Journal of Statistical Planning and Inference 191 (2017) 81–87

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

The formal relationship between analytic and bootstrap approaches to parametric inference

T.J. DiCiccio [a], T.A. Kuffner [b,*], G.A. Young [c]

[a] Cornell University, Ithaca, NY 14853, USA
[b] Washington University in St. Louis, St. Louis, MO 63130, USA
[c] Imperial College London, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 5 August 2016
Received in revised form 28 March 2017
Accepted 15 May 2017
Available online 1 June 2017

Keywords:
p-value
Bootstrap
Likelihood
Signed root likelihood ratio statistic
Normal pivot

ABSTRACT

Two routes most commonly proposed for accurate inference on a scalar interest parameter in the presence of a (possibly high-dimensional) nuisance parameter are parametric simulation ('bootstrap') methods, and analytic procedures based on normal approximation to adjusted forms of the signed root likelihood ratio statistic. Under some null hypothesis of interest, both methods yield *p*-values which are uniformly distributed to error of third-order in the available sample size. But, given a specific dataset, what is the formal relationship between *p*-values calculated by the two approaches? We show that the two methodologies give the same inference to second order in general: the analytic *p*-value calculated from a dataset will agree with the bootstrap *p*-value constructed from that same dataset to $O(n^{-1})$, where n is the sample size. In practice, the agreement is often startling.

1. Introduction

We are concerned with inference, primarily using the signed root likelihood ratio statistic R, on a scalar interest parameter ψ, in the presence of a (possibly high-dimensional) nuisance parameter ϕ, based on a random sample of size n from an assumed parametric distribution depending on $\theta = (\psi, \phi)$. Two routes most commonly proposed for accurate inference on ψ are parametric simulation ('bootstrap') methods, (see DiCiccio et al., 2001; Lee and Young, 2005) and analytic procedures based on normal approximation to adjusted forms of R, obtained via small-sample asymptotics. Prominent among analytic procedures is use of a normal approximation to the R^* statistic introduced by Barndorff-Nielsen (1986, 1991). Our purpose here is to elucidate the formal relationship between the bootstrap approach to inference, specifically as applied to the signed root statistic R, and the analytic approach based on R^*. In this paper, we examine the specific relationships between the bootstrap and analytic methods for estimation of *p*-values for inference on ψ: particular focus in our numerical illustrations will be with estimation of *p*-values under the null hypothesis. We use results from DiCiccio et al. (2015a, b) to show that from a theoretical perspective, quite generally, analytic and bootstrap *p*-values are equivalent to $O(n^{-1})$: the two *p*-values constructed from the same dataset agree to that order. Several examples showing close empirical agreement of *p*-values, even for very small sample sizes n, are provided.

2. Problem setting

Suppose $Y = (Y_1, \ldots, Y_n)$ is a continuous random vector whose distribution depends on a parameter $\theta = (\theta^1, \ldots, \theta^d) = (\psi, \phi)$, where ψ is a scalar parameter of interest and ϕ is a vector of nuisance parameters, of dimension $d - 1$. Further

* Corresponding author.
E-mail addresses: tjd9@cornell.edu (T.J. DiCiccio), kuffner@wustl.edu (T.A. Kuffner), alastair.young@imperial.ac.uk (G.A. Young).

http://dx.doi.org/10.1016/j.jspi.2017.05.007
0378-3758/

suppose that it is required to test the null hypothesis $H_0 : \psi = \psi_0$ against a one-sided alternative. We wish to compare, for a given dataset, the p-values derived from analytic approximation to the distribution of R^* with the p-values derived from the bootstrap distribution of R.

For testing the null hypothesis against one-sided alternatives, we may use the signed root of the usual likelihood ratio statistic

$$R(\psi) = \mathrm{sgn}(\hat{\psi} - \psi)[2\{L(\hat{\theta}) - L(\hat{\theta}_\psi)\}]^{1/2} = \mathrm{sgn}(\hat{\psi} - \psi)[2\{M(\hat{\psi}) - M(\psi)\}]^{1/2},$$

where $L(\theta)$ is the log-likelihood function, $\hat{\theta} = (\hat{\psi}, \hat{\phi})$ is the global maximum likelihood estimator, $\hat{\theta}_\psi = (\psi, \hat{\phi}_\psi)$ is the constrained maximum likelihood estimator given ψ, and $M(\psi) = L(\hat{\theta}_\psi)$ is the log profile likelihood function for ψ. Under the null hypothesis, the repeated sampling distribution of R is standard normal to error of order $O_p(n^{-1/2})$. The analytic route to achieve higher-order accuracy in such an inferential setting is based on a standard normal approximation to an adjusted version of $R(\psi)$, denoted by $R^*(\psi)$.

