Journal of Multivariate Analysis 143 (2016) 424–439

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Higher order density approximations for solutions to estimating equations

Anthony Almudevar

Department of Biostatistics and Computational Biology, University of Rochester, 265 Crittenden Blvd, Rochester, NY, 14642, United States

ARTICLE INFO

Article history:
Received 12 April 2015
Available online 7 October 2015

AMS subject classifications:
62E20
62F12

Keywords:
Small sample asymptotics
Rice formula
Estimating equations
Generalized linear models

ABSTRACT

General formulae for the intensity function of a point process defined by the solution set of a system of smooth random equations are widely available in the literature, offering a precise characterization of a type of random process arising naturally in many fields. Almost all are modifications or generalizations of an original formula derived by Rice (1945) and share the same relatively simple structure. Related methods have been applied to the evaluation of the density of the solution to multidimensional estimating equations arising in statistical inference (Skovgaard, 1990; Jensen and Wood, 1998; Almudevar et al., 2000). This approach has been able to verify or extend a variety of known approximation methods, but has otherwise not been commonly used in the area of small sample asymptotic theory, despite its potential for the development of approximation methods of considerable generality. This article develops a general order $O(1/n)$ density approximation method for solutions to multidimensional estimating equations which are sums of continuous independent, non-identically distributed random vectors. Two issues in particular which arise in the application of the Rice formula are addressed. Validation of this formula is often technically challenging, so a set of general conditions motivated specifically by the application to estimating equations is developed. In addition, the Rice formula includes a conditional expectation which would be difficult to evaluate for non-Gaussian processes. To address this issue, a general order $O(1/n)$ approximation for expectations conditioned on random sums is derived, which may be directly used in the Rice formula under the hypothesis considered here. The method is demonstrated using the negative exponential regression model, a type of non-canonical generalized linear model.

1. Introduction

Systems of random equations appear in a number of areas in probability and statistics. Much early work was motivated by the problem of characterizing level crossings of random processes. A general solution was first derived in [20], interpreted as the temporal distribution of the level crossings of a smooth stochastic process, related to the study of exceedances of reflected waves in transmission lines. A related problem is the characterization of the distribution of the roots of polynomials with random coefficients, solved in [14] for one equation, and generalized to multiple equations in, for example, [23]. See [19,17] for surveys of the early history of this model. The problem is, of course, ubiquitous in statistical methodology, in which estimators are frequently constructed as solutions to systems of equations which are dependent on random quantities X, typically a random sample from sample space $\mathcal{X}$. The analysis which follows is relevant to all these applications.

E-mail address: anthony_almudevar@urmc.rochester.edu.

http://dx.doi.org/10.1016/j.jmva.2015.09.014
0047-259X/

Suppose we have mapping $\Psi : \mathcal{X} \times \Theta \to \mathbb{R}^p$ for a parameter space $\Theta \subset \mathbb{R}^p$. Let $X \in \mathcal{X}$ be some random quantity. Then define the system of equations

$$\Psi(X, \theta) = u. \tag{1}$$

It is usually convenient to set $u = 0$, however, there is sometimes an advantage in allowing u to vary. Eq. (1) may be interpreted as a random system of p equations in p variables $\theta \in \Theta$, with X serving as the random index. The set of solutions then defines a point process Q on Θ. The problem considered here is the characterization of the distribution of Q. The *Rice formula* for the intensity function of Q, in its multivariate extension, is given by

$$\lambda^A(\theta) = f_{\Psi(X,\theta)}(u) E\left[\text{abs}(\det(\Psi'(X, \theta)))I_E \mid \Psi(X, \theta) = u\right]. \tag{2}$$

Here, an indicator function I_E is included to restrict Q to solutions of interest. For example, if $\Psi(X, \theta)$ is a stochastic process in time θ, then Q is the set of level crossings of u. If I_E indicates positive signs of $\det(\Psi'(X, \theta))$ then λ_A is the intensity function for up-crossings only. Alternatively, $\Psi(X, \theta)$ may be the derivative vector for an objective function $\psi : \mathcal{X} \times \Theta \to \mathbb{R}$, so that the sign of $\det(\Psi'(X, \theta))$ distinguishes between maxima and minima. It may also be of interest to incorporate this distinction into I_E.

