BAYESIAN ANALYSIS USING MONTE CARLO INTEGRATION WITH AN EXAMPLE OF THE ANALYSIS OF SURVIVAL DATA

By:
Leland Stewart

October 1984

Advanced Systems Studies
Lockheed Palo Alto Research Laboratory
Department 92-20, Building 205
3251 Hanover Street
Palo Alto, California 94304
(415) 424-2710

SUMMARY

Bayesian analysis using Monte Carlo integration is an effective approach for handling rich multiparameter families of distributions; nonconjugate priors; censored data; extrapolation uncertainty; and the computation of posterior distributions for parameters or predictions of interest. In the example, posterior percentile curves for a rate function are computed from survival data.

Keywords: Bayesian statistics; Monte Carlo integration; posterior distributions; censored data; survival analysis.

1. INTRODUCTION

Bayesian statistics and numerical methods form a particularly effective partnership. There are at least two reasons for this:

- o It is often easy to compute the values of the likelihood function and the prior density (and therefore the posterior density up to a multiplicative constant) on any specified set of points in the parameter space. This is true even for many models and problems that would be very difficult to handle with conventional methods.
- o The principal computational problem in making a Bayesian inference or decision is the evaluation of integrals involving the posterior density.

Thus the solution to many difficult inference and decision problems can be obtained by numerical integration using only the values of the likelihood function and the prior density on a finite set of parameter points. The suitable choice of this set is the most important, and sometimes the most difficult, aspect of the computation.

2. BAYESIAN ANALYSIS USING MONTE CARLO INTEGRATION

Bayesian inference and decision making usually require the computation of posterior cumulative distribution functions and/or posterior moments of <u>functions</u> of the parameter vector

$$\theta = (\theta_1, \theta_2, \ldots, \theta_K)' \in \Theta$$

This requires evaluation of integrals of the form

$$I = \int_{\Theta} h(\theta) \; \xi(\theta|D) \, d\theta \tag{1}$$

where h is a real valued function. The posterior density $\xi(\theta|D)$ is given by Bayes' theorem,

$$\xi(\theta|D) \propto \xi(\theta) \; L(\theta|D)$$

where $\xi(\theta)$ is the prior density and $L(\theta|D)$ is the likelihood function corresponding to the observed data D.

The Monte Carlo approximation to I is given by

$$\hat{I} = \sum_{m=1}^{M} h(\theta_m) \, W(\theta_m) \Bigg/ \sum_{m=1}^{M} W(\theta_m) \tag{2}$$

The "weights" $W(\theta_m)$, m = 1, 2, ..., M, are computed from

$$W(\theta) = \xi(\theta) \; L(\theta|D)/g^*(\theta) \tag{3}$$

The θ_m, m = 1, 2, ..., M, are generated independently from the K-variate density $g^*(\theta)$. This is known as "importance sampling." The density $g^*(\theta)$ is called the "importance function" or "generating density" and is usually chosen to approximate the posterior density, with the restriction that its form must be such that the θ_m's can be easily generated.

Details about the techniques used in applying this methodology and discussions of Monte Carlo integration error and its reduction can be found in Stewart 1968, 1970, 1977, 1979, 1983; Stewart and Johnson 1971, 1972; Johnson and Stewart 1971; Kloek and van Dijk 1975, 1978; van Dijk and Kloek 1978, 1980, 1983; McGhee and Walford 1965, 1967, 1968; Heiberger 1976; Zellner and Rossi 1982.

3. EXAMPLE

Consider data of the type shown in Fig. 1. Subjects enter the study at different ages and remain in the study until death, withdrawal or the termination of the study. (There may be some subjects who would have been in the study but died before their entry date. Since these occurrences are not included in the data, the term "left truncation" is sometimes used with this type of data.)

