THE COMPONENTS OF VARIANCE AND THE CORRELATIONS BETWEEN RELATIVES IN SYMMETRICAL RANDOM MATING POPULATIONS^{1*}

T. W. HORNER AND O. KEMPTHORNE²

Statistical Laboratory, Iowa State College, Ames, Iowa

Received August 10, 1954

THE general problem of the partition of genotypic variance and the correlation between non-inbred relatives in random mating populations was solved by KEMPTHORNE (1954, 1955). The purpose of the present paper is to apply these results to the case of random mating populations for which the genotypic value of a genotype is symmetrical with respect to the components, one from each locus, of the genotype. In the present paper it is supposed that there are only two alleles at each locus with three phases as a consequence, these being specified as the nulliplex (—), simplex (+—) and duplex (++) conditions. Gene frequency will in the present paper be taken to be 1/2 for every locus. Additional assumptions are normal diploid behavior at meiosis, and no linkage. Attention will be paid particularly to certain classical gene models, which give the genotypic value in terms of the genotype. The ones to be considered are the following: complementary, duplicate factor, multiplicative and optimum number.

THE COMPONENTS OF THE GENOTYPIC VALUE IN SYMMETRICAL POPULATIONS

In general, a genotypic value is expressed in terms of the population mean, the average effect of each gene, a contribution from each pair of allelic genes, a contribution from each pair of non-allelic genes, a contribution from each possible set of three genes and so on.

Following the notation of KEMPTHORNE the n loci of a random mating population will be denoted by superfixes running from 1 to n . At the i th locus it will be supposed that there are m_i alleles which will be denoted by A_j^i where j runs from 1 to m_i . The gene frequency of the j th allele at the i th locus will be denoted by p_j^i . Because of random mating, the genotypic array of the population is

$$\prod_{i=1}^n \{\sum p_j^i A_j^i\}^2$$

The symbol $\prod_{i=1}^n A_{j_i}^i A_{k_i}^i$ will denote both the genotype of an individual and the genotypic value of that individual. Further definitions follow:

$$\mu = \prod_{i=1}^n [\sum_m p_m^i A_m^i] [\sum_n p_n^i A_n^i] = \text{mean of population}$$

* Part of the cost of mathematical formulae has been paid by the GALTON and MENDEL MEMORIAL FUND.

¹ Journal paper no. J-2701 of the Iowa Experiment Station, Ames, Iowa Project 890 and 169.

² Assistant professor and professor of statistics, respectively.

$$a_{j_i}^i = [A_{j_i}^i - \sum_m p_m^i A_m^i] [\sum_n p_n^i A_n^i] \prod_{i' \neq i} [\sum_m p_m^{i'} A_m^{i'}] [\sum_n p_n^{i'} A_n^{i'}]$$

= additive effect of gene $A_{j_i}^i$ in the population

$$d_{j_i k_i}^i = [A_{j_i}^i - \sum_m p_m^i A_m^i] [A_{k_i}^i - \sum_n p_n^i A_n^i] \prod_{i' \neq i} [\sum_m p_m^{i'} A_m^{i'}] [\sum_n p_n^{i'} A_n^{i'}]$$

= dominance deviation due to genes $A_{j_i}^i, A_{k_i}^i$ in the population

$$(a^i a^{i'})_{j_i j_{i'}} = [A_{j_i}^i - \sum_m p_m^i A_m^i] [\sum_n p_n^i A_n^i] [A_{j_{i'}}^{i'} - \sum_m p_m^{i'} A_m^{i'}],$$

$$[\sum_n p_n^{i'} A_n^{i'}] \prod_{i'' \neq i \neq i'} [\sum_m p_m^{i''} A_m^{i''}] [\sum_n p_n^{i''} A_n^{i''}]$$

= additive $\times$ additive deviation arising from gene $A_{j_i}^i$ at locus i and gene $A_{j_{i'}}^{i'}$ at locus i' ,

$$(a^i d^{i'})_{j_i k_i j_{i'} k_{i'}} = [A_{j_i}^i - \sum_m p_m^i A_m^i] [\sum_n p_n^i A_n^i] [A_{j_{i'}}^{i'} - \sum_m p_m^{i'} A_m^{i'}]$$

$$[A_{k_{i'}}^{i'} - \sum_n p_n^{i'} A_n^{i'}] \prod_{i'' \neq i \neq i'} [\sum_m p_m^{i''} A_m^{i''}] [\sum_n p_n^{i''} A_n^{i''}]$$