At first glance, application of this model to statistics would appear to differ in one important respect from the other applications cited above in the sense that (1) is intended to have exactly one solution, say $\theta(X)$. Interestingly, this distinction need not play an important role in the development of a formula such as λ^A. If a unique solution exists with probability 1, equivalently $P(|Q| = 1) = 1$, then $\lambda^A(\theta)$ is simply the density of $\theta(X)$. In practice, this situation often does not strictly hold, and work on this problem within this context has proceeded by accepting the variable cardinality of Q, and developing regularity conditions under which suitable forms of the Rice formula hold [21,13,1].

This type of method is able to extend, or more easily derive, a number of known approximations. In [21] Barndorff-Nielsen's formula for the density of a maximum likelihood estimator conditioned on an ancillary statistic [4,5] is directly derived using λ^A (see also [22]). Saddlepoint density approximations for elliptical contrast functions are rigorously obtained in [13], while in [1] saddlepoint density approximations for regression and scale estimates obtained from Huber's M-estimation method [12] are derived, extending results reported in [9]. Regularity conditions used in [1] were weakened in [10] to permit application to higher order bootstrap approximations. It seems possible, therefore, that the Rice formula can form the basis for density evaluation methods of considerable generality.

Despite the ubiquitous appearance of formulae related to (2), verifying its validity may pose significant technical challenges. It holds under general conditions for Gaussian processes [3] and validation has been extended to stable processes (see [17] for discussion). Another approach was developed by Brillinger [7], in which the Rice formula was shown to hold under general conditions for almost all u (Leadbetter and Spaniolo [17] surveys more recent results of this type). This measure theoretic qualification may be quite restrictive, since in many applications, especially in statistical inference, we wish to evaluate (2) at a fixed value of u (the results of [21,13,1] hold in this stronger sense). To address this problem, the argument in [7] was extended to the fixed u case by noting that (2) must at least hold at some u' arbitrarily close to u, then developing continuity conditions exploiting this fact.

We will address in this article two issues which may hamper more widespread use of the Rice formula. The first is to develop general validation conditions in the context of (1). Similar to [7], the essential requirement is that the process be continuous in some sense. However, we are able to eliminate conditions imposed by Brillinger on the counting process for Q, including the existence of higher order moments and continuity in u, at the price of introducing additional assumptions on Ψ which are purely functional. In particular, we will assume that Ψ is a smoothly differentiable function of both θ and $X \in \mathbb{R}^m$. The remaining conditions on the probability measure of the process are quite minimal. Define $\Psi^*(X, \theta) = \Psi'(X, \theta)^{-1}\Psi(X, \theta)$ (a more precise definition will follow). We will require only the existence and boundedness of the density of Ψ^* near u, as well as a tightness condition on the marginal density of Ψ^* over $\mathcal{X}$. This type of regularity condition will be much more natural for problems in statistical inference.

Our second task will be to consider the problem of developing a high order approximation method for the conditional expectation in (2) (the difficulty of evaluating this formula is noted in [21]). A closed form solution is well known in the multivariate Gaussian case, but can otherwise be quite difficult to calculate. We will develop an order $O(1/n)$ approximation when Ψ is a sum of n independent but not identically distributed terms. The intuition behind this is that when Ψ is approximately normal, the conditional expectation approximately behaves as in the Gaussian case. The method is demonstrated on a non-canonical exponential family regression model discussed in [15].

Software is available at www.urmc.rochester.edu/biostat/people/faculty/almudevar.cfm.

2. Validation of the rice formula for random systems of equations

In [1] an alternative form of the Rice formula was developed, given in its simplest form by

$$\lambda^B(\theta) = f_{\Psi^*(X,\theta)}(0), \tag{3}$$

(we may consider, without loss of generality, u in (1) to be fixed at 0, otherwise incorporate u into Ψ). With suitable regularity conditions (discussed below) the intensity function for Q will be λ^B. Using a standard transformation argument λ^A can