UNIVERSIDADE DE SÃO PAULO

Some Practical Aspects of Approximate Bayesian Inference

JORGE ALBERTO ACHCAR

Nº 003

N O T A S

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

Some Practical Aspects of Approximate Bayesian Inference

JORGE ALBERTO ACHCAR

Nº 003

N O T A S D O I C M S C
Série Estatística

São Carlos
Jun. / 1993

Some Practical Aspects of Approximate Bayesian Inference

Jorge Alberto Achcar
Universidade de São Paulo
ICMSC,USP, Caixa Postal 668
13.560, São Carlos, S.P., Brasil

Abstract

In this paper, we show in some selected applications a comparative study among some popular existing approximation and numerical methods being used in Bayesian inference. In each application, we show the practical aspects and advantages of each procedure.

Keywords: Approximation methods, numerical methods, Bayesian inference, comparative study.

1 Introduction

The use of numerical or approximation methods to get posterior moments or marginal posterior densities of interest is becoming very popular in Bayesian statistics. Usually, the Bayesian statistician has interest to find posterior moments of the form

$$E\{g(\Theta)|\mathcal{D}\} = \frac{\int g(\Theta)\pi(\Theta)l(\Theta|\mathcal{D})d\Theta}{\int \pi(\Theta)l(\Theta|\mathcal{D})d\Theta} \tag{1}$$

where $g(\Theta)$ is a selected function of $\Theta \in R^m, \pi(\Theta)$ is a prior density, $l(\Theta|\mathcal{D})$ is the likelihood function for Θ given a data set $\mathcal{D}$, and marginal posterior densities of the form

$$\pi(\Theta_1|\mathcal{D}) = \int \pi(\Theta_1, \Theta_2|\mathcal{D})d\Theta_2 \tag{2}$$

where $\pi(\Theta_1, \Theta_2, |\mathcal{D})$ is the joint posterior density for $\Theta = (\Theta_1, \Theta_2), \Theta_1 \in R^k$ and $\Theta_2 \in R^{m-k}$.

When it is not possible to get exact analytical solutions for the integrals in (1) and (2), the Bayesian statistician chooses one among the different existing strategies: the use of numerical methods (see for example, Naylor and Smith, 1982); the use of approximation methods for integrals (see for example, Lindley, 1980; or Tierney and Kadane, 1986); or the use of Monte Carlo procedures or Gibbs sampling (see for example Kloek and Van Dijk, 1978; or Gelfand and Smith, 1990).

In some applications, the Bayesian statistician could be undecided by the best strategy in terms of computational cost and accuracy of the obtained results. In this paper, we show in one comparative study with some selected applications, the practical aspects of each integration procedure available to solve the integrals in (1) and (2).

2 Posterior Moments

In this section, we present a brief summary of some well known procedures available to solve integration problems in Bayesian statistics.

2.1 Tierney and Kadane Approximation

The method of approximation for posterior moments introduced by Tierney and Kadane (1986) is based on Laplace's approximations for the integrals in the numerator and denominator of (1). Laplace's method for approximation of integrals is used to solve integrals of the form

$$I = \int f(\Theta)exp\{-nh(\Theta)\}d\Theta \tag{3}$$

where $-nh(\Theta)$ is a function having a maximum at $\hat{\Theta}$ and which satisfies the usual regularity conditions.

To approximate integrals of the form (3), Laplace's method assumes an expansion of h and f in Taylor series about $\hat{\Theta}$ (see Tierney and Kadane, 1986; or Tierney, Kass and Kadane, 1989a, 1989b).

With Θ one-dimensional, Laplace's approximation for I is given by

$$\hat{I} \cong \left(\frac{2\pi}{n}\right)^{1/2} \sigma\ f(\hat{\Theta}) exp\{-nh(\hat{\Theta})\} \tag{4}$$

where $\sigma = \{h''(\hat{\Theta})\}^{-1/2}$.

In the m-dimensional case,

$$\hat{I} \cong (2\pi)^{m/2} \left\{det\left(n\ \Sigma_h^{-1}(\hat{\Theta})\right)\right\}^{-1/2} f(\hat{\Theta}) exp\left\{-nh(\hat{\Theta})\right\} \tag{5}$$

where $\Sigma_h^{-1}(\hat{\Theta})$ is the Hessian matrix of h at $\hat{\Theta}$, given by $\Sigma_h^{-1}(\hat{\Theta}) = \left(\frac{\partial^2 h}{\partial\Theta_i \partial\Theta_j}\right)\Big|_{\hat{\Theta}} ; i,j = 1,2,\ldots,m$.

To approximate the posterior moment (1) using Laplace's method, we could consider $\pi(\Theta)l(\Theta|\mathcal{D}) = exp\{-nh(\Theta)\}$ for the numerator and denominator of (1) with f equals to g and 1, respectively. Thus, we get the modal approximation

$$\hat{E}\{g(\Theta)|\mathcal{D}\} \ \cong\ g(\hat{\Theta})\left\{1 + 0\left(n^{-1}\right)\right\} \tag{6}$$

where $\hat{\Theta}$ is the mode of $\pi(\Theta|\mathcal{D})$.

We can consider other choices for f in the integrals in the numerator and denominator of (1) to get more accurate approximations for the posterior moment (1). Considering $f = 1$ in both integrals in (1), we have:

$$E\{g(\Theta)|\mathcal{D}\} \ = \ \frac{\int e^{-nh^*(\Theta)} d\Theta}{\int e^{-nh(\Theta)} d\Theta} \tag{7}$$

where $g(\Theta)$ is a positive function, $-nh(\Theta) = ln\pi(\Theta) + lnl(\Theta|\mathcal{D})$ and $-nh^*(\Theta) = lng(\Theta) - nh(\Theta)$.

Using Laplace's method for both integrals in (7), we get

$$\hat{E}\{g(\Theta)|\mathcal{D}\} \ \cong\ \left(\frac{\sigma^*}{\sigma}\right) exp\left\{-[h^*\left(\hat{\Theta}^*\right) - h(\hat{\Theta})]\right\} \tag{8}$$

where $\hat{\Theta}$ maximizes $-nh(\Theta), \hat{\Theta}^*$ maximizes $-nh^*(\Theta), \sigma = \left\{det\left(n\ \Sigma_h^{-1}(\hat{\Theta})\right)\right\}^{-1/2}$ and $\sigma^* = \left\{det\left(n\ \Sigma_{h^*}^{-1}\left(\hat{\Theta}^*\right)\right)\right\}^{-1/2}$.

This approximation satisfies,

$$E\{g(\Theta)|\mathcal{D}\} = \hat{E}\{g(\Theta)|\mathcal{D}\}\left(1 + 0(n^{-2})\right) . \qquad (9)$$

Usually, we get good accuracy for the approximation (8), especially for large values of n, but this approximation could be improved in many cases, considering an appropriate reparametrization (see for example, Achcar and Smith, 1990; or Kass and Slate, 1992).

2.2 Numerical Methods

The use of numerical methods is another strategy explored by many Bayesian statisticians (see for example, Naylor and Smith, 1982). Usually, these methods are appropriate for problems with small number of parameters. One of these methods is given by Simpson's rule, where for one-parameter case is given by

$$\int_a^b f(\Theta)d\Theta = \frac{h}{3}\left(f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots + 2f_{2n-2} + 4f_{2n-1} + f_{2n}\right) \qquad (10)$$

where $fr = f(\Theta_r)$, and the finite interval $[a, b]$ is divided in 2n (even number) of sub intervals, each one of length h, such that $b - a = 2nh$. The error for (10) is $R_n = -\frac{1}{180}(b-a)h^4 f^{(4)}(\xi)$, where $a < \xi < b$.

Simpson's rule is appropriate when the interval $[a, b]$ is finite and $f(\Theta)$ is differentiable, which is not possible to occur in many Bayesian applications, or $f(\Theta)$ could be very complicated.

Naylor and Smith (1982) use Gaussian quadrature procedures to solve the integration problems in Bayesian statistics. To use Gaussian rules, we should choose a non-negative integrable function $w(\Theta)$, such that.

$$\int_a^b f(\Theta)d\Theta = \int_a^b g_G(\Theta)w(\Theta)d\Theta \qquad (11)$$

where $g_G(\Theta) = f(\Theta)/w(\Theta)$.

One possible choice is given by orthogonal polynomials (see for example, Conte and De Boor, 1965) to find efficient methods to approximate (1).

