Kelvin Niu

Question 1

$y_i \mid c_i=1 \sim LN(\mu_1, \sigma_1^2)$

$y_i \mid c_i=0 \sim LN(\mu_2, \sigma_2^2)$

$c_i = \begin{cases} 1, & \text{if weekday} \\ 0, & \text{if weekend} \end{cases}$

$c_i \sim Bern(p)$

$p \sim Beta(5, 2)$

$\mu_2 \sim N(m, v)$

$\mu_1 = \mu_2 + \varepsilon$

$\varepsilon \sim N(\mu_\varepsilon, \sigma_\varepsilon^2) I(\varepsilon > 0)$

$\sigma_1^{-2} \sim Ga(a, b)$

$\sigma_2^{-2} \sim Ga(a, b)$

$$\pi(c, p, \mu_2, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2} \mid y) \propto \prod_{i=1}^{n} \left[LN(y_i; \mu_2+\varepsilon, \sigma_1^2)^{c_i} LN(y_i; \mu_2, \sigma_2^2)^{1-c_i} p^{c_i} (1-p)^{1-c_i} \right]$$

$$\cdot\, Beta(p; 5, 2) N(\mu_2; m, v) N(\varepsilon; \mu_\varepsilon, \sigma_\varepsilon^2) I(\varepsilon > 0) (\sigma_1^{-2})^{a-1} e^{-b\sigma_1^{-2}} (\sigma_2^{-2})^{a-1} e^{-b\sigma_2^{-2}}$$

$$\pi(c_i \mid c_{(-i)}, p, \mu_2, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y) \propto \left[p\, LN(y_i; \mu_2+\varepsilon, \sigma_1^2) \right]^{c_i} \left[(1-p) LN(y_i; \mu_2, \sigma_2^2) \right]^{1-c_i}$$

$$P(c_i = 1 \mid c_{(-i)}, p, \mu_2, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y) = \frac{p\, LN(y_i; \mu_2+\varepsilon, \sigma_1^2)}{p\, LN(y_i; \mu_2+\varepsilon, \sigma_1^2) + (1-p) LN(y_i; \mu_2, \sigma_2^2)}$$

$$\boxed{c_i \mid c_{(-i)}, p, \mu_2, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y \sim Bern\left(\frac{p\, LN(y_i; \mu_2+\varepsilon, \sigma_1^2)}{p\, LN(y_i; \mu_2+\varepsilon, \sigma_1^2) + (1-p) LN(y_i; \mu_2, \sigma_2^2)} \right)}$$

$$\pi(p \mid c, \mu_2, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y) \propto p^{\sum_{i=1}^n c_i} (1-p)^{n - \sum_{i=1}^n c_i} p^4 (1-p)$$

$$\propto p^{4 + \sum_{i=1}^n c_i} (1-p)^{1 + n - \sum_{i=1}^n c_i}$$

$$\boxed{p \mid c, \mu_2, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y \sim Beta\left(5 + \sum_{i=1}^n c_i,\ 2 + n - \sum_{i=1}^n c_i\right)}$$

$$\pi(\mu_2 \mid c, p, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y) \propto \prod_{c_i=1} \left[LN(y_i; \mu_2+\varepsilon, \sigma_1^2) \right] \prod_{c_i=0} \left[LN(y_i; \mu_2, \sigma_2^2) \right] N(\mu_2; m, v)$$

$$\propto \exp\left[-\frac{1}{2\sigma_1^2} \sum_{c_i=1} (-2\mu_2 \ln y_i + 2\mu_2 \varepsilon + \mu_2^2) \right] \exp\left[-\frac{1}{2\sigma_2^2} \sum_{c_i=0} (-2\mu_2 \ln y_i + \mu_2^2) \right] \exp\left[-\frac{1}{2v} (\mu_2^2 - 2\mu_2 m) \right]$$

$$\propto \exp\left[-\frac{1}{2} \left[\left(\sigma_1^{-2} \sum_{c_i=1} 1 + \sigma_2^{-2} \sum_{c_i=0} 1 + v^{-1} \right) \mu_2^2 - 2 \left(\sigma_1^{-2} \sum_{c_i=1} \ln y_i - \sigma_1^{-2} \varepsilon \sum_{c_i=1} 1 + \sigma_2^{-2} \sum_{c_i=0} \ln y_i + v^{-1} m \right) \mu_2 \right] \right]$$

