Objective Bayes Model Comparisons for Categorical Data

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Ioannis Ntzoufras

(ntzoufras@aueb.gr)

22–23 October 2018: Challenges in Categorical Data Analysis @ Aachen

Objective & Bayes?

1. This is an “oxymoron” since Bayes is by definition subjective.
2. It is a “marketing” term for the implementation of the Bayesian methods under the absence of prior information; the other alternative is “Default Bayes”.
3. Even I do not like the term, but I wrote a review paper in *Bayesian Analysis* in 2018 co-authored with G. Consonni, D. Fouskakis and B. Liseo.
4. O’Bayes has long tradition within ISBA (13 biannual meeting with over 100 participants per meeting $\Rightarrow$ so it is a real thing).
5. Research focuses on Default priors for inference, for Model comparisons, Prior combatibility across models, Bayesian Non Parametrics, Shrinkage methods for large p small n problems.

Bayesian Model Comparison

Posterior Odds (PO) between models M_0 and M_1 is given by

$$PO_{01} \equiv \frac{\pi(M_0|\boldsymbol{y})}{\pi(M_1|\boldsymbol{y})} = \frac{m_0(\boldsymbol{y})}{m_1(\boldsymbol{y})} \times \frac{\pi(M_0)}{\pi(M_1)} = BF_{01} \times O_{01} \tag{1}$$

which is a function of the **Bayes Factor** (BF_{01}) and the **Prior Odds** (O_{01}).

In the above $m_\ell(\boldsymbol{y})$ is the marginal likelihood under model M_ℓ and $\pi(M_\ell)$ is the prior probability of model M_ℓ given by

$$m_\ell(\boldsymbol{y}) = \int f_\ell(\boldsymbol{y}|\boldsymbol{\theta}_\ell)\pi_\ell(\boldsymbol{\theta}_\ell)d\boldsymbol{\theta}_\ell, \tag{2}$$

where $f_\ell(\boldsymbol{y}|\boldsymbol{\theta}_\ell)$ is the likelihood under model M_ℓ with parameters $\boldsymbol{\theta}_\ell$ and $\pi_\ell(\boldsymbol{\theta}_\ell)$ is the prior distribution of model parameters given model M_ℓ.

The Lindley-Bartlett-Jeffreys Paradox

For a single model inference $\Rightarrow$ a highly diffuse prior on the model parameters is often used (to represent ignorance).

$\Rightarrow$ Posterior density takes the shape of the likelihood and is insensitive to the exact value of the prior density function.

For multiple models inference $\Rightarrow$ BFs (and POs) are quite sensitive to the choice of the prior variance of model parameters.

$\Rightarrow$ For nested models, we support the simplest model with the evidence increasing as the variance of the parameters increase ending up to support of more parsimonious model no matter what data we have.

$\Rightarrow$ Under this approach, the procedure is quite informative since the data do not contribute to the inference.

$\Rightarrow$ Improper priors cannot be used since the BFs depend on the undefined normalizing constants of the priors.

Principles for O'Bayes Model Comparisons

- Compatibility of priors.
- Validation of Bayesian approaches.
- Methods with good frequentist properties

  (FDR control - application in Quality control and clinical trials).
- Criteria for objective Bayesian model choice (Bayarri et al., 2012; Annals Stat.).
  - Propriety;
  - Model Selection Consistency; Information consistency; intrinsic consistency;
  - Predictive matching;
  - Measurement Invariance; Group Invariance

Tools for O'Bayes Model Comparisons

- The Unit information principle.
- Training Samples $\Rightarrow$ Intrinsic Bayes Factors $\Rightarrow$ Intrinsic Priors.
- Imaginary Data.
  - Fixed Imaginary Data $\Rightarrow$ power prior $\Rightarrow$ g-prior $\&$ its mixtures.
  - Random Imaginary Data $\Rightarrow$ Expected posterior prior $\&$ Power-EPPs.
- Emprirical Bayes approaches.
- Non-local prior approaches.

Examples of O'Bayes Model Comparisons for Categorical Data

- Priors developed (g-priors, hyper-g priors) can be directly implemented in Poisson log-linear formulations for contingency tables.
- Power-priors and imaginary data can be also extended for models for contingency tables (work in progress for association models with K.Mantzouni and Maria Kateri).
- Power-priors and imaginary data can lead to several well known default prior formulations for models for contingency tables (work in progress with K.Mantzouni and C. Tarantola).
- Approaches based on training samples (i.e. Intrinsic Bayes Factors) are more difficult to implemented due to sparsity and no clear definition of minimal training samples.


Phew! Did I manage to say all this in 5 minutes?


Don't go! There is more...

- **Statistics**
- **Greeks**
- **Italians**
- **In an Island!**
- **WHAT Can you Ask for more?**

5th Meeting on Statistics

7-9 September 2019
Aigina or Naxos ?