In practical work, we could use different systems of orthogonal polynomials. One possibility, explored by Naylor and Smith (1982) is given by the class of orthogonal polynomials which are orthogonals relative to the weight function $w(\Theta) = exp\{-\Theta^2\}$, in the interval $[a, b] = [-\infty, \infty]$.

Thus, we have the approximation,

$$\int_{-\infty}^{\infty} e^{-\Theta^2} g_G(\Theta) d\Theta \cong \sum_{k=1}^{n} \alpha_k g_G(\Theta_k) + R_n \qquad (12)$$

where $\Theta_1, \Theta_2, \ldots, \Theta_n$ are the roots of the polynomial equation of Hermite $H_n(\Theta) = 0$, and

$$\alpha_k = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 \left[H_{n-1}(\Theta_k)\right]^2} . \qquad (13)$$

The values of the roots Θ_k and of the coeficients α_k are given in tables (see for example, Abramowitz and Stegun, 1965, p. 924). In table 1, we have some values of Θ_k and α_k for Gauss-Hermite quadrature with $n \leq 5$.

Table 1: Roots of polynomial equation of Hermite $H_n(\Theta)$ and coefficients for Gauss-Hermite quadrature ($n \leq 5$).

n	Θ_k	α_k
1	0	1.772454
2	± 0.707107	0.886227
3	0	1.181636
	± 1.224745	0.295410
4	± 0.524648	0.804914
	± 1.650680	0.081313
5	0	0.945309
	± 0.958572	0.393619
	± 2.020183	0.019953

2.3 Monte Carlo Procedure

The approximation and numerical methods of integration considered above, usually are not appropriate for high dimensional problems ($m \geq 5$). An alternative way for these problems is to use Monte Carlo procedure.

To use Monte Carlo procedure for integrations to approximate the posterior moment (1), we should choose a generating density $h_M^*(\Theta)$ defined as an " inportance density" (see for example, Kloek and Van Dijk, 1978; or Geweke, 1989). The choice of the "importance density" $h_M^*(\Theta)$, usually follows two important points:

(*i*) $h^*_M(\Theta)$ approximates the posterior density for Θ.

(*ii*) A sample of size M is easily generated by $h^*_M(\Theta)$.

We can rewrite (1) in the form

$$E\{g(\Theta)|\mathcal{D}\} = \int g(\Theta)w_M(\Theta)d\Theta \tag{14}$$

where $w_M(\Theta) = \dfrac{\pi(\Theta)l(\Theta|\mathcal{D})}{\int \pi(\Theta)l(\Theta|\mathcal{D})d\Theta}$.

Using Monte Carlo, we find

$$\hat{E}\{g(\Theta)|\mathcal{D}\} \cong \sum_{m=1}^{M} \hat{w}_M(\Theta_m)g(\Theta_m) \tag{15}$$

where the weights $\hat{w}(\Theta_m)$ are given by

$$\hat{w}_M(\Theta_m) = \frac{\pi(\Theta_m)l(\Theta_m|\mathcal{D})/h^*_M(\Theta_m)}{\sum_{j=1}^{M}[\pi(\Theta_j)l(\Theta_j|\mathcal{D})/h^*_M(\Theta_j)]}$$

Observe that,

(*i*) The sum of weights $\hat{w}_M(\Theta_m)$ is equals to one.

(*ii*) Since $l(\Theta_m/\mathcal{D})$ and $\pi(\Theta_m)$ are given in ratio form, we only need their Kernels in place of the complete densities.

(*iii*) The accuracy of the approximation (15) depends on the choice of $h^*_M(\Theta)$.

3 Marginal Posterior Densities

To approximate marginal posterior densities of one or some coordinates of $\Theta \in R^m$, we could use directly one of the existing approximation or numerical methods to get the marginal posterior density of interest. If we have interest to approximate marginal posterior densities of nonlinear functions of the form $\eta = g(\Theta)$, we could find the marginal posterior density of η based on a complete specification of a transformation. or we could use one of the approximation procedures for nonlinear functions proposed by Tierney, Kass and Kadane (1989) and by Leonard, Hsu and Tsui (1989).

3.1 Tierney, Kass and Kadane Approximation

When we consider a partition of $\Theta \in R^m$ in the form $\Theta = (\Theta_1, \Theta_2), \Theta_1 \in R^k$ and $\Theta_2 \in R^{m-k}$, we could use Laplace's method directly to get approximations for the marginal posterior density of $\eta = g(\Theta) = \Theta_1$, given by

$$\pi(\Theta_1|\mathcal{D}) = \int \pi(\Theta_1, \Theta_2|\mathcal{D}) d\Theta_2 \tag{16}$$

where $\pi(\Theta_1, \Theta_2|\mathcal{D}) = c\, f(\Theta_1, \Theta_2) exp\{-nh(\Theta_1, \Theta_2)\}$.

Using Laplace's method, we get

$$\hat{\pi}_1(\Theta_1|\mathcal{D}) \cong (2\pi)^{-\frac{1}{2}k} \left\{ \frac{det\ \Sigma(\Theta_1)}{det\ \Sigma} \right\}^{1/2} \frac{\pi(\Theta_1, \hat{\Theta}_2(\Theta_1)|\mathcal{D})}{\pi(\hat{\Theta}_1, \hat{\Theta}_2|\mathcal{D})} \tag{17}$$

where Σ is the inverse of the Hessian matrix of $nh(\Theta_1, \Theta_2)$ at $\hat{\Theta} = (\hat{\Theta}_1, \hat{\Theta}_2)$, the local maximum of $-nh(\Theta_1, \Theta_2), \hat{\Theta}_2(\Theta_1)$ maximizes $-nh_{\Theta_1}(\Theta_2) = -nh(\Theta_1, \Theta_2)$ with Θ_1 fixed and $\Sigma(\Theta_1)$ is the inverse of the Hessian matrix of $nh_{\Theta_1}(\Theta_2)$ at $\hat{\Theta}_2(\Theta_1)$. Usually, we consider $f(\Theta_1, \Theta_2) = 1$ in (16).

In many applications, we have interest to approximate marginal posterior densities of more general functions $\eta = g(\Theta), \Theta \in R^m$, where g is a nonlinear real-valued or k-dimensional vector-valued function and it is assumed to be smooth and to have a gradient different of zero, or a Jacobian that is of rank k, near the mode of the joint posterior density for Θ. In some applications, it is possible to use a global transformation $\phi = \phi(\Theta)$, such that $\phi = (g(\Theta), \Theta_2)$, but in many cases this is difficult or impossible to obtain.

Tierney, Kass and Kadane (1989) consider an approximation that does not depend on an explicit specification of a transformation. Considering $g(\Theta)$ a k-dimensional vector-valued function, the TKK approximation is given by

$$\hat{\pi}_{TKK}(\eta|\mathcal{D}) \cong$$
$$\cong (2\pi)^{-\frac{1}{2}k} \left\{ \frac{det\ \Sigma(\eta)}{det(\ \Sigma) det \left[(Dg)' \ \Sigma(\eta)(Dg) \right]} \right\}^{1/2} \frac{\pi\left(\hat{\Theta}(\eta)|\mathcal{D}\right)}{\pi\left(\hat{\Theta}|\mathcal{D}\right)} \tag{18}$$

where Σ is the inverse of the Hessian matrix of $nh(\Theta)$ at $\hat{\Theta}, \hat{\Theta}(\eta)$ maximizes $-nh(\Theta)$ subject to the constraint $\eta = g(\Theta)$, $\Sigma(\eta)$ is the inverse of the Hessian matrix of $nh(\Theta)$ at $\hat{\Theta}(\eta)$ and Dg is the gradient or Jacobian of g evaluated at $\hat{\Theta}(\eta)$.

Observe that:

(i) If g is one-dimensional then Dg is a column vector with components $\partial g\left\{\hat{\Theta}(\eta)\right\}/\partial\Theta_i$, for $i = 1, 2, \ldots, m$.

(*ii*) If g is k-dimensional, then Dg is an $m \times k$ matrix with elements $\partial gj \left\{\hat{\Theta}(\eta)\right\} / \partial \Theta_i$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, k$.

We can write (18) in the form,

$$\hat{\pi}_{TKK}(\eta|\mathcal{D}) \propto \left\{ \frac{det\ \Sigma(\eta)}{det\left[(Dg)'\ \Sigma(\eta)(Dg)\right]} \right\}^{1/2} \pi(\hat{\Theta}(\eta)/\mathcal{D}) \tag{19}$$

3.2 Leonard, Hsu and Tsui Approximation

Leonard, Hsu and Tsui (1989) propose another form to approximate marginal posterior densities of nonlinear functions of the form $\eta = g(\Theta)$.

