Resemblance Between Relatives (Part 1)

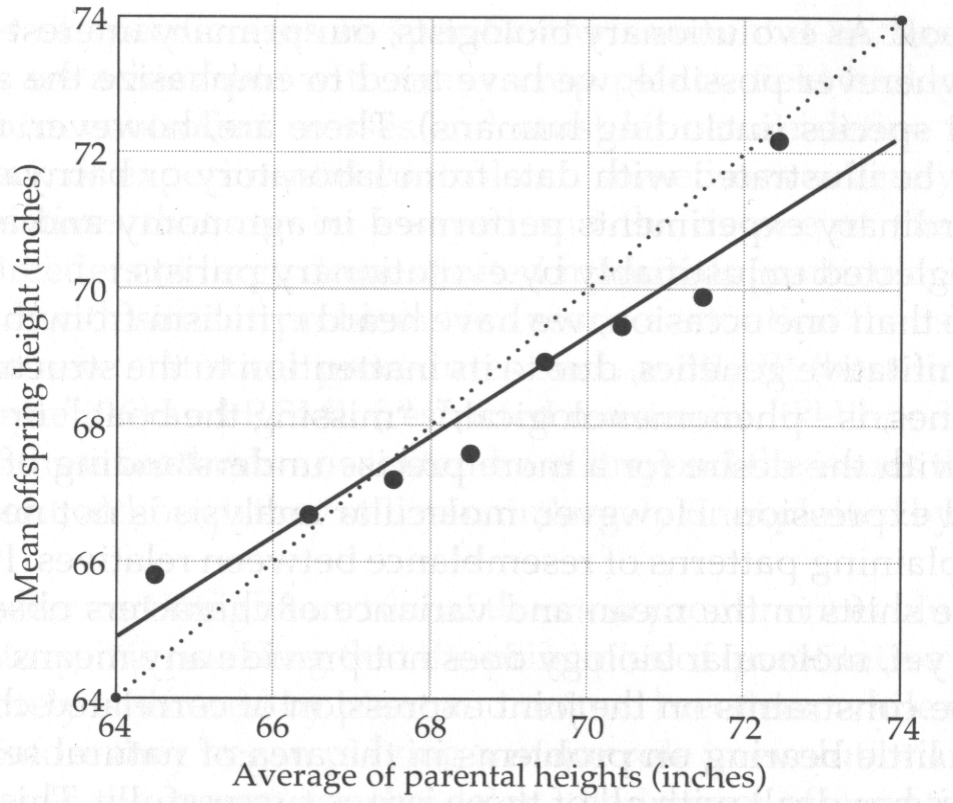
Resemblance of Relatives Intro

- Individuals in a population can vary considerably on physical, cognitive, and behavioral traits. Examples include height, weight, intelligence, personality, etc.
- Individual differences may arise from variation in genes, environmental experiences, or a combination of both
- It is generally accepted that individual differences in most traits are due to both genetic and environmental factors.
- To the extent that a trait is influenced by genetic (heritable) effects, phenotypic similarity is correlated with genetic relatedness.

Resemblance of Relatives Intro

- The degree of resemblance among relatives allows for the estimation of genetic variance as well as heritability of a trait.
- We will now focus on how phenotypic correlation of relatives can be used to estimate the relative magnitude of genetic and environmental influences on trait variation

Galton 1889 Paper on Resemblance of Relatives



[Galton, 1889]

5

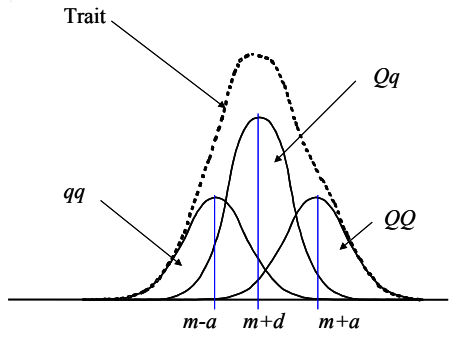
Fisher 1918 Paper on Resemblance of Relatives

XV.—**The Correlation between Relatives on the Supposition of Mendelian Inheritance.** By **R. A. Fisher**, B.A. *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.

Several attempts have already been made to interpret the well-established results of biometry in accordance with the Mendelian scheme of inheritance. It is here attempted to ascertain the biometrical properties of a population of a more general type than has hitherto been examined, inheritance in which follows this scheme. It is hoped that in this way it will be possible to make a more exact analysis of the causes of human variability. The great body of available statistics show us that the deviations of a human measurement from its mean follow very closely the Normal Law of Errors, and, therefore, that the variability may be uniformly measured by the standard deviation corresponding to the square root of the mean square error. When there are two independent causes of variability capable of producing in an otherwise uniform population distributions with standard deviations σ_1 and σ_2, it is found that the distribution, when both causes act together, has a standard deviation $\sqrt{\sigma_1^2+\sigma_2^2}$. It is therefore desirable in analysing the causes of variability to deal with the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the normal population to which it refers, and we may now ascribe to the constituent causes fractions or percentages of the total variance which they together produce. It



RA Fisher (1918). *Transactions of the Royal Society of Edinburgh* **52**: 399-433.

7

Partitioning of Phenotypic Values

- We previously introduced the general model of $Y = G + E$ where Y is the trait value, G is the genotypic value, and E is the environmental deviation that is assumed to have a mean of 0 so that $E(Y) = E(G)$.
- If we assume that the genetic and environmental components are independent, we have that the variance of Y is $Var(Y) = \sigma^2_Y = \sigma^2_G + \sigma^2_E$.
- We previously focused on the decomposition of the genetic variance σ^2_G into additive and dominance components, i.e., $G = A + D$ where

$$Var(G) = \sigma^2_G = \sigma^2_A + \sigma^2_D$$

- For multi-locus traits, we showed σ^2_G can also include interaction variance components for the additive and dominance effects. Assume no interaction effects for now.

Partitioning of Phenotypic Values

- So, for any random individual U from the population we have that $Var(Y_U) = \sigma^2_Y = \sigma^2_G + \sigma^2_E = \sigma^2_A + \sigma^2_D + \sigma^2_E$.
- We also previously showed that the covariance of the genotypic values for non-inbred individuals U and V that

$$Cov(G_U, G_V) = 2\theta\sigma^2_A + k_2\sigma^2_D$$

  where θ and k_2 and the kinship coefficient and probability of sharing two alleles IBD for individuals U and V
- If we assume that there is no environmental correlation among relatives U and V for the trait (potentially a very strong assumption for relatives that lived or were reared in the same household!), we have that

$$Cov(Y_U, Y_V) = Cov(G_U, G_V) = 2\theta\sigma^2_A + k_2\sigma^2_D$$

Genetic Covariance

- Genetic Covariance for a few outbred relative pairs

Relationship	Genetic Covariance
Identical Twins	$\sigma_A^2 + \sigma_D^2$
Full Siblings	$\frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_D^2$
Parent-Offspring	$\frac{1}{2}\sigma_A^2$
Half Siblings*	$\frac{1}{4}\sigma_A^2$
Uncle-Nephew	$\frac{1}{4}\sigma_A^2$
First Cousins	$\frac{1}{8}\sigma_A^2$
Double First Cousins	$\frac{1}{4}\sigma_A^2 + \frac{1}{16}\sigma_D^2$
Second Cousins	$\frac{1}{32}\sigma_A^2$
Unrelated	0

* Also grandparent-granchild

Resemblance of Parent and Offspring

- Consider the trait values for a parent-offspring pairs. Let Y_O be the phenotype value for the offspring and let Y_P be the phenotype value for a parent.
- For a sample of parent-offspring pairs, one can use linear regression to predict offspring phenotype from the parent's phenotype .
- The linear regression model of Y_O on Y_P is $Y_O = \beta_0 + \beta_1 Y_P$ where

$$\beta_1 = \frac{Cov(Y_P, Y_O)}{Var(Y_P)}$$

- From the previous slide we have that the genetic covariance for parent-offspring pairs is $\frac{1}{2}\sigma_A^2$. What is β_1 in terms of the additive variance and the variance of the trait for parent-offspring pairs? Assume that there is no correlation among the genetic and environmental effects.

