ECONOMETRIC INSTITUTE

TWO ALGORITHMS FOR THE COMPUTATION OF POSTERIOR MOMENTS AND DENSITIES USING MONTE CARLO INTEGRATION

J.P. HOP AND H.K. VAN DIJK

REPORT 9174/A

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM-THE NETHERLANDS

SISAM AND MIXIN:
TWO ALGORITHMS FOR THE COMPUTATION OF POSTERIOR MOMENTS AND DENSITIES USING MONTE CARLO INTEGRATION[1]

by

J. Peter Hop and Herman K. van Dijk

Abstract

Two algorithms, and corresponding FORTRAN computer programs, for the computation of posterior moments and densities using the principle of importance sampling are described in detail. The first algorithm makes use of a multivariate Student t importance function as approximation of the posterior. It can be applied when the integrand is moderately skew. The second algorithm makes use of a decomposition: a multivariate normal importance function is used to generate directions (lines) and one dimensional classical quadrature is used to evaluate the integrals defined on the generated lines. The second algorithm can be used in cases where the integrand is possibly very skew in any direction.

Contents

E-MAIL: hop@cvx.eur.nl or hkvdijk@ect.few.eur.nl

---

[1] We thank Luc Bauwens, Frank Kleibergen and Suzanne Bijl for helpful comments. This paper was presented as an invited paper at the 48th session of the International Statistical Institute at Cairo, September 1991 and will appear in de Journal: Computer Science in Economics and Management. A more extensive version of this paper, Hop and Van Dijk (1990), with computer programs is available as report 9031/A of the Econometric Institute.

1. INTRODUCTION

DEFINITION OF INTEGRALS; EXAMPLES.

The multivariate integrals that we consider may be briefly described as follows. Let θ be an ℓ-vector of parameters of interest and let $g(\theta)$ be a function of θ that satisfies certain mathematical regularity properties (so that the integrals to be computed exist). The posterior mean of $g(\theta)$ is defined as

$$Eg(\theta) = \frac{\int g(\theta)p(\theta|I)d\theta}{\int p(\theta|I)\ d\theta} \tag{1.1}$$

where $p(\theta|I)$ is a kernel of a posterior density function, given the information set I. The term kernel means that $p(\theta|I)$ is proportional but not equal to a density function. The denominator of (1.1) plays the role of integrating constant similar to the role of $\sqrt{2\pi}$ in the case of the normal distribution and the role of $\Gamma(\alpha)$ in the case of the Gamma distribution with parameter α. Further, the information set I contains the data and some a priori conditioning assumptions with respect to the stochastic process that is defined by the posterior kernel $p(\theta|I)$. For convenience, we make use of the brief notation $p(\theta)$ instead of the more cumbersome notation $p(\theta|I)$.

As mentioned above, we assume that the integrals to be computed exist, that is, the function $g(\theta)p(\theta)$ is integrable. Simple examples of $g(\theta)$ are $g(\theta) = \theta$ and $g(\theta) = \theta\theta'$. Note that g may be a vector or a matrix of functions of θ. We emphasize that g may be a nonlinear function of θ. Two examples of nonlinear functions are the implied reduced form parameters of a system of simultaneous equations and the roots of the characteristic polynomial of a system of linear difference equations.

There are several cases where the evaluation of the integrals defined in (1.1) is a nontrivial matter. First, the integrating constant (the denominator of (1.1)) is not known in terms of elementary functions (e.g., $\sqrt{2\pi}$). Second, even when the integrating constant is known, the functional form of $g(\theta)$ is such that $Eg(\theta)$ is not known in terms of elementary functions. An example of such a $g(\theta)$ is the case of the roots of a characteristic polynomial of a linear difference equation. For a case study where the integrating constant is not known **and** $g(\theta)$ is a nonlinear function of θ we refer to Van Dijk and Kloek (1980).

MONTE CARLO NUMERICAL INTEGRATION.

For the cases mentioned above one may evaluate (1.1) by means of numerical integration methods that are known as Monte Carlo integration methods, henceforth referred to as MC. MC integration methods make use of the following two properties:

(1) *Generating a large sample of pseudo random numbers is very easy using a computer procedure.*

There exist physical devices that generate pseudo random numbers. However, the use of MC integration methods involves usually a computer procedure for the generation of these random numbers. Pseudo random numbers are generated on a computer by means of a deterministic method. Thus, a sequence of pseudo random numbers is perfectly reproducible. This explains the use of the term **pseudo** random number. For convenience, we delete, henceforth, this qualification.

(2) *An integral may be interpreted as the expectation of a random variable. This expected value is estimated using generated random numbers.*

The accuracy of the estimation procedure is measured using standard results from large sample theory. Alternatively, one may interpret the estimation problem as a Bayesian estimation problem using a uniform prior with respect to the value of the integral and a large sample of random drawings. For convenience, we make use of standard large sample results.

EXAMPLE.

Consider (1.1). Let $\ell = 1$ and $p(\theta|I) = 1$ on the interval $[0, 1]$ and $p(\theta|I) = 0$ elsewhere. The problem is to evaluate $\int g(\theta)d\theta$ on $[0, 1]$. Interpret θ as a random variable with a uniform distribution on $[0, 1]$. Use a computer procedure and generate a sample of uniformly distributed random numbers $\theta^{(1)}, \ldots, \theta^{(N)}$. Given that $g(\theta)$ has a mathematical regularity property, $g(\theta^{(1)}), \ldots, g(\theta^{(N)})$ is also a sample of random numbers and one can make use of the following approximation

$$\int g(\theta)d\theta \approx \frac{1}{N} \sum_{i=1}^{N} g(\theta^{(i)})$$

The sample mean is an estimator based on N pseudo random drawings. In the next section we discuss the case where one can generate from the posterior distribution with density proportional to $p(\theta|I)$ and in the remainder of this report we discuss the case where one cannot draw from the posterior distribution.

2. DIRECT SIMULATION

SAMPLING FROM POSTERIOR.

Consider the following problem:

Compute the integrals in equation (1.1) *by means of MC for the case where a computer procedure is available that enables one to generate a sample of random drawings from a distribution function* $F(\theta)$ *with a density function that is equal or proportional to the posterior kernel* $p(\theta)$.

Let $\theta^{(1)}, \ldots, \theta^{(N)}$ denote the generated random sample. Given that $g(\theta)$ has a certain regularity property (i.e., g is measurable), it follows that $g(\theta^{(1)}), \ldots, g(\theta^{(N)})$ is also a random sample. Then we may approximate the posterior mean (1.1) by the sample mean $\bar{g}$, which is given as

$$\bar{g} = \frac{1}{N} \sum_{i=1}^{N} g(\theta^{(i)}) \tag{2.1}$$

The computation of $Eg(\theta)$ by means of this procedure is referred to in the literature as **direct simulation** since one is able to **simulate** a random sample **directly** from the distribution studied, using a computer procedure.

[For more examples on direct simulation and for references on computer procedures that generate sequences of random numbers for many families of distributions we refer to, e.g., Hammersley and Handscomb (1964, Chapter 3). Some recent references are Kinderman and Ramage (1976), Kinderman and Monahan (1980), Rubinstein (1981, Chapter 3) and Marsaglia (1984)].

FLOW DIAGRAM.

A flow diagram for direct simulation is given in Figure 1. Note that, in accordance with standard usage in Fortran, we make use of an arrow sign instead of an equality sign. For instance, one interprets $S^{(0)} \leftarrow 0$ as: the value zero is assigned to the variable $S^{(0)}$.

$$S^{(0)} \leftarrow 0 \tag{2.2}$$

$$S^{(i)} \leftarrow S^{(i-1)} + g(\theta^{(i)}) \quad (i = 1, \ldots, N) \tag{2.3}$$

$$\bar{g} \leftarrow \frac{S^{(N)}}{N} \tag{2.4}$$

The symbol $S^{(0)}$ stands for the initial zero value of the sum of the sequence of random variables $g(\theta^{(1)}), \ldots, g(\theta^{(N)})$; $S^{(i)}$ denotes the i-th partial sum of this sequence, given as $S^{(i)} = g(\theta^{(1)}) + \ldots + g(\theta^{(i)})$. Equation (2.3) illustrates that one does not have to store the large set of random numbers $g(\theta^{(1)}), \ldots, g(\theta^{(N)})$ as suggested by the i-th partial sum, but one can make repeatedly use of a computer procedure that generates a random number.

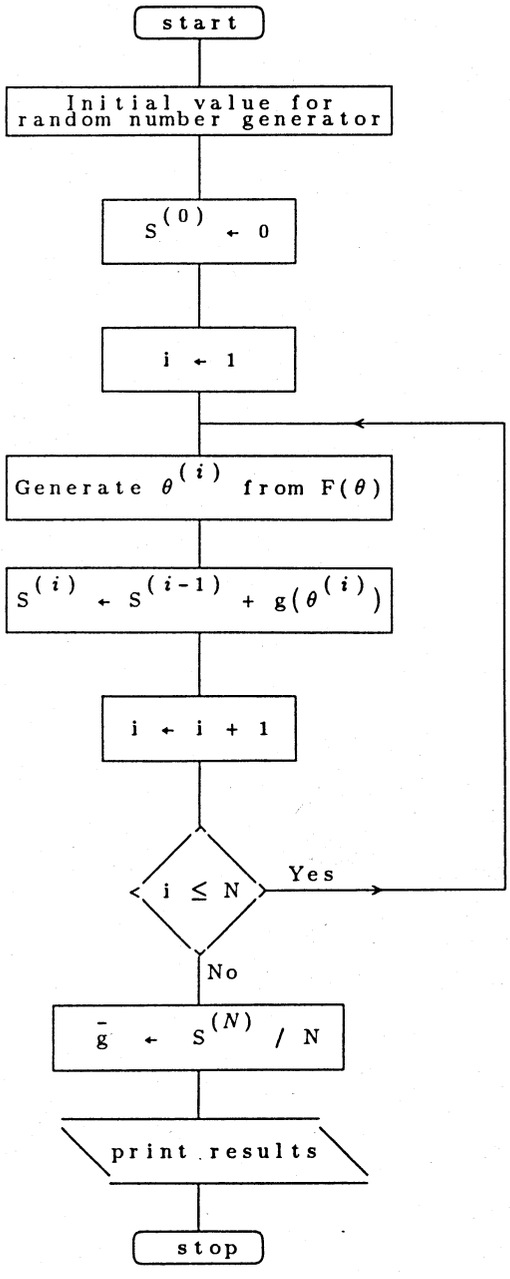


Figure 1. Flow diagram for direct simulation.

STATISTICAL MEASURES OF NUMERICAL ACCURACY.

Convergence with probability one. The accuracy of the approximation (2.1) or (2.4) may be studied by increasing the size of the sample from N to $2N$, $3N$, ... ,MN. The results can be printed at each value of jN, with $j = 1, \ldots, M$. Given certain regularity conditions (e.g., the integral $Eg(\theta)$ should exist), the Monte Carlo estimator $\bar{g}$ converges with probability one to $Eg(\theta)$. Therefore the estimator is also consistent. Proofs can be found in the statistical literature.

Recursive update of $\bar{g}$. One may study the convergence of $\bar{g}$ to $Eg(\theta)$ using a recursive method for the computation of $Eg(\theta)$. Let $\bar{g}_j$ denote the estimated value of the integral $Eg(\theta)$ after j samples of size N. The values $\bar{g}_j$, $j = 1, \ldots, M$, are related in a recursive way, i.e., we can write

$$\bar{g}_0 = 0 \tag{2.5}$$

$$j\bar{g}_j = (j - 1)\bar{g}_{j-1} + \Delta\bar{S}_{j,N} \quad (j = 1, \ldots, M) \tag{2.6}$$

where

$$\Delta\bar{S}_{j,N} = \frac{S^{(jN)} - S^{(j-1)N}}{N} \tag{2.7}$$

and $S^{(jN)}$ is defined in equations (2.2) and (2.3). Note that for $j = 1$, it follows that (2.7) is equal to (2.4). Using (2.5) and (2.6), it is seen that $\bar{g}_j$ can be computed in a recursive way as

$$\bar{g}_j = \left[\frac{j - 1}{j}\right]\bar{g}_{j-1} + \frac{\Delta S_{j,N}}{j} \quad (j = 1, \ldots, M) \tag{2.8}$$

Equation (2.8) has the structure of a linear difference equation in $\bar{g}$ with coefficients $(j - 1) / j$ and and $\Delta S_{j,N} / j$. This indicates that $\bar{g}_j$ tends to a stationary value when j tends to infinity since $((j - 1) / j) < 1$. However, $\Delta\bar{S}_{j,N} / j$ is not a constant but the realized value of a random variable. Even at a large value of j the additional term $\Delta\bar{S}_{j,N} / j$ may throw the sequence temporarily off-track, but it will be less and less probable that this event has an effect as j increases. Some examples of sequences of $\bar{g}_j$ are given in Van Dijk (1987).

