

LETTER TO THE EDITOR

ON A FAMILY OF PRIOR DISTRIBUTIONS FOR A CLASS OF BAYESIAN SEARCH MODELS

K. D. GLAZEBROOK,* *University of Newcastle upon Tyne*

Abstract

We propose a two-parameter family of conjugate prior distributions for the number of undiscovered objects in a class of Bayesian search models. The family contains the one-parameter Euler and Heine families as special cases. The two parameters may be interpreted respectively as an overall success rate and a rate of depletion of the source of objects. The new family gives enhanced flexibility in modelling.

BAYES SEQUENTIAL DECISION RULE; GITTINS INDEX; OPTIMAL STOPPING; SEARCH MODELS

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 62C10
SECONDARY 60G40

Benkherouf and Bather (1988) and Benkherouf et al. (1992) considered Bayesian sequential decision models for a class of search problems modelling oil exploration. In these models $\pi = (\pi_0, \pi_1, \dots)$ is a prior on the number of undiscovered objects, v is the value of a discovery, c is the cost of a search, where $v > c > 0$, and α is a discount rate. A search is conducted at times $t = 0, 1, 2, \dots$ until further searching is uneconomic. Each search finds either a single object or nothing. The conditional probability of a search failing to find an object given that there are n undiscovered objects to be found is q^n , $0 < q < 1$. This geometric form is necessary to ensure that the numbers of successes and failures to date yield a sufficient statistic for the problem. Bather (1992) has proposed a continuous-time version of this model. The goal is to find a stopping rule which maximises the total net (discounted) value of searching.

Benkherouf and Bather (1988) proposed two conjugate families of priors for the problem, which they called Euler and Heine. They are both one-parameter families, the parameter representing an initial success rate in each case. These families give limited flexibility in modelling. We propose a conjugate two-parameter family, which contains the Euler and Heine families as special cases. The first parameter (λ) is again a success rate, whereas the second (θ) may be thought of as a rate of depletion of the source of objects. Together, they give us enhanced flexibility in modelling how rapidly the success rate decreases as objects are found.

Kemp (1992) discusses the Euler and Heine distributions from a different point of view. She identifies them as q -series analogues of the Poisson distribution and introduces a new distribution (the pseudo-Euler) which shares this property. The pseudo-Euler is not a member of our two-parameter family. The question of whether our two-parameter family can

Received 17 February 1993; revision received 27 April 1993.

* Postal address: Department of Mathematics and Statistics, The University, Newcastle upon Tyne NE1 7RU, UK.

be helpfully studied from the Kemp (1992) perspective is a topic for further investigation.

Suppose we begin our development with the requirement that the posterior following s successes should be the same as the posterior following f failures. This yields via Bayes' theorem that

$$\pi_{n+s}/\pi_n \propto q^{nf} \left\{ \prod_{i=1}^s (1 - q^{n+i}) \right\}^{-1}, \quad n \in \mathbb{N},$$

which suggests the form

$$(1) \quad \pi_{n+1}/\pi_n \propto q^{nf/s} (1 - q^{n+1})^{-1}, \quad n \in \mathbb{N}.$$

If we write $\theta = f/s$ then (1) implies the prior $\pi(\lambda, q, \theta)$ given by

$$(2) \quad \pi_n(\lambda, q, \theta) = \pi_0(\lambda, q, \theta) \lambda^n q^{\theta n(n-1)/2} \left\{ \prod_{i=1}^n (1 - q^i) \right\}^{-1}, \quad n \in \mathbb{N},$$

with associated parameter space $\{\theta = 0, 0 < \lambda < 1\} \cup \{\theta > 0, \lambda > 0\}$. This is our two-parameter family. (Recall that q is fixed.) We obtain the Euler distribution $E(\lambda, q)$ and the Heine distribution $H(\lambda, q)$ by taking $\theta = 0$ and 1 respectively in (2).

That this is a conjugate family emerges from the following lemma, which follows from Bayes' theorem. In the result, π^S and π^F denote the posterior following a single success or failure respectively.

Lemma 1. $\pi = \pi(\lambda, q, \theta) \Rightarrow \pi^S = \pi(\lambda q^\theta, q, \theta)$ and $\pi^F = \pi(\lambda, q, \theta)$.

In Lemma 2, ' $>$ ' denotes the usual likelihood ratio ordering. The proof is straightforward.

