KEYFINDER - A COMPLETE TOOLKIT FOR GENERATING BLOCKED AND/OR FRACTIONAL-REPLICATE FACTORIAL DESIGNS

VERSION 3.3 - OVERVIEW

1. Introduction

KEYFINDER is a menu-driven interactive program that assists statisticians in the difficult and time consuming tasks of generating, randomizing and tabulating factorial designs in completely general situations. Its particular forte is the use of search procedures to generate blocked and/or fractional-replicate designs with user-specified confounding and aliasing properties. At present, there is no other system on the market with comparable capabilities.

To obtain a free copy of the KEYFINDER program and manual, simply send an email request to the custodians. Full details are given at the end of this overview. KEYFINDER is distributed as an executable program file and runs under all versions of Windows. The source code was written in Prolog

Chapters 2 and 3 below give a general overview of the facilities available in KEYFINDER. For detailed instructions on how to use the program to generate experimental designs, you should consult the comprehensive KEYFINDER manual. The latter document also includes a theoretical treatment of the basic principles underlying KEYFINDER.

KEYFINDER is a system which is intended to help statisticians with those parts of experimental design construction that are difficult or tedious. It is not a system for those with little or no statistical knowledge. A basic understanding of the principles and underlying assumptions of experimental design and analysis of variance is assumed throughout this document and in the KEYFINDER manual.

A list of KEYFINDER publications follows Chapter 3.

2. Design Generation

KEYFINDER generates factorial designs involving a number of treatment factors (denoted A, B, ...), each of which is tested at a number of distinct "levels". KEYFINDER can generate

- multiple, single or fractional replicates

In multiple-replicate designs, each combination of treatment levels is tested several times, in single-replicate designs just once. In fractional-replicate designs, only a subset of the possible combinations of treatment levels is tested.

Designs may be

- symmetrical or asymmetrical

A symmetrical design is one in which all the block and treatment factors have the same number of levels; in an asymmetrical design, the numbers of levels may differ.

Designs may be generated by a number of methods, the most important being

- design keys

The design key was originally conceived as a device both for constructing fractional-replicate and/or blocked factorial designs and for identifying the aliasing and/or confounding patterns therein (Patterson, 1965, 1976; Patterson and Bailey, 1978). KEYFINDER reverses this process by searching for design keys yielding designs matching the user's *a priori* aliasing and confounding requirements (Zemroch, Lunn, Baines and Clithero, 1989). In particular, KEYFINDER can generate designs of

- fixed resolution

guaranteeing that only interactions above a given order are aliased with main effects (Box and Hunter, 1961) and also

- compromise designs

which additionally protect user-defined lists of higher-order interactions against aliasing.

The experimental units in all the design types described above can be arranged into blocks. The block factors, denoted P, Q, ... may be arranged into the following types of block structure:

- unblocked
- single block factor (P)
- row-and-column (P + Q + ...)
- crossed (P * Q * ...)
- nested (P / Q / ...) .

The notation used is that of Wilkinson and Rogers (1973) and that used in the GenStat statistical package (see http://www.vsni.co.uk/). In the subsequent analysis of the resulting experimental data, treatment factors will normally be assumed to have "fixed" effects, and block factors "random" effects (Scheffé, 1959).

When the block factors are nested, KEYFINDER can generate

- split-plot designs

with the allocation of treatments to the various error strata, P, P.Q, ..., P.Q..(units), completely under the user's control (Nelder, 1965).

KEYFINDER generates designs of

- fixed confounding limit

guaranteeing that treatment interactions above a given order only are confounded with blocks (Zemroch *et al.*, 1989). Again compromise designs can be generated which additionally protect user-defined lists of higher-order treatment interactions against confounding with blocks.

Design keys generate designs in which the numbers of levels of the various block and treatment factors, and the number of "experimental units" or "design points", are all powers of the same prime number p. To improve KEYFINDER's flexibility, a wide variety of further procedures are provided for generating classes of fractional-replicate designs with numbers of units n that are not prime powers p^q.

The most important classes of symmetrical fractional-replicate designs available in KEYFINDER with $n \neq p^q$ are the resolution-3

- Plackett-Burman designs

for testing a number of 2-level factors with n a multiple of 4 (Plackett and Burman, 1946), and the

- Addelman-Kempthorne designs

for testing several s-level factors, with $n = 2s^2$ and s a prime or prime power (Addelman and Kempthorne, 1961). Procedures are also provided for generating various general classes of asymmetrical fractional-replicate design and these are described in the book of Dey (1985). Usually, non-key designs can only be arranged into simple block structures with either a single block factor, or a greater number of block factors in a row-and-column arrangement.

