
A Fully Bayesian Model for RNA-seq Data

Will Landau

Department of Statistics
Iowa State University

May 30, 2013

Contents

1	The Model	1
2	Full Conditional Distributions	2
3	Simplifying the Full Conditionals	3
3.1	$p(c_n \mid \cdots)$	3
3.2	$p\left(\frac{1}{\sigma_c^2} \mid \cdots\right)$	4
3.3	$p(\varepsilon_{g,n} \mid \cdots)$	4
3.4	$p\left(\frac{1}{\sigma_g^2} \mid \cdots\right)$	5

1 The Model

Let $y_{g,n}$ be the expression level of gene g ($g = 1, \dots, G$) in library n ($n = 1, \dots, N$). Let $\mu(n, \phi_g, \alpha_g, \delta_g)$ be the function given by:

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{library } n \text{ is in treatment group 1} \\ \phi_g + \delta_g & \text{library } n \text{ is in treatment group 2} \\ \phi_g + \alpha_g & \text{library } n \text{ is in treatment group 3} \end{cases}$$

Then:

$$\begin{aligned} y_{g,n} &\sim \text{Poisson}(\exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \\ c_n &\sim \text{N}(0, \sigma_c^2) \\ \sigma_c &\sim \text{U}(0, \sigma_{c0}) \\ \varepsilon_{g,n} &\sim \text{N}(0, \sigma_g^2) \\ \sigma_g^2 &\sim \text{Inv-Gamma}\left(\text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2}\right) \\ d &\sim \text{U}(0, d_0) \\ \tau^2 &\sim \text{Gamma}(\text{shape} = a_\tau, \text{rate} = b_\tau) \\ \phi_g &\sim \text{N}(\theta_\phi, \sigma_\phi^2) \\ \theta_\phi &\sim \text{N}(0, \gamma_\phi^2) \\ \sigma_\phi &\sim \text{U}(0, \sigma_{\phi 0}) \\ \alpha_g &\sim \pi_\alpha \text{I}(\alpha_g = 0) + (1 - \pi_\alpha) \text{I}(\alpha_g \neq 0) \text{N}(\theta_\alpha, \sigma_\alpha^2) \\ \theta_\alpha &\sim \text{N}(0, \gamma_\alpha^2) \\ \sigma_\alpha &\sim \text{U}(0, \sigma_{\alpha 0}) \\ \pi_\alpha &\sim \text{Beta}(a_\alpha, b_\alpha) \\ \delta_g &\sim \pi_\delta \text{I}(\delta_g = 0) + (1 - \pi_\delta) \text{I}(\delta_g \neq 0) \text{N}(\theta_\delta, \sigma_\delta^2) \\ \theta_\delta &\sim \text{N}(0, \gamma_\delta^2) \\ \sigma_\delta &\sim \text{U}(0, \sigma_{\delta 0}) \\ \pi_\delta &\sim \text{Beta}(a_\delta, b_\delta) \end{aligned}$$

where:

- Independence is implied unless otherwise specified.
- The parameters to the left of the “ \sim ” are implicitly conditioned on the parameters to the right.

2 Full Conditional Distributions

Let $k(n)$ be the treatment group of library n . Then:

