DPGLM – Posterior Sampling

1 DPGLM

Consider a GLM

$$y \sim p(y \mid x) = p_x(y) \propto \exp(\theta_x y)\widetilde{H}(y) \tag{1}$$

with response $y \in \mathcal{Y} \subset \mathcal{R}$ and a p-dimensional covariate vector $x \in \mathcal{X}$. Here, $T(\theta_x)$ is the normalization constant for p_x with

$$T(\theta_x) = \int_{\mathcal{Y}} \exp(\theta_x y)\widetilde{H}(dy) \ . \tag{2}$$

Hence $\frac{\widetilde{H}}{T(0)}$ is the baseline density. Let $b(\theta) = \ln T(\theta)$. In the classical GLM, the baseline distribution is assumed to be in a parametric family—in the proposed semi-parametric model it becomes an unknown parameter. As in the classical GLM, $\eta = x^T\beta$ is a linear predictor, g is a link function, and $\mu = \mathbb{E}(y \mid x) = g^{-1}(\eta)$. For fixed $\widetilde{H}$, the expectation μ implicitly determines θ by the equation

$$\mu = \mathbb{E}(y \mid x) = b'(\theta_x) = \frac{\int_{\mathcal{Y}} y \exp(\theta_x y)\widetilde{H}(dy)}{\int_{\mathcal{Y}} \exp(\theta_x y)\widetilde{H}(dy)} \ , \tag{3}$$

where we note that $b'(\theta)$ is a strictly increasing function of θ. The free parameters in the model are β and $\widetilde{H}$. By contrast, $\theta_x = \theta(\beta, \widetilde{H}; x)$ is a derived parameter based on (3). We can write the solution for θ as a function of μ, denoted as $b'^{-1}(\mu; \widetilde{H})$, which additionally depends on $\widetilde{H}$, in addition to depending on β, implicitly through $\mu = x^T\beta$.

We put a multivariate normal prior on β and a completely random measure (CRM) on $\widetilde{H}$ as:

$$\widetilde{H} \sim \text{CRM}(\widetilde{\nu}), \text{ with Lévy intensity, } \widetilde{\nu}(dy, ds) = \rho(ds \mid y)H(dy), \tag{4}$$

where $\rho(\cdot \mid y)$ is a measure on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ for y in $\mathcal{Y}$ and H is a measure on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$. If $\rho(\cdot \mid y)$ does not depend on y i.e., $\rho(\cdot \mid y) = \rho(\cdot)$ for all y, then both $\widetilde{\nu}$ and $\widetilde{H}$ are termed homogeneous. Otherwise, $\widetilde{\nu}$ and $\widetilde{H}$ are termed non-homogeneous. A CRM on $\widetilde{H}$ implies a prior on $\mathcal{F} = \{p_x : x \in \mathcal{X}\}$. In case of gamma CRM on $\widetilde{H}$, with concentration parameter α, the prior on $\mathcal{F}$ becomes a Dependent Dirichlet Process (DDP) prior. We can express $\widetilde{H}(\cdot) = \sum_{\ell=1}^{\infty} s_\ell \delta_{y_\ell}(\cdot)$ with

Lévy intensity, $\widetilde{\nu}(dy, ds) = \alpha \frac{e^{-s}}{s} ds H(dy)$. Therefore $p_x(y)$ can be expressed as follows:

$$\begin{aligned} p_x(y) &\propto \exp(\theta_x y) \sum_{\ell=1}^{\infty} s_\ell \delta_{y_\ell}(y) \\ &= \sum_{\ell=1}^{\infty} \{\exp(\theta_x y)\, s_\ell\} \delta_{y_\ell}(y) \\ &= \sum_{\ell=1}^{\infty} s_\ell(x; y) \delta_{y_\ell}(y), \end{aligned}$$

where $s_\ell(x; y) = \exp(\theta_x y)\, s_\ell$, depends on x implicitly through θ_x. Hence, this Bayesian framework falls nicely into the category of varying weights DDP with the atoms being constant across x.

