

We borrow some function names from R to indicate the corresponding probability functions ***dbinom*** (for the binomial distribution) and ***dhyper*** (for the hypergeometric distribution), and a density function ***dbeta*** (for the beta distribution).

The prior distribution for the probabilities \mathbf{p} is dirichlet with parameter $\mathbf{a} = (a_1, a_2, a_3, a_4)$. We take $a_1 = a_2 = a_3 = a_4 = 1$, a “non-informative” (or flat, or “non-opinionated”) prior distribution. Here the dirichlet will factor into a beta distribution for (p_D, P_L) and another beta distribution for (p_R, P_N) . This is handy here, because the data distribution factors into a binomial distribution for (n_D, n_L) , another binomial distribution for (n_R, n_N) , and a hypergeometric distribution for n_{RD} .

We let $Beta$ denote the beta distribution normalizer constant, defined by

$$Beta(a_1, a_2) = \frac{Gamma(a_1)Gamma(a_2)}{Gamma\left(\sum_i a_i\right)} = \frac{(a_1 - 1)!(a_2 - 1)!}{\left(\sum_i a_i - 1\right)!},$$

where $Gamma$ is the standard gamma function, with the convenient property $Gamma(n) = (n - 1)!$.

$$\begin{aligned} m_{Lump} &= \Pr(\mathbf{n}|\phi = Lump) = \int \Pr(\mathbf{n}|\mathbf{p}, \phi = Lump) \Pr(\mathbf{p}) \\ &= \int \Pr(n_R|\mathbf{p}, \phi = Lump) \Pr(n_D|\mathbf{p}, \phi = Lump) \Pr(n_{RD}|n_R, n_D, \phi = Lump) \Pr(\mathbf{p}) \\ &= \int_{\mathbf{p}} dbinom(n_R, n, p_R) dbeta(n_D, n, p_D) dhyper(n_{RD}, n_R, n_N, n_D) ddirichlet(\mathbf{p}, \mathbf{a}) \\ &= \int_{\mathbf{p}} \frac{n!}{n_R!n_N!} p_R^{n_R} p_N^{n_N} \cdot \frac{n!}{n_D!n_L!} p_D^{n_D} p_L^{n_L} \cdot \frac{\binom{n_R}{n_{RD}} \binom{n_L}{n_{LD}}}{\binom{n}{n_D}} \cdot \frac{p_R^{a_R-1} p_N^{a_N-1}}{Beta(a_R, a_N)} \cdot \frac{p_D^{a_D-1} p_L^{a_L-1}}{Beta(a_D, a_L)} \\ &= \frac{n!}{n_R!n_N!} \frac{n!}{n_D!n_L!} \frac{\binom{n_R}{n_{RD}} \binom{n_N}{n_{ND}}}{\binom{n}{n_D}} \frac{Beta(n_R+a_R, n_N+a_N)}{Beta(a_R, a_N)} \cdot \frac{Beta(n_D+a_D, n_L+a_L)}{Beta(a_D, a_L)} \end{aligned}$$