Kelvin Niu

Homework 5

Dunson: y = 3, y ~ Pois(θ)
Larry: x = 6, x ~ Pois(θγ)

Question 1

I chose a Gamma(1, 1) prior for γ, so that E[γ] = 1, which does not favor either group. For θ, I chose a Gamma(4, 1) prior, so that E[θ] = 4, which seems like a reasonable average number of papers published by a group in a year.
θ ~ Gamma(a, b) = Gamma(4, 1)
γ ~ Gamma(c, d) = Gamma(1, 1)

Question 2

For "perfect" sampling, I ran 1000 draws where I first obtained γ from π(γ | y, x) and then used this value of γ to obtain θ from π(θ | γ, y, x) where
π(γ | y, x) = Gamma(c + x, d) = Gamma(7, 1)
π(θ | γ, y, x) = Gamma(a + x + y, b + 1 + γ) = Gamma(13, 2 + γ)

For Gibbs sampling, I first set an initial value of $\theta^{(0)}$ = 4. Then I ran 1000 iterations where I used the value of $\theta^{(t-1)}$ to get $\gamma^{(t)}$ from π(γ | θ, y, x) and then used this value of $\gamma^{(t)}$ to get $\theta^{(t)}$ from π(θ | γ, y, x) where
π(γ | θ, y, x) = Gamma(c + x, d + θ) = Gamma(7, 1 + θ)
π(θ | γ, y, x) = Gamma(a + y + x, b + 1 + γ) = Gamma(13, 2 + γ)
The trace plots for θ and γ both appeared to stabilize very quickly, perhaps because the $\theta^{(0)}$ was chosen close to the resulting mean estimate, thus no burn-in values were thrown out.

Question 3

Table: Comparing the posterior estimates of θ and γ from "perfect" and Gibbs sampling

	θ		γ	
	Mean	95% CI	Mean	95% CI
"Perfect"	1.5754	[0.6587, 3.0874]	7.0525	[2.9725, 12.8292]
Gibbs	3.8063	[1.7520, 6.7432]	1.5781	[0.5486, 3.3840]

Question 4

The results from "perfect" sampling and Gibbs sampling are very different when theoretically they should be similar. Upon closer examination, we see that we do not have the actual distribution of γ | y, x for "perfect" sampling. The actual distribution of γ | y, x is

$$P(\gamma \mid y, x) = \int P(\theta, \gamma \mid y, x) d\theta$$

$$= \frac{b^a}{\Gamma(a)} \frac{d^c}{\Gamma(c)} \frac{1}{y!\,x!} \gamma^{c+x-1} e^{-d\gamma} \int_0^\infty \theta^{a+y+x-1} e^{-(b+1+\gamma)\theta} d\theta$$

$$= \frac{b^a}{\Gamma(a)} \frac{d^c}{\Gamma(c)} \frac{1}{y!\,x!} \gamma^{c+x-1} e^{-d\gamma} \frac{\Gamma(a+y+x)}{(b+1+\gamma)^{a+y+x}}$$

$$\propto \frac{\gamma^{c+x-1}e^{-d\gamma}}{(b+1+\gamma)^{a+y+x}}$$
$$\neq \pi(\gamma \mid y, x)\ from\ Question\ 2$$
which is not a recognizable distribution and not equal to the proposed $\pi(\gamma \mid y, x)$ in Question 2.

Thus, we must base our conclusions off of the results from Gibbs sampling. Dunson's annual paper output is estimated to be 3.81 with a 95% credible interval of [1.75, 6.74]. The multiplicative factor γ (how many times more productive Larry is than Dunson) is estimated to be 1.58. However, looking at the 95% credible interval for γ, we see that this interval contains values both less than 1 and greater than 1, thus we cannot conclude strongly that Larry is more productive than Dunson in producing papers. This makes sense because we only have one data point for each person. However, there is still some evidence if we estimate the posterior probability of Larry being more productive than Dunson from the samples:
$P(\gamma > 1 \mid y, x) \approx 0.775$

<u>Question 5</u>
Dunson: $y = 3$, $y \sim \text{Pois}(\theta)$
Larry: $x = 6$, $x \sim \text{Pois}(\theta\gamma)$
Katherine: $z = 2$, $z \sim \text{Pois}(\theta\psi)$
Again, I will choose a Gamma(4, 1) prior for θ because $E[\theta] = 4$ papers seems like a reasonable average number of papers published by a group in the year. In order not to favor any group, I will choose a Gamma(1, 1) prior for both γ and ψ so that $E[\gamma] = E[\psi] = 1$.
$\theta \sim \text{Gamma}(a, b) = \text{Gamma}(4, 1)$
$\gamma \sim \text{Gamma}(c, d) = \text{Gamma}(1, 1)$
$\psi \sim \text{Gamma}(c, d) = \text{Gamma}(1, 1)$
For Gibbs sampling, I first set an initial value of $\theta^{(0)} = 4$. Then I ran 1000 iterations where I used the value of $\theta^{(t-1)}$ to get $\gamma^{(t)}$ and $\psi^{(t)}$ from $\pi(\gamma \mid \psi, \theta, y, x, z)$ and $\pi(\psi \mid \gamma, \theta, y, x, z)$, respectively, and then used these values of $\gamma^{(t)}$ and $\psi^{(t)}$ to get $\theta^{(t)}$ from $\pi(\theta \mid \gamma, \psi, y, x, z)$ where
$\pi(\gamma \mid \psi, \theta, y, x, z) = \text{Gamma}(c + x, d + \theta) = \text{Gamma}(7, 1 + \theta)$
$\pi(\psi \mid \gamma, \theta, y, x, z) = \text{Gamma}(c + z, d + \theta) = \text{Gamma}(3, 1 + \theta)$
$\pi(\theta \mid \gamma, \psi, y, x, z) = \text{Gamma}(a + y + x + z, b + 1 + \gamma + \psi) = \text{Gamma}(15, 2 + \gamma + \psi)$

Table: Posterior estimates of θ, γ, and ψ from Gibbs sampling

	θ		γ		ψ	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
Gibbs	3.6805	[1.8149, 6.2566]	1.5941	[0.5775, 3.4614]	0.7020	[0.1298, 1.8928]

Dunson's annual paper output is estimated to be 3.68 with a 95% credible interval of [1.81, 6.26]. The multiplicative factor γ (how many times more productive Larry is than Dunson) is estimated to be 1.59 and the multiplicative factor ψ (how many times more productive Katherine is than Dunson) is estimated to be 0.70. However, like before, the 95% credible intervals of γ and ψ both contain values less than 1 and greater than 1, thus there is not enough evidence to conclude strongly that Larry is more productive than Dunson or that Katherine is less productive than Dunson. In addition, the intervals for γ and ψ do not overlap, so we cannot conclude strongly that Larry is more productive than Katherine. However, if we estimate the posterior probabilities from the samples, we see that there is still some evidence:

The posterior probability of Larry being more productive than Dunson is approximately $P(\gamma > 1 \mid y, x, z) \approx 0.753$

The posterior probability of Katherine being less productive than Dunson is approximately $P(\psi < 1 \mid y, x, z) \approx 0.802$

The posterior probability of Larry being more productive than Katherine is approximately $P(\gamma > \psi \mid y, x, z) \approx 0.896$