Boletín de Estadística e Investigación Operativa
Vol. 25, No. 1, Febrero 2009, pp. 14-21

ESTADÍSTICA

A critical overview on optimal experimental designs

Jesús López-Fidalgo
Department of Mathematics - Institute of Mathematics Applied to Science and Engineering
University of Castilla-La Mancha
✉ jesus.lopezfidalgo@uclm.es

Abstract

Although the classic theory of experimental design is well known, this is not the case when the word "optimal" is added. A short introduction to optimal experimental design is made through the consideration of some typical criticisms to this theory.

Keywords: Approximate design, equivalence theorem, exact design, information matrix, optimality criteria.
AMS Subject classifications: 62K05.

1. Introduction

Everybody knows what the classic experimental design theory is, but it can be said that the optimal experimental design theory is not well known even among statisticians. The classic experimental design is looking for good designs, studying their properties and trying to generate some types of designs. Meanwhile, the optimal design theory aims the best possible design, mainly from a regression setup, including somehow classic designs (e.g. Dorta-Guerra, González-Dávila and Ginebra, 2008). On the other hand there is a number of criticisms around this theory, some of them coming from practitioners, some coming from statisticians. The aim of this paper is to offer a simple introduction to the topic through these critical views.

The main bases of the theory are being introduced briefly in what follows. So far, let the response of an experiment be expressed with a linear model,

$$E(y) = f^T(x)\theta, \quad \mathrm{var}(y) = \sigma^2, \quad x \in \chi,$$

where $f^T(x) = (f_1(x), \ldots, f_k(x))$ is a vector of continuous linearly independent known functions defined in a compact set χ, $\theta^T = (\theta_1, \ldots, \theta_k)$ are the unknown parameters to be estimated and σ^2 is the constant variance. The responses are assumed normal and uncorrelated.

If the conditions of the experiment, x, are under the control of the practitioner an experimental design can be built in advance. Thus, an *exact design* will be a sequence of experimental conditions, $x_1, \ldots, x_n$, from a compact set χ. Assuming that only m of these points are different, a probability measure may be assigned to the design. If point x_i appears n_i times then $p_i = n_i/n$ will be the probability of x_i, that is the proportion of experiments to be made under these conditions.

Using this idea Kiefer (1959) gave a more general definition of design (*approximate design*) as any probability distribution ξ on the space χ. The *information matrix* for a linear model is defined as:

$$M(\xi) = \int_\chi f(x) f^T(x) \xi(dx).$$

The inverse of the information matrix is proportional to the covariance matrix of the least square estimates. Thus, an experimental design "minimizing", in some sense, the inverse of the information matrix, should be found. An *optimality criterion* for choosing a design is given by a function $\Phi : \mathcal{M} \longrightarrow \mathbf{R} \cup \{+\infty\}$, such that Φ is convex, non–increasing (if $M - N$ is non negative definite then $\Phi(M) \leq \Phi(N)$) and positively homogeneous in the sense that $\Phi(\delta M) = \delta^{-1}\Phi(M)$, $\delta > 0$. A Φ*–optimal design* is a design ξ^* minimizing Φ.

The set of the information matrices, $\mathcal{M}$, is convex and compact. The Caratheodory's theorem guarantees that, given an information matrix, there always exists a design with this information matrix and no more than $k(k+1)/2+1$ points in its support. This is very convenient from the practical point of view since the experimentation may be performed just in a finite number of cases. Thus, the search for optimal designs may be restricted to designs ξ, with a finite support,

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \ldots & x_m \\ p_1 & p_2 & \ldots & p_m \end{array} \right\},$$

where $\xi(x_i) = p_i$.

The efficiency of a design ξ with respect to the criterion Φ will be

$$\mathrm{eff}(\xi, \Phi) = \frac{\Phi(M(\xi^*))}{\Phi(M(\xi))}.$$

Thus, if a design has 50% efficiency then a half of the observations with the optimal design will produce the same results with respect to the criterion Φ. This is something a practitioner can understand much better that p-values or test power.

