# 0 Introduction

The goal of this document is to fully characterize the Dunson and Stanford day-specific probabilities model. In its current state it tries to provide full detail of the derivations described in *Bayesian Inferences on Predictors of Conception Probabilities*.

# 1 The day-specific probabilities model

## 1.1 Model specification

We wish to model the probability of a woman becoming pregnant for a given menstrual cycle as a function of her covariate status across the days of the cycle. Consider a study cohort and let us index

$$
\begin{aligned}
&\text{woman } i, && i = 1, \dots, n \\
&\text{cycle } j, && j = 1, \dots, n_i \\
&\text{day } k, && k = 1, \dots, K
\end{aligned}
$$

where day $k$ refers to the $k^{\text{th}}$ day out of a total of $K$ days in the fertile window. Let us write day $i, j, k$ as a shorthand for individual $i$, cycle $j$, and day $k$ and similarly for cycle $j, k$. Then we observe that

$$
\begin{aligned}
&\mathbb{P}\Big(\text{yes a pregnancy for cycle i,j} \;\Big|\; \text{intercourse status across cycle}\Big) \\
&\quad = 1 - \mathbb{P}\Big(\text{not a pregnancy for cycle } i, j \;\Big|\; \text{intercourse status across cycle}\Big) \\
&\quad = 1 - \mathbb{P}\Big(\text{didn't become pregnant on any of days } 1, \dots, K \;\Big|\; \text{intercourse status across cycle}\Big) \\
&\quad = 1 - \prod_{k=1}^{K} \mathbb{P}\Big(\text{didn't become pregnant on day } i, j, k \;\Big|\; \text{intercourse status across cycle,} \\
&\qquad\qquad\qquad\qquad \text{didn't become pregnant on days 1 through } k-1 \text{ of cycle } i, j\Big) \\
&\quad = 1 - \prod_{k=1}^{K} \Big\{ 1 - \mathbb{P}\Big(\text{became pregnant on day } i, j, k \;\Big|\; \text{intercourse status across cycle,} \\
&\qquad\qquad\qquad\qquad \text{didn't become pregnant on days 1 through } k-1 \text{ of cycle } i, j\Big) \Big\} \\
&\quad = 1 - \prod_{k=1}^{K} \Big\{ 1 - I\Big(\text{yes intercourse on day i,j,k}\Big) \\
&\qquad\qquad \times \mathbb{P}\Big(\text{became pregnant on day } i, j, k \;\Big|\; \text{yes intercourse on day i,j,k,} \\
&\qquad\qquad\qquad\qquad \text{didn't become pregnant on days 1 through } k-1 \text{ of cycle } i, j\Big) \Big\} \\
&\quad = 1 - \prod_{k=1}^{K} \Big\{ 1 - \mathbb{P}\Big(\text{became pregnant on day } i, j, k \;\Big|\; \text{yes intercourse on day i,j,k,} \\
&\qquad\qquad \text{didn't become pregnant on days 1 through } k-1 \text{ of cycle } i, j\Big) \Big\}^{I\left(\text{yes intercourse on day i,j,k}\right)}
\end{aligned}
$$

With this result in mind, we now consider the Dunson and Stanford day-specific probabilities model. Using the same indexing scheme as above, denote

| | |
|---|---|
| $Y_{ij}$ | an indicator of conception for woman $i$, cycle $j$ |
| $X_{ijk}$ | an indicator of intercourse for woman $i$, cycle $j$, day $k$ |
| $\boldsymbol{u}_{ijk}$ | a covariate status vector of length $q$ for woman $i$, cycle $j$, day $k$ |

Then writing $\boldsymbol{X}_{ij} = \left(X_{ij1}, \ldots, X_{ijK}\right)$ and $\boldsymbol{U}_{ij} = \left(\boldsymbol{u}'_{ijk}, \ldots, \boldsymbol{u}'_{ijk}\right)'$, Dunson and Stanford propose the model:

$$\begin{aligned}
\mathbb{P}\left(Y_{ij} = 1 \mid \xi_i, \boldsymbol{X}_{ij}, \boldsymbol{U}_{ij}\right) &= 1 - \prod_{k=1}^{K} (1 - \lambda_{ijk})^{X_{ijk}} \\
\lambda_{ijk} &= 1 - \exp\left\{-\xi_i \exp\left(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\right)\right\} \\
\xi_i &\sim \mathcal{G}(\phi, \phi)
\end{aligned} \tag{1}$$

From our previous derivation, we see that we may interpret $\lambda_{ijk}$ as the day-specific probability of conception in cycle $j$ from couple $i$ given that conception has not already occured, or in the language of Dunson and Stanford, given intercourse only on day $k$.

