Computational Statistics and Data Analysis 71 (2014) 849–858

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Computational issues of generalized fiducial inference

Jan Hannig [a], Randy C.S. Lai [b], Thomas C.M. Lee [b,*]

[a] Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260, USA

[b] Department of Statistics, University of California at Davis, 4118 Mathematical Sciences Building, One Shields Avenue, Davis, CA 95616, USA

ARTICLE INFO

Article history:
Received 10 April 2012
Received in revised form 1 March 2013
Accepted 4 March 2013
Available online 14 March 2013

Keywords:
Confidence distributions
Dempster–Shafer calculus
Generalized confidence intervals
Generalized fiducial inference

ABSTRACT

Generalized fiducial inference is closely related to the Dempster–Shafer theory of belief functions. It is a general methodology for constructing a distribution on a (possibly vector-valued) model parameter without the use of any prior distribution. The resulting distribution is called the generalized fiducial distribution, which can be applied to form estimates and confidence intervals for the model parameter. Previous studies have shown that such estimates and confidence intervals possess excellent frequentist properties. Therefore it is useful and advantageous to be able to calculate the generalized fiducial distribution, or at least to be able to simulate a random sample of the model parameter from it. For a small class of problems this generalized fiducial distribution can be analytically derived, while for some other problems its exact form is unknown or hard to obtain. A new computational method for conducting generalized fiducial inference without knowing the exact closed form of the generalized fiducial distribution is proposed. It is shown that this computational method enjoys desirable theoretical and empirical properties. Consequently, with this proposed method the applicability of generalized fiducial inference is enhanced.

1. Introduction

Fisher (1930) introduced the idea of fiducial probability and fiducial inference as an attempt to overcome what he saw as a serious deficiency of the Bayesian approach to inference: the use of a prior distribution on model parameters even when no prior information is available. Fiducial inference created some controversy once Fisher's contemporaries realized that, unlike earlier simple applications involving a single parameter, fiducial inference often led to procedures that were not exact in the frequentist sense and did not possess other properties claimed by Fisher (Lindley, 1958; Zabell, 1992). An interested reader can consult Section 2 of Hannig (2009) for a discussion of the history of fiducial inference and a more complete list of references.

Tsui and Weerahandi (1989) and Weerahandi (1993) proposed a new approach for constructing hypothesis tests using the concept of generalized P-values and generalized confidence intervals. Hannig et al. (2006) established a direct connection between fiducial intervals and generalized confidence intervals and proved the asymptotic frequentist correctness of such intervals. These ideas were unified for parametric problems in Hannig (2009) without requiring any group structure related to the model. This unification is termed *generalized fiducial inference* and has been found to have excellent theoretical and empirical properties for a number of practical applications (E et al., 2008; Hannig and Lee, 2009; Wandler and Hannig, 2011).

* Corresponding author. Tel.: +1 530 554 1357; fax: +1 530 752 7099.
E-mail addresses: jan.hannig@unc.edu (J. Hannig), rcslai@ucdavis.edu (R.C.S. Lai), tcmlee@ucdavis.edu (T.C.M. Lee).

0167-9473/$ – see front matter
http://dx.doi.org/10.1016/j.csda.2013.03.003

The Dempster–Shafer theory of belief functions (Dempster, 2008; Shafer, 2011) is closely related to the generalized fiducial inference. In both cases a relationship between the data, $\mathbf{X}$, and the parameters, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is expressed in a functional form; i.e.,

$$\mathbf{X} = \mathbf{G}(\boldsymbol{\theta}, \mathbf{U}), \tag{1.1}$$

where $\mathbf{U}$ is the random component of the structural equation, a random variable or vector whose distribution is completely known and independent of any parameters. The Eq. (1.1) is typically called data-generating equation, structural equation or a-equation.

After observing a fixed realized value of $\mathbf{X}$, say $\mathbf{x}_0$, one can invert the Eq. (1.1) to define a belief function on the parameter space $\boldsymbol{\Theta}$. To explain the idea behind the formal definition of this belief function, suppose first that the structural relation (1.1) can be inverted, solved for θ, and the inverse $\mathbf{Q}(\mathbf{x}_0, \mathbf{u})$ always exists. That is, for any observed $\mathbf{x}_0$ and for all $\mathbf{u}$, there is always a unique $\boldsymbol{\theta}$ solving $\mathbf{x}_0 = \mathbf{G}(\boldsymbol{\theta}, \mathbf{u})$. (The more realistic case where the inverse $\mathbf{Q}$ does not exist will be discussed in the next section.) Since the distribution of $\mathbf{U}$ is completely known, one can always generate a random sample $\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_M$ from it. This random sample of $\mathbf{U}$ is transformed into a random sample of $\boldsymbol{\theta}$ via the inverse $\mathbf{Q}$: $\{\tilde{\boldsymbol{\theta}}_1 = \mathbf{Q}(\mathbf{x}_0, \tilde{\mathbf{u}}_1), \ldots, \tilde{\boldsymbol{\theta}}_M = \mathbf{Q}(\mathbf{x}_0, \tilde{\mathbf{u}}_M)\}$, and the resulting random sample $\tilde{\boldsymbol{\theta}}_1, \ldots, \tilde{\boldsymbol{\theta}}_M$ can be used to obtain estimates and approximate confidence intervals for $\boldsymbol{\theta}$. From this one can see that a probability density function $r(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$ is implicitly defined. The corresponding distribution is termed the *generalized fiducial distribution*, which, in this case, is equivalent to the belief function.