Computing optimal designs is not an easy task. For approximate designs and uncorrelated observations there is the so called *equivalence theorem*, which is the basis of most of the algorithms for computing designs as well as checking how efficient a particular design is. This theorem, based on checking the directional derivatives, has been extended to some other cases. There is some current work trying to do something similar for exact designs.

The most popular criterion is D-optimality, which looks for the maximization of the determinant of the information matrix. The so called D-optimal design minimizes the volume of the confidence ellipsoide of the parameters in the model. This criterion is invariant for re-parameterizations. Another popular criterion is A-optimality, which considers the mean of the variances of the estimates,

$$\Phi_A(M(\xi)) = \frac{1}{k}\mathrm{tr}M^{-1}(\xi).$$

There are quite a few more, but there is not space here to define all of them.

All this theory has been formulated in this paper for linear models, but extensions to nonlinear models can be done through the Fisher Information Matrix and the asymptotic approximation of its inverse to the covariance matrix. For an introduction to the topic there is a number of classic books (Fedorov 1972, Silvey 1980, Pázman 1986, Atkinson and Donev 1992, Pukelsheim 1993, Fedorov and Hackl 1997) as well as more recent books devoted to particular aspects of the theory, not mentioned here.

2. Some criticisms

2.1. Model dependence

One of the main criticisms to the optimal experimental design theory is that a model has to be selected a priori without any data yet. This is a deeper problem than it seems to be at the first glance. Frequently there is a strong dependence between the model and the optimal design. Thus, a design may be rather good for a particular model and pathetic for a different one, which may finally be proved more appropriate for the current data.

Box used to write frequently statements like: “Models, of course, are never true, but fortunately it is only necessary that they be useful” (Box, 1979). This idea does not solve the problem mentioned in the last paragraph, but stresses the truth that this is not only a problem of designing an experiment. In any case, the experiment needs to be designed before having the observations and the problem has to be undertaken with the tools at hand at that moment. In practice, there is always some experience or retrospective data one can trust. There are also the intuitions of the practitioner. Even some models are analytically derived, e.g. solving a differential equation system.

Moreover there is a particular interest nowadays in developing optimality criteria to discriminate between rival models (e.g. López-Fidalgo, Tommasi and

Trandafir, 2007, 2008) when there are several candidates to be the best model.

2.2. Information matrix for nonlinear models

The optimal design theory is clean for linear models, but most of the models used in practice are not linear. The theory, including the equivalence theorem, still applies for a variety of nonlinear models, but the information matrix will depend on the "no yet" estimated parameters. This is not a negligible objection at all, but there are several reasonable ways to deal with the problem:

- Locally Φ–optimal designs may be computed depending on nominal values of the parameters. Some times explicit optimal designs may be given depending on generic values of the parameters, but most in the cases numerical calculus is performed and some explicit numerical values of the parameters have to be used in the computations. In any case, a sensitivity analysis should be carried out to be sure the design is not going to change much with possible errors in the choice of the parameters.

- Going further with the last idea minimax designs may be computed to be safer.

- An approach of increasing interest is the adaptive design idea (e.g Moler, Plo and San Miguel, 2006), were at each step the designs used the observations obtained previously. The parameters are estimated in each step and thus the dependence on the nominal values of the parameters becomes less and less important.

- Another typical approach to this problem is the use of some kind of Bayesian designs.

2.3. Criterion selection

As mentioned in the Introduction, there is a number of different criteria, sometimes even parametric classes of them (see e.g. López-Fidalgo and Rodríguez–Díaz, 2004) pursuing different aims. In practice, a few optimality criteria are really used and the choice of one of them is not a big deal. Moreover, the original equivalence theorem, much more restricted than the general one, proved the equivalence (so the name) of D-optimality and a criterion to minimize the variance of the predictions (G–optimality). This means the optimal designs are not always so far for different criteria. Even more, sometimes there exist designs universally optimal (Harman, 2008). Nevertheless, if there is interest in different criteria producing different optimal designs, compound criteria may be used in order to come to a compromise for the optimality (Cook and Wong, 1994).

