ARTICLE IN PRESS

Journal of Statistical Planning and Inference ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Combining independent Bayesian posteriors into a confidence distribution, with application to estimating climate sensitivity

Nicholas Lewis

Bath, United Kingdom

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Combining evidence
Confidence distribution
Objective Bayesian
Likelihood ratio
Bayesian updating
Jeffreys prior

ABSTRACT

Combining estimates for a fixed but unknown parameter to obtain a better estimate is an important problem, but even for independent estimates not straightforward where they involve different experimental characteristics. The problem considered here is the case where two such estimates can each be well represented by a probability density function (PDF) for the ratio of two normally-distributed variables. Two different statistical methods – objective Bayesian and frequentist likelihood-ratio – are employed and compared. Each probabilistic estimate of the parameter value is represented by a fitted three-parameter Bayesian posterior PDF providing a close approximation to the ratio of two normals, that can legitimately be factored into a likelihood function and a noninformative prior distribution. The likelihood functions relating to the parameterised fits to the probabilistic estimates are multiplicatively combined and a prior is derived that is noninformative for inference from the combined evidence. An objective posterior PDF that incorporates the evidence from both sources is produced using a single-step approach, which avoids the order-dependency that would arise if Bayesian updating were used. The frequentist signed root likelihood-ratio method is also applied. The probability matching of credible intervals from the posterior distribution and of approximate confidence intervals from the likelihood-ratio method is tested, showing that both methods provide almost exact confidence distributions. The approach developed is applied in the important case of the Earth's equilibrium climate sensitivity, by combining an estimate from instrumental records with an estimate representing largely independent paleoclimate proxy evidence, resulting in a median estimate of 2.0 °C and a 5%–95% confidence/credibility interval of (1.1, 4.5) °C.

1. Introduction

Often, existing estimates of a fixed but unknown parameter are poorly constrained and combining evidence from them offers the most obvious route to obtaining a less imprecise estimate. But, even assuming uncertainties in two estimates are independent, in most cases statistical theory does not provide a unique, optimal method of combining them. Multiplicatively combining the likelihood functions underlying the two estimates, if they are available, provides one obvious starting point, as that is a standard method for combining parametric information from independent sources, applicable in both frequentist and Bayesian paradigms. In the frequentist paradigm, a simple and straightforward approach, yielding approximate confidence distributions, is to apply standard likelihood-ratio methods (Pawitan, 2001, p. 36–37) to the combined likelihood. If there are unwanted (nuisance) parameters, the likelihoods can be reduced, for instance by forming a profile-likelihood, either before or after combining them. However, the standard Bayesian method for using likelihood

E-mail address: nhlewis@btinternet.com.

https://doi.org/10.1016/j.jspi.2017.09.013
0378-3758/

Please cite this article in press as: Lewis N., Combining independent Bayesian posteriors into a confidence distribution, with application to estimating climate sensitivity. J. Statist. Plann. Inference (2017), https://doi.org/10.1016/j.jspi.2017.09.013.

information to combine evidence from independent sources (Bayesian updating) differs from frequentist approaches, and leaves unresolved the problem of what prior distribution to use.

Efron (1998) wrote of the "250-year search for a dependable objective Bayesian theory" and suggested that the development of confidence distributions and approximations thereto might hold a key to it, points echoed by Schweder and Hjort (2002) and Singh et al. (2005). Schweder and Hjort (2002) also wrote that "the confidence distribution may serve as the frequentist analogue of the Bayesian's posterior density, and together with the reduced likelihood a frequentist apparatus for information updating is available as a competitor to the Bayesian methodology". Fraser et al. (2010) stressed that, in absence of information about how the parameter value was generated, to avoid Bayesian inference giving misleading results it was necessary that a prior that provided correct calibration of posterior probabilities to frequencies, at least approximately, be used. Such an approach, which reflects a so-called "objective Bayesian" way of thinking, suggests a possible convergence in many cases between the resulting Bayesian posterior distributions and frequentist confidence distributions.

