Ann. Hum. Genet., Lond. (1980), **44**, 95
Printed in Great Britain

Further remarks on estimating genetic correlations

BY CEDRIC A. B. SMITH
Galton Laboratory, University College London

1. INTRODUCTION

In a previous paper (Smith 1980) a method was suggested for estimating correlations between relatives. The purpose of this paper is to explore the matter further, with additional examples.

2. CORRECTIONS AND EXPLANATIONS OF FORMULAS

I first take the opportunity of apologizing for and correcting some errors in the previous paper. The most serious is in equation (5.5), which is clearly nonsensical as it stands. Suppose that we take as estimate of the population mean μ a weighted sum of the means in individual families; that is,

$$\mu^{\text{est}} = (\sum_F \omega_F \bar{x}_F)/\sum_F \omega_F. \tag{2.1}$$

Here we introduce a convention, followed in the rest of the paper, that to simplify notation any suffix (such as F here) which is summed over is written with a capital letter. Here $\bar{x}_F$ is the mean in family number F, and ω_F is some arbitrary (positive) weight attached to $\bar{x}_F$. We suppose that the weights ω_F attached to the different families are all of comparable order of magnitude, and we write Ω for their sum, $\Sigma\omega_F$ (note that in equation (4.1) of Smith (1980) the bar over x_F was accidentally omitted).

Suppose also that, as in equations (5.1) and (5.2) of the previous paper we write

$$\Delta_B = \Sigma\, w_F\, \xi_F^2 = \Sigma\, w_F(\bar{x}_F - \mu^{\text{est}})^2,$$
$$W = \Sigma\, w_F,$$

where the w_f are another set of positive weights (which may or may not be different from the ω_f). Then if the ω_f and w_f are constant, the correct formula for the expectation of Δ_B is

$$\mathscr{E}(\Delta_B) = \Sigma[(An_F^{-1} + B)(w_F - 2\Omega^{-1} w_F\, \omega_F + W\Omega^{-2}\omega_F^2)] \tag{2.2}$$

and not formula (5.5) of Smith (1980).

Here A and B are respectively the components of variance within and between families. If the sample contains a large number of families, the first term $\Sigma(An_F^{-1} + B)\, w_F$ is the dominant one, and the remaining terms may be neglected without much error.

If we take ω_f and w_f to be equal, $\omega_f = w_f$, then (2.2) simplifies to

$$\mathscr{E}(\Delta_B) = \Sigma[(An_F^{-1} + B)(w_F - W^{-1} w_F^2)]. \tag{2.3}$$

In particular, if we take $\omega_f = w_f = 1$, and the number of families is N_f, then

$$\mathscr{E}(\Delta_B) = (\text{N}_f - 1)[A\text{N}_F^{-1}\Sigma n_F^{-1} + B], \tag{2.4}$$

while if $\omega_f = w_f = n_f$, the number of individuals in family f, and we write $\Sigma n_F = N_1$, $\Sigma n_f^2 = N_2$, then

$$\mathscr{E}(\Delta_B) = A(\text{N}_f - 1) + B(N_1 - N_2/N_1). \tag{2.5}$$

0003-4800/80/0000-4410 $02.00 © 1980 University College London

These are well-known formulas. If we have already calculated Δ_B and found some reasonable estimate of A, then they can be used to find an estimate of B by replacing $\mathscr{E}(\Delta_B)$ on the left hand side by Δ_B itself and solving the resulting equation.

The explanation of equations (5.8) and (5.9) of the previous paper is also confused and incorrect (although the basic idea is sound). The following argument seems better (but would need detailed examination to state the precise conditions under which it would hold). Suppose that the distribution being considered depends on N_θ parameters θ_i, forming a vector $\boldsymbol{\theta}$. The true value of $\boldsymbol{\theta}$ in the population will be called $\boldsymbol{\theta}_0$. In each family f introduce a set of N_θ random functions $z_{fi}(\boldsymbol{\Phi}, \boldsymbol{\Theta})$, where the vector arguments $\boldsymbol{\Phi}$, $\boldsymbol{\Theta}$, to be specified more exactly later, have respectively N_Φ and N_θ components. Let $\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\Theta})$ be the vector with components

$$z_i(\boldsymbol{\Phi}, \boldsymbol{\Theta}) = \Sigma z_{Fi}(\boldsymbol{\Phi}, \boldsymbol{\Theta}) \tag{2.6}$$

(summed over all families F). By the law of large numbers, this vector will (under suitable conditions) have a variance of the order of magnitude N_f, the number of families. Let us suppose that there is some fixed vector $\boldsymbol{\Phi}_0$ such that $\boldsymbol{\Phi} - \boldsymbol{\Phi}_0$ is of order $\mathrm{N}_f^{-\frac{1}{2}}$. Then if $\boldsymbol{\Theta} - \boldsymbol{\theta}_0$ is also of order $\mathrm{N}_f^{-\frac{1}{2}}$, we have on expansion in a Taylor series and neglect of second order terms (of order $\mathrm{N}_f^{-\frac{1}{2}}$)

$$\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\Theta}) = \mathbf{z}(\boldsymbol{\Phi}_0, \boldsymbol{\Theta}_0) + \mathbf{a}(\boldsymbol{\Phi} - \boldsymbol{\Phi}_0) + \mathbf{b}(\boldsymbol{\Theta} - \boldsymbol{\theta}_0). \tag{2.7}$$