The LHT approximation for the marginal posterior density of $\eta = g(\Theta)$ is considered by expanding the logarithm of the joint posterior density for Θ about $\hat{\Theta}(\eta)$, the maximum of $-nh(\Theta)$ subject to the constraint $\eta = g(\Theta)$, up to second order. The LHT approximation is given by

$$\begin{aligned} \hat{\pi}_{LHT}(\eta|\mathcal{D}) &\propto \left\{det\ \Sigma(\eta)\right\}^{1/2} \pi\left(\hat{\Theta}(\eta)|\mathcal{D}\right) \\ &\times exp\left\{\frac{1}{2}\ l'_\eta\ \Sigma(\eta) l_\eta\right\} f\left(\eta/\tilde{\Theta}(\eta),\ \Sigma(\eta)\right) \end{aligned} \tag{20}$$

where l_η is the gradient of $nh(\Theta)$ at $\hat{\Theta}(\eta), \tilde{\Theta}(\eta) = \hat{\Theta}(\eta) - \Sigma(\eta) l_\eta$ and $f(\eta/\mu, \mathbb{C})$ is the density of $\eta = g(\Theta)$ at η when Θ has a normal distribution with mean vector μ and covariance matrix C. Observe that f, in general, needs to be approximated.

Leonard, Hsu and Tsui (1989) suggest to approximate $f\left(\eta/\tilde{\Theta}(\eta),\ \Sigma(\eta)\right)$ by the density of a normal distribution with mean $g\left(\tilde{\Theta}(\eta)\right)$ and variance

$$\left[Dg\left(\tilde{\Theta}(\eta)\right)\right]'\ \Sigma(\eta)\left[Dg\left(\tilde{\Theta}(\eta)\right)\right].$$

3.3 Gibbs Sampler

The Gibbs sampler is a procedure for generating random variables from a marginal distribution indirectly, without having to calculate the density (see for example, Casella and George, 1992).

Consider $\pi(\Theta_1, \Theta_2, \ldots, \Theta_m|\mathcal{D})$ a joint posterior density for $\Theta_1, \Theta_2, \ldots, \Theta_m$ and assume that we have interest to find some characteristics of the marginal posterior density

$$\pi(\Theta_1|\mathcal{D}) = \int \ldots \int \pi(\Theta_1, \Theta_2, \ldots, \Theta_m|\mathcal{D}) d\Theta_2 \ldots d\Theta_m \ , \tag{21}$$

like the posterior mean $E\{\Theta_1|\mathcal{D}\}$ or the posterior variance var $\{\Theta_1|\mathcal{D}\}$.

Rather than compute or approximate $\pi(\Theta_1|\mathcal{D})$ directly, the Gibbs sampler generates a sample $\Theta_{11}, \Theta_{12}, \ldots, \Theta_{1n}$ of $\pi(\Theta_1|\mathcal{D})$ without requiring it. From this simulated sample, we can calculate the mean, variance and other characteristics of $\pi(\Theta_1|\mathcal{D})$. For example, to calculate $E\{\Theta_1|\mathcal{D}\}$, we could use $\frac{1}{n}\sum_{i=1}^{n}\Theta_{1i}$.

Consider the special two-dimensional case, that is, with $\Theta = (\Theta_1, \Theta_2), \Theta \in R^2$. Generate a sample of $\pi(\Theta_1|\mathcal{D})$ from samples of the conditional distributions $\pi(\Theta_1|\Theta_2, \mathcal{D})$ and $\pi(\Theta_2|\Theta_1, \mathcal{D})$, usually known. The "Gibbs" sequence is given by

$$\Theta'_{20}, \Theta'_{10}, \Theta'_{21}, \Theta'_{11}, \Theta'_{22}, \Theta'_{12}, \ldots, \Theta'_{2k}, \Theta'_{1k} \ . \tag{22}$$

The initial value Θ'_{20}, is specified, and the rest of (22) is obtained iteratively by alternately generating values

$$\begin{aligned} \Theta'_{1j} &\sim \pi(\Theta_1|\Theta'_{2j}, \mathcal{D}) \\ \text{and} \qquad & \\ \Theta'_{2(j+1)} &\sim \pi(\Theta_2|\Theta'_{1j}, \mathcal{D}) \ . \end{aligned} \tag{23}$$

The distribution of Θ_{1k} converges to $\pi(\Theta_1|\mathcal{D})$ when $k \to \infty$. A simple proof of this result is given by Casella and George (1992).

Gelfand and Smith (1990) suggest generating m independent Gibbs sequences of length k, and then using the final value of each sequence. If k is large enough, we have an approximate *iid* sample of $\pi(\Theta_1|\mathcal{D})$.

4 Some Selected Applications

4.1 An Application Using the Cauchy Distribution

Let 11.4, 7.3, 9.8, 13.7 and 10.6 be a random sample of size n of a Cauchy distribution with density

$$f(y|\Theta) = \pi^{-1}\left[1 + (y - \Theta)^2\right]^{-1} \tag{24}$$

where $-\infty < y < \infty$ (data set given in Box and Tiao, 1973, p. 64). Considering a locally uniform noninformative prior density for Θ, the posterior density for Θ is given by

$$\pi(\Theta|\mathcal{D}) = c\ H(\Theta) \tag{25}$$

where $H(\Theta) = 10^5[1+(11.4-\Theta)^2]^{-1}[1+(7.3-\Theta)^2]^{-1}\ldots[1+(10.6-\Theta)^2]^{-1}$, and $c^{-1} = \int_{-\infty}^{\infty} H(\Theta)d\Theta$.

To find c^{-1}, we could use an integration procedure given in section 2. Observe that $H(\Theta)$ is different of zero for values between 6 and 14 (see figure 1).

Considering Simpson's rule with $h = 0.5, a = 6, b = 15$ and $2n = 18$ (see (10)), we find $c^{-1} = 536.34$, that is, $c = 0.00186$. Also, $E\{\Theta|\mathcal{D}\} = 10.618$ and $E\{\Theta^2|\mathcal{D}\} = 113.433$.

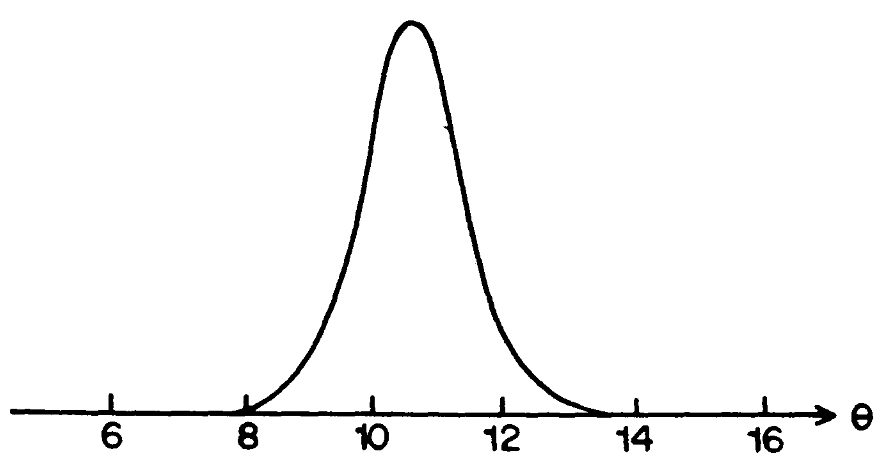

Figure 1: Posterior Density for Θ

Considering a numerical integration procedure based on Gaussian quadrature with Hermite polynomials, we find approximations for $c^{-1}, E\{\Theta|\mathcal{D}\}$ with $g_G(\Theta) = \epsilon^{\Theta^2}H(\Theta)$ in (12). With Hermite polynomials of degree $n = 5$ (see table 1) and since the variation of Θ in $H(\Theta)$ is concentrated between 6 and 16, we have $\epsilon^{\Theta^2}H(\Theta) \cong 0$ for values between -2.020183 and 2.020183. Therefore, we consider a transformation $\xi = \Theta - 11$, that is, $\Theta = \xi + 11$ (observe that $\overline{y} = 10.56$). Thus,

$$\begin{aligned} H(\xi) &= 10^5[1+(0.4-\xi)^2]^{-1} \cdot [1+(-3.7-\xi)^2]^{-1} \cdot [1+(-1.2-\xi)^2]^{-1} \cdot \\ &\cdot \; [1+(2.7-\xi)^2]^{-1}\,[1+(-0.4-\xi)^2]^{-1} \end{aligned} \tag{26}$$

In this parametrization we find $c^{-1} = 531.68$ considering Gauss-Hermite with $n = 5$. With $n = 12$, we find $c^{-1} = 536.30$ (observe that $c^{-1} = 536.34$ considering Simpson's rule).