Here I compare simple objective Bayesian and frequentist likelihood-ratio methods for combining independent probabilistic estimates of an uncertain but fixed parameter, that can each be well represented by a (differing) ratio-normal distribution (that of the ratio of two normally-distributed random quantities, independent if not stated otherwise). In doing so, I take advantage of a transformed-normal approximation to ratio-normal distributions (Raftery and Schweder, 1993). I show that the two methods provide almost identical inference, approximating closely to confidence distributions. I also show that the standard Bayesian methodology for information updating produces different results from those using the proposed objective Bayesian approach, with much worse probability matching and results that, inconsistently, depend on the order in which information is analysed. The methods developed here are exemplified by application to the important case of estimating the Earth's equilibrium climate sensitivity (ECS).

The rest of this paper is organised as follows. Section 2 discusses Bayesian and likelihood-ratio parameter inference and the methods used in this study. Section 3 discusses the physical relevance of the ratio-normal distribution, selects a suitable parameterised approximate distribution to use in this case and tests inference using it. Section 4 explains the methodology for combining the information embodied in two independent such parameterised estimates. Section 5 applies the methods developed to combine evidence regarding ECS. Section 6 summarises, and discusses issues raised.

2. Bayesian vs. likelihood-ratio based parameter inference

2.1. Bayesian inference

Bayes theorem (Bayes, 1763), for a univariate continuous parameter θ on which observed data $\boldsymbol{y}$ depend, states that the (posterior) probability density function (PDF), $p_\theta(\theta|\boldsymbol{y})$, for θ is proportional to the probability density of the data $p_{\boldsymbol{y}}(\boldsymbol{y}|\theta)$ (the "likelihood" when considered as a function of θ, with $\mathbf{y}$ fixed) multiplied by the density of a prior distribution (prior) for θ, $p_\theta(\theta)$:

$$p_\theta(\theta|\boldsymbol{y}) \propto p_{\boldsymbol{y}}(\boldsymbol{y}|\theta)p_\theta(\theta) \tag{1}$$

(the subscripts indicating which variable density is for). The constant of proportionality is such that the posterior PDF integrates to unit probability. Under subjective Bayesian interpretations, the prior (and hence the posterior) represents the researcher's own degree of belief regarding possible parameter values, prior and posterior reflecting respectively relevant prior knowledge and that knowledge updated by the observed data. If the data are weak, an informative prior is likely to strongly influence parameter estimation, and the resulting posterior cumulative distribution function (CDF) is unlikely to approximate a confidence distribution.

Objective Bayesian approaches eschew use of a prior that reflects beliefs regarding the parameter value, and are usually used for inference in the absence of generally-agreed existing knowledge about parameter values. The aim is for the results to be – as for frequentist results – a function only of the data from which they are derived and the experimental setup, which determines both the model and the sampling plan. In order to achieve this, a "noninformative prior", which is mathematically derived from the assumed statistical model and has no probabilistic interpretation, must be used (Bernardo and Smith, 1994, p. 298 and p. 306; Kass and Wasserman, 1996; Bernardo, 2011). A noninformative prior is merely a tool for the generation of the desired posterior PDFs for the parameter(s) of interest (Bernardo and Smith, 1994, p. 306), and may be viewed as an appropriate mathematical weight function (Fraser et al., 2010). It is common to judge the merits of a noninformative prior by its "probability matching", i.e. how closely the resulting posterior probabilities agree with repeated sampling frequencies (Berger and Bernardo, 1992, p. 36; Kass and Wasserman, 1996). Although in some common cases parameter inference using a noninformative prior produces a posterior PDF that is an exact confidence density, in general this is not possible (Lindley, 1958).

Noninformative priors were originally developed using invariance considerations (Jeffreys, 1946). Jeffreys prior is the square root of the (expected) Fisher information – or of its determinant in multi-parameter cases. Fisher information – the expected value of the negative second derivative of the log-likelihood function with respect to the parameters – is a measure of the amount of information that the data, on average, carries about the parameter values. The more sophisticated reference analysis approach (Bernardo, 1979; Berger and Bernardo, 1992) uses information-theoretic concepts to derive a minimally-informative prior. In the univariate parameter continuous case, Jeffreys prior is normally the reference prior and is known