Distribuciones Globales de Referencia

José M. Bernardo
Universidad de Valencia, Spain
jose.m.bernardo@uv.es

SEIO 2012
XXXIII Congreso Nacional de Estadística e I.O.
Madrid, 17 de Abril de 2012

Summary

(i) **Approximate Reference Priors**

Motivation

Definition

(ii) **Hardy-Weinberg Equilibrium**

(iii) **Multinomial Model**

(iv) **Basic References**

Global Reference Priors

- Reference priors are defined for an ordered parameterization. Given model $\mathcal{M}_{\boldsymbol{z}} = \{p(\boldsymbol{z} \,|\, \boldsymbol{\omega}), \boldsymbol{z} \in \boldsymbol{\mathcal{Z}}, \boldsymbol{\omega} \in \boldsymbol{\Omega}\}$ with m parameters, the (joint) reference prior $\pi_{\phi_1}(\boldsymbol{\phi})$ required to obtain the marginal reference posterior of ϕ_1 (via Bayes theorem and appropriate integration) is sequentially obtained as

$$\pi_{\phi_1}(\boldsymbol{\phi}) = \pi(\phi_m \,|\, \phi_{m-1}, \ldots, \phi_1) \times \cdots \times \pi(\phi_2 \,|\, \phi_1)\, \pi(\phi_1).$$

- However, one is often simultaneously interested in several functions of the parameters. Given $\mathcal{M}_{\boldsymbol{z}} = \{p(\boldsymbol{z} \,|\, \boldsymbol{\omega}), \boldsymbol{z} \in \boldsymbol{\mathcal{Z}}, \boldsymbol{\omega} \in \boldsymbol{\Omega} \subset \Re^m\}$ with m parameters, consider a set $\boldsymbol{\theta}(\boldsymbol{\omega}) = \{\theta_1(\boldsymbol{\omega}), \ldots, \theta_r(\boldsymbol{\omega})\}$ of $r > 1$ functions of interest. A global reference prior, which may be proved to provide good approximate reference posteriors for each of the θ_i's is then required.

- If there is a single joint prior $\pi_{\boldsymbol{\theta}}(\boldsymbol{\omega})$ whose corresponding marginal posteriors are precisely equal to the reference posteriors for each of the θ_i's so that, for all $\boldsymbol{z}$ values, $\pi_{\boldsymbol{\theta}}(\theta_i \mid \boldsymbol{z}) = \pi(\theta_i \mid \boldsymbol{z})$, then this should be a solution. There may be may other priors which satisfy this condition.
- If the joint reference priors for the θ_i are all equal, then $\pi_{\boldsymbol{\theta}}(\boldsymbol{\omega}) = \pi_{\theta_i}(\boldsymbol{\omega})$ is defined to be *the* solution to the problem posed. For instance, in the univariate normal model, this implies that $\pi(\mu, \sigma) = \sigma^{-1}$, which is the reference prior when either μ or σ are of interest, should also be used to make joint inferences for (μ, σ), or to obtain a reference predictive distribution.
- Generally there will not exist a single joint prior $\pi_{\boldsymbol{\theta}}(\boldsymbol{\omega})$ which would yield marginal posteriors for each of the θ_i's which are all equal to the corresponding reference posteriors. Hence, an approximate solution must be found. This may be implemented using intrinsic discrepancies:

Definition

- Berger, Bernardo and Sun (work in progress) suggest a procedure to select a joint prior $\pi_{\boldsymbol{\theta}}(\boldsymbol{\omega})$ whose corresponding marginal posteriors $\{\pi_{\boldsymbol{\theta}}(\theta_i \mid \boldsymbol{z})\}_{i=1}^r$ will be close, for all possible data sets $\boldsymbol{z} \in \boldsymbol{\mathcal{Z}}$, to the set of reference posteriors $\{\pi(\theta_i \mid \boldsymbol{z})\}_{i=1}^r$ yielded by the set of reference priors $\{\pi_{\theta_i}(\boldsymbol{\omega})\}_{i=1}^r$ derived under the assumption that each of the θ_i's is of interest.
- The idea behind the Definition is to select some mathematically tractable family of prior distributions for $\boldsymbol{\omega}$, and to choose that element within the family which minimizes the average expected intrinsic discrepancy between the marginal posteriors for the θ_i's obtained from that prior and the corresponding reference posteriors.