Here $\mathbf{a} = d\mathbf{z}/d\boldsymbol{\Phi}$ and $\mathbf{b} = d\mathbf{z}/d\boldsymbol{\Theta}$ are, strictly speaking, random variables, but to the order of approximation used here they can be replaced by their expected values, and so treated as constants. Define $\boldsymbol{\zeta}(\boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\theta})$ to be the expected value of the random variable $\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\Theta})$ on the assumption that $\boldsymbol{\Phi}$ is held constant, and that the parameter of the distribution takes the value $\boldsymbol{\theta}$. On taking expectations in (2.7) we find

$$\boldsymbol{\zeta}(\boldsymbol{\Phi}, \boldsymbol{\Theta}, \boldsymbol{\theta}) = \boldsymbol{\zeta}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}) + \mathbf{a}(\boldsymbol{\Phi} - \boldsymbol{\Phi}_0) + \mathbf{b}(\boldsymbol{\Theta} - \boldsymbol{\theta}_0). \tag{2.8}$$

If $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ is of order $\mathrm{N}_f^{-\frac{1}{2}}$, and we again neglect second order terms, we have

$$\boldsymbol{\zeta}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}) = \boldsymbol{\zeta}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}_0) + \mathbf{c}(\boldsymbol{\theta} - \boldsymbol{\theta}_0), \tag{2.9}$$

where $\mathbf{c}$ is a square matrix, of order of magnitude $\mathrm{N}_f^{-\frac{1}{2}}$, which in general will be non-singular. On combining (2.7), (2.8) and (2.9) we get

$$\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\theta}) - \boldsymbol{\zeta}(\boldsymbol{\Phi}, \boldsymbol{\theta}, \boldsymbol{\theta}) = [\mathbf{z}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0) - \boldsymbol{\zeta}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}_0)] + \mathbf{c}(\boldsymbol{\theta} - \boldsymbol{\theta}_0). \tag{2.10}$$

Now choose for $\boldsymbol{\Phi}$ some (reasonable) function $\boldsymbol{\Phi} = \boldsymbol{\phi}(\boldsymbol{\theta})$ of $\boldsymbol{\theta}$, and set $\boldsymbol{\Phi}_0 = \boldsymbol{\phi}(\boldsymbol{\theta}_0)$. Then, provided that $(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ is of order of magnitude $\mathrm{N}_f^{-\frac{1}{2}}$, we will also have $(\boldsymbol{\Phi} - \boldsymbol{\Phi}_0)$ of order $\mathrm{N}_f^{-\frac{1}{2}}$, as we required in deriving the above results. Let us choose for the equation of estimation of $\boldsymbol{\theta}$ the relation

$$\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\theta}) = \boldsymbol{\zeta}(\boldsymbol{\Phi}, \boldsymbol{\theta}, \boldsymbol{\theta}). \tag{2.11}$$

In effect, this means that we are setting the random variable $\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\theta})$ equal to its expectation, except that the expectation is calculated assuming that $\boldsymbol{\Phi}$ is constant, and that the parameter of the distribution is the same as the argument $\boldsymbol{\theta}$ of the function $\mathbf{z}(\boldsymbol{\Phi}, \boldsymbol{\theta})$. Then from (2.10) we get

$$\boldsymbol{\theta} = \boldsymbol{\theta}_0 - \mathbf{c}^{-1}[\mathbf{z}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0) - \boldsymbol{\zeta}(\boldsymbol{\Phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\theta}_0)]. \tag{2.12}$$

But the expression in square brackets is, by definition, a random variable of expectation zero (in the actual population, which has parameter $\boldsymbol{\theta}_0$) and order of magnitude $\mathrm{N}_f^{-\frac{1}{2}}$. Hence there is a solution of (2.11) which differs from $\boldsymbol{\theta}_0$ by a random variable which has order of magnitude $\mathrm{N}^{-\frac{1}{2}}$, as was required in the derivation of these results. Also the expectation of $\boldsymbol{\theta}$ is, to our order of approximation, equal to $\boldsymbol{\theta}_0$, i.e. the estimate is unbiased.

There are three other small errors in Smith (1980) which may also be mentioned here. On page 272, line 6 up, the definition should read '$\xi_f = \bar{x}_f - \mu$'. On page 277, in (8.1) the second factor in the denominator should read '$A_{qq}+B_{qq}$'. The data of Table 1 concern pattern intensity on soles, not on fingers.

3. GENERAL THEORY

We now consider the following situation. We suppose that we have collected from the population a random (or otherwise unbiased) sample divided into families, numbered off as $f = 1, 2, \ldots, \text{N}_f$. Two distinct families are supposed to have statistically independent measurements.

Each individual in a family is assigned to some 'position' π in that family (such as 'proband', 'mother (of proband)', 'father', 'sister', etc.). On each individual we measure a set of 'characters' y_u ($u = 1, 2, \ldots, \text{N}_y$), including possibly presence-or-absence characters which can be scored as 1 or 0 respectively. To avoid complications, we insist that every individual is measured for every character.

