GSK MCM formula

Dong Dai

1 At DMA level

$\bar{y}^{(d)}(t)$ is assumed with Gaussian distribution with mean $\bar{\mu}^{(d)}(t)$ and variance v, where

$$\bar{\mu}^{(d)}(t) = \alpha + \sum_{k=1}^{K} \beta_k \cdot \log\left(1 + \sum_{\ell \leq t} \bar{N}_k^{(d)}(\ell) \cdot \lambda_k^{t-\ell}\right)$$

and

- d: d-th DMA group.
- t: Time t.
- $\bar{y}^{(d)}(t)$: average number of NRx at time t of all HCPs of d-th DMA group.
- k: k-th promotion channel, and $1 \leq k \leq K$ where K is total number of promotion channels.
- $\bar{N}_k^{(d)}(t)$: average number of promotions from k-th promotion channel at time t of all HCPs of d-th DMA group.
- α: Intercept.
- λ_k: Time-decay parameter of channel k. The smaller λ_k, the faster promotion effect decays. And denote $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$.
- $\beta_k^{(d)}$: Promotion impact parameter of channel k. The higher β_k, the more effective of the promotion channel. And denote $\boldsymbol{\beta}^{(d)} = (\beta_1^{(d)}, \ldots, \beta_K^{(d)})$.

The conditional distribution of $\bar{y}^{(d)}(t)$ can be written as

$$f(\bar{y}^{(d)}(t)|\boldsymbol{\beta}^{(d)}, \boldsymbol{\lambda}) = (2\pi v)^{-1/2} \exp\left[-\frac{\left(\bar{y}^{(d)}(t) - \bar{\mu}^{(d)}(t)\right)^2}{2v}\right]$$

Further assume the prior distribution of $\boldsymbol{\beta}^{(d)} \sim N(\boldsymbol{\beta}_0, \tau_1^2)$ and $\boldsymbol{\lambda} \sim N(\boldsymbol{\lambda}, \tau_2^2)$, then the joint distribution can be expressed as

$$\begin{aligned} f(\bar{y}^{(d)}(t), \boldsymbol{\beta}^{(d)}, \boldsymbol{\lambda}) \quad = \quad & (2\pi v)^{-1/2} \exp\left[-\frac{\left(\bar{y}^{(d)}(t) - \bar{\mu}^{(d)}(t)\right)^2}{2v}\right] \\ & \cdot (2\pi\tau_1^2)^{-1/2} \exp\left(-\frac{\|\boldsymbol{\beta}^{(d)} - \boldsymbol{\beta}_0\|_2^2}{2\tau_1^2}\right) \cdot (2\pi\tau_2^2)^{-1/2} \exp\left(-\frac{\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_0\|_2^2}{2\tau_2^2}\right) \end{aligned}$$

We use the Maximum A Posteriori (MAP) Estimation to estimate parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}^{(d)}$. Specifically,

$$\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\beta}^{(d)}} = \underset{\boldsymbol{\lambda}, \boldsymbol{\beta}^{(d)}}{\operatorname{argmax}} \left(\sum_d \sum_t \left[-\frac{1}{2v} \left(\bar{y}^{(d)}(t) - \bar{\mu}^{(d)}(t) \right)^2 \right] W(t) - \frac{\|\boldsymbol{\beta}^{(d)} - \boldsymbol{\beta}_0\|_2^2}{2\tau_1^2} - \frac{\|\boldsymbol{\lambda} - \boldsymbol{\lambda}_0\|_2^2}{2\tau_2^2} \right)$$

where the additional term $W(t)$ is a time dependant function to assign more weight to more recent data.

2 At individual HCP level

$y^{(i)}(t)$ is assumed with Gaussian distribution with mean $\mu^{(i)}(t)$ and variance $v^{(i)}$, where

$$\mu^{(i)}(t) = \alpha^{(i)} + \sum_{k=1}^{K} \beta_k^{(i)} \cdot \log \left(1 + \sum_{\ell \le t} N_k^{(i)}(\ell) \cdot \left(\lambda_k^{(i)} \right)^{t-\ell} \right)$$

and

- i: i-th HCP.
- t: Time t.
- $y^{(i)}(t)$: number of NRx from i-th HCP at time t.
- k: k-th promotion channel, and $1 \le k \le K$ where K is total number of promotion channels.
- $N_k^{(i)}(t)$: number of promotions from k-th promotion channel at time t.
- $\alpha^{(i)}$: Intercept.
- $\beta_k^{(i)}$: Promotion impact parameter of channel k, it has a prior with mean $\hat{\beta}_k^{(d_i)}$, where d_i is the DMA code of i-th HCP. And denote $\boldsymbol{\beta}^{(i)} = (\beta_1^{(i)}, \ldots, \beta_K^{(i)})$.

The Maximum A Posteriori (MAP) Estimation of parameters $\boldsymbol{\beta}^{(i)}$ is

$$\hat{\boldsymbol{\beta}^{(i)}} = \underset{\boldsymbol{\beta}^{(i)}}{\operatorname{argmax}} \left(\sum_i \sum_t \left[-\frac{1}{2v^{(i)}} \left(y^{(i)}(t) - \mu^{(i)}(t) \right)^2 \right] W(t) - \frac{\|\boldsymbol{\beta}^{(i)} - \hat{\boldsymbol{\beta}^{(d_i)}}\|_2^2}{2\tau_i^2} \right)$$