Also, with $n = 12$, we find $E\{\xi|\mathcal{D}\} = -0.381407$, that is, $E\{\Theta|\mathcal{D}\} = 10.619$ and $E\{\xi^2|\mathcal{D}\} = 0.84294$. Since $E\{\Theta^2|\mathcal{D}\} = E\{\xi^2|\mathcal{D}\} + 22E\{\xi|\mathcal{D}\} + 121$, we get $E\{\Theta^2|\mathcal{D}\} = 113.452$.

Using Monte Carlo procedure, we choose an "importance density" $h_M^*(\Theta)$ equals to a normal density $N(11;4)$ since $H(\Theta)$ is different of zero for values of Θ between 6 and 16.

Generating $M = 100$ observations of the Normal distribution $N(11;4)$, we find approximations for $E\{g(\Theta)|\mathcal{D}\}$ using (15) with

$$\hat{w}_M(\Theta_m) = \frac{H(\Theta_m)/exp\left\{-\frac{1}{8}(\Theta_m - 11)^2\right\}}{\sum_{j=1}^{100} H(\Theta_j)/exp\left\{-\frac{1}{8}(\Theta_j - 11)^2\right\}} \tag{27}$$

With $g_i(\Theta) = \Theta^i$, for $i = 1, 2$, we obtain $\hat{E}\{\Theta|\mathcal{D}\} \cong 10.626$ and $\hat{E}\{\Theta^2|\mathcal{D}\} \cong 113.483$.

Considering $M = 80$ observations generated from $h_M^*(\Theta)$, we get $\hat{E}\{\Theta|\mathcal{D}\} \cong 10.641$ and $\hat{E}\{\Theta^2|\mathcal{D}\} \cong 113.989$.

We also find Tierney and Kadane approximations for $E\{\Theta|\mathcal{D}\}$ and $E\{\Theta^2|\mathcal{D}\}$. For $g_i(\Theta) = \Theta^i, i = 1, 2$, we write,

$$E\{g_i(\Theta)|\mathcal{D}\} = \frac{\int_{-\infty}^{\infty} e^{-nh_i^*(\Theta)} d\Theta}{\int_{-\infty}^{\infty} e^{-nh(\Theta)} d\Theta} \tag{28}$$

where $-nh(\Theta) = 5ln10 - ln\left[1 + (11.4 - \Theta)^2\right] - ln\left[1 + (7.3 - \Theta)^2\right] - ln\left[1 + (9.8 - \Theta)^2\right] - ln\left[1 + (13.7 - \Theta)^2\right] - ln\left[1 + (10.6 - \Theta)^2\right]$, and $-nh_i^*(\Theta) = lng_i(\Theta) - nh(\Theta)$.

From (8), we get

$$\hat{E}\{\Theta|\mathcal{D}\} \cong \frac{\left\{h''(\hat{\Theta})\right\}^{1/2} \hat{\Theta}_1^* H(\hat{\Theta}_1^*)}{\left\{h_1^{*\prime\prime}(\hat{\Theta}_1^*)\right\}^{1/2} H(\hat{\Theta})}$$

and

$$\hat{E}\{\Theta^2|\mathcal{D}\} \cong \frac{\left\{h''(\hat{\Theta})\right\}^{1/2} \hat{\Theta}_2^* H(\hat{\Theta}_2^*)}{\left\{h_2^{*\prime\prime}(\hat{\Theta}_2^*)\right\}^{1/2} H(\hat{\Theta})} \tag{29}$$

where $\hat{\Theta}, \hat{\Theta}_1^*$ and $\hat{\Theta}_2^*$ maximize $-nh(\Theta), -nh_1^*(\Theta)$ and $-nh_2^*(\Theta)$, respectively. Using Newton-Raphson method, we find $\hat{\Theta} = 10.610, \hat{\Theta}_1^* = 10.650$ and $\hat{\Theta}_2^* = 10.700$. Thus, $\hat{E}\{\Theta|\mathcal{D}\} \cong 10.625$ and $\hat{E}\{\Theta^2|\mathcal{D}\} = 113.584$. In table 2, we have a summary of the obtained approximations.

Table 2: Approximations for $E\{\Theta|\mathcal{D}\}$ and $E\{\Theta^2|\mathcal{D}\}$

	$E\{\Theta\|\mathcal{D}\}$	$E\{\Theta^2\|\mathcal{D}\}$
Simpson with $h = 0.5$	10.618	113.433
Gaussian Quadrature (Gauss-Hermite with $n = 12$)	10.619	113.452
Monte Carlo with $M = 100$	10.626	113.483
Tierney and Kadane Approximation	10.625	113.584

Some Important conclusions:

(*i*) We find similar results considering different integration procedures, especially in the approximation of $E\{\Theta|\mathcal{D}\}$ (see table 2).

(*ii*) To apply Gaussian quadrature rules, is is important to find the region of variation of Θ, given a data set. An appropriate transformation could be very important to get good accuracy of the approximations.

(*iii*) Monte Carlo procedure requires the choice of an appropriate "importance density" for each problem depending on a carefully preliminary analysis of the data.

(*iv*) Tierney and Kadane approximations do not depend on sophisticated computational expertize or a very elaborate preliminary abalysis of the data, and only requires maximuns and second derivatives of functions. It is important, to point out that usually, a good reparametrization could improve the accuracy of these approximations, especially for small sample sizes.

4.2 An Application Considering the Feigl and Zelen Survival Time Data

Let us consider the lenkemia survival time data of Feigl and Zelen (1965) considering a concomitant variable WBC, the white blood cell count of a patient. They consider a model in which survival times are assumed to be exponentially distributed with a mean survival time of the form $\Theta_1 e^{\Theta_2 x}$, where x is the natural logarithm of the white blood cell count measured in units of 10000. Thus Θ_1 represents the mean survival time of a patient with a white blood cell count of 10000 and Θ_2 represents the approximate percentage change in mean survival time corresponding to a one percent increase in the white blood cell count. Their sample consisted of patients classified as AG positive or AG negative based on examination of the leukemia cells.

In table 3, we have the data of 17 AG positive patients. We want to find the marginal posterior density for the two year (t = 96 weeks) survival probability of patients with a white blood cell count of 50000, given by

$$\eta = g(\Theta_1, \Theta_2) = exp\left\{-\frac{96}{\Theta_1 5^{\Theta_2}}\right\} \tag{30}$$

Considering the data of table 3, the likelihood function for Θ_1 and Θ_2 is given by

$$l(\Theta_1, \Theta_2|\mathcal{D}) = \Theta_1^{-17} exp\left\{-3.7552\Theta_2 - \frac{1}{\Theta_1}\sum_{i=1}^{17} t_i e^{-\Theta_2 x_i}\right\}. \tag{31}$$

Assuming a noninformative Jeffreys prior density for Θ_1 and Θ_2 given by $\pi(\Theta_1, \Theta_2) \propto 1/\Theta_1$, we find the joint posterior density for Θ_1 and Θ_2,

$$\pi(\Theta_1, \Theta_2|\mathcal{D}) \ \propto \ \Theta_1^{-18} exp\left\{-3.7552\Theta_2 - \frac{1}{\Theta_1}\sum_{i=1}^{17} t_i e^{\Theta_2 x_i}\right\} \tag{32}$$

where $\Theta_1 > 0$ and $-\infty < \Theta_2 < \infty$.

Considering the transformation of variables $\eta = g(\Theta_1, \Theta_2)$ (see (30)) and $\phi_2 = \Theta_2$ (Jacobian is $96/[\eta 5^{\phi_2}(-ln\eta)^2]$), the joint posterior density for η and ϕ_2 is given by

$$\pi(\eta, \phi_2|\mathcal{D}) \ \propto \ \frac{(-ln\eta)^{16}(5^{\phi_2})^{17}}{\eta} \ \times \tag{33}$$

$$\times \ exp\left\{-3.7552\phi_2 \ - \ \frac{(-ln\eta)5^{\phi_2}}{96} \sum_{i=1}^{17} t_i e^{-x_i\phi_2}\right\}$$

where $0 \leq \eta \leq 1$ and $-\infty < \phi_2 < \infty$.

