COMP 364: Computer Tools for the Life Sciences

Solution: Assignment 4, Question 5

Winter 2016

Recall the model described in lecture:

$$\mathbb{P}[b_i|e_i, r, m, f] = \begin{cases} f \cdot \frac{e_i}{3} + (1-f)(1-e_i) & : b_i = r \\ f(1-e_i) + (1-f)\frac{e_i}{3} & : b_i = m \\ \frac{e_i}{3} & : \text{otherwise} \end{cases}$$

For ease of exposition, assume that $r = A, m = T$. There are three cases. The otherwise case occurs when either $b_i = G$ or $b_i = C$. Hence, there are two such situations contributing $\frac{e_i}{3} \cdot 2 = \frac{2}{3}e_i$.

The $b_i = r$ case contributes $(1-f)(1-e_i) + f(\frac{e_i}{3})$. The $b_i = m$ case contributes $f(1-e_i) + (1-f)(\frac{e_i}{3})$.

Proof *The sum of all terms in the MuTect model, is equal to 1.* We want to show:

$$\frac{2}{3} \cdot e_i + (1-f)(1-e_i) + f(\frac{e_i}{3}) + f(1-e_i) + (1-f)(\frac{e_i}{3}) = 1$$

Re-arranging terms, we have:

$$(1-f)(1-e_i) + (1-f)(\frac{e_i}{3}) + f(\frac{e_i}{3}) + f(1-e_i) + \frac{2}{3} \cdot e_i = 1$$

Now, factoring:

$$(1-f)(1-e_i + \frac{e_i}{3}) + f(\frac{e_i}{3} + 1 - e_i) + \frac{2}{3} \cdot e_i = 1$$

Simplifying:

$$(1-f)(1 - \frac{2}{3} \cdot e_i) + f(1 - \frac{2}{3} \cdot e_i) + \frac{2}{3} \cdot e_i = 1$$

Factoring again:

$$(1 - \frac{2}{3} \cdot e_i)(1 - f + f) + \frac{2}{3} \cdot e_i = 1$$

And, finally, simplifying further:

$$\begin{aligned} (1 - \frac{2}{3} \cdot e_i) + \frac{2}{3} \cdot e_i &= 1 \\ 1 &= 1 \end{aligned}$$

$\square$