Let the subscript i denote the i_{th} subject and

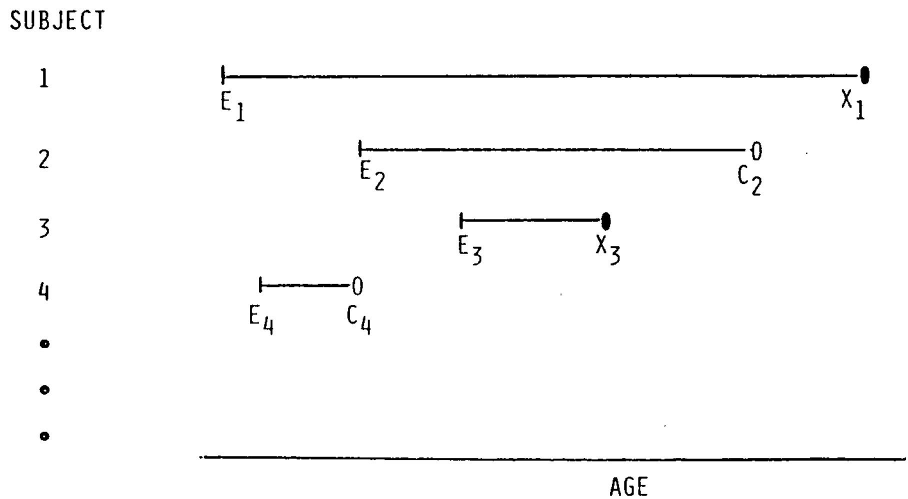
- E_i -age at entering the study
- X_i - age at death
- C_i - age at exit from study due to withdrawal or termination of study.

For each subject we observe E_i and min (X_i, C_i)

We define a rate function r(x) by:

$$r(x)\ dx \quad = \quad P\left[x<X_i<x+dx \,\middle|\, E_i \le x < X_i,\ x+dx \le C_i\right] \tag{4}$$

SURVIVAL DATA



I - E_i - AGE AT ENTRY

● - X_i - AGE AT DEATH

0 - C_i - AGE AT EXIT FROM STUDY DUE TO WITHDRAWAL OR TERMINATION OF STUDY

Fig. 1

and assume that $r(x)$ does not depend on i or E_i. (We also assume that the mechanisms that determine E_i and C_i are non-informative for inferences about $r(x)$.)

$r(x)$ is the hazard rate (or force of mortality) function corresponding to a density $f(x)$ and cumulative distribution function, $F(x)$. They are related by:

$$r(x) = \frac{f(x)}{1-F(x)} \quad (5)$$

The conditional density of X given E and that $C = \infty$ is

$$f(x|E;\ c=\infty) = \frac{f(x)}{1-F(E)} \quad x > E \quad (6)$$

The likelihood function given the data D is

$$L(\theta|D) = \prod_{i:X_i<C_i} \left[\frac{f(X_i|\theta)}{1-F(E_i|\theta)}\right] \prod_{j:X_j>C_j} \left[\frac{1-F(C_j|\theta)}{1-F(E_j|\theta)}\right]$$

$$= \frac{\prod_{i:X_i<C_i} f(X_i|\theta) \prod_{j:X_j>C_j} [1-F(C_j|\theta)]}{\prod_{\text{all } i} [1 - F(E_i|\theta)]} \quad (7)$$

The data used in this example were from a survival study of the male residents of Channing House, a retirement center located in Palo Alto. The data and further information are given in Hyde 1977, 1980. Of the total of 97 men in the study, 46 died, 5 withdrew and 46 survived to the end of the study. The ages at entry varied from 62 to 89 years.

The prior distribution and the eight parameter (K=8) state space of rate functions were defined in a manner similar to Example 1 of Stewart 1979 but with $K^* = 16$ as defined in Section 3.1. An analyst with experience derived from similar situations may have chosen a prior with a different structure, but this should cause no problems because of the great flexibility allowed in choosing priors when using Monte Carlo integration.