Definition 1 *Consider model* $\mathcal{M}_{\boldsymbol{z}} = \{p(\boldsymbol{z}\,|\,\boldsymbol{\omega}), \boldsymbol{z} \in \mathcal{Z}, \boldsymbol{\omega} \in \boldsymbol{\Omega}\}$ *and* $r > 1$ *functions of interest,* $\{\theta_1(\boldsymbol{\omega}), \ldots, \theta_r(\boldsymbol{\omega})\}$*. Let* $\{\pi_{\theta_i}(\boldsymbol{\omega})\}_{i=1}^r$ *be the relevant reference priors, and* $\{\pi_{\theta_i}(\boldsymbol{z})\}_{i=1}^r$ *and* $\{\pi(\theta_i\,|\,\boldsymbol{z})\}_{i=1}^r$ *the corresponding prior predictives and marginal posteriors. Let* $\mathcal{F} = \{\pi(\boldsymbol{\omega}\,|\,\boldsymbol{a}), \boldsymbol{a} \in \mathcal{A}\}$ *be a family of prior functions. For each* $\boldsymbol{\omega} \in \boldsymbol{\Omega}$*, the best approximate joint reference prior within* $\mathcal{F}$ *is that which minimizes the average expected intrinsic loss*

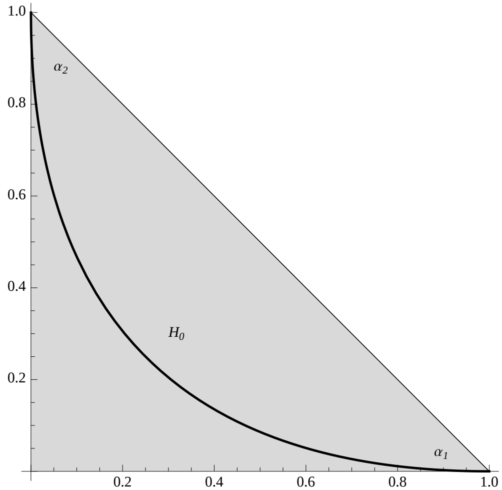
$$d(\boldsymbol{a}) = \frac{1}{r}\sum_{i=1}^{r}\int_{\mathcal{Z}} \delta\{\pi_{\theta_i}(\cdot\,|\,\boldsymbol{z}),\, p_{\theta_i}(\cdot\,|\,\boldsymbol{z},\boldsymbol{a})\}\,\pi_{\theta_i}(\boldsymbol{z})\,d\boldsymbol{z}, \quad \boldsymbol{a} \in \mathcal{A}.$$

• Some illustrative examples are now described, the analysis of the Hardy-Weinberg equilibrium, and the multinomial model (with important applications in the analysis of contingency tables).

Hardy-Weinberg Equilibrium

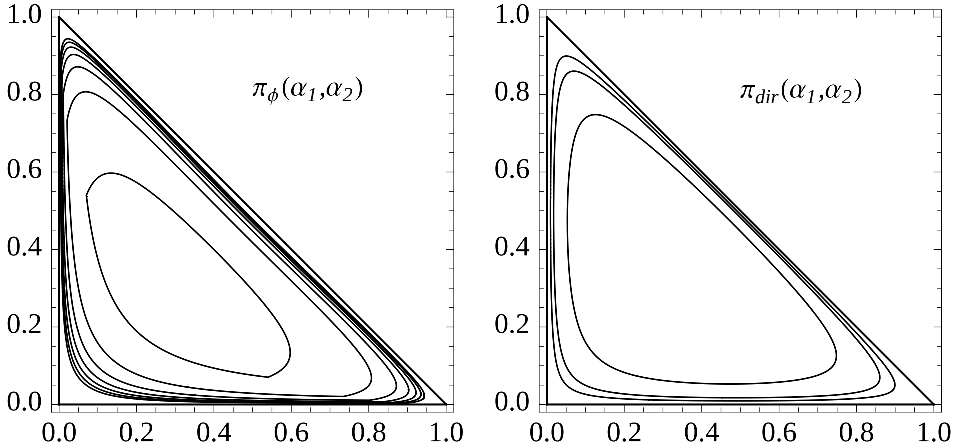
- To determine whether or not a population mates randomly:
- At a single autosomal locus with two alleles, a diploid individual has three possible genotypes, $\{AA, aa, Aa\}$, with (unknown) population frequencies $\{\alpha_1, \alpha_2, \alpha_3\}$, where $0 < \alpha_i < 1$ and $\sum_{i=1}^{3} \alpha_i = 1$.
- Hardy-Weinberg (HW) equilibrium iff $\exists\, p = \Pr(A)$, such that $\{\alpha_1, \alpha_2, \alpha_3\} = \{p^2, (1-p)^2, 2p(1-p)\}$.
- Given a random sample of size n from the population, and observed $\boldsymbol{z} = \{n_1, n_2, n_3\}$ individuals (with $n = n_1 + n_2 + n_3$) from each of the three possible genotypes $\{AA, aa, Aa\}$, the question is whether or not these data support the hypothesis of HW equilibrium.
- This is a good example of *precise* hypothesis in the sciences:

• The null is $H_0 = \{(\alpha_1, \alpha 2);\ \sqrt{\alpha_1} + \sqrt{\alpha_2} = 1\}$, a zero measure set within the (simplex) parameter spate of a trinomial distribution.



