APPROACH 1

let π_{ijk} , where $(i, j, k) \in \{0,1\}^3$,

~~indicate the frequency of allelic pair (i, j)~~

be the proportion of gametes ~~individuals~~, over the 2 populations, who have allelic pair (i, j) and are in population k.

$$\pi_{ijk} = \frac{1}{2N} \sum_{n=1}^{2N} \mathbb{1}\{\text{gamete } \cancel{\text{individual}}\ n \text{ has alleles } (i,j) \text{ and is in population } k\}$$

The null H_0 is that the 2 populations are in equilibrium

H_0 : $\pi_{ijk} = \alpha_i \beta_j \gamma_k$ s.t. $\alpha_0 + \alpha_1 = 1$, $\beta_0 + \beta_1 = 1$, $\gamma_0 + \gamma_1 = 1$

$\Rightarrow \dim H_0 = 3$

H_1 : π_{ijk} for $(i,j,k) \in \{0,1\}^3$ are unconstrained

i.e. $\sum_{i,j,k} \pi_{ijk} = 1$

$\Rightarrow \dim H_1 = 7$

Let's use a likelihood ratio statistic :

$$\Lambda = 2\left(\sup_{H_1} \ell(\pi) - \sup_{H_0} \ell(\pi) \right) \approx \chi^2_{7-3}$$

Plugging in MLE estimates $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$:

$$\Lambda = 2 \left(\sum_{i=0}^{1} \sum_{j=0}^{1} \sum_{k=0}^{1} n_{ijk} \log\left(\frac{n_{ijk}}{2N}\right) - \right.$$

$$\left. - \sum_{i} \sum_{j} \sum_{k} n_{ijk} \log\left(\frac{n_{i\cdot\cdot}\, n_{\cdot j \cdot}\, n_{\cdot\cdot k}}{(2N)^3}\right) \right)$$

where n_{ijk} = # of the 2N gametes in category (i, j, k)

$n_{\cdot j \cdot}$ = # of the 2N gametes with 2nd variable equal to j

$$= 2 \sum_{i} \sum_{j} \sum_{k} O_{ijk} \log\left(O_{ijk} / E_{ijk} \right)$$

where O_{ijk} = observed # gametes in category (i, j, k)

E_{ijk} = expected # " " " " under H_0 using $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$

Using the fact that $x \log\left(\frac{x}{a}\right) \approx (x-a) + \frac{(x-a)^2}{2a}$

$$\Lambda \approx \sum_{i} \sum_{j} \sum_{k} \frac{(O_{ijk} - E_{ijk})^2}{2 E_{ijk}} \approx \chi^2_{7-3} \text{ under } H_0$$

$\Rightarrow$ joint $r^2 := \mathbb{P}(\Lambda \leq \chi^2_4)$