Computational Statistics
5. exercise session, 19.3.2010

1. In this problem the posterior distribution is $\mathrm{Gam}(a, b)$ with $a = 11$ and $b = 5.1$.

a) Calculate 90 % equal tail posterior interval using the quantile function of the gamma distribution (R function `qgamma()`).

a) Calculate 90 % HPD interval.

Hint for part b: You can eliminate the coverage requirement

$$\int_L^U \mathrm{Gam}(\theta \mid a, b)\, \mathrm{d}\theta = 0.9$$

by using the quantile function of the $\mathrm{Gam}(a, b)$ distribution. (If $L = L(t)$ is the t quantile $q(t)$, then $U = U(t)$ must be ...) Then search iteratively for t such that $f(U(t)) - f(L(t))$ is approximately zero, where f is the pdf of the posterior. You can do this either graphically or by using some zero-finding routine such as the R function `uniroot()`.

2. (Using importance sampling to change the prior.) Conditionally on θ, let $Y_i \sim \mathrm{Exp}(\theta)$ independently. Suppose that $n = 10$, and that the observed sum of the Y:s is $s = \sum_{i=1}^n y_i = 5$. (The sum is a sufficient statistic.) We want to calculate the posterior mean, median and 5 % and 95 % quantiles under both of the following priors.

a) The conjugate prior $\mathrm{Gam}(a, b)$, with $a = 1$, $b = 0.1$.

b) The Weibull prior $\mathrm{Weib}(2, 10)$.

Under the conjugate gamma prior of part a, the analysis is simple. The posterior is a gamma distribution (see Sec. 1.4 or Sec. 5.3.2). Its mean can be calculated from a known formula and its quantile function is available (`qgamma()` in R).

To analyze the posterior under the prior of part b, first generate a large sample (say, of size $N_1 = 10^5$) from the posterior of part a, and estimate the posterior mean by importance sampling. Next, estimate the 5 %, 50 % and 95 % quantiles by using SIR: take a smaller sample (say, of size $N_2 = 10^4$) from the initial sample by using the importance weights as probability weights (R function `sample()`). Finally, calculate empirical quantiles (function `quantile()`).

3. Consider normal observations with known mean μ and unknown precision θ, i.e.,

$$Y_i \mid \theta \overset{\text{i.i.d.}}{\sim} N(\mu, \frac{1}{\theta}), \qquad i = 1, \dots, n.$$

Take for θ the imporper prior

$$p(\theta) \propto \frac{1}{\theta}, \qquad \theta > 0.$$

Determine the posterior. Under what conditions is it proper?

4. (Linear regression with a non-conjugate prior). Consider the model

$$\begin{aligned}[Y \mid \beta, \tau] &\sim N_n(X\beta, \frac{1}{\tau} I) \\ p(\beta, \tau) &= N_p(\beta \mid \mu_0, Q_0^{-1}) \ \text{Gam}(\tau \mid a, b).\end{aligned}$$

Here X is a known $n \times p$ model matrix, $\tau > 0$ (a scalar) is the precision parameter of the error distribution, the coefficient vector β has p components, and the parameters β and τ are assumed to be independent in their joint prior distribution. (This is **not** any standard conjugate prior.) An alternative (and perhaps more familiar) way of writing the likelihood of the model is to state that

$$Y_i = x_i^T \beta + \epsilon_i, \quad i = 1, \dots, n,$$

where $\epsilon_i \sim N(0, \sigma^2)$ independently of each other and independently of β, the error variance $\sigma^2 \equiv 1/\tau$, and $x_i^T = X(i, :)$ is the ith row vector of the model matrix.

a) Write the joint density $p(y, \beta, \tau)$ (including all the normalizing constants).

b) Show that the full conditional $p(\tau \mid \beta, y)$ is a gamma density $\text{Gam}(\tau \mid a_1, b_1)$, and give formulas for the hyperparameters a_1 and b_1. (Recall that $\det(sA) = s^n \det(A)$ whenever s is a scalar and A is a $n \times n$ matrix.)

5. We continue with the linear regression model of the previous problem. Show that the full conditional $p(\beta \mid \tau, y)$ is a normal distribution $N(\mu_1, Q_1^{-1})$, and give formulas for μ_1 and Q_1.