Biometrika (1989), 76, 1, *p.* 183
Printed in Great Britain

A candidate's formula: A curious result in Bayesian prediction

BY JULIAN BESAG

Department of Mathematical Sciences, University of Durham, Durham DH1 3LE, U.K.

SUMMARY

This note provides a simple formula for the Bayesian predictive density of future observations.

Some key words: Bayesian prediction; Predictive density.

Let $y=(y_1, \ldots, y_n)$ denote a random sample from a population having probability density function $f(.|\theta)$, where θ is an unknown, possibly vector, parameter. Let $\pi_0(\theta)$ denote a prior density for θ, so that the corresponding posterior density is

$$\pi_n(\theta|y) \propto \pi_0(\theta) f(y_1|\theta) \ldots f(y_n|\theta). \tag{1}$$

Then the predictive density $p_n(y^*|y)$ of a future independent observation from $f(.|\theta)$ is usually calculated from

$$p_n(y^*|y) = \int f(y^*|\theta)\pi_n(\theta|y)\, d\theta;$$

see, for example, Aitchison & Dunsmore (1975, p. 24).

The purpose of the present note is to point out that expanding the joint density of y^* and θ, given y, in two different ways yields the surprising alternative formula:

$$p_n(y^*|y) = f(y^*|\theta)\pi_n(\theta|y)/\pi_{n+1}(\theta|y, y^*), \tag{2}$$

where $\pi_{n+1}(\theta|y, y^*)$ is the posterior density of θ, with y augmented by an additional observation y^*. In examples, $\pi_0(\theta)$ is often chosen to be conjugate to $f(.|\theta)$, in which case the normalizing constant in (1) is immediate, as is $\pi_{n+1}(\theta|y, y^*)$, and (2) gives $p_n(y^*|y)$ directly, without any need for integration.

The result (2) is trivially amended for the prediction of several future observations and is easily extended to cater for time series in which the density of y^* given y and θ depends additionally on some of the previous y_i's. Note also that the terms on the right-hand side of (2) need only be known for a single value of θ: this has prompted the interesting suggestion from a referee that the formula is also potentially useful when the posterior distribution of θ is known only approximately but where certain values of θ have special significance (Tierney & Kadane, 1986).

Equation (2) appeared without explanation in a Durham University undergraduate final examination script of 1984. Regrettably, the student's name is no longer known to me.

REFERENCE

AITCHISON, J. & DUNSMORE, I. R. (1975). *Statistical Prediction Analysis*. Cambridge University Press.
TIERNEY, L. & KADANE, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *J. Am. Statist. Assoc.* **81**, 82–6.

[*Received April* 1988. *Revised October* 1988]