Experimental Design:
Recent Advances and Challenges for the Future

Paul Darius and Kenneth Portier

As reflected in the title, this talk has two parts. The first part reviews recent developments in experimental design, based on research published over the last three years. Given the amount and the diversity of the literature, it is a personal (and hence biased) general overview, which extends the material in section 5.3 of the Research Methodology book. The second part offers reflections on new roles for experimental design at the beginning of the twenty-first century: extrapolating trends that are already visible and identifying challenges that will have to be met. This material partly extends section 5.6 of the book.

1. Recent Developments

Because of the diversity of recent developments, we group them under five general topics. Many of the issues discussed have aspects linking them to several other topics as well. Some equally important but less mainstream issues are not mentioned here but since they are closely related to other research methodologies they may be covered in one of the other talks.

1.1. Linear/normal models.

Within the classical linear model (with i.i.d. normal errors) framework, there is still considerable interest in some types of classical designs. For fractional factorial designs, for instance, several studies discuss the existence of orthogonal arrays with particular properties, while other research examines the properties of supersaturated designs (designs with a very small number of observations, compared to the number of factors). Incomplete block designs continue to be heavily studied. Research topics include: the existence of incomplete block designs satisfying additional requirements, the construction of group divisible designs (a particular type of partially balanced incomplete block designs) and the use of partially balanced incomplete block designs in quantitative genetics and survey sampling.
Crossover designs have (as discussed in section 5.3.4) many advantages while also posing many design challenges. Since within each unit the responses are likely to be correlated, attempts are made to model this correlation structure, and to find designs that are optimal for the chosen model. Furthermore, the information on the dependence comes from many very short time series, which is rather unusual and leads to estimation problems.

The classical designs were chosen because of their ability to avoid bias and to make efficient comparisons between group means under a given model. Recently, attention has gone to developing designs for other interesting purposes. Examples are designs for simultaneously estimating both mean and variance functions, and designs for discriminating between models.
Results on sequential and group-sequential designs (including the development of early stopping rules) have also appeared. Most (but not all) of the developments in this area are related to the design of clinical trials.
In agriculture, designs are being developed to efficiently control for potential intra and inter plot competition in field experiments. We find renewed interest in the design and analysis of intercropping (tropical agriculture) experiments, in which two or more crops are grown simultaneously, one providing shelter or protection for the others.

1.2. Optimal designs

Optimal designs are experimental layouts that optimize some function of the variance of the parameter estimates. Since typically multiple parameters are to be estimated, for each method of defining overall parameter precision (variance), there is a corresponding optimal design. The general mathematical theory of optimality has been developed in great detail. Optimality criteria have been used to construct ad-hoc designs for situations where physical constraints made it impossible to use standard designs. Computer-intensive methods to construct optimal designs from a set of candidate points have also been developed but have been slow in reaching general use. In addition, in recent years more research has been devoted to ways of finding optimal designs for specific types of models. While optimality ideas have permeated a large part of the design research literature, we find little of this material in textbooks on experimental design. Examples of situations for which optimal designs have been published are: dose-response studies, particularly those utilizing logit and probit analysis models; growth studies; immunoassay studies; studies to determine the order of a series of chemical reactions; toxicokinetic studies; obtaining data for calibrating kinetic environmental models; replication-free regression studies; single and multiple blocking factor studies where block effects are assumed linear; compartmental analysis and accelerated life testing.

1.3. **Designs for non-normal/nonlinear models.**

Research on non-linear/non-normal models, and in particular the whole area of generalized linear models, has concentrated primarily on analysis problems rather than study design problems. Many of the optimal design approaches mentioned in the previous section can and have been extended to non-linear/non-normal situations. A particularly active area of research is in the construction of response surface designs for binary or Poisson data.

1.4. Designs for models with correlated error structures.

A number of studies have examined the effects of correlated errors on various design elements, and in particular on model parameter estimates when no correlation is assumed. Correlated errors can often be the consequence of random factors not being accounted for in the study design. Some attention has been directed toward the ability of staggered nested designs to mitigate this problem. Correlated errors can also be a consequence of randomization constraints, as for example occurs with bi-randomization in response surface designs. Finally, experimental units that are close together in time or space typically will display moderate to high correlations. Research on designs that best deal with spatial variability, by eliminating it at the design phase or facilitating estimation in the analysis model, continues, particularly with large agricultural field experiments.

1.5. **Bayesian approaches**.

The Bayesian approach to data analysis has evolved over thirty years from a highly theoretical subject to a practical analytic tool. Bayesian ideas have also been introduced in recent design literature. A Bayesian approach to design requires specification of a prior distribution for the parameters of the model. An optimal Bayesian design is a design that minimizes the expected value of a loss function. There is an obvious link with the optimality ideas from section 1.2 but here the optimality depends not only on the choice of the design space, but also on the prior chosen. Most of the models mentioned in section 1.2 have been studied from a Bayesian point of view. The Bayesian approach is especially appealing for nonlinear problems, where the usual optimality approach is simply not useful, but where the introduction of a prior allows construction of designs that are locally optimal.

