Sved framework :

Suppose we start with $2N_e$ i.i.d. gametes

$$\left(Z_i\right)_{i=1}^{2N_e} = \left(\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \right)_{i=1}^{2N_e}$$

where $\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \overset{iid}{\sim}$ Multinom $\left(p_{00}, p_{01}, p_{10}, p_{11} \right)$

i.e. $\mathbb{P}\left(\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = p_{00}$ etc.

Suppose we sample Z and $\tilde{Z}$ uniformly from

$\{ Z_1, Z_2, \dots, Z_{2N_e} \}$

Then because $\uparrow$ are iid :

$\text{corr}(Z, \tilde{Z}) = 1 \iff$ we sampled the same r.v.

$\text{corr}(Z, \tilde{Z}) = 0 \iff$ otherwise

Therefore, if we assume $\underline{\underline{c = 0}}$, at any generation, if we sample $Z, \tilde{Z}$ uniformly from the gamete pool : $\text{corr}(Z, \tilde{Z}) = \mathbb{P}(\text{IBD})$

$\downarrow$

"identity by descent"

Counter example of $r^2 \neq \mathbb{P}(\text{IBD})$

~~[illegible]~~ Explanation: r^2 and $\mathbb{P}(\text{IBD})$ are conceptually VERY DIFFERENT

- r^2

Let X be the r.v. that indicates whether a randomly picked gamete has allele 1 at the 1st locus

Let Y be the analogous for the 2nd locus

$$\Rightarrow \quad r = \text{corr}(X, Y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

where $x_i = \mathbb{1}\{\text{gamete } i \text{ has allele 1 at locus 1}\}$

$y_i = \mathbb{1}\{$ " " " locus 2 $\}$

- $\mathbb{P}(\text{IBD})$ is explained in the previous page

Suppose we have only 2 gametes, which at generation 0 we sample independently from this contingency table in equilibrium

	1/2	1/2
1/3	1/6	1/6
2/3	1/3	1/3

Gen 0 : $\mathbb{E}(r^2) = 0$

$\mathbb{P}(\text{IBD}) = 0$

Gen 1 : Case 1 two children gametes come from same parent

$\Rightarrow \quad \mathbb{P}(\text{IBD}) = 1$

$\mathbb{E}(r^2) = 1$

Case 2 two children gametes come from different parents

$\Rightarrow \quad \mathbb{P}(\text{IBD}) = 0$

$\mathbb{E}(r^2) > 0$