The

- direct product

method combines sub-designs with n_1, n_2, ... units into a final design with $n_1 \times n_2 \times \ldots$ units. The sub-designs may be any combination of full factorials, designs built from keys, and non-key designs. The direct product method gives true flexibility in that asymmetrical designs can be built with no constraints on the numbers of levels of the various factors. Direct product designs are used in the so-called "Taguchi" method for quality control (Taguchi, 1976).

Designs produced by all the above methods can be manipulated in a number of ways and this further widens the flexibility of KEYFINDER. Several

- repeat

copies of a base design (usually a fractional-replicate) can be juxtaposed, allowing irregular fractions such as $^2/_3$- or $^3/_5$-replicates to be formed. Different row vectors can be added to each copy of the base design to ensure that the treatment combinations tested in the final design are all different.

Design dimensions may also be modified by the

- collapse

of factors. A 12-level factor, for example, can be collapsed into 2- and 6-level subfactors or into 3- and 4-level subfactors. If the collapsed design is to be a fractional-replicate, the options are even more varied. The 4-level subfactor might be further collapsed into three 2-level subfactors. One of these could then be recombined with the 3-level factor to give a 6-level factor.

KEYFINDER only produces designs which are

- balanced

in the sense that all the s_i levels of any particular block or treatment factor will appear an equal number of times (n/s_i) in the design. Moreover, the levels of any particular treatment factor will appear an equal number of times in each block (except in split-plot designs where a treatment estimated in a between-block stratum will always take the same level in any particular block). KEYFINDER, therefore, cannot generate balanced or partially balanced incomplete block designs (BIBs and PBIBs).

All the designs generated by KEYFINDER yield experimental data which can be analyzed in a straightforward way by leading statistical packages such as GenStat, SAS, JMP or R.

KEYFINDER has further facilities for the input of

- user-supplied designs

opening its extensive randomization and tabulation facilities to external designs provided by the user. Such designs can also be used as subdesigns of direct product designs.

For completion, facilities are provided for relabelling and omitting factor levels. By assigning the same labels to two different levels of a particular factor it is possible to produce fractional-replicate designs which, although unbalanced, obey the "proportional frequency" condition (Addelman, 1962). The user is warned, however, of difficulties in the interpretation of the subsequent analysis if unexpected interactions appear in the causal model (Lewis and John, 1976). Omitting factor levels is a useful device in constructing, for example, BIB designs based on Youden squares (Yates, 1936). However, this again is a procedure to be used with caution as main effects can easily be rendered non-orthogonal.

3. Other Facilities

KEYFINDER's design generation procedures all yield "base" designs whose rows (experimental units) appear in some sort of systematic order. KEYFINDER has extensive facilities for the subsequent

- randomization

of such designs. These include procedures for (i) randomizing the symbols 0, 1, 2, ..., s_i-1 labelling the s_i levels of each block and treatment factor, i.e. randomizing the symbols in each column of the design matrix, and (ii) randomizing the order of the experimental units, i.e. randomizing the order in which the rows of the design matrix appear. In blocked designs, separate randomizations of the experimental units may be performed within each block or within each block factor combination. If appropriate, the user may then

- sort

the experimental units into the actual order or layout in which the tests are to be performed; this order will usually be determined by the levels of the block factors.

Two designs may be stored in KEYFINDER's "memory" at any one time. These are usually the unrandomized base design and the most recent randomized/sorted plan - the working design. A design key may also be stored in memory. Either of the two designs or the design key may be

- displayed on screen

or

- stored on file

at any time.

Designs may be displayed and stored using a variety of labels and formats under the complete control of the user. This enables the user to produce test sequences ready for use by the experimenter and to export designs to statistical programs and/or Excel for subsequent data presentation and analysis.

Designs and keys may also be stored in formats suitable for subsequent

- retrieval from file

Designs can thus be (i) recreated or (ii) used as subdesigns of direct product designs. The user may also

- list files stored on disk
- inspect disk files without retrieval

and

- execute general host system commands

Once a design key is stored in memory, KEYFINDER has facilities for deriving its

- defining relation

either in whole or in part. The defining relation may then be used to derive sets of

- aliases

with the user having control both on the type (block or treatment) of aliases that are sought, and on their maximum length. The defining relation and sets of aliases may also be stored on file.