Using Laplace's method, we obtain the marginal posterior density for η given by

$$\widehat{\pi}(\eta|\mathcal{D}) \ \propto \ \frac{(-ln\eta)^{15.5}(5^{\widehat{\phi}_2})^{17}}{\eta\left\{\sum_{i=1}^{17}(5e^{-x_i})^{\widehat{\phi}_2}[ln(5e^{-x_i})]^2 t_i\right\}^{1/2}} \ \times$$

$$\times \ exp\left\{-3.7552\widehat{\phi}_2 - \frac{(-ln\eta)5^{\widehat{\phi}_2}}{96}\sum_{i=1}^{17} t_i e^{-\widehat{\phi}_2 x_i}\right\} \tag{34}$$

Table 3: Feigl and Zelen data (AG positive)

WBC/10000	Survival Times (In weeks)
0.230	65
0.075	156
0.430	100
0.260	134
0.600	16
1.050	108
1.000	121
1.700	4
0.540	39
0.700	143
0.940	56
3.200	26
3.500	22
10.000	1
10.000	1
5.200	5
10.000	65

where $0 \leq \eta \leq 1$, and $\widehat{\phi}_2$ is obtained by using an iterative procedure (e.g., Newton-Rapson) for each value of η, such that,

$$\frac{(ln\eta)}{96}\sum_{i=1}^{17}(5e^{-x_i})^{\widehat{\phi}_2} ln[5e^{-x_i}]t_i + 23.6052 = 0.$$

The mode of the marginal posterior density for η (34) is given by $\tilde{\eta} \cong 0.003$ (see figure 2).

Considering Tierney, Kass and Kadane (1989) approximation for nonlinear functions (see (19)). the marginal posterior density for η is given by

$$\hat{\pi}_{TKK}(\eta/\mathcal{D}) \;\propto\; e^{-nh\left(\widehat{\Theta}_1(\eta),\widehat{\Theta}_2(\eta)\right)}\left\{\frac{det\ \Sigma(\eta)}{det\left[(Dg)'\ \Sigma(\eta)(Dg)\right]}\right\}^{1/2} \tag{35}$$

where $0 \leq \eta \leq 1, \widehat{\Theta}(\eta) = \left(\widehat{\Theta}_1(\eta), \widehat{\Theta}_2(\eta)\right)$ maximizes $-nh(\Theta_1, \Theta_2) = -18ln\Theta_1 - 3.7552\Theta_2 - \frac{1}{\Theta_1}\sum_{i=1}^{17} t_i e^{-\Theta_2 x_i}$ subject to the constraint $\eta = g(\Theta_1, \Theta_2) = exp\left\{-\frac{96}{\Theta_1 5^{\Theta_2}}\right\}$ (or $ln(-ln\eta) = ln96 - ln\Theta_1 - \Theta_2 ln5$), $\Sigma(\eta)$ is the inverse of the Hessian of $nh(\Theta_1, \Theta_2)$ at $\widehat{\Theta}(\eta)$ and Dg is the gradient of g evaluated at $\widehat{\Theta}(\eta)$.

To maximize $-nh(\Theta_1, \Theta_2)$ subject to the constraint $ln(-ln\eta) = ln96 - ln\Theta_1 - \Theta_2 ln5$, we consider the function

$$Q(\Theta_1,\Theta_2) = -nh(\Theta_1,\Theta_2) - \mu\{ln(-ln\eta) - ln96 + ln\Theta_1 + \Theta_2 ln5\} \quad (36)$$

where μ is the Lagragian multiplier.

Using an iterative procedure (e.g., Newton-Raphson), we find $\hat{\phi}_2(\eta)$ for each value of η, such that

$$f(\hat{\Theta}_2) = -0.03966\sum_{i=1}^{17} t_i x_i e^{-\hat{\Theta}_2 x_i} + 0.06383\sum_{i=1}^{17} t_i e^{-\hat{\Theta}_2 x_i} + \frac{96}{(ln\eta)5^{\hat{\Theta}_2}} = 0 \quad (37)$$

Also,

$$\hat{\Theta}_1(\eta) = 96/\left[(-ln\eta)5^{\hat{\Theta}_2}\right] \quad (38)$$

The inverse of the Hessian matrix for $nh(\Theta_1,\Theta_2)$ at $\hat{\Theta}(\eta)$ is given by

$$\Sigma(\eta) = \begin{pmatrix} a_3/\Delta & -a_2/\Delta \\ -a_2/\Delta & a_1/\Delta \end{pmatrix} \quad (39)$$

where $\Delta = a_1a_3 - a_2^2, a_1 = n\frac{\partial^2 h}{\partial\Theta_1^2}/_{\hat{\Theta}(\eta)} = -\frac{18}{\hat{\Theta}_1^2} + \frac{2}{\hat{\Theta}_1^3}\sum_{i=1}^{17} t_i e^{-\hat{\Theta}_2 x_1}$, $a_2 = n\frac{\partial^2 h}{\partial\Theta_1\partial\Theta_2}/_{\hat{\Theta}(\eta)} = \frac{1}{\hat{\Theta}_1^2}\sum_{i=1}^{17} t_i x_i e^{-\hat{\Theta}_2 x_i}$, and $a_3 = n\frac{\partial^2 h}{\partial\Theta_2^2}/_{\hat{\Theta}(\eta)} = \frac{1}{\hat{\Theta}_1}\sum_{i=1}^{17} t_i x_i^2 e^{-\hat{\Theta}_2 x_i}$.

We are using $(\hat{\Theta}_1,\hat{\Theta}_2)$ rather than $\left(\hat{\Theta}_1(\eta),\hat{\Theta}_2(\eta)\right)$ as a simplification of notation.

From (39), we have $det\ \Sigma(\eta) = \Delta^{-1}, (Dg)'\ \Sigma(\eta)(Dg) = \Delta^{-1}(a_1b_2^2 + a_3b_1^2 - 2a_2b_1b_2)$, where $(Dg)' = (b_1,b_2), b_1 = \frac{\partial g}{\partial\Theta_1}/_{\hat{\Theta}(\eta)} = \frac{\eta(-ln\eta)}{\hat{\Theta}_1}$, and $b_2 = \frac{\partial g}{\partial\Theta_2}/_{\hat{\Theta}(\eta)} = \eta(-ln\eta)ln5$.

Thus, the TKK approximation (35) is given by

$$\hat{\pi}_{TKK}(\eta|D) \propto \frac{\hat{\Theta}_1^{18} exp\left\{-3.7552\hat{\Theta}_2 - \frac{1}{\hat{\Theta}_1}\sum_{i=1}^{17} t_i e^{-\hat{\Theta}_2 x_i}\right\}}{\{a_1b_2^2 + a_3b_1^2 - 2a_2b_1b_2\}^{1/2}} \quad (40)$$

where $0 \le \eta \le 1$. The mode of (40) is given by $\tilde{\eta} \cong 0.004$. (see figure 2).

In figure 2, we observe very similar approximations for the marginal posterior density of η considering (34) and (40). We also have in figure 2, the graph of the marginal posterior density for η obtained by using Monte Carlo procedure with a generated sample of size $M = 150$. There is close agreement among the approximations.