2 Posterior Sampling

Let $\mathcal{D}$ denote the observed data $\{x_i, y_i\}_{i=1}^n$. From (2), we denote $T_i = T_i(\mathcal{Y}) = T(\theta_{x_i}) = \int_{\mathcal{Y}} \exp(\theta_{x_i} y) \widetilde{H}(dy)$. For simplicity, we consider no ties in $\{y_i\}_{i=1}^n$; we extend it to the general case later. Consider n disjoint subsets $C_1, \ldots, C_n$ of $\mathcal{Y}$, where we take $C_i := \{y \in \mathcal{Y} : d(y, y_i) < \epsilon\}$, where d is a distance function and $C_{n+1} = \mathcal{Y} \setminus \cup_{i=1}^n C_i$. We next denote $\widetilde{T}_i = T_i(C_i) = \int_{C_i} \exp(\theta_{x_i} y) \widetilde{H}(dy)$. The conditional Laplace functional of $\widetilde{H}$ (note: literally, Laplace functional for posterior $\widetilde{H} \mid y_1, \ldots, y_n$, when we push $\epsilon \to 0$) is given by,

$$\begin{aligned} E\left(e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \mid y_1 \in C_1, \ldots, y_n \in C_n\right) &= \int e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \Pr(\widetilde{H} \mid y_1 \in C_1, \ldots, y_n \in C_n) d(\widetilde{H}) \\ &= \int e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \frac{\Pr(\widetilde{H}, y_1 \in C_1, \ldots, y_n \in C_n)}{\Pr(y_1 \in C_1, \ldots, y_n \in C_n)} d(\widetilde{H}) \\ &= \frac{\int e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \Pr(\widetilde{H}, y_1 \in C_1, \ldots, y_n \in C_n) d(\widetilde{H})}{\Pr(y_1 \in C_1, \ldots, y_n \in C_n)} \\ &= \frac{\int e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \Pr(\widetilde{H}, y_1 \in C_1, \ldots, y_n \in C_n) d(\widetilde{H})}{\int \Pr(\widetilde{H}, y_1 \in C_1, \ldots, y_n \in C_n) d(\widetilde{H})} \\ &= \frac{\int e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \Pr(y_1 \in C_1, \ldots, y_n \in C_n \mid \widetilde{H}) \Pr(\widetilde{H}) d(\widetilde{H})}{\int \Pr(y_1 \in C_1, \ldots, y_n \in C_n \mid \widetilde{H}) \Pr(\widetilde{H}) d(\widetilde{H})} \\ &= \frac{\int e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \frac{T_i(C_i)}{T_i(\mathcal{Y})} \Pr(\widetilde{H}) d(\widetilde{H})}{\int \prod_{i=1}^n \frac{T_i(C_i)}{T_i(\mathcal{Y})} \Pr(\widetilde{H}) d(\widetilde{H})} \\ &= \frac{E\left(e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \prod_{i=1}^n \frac{\widetilde{T}_i}{T_i}\right)}{E\left(\prod_{i=1}^n \frac{\widetilde{T}_i}{T_i}\right)} \end{aligned}$$

[not a conditional expectation anymore :)]

2.1 Numerator calculation

We have $\frac{1}{T_i} = \int_{\mathbb{R}^+} e^{-T_i u_i} du_i = \int_{\mathbb{R}^+} e^{-\int_{\mathcal{Y}} \exp(\theta_{x_i} y)\widetilde{H}(dy) u_i} du_i = \int_{\mathbb{R}^+} e^{-\int_{\mathcal{Y}} u_i \exp(\theta_i y)\widetilde{H}(dy)} du_i$. Hence,

$$\begin{aligned}
\prod_{i=1}^{n} \frac{1}{T_i} &= \prod_{i=1}^{n} \int_{\mathbb{R}^+} e^{-\int_{\mathcal{Y}} u_i \exp(\theta_i y)\widetilde{H}(dy)} du_i \\
&= \int_{\mathbb{R}^{+n}} e^{-\int_{\mathcal{Y}} \sum_{i=1}^{n} u_i \exp(\theta_i y)\widetilde{H}(dy)} du_1 \ldots du_n \\
&= \int_{\mathbb{R}^{+n}} \prod_{j=1}^{n+1} e^{-\int_{C_j} \sum_{i=1}^{n} u_i \exp(\theta_i y)\widetilde{H}(dy)} du_1 \ldots du_n
\end{aligned}$$

