1 Exercises and Complements

In this section we give some further worked examples on priors, likelihoods and posterior distributions and ways to study them and we suggest some exercises. T

1.0.1 Bayesian Bootstrap

Exercise: The Bayesian bootstrap for the mean can be applied by simulating the probabilities in the moment equation

$$\Sigma_{i=1}^{n} p_i x_i - \mu = 0.$$

Under the multinomial model for the data the $\{p_i\}$ can be simulated as the ratio of unit exponential variates to their sum. This leads to

$$\mu \sim \frac{\Sigma g_i x_i}{\Sigma_i g_i}$$

where the $\{g_i\}$ are n independent unit exponentials. In R this can be implemented as

g = rexp(n)
mu=sum(g*x)/sum(g).

Generate some data and simulate from the posterior distribution of μ 10,000 times. Study the resulting distribution graphically. Is it nearly normal? Where is it centered? What is an hpd region? The following commands show an R session implementing this exercise.

```
> n=30;nrep=10000;m=rep(0,nrep)
> x=rgamma(n,0.5,0.1); # mean = 0.5/0.1 = 5
> for(i in 1:nrep){g=rexp(n);m[i]=sum(g*x)/sum(g)}
> summary(x) # look at x
Min. 1st Qu. Median Mean 3rd Qu. Max.
5.438e-05 3.000e-01 1.256e+00 3.148e+00 4.365e+00 1.799e+01
> summary(m) # look at m
Min. 1st Qu. Median Mean 3rd Qu. Max.
1.122 2.601 3.049 3.143 3.604 7.497
> hist(m,nclass=40)  # inspect the posterior
```

1.0.2 Bayesian Bootstrapping the Linear Model – page 144

It's easy to BB the linear model by using the "weights" facility normally supplied with pachaged software. For example, in R the function `lm` – look at it by doing `?lm` — allows for weighted regression in which $\Sigma_{i=1}^{n} w_i e_i^2$ is minimized with respect to variation in β. But the coefficients in such a regression are precisely what is needed for a BB realization. So to do a single BB realization the command

```
g<-rexp(n); b <- lm(y ~ x,weights=g)$coef
```

will do the trick.

Exercise: Find or generate some y, x data with x an $n \times k$ matrix (without the unit vector) and form nrep BB realizations of the β posterior and compare the coefficient distribution with the normal linear model result.

1.0.3 Gibbs Sampler

Exercise: Sample from a bivariate normal distribution with moments

$$\mu = \begin{bmatrix} 1 \\ 3 \end{bmatrix}; \quad \Sigma = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}$$

using the component conditionals

$$\begin{aligned} f(y|x) &= \mu_y + (\sigma_{xy}/\sigma_{xx})(x - \mu_x), \sigma_{yy}(1 - \rho_{xy}^2) \\ f(x|y) &= \mu_x + (\sigma_{xy}/\sigma_{yy})(y - \mu_y), \sigma_{xx}(1 - \rho_{yx}^2). \end{aligned}$$

Hint: Start from, say, $x = \mu_x$, then sample from $y|x$ then sample $x|y$. Repeat. Run the chain through 2000 steps; discard the first 200; plot the output. Are the $X(Y)$ realizations marginally normal?

Exercise: Sample from a normal $n(1, 1)$ distribution truncated on the left at zero using the fact that the distribution function is

$$G(x) = \frac{\Phi(x-1) - \Phi(-1)}{1 - \Phi(-1)}, \quad x > 0$$

and that $G(X)$ is distributed uniformly on $(0, 1)$.

1.0.4 Prediction

Suppose the data follow a normal linear model $y = x'\beta + \varepsilon$ with $\varepsilon \sim n(0, \tau)$. To predict the value of Y when $x = x_f$ a Bayesian would use the predictive distribution

$$p(y_f|x_f, x, y).$$

where x, y are the sample data. But this is equal to

$$\begin{aligned} p(y_f|x_f, x, y) &= \int p(y_f, \beta, \tau|x_f, x, y) d\beta d\tau \\ &= \int p(y_f|\beta, \tau, x_f, x, y) p(\beta, \tau|x_f, x, y) d\beta d\tau. \end{aligned}$$

This is the normal conditional distribution of Y_f averaged with respect to the posterior distribution of β, τ. You can do this integral analytically but its far easier to sample from the predictive distribution. This would be done in 3 steps.

1. Sample τ from its (gamma) posterior.

2. Sample a β from its normal $(b, \tau(X'X))$ distribution given τ. (Use *mvrnorm* in `library(MASS)`).

3. Sample y_f from its normal $x_f'\beta, \tau$ distribution given β, τ.

Exercise: Choose a value x_f at which to predict Y. Then use the β, τ values simulated in the second exercise of section 1 to sample from the predictive distribution of Y when $x = x_f$.

1.0.5 Schennach's Method.

For a description of the method see my paper *Bayesian Quantile Regression* on the course web page (or on mine).

To experiment with it consider a posterior distribution for the mean μ using the single moment

$$g_i = y_i - \mu$$

The Lagrange multiplier satisfies

$$\lambda = \arg\min_{\eta} \Sigma_{i=1}^n \exp\{\eta\gamma_i\}$$

and the posterior density of μ is

$$p(\mu|y) \propto \Pi_{i=1}^n e^{\lambda g_i} / (\Sigma_{i=1}^n e^{\lambda g_i})^n$$

So the steps might be:

1. Generate some data e.g. `y<-rnorm(30,1,1)`.

2. Define a function, e.g. `fun<-function(l){g<-y-mu;return(sum(exp(l*g)))}`

3. To evaluate the posterior at some point mu set mu as, say, `mu<-1.0`; then use the R function nlm as, say `l<-nlm(fun,0)$est` which gives you the minimizing λ for that value of μ. (The 0 is the start value for the iterative minimization).

4. Calculate the log posterior as, say, `g<-y-mu`; `lp<- l*g-n*log(sum(exp(l*g)))`.

5. If you do this repeatedly for different values of mu and store the results in, say, val you can then plot exp(val) to see the posterior.

Two practical points here: You have to watch for overflow in calculating exp(val). And you must choose a sensible set of mu values at which to calculate the posterior. To decide on the latter you might look at mean(y)±3 standard deviations.

Here is a transcript of an R session that compares the Bayesian bootstrap and the Betel posteriors with each other and with the asymptotic normal form:

n<-40;val<-rep(0,100);y<-rnorm(n); # normal data; smallish sample

m<-seq(mean(y)-3/sqrt(n),mean(y)+3/sqrt(n),length=100); # plot the posterior over a reasonable range

> for(i in 1:100){mu<-m[i];g<-y-mu;l<-nlm(fun,0)$est;val[i]<-sum(l*g)-n*log(sum(exp(l*g)))} #evaluate the log BETEL posterior

> d<-m[7]-m[6];s<-sum(d*exp(val));plot(m,exp(val)/s); # normalize and plot the BETEL posterior

> lines(m,dnorm(m,mean(y),1/sqrt(n)),col="red") # superimpose the asymptotic distribution

> abline(v=mean(y)) # indicate the sample mean

> nval<-rep(0,10000);for(i in 1:10000){g<-rexp(n);nval[i]<-sum(y*g)/sum(g)} # 10000 realizations from the BB posterior

> lines(density(nval),col="blue") # superimpose a kernel smooth of these realizations.

Conclusion: for n=40 and normal data BETEL and BB are the same and equal to their asymptotic form.