2.4. Controversy exact versus approximate designs

Approximate designs are quite convenient from a theoretical and computational point of view. But to be implemented in practice rounding off is needed with the corresponding loss of efficiency of the design. On the contrary, exact designs are practicable but of very difficult computation. Box has never accepted the use of approximate designs of Kiefer (1959). This controversy should not affect the development of the theory. After years of experience in the area it may be said that exact designs are needed, and less difficult to compute, for small sample sizes (Pukelsheim and Rieder 1992, Imhof, López–Fidalgo and Wong, 2001). For large samples the approximate designs may be rounded off in an efficient way.

2.5. Frequently, optimal designs demand extreme conditions

A typical situation in statistics is that extreme conditions in the experiments offer, "theoretically", more information to make decisions. But this may be unaffordable, toxic, dangerous or just awful for the practitioners. Even more, if the optimal design reduces to a few points, less than what they would like, they would reject any use of it. This is absolutely true and the statistician has to be very careful with this aspect. Frequently, the optimal design has to be considered as a reference to measure the efficiency of the designs they use in practice or to choose the best among a class of designs they like. On the other hand, the criteria used may restrict the search to a class of designs in order to preserve the requirements of the experimentalists.

2.6. Difficult computation

Computation of optimal designs is not an easy task in general. As a matter of fact, the search for optimal designs frequently restricts to one-dimensional models, although some work has been performed with more complex models (e.g. López–Fidalgo and Garcet–Rodríguez, 2004). There is an increasing interest in developing good algorithms to compute designs, either exact or approximate designs (see e.g. López–Fidalgo and Rodríguez–Díaz, 2004 or Martín–Martín, Torsney and López–Fidalgo, 2007). One may think the people working on optimal design must be good in optimization. They are not bad, but they are not experts in the topic. At the same time, people in optimization are sometimes far from statistics and even more from experimental designs. Therefore, there is a need of more cooperation between them.

2.7. Scale problem

Some criteria are not invariant with respecto to re-parameterizations. This means that the scale of a parameter may be much bigger than the scale of another parameter in the model, causing different magnitudes of the variances of their estimators. Therefore, the criterion may not pay attention enough to the small variance in magnitude, but equally important in the inference. This is the case of A–optimality, among other criteria. This problem requires special care and some standardization. Different solutions have been given to this problem. For instance, standardized optimality criteria by the efficiencies of each parameter (Dette, 1997) produce similar final efficiencies for estimating each parameter of the model regardless the magnitude of the variances of the estimators. Another possible standardization is by the coefficient of variation (López–Fidalgo and Rivas-López, 2007; López–Fidalgo, Rivas-López and Fernández-Garzón, 2007). This last approach adds a dependence on the parameters, that for non-linear models is already there.

3. Concluding remarks

One of the main advantages for an optimal design research fellow is that he or she walks around many different topics in Statistics and even in other areas of Mathematics. Thus, optimal designs may be an objective for any kind of model such as survival analysis or reliability, models with correlated observations, kinetics and compartmental models, mixed models, mixture of distributions, models of mixtures, censoring and potential missing data, model discrimination, mathematical programming, even algebraic geometry,... This means one is frequently introduced in a new area without changing the main area of research. Even if the optimal designs are not finally used this theory helps in understanding more deeply the estimation errors and correlations between them. But one of the main tasks of the statisticians is to convince the practitioner they need to design their experimentation correctly and efficiently apart from doing the correct statistical analysis.

Acknowledgments

This work was sponsored by Ministerio de Educación y Ciencia and fondos FEDER (MTM2007 67211 C03-01), Junta de Comunidades de Castilla-La Mancha (PAI07-0019-2036) and the University of Castilla-La Mancha.

References

[1] Atkinson, A.C. and Donev, A.N. (1992). *Optimum experimental design*, Oxford Science Publications, Oxford.

[2] Box, G.E.P. (1979). Some problems of statistics and everyday life. *Journal of the American Statistical Association*, **74**, 1-4.

