Journal of Statistical Planning and Inference 149 (2014) 1–15

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Three-phase optimal design of sensitivity experiments

C.F. Jeff Wu [a,*], Yubin Tian [b]

[a] Georgia Institute of Technology, United States
[b] Beijing Institute of Technology, China

ARTICLE INFO

Article history:
Received 5 December 2012
Received in revised form
7 May 2013
Accepted 16 October 2013
Available online 29 October 2013

Keywords:
D-optimal designs
Quantile estimation
Robbins–Monro
Sensitivity testing
Stochastic approximation
Up-and-down method

ABSTRACT

In sensitivity testing the test specimens are subjected to a variety of stress levels to generate response or nonresponse. These data are used to estimate the critical stimulus (or threshold) level of the experimental object. Because of its versatile applications, several sensitivity testing procedures have been proposed and used in practice. There remains the outstanding question of finding an efficient procedure, especially when the sample size is small and the interest lies in the extreme percentiles. In the paper we propose a novel three-phase procedure, dubbed 3pod, which can be described as a trilogy of "search-estimate-approximate". A core novel idea is to choose the stress levels to quickly achieve an overlapping data pattern which ensures the estimability of the underlying parameters. Simulation comparisons show that 3pod outperforms existing procedures over a range of scenarios in terms of efficiency and robustness.

1. Introduction

In some experimental investigations each experimental unit has a critical stimulus level that cannot be observed directly. The distribution of these critical levels (or thresholds) over test specimens is of primary interest. In pyrotechnic applications the stimulus (or stress) level, denoted by x, may be the drop height of an explosive, or the pressure on a pile of ammunition, and the response, denoted by y, is binary (e.g., explosion/non-explosion). In industrial applications, x may be the wind speed in a wind-tunnel experiment or the dosage of a compound in a toxicity study. Sensitivity tests are commonly used to estimate some aspects of this distribution. Numerous methods have been proposed for conducting sequential sensitivity testing. But there remains the outstanding question of finding a sequential procedure that works well when the test sample size is small, especially for estimating *extreme* quantiles. For example, there is no clear winner among the competing methods in a comprehensive simulation study reported by Young and Easterling (1994). The median is considered because it is easier to estimate (i.e., more information in the data). And engineers often use it as a benchmark, especially during product development. However, to assess the reliability of a product for field use, the interest and focus is often on the extreme percentiles with $p \geq 0.9$ or even $p \geq 0.99$.

The problem can be formulated as follows. Let $y=1$ or 0 denote the binary outcome, response or nonresponse respectively and $F(x)$ denote the probability Prob $(y=1|x)$ of response at a given stimulus level x. Usually we consider the location-scale model:

$$F(x,\mu,\sigma) = G((x-\mu)/\sigma), \tag{1}$$

* Corresponding author. Tel.: +1 404 352 0885.
E-mail address: jeffwu@isye.gatech.edu (C.F.J. Wu).

0378-3758/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jspi.2013.10.007

where μ and $\sigma>0$ are unknown parameters and G is a known distribution function. Define x_p to be the pth quantile of the distribution, i.e., $F(x_p)=p$. Hence $x_p=\mu+\sigma G^{-1}(p)$. Many methods have been proposed for quantile estimation. A review will be given in Section 2. However, when p is large, e.g., $p=0.99$ or 0.999, and the sample size is small to moderate, accurate estimation of extreme quantiles is still a challenging problem.

The main purpose of this paper is to develop a new *three-phase sequential procedure* that can find the desired stimulus levels more precisely by quickly and efficiently exploiting the information in the test data and expert knowledge. In Section 2, we will give a brief review of some existing methods that are known to be competitive or used in practice. In Section 3, we describe the proposed procedure. Its phase I is to generate some response and nonresponse values, to identify a reasonable experimental range, and to move the stress levels to achieve an overlapping pattern. Phase II is to facilitate the maximum likelihood estimation of the parameters in the assumed model and to spread the stress levels for optimal parameter estimation. Once the data pattern is in good shape, we move to Phase III which uses an efficient approximation scheme to get the stress levels converge to the unknown quantile quickly. This three-phase procedure can be viewed as a trilogy of "search–estimate–approximate". Section 4 gives one example to illustrate the key steps of the three-phase procedure. Simulation comparisons of the procedure with some existing methods are given in Section 5. Concluding remarks and further questions are given in Section 6.