Posterior percentile curves for the values of the rate functon, $r(x)$, are shown in Fig. 2. Also shown are two dashed curves, designated by M and X, which are, respectively, the posterior mean and the $r(x)$ that maximizes $\xi(\theta|D)$. To obtain these curves, Monte Carlo integration was used to compute the posterior distribution of the value of the function at each of the 15 evenly spaced values of X. Each posterior was evaluated at 250 points along the vertical axis. This required the simultaneous Monte Carlo evaluation of 3765 integrals. The same set $\{\theta_m, W(\theta_m):m=1,2,\ldots,M=2000\}$ was used for all of the integrals.

Displaying a sample of rate functions from the posterior is another way of illustrating the properties of the posterior. Twenty rate functions generated from the posterior are displayed in Fig. 3. The sample was generated by the "acceptance-rejection" method described in Stewart 1983.

4. ACKNOWLEDGEMENT

Work on Bayesian analysis using Monte Carlo integration has been supported by the Lockheed Independent Research Program.

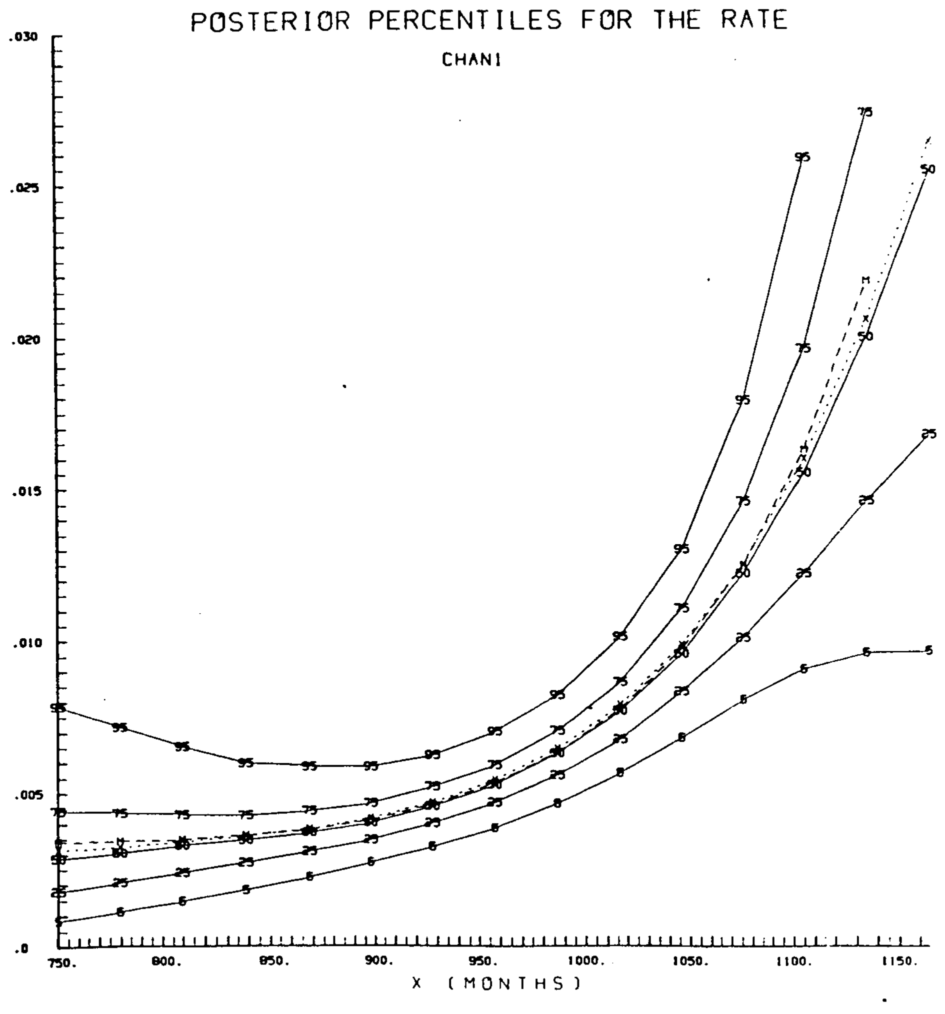

POSTERIOR PERCENTILES FOR THE RATE
CHANI
.030
.025
.020
.015
.010
.005
.0
750.
800.
850.
900.
950.
1000.
1050.
1100.
1150.
X (MONTHS)

