Likelihood in Griffin et al. 2014

1 Model brief from SI of Griffin et al. 2014

Suppose that the studies are indexed by $l = 1, ..., N$. Then the total likelihood of all the studies are just the product of the contribution from each study,

$$L_{total} = \prod_{l=1}^{N} L_l.$$

For easy notation, Griffin et al. 2014 drops the index of study and consider the likelihood of each study, dented by L.

For each site within a study, there can be incidence data and/or prevalence data whose contributions to the likelihood of the study are denoted by L_c and L_p respectively.

1.1 Incidence data

Suppose that within a study, there are sties indexed $j = 1, ..., m$. Let $k = 1, ..., 4$ index the age groups $0 \sim 2$ years, $2 \sim 5$ years, $5 \sim 15$ years and 15 plus. Within these groups, assume that the study presented data in age groups $i = 1, ..., s_k$.

For a given set of parameters, including the EIR for each site and treated proportion, let the model-predicted incidence of clinical malaria in site j and age group i by μ_{ji}, and the person-time at risk and number of events by T_{ji} and y_{ji} respectively. The parameter r denotes the effect of the case detection method, with $r = 1$ for daily ACD and less than 1 for weekly ACD or PCD. Then it is assumed that

$$y_{ji} \sim Poisson(re^u v_{jk(i)} T_{ji} \mu_{ji}),$$

where u denotes the study-level random effect and v_{jk} denotes the age-specific random effects for age group k in site j, with the priors

$$u \sim Normal(0, \sigma_c^2), \quad v_{jk} \sim Gamma(1/\alpha_c, \alpha_c).$$

Letting $\mathbf{y} = \{y_{ji}; j = 1, ..., m, i = 1, ..., s_k, k = 1, ..., 4\}$ and $\mathbf{v} = \{v_{jk}; j = 1, ..., m, k = 1, ..., 4\}$, then the overall likelihood for clinical incidence data in a study is

$$L_c(\mathbf{y}, u, \mathbf{v}) = f(u| - \sigma_c^2/2, \sigma_c^2) \prod_{j=1}^{m} \prod_{k=1}^{4} g(v_{jk}|1/\alpha_c, \alpha_c) \prod_{i=1}^{s_k} P(y_{ji}|ruv_{jk}T_{ji}\mu_{ji}),$$

where f is a log-Normal probability density, g is a Gamma probability density and P is a Poisson probability.

Remark: This is the formula (10) in the SI of Griffin et al. 2014. The notation for u here and in the SI looks confusing, since it represents a Normal random effect earlier but a log-Normal random effect in this formula. But in the Poisson part, e^u earlier is replaced with u in this formula, so this is fine. We shall use u as a Normal random effect throughout this document later on. However, there was also confusion with its mean and standard deviation in the prior specification. Note that, it was specified with $u \sim Normal(0, \sigma_c^2)$ earlier. But $f(u| - \sigma_c^2/2, \sigma_c^2)$ in this formula using a log-Normal density, as a convention, means that in the Normal scale, $u \sim Normal(-\sigma_c^2/2, \sigma_c^2)$. We shall use $u \sim Normal(-\sigma_c^2/2, \sigma_c^2)$ for the likelihood calculation since it leads to an average random effect of $E(e^u) = 1$.

1.2 Prevalence data

Suppose that the data are stratified into 8 age groups with boundaries $0, 2, 5, 10, 15, 20, 30, 40$ years. Let the model-predicted prevalence of infection in site j and age group i be p_{ji}, and the number sampled and positive be n_{ji} and x_{ji} respectively. It is assumed that there is a study-level random effect $w \sim Normal(0, \sigma_p^2)$, and

$$x_{ji} \sim Beta - binomial(q_{ji}\theta, (1 - q_{ji})\theta, n_{ji}),$$

where $q_{ji} = invlogit(logit(p_{ji}) + w)$, with $logit(p) = log(p/(1 - p))$ and $invlogit$ its inverse function, and θ is an over-dispersion parameter.