Asymptotic normality; Absolute numerical error ($\bar{\sigma}/\sqrt{N}$); Relative numerical error ($1/\sqrt{N}$). An other measure of numerical accuracy can also be derived from large sample theory. Under certain regularity conditions it follows from central limit theory that the estimator $\bar{g}$, equation (2.4), is

approximately **normally** distributed with mean $Eg(\theta)$ and variance σ^2/N, where σ^2 is the variance of $g(\theta)$ and is given as

$$\sigma^2 = Eg^2(\theta) - [Eg(\theta)]^2 \tag{2.9}$$

The normal approximation tends to become more accurate when the sample size N tends to become larger. Under the assumption that the integrals in (2.9) exist, one may estimate σ^2 by the sample variance

$$\bar{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} g^2(\theta^{(i)}) - \bar{g}^2 \tag{2.10}$$

Given an estimator $\bar{\sigma}^2$ for σ^2, one can define a 95 percent confidence interval for $Eg(\theta)$ in the usual way as $[\bar{g} - 1.96\bar{\sigma}/\sqrt{N}, \bar{g} + 1.96\bar{\sigma}/\sqrt{N}]$. A value of $\bar{\sigma}/\sqrt{N}$ will be defined as an **absolute numerical error** and a value of $\bar{\sigma}/(\sigma\sqrt{N})$ may be interpreted as a **relative numerical error**, *i.e.*, the numerical error $\bar{\sigma}/\sqrt{N}$ scaled with the standard deviation of random variable $g(\theta)$. Note that in this case σ is the posterior standard deviation of the function $g(\theta)$. Clearly, the estimated relative error is equal to $1/\sqrt{N}$. [For an introduction to the sampling theory results that we use, we refer to Mood, Graybill and Boes (1974, Chapters 2 and 6) and for a more advanced treatment we refer to Cramér (1946, Chapters 25 and 27), and Billingsley (1979).]

As an exercise one can determine the required size of the sample for a pre-assigned level of numerical accuracy. Suppose one is interested in estimating a probability, i.e., $Eg(\theta)$ is a number P in the interval $[0, 1]$. Suppose further that one is satisfied with a two digit accuracy at a 95 per cent confidence level. That is, the required confidence interval bounds are given as $P - .005$ and $P + .005$ and we impose $1.96\sigma_P/\sqrt{N} \leq .005$. This implies that N must be greater or equal to $153{,}664\sigma_P^2$.

3. SIMPLE IMPORTANCE SAMPLING (SISAM)

3.1. Basics of Importance Sampling

THE IMPORTANCE FUNCTION AND WEIGHT FUNCTION.

Consider the following problem:

Compute the integrals in equation (1.1) *by means of MC for the case where it is not known how one can generate a sample of random drawings from a distribution function with density equal or proportional to* $p(\theta)$ *but it is known how one can generate a random sample from a distribution with a density equal (or proportional) to* $i(\theta)$, *which is an approximation of* $p(\theta)$.

The function $i(\theta)$, which is known as **importance** function, is a density function with **convenient** Monte Carlo properties that is used as an **approximation** of $p(\theta)$. One can replace $p(\theta)$ in (1.1) by $w(\theta)i(\theta)$ where the **weight** function $w(\theta)$ is defined as

$$w(\theta) = \frac{p(\theta)}{i(\theta)} \tag{3.1}$$

This yields

$$Eg(\theta) = \frac{\int g(\theta)w(\theta)i(\theta)d\theta}{\int w(\theta)i(\theta)d\theta} \tag{3.2}$$

where $i(\theta)$ is restricted to be positive on the region of integration. In this section we make use of a simple choice with respect to the class of importance functions. That is, we opt for the **multivariate Student t** class of density functions. This choice is justified when the integrand is moderately skew. We make use of the term **Simple Importance Sampling (SISAM)** in this case. For more details on the choice of an importance function and for some alternatives to simple importance sampling we refer to Van Dijk (1984, Chapter 3), Bauwens (1984), and Geweke (1989). An alternative method that can deal with integrands that are arbitrarily skew is explained in Section 4.

Next, let $\theta^{(1)}, \ldots, \theta^{(N)}$ be a random sample from a distribution with a density function equal (or proportional) to $i(\theta)$. Then, as stated in Section 2, $g(\theta^{(1)}), \ldots, g(\theta^{(N)})$ is also a sequence of independently distributed random variables with a common distribution function. Let $g(\theta^{(i)})$ be the typical i-th element of this sequence. The importance sampling estimator of the vector $Eg(\theta)$ is given as

$$\hat{g}(\theta) = \frac{\frac{1}{N}\sum_{i=1}^{N} g(\theta^{(i)})w(\theta^{(i)})}{\frac{1}{N}\sum_{i=1}^{N} w(\theta^{(i)})} \qquad (3.3)$$

This estimator may be interpreted as a weighted sample mean of the above mentioned random sample where $w(\theta^{(1)}), \ldots, w(\theta^{(N)})$ are the weights. If we define $g(\theta) = \theta_j$ and $g(\theta) = \theta_j\theta_k$ where $j,k = 1, \ldots, \ell$, one can write the importance sampling estimators for the marginal posterior first and second order moments of the elements of θ in a straightforward way.

STATISTICAL MEASURES OF NUMERICAL ACCURACY FOR A RATIO OF INTEGRALS.

The weighted sample mean is a good approximation of $Eg(\theta)$ if the sample size N is sufficiently large and the variation in the weights is bounded. In order to evaluate the numerical accuracy of (3.3) explicitly, we make use of results from large sample theory.

Convergence with probability one, recursive update. In a similar way as explained in Section 2 one may increase the size of the sample from N to $2N$, $3N, \ldots, MN$ and study the convergence in probability of $\hat{g}(\theta)$. Given certain mathematical conditions, it follows that $\hat{g}(\theta)$ converges with probability one to $Eg(\theta)$; details are omitted. The formulas for a recursive update of $\hat{g}(\theta)$ are the same as given in (2.5) - (2.8) with $\bar{g}$ replaced by $\hat{g}(\theta)$.

Asymptotic normality; Absolute numerical error $(\bar{\sigma}_{mc}/\sqrt{N})$; **Relative numerical error.** Given certain regularity conditions from central limit theory, it follows that the estimator $\hat{g}(\theta)$ is **approximately normally** distributed with mean $Eg(\theta)$ and variance σ^2_{mc}/N, where σ^2_{mc} is the variance of the ratio of the correlated random variables $g(\theta)w(\theta)$ and $w(\theta)$. [For convenience, we take $g(\theta)$ as a scalar function and make use of the index mc to indicate that σ^2_{mc}/N is the variance of the Monte Carlo estimator.] The normal approximation becomes more accurate as N becomes larger.

Since the estimator $\hat{g}(\theta)$ is the ratio of the **correlated** random variables given in the numerator and denominator of (3.3), the formula for σ^2_{mc} is more complicated than the usual definition of a variance [compare Section 2, equation (2.9)]. For the case of the j-th element of the posterior mean one may proceed as follows. Define in equation (3.3) $g(\theta) = \theta_j$ and $\hat{g}(\theta) = \hat{\theta}_j$ and define

$$\hat{t}_j = \frac{1}{N}\sum_{i=1}^{N} \theta_j^{(i)} w(\theta^{(i)}) \quad (j = 1, \ldots, \ell) \qquad (3.4)$$

$$\hat{t}_0 = \frac{1}{N}\sum_{i=1}^{N} w(\theta^{(i)}) \qquad (3.5)$$

Then (3.3) can be rewritten for the case of the j-th element of the posterior mean as

$$\hat{\theta}_j = \frac{\hat{t}_j}{\hat{t}_0}$$

Given certain regularity conditions, in particular, the condition that plim $\hat{t}_0$ exists and given that $\hat{\theta}_j$ is a continuous function of $(\hat{t}_j, \hat{t}_0)$, say $h(\hat{t}_j, \hat{t}_0)$, one may use the approximation for $\sigma^2_{mc,j}$, given as

$$\begin{aligned} \sigma^2_{mc,j} &\approx \frac{\partial h}{\partial(\hat{t}_j\ ,\ \hat{t}_0)} \operatorname{var}(\hat{t}_j\ ,\ \hat{t}_0) \frac{\partial h}{\partial(\hat{t}_j\ ,\ \hat{t}_0)'} \\ &\approx \left[\frac{\partial \hat{\theta}_j}{\partial \hat{t}_j}\ ,\ \frac{\partial \hat{\theta}_j}{\partial \hat{t}_0}\right] \begin{bmatrix} \operatorname{var}(\hat{t}_j) & \operatorname{cov}(\hat{t}_j\ ,\ \hat{t}_0) \\ \operatorname{cov}(\hat{t}_0\ ,\ \hat{t}_j) & \operatorname{var}(\hat{t}_0) \end{bmatrix} \begin{bmatrix} \frac{\partial \hat{\theta}_j}{\partial \hat{t}_j} \\ \frac{\partial \hat{\theta}_j}{\partial \hat{t}_0} \end{bmatrix} \end{aligned} \tag{3.6}$$

where the first column is the vector of first order partial derivatives of the function h with respect to $(\hat{t}_j, \hat{t}_0)'$. For a more detailed explanation of equation (3.6) we refer to Dhrymes (1971, pp. 112–114) or Geweke (1989). The evaluation of the right hand side of (3.6) yields [see Mood, Graybill and Boes (1974), page 181]

$$\sigma^2_{mc,j} \approx \frac{1}{\hat{t}_0^2} \operatorname{var}(\hat{t}_j) - 2\frac{\hat{t}_j}{\hat{t}_0^3} \operatorname{cov}(\hat{t}_j\ ,\ \hat{t}_0) + \frac{\hat{t}_j^2}{\hat{t}_0^4} \operatorname{var}(\hat{t}_0) \tag{3.7}$$

An estimator for $\sigma^2_{mc,j}$ follows directly from (3.7) once estimators for $\operatorname{var}(\hat{t}_j)$, $\operatorname{var}(\hat{t}_0)$ and $\operatorname{cov}(\hat{t}_j, \hat{t}_0)$ are determined. By making use of (3.4) and (3.5) and standard theory on sample moments [compare, e.g., Mood, Graybill and Boes (1974, Chapters 2 and 6)] one can write the Monte Carlo estimators, using importance sampling, for the moments given at the right hand side of (3.7) as

$$\hat{\operatorname{var}}(\hat{t}_j) = \frac{1}{N} \sum_{i=1}^{N} \left[\theta_j^{(i)} w\left(\theta^{(i)}\right)\right]^2 - \left(\hat{t}_j\right)^2 \tag{3.8}$$

$$\hat{\operatorname{var}}(\hat{t}_0) = \frac{1}{N} \sum_{i=1}^{N} \left[w\left(\theta^{(i)}\right)\right]^2 - \left(\hat{t}_0\right)^2 \tag{3.9}$$

$$\hat{\operatorname{cov}}(\hat{t}_j\ ,\ \hat{t}_0) = \frac{1}{N} \sum_{i=1}^{N} \theta_j^{(i)} \left[w\left(\theta^{(i)}\right)\right]^2 - \hat{t}_j \hat{t}_0 \tag{3.10}$$

Given an estimator $\hat{\sigma}^2_{mc,j}$ for $\sigma^2_{mc,j}$, one can define in the usual way a 95 percent confidence interval for $E\theta_j$. Let $B_j = 1.96\hat{\sigma}_{mc,j}/\sqrt{N}$, where 1.96 is taken from the table of the standard normal integral that is listed in most textbooks on statistics [see, e.g., Mood, Graybill, and Boes (1974, p. 522)]. Then the interval $[E\hat{\theta}_j - B_j\,,\, E\hat{\theta}_j + B_j]$ contains the value of $E\theta_j$ with a probability equal to 0.95. A value for $\hat{\sigma}_{mc,j}/\sqrt{N}$ will be defined as an **absolute numerical error** and $(\hat{\sigma}_{mc,j}/\sqrt{N})/\sigma_j$, where σ_j is the posterior standard deviation of θ_j, will be defined as a **relative numerical error**. As an example we refer to the output file of SISAM given in Section 6.3, where the relative error are given at the bottom line and the estimated posterior standard deviations are given below the estimates of the posterior means.

DIAGNOSTICS FOR WEIGHT FUNCTION

A necessary condition for numerically stable estimates of the integrals is that the variation of the weights (3.1) is bounded. Simple diagnostics are the ten drawings with largest weight and the frequency distribution of all the weights. In section 6.3 an example is presented of such a frequency distribution with interval bounds equal to powers of ten of the weights.

PARTIAL SUMS FOR MOMENTS AND ERROR ESTIMATES.

Summarizing, for importance sampling estimates of the posterior first and second order moments we have to compute the following sums:

$$\boxed{\sum_{i=1}^{N} w\left(\theta^{(i)}\right),\ \sum_{i=1}^{N} \theta_j^{(i)} w\left(\theta^{(i)}\right),\ \sum_{i=1}^{N} \theta_j^{(i)}\theta_k^{(i)} w\left(\theta^{(i)}\right)} \qquad (3.11)$$

$$(j\ ,\ k\ =\ 1,\ \ldots\ ,\ell)$$

[compare (3.3) and the text below (3.3)]. For the evaluation of numerical errors of the posterior first order moments we have to compute, in addition to (3.11), the sums:

$$\boxed{\sum_{i=1}^{N} \left[w\left(\theta^{(i)}\right)\right]^2,\ \sum_{i=1}^{N} \left[\theta_j^{(i)} w\left(\theta^{(i)}\right)\right]^2,\ \sum_{i=1}^{N} \theta_j^{(i)} \left[w\left(\theta^{(i)}\right)\right]^2} \qquad (3.12)$$

$$(j\ =\ 1,\ \ldots\ ,\ell)$$

[compare (3.8) - (3.10)].