Another possibility is to consider Leonard, Hsu and Tsui (1989) approximation (see (20)). The maximum of $-nh(\Theta_1,\Theta_2)$ given in (35) subject to the constraint $\eta = g(\Theta_1,\Theta_2) =$

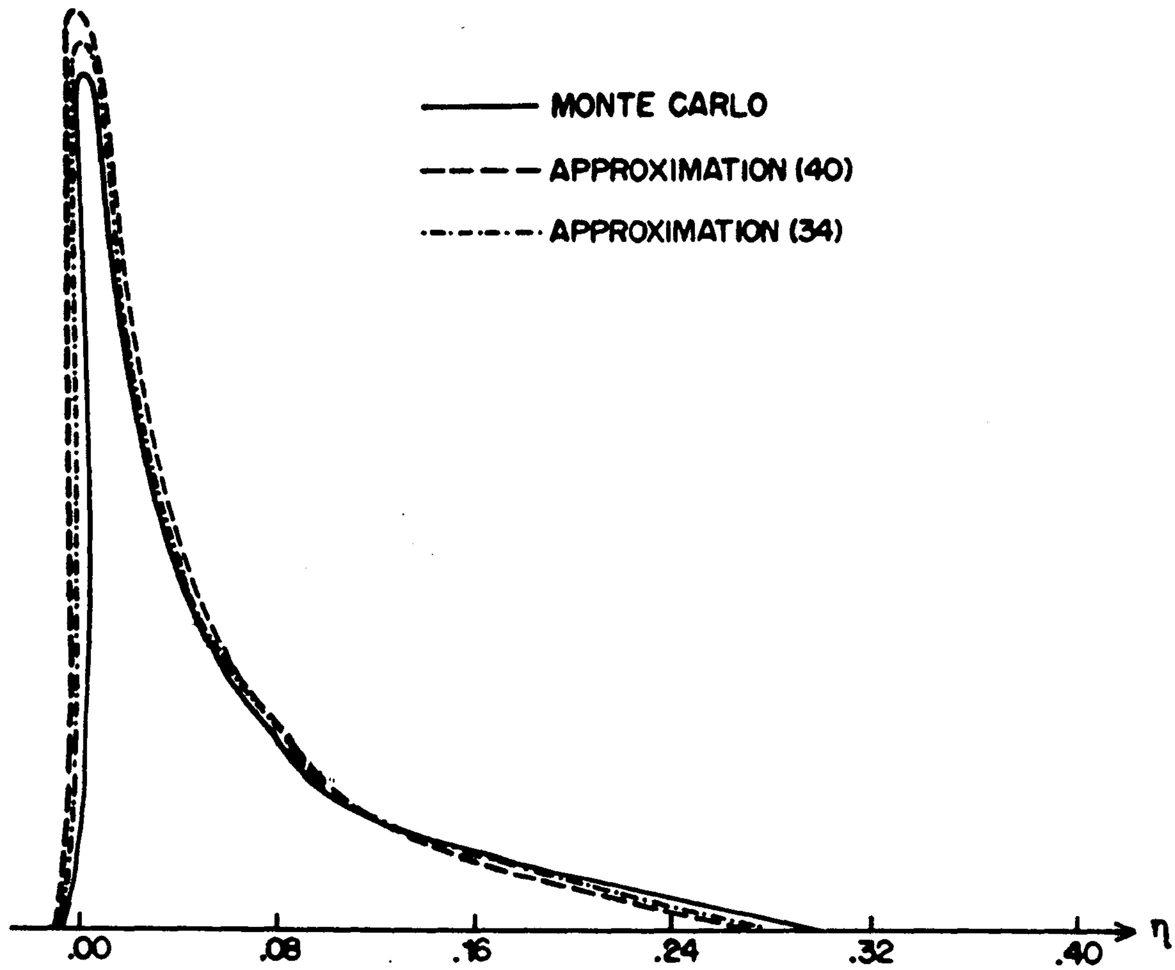

Figure 2: Marginal Posterior Density for η

$exp\left\{\frac{96}{\Theta_1 5^{\Theta_2}}\right\}$ are given in (37) and (38). The gradient of $nh(\Theta_1, \Theta_2)$ at $\widehat{\Theta}(\eta)$ is given by $l'_\eta = (c_1, c_2)$, where

$$c_1 = n\frac{\partial h}{\partial \Theta_1} \mid_{\widehat{\Theta}(\eta)} = \frac{18}{\widehat{\Theta}_1} - \frac{1}{\widehat{\Theta}_1^2} \sum_{i=1}^{17} t_i \epsilon^{-\widehat{\Theta}_2 x_i} \; .$$

and

$$c_2 = n\frac{\partial h}{\partial \Theta_2} \mid_{\widehat{\Theta}(\eta)} = 3.7552 - \frac{1}{\widehat{\Theta}_1} \sum_{i=1}^{17} t_i x_i \epsilon^{\widehat{\Theta}_2 x_i} \; . \tag{41}$$

Also, $\tilde{\Theta}(\eta) = \left(\tilde{\Theta}_1(\eta), \tilde{\Theta}_2(\eta)\right)$, is given (see (20)) by $\hat{\Theta}_1(\eta) = \widehat{\Theta}_1(\eta) - \Delta^{-1}(a_3 c_1 - a_2 c_1)$ and $\hat{\Theta}_2(\eta) = \widehat{\Theta}_2(\eta) - \Delta^{-1}(a_1 c_2 - a_2 c_1)$ where c_1 and c_2 are given in (41), a_1, a_2, a_3 and Δ are given in (39).

Considering a normal approximation $N(d_2, d_1)$ for $f\left(\eta/\tilde{\Theta}(\eta), \not{\Sigma}(\eta)\right)$ (see (20)), where $d_1 = \left(Dg(\tilde{\Theta}(\eta))\right)' \not{\Sigma}(\eta) \left(Dg(\tilde{\Theta}(\eta))\right) = \Delta^{-1}\left(a_1 \tilde{b}_2^2 + a_3 \tilde{b}_1^2 - 2a_2 \tilde{b}_1 \tilde{b}_2\right), \tilde{b}_1 = \partial g/\partial\Theta_1 | \tilde{\Theta}(\eta), \tilde{b}_2 = \partial g/\partial\Theta_2 | \tilde{\Theta}(\eta)$ and $d_2 = exp\left\{-\frac{96}{\tilde{\Theta}_1 5^{\tilde{\Theta}_2}}\right\}$, and since $l_\eta \not{\Sigma}(\eta) l_\eta = \Delta^{-1}(a_1 c_2^2 + a_3 c_1^2 - 2a_2 c_1 c_2)$, we get the LHT approximation,

$$\hat{\pi}_{LHT}(\eta|\mathcal{D}) \propto \frac{exp\left\{-\frac{1}{2d_1}(\eta - d_2)^2\right\}}{(\Delta d_1)^{1/2}\hat{\Theta}_1^{18}} \times$$

$$\times\ exp\left\{-3.7552\hat{\Theta}_2 - \frac{1}{\hat{\Theta}_1}\sum_{i=1}^{17} t_i e^{-\hat{\Theta}_2 x_i} + \Delta^{-1}(a_1c_2^2 + a_3c_1^2 - 2a_2c_1c_2)\right\} \tag{42}$$

where $0 \leq \eta \leq 1$. In figure 3, we have the graph of (42). The mode of (42) is given by $\tilde{\eta} \cong 0.03$.