Most of the non-key design construction procedures in KEYFINDER call a search algorithm which generates

- Galois fields

and thereby

- Hadamard matrices

(Zemroch, 2015).

These facilities may be used independently of any design construction procedure. KEYFINDER may be used in interactive mode as a Galois Field calculator, removing much of the drudgery from working with these abstract structures. Square and rectangular Hadamard matrices of arbitrary dimensions may also be generated. These may subsequently be displayed on the screen, or stored on file, in a wide variety of notations.

KEYFINDER PUBLICATIONS

Zemroch, P.J. (1988). Strategies for generating blocked fractional replicate designs by computer. Computational Statistics Quarterly, **4**, 43 - 57.

Zemroch, P.J., Lunn, K., Baines, A. and Clithero, D.T. (1989). Finding design keys using Prolog. Computational Statistics Quarterly, **4**, 311 - 332.

Zemroch, P.J. (1990). KEYFINDER: A Prolog system for finding design keys - Version 2. In "COMPSTAT 1990 Software Catalogue", Atlas - Congress Department, Lastovska 23, 41000 Zagreb, 5 - 6.

Zemroch, P.J. (1991). KEYFINDER - a Prolog program for generating experimental designs. In "Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface", Interface Foundation of North America, Fairfax Station VA, 348-351.

Zemroch, P.J. (2015). The computerized generation of fractional-replicate designs using Galois fields and Hadamard matrices. Qual. Reliab. Engng. Int., 31, 1197-1207. DOI: 10.1002/qre.1846.

Zemroch, P.J. (2016). KEYFINDER – a complete toolkit for generating blocked and/or fractional-replicate factorial designs. Technical description and user guide (Version 3.3, 2016 update). Shell Global Solutions Report GS.08.52610.

OTHER REFERENCES

Addelman, S. (1962). Orthogonal main-effect plans for asymmetrical factorial experiments. Technometrics, **4**, 21-46.

Addelman, S. and Kempthorne, O. (1961). Some main effect plans and orthogonal arrays of strength two. Ann. Math. Stat., **32**, 1167-1176.

Box, G.E.P. and Hunter, J.S. (1961). The 2^{k-p} fractional factorial designs, part I. Technometrics, **3**, 311 - 351.

Dey, A. (1985). Orthogonal fractional factorial designs. Wiley Eastern, New Delhi.

Lewis, S.M. and John, J.A. (1976). Testing main effects in fractions of asymmetrical factorial experiments. Biometrika, **63**, 678-680.

Nelder, J.A. (1965). The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. Proc. R. Soc. **A**, **283**, 147 - 162.

Patterson, H.D. (1965). The factorial combination of treatments in rotation experiments. J. Agric. Sci., **65**, 171 - 182.

Patterson, H.D. (1976). Generation of factorial designs. J. Roy. Statist. Soc. **B**, **38**, 175-179.

Patterson, H.D. and Bailey, R.A. (1978). Design keys for factorial experiments. Appl. Statist., **27**, 335 - 343.

Plackett, R.L. and Burman, J.P. (1946). The design of optimum multifactorial experiments. Biometrika, **33**, 305 - 325.

Scheffé, H. (1959). The Analysis of Variance. Wiley, New York.

Taguchi, G. (1976). An Introduction to Quality Control. Central Japan Quality Control Association, Nagoya.

Wilkinson, G.N. and Rogers, C.E. (1973). Symbolic description of factorial models for analysis of variance. Appl. Statist., **22**, 392 - 399.

Yates, F. (1936). Incomplete latin squares. J. Agric. Sci., **26**, 301-315.

HOW TO GET KEYFINDER

KEYFINDER is distributed as an executable program file and runs under all versions of Windows.

Users have the option of using either the *Windows* or the *DOS* version of KEYFINDER. Both have some minor limitations. In particular, it is not possible interrupt either version during unproductive searches without aborting the program and restarting from scratch.

The Windows version is limited in that system commands cannot be issued from the KEYFINDER window, nor can lists of files be obtained. These tasks need to be performed in external Explorer or DOS windows.

To obtain a free copy of the program and manual, simply send an email request to

Statistics-Chemometrics@shell.com

stating which version you would like and your full name, affiliation and business address.