MARGINAL POSTERIOR DENSITIES.

Univariate marginal posterior densities of θ_j, $j = 1, \ldots, \ell$, can be approximated by so-called frequency **histograms** or frequency **polygons** using MC methods. We start by defining (a_{k-1}, a_k), $k = 1, \ldots, K$, as a bounded interval for the parameter θ_j, $j = 1, \ldots, \ell$. Further, let $g(\theta)$ from (3.2) now be an indicator function defined as

$$\begin{aligned} g(\theta) &= 1 \quad \text{if} \quad a_{k-1} < \theta_j < a_k \\ &= 0 \quad \text{elsewhere} \end{aligned} \tag{3.13}$$

The probabilities $g_1, \ldots, g_k, \ldots, g_K$, where the typical element g_k is defined as $g_k = g[a_{k-1} < \theta_j < a_k]$, can be used for the construction of a frequency histogram using MC. The posterior density of θ_j evaluated at $\frac{1}{2}(a_{k-1} + a_k)$ can be approximated by $g_k/(a_k - a_{k-1})$ if the interval (a_{k-1}, a_k) is sufficiently small. This approximation can be used for the construction of a frequency polygon using importance sampling.

An importance sampling estimator for g_k may be derived as follows. Let $\theta^{(1)}, \ldots, \theta^{(N)}$ be a random sample generated from a distribution with density $i(\theta)$. Further, let $\bar{\theta}^{(h)} = \theta^{(h(i))}$, where $h = h(i)$ is generated by the following rule

$$\begin{aligned} h(0) &= 0 \\ h(i) &= h(i-1) + d(i) \quad (i = 1, \ldots, N) \end{aligned} \tag{3.14}$$

where

$$\begin{aligned} d(i) &= 1 \quad \text{if } a_{k-1} < \theta_j^{(i)} < a_k \\ &= 0 \quad \text{elsewhere} \end{aligned}$$

Finally, let N_1 be defined as $N_1 = h(N)$. Then an importance sampling estimator for g_k is given as

$$\bar{g}_k = \frac{\frac{1}{N} \sum_{h=1}^{N_1} w(\bar{\theta}^{(h)})}{\frac{1}{N} \sum_{i=1}^{N} w(\theta^{(i)})} \tag{3.15}$$

The definition of the estimator $\bar{g}_K$ is a bit tedious, but the computation of $\bar{g}_k$ is very simple. In fact, one has only to determine the particular interval

to which a weight $w(\theta^{(i)})$ belongs. This is especially simple when the interval width $a_k - a_{k-1}$ is the same for all k. Let b be the common interval width for $k = 1, \ldots, K$. Let r be a real number given as

$$r = (\theta_j^{(i)} - a_0) / b \tag{3.16}$$

That is, r is a real number in the interval $[1, K]$. Truncate r at its decimal point in order to make r an integer, defined as ir. Then it follows that $ir + 1$ is the interval to which a particular weight $w(\theta^{(i)})$ belongs.[2] So, estimates for g_k are computed by adding the weights that belong in each interval and by dividing the sum of the weights in each interval by the total sum of the weights. Details are presented in the computer program. Minor modifications of the procedure described here are necessary when the intervals have unequal width. Further, the extreme values a_0 and a_K may be equal to minus and plus infinity. (Not implemented in our programs). Finally, we note that the computation of bivariate marginal posterior densities proceeds in a similar way as the computation of the univariate marginal posterior densities.

3.2. Outline of Subroutine SISSUB

A summary of the structure of the computer program SISAM is presented in the flow diagram given in Figure 2. SISAM consists of a main program and a collection of routines. In the main program the directives for the model studied are read while the actual calculations are performed in the subroutine SISSUB. For convenience we review in this subsection the structure of SISSUB.

INITIAL VALUES.

After some declarations of variables, the square root of the scaling matrix of the importance function is calculated by means of a Householder reduction technique. Next, initial zero values are assigned to a number of arrays.

---

[2] The case where ir + 1 is exactly equal to K is not important, since it has probability measure zero.

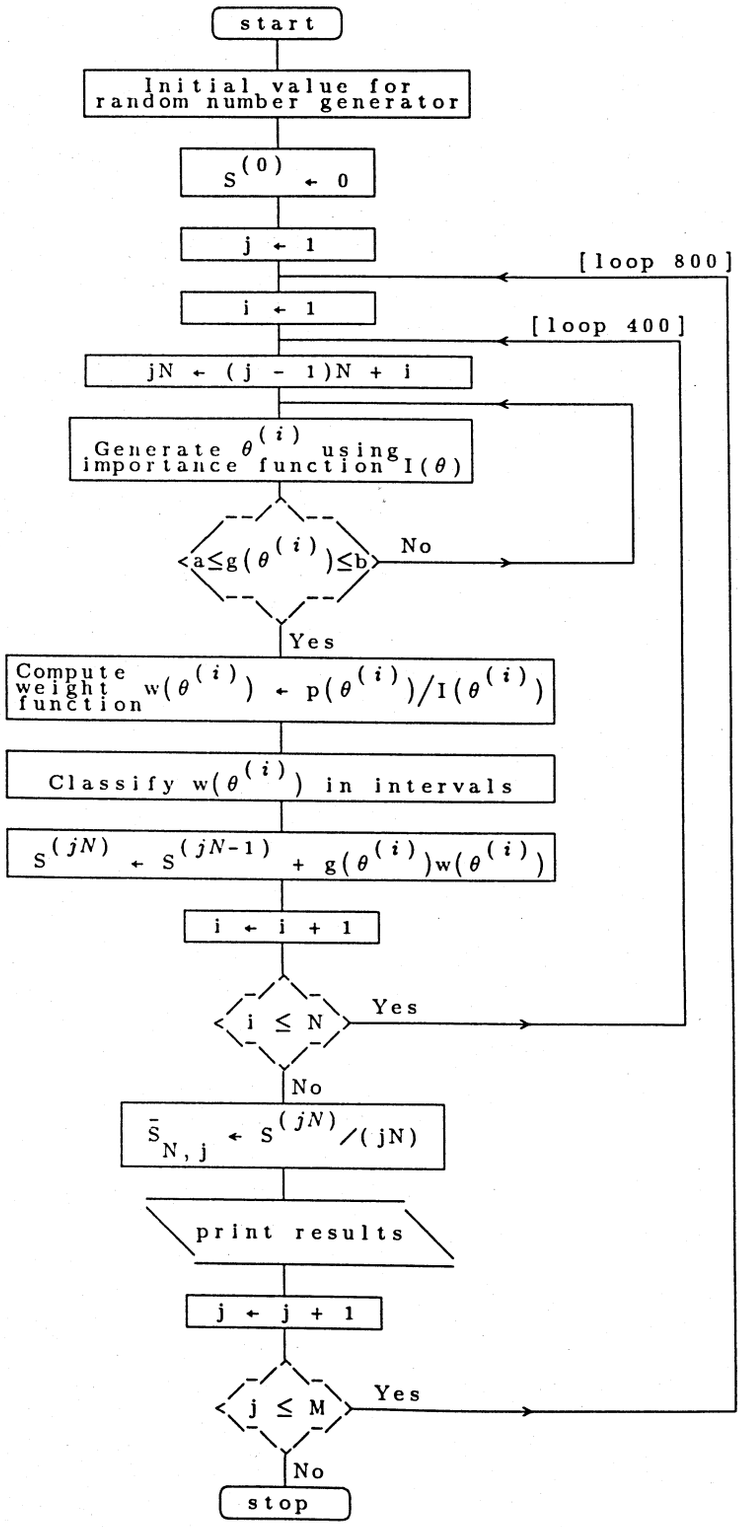


Figure 2. Flow diagram for simple importance sampling.

INTERMEDIATE AND FINAL RESULTS: LOOPS 400 AND 800.

There are two major loops in the subroutine SISSUB, see Figure 2. The first loop is the 800-loop with increment variable J. For every value of J results are send to the print output file. The second loop, within the 800-loop, is the 400-loop with increment variable I. In this loop the actual sampling from the importance distribution takes place.

GENERATION OF RANDOM DRAWING AND TESTS FOR RESTRICTIONS.

After a random vector is generated from the multivariate t importance function, it is tested whether the random drawing falls within the region of integration and whether it violates the restriction(s) in the by the user supplied subroutine RSTRCT.

WEIGHT FUNCTION AND DIAGNOSTICS.

The accepted random drawing is used to compute the weight function. The ten drawings with the largest weight are registered in order to discover whether the posterior results are heavily influenced by a few drawings with an very large weight. If a few relative large drawings occur, is this an indication that the importance function is a poor approximation to the posterior. Further, the frequency distribution of ten-powers of the weights is shown.

PARTIAL SUMS.

The weight function is used in a number of summation arrays.

POSTERIOR MOMENTS AND DENSITIES AFTER N DRAWINGS AND M ROUNDS.

After N accepted drawings the results are written to the print output file. After completing M rounds results are written to the plot output file also, until at the end of the subroutine the closing statement of the 800-loop is reached and control is returned to the main program.

NONLINEAR TRANSFORMATIONS OF PARAMETERS.

The computations in the program are done for $g(\theta) = \theta$ and $g(\theta) = \theta\theta'$. If one is interested in a nonlinear function, say $g(\theta) = 1/\theta$, a subsequent call of the subroutine SISSUB is performed, after resetting the random number generator. The program executes all calculations for a second time with $g(\theta)$ in stead of θ and $\theta\theta'$.

ROTATIONS WITH UPDATES OF THE IMPORTANCE MOMENTS.

After control is returned to the main program the input variable M is checked whether an other call of the subroutine SISSUB is requested. If so, the last results for the posterior mean and covariance matrix will be used as moments for the importance function in the next call of SISSUB. Finally, after M rotations, the mean and covariance matrix of the vector θ are written to a separate save file, if requested.

ERROR MESSAGES.

On several places in the program checks are performed to prevent the sampling process from running an extremely long CPU-time. If one of the checks is positive, an error message is printed and control is returned to the main program through a parameter in the heading of the subroutine SISSUB. The main program closes the open output files and terminates the program.

GRAPHICS GRID

The number of cells for the computation of the bivariate marginals is the same for the printer output file and the plot output file, but the user can freely modify the number of cells, if one wishes a plot output file with a finer grid.

4. MIXED INTEGRATION (MIXIN)

4.1 Basics of Mixed Integration

INTRODUCTION, MULTIVARIATE SKEWNESS.

Mixed integration, henceforth referred to as MIN, is a numerical integration method for the evaluation of multivariate integrals where the integrand is **multivariate skew**, that is, the integrand has different tail behavior in different directions. Such integrands contrast with integrands that have symmetric tail behavior, e.g., the multivariate normal density function and the multivariate Student-t density function.

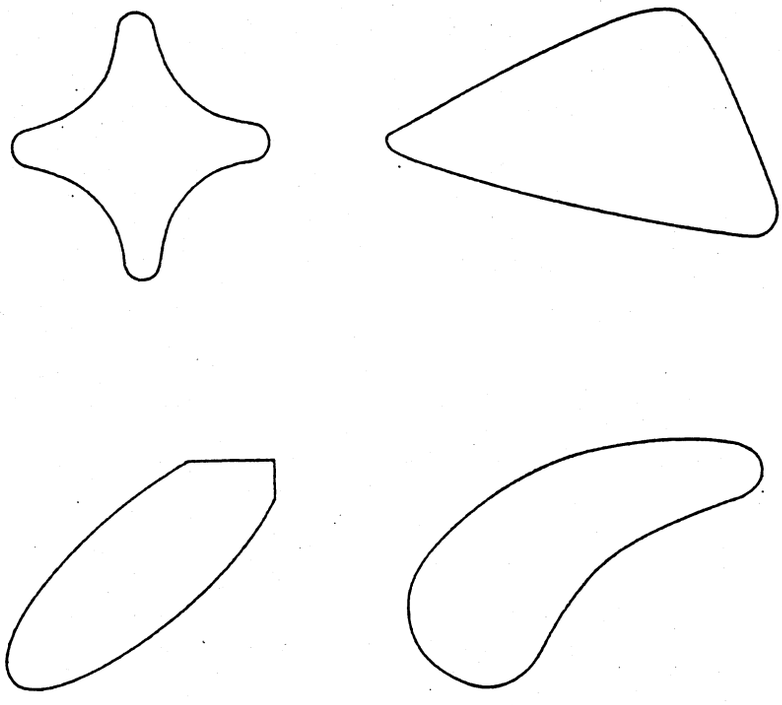


Figure 3. Examples of contours of multivariate skew functions.

TWO MAIN COMPUTATIONAL STEPS.

