Journal of Statistical Planning and Inference 145 (2014) 204–221

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

A prior-free framework of coherent inference and its derivation of simple shrinkage estimators

David R. Bickel [a,b,*], Marta Padilla [a]

[a] Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology, and Immunology, University of Ottawa; 451 Smyth Road; Ottawa, Ontario, Canada K1H 8M5

[b] Department of Mathematics and Statistics, University of Ottawa; 451 Smyth Road; Ottawa, Ontario, Canada K1H 8M5

ARTICLE INFO

Article history:
Received 28 June 2012
Received in revised form
22 July 2013
Accepted 6 August 2013
Available online 22 August 2013

Keywords:
Confidence distribution
Confidence curve
Confidence measure
Confidence posterior distribution
Fiducial inference
Fiducial shrinkage
Observed confidence level

ABSTRACT

The reasoning behind uses of confidence intervals and p-values in scientific practice may be made coherent by modeling the inferring statistician or scientist as an idealized intelligent agent. With other things equal, such an agent regards a hypothesis coinciding with a confidence interval of a higher confidence level as more certain than a hypothesis coinciding with a confidence interval of a lower confidence level. The agent uses different methods of confidence intervals conditional on what information is available. The coherence requirement means that all levels of certainty of hypotheses about the parameter agree with the same distribution of certainty over parameter space. The result is a unique and coherent fiducial distribution that encodes the post-data certainty levels of the agent.

While many coherent fiducial distributions coincide with confidence distributions or Bayesian posterior distributions, there is a general class of coherent fiducial distributions that equates the two-sided p-value with the probability that the null hypothesis is true. The use of that class leads to point estimators and interval estimators that can be derived neither from the dominant frequentist theory nor from Bayesian theories that rule out data-dependent priors. These simple estimators shrink toward the parameter value of the null hypothesis without relying on asymptotics or on prior distributions.

1. Introduction

In the years following the oracle that some form of fiducial inference may play a pivotal role in the 21st-century statistics (Efron, 1998), there has been an ongoing resurgence of interest in fiducial distributions that generate confidence intervals (e.g., Schweder and Hjort, 2002; Singh et al., 2005, 2007; Polansky, 2007; Tian et al., 2011; Bityukov et al., 2011; Kim and Lindsay, 2011; Bickel, 2011, 2012b) and in other distributions of the fiducial type (e.g., Hannig et al., 2006; Hannig, 2009; Xiong and Mu, 2009; Gibson et al., 2011; Wang et al., 2012; Zhao et al., 2012; Balch, 2012). Fiducial inference initially promised an objective alternative to Bayesianism as a form of inductive reasoning (Fisher, 1973) but has historically suffered from problems of understanding the meaning of fiducial probability and from the ability to derive conflicting fiducial probabilities from the same family of sampling distributions (see Wilkinson, 1977). This paper addresses both difficulties by interpreting fiducial probability in terms of the theories of coherent decision making that also undergird Bayesian inference.

* Corresponding author at: Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology, and Immunology, University of Ottawa; 451 Smyth Road; Ottawa, Ontario, Canada K1H 8M5, Tel.: +1 613 562 5800.

E-mail address: dbickel@uottawa.ca (D.R. Bickel).

0378-3758/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jspi.2013.08.011

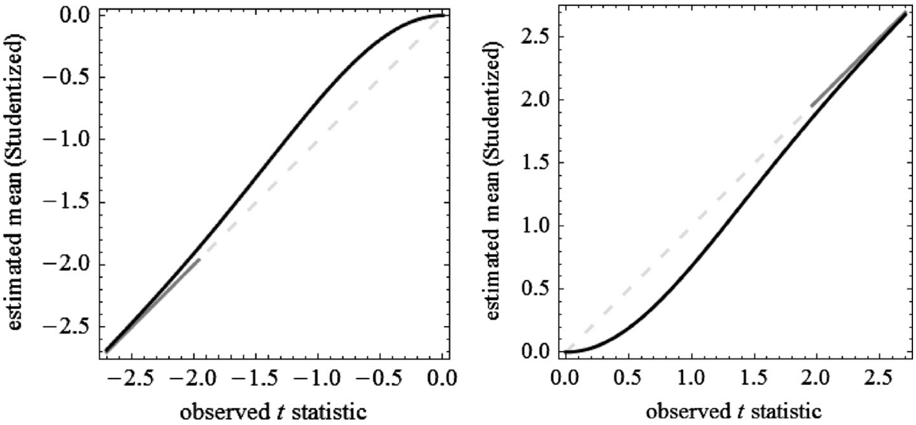

Fig. 1. Estimates of the normal mean relative to its standard error as a function of the observed number of standard errors from 0, the null hypothesis value. The black curve is the posterior mean with respect to the fiducial distribution, and the gray line is the maximum-likelihood estimate (MLE), plotted as a solid line wherever the null hypothesis is rejected at the 5% significance level and as a dashed line elsewhere. See Example 9 for details.