Delving further, we see that $\lambda_{ijk}$ is strictly increasing in $u_{ijkh}\beta_h$, where we are denoting $u_{ijkh}$ to be the $h^{\text{th}}$ term in $\boldsymbol{u}_{ijk}$ and similarly for $\beta_h$. When $\beta_h = 0$ then the $h^{\text{th}}$ covariate has no effect on the day-specific probability of conception.

$\lambda_{ijk}$ is also strictly increasing in $\xi_i$ which as Dunson and Stanford suggest may be interpreted as a woman-specific random effect. The authors state that specifying the distribution of the $\xi_i$ with a common parameters prevents nonidentifiability between $\mathbb{E}[\xi_i]$ and the day-specific parameters. Since $\text{Var}[\xi_i] = 1/\phi$ it follows that $\phi$ may be interpreted as a measure of variability across women.

### 1.1.1 Computation consideration

As an aside, we note that it may be more computationally convenient to calculate

$$\begin{aligned}
&\mathbb{P}\left(Y_{ij} = 1 \mid \xi_i, \boldsymbol{X}_{ij}, \boldsymbol{U}_{ij}\right) \\
&\quad = 1 - \prod_{k=1}^{K} (1 - \lambda_{ijk})^{X_{ijk}} \\
&\quad = 1 - \prod_{k=1}^{K} \left[ \exp\left\{-\xi_i \exp\left(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\right)\right\} \right]^{X_{ijk}} \\
&\quad = 1 - \prod_{k=1}^{K} \exp\left\{-X_{ijk}\xi_i \exp\left(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\right)\right\}
\end{aligned}$$

## 1.2 Marginal probability of conception

The marginal probability of conception, obtained by integrating out the couple-specific frailty $\xi_i$, has form as follows.

$$
\begin{aligned}
&\mathbb{P}(Y_{ij} = 1 \,|\, \boldsymbol{X}_{ij}, \boldsymbol{U}_{ij}) \\
&\quad = \int_0^\infty \mathbb{P}\Big(Y_{ij}, \xi_i \,|\, \boldsymbol{X}_{ij}, \boldsymbol{U}_{ij}\Big)\, d\xi_i \\
&\quad = \int_0^\infty \mathbb{P}\Big(Y_{ij}, \xi_i \,|\, \boldsymbol{X}_{ij}, \boldsymbol{U}_{ij}\Big)\, \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = \int_0^\infty \left[1 - \prod_{k=1}^{K} (1-\lambda_{ijk})^{X_{ijk}}\right] \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = 1 - \int_0^\infty \prod_{k=1}^{K} (1-\lambda_{ijk})^{X_{ijk}}\, \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = 1 - \int_0^\infty \prod_{k=1}^{K} \Big[ \exp\Big\{-\xi_i \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\Big\} \Big]^{X_{ijk}} \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = 1 - \int_0^\infty \prod_{k=1}^{K} \exp\Big\{ -\xi_i X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\Big\} \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = 1 - \int_0^\infty \exp\left\{ -\xi_i \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\right\} \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = 1 - \left[ \frac{\phi}{\phi + \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)} \right]^{\phi}
\end{aligned}
$$

since

$$
\begin{aligned}
&\int_0^\infty \exp\left\{ -\xi_i \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\right\} \mathcal{G}(\xi_i;\, \phi, \phi)\, d\xi_i \\
&\quad = \int_0^\infty \exp\left\{ -\xi_i \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\right\} \frac{\phi^{\phi}}{\Gamma(\phi)}\, \xi_i^{\phi-1} d\xi_i \\
&\quad = \int_0^\infty \frac{\phi^{\phi}}{\Gamma(\phi)}\, \xi_i^{\phi-1} \exp\left\{ -\xi_i \left[ \phi + \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\right]\right\} d\xi_i \\
&\quad = \left[ \frac{\phi}{\phi + \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)} \right]^{\phi} \int_0^\infty \frac{\Big[\phi + \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\Big]^{\phi}}{\Gamma(\phi)} \\
&\qquad\qquad \times \xi_i^{\phi-1} \exp\left\{ -\xi_i \left[ \phi + \sum_{k=1}^{K} X_{ijk} \exp\Big(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\Big)\right]^{\phi} \right\} d\xi_i
\end{aligned}
$$

and the function inside the integral is a gamma density function.

## 2 Posterior computation

Express the data augmentation model as

$$
\begin{aligned}
Y_{ij} &= I\left(\sum_{k=1}^{K} X_{ijk} Z_{ijk} > 0\right), \\
Z_{ijk} &\sim \text{Poisson}\left(\xi_i \exp\left(\boldsymbol{u}'_{ijk}\boldsymbol{\beta}\right)\right), \qquad k = 1, \ldots, K
\end{aligned} \tag{2}
$$

Let us further define $W_{ijk} = X_{ijk} Z_{ijk}$ for all $i, j, k$.

### 2.1 Verifying the equivalence of the data augmentation model