The distinctive feature of MIN is that it employs a mixture of one dimensional classical numerical quadrature and importance sampling. The method consists of the following two main steps:

(1) *Generate a point* $\theta^{(i)}$ *from a multivariate normal distribution that has the posterior mode* θ^0 *as its center and minus the inverse of the Hessian of the logposterior, evaluated at the posterior mode, as its covariance matrix* V. *(We assume that* θ^0 *has been estimated by a preliminary optimization procedure). A generated point* $\theta^{(i)}$ *defines a line through* $\theta^{(i)}$ *and* θ^0.

(2) *Perform one dimensional numerical integration along the line mentioned above, where the integrand is the posterior kernel multiplied by a particular factor that is specified below.*

These two steps are repeated a sufficiently large number of times.

CONDITIONAL AND MARGINAL STEP.

One may argue that the MIN-technique conditions on skewness. That is, generating lines, or, more precisely stated, generating directions, by means of a multivariate normal sampling procedure occurs in a symmetric way. This is the marginal step in the computation. Given a generated direction, one performs a one dimensional numerical integration step, which takes account of the possible skewness in the integrand. Thus, classical numerical quadrature occurs in the conditional step of the two-step integration procedure.

NOTE.

The background of mixed integration is sketched below. More details are presented in Van Dijk, Kloek, and Boender (1985) and Van Dijk (1987). The reader not interested in this background information can restrict attention to some formulas and to the structure of the computer program that is given in Section 4.2.

BACKGROUND, TRANSFORMATION OF VARIABLES.

MIN is based on a transformation of variables. Let θ^0 denote the posterior mode, which has been estimated by a numerical optimization method. The ℓ-vector of parameters of interest θ is changed into a pair (η, ρ), where the $(\ell - 1)$-vector η is defined as $\theta - \theta^0$ with the last element deleted. The vector η describes the direction of the vector $\theta - \theta^0$. The scalar ρ satisfies $\rho^2 = (\theta - \theta^0)'V^{-1}(\theta - \theta^0)$. It describes the distance between θ and the posterior mode θ^0 in a metric that makes use of a covariance matrix V [Where V is given as minus the inverse of the Hessian of the natural logarithm of the posterior density evaluated at the posterior mode]. A sign convention for ρ is added in order to guarantee that the transformation of θ into (η, ρ) is one-to-one. The advantage of the

transformation employed is that the two computational steps sketched above can be performed.

COMPUTATION OF ZERO-TH ORDER MOMENT.

We start with rewriting θ in the following way. Let y be an auxiliary random variable, defined as

$$y := \frac{\theta - \theta^0}{\rho} \tag{4.1}$$

so that

$$\theta = \theta^0 + \rho y \tag{4.2}$$

As a next step we make use of a particular feature of mixed integration. That is, it follows from the properties of the employed transformation of variables that random drawings $\eta^{(i)}$ can be generated by generating θ drawings **from a multivariate normal distribution with mean θ^0 and covariance V** and then applying the transformation involved. For details we refer Van Dijk, Kloek, and Boender (1985) and Van Dijk (1987). In the actual computations we make use of the property that the generation of directions occurs in a symmetric way. That is, if $\theta^{(i)} - \theta^0$ is a generated point, one can take $\theta^0 - (\theta^{(i)} - \theta^0)$ as a next generated point since these two points are symmetric around θ^0. The effect of such a sampling scheme, which is called **antithetic** sampling, is that each generated point is used **twice** and that the one dimensional integrals are computed on the **entire** real line. Given a random sample $\eta^{(1)}, \ldots, \eta^{(N)}$ we estimate $\int p(\theta)d\theta$ by

$$\hat{\phi}_0 \propto \frac{1}{2N} \sum_{i=1}^{N} w_0(\eta^{(i)}) \tag{4.3}$$

where

$$w_0(\eta^{(i)}) = \int p(\theta^0 + \rho y^{(i)}) \, |\rho^{\ell-1}| \, d\rho \tag{4.4}$$

Due to the antithetic sampling method one divides by $2N$ in (4.3). Further, we make use of a proportionality constant since the random sample $\eta^{(1)}, \ldots, \eta^{(N)}$ stems from a distribution with a density where the normalizing constant is not written in an explicit way. Since we deal always with ratios of integrals these proportionality constants cancel.

COMPUTATION OF FIRST AND SECOND ORDER MOMENTS.

Since θ^0 may be interpreted as a vector of known constants, it follows from (4.2) that

$$E\theta = \theta^0 + E\rho y \tag{4.5}$$

Equation (4.5) illustrates that the term $E\rho y$ represents skewness. In the case of symmetry, when $E\theta = \theta^0$, one has the property $E\rho y = 0$. The computation of the posterior mean in case of skewness proceeds as follows. We write

$$E\theta = \theta^0 + \frac{\int \rho y p(\theta) d\theta}{\int p(\theta) d\theta} \tag{4.6}$$

A mixed integration estimator for the integral in the numerator of equation (4.6) can be derived by going through the same transformation of variables as for the case of the denominator. By making use of results that are similar to the ones given in equation (4.3) - (4.4) it follows that the vector of integrals in the numerator of (4.6) can be approximated by

$$\hat{\phi}_j \propto \frac{1}{2N} \sum_{i=1}^{N} y^{(i)} w_1\left(\eta^{(i)}\right) \quad (j = 1, \ldots, \ell) \tag{4.7}$$

where

$$w_1\left(\eta^{(i)}\right) = \int \rho p\left(\theta^0 + \rho y^{(i)}\right) |\rho^{\ell-1}| d\rho \tag{4.8}$$

Note that the integrand in (4.8) is equal to the integrand of (4.4) premultiplied by a factor ρ. A MIN-estimator for $E\theta_j$, $j = 1, \ldots, \ell$, can now be written as

$$\hat{\theta}_j = \theta_j^0 + \frac{\hat{\phi}_j}{\hat{\phi}_0} \quad (j = 1, \ldots, \ell) \tag{4.9}$$

Note that we made use of a proportionality sign in (4.3) and (4.7), but that we use an equality sign in (4.9). Since the numerator and denominator in (4.9) have been estimated using the same random sample, it follows that the numerical constants, which were omitted in (4.3) and (4.7), cancel and one can make use of an equality sign in (4.9).

For the matrix of second order moments around the mean one can make use of (4.2) and derive that

$$E(\theta - E\theta)(\theta - E\theta)' = E(\rho y - E\rho y)(\rho y - E\rho y)'$$
$$= E\rho^2 yy' - E\rho y E\rho y' \qquad (4.10)$$

The first term on the second line of (4.10) can be approximated by

$$\hat{E}\rho^2 y_j y_k = \frac{\sum_{i=1}^{N} y_j^{(i)} y_k^{(i)} w_2(\eta^{(i)})}{\sum_{i=1}^{N} w_0(\eta^{(i)})} \qquad (j,k = 1, \ldots, \ell) \qquad (4.11)$$

where

$$w_2(\eta^{(i)}) = \int \rho^2 p(\theta^0 + \rho y^{(i)}) |\rho^{\ell-1}| d\rho \qquad (4.12)$$

The integrand in (4.12) is equal to the integrand of (4.4) premultiplied by a factor ρ^2. The results, given in [(4.4), (4.8), (4.12)] indicate that for the case of the first order and second order moments of an ℓ-vector of parameters θ one has to compute three one dimensional integrals for each generated direction $\eta^{(i)}$.

NUMERICAL ERROR ESTIMATES OF POSTERIOR MEAN.

The computation of $\hat{\theta}_j$ gives rise to two types of numerical integration errors. One type of error is due to the application of a particular one dimensional classical quadrature formula and the second type of error is a sampling error due to the application of Monte Carlo integration. We shall discuss the sampling error given that the classical numerical integration is used in such a way that it's error has a much smaller effect on the value of the integrals than the sampling error.

The sampling error in the MIN-estimate of the posterior mean can be evaluated by making use of the formula for the variance $\sigma^2_{mc,j}$ of the posterior mean estimate, given in Section 3, equation (3.7). Instead of equations (3.4) and (3.5), we make use of (4.7) and (4.3) respectively, that is, we replace $\hat{t}$ in (3.7) by $\hat{\phi}$. As a next step we have to determine the variances of $\hat{\phi}_j$ and $\hat{\phi}_0$ and the covariance between $\hat{\phi}_j$ and $\hat{\phi}_0$. However, we cannot apply equations (3.8) - (3.10) directly to the case of MIN, because the estimators $\hat{\phi}_j$ and $\hat{\phi}_0$ are computed by making use of antithetic sampling.

In principle, one could evaluate $\hat{\phi}_0$ by computing one dimensional integrals separately on R^+ and R^-, using antithetic sampling. If we write $\hat{\phi}_{01}$ as

$$\hat{\phi}_{01} \propto \frac{1}{2N} \sum_{i=1}^{N} \int_{R^+} p\left(\theta^0 + \rho y^{(i)}\right) |\rho^{\ell-1}| \, d\rho \tag{4.15}$$

and $\hat{\phi}_{02}$ as the same expression over R^-, then $\hat{\phi}_0 = \hat{\phi}_{01} + \hat{\phi}_{02}$ is an estimator for $\int p(\theta) d\theta$ and

$$\text{var}(\hat{\phi}_0) = \text{var}(\hat{\phi}_{01}) + \text{var}(\hat{\phi}_{02}) + 2\text{cov}(\hat{\phi}_{01}, \hat{\phi}_{02}) \tag{4.16}$$

Since $\hat{\phi}_{01}$ and $\hat{\phi}_{02}$ are negatively correlated, it follows that one obtains a reduction in the variance of $\hat{\phi}_0$. The gain in computational efficiency depends on the size of the correlation between $\hat{\phi}_{01}$ and $\hat{\phi}_{02}$, which varies for each particular problem. However, we cannot make use of (4.16) since we compute a single one dimensional integral on the entire real line instead of two one dimensional integrals on, respectively, the positive and negative real line. The latter approach is computationally much less efficient. As a consequence of our procedure we cannot evaluate the increase in computational efficiency due to the use of antithetic sampling. We shall make use of

$$\hat{\text{var}}(\hat{\phi}_0) \propto \frac{1}{4N} \sum_{i=1}^{N} \left[w_0\left(\eta^{(i)}\right)\right]^2 - \hat{\phi}_0^2 \tag{4.17}$$

This estimator of $\text{var}(\hat{\phi}_0)$ overestimates, probably, the true variance. A similar conclusion holds for the estimators for $\text{var}(\hat{\phi}_j)$ and $\text{cov}(\hat{\phi}_j, \hat{\phi}_0)$. Details are left to the interested reader. We shall make use of the following formulas

$$\hat{\text{var}}(\hat{\phi}_j) \propto \frac{1}{4N} \sum_{i=1}^{N} \left[y^{(i)} w_1\left(\eta^{(i)}\right)\right]^2 - \hat{\phi}_j^2 \tag{4.18}$$

$$\hat{\text{cov}}(\hat{\phi}_j, \hat{\phi}_0) \propto \frac{1}{4N} \sum_{i=1}^{N} y^{(i)} w_1\left(\eta^{(i)}\right) w_0\left(\eta^{(i)}\right) - \hat{\phi}_j \hat{\phi}_0 \tag{4.19}$$

and substitute these in (3.7). As a consequence of our approach, one should interpret the reported **absolute** and **relative numerical** errors as upper bounds for the actual errors.

PARTIAL SUMS FOR POSTERIOR MOMENTS AND ERROR ESTIMATES.

Summarizing, for the MIN-estimates of the posterior first order and second order moments we have to compute the following sums:

$$\boxed{\sum_{i=1}^{N} w_0(\eta^{(i)}),\ \sum_{i=1}^{N} y_j^{(i)} w_1(\eta^{(i)}),\ \sum_{i=1}^{N} y_j^{(i)} y_k^{(i)} w_2(\eta^{(i)})} \qquad (4.20)$$

$$(j\ ,\ k = 1,\ \ldots\ ,\ell)$$

and for a rough estimate of the numerical error of the posterior mean estimate we have to compute, in addition to the sums reported in (4.20), the following sums:

$$\boxed{\sum_{i=1}^{N} [w_0(\eta^{(i)})]^2,\ \sum_{i=1}^{N} [y_j^{(i)} w_1(\eta^{(i)})]^2,\ \sum_{i=1}^{N} y_j^{(i)} w_0(\eta^{(i)}) w_1(\eta^{(i)})} \qquad (4.21)$$

$$(j = 1,\ \ldots\ ,\ell)$$

These sums are listed in the computer program. Note that the sample size is $2N$ instead of N. As a consequence, one obtains a reduction in the numerical error estimates.

COMPUTATION OF POSTERIOR DENSITIES

The computation of marginal posterior densities using mixed integration requires, in principle, the computation of K one dimensional integrals P_k, $(k = 1, \ldots, K)$, which have been defined in Section 3, equation (3.13). That is, given a generated direction $\eta^{(i)}$ one has to apply K times a one dimensional numerical integration procedure. This is very cumbersome and in most cases computationally rather expensive. We shall follow a pragmatic approach that does not require the computation of any one dimensional integral beyond the one that has been computed for the normalizing constant $[w_0(\eta)$, see equation (4.4)].

