About Section 4.6 of Monograph (Thompson, 2000)

Anne-Louise Leutenegger

March 28, 2001

Homozygosity Mapping with a Single Marker

The expected lod score for some ρ, when the true value of the recombination fraction is r, is computed as a weighted average of the log-likelihood ratios (lod scores) for each possible outcome:

$$\begin{aligned} Elod(\rho) =& \mathrm{E}_r\left(\mathrm{lod}(\rho)\right) \\ =& \mathrm{E}_r\left(\log(\frac{L(\rho)}{L(\rho=\frac{1}{2})})\right) \\ =& \sum_{\text{data}} \Pr(\text{data}; r) \log\left(\frac{L(\rho)}{L(\rho=\frac{1}{2})}\right) \end{aligned}$$

The weights are the probabilities of each outcome computed with the true value of the recombination fraction.

For each affected individual with inbreeding coefficient f, the contribution to the $Elod$ is outlined in what follows. If the individual has heterozygous marker phenotype A_jA_l, the data probabilities are:

$$\begin{aligned} \Pr(\text{data}; r) =& \frac{\Pr(A_jA_l\,\&\,\text{affected individual}; r)}{\Pr(\text{Affected individual})} \\ =& \frac{2p_jp_l(q(f-f_2(r)) + q^2(1-2f+f_2(r)))}{qf + q^2(1-f)} \end{aligned} \tag{1}$$

and, the likelihood ratio is as presented in equation (4.6) of the monograph (Thompson, 2000):

$$\frac{L(\rho)}{L(\rho=\frac{1}{2})} = \frac{(f-f_2(\rho)) + q(1-2f+f_2(\rho))}{(1-f)(f+q(1-f))} \tag{2}$$

Similarly, if the individual has homozygous marker phenotype A_jA_j:

$$\Pr(\text{data}; r) = \frac{qp_jf_2(r) + q^2p_j(f-f_2(r)) + qp_j^2(f-f_2(r)) + q^2p_j^2(1-2f+f_2(r))}{qf+q^2(1-f)} \tag{3}$$

and, the likelihood ratio is as shown in equation (4.7) of the monograph (Thompson, 2000):

$$\frac{L(\rho)}{L(\rho=\frac{1}{2})} = \frac{f_2(\rho)+q(f-f_2(\rho))+p_j(f-f_2(\rho))+qp_j(1-2f+f_2(\rho))}{f^2+qf(1-f)p_jf(1-f)+qp_j(1-f)^2} \tag{4}$$

Consider the *Elod* at $\rho = 0$ when zero is the true value of the recombination (i.e. $\rho = r = 0$). Then, data probabilities (1) are:

$$\Pr(\text{data}; 0) = \frac{2p_jp_lq^2(1-f)}{qf+q^2(1-f)}$$

and, the likelihood ratio (2) is:

$$\frac{q(1-f)}{(1-f)(f+q(1-f))}$$

Similarly, data probabilities (3) are:

$$\Pr(\text{data}; 0) = \frac{qp_jf+q^2p_j^2(1-f)}{qf+q^2(1-f)}$$

and, the likelihood ratio (4) is:

$$\frac{f+qp_j(1-f)}{(f+(1-f)q)(f+(1-f)p_j)}$$

Hence, the *Elod* for each affected individual is:

$$\begin{aligned}
Elod(0) &= \sum_j p_j \frac{f+qp_j(1-f)}{f+q(1-f)} \log\left(\frac{f+qp_j(1-f)}{(f+(1-f)q)(f+(1-f)p_j)}\right) \\
&\qquad + \sum_{j<l} 2p_jp_l \frac{q(1-f)}{f+q(1-f)} \log\left(\frac{q}{f+q(1-f)}\right) \\
&= \sum_j p_j \frac{f+qp_j(1-f)}{f+q(1-f)} \log\left(\frac{f+qp_j(1-f)}{(f+(1-f)q)(f+(1-f)p_j)}\right) \\
&\qquad + \frac{q(1-f)}{f+q(1-f)} \log\left(\frac{q}{f+q(1-f)}\right)(1-\sum_j p_j^2)
\end{aligned}$$

As $q \to 0$, $Elod(0) \to -\sum_j p_j \log(f+(1-f)p_j)$.

References

Thompson, E. A. (2000). *Statistical Inferences from Genetic Data on Pedigrees.* I.M.S.