Solution of assignment 5, ST2304

Problem 2 The likelihood function can be simplified to

$$\begin{aligned} L(p) &= \frac{n!}{x_{AA}!x_{Aa}!x_{aa}!} p^{2x_{AA}} 2p^{x_{Aa}}(1-p)^{x_{Aa}}(1-p)^{2x_{aa}} \\ &= \frac{n!}{x_{AA}!x_{Aa}!x_{aa}!} 2p^{2x_{AA}+x_{Aa}}(1-p)^{x_{Aa}+2x_{aa}}. \end{aligned} \tag{1}$$

Taking logs, the log likelihood becomes

$$\begin{aligned} \ln L(p) = \ln n! - \ln x_{AA}! - \ln x_{Aa}! - \ln x_{aa}! - \ln 2 \\ + (2x_{AA} + x_{Aa}) \ln p + (x_{Aa} + 2x_{aa}) \ln(1-p) \end{aligned} \tag{2}$$

The likelihood has its maximum when

$$\begin{aligned} \frac{d}{dp} \ln L(p) &= 0 \\ \frac{2x_{AA} + x_{Aa}}{p} - \frac{x_{Aa} + 2x_{aa}}{1-p} &= 0 \end{aligned} \tag{3}$$

or, letting x_A and x_a denote the total number of A and a-alleles in the sample,

$$\begin{aligned} \frac{x_A}{p} - \frac{x_a}{1-p} &= 0 \\ x_A(1-p) &= x_a p \\ x_A &= (x_a + x_A)p \\ \hat{p} &= \frac{x_A}{x_A + x_a} = \frac{x_A}{2n}. \end{aligned} \tag{4}$$

that is, provided that the population is in Hardy-Weinberg equilibrium, the MLE of p is equal to the sample frequency of A.