

THE GAMETIC DISTRIBUTION IN MENDELIAN HEREDITY

By B. I. HAYMAN*

[*Manuscript received August 14, 1961*]

Summary

The expectation of a gametic frequency is expressed in terms of Jones' (1961) linkage probability. The derivation illustrates a use of set notation with wide applicability in quantitative genetics.

Schnell's (1961) alternative linkage parameter is related to Jones' and the relative merits of the two parameters discussed.

1. Introduction

The theory of quantitative inheritance has developed in three stages and by two processes. One process has been the construction of mathematical representations of gene action and deterministic gene transmission in simple genetical models. The other process has been the development of mathematical and statistical techniques of manipulating these representations in the generality required for the multifactorial systems controlling quantitative characters.

The first stage of development was Fisher's (1918) generalization by quantification of the concept of dominance of major genes. At the same time Fisher showed statistically that his quantitative model of the action of an allelic pair of genes could be applied simultaneously to a large number of gene pairs so as to separate the combined additive action of these genes from their combined dominance action. Mather (1949) extended the application of this model and showed how these two components of genetic variation could be estimated.

Fisher (1918) also gave a model of epistatic action between two pairs of genes but did not generalize it to cover epistasis between any number of genes. Consequently the theory of quantitative inheritance remained restricted for some time to that of additive and dominance action. The second stage, of a general theory of epistasis, had to wait for work by Anderson and Kempthorne (1954), Cockerham (1954), Hayman (1954), and Hayman and Mather (1955).

As soon as quantitative theory involved variances and higher-order statistics and epistatic gene action, linkage parameters had to be introduced. In the simpler models the parameters describing crossing-over between pairs of major genes sufficed, but in the more complex models linkage was generally assumed to be absent, though some tests for linkage were devised (Mather 1949). Geiringer (1944) proposed a set of linkage parameters for more than two genes but her system had disadvantages.

The third stage of development of the theory of quantitative inheritance was not reached until Jones (1961) devised a set of independent linkage parameters and gave some examples of its use in lower-order statistics with dual epistacy. This

* Applied Mathematics Laboratory, D.S.I.R., Christchurch, N.Z.

paper further develops this third stage by presenting a method for handling these linkage parameters in full generality in diploid inheritance. The method was originally devised by Hayman (1954) to handle epistasis of all orders, but in the absence of linkage. More recently Schnell (1961) has constructed an alternative set of linkage parameters and also uses this method.

The particular problem solved here is the expression of expected gametic frequencies in terms of linkage probabilities. Representation of gene transmission and gene action will be discussed in later publications.

Although this paper is intended to be a basis for further studies in quantitative genetics, the results are equally applicable to qualitative genetics. This is because it is concerned solely with the theory of gene transmission and not of gene action. The group of genes whose linkage relations are to be considered may jointly control some quantitative character or they may control distinct qualitative characters.

2. Set Notation

2.1 Parametric systems in quantitative genetics have a hierarchical structure. For example, most models of gene action contain parameters which measure the additive and dominance actions of each gene pair supplemented by parameters measuring the residual actions of the genes taken two pairs at a time, then three pairs at a time and so on. Linkage parameters were originally confined to measures of crossing-over between pairs of loci and then Jones (1961) added further parameters involving four loci, six loci, and so on.

A set of k loci may be numbered $1, 2, \dots, k$, a general locus being specified by i with $1 \leq i \leq k$. Parameters which are functions of one, two, three, etc. loci have the forms $f(i)$, $f(ij)$, $f(ijl)$, etc. respectively. A linear function of such parameters over k loci may take the form

$$F(1, 2, \dots, k) = f + \sum f(i) + \sum f(ij) + \sum f(ijl) + \dots,$$

where summation is over all values of the variables from 1 to k with the restriction that no two variables in the same parameter are equal in value.

A more concise notation is based on considering the numbers $1, 2, \dots, k$ as a set κ and any selection from these numbers as a subset $\alpha \subset \kappa$. Then one definition $f(\alpha)$ suffices for the parameters and the above linear relation becomes

$$F(\kappa) = \sum_{\alpha \subset \kappa} f(\alpha),$$

summation being over all subsets in κ including the empty set and κ itself. The twofold advantages of this set notation are the brevity it permits and the absence of any need for the awkward restriction that the values of the variables be unequal.

The combination of sets α and β follows the rules:

$$\alpha + \beta = \text{set containing all elements in } \alpha \text{ or } \beta.$$

$$\alpha - \beta = \text{set of elements in } \alpha \text{ which do not occur in } \beta.$$

$$\alpha\beta = \text{set of elements in both } \alpha \text{ and } \beta.$$

An important function is $n(\alpha)$, the number of elements in, or the *size* of, the set α . The same number is used for the *order* of the parameter $f(\alpha)$.

2.2 The binomial theorem provides a good example of the brevity of this notation

$$(1+x)^k = \sum_{a=0}^k \binom{k}{a} x^a,$$

becomes

$$(1+x)^{n(\kappa)} = \sum_{a \subset \kappa} x^{n(a)}.$$

The special case when $x = -1$ provides an analogy to the delta function.

$$\begin{aligned} d(\kappa) &= \sum_{a \subset \kappa} (-1)^{n(a)} \\ &= 1 && \text{if } \kappa = 0 \\ &= 0 && \text{if } \kappa \neq 0. \end{aligned}$$

2.3 An intermediate notation is the average of $f(\alpha)$ of equal order over the set κ of loci. Define

$$f(a; \kappa, k) = \sum'_{a \subset \kappa} f(\alpha) / \binom{k}{a},$$

where

$$n(\alpha) = a \leq k = n(\kappa),$$

and the prime indicates summation over sets of constant size. Then

$$F(\kappa) = \sum_{a=0}^k \binom{k}{a} f(a; \kappa, k).$$

In the expansion of the binomial theorem

$$f(a; \kappa, k) = x^a,$$

and so reverts to the usual form.

There are two advantages of these average functions. Firstly, the coefficients in relations involving average linkage functions are often independent of the total set κ of loci. This means that $f(a; \kappa, k)$ may be abbreviated to $f(a)$ in such relations and then manipulated regardless of κ . Secondly, and as a consequence, the facility of generating functions becomes available.

$$G_f(x) = \sum_{a=0}^{\infty} f(a) x^a / a!$$

can be constructed, at least formally, if κ is ignored.

3. Genetical Parameters

Genes in a diploid individual form two homologous groups, one derived from each parent. Gametes produced by an individual contain a haploid group of genes, each gene being a copy of one of a pair of homologues. The chance that the i th and j th genes in a gamete are copies of genes in the same parental group of genes is q_{ij} . The chance that they are copies of genes in different parental groups, or the chance of crossing-over in meiosis, is $p_{ij} = 1 - q_{ij}$. Following Jones (1961), but using set notation for generality, the linkage probability $q(\alpha)$ is the chance of no crossing-over in the set α of loci, or the chance that a gamete contains copies of genes in the same parental group at loci in the set α . The order of $q(\alpha)$ is $n(\alpha)$. The phenomenon of affinity (Wallace 1959) is assumed to be negligible or absent.

Suppose that a gamete contains genes at loci α originating from one parent and at loci β originating from the other parent. Define $l(\alpha, \beta)$ to be the expected frequency of this gamete amongst all gametes which differ at loci $\alpha + \beta$. If the parents are ordered so as to distinguish homologous gametes then

$$l(\alpha, \beta) = l(\beta, \alpha).$$

4. A Relation between Gametic and Linkage Probabilities

4.1 Theorem

An expression for the gametic probability in terms of linkage probabilities is

$$l(\alpha, \beta) = \frac{1}{2} \sum_{\rho \subset \beta} (-1)^{n(\rho)} q(\alpha + \rho).$$

Proof

Let

$$n(\alpha) = a.$$

Then

$$\begin{aligned} l(\alpha, \beta) &= \text{probability that } a \text{ genes arise from one particular parent} \\ &= \frac{1}{2} \{ \text{prob. at least } a \text{ from one parent} \\ &\quad - \text{prob. at least } a+1 \text{ from one parent} \\ &\quad + \