Population Genetics – Chapters 1–4 Cheat Sheet

Yongyan Liu

December 13, 2025

2. Hardy–Weinberg Equilibrium (HWE) (Ch. 2)

Assumptions: Infinite size, random mating, no selection/mutation/migration. Two alleles A, a with $p, q = 1 - p$

Genotype Frequency: $P(AA) = p^2, \quad P(Aa) = 2pq, \quad P(aa) = q^2$

Allele Frequency: $p = P_{AA} + \frac{1}{2}P_{Aa}, \quad q = 1 - p$

Rare recessive disease: If disease prevalence $x = P(aa)$, then $q = \sqrt{x}$. Carrier freq. $\approx 2q$ when $q \ll 1$.

Multiple alleles (K): $P(A_iA_i) = p_i^2, \quad P(A_iA_j) = 2p_ip_j \ (i \neq j)$

Goodness-of-fit test for HWE: $X^2 = \sum \dfrac{(O_i - E_i)^2}{E_i}, \qquad df = (\#\text{genotypes} - 1) - (\#\text{alleles} - 1)$

Why HWE may fail: Wahlund effect (population stratification), inbreeding/assortative mating, selection by genotype, mutation/migration.

3. Gametic & Linkage Disequilibrium (LD) (Ch. 3)

Recombination fraction: $\theta = g_{Ab} + g_{aB}$, with $0 \leq \theta \leq 0.5$. If $\theta < 0.5 \rightarrow$ linked.

LD parameter: $D = g_{AB} - p_Ap_B = g_{AB}g_{ab} - g_{Ab}g_{aB}$

Gamete frequencies in terms of D

$$g_{AB} = p_Ap_B + D, \qquad g_{Ab} = p_Ap_b - D,$$
$$g_{aB} = p_ap_B - D, \qquad g_{ab} = p_ap_b + D.$$

LD decay across generations: $D_n = D_0(1 - \theta)^n$.

Evolutionary Significance: $g^*_{AB} = g_{AB} - \theta D$

Normalized LD (Lewontin): $D' = \begin{cases} D/min(p_Ap_b, p_ap_B), D \geq 0 \\ D/min(p_Ap_B, p_ap_b), D < 0 \end{cases}$

Correlation measure: $r^2 = D^2/(p_Ap_ap_Bp_b)$.

Estimation of Gamete freq. from Genotype freq.: $g_{AB} = \dfrac{2N_{AABB} + N_{AaBB} + N_{AABb}}{2(N_{total} - N_{AaBb})}$ ignore Ab/aB

Test $H_0 : D = 0$: $\chi^2 = \dfrac{ND^2}{p_Ap_ap_Bp_b} \sim \chi^2_1$.

4. Quantitative Genetics (Ch. 4)

4.1 Model and Scaling

- Trait model: $P = G + E$, with $G = A + D + I$ (additive, dominance, epistasis).
- Single-locus scaling (random mating, no sel./mut.):

- Scaled Value: $A_1A_1 \to +a$, $A_2A_2 \to -a$, $A_1A_2 \to d$.
- Genotype freqs: $p^2, 2pq, q^2$ (HWE).
- Population mean: $\mu = a(p-q) + 2pqd$.

Genotypic effects: G_{ij} = scaled value $- \mu$.

4.2 Average Effects & Breeding Values

- Average effect: $\alpha = a + d(q-p)$, A1: $\alpha_1 = q\alpha$, A_2: $\alpha_2 = -p\alpha$.
- Additive (breeding) value: $A_{11} = 2q\alpha, \quad A_{12} = (q-p)\alpha, \quad A_{22} = -2p\alpha$
- Dominance deviations: $D_{11} = -2q^2d, \quad D_{12} = 2pqd, \quad D_{22} = -2p^2d$
- Relationship: $G = A + D$.

4.3 & 4.4 Variance Components & Heritability

Phenotypic Variance: $V_P = V_G + V_E = V_A + V_D + V_E$

Additive genetic variance: $V_A = 2pq\,[a + d(q-p)]^2 = 2pq\alpha^2$

Within-locus interaction (dominance): $V_D = (2pqd)^2$

Broad-sense Heritability: $H^2 = V_G/V_P$ (all genetic).

Narrow-sense Heritability: $h^2 = V_A/V_P$ (additive only).

Interpretation: h^2 is the fraction of phenotypic variance explained by additive genetics; it predicts response to selection.

4.5 Estimating h^2 with Relatives

Let $P_X = G_X + E_X,\ P_Y = G_Y + E_Y$. With independent E's: $\text{Cov}(P_X, P_Y) = \text{Cov}(G_X, G_Y) = (\frac{r_1}{2})V_A + r_2(V_A + V_D)$

- **Parent–offspring**: $\text{Cov}(O, P) = \frac{1}{2}V_A$.
  Regression slope $b_{OP} = \text{Cov}(O,P)/V_P = \frac{1}{2}h^2$ -> $\boxed{h^2 = 2b_{OP}}$.
- **Mid-parent** $\bar{P} = (P_{father} + P_{mother})/2$: slope $b_{O\bar{P}}$ -> $\boxed{h^2 = b_{O\bar{P}}}$.
- **Full sibs**: Cov $= \frac{1}{2}V_A + \frac{1}{4}V_D$.
- **Half sibs**: Cov $= \frac{1}{4}V_A$.

Predicting offspring deviation: If parent deviation is x, expected offspring deviation $= h^2x$ (mid-parent) or $= \frac{1}{2}h^2x$ (single parent).