= additive $\times$ dominance deviation arising from gene $A_{j_i}^i$ at locus i , and genes $A_{j_{i'}}^{i'}$, $A_{k_{i'}}^{i'}$ at locus i' .

$$(d^i d^{i'})_{j_i k_i j_{i'} k_{i'}} = [A_{j_i}^i - \sum_m p_m^i A_m^i] [A_{k_i}^i - \sum_n p_n^i A_n^i] [A_{j_{i'}}^{i'} - \sum_m p_m^{i'} A_m^{i'}]$$

$$[A_{k_{i'}}^{i'} - \sum_n p_n^{i'} A_n^{i'}] \prod_{i'' \neq i \neq i'} [\sum_m p_m^{i''} A_m^{i''}] [\sum_n p_n^{i''} A_n^{i''}]$$

= dominance $\times$ dominance deviation arising from genes $A_{j_i}^i, A_{k_i}^i$ at locus i and genes $A_{j_{i'}}^{i'}, A_{k_{i'}}^{i'}$ at locus i' .

$(a^i a^{i'} a^{i''})_{j_i j_{i'} j_{i''}} =$ additive $\times$ additive $\times$ additive deviation arising from genes $A_{j_i}^i$ at locus i , $A_{j_{i'}}^{i'}$ at locus i' , and $A_{j_{i''}}^{i''}$ at locus i'' , etc.

When we take the special case of symmetry with gene frequency equal to 1/2 at each locus, we may omit the superscripts for the alleles. We note that any contribution may be expressed formally in terms of genotypic values by multiplying together in appropriate ways factors like

$$\sum p_m^i A_m^i = \left(\frac{A_0 + A_1}{2} \right)$$

and

$$A_0^i - \sum p_m^i A_m^i = \left(\frac{A_0 - A_1}{2} \right)$$

or

$$A_1^i - \sum p_m^i A_m^i = -\left(\frac{A_0 - A_1}{2} \right)$$

Hence any contribution to the genotypic values is obtained by multiplying factors $\left(\frac{A_0 - A_1}{2}\right)$ and $\left(\frac{A_0 + A_1}{2}\right)$ and adding the appropriate sign. Hence in the special case being considered,

$$\begin{aligned}\mu &= \left[\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right]^n \\ a_{0,i}^i &= -a_{1,i}^i = \frac{1}{2} \left(\frac{A_0 A_0}{2} - \frac{A_1 A_1}{2} \right) \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-1} \\ d_{0,0,i}^i &= -d_{0,1,i}^i = -d_{1,0,i}^i = d_{1,1,i}^i \\ &= \left(\frac{A_0 A_0}{4} - \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right) \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-1} \\ (a^i a^{i'})_{0,0,i'} &= -(a^i a^{i'})_{0,1,i'} = -(a^i a^{i'})_{1,0,i'} = (a^i a^{i'})_{1,1,i'} \\ &= \left(\frac{1}{2} \right)^2 \left(\frac{A_0 A_0}{2} - \frac{A_1 A_1}{2} \right)^2 \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-2} \\ (a^i d^{i'})_{0,0,0,i',0,i'} &= \pm \text{similar quantities} \\ &= \left(\frac{1}{2} \right) \left(\frac{A_0 A_0}{2} - \frac{A_1 A_1}{2} \right) \left(\frac{A_0 A_0}{4} - \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right) \\ &\quad \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-2} \\ (d^i d^{i'})_{0,0,0,i',0,i'} &= \pm \text{similar quantities} \\ &= \left(\frac{A_0 A_0}{4} - \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^2 \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-2} \\ (a^i a^{i'} a^{i''})_{0,0,0,i',0,i',0,i''} &= \left(\frac{1}{2} \right)^3 \left(\frac{A_0 A_0}{2} - \frac{A_1 A_1}{2} \right)^3 \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-3}\end{aligned}$$

In general letting $(a^{t_a} d^{t_d})$ be an effect which in its symbolization has a entering t_a times and d entering t_d times, and dropping the subscripts which determine the sign only, we have

$$\begin{aligned}(a^{t_a} d^{t_d}) &= \pm \left(+\frac{1}{2} \right)^{t_a} \left(-\frac{A_0 A_0}{2} + \frac{A_1 A_1}{2} \right)^{t_a} \left(\frac{A_0 A_0}{4} - \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{t_d} \\ &\quad \left(\frac{A_0 A_0}{4} + \frac{2A_0 A_1}{4} + \frac{A_1 A_1}{4} \right)^{n-t_a-t_d}\end{aligned}$$

The sign to be attached is essentially irrelevant because we are concerned only with the sum of squares of quantities like $(a^{t_a} d^{t_d})$ with subscripts. A rule can be specified easily.

