Computational Statistics
Second course exam, 13 May 2011.

1. Consider a statistical model with a univariate parameter θ which has a continuous distribution. We want to estimate $E[k(\Theta) \mid Y = y]$, i.e., the posterior expectation of $k(\theta)$. Explain how it can be approximated if we set up an evenly spaced grid on the θ-axis and evaluate the prior, the likelihood and $k(\theta)$ on the grid points. Explain how the approximation is related to the midpoint rule for numerical integration.

2. Consider a statistical model with a positive parameter $\theta > 0$ which has a continuous distribution. We want to simulate the posterior distribution. Describe (in pseudocode) a Metropolis–Hastings sampler which is based on the following simulation idea. If $\theta > 0$ is the current state of the chain, then generate a value v from the Exp(1) distribution and propose the state $\theta' = v\,\theta$.

3. Consider the model, where Y_i:s are independent conditionally on the parameter vector (φ, γ, τ) and where

$$\begin{aligned} [Y_i \mid \varphi, \gamma, \tau] &\sim \mathrm{Bin}(k, \varphi), \qquad 1 \le i \le \tau \\ [Y_i \mid \varphi, \gamma, \tau] &\sim \mathrm{Bin}(k, \gamma), \qquad \tau + 1 \le i \le n. \end{aligned}$$

The sample size $k \ge 1$ of the binomial is a known constant. The change point $1 \le \tau \le n-1$ as well as the probability parameters $0 < \varphi < 1$ and $0 < \gamma < 1$ are unknown. The priors are: the discrete uniform distribution on $1, \dots, n-1$ for τ; the beta distribution $\mathrm{Be}(a_1, b_1)$ for φ; and the beta distribution $\mathrm{Be}(a_2, b_2)$ for γ. Here a_1, b_1, a_2, b_2 are fixed values. The parameters are independent in the joint prior.

Present one step (where you update all three of the parameters) for the systematic scan Gibbs sampler, which tries to simulate from the posterior. Denote the state of the chain at the beginning of the iteration by $(\varphi_i, \gamma_i, \tau_i)$, and at its end by $(\varphi_{i+1}, \gamma_{i+1}, \tau_{i+1})$. You should work out all the distributions from which one needs to simulate.

4. We consider a k-variate statistical model whose likelihood function and prior density we have implemented in a computer program. We know that the posterior is roughly of multivariate normal form. We have available an optimization routine, which can calculate the maximum or the minimum point of any multivariate function we care to program. What is more, we have available a numerical differentiation routine which can calculate reliably the gradient (the vector of first derivatives) and the Hessian (the matrix of second derivatives) of any smooth multivariate function at any point on its domain. (If you do not feel comfortable with multivariate calculus, you may assume that $k = 1$.)

a) Given these tools, explain how we can calculate the center b and precision matrix Q of a multivariate normal distribution $N_k(b, Q^{-1})$ which approximates the posterior.

b) Explain how we can calculate an approximation to the marginal likelihood of the model with Laplace approximation.

c) We can check the accuracy of the Laplace approximation by importance sampling, where we draw values from the k-variate $t_k(\nu, b, Q^{-1})$ distribution with degrees of freedom parameter $\nu = 4$, center b and dispersion parameter Q^{-1}. Explain how this importance sampling estimator for the marginal likelihood works. Assume that we have available functions for generating i.i.d. samples and for evaluating the density function for this multivariate t distribution.

Familiar distributions

Binomial distribution $\mathrm{Bin}(n, p)$, n positive integer, $0 \le p \le 1$, has pmf

$$\mathrm{Bin}(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n.$$

Poisson distribution $\mathrm{Poi}(\theta)$ with parameter $\theta > 0$ has pmf

$$\mathrm{Poi}(x \mid \theta) = \mathrm{e}^{-\theta} \frac{\theta^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Beta distribution $\mathrm{Be}(a, b)$ with parameters $a > 0, b > 0$ has pdf

$$\mathrm{Be}(x \mid a, b) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}, \quad 0 < x < 1.$$

$B(a, b)$ is the beta function with arguments a and b,

$$B(a, b) = \int_0^1 u^{a-1} (1-u)^{b-1} \, \mathrm{d}u = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}.$$

Exponential distribution $\mathrm{Exp}(\lambda)$ with rate $\lambda > 0$ has pdf

$$\mathrm{Exp}(x \mid \lambda) = \lambda \, \mathrm{e}^{-\lambda x}, \quad x > 0.$$

Gamma distribution $\mathrm{Gam}(a, b)$ with parameters $a > 0, b > 0$ has pdf

$$\mathrm{Gam}(x \mid a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} \mathrm{e}^{-bx}, \quad x > 0.$$

$\Gamma(a)$ is the gamma function,

$$\Gamma(a) = \int_0^\infty x^{a-1} \mathrm{e}^{-x} \, \mathrm{d}x, \qquad a > 0.$$

It satisfies $\Gamma(a+1) = a\Gamma(a)$ for all $a > 0$, and $\Gamma(1) = 1$, and therefore $\Gamma(n) = (n-1)!$, when $n = 1, 2, 3, \ldots$

Normal distribution $N(\mu, \sigma^2)$ with mean μ and variance $\sigma^2 > 0$ has pdf

$$N(x \mid \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right).$$

Multivariate normal distribution (in d dimensions), $N_d(\mu, Q^{-1})$ with mean $\mu \in \mathbb{R}^d$ and precision matrix Q (a symmetric, positive definite $d \times d$ matrix) has pdf

$$N_d(x \mid \mu, Q^{-1}) = (2\pi)^{-d/2} (\det Q)^{1/2} \exp\left(-\frac{1}{2}(x-\mu)^T Q (x-\mu)\right).$$

(Precision matrix is by definition the inverse of the covariance matrix of the distribution.)