Constructing BIBDs by an Expert System

Dieter Rasch

Research Centre of Animal Production Dummerstorf-Rostock
of the Academy of Agricultural Sciences of the GDR
and
McMaster University
Department of Mathematics and Statistics
Hamilton, Ontario
CANADA

1. Introduction.

In the last two decades, papers on the construction of BIBDs and tables of such designs have appeared with an increasing frequency in combinatorial journals rather than, as was often the case in the past, in statistical journals. The users of BIBDs are, for the most part, not mathematicians; they are scientists working in research laboratories, in particular, in agriculture in variety testing. As a result of this development, the access of these users to recent results in design theory became more difficult, sometimes impossible. Even applied statisticians working in experimental design are seldom aware of the most recent results. For instance, nobody can expect a research worker in variety testing to be able to construct the smallest (v, k)-design for $k = 6$ using Hanani's 1975 paper [1].

To fill the gap between well-known theoretical results and the demand of the potential user, a computer aided approach is proposed.

2. The expert system CADEMO.

CADEMO is an abbreviation of Computer Aided Design of Experiments and Modelling. It supports research workers and statisticians acting as consultants in choosing an appropriate mathematical model for the experimental results, and also in designing their experiments. Experimental design in the broader sense (see for details Rasch and Herrendörfer [3]) means:

(i) Sample size determination for given precision requirement; in BIBD's this is the determination of r.
(ii) Optimum allocation of measurements in linear and nonlinear regression by nonlinear optimization methods.
(iii) Construction of designs with one block factor (BIBDs, PBIBDs), with more block factors (row-column designs) as well as of factorial designs.

CADEMO is a modular system which runs on IBM-compatible PCs under MS-DOS 3.2 and higher, and is described in more detail in Rasch *et al* [4]. CADEMO assumes nearly no knowledge of computers. If it is started, it leads the user interactively to the solution of his problem. A basic course in statistics is all what is needed. Many help functions and the possibility to call for an explanation from a dictionary of words like "factor", "character", "block", assist the user. One of the more than 20 module complexes of the system is the complex "Selection and Construction of Experimental Designs" — ANLA. In the present stage of the development ANLA contains three modules:

(1) Complete randomized designs and random assignment of treatments, blocks, etc. (ZUZU).
(2) Construction of balanced incomplete block designs (KBUB).
(3) Construction of factorial designs (FAK2).

3. The module "Construction of BIBDs" of CADEMO.

In practical applications v is fixed in advance, and the block size k (for instance the litter size of pigs) is also more or less fixed. The research worker is either looking for the smallest BIBD for a given (v, k)-combination or he knows from an a priori sample size determination the value of r and wants to construct the smallest BIBD with parameters v, k, r^* and $r^* \geq r$. CADEMO in the first version gives solutions for the first problem; solutions of the second problem will be available in 1990. Because computing time of modern PC-AT even in an interactive system is of no importance (it needs seconds to generate a $PG(n, q)$ even for larger n and q), the constructed designs are not stored. In CADEMO the algorithms for generating the smallest BIBD for a given pair (v, k) are programmed using the following tools:

- euclidian geometries
- projective geometries
- difference sets
- initial blocks (for cyclic designs)
- reducing designs
- complementary design
- dropping a column from a Latin Square (if $k = v - 1$)
- generating all pairs of v elements (if $k = 2$).

In selecting the suitable method the paper by Hanani [1] and the tables of Mathon and Rosa [2] were very useful. For $v \leq 25$ the user obtains a BIBD. At first he is asked to put in the value of v and the value of k. After the design is computed the values of r, b and λ are shown on the screen. Now the user can decide between an output

- on the screen (subsets of 14 blocks each)
- by printing
- to a file.

A more detailed description is available from the author.

Acknowledgement.

I would like to thank Dr. A. Rosa for valuable comments and suggestions during

my stay at McMaster. After these discussions some gaps of (v, k)-combinations could be filled.

References

1. H. Hanani, *Balanced incomplete block designs and related designs*, Discrete Mathematics **11** (1975), 255–369.
2. R. Mathon and A. Rosa, *Tables of parameters of BIBDs with* $r \leq 41$ *including existence, enumeration, and resolvability results*, Annals of Discrete Mathematics **26** (1985), 275–308.
3. G. Rasch and G. Herrendörfer, "Statistical Experimental Design — Sample Size Determination and Block Designs", Reidel Publ. Co., Dortrecht, Boston, Lancaster, Tokyo, 1986.
4. D. Rasch, V. Guiard, G. Nürnberg, E. Rudolph, and F. Teuscher, *The expert system CADEMO — Computer-aided design of experiments and modelling*, Statistical Software Newsletter **13** (1987), 107–114.