THE GENE MODELS TO BE CONSIDERED

These are specified as follows, where y_0 denotes the number of $(--)$ loci, y_1 the number of $(+-)$ loci and y_2 the number of $(++)$ loci in the genotype.

Complementary model

$$\text{Genotypic value} = 0^{y_0} = \begin{cases} 0 & \text{when } y_0 > 0 \\ 1 & \text{when } y_0 = 0 \end{cases}$$

An effect is not produced if there are any (—) loci in the genotype. Conversely an effect is produced if the plus gene is present in the simplex (+) or duplex (++) condition at all loci. A special case of this model (the case in which the number of interacting loci, say n , is equal to two) is shown below.

Values of genotypes for complementary model ($n = 2$)

Phase at second locus	Phase at first locus		
	++	+-	--
++	1	1	0
+-	1	1	0
--	0	0	0

Duplicate Factors Model

$$\text{Genotypic value} = 1 - 0^{y_1+y_2} = \begin{cases} 0 & \text{when } y_1 + y_2 = 0 \\ 1 & \text{when } y_1 + y_2 > 0 \end{cases}$$

According to this equation an effect is produced when any locus is in the simplex or duplex phase.

Multiplicative Model

$$\text{Genotypic value} = e^{y_1 y_2}$$

Each locus in the simplex phase results in a multiplication of the genotypic value by e and each locus in the duplex phase results in a multiplication of the genotypic value by b .

Optimum Number Model

The optimum number model is represented by

$$\begin{aligned} \text{Genotypic value} &= -[y_1 + 2y_2 - (n + d)]^2 \\ &= -[d + y_0 - y_2]^2 \end{aligned}$$

where y_0 , y_1 , y_2 and n are used as defined elsewhere so that $(y_1 + 2y_2)$ is the number of plus genes of the genotype. The value of the genotype as given above is greatest when $y_1 + 2y_2 = n + d$. Thus $(n + d)$ is an optimum number of plus genes. Since the possible range in number of plus genes is from 0 to $2n$, d may have any value in the range $-n$ to n . The above model represents a special case of a more general one considered by WRIGHT (1935). He distinguished between a primary and a secondary scale for value of a genotype, the latter being a scale of desirability which would measure value in terms of net advantage to the organism. He postulated additivity

of locus effects relative to value on the primary scale, but that maximum value on the secondary scale was associated with an intermediate value on the primary scale.

THE CONTRIBUTIONS TO GENOTYPIC VALUE UNDER THE PARTICULAR MODELS

Any arbitrary contribution apart from sign has been shown above as

$$(a^{t_a} d^{t_d}) = \pm \left(\frac{1}{2}\right)^{t_a} \left(-\frac{A_0 A_0 + A_1 A_1}{2}\right)^{t_a} \left(\frac{A_0 A_0 - 2A_0 A_1 + A_1 A_1}{4}\right)^{t_d} \\ \times \left(\frac{A_0 A_0 + 2A_0 A_1 + A_1 A_1}{4}\right)^{n-t_a-t_d}.$$

