Segmenter equations

Contents

Clonal - Likelihood

π = purity, σ = ploidy, n_A = N Allele A, n_B = N Allele B

BAF

BAF likelihood of SNPs in a segment S divided in J bins:

$$L(b|\phi) = \prod_{j=1}^{J} Beta(b_j|\phi)$$

$\phi = (n_A, n_B, \pi, n_{SNP_j})$

the distribution has 2 parameters (shape) α_{BAF} and β_{BAF} defined as:

$$\alpha_{BAF} = \frac{(DP_j - 2) \times E_{BAF} + 1}{1 - E_{BAF}}, \beta_{BAF} = DP_j$$

where

$$E_{BAF} = \frac{n_B \pi + (1 - \pi)}{(n_A + n_B)\pi + 2(1 - \pi)}$$

if $\pi = 1$, $E_{BAF} = \frac{n_B}{n_A + n_B}$

DR

DR likelihood for SNPs in segment S divided in J bins:

$$L(d|\phi, \sigma) = \prod_{j=1}^{J} \Gamma(d_j|\phi, \sigma)$$

$\phi = (n_A, n_B, \pi, n_{SNP_j})$

where:

- shape, k or $\alpha = E_{DR} \times \sqrt{DP_j} + 1$
- scale $\sigma = \frac{1}{\sqrt{DP_j}}$ or rate $\beta = \sqrt{DP_j}$

$$E_{DR} = \frac{(n_A + n_B)\pi + 2(1-\pi)}{\sigma}$$

if $\pi = 1$, $E_{DR} = \frac{N_A+N_B}{\sigma}$

VAF

The likelihood for the number of reads nv_j mapping on a SNV j with coverage dp_j in a segment S divided in J bins:

$$L(nv|dp, v) = \prod_{j=1}^{J} \sum_{m=1}^{M} v_m Bin(nv_j|dp_j, v_m)$$

$$v_m = \frac{m\pi}{(n_A + n_B)\pi + 2(1-\pi)}$$

if $\pi = 1$, clonal_peaks $= \frac{m}{n_A+n_B}$ where $m \in 1, 2$ is the multiplicity of the SNV.

In binomial: $dp = n$, $nv = k$, $\phi_p = p$

MAF

Subclonal - Likelihood

$$n_{A1}, n_{B1}, n_{A2}, n_{B2}, \rho, \sigma$$

ρ = CFF of sub-clonal, σ = ploidy

BAF

$$E_{BAF} = \frac{min(n_{A1}\rho * n_{A2}(1-\rho), n_{B1}\rho * n_{B2}(1-\rho))\pi + (1-\pi)}{(\rho(n_{A1} + n_{B1}) + (1-\rho)(n_{A2} + n_{B2}))\pi + 2(1-\pi)}$$

DR

$$E_{DR} = \frac{(\rho(n_{A1} + n_{B1}) + (1-\rho)(n_{A2} + n_{B2}))\pi + 2(1-\pi)}{\sigma}$$

VAF

Shared mutations:

$$v_{m1,m2} = \frac{(m_1\rho + m_2(1-\rho))\pi}{[\rho(n_{A1} + n_{B1}) + (1-\rho)(n_{A2} + n_{B2})]\pi + 2(1-\pi)}$$

Private mutations of clone i:

$$v_{m_i} = \cdot \frac{m_i\rho_i\pi}{[\rho(n_{A1} + n_{B1}) + (1-\rho)(n_{A2} + n_{B2})]\pi + 2(1-\pi)}$$