2. **Challenges for the future**.

How will the field of statistics look like in the 21st century, and what will be the role of experimental design? Research and its associated technology are changing at a very rapid rate and statistics, as a set of tools and procedures to support this research, will have to adapt and grow to meet these changes. The direction and magnitude of change is already visible, and a few of the major directions are introduced in the next paragraphs. It is clear that new challenges will, no doubt, emerge in the years to come.

In chemistry, the area of chemometrics is developing rapidly. Chemometrics could be defined as statistical methodology specially adapted to the intricacies of analytical chemistry experimentation and data. Although most of the attention has been directed on data analysis, some new design research has been published. The nature of some of the data and the acquisition constraints make this area a fertile ground for new design research.

In biology, sequence and genome data are being collected at a fantastic rate. Computational biologists analyze these data by studying pattern count statistics and comparing observed sequences in order to find "similar" patterns. Transcription errors (insertions/deletions) lead to "error" or "noise" which in turn create special statistical problems. Limits on levels of control and manipulation in this research make the role of experimental design in this area unclear.

In agriculture, advances in sensor technology and satellite communication have led to the new field of study generally classified as precision agriculture. Using GPS technology, current agricultural machines (e.g. spreaders, sprayers, graders, harvesters) can record their location in the field while simultaneously measuring and controlling process parameters (e.g. rate of fertilizer applied). Massive amounts of data are generated, often contaminated with errors from multiple sources, that are typically summarized geographically in a map. If the precision of these machines can be made high enough, we will see a revolution in the area of field experimentation. Experiments will be less expensive and capable of being performed on the farm, providing the producer the ability to fine tune their operation and optimize it for their local microclimate and soil conditions.

Data mining is the major new analysis tool in marketing and management sciences. Data mining refers to the search for previously unsuspected relationships of interest in the very large databases typically created as a part of standard business operations, e.g. financial consumer transactions, healthcare data, and internet website clickstream data . Much of this activity has centered around data not typically collected as part of a formal study, hence the usefulness of experimental design in this context is not clear.

Although a continuous stream of new methodological advancements and new areas of application is vital, the ultimate role of experimental design techniques in the 21st century will depend on the willingness of the future researcher to apply them. Here we see particular challenges in education and software development.

Education has a very long-range effect. Many of the people who should be using experimental design in the first part of the 21^{st} century have already received their formal statistical education. Since much of today's teaching is done with textbooks that were written 10 to 20 years ago, much of the new methodologies are not being introduced to students. Many of the currently available experimental design textbooks are cookbooks, depicting study design as choosing from a suite of recipes that are amenable to known, simple, and straight-forward analysis. Second, in many introductory textbooks, statistics is still presented (at least implicitly) as an activity that is only introduced once data have been collected. Some of the more recent texts spend more effort on discussing the basic components and characteristics of a good experimental design, the need to meet study objectives and less on the standard patterns. Since more flexible data analysis tools (general linear mixed models, generalized models, generalized estimating equations, etc.) are now available, the need for researchers to stick to just the standard design prototypes is lessened. This makes the work of teachers of experimental design, in some sense, more difficult because the choice of design is no longer made on the criterion of ease of analysis, but more on the objectives and resources available for the research. The educational challenge then is to continue to combine study of experimental design with that of data analysis approaches while not having the analysis drive the design decision.

The other challenge is software development: without the availability of good and easy to use software, even the best design methods will not see wide use in research. While we have already elaborated on the requirements for experimental design software in section 5.6 of the book, here we add two more requirements.

The first requirement has to do with better communication facilities. The Internet has thoroughly influenced the way researchers use computers, and will certainly continue to do so in the future. Eventually this will influence the way statistical software is used as well. We can imagine situations where we might need to send a dataset to a particular server for a specialized analysis, or even where a complete statistical consultation is performed over the Internet. The whole success of the Internet is based on the acceptance of standards for communication. Statistics, on the other hand has had its success despite failing to adopt clear standards for communicating statistical data, and not providing standards for communicating associated metadata , including study design parameters, analysis models, assumptions, findings, etc. While we do not see much progress being made in this direction, we hope that new emerging web standards such as the XML protocols will offer opportunities for development of these standards.

Even harder than the adoption of communication standards for statistical data and analysis is the requirement for codifying and interpreting the background knowledge about the subject of investigation in such a way that it can be used to guide the experimental design. It would be extremely useful to have software that not only can store such knowledge (in the form of metadata), but also actively uses this knowledge in aiding the researcher in choosing among candidate designs. Such expert design systems seem inevitable but only feasible in the distant future.