Let $w_1^*, \ldots, w_M^*$ be the weights of the Gauss-Legendre quadrature that is used as one dimensional numerical integration method. Then one can write

$$w_0(\eta^{(i)}) \approx \sum_{m=1}^{M} p(\theta^0 + \rho_m y^{(i)}) \,|\rho_m^{\ell-1}|\, w_m^* \qquad (4.22)$$

[compare, e.g., Abramowitz and Stegun (1964, Table 25.4)]. For each term on the right hand side of (4.22), we determine to which interval it belongs. [compare the comment after equation (3.15)]. Additional details are presented in the computer program. We note that the error in this short-cut procedure is not known. That is, it is not known whether the approximation of the integral in a particular interval $[a_{k-1}, a_k]$ on a generated line by means of

one or more terms given at the right hand side of (4.22) is as accurate as the value of the integral on the entire real line. Our practical experience with this short-cut procedure is very good.

BOUNDED REGIONS.

We emphasize that in our applications the region of integration is bounded. As a consequence, the line integrals with respect to ρ are computed on a bounded interval. The upper and lower bound of this interval may be determined as follows. Given that $a_j < \theta_j < b_j$, $j = 1, \ldots, \ell$, one can make use of (4.2) and write

$$a_j - \theta^0 < \rho y_j < b_j - \theta^0 \quad (j = 1, \ldots, \ell) \tag{4.23}$$

For each of the ℓ-dimensions one can compute two values of ρ such that the inequalities are binding constraints. That is, if $y_j > 0$, then we define

$$\rho_j^* = \frac{a_j - \theta_j^0}{y_j}, \quad \rho_j^{**} = \frac{b_j - \theta_j^0}{y_j} \quad (j = 1, \ldots, \ell) \tag{4.24}$$

If $y_j < 0$, then ρ_j^* and ρ_j^{**} are interchanged in (4.24). As a next step one determines the minimum value of $(\rho_1^{**}, \ldots, \rho_l^{**})$ and the maximum value of $(\rho_1^*, \ldots, \rho_l^*)$. These extreme values are the limits of integration for the line integrals. Practical details are presented in the computer program.

NONLINEAR FUNCTIONS OF PARAMETERS.

Mixed integration can also be used for the computation of moments of nonlinear functions of θ, such as implied multipliers of simultaneous equations models. The restriction of mixed integration is that for each nonlinear function one has to compute a one dimensional integral with respect to ρ, given a generated direction $\eta^{(i)}$.

4.2 Outline of Subroutine MIXSUB

A summary of the structure of the computer program MIXIN is presented in the flow diagram given in Figure 4. MIXIN consists of a main program and a collection of routines. In the main program the directives for the model studied are read while the actual calculations are performed in the subroutine MIXSUB. In the comment of MIXSUB the meaning of the formal

parameters of the subroutine is explained. For convenience we review in this subsection the structure of MIXSUB. The computational steps that are similar to the ones in the program SISAM are only briefly mentioned.

INITIAL VALUES.

These are similar to the initial values of SISSUB.

INTERMEDIATE AND FINAL RESULTS: LOOPS 400 AND 800.

With reference to figure 4 the same remarks as in the subsection for SISSUB, hold here too.

GENERATION OF RANDOM DIRECTIONS (OR LINES) AND TESTS FOR RESTRICTIONS.

Define lines by generating $\theta^{(i)}$ *from* $N(\theta^0,V)$. A point $\theta^{(i)}$ is generated from a multivariate normal distribution that has the posterior mode θ^0, or an other location estimate, as its center and minus the inverse of the Hessian matrix of the logposterior, evaluated at the mode, as its covariance matrix V. (One may start, of course, with an other scaling matrix than the Hessian.) In many cases θ^0 has been estimated by a preliminary optimization procedure.

The result of this step is as follows. A generated point $\theta^{(i)}$ **defines a line through** $\theta^{(i)}$ **and** θ^0. We note that the program tests whether the generated point $\theta^{(i)}$ is within the bounded region of integration. If this is not the case, an other point is generated. In principle, this test is not necessary for the generation of lines in the region of integration, but it appears to give better results in practice for the cases that we studied. Apparently more lines are generated in directions where the posterior has substantial probability mass.

$$\textit{Compute } y^{(i)} = \frac{\theta^{(i)} - \theta^0}{\rho}, \quad \textit{with} \quad \rho := [(\theta^{(i)} - \theta^0)'V^{-1}(\theta^{(i)} - \theta^0)]$$

The l-vector $y^{(i)}$ is an auxiliary random vector; compare the discussion in the preceding subsection. The scalar ρ measures the distance between $\theta^{(i)}$ and θ^0 in a metric that makes use of V^{-1}.

Deleting the index i, one can write $\theta = \theta^0 + \rho y$. Clearly, if one increases ρ from 0 to ∞ one moves on the line that starts in θ^0 and extends through the point $\theta^{(i)}$ generated in the earlier step. One may also use ρ on the interval from 0 to $-\infty$. The purpose of MIXSUB is to compute one dimensional integrals on the lines described above.

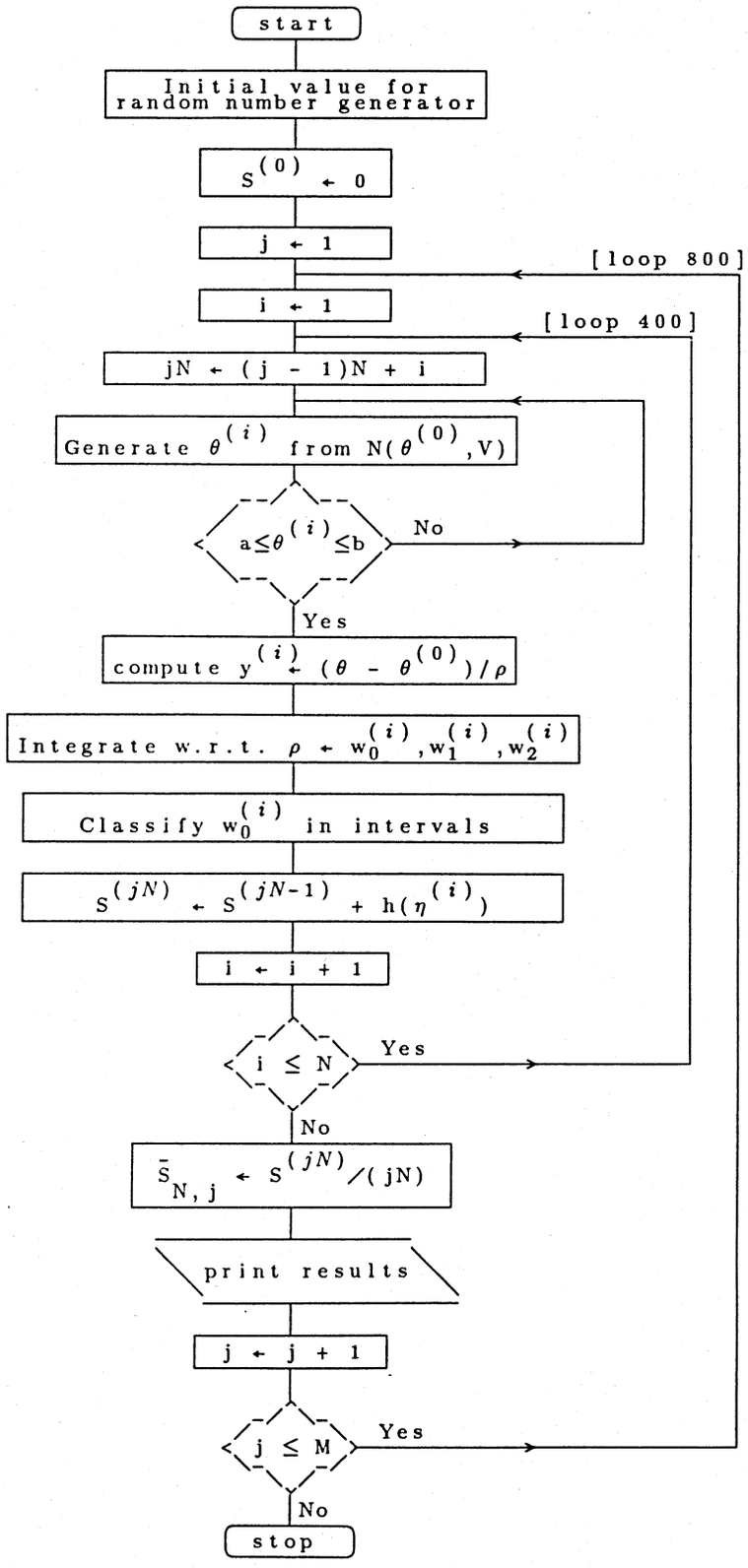


Figure 4. Flow diagram for mixed integration.

COMPUTE UPPER AND LOWER BOUNDS (a , b) OF THE LINE THROUGH θ^0 AND $\theta^{(i)}$.

This is explained in the preceding subsection.

COMPUTE ONE DIMENSIONAL LINE INTEGRALS.

As explained in the preceding subsection we are interested in the following three line integrals

$$w_k^{(i)} = \int p(\theta^{(0)} + \rho y^{(i)}) |\rho^{\ell-1}| \rho^k \, d\rho, \quad (k = 0,1,2)$$

For $k = 0$ one obtains an estimator for the integral $\int p(\theta) d\theta$. For $k = 1$ and $k = 2$ the computed line integrals can be used for the computation of the posterior first and second order moments. We note that the ten drawings with the largest integral value do not have such a clear cut interpretation as in the case of SISAM. One can suppress the evaluation of these ten drawings easily.

PARTIAL SUMS.

The partial sums for the computation of the posterior first and second order moments and the error estimates are in the program MIXIN used in a different way as in the case of SISAM. In MIXIN, one deals with three different weights, w_0, w_1, w_2, for the zero-th, first and second order moments, respectively, while in SISAM one has the same weight for all values of $g(\theta)$. For the case of MIXIN, we summarize the update of the partial sums as follows

$$S^{(jN)} = S^{(jN-1)} + h(\eta^{(i)})$$

where $h(\eta^{(i)})$ is the typical term in the three summations given in (4.20). Note that in MIXIN the integrand in the line integral is the transformed posterior density possibly multiplied by a factor in ρ. Given a certain accuracy of the one dimensional integration procedure, the approximation error in MIXIN depends on the efficient generation of lines. That is, it increases the accuracy when lines are generated where the posterior density has a substantial probability mass. In contrast, SISAM makes use of an importance function which is an approximation to the posterior density. Numerical accuracy depends on whether the importance function has its probability mass concentrated in the same region where the posterior has its probability mass.

POSTERIOR MOMENTS AND DENSITIES AFTER N DRAWINGS AND M ROUNDS.

The same remarks as for SISSUB can be made here.

ROTATIONS WITH UPDATES OF THE IMPORTANCE MOMENTS.

The same remarks as for SISSUB hold, except one has to read MIXSUB and random drawings are generated from the Normal distribution.

ERROR MESSAGES.

The error messages are the same as for SISSUB.

GRAPHICS GRID

The remarks on the graphics grid are the same as for SISSUB.

5. EXAMPLE OF AN ECONOMETRIC MODEL: JOHNSTON'S MODEL

5.1 Model and Posterior Kernel

In this subsection we briefly discuss Johnston's model [see J.J. Johnston, 1963, Econometric methods, first edition, McGraw-Hill New York]. This model is a simple example of a macro-economic expenditure model and is used for illustrative purposes only. In particular, the input, the output, and the functioning of the programs SISAM and MIXIN is illustrated.

Johnston's model consists of two stochastic equations and one accounting identity.

$$C_t = \alpha_1 + \beta_1 Y_t + u_t$$

$$I_t = \alpha_2 + \beta_2 Y_t + \gamma_2 I_{t-1} + v_t \quad (t = 1, \ldots, T) \tag{5.1}$$

$$Y_t = C_t + I_t + Z_t$$

where the variables are defined as

C_t: consumer expenditures in period t,

I_t: investment expenditures in period t,

Z_t: exogenous expenditures in period t,

Y_t: total expenditures in period t.

The parameters of interest are given as

β_1: marginal propensity to consume out of total expenditures,

β_2: marginal propensity to invest out of total expenditures,

γ_2: adjustment parameter.