The combination of a particular position, π, and a particular character y_u measured on individuals at that position, will be called a 'positioned character', and will in general be denoted by a single suffix, as x_p. Thus 'father's height' or 'mother's arm length' or 'sister's ridge count' are positioned characters. Two positioned characters (π, y_u) and (π, y_v) sharing a position π will be called 'co-positioned characters'. Thus, since we assume that every character is measured on every individual, co-positioned characters are those which are measured on the same individual, such as 'sister's height' and 'sister's arm length'. We introduce a symbol s_{pq} which takes the value 1 when p and q are co-positional, and 0 otherwise.

We make the following assumptions about the distributions of the characters:

(3A) For any positioned character x_p and family f, the n_{fp} observed values of x_p are a random sample from some distribution, with true mean μ_{fp} and variance v_{fpp}. This assumption might break down if, for example, there is a maternal age effect. We further assume that the mean μ_{fp} and variance v_{fpp} are independent of the number of observations n_{fp}.

Further, if x_p and x_q are co-positioned, we assume that the pairs of values (x_p, x_q) measured on the same individual are drawn from a bivariate distribution with covariance v_{fpq}. If x_p and x_q are not co-positioned, we assume that $v_{fpq} = 0$.

(3B) With certain positions π there is at most one individual occupying this position in each family. For example, if one individual is identified as 'proband' in each family, there is at most one 'proband's mother'. If x_p and x_q are positioned characters both associated with such a position π (with possibly $p = q$), then we will set conventionally

$$\mu_{fp} = \bar{x}_{fp}; \quad \mu_{fq} = \bar{x}_{fq}; \quad v_{fpp} = v_{fpq} = v_{fqq} = 0; \tag{3.1}$$

which is not unreasonable since there is no way of defining a within-family variance or covariance for such a character.

(3C) For fixed p, the family true means μ_{fp} are assumed to be drawn from a distribution with mean μ_p and variance B_{pp}. Similarly, the pairs of means μ_{fp}, μ_{fq} are assumed to be drawn from a bivariate distribution with covariance B_{pq}. The B_{pp}, B_{pq} are the 'between family components of variance and covariance'.

(3D) Similarly the within family covariance v_{fpq} (or variance if $p = q$) may vary between families f, and is assumed to be drawn from a distribution with true mean A_{pq}, the 'within family component of covariance (or variance)'. The covariance between x_p and x_q in the population as a whole is then $(A_{pq}+B_{pq})$, so that the correlation between x_p and x_q is

$$\rho_{pq} = \frac{A_{pq}+B_{pq}}{\sqrt{[(A_{pp}+B_{pp})(A_{qq}+B_{qq})]}}. \tag{3.2}$$

(3E) We can sometimes assume that there are relations between the means of different positioned characters x_p, x_q, and also between their components of variance and covariance. Thus suppose that daughter's height is denoted by x_1, and mother's height by x_5. Then in the absence of selection or secular change it would be reasonable to assume that the true mean of each in the population was equal to the mean for females in general, so that $\mu_1 = \mu_5$. Even if we were not willing to assume this for heights, we could assume it for finger-ridge counts, at least until there is evidence to the contrary. Similarly, we could assume that the total variance is the same in daughters and mothers,

$$A_{11}+B_{11} = A_{55}+B_{55} = B_{55},$$

since there is no intra-family variability in mothers, so that $A_{55} = 0$. In any case, since the matrices A_{pq} and B_{pq} are symmetric by the way they are defined, we must have the relation $B_{15} = B_{51}$.

To accommodate these relations, we introduce a set of independent 'mean parameters' M_i such that any population character mean μ_p can be expressed as a known linear combination of the M_i:

$$\mu_p = \Sigma m_{Ip} M_I, \tag{3.3}$$

(recall that by the summation convention previously established, the use of a capital suffix I means that the summation is over all possible values of I, say from $I = 1$ to N_M). For example, we might take μ_1 = mean daughter's height $= \mu_5$ = mean mother's height $= M_1$, so that we would have

$$m_{11} = 1, \quad m_{i1} = 0 \quad \text{for} \quad i \neq 1;$$
$$m_{15} = 1, \quad m_{i5} = 0 \quad \text{for} \quad i \neq 1.$$

Similarly we take a set of independent variance parameters, C_j ($j = 1$ to N_C), such that each of the A_{pq} and B_{pq} can be expressed as known linear functions of the C_j:

$$A_{pq} = \Sigma a_{Jpq} C_J; \quad B_{pq} = \Sigma b_{Jpq} C_J. \tag{3.4}$$

Thus if, as above, we assume that $B_{55} = A_{11}+B_{11}$ and $A_{55} = 0$, we have on setting $C_1 = A_{11}$ and $C_2 = B_{11}$,

$$a_{111} = 1, \quad a_{211} = 0, \quad b_{111} = 0, \quad b_{211} = 1,$$
$$a_{155} = 0, \quad a_{255} = 0, \quad b_{155} = 1, \quad b_{255} = 1.$$

4. PRELIMINARY ESTIMATES OF μ_p ,A_{pq}, B_{pq}

A simple method of finding preliminary estimates of μ_p, A_{pq}, B_{pq} is as follows. Let π denote any position in the family, and let x_p, x_q be positioned characters associated with this position (with possibly $p = q$). Let $n_{f\pi} = n_{fp} = n_{fq}$ be the number of individuals in family f in position π. If

$n_{f\pi} > 0$, we calculate the following, where in each case the summation is over these individuals in family f and position π:

$$\text{total,}\quad \mathrm{T}_{fp} = \Sigma x_p,$$

$$\text{crude sum of products,}\quad \mathrm{U}_{fpq} = \Sigma x_p x_q \tag{4.1}$$

(= crude sum of squares when $p = q$). Then in the usual way we find, within the family f, the sample values of the following:

$$\begin{aligned} &\text{mean,} && \bar{x}_{fp} = \mathrm{T}_{fp}/n_{fp}, \\ &\text{codeviance,} && \Delta_{fpq} = \mathrm{U}_{fpq} - \mathrm{T}_{fp}\bar{x}_{fq}, \\ &\text{degrees of freedom,} && \nu_{f\pi} = \nu_{fp} = \nu_{fq} = n_{f\pi} - 1 = n_{fp} - 1. \end{aligned} \tag{4.2}$$

When $n_{f\pi} = 0$, we cannot find $\bar{x}_{fp}$ or Δ_{fpq} in this way, but we will conventionally set $\bar{x}_{fp} = 0$, $\nu_{fp} = 0$, and $\Delta_{fpq} = 0$ for all p, q associated with π. We also set $\Delta_{fpq} = 0$ whenever x_p, x_q are not co-positioned. Summing over all families we then get for positioned characters x_p, x_q, the following values,

$$\begin{aligned} &\text{total number of observations,} && n_p = \Sigma\, n_{Fp}, \\ &\text{grand total,} && \mathrm{T}_p = \Sigma\, \mathrm{T}_{Fp}, \\ &\text{total codeviance,} && \Delta_{pq} = \Sigma\, \Delta_{Fpq}. \end{aligned}$$

If p, q are co-positioned, this codeviance has total degrees of freedom

$$\nu_\pi = \nu_p = \nu_q = \Sigma\, \nu_{F\pi}.$$

One obvious unbiased estimate of μ_p is then the grand mean, $\bar{x}_p = \mathrm{T}_p/n_p$. (One could also use the unweighted mean of the $\bar{x}_{fp}$, taken over those families f for which $n_{fp} > 0$.) An unbiased estimate of A_{pq} is Δ_{pq}/ν_p, the 'mean sum of products within families' of an analysis of covariance.

It is a complicated matter to find formulas giving a strictly unbiased estimate of B_{pq}. But if the object is mainly to find a preliminary value to be improved by further calculation, strict unbiasedness is not important, and any reasonable estimate will do. Consider therefore only those families for which n_{fp} and n_{fq} are both positive; suppose that there are N_f^{pq} such families in all. Calculate the 'unweighted codeviance'

$$\Delta_{pq}^B = \Sigma(\bar{x}_{Fp} - \bar{x}_p)(\bar{x}_{Fq} - \bar{x}_q), \tag{4.3}$$

summed over these families. To a first approximation

$$\mathscr{E}(\Delta_{pq}^B) = \Sigma\, n_{Fp}^{-1}\, s_{pq} A_{pq} + \mathrm{N}_f^{pq} B_{pq}, \tag{4.4}$$

(where, as defined above, $s_{pq} = 1$ when p, q are co-positioned, otherwise 0). Hence

$$B_{pq}^{\text{est}} = (\Delta_{py}^B - \Sigma\, n_{Fp}^{-1}\, s_{pq} A_{pq})/\mathrm{N}_f^{pq} \tag{4.5}$$

is a reasonable (consistent and approximately unbiased) estimate of B_{pq}.

5. ITERATIVE IMPROVEMENTS OF THE ESTIMATES

We now proceed as in the previous paper (Smith, 1980) but with minor modifications.

We suppose that we have already found from the sample the values of n_{fp} (= number of measurements of positioned character p in family f = number of individuals of position π in family f, where p belongs to π), so that $n_{fp} = n_{fq}$ if p and q are co-positioned; $\bar{x}_{fp}$ = mean of positioned character p in family f; Δ_{fpq} = codeviance (sum of products of deviations) of positioned characters x_p, x_q within family f (= 0 when p, q are not co-positioned); ν_{fp} = number of degrees of freedom of Δ_{fpq} (= $n_{fp}-1$, except that it is 0 when $n_{fp}=0$); $\Delta_{pq}=\Sigma\Delta_{Fpq}$; $\nu_p=\Sigma\nu_{Fp}$. We also have obtained preliminary estimates of A_{pq} and B_{pq}. The symbol $\mathbf{A}$ will denote the matrix with components A_{pq}, $\bar{\mathbf{x}}_f$ the vector with components $\bar{x}_{fp}$, and similarly wherever applicable. We will write

$$\boldsymbol{\alpha} = \mathbf{A}^{-1}. \tag{5.1}$$

We suppose that in the model we are considering, μ_p, the true (population) mean of positioned character x_p, is a known linear function of some 'mean parameters' M_i:

$$\mu_p = \Sigma m_{Ip} M_I, \tag{5.2}$$

(summed over I, in accordance with our convention that summation is over suffixes denoted by capital letters), where the m_{ip} are known by hypothesis, and similarly there are 'variance parameters' C_i such that for the population values (though not necessarily for sample estimates)

$$A_{pq} = \Sigma a_{Ipq} C_I; \quad B_{pq} = \Sigma b_{Ipq} C_I. \tag{5.3}$$

For each family f we would now like to calculate the matrix $\mathbf{u}_f$ with elements

$$u_{fpq} = n_{fp}^{-1} s_{pq} A_{pq} + B_{pq},$$

and then invert it to get a matrix $\boldsymbol{\omega}_f$. However, if $n_{fp}=0$ for some p (as will often happen) this means that some elements of u_f will be infinite. The inverse will still be finite, but the calculation may not be entirely simple on a computer, which is not equipped to deal with infinite elements. We can overcome the problem by defining the 'generalized reciprocal' r_{fp} of n_{fp} as

$$r_{fp} = \begin{cases} 1/n_{fp} & \text{if} \quad n_{fp} \neq 0, \\ 0 & \text{if} \quad n_{fp} = 0. \end{cases} \tag{5.4}$$

We set

$$u_{fpq} = r_{fp} s_{pq} A_{pq} + B_{pq}, \tag{5.5}$$

and then perform a 'controlled inversion' of $\mathbf{u}_f$ as follows:

(α) for all p for which $n_{fp}=0$, delete the pth row and column,

(β) invert the matrix which remains,

(γ) in the places where rows and columns were deleted under (α), reinsert rows and columns with all elements zero. The resulting matrix is $\boldsymbol{\omega}_f$.

Now find (for all $i, j = 1$ to N_M, where N_M is the number of mean parameters M_i)

$$\left.\begin{aligned} X_i &= \Sigma\, m_{iP}\, \omega_{FPQ}\, \bar{x}_{FQ}, \\ \Omega_{ij} &= \Sigma\, m_{iP}\, m_{jQ}\, \omega_{FPQ}, \\ \boldsymbol{\Psi} &= \boldsymbol{\Omega}^{-1}, \\ \mathbf{M}^{\text{est}} &= \boldsymbol{\Psi}\mathbf{X}. \end{aligned}\right\} \tag{5.6}$$

$\mathbf{M}^{\text{est}}$ is then an estimate of the mean parameter vector $\mathbf{M}$, and from it we find

$$\boldsymbol{\mu}^{\text{est}} = \mathbf{m}^T\mathbf{M}^{\text{est}} \tag{5.7}$$

(^{T} denoting transposition), which is an estimate of the population mean vector $\boldsymbol{\mu}$. Let

$$\begin{aligned}
\xi_{fp} &= \bar{x}_{fp} - \mu_p^{\text{est}}, \\
S_{pq}^A &= \Sigma\, \Delta_{QP}\, \alpha_{Qp}\, \alpha_{qP} + \Sigma\, \omega_{FQp}\, \omega_{FqP}\, r_{Fp}\, s_{pq}\, \xi_{FP}\, \xi_{FQ}, \\
S_{pq}^B &= \Sigma\, \omega_{FQp}\, \omega_{FqP}\, \xi_{FP}\, \xi_{FQ}, \\
S_i &= \Sigma\, a_{iPQ}\, S_{PQ}^A + \Sigma\, b_{iPQ}\, S_{PQ}^B, \\
T_{pq,\,kl}^{AA} &= \nu_p\, s_{pq}\, \alpha_{lp}\, \alpha_{qk} + \Sigma\, r_{Fp}\, r_{Fk}\, s_{pq}\, s_{kl}\, \omega_{Flp}\, \omega_{Fqk}, \\
T_{pq,\,kl}^{AB} &= \Sigma\, r_{Fp}\, s_{pq}\, \omega_{Flp}\, \omega_{Fqk}, \\
T_{pq,\,kl}^{BA} &= \Sigma\, r_{Fk}\, s_{kl}\, \omega_{Flp}\, \omega_{Fqk} \\
T_{pq,\,kl}^{BB} &= \Sigma\, \omega_{Flp}\, \omega_{Fqk}, \\
T_{ij} &= \Sigma\, a_{iPQ}\, a_{jKL}\, T_{PQ,KL}^{AA} + \Sigma\, a_{iPQ}\, b_{jKL}\, T_{PQ,KL}^{AB} \\
&\quad + \Sigma\, b_{iPQ}\, a_{jKL}\, T_{PQ,KL}^{BA} + \Sigma\, b_{iPQ}\, b_{jKL}\, T_{PQ,KL}^{BB}, \qquad (5.8)
\end{aligned}$$

$$\mathbf{t} = \mathbf{T}^{-1}, \qquad \mathbf{C}^{\text{est}} = \mathbf{tS}. \qquad (5.9)$$

Then $\mathbf{C}^{\text{est}}$ is an estimate of the variance parameter vector $\mathbf{C}$. From it we can find new estimates of $\mathbf{A}$ and $\mathbf{B}$ using equation (5.3). The whole calculation can be repeated taking the previous estimates as new provisional values and obtaining new estimates of $\mathbf{M}$, $\boldsymbol{\mu}$, $\mathbf{C}$, $\mathbf{A}$, $\mathbf{B}$. These can in turn be taken as provisional values, and so on iteratively. It seems that in general, in large samples, the process will converge rapidly to some final estimates. A later example shows that occasionally troublesome samples can occur with no obvious convergence. Leaving such exceptional cases aside, the arguments of Section 2 show that such estimates are nearly unbiased. For since $\mathscr{E}(\bar{x}_{fq}) = \mu_q = \Sigma\, m_{Iq} M_I$, a substitution of this in the expressions for X_i, Ω_{ij} in (5.6) gives

$$\mathscr{E}(\mathbf{X}) = \mathbf{\Omega M}, \qquad (5.10)$$

provided that the $\boldsymbol{\omega}_f$ are kept constant, i.e. provided that $\mathbf{A}$ and $\mathbf{B}$ are held constant. Hence the equation $\mathbf{M}^{\text{est}} = \mathbf{\Psi X}$, which is equivalent to $\mathbf{X} = \mathbf{\Omega M}^{\text{est}}$, amounts to setting $\mathbf{X}$ equal to its expectation. (Formally $\mathbf{X}$ corresponds to $\mathbf{z}$ of (2.11), $\mathbf{A}$ and $\mathbf{B}$ to $\mathbf{\Phi}$, and $\mathbf{M}$ to $\boldsymbol{\theta}$.) Thus we conclude that $\mathbf{M}^{\text{est}}$ is a nearly unbiased estimate of $\mathbf{M}$. Similarly, if $\boldsymbol{\mu}$ is held fixed we find that

$$\mathscr{E}(\xi_{fp}\, \xi_{fq}) = u_{fpq},$$

and hence that $\mathscr{E}(\mathbf{S}) = \mathbf{TC}$. Hence the equation $\mathbf{C}^{\text{est}} = \mathbf{tS}$, equivalent to $\mathbf{S} = \mathbf{TC}^{\text{est}}$, sets $\mathbf{S}$ equal to its expectation, and hence its solution $\mathbf{C}^{\text{est}}$ is a nearly unbiased estimate of $\mathbf{C}$. Furthermore, if we assume that the distributions are homoscedastic and normal, then an argument like that of Smith (1980) shows that $\mathbf{M}^{\text{est}}$, $\mathbf{C}^{\text{est}}$ are the evaluates (maximum likelihood estimates) of $\mathbf{M}$ and $\mathbf{C}$. In fact, the efficient score vectors for $\mathbf{M}$ and $\mathbf{C}$ are $\mathbf{U}^M = \mathbf{X} - \mathbf{\Omega M}$, $\mathbf{U}^C = \frac{1}{2}(\mathbf{S} - \mathbf{TC})$ respectively, with corresponding expected information matrices $\mathbf{\Omega}$ and $\frac{1}{2}\mathbf{T}$, and error variance matrices $\mathbf{\Psi}$ and $2\mathbf{t}$. Apart from the use of 'expected' rather than 'observed' information, the successive estimates $\mathbf{M}^{\text{est}} = \mathbf{\Psi X}$, $\mathbf{C}^{\text{est}} = \mathbf{tS}$ are those got by the Newton–Raphson solution of the maximum-likelihood equation.

6. EXAMPLES OF VARIANCE COMPONENT AND CORRELATION ESTIMATES

Example 1. Daughter–mother correlation for the data of Table 1 of Smith (1980) (there called 'sister–mother correlation'). Let positions $p = 1$, 2 correspond to daughters ('sisters') and mothers respectively; since these are different, $A_{12} = 0$. Since each family has only one mother,

Table 1. *Successive estimates relating to pattern intensity on soles in daughters and mothers*

Iteration	$\mu_1 = \mu_2$ $= M_1$	$A_{11} = C_1$	$B_{11} = C_2$	$B_{12} = C_3$	B_{22}
0	—	1·010	0·830	1·092	2·424
1	3·21	0·931	1·042	0·960	1·972
2	3·31	0·962	1·173	1·102	2·135
3	3·32	0·953	1·190	1·115	2·146
4	3·32	0·953	1·193	1·119	2·146
5	3·32	0·952	1·194	1·119	2·146

Table 2. *Data on PIP and PIF in sibships*

	Sisters			Brothers		
Family number	No.	Mean PIP	Mean PIF	No.	Mean PIP	Mean PIF
f	$n_{f1} = n_{f2}$	$\bar{x}_{f1}$	$\bar{x}_{f2}$	$n_{f3} = n_{f4}$	$\bar{x}_{f3}$	$\bar{x}_{f4}$
1	4	2·5	13·5	2	1·5	13·5
2	4	4·75	11·75	1	2·0	11·0
3	1	3·0	9·0	2	5·0	15·5
4	2	3·0	15·0	1	4·0	10·0
5	3	3·33	10·33	3	2·33	6·0
6	1	3·0	16·0	1	4·0	16·0
7	1	2·0	18·0	2	2·5	10·5
8	2	4·5	13·0	2	4·0	12·0
9	2	3·5	12·0	3	3·0	10·0
10	3	4·33	11·0	1	8·0	15·0
11	4	2·75	11·0	2	4·0	10·5
12	2	8·0	15·0	3	3·67	15·67
13	2	3·5	11·5	1	3·0	11·0
14	—	—	—	1	3·0	14·0
15	3	3·0	9·0	1	3·0	16·0
Total (grand mean)	34	(3·68)	(12·12)	26	(3·39)	(11·58)

$$\mathbf{\Delta} = \begin{bmatrix} 35·33 & 9·92 & 0 & 0 \\ 9·92 & 98·92 & 0 & 0 \\ 0 & 0 & 16·33 & 9·17 \\ 0 & 0 & 9·17 & 104·17 \end{bmatrix}; \quad \mathbf{v} = \begin{bmatrix} 20 \\ 20 \\ 11 \\ 11 \end{bmatrix}; \quad \mathbf{A} \simeq \begin{bmatrix} 1·77 & 0·50 & 0 & 0 \\ 0·50 & 4·94 & 0 & 0 \\ 0 & 0 & 1·49 & 0·83 \\ 0 & 0 & 0·83 & 9·47 \end{bmatrix};$$

$$\mathbf{B} \simeq \begin{bmatrix} 1·09 & 0·15 & 0·38 & 1·07 \\ 0·15 & 4·20 & -0·66 & 0·23 \\ 0·38 & -0·66 & 1·19 & 1·79 \\ 1·07 & 0·23 & 1·79 & 2·15 \end{bmatrix}$$

we set $A_{22} = 0$. In our model we assume, as seems reasonable, that daughters and mothers have the same mean, $\mu_1 = \mu_2$, and the same total variance, $A_{11} + B_{11} = A_{22} + B_{22} = B_{22}$. Thus there is just one mean parameter, $M_1 = \mu_1 = \mu_2$, and three (co)variance parameters, $C_1 = A_{11}$, $C_2 = B_{11}$, $C_3 = B_{12} = B_{21}$, with $B_{22} = C_1 + C_2$. Hence the m_{ip}, a_{ipq} and b_{ipq} take the values

$$m_{11} = m_{12} = 1 = a_{111} = b_{211} = b_{312} = b_{321} = b_{122} = b_{222},$$

all others being zero. As first provisional estimate of A_{11} we take the mean square within

Table 3. *Successive estimates relating to sisters' PIP, PIF*

Iter-ation	μ_1	μ_2	A_{11}	$A_{12} = A_{21}$	A_{22}	B_{11}	$B_{12} = B_{21}$	B_{22}
0	—	—	1·77	0·50	4·94	1·09	0·15	4·20
1	3·69	12·78	1·78	0·43	5·23	1·10	0·27	2·62
2	3·70	12·32	1·78	0·40	5·43	1·10	0·32	2·20
3	3·70	12·30	1·78	0·39	5·52	1·09	0·34	2·04
4	3·70	12·29	1·78	0·38	5·57	1·09	0·35	1·97
5	3·70	12·28	1·78	0·38	5·58	1·09	0·36	1·94
6	3·70	12·28	1·78	0·38	5·59	1·09	0·36	1·93

sisterships, 1·010. Equation (4.5) gives provisional estimates $B_{11} = 0{\cdot}830$, $B_{12} = B_{21} = 1{\cdot}092$, $B_{22} = 2{\cdot}424$. (Note that the relationship $B_{22} = A_{11} + B_{11}$ does not hold for these first provisional estimates, but it will do so for all subsequent ones.) The sequences of successive estimates obtained using our iterative method is shown in Table 1. Clearly after 5 iterations we have found the final values to 3 places of decimals. Also we find

$$\mathbf{\Psi} = [0{\cdot}11]; \quad \mathbf{t} = \begin{bmatrix} 0{\cdot}081 & -0{\cdot}055 & 0{\cdot}003 \\ -0{\cdot}055 & 0{\cdot}223 & 0{\cdot}166 \\ 0{\cdot}003 & 0{\cdot}166 & 0{\cdot}203 \end{bmatrix},$$

so that the standard error of $M_1 = \mu_1 = \mu_2$ is $\surd 0{\cdot}11 = 0{\cdot}33$, while the standard errors of C_1, C_2, C_3 (assuming normality) are respectively

$$\surd(2\times 0{\cdot}081) = 0{\cdot}40, \quad \surd(2\times 0{\cdot}223) = 0{\cdot}67 \quad \text{and} \quad \surd(2\times 0{\cdot}203) = 0{\cdot}64.$$

The estimate of daughter–mother correlation is therefore

$$\rho_{12} = \frac{B_{12}}{\surd[(A_{11}+B_{11})B_{22}]} = \frac{B_{12}}{A_{11}+B_{11}} = \frac{1{\cdot}119}{2{\cdot}146} = 0{\cdot}52.$$

As ρ_{12} is here a function of A_{11}, B_{11} and B_{12}, its approximate standard error can be obtained by the usual 'delta method' from the error variances and covariances of A_{11}, B_{11} and B_{12}. It comes to S.E. $(\rho_{12}) = 0{\cdot}19$.