2. Review of some existing methods

The literature contains many studies that compare a large number of methods. An early one with a good review is Wetherill (1963). Some of the papers mentioned in this section also have good reviews. Therefore we will confine our review in this section to those that will be compared with our method in Section 3. First is the up-and-down method (Dixon and Mood, 1948), a.k.a. the Bruceton test. It increases (and respectively decreases) the x value by a step length d if $y=0$ (and resp. $y=1$). Because of its simplicity, it is still widely used, even though it has been known among researchers to be inefficient in most situations. It is only for estimating the median $x_{0.5}$. The Robbins and Monro (1951) stochastic approximation procedure and its adaptive version (Lai and Robbins, 1979) are known to be much more efficient than the up-and-down method. Joseph (2004) recognized that the procedure, originally developed for continuous data, is not well suited for binary data. He then modified the Robbins–Monro (RM) procedure as follows. Let $\theta=x_p$ and assume the prior distribution of θ is $\mathrm{N}(x_1,\tau_1^2)$. Consider the following stochastic approximation scheme:

$$x_{i+1}=x_i-a_i(y_i-b_i),\quad i\geq 1, \tag{2}$$

where a_i and b_i are two sequences of constants. Let $Z_i=x_i-\theta,\ i\geq 1$. He proposed to choose a_i and b_i such that $\mathrm{E}(Z_{i+1}^2)$ is minimized subject to the condition $E(Z_{i+1})=0$. Under the normal approximation of the distribution of Z_i by $\mathrm{N}(0,\tau_i^2)$, he showed that the solution is given by

$$b_i=\Phi\left\{\frac{\Phi^{-1}(p)}{(1+\beta^2\tau_i^2)^{1/2}}\right\},\quad a_i=\frac{1}{b_i(1-b_i)\frac{\beta\tau_i^2}{(1+\beta^2\tau_i^2)^{1/2}}\phi\left\{\frac{\Phi^{-1}(p)}{(1+\beta^2\tau_i^2)^{1/2}}\right\}},$$
$$\tau_{i+1}^2=\tau_i^2-b_i(1-b_i)a_i^2,\quad \beta=\frac{G'(G^{-1}(p))}{\phi(\Phi^{-1}(p))}\cdot\frac{1}{\sigma}, \tag{3}$$

where $\Phi(\cdot)$ is the stand normal CDF and $\phi(\cdot)$ is its density function. For the sake of brevity, we shall refer to this binary version of the RM procedure as the Robbins–Monro–Joseph (RMJ) procedure. If the true distribution is normal ($G=\Phi$), then β in (3) reduces to σ^{-1}. Joseph (2004) recommended the choice

$$\tau_1=\frac{c}{\Phi^{-1}(0.975)}. \tag{4}$$

We choose $c=5$ in the simulation study for RMJ. As is common in stochastic approximation, the last x value is used as the estimate of the unknown value x_p. That is, for N iterations, use x_{N+1} as the estimate.

Another alternative is the MLE recursive method due to Wu (1985). Approximate the true unknown model by using a parametric model $F(x|\gamma)$ like the logit or probit indexed by a few parameters, e.g., $\gamma=(\mu,\sigma)$. After n runs, let $\hat{\gamma}_n$ be the MLE of γ. The next run is chosen at the level x_{n+1} defined as $F(x_{n+1}|\hat{\gamma}_n)=p$. However the method can only be used if the MLE of (μ,σ) in the model (1) exists. According to Silvapulle (1981), a necessary and sufficient condition for the existence of MLE is to have an *overlapping pattern* in the data. That is, the largest x value among the y_i's with $y=0$, denoted by M_0, should be larger than the lowest x value among the y_i's with $y=1$, denoted by m_1, namely

$$M_0>m_1. \tag{5}$$

Wu (1985) recognized that his method needs to start with an initial design that satisfies this condition. A Bayesian extension of Wu's method was proposed by Joseph et al. (2007). We refer to $[m_1,M_0]$ as the *overlapping interval*. If $M_0\leq m_1$, we refer to $[M_0,m_1]$ as the *separation interval* because the region for nonresponse ($y=0$) is separated from the region for response ($y=1$). If a separation pattern is observed, taking the next stress level x within the separation interval will not change the separation pattern (but the interval will get shorter). We refer to this as "trapped in separation". At a certain point, the next stress level should be taken outside the separation interval in order to break the logjam.