Journal of Statistical Planning and Inference 139 (2009) 4179 – 4192

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Bayesian inference for circular distributions with unknown normalising constants

Sourabh Bhattacharya[a,*], Ashis SenGupta[b]

[a] Bayesian and Interdisciplinary Research Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700108, India
[b] Applied Statistics Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700108, India

ARTICLE INFO

Article history:
Received 27 June 2008
Received in revised form
16 June 2009
Accepted 17 June 2009
Available online 21 June 2009

Keywords:
Bayesian inference
Circular statistics
Importance sampling
Markov chain Monte Carlo
Umbrella sampling
von-Mises distribution

ABSTRACT

Very often, the likelihoods for circular data sets are of quite complicated forms, and the functional forms of the normalising constants, which depend upon the unknown parameters, are unknown. This latter problem generally precludes rigorous, exact inference (both classical and Bayesian) for circular data.

Noting the paucity of literature on Bayesian circular data analysis, and also because realistic data analysis is naturally permitted by the Bayesian paradigm, we address the above problem taking a Bayesian perspective. In particular, we propose a methodology that combines importance sampling and Markov chain Monte Carlo (MCMC) in a very effective manner to sample from the posterior distribution of the parameters, given the circular data. With simulation study and real data analysis, we demonstrate the considerable reliability and flexibility of our proposed methodology in analysing circular data.

1. Introduction

Circular data arise in diverse scientific investigations; in biology and medicine, where biological rhythms exhibited by blood pressure, body temperature, etc., are of interest; migration directions taken by birds and insects, orientations of biological organisms; in ecology where interest often lies in wind direction and concentration of pollutants, such as ozone, orientations of rock cores in geology, in palaeoecology, where interest may lie in the study of palaeocurrents to infer about the direction of river flows in the past. Moreover, all periodic physical phenomena may be analysed using circular statistics—these may include arrival times of patients in a hospital in a day, the occurrence of airplane accidents over the year (uniform distribution of the occurrence of accidents indicates that the accidents are really accidents). Circular statistics finds application in physics too; indeed, the celebrated circular normal distribution (more on this subsequently) was the result of conversion of the fractional parts of the atomic weights of the 24 lightest elements into angles. Interestingly, it seems that circular data analysis may also be applied to sports like cricket; the bowling side may prevent an expert batsman from scoring too many runs by placing fielders in the direction at which the batsman is most likely to hit—the most likely direction may be obtained by analysing past records of the batsman, using available models for circular data. For many more examples, see Fisher (1993), Mardia and Jupp (1999), Jammalamadaka and SenGupta (2001).

However, it is not straightforward to extend conventional statistical approaches to analyse circular data. The difficulties stem primarily from the disparate topologies of the circle and the straight line; the former being concerned with circular data and the latter associated with conventional, linear data. The likelihoods of the circular distributions usually have complicated forms, with the normalising constant being unknown. This causes problems in inference, since the normalising constant invariably

* Corresponding author.
E-mail addresses: sourabh@isical.ac.in (S. Bhattacharya), ashis@isical.ac.in (A. SenGupta).

0378-3758/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2009.06.008

depends upon the model parameters. So far practical applications of circular statistical theory have been restricted to very low-dimensional circular models (typically, one-dimensional models). But such methods do not generalise to realistic, challenging, and high-dimensional situations.

Recognising the dearth of literature on Bayesian circular data analysis, in spite of the overwhelming influence of the Bayesian paradigm in other areas of scientific data analysis, in this paper we adopt the Bayesian viewpoint and introduce a new algorithm that combines importance sampling (IS) and Markov chain Monte Carlo (MCMC) to draw samples from the posterior of interest. Typically, at each iteration, the unknown normalising constant is estimated reliably by IS, irrespective of the dimensionality. Since the support of circular distributions is compact, the estimate based on IS is reliable in the sense that it has finite variance. We also propose a novel dynamic version of IS, where the IS density varies dynamically with each iteration. We argue that the latter captures the dynamic nature of the normalising constant, which being a function of model parameters, changes with iteration as the parameters are updated. Since our methodology is a combination of IS and MCMC, we refer to this as ISMCMC.