Resemblance of Parent and Offspring

- We have that

$$\beta_1 = \frac{Cov(Y_P, Y_O)}{Var(Y_P)} = \frac{\frac{1}{2}\sigma_A^2}{\sigma_Y^2}$$

- From the previous slide we have that the genetic covariance for parent-offspring pairs is $\frac{1}{2}\sigma_A^2$.
- Remember that the narrow-sense heritability is

$$h^2 = \frac{\sigma_A^2}{\sigma_Y^2}$$

- So, from the slope of the regression or parent on offspring trait value an estimate of the h^2 can be obtained from

$$\hat{h}^2 = 2\hat{\beta}_1$$

Resemblance of Mid-Parent and Offspring

- Now consider the trait values for the mid-parent trait value of the parents and the offspring pairs. Let Y_O be the phenotype value for the offspring and let $\bar{Y}_P = \frac{Y_{P_1} + Y_{P_2}}{2}$ be the mid-parent (or mean of both parents) phenotype value.
- Now obtain an estimate of the narrow-sense heritability of the trait from the slope. Assume that the different sexes have the same variance
- Which heritability estimator would be more efficient, the one based on the linear regression of offspring on a single parent or the estimator based on the regression of offspring on mid-parent?

Resemblance of Mid-Parent and Offspring

- We have that the slope of the regression line of offspring on mid-parent is

$$\beta_1 = \frac{Cov(\bar{Y}_P, Y_O)}{Var(\bar{Y}_P)} = \frac{\frac{1}{2}\sigma_A^2}{\frac{1}{2}\sigma_P^2} = \frac{\sigma_A^2}{\sigma_P^2} = h^2$$

- So the slope of the regression of offspring on mid-parent provides an estimate of h^2
- The heritability estimator using the offspring and mid-parents has a smaller variance than the estimator based on the offspring and a single parent.
- For some traits, there could be a sex-effect, so using the mid-parent value may not be advisable for heritability estimation unless some adjustment is made to the phenotype values to account for relevant influential confounders such as age, sex, etc.

Resemblance of Half-Siblings

- Now consider the trait values for half-sibling pairs and obtain pairs and obtain
- From the previous slide we have that the genetic covariance for half-sibling pairs is $\frac{1}{4}\sigma_A^2$
- Obtain an estimate of the narrow-sense heritability of the trait from the slope of the regression.

General Formula

- Note that the correlation of the phenotypes for a pair of non-inbred individuals U and V who have a dominance variance component equal to 0 (and assuming either no or negligible gene-environmental interaction effects for the trait) is

$$\rho_{Y_U Y_V} = 2\theta_{UV} h^2$$

- This relationship between the correlation of the phenotypes of relatives and heritability is why the slope of a linear regression model can be used for inference on heritability. For those relationship types that only have an additive variance component (and a dominance variance component that is 0), we have the following:

$$h^2 = \frac{\rho_{Y_U Y_V}}{2\theta_{UV}} = \frac{\beta_1}{2\theta_{UV}}$$

where it is assumed that $Var(Y_U) = Var(Y_V) = \sigma^2_Y$. It is common to standardized the phenotype so that $\sigma^2_Y = 1$ for heritability estimation.

Resemblance of Full-Siblings

- What is the slope of the regression line for trait values of full-siblings?
- Under what conditions/assumptions can the slope for this relative pair type provide an reasonable estimate for the heritability of a trait?

Resemblance of Full-Siblings

- We have that the slope of the regression line for trait values of full-sibling pairs is

$$\beta_1 = \frac{\frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_D^2}{\sigma_Y^2}$$

- So the narrow-sense heritability can not be estimated from the slope of the regression line (or correlation of phenotype values) with full siblings. This relative pair type will only provide an unbiased of heritability using the slope of the regression line if the trait is completely additive, i.e., $\sigma_D^2 = 0$