Exercises II

1. Explain each step in the equations below (also p.13 in lecture text) showing that posterior probability is the same either way: (1) with original prior $P(r)$ & the whole data X_1, X_2 taken at once, or (2) with original prior first applied to compute posterior probability using a part of the data X_1, and then using the obtained posterior as the next prior for r when computing final posterior using remaining data X_2. (Hint: use rules of probability calculus, and conditional independence, found in 'Preliminary material and notations' section, and focus on what is a function of r).

$$P(r \mid X_1, X_2) = \frac{P(X_1, X_2 \mid r)P(r)}{P(X_1, X_2)} = \frac{P(X_2 \mid X_1, r)P(X_1 \mid r)P(r)}{P(X_1, X_2)} = \frac{P(X_2 \mid r)P(X_1 \mid r)P(r)}{P(X_1, X_2)}$$

$$= \frac{P(X_2 \mid r)P(r \mid X_1)P(X_1)}{P(X_2 \mid X_1)P(X_1)} = \frac{P(X_2 \mid r)P(r \mid X_1)}{P(X_2 \mid X_1)} \propto P(X_2 \mid r)P(r \mid X_1)$$

2. Solve (and plot in R) the posterior density of proportion r based on $x = 3$ positives in a sample of $n = 8$ assuming model $x \sim \text{Binom}(n, r)$ and prior $\pi(r) = \text{Uniform}(0, 1)$. Compare posterior density with those obtained with different priors: $\pi(r) = \text{Beta}(0.5, 0.5)$, $\pi(r) = \text{Beta}(\alpha \approx 0, \beta \approx 0)$. Compare posterior modes, posterior means, and sample mean. What if the data had been different, for example $x = 0, n = 8$? For help and options with plot function, type `help(plot)` on command prompt. (Create a sequence of values `p<-seq(0,1,by=0.005)` for plotting and use `plot(p,dbeta(p,a,b),type="l")`). You can find info about distributions from the file in Moodle pages ('Standard distributions'), or even in Wikipedia.

3. In estimating population prevalence (proportion) p of 'bad eggs', show that the posterior distribution of p becomes the same in the following cases: (1) we collect N eggs and observe x of them to be bad. (2) We keep collecting eggs until we find x bad eggs - and the total sample happens to become N. We assume the prior distribution $\pi(p)$ is the same - whatever it is - in both cases.

4. Let $Z \sim \text{N}(0, 1)$ and $Y = (2Z + 1)^3$. Solve the density function of Y and plot in R. (Hint: transformation of variables, p.4 in text, also p.25). Simulate the distribution of Y in R using `Z <- rnorm(10000,0,1)` and plot the density function of Y (approximated as a histogram) using simulated sample of Z for calculating Y. (This example also in BUGS Book p.24).

5. Life times of light bulbs are modeled as $X_i \sim \text{Exp}(\theta)$, so $E(X \mid \theta) = 1/\theta$. One light bulb lasted 5 and one 7 months, and one was still shining after 9 months of use. Solve posterior distribution for parameter θ based on the two observed exact life times using Gamma(α, β)-prior when $\alpha \approx \beta \approx 0$, and compute the posterior probability $P(E(X \mid \theta) > 10 \mid \text{data})$ using R-function for computing the needed cumulative probability for this (hint: help(pgamma)). Use Monte Carlo simulation to compute the same (hint: help(rgamma)). Then solve the posterior using full data (i.e. with 'full likelihood' including all observations, not just the two exactly observed times). What if $T > 9$ had been the only observed data? What results you would get then?