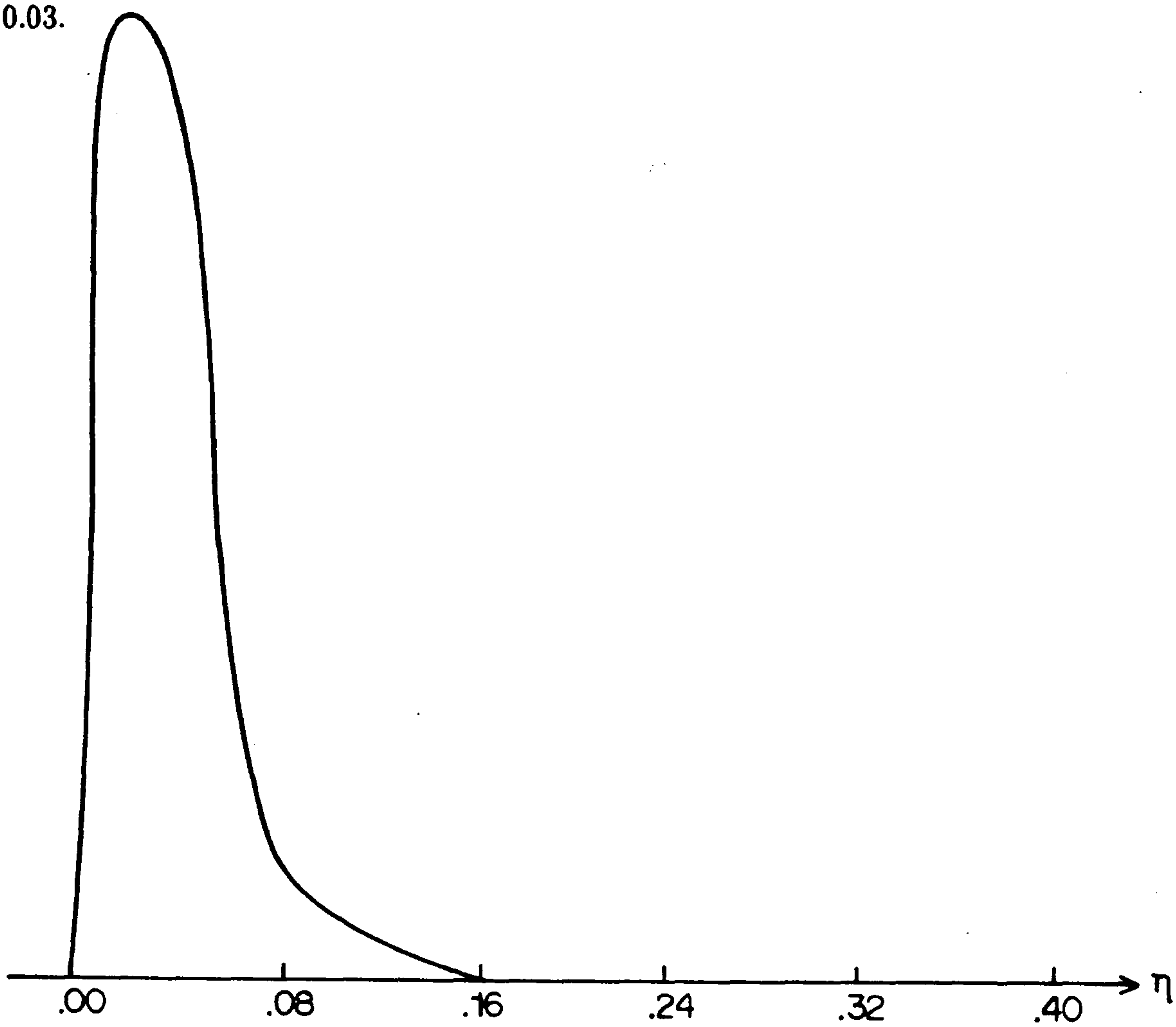

Figure 3: LHT Approximation (42) for the Marginal Posterior Density for η

Some conclusions

(*i*) In (42), the normal approximation for $f\left(\eta/\hat{\Theta}(\eta),\ \hat{\Sigma}(\eta)\right)$ can not be very good, since we are getting a different approximation for the marginal posterior density of η (see figures 2 and 3).

(*ii*) We could consider other choices to approximate $f\left(\eta/\hat{\Theta}(\eta),\ \hat{\Sigma}(\eta)\right)$ in (42) or to consider different parametrizations (e.g., the logit, the Guerrero-Johnson (1982) or the Aranda-Ordaz (1981) transformation, since η is very close to zero).

(*iii*) The accuracy of approximations (34) and (40) are very good in comparison with Monte Carlo procedure.

(*iv*) One great advantage of TKK approximation (40): it is not required the explicit specification of a transformation.

4.3 An Application in the Comparison of two Treatments

The data of table 4, represent failure times, in minutes, for two types of electrical insulation in an experiment in which the insulation was subjected to a continuously increasing voltage stress (data given in Lawless, 1982, p. 138).

Table 4: Electrical Insulation Data

Type A	219.3	79.4	86.0	150.2	21.7	18.5
($n_1 = 12$)	121.9	40.4	147.1	35.1	42.3	48.7
Type B	21.8	70.7	24.4	138.6	151.9	75.3
($n_2 = 12$)	12.3	95.5	98.1	43.2	28.6	46.9

From table 4, we have $n_1 = n_2 = 12$ (number of units in each treatment), $n = n_1 + n_2 = 24, \sum_{i=1}^{n_1} t_{1i} = 1010.7$ and $\sum_{i=1}^{m_2} t_{2i} = 807.3$.

Considering the Feigl and Zelen (1965) model of section 4.2, where the failure times follow the exponential distribution with mean $\Theta_1 e^{\Theta_2 x}$ with $x = 1$ for type A group and $x = 0$ for type B group, and assuming the same noninformative prior $\pi(\Theta_1, \Theta_2) \propto 1/\Theta_1$, the joint posterior density for Θ_1 and Θ_2 is given by

$$\pi(\Theta_1, \Theta_2/\mathcal{D}) \propto \Theta_1^{-25} exp\left\{-12\Theta_2 - \frac{1}{\Theta_1}(1010.7e^{-\hat{\Theta}_2} + 807.3)\right\} \tag{43}$$

where $\Theta_1 > 0$ and $-\infty < \Theta_2 < \infty$.

To find the marginal posterior density for the $t = 100$ minutes reliability $\eta = g(\Theta_1 . \Theta_2) = exp\left\{-\frac{100}{\Theta_1 e^{\Theta_2 x}}\right\}$ where $x = 1$ (type A) or $x = 0$ (type B), we could use one of the existing approximation methods. Considering the transformation of variables $\eta = exp\left\{-\frac{100}{\Theta_1 e^{\Theta_2 x}}\right\}$ and $\phi_2 = \Theta_2$, we use Laplace's method to get

$$\hat{\pi}(\eta|\mathcal{D}) \propto \begin{cases} \eta^{9.107}(-ln\eta)^{11} & \text{for } x = 1 \\ \eta^{7.073}(-ln\eta)^{11} & \text{for } x = 0 \end{cases} \tag{44}$$

where $0 \leq \eta \leq 1$. The mode of (44) is given by $\tilde{\eta} = 0.2988$ for $x = 1$ (type A group) and $\tilde{\eta} = 0.2111$. for $x = 0$ (type B group).

Considering the TKK approximation (19), we get

$$\hat{\pi}_{TKK}(\eta|\mathcal{D}) \propto \begin{cases} \dfrac{\eta^{9.107}(-ln\eta)^{11.5}}{\{10.107(-ln\eta)+1\}^{1/2}} & \text{for } x = 1 \\ \eta^{7.073}(-ln\eta)^{11} & \text{for } x = 0 \end{cases} \tag{45}$$

Observe that both approximations (44) and (45) coincides for $x = 0$ (type B group).

In figures 4 and 5, we have the graphs of the approximate marginal posterior densities for the reliability function at $t = 100$ minutes. We also have in figures 4 and 5, the graphs of the marginal posterior densities of η with $x = 0$ and $x = 1$, considering a numerical procedure based on Gaussian quadrature (Gauss-Hermite with $n = 12$). We observe close agreement among the approximations.