Each of the factors on the right hand side may be expanded by the multinomial theorem thus:

$$\left(-\frac{A_0 A_0 + A_1 A_1}{2}\right)^{t_a} = \sum \frac{t_a!}{\lambda_{00}! \lambda_{02}!} \left(\frac{A_0 A_0}{2}\right)^{\lambda_{00}} \left(\frac{A_1 A_1}{2}\right)^{\lambda_{02}} (-1)^{\lambda_{00}}, \\ \left(\frac{A_0 A_0 - 2A_0 A_1 + A_1 A_1}{4}\right)^{t_d} = \sum \frac{t_d!}{\lambda_{10}! \lambda_{11}! \lambda_{12}!} \left(\frac{A_0 A_0}{4}\right)^{\lambda_{10}} \left(\frac{2A_0 A_1}{4}\right)^{\lambda_{11}} \\ \times \left(\frac{A_1 A_1}{4}\right)^{\lambda_{12}} (-1)^{\lambda_{11}},$$

and

$$\left(\frac{A_0 A_0 + 2A_0 A_1 + A_1 A_1}{4}\right)^{n-t_a-t_d} = \sum \frac{(n-t_a-t_d)!}{\lambda_{20}! \lambda_{21}! \lambda_{22}!} \left(\frac{A_0 A_0}{4}\right)^{\lambda_{20}} \left(\frac{2A_0 A_1}{4}\right)^{\lambda_{21}} \\ \times \left(\frac{A_1 A_1}{4}\right)^{\lambda_{22}}$$

where summation occurs over all portions of the primary exponent, t_a , t_d or $n - t_a - t_d$, as the case may be.

Hence we have

$$(a^{t_a} d^{t_d}) = \pm \left(\frac{1}{2}\right)^{t_a} E(-1)^{\lambda_{00}+\lambda_{11}} (A_0 A_0)^{\lambda_{00}+\lambda_{10}+\lambda_{20}} (A_0 A_1)^{\lambda_{11}+\lambda_{21}} (A_1 A_1)^{\lambda_{02}+\lambda_{12}+\lambda_{22}}$$

where E denotes mathematical expectation over the product of the multinomials specified above, and the symbol for a genotype is replaced by the genotypic value corresponding to that genotype. This product of the multinomials can be written as

$$f(\lambda_{00}, \lambda_{02} \text{ etc.}, \lambda_{22}) = \prod_{i=0}^2 \left[\frac{\theta_i!}{\prod_j \lambda_{ij}!} \prod_j \Delta_{ij}^{\lambda_{ij}} \right]$$

in which

$$\begin{array}{lll} \theta = t_a, & \theta_1 = t_d, & \theta_2 = n - t_a - t_d \\ \Delta_{02} = \frac{1}{2}, & \Delta_{12} = \frac{1}{4}, & \Delta_{22} = \frac{1}{4} \end{array}$$

$$\begin{array}{lll} \Delta_{00} = \frac{1}{2}, & \Delta_{11} = \frac{1}{2}, & \Delta_{21} = \frac{1}{2} \\ & \Delta_{10} = \frac{1}{4}, & \Delta_{20} = \frac{1}{4} \end{array}$$

Also we have that

$$(A_0A_0)^{\lambda_{00}+\lambda_{10}+\lambda_{20}}(A_0A_1)^{\lambda_{11}+\lambda_{21}}(A_1A_1)^{\lambda_{02}+\lambda_{12}+\lambda_{22}}$$

interpreted as the genotypic value instead of the genotype, is the genotypic value of a genotype having

$(\lambda_{00} + \lambda_{10} + \lambda_{20})$ nullplex loci

$(\lambda_{11} + \lambda_{21})$ simplex loci

$(\lambda_{02} + \lambda_{12} + \lambda_{22})$ duplex loci

and may be written as $Y_{y_0y_1y_2}$ where

$$y_2 = \lambda_{02} + \lambda_{12} + \lambda_{22}$$

$$y_1 = \lambda_{11} + \lambda_{21}$$

$$y_0 = \lambda_{00} + \lambda_{10} + \lambda_{20}$$

In the case of the complementary model, for instance

$$Y_{y_0y_1y_2} = 0^{y_0}.$$

We may now use elementary methods or the convenient multinomial expectation identity

$$E \prod_{i,j} \rho_{ij} \lambda_{ij} = \prod_i \left(\sum_j \rho_{ij} \Delta_{ij} \right)^{\theta_i}$$

given by HORNER (1952) to obtain the expectations. We have:

i	j	Δ_{ij}	ρ_{ij}		
2	2	$\frac{1}{4}$	1	$\theta_2 = n - t_a - t_d$	$\sum_j \rho_{2j} \Delta_{2j} = \frac{3}{4}$
	1	$\frac{1}{2}$	1		
	0	$\frac{1}{4}$	0		
1	2	$\frac{1}{4}$	1	$\theta_1 = t_d$	$\sum_j \rho_{1j} \Delta_{1j} = \frac{1}{4}$
	1	$\frac{1}{2}$	-1		
	0	$\frac{1}{4}$	0		
0	2	$\frac{1}{2}$	1	$\theta_0 = t_a$	$\sum_j \rho_{0j} \Delta_{0j} = \frac{1}{2}$
	0	$\frac{1}{2}$	0		

The results are shown in table 1.

TABLE 1
Values for $\langle a^{t_a t_d} \rangle$ under various models

Model	
Complementary	$\pm (\frac{1}{4})^n 3^{n-t_a-t_d} (-1)^{t_d}$
Duplicate factor	$\pm (-1)^{t_a} [0^{t_a+t_d} - (\frac{1}{4})^n]$
Multiplicative	$\pm (\frac{1}{4})^n (b-1)^{t_a} (b-2e+1)^{t_d} (b+2e+1)^{n-t_a-t_d}$
Optimum number	$\begin{cases} a^0 d^0 = -(d^2 + \frac{1}{2}n), a^1 d^0 = \pm d, a^0 d^1 = \pm \frac{1}{2} \\ a^2 d^0 = \pm \frac{1}{2}, a^1 d^1 = a^0 d^2 = 0, \\ a^{t_a} d^{t_d} = 0 \text{ for } t_a + t_d > 2 \end{cases}$

COMPONENTS OF GENOTYPIC VARIANCE

KEMPTHORNE (1954, 1955) defines the components of genetic variance as shown below.

$$\begin{aligned}
 E \left\{ \sum_i [a_{ji}^i]^2 \right\} &= \frac{1}{2} \sigma_A^2 \\
 E \left\{ \sum_i [d_{jki}^i]^2 \right\} &= \sigma_D^2 \\
 E \left\{ \sum_{i < i'} [(a^i a^{i'})_{jij_i'}]^2 \right\} &= \frac{1}{4} \sigma_{AA}^2 \\
 E \left\{ \sum_{i < i'} [(a^i d^{i'})_{jij_i'k_i'}]^2 \right\} &= \frac{1}{2} \sigma_{AD}^2 \\
 E \left\{ \sum_{i < i'} [(d^i d^{i'})_{jij_i'k_i'}]^2 \right\} &= \sigma_{DD}^2 \\
 E \left\{ \sum_{i < i' < i''} [(a^i a^{i'} a^{i'')_{jij_i'j_i''}}]^2 \right\} &= \frac{1}{8} \sigma_{AAA}^2
 \end{aligned}$$

Under the symmetry assumption,

$$E(a_{ji}^i)^2 = Ea_j^2 = p_0 a_0^2 + p_1 a_1^2 = p_0 a_0^2 + p_1 (-a_0)^2 = a^2, \text{ say.}$$

Likewise,

$$\begin{aligned}
 E(d_{jki}^i)^2 &= Ed_{jk}^2 = p_0 p_0 d_{00}^2 + p_0 p_1 d_{01}^2 + p_1 p_0 d_{10}^2 + p_1 p_1 d_{11}^2 \\
 &= p_0 p_0 d_{00}^2 + p_0 p_1 (-d_{00})^2 + p_1 p_0 (-d_{00})^2 + p_1 p_1 d_{00}^2 \\
 &= d^2, \text{ say.}
 \end{aligned}$$

The components, then, in the symmetrical case reduce as shown below.

$$\begin{aligned}
 \frac{1}{2} \sigma_A^2 &= na^2 = \frac{n!}{1! 0!(n-1)!} a^2 \\
 \sigma_D^2 &= nd^2 = \frac{n!}{0! 1!(n-1)!} d^2 \\
 \frac{1}{4} \sigma_{AA}^2 &= \frac{n(n-1)}{2} (aa)^2 = \frac{n!}{2! 0!(n-1)!} (aa)^2 \\
 \frac{1}{2} \sigma_{AD}^2 &= n(n-1)(ad)^2 = \frac{n!}{1! 1!(n-2)!} (ad)^2 \\
 \sigma_{DD}^2 &= \frac{n(n-1)}{2} (dd)^2 = \frac{n!}{0! 2!(n-2)!} (dd)^2 \\
 \frac{1}{8} \sigma_{AAA}^2 &= \frac{n(n-1)(n-2)}{6} (aaa)^2 = \frac{n!}{3! 0!(n-3)!} (aaa)^2 \\
 \frac{1}{4} \sigma_{AAD}^2 &= \frac{n(n-1)(n-2)}{2} (aad)^2 = \frac{n!}{2! 1!(n-3)!} (aad)^2 \\
 \frac{1}{2} \sigma_{ADD}^2 &= \frac{n(n-1)(n-2)}{2} (add)^2 = \frac{n!}{1! 2!(n-3)!} (add)^2
 \end{aligned}$$