• The parameter of interest is is the intrinsic divergence of H_0 from the model, $\phi(\alpha_1, \alpha_2) = \delta\{H_0, \mathrm{Tri}(r_1, r_2, r_3 \,|\, \alpha_1, \alpha_2)\}$

• The closest Dirichlet prior to the reference prior $\pi_\phi(\alpha_1, \alpha_2)$ when $\theta(\alpha_1, \alpha_2)$ is of interest is is $\text{Di}[\alpha_1, \alpha_2 \,|\, 1/3, 1/3, 1/3]$ (Bernardo and Tomazella, 2010).



• The two priors are pretty similar, and produce qualitatively similar results when testing the null that the population is is HW equilibrium.

Multinomial Model

- Consider a multinomial model with m categories and parameters $\{\theta_1, \ldots, \theta_{m-1}\}$, define $\theta_m = 1 - \sum_{i=1}^{m-1} \theta_i$, and let the functions of interest be the m probabilities $\{\theta_1, \ldots, \theta_m\}$. Let $\boldsymbol{z} = \{n_1, \ldots, n_m\}$ be the results observed from a random sample of size n.
- The reference prior for θ_i depends on i, and the corresponding reference posterior of θ_i is the beta distribution $\pi(\theta_i \,|\, \boldsymbol{z}) = \pi(\theta_i \,|\, n_i, n) = \text{Be}(\theta_i \,|\, n_i + 1/2, n - n_i + 1/2)$ (Berger and Bernardo, 1992a) which only depends on the number of observations n_i which fall in category i and on the total number n of observations (therefore avoiding the partition paradox which occurs when the posterior for θ_i depends on the total number m of categories considered).

- Consider the family of (proper) Dirichlet priors of the form $p(\theta \mid a) \propto \prod_{i=1}^{m} \theta_i^{a-1}$, with $a > 0$. The corresponding marginal posterior distribution of θ_i is $\text{Be}(\theta_i \mid n_i + a, n - n_i + (m-1)a)$ (notice the dependence on the number m of categories).
- The intrinsic discrepancy $\delta_i\{a \mid n_i, m, n\}$ between this distribution and the corresponding reference prior is naturally depends on n_i and has an explicit expression in terms of Gamma functions (Bernardo, 2011). The reference predictive for n_i is

$$\pi(n_i \mid n) = \int_0^1 \text{Bi}(n_i \mid n, \theta_i)\, \text{Be}(\theta_i \mid 1/2, 1/2)\, d\theta_i \cdot$$

and the average expected intrinsic loss of using a joint Dirichlet prior with parameter a with a sample of size n is obtained as

$$d(a \mid m, n) = \sum\nolimits_{n_i=0}^{n} \delta\{a \mid n_i, m, n\}\, \pi(n_i \mid n).$$

- By the symmetry of the problem, the m parameters $\{\theta_1, \ldots, \theta_m\}$ yield all the same expected loss. It may be verified that the function $d(a \mid m, n)$ is concave, with a unique minimum numerically found to be at $a^* \approx 1/m$.
- Thus, the best global Dirichlet prior when one is interested in all the cells of a multinomial model (and therefore in all the cells of a contingency table) is that with parameter $a = 1/m$, yielding an approximate marginal reference posterior $\mathrm{Be}(\theta_i \mid n_i + 1/m, n - n_i + (m-1)/m)$, with mean $(n_i + 1/m)/(n + 1)$.
- This is an important result in the analysis of sparse frequency and contingency tables.

Basic References

(In chronological order)

Available on line at **www.uv.es/bernardo**

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc. B* **41**, 113–147 (with discussion).

Berger, J. O. and Bernardo, J. M. (1992a). Ordered group reference priors with applications to a multinomial problem. *Biometrika* **79**, 25–37.

Berger, J. O. and Bernardo, J. M. (1992b). On the development of reference priors. *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) Oxford: University Press, 35–60 (with discussion).

Bernardo, J. M. (2005a). Reference analysis. *Bayesian Thinking: Modeling and Computation, Handbook of Statistics* **25** (Dey, D. K. and Rao, C. R., eds). Amsterdam: Elsevier, 17–90.

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

Bernardo, J. M. and Tomazella, V. (2010). Bayesian reference analysis of the Hardy-Weinberg equilibrium. *Frontiers of Statistical Decision Making and Data Analysis. In Honor of James O. Berger* (M.-H. Chen, P. Müller, D. Sun, K. Ye, and D. K. Dey, eds.) New York: Springer, 31–43.

Bernardo, J. M. (2011). Objective Bayesian estimation and hypothesis testing. *Bayesian Statistics 9* (J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds.) Oxford: University Press, 1–68 (with discussion).

Berger, J. O., Bernardo, J. M. and Sun, D. (2012). Objective priors for discrete parameters. *J. Amer. Statist. Assoc.* **107** (to appear).

- **Many thanks for your attention!**

jose.m.bernardo@uv.es
www.uv.es/bernardo

- **Recently published (October 2011)**

Bayesian Statistics 9

- J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, (eds.) Oxford, UK: Oxford University Press, October 2011.
- ISBN 978-0-19-969458-7, Price: £120.00 $\approx$ 143 Eur.
- http://ukcatalogue.oup.com/product/9780199694587.do