[3] Cook, R.D. and Wong, W.K. (1994) On the Equivalence of Contrained and Compound Optimal Designs. *Journal of the American Statistical Association*, **89(426)**, 687–692.

[4] Dette, H. (1997). Designing Experiments with Respect to “Standardized” Optimality Criteria. *Journal of the Royal Statistical Society, Series B*, **59(1)**, 97–110.

[5] Dorta-Guerra, R., González-Dávila, E. and Ginebra, J. (2008). Two-level experiments for binary response data. *Computational Statistics and Data Analysis*, **53(1)**, 196–208.

[6] Fedorov, V.V. (1972). *Theory of optimal experiments*, Academic Press, New York.

[7] Fedorov, V.V. and Hackl, P. (1997). *Model-Oriented Design of Experiments*, Springer, New York.

[8] Harman, R. (2008). Equivalence theorem for Schur optimality of experimental designs. *Journal of Statistical Planning and Inference*, **138**, 1201–1209.

[9] Imhof, L., López–Fidalgo, J. and Wong, W. K. (2001). Efficiencies of rounded optimal approximate designs for small samples. *Statistica Neerlandica*, **55**, 301–315.

[10] Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society, Series B*, **21**, 272-319.

[11] López–Fidalgo, J. and Garcet–Rodríguez, S. (2004). Optimal Experimental Designs When Some Independent Variables Are not Subject to Control. *Journal American Statistical Association*, **99(468)**, 1190–1199.

[12] López–Fidalgo, J. and Rivas–López, M.J. (2007). MV-Optimality Standardized through the Coefficient of Variation. *Journal of Statistical Planning and Inference*, **137**, 2680–2007.

[13] López–Fidalgo, J., Rivas–López, M.J. and Fernández-Garzón, B. (2007). A-Optimality Standardized Through the Coefficient of Variation. *Communications in Statistics. Theory and Methods*, **36(4)**, 781–792.

[14] López–Fidalgo, J. and Rodríguez–Díaz, J.M. (2004). Elfving method for m-dimensional models. *Metrika*, **59**, 235–244.

[15] López–Fidalgo, J., Tommasi, Ch. and Trandafir, C. (2007). An optimal experimental design criterion for discriminating between non-Normal models. *Journal of the Royal Statistical Society, Series B*, **69(2)**, 231–242.

[16] López–Fidalgo, J., Tommasi, Ch. and Trandafir, C. (2008). Optimal designs for discriminating between some extensions of the Michaelis–Menten model. *Journal of Statistical Planning and Inference*, **138**, 3797 – 3804.

[17] Martín–Martín, R., Torsney, B. and López–Fidalgo, J. (2007). Construction of Marginally and Conditionally Restricted Designs using Multiplicative Algorithms. *Computational Statistics and Data Analysis*, **51**, 5547–5561.

[18] Moler, J.A., Plo, F. and San Miguel, M. (2006). An adaptive design for clinical trials with non-dichotomous response and prognostic factors. *Statistics and Probability Letters*, **76**, 1940–1946.

[19] Pazman, A. (1986). *Foundations of optimum experimental design*, D. Reidel publishing company, Dordrecht.

[20] Pukelsheim, F. (1993). *Optimal Designs of Experiments*, John Wiley and Sons, New York.

[21] Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika*, **79(4)**, 763–770.

[22] Silvey, S.D. (1980). *Optimal design*, Chapman and Hall, London.

About the author

Jesús López Fidalgo is full professor at the University of Castilla-La Mancha. He has been Postdoctoral Fellow at the University of Manchester (1992), Visiting Scholar in the Department of Biostatistics at the University of California, Los Angeles (1998/99) and Visiting Professor in the Department of Statistics at the University of California, Riverside (2005). He is elected member of the ISI and he was the last Editor of the Spanish Statistical and Operational Research Bulletin. Recently, he was elected as Dean of the High Industrial Engineering School of his University. His main interest is in Optimal Experimental Design as well as Applied Statistics. He has published papers in high standard journals as JASA, JRSSB or Bioinformatics.