$$\boxed{\mu_2 \mid c, p, \varepsilon, \sigma_1^{-2}, \sigma_2^{-2}, y \sim N\left(\left(\sigma_1^{-2}\sum_{c_i=1} 1 + \sigma_2^{-2}\sum_{c_i=0} 1 + v^{-1}\right)^{-1}\left(\sigma_1^{-2}\sum_{c_i=1}\ln y_i - \sigma_1^{-2}\varepsilon\sum_{c_i=1} 1 + \sigma_2^{-2}\sum_{c_i=0}\ln y_i + v^{-1}m\right), \left(\sigma_1^{-2}\sum_{c_i=1} 1 + \sigma_2^{-2}\sum_{c_i=0} 1 + v^{-1}\right)^{-1}\right)}$$

$$\pi(\varepsilon \mid c, p, \mu_2, \sigma_1^{-2}, \sigma_2^{-2}, y) \propto \prod_{c_i=1}\left[LN(y_i; \mu_2+\varepsilon, \sigma_1^{2})\right] N(\varepsilon; \mu_\varepsilon, \sigma_\varepsilon^2)\, I(\varepsilon > 0)$$

$$\propto \exp\left[-\frac{1}{2\sigma_1^2}\sum_{c_i=1}(-2\varepsilon\ln y_i + 2\mu_2\varepsilon + \varepsilon^2)\right]\exp\left[-\frac{1}{2\sigma_\varepsilon^2}(\varepsilon^2 - 2\varepsilon\mu_\varepsilon)\right] I(\varepsilon > 0)$$

$$\propto \exp\left[-\frac{1}{2}\left[\left(\sigma_1^{-2}\sum_{c_i=1} 1 + \sigma_\varepsilon^{-2}\right)\varepsilon^2 - 2\left(\sigma_1^{-2}\sum_{c_i=1}\ln y_i - \sigma_1^{-2}\mu_2\sum_{c_i=1} 1 + \sigma_\varepsilon^{-2}\mu_2\right)\varepsilon\right]\right] I(\varepsilon > 0)$$

$$\boxed{\varepsilon \mid c, p, \mu_2, \sigma_1^{-2}, \sigma_2^{-2}, y \sim N\left(\left(\sigma_1^{-2}\sum_{c_i=1} 1 + \sigma_\varepsilon^{-2}\right)^{-1}\left(\sigma_1^{-2}\sum_{c_i=1}\ln y_i - \sigma_1^{-2}\mu_2\sum_{c_i=1} 1 + \sigma_\varepsilon^{-2}\mu_2\right), \left(\sigma_1^{-2}\sum_{c_i=1} 1 + \sigma_\varepsilon^{-2}\right)^{-1}\right) I(\varepsilon > 0)}$$

$$\pi(\sigma_1^{-2} \mid c, p, \mu_2, \varepsilon, \sigma_2^{-2}, y) \propto \prod_{c_i=1}\left[LN(y_i; \mu_2+\varepsilon, \sigma_1^{2})\right](\sigma_1^{-2})^{a-1} e^{-b\sigma_1^{-2}}$$

$$\propto \sigma_1^{-\sum_{c_i=1} 1}\exp\left[-\frac{1}{2\sigma_1^2}\sum_{c_i=1}(\ln y_i - (\mu_2+\varepsilon))^2\right](\sigma_1^{-2})^{a-1} e^{-b\sigma_1^{-2}}$$

$$\propto (\sigma_1^{-2})^{a+\frac{1}{2}\sum_{c_i=1} 1 - 1}\exp\left[-\sigma_1^{-2}\left(b + \frac{1}{2}\sum_{c_i=1}(\ln y_i - (\mu_2+\varepsilon))^2\right)\right]$$

$$\boxed{\sigma_1^{-2} \mid c, p, \mu_2, \varepsilon, \sigma_2^{-2}, y \sim Ga\left(a + \frac{1}{2}\sum_{c_i=1} 1, b + \frac{1}{2}\sum_{c_i=1}(\ln y_i - (\mu_2+\varepsilon))^2\right)}$$

$$\pi(\sigma_2^{-2} \mid c, p, \mu_2, \varepsilon, \sigma_1^{-2}, y) \propto \prod_{c_i=0}\left[LN(y_i; \mu_2, \sigma_2^{2})\right](\sigma_2^{-2})^{a-1} e^{-b\sigma_2^{-2}}$$

