FCDs for mixed-effect model

Denis Valle

January 2018

To implement a Gibbs sampler for this model, we will need the full conditional distributions (FCD). These FCDs are not fundamentally different from the FCDs that you have already derived previously. The key thing is really to be very careful in terms of which distributions should be included. Here they are:

- For β_{0k}. For this parameter, we only look at the data that come from the k-th county. The other data do not depend on β_{0k}:

$$p(\beta_{0k}|\ldots) \propto \left\{\prod_{i=1}^{n_k} N(y_{ik}|\beta_{0k}+\beta_1 x_{ik1}+\beta_2 x_{ik2},\sigma^2)\right\} N(\beta_{0k}|\gamma,\tau^2)$$

$$\propto \left\{\prod_{i=1}^{n_k} \exp\left(-\frac{1}{2\sigma^2}(y_{ik}-\beta_1 x_{ik1}-\beta_2 x_{ik2}-\beta_{0k})^2\right)\right\} \exp\left(-\frac{1}{2\tau^2}(\beta_{0k}-\gamma)^2\right)$$

$$\propto \exp\left(-\frac{1}{2\left\{\frac{n_k}{\sigma^2}+\frac{1}{\tau^2}\right\}^{-1}}\left(\beta_{0k}-\left[\frac{\frac{\sum_{i=1}^{n_k}(y_{ik}-\beta_1 x_{ik1}-\beta_2 x_{ik2})}{\sigma^2}+\frac{\gamma}{\tau^2}}{\frac{n_k}{\sigma^2}+\frac{1}{\tau^2}}\right]\right)^2\right)$$

Thus, $p(\beta_{0k}|\ldots) = N\left(\left\{\frac{n_k}{\sigma^2}+\frac{1}{\tau^2}\right\}^{-1}\left\{\frac{\sum_{i=1}^{n_k}(y_{ik}-\beta_1 x_{ik1}-\beta_2 x_{ik2})}{\sigma^2}+\frac{\gamma}{\tau^2}\right\}, \left\{\frac{n_k}{\sigma^2}+\frac{1}{\tau^2}\right\}^{-1}\right)$.

- For $\frac{1}{\tau^2}$:

$$p\left(\frac{1}{\tau^2}\middle|\ldots\right) \propto \left\{\prod_{k=1}^{K} N(\beta_{0k}|\gamma,\tau^2)\right\} Gamma\left(\frac{1}{\tau^2}|a,b\right)$$

$$\propto \left\{\prod_{k=1}^{K}\left(\frac{1}{\tau^2}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2\tau^2}(\beta_{0k}-\gamma)^2\right)\right\}\left(\frac{1}{\tau^2}\right)^{a-1}\exp\left(-\frac{1}{\tau^2}b\right)$$

$$p\left(\frac{1}{\tau^2}\middle|\ldots\right) = Gamma\left(\frac{K+2a}{2}, \frac{\sum_k^K(\beta_{0k}-\gamma)^2}{2}+b\right)$$

As expected, this FCD shows that, if our prior is uninformative (i.e., $a = b \approx 0$), then $E\left[\frac{1}{\tau^2}\right] = \frac{K/2}{\frac{\sum_k^K (\beta_{0k}-\gamma)^2}{2}} = \frac{K}{\sum_k^K (\beta_{0k}-\gamma)^2}$, which is the inverse of the formula we typically think of when calculating the sample variance of the β_{0k} $(i.e., \frac{\sum_k^K (\beta_{0k}-\gamma)^2}{K})$.

- For γ:

$$p(\gamma|\ldots) \propto \left\{\prod_{k=1}^{K} N(\beta_{0k}|\gamma,\tau^2)\right\} N(\gamma|0{,}100)$$

$$p(\gamma|\ldots) = N\left(\left[\frac{K}{\tau^2}+\frac{1}{100}\right]^{-1}\frac{\sum_k^K \beta_{0k}}{\tau^2}, \left[\frac{K}{\tau^2}+\frac{1}{100}\right]^{-1}\right)$$

If we had set a very large prior variance for γ, then $E(\gamma|\ldots) \approx \left[\frac{K}{\tau^2}\right]^{-1}\frac{\sum_k^K \beta_{0k}}{\tau^2} = \frac{\sum_k^K \beta_{0k}}{K}$ (i.e., the mean of all the county intercepts)

You should be able to derive the FCDs for the other parameters. Here is what they should be:

- For β_1

$$p(\beta_1|\ldots) \propto \left\{\prod_{k=1}^{K}\prod_{i=1}^{n_k} N(y_{ik}|\beta_{0k}+\beta_1 x_{ik1}+\beta_2 x_{ik2},\sigma^2)\right\} N(\beta_1|0{,}100)$$

$$p(\beta_1|\ldots) = N\left(\left\{\frac{\sum_k\sum_i x_{ik1}^2}{\sigma^2}+\frac{1}{100}\right\}^{-1}\frac{\sum_k\sum_i x_{ik1}(y_{ik}-\beta_{0k}-\beta_2 x_{ik2})}{\sigma^2}, \left\{\frac{\sum_k\sum_i x_{ik1}^2}{\sigma^2}+\frac{1}{100}\right\}^{-1}\right)$$

- For β_2

$$p(\beta_2|\ldots) \propto \left\{\prod_{k=1}^{K}\prod_{i=1}^{n_k} N(y_{ik}|\beta_{0k}+\beta_1 x_{ik1}+\beta_2 x_{ik2},\sigma^2)\right\} N(\beta_2|0{,}100)$$

$$p(\beta_2|\ldots) = N\left(\left\{\frac{\sum_k\sum_i x_{ik2}^2}{\sigma^2}+\frac{1}{100}\right\}^{-1}\frac{\sum_k\sum_i (y_{ik}-\beta_{0k}-\beta_1 x_{ik1})x_{ik2}}{\sigma^2}, \left\{\frac{\sum_k\sum_i x_{ik2}^2}{\sigma^2}+\frac{1}{100}\right\}^{-1}\right)$$

- For $\frac{1}{\sigma^2}$

$$p\left(\frac{1}{\sigma^2}\middle|\ldots\right) \propto \left\{\prod_{k=1}^{K}\prod_{i=1}^{n_k} N(y_{ik}|\beta_{0k} + \beta_1 x_{ik1} + \beta_2 x_{ik2}, \sigma^2)\right\} Gamma\left(\frac{1}{\sigma^2}|a, b\right)$$

$$p\left(\frac{1}{\sigma^2}\middle|\ldots\right) = Gamma\left(\frac{N+2a}{2}, \frac{\sum_k \sum_i (y_{ik} - \beta_{0k} - \beta_1 x_{ik1} - \beta_2 x_{ik2})^2}{2} + b\right)$$

IMPORTANT: note that we can combine a multi-level regression like this with some of the other models we talked about in the previous class that rely on a latent continuous z variable (e.g., detection limit, probit, and ordered multinomial response), resulting in models that are relatively sophisticated although the individual pieces need not be complicated.