Example 2. Cross-correlation between pattern intensity on palms (PIP) in one female and pattern intensity on fingers (PIF) in her sister. Some values of x_1 = PIP and x_2 = PIF in sisters, x_3 = PIP and x_4 = PIF for brothers in a sample of 15 sibships extracted from Loesch's data are summarized in Table 2, in the forms of numbers of sisters and brothers and their sibship means. This table also gives the deviance matrix $\mathbf{\Delta}$ (sums of squares and products within sibships), the vector $\mathbf{v}$ of corresponding degrees of freedom, and the first provisional estimates of A and B. As we are interested only in sisters in this example, rows and columns relating to brothers (i.e. to x_3 and x_4) can be ignored. Because x_1 and x_2 are measured on the same individuals, we have non-zero A_{12}, and therefore 6 variance parameters, $A_{11}, A_{12} = A_{21}, \ldots$, which we can number off as C_1 to C_6. The successive estimates are given in Table 3. After 6 iterations the values have stabilized to 2 decimals. The correlation between x_1 in one individual and x_2 in her sister is

$$\rho_{12}^{\text{sis}} = \frac{B_{12}}{\surd[(A_{11}+B_{11})(A_{22}+B_{22})]} = \frac{0{\cdot}36}{\surd[2{\cdot}87\times 7{\cdot}25]} = 0{\cdot}08,$$

whereas the correlation between x_1 and x_2 in the *same* individual is

$$\rho_{12}^{\female} = \frac{A_{12}+B_{12}}{\surd[(A_{11}+B_{11})(A_{22}+B_{22})]} = \frac{0{\cdot}74}{\surd[2{\cdot}87\times 7{\cdot}25]} = 0{\cdot}16.$$

Table 4. *Successive estimates relating to pattern-intensity on soles in sisters and brothers*

Iteration	μ_1	μ_2	A_{11}	A_{22}	B_{11}	$B_{12} = B_{21}$	B_{22}
0	—	—	1·01	0·83	0·83	1·40	1·36
1	0·88	5·51	7·24	4·85	10·80	1·66	−9·89
2	3·01	3·70	1·02	0·85	0·56	1·69	2·32
3	3·58	3·28	0·11	−0·73	0·76	1·48	3·51
4	3·02	3·52	0·94	−0·05	0·59	1·27	2·17
5	2·49	3·76	1·30	0·87	−0·52	0·38	2·34

Example 3. Sister–brother correlation for one character, e.g. PIP, such as the correlation between x_1 and x_3 in Table 3. Since sisters and brothers have different positions, $A_{13} = 0$, and we have only 5 variance parameters. Otherwise the calculation proceeds very much as in the previous example. After 6 iterations we find

$$\boldsymbol{\mu} = \begin{bmatrix} 3{\cdot}70 \\ 3{\cdot}42 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1{\cdot}79 & 0 \\ 0 & 1{\cdot}79 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1{\cdot}07 & 0{\cdot}33 \\ 0{\cdot}33 & 0{\cdot}54 \end{bmatrix},$$

$$\rho = 0{\cdot}33/\sqrt{[(1{\cdot}79+1{\cdot}07)(1{\cdot}79+0{\cdot}54)]} = 0{\cdot}13.$$

Example 4. Sister–brother correlation for the data of Table 1 of Smith (1980). This might be expected to behave just like Example 3 above, but instead successive estimates fluctuate wildly; the first five iterations are shown in Table 4. This seems to be related to the fact that for the first provisional estimates (iteration 0) $B_{12}^2 > B_{11}B_{22}$, which cannot happen in a true variance matrix. Presumably this is a random fluctuation due to the smallness of the sample, and convergence would be restored in a sufficiently large sample. Otherwise one would have to question the correctness of the assumed model.

7. DISCUSSION

Although in all our examples we have for simplicity considered only two characters at one time, there is no difficulty of principle in dealing with many characters simultaneously, e.g. those of Table 1 of Smith (1980) or Table 2 of this paper. But the calculations become considerably heavier. It is also possible in principle to include cousins. In the simplest one-character model in which sex-differences are ignored there are 3 variance parameters, $C_1 = A =$ the variance within sibships, $C_2 = B =$ the component between sibships within cousinships, $C_3 = \Gamma =$ the component between cousinships. Thus if μ_p denotes the true mean of a sibship p, we have

$$B_{pq} = \operatorname{cov}(\mu_p, \mu_q) = B + \Gamma \text{ when } p = q,$$

$= \Gamma$ when p and q are distinct sibships within the same cousinship, and 0 otherwise. Since the B_{pq} are known linear functions of the C_i, the method described here can be directly applied. The correlations are then:

$$\text{sib–sib correlation} = (B+\Gamma)/(A+B+\Gamma),$$

$$\text{cousin–cousin correlation} = \Gamma/(A+B+\Gamma). \tag{7.1}$$

The references given in the previous paper are relevant here, but it does not seem necessary to repeat them in detail; the reader can find the list in Smith (1980). But since that paper went to press an extensive (but still incomplete) bibliography of about a thousand related papers has been published (Sahai, 1979).

SUMMARY

A method of estimating correlations between relatives given in a previous paper in *Annals of Human Genetics* has been further explored, explained, and illustrated by examples.

I am indebted to Dr Danuta Z. Loesch for permission to use her data.

REFERENCES

Sahai, Hardeo (1979). A bibliography on variance components. *International Statistical Review* **47**, 177–222.

Smith, C. A. B. (1980). Estimating genetic correlations. *Annals of Human Genetics* **43**, 265–284.