Introduction to Quantile Dispersion Graphs for Evaluating and Comparing Designs for Logistic Regression Models

by

Kevin Robinson

and

André I. Khuri

- The key ideas of *response surface methodology* (RSM) were initially developed using classical linear models under the normality assumption (Box and Wilson, 1951).
- The rapidly growing field of biostatistics and the pharmaceutical industry, where nonnormal response variables occur frequently, have caused significant changes in RSM making it more adaptable to the new environments.

  A key element affecting these changes is the introduction of *generalized linear models* (GLMs).
- Unfortunately, the use of GLMs in RSM, particularly in the design area, is still in its infancy.

Reasons

1. The problem of dependency of the design on the unknown parameters of the model.

The Design Dilemma Concerning Generalized Models (GLMs)

Unlike linear models, designs for GLMs depend on the unknown parameters of the fitted model.

Dilemma:
The purpose of a design is to get "good" estimates of the parameters.

"You tell me the value of $\boldsymbol{\beta}$ and I promise to design the best estimate for estimating $\boldsymbol{\beta}$" *Cochran (1973).*

Common approaches to solving the design dependence problem:

(a) Supply "good" initial values of the parameters, and use them in an appropriate design criterion (e.g., local D-optimality).

(b) Apply a sequential procedure. The experiment is run in stages. Updated estimates of the parameters are obtained at the end of each stage.

(c) Apply a Bayesian approach by assuming a prior distribution on the parameters. Incorporate this prior into an appropriate design criterion by integrating it over the prior distribution (see Chaloner and Verdinelli, 1995).

2. Most available design criteria

(a) are for models with one control variable

(b) rely on a single-valued index that measures variance-based efficiencies such as D-efficiency and G-efficiency (bias is ignored)

(c) do not consider the quality of prediction throughout the experimental region

(d) do not provide information concerning the degree of robustness of a design to the true values of the unknown parameters in the model.