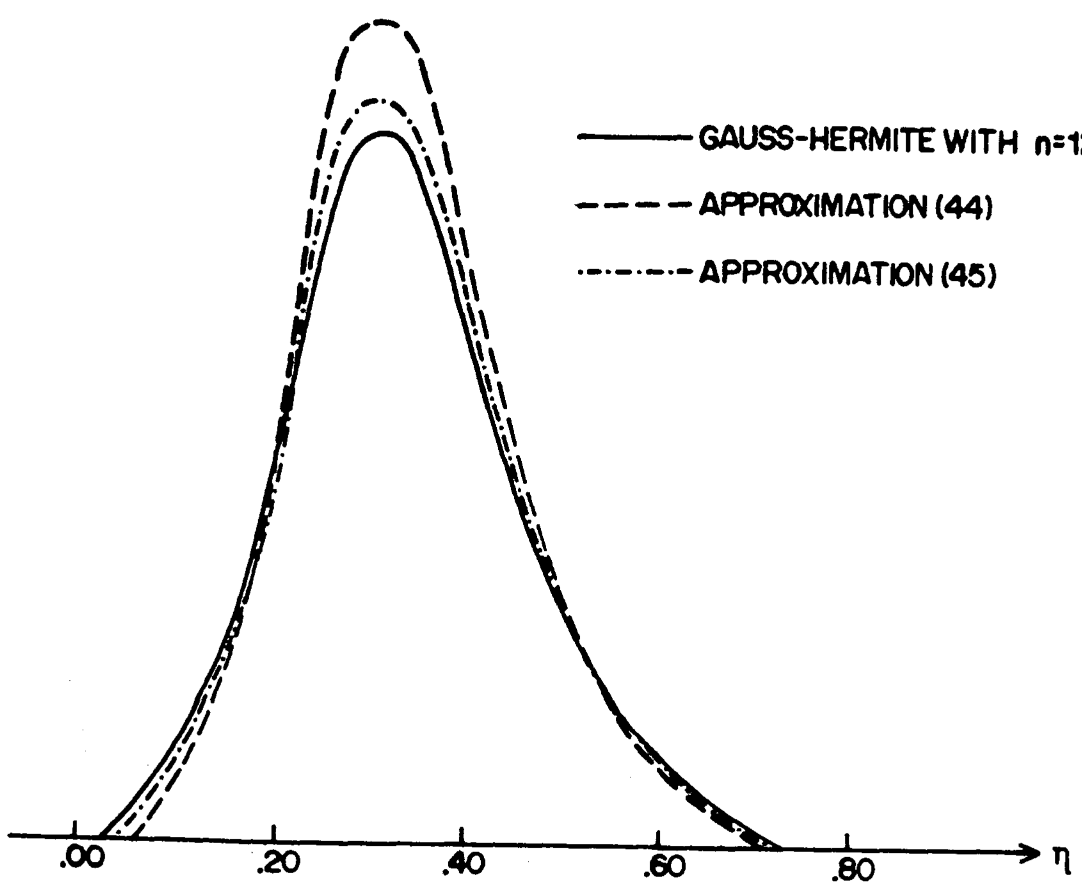

Figure 4: Marginal Posterior Density for η with $x = 1$

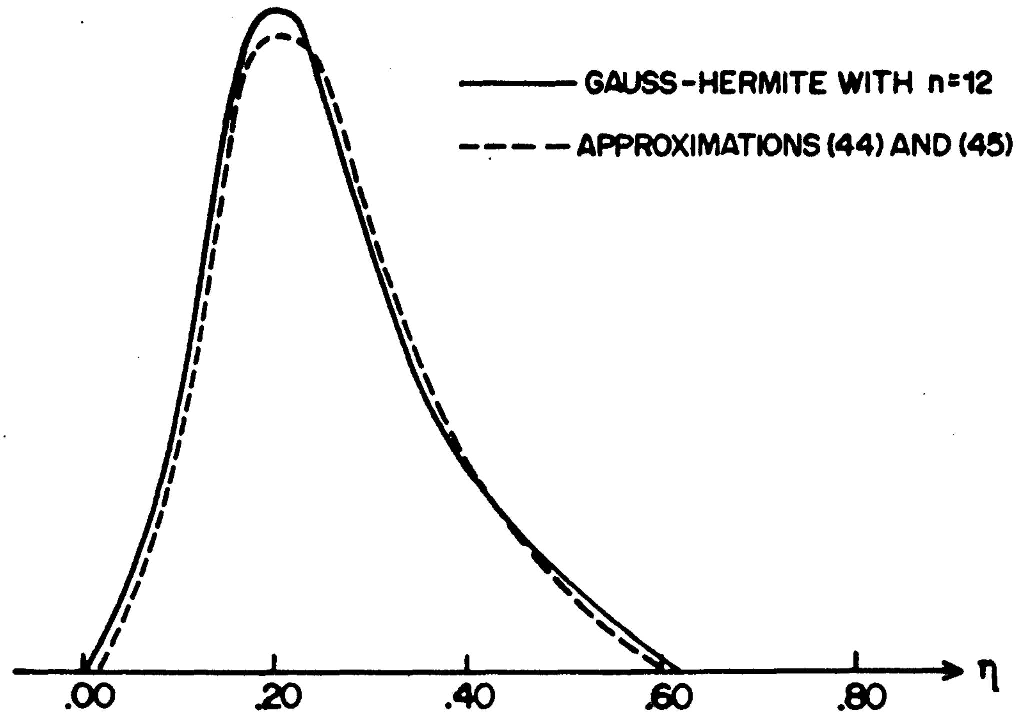

Figure 5: Marginal Posterior Density for η with $x = 0$

Some Conclusions:

(*i*) For $x = 1$ (type A group), we observe better accuracy for the TKK approximation (45) in comparison with Gauss-Hermite procedure ($n = 12$) (see figure 4).

(*ii*) For $x = 0$ (type B group), we observe similar results considering the approximations (44), (45) and the Gauss-Hermite procedure with $n = 12$ (see figure 5).

References

[ABRAMOWITZ, M.; STEGUN, I. (1965)] *Handbook of Mathematical Functions*, Dover.

[ACHCAR, J.A.; SMITH, A.F.M. (1990)] "Aspects of reparametrization in approximate Bayesian inference", *in Essays in honor of George A. Barnard*, ed. J. Hodges, Amsterdam: North - Holland, 431-452.

[ARANDA-ORDAZ, F.J. (1981)] "On two families of transformations to additivity for binary response data", Biometrika, 68, 357-363.

[BOX, G.E.P.; TIAO, G.C. (1973)] *Bayesian inference in statistical analysis*, MA:Addison-Wesley.

[CASELLA, G.; GEORGE, E.I. (1992)] "Explaining the Gibbs sampler", The American Statistician, 46, 3, 167-174.

[CONTE, S.D.; DE BOOR, C. (1965)] *Elementary numerical Analysis: an algorithmic approach*, NY: McGraw-Hill.

[FEIGL, P.; ZELEN, M. (1965)] "Estimation of exponential survival probabilities with concomitant information", Biometrics, 21, 826-837.

[GELFAND, A.E.; SMITH, A.F.M. (1990)] "Sampling-based approaches to calculating marginal densities", Journal of the American Statistical Association, 85, 398-409.

[GEWEKE, J. (1989)] "Bayesian inference in econometric models using Monte Carlo integration", Econometrica, 57, 1317-1339.

[GUERRERO, V.M.; JOHNSON, R.A. (1982)] "Use of the Box-Cox transformation with binary response models", Biometrika, 69, 309-314.

[KASS, R.E.; SLATE, E.H. (1992)] "Reparametrization and diagnostics of posterior nonnormality", in *Bayesian Statistics 4* (eds J.M. Bernardo. J. Berger. A.P. Dawid and A.F.M. Smith), Oxford: Oxford University Press.

[KLOEK, T.; VAN DIJK, H.K. (1978)] "Bayesian estimates of equation system parameters: an application of integration by Monte Carlo", Econometrica, 46, 1-19.

[LAWLESS, J.F. (1982)] *Statistical models and methods for lifetime data*, NY: John Wiley & Sons.

[LEONARD, T.; HSU, J.S.J.; TSUI, K.W. (1989)] "Bayesian marginal inference", Journal of the American Statistical Association, 84, 408, 1051-1058.

[LINDLEY, D.V. (1980)] "Approximate Bayesian methods", *Baycsian Statistics*, proceedings of the first international meeting held in Valencia (eds. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith). Valencia: University press.

[NAYLOR, J.C.; SMITH, A.F.M. (1982)] "Applications of a method for the efficient computation of posterior distribution", Applied Statistics, 31, 214-225.

[TIERNEY, L.; KADANE, J.B. (1986)] "Accurate approximations for posterior moments and marginal densities", Journal of the American Statistical Association, 81, 82-86.

[TIERNEY, L.; KASS, R.E.; KADANE, J.B. (1989a)] "Approximate marginal densities for nonlinear functions", Biometrika, 76, 425-433.

[TIERNEY, L.; KASS, R.E.; KADANE, J.B. (1989b)] "Fully exponential Laplace approximations to expectations and variances of nonpositive functions", Journal of the American Statistical Association, 84, 710-716.

N O T A S   D O   I C M S C

SERIE: Estatistica

Nº 001/93 - ACHCAR, J.A. - Some aspects of reparametrization for statistical models

Nº 002/93 - RODRIGUES, J.; BABA, M.Y. - Bayesian estimation of a simple, regression model with measurement errors

SERIE: Computacao

Nº 001/93 - MORABITO, R.; ARENALES, M.N. - An and/or graph to the container loading problem

Nº 002/93 - NICOLETTI, M.C.; MONARD, M.C. - Herbrand interpretation model and least model within the framework of logic programming

SERIE: Matematica

Nº 001/93 - MICALI, A.; CHIBLOUN, R. - Dimension de Krull des algebres graduees II

Nº 002/93 - MICALI, A.; DIAS, I. - Singularites et algebres symetriques

Nº 003/93 - CARVALHO, A.N.; OLIVEIRA, L.A.F. - Delay-partial differential equations with some large diffusions

Nº 004/93 - CARVALHO, A.N. - Infinite dimensional dynamics described by ordinary differential equations

Nº 005/93 - GALANTE, L.F.; RODRIGUES,H.M. - On bifurcation and symmetry of solutions of symmetric nonlinear equations with odd-harmonic forcings

Nº 006/93 - CROMWELL, P.R.; MARAR, W.L. - Semiregular surfaces with a single triple-point