A restriction on the space of genetic relationships

By E. A. THOMPSON

King's College Research Centre, King's College, Cambridge CB2 1ST

The autosomal genetic consequences of a genealogical relationship between two non-inbred individuals are determined by the Cotterman k-coefficients, \((k_0, k_1, k_2)\), corresponding to the relationship (see, for example, Crow & Kimura (1970)). Thompson (1975) considered a modification of these coefficients,

\[
K = (K_0, K_1, K_2) = (k_0, 2k_1, k_2); \quad K_0 + K_1 + K_2 = 1, \tag{1}
\]

where \(K_i\) is the probability that the two individuals have precisely \(i\) genes in common at a given autosomal locus of interest. (In this context, genes are considered to be the same if and only if they are identical by descent from some common ancestor.) The use of the coefficients \(K_0\), \(K_1\) and \(K_2\), corresponding to the probabilities \(\Delta_0\), \(\Delta_1\) and \(\Delta_2\) in the more general formulation of Jacquard (1974), enables relationships to be represented as points in an equilateral triangle of unit height.

Van Aarde (1975) has recently considered non-inbred relatives, and derived, for a different purpose, expressions for \(K_2\) and \(\Phi\), the coefficient of kinship between the two individuals, in terms of the coefficients of kinship of the four parents, regardless of whether or not these are inbred. We use his expressions to derive a restriction on the space (1) of possible vectors \(K\), and to obtain an alternative expression of a relationship in terms of two other parameters \(p\) and \(q\).

Suppose first that two collateral non-inbred relatives are such that their fathers are related and their mothers are related, but their fathers are not related to their mothers. Then the event that they receive identical genes from their fathers is independent of the event that they receive identical genes from their mothers, where again identical means identical by descent. The probability of the former event is \(\phi_{ff}\), the coefficient of kinship between fathers, while that of the latter is \(\phi_{mm}\), the coefficient of kinship between mothers.

We thus find that the vector \(K\) takes the form

\[
(K_0, K_1, K_2) = ((1 - \phi_{mm}) (1 - \phi_{ff}), (1 - \phi_{ff}) \phi_{mm} + (1 - \phi_{mm}) \phi_{ff}, \phi_{mm} \phi_{ff}). \tag{2}
\]

This is the case regardless of whether or not the fathers and mothers are inbred. (We could clearly equally consider the case where the two fathers are unrelated and the two mothers are unrelated, but the father of each individual is related to the mother of the other. For terminological convenience we shall specify the above case, but it is to be understood that the latter is also implicitly included.)

Van Aarde (1975) has recently pointed out (as also did Trustrum (1961)) that there are genealogical relationships for which each of the father and mother of each individual is related to both the father and the mother of the other, without the individuals being themselves inbred. The simplest case is that of quadruple-half-first-cousins—a far from pathological relationship. In such cases the four parental kinship coefficients \(\phi_{ff}, \phi_{mm}, \phi_{fm}\) and \(\phi_{mf}\) may all be simultaneously non-zero, where \(\phi_{fm}\) and \(\phi_{mf}\) denote kinships between the father of one individual and the mother of the other. The identity state probabilities no longer take the form (2), and, as discussed by
van Aarde, standard formulae for the covariance between non-inbred relatives, which have sometimes assumed the form (2), may not be applicable.

However, we shall show that for all relationships between non-inbred relatives we have

$$K = (K_0, K_1, K_2) = ((1 - p)(1 - q), p(1 - q) + q(1 - p), pq),$$

for some unique pair $p$ and $q$ between 0 and 1, although $p$ and $q$ may not in general have any interpretation as parental kinships. For example, for quadruple-half-first-cousins a simple direct computation gives

$$(K_0, K_1, K_2) = (17/32, 14/32, 1/32)$$

which is of the form (3) if

$$p = (2 + \sqrt{2})/8 \text{ and } q = (2 - \sqrt{2})/8.$$  

This particular $p$ and $q$ clearly cannot have any interpretation as kinship coefficients as they are not dyadic rationals, or even rationals.

Suppose first that $K_2 = 0$. Then

$$K = ((1 - K_1), K_1, 0) \text{ since } K_0 + K_1 + K_2 = 1.$$  

If we take $q = 0$ and $p = K_1$ then $K$ is of the form (3). So suppose $K_2 > 0$; then if there exists a representation of the form (3), $p$ and $q$ are strictly positive. Then

$$K_2 = pq \text{ or } q = K_2/p,$$

and

$$K_0 = (1 - p)(1 - q) = (1 - p)(1 - K_2/p)$$

or

$$g(p) = p^2 - p(1 + K_2 - K_0) + K_2 = 0.$$  

Now

$$g(0) > 0 \text{ and } g(1) > 0$$

and

$$g'(p) = 0 \text{ at } p = \frac{1}{2}(1 + K_2 - K_0) = p^* \text{ say, where } 0 \leq p^* \leq 1.$$  

Thus the necessary and sufficient condition for a representation (3), with $p$ and $q$ in the required range, is that (4) has real roots, or

$$(1 + K_2 - K_0)^2 \geq 4K_2$$

which reduces to

$$4K_2K_0 \leq K_1^2.$$  

Relationships satisfying (5) are those lying outside a parabolic sector on the side $K_1 = 0$ of the triangular space of vectors $K$ (figure 1).

We now show the more surprising result that in fact all relationships must satisfy (5), or that in terms of the Cotterman $k$-coefficients

$$k_1^2 \geq k_0k_2.$$  

Van Aarde (1975) shows that the general form for the relationship between any two non-inbred individuals, whose parents may be inbred, is

$$K_2 = \phi_{mn}\hat{\Phi} + \phi_{mf}\hat{\Phi} + \phi_{mf}\hat{\Phi} + \phi_{mm}\hat{\Phi},$$

and

$$\Phi = (\phi_{mn} + \phi_{mf} + \phi_{mf} + \phi_{mm})/4,$$

where $\Phi$ is the coefficient of kinship between the two individuals, and $\phi_{ff}$, etc., are the parental coefficients defined above. Since the individuals are not inbred

$$\Phi = (2K_2 + K_1)/4 \quad (\text{Thompson (1975)}).$$
A restriction on the space of genetic relationships

Fig. 1. Restriction on the space of gene identity coefficients. No genealogical relationship gives coefficients within the parabola on side UR of the triangular space. Vertices U, Q and R denote respectively the positions for relationships unrelated, parent-offspring and identical twins. B and P denote the positions for sibs and quadruple-half-first-cousins while P, P and P denote double rth cousins for r = 1, 3/2 and 2. Sibs whose two parents are equally inbred give points on the parabolic arc BR.

(Note that, since the parents may be inbred, \( \phi_{ff} \), etc., may not be written in this form; Jacquard (1974) gives the more general form of the kinship between two inbred individuals in terms of his extended set of identity state probabilities.)

Now the arithmetic-geometric mean inequality gives

\[ \phi_{mm} \phi_{ff} \leq \left( \frac{1}{2}(\phi_{mm} + \phi_{ff}) \right)^2. \]

Thus

\[ K_2 \leq \left( (\phi_{mm} + \phi_{ff})^2 + (\phi_{mf} + \phi_{fm})^2 \right)/4. \]  
(7)

If we write \( \phi_{mm} + \phi_{ff} = x \), then \( \phi_{fm} + \phi_{mf} = 4\Phi - x \), and

\[ K_2 \leq (x^2 + (4\Phi - x)^2)/4 \]

\[ = 4\Phi^2 - \frac{1}{2}x(4\Phi - x) \]

\[ \leq 4\Phi^2, \text{ since } 0 \leq x \leq 4\Phi. \]  
(8)

That is,

\[ K_2 \leq (K_1 + 2K_3)^2/4 \]

from (6), which reduces to \( 4K_2K_0 \leq K_1^2 \), which is the same constraint (5) derived above. Equality is achieved in (7) if \( \phi_{mm} = \phi_{ff} \) and \( \phi_{fm} = \phi_{mf} \), and in (8) if \( \phi_{mm} = \phi_{ff} = 0 \) or \( \phi_{fm} = \phi_{mf} = 0 \).
Thus any relationship between two non-inbred individuals lies in a subspace of the triangular space (1). The space within a parabolic curve on base $K_1 = 0$ is not attainable (for non-inbred individuals) by any genealogical structure whatsoever. Points on the parabola correspond to relationships where mothers and fathers have the same relationship, and fathers are unrelated to mothers. These include identical twins, sibs, double-first-cousins, double-$r$th-cousins for any $r$, and unrelated pairs. Fig. 1 demonstrates this restriction; the position of quadruple-half-first cousins is also shown. Thompson (1976) shows the positions of other standard relationships in the triangular space.

We have also shown that all relationships between non-inbred relatives may be expressed in the form (3) for some uniquely determined unordered pair $(p, q)$ lying between 0 and 1. Note that

$$\Phi = (K_1 + 2K_2)/4 = (p + q)/4.$$ 

Thus van Aarde’s criticism that in general Malécot’s form for the covariance between non-inbred relatives is not applicable, is true only in that the component parts are not parent kinships, not in that such a form is incorrect. There are problems in which a reparametrization in terms of the pair $(p, q)$ may be useful, in particular, those in which $K_2$ and $\Phi$ are principally of interest, these being, respectively, the product and a simple multiple of the sum of the two elements in the pair.

Finally we note that the parabola (5) is tangential to the side $K_2 = 0$ at the vertex $K_2 = K_1 = 0$ and tangential to $K_0 = 0$ at the vertex $K_1 = K_0 = 0$. There are thus no bilineal relationships approaching arbitrarily closely the state of unrelatedness, and no relationships with positive $K_0$ approaching the state of identical twins. The maximal level of $K_2$ is of order $\Phi^2$, and hence for small $\Phi$ additive genetic effects necessarily predominate in the correlation between relatives. Regardless of inbreeding levels and complexity of relationship between parents, there can never exist a relationship

$$(K_0, K_1, K_2) = (\frac{1}{4}, 0, \frac{1}{2}),$$

and indeed amongst all relationships with coefficient of kinship $1/4$ the relationship between sibs whose parents are not inbred is the one with maximal $K_2$. (Sibs whose parents are inbred to the same degree have coefficients lying on the section $BR$ of the parabolic curve of figure 1.)

**SUMMARY**

It was pointed out by Trustrum (1961) that even for non-inbred pairs of relatives it is possible for all four cross-parental kinship coefficients to be non-zero, and hence that the expression often assumed for the correlation between such relatives is not completely general. Van Aarde (1975) has recently made the same comment. We derive a restriction on the space of attainable Cotterman coefficients for a relationship between two arbitrary non-inbred relatives. This restriction implies that the form of the expression for the correlation is in fact general, although the components cannot always be interpreted as parental kinships.

**REFERENCES**