This recipe defines a joint distribution on the parameter space. When making inference about individual parameters, we then use the marginal of this joint distribution. It is well known that marginalization together with different structural equations can lead to non-uniqueness; i.e., several different marginal fiducial distributions, (e.g., Dawid et al., 1973). We do not see this as a problem because our goal is not to define the unique fiducial distribution. Instead, we are aiming to define a distribution on the parameter space with good inferential properties. Approximate confidence intervals based on the marginal fiducial distribution lead often to asymptotically correct coverage (Hannig, 2009, 2013). More importantly, such confidence intervals have been shown to have very good small sample properties in a number of applied problems. For computational reasons, we also recommend using structural equations that have simple form so that the calculations described in Section 3 can be done in a closed form; see for example Section 4.3. More discussion on the choice of structural equation can be found in Section 5 of Hannig (2013).

We remark that there is a new exciting inferential approach developed by Zhang and Liu (2011) that in some situations does not require a computation of the whole distribution, making it attractive for various problems. However, this method seems to have a requirement that the structural equation is invertible almost surely (Zhang and Liu, 2011, Theorem 3.1). Such an assumption is not reasonable when the dimension of the minimal sufficient statistics is larger than the number of parameters; e.g., the Cauchy regression problem to be discussed below.

The density function $r(\boldsymbol{\theta})$ thus plays an important role in the generalized fiducial approach to data analysis: it can be applied to derive estimates and construct confidence intervals for the parameter $\boldsymbol{\theta}$, in a similar manner as the posterior density function in the Bayesian paradigm. However, for many statistical problems, the exact form of $r(\boldsymbol{\theta})$ cannot be easily calculated. A major contribution of this article is to propose practical methods for computing integrals of $r(\boldsymbol{\theta})$ when a closed form expression for $r(\boldsymbol{\theta})$ is not readily available. Consequently, this article greatly enhances the applicability of generalized fiducial inference for statistical problems.

The rest of this article is organized as follows. Firstly some background material is presented in Section 2. Then some computational ideas are developed in Section 3 for simulating a random sample from $r(\boldsymbol{\theta})$ when the exact form of $r(\boldsymbol{\theta})$ is unknown. Section 4 provides some simulation studies to demonstrate the advantages of the proposed computational approach. Lastly, concluding remarks are offered in Section 5.

2. Background

This section provides some essential background material. First recall that in the discussion above the inverse $\mathbf{Q}$ is assumed to exist. In practice this is likely to be wrong and the inverse $\mathbf{Q}$ does not exist. This can happen for two opposing reasons: for some value of $\mathbf{x}_0$ and $\mathbf{u}$, either there is more than one $\boldsymbol{\theta}$, or there is no $\boldsymbol{\theta}$ satisfying $\mathbf{x}_0 = \mathbf{G}(\boldsymbol{\theta}, \mathbf{u})$. The first situation can be dealt with by using the mechanics of Dempster–Shafer calculus (Dempster, 2008); see also Section 4 of Hannig (2009). The main idea is that the belief function for any set has three components summarizing (i) the strength evidence for the set, (ii) the strength of the evidence against the set, and (iii) the strength of the evidence that is inconclusive. Hannig (2013) shows that in many statistical problems of practical interest the portion of the evidence that is inconclusive is asymptotically negligible.

For the second situation where no $\boldsymbol{\theta}$ satisfies $\mathbf{x}_0 = \mathbf{G}(\boldsymbol{\theta}, \mathbf{u})$, Hannig (2009) suggests removing the values of $\mathbf{u}$ for which there is no solution from the sample space and then re-normalizing the probabilities; i.e., using the distribution of $\mathbf{U}$ conditional on the event that "there is at least one $\boldsymbol{\theta}$ solving the equation $\mathbf{x}_0 = \mathbf{G}(\boldsymbol{\theta}, \mathbf{U})$". The rationale for this choice is that we know that the observed data $\mathbf{x}_0$ were generated using some fixed unknown $\boldsymbol{\theta}_0$ and $\mathbf{u}_0$; i.e., $\mathbf{x}_0 = \mathbf{G}(\boldsymbol{\theta}_0, \mathbf{u}_0)$. The information that the solution of the equation $\mathbf{x}_0 = \mathbf{G}(\boldsymbol{\theta}, \mathbf{U})$ exists for the true $\mathbf{U} = \mathbf{u}_0$ is available to us in addition to knowing the distribution of $\mathbf{U}$. The values of $\mathbf{u}$ for which $\mathbf{x}_0 = \mathbf{G}(\cdot, \mathbf{u})$ does not have a solution could not be the true $\mathbf{u}_0$ hence only the values of $\mathbf{u}$ for which there is a solution should be considered in the definition of the generalized fiducial