ASYMPTOTICS FOR ROBUST SEQUENTIAL DESIGNS IN MISSPECIFIED REGRESSION MODELS

SANJOY SINHA
University of Winnipeg

DOUGLAS P. WIENS
University of Alberta

We revisit a proposal, for robust sequential design in the presence of uncertainty about the regression response, previously made by these authors. We obtain conditions under which a sequence of designs for nonlinear regression models leads to asymptotically normally distributed estimates. The results are illustrated in a simulation study. We conclude that estimates computed after the experiment has been carried out sequentially may, in only moderately sized samples, be safely used to make standard normal-theory inferences, ignoring the dependencies arising from the sequential nature of the sampling. The quality of the normal approximation deteriorates somewhat when the random errors are heteroscedastic.

1. Introduction

In a recent article (Sinha and Wiens 2002, henceforth referred to as SW), we applied notions of robustness of design in the presence of response uncertainty in a nonlinear regression setting. We developed and implemented an algorithm for the sequential selection of design points $\mathbf{x}$, from a specified "design space" $\mathcal{S} \subset \mathbb{R}^q$, at which to observe a random variable Y. This random variable was assumed to follow a regression model with a nonlinear and possibly misspecified response function. The sequential sampling scheme used in SW—and described below—is somewhat involved. Asymptotic normality of the resulting estimates was posited, and tested in a simulation study. In this article we shall fill in some of the theoretical gaps left by SW, and then revisit the aforementioned nonlinear regression application.

We begin by describing in detail the application which motivates the current study. We entertain a sequence of nonlinear regression problems indexed by the sample size n. At the nth stage it is supposed that one samples from a distribution P^n of r.v.s Y_n, with means depending on $\mathbf{x}$ through an unknown parameter vector $\boldsymbol{\theta}_n \in \mathbb{R}^p$ and a nonlinear function of $\mathbf{x}$ and $\boldsymbol{\theta}_n$ ranging over a neighbourhood of a tentative choice f:

$$E[Y_n \mid \mathbf{x}] \approx f(\mathbf{x}; \boldsymbol{\theta}_n). \tag{1.1}$$

For instance the experimenter may fit a Michaelis-Menten response $f(x; \boldsymbol{\theta}) = \theta_0 x/(\theta_1 + x)$ when in fact the true response is exponential: $E[Y \mid \mathbf{x}] = \theta_0(1 - e^{-\theta_1 x})$. (We shall return to this particular example in §3.1 below.)

AMS subject classifications: Primary 62L05, 60F05; Secondary 62J02..

Keywords and phrases: asymptotic normality; biased regression; mixing; nonlinear regression; sequential design..