Exercises IV

1. To say 'variables are exchangeable' is equivalent to saying that 'they are *conditionally independent*, given a parameter p'. They are not unconditionally independent, i.e. covariance $Cov(X_i, X_j) > 0$. Show this with binary variables X_i which are conditionally independently distributed as Bernoulli(p), and where p has Beta(α, β)-distribution. Hint: Use the formula for covariance below, similar to the formula of total variance, to show $\text{Cov}(X_i, X_j) > 0$:

$$Cov(X_i, X_j) = Cov(E(X_i \mid p), E(X_j \mid p)) + E(Cov(X_i, X_j \mid p))$$

(This exemplifies the fact that although X_i are independent of each other conditionally, when p is given: $P(X_i \mid p, X_j, j \neq i) = P(X_i \mid p)$, they are not independent when p is unknown. Then, for predicting next X_i we can obviously learn from observing previous X_j. Only **if** p would be perfectly known, observed data would have no role in our predictions).

2. Bayes factor measures the *change* between prior odds and posterior odds, due to data. Posterior probability simply tells the current degree of belief, given all data and priors we had. Assume the model for X as Normal distribution $\text{N}(\mu, 1)$. A null hypothesis is $H_0 : \mu = 0$, versus alternative hypothesis $H_1 : \mu = 2$. Given that we observed $X = 1$, what is Bayes factor? Assume priors $q_0 = P(\mu = 0) = 0.5 = P(\mu = 2) = q_1$, what is posterior probability of each hypothesis? How large should X increase before alternative hypothesis has posterior probability over 90%? What the Bayes factor would then be?

3. Assume variables X_i have Poisson(λ) model, and that the prior of λ is Gamma(α, β). In lecture notes, the posterior distribution of λ was calculated in the case of single observation X. Compute the posterior when we have $X_1, \ldots, X_n$ observations.

4. Solve the posterior predictive distribution for exponential model $\pi(X) = \theta \exp(-\theta X)$, assuming the posterior distribution $\pi(\theta \mid \text{data})$ is represented by Gamma(α, β)-density: $(\beta^\alpha / \Gamma(\alpha))\theta^{\alpha-1} \exp(-\beta\theta)$. In other words, calculate $\int_0^\infty \pi(X \mid \theta)\pi(\theta \mid \alpha, \beta)\mathbf{d}\theta$ with these functions. (In the manipulations, you may need this result of *gamma-functions*: $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$).

5. Intro to Monte Carlo: posterior distributions can be approximated by drawing large number of random draws from the distribution, then drawing histograms. Assume the posterior of p is Beta($x + 1, n - x + 1$), and choose some values for n, x. A sample of 1000 draws is generated in R by `p <- rbeta(1000,x+1,n-x+1)` . With this sample, we could easily visualize the distribution of any transformation $\sqrt{p}$ or p^2, etc, by plotting the histogram of this from the sample, e.g. `hist(sqrt(p))` or `hist(p^2)`. Try this in the computer class with R. **Optional:** if you still have time, open the OpenBUGS-software, locate the menu bar, and click `File`, and `New` to open a window. Write the following lines there

```
model{
p ~ dbeta(alpha,beta); alpha <- x+1; beta <- n-x+1
x <- 7; n <- 20
p2 <- p*p
}
```

Then click `Model` and within that `Specification` to open 'specification tool window'. Click `check model`, then `compile`, then `gen inits`. Click again `Model` and within that `Update` to open 'Update Tool window' and click `update`. See iteration counter running. Click `Inference` from the menu bar and within that `Samples` to open 'Sample Monitor Tool window'. Write `p` in the 'node' and click `set`. Do the same for `p2`. Click again `update` from 'Update Tool'. Go back to 'Sample Monitor Tool' and select `p` from the 'node'. Click `density` to plot the approximated density from the generated sample. Select and plot `p2` too.