The main premiss is that many of the usual applications of confidence intervals in science lead to reasonable inferences that can be improved by enforcing self-consistency in the technical sense of probabilistic coherence, which does not in itself require Bayesian posterior distributions (Hacking, 1967; Goldstein, 1997; Bickel, 2012a).

Section 2 provides preliminary concepts and propositions, demonstrating that interpreting confidence levels as certainty levels or hypothetical levels of belief leads either to non-coherent estimates and hypothesis testing or to inference on the basis of a confidence distribution of the parameter as if it were a Bayesian posterior distribution. Iterating that reasoning along the lines of Fisher's fiducial argument for multiple parameters leads to merging confidence distributions into a parameter distribution that is coherent; thus, it is a probability measure.

This is fiducial inference in the sense that it is a modern development of fiducial reasoning but without the often impractical requirements involving aspects of conditional inference and without violating the rules of ordinary probability theory (that of the Kolmogorov axioms). The framework proposed in Section 2 also differs from Fisher's in its incorporation of nested confidence sets of vector parameters. Thus, the proposed framework for inference is presented as a realization of the core ideas behind the original fiducial argument, Neyman–Pearson confidence intervals, and theories of coherent decision-making that prescribe minimizing expected loss with respect to a posterior distribution (e.g., von Neumann and Morgenstern, 1953; Savage, 1954). (Following the usage in Dempster, 2008; Eaton and Sudderth, 2010; Bickel, 2012a, the term "posterior" herein means data-dependent and thus includes but is not limited to a Bayesian posterior relative to some prior.)

Section 3 demonstrates that the resulting framework of fiducial inference can lead to shrinkage in point and interval estimates toward ϕ_0, a null hypothesis value, in a way that is not possible in the standard frequentist and Bayesian approaches. For example, Fig. 1 displays the shrunken parameter estimate as an alternative to the usual frequentist estimate computed after testing the null hypothesis. Given the two-sided p-value PV, the MLE $\widehat{\phi}$ is simply shrunk to $(1-\mathrm{PV})\widehat{\phi}$. That value would only be available from Bayes's theorem if the prior depended on the sample size such that the posterior probability of the null hypothesis were equal to PV.

The shrinkage estimates improve performance only to the extent that ϕ_0 is close to the true value of the interest parameter, as is illustrated by simulation in Section 4. Thus, ϕ_0 is best specified using previous knowledge of the application domain, as in Willink (2008) and work on frequentist shrinkage cited therein. Fortunately, the required knowledge is slight enough to be readily available in nearly all applications in which p-values are reported since a ϕ_0 must be specified to conduct a statistical test of any simple null hypothesis that $\phi=\phi_0$. Typically, $\phi_0=0$ or another value that corresponds to the absence of something: no effect, no association, no difference, etc. For example, Montazeri et al. (2010) applied methods from Willink (2008) using the value of ϕ_0 that encodes the absence of differential gene expression. By contrast, Bayesian methods that allow nonzero posterior probability of a null hypothesis not only require the specification of ϕ_0 but also a joint prior distribution of all unknown parameters, including the assignment of a prior probability that $\phi=\phi_0$. Thus, the previous information needed to specify ϕ_0 for the point and interval estimators in Section 3 is far less than that needed to apply comparable Bayesian estimation.

Remarks elaborating on technical points appear in Section 5, a brief discussion in Section 6, and deferred proofs in Appendix A.

2. Coherent fiducial distributions

The concept of a coherent fiducial distribution will be introduced in order to ground coherent decision making in a procedure of confidence intervals or more general confidence sets. Specifically, the coherence condition will be supplemented with a confidence-based condition in order to prescribe point estimates, interval estimates, hypothesis