Examining the Role of a Non-informative Prior Function Through Weakly Informative Prior Densities

Takemi Yanagimoto[1,3] and Toshio Ohnishi[2]

[1]Department of Industrial and Systems Engineering, Chuo University, Japan

[2]Faculty of Economics, Kyushu University, Japan

[3] Corresponding author: Takemi Yanagimoto, e-mail: yanagmt@indsys.chuo-u.ac.jp

Abstract

A non-informative prior function is discussed from the standpoint of a weakly informative prior density, though it has been pursued in relation only to the sampling densities. We note that a weakly informative prior density is useful for examining a no-informative prior function. Some disadvantages of a non-informative prior function are pointed out. On the other hand, a non-informative prior function is compared favorably with a degenerated prior density on an unknown point. The role of the posterior density compared with the marginal likelihood is discussed.

Keywords: Degenerated density, Haldane prior, Poisson mean, proper prior.

1. Introduction

Non-informative prior functions (densities), such as Jeffreys' and the reference prior ones, have been discussed in relation only to the sampling densities $\{p(x|\theta); \theta \in \Theta\}$, see Berger *et al.* (2009), for example. Defining a family of proper prior densities having a parameter representing the degree of our belief of prior information, we characterize a non-informative function as the rescaled lower limit of a sequence in the family. This approach allows us to shed new light on the role of a non-informative prior.

Examining the behavior of the induced posterior density and that of the credible region, we observe that some of familiar non-informative priors in the literature are informative in some sense. We observe also that the choice of a non-informative prior function among multiple candidates can be sensitive with the comparison of credible regions. In contrast, the role of a non-informative prior is positively evaluated as an alternative of a degenerated prior at an unknown point, which is largely different from a proper prior density.

2. Weakly informative density

Consider a family of proper prior densities based on a non-informative prior function $n(\theta)$ of the form

$$\mathcal{P} = \{\pi(\theta; c, g) = f(\theta;\, c, g)n(\theta) \cdot K(c, g) |\, g \in R^+\}. \tag{2.1}$$

An option of the choice of $f(\theta;\, c, g)$ is $p^g(m|\theta)$, as is in Ibrahim *et al.* (2001). Suppose that $\pi(\theta; c, g)$ converges weakly to $n(\theta)$ as g tends to 0. Then we may call $\pi(\theta; c, g)$ a weakly informative prior density, if g is close to 0. It is a proper prior density, and is regarded as a superior one to a non-informative prior function. Another promising choice is that $f(\theta;\, c, g) = \exp -g\mathrm{D}(\theta, c)$ with the Kullback-Leibler divergence.

3 Induced posterior density.

This prior density can be useful for examining non-informative prior functions. This is because a weakly informative prior density is actually proper, and standard treatments of the Bayesian procedure can be applied. We can examine a prior density through the posterior density. When the posterior density induced from a prior density becomes degenerated, such a prior density is to be avoided. In fact, the credible region results in a singleton. This indicates that the familiar Haldane prior for the binomial incidence probability is discouraging. The distance between the induced posterior density and that induced from a weakly informative density provides us with an approach to this subject, as was observed in Berger *et al.* (2009).

Another view is available from the fact that the stochastic ordering is often observed between the posterior densities induced from two existing non-informative prior functions. This indicates that at least one of the two non-informative prior functions is actually informative. In other words, it is not easy to define a non-informative prior function in its own meaning. Further, the two induced posterior densities are sensitive with this difference between two prior densities, as is claimed in Gelman (2006).

4 Degenerated posterior density.

A weakly prior density is also useful for examining the role of the degenerated prior density on an unknown point. Recall that the comparison study between a non-informative prior function and the degenerated prior density is often complicated. When this prior density is employed, the marginal likelihood is used. Note that the marginal likelihood takes a small value, when a weakly informative prior density is based on an improper prior function. This means that the marginal likelihood is a criterion evaluating a non-informative prior function better that a proper prior density. This indicates the necessity of a criterion based on the posterior density, such as DIC in Spiegelhalter *et al.* (2002).

The other reservation about this prior density is the fact that it yields an improper posterior density, which is discussed Yanagimoto and Ohnishi (2012). In light of the important role of the posterior density this fact is discouraging

As many authors stressed, including Yanagimoto and Ohnishi (2011), the use of degenerated prior density should be minimized in Bayesian inference. This is a main issue of the compromise between Bayesian and the frequentist approches.

References

Berger, J.O., Bernardo, J.M. and Sun, D. (2009). The formal definition of reference priors. *Ann. Statist.* 37, 905-938.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian anal.*, 1, 515-534.

Ibrahim, J.G. and Chen, M-H. and Sinha, D. (2001). *Bayesian Survival Analysis.* Springer-Verlag, New York.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussions). *J. Roy. Statist. Soc., B*, **64**: 583 - 639.

Yanagimoto, T. and Ohnishi, T. (2011). Saddlepoint condition on a predictor to reconfirm the need for the assumption of a prior distribution. *J. Statist. Plann. Inf.*, **41**, 1990-2000.

Yanagimoto T. and Ohnishi T. (2012). Permissible boundary prior function as a virtually proper prior density. ISM-RM 1158R.