$$\begin{aligned}
p(c_n | \dots) &= \prod_{g=1}^G \text{Poisson}(y_{g,n} | \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot N(c_n | 0, \sigma_c^2) \\
p(\sigma_c | \dots) &= \prod_{n=1}^N \prod_{g=1}^G N(c_n | 0, \sigma_c^2) \cdot U(\sigma_c | 0, \sigma_{c0}) \\
p(\varepsilon_{g,n} | \dots) &= \text{Poisson}(y_{g,n} | \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot N(\varepsilon_{g,n} | 0, \sigma_g^2) \\
p(\sigma_g^2 | \dots) &= \prod_{n=1}^N N(\varepsilon_{g,n} | 0, \sigma_g^2) \cdot \text{Inv-Gamma}\left(\sigma_g^2 | \text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2}\right) \\
p(d | \dots) &= \prod_{n=1}^N \prod_{g=1}^G \text{Inv-Gamma}\left(\sigma_g^2 | \text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2}\right) \cdot U(d | 0, d_0) \\
p(\tau^2 | \dots) &= \prod_{n=1}^N \prod_{g=1}^G \text{Inv-Gamma}\left(\sigma_g^2 | \text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2}\right) \cdot \text{Gamma}(\text{shape} = a_\tau, \text{rate} = b_\tau) \\
p(\phi_g | \dots) &= \prod_{n=1}^N \text{Poisson}(y_{g,n} | \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot N(\phi_g | \theta_\phi, \sigma_\phi^2) \\
p(\theta_\phi | \dots) &= \prod_{n=1}^N \prod_{g=1}^G N(\phi_g | \theta_\phi, \sigma_\phi^2) \cdot N(\theta_\phi | 0, \gamma_\phi^2) \\
p(\sigma_\phi | \dots) &= \prod_{n=1}^N \prod_{g=1}^G N(\phi_g | \theta_\phi, \sigma_\phi^2) \cdot U(\sigma_\phi | 0, \sigma_{\phi 0}) \\
p(\alpha_g | \dots) &= \prod_{k(n) \neq 2} \text{Poisson}(y_{g,n} | \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \\
&\quad \times (\pi_\alpha I(\alpha_g = 0) + (1 - \pi_\alpha) I(\alpha_g \neq 0) N(\alpha_g | \theta_\alpha, \sigma_\alpha^2)) \\
p(\theta_\alpha | \dots) &= \prod_{k(n) \neq 2} \prod_{g=1}^G (\pi_\alpha I(\alpha_g = 0) + (1 - \pi_\alpha) I(\alpha_g \neq 0) N(\alpha_g | \theta_\alpha, \sigma_\alpha^2)) \cdot N(\theta_\alpha | 0, \gamma_\alpha^2) \\
p(\sigma_\alpha | \dots) &= \prod_{k(n) \neq 2} \prod_{g=1}^G (\pi_\alpha I(\alpha_g = 0) + (1 - \pi_\alpha) I(\alpha_g \neq 0) N(\alpha_g | \theta_\alpha, \sigma_\alpha^2)) \cdot U(\sigma_\alpha | 0, \sigma_{\alpha 0}) \\
p(\pi_\alpha | \dots) &= \prod_{k(n) \neq 2} \prod_{g=1}^G (\pi_\alpha I(\alpha_g = 0) + (1 - \pi_\alpha) I(\alpha_g \neq 0) N(\alpha_g | \theta_\alpha, \sigma_\alpha^2)) \cdot \text{Beta}(\pi_\alpha | a_\alpha, b_\alpha)
\end{aligned}$$

$$\begin{aligned}
p(\delta_g \mid \cdots) &= \prod_{k(n)=2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \delta_g, \delta_g))) \\
&\quad \times (\pi_\delta \mathbf{I}(\delta_g = 0) + (1 - \pi_\delta) \mathbf{I}(\delta_g \neq 0) \mathbf{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \\
p(\theta_\delta \mid \cdots) &= \prod_{k(n)=2} \prod_{g=1}^G (\pi_\delta \mathbf{I}(\delta_g = 0) + (1 - \pi_\delta) \mathbf{I}(\delta_g \neq 0) \mathbf{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \cdot \mathbf{N}(\theta_\delta \mid 0, \gamma_\delta^2) \\
p(\sigma_\delta \mid \cdots) &= \prod_{k(n)=2} \prod_{g=1}^G (\pi_\delta \mathbf{I}(\delta_g = 0) + (1 - \pi_\delta) \mathbf{I}(\delta_g \neq 0) \mathbf{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \cdot \mathbf{U}(\sigma_\delta \mid 0, \sigma_{\delta 0}) \\
p(\pi_\delta \mid \cdots) &= \prod_{k(n)=2} \prod_{g=1}^G (\pi_\delta \mathbf{I}(\delta_g = 0) + (1 - \pi_\delta) \mathbf{I}(\delta_g \neq 0) \mathbf{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \cdot \text{Beta}(\pi_\delta \mid a_\delta, b_\delta)
\end{aligned}$$

3 Simplifying the Full Conditionals

3.1 $p(c_n \mid \cdots)$

$$p(c_n \mid \cdots) = \prod_{g=1}^G \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot \mathbf{N}(c_n \mid 0, \sigma_c^2)$$

which is nothing immediately recognizable. I will sample from this distribution in a metropolis step using the kernel of this distribution (taking $\lambda_{g,n} = \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))$):

$$\begin{aligned}
&\prod_{g=1}^G \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \exp\left(-\frac{c_n^2}{2\sigma_c^2}\right) \\
&= \prod_{g=1}^G \exp\left(y_{g,n} \log \lambda_{g,n} - \lambda_{g,n} - \frac{c_n^2}{2\sigma_c^2}\right) \\
&= \exp\left(\sum_{g=1}^G [y_{g,n} \log \lambda_{g,n} - \lambda_{g,n}] - \frac{G c_n^2}{2\sigma_c^2}\right)
\end{aligned}$$

where the sum inside the exponent can be parallelized on the GPU.