The parameters α_1 and α_2 are not of particular interest (so-called nuisance parameters). The set of equations (5.1) can be written in vector-matrix notation as

$$(C\ I\ Y)\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -\beta_1 & -\beta_2 & 1 \end{bmatrix} + (\iota\ Y_{-1}\ Z)\begin{bmatrix} -\alpha_1 & -\alpha_2 & 0 \\ 0 & -\gamma_2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = (u\ \ v\ \ 0) \tag{5.2}$$

the 10×6 matrix $D := (C\ I\ Y\ \iota\ Y_{-1}\ Z)$ consists of 10 observations on the three endogenous variables C_t, I_t, and Y_t and the three predetermined variables 1, Y_{t-1} and Z_t. The equation system (5.2) is an example of the well-known linear simultaneous equation model (SEM); for details we refer to e.g., Judge et al., (1988). The rows of $(u\ v)$ are independently normally distributed with

zero mean and nonsingular covariance matrix Σ. Further, we assume the standard set of assumptions of the linear SEM. If we take the prior information from a non-informative approach in particular we take $p(\beta_1, \beta_2, \gamma_2, \Sigma) \propto |\Sigma|^{-2}$, then one can derive that the marginal posterior density of the parameters of interest can be written as

$$p(\beta_1, \beta_2, \gamma_2|D) \propto |1 - \beta_1 - \beta_2|^T |U'NU|^{-\frac{1}{2}T} \qquad (5.3)$$

with $T = 10$ annual observations. For details see Van Dijk (1984, chapter 2). We note that in this case (5.3) can also be interpreted as the concentrated likelihood function. The 2×2 matrix $U'NU$ is given as

$$U'NU = (c - \beta_1 y, \; i - \beta_2 y - \gamma_2 i_{-1})'(c - \beta_1 y, \; i - \beta_2 y - \gamma_2 i_{-1}) \qquad (5.4)$$

The data matrix (c, i, y, i_{-1}) measures variables as deviations from their means. The subroutine PSTROR which computes the natural log of the right hand side of (5.3) for specific values of $(\beta_1, \beta_2, \gamma_2)$ has to be written in FORTRAN and the declaration header is shown in the next subsection.

We end this section with a remark. The prior density of $(\beta_1, \beta_2, \gamma_2)$ is uniform on the three dimensional space, defined as the Cartesian product of the intervals [-2 , .8], [-1.7 , .25] and [-.4 , 1]. The bounds of this region have been chosen as rather wide in order to investigate the performance of the MC integration routines. Of course, large negative values of, for instance, β_1 have no economic sense.

5.2 The user supplied routines PSTROR, RSTRCT and GTHETA

THE POSTERIOR ROUTINE (PSTROR)

To keep the programs SISAM and MIXIN applicable for all kinds of models, one has to keep those parts that are typical for a particular model external to the program. In this way the programs need not be altered every time one uses them for a different model or for an other specification of the same model.

The user has to provide a subroutine called PSTROR with the same number and type of formal parameters as shown in listing 5.1 where the log-posterior kernel value, denoted by FPOST, is calculated for the case of Johnston's model. The programs call this routine twice. The first time in the main

program, with ICALL = *true*, the second time in the subroutines SISSUB and MIXSUB with ICALL = *false* for every not rejected drawing. (In the case of MIXSUB the subroutine PSTROR is actually called twice in the integration subroutine). The call from the main program is meant as an initial call in order to read data and to make other necessary preparations for the calculation of FPOST. When the subroutine is called from SISSUB and MIXSUB, the actual calculation of FPOST takes place.

We note that in the example of Johnston's model a scale-value is used (SCFAC) to prevent problems with the machine precision (overflow). The use of this scaling factor has no effect on the final results.

After execution of the routine control is returned to SISSUB or MIXSUB.

Listing 5.1

```
      SUBROUTINE PSTROR(PARAM,NDIM,FPOST,IFCNT,ICALL)
      LOGICAL ICALL
      INTEGER NDIM,IFCNT
      DOUBLE PRECISION PARAM(NDIM),FPOST
C
C
C     LOG-POSTERIOR, JOHNSTON-MODEL.
C
C     PARAM(NDIM): THE VECTOR WITH PARAMETER VALUES
C     NDIM       : THE NUMBER OF PARAMETERS
C     FPOST      : THE VALUE OF THE LOG-POSTERIOR-FUNCTION
C     IFCNT      : COUNT OF THE NUMBER OF CALLS OF THIS ROUTINE
C     ICALL      : IF "TRUE" (CALL FROM THE MAIN-PROGRAM) ONLY THE DATA ARE
C                    READ AND SOME HELP VARIABLES ARE INITIATED,
C                    IF "FALSE" (CALL FROM SUBROUTINE) CALCULATION OF "FPOST"
C                    IS DONE AND "IFCNT" IS INCREASED BY ONE
C
C     For spatial reasons we omitted the body of the routine
C
      RETURN
      END
```

THE RESTRICTION ROUTINE (RSTRCT)

The same remark as made in the first sentence of the description of the subroutine PSTROR holds for this routine also.

The routine has to be specified exactly as shown in Listing 5.2. In general the user is free to fill in the body of the routine. In our example of Johnston's model a test is performed on the Jacobian term. Specifically,

after a random drawing is generated from a multivariate Student-t distribution (SISSUB) or the multivariate Normal distribution (MIXSUB) and after it is tested whether this drawing is in the region specified by the upper and lower bounds the routine RSTRCT tests whether the Jacobian is less than 0.01 for the given values of the parameters. If the tests holds the random drawing is rejected.

After execution of this routine control is returned to SISSUB or MIXSUB.

Listing 5.2

```
      SUBROUTINE RSTRCT(PARAM,NDIM,FAIL)
      LOGICAL FAIL
      INTEGER NDIM
      DOUBLE PRECISION PARAM(NDIM)
C
C
C     JOHNSTON-MODEL, RESTRICTIONS ON PARAMETERS: TEST ON JACOBIAN.
C
C     PARAM(NDIM): THE VECTOR WITH PARAMETER-VALUES
C     NDIM       : THE NUMBER OF PARAMETERS
C     FAIL       : IF "TRUE" THE PARAMETERS MEET THE RESTRICTION(S)
C                     THE DRAWING WILL BE REJECTED,
C                     IF "FALSE" THE PARAMETERS DO NOT MEET THE RESTRICTION(S)
C                     THE DRAWING WILL BE ACCEPTED
C
C     For spatial reasons we omitted the body of the routine
C
      RETURN
      END
```

THE TRANSFORMATION ROUTINE (GTHETA)[3]

This routine is meant to give the user the freedom to calculate means, variances and correlations of some function g of the parameter vector θ. Although the use of this routine is optional (see listing 6.1 in the next section) one has to provide the declaration of the routine as in listing 5.3. In the example of the Johnston's model this options is not selected. The input vector PARAM(NDIM) is the parameter of an accepted drawing, FPOST is the value of the log-posterior kernel, WEIGHT is the value of the weight of this accepted drawing and GPARAM(MDIM) is the output vector of the

[3] In the current operational versions of MIXIN this optional routine is not yet implemented, but will be soon operational.

transformation $g(\theta)$.

After execution of this routine control returned to SISSUB or MIXSUB.

Listing 5.3

```
      SUBROUTINE GTHETA(PARAM,NDIM,FPOST,WEIGHT,GPARAM,MDIM)
      INTEGER NDIM,MDIM
      DOUBLE PRECISION FPOST,WEIGHT
      DOUBLE PRECISION PARAM(NDIM),GPARAM(MDIM)
C
C
C     JOHNSTON-MODEL, NO "G" OF "THETA"
C
C     PARAM(NDIM) : THE VECTOR WITH PARAMETER-VALUES
C     NDIM        : THE NUMBER OF PARAMETERS OF "PARAM"
C     FPOST       : THE LOG-POSTERIOR-KERNEL-VALUE
C     WEIGHT      : THE VALUE OF THE WEIGHT
C     GPARAM(MDIM): THE VECTOR THE VALUES OF FUNCTION "G" OF "THETA"
C     MDIM        : THE NUMBER OF PARAMETERS OF"GPARAM"
C
C     This option is not selected, so the body of the routine is left empty.
C
      RETURN
      END
```

6 AN EXAMPLE OF INPUT AND OUTPUT OF SISAM

6.1. The Parameter File

This file provides the program SISAM with information on the model studied; what input the program expects and what kind of output the program yields. An example of the parameter file that corresponds with the model of Section 5 is given below.

Listing 6.1

File: JSISAM.PAR, input-directives for the program SISAM, running the Johnston-model

```
JOHNSTON MODEL                                  ! TITLE
3                                ! DIMENSION OF THE VECTOR "THETA"
0                                ! DIMENSION OF THE VECTOR "THETAG": =0 IF NO "G"
JSISAM.RES                       ! NAME OF THE PRINT-OUTPUT-FILE
JOHNSTON.INP                     ! NAME OF THE INPUT-FILE: "PMODE & HINV"
-2.00 -1.70 -0.40                ! LOWER BOUNDS FOR "THETA"
 0.80  0.25  1.00                ! UPPER BOUNDS FOR "THETA"
1                                ! AUXILIARY OUTPUT (0=NO, 1=YES)
1                                ! UNIVARIATE MARGINALS (0=NO, 1=YES)
1                                ! BIVARIATE MARGINALS (0=NO, 1=YES)
1                                ! PLOT-FILE BIV.MAR.POSTERIOR (0=NO, 1=YES)
JSISAM.PLP                       ! NAME OF THE PLOT-FILE
0                                ! PLOT-FILE BIV.MAR.IMPORTANCE (0=NO, 1=YES)
79                               ! INITIAL VALUE OF RANDOM NUMBER GENERATOR
1                                ! DEGREES OF FREEDOM OF STUDENT-T
2                                ! NUMBER OF ROUNDS
20000                            ! NUMBER OF DRAWINGS
2                                ! NUMBER OF ROTATIONS
1                                ! SAVE "MEANPO & COVPO" (0=NO, 1=YES)
JSISAM.SAV                       ! NAME OF THE SAVE-FILE
```

If one starts the program in batch mode it will start reading this parameter file. If one starts the program interactively, the user will be prompted for every piece of input information separately. Most of the directives are self explanatory. For convenience, we make a few remarks.

AUXILIARY OUTPUT:

If the auxiliary output option is switched on, the program will print a

frequency distribution of the weights and, for the ten drawings with the largest weights, the program will print the values of the importance function, the posterior kernel, and the parameters.

PLOTS FOR BIVARIATE MARGINALS:

If the directive for bivariate marginals is switched on, additional directives are required to control whether or not plot files should be created for later use with an other program. Otherwise only line printer plots will be written to the output print file. Only after the maximum number of rounds is reached, data will be written to either plot files. If more then one rotation is requested the data of the plot file for the next rotation are appended to the former in the respective files.

By default the parameter values are divided in 15 classes. So every plot consists of 225 data lines in the plot file.

NUMBER OF DRAWINGS:

The number of drawings stands for the number of not rejected drawings. That is drawings within the bounds and not meeting the restrictions specified in routine RSTRCT (see Section 5)

NUMBER OF ROUNDS:

The number of times that intermediate results are printed. For instance, with 2 rounds and 20.000 drawings, results are printed at 20.000 drawings and 40.000 drawings.

SAVE FILE:

After all rotations are performed, one can request to save the last calculated moments of the vector θ in a file specified by the user. This is useful if one wants to rerun the program after inspection of the posterior results.

6.2 The Input Data and Starting Values

The input file JOHNSTON.INP (see listing 6.1) is read in the main program. It contains the mode, which is used as location estimate and the scaling matrix for the multivariate Student-t density in the first rotation. These starting values are usually obtained by maximization of the log

posterior density (mode), and by computing minus the inverse of the log posterior density evaluated at the mode (scaling matrix). (see the first page of the output file in listing 6.2)

The file JOHNSTON.DAT contains the data with which the model is estimated. This file is read once in the user supplied routine PSTROR at the first call of that routine in the main program (ICALL = *true*).

6.3 The Output file

The file JSISAM.RES contains the output for the printer according to the directives listed in the file JSISAM.PAR. Only the first page of the output is listed in listing 6.2 and the final results of rotation two, round two.

The title of each page shows the title line of the parameter file and the current rotation and round number. The subtitle shows information on the number of drawings from the Student-t distribution and the number of calls of the user supplied routine PSTROR.

As requested, the data of the bivariate marginal posterior densities are sent to a separate output file called JSISAM.PLP. (see figure 5)

Postscript

Recently many new applications of Bayesian statistical analysis using Momte Carlo integration have appeared in the literature. A necessarily incomplete list of references includes the following authors: Bauwens and Richard (1985), Boender and Van Dijk (1991), DeJong (1991), DeJong and Whiteman (1989, 1990, 1991a, 1991b), Geweke (1986, 1988a, 1988b, 1988c, 1989a, 1989b), Geweke, Marshall and Zarkin (1986), Kim and Maddala (1991), Oh and Berger (1989), Richard and Steel (1988), Steel (1987), Van Dijk and Kloek (1983), West (1990), Zellner and Rossi (1984), Zellner, Bauwens and Van Dijk (1988).