2 Likelihood calculation

For a little bit more general and easier notations, we assume $u \sim Normal(\mu_c, \sigma_c^2)$, $v_{jk} \sim Gamma(\alpha, \beta)$ and $w \sim Normal(\mu_p, \sigma_p^2)$. Letting $\mathbf{x} = \{x_{ji}; j = 1, ..., m, i = 1, ..., 8\}$, then the likelihood for each study can then be easily calculated as follows,

$$L(\mathbf{y}, \mathbf{x}, u, \mathbf{v}, w) = L_c(\mathbf{y}, u, \mathbf{v})L_p(\mathbf{x}, w),$$

where

$$L_c(\mathbf{y}, u, \mathbf{v}) = f(u|\mu_c, \sigma_c^2) \prod_{j=1}^{m} \prod_{k=1}^{4} g(v_{jk}|\alpha, \beta) \prod_{i=1}^{s_k} P(y_{ji}|re^u v_{jk} T_{ji} \mu_{ji}),$$

with f a Normal probability density, g a Gamma density and P a Poisson density, and

$$L_p(\mathbf{x}, w) = f(w|\mu_p, \sigma_p^2) \prod_{j=1}^{m} \prod_{i=1}^{8} bb(x_{ji}|q_{ji}\theta, (1 - q_{ji})\theta, n_{ji}),$$

with bb a Beta-binomial probability density function as

$$bb(x|a, b, n) = \binom{n}{x} \frac{B(x + a, n - x + b)}{B(a, b)},$$

where B is a beta function.

As suggested in the SI of Griffin et al. 2014, the random effect can be integrated out to work on the marginal likelihood of $\mathbf{y}$ and $\mathbf{x}$ only, i.e.

$$L(\mathbf{y}, \mathbf{x}) = \int_u \int_{\mathbf{v}} \int_w L(\mathbf{y}, \mathbf{x}, u, \mathbf{v}, w) dw d\mathbf{v} du = \int_u \int_{\mathbf{v}} L_c(\mathbf{y}, u, \mathbf{v}) d\mathbf{v} du \int_w L_p(\mathbf{x}, w) dw.$$

Note that we are able to integrate out of Gamma random effect $\mathbf{v}$ analytically using the Poisson-Gamma conjugacy, and the other two Normal random effects numerically using adaptive Gauss-Hermit quadrature.

2.1 Analytical integration over v

Letting $\mathbf{y}_{jk} = \{y_{ji}; i = 1, ..., s_k\}$ and $h(\mathbf{y}_{jk}, v_{jk}) = g(v_{jk}|\alpha, \beta) \prod_{i=1}^{s_k} P(y_{ji}|re^u v_{jk} T_{ji}\mu_{ji})$, we have

$$\begin{aligned}\int_{\mathbf{v}} L_c(\mathbf{y}, u, \mathbf{v}) d\mathbf{v} &= \int_{\mathbf{v}} f(u|\mu_c, \sigma_c^2) \prod_{j=1}^{m} \prod_{k=1}^{4} h(\mathbf{y}_{jk}, v_{jk}) d\mathbf{v} \\ &= f(u|\mu_c, \sigma_c^2) \prod_{j=1}^{m} \prod_{k=1}^{4} \int_{v_{jk}} h(\mathbf{y}_{jk}, v_{jk}) dv_{jk}.\end{aligned}$$

It is easy to show that,

$$\begin{aligned}h(\mathbf{y}_{jk}, v_{jk}) &= g(v_{jk}|\alpha, \beta) \prod_{i=1}^{s_k} P(y_{ji}|re^u v_{jk} T_{ji}\mu_{ji}) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} v_{jk}^{\alpha-1} e^{-\beta v_{jk}} \prod_{i=1}^{s_k} \left[e^{-re^u v_{jk} T_{ji}\mu_{ji}} \frac{(re^u v_{jk} T_{ji}\mu_{ji})^{y_{ji}}}{y_{ji}!} \right] \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} v_{jk}^{\alpha-1} e^{-\beta v_{jk}} e^{-\sum\limits_{i=1}^{s_k} \left(re^u v_{jk} T_{ji}\mu_{ji}\right)} \prod_{i=1}^{s_k} v_{jk}^{y_{ji}} \prod_{i=1}^{s_k} \left[\frac{(re^u T_{ji}\mu_{ji})^{y_{ji}}}{y_{ji}!} \right] \\ &= v_{ji}^{\alpha + \sum\limits_{i=1}^{s_k} y_{ji} - 1} e^{-\left[\beta + \sum\limits_{i=1}^{s_k} (re^u T_{ji}\mu_{ji})\right] v_{jk}} \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^{s_k} \left[\frac{(re^u T_{ji}\mu_{ji})^{y_{ji}}}{y_{ji}!} \right]\end{aligned}$$

