In class, we have demonstrated Gibbs sampling for a two-parameter Normal model, where both $\mu$ and $\sigma^2$ are unknown. In fact, the Gibbs sampling algorithm works for any two-parameter model, or multi-parameter model when the number of parameters is more than one. We will see another one here.

Recall the Gamma-Poisson conjugate model, where the sampling model is $Y_1, \cdots, Y_n|\theta \overset{i.i.d.}{\sim}$ Poisson($\theta$), and the conjugate prior for $\theta$ is $\theta \sim$ Gamma($a, b$). We know that the posterior distribution of $\theta$ is $\theta \mid \mathbf{y} \sim \text{Gamma}(a + \sum_{i=1}^n y_i, b + n)$.

In previous work, we have treated $a$ and $b$ as known/chosen ahead of time by the researcher (you). Now, suppose that $a$ is known but $b$ is unknown. In this case, we will treat $b$ as an unknown parameter. So it needs a prior. Let's use the following prior for $b$: $b \sim \text{Gamma}(1, 1)$.

We come to the following statistical model:

$$
\begin{aligned}
Y_1, \cdots, Y_n \mid \theta, b &\overset{\text{iid}}{\sim} \text{Poisson}(\theta) && (1)\\
\theta \mid b &\sim \text{Gamma}(a, b) && (2)\\
b &\sim \text{Gamma}(1, 1) && (3)
\end{aligned}
$$

*Notice in (1) and (2) that we now condition on $b$ to demonstrate that it's an unknown parameter!*

How will we sample the joint posterior distribution of $(\theta, b)$? Gibbs sampler!

- Step 1: Write out the likelihood function $f(\mathbf{y}|\theta, b)$. Then write the likelihood up to proportionality (with respect to $\theta$ and $b$) using the $\propto$ symbol. Even better: use $\overset{\theta,b}{\propto}$.
- Step 2: Write out the functional form of the joint prior distribution $f(\theta, b)$. Then write the prior up to proportionality (with respect to $\theta$ and $b$) using the $\propto$ (as in Step 1).
- Step 3: Write out the joint posterior distribution $f(\theta, b \mid \mathbf{y})$ up to proportionality using the $\propto$ symbol (as in Step 1).
- Step 4: Obtain the full conditional distribution for each unknown parameter. This means grabbing the "relevant" quantities from Step 3. In this problem, you should be able to spot kernels. If you do, then you know exactly what the distribution should be! *If you don't spot a kernel, then we actually cannot drop things up to proportionality; we would have to carry all the terms with us to ensure that our distribution is a proper PDF that integrates to 1.*
- Step 4.5: Intuition check: how does the full conditional distribution for $\theta$ you found in Step 4 compare to the posterior distribution for $\theta$ when we treated $b$ as fixed? Why does this "make sense"?
- Step 5: Code your Gibbs sampler! *You'll finish this on the homework!*