3.2 $p\left(\frac{1}{\sigma_c^2} \mid \dots\right)$

$$\begin{aligned}
p(\sigma_c \mid \dots) &= \prod_{n=1}^N \prod_{g=1}^G \mathcal{N}(c_n \mid 0, \sigma_c^2) \cdot \mathcal{U}(\sigma_c \mid 0, \sigma_{c0}) \\
&\propto \prod_{n=1}^N \prod_{g=1}^G \frac{1}{\sqrt{\sigma_c^2}} \exp\left(-\frac{c_n^2}{2\sigma_c^2}\right) \cdot \mathcal{I}(0 < \sigma_c < \sigma_{c0}) \\
&= (\sigma_c^2)^{-NG/2} \exp\left(-\frac{1}{\sigma_c^2} \frac{G}{2} \sum_{n=1}^N c_n^2\right) \cdot \mathcal{I}(0 < \sigma_c < \sigma_{c0})
\end{aligned}$$

which, for constants a and $b = \frac{G}{2} \sum_{n=1}^N c_n^2$, can be written as

$$a \cdot (\sigma_c^2)^{-NG/2} \exp\left(-\frac{1}{\sigma_c^2} b\right) \mathcal{I}(0 < \sigma_c^2 < \sigma_{c0}^2)$$

Transformation: let $z = g(\sigma_c) = \sigma_c^2$ so that $g^{-1}(z) = \sqrt{z}$ and:

$$\begin{aligned}
p(\sigma_c^2 = z \mid \dots) &= p(\sigma_c = g^{-1}(z) \mid \dots) \left| \frac{dg^{-1}(z)}{dz} \right| \\
&= a \cdot z^{-NG/2} \exp\left(-\frac{1}{(\sqrt{z})^2} b\right) \mathcal{I}(0 < z < \sigma_{c0}^2) \left| \frac{1}{2} z^{-1/2} \right| \\
&= \frac{a}{2} z^{-(NG/2-1/2+1)} \exp\left(-\frac{1}{z} b\right) \mathcal{I}(0 < z < \sigma_{c0}^2) \\
&= \text{Inv-Gamma}\left(z \mid \text{shape} = \frac{NG-1}{2}, \text{rate} = \frac{1}{b}\right) \mathcal{I}(0 < z < \sigma_{c0}^2)
\end{aligned}$$

Recalling that $\frac{1}{b} = \frac{2}{G} \frac{1}{\sum_{n=1}^N c_n^2}$,

$$p\left(\frac{1}{\sigma_c^2} \mid \dots\right) = \text{Gamma}\left(\frac{1}{\sigma_c^2} \mid \text{shape} = \frac{NG-1}{2}, \text{rate} = \frac{2}{G} \frac{1}{\sum_{n=1}^N c_n^2}\right) \mathcal{I}\left(\frac{1}{\sigma_c^2} < \frac{1}{\sigma_{c0}^2}\right)$$

3.3 $p(\varepsilon_{g,n} \mid \dots)$

$$p(\varepsilon_{g,n} \mid \dots) = \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot \mathcal{N}(\varepsilon_{g,n} \mid 0, \sigma_g^2)$$

which is nothing immediately recognizable. I will sample from this distribution in a metropolis step using the kernel:

$$\begin{aligned} & \lambda_{g,n}^{y_{g,n}} \exp(-\lambda_{g,n}) \exp\left(-\frac{\varepsilon_{g,n}^2}{2\sigma_g^2}\right) \\ &= \exp\left(y_{g,n} \log \lambda_{g,n} - \lambda_{g,n} - \frac{\varepsilon_{g,n}^2}{2\sigma_g^2}\right) \end{aligned}$$

where $\lambda_{g,n} = \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))$. The $\varepsilon_{g,n}$'s will be sampled in parallel across genes on the GPU.

3.4 $p\left(\frac{1}{\sigma_g^2} \mid \dots\right)$

$$\begin{aligned} p(\sigma_g^2 \mid \dots) &= \prod_{n=1}^N \mathcal{N}(\varepsilon_{g,n} \mid 0, \sigma_g^2) \cdot \text{Inv-Gamma}\left(\sigma_g^2 \mid \text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2}\right) \\ &\propto \prod_{n=1}^N (\sigma_g^2)^{-1/2} \exp\left(-\frac{1}{\sigma_g^2} \frac{\varepsilon_{g,n}^2}{2}\right) \cdot (\sigma_g^2)^{-(d\tau^2/2+1)} \exp\left(-\frac{1}{\sigma_g^2} \frac{d}{2}\right) \\ &= \prod_{n=1}^N (\sigma_g^2)^{-(d\tau^2+3)/2} \exp\left(-\frac{1}{\sigma_g^2} \left(\frac{\varepsilon_{g,n}^2}{2} + \frac{d}{2}\right)\right) \\ &= (\sigma_g^2)^{-N(d\tau^2+3)/2} \exp\left(-\frac{1}{\sigma_g^2} \left(\frac{2N}{d} + \frac{1}{2} \sum_{n=1}^N \varepsilon_{g,n}^2\right)\right) \end{aligned}$$