, with $\mathcal{Y} = \cup_{j=1}^{n+1} C_j$. We have, $e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \prod_{i=1}^{n} \frac{\widetilde{T}_i}{T_i}$

$$\begin{aligned}
&= \prod_{j=1}^{n+1} e^{-\int_{C_j} h(y)\widetilde{H}(dy)} \prod_{i=1}^{n} \widetilde{T}_i \int_{\mathbb{R}^{+n}} \prod_{j=1}^{n+1} e^{-\int_{C_j} \sum_{i=1}^{n} u_i \exp(\theta_i y)\widetilde{H}(dy)} du_1 \ldots du_n \\
&= \prod_{i=1}^{n} \widetilde{T}_i \int_{\mathbb{R}^{+n}} \prod_{j=1}^{n+1} e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} du_1 \ldots du_n \\
&= \prod_{i=1}^{n} \widetilde{T}_i \int_{\mathbb{R}^{+n}} e^{-\int_{C_{n+1}} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \prod_{j=1}^{n} e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} du_1 \ldots du_n \\
&= \int_{\mathbb{R}^{+n}} e^{-\int_{C_{n+1}} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \prod_{j=1}^{n} \left\{ -\frac{d}{du_j} e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \right\} du_1 \ldots du_n \\
&\left[\text{Note that : } \frac{d}{du_j} e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} = -\widetilde{T_j} e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \right]
\end{aligned}$$

Hence, $\mathbb{E}_{\widetilde{H}} \left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \prod_{i=1}^{n} \frac{\widetilde{T}_i}{T_i} \right]$

$$\begin{aligned}
&\overset{Fubini}{=} \int_{\mathbb{R}^{+n}} \mathbb{E}_{\widetilde{H}} \left[e^{-\int_{C_{n+1}} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \prod_{j=1}^{n} \left\{ -\frac{d}{du_j} e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \right\} \right] du_1 \ldots du_n \\
&= \int_{\mathbb{R}^{+n}} \mathbb{E}_{\widetilde{H}} \left[e^{-\int_{C_{n+1}} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \right] \prod_{j=1}^{n} -\frac{d}{du_j} \mathbb{E}_{\widetilde{H}} \left[e^{-\int_{C_j} \left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)\widetilde{H}(dy)} \right] du_1 \ldots du_n \\
&\left[\text{We choose } \epsilon \text{ very close to 0 such that } C_j \cap C_{j'} = \phi \overset{CRM}{\Longrightarrow} \widetilde{H}(C_j) \text{ is independent over } j \right]
\end{aligned}$$

Let us define $\eta(y, \mathbf{u}) = \eta(y, u_1, \ldots, u_n) := h(y) + \sum_{i=1}^{n} u_i \exp(\theta_i y)$ and $S = \mathbb{R}^+ \times \mathcal{Y} = \cup_{j=1}^{n+1} S_j$, with $S_j = \mathbb{R}^+ \times C_j$. From (4), $\widetilde{\nu}(ds, dy) = \rho(ds \mid y) H(dy)$. Next using Lévy–Khintchine representation,

$$
\begin{aligned}
&\mathbb{E}_{\widetilde{H}}\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \prod_{i=1}^{n} \frac{\widetilde{T}_i}{T_i}\right] \\
&= \int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+ \times C_{n+1}} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\} \prod_{j=1}^{n} -\frac{d}{du_j} \exp\left\{-\int_{\mathbb{R}^+ \times C_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\} du_1 \dots du_n \\
&= \int_{\mathbb{R}^{+n}} \exp\left\{-\int_{S_{n+1}} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\} \prod_{j=1}^{n} -\frac{d}{du_j} \exp\left\{-\int_{S_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\} du_1 \dots du_n \\
&= \int_{\mathbb{R}^{+n}} \exp\left\{-\int_{S} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\} \cdot \prod_{j=1}^{n} \left[\left\{-\frac{d}{du_j} \exp\left(-\int_{S_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right)\right\} \cdot \right. \\
&\qquad\qquad \left. \exp\left\{\int_{S_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\}\right] du_1 \dots du_n \\
&= \int_{\mathbb{R}^{+n}} \exp\left\{-\int_{S} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\} \cdot \prod_{j=1}^{n} V_{C_j}^{(1)}(\mathbf{u})\ du_1 \dots du_n
\end{aligned}
$$

, with $V_{C_j}^{(1)}(\mathbf{u}) = \left\{-\frac{d}{du_j} \exp\left(-\int_{S_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right)\right\} \cdot \exp\left\{\int_{S_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right\}$.
We have,