Fig. 2

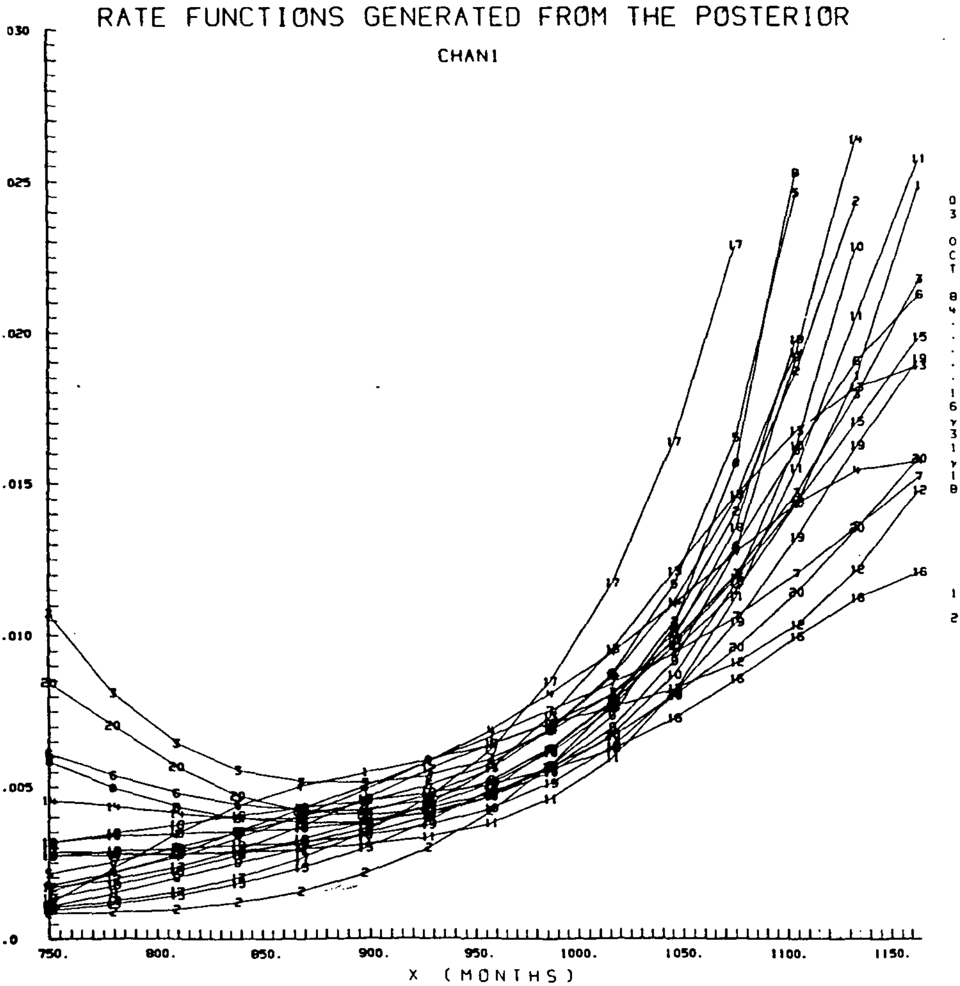

RATE FUNCTIONS GENERATED FROM THE POSTERIOR
CHANI
030
025
.020
.015
.010
.005
.0
750.
800.
850.
900.
950.
1000.
1050.
1100.
1150.
X (MONTHS)

Fig. 3

5. REFERENCES

Heiberger, R. M. (1976). "Posterior moments of scalar functions of a covariance matrix," Communications in Statistics - Simulation and Computation, Vol. B5 (2 and 3), 107-140.

