



ELSEVIER

Computational Statistics & Data Analysis 43 (2003) 47–62

COMPUTATIONAL
STATISTICS
& DATA ANALYSIS

www.elsevier.com/locate/csda

Quantile dispersion graphs for evaluating and comparing designs for logistic regression models

Kevin S. Robinson^a, André I. Khuri^{b,*}

^aDepartment of Statistics, The University of Akron, Akron, OH 44325-1913, USA

^bDepartment of Statistics, University of Florida, P.O. Box 118545, 103 Griffin-Floyd Hall, Gainesville, FL 32611-8545, USA

Received 1 May 2001; accepted 1 June 2002

Abstract

Designs for fitting a generalized linear model depend on the unknown parameters of the model. The use of any design optimality criterion would therefore require some prior knowledge of the parameters. In this article, a graphical technique is proposed for comparing and evaluating designs for a logistic regression model. Quantiles of the scaled mean-squared error of prediction are obtained on concentric surfaces inside a region of interest, R . For a given design, these quantiles depend on the model's parameters. Plots of the maxima and minima of the quantiles, over a subset of the parameter space, produce the so-called quantile dispersion graphs (QDGs). The plots provide a comprehensive assessment of the overall prediction capability of the design within the region R . They also depict the dependence of the design on the model's parameters. The QDGs can therefore be conveniently used to compare several candidate designs. Two examples are presented to illustrate the proposed methodology.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Binary response; Design dependence on unknown parameters; Generalized linear models; Prediction bias; Response surface methodology; Scaled mean-squared error of prediction

1. Introduction

Since the introduction of generalized linear models (GLMs) by Nelder and Wedderburn (1972), the focus of attention in optimal design theory has shifted from the traditional linear model to these more general models. However, the main design optimality criteria, such as A -, D -, E -, and G -optimality, that is, the so-called alphabetic

* Corresponding author. Tel.: +1-352-392-1941x238; fax: +1-352-392-5175.

E-mail address: ufakhuri@stat.ufl.edu (A.I. Khuri).

optimality, remain the same as in linear models. These criteria focus entirely on the precision of parameter estimates, which is measured by the asymptotic variance–covariance matrix of the estimates. But, unlike linear models, designs for GLMs depend on the unknown parameters of the fitted model. This dependence problem causes great difficulty in the construction and evaluation of designs. Common approaches to solving this problem include sequential generation of designs and the use of the Bayesian methodology. In the first approach, initial values of the parameters are used as “best guesses” to determine a “locally” optimal design. Response values are then obtained on the basis of the generated design. In the next stage, updated estimates of the parameters are developed and then used to determine additional design points, and so on. The implementation of such a strategy is feasible provided that the response values in a given stage can be obtained in a short time, as in sensitivity testing (see Young and Easterling, 1994). In the Bayesian approach, a prior distribution is assumed on the parameters, which is then incorporated into an appropriate design criterion by integrating it over the prior distribution. For example, one criterion maximizes the average over the prior distribution of the logarithm of the determinant of Fisher information matrix. This criterion is equivalent to D -optimality in linear models. Bayesian versions of other alphabetic optimality criteria can also be used such as A -optimality. The Bayesian approach was discussed by several authors including Chaloner (1987), Chaloner and Larntz (1989, 1991), Mukhopadhyay and Haines (1995), and Chaloner and Verdinelli (1995).

In this article, we consider the problem of discriminating among designs for logistic regression models. These models are appropriate for binary data situations which are frequently encountered in dose–response, quantal response, and success–failure experiments (see, for example, Carter et al., 1986; Vidmar et al., 1992; Piegorsch, 1998).

Since in small samples the parameter estimates for a logistic regression model are biased, we consider the mean-squared error of prediction (MSEP) as a criterion for comparing designs. The MSEP incorporates both the variance and bias associated with the estimated mean response. As in any other design criterion for GLMs, the MSEP depends on the unknown parameters of the model. A graphical procedure is introduced for comparing designs using the so-called quantile dispersion graphs (QDGs) of the MSEP. These graphs provide an assessment of the overall prediction capability of a given design through a visual display of the MSEP. They also give a clear depiction of the dependence of the design on the model’s parameters. Khuri and Lee (1998) used a graphical procedure for comparing designs for nonlinear models. The proposed graphical procedure generalizes the work of Khuri and Lee (1998) by addressing non-normality and nonconstant error variance.

The remaining portion of this article is organized as follows: in Section 2, a general expression for the MSEP is developed for GLMs. An application of this expression to logistic regression models is presented in Section 3. The graphical approach based on the QDGs is described in Section 4. Two examples are presented in Section 5 to illustrate the use of the QDGs. Finally, concluding remarks are given in Section 6.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات