

9

Inheritance of Quantitative Traits

9.1: Learning Objectives

By the end of this chapter, the reader should be able to distinguish qualitative from quantitative traits, partition phenotypic variance into genetic and environmental components, define and estimate broad- and narrow-sense heritability, and apply the polygenic model to Nigerian quantitative traits such as height, birth weight, and packed cell volume (PCV).

9.2: Qualitative versus Quantitative Traits

The traits considered in Chapters 1–8 (blood group, sickle-cell status, disease presence/absence) are **qualitative** (categorical or discrete): each individual falls into one of a small number of distinct classes, and single-locus Mendelian rules often describe their transmission well. Many traits of central importance in Nigerian public health and agriculture, however, are **quantitative** (continuous or nearly continuous): height, blood pressure, birth weight, packed cell volume (PCV/haematocrit), and crop yield, for example, vary smoothly over a range of values and do not sort into a handful of discrete classes.

Quantitative traits are typically influenced by many genes simultaneously (the **polygenic model**), each of small individual effect, together with environmental factors — a framework first formalised by R. A. Fisher (1918), reconciling Mendelian inheritance with the continuous variation and approximately

normal (Gaussian) distributions observed for such traits in populations, via the **Central Limit Theorem**: the sum of many independent, small-effect genetic and environmental contributions tends toward a normal distribution.

9.3: Partitioning Phenotypic Variance

The basic model of quantitative genetics decomposes an individual's phenotypic value, P , as:

$$P = G + E$$

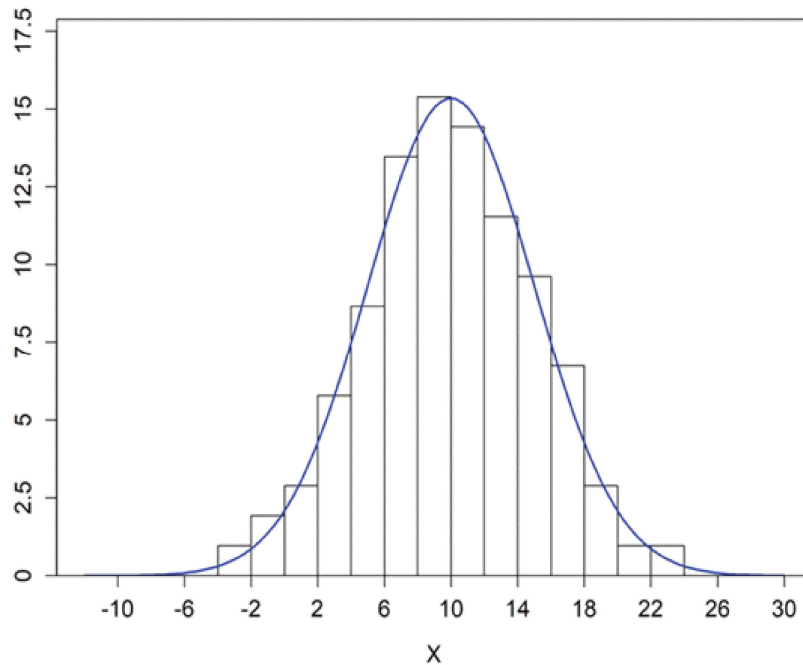
where G is the individual's genotypic value (the net effect of all genes influencing the trait) and E is the environmental deviation (all non-genetic influences, including measurement error). Assuming G and E are statistically independent (a reasonable first approximation, though gene–environment correlation and interaction are recognised complications), the phenotypic variance decomposes as:

$$V_P = V_G + V_E$$

The genotypic variance, V_G , is itself further partitioned:

$$V_G = V_A + V_D + V_I$$

where V_A is the **additive genetic variance** (variance due to the average, independently transmissible effects of individual alleles — the component that determines resemblance between parents and offspring, and the component selection acts on most efficiently), V_D is the **dominance variance** (variance due to interaction between the two alleles at the same locus), and V_I is the **epistatic (interaction) variance** (variance due to interaction between alleles at different loci).



The distribution of a quantitative trait (here, a simulated genotypic-value distribution, x) across a population, closely approximated by a normal curve — the empirical signature of a polygenic trait influenced by many genes of small effect, each contributing a small increment to G (Section 9.2).

9.4: Heritability

Broad-sense heritability, H^2 , is the proportion of total phenotypic variance attributable to all genetic sources:

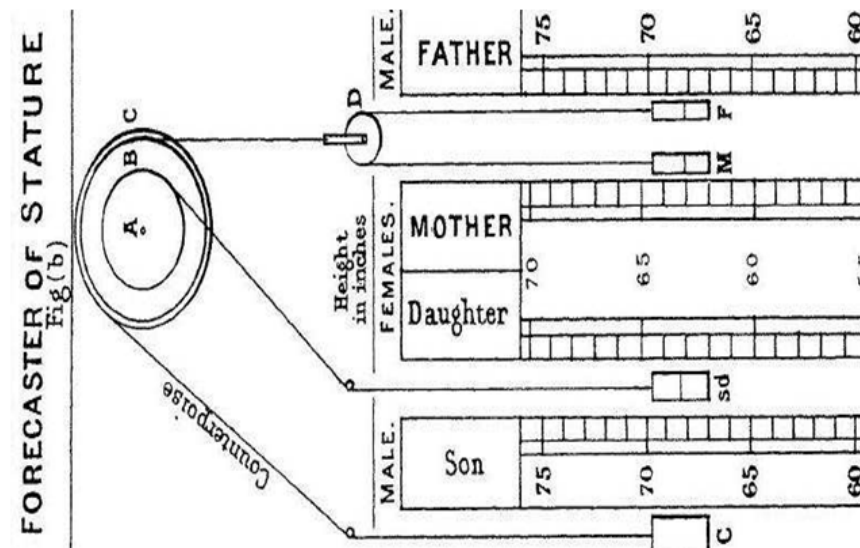
$$H^2 = V_G/V_P$$

Narrow-sense heritability, h^2 , is the proportion of total phenotypic variance attributable to *additive* genetic variance alone:

$$h^2 = V_A/V_P$$

Narrow-sense heritability is the more important quantity for predicting resemblance between relatives and the response to selection, because only the additive component is faithfully transmitted, on average, from parent to offspring (dominance and epistatic effects depend on particular allele *combinations*, which are broken up by segregation and recombination at each generation, Chapters 2 and 7).

9.5: Estimating Heritability from Parent–Offspring Regression



A pedigree recording a quantitative trait (stature) across generations, the classical data structure for parent–offspring heritability estimation.

A classical method for estimating h^2 uses the regression of offspring phenotype on (mid-) parent phenotype. If $\bar{P}_{\text{mid}} = (P_{\text{mother}} + P_{\text{father}})/2$ is the mean of the two parents' phenotypic values, and $P_{\text{offspring}}$ the offspring's value, then under the additive polygenic model, the regression slope of

offspring on mid-parent value directly estimates narrow-sense heritability:

$$b(O, mid - parent) = h^2$$

(If only a single-parent regression is used — offspring on one parent only — the slope estimates $h^2/2$, because the offspring shares, on average, only half the parent's genes.)

Theorem 9.1: (Falconer's Parent–Offspring Regression Theorem)

Under the additive polygenic model, with no dominance, no epistasis, no genotype–environment correlation, and random mating, the regression coefficient of offspring phenotype on mid-parent phenotype equals the narrow-sense heritability, h^2 .

Proof. By definition, the regression slope of Y (offspring phenotype) on X (mid-parent phenotype) is $b = \text{Cov}(X, Y) / \text{Var}(X)$. Write each individual's phenotype as $P = A + E$ (additive genetic value plus environmental deviation, Section 9.3), so the mid-parent value is $\bar{P}_{mid} = \bar{A}_{mid} + \bar{E}_{mid}$ where $\bar{A}_{mid} = (A_{mother} + A_{father})/2$. Assuming environmental deviations are independent across family members and independent of genotype (no shared-environment or gene–environment correlation, part of the model's assumptions), $\text{Cov}(\bar{P}_{mid}, P_{offspring}) = \text{Cov}(\bar{A}_{mid}, A_{offspring})$, since all cross-terms involving E vanish. Under random mating and Mendelian transmission, the offspring's additive genetic value is $A_{offspring} = 1/2 A_{mother} + 1/2 A_{father} + M$, where M is the within-family Mendelian sampling deviation (independent of both

parents' genotypes, since it represents the randomness of which specific parental alleles are transmitted). Hence:

$$\begin{aligned} & \text{Cov}(\bar{A}_{mid}, A_{offspring}) \\ &= \text{Cov}(1/2 A_m + 1/2 A_f, 1/2 A_m + 1/2 A_f \\ &+ M) = \text{Var}(1/2 A_m + 1/2 A_f) \\ &= 1/4 V_A + 1/4 V_A = 1/2 V_A, \end{aligned}$$

using independence of A_m , A_f (random mating) and of M from both. Meanwhile $\text{Var}(\bar{P}_{mid}) = \text{Var}(1/2 P_m + 1/2 P_f) = 1/4 V_P + 1/4 V_P = 1/2 V_P$ (again by independence of the two unrelated parents). Therefore:

$$b = \frac{\text{Cov}(\bar{P}_{mid}, P_{offspring})}{\text{Var}(\bar{P}_{mid})} = \frac{1/2 V_A}{1/2 V_P} = \frac{V_A}{V_P} = h^2. \quad \square$$

Worked Example 9.3: (Parent–offspring height regression in a Nigerian university cohort)

A biometrics class at a Nigerian university measures adult height for 150 mother–father–offspring trios drawn from staff and students. The mean mid-parent height is 168.2 cm and mean offspring height is 169.5 cm; the sample covariance between mid-parent and offspring height is 18.4 cm², and the sample variance of mid-parent height is 26.3 cm². By Theorem 9.1, the estimated narrow-sense heritability is

$$\hat{h}^2 = \frac{18.4}{26.3} \approx 0.70,$$

consistent with the substantial (though not exclusive) genetic contribution to adult height documented across many human populations; the regression's intercept also allows a prediction of expected offspring height for a given mid-parent value, and — because $\hat{h}^2 < 1$ — Theorem 9.1's underlying model additionally predicts the classical phenomenon of

regression to the mean: offspring of unusually tall or unusually short mid-parent pairs are expected, on average, to lie closer to the population mean than their parents did, precisely because environmental deviations (the $1 - h^2$ share of variance) are not transmitted from parent to offspring.

Worked Example 9.1 (Height, Nigerian cohort)

A study of 200 father–mother–adult-child trios in a South-Western Nigerian community records mid-parent height and adult offspring height (in cm). A regression of offspring height on mid-parent height yields a slope $b = 0.78$, with standard error 0.06.

(a) Interpret b as an estimate of heritability: Since $b = h^2$ under the mid-parent regression model, $\hat{h}^2 = 0.78$, i.e., approximately 78% of the variance in adult height in this population is estimated to be attributable to additive genetic effects — consistent with the generally high heritability of stature (commonly estimated at 0.7–0.9) reported in human populations worldwide.

(b) Construct an approximate 95% confidence interval for h^2 . $\hat{h}^2 \pm 1.96 \times \text{SE} = 0.78 \pm 1.96(0.06) = 0.78 \pm 0.118$, giving an interval of approximately (0.66, 0.90).

9.6: Estimating Heritability from Sibling Data

An alternative classical approach uses the **intraclass correlation** among full siblings, t , obtained from a one-way analysis of variance (ANOVA) of trait values with “family” as the grouping factor:

$$t = \sigma_B^2 / (\sigma_B^2 + \sigma_W^2)$$

where σ_B^2 is the between-family variance component and σ_W^2 the within-family variance component. Since full siblings share, on average, one-half of their segregating genes

(coefficient of relationship $r = 1/2$, Chapter 6), and assuming no dominance variance, shared environment, or assortative mating, the intraclass correlation among full sibs estimates one-half of narrow-sense heritability:

$$t \approx \frac{1}{2}h^2 \Rightarrow \hat{h}^2 = 2t$$

Worked Example 9.2: (Packed Cell Volume, PCV)

A one-way ANOVA of PCV (%) measured in 60 Nigerian sibling pairs (families) gives $\sigma^2_B = 9.6$ and $\sigma^2_W = 14.4$.

$$t = 9.6/(9.6 + 14.4) = 9.6/24.0 = 0.40 \Rightarrow \hat{h}^2 = 2(0.40) = 0.80$$

The estimated narrow-sense heritability of PCV in this sibling sample is 0.80, indicating substantial genetic control of this widely measured Nigerian clinical trait, alongside the well-known environmental influences of iron status, malaria infection, and nutrition.

9.7: The Polygenic (Multifactorial) Threshold Model

Some traits that are themselves discrete (affected/unaffected) — for example, cleft lip/palate, some forms of congenital heart disease, and neural tube defects, all of documented occurrence in Nigerian clinical genetics — are believed to result from an underlying continuous, normally distributed **liability** determined by many genes and environmental factors jointly, with the disease phenotype appearing only when an individual's liability exceeds a threshold, T . Under this **multifactorial threshold model**, the population prevalence of the disease corresponds to the area under the standard normal curve beyond T , and the elevated recurrence risk in relatives of an affected individual (compared to population prevalence) can be used to estimate the heritability of liability — a method of particular relevance to Nigerian family-counselling

settings for congenital anomalies, where empirical recurrence risks (rather than simple Mendelian ratios) are typically what is quoted to parents.

Theorem 9.2: (Parent-Offspring Covariance Under an Additive Genetic Model)

Under a purely additive genetic model with no dominance or epistasis, and random mating, the covariance in phenotype between a parent and their offspring equals one-half the additive genetic variance: $\text{Cov}(P_{\text{parent}}, P_{\text{offspring}}) = (1/2) V_A$.

Proof. Write each individual's phenotype as $P = A + E$, where A is the additive genetic value (a sum of average allelic effects across loci) and E is an independent environmental deviation with $E[E] = 0$. An offspring's additive genetic value is the sum of one randomly transmitted allelic effect from each parent at each locus; by the additivity assumption, $A_{\text{offspring}} = (1/2)A_{\text{parent1}} + (1/2)A_{\text{parent2}} + M$, where M is a Mendelian sampling deviation with mean zero that is independent of both parents' genetic values (it reflects which of the two parental alleles was transmitted at each locus, an independent random event by Theorem 1.1). Then $\text{Cov}(A_{\text{parent1}}, A_{\text{offspring}}) = \text{Cov}(A_{\text{parent1}}, (1/2)A_{\text{parent1}} + (1/2)A_{\text{parent2}} + M) = (1/2)\text{Var}(A_{\text{parent1}}) + (1/2)\text{Cov}(A_{\text{parent1}}, A_{\text{parent2}}) + \text{Cov}(A_{\text{parent1}}, M)$. Under random mating, $\text{Cov}(A_{\text{parent1}}, A_{\text{parent2}}) = 0$, and $\text{Cov}(A_{\text{parent1}}, M) = 0$ by the independence of Mendelian sampling from the parental genotype. Hence $\text{Cov}(A_{\text{parent1}}, A_{\text{offspring}}) = (1/2)\text{Var}(A_{\text{parent1}}) = (1/2)V_A$. Since the environmental deviations are independent of genotype and of each other across generations, $\text{Cov}(P_{\text{parent}}, P_{\text{offspring}}) = \text{Cov}(A_{\text{parent}}, A_{\text{offspring}}) = (1/2)V_A$, which is exactly the

theoretical basis for the parent-offspring regression method of Section 9.5.

9.8: Chapter Summary

- Quantitative traits are influenced by many genes of small effect (polygenic model) plus environment; phenotypic variance decomposes as $V_P = V_A + V_D + V_I + V_E$.
- Narrow-sense heritability, $h^2 = V_A/V_P$, governs resemblance between relatives and response to selection.
- h^2 can be estimated from the offspring-on-mid-parent regression slope, or from twice the sibling intraclass correlation.