$$\propto \sigma_2^{-\sum_{c_i=0} 1}\exp\left[-\frac{1}{2\sigma_2^2}\sum_{c_i=0}(\ln y_i - \mu_2)^2\right](\sigma_2^{-2})^{a-1} e^{-b\sigma_2^{-2}}$$

$$\propto (\sigma_2^{-2})^{a+\frac{1}{2}\sum_{c_i=0} 1 - 1}\exp\left[-\sigma_2^{-2}\left(b + \frac{1}{2}\sum_{c_i=0}(\ln y_i - \mu_2)^2\right)\right]$$

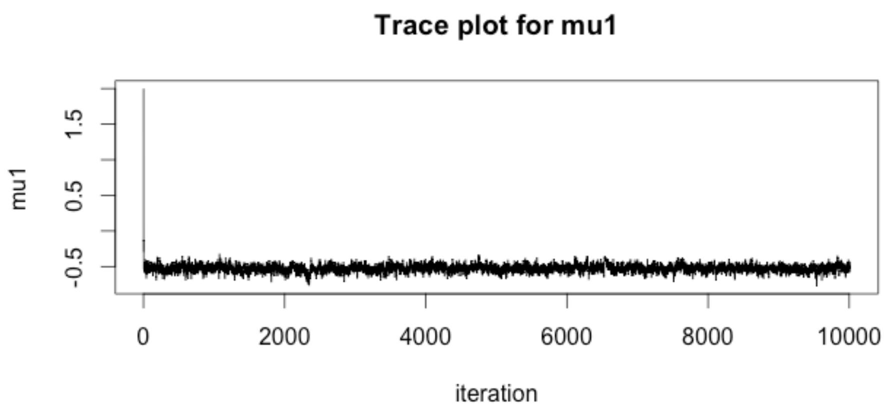
$$\boxed{\sigma_2^{-2} \mid c, p, \mu_2, \varepsilon, \sigma_1^{-2}, y \sim Ga\left(a + \frac{1}{2}\sum_{c_i=0} 1, b + \frac{1}{2}\sum_{c_i=0}(\ln y_i - \mu_2)^2\right)}$$

