

4. Examples

4.1 2^N Factorial Experiments

4.1.1 Example I

Consider one fourth replicate of a 2^5 factorial plan that is composed of 8 observations. As a n example, generate a fraction defined by I=ABC=CDE=ABDE.

SAS program:

```
data all ;
 do i = 0 to 31;
      A = mod(i,2);
      B = int(mod(i,4)/2);
      C = int(mod(i,8)/4);
      D = int(mod(i,16)/8);
      E = int(mod(i,32)/16);
  output;
 end;

data frac; set all;
 if (mod(A+B+C,2) ne 1) then delete;
 if (mod(C+D+E,2) ne 1) then delete;
proc print; var  A B C D E; run;
```

SAS output:

TRT	A	B	C	D	E
1	0	0	1	0	0
2	1	1	1	0	0
3	1	0	0	1	0
4	0	1	0	1	0
5	1	0	0	0	1
6	0	1	0	0	1
7	0	0	1	1	1
8	1	1	1	1	1

4.1.2 Example II

Consider the problem of putting a 2^5 factorial in 8 blocks of size 4 in a complete replicate. A reasonable choice for the first contrast to be sacrificed is ABCD. For the second generator, one can select BDE. This implies that ABCD × BDE=ACE is also confounded. Now for the third generator, select ADE. This implies that BCE, AB, and CD are also confounded. In order to put the 2^5 factorial into eight blocks of size four we need to confound a total of seven contrasts. These result from selecting three generators and then finding all their generalized interactions.

SAS program:

```
data all;
 do i = 0 to 31;
      A = mod(i,2);
      B = int(mod(i,4)/2);
      C = int(mod(i,8)/4);
      D = int(mod(i,16)/8);
      E = int(mod(i,32)/16);
   output;
  end;
data frac; set all;
 block=mod(A+B+C,2)+2*mod(C+D+E,2)+1;
proc sort; by block;
proc print; var block A B C D E; run;
```

SAS output:

TRT	BLOCK	A	B	C	D	E
1	1	0	0	0	0	0
2	1	1	1	0	0	0
3	1	1	0	1	1	0
4	1	0	1	1	1	0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
32	4	1	1	1	1	1

4.2 3^N Factorial Experiments

4.2.1 Example I

Consider a 3^2 factorial experiment in three blocks of size three for a one complete replicate by confounding AB^2. In order to obtain the intra block subgroup, consider the following program.

SAS program:

```
data all;
  do i = 0 to 8;
     A = mod(i,3);
     B = int(mod(i,9)/3);
   output;
  end;
data frac; set all;
block=mod(A+2*B,3)+1;
proc sort; by block;
proc print; var block A B; run;
```

SAS output:

TRT	BLOCK	A	B
1	1	0	0
2	1	1	1
3	1	2	2
4	2	1	0
5	2	2	1
6	2	0	2
7	3	2	0
8	3	0	1
9	3	1	2

4.2.2 Example II

In order to find one-third fraction of a 3^2 factorial experiment, a simple modification of the above SAS program is needed.

SAS program:

```
data all;
  do i = 0 to 8;
    A = mod(i,3);
    B = int(mod(i,9)/3);
    output;
    end;
data frac; set all;
 if (mod(A+2*B,3) ne 0) then delete;
proc print; var A B; run;
```

SAS output:

TRT	A	B
1	0	0
2	1	1
3	2	2

Acknowledgement

The author thanks Dr. Francis Giesbrecht for various helpful discussions and examples.

References

[1] Anderson, V. L. and McLean, R. A. (1974). *Design of Experiments, A Realistic Approach*, New York: Marcel Dekker.

[2] Bose, R. C. (1947). ''Mathematical Theory of the Symmetrical Factorial Designs,'' *Sankhya*, **8**, 107--166.

[3] Bose, R. C., and Kishen, K. (1940). ''On the Problem of Confounding in the General Symmetric Factorial Designs,'' *Sankhya*, 5, 21--36.

[4] Das, M. N. and Giri, N. C. (1986). *Design and Analysis of Experiments*, New York: Wiley.

[5] Davies, O. L. ed. (1954). *Design and Analysis of Industrial Experiments*, Edinburgh: Oliver and Boyd.

[6] Finney, D. J. (1945). ''The Fractional Replication of Factorial Arrangements,'' *Annals of Eugenics*, **12**, 291--301.

[7] Fisher, R. A. (1935). *Design of Experiments*, Edinburgh: Oliver and Boyd.

[8] Giesbrecht F. G. (1995). Lectures on Experimental Designs, Unpublished, Department of Statistics, North Carolina State University.

[9] John, J. A., and Quenouille, M. H. (1977). *Experiments: Design and Analysis*, London:Charles Griffin.

[10] Kempthorne, O. (1952). *The Design and Analysis of Experiments*, New York: Wiley.

[11] Raghavarao, D. (1971). *Constructions and Combinatorial Problems in Design of Experiments*, New York: Wiley.

[12] SAS Institute Inc., SAS/QC Software, Version 6.12, The Factex Procedure.

[13] Stewart, I. (1989). *Galois Theory*, Chapman and Hall Mathematics.

[14] Street, A. P. and Street, D. J. (1987). *Combinatorics of Experimental Design*, Oxford Science Publications

[15] Yates, F. (1936). ''A New Method of Arranging Variety Trials Involving A Large Number of Varieties,'' *Journal of Agricultural Science*, **26**, 424--455.

[16] ______ (1937). ''The Design and Analysis of Factorial Experiments,'' *Imperial Bureau of Soil Science*, Harpenden, England.