On the Problem of Evaluating Formal Posterior Distributions

Morris L. Eaton
Professor Emeritus

School of Statistics
University of Minnesota

Acknowledgement: This presentation is based in part on joint work with J. P. Hobert, G. L. Jones and Wen-Lin Lai. Here are two relevant references:

- Eaton, M. L., Hobert, J. P., Jones, G. L., and Wen-Lin Lai(2008).”Evaluation of formal posterior distributions via Markov chain arguments”. *Ann. Statist.*, Vol 36, p. 2423-2452.

- Eaton, M. L., Hobert, J. P., and Jones, G. L.(2007). ”On perturbations of strongly admissible prior distributions”. *Ann. Inst. H. Poincaré Probab. Statist.*, Vol 43, p. 633-653.

Notation and Assumptions

- A model $P(dx|\theta)$ for data X is given

- A sigma finite improper prior ν is specified

- Assume the marginal measure for X given by
$$M(dx) = \int P(dx|\theta)\,\nu(d\theta)$$
is sigma finite. Then the formal posterior $Q(d\theta|x)$ exists and satisfies
$$P(dx|\theta)\nu(d\theta) = Q(d\theta|x)M(dx)$$

It is the formal posterior Q that we wish to evaluate.

Estimation with Quadratic Loss

- Consider estimating $\phi(\theta)$, which is a k dimensional vector function, when the loss function is

$$L(a, \theta) = ||a - \phi(\theta)||^2$$

- The formal Bayes estimator is just $\widehat{\phi}(x)$ which is the posterior mean (under $Q(.|x)$) of $\phi(\theta)$

- If $\widehat{\phi}$ is almost-ν-admissible for all bounded ϕ, we say that Q, or equivalently ν, is *strongly admissible*(SA).

Under what conditions is Q SA?

Example - Multivariate Normal

- Suppose X is p-dimensional multivariate normal with mean vector θ and covariance the identity.

- Assume the improper prior has the form

  $$\nu(d\theta) = g(||\theta||^2)d\theta$$

  where $d\theta$ is Lebesgue measure

- Under what conditions on g will Q be SA? When $p = 1$ or $p = 2$, we suspect that if $g = 1$ then perhaps Q will be strongly admissible. But when p is greater than or equal to 3, we suspect that $g = 1$ will perhaps not yield an SA Q because of Stein shrinkers. However other g's may yield SA Q's.

Example Continued: Here are some things that we know—we will discuss the proofs as we proceed.

- For $p = 1, 2$ and $g = 1$, Q is SA.

- For nonnegative α, β and t, consider

$$g_\alpha(t) = (\alpha + t)^{-\beta}$$

  and

$$g_\alpha(||\theta||^2)d\theta = \nu(d\theta)$$

  For $p > 2$, consider the two intervals

  $I_+ = [(p-2)/2, p/2]$ and $I_0 = [(p-2)/2, p/2)$

- **Theorem:** Assume $p > 2$.

  (i) If $\alpha = 0$ and $\beta \in I_0$, Q is SA.

  (ii) If $\alpha > 0$ and $\beta \in I_+$, Q is SA.

Basic Idea of Proofs -General Theory

The basic idea is to introduce a Markov Chain (MC) associated with the model and the improper prior. Then we show recurrence of the MC implies that the formal posterior Q is SA.

The transition function of the MC is the expected value of the formal posterior

$$R(C|\theta) = E_\theta Q(C|X)$$

where expectation is computed under the model when θ is the truth.

A basic object in our study is the measure

$$S(d\theta, d\eta) = R(d\eta|\theta)\nu(d\theta)$$

The measure S is symmetric and has ν as its marginals. Further ν is a stationary measure for R.

Some Notation A subset of Θ, say C, is ν-proper if the ν measure of C is positive and finite. For such a set, let

$$V(C) = \{\, g | I_C \le g,\, g \in L_1(\nu)\,\}$$

Associated with the measure S is the Dirichlet form

$$\Delta(h) = (1/2) \iint (h(\theta) - h(\eta))^2 \, S(d\theta, d\eta)$$

where h is in $L_2(\nu)$.

Note that $\Delta(\sqrt{g})$ is well defined for $g \in V(C)$.

A ν proper set C is locally-ν-recurrent (lνr) if the MC associated with R returns to C wp1 when it starts in C (except for a set of starting values of ν measure zero).

Some Main Theorems

Theorem 1: If for all ν proper C,

$$(1) \quad \inf_{g \in V(C)} \Delta(\sqrt{g}) = 0$$

then Q is SA.

Theorem 2: The ν proper set C is lνr if and only if (1) holds.

The main idea behind Theorem 1 is that the Blyth-Stein condition for admissibility holds when (1) holds because the integrated risk difference based on a prior $g(\theta)\nu(d\theta)$ is , for each bounded ϕ, bounded above by a constant times $\Delta(\sqrt{g})$.

These two results are what provide a connection between the behavior of the MC and admissibility. The results are hard to use because (1) needs to be verified for all ν proper sets C.

Some helpful technical results

- A set C is *recurrent* if, after time 0, the MC hits C wp1 for all starting values.

  Proposition: If the chain has one ν proper recurrent set, then all ν proper sets are lνr and Q is SA.

- In some situations it is possible to partition Θ as $\cup\{\,\Theta_t|0 \le t < \infty\,\}$ and write

  $$\nu(d\theta) = \xi(d\theta|t)\, m(dt)$$

  where m is a sigma finite measure on $[0, \infty)$ and $\xi(.|t)$ is a probability measure on Θ_t for each t. With the new model

  $$P^*(dx|t) = \int P(dx|\theta)\, \xi(d\theta|t)$$

  and improper prior m, recurrence of the chain on $[0, \infty)$ implies recurrence of the original chain on Θ and hence SA. (See EHJL(2008)).

Normal example continued:

- For $p = 1, 2$, the MC is a symmetric random walk on Θ and is known to be recurrent. Hence SA obtains.

- For $p \geq 3$ and g_α indicated earlier, we reduce the MC on $\Theta = R^p$ to a MC on $[0, \infty)$, and then verify moment conditions (3 of them) that give recurrence —this is where the restrictions on α and β arise. On $[0, \infty)$ we show that, for appropriate α and β, there is one recurrent set. Hence all ν proper sets are lνr and so SA obtains. The calculations are not beautiful (see EHJL(2008)).

An Extension:

From EHJ(2007) we have the following general result:

- Theorem: Suppose ν is SA. Let ψ be a bounded non-negative real valued function on Θ. Consider the new prior $\psi(\theta)\nu(d\theta)$. Under certain regularity conditions, the new prior is also SA

- The Theorem above applies to our normal example so we have a very wide class of priors for which SA obtains. Namely any prior which is a bounded non-negative function times an appropriate g_α is SA. The basic restriction is that the prior behave like a bounded function times a prior with specific tail behavior.

The Unbounded Case

- A few general things are known when ϕ is unbounded (see Eaton(2001)). In particular the transition function of the relevant MC now depends on ϕ, but recurrence of this chain does imply "admissibility" of the formal Bayes estimator (with regularity of course).

- For the normal example and the particular function $\phi(\theta) = \theta$, Brian Shea (a graduate student at Minnesota) has some interesting results.