The development of $R^*(\psi)$ is as follows. Suppose that the log-likelihood function is written as $L(\theta; \hat{\theta}, a)$, with $(\hat{\theta}, a)$ minimal sufficient, where a is ancillary, having a distribution which, at least approximately, does not depend on θ. Such a decomposition holds, trivially, in full exponential families, where $\hat{\theta}$ is minimal sufficient and no ancillary a is required, and in transformation models, where the maximal invariant serves as the ancillary a. As noted by Severini (2000, §6.5), beyond the exponential family and transformation model contexts, it may be difficult to establish that such a decomposition holds, but general approximations, in particular constructions of approximate ancillaries, are possible which still allow validity of the properties discussed here for analytic methods of inference. A drawback of such constructions is that explicit expression of the log-likelihood in terms of $(\hat{\theta}, a)$ may then be intractable. This does not affect the calculation of a bootstrap p-value, but would require approximation to the R^* statistic, which we now describe.

The R^* statistic is defined (Barndorff-Nielsen, 1986, 1991) as

$$R^*(\psi) = R(\psi) + R(\psi)^{-1} \log(U(\psi)/R(\psi)),$$

with

$$U(\psi) = \begin{vmatrix} L_{;\hat{\theta}}(\hat{\theta}) - L_{;\hat{\theta}}(\hat{\theta}_\psi) \\ L_{\phi;\hat{\theta}}(\hat{\theta}_\psi) \end{vmatrix} / \{|j_{\phi\phi}(\hat{\theta}_\psi)|^{1/2} |j(\hat{\theta})|^{1/2}\}.$$

Here $j(\theta) = (-L_{rs}(\theta))$ denotes the observed information matrix, with $L_{rs}(\theta) = \partial^2 L(\theta)/\partial\theta^r \partial\theta^s$, where the indices r, s range from $1, \ldots, d$, and $j_{\phi\phi}$ denotes the $(d-1) \times (d-1)$ sub-matrix corresponding to components of the nuisance parameter ϕ. Also,

$$L_{;\hat{\theta}}(\theta) \equiv L_{;\hat{\theta}}(\theta; \hat{\theta}, a) = \frac{\partial}{\partial\hat{\theta}} L(\theta; \hat{\theta}, a), \qquad L_{\phi;\hat{\theta}}(\theta) \equiv L_{\phi;\hat{\theta}}(\theta; \hat{\theta}, a) = \frac{\partial^2}{\partial\phi\partial\hat{\theta}} L(\theta; \hat{\theta}, a).$$

The conditional distribution of the test statistic $R^*(\psi) = R + R^{-1} \log(U/R)$ given a, and hence the unconditional distribution under repeated sampling, is standard normal to error of order $O_p(n^{-3/2})$. An alternative to the standard normal distribution for approximating tail probabilities of $R^*(\psi)$ is the generalized Lugannani–Rice formula (Barndorff-Nielsen, 1991); to error of order $O(n^{-3/2})$,

$$\mathrm{pr}(R^* \le r^* \big| a; \theta) = \Phi(r^*) = \Phi(r) + \varphi(r)(1/r - 1/u), \tag{1}$$

where $r^* = r + r^{-1} \log(u/r)$. The simulation route to inference in this setting is based on the parametric bootstrap approximation to the marginal distribution of $R(\psi)$. This is defined as the sampling distribution of $R(\psi)$ under the model specified by parameter value $\hat{\theta}_\psi$, the constrained maximum likelihood estimator for the observed data sample: see DiCiccio et al. (2001) and Lee and Young (2005). This parametric bootstrap yields p-values which are, under repeated sampling and supposing ψ is the true value of the interest parameter, uniformly distributed to error of order $O(n^{-3/2})$.

We consider first a motivating example.

Example 1 (*Extreme Value Location-Scale*)**.** Let $\{X_1, \ldots, X_n\}$ be a random sample from the Weibull density

$$f(x; \beta, \gamma) = \gamma\beta(\gamma x)^{\beta-1} \exp\{-(\gamma x)^\beta\}, \quad x > 0,$$

with interest parameter β. Defining $Y_i = \log X_i$, the Y_i are random samples from an extreme value distribution $EV(\mu, \psi)$, a location-scale family, with scale and location parameters $\psi = \beta^{-1}$, $\mu = -\log\gamma$. Interest is in inference on the scale parameter of the extreme value distribution. This distribution constitutes an ancillary statistic model: inference for ψ conditions on the observed data value of the ancillary $a = (a_1, \ldots, a_n)$, with $a_i = (y_i - \hat{\mu})/\hat{\psi}$. Exact conditional inference is analytically straightforward, but requires numerical integration for its calculation: see, for instance, Pace and Salvan (1997, §7.6). Here, it is easily verified that the conditional distribution of $R(\psi)$ given a does not depend on the nuisance parameter μ, so the exact conditional inference is equivalent to a 'conditional bootstrap', which would be based on simulating the conditional distribution of $R(\psi)$ given a, modulo the error introduced by the finite simulation required in practice. It is of interest to see how well this exact conditional inference is approximated by a marginal bootstrap, which ignores the conditioning and is based on simulation of the marginal distribution of $R(\psi)$.