Listing 6.2

Print output file JSISAM.RES

```
*************************************************************************
*                                                                       *
*   --- SIMPLE IMPORTANCE SAMPLING ---  (SISAM)          DATE: 31-Oct-91 *
*  JOHNSTON MODEL                                        ROTATION:     1 *
*                                                                       *
*************************************************************************

             1          2          3

 BOUNDS OF THE PARAMETERS

      -2.0000    -1.7000    -0.4000

       0.8000     0.2500     1.0000

 IMPORTANCE MEANS AND STANDARDDEVIATIONS

       0.4579     0.0893     0.3629

       0.1013     0.0354     0.1124

 IMPORTANCE CORRELATION MATRIX

 1     1.0000

 2     0.8769     1.0000

 3     0.1738    -0.1616     1.0000

 IMPORTANCE COVARIANCE MATRIX

 1     0.0103

 2     0.0031     0.0013

 3     0.0020    -0.0006     0.0126


 EIGENVALUES OF MIN-INVERSE-HESSIAN-MATRIX
    0.000145   0.010156   0.013850


 INITIAL VALUE OF RANDOM NUMBER GENERATOR         :   79

 DEGREES OF FREEDOM FOR THE STUDENT-T DISTRIBUTION:    1
```

```
*****************************************************************************
*                                                                           *
*   --- SIMPLE IMPORTANCE SAMPLING ---  (SISAM)                DATE: 31-Oct-91 *
*  JOHNSTON MODEL                                              ROTATION:     2 *
*                                                              ROUND:        2 *
*                                                                           *
*****************************************************************************

  NUMBER OF ACCEPTED RANDOM DRAWINGS   40000
  NUMBER OF REJECTED RANDOM DRAWINGS   22792
  NUMBER OF FUNCTION EVALUATIONS       80000

  IMPORTANCE MEANS AND STANDARDDEVIATIONS (TRUNCATED)

       -0.6653    -0.3314     0.3178
        0.6307     0.2590     0.1780

  IMPORTANCE CORRELATION MATRIX (TRUNCATED)

  1     1.0000
  2     0.8485     1.0000
  3     0.0617     0.1785     1.0000

  IMPORTANCE COVARIANCE MATRIX (TRUNCATED)

  1     0.3977
  2     0.1386     0.0671
  3     0.0069     0.0082     0.0317


  POSTERIOR MEAN AND STANDARDDEVIATION OF THE PARAMETER VECTOR "THETA"

       -0.5993    -0.3092     0.3125
        0.7862     0.3289     0.1465

  POSTERIOR CORRELATION MATRIX OF THE PARAMETER VECTOR "THETA"

  1     1.0000
  2     0.9174     1.0000
  3     0.1680     0.3088     1.0000

  POSTERIOR COVARIANCE MATRIX OF THE PARAMETER VECTOR "THETA"

  1     0.6181
  2     0.2373     0.1082
  3     0.0193     0.0149     0.0215


  NUMERICAL ERROR ESTIMATES OF POSTERIOR MEAN

   ERROR           :
      0.010348   0.004266   0.001478
   RELATIVE ERROR :
      0.013163   0.012970   0.010088
   CORREL. COEFF. :
      0.158431  -0.127086   0.914430
```

```
*************************************************************************
*                                                                       *
*   --- SIMPLE IMPORTANCE SAMPLING ---  (SISAM)            DATE: 31-Oct-91 *
*  JOHNSTON MODEL                                          ROTATION:     2 *
*                                                          ROUND:        2 *
*                                                                        *
*************************************************************************

  NUMBER OF ACCEPTED RANDOM DRAWINGS   40000     DENOMINATOR:
  NUMBER OF REJECTED RANDOM DRAWINGS   22792     MEAN      =   0.151372E-04
  NUMBER OF FUNCTION EVALUATIONS       80000     STD. DEV. =   0.281954E-04

  FREQUENCIES OF "IPOW": (WEIGHT = 0.*****E+IPOW)

      -24      -23      -22      -21      -20      -19      -18      -17
        0        0        1        0        4        7        9       11

      -16      -15      -14      -13      -12      -11      -10       -9
       16       13       24       49       67      122      154      283

       -8       -7       -6       -5       -4       -3       -2       -1
      444      720     1604    18715    17144      613        0        0


  TEN DRAWINGS WITH LARGEST WEIGHT

         W            LN(IMP)         LN(POS)          (THETA(I), I = 1,NDIM)

  0.5527297E-03     -3.3737669    -10.8744084     0.46103   0.06702   0.53429
  0.5460985E-03     -4.2693571    -11.7820682     0.51246   0.07082   0.62155
  0.5386072E-03     -2.6889069    -10.2154309     0.51115   0.09543   0.44738
  0.5366733E-03     -2.8858177    -10.4159388     0.44728   0.07270   0.48715
  0.5352066E-03     -2.9134306    -10.4462884     0.51777   0.09146   0.47298
  0.5312612E-03     -2.7768370    -10.3170938     0.47654   0.08802   0.47074
  0.5143080E-03     -2.3700850     -9.9427733     0.45672   0.08229   0.41576
  0.4986392E-03     -2.4400944    -10.0437222     0.48720   0.09674   0.41912
  0.4985223E-03     -2.2482296     -9.8520918     0.46191   0.08886   0.38700
  0.4862944E-03     -2.9264063    -10.5551026     0.42811   0.06546   0.49489
```

```
*****************************************************************************
*                                                                           *
*   --- SIMPLE IMPORTANCE SAMPLING ---  (SISAM)            DATE: 31-Oct-91 *
*  JOHNSTON MODEL                                          ROTATION:      2 *
*                                                          ROUND:         2 *
*                                                                           *
*****************************************************************************
```

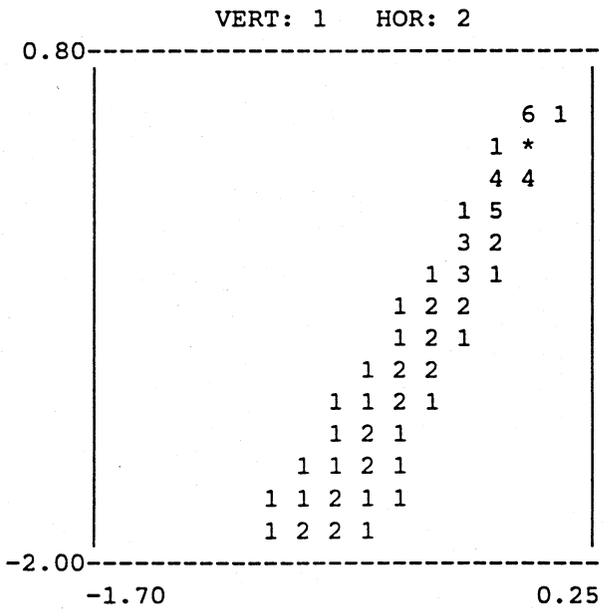
MARGINAL POSTERIOR DENSITIES "P", AND
MARGINAL IMPORTANCE DENSITIES (TRUNCATED) "I"

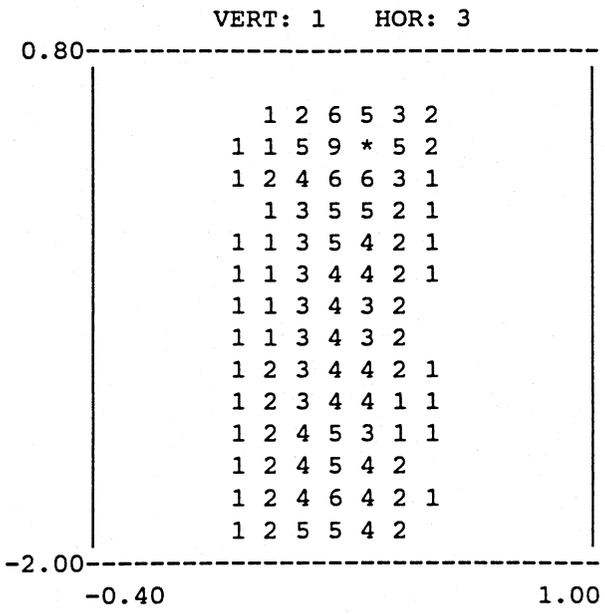
PARAMETER 1			PARAMETER 2		
	P	I		P	I
(-2.00 , -1.81)	0.072	0.034	(-1.70 , -1.57)	0.001	0.000
(-1.81 , -1.63)	0.070	0.041	(-1.57 , -1.44)	0.001	0.000
(-1.63 , -1.44)	0.063	0.052	(-1.44 , -1.31)	0.003	0.001
(-1.44 , -1.25)	0.064	0.066	(-1.31 , -1.18)	0.005	0.002
(-1.25 , -1.07)	0.061	0.082	(-1.18 , -1.05)	0.010	0.003
(-1.07 , -0.88)	0.059	0.096	(-1.05 , -0.92)	0.022	0.008
(-0.88 , -0.69)	0.060	0.113	(-0.92 , -0.79)	0.047	0.026
(-0.69 , -0.51)	0.059	0.113	(-0.79 , -0.66)	0.076	0.063
(-0.51 , -0.32)	0.061	0.104	(-0.66 , -0.53)	0.100	0.112
(-0.32 , -0.13)	0.065	0.089	(-0.53 , -0.40)	0.108	0.165
(-0.13 , 0.05)	0.070	0.068	(-0.40 , -0.27)	0.113	0.204
(0.05 , 0.24)	0.095	0.055	(-0.27 , -0.14)	0.120	0.180
(0.24 , 0.43)	0.123	0.041	(-0.14 , -0.01)	0.151	0.126
(0.43 , 0.61)	0.077	0.028	(-0.01 , 0.12)	0.233	0.073
(0.61 , 0.80)	0.001	0.017	(0.12 , 0.25)	0.010	0.036

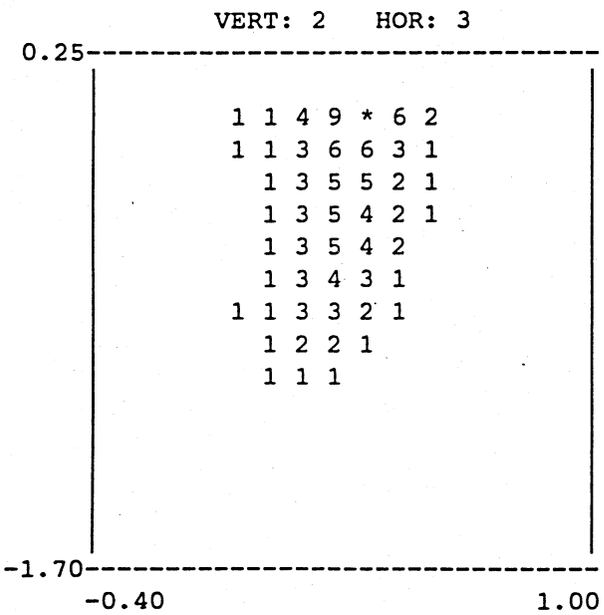
PARAMETER 3		
	P	I
(-0.40 , -0.31)	0.001	0.003
(-0.31 , -0.21)	0.001	0.006
(-0.21 , -0.12)	0.004	0.009
(-0.12 , -0.03)	0.011	0.017
(-0.03 , 0.07)	0.032	0.033
(0.07 , 0.16)	0.081	0.072
(0.16 , 0.25)	0.187	0.165
(0.25 , 0.35)	0.272	0.279
(0.35 , 0.44)	0.239	0.225
(0.44 , 0.53)	0.115	0.102
(0.53 , 0.63)	0.041	0.045
(0.63 , 0.72)	0.010	0.021
(0.72 , 0.81)	0.004	0.012
(0.81 , 0.91)	0.002	0.007
(0.91 , 1.00)	0.000	0.004

```
*****************************************************************************
*                                                                           *
*   --- SIMPLE IMPORTANCE SAMPLING ---  (SISAM)              DATE: 31-Oct-91 *
*  JOHNSTON MODEL                                           ROTATION:      2 *
*                                                           ROUND:         2 *
*                                                                           *
*****************************************************************************
```

BIVARIATE MARGINAL POSTERIOR DENSITIES

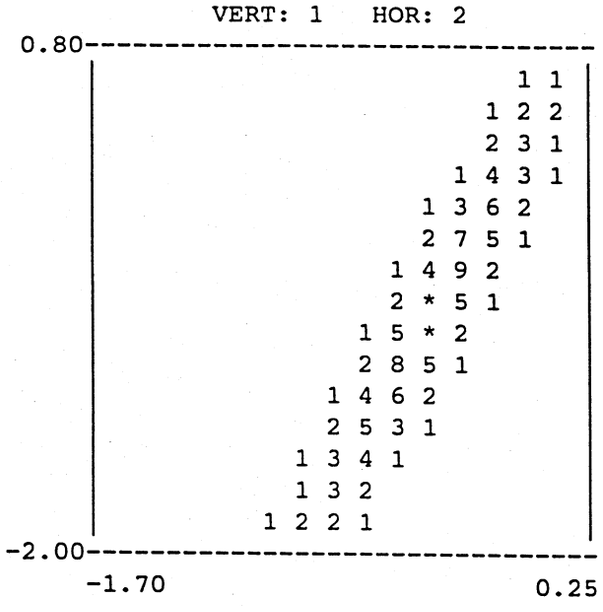


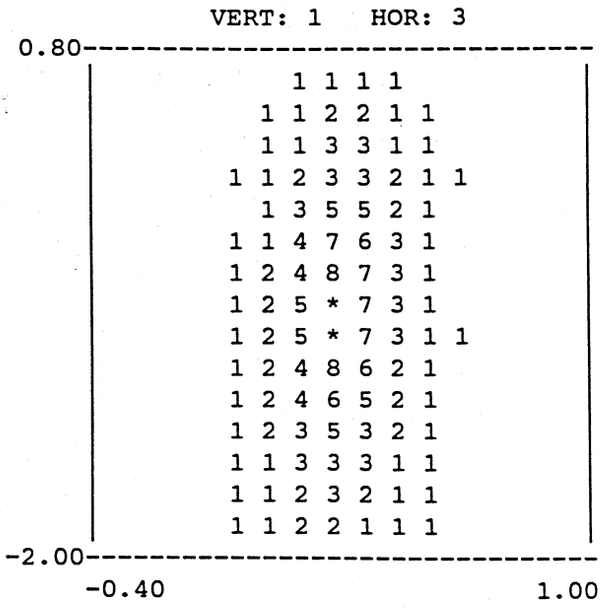




```
**********************************************************************************
*                                                                                *
*   --- SIMPLE IMPORTANCE SAMPLING ---  (SISAM)              DATE: 31-Oct-91 *
*  JOHNSTON MODEL                                            ROTATION:      2 *
*                                                            ROUND:         2 *
*                                                                                *
**********************************************************************************
```

