Assignment 5

1. Let us assume that your measurement data consists of $n = 20$ observations, the average of which is $\bar{x} = 30$ and variance $s^2 = 20$; this is sufficient statistics for the data. A judicious assumption is that measured values are normally distributed. The posterior distribution for the mean μ and precision $\tau \left(= \frac{1}{\sigma^2}\right)$ was found to be of the form

$$\pi(\mu, \tau|x) \propto \tau^{a+\frac{n}{2}-1} \exp\left\{-\frac{\tau}{2}[(n-1)s^2 + n(\bar{x}-\mu)^2] - \frac{d}{2}(\mu - c)^2 - b\tau\right\}.$$

(Lecture 5, p. 14.)

In the book (Wilkinson, SMSB), the full conditionals for the parameters of the normal distribution are "derived" (in fact, given) (Lecture 5, p. 17):

$$\tau|\mu, x \sim Ga\left(a + \frac{n}{2},\ b + \frac{1}{2}[(n-1)s^2 + n(\bar{x}-\mu)^2]\right),$$

$$\mu|\tau, x \sim N\left(\frac{cd + n\tau\bar{x}}{n\tau + d}, \frac{1}{n\tau + d}\right).$$

Implement the **Gibbs sampler** algorithm and use it to plot the joint distribution for μ and $\tau = 1/\sigma^2$ and the marginal distributions for μ and τ. Choose parameter values $a = b = 1, c = 3, d = 1/100$. As initial values (called prior means, often used as the starting point for sampling) for mean and precision use $\mu = a/c$ and $\tau = c$. **Do 100 000 iterations (= samples)**. **Plot (i) marginal values of μ vs iteration steps, (ii) marginal distribution for μ, (iii) marginal values of τ vs iteration steps, (iv) marginal distribution for τ, and (v) the joint posterior distribution for μ and τ. Compute and print the mean values for (vi) μ and (vii) τ.** Each task (i) – (vii) is of equal weight in peer grading.
(In other words, produce results similar (but not identical) to those in the SMSB book: Fig. 10.4 (3rd edition, the pdf of the book).)
This should be reasonably straightforward a task with the hint given below, so in order for it to be useful for you, please refer to Lecture 5, especially page 18, and make sure you understand how your algorithm relates to the described Gibbs sampling procedure and the used notation. (No need to report anything about this last part.)

Hint: This would be something to write an algorithm for without any model, but it would be unfair if some found it and others did not. So, take a peek at the function implemented in R in Figure 10.2 in Wilkinson.
Note: There is a 'burn-in' period before convergence is reached, the corresponding values should be discarded before analysis takes place. In general, one should also check the burn-in time. Here, the burn-in time is negligible, so it should not affect your results. You can easily check for this by discarding different numbers of initial steps and see if this has any effect. As said, in this case it should not. (No need to report anything about this part.)
Note for those using random.normal: Please note that, unlike in the standard notation used by the book, where the normal distribution is defined as $N(\text{mean}, \text{variance})$, python's random.normal is defined as random.normal(mean, standard deviation), so take care of using correct parameter values in order for you to belong to the lucky FF group having identical plots.
Note for those using random.gamma: (The same sort of note as for random.normal.) As all the assignments, this one also uses the book's and lecture notes' notation $Ga(\alpha, \beta)$. If you

are using python's random.gamma, then take note that this function is defined as $Ga(k, \theta)$, where $k = \alpha$ and $\theta = 1/\beta$. (You can either write the gamma function without using library routines. That's short and easy, but you are also allowed to use the library gamma here.)