9.9: Computational Corner: Estimating Heritability by Regression

R.

```
# Parent-offspring regression heritability estimate (Section
# 10.5), PCV example
set.seed(101)
n <- 200
mid_parent_pcv <- rnorm(n, mean = 38, sd = 3)      # mean
# parental PCV (%)
h2_true <- 0.45
offspring_pcv <- 38 + h2_true * (mid_parent_pcv - 38)
# + rnorm(n, 0, 2.5)

fit <- lm(offspring_pcv ~ mid_parent_pcv)
slope <- coef(fit)[2]
h2_hat <- slope      # regression on MID-parent value
# estimates h2 directly
summary(fit)
```

```
cat("Estimated heritability (mid-parent regression):", h2_hat,
"\n")
```

Python.

```
import numpy as np
import statsmodels.api as sm

np.random.seed(101)
n=200
mid_parent_pcv = np.random.normal(38, 3, n)
h2_true=0.45
offspring_pcv = 38 + h2_true * (mid_parent_pcv - 38) +
np.random.normal(0, 2.5, n)

X = sm.add_constant(mid_parent_pcv)
model = sm.OLS(offspring_pcv, X).fit()
h2_hat = model.params[1]
print(model.summary())
print("Estimated heritability:", h2_hat)
```

The recovered slope should lie close to the simulated true heritability of 0.45, illustrating the parent–offspring regression estimator of Section 10.5 on a simulated packed-cell-volume (PCV) trait of the kind measured in Nigerian cohort studies; repeating the simulation with a different random seed shows the sampling variability of h^2 even when the underlying model is exactly correct.

Mini-Project

Mini-Project 9: Estimating Heritability Two Ways and Comparing Results

Simulate a quantitative trait with a known heritability and recover that heritability using both parent-offspring regression and sibling correlation.

Step 1: Simulate a quantitative trait for 200 parent-offspring pairs and 200 sibling pairs, generated from a model with a known true heritability (e.g., $h^2 = 0.5$).

Step 2: Estimate heritability via parent-offspring regression (Section 9.5).

Step 3: Estimate heritability via sibling correlation (Section 9.6).

Step 4: Compare both estimates to the true simulated value and to each other.

Step 5: Write a short paragraph discussing which estimate was closer to the truth in your simulation, and one reason the two methods can disagree in real data.

Deliverable: *Simulation code, both heritability estimates with their standard errors, and a short written comparison.*

9.10 Practice Questions

Q1. A regression of offspring birth weight on mid-parent birth weight (both measured in kg, in a Lagos maternity cohort) yields a slope of 0.35. Estimate h^2 for birth weight in this population, and comment on why it might be lower than the h^2 typically estimated for adult height.

- Q2.** A one-way ANOVA of adult height among 80 Nigerian sibships gives $\sigma^2_B = 45$ and $\sigma^2_W = 15$. Estimate the intraclass correlation t and the narrow-sense heritability h^2 .
- Q3.** Explain why dominance variance (V_D) and epistatic variance (V_I) contribute to broad-sense heritability (H^2) but are largely “invisible” to standard parent–offspring regression estimates of narrow-sense heritability (h^2).
- Q4.** The population prevalence of a congenital heart defect is 0.5%, but among full siblings of an affected child, the observed prevalence rises to 3%. Explain qualitatively (no calculation required) how the multifactorial threshold model accounts for this elevated recurrence risk without invoking a single Mendelian gene.
- Q5.** Explain, in your own words and with reference to the Central Limit Theorem (Section 9.2), why a trait influenced by hundreds of genes of small individual effect tends to show an approximately normal (bell-shaped) distribution in the population, even though each individual gene’s effect on its own need not be normally distributed.
- Q6.** A researcher partitions the phenotypic variance of birth weight in a Nigerian cohort and estimates $V_A = 0.35$, $V_D = 0.05$, $V_I = 0.02$, $V_E = 0.58$ (all in the same units of variance). Compute the broad-sense heritability ($H^2 = V_G/V_P$) and the narrow-sense heritability ($h^2 = V_A/V_P$), and explain why $h^2 \leq H^2$ always.