BIVARIATE MARGINAL DENSITIES OF IMPORTANCE FUNCTION (TRUNCATED)





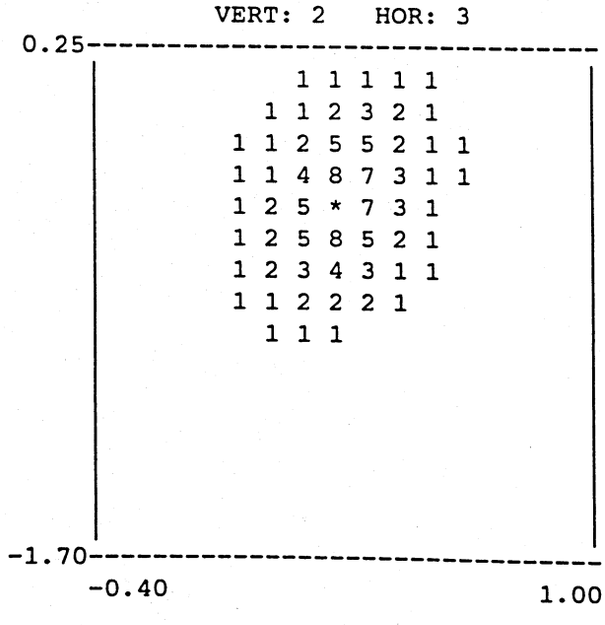
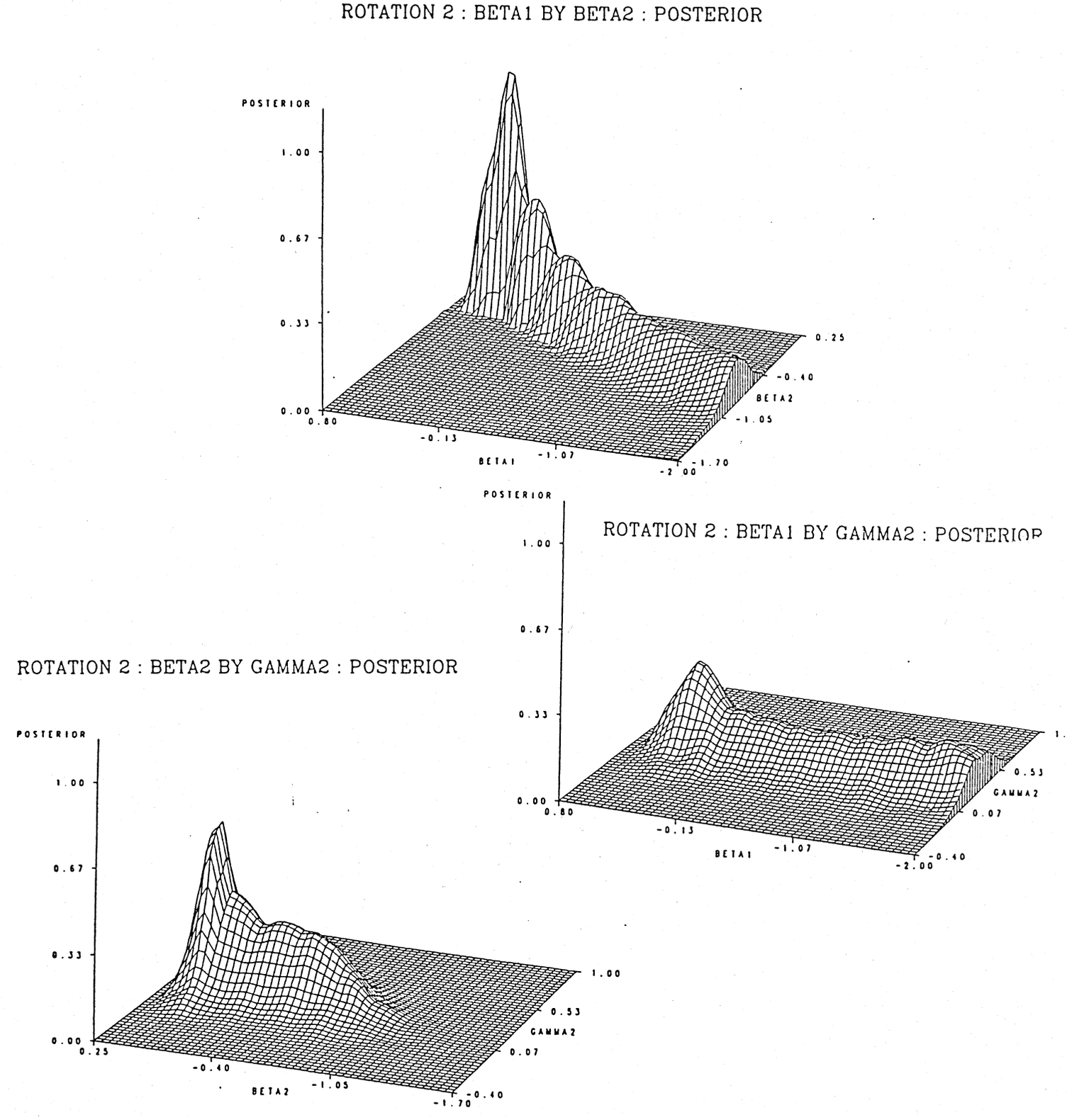


Figure 5. Plot output from JSISAM.PLP.



References

Abramowitz, M and I.A. Stegun (1964) *Handbook of Mathematical Functions*, Washington, DC: National Bureau of Standards.

Bauwens, L. (1984) *Bayesian Full Information Analysis of Simultaneous Equation Models Using Integration by Monte Carlo*, New York: Springer Verlag.

Bauwens, L. and J.F. Richard (1985) "A 1-1 Poly-*t* Random Variable Generator with Applications to Monte Carlo Integration" *Journal of Econometrics* 29, 19-46.

Boender, C.G.E. and H.K. Van Dijk (1991) "Bayes Estimates of Multi-Criteria Decision Alternatives Using Monte Carlo Integration" Report: 9144/A, Econometric Institute, Erasmus University, Rotterdam. (forthcoming in *Statistica Neerlandica*)

Billingsley, P. (1979) *Probability and Measure*, New York: Wiley.

Cramér, H. (1946) *Mathematical Methods of Statistics*, Princeton, NJ: Princeton University Press.

DeJong, D.N. (1991) "Bayesian Inference in Limited Dependent Variable Models: An Application to Measuring Strike Duration" Manuscript, Department of Economics, University of Pittsburgh.

DeJong, D.N. and C.H. Whiteman (1989) "Trends and Cycles as Unobserved Components in U.S. Real GNP: A Bayesian Perspective" *Proceedings of the American Statistcal Association, Business and Economics Section*, 63-70.

———, (1990) "Are Output Fluctuations Transitory?: A Bayesian Perspective" Working paper No. 90-26, Department of Economics, University of Iowa.

———, (1991a) "Reconsidering 'Trends and Random Walks in Macroeconomic Time Series'" *Journal of Monetary Economics* (forthcoming).

———, (1991b) "The Temporal Stability of Dividends and Stok Prices: Evidence from the Likelihood Function" *American Economic Review* 81, 600 – 617.

Dhrymes, P.J. (1971) *Econometrics, Statistical Foundations and Applications*, New York: Harper and Row.

Geweke, J. (1986) "Exact Inferences in the Inequality Constrained Normal Linear Regression Model" *Journal of Applied Econometrics* 1, 127-141.

———, (1988a) "Exact Inference in Models with Autoregressive Conditional Heteroscedasticity" in E. Berndt, H. White and W. Barnett, eds. *Dynamic Econometric Modeling*, University Press, Cambridge.

———, (1988b) "The Secular and Cyclical Behavior of Real GDP in Nineteen OECD Countries" *Journal of Business and Economic Statistics* 6, 479-486.

———, (1988c) "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference" *Journal of Econometrics* 38, 73-89.

———, (1989a) "Exact Predictive Densities in Linear Models with ARCH Disturbances" *Journal of Econometrics* 40, 63-86.

———, (1989b) "Bayesian Inference in Econometric Models Using Monte Carlo Integration" *Econometrica* 57, 1317-1339.

Geweke, J., R.C. Marchal and G.A. Zarkin (1986) "Exact Inference for Continuous Time Markov Chain Models" *Review of Economic Studies LIII*, 653-669.

Hammersley, J.M., and D.C. Handscomb (1964) *Monte Carlo Methods*, 1st ed. London: Methuen.

————, (1979) *Monte Carlo Methods*, 2nd ed. London: Chapman and Hall.

Hop, J.P. and H.K. van Dijk (1990) "SISAM and MIXIN: Two Algorithms for the Computation of Posterior Moments and Densities Using Monte Carlo Integration" Report 9031/A, Econometric Institute, Erasmus University, Rotterdam, pp. 150.

Judge, G.C., W.E. Griffiths, H. Lutkepohl and T.C. Lee (1988) *Introduction to the Theory and Practice of Econometrics*, 2nd ed. New York: Wiley.

Kim, In-Moo and G.S. Maddala (1991) "Structural Change and Unit Roots" Manuscript, Department of Economics, The University of Florida, Gainesville.

Kinderman, A.J. and J.F. Monahan (1980) "New Methods for Generating Students t and Gamma Variables" *Computing* 25, 369-377.

Kinderman, A.J. and J.G. Ramage (1976) "Computer Generation of Normal Random Variables" *Journal of the American Statistical Association* 71, 893-896.

Marsaglia, G. (1984) "The Exact-Approximation Method for Generating Random Variables in a Computer" *Journal of the American Statistical Association* 79, 218-221.

Mood, A.M., F.A. Graybill and D.C. Boes (1974) *Introduction to the Theory of Statistics*, 3rd ed. New York: McGraw-Hill.

Oh, Man-Suk and J.O. Berger (1989) "Adaptive Importance Sampling in Monte Carlo Integration" Manuscript, Purdue University and Duke University.

Richard, J.F. and M.F.J. Steel (1988) "Bayesian Analysis of Systems of Seemingly Unrelated Regression Equations under a Recursive Extended Natural Conjugate Prior Density" *Journal of Econometrics* 38, 7-37.

Rubinstein, R.Y. (1981) *Simulation and the Monte Carlo Method*, New York: Wiley.

Steel, M.F.J. (1987) "A Bayesian Analysis of Multivariate Exogeneity: A Monte Carlo Approach" Doctoral dissertation, Université Catholique de Louvain.

Van Dijk, H.K. (1984) "Posterior Analysis of Econometric Models Using Monte Carlo Integration" Doctoral dissertation, Erasmus University, Rotterdam.

Van Dijk, Herman K. (1987) "Some Advances in Bayesian Estimation Methods Using Monte Carlo Integration" in T.B. Fomby and G.F. Rhodes, eds. *Advances in Econometrics* 6, 215-261, Greenwich Connecticut: JAI Press.

Van Dijk, H.K., J.P. Hop and A.S. Louter (1987) "An Algorithm for the Computation of Posterior Moments and Densities Using Simple Importance Sampling" *The Statistician* 36, 83-90.

Van Dijk, H.K. and T. Kloek (1980) "Further Experience in Bayesian Analysis Using Monte Carlo Integration" *Journal of Econometrics* 14, 307-328.

————, (1983) "Experiments with some Alternatives for Simple Importance Sampling in Monte Carlo Integration" in J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds. *Bayesian Statistics* 2, Amsterdam: North Holland, 511-530.

Van Dijk, H.K., T. Kloek and C.G.E. Boender (1985) "Posterior Moments Computed by Mixed Integration" *Journal of Econometrics* 29, 3-18.

West, M. (1990) "Bayesian Kernel Density Estimation" ISDS Discussion paper #90-A12, Duke University.

Zellner, A., L. Bauwens and H.K. Van Dijk (1988) "Bayesian Specification Analysis and Estimation of Simultaneous Equation Models Using Monte Carlo Methods" *Journal of Econometrics* 38, 39-72.

Zellner, A. and P. Rossi (1984) "Bayesian Analysis of Dichotomous Quantal Response Models" *Journal of Econometrics* 25, 365-394.