Homework Set
AMS 207   Spring 2018
Wyara Vanesa Moura e Silva

1. Obtain the mean and the variance for the beta-binomial distribution. Show that it tackles the overdispersion problem. Hint: use the formulas for conditional expectations and variances.

2. Obtain the Laplace approximation for the posterior expection of logit(mu) and log(tau) in the cancer mortality rate example (data available from the package LearnBayes). The failure time of a pump follows a two-parameter exponential distribution, f(y—b,m) = 1/b exp(-(y-m)/b, when y¿=m.

3. Obtain the likelihood for b and m based on an i.i.d. sample of size n

4. Consider a suitable transformation that maps the parameters b and m to the plane

5. Consider a sample of 8 pumps, where all pumps failed at some time. The smallest failure time was 23721 minutes and the total testing time for all pumps was 15962989 minutes. Assume a sensible prior for the transformed parameters and explore the contours of the posterior distribution.

6. Find a normal approximation to the posterior distribution of the transformed parameters

7. Use rejection sampling and SIR to approximate the posterior distribution. Compare. Use importance sampling as well as a Laplace approximation to estimate the posterior mean and variance of the transformed parameters.

8. Define the reliability at time t_0 as R(t_0) = $\exp(-(t_0 - m)/b)$. Describe the posterior moments and the posterior dsitribution of R(10^6).

9. Problems 3.1,3. 2, 3.5; 4.1, 4.2

10. Problems 5.7,5.8,5.10, 5.12

11. Repeat the SAT example with: Direct sampling from the posterior Gibbs sampling from the posterior Abrams and Sansó '98 approximations for the posterior moments Problems 5.14, 5.15

12. Write the Bayes factor, BIC, DIC and Gelfand and Ghost criterion to compare a model where n observations are assumed to be sampled with a poisson distribution with a gamma prior, to a model where the observations are sampled from a binomial distribution, with a fixed, large, number of trials and beta prior for the probability of success.

13. Consider the data on deaths by horsekicks in the Prussian army. Fit the data using the two different models.

14. Perform a prior sensitivity analysis.

15. Present a model comparison analysis using the criteria mentioned above.

16. Consider the SAT example, use the DIC to compare the models with no pooling, total pooling and partial pooling based on a hierarchical model with unknown variance.

17. Problems 6.2,6.6; 7.4, 7.5

18. Problems 8.11, 8.14

19. For each of the examples considered in class regarding the censored and truncated weights data develop an approach based on MCMC with auxiliary variables and write the full conditionals.

20. Consider the following data regarding the heights in inches of male students at a college. First interval: less than 66, counts 14; second interval 66 to 68, counts 30; third interval 68 to 70, counts 49; fourth interval 70 to 72, counts 70; fifth interval 72 to 74, counts 33; sixth interval greater than 74, counts 15. Assume that the height of students is normally distributed, and assume a non-informative distribution for the parameters of the normal. Use Metroplis Hastings to estimate the parameters of the normal using a multinomial likelihood. Introduce latent variables and use Gibbs sampling to do the estimation. Compare.

21. Find the marginal distribution of the regression coefficients in a linear normal model. Show that the posterior predictive distribution of a new observation corresponding to a regression model is a student distribution as indicated in the slides.

22. Consider the data available as "birthweight" form the package "LearnBayes". Fit a linear regression that considers age and gender as explanatory variables for birth weight. Describe the posterior distribution of the regression parameters using a sample-based approach. Explore the predictive posterior distribution for the birth weight of children in the following four cases: (a) 36 week female/male; (b) 40 week female/male. Compare.

23. Show that the posterior distribution of the regression parameters using an informative prior is the same when the quadratic forms are completed directly and when the prior is considered as additional data.

24. Consider a normal linear model where the errors have a covariance matrix that is a multiple of the identity. Is it possible to obtain a conjugate prior (informative) for the regression coefficients and the variance? If so, find the posterior distribution.

25. Obtain the expressions for the marginal distributions of the data and the Bayes factors for linear models using g-priors.

26. Problems 14.10.1 and 14.10.10

27. Perform a re-analysis of the cancer mortality data using Gibbs sampling introducing appropriate latent variables.

28. Show that a Beta-Binomial likelihood tackles the overdispersion problem in models for Poisson count data.

29. Write the full conditionals for a robust regression where the errors are assumed to correspond to a student distribution with known degrees of freedom.

30. Consider data corresponding to a mixture of M exponential densities. Consider appropriate conjugate priors for the parameters of the model. Find the full conditionals needed to explore the posterior distribution using a Gibbs sampler.

31. Problems 13.5, 13.6, 13.9

32. Use the EM algorithm to estimate the mode of the posterior distribution of the location and scale parameters in a model where the observations follow a student distribution with fixed degrees of freedom and the prior is non-informative.