Discussion of the papers by J. Ghosh and N. Reid

Brunero Liseo
Sapienza Università di Roma
Saturday June 6, 2009

Alternative titles

- Ghosh et al. : "(Almost) all roads lead to Jeffreys' (and reference) prior.

- Reid et al.: To what extent objective Bayes and "good" frequentist estimation procedures may match?

or

- Is it possible to find a method for objective prior which is harder to compute than reference prior ?!?

Main goal of OB inference

The search of a prior distribution to complete a "classical" statistical model

The search must consider several desiderata (Bernardo and Ramòn, 1998, The statistician)

1) Invariance
2) Consistent marginalization
3) Consistent sampling probabilities
4) Generality
5) Admissibility

The Jeffreys' and reference priors seems to satisfy to a large extent the above desiderata.

Ghosh

- Ghosh’s talk reverses the perspective and try to answer (among others) the question:

What are the most *general* assumptions which produce, as a “good“ OB prior, the Jeffreys' (reference) prior?

(at least…) 2 different answers

A) Probability matching priors (Frequentist road, point 3)

B) Different Information measures (Bayesian road, p.2)

Reid's talk follows Road A

Keystone relation:

$$dF(y^0;\theta) = 0$$

or

$$dy = -\frac{F_\theta(y^0;\theta)}{F_y(y^0;\theta)}d\theta = V(\theta)d\theta$$

Reid's talk follows Road A

Main messages in the conclusions:

1. To obtain good B/F matching beyond 1° order approximation, prior needs to depend on the data

1. Under special conditions, and taking the observed value y_0 as fixed, one can reparameterize the model in order to have approximate location parameters.

1. The "Jacobian" of the reparameterization, in a sense, produces the "weight" function which can be used as a default prior

- This is an alternative (I would not say simpler ...) to the reference prior approach which is based on information divergence arguments …

- How easy to check sampling properties of the OB procedure?

- Data-dependence (dangerous…)

Road B

- There is a class of invariant measures of information (including KL divergence, L^1 , Hellinger)
- L^2 does not work (is this related to the lack of invariance of the measure?)
- How about χ^2 distance?

… in general KL, Hellinger and χ^2 have a similar behaviour (Borovkov, 1998) and all of them are related (at least locally) to Fisher-Rao metrics

$$\rho(\theta, \theta + d\theta) = \sum_i \sum_j I_{ij}(\theta) \partial\theta_i \partial\theta_j \,(1 + o(1))$$

Ghosh Final statement

"Information in a prior cannot be measured except in the context of a model …"

Deep Implications…:

1) !!!!!!

1) The same prior can have different relative effects in two different models! (see also Reid's talk!)
2) The statement is not entirely compatible with the subjective Bayesian way of thinking about the parameter as a real quantity and not a Greek letter!! (Lindley 1990, Stat. Sci)

I have used Lindley's statement several times to convince frequentist statisticians that Bayes was better...

I have used Lindley's statement several times to convince frequentist statisticians that Bayes was better…

But:

To become obsessed with defining real-world things is to miss completely the subtlety of Mathematics (D. Williams, 2001, Weighing the odds, Cambridge, pag.220)

A new agenda for OB

Based on a suggestion of prof. Ghosh:

- *What are the most important issues to consider in the construction of objective priors?*

A new agenda for OB

1) *O'Bayes evaluates an objective prior by the posterior it produces.* (ON AVERAGE)

2) *Only those improper priors can be used which produce proper posterior.* *(YES!)*

3) *Information in a prior has a representation of the form* $A_n + B_n + C_n(1+o(1))$ *in diminishing orders of magnitude.* *(Role of the prior, Reid talk)*

A new agenda for OB

4) Information in a prior cannot be defined except in the context of an experiment: with data from an experiment, we define information in the prior by comparing prior and posterior. In the planning stage we take an average over all posteriors. (This is crucial and already commented)

A new agenda for OB

5) Coherence (YES!)

6) Shouldn't there be some connections between OB and Robust Bayes?

Sure! Ideally one could use his/her elicitation to produce

a class of priors and then derive a sort of "constrained" objective prior.

This would be great, although the math is going to be in general intractable …

7) Distinction between parameter of interest and nuisance parameters