$$
\begin{aligned}
V_{C_j}^{(1)}(\mathbf{u}) &= -\frac{d}{du_j}\left(-\int_{S_j} \left[1-e^{-s\eta(y,\mathbf{u})}\right] \widetilde{\nu}(ds,dy)\right) \\
&= \int_{S_j} e^{-s\eta(y,\mathbf{u})}\ s\frac{d}{du_j}\left\{h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right\} \widetilde{\nu}(ds,dy) \\
&= \int_{S_j} e^{-s\eta(y,\mathbf{u})}\ s\exp(\theta_j y)\ \widetilde{\nu}(ds,dy) \\
&= \int_{C_j} \exp(\theta_j y)\left\{\int_{\mathbb{R}^+} s e^{-s\eta(y,\mathbf{u})}\ \rho(ds \mid y)\right\} H(dy) \\
&= \int_{C_j} \exp(\theta_j y)\phi_1(\mathbf{u},y) V_{C_j}^{(0)}(\mathbf{u}) H(dy) = \int_{C_j} \exp(\theta_j y)\phi_1(\mathbf{u},y) H(dy),
\end{aligned}
$$

with $V_{C_j}^{(0)}(\mathbf{u}) = 1$ and $\phi_1(\mathbf{u},y) = \int_{\mathbb{R}^+} s e^{-s\left(h(y)+\sum_{i=1}^{n} u_i \exp(\theta_i y)\right)} \rho(ds \mid y)$. Let us denote $\Delta_{\theta_j}^{(1)}(\mathbf{u},y) := \exp(\theta_j y)\phi_1(\mathbf{u},y)$. Note that: we do not need to introduce $V_{C_j}^{(0)}(\mathbf{u})$ for our proof.

2.2 Denominator calculation

$\mathbb{E}_{\widetilde{H}}\left[\prod_{i=1}^{n} \frac{\widetilde{T}_i}{T_i}\right]$ = same expression as numerator with $\eta(y,\mathbf{u}) = \sum_{i=1}^{n} u_i \exp(\theta_i y)$ (i.e., $h(y) = 0$).

2.3 Final expression

$$
\begin{aligned}
&E\left(e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \mid y_1 \in C_1, \ldots, y_n \in C_n\right) \\
&= \frac{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left(h(y)+\sum_{i=1}^n u_i \exp(\theta_i y)\right)}\right]\rho(ds \mid y)H(dy)\right\} \cdot \prod_{j=1}^n \int_{C_j} \Delta_{\theta_j}^{(1)}(\mathbf{u}, y)H(dy)\ du_1 \ldots du_n}{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\right]\rho(ds \mid y)H(dy)\right\} \cdot \prod_{j=1}^n \int_{C_j} \Delta_{\theta_j}^{\star(1)}(\mathbf{u}, y)H(dy)\ du_1 \ldots du_n} \\
&\overset{\epsilon\to 0}{=} \frac{\int_{\mathbb{R}^{+n}} A \cdot \prod_{j=1}^n \left\{\int_{\mathcal{Y}} \exp(\theta_j y)\left[\int_{\mathbb{R}^+} se^{-s\left\{h(y)+\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\rho(ds \mid y)\right]\delta_{y_j}(y)H(dy)\right\}\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} B \cdot \prod_{j=1}^n \left\{\int_{\mathcal{Y}} \exp(\theta_j y)\left[\int_{\mathbb{R}^+} se^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\rho(ds \mid y)\right]\delta_{y_j}(y)H(dy)\right\}\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} A \cdot \prod_{j=1}^n \left\{\exp(\theta_j y_j)\left[\int_{\mathbb{R}^+} se^{-s\left\{h(y_j)+\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j)\right]\right\}\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} B \cdot \prod_{j=1}^n \left\{\exp(\theta_j y_j)\left[\int_{\mathbb{R}^+} se^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j)\right]\right\}\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} A \cdot \prod_{j=1}^n \int_{\mathbb{R}^+} e^{-h(y_j)s}\ se^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j)\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} B \cdot \prod_{j=1}^n \int_{\mathbb{R}^+} se^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j)\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} A \cdot C\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} B \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left(h(y)+\sum_{i=1}^n u_i \exp(\theta_i y)\right)}\right]\rho(ds \mid y)H(dy)\right\} \cdot C\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\right]\rho(ds \mid y)H(dy)\right\} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[e^{\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s} - e^{-h(y)s} + 1 - 1\right] e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\rho(ds \mid y)H(dy)\right\} \cdot C\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\right]\rho(ds \mid y)H(dy)\right\} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left(\left[1-e^{-h(y)s}\right] e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s} + \left[1-e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\right]\right)\rho(ds \mid y)H(dy)\right\} \cdot C\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\right]\rho(ds \mid y)H(dy)\right\} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-h(y)s}\right]\nu^{(\mathbf{u})}(ds, dy)\right\} \cdot \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\right]\nu(ds, dy)\right\} \cdot C\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\right]\rho(ds \mid y)H(dy)\right\} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}} \left[1-e^{-h(y)s}\right]\nu^{(\mathbf{u})}(ds, dy)\right\} \cdot e^{-\psi(\mathbf{u})} \cdot C\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} e^{-\psi(\mathbf{u})} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} E_{\widetilde{H}^{(\mathbf{u})}}\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}^{(\mathbf{u})}(dy)}\right] \cdot \prod_{j=1}^n E\left[e^{-h(y_j)}J_j^{(\mathbf{u}, y_j)}\right] \int_{\mathbb{R}^+} se^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j) \cdot e^{-\psi(\mathbf{u})}\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} e^{-\psi(\mathbf{u})} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} E_{\widetilde{H}^{(\mathbf{u})}}\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}^{(\mathbf{u})}(dy)}\right] \cdot \prod_{j=1}^n E\left[e^{-h(y_j)}J_j^{(\mathbf{u}, y_j)}\right] \prod_{j=1}^n \int_{\mathbb{R}^+} se^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j) \cdot e^{-\psi(\mathbf{u})}\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} e^{-\psi(\mathbf{u})} \cdot D\ d\mathbf{u}} \\
&= \frac{\int_{\mathbb{R}^{+n}} E_{\widetilde{H}^{(\mathbf{u})}}\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}^{(\mathbf{u})}(dy)}\right] \cdot \prod_{j=1}^n E\left[e^{-h(y_j)}J_j^{(\mathbf{u}, y_j)}\right] \cdot D \cdot e^{-\psi(\mathbf{u})}\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} e^{-\psi(\mathbf{u})} \cdot D\ d\mathbf{u}} \\
&= \int_{\mathbb{R}^{+n}} E_{\widetilde{H}^{(\mathbf{u})}}\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}^{(\mathbf{u})}(dy)}\right] \cdot \prod_{j=1}^n E\left[e^{-h(y_j)}J_j^{(\mathbf{u}, y_j)}\right] \cdot \frac{D \cdot e^{-\psi(\mathbf{u})}\ d\mathbf{u}}{\int_{\mathbb{R}^{+n}} e^{-\psi(\mathbf{u})} \cdot D\ d\mathbf{u}}
\end{aligned}
$$

where $\Delta^{\star(1)}_{\theta_j}(\mathbf{u}, y) = \exp(\theta_j y)\int_{\mathbb{R}^+} s e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}}\rho(ds \mid y)$; $\quad \nu^{(\mathbf{u})}(ds, dy) = \rho^{(\mathbf{u})}(ds \mid y)H(dy) = e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\rho(ds \mid y)H(dy)$; $\quad e^{-\psi(\mathbf{u})} = \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}}\left[1 - e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\right]\nu(ds, dy)\right\}$.

Hence, finally, $E\left(e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \mid y_1 \in C_1, \ldots, y_n \in C_n\right)$

$$
\begin{aligned}
&= \int_{\mathbb{R}^{+n}} E_{\widetilde{H}^{(\mathbf{u})}}\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}^{(\mathbf{u})}(dy)}\right] \cdot \prod_{j=1}^n E\left[e^{-h(y_j)\,J_j^{(\mathbf{u},y_j)}}\right] \cdot \frac{D \cdot e^{-\psi(\mathbf{u})}\, d\mathbf{u}}{\int_{\mathbb{R}^{+n}} e^{-\psi(\mathbf{u})} \cdot D\, d\mathbf{u}} \\
&= \int_{\mathbb{R}^{+n}} E\left[e^{-\int_{\mathcal{Y}} h(y)\widetilde{H}(dy)} \middle| \mathcal{D}_n, \mathbf{u}\right] \cdot \Pr(\mathbf{u} \mid \mathcal{D}_n)\, d\mathbf{u}
\end{aligned}
$$

Notes: For this proof, we do not need Result 3.1 and 3.2. I think they will be needed when we consider ties in our proof.

2.4 Theorem 1 (with no ties)

Hence, $\widetilde{H} \mid \mathbf{u}, \mathcal{D}_n := \widetilde{H}^{(\mathbf{u},\mathcal{D}_n)} \overset{d}{=} \widetilde{H}^{(\mathbf{u})} + \sum_{i=1}^n J_i^{(\mathbf{u},y_i)}\delta_{y_i}$, where

1. $\widetilde{H}^{(\mathbf{u})} \sim$ CRM $\left(\nu^{(\mathbf{u})}\right)$ with $\nu^{(\mathbf{u})}(ds, dy) = e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\rho(ds \mid y)H(dy)$
2. $\Pr\left(J_j^{(\mathbf{u},y_j)}\right) \propto s e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j)$

2.5 Proposition 1 (with no ties)

Next $\Pr\left(\mathbf{u} \mid \mathcal{D}_n, \widetilde{H}\right) \propto D \cdot e^{-\psi(\mathbf{u})} = \left[\prod_{j=1}^n \int_{\mathbb{R}^+} s e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i y_j)\right\}}\rho(ds \mid y_j)\right] e^{-\psi(\mathbf{u})}$, with $e^{-\psi(\mathbf{u})} = \exp\left\{-\int_{\mathbb{R}^+\times\mathcal{Y}}\left[1 - e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\right]\nu(ds, dy)\right\}$.

Hence, we have a Gibbs sampler ready!

For DP (i.e., gamma CRM on $\widetilde{H}$), calculate $\psi(\mathbf{u})$, ... See 'Posterior sampling steps' document.

2.6 Theorem 1 (with ties)

Suppose, $(z_1, \ldots, z_{\widetilde{n}})$ denotes the unique observations in the data and there are n_j repetitions for z_j, $j = 1, \ldots, \widetilde{n}$. Hence, $\sum_{j=1}^{\widetilde{n}} n_j = n$.

Hence, $\widetilde{H} \mid \mathbf{u}, \mathcal{D}_n := \widetilde{H}^{(\mathbf{u},\mathcal{D}_n)} \overset{d}{=} \widetilde{H}^{(\mathbf{u})} + \sum_{i=1}^{\widetilde{n}} J_i^{(\mathbf{u},z_i)}\delta_{z_i}$, where

1. $\widetilde{H}^{(\mathbf{u})} \sim$ CRM $\left(\nu^{(\mathbf{u})}\right)$ with $\nu^{(\mathbf{u})}(ds, dy) = e^{-\left\{\sum_{i=1}^n u_i \exp(\theta_i y)\right\}s}\rho(ds \mid y)H(dy)$
2. $\Pr\left(J_j^{(\mathbf{u},z_j)}\right) \propto s^{n_i} e^{-s\left\{\sum_{i=1}^n u_i \exp(\theta_i z_j)\right\}}\rho(ds \mid z_j)$

2.7 Proposition 1 (with ties)

Next $\Pr\left(\mathbf{u} \mid \mathcal{D}_n, \widetilde{H}\right) \propto D \cdot e^{-\psi(\mathbf{u})} = \left[\prod_{i=1}^{\widetilde{n}} \cancel{u_i^{n_i-1}} \int_{\mathbb{R}^+} s^{n_i} e^{-s\left\{\sum_{i=1}^{n} u_i \exp(\theta_i y_i)\right\}} \rho(ds \mid z_i)\right] e^{-\psi(\mathbf{u})}$, with $e^{-\psi(\mathbf{u})} = \exp\left\{-\int_{\mathbb{R}^+ \times \mathcal{Y}} \left[1 - e^{-\left\{\sum_{i=1}^{n} u_i \exp(\theta_i y)\right\} s}\right] \nu(ds, dy)\right\}$. [Instead of $\prod_{i=1}^{\widetilde{n}} u_i^{n_i-1}$ should we have $\left(\sum_{i=1}^{n} u_i \exp(\theta_i y_i)\right)^{n-1}$?]

2.8 The u business

$u_j = \frac{\Gamma}{T_j}$, where $\Gamma \sim Gamma(1,1)$ independently of $T_j = \int_{\mathcal{Y}} \exp(\theta_j y) \widetilde{H}(dy)$. Next we use this transformation $(\Gamma, T_j) \to (u_j, T_j)$. Jacobian is $T_j \implies P(u_j, T_j) = T_j \exp(-T_j u_j) P(T_j)$. This further implies $u_j \mid T_j \sim Gamma(1, T_j) \implies \int_{R^+} T_j \exp(-T_j u_j) du_j = 1$ and we started with $\frac{1}{T_j} = \int_{R^+} \exp(-T_j u_j) du_j$.

3 θ updating

In Section 2, we assume θ_i is fixed through out. We can add one more layer to our existing model as, $\theta_i \mid \beta, \widetilde{H} \sim N(\xi_i, c)$ with very small c and $\xi_i = b'^{-1}(\mu_i; \widetilde{H})$ with $g(\mu_i) = x_i'\beta$. In this case, we can do Gibbs update: $\beta \to \boldsymbol{\theta} \to \mathbf{u} \to \widetilde{H} \to \beta \to \ldots$