Statistics and Applications
Vol. 11, Nos. 1 & 2, 2013 (New Series), pp. 1-13

Some finer aspects of the de la Garza phenomenon: a study of exact designs in linear and quadratic regression models

Bikas K. Sinha[1], Nripes K. Mandal[2], Manisha Pal[3]
[1] *Retired Faculty, Indian Statistical Institute, Kolkata, India*
[2] *Calcutta University, Kolkata, India*
[3] *Calcutta University, Kolkata, India*

Abstract

The well-known de la Garza Phenomenon [de la Garza, 1954] relates to the information matrix of the parameters in a standard Gauss-Markov linear model involving a single covariate in polynomial regression. It works well in the framework of approximate or continuous designs. For discrete or exact designs, one has to be careful in extracting its full spirit. We propose to discuss some features of this highly fascinating area of research.

Key words Linear models, regression designs, continuous designs, exact designs, information matrix, equivalence, domination.

1 Introduction

In the context of optimality studies in regression designs, the importance of de la Garza phenomenon has been emphasized repeatedly in various forms and settings. The most general known application is in the context of polynomial regression models involving a single covariate. Vide Pukelsheim (1993, 2006). Less known are most recent studies in non-linear parameter settings. Vide Yang (2010). In between, Liski et al (2002), Pukelsheim (1993, 2006), Khuri et al (2006) and Yang and Stufken (2009) have dealt with this phenomenon and discussed its importance in the search for optimal designs. However, all these studies [as also the original article by de la Garza in 1954, at least in spirit] relate to what is known as approximate or, continuous design set-up. In this article our focus is exclusively on exact or discrete design set-up.

Consider the problem of selecting an experimental design to furnish information on the parameters in a linear model $y = f'(x)\beta + e$, where y is the observable response variable, β is a $p \times 1$ vector of unknown parameters; the errors are assumed to be independent and identically distributed with mean 0 and constant variance σ^2, and f is a $p \times 1$ vector-valued known continuous function of the $q \times 1$ vector of design variables x that is constrained to lie in a compact subset $\chi(x)$ of R^q. An experimental design can be described as a probability mass function ξ that places total mass on a finite collection of s points in the design region $\chi(x)$. In exact [or, small sample i.e., discrete] design, $n(x_i)$ is required to be an integer for $i = 1, ..., s$, where n is the total sample size and $n(x_i)$ is the number of

Soc. Series B. **54**, 569-583.

Fornius, E. F. and Nyquist, H. (2009). Using the Canonical Design Space to Obtain c-optimal Designs for the Quadratic Logistic Model. *Comm. Statist. Th. and Meth.*, **39**, 144-157.

Guiard, V. and Pichlmeier,U. (1999). Critical assessment of the c-optimality design criteria for estimating the median effective dose in quantal dose-response curves. *Biometrical J.* **41**, 799-816.

Hedayat, A. S., Yan, B. and Pezzuto, J. M. (1997). Modeling and identifying optimum designs for fitting dose-response curves based on raw optical density data. *J. Amer. Statist. Assoc.* **92**, 1132-1140.

Khan, M. K. and Yazdi, A. A. (1988). On D-optimal designs for binary data. *J. Statist. Plann. Inf.* **18**, 83-91.

Khuri, A. I., Mukherjee, B., Sinha, B. K. and Ghosh, M. (2006). Design issues for generalized linear models: A review. *Statist. Sci.* **21**, 376-399.

Liski, E., Mandal, N. K., Shah, K. R. and Sinha, Bikas K.(2002). *Topics in optimal design.* Springer-Verlag Lecture Notes in Statistics Series. **163**.

Mathew, T. and Sinha, B. K. (2001). Optimal designs for binary data under logistic regression. *J. Statist. Plann. Inf.* **93**, 295-307.

Minkin, S. (1987). Optimal design for binary data. *J. Amer. Statist. Assoc.* **82**, 1098-1103.

Nandy, K. and Nandy, R. (2013a). A study of the performance of two-stage optimal designs in a logistic dose-response model. *Private communication.*

Nandy, K, and Nandy, R. (2011, 2013b). Some aspects of optimal designs for quadratic dose-response models. Presented at Statistical Meeting in Halifax, Canada. *Private Communication.*

Pukelsheim, F. (1993). *Optimal Design of Experiments.* Wiley, New York; Also (2006). SIAM, Philadelphia, PA.

Silvey, S. D. (1980). *Optimal Design.* Chapman and Hall, London.

Sitter, R. R. and Wu, C.F.J. (1993). Optimal designs for binary response experiments: D- and A- criteria. *Scand. J. Statist.* **20**, 329-342.

Wu, C.F.J. (1988). Optimal design for percentile estimation of a quantal response curve. In: Dodge, Y., Federov, V. and Wynn, H.P. (Eds.), *Optimal Design and Analysis of Experiments.* Elsevier, Amsterdam, 213-223.

Yang, M. and Stufken, J. (2009). Support points of locally optimal designs for non-linear models with two parameters. *Ann. Statist.* **37**, 518-541.

Yang, M. (2010). On the de la Garza phenomenon. *Ann. Statist.* **38**, 2499-2524.