$$\sigma_{DD}^2 = \frac{n(n-1)(n-2)}{3!} (ddd)^2 = \frac{n!}{0!3!(n-3)!} (ddd)^2$$

Letting $\sigma_{A^t_a D^t_d}^2$ be the component which has the subscript A appearing t_a times and the subscript D appearing t_d times we have

$$\sigma_{A^t_a D^t_d}^2 = \frac{n!}{t_a! t_d! (n - t_a - t_d)!} 2^{t_a} (a^{t_a} d^{t_d})^2$$

Hence we have the results shown in table 2.

TABLE 2
Formulae for components of genotypic variance

Model	Formula
Complementary	$\frac{n!}{t_a! t_d! (n - t_a - t_d)!} \left(\frac{1}{16}\right)^{t_a} \left(\frac{1}{16}\right)^{t_d} \left(\frac{9}{16}\right)^{n-t_a-t_d}$
Duplicate factor	$\frac{n!}{t_a! t_d! (n - t_a - t_d)!} 2^{t_a} [0^{t_a+t_d} - \left(\frac{1}{4}\right)^n]^2$
Multiplicative	$\frac{n!}{t_a! t_d! (n - t_a - t_d)!} \left[\frac{(b-1)^2}{8} \right]^{t_a} \left(\frac{b-2e+1}{4} \right)^{2t_d} \left(\frac{b+2e+1}{4} \right)^{2(n-t_a-t_d)}$
Optimum number	$\sigma_A^2 = 2nd^2, \sigma_D^2 = \frac{n}{4}, \sigma_{AA}^2 = \frac{n(n-1)}{2}$

DISCUSSION OF MODELS

Complementary Model

The components of genotypic variance can be written down for any particular case from the formula in table 2. To illustrate the results we present the actual components for the case of two and three loci.

For the case of two loci, we find $\sigma_A^2 = 36/256$, $\sigma_D^2 = 18/256$, $\sigma_{AA}^2 = 4/256$, $\sigma_{AD}^2 = 4/256$, $\sigma_{DD}^2 = 1/256$ so that σ_g^2 is equal to $63/256$. This conforms to the general result [see for example HORNER 1952 and HORNER, COMSTOCK and ROBINSON unpublished] that σ_g^2 equals $(3/4)^n - (3/4)^{2n}$, where n is the number of loci. In the case of two loci, the proportion of genotypic variance which is additive is $36/63 = 4/7$. The parent-offspring covariance is $19/256$ in accordance with the general formula for n loci of $(10/16)^n - (9/16)^n$ (HORNER 1952). A detailed discussion of some aspects of these formulae is given by HORNER (1952) and HORNER, COMSTOCK and ROBINSON (unpublished). The full-sib covariance is $(41/64)^n - (9/16)^n$. For two loci it is about $24/256$, so that twice the full-sib covariance over-estimates the additively genetic variance by about $\frac{1}{3}$. The parent-offspring covariance leads to an over-estimate by 6%.

For the case of three loci, we find $\sigma_A^2 = 486/4096$, $\sigma_D^2 = 243/4096$, $\sigma_{AA}^2 = 108/4096$, $\sigma_{AD}^2 = 108/4096$, $\sigma_{DD}^2 = 27/4096$, $\sigma_{AAA}^2 = 8/4096$, $\sigma_{AAD}^2 = 12/4096$, $\sigma_{ADD}^2 = 6/4096$, $\sigma_{DDD}^2 = 1/4096$ with $\sigma_g^2 = 999/4096$ in accordance with the general formula quoted above. The proportion of genotypic variance which is additive is $486/999$ or essentially 49%. The parent-offspring covariance is $271/4096$. Twice the parent-offspring covariance over-estimates the additively genetic variance by 12%. The full-sib

covariance is approximately 348/4096, so that twice the full-sib covariance overestimates the additively genetic variance by about 44%.

Duplicate Factor Model

Hence for the case of two loci, we find $\sigma_A^2 = 4/256$, $\sigma_D^2 = 2/256$, $\sigma_{AA}^2 = 4/256$, $\sigma_{AD}^2 = 4/256$, $\sigma_{DD}^2 = 1/256$, and $\sigma_G^2 = 15/256$. Twice the parent-offspring covariance is 6/256 which is 50% greater than σ_A^2 . Twice the full-sib covariance is essentially 8/256, which is twice times σ_A^2 . Thus in this case the additively genetic variance is widely overestimated by twice the parent-offspring covariance or twice the full-sib covariance.

For the case of three loci, we find $\sigma_A^2 = 6/4096$, $\sigma_D^2 = 3/4096$, $\sigma_{AA}^2 = 12/4096$, $\sigma_{AD}^2 = 12/4096$, $\sigma_{DD}^2 = 3/4096$, $\sigma_{AAA}^2 = 8/4096$, $\sigma_{AAD}^2 = 12/4096$, $\sigma_{ADD}^2 = 6/4096$, $\sigma_{DDD}^2 = 1/4096$, with $\sigma_G^2 = 63/4096$. In general, σ_G^2 for the duplicate factor model is $(1/4)^n - (1/16)^n$, the parent-offspring covariance is $(2/16)^n - (1/16)^n$, and the full-sib covariance is $(9/64)^n - (1/16)^n$. Twice the parent-offspring covariance for three loci is 14/4096 which over-estimates the additive genetic variance by 133%. Twice the full-sib covariance is essentially 21/4096 which is three and one half times the additive genetic variance.

Multiplicative Model

In this case, the results can be illustrated only by choosing values for b and e . For purposes of illustration we consider the cases considered by HORNER (1952) and

TABLE 3
Component and covariance values for the multiplicative model

Component	$b = e = 1.06$ 20 loci	$b = e = 1.1$		$b = e = 1.4$	
		2 loci	3 loci	2 loci	3 loci
<i>A</i>	4.79×10^{-2}	2.89×10^{-3}	5.01×10^{-3}	6.76×10^{-2}	1.71×10^{-1}
<i>D</i>	2.40×10^{-2}	1.44×10^{-3}	2.50×10^{-3}	3.38×10^{-2}	8.57×10^{-2}
<i>AA</i>	1.88×10^{-1}	1.56×10^{-6}	5.42×10^{-6}	4.0×10^{-4}	2.03×10^{-3}
<i>AD</i>	1.88×10^{-1}	1.56×10^{-6}	5.42×10^{-6}	4.0×10^{-4}	2.03×10^{-3}
<i>DD</i>	4.69×10^{-6}	3.91×10^{-7}	1.35×10^{-6}	1.0×10^{-4}	5.07×10^{-4}
<i>AAA</i>	4.64×10^{-7}		1.95×10^{-9}		8.00×10^{-6}
<i>AAD</i>	6.96×10^{-7}		2.93×10^{-9}		1.2×10^{-5}
<i>ADD</i>	3.48×10^{-7}		1.46×10^{-9}		6.00×10^{-6}
<i>DDD</i>	5.80×10^{-8}		2.44×10^{-10}		1.00×10^{-6}
σ_G^2	7.23×10^{-2}	4.34×10^{-3}	7.52×10^{-3}	1.02×10^{-1}	2.62×10^{-1}
$\frac{\sigma_A^2}{\sigma_G^2}$	.663	.666	.665	.661	.655
$\frac{\sigma_A^2 + \sigma_D^2}{\sigma_G^2}$	.994	.999	.998	.991	.982
Cov (Parent-offspring)	2.40×10^{-2}	1.44×10^{-3}	2.51×10^{-3}	3.39×10^{-2}	8.62×10^{-2}
Full-sib Cov.	3.00×10^{-2}	1.81×10^{-3}	3.13×10^{-3}	4.24×10^{-2}	1.08×10^{-1}
2 Cov (Parent-Offspring)/ σ_A^2	1.0025	1.0003	1.0005	1.003	1.006
2 Cov (Full-sib)/ σ_A^2	1.25	1.25	1.25	1.25	1.26

HORNER, COMSTOCK and ROBINSON (unpublished), i.e. (i) $b = e = 1.1$ (ii) $b = e = 1.4$ and an additional case (iii) $b = e = 1.06$. This latter value was selected by supposing that the range in bushels per acre for a certain variety of corn was 96 to 30 and that twenty pairs of multiplicative genes with complete dominance were responsible for this range. The ratio of the largest to the smallest genotypic value would be then $b^{20} = 3.2$ which yields a value of $b = 1.06$.

