Compound Design Criteria : Model and Variance

Youngil Kim
Dept. of Information System
ChungAng University
Ahnsung, Kyunggi-Do, Korea
yik01@cau.ac.kr

1. Introduction

Box and Draper(1975) listed some properties of a design that should be considered in design selection. But it is impossible that one design criterion from optimal experimental design theory reflects many potential objectives of an experiment, because the theory was originally based on the underlying model and its strict assumption about the error structure. Therefore, when it is necessary to implement multi-objective experimental design, it is a common practice to balance out the several optimal design criteria so that each design criteria involved benefits in terms of its relative "high" efficiency. In this study we proposed several composite design criteria taking the case of heteroscedastic model. When the heteroscedasticity is present in the model, the well known equivalence theorem between D- and G-optimality no longer exists and furthermore their design characteristics are sometimes drastically different. We introduced three different design criteria, constrained design, combined design, and mini-max design criteria for this purpose. While the first two methods do reflect the prior belief of experimenter, the last one does not take it into account, which is sometimes desirable.

2. Preliminaries

We assume the general linear regression model. Specifically, let $f_1, \cdots, f_p$ be p given linearly independent continuous regression function defined on some compact design subspace Ω of R^k and let $f^T(x) = (f_1(x), f_2(x), \cdots, f_p(x))$. For each x in Ω, a univariate response variable $y(x)$ is observed under the model

$$y(x) = f^T(x)\beta + e/\lambda(x)^{1/2}.$$

where $\beta^T = (\beta_1, \beta_2, \cdots, \beta_p)$ are the parameters and the $e's$ are uncorrelated real-valued error random variables having mean zero and constant variance. It is also assumed that $\lambda(x)$ is a known, bounded, positive real-valued continuous function defined on Ω. The function $\lambda(x)$ is commonly called the efficiency function and it reflects the heteroscedastic structure in the model. When $\lambda(x) = 1$, we have a homoscedastic model.

A design problem can be described as a probability mass function ξ that places total mass on a finite collection of s points in the design region Ω If the design has a mass ξ_i at x_i and

n observations are allowed for the experiment, approximately $n\xi_i$ observations are taken at x_i. The x_i's are called the support points. We ignore the constraint that the $n\xi_i$ be an integer value here in this paper. A common measure of the information contained in a design ξ is the information matrix of the design defined by

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

We are primarily concerned with nonsingular designs, designs whose information matrices are nonsingular. A design problem is characterized by the triplet $(\Omega, f(x), \lambda(x))$ together with an optimality criterion function ϕ, which is selected to reflect the experimenter's interest. Virtually all optimality criteria characterize the worth of a design through a concave functional ϕ that depends only on the information matrix defined above. The problem confronting the experimenter is how to select a design that maximizes ϕ over the space of the information matrices. Perhaps the most common criterion is D-optimality, which corresponds to $\phi(\xi) = \log|M(\xi)|$. The unknown parameters β are estimated by the least squares method and the variance function of the fitted value at the point x using design ξ is,

$$d(x,\xi) = f^T(x)M^{-1}(\xi)f(x).$$

There is a group of criteria which provide designs minimizing some functions of the expected square error of the fitted curve. The most popular one is G-optimality minimizing the maximum variance function, $d(x,\xi)$. Kiefer and Wolfowitz (1959) provided an equivalence theorem between D- and G-optimality when $\lambda(x)=1$. But the equivalence theorem is irrelevant when the error structure is not constant. Recently, Cook and Wong(1993) listed the sufficient conditions for the heteroscedastic G-optimality, which is still not practical to be implemented computationally. In other words, when the experimenter has two objectives in mind, D- and G-optimality under the circumstances of heteroscedastic error structure, there is a need to balance out those two criteria.

3. Main Results

The approach suggested here, mini-max one is conservative, compared to the previously defined approaches, constrained and composite designs. But the experimenter does not need to provide the user-specified constant. This result can be extended to the case when the experimenter has more than two objectives, which is under investigation.

REFERENCES

Kiefer, J. and J. Wolfowitz (1960), The equivalence to two extreme problems, *Canadian J. of math.*, **12**, 363-366.

Cook, R. D. and W. K. Wong (1993), On the equivalence of constrained and compound optimal designs, *J. of the American Statistical Association*, **89**, 687-692.

Box, G. E. P. and N. R. Draper (1975), Robust Designs, *Biometrika*, **62**, 347-352.