Supplement to "Sex, lies and self-reported counts: Bayesian analysis of longitudinal heaped count data"

Forrest W. Crawford* Robert E. Weiss† Marc A. Suchard‡

September 9, 2015

1 Laplace transform of transition probabilities for a general BDP

Here we provide expressions for the Laplace transform of the transition probability $P_{mn}(t)$ for any $m, n \geq 0$, based on Crawford and Suchard (2012), who provide detailed derivations and a numerical method for inverting the Laplace transforms to obtain time-domain transition probabilities. Before stating the main result, we establish some notation. Consider the Laplace transform of $P_{00}(t)$,

$$h_{00}(s) = \cfrac{1}{s + \lambda_0 - \cfrac{\lambda_0 \mu_1}{s + \lambda_1 + \mu_1 - \cfrac{\lambda_1 \mu_2}{s + \lambda_2 + \mu_2 - \cdots}}}. \tag{1}$$

Define the partial numerators in (1) to be $a_1 = 1$ and $a_n = -\lambda_{n-2}\mu_{n-1}$. Let the partial denominators be $b_1 = s + \lambda_0$ and $b_n = s + \lambda_{n-1} + \mu_{n-1}$ for $n \geq 2$. Then (1) becomes

$$f_{00}(s) = \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3 + \cdots}}}. \tag{2}$$

A more compact representation of this continued fraction is

$$f_{00}(s) = \frac{a_1}{b_1+} \frac{a_2}{b_2+} \frac{a_3}{b_3+} \cdots . \tag{3}$$

Now the kth approximant of $f_{00}(s)$ is the truncated continued fraction

$$f_{00}^{(k)}(s) = \frac{a_1}{b_1+} \frac{a_2}{b_2+} \cdots \frac{a_k}{b_k} = \frac{A_k(s)}{B_k(s)}. \tag{4}$$

where $A_k(s)$ and $B_k(s)$ are the kth partial numerator and denominator of the continued fraction (2).

*Department of Biostatistics, Yale School of Public Health, New Haven, CT 06510-8043 USA; forrest.crawford@yale.edu

†Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, CA 90095-1772 USA; robweiss@ucla.edu

‡Departments of Biomathematics, Biostatistics and Human Genetics, University of California, Los Angeles, Los Angeles, CA 90095-1766 USA; msuchard@ucla.edu

Theorem 1 *The Laplace transform of the transition probability $P_{m,n}(t)$ is given by*

$$f_{m,n}(s) = \begin{cases} \left(\prod_{j=n+1}^{m} \mu_j\right) \dfrac{B_n(s)}{B_{m+1}(s)+} \dfrac{B_m(s)a_{m+2}}{b_{m+2}+} \dfrac{a_{m+3}}{b_{m+3}+} \cdots & \text{for } n \le m, \\ \\ \left(\prod_{j=m}^{n-1} \lambda_j\right) \dfrac{B_m(s)}{B_{n+1}(s)+} \dfrac{B_n(s)a_{n+2}}{b_{n+2}+} \dfrac{a_{n+3}}{b_{n+3}+} \cdots & \text{for } m \le n, \end{cases} \tag{5}$$

where a_n, b_n, and B_n are as defined above (Murphy and O'Donohoe, 1975; Crawford and Suchard, 2012).

2 Full posterior

Let $\mathbf{X}$ and $\mathbf{Y}$ be the collection of all true and reported counts respectively for all subjects and timepoints. Likewise, let $\mathbf{Z}$ and $\mathbf{W}$ be the collection of all subject- and timepoint-specific covariate vectors, and let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be the corresponding fixed and random effects. The posterior density falls out as

$$\begin{aligned} p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_\beta, \mathbf{X} \mid \mathbf{Y}, \mathbf{Z}, \mathbf{W}) &\propto p(\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\gamma}) p(\mathbf{X} \mid \mathbf{Z}, \mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &\quad \times \Pr(\boldsymbol{\alpha}) p(\boldsymbol{\beta} \mid \boldsymbol{\Sigma}_\beta) p(\boldsymbol{\theta}) p(\boldsymbol{\gamma}) \Pr(\boldsymbol{\Sigma}_\beta) \\ &= \left[\prod_{i=1}^{N} \prod_{t=1}^{n_i} p(Y_{it} \mid X_{it}, \boldsymbol{\theta}, \boldsymbol{\gamma}) p(X_{it} \mid \mathbf{Z}_{it}, \mathbf{W}_{it}, \boldsymbol{\alpha}, \boldsymbol{\beta}_i)\right] \\ &\quad \times p(\boldsymbol{\alpha}) p(\boldsymbol{\beta} \mid \boldsymbol{\Sigma}_\beta) p(\boldsymbol{\theta}) p(\boldsymbol{\gamma}) p(\boldsymbol{\Sigma}_\beta). \end{aligned} \tag{6}$$

3 Sampling $\boldsymbol{\alpha}$

The full conditional distribution of $\boldsymbol{\alpha}$ is

$$\begin{aligned} \Pr(\boldsymbol{\alpha} \mid \mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\beta) &= \Pr(\boldsymbol{\alpha} \mid \mathbf{X}, \mathbf{Z}, \mathbf{W}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\beta) \\ &\propto \Pr(\mathbf{X} \mid \mathbf{Z}, \mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \Pr(\boldsymbol{\alpha}) \\ &= \left[\prod_{i=1}^{N} \prod_{t=1}^{n_i} \Pr(X_{it} \mid \mathbf{Z}_{it}, \mathbf{W}_{it}, \boldsymbol{\alpha}, \boldsymbol{\beta}_i)\right] \Pr(\boldsymbol{\alpha}). \end{aligned} \tag{7}$$

Since X_{it} is distributed according to the GLMM with mean $\eta_{it} = \exp(\mathbf{W}_{it}\boldsymbol{\alpha} + \mathbf{Z}_{it}\boldsymbol{\beta}_i)$, there is no closed-form expression for the full conditional distribution of $\boldsymbol{\alpha}$ in general. Therefore, we specify a proposal distribution as the normal approximation to the likelihood at the previous value of $\boldsymbol{\alpha}$ times the conjugate normal prior for $\boldsymbol{\alpha}$, and we perform a Metropolis accept/reject step to obtain the next sample of $\boldsymbol{\alpha}$.

