Kinship Coefficients and Covariance Between Relatives

Biostatistics 666

Kinship Coefficients

- Measure:
  - Relatedness between two individuals
  - Inbreeding in a single individual
- Useful predictors of covariance and correlation between relatives

Definition

- Given two individuals
  - One with genes (g_i, g_j)
  - The other with genes (g_k, g_l)

- The kinship coefficient is:
  - $\frac{1}{4}P(g_i \equiv g_k) + \frac{1}{4}P(g_i \equiv g_l) + \frac{1}{4}P(g_j \equiv g_k) + \frac{1}{4}P(g_j \equiv g_l)$

- Probability that two genes sampled at random from each individual are identical

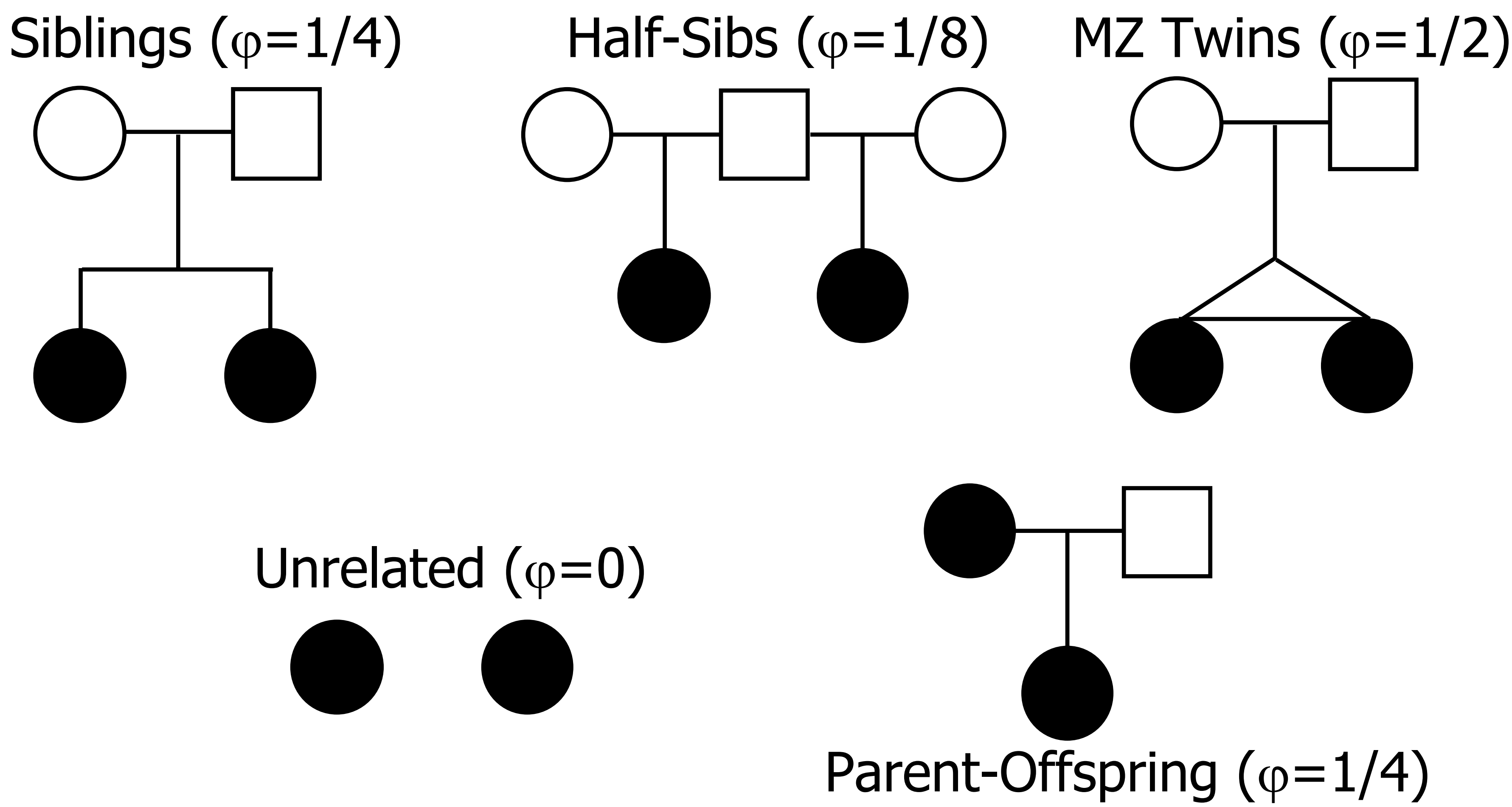
Inbreeding Coefficient

- For any given individual…
- The inbreeding coefficient is the kinship coefficient between the individual's parents
- Should be zero if individual is not inbred