Using the fact that

$$\int_x g(x|\alpha, \beta) dx = \int_x \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = 1,$$

we have

$$\int_x x^{\alpha-1} e^{-\beta x} dx = \frac{\Gamma(\alpha)}{\beta^\alpha}.$$

Therefore,

$$\begin{aligned}\int_{v_{jk}} h(\mathbf{y}_{jk}, v_{jk}) dv_{jk} &= \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^{s_k} \left[\frac{(re^u T_{ji}\mu_{ji})^{y_{ji}}}{y_{ji}!} \right] \int_{v_{jk}} v_{ji}^{\alpha + \sum\limits_{i=1}^{s_k} y_{ji} - 1} e^{-\left[\beta + \sum\limits_{i=1}^{s_k} (re^u T_{ji}\mu_{ji})\right] v_{jk}} dv_{jk} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \prod_{i=1}^{s_k} \left[\frac{(re^u T_{ji}\mu_{ji})^{y_{ji}}}{y_{ji}!} \right] \frac{\Gamma(\alpha + \sum\limits_{i=1}^{s_k} y_{ji})}{\left[\beta + \sum\limits_{i=1}^{s_k} (re^u T_{ji}\mu_{ji})\right]^{\alpha + \sum\limits_{i=1}^{s_k} y_{ji}}}.\end{aligned}$$

Remark: One implication of the above is that the posterior distribution of v_{jk} given all the other parameters and values is just a Gamma distribution,

$$v_{jk}|\mathbf{y}_{jk}, u, ... \sim Gamma(\alpha + \sum_{i=1}^{s_k} y_{ji}, \beta + \sum_{i=1}^{s_k} (re^u T_{ji}\mu_{ji}))$$

2.2 Numerical integration over u and w

We know that the standard Gauss-Hermit quadrature can approximate the following function

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i),$$

where $\{(x_i, w_i); i = 1, ..., n\}$ are n pairs of Gauss-Hermit quadrature nodes and weights.

Since u and w are both specified with Normal prior, they can be integrated out using adaptive Gauss-Hermit quadrature. In a general form, we aim to approximate an integration in the following form,

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} f(x)dx.$$

Letting $x' = \frac{x-\mu}{\sqrt{2}\sigma}$, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} f(x)dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\{-x'^2\} f(\sqrt{2}\sigma x' + \mu)\sqrt{2}\sigma dx' \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\{-x'^2\} f(\sqrt{2}\sigma x' + \mu) dx' \\
&\approx \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} w_i f(\sqrt{2}\sigma x_i + \mu),
\end{aligned}$$

where $\{(x_i, w_i); i = 1, ..., n\}$ are n pairs of Gauss-Hermit quadrature nodes and weights.

Therefore, we just need to transform the standard Gauss-Hermit quadrature nodes and weights as follows,

$$w_i' = \frac{1}{\sqrt{\pi}} w_i, \quad x_i' = \sqrt{2}\sigma x_i + \mu.$$

In this way, we can numerically integrate out the Normal random effect from the likelihood by adapting the standard Gauss-Hermit quadrature using their mean and standard deviation.

Remark: In the code implementation, we used the nodes and weights generated by calling the *gauss.quad.prob(...)* function in the *statmod* package, e.g.

```
gh <- gauss.quad.prob(100, dist="normal")
```

Note that the generated nodes and weights are already adapted to standard Normal distribution. Thus the transformation is slightly different.

In fact, the nodes and weights generated from the above code, denoted by $\{(x_i', w_i'); i = 1, ..., n\}$, are used to approximate the integration in the following form,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} f(x)dx \approx \sum_{i=1}^{n} w_i' f(x_i').$$

In order to approximate

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} f(x)dx,$$

we just take transformation $x'' = \frac{x-\mu}{\sigma}$, then

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} f(x)dx$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{x''^2}{2}\} f(\sigma x'' + \mu)\sigma dx''$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x''^2}{2}\} f(\sigma x'' + \mu) dx''$$
$$\approx \sum_{i=1}^{n} w_i' f(\sigma x_i' + \mu).$$

Therefore, the transformation on the weights and nodes are

$$w_i'' = w_i', \quad x_i'' = \sigma x_i' + \mu.$$

3 Prior specification for random effect

The prior specifications for some hyper-parameters in the likelihood are presented below.

$$\alpha_c/5 \sim half\ t_6,$$

$$\sigma_c/5 \sim half\ Normal(0,1),$$
$$\sigma_p/5 \sim half\ Normal(0,1),$$

and

$$\alpha_p = 1/\theta \sim half\ t_6,$$

where t_6 is the t-distribution with $df = 6$, and "half" means the truncated distribution on $(0, \infty)$.