Component and covariance values, along with certain ratios are set out in table 3.

The genotypic variance in general is $\left(\frac{b^2 + 2e^2 + 1}{4}\right)^n - \left(\frac{b + 2e + 1}{4}\right)^{2n}$, the parent-offspring covariance (complete dominance) is $\left(\frac{10b^2 + 4b + 2}{16}\right)^n - \left(\frac{3b + 1}{4}\right)^{2n}$ and the full-sib covariance (complete dominance) is $\left(\frac{41b^2 + 14b + 9}{64}\right)^n - \left(\frac{3b + 1}{4}\right)^{2n}$.

An examination of table 3 reveals that the epistatic portion of the genotypic variance is rather small. Indeed the additive portion of this variance amounts to about 66% and the additive and dominance portions together account for 99% in the cases examined. The over-estimation of the additive genetic variance by twice the parent-offspring covariance is trivial. The over-estimation in the case of twice the full-sib covariance is 25%, but this is almost entirely due to dominance. The smallness of the epistatic portion of the genotypic variance in the case of the multiplicative model raises serious questions concerning the usefulness of logarithmic scaling in removing epistatic effects. It does not appear that the use of a logarithmic scale instead of an arithmetic scale when there are multiplicative effects of the order here considered will affect appreciably the estimate of heritability in a population, when gene frequency is one half.

Optimum number model

The components for the optimum number model are $\sigma_A^2 = 2nd^2$, $\sigma_D^2 = \frac{n}{4}$, and $\sigma_{AA}^2 = \frac{n(n-1)}{2}$, all others being zero. Twice the parent-offspring covariance over-estimates the additive genetic variance by the fraction $\frac{n-1}{8d^2}$ which is to be contrasted to $\frac{2n-1}{16d^2}$ for twice the full-sib covariance. These fractions will be large if d is close to zero.

GENERAL DISCUSSION

Knowledge of the relative magnitudes of the components of genotypic variance, when epistasis is present, is useful in ascertaining the importance of epistasis in the hereditary mechanism, in computing the bias in estimation and prediction procedures which ignore the effects of epistasis and in determining optimum breeding and selection procedures when epistasis is present. The consequences of the simple epistatic models of this paper might be considered as consequences which might be approached as a limit by the more complex models of nature. An important model omitted

from this paper is the threshold model. With this model an effect is produced if the number of plus genes of a genotype exceeds some threshold number, otherwise it is not produced. The threshold model, although it can be expressed mathematically, appears to be more mathematically intractable than the models considered herein. The assumption of a gene frequency of one-half at every locus makes the results applicable strictly only to an F_2 population. The generalization to an arbitrary gene frequency at any locus has been done and is a step toward determining the consequences of selection when epistasis is present.

SUMMARY

KEMPTHORNE (1954, 1955) solved the general problem of the partition of genotypic variance and the correlation between non-bred relatives in random mating populations. These results are applied in the present paper to the case of random mating populations for which the genotypic value of a genotype is symmetrical with respect to the components, one from each locus, of the genotype, and for which gene frequency at each locus is one-half. The following classical gene models were given particular attention and the results discussed for each: complementary, duplicate factor, multiplicative, and optimum number.

LITERATURE CITED

- KEMPTHORNE, O., 1954 The correlation between relatives in a random mating population. *Proc. Roy. Soc. London B* **143**: 103-113.
- KEMPTHORNE, O., 1955 The theoretical values of correlations between relatives in random mating populations. *Genetics* **40**: 153-167.
- WRIGHT, S., 1935 The analysis of variance and the correlations between relatives with respect to deviations from an optimum. *J. Genet.* **30**: 243-256.
- HORNER, T. W., 1952 Non-allelic gene interaction and the interpretation of quantitative genetic data. Ph.D. Thesis. Department of Exp. Stat., N. C. State College.