The rest of our paper is organised as follows. In Section 2 we provide examples of models for circular data, and note the challenges the unknown normalising constants pose to Bayesian inference for circular data. Our proposal, which combines IS and MCMC, and designed to tackle this problem, is introduced in Section 3. An overview of the literature on estimating (ratios of) unknown normalising constants and the connection with our proposal, is discussed in Section 4. In Section 5 we discuss issues related to effective choice of IS densities needed in our approach. Sequential and block updating of the parameters concerned is discussed in Section 6. Simulation study and real applications of our methodology are discussed in Sections 7 and 8, respectively. Conclusions and future work are discussed in Section 9.

2. Typical examples of models for directional data with unknown normalising constants

In the typical examples on circular data models that follow, $1/C$, as a function of unknown parameters, denotes the unknown normalising constant.

Univariate circular models:

(a)
$$f(\theta)=\frac{1}{C(\kappa,\nu)}\exp\{\kappa\cos(\theta+\nu\sin(\theta))\},\quad \theta,\nu\in[-\pi,\pi),\ \ \kappa\geq 0.$$

This model is symmetric, and can be used to model circular data that exhibit symmetry. Observe that if $\nu=0$, then the distribution is symmetric about zero; in fact, the model reduces to the well-known von-Mises (circular normal) distribution. Hence, the von-Mises distribution is a special case of model (a).

(b)
$$f(\theta)=\frac{1}{C(\kappa,\nu)}\exp\{\kappa\cos(\theta+\nu\cos(\theta))\},\quad \theta,\nu\in[-\pi,\pi),\ \ \kappa\geq 0.$$

Model (b) is suitable for modelling asymmetric circular data. As in (a), the von-Mises distribution is a special case of model (b) as well. Additional details of models (a) and (b) can be found in Batschelet (1981).

(c)
$$f(\theta)=\frac{1}{C(\kappa_1,\kappa_2,\mu_1,\mu_2)}\exp\{\kappa_1\cos(\theta-\mu_1)+\kappa_2\cos 2(\theta-\mu_2)\},\quad \theta,\nu_1,\nu_2\in[-\pi,\pi),\ \ \kappa_1,\kappa_2\geq 0.$$

Model (c), which is often referred to as the generalised von-Mises distributions in the literature, is most suitable for circular data that exhibit bimodality with different concentration around each mode (that is, densities at the two modes are different).

(d)
$$f(\theta)=\frac{1}{C(\kappa_n,\lambda_n,\mu,\nu)}\left[1-\frac{2}{n}\{k_n\cos(\theta-\mu)+\lambda_n\cos 2(\theta-\nu)\}\right]^{-(n+2)/2}\quad \text{where } \theta,\mu,\nu\in[-\pi,\pi),$$
$$\kappa_n=\frac{k}{1-2\dfrac{c}{n}},\quad \kappa\geq 0,\ \ c<0\ \text{ and }\ \lambda_n=\frac{\lambda}{1-2\dfrac{c}{n}},\quad \lambda\geq 0,\ \ \lambda+\kappa<\frac{n-2c}{2}.$$

Model (d) is quite general in that if the data fail to provide sufficient information whether or not bimodality should be modelled explicitly, then (d) may be used; small values of the parameter n makes bimodality pronounced, while for high values of n, bimodality fades out.

(e)
$$f_{\psi,\nu,\kappa}(\theta)=\frac{1}{C(\psi,\nu,\kappa)}\{1+\tanh(\kappa\psi)\cos(\theta-\nu)\}^{1/\psi},\quad \nu-\pi\leq\theta\leq\nu+\pi,\ \ \kappa\geq 0,\ \ -\infty<\psi<\infty.$$

Model (e) is a very general model proposed by Jones and Pewsey (2005). Details of this model and illustration of our proposed methodology on a real data application centred on model (e) have been provided in Section 8.1.

Bivariate circular (toroidal) models:

(f)
$$f(\theta,\phi)=\frac{1}{C(\kappa_1,\kappa_2,\kappa_3,\mu,\nu)}\exp\{\kappa_1\cos(\theta-\mu)+\kappa_2\cos(\phi-\nu)-\kappa_3\cos(\theta-\mu-\phi+\nu)\}$$

where $\theta,\mu,\phi,\nu\in[-\pi,\pi)$, and $\kappa_1,\kappa_2,\kappa_2\geq 0$.