The last line is the kernel of an inverse gamma distribution with shape parameter $N\frac{d\tau^2+3}{2} - 1$ and rate parameter $\left(\frac{2N}{d} + \frac{1}{2} \sum_{n=1}^N \varepsilon_{g,n}^2\right)^{-1}$. Hence:

$$p(\sigma_g^2 \mid \dots) = \text{Gamma}\left(\text{shape} = N\frac{d\tau^2+3}{2} - 1, \text{rate} = \left(\frac{2N}{d} + \frac{1}{2} \sum_{n=1}^N \varepsilon_{g,n}^2\right)^{-1}\right)$$

The σ_g^2 's will be sampled in parallel on the GPU.

$$\begin{aligned}
p(d \mid \dots) &= \prod_{n=1}^N \prod_{g=1}^G \text{Inv-Gamma} \left(\sigma_g^2 \mid \text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2} \right) \cdot \text{U}(d \mid 0, d_0) \\
p(\tau^2 \mid \dots) &= \prod_{n=1}^N \prod_{g=1}^G \text{Inv-Gamma} \left(\sigma_g^2 \mid \text{shape} = \frac{d \cdot \tau^2}{2}, \text{rate} = \frac{d}{2} \right) \cdot \text{Gamma}(\text{shape} = a_\tau, \text{rate} = b_\tau) \\
p(\phi_g \mid \dots) &= \prod_{n=1}^N \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \cdot \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \\
p(\theta_\phi \mid \dots) &= \prod_{n=1}^N \prod_{g=1}^G \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \cdot \text{N}(\theta_\phi \mid 0, \gamma_\phi^2) \\
p(\sigma_\phi \mid \dots) &= \prod_{n=1}^N \prod_{g=1}^G \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \cdot \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0}) \\
p(\alpha_g \mid \dots) &= \prod_{k(n) \neq 2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \\
&\quad \times (\pi_\alpha \text{I}(\alpha_g = 0) + (1 - \pi_\alpha) \text{I}(\alpha_g \neq 0) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)) \\
p(\theta_\alpha \mid \dots) &= \prod_{k(n) \neq 2} \prod_{g=1}^G (\pi_\alpha \text{I}(\alpha_g = 0) + (1 - \pi_\alpha) \text{I}(\alpha_g \neq 0) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)) \cdot \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2) \\
p(\sigma_\alpha \mid \dots) &= \prod_{k(n) \neq 2} \prod_{g=1}^G (\pi_\alpha \text{I}(\alpha_g = 0) + (1 - \pi_\alpha) \text{I}(\alpha_g \neq 0) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)) \cdot \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0}) \\
p(\pi_\alpha \mid \dots) &= \prod_{k(n) \neq 2} \prod_{g=1}^G (\pi_\alpha \text{I}(\alpha_g = 0) + (1 - \pi_\alpha) \text{I}(\alpha_g \neq 0) \text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)) \cdot \text{Beta}(\pi_\alpha \mid a_\alpha, b_\alpha) \\
p(\delta_g \mid \dots) &= \prod_{k(n)=2} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \delta_g, \delta_g))) \\
&\quad \times (\pi_\delta \text{I}(\delta_g = 0) + (1 - \pi_\delta) \text{I}(\delta_g \neq 0) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \\
p(\theta_\delta \mid \dots) &= \prod_{k(n)=2} \prod_{g=1}^G (\pi_\delta \text{I}(\delta_g = 0) + (1 - \pi_\delta) \text{I}(\delta_g \neq 0) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \cdot \text{N}(\theta_\delta \mid 0, \gamma_\delta^2) \\
p(\sigma_\delta \mid \dots) &= \prod_{k(n)=2} \prod_{g=1}^G (\pi_\delta \text{I}(\delta_g = 0) + (1 - \pi_\delta) \text{I}(\delta_g \neq 0) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \cdot \text{U}(\sigma_\delta \mid 0, \sigma_{\delta 0}) \\
p(\pi_\delta \mid \dots) &= \prod_{k(n)=2} \prod_{g=1}^G (\pi_\delta \text{I}(\delta_g = 0) + (1 - \pi_\delta) \text{I}(\delta_g \neq 0) \text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)) \cdot \text{Beta}(\pi_\delta \mid a_\delta, b_\delta)
\end{aligned}$$