1 Model in Ash

The model considered in Ash is for each i,

$$x_i \sim \mathcal{N}(\alpha_i, y_i^2). \tag{1}$$

Note that in a PoissonBinomial model, $x_i = \hat{\alpha}_i$ and $y_i = se(\hat{\alpha}_i)$ are estimate (MLE) for $\alpha_i = logit(p_i)$ and its standard error which are estimated by using a glm function in R. Ash considers a mixture of normal distributions as a prior on α_i. Specifically, for each i,

$$\alpha_i \mid \pi, \sigma^2 \sim \sum_{m=1}^{M} \pi_m \mathcal{N}(0, \sigma_m^2), \tag{2}$$

where $\pi = (\pi_1, \ldots, \pi_M)$ are the mixture proportions which are constrained to be non-negative and sum to one and $\sigma^2 = (\sigma_1^2, \ldots, \sigma_M^2)$ are the variances for each normal distribution. For now, Ash assumes that σ^2 is known and estimates π by using an empirical Bayes procedure.

2 EM algorithm in Ash

The MLE for π can be obtained by using the following EM algorithm. Let $D_i = (x_i, y_i)$ and $D = (D_1, \ldots, D_n)$. Consider unobserved latent variables $Z = (Z_1, \ldots, Z_n)$, where $Z_i \in \{1, \ldots, M\}$ and $\mathsf{P}(Z_i = m) = \pi_m$. Then, a complete data likelihood can be written

$$\begin{aligned} \mathsf{P}(D, Z \mid \pi) &= \prod_{i=1}^{n} \mathsf{P}(D_i, Z_i \mid \pi) & (3)\\ &= \prod_{i=1}^{n} \prod_{m=1}^{M} \mathsf{P}(D_i, Z_i = m \mid \pi)^{I(Z_i=m)} & (4)\\ &= \prod_{i=1}^{n} \prod_{m=1}^{M} [\mathsf{P}(D_i, \mid Z_i = m, \pi)\pi_m]^{I(Z_i=m)}, & (5) \end{aligned}$$

yielding a log likelihood

$$\log \mathsf{P}(D, Z \mid \pi) = \sum_{i=1}^{n} \sum_{m=1}^{M} I(Z_i = m)[\log \mathsf{P}(D_i, \mid Z_i = m) + \log \pi_m]. \tag{6}$$

E-step: For each i and m,

$$\begin{aligned} \mathsf{P}(Z_i = m \mid D_i, \pi^l) &= \frac{\mathsf{P}(Z_i = m, D_i \mid \pi^l)}{\sum_{n=1}^{M} \mathsf{P}(Z_i = n, D_i \mid \pi^l)}, & (7)\\ &= \frac{\pi_m^l \mathsf{P}(D_i \mid Z_i = m)}{\sum_{n=1}^{M} \pi_n^l \mathsf{P}(D_i \mid Z_i = n)}, & (8)\\ &= \frac{\pi_m^l \mathrm{BF_i}(\sigma_\mathrm{m}^2)}{\sum_{n=1}^{M} \pi_n^l \mathrm{BF_i}(\sigma_\mathrm{n}^2)}, & (9) \end{aligned}$$

where

$$\mathrm{BF_i}(\sigma_\mathrm{m}^2) = \frac{\mathsf{P}(\mathrm{D_i} \mid \mathrm{Z_i} = \mathrm{m})}{\mathsf{P}(\mathrm{D_i} \mid \alpha_\mathrm{i} = 0)}. \quad (10)$$

M-step: Find the parameters π which maximizes $\mathrm{E}_{Z|D,\pi_l}[\log \mathsf{P}(D, Z \mid \pi)]$.

$$\begin{aligned} \pi^{l+1} &= \underset{\pi}{\operatorname{argmax}} \, \mathrm{E}_{Z|D,\pi^l}[\log \mathsf{P}(D, Z \mid \pi)], & (11)\\ &= \underset{\pi}{\operatorname{argmax}} \sum_{i=1}^{n} \sum_{m=1}^{M} A_{im}[\log \mathrm{BF_{im}} + \log \pi_m], & (12)\\ &= \underset{\pi}{\operatorname{argmax}} \, Q(\pi \mid \pi^l), & (13) \end{aligned}$$

where $A_{im} = \mathsf{P}(Z_i = m \mid D_i, \pi^l)$ and $\mathrm{BF_{im}} = \mathrm{BF_i}(\sigma_\mathrm{m}^2)$. For each $m = 1, \ldots, M-1$,