Hyde, John (1977). "Testing survival under right censoring and left truncation," Biometrika, 64, 225-30.

------, (1980). "Survival analysis with incomplete observations," in Biostatistics Casebook edited by R.G. Miller, B. Efron, B.W. Brown and L.E. Moses, New York, John Wiley.

Johnson, Jerry D., and Leland Stewart (1971). "Failure prediction from interval data," Annals of Assurance Sciences - Proceedings 1971 Annual Symposium on Reliability, The Institute of Electrical and Electronics Engineers, Inc., 356-61.

Kloek, T., and H. K. van Dijk (1975). Bayesian estimates of equation system parameters - an unorthodox application of Monte Carlo, Report 7511/E, Econometric Institute, Erasmus University, Rotterdam, The Netherlands.

------, and H. K. van Dijk (1978). "Bayesian estimates of equation system parameters; an application of integration by Monte Carlo," Econometrica 46, p. 1-19. Reprinted in: A. Zellner, ed., (1980) Bayesian analysis in econometrics and statistics, essays in honor of Harold Jeffreys, North-Holland, Amsterdam, The Netherlands.

McGhee, R. B., and R. B. Walford (1965). A Bayesian solution to the reentry tracking problem, SSD-TR-65-165, USCEE Report 151, Department of Electrical Engineering, University of Southern California, Los Angeles, CA.

------, and R. B. Walford (1967). A Bayesian approach to nonlinear system identification, SSD-TR-67-126, USCEE Report 193, Department of Electrical Engineering, University of Southern California, Los Angeles, CA.

------, and R. B. Walford (1968). "A Monte Carlo approach to the evaluation of conditional expectation parameter estimates for nonlinear dynamic systems," IEEE Transactions on Automatic Control, AC-13, 29-37.

Stewart, Leland T. (1968). "Inferences about the tail of a function using Bayesian extrapolation," handout for presentation at the Annual Meeting of the American Statistical Association, LMSC-D630093, Lockheed Palo Alto Research Laboratories, Palo Alto, CA.

------, (1970). "Bayesian analysis of multiparameter univariate problems using Monte Carlo integration, "handout for presentation at the Conference on Bayesian Inference, Oregon State University, LMSC-D876830, Lockheed Palo Alto Research Laboratories, Palo Alto, CA.

------, (1977). Bayesian analysis using Monte Carlo integration, LMSC-D560641, Lockheed Palo Alto Research Laboratories, Palo Alto, CA.

------, (1979). "Multiparameter univariate Bayesian analysis," Journal of the American Statistical Association, Vol. 74, 684-93.

------, (1983). "Bayesian analysis using Monte Carlo integration - a powerful methodology for handling some difficult problems," The Statistician, Vol. 32, 195-200.

------, and Jerry D. Johnson (1971). "An example of the use of Monte Carlo integration in Bayesian decision problems," Annals of Reliability and Maintainability - 1971, 10, New York: The American Society of Mechanical Engineers, 19-23.

------, and Jerry D. Johnson (1972). "Determining optimum burn-in and replacement times using Bayesian decision theory," IEEE Transactions on Reliability, R-21, 170-175.

van Dijk, H. K., and T. Kloek (1978). Posterior analysis of Klein's Model I, Report 7824/E, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, Rotterdam, The Netherlands.

------, and T. Kloek (1980). "Further experience in Bayesian analysis using Monte Carlo integration," Journal of Econometrics, Vol. 14, 307-328.

------, and T. Kloek (1983). "Monte Carlo analysis of skew posterior distributions: an illustrative econometric example," The Statistician, Vol. 32, 216-223.

Zellner, A., and P. E. Rossi (1982). "Bayesian analysis of dichotomous qualitative response models," H. G. B. Alexander Research Foundation paper, Graduate School of Business, University of Chicago, presented at the 24th meething of NSF-NBER Seminar on Bayesian Inference in Econometrics, SUNY-Albany, May 1982.