4 Sampling $\boldsymbol{\beta}_i$

Similarly, the full conditional distribution of $\boldsymbol{\beta}_i$ is

$$\begin{aligned} \Pr(\boldsymbol{\beta}_i \mid \mathbf{X}_i, \mathbf{Z}_i, \mathbf{W}_i, \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\Sigma}_\beta) &\propto \Pr(\mathbf{X}_i \mid \mathbf{Z}_i, \mathbf{W}_i, \boldsymbol{\alpha}, \boldsymbol{\beta}_i) \Pr(\boldsymbol{\beta}_i \mid \boldsymbol{\Sigma}_\beta) \\ &= \left[\prod_{t=1}^{n_i} \Pr(X_{it} \mid \mathbf{Z}_{it}, \mathbf{W}_{it}, \boldsymbol{\alpha}, \boldsymbol{\beta}_i)\right] \Pr(\boldsymbol{\beta}_i \mid \boldsymbol{\Sigma}_\beta). \end{aligned} \tag{8}$$

We also sample each $\boldsymbol{\beta}_i$ using a Metropolis-Hastings step with a proposal density derived from a normal approximation as above for $\boldsymbol{\alpha}$.

5 Sampling $\boldsymbol{\theta}$

The full conditional distribution of $\boldsymbol{\theta}$ is

$$\begin{aligned}\Pr(\boldsymbol{\theta} \mid \mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}, \gamma) &\propto \Pr(\boldsymbol{\theta} \mid \mathbf{X}, \mathbf{Y}, \gamma) \\ &\propto \Pr(\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\theta}, \gamma) \Pr(\boldsymbol{\theta}) \\ &= \left[\prod_{i=1}^{N} \prod_{t=1}^{n_i} \Pr(Y_{it} \mid X_{it}, \boldsymbol{\theta}, \gamma)\right] \Pr(\boldsymbol{\theta}).\end{aligned} \tag{9}$$

To keep the elements of $\boldsymbol{\theta}$ non-negative, we use a truncated normal proposal distribution

$$\theta_\ell^\star \sim \text{Normal}_+\left(\theta_\ell^{(j)},\ \sigma_\theta^2\right) \tag{10}$$

centered at the current value $\theta_\ell^{(j)}$ and where σ_θ^2 is a tuning parameter chosen to optimize mixing of the chain. $\text{Normal}_+(\cdot)$ denotes the distribution with density

$$\frac{1}{\sqrt{2\pi}\sigma_\theta} \exp\left[-\frac{\left(\theta - \theta^{(j)}\right)^2}{2\sigma_\theta^2}\right] \Big/ \left(1 - \Phi\big(-\theta^{(j)}/\sigma_\theta^2\big)\right) \tag{11}$$

for $\theta > 0$ and 0 otherwise, where $\Phi(\cdot)$ is the standard normal distribution function. We perform a Metropolis-Hastings accept/reject step to find the next sample of $\boldsymbol{\theta}$.

6 Sampling $\boldsymbol{\gamma}$

The full conditional distribution of $\boldsymbol{\gamma}$ is similar to that of $\boldsymbol{\theta}$:

$$\begin{aligned}\Pr(\gamma \mid \mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}, \boldsymbol{\theta}) &\propto \Pr(\gamma \mid \mathbf{X}, \mathbf{Y}, \boldsymbol{\theta}) \\ &\propto \Pr(\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\theta}, \gamma) \Pr(\gamma) \\ &= \left[\prod_{i=1}^{N} \prod_{t=1}^{n_i} \Pr(Y_{it} \mid X_{it}, \boldsymbol{\theta}, \gamma)\right] \Pr(\gamma).\end{aligned} \tag{12}$$

To arrive at a valid proposal for $\boldsymbol{\gamma}$, we take the monotonicity constraint $\gamma_0 < \cdots < \gamma_J$ into account using a rejection sampling method. We propose a new sample $\boldsymbol{\gamma}^*$ from the previous value of $\boldsymbol{\gamma}$ using a multivariate Normal kernel:

$$\boldsymbol{\gamma}^\star \sim \text{Normal}\left(\boldsymbol{\gamma}^{(j)}, \Sigma_\gamma\right) \tag{13}$$

centered at the current value of $\gamma_\ell^{(j)}$ and reject the sample if the constraint is not satisfied. We choose the tuning parameter Σ_γ to optimize mixing. We perform an accept/reject Metropolis-Hastings step to choose the next value of $\boldsymbol{\gamma}$.

References

Crawford FW, Suchard MA (2012) Transition probabilities for general birth-death processes with applications in ecology, genetics, and evolution. Journal of Mathematical Biology 65:553–580

Murphy JA, O'Donohoe MR (1975) Some properties of continued fractions with applications in Markov processes. IMA Journal of Appled Mathematics 16:57–71