$$\begin{aligned} \frac{\partial Q(\pi \mid \pi^l)}{\partial \pi_m} &= \sum_{i=1}^{n} [\frac{A_{im}}{\pi_m} + \frac{-A_{iM}}{\pi_M}], & (14)\\ &= \sum_{i=1}^{n} [\frac{A_{im}\pi_M - A_{iM}\pi_m}{\pi_m \pi_M}], & (15)\\ &= \frac{\pi_M \sum_{i=1}^{n} A_{im} - \pi_m \sum_{i=1}^{n} A_{iM}}{\pi_m \pi_M}, & (16) \end{aligned}$$

where

$$\begin{aligned} A_{iM} &= 1 - (A_{i1} + \ldots, +A_{i(M-1)}), & (17)\\ \pi_{iM} &= 1 - (\pi_{i1} + \ldots, +\pi_{i(M-1)}). & (18) \end{aligned}$$

Then,

$$\pi_M \sum_{i=1}^{n} A_{im} = \pi_m \sum_{i=1}^{n} A_{iM} \quad \text{for} \quad m = 1, \ldots, M-1. \quad (19)$$

Summing $M-1$ equations in (19) leads to

$$\pi_M[\sum_{i=1}^{n}\sum_{m=1}^{M-1} A_{im}] = (1-\pi_M)\sum_{i=1}^{n} A_{iM}. \tag{20}$$

Then,

$$\begin{aligned}\pi_M &= \frac{\sum_{i=1}^{n} A_{iM}}{\sum_{i=1}^{n}\sum_{m=1}^{M} A_{im}}, &(21)\\ &= \frac{\sum_{i=1}^{n} A_{iM}}{n}, &(22)\end{aligned}$$

and for each $m = 1, \ldots, M-1$,

$$\pi_m = \frac{\sum_{i=1}^{n} A_{im}}{n}. \tag{23}$$

3 Likelihood approximation in a PoissonBinomial Model

Under a PoissonBinomial model, a log likelihood function for $\alpha_i = logit(p_i)$ can be written as

$$\begin{aligned}f(\alpha_i) &= \log P(D_i \mid \alpha_i), &(24)\\ &= \log \binom{n_i}{x_i}\left(\frac{1}{1+\exp^{-\alpha}}\right)^{x_i}\left(\frac{\exp^{-\alpha}}{1+\exp^{-\alpha}}\right)^{n_i-x_i}. &(25)\end{aligned}$$

Taking three elements in Taylor series of $f(\alpha_i)$ about a MLE $\hat{\alpha}_i$, $f(\alpha)$ can be approximated by

$$\begin{aligned}f(\alpha_i) &\approx f(\hat{\alpha}_i) + \frac{f''(\hat{\alpha}_i)(\alpha_i-\hat{\alpha}_i)^2}{2}, &(26)\\ &\approx f(\hat{\alpha}_i) - \frac{(\alpha_i-\hat{\alpha}_i)^2}{2se(\hat{\alpha}_i)^2}. &(27)\end{aligned}$$

Then, a likelihood function for α_i, $l(\alpha_i)$ can be approximated by

$$l(\alpha_i) \approx \mathcal{N}(\hat{\alpha}_i, se(\hat{\alpha}_i)^2). \tag{28}$$

4 BF and posterior prob approximation in Ash

This section describes derivations for the approximate Bayes Factor (ABF) and posterior on α_i when $Z_i = m$ in the prior from equation (2).