Lemma 2. $\lambda > \mu \Rightarrow \pi(\lambda, q, \theta) > \pi(\mu, q, \theta)$.

Now write

$$p(\lambda, q, \theta) \triangleq \sum_{n>0} (1 - q^n) \pi_n(\lambda, q, \theta)$$

for the (unconditional) probability of a successful search when the prior is $\pi(\lambda, q, \theta)$. Simple algebra yields

$$p(\lambda, q, \theta) = \lambda \pi_0(\lambda, q, \theta) (\pi_0(\lambda q^\theta, q, \theta))^{-1}.$$

The following result is a consequence of Lemma 2.

Corollary 1. $\lambda \cong \mu \Rightarrow p(\lambda, q, \theta) \cong p(\mu, q, \theta)$.

Proof. The likelihood ratio ordering of Lemma 2 implies a stochastic ordering of the distributions concerned. Hence we conclude that

$$\lambda \cong \mu \Rightarrow \sum_{k \cong n} \pi_k(\lambda, q, \theta) \cong \sum_{k \cong n} \pi_k(\mu, q, \theta), \quad n \in \mathbb{N}.$$

The result now follows from standard properties of stochastic ordering and the fact that $1 - q^n$ is increasing in n .

With the above results in place, we can now follow the analysis of Benkherouf and Bather (1988). Note that with prior $\pi(\lambda, q, \theta)$ the expected (undiscounted) net value of a single search is

$$(3) \quad -c + vp(\lambda, q, \theta) = -c + \lambda v \pi_0(\lambda, q, \theta) \{ \pi_0(\lambda q^\theta, q, \theta) \}^{-1}.$$

It follows from Lemma 1 and Corollary 1 that this quantity decreases almost surely as the system evolves, at a rate which depends upon θ and the proportion of successes to date. In Theorem 1, let (s, f) denote a state in which $s + f$ searches have been conducted of which s have been successes and f have been failures.

Theorem 1. In the stopping problem, if $\pi = \pi(\lambda, p, \theta)$ then (s, f) is a continuation point if and only if

$$(4) \quad -c + \lambda q^{(s+\theta+f)} v \pi_0 \{ \lambda q^{(s+\theta+f)}, q, \theta \} (\pi_0 [\lambda q^{(s+1)+\theta+f}, q, \theta])^{-1} > 0.$$

Note that if we have equality in (4) then both stopping and continuing to search are optimal at (s, f) . Note also that the optimal stopping rule is a 'one-step look-ahead' or 'myopic' rule. We see from Theorem 1 that the key statistic determining optimal decisions is $s\theta + f$.

Note further that, following Benkherouf et al. (1992), the one-step value in (3) divided by $(1 - \alpha)$ may be thought of as a Gittins index. In this way we may construct optimal policies for choosing how to search in K different areas each of which is modelled as above, where there are no between-area dependencies. In Theorem 2, $i = 1, 2, \dots, K$ denotes area.

Theorem 2. In the K -area problem, if $\pi^i = \pi(\lambda_i, q_i, \theta_i)$ and (s_i, f_i) is the record of past searches of area i , then search number $1 + \sum_{i=1}^K (s_i + f_i)$ will be in whichever area has maximal value of the index

$$-c_i + \lambda_i q_i^{(s_i+\theta_i+f_i)} v_i \pi_0^i \{ \lambda_i q_i^{(s_i+\theta_i+f_i)}, q_i, \theta_i \} (\pi_0^i [\lambda_i q_i^{(s_i+1)+\theta_i+f_i}, q_i, \theta_i])^{-1},$$

provided this is positive. If all indices are negative, it is optimal to stop searching altogether.

References

- BATHER, J. A. (1992) Search models. *J. Appl. Prob.* **29**, 605–615.
 BENKHEROUF, L. AND BATHER, J. A. (1988) Oil exploration: sequential decisions in the face of uncertainty. *J. Appl. Prob.* **25**, 529–543.
 BENKHEROUF, L., GLAZEBROOK, K. D. AND OWEN, R. W. (1992) Gittins indices and oil exploration. *J. R. Statist. Soc.* **B54**, 229–241.
 KEMP, A. W. (1992) Heine–Euler extensions of the Poisson distribution. *Commun. Statist. Theory Meth.* **21**, 571–588.