Notation

- φ_{ij}
  - Kinship between individuals i and j as
- f_i
  - Inbreeding for individual i
- Note
  - $\varphi_{ii} = \frac{1}{2}(1 + f_i)$
  - $f_i = \varphi_{mother(i)father(i)}$

Some kinship coefficients

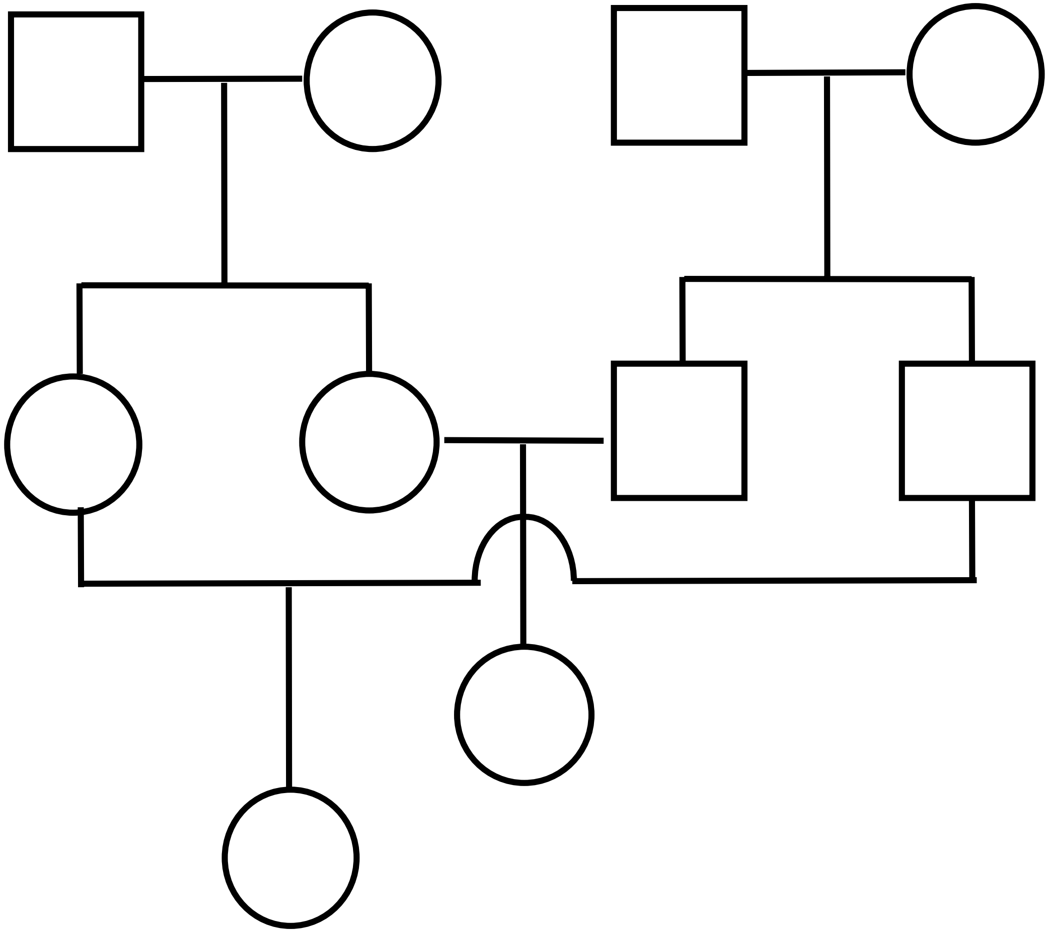


Computing Kinship Coefficients

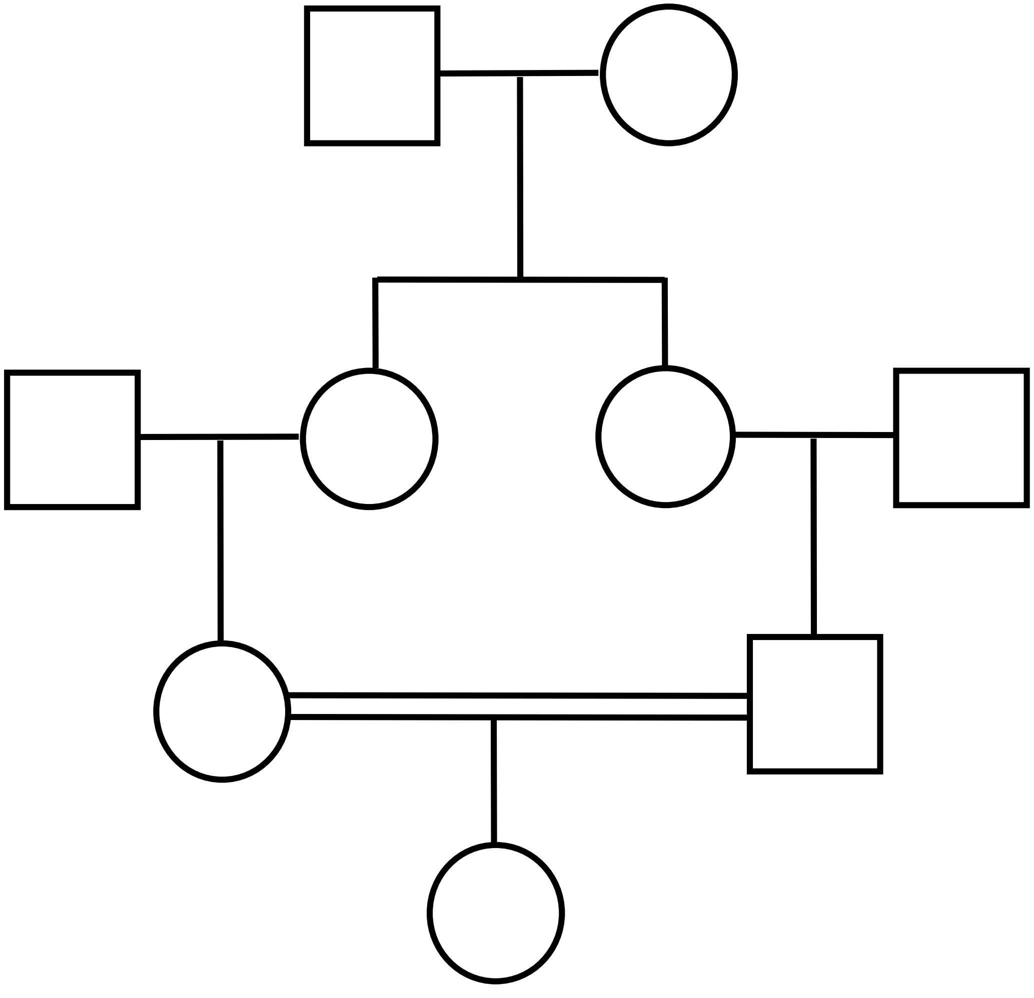
- φ_{ij} – kinship coefficient between *i* and *j*
  - Order individuals so ancestors always precede their descendants
    - If $i>j$, then *i* is not ancestor of *j*

$$\varphi_{ij} = \begin{cases} 0 & i \text{ and } j \text{ are founders} \\ \tfrac{1}{2} & i = j, i \text{ is a founder} \\ \tfrac{1}{2}(\varphi_{mother(i)\,j} + \varphi_{father(i)\,j}) & i \neq j \\ \tfrac{1}{2}(1 + \varphi_{mother(i)\,father(i)}) & i = j \end{cases}$$

More Kinship Coefficients…



Inbred Example



Relationship to IBD probabilities

- In non-inbred pedigrees, kinship coefficients can be derived from IBD probabilities:

$$\varphi_{ij} = \tfrac{1}{4} P(IBD_{ij} = 1) + \tfrac{1}{2} P(IBD_{ij} = 2)$$

Model Setup

- Consider a gene with two alleles
  - Define g_i as an indicator variable
    - 1 if allele A is present
    - 0 if another allele is present

- For a single gene, calculate
  - $E(g_i)$
  - $V(g_i)$

Now consider a pair of genes

- If g_i and g_j are sampled independently…
  - $E(g_i + g_j)$
  - $V(g_i + g_j)$
  - $Cov(g_i, g_j)$
- What if g_i and g_j are sampled from related individuals?
  - Results depend on kinship…

A Model for Quantitative Traits

- Observed trait is sum of multiple effects:
  - Population mean
    - $\mu = 0$
  - Environmental effects
    - $N(0, \sigma_e^2)$
  - Genetic effects
    - $N(0, \sigma_g^2)$
    - Each allele has effect $\sim N(0, \sigma_g^2 / 2)$

Covariance between Relatives

- Consider:
  - Genes for individual i are (g_{i1}, g_{i2})
  - Genes for individual j are (g_{j1}, g_{j2})

- Contributions to phenotype:
  - For individual i are $(g_{i1} + g_{i2})$
  - For individual j are $(g_{j1} + g_{j2})$

Then…

$$E\left((g_{i1}+g_{i2})(g_{j1}+g_{j2})\right)=$$

$$=E\left(g_{i1}g_{j1}+g_{i1}g_{j2}+g_{i1}g_{j2}+g_{i2}g_{j2}\right)$$

$$=P(g_{i1}\equiv g_{j1})V(g)+P(g_{i1}\equiv g_{j2})V(g)+$$

$$P(g_{i2}\equiv g_{j1})V(g)+P(g_{i2}\equiv g_{j2})V(g)$$

$$=4\varphi_{ij}V(g)$$

$$=\ 2\varphi_{ij}\sigma_g^2$$

Implications…

- In a simple model:
  - Genes are additive the covariance
  - No shared environment
- Covariance between relatives is a simple function of kinship
- Total variance due to genes could be estimated by relating the two