$$
\begin{aligned}
\mathsf{P}(D_i \mid Z_i = m) &= \int \frac{C}{\sqrt{2\pi se(\hat{\alpha}_i)^2}} \exp[-\frac{(\alpha_i - \hat{\alpha}_i)^2}{2se(\hat{\alpha}_i)^2}] \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp[-\frac{\alpha_i^2}{2\sigma_m^2}] \mathrm{d}\alpha_i, && (29)\\
&= \frac{C}{2\pi\sqrt{se(\hat{\alpha}_i)^2\sigma_m^2}} \int \exp[-\frac{\sigma_m^2 (\alpha_i - \hat{\alpha}_i)^2 + se(\hat{\alpha}_i)^2\alpha_i^2}{2se(\hat{\alpha}_i)^2\sigma_m^2}] \mathrm{d}\alpha_i, && (30)\\
&= \frac{C}{2\pi\sqrt{se(\hat{\alpha}_i)^2\sigma_m^2}} \int \exp[-\frac{(\sigma_m^2 + se(\hat{\alpha}_i)^2)\, \alpha_i^2 - 2\sigma_m^2\hat{\alpha}_i\alpha_i + \sigma_m^2\hat{\alpha}_i^2}{2se(\hat{\alpha}_i)^2\sigma_m^2}] \mathrm{d}\alpha_i, && (31)\\
&= \frac{C}{2\pi\sqrt{se(\hat{\alpha}_i)^2\sigma_m^2}} \int \exp[-\frac{(\sigma_m^2 + se(\hat{\alpha}_i)^2)\, [\alpha_i - \frac{\sigma_m^2\hat{\alpha}_i}{\sigma_m^2 + se(\hat{\alpha}_i)^2}]^2 + \frac{\sigma_m^2\hat{\alpha}_i^2 se(\hat{\alpha}_i)^2}{\sigma_m^2 + se(\hat{\alpha}_i)^2}}{2se(\hat{\alpha}_i)^2\sigma_m^2}] \mathrm{d}\alpha_i, && (32)\\
&= \frac{C\sqrt{2\pi\frac{\sigma_m^2 se(\hat{\alpha}_i)^2}{\sigma_m^2 + se(\hat{\alpha}_i)^2}}}{2\pi\sqrt{se(\hat{\alpha}_i)^2\sigma_m^2}} \exp[-\frac{\hat{\alpha}_i^2}{2(se(\hat{\alpha}_i)^2 + \sigma_m^2)}], && (33)\\
&= \frac{C}{\sqrt{2\pi(\sigma_m^2 + se(\hat{\alpha}_i)^2)}} \exp[-\frac{\hat{\alpha}_i^2}{2(se(\hat{\alpha}_i)^2 + \sigma_m^2)}], && (34)
\end{aligned}
$$

and

$$
\mathsf{P}(D_i \mid \alpha_i = 0) = \frac{C}{\sqrt{2\pi se(\hat{\alpha}_i)^2}} \exp[-\frac{\hat{\alpha}_i^2}{2se(\hat{\alpha}_i)^2}]. \qquad (35)
$$

Then, ABF can be written as

$$
\begin{aligned}
\mathrm{BF_i}(\sigma_\mathrm{m}^2) &= \frac{\mathsf{P}(D_i \mid Z_i = m)}{\mathsf{P}(D_i \mid \alpha_i = 0)}, && (36)\\
&= \sqrt{\frac{se(\hat{\alpha}_i)^2}{\sigma_m^2 + se(\hat{\alpha}_i)^2}} \exp[\frac{\hat{\alpha}_i^2}{2se(\hat{\alpha}_i)^2} \frac{\sigma_m^2}{\sigma_m^2 + se(\hat{\alpha}_i)^2}], && (37)\\
&= \sqrt{\lambda} \exp[T^2(1-\lambda)/2], && (38)
\end{aligned}
$$

where

$$
\begin{aligned}
\lambda &= \frac{se(\hat{\alpha}_i)^2}{se(\hat{\alpha}_i)^2 + \sigma_m^2}, && (39)\\
T &= \frac{\hat{\alpha}_i}{se(\hat{\alpha}_i)}. && (40)
\end{aligned}
$$

And a posterior on α_i is

$$\begin{aligned}
\mathsf{P}(\alpha_i \mid D_i, Z_i = m) &\propto l(\alpha_i)\mathsf{P}(\alpha_i \mid Z_i), & (41)\\
&\propto \exp\,[-\frac{(\alpha_i - \hat{\alpha}_i)^2}{2se(\hat{\alpha}_i)^2}]\exp\,[-\frac{\alpha_i^2}{2\sigma_m^2}], & (42)\\
&\propto \exp\,[-\frac{(\sigma_m^2 + se(\hat{\alpha}_i)^2)\,[\alpha_i - \frac{\sigma_m^2\hat{\alpha}_i}{\sigma_m^2+se(\hat{\alpha}_i)^2}]^2}{2se(\hat{\alpha}_i)^2\sigma_m^2}], & (43)
\end{aligned}$$

leading to

$$\alpha_i \mid D_i, Z_i = m \sim \mathcal{N}(\frac{\sigma_m^2\hat{\alpha}_i}{\sigma_m^2 + se(\hat{\alpha}_i)^2}, \frac{\sigma_m^2 se(\hat{\alpha}_i)^2}{\sigma_m^2 + se(\hat{\alpha}_i)^2}). \quad (44)$$

References