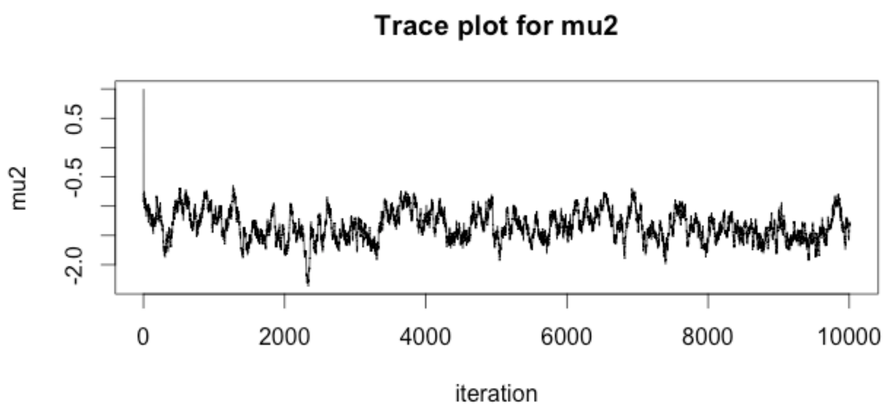
Kelvin Niu

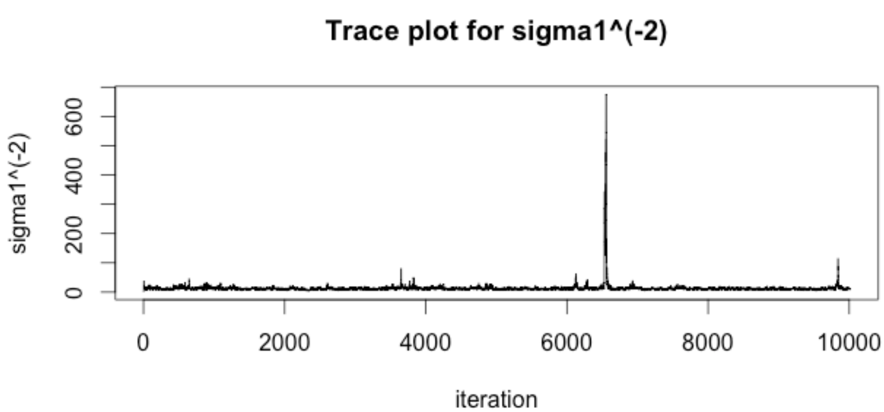
Lab 10

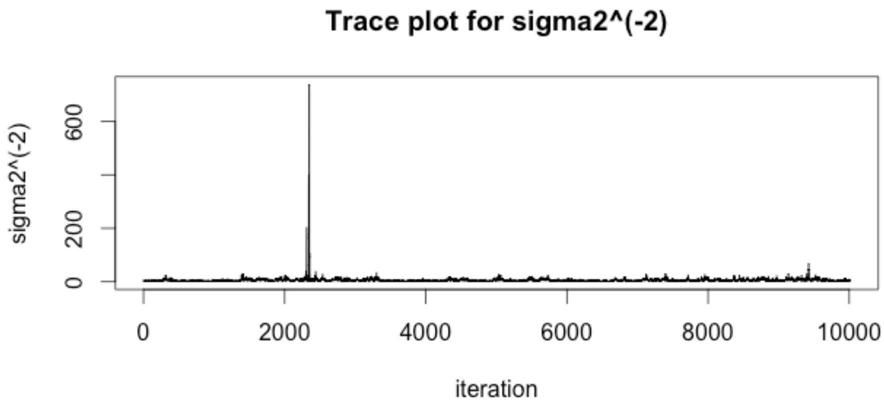
<u>Question 2</u>

Using the full conditionals and weak priors ($m = 1$, $v = 100$, $\mu_\varepsilon = 1$, $\sigma_\varepsilon^2 = 100$, $a = 0.001$, $b = 0.001$), I ran 10,000 iterations of Gibbs sampling. Examining the trace plots, we see that μ_1 and μ_2 appear to converge, but σ_1^{-2} and σ_2^{-2} have some spikes that may affect the posterior estimates of those parameters. I threw out 10 burn-ins.

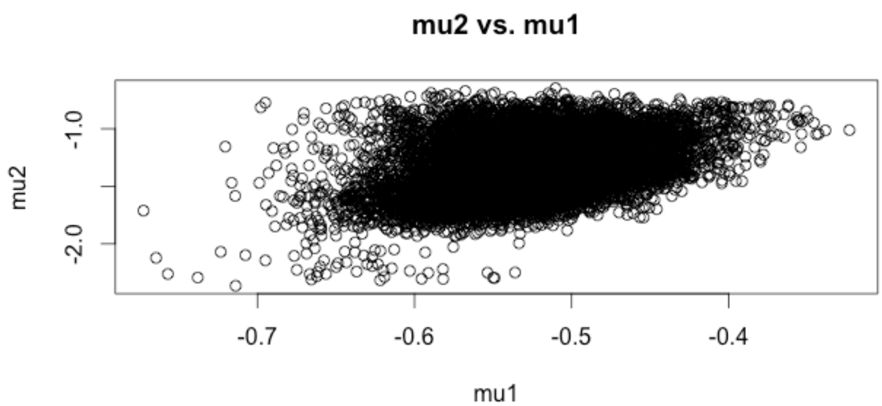








Question 3



Question 4
To estimate the number of days in the sample that are weekdays, I counted the total number of weekdays at each iteration and found the mean and 95% credible interval:
Mean: **74.2965**
95% CI: **[47, 89]**
To estimate the probability that the technician is coming in less often on weekends, I found the proportion of iterations for which weekends made up less than 2/7 of the days: **0.6800**

Question 5
Table: Posterior point estimates and credible intervals for parameters

	Mean	95% CI
μ_1	-0.5262	[-0.6279, -0.4260]
μ_2	-1.3375	[-1.7700, -0.8478]
σ_1^2	0.0769	[0.0431, 0.1496]
σ_2^2	0.1770	[0.0661, 0.5737]

The posterior point estimates for μ_1 and μ_2 are relatively similar to what I got for last week's lab. The estimate for σ_1^2 is close as well but the estimate for σ_2^2 is a little farther away. Since the estimated number of weekdays in this dataset is approximately 74 and last week's dataset had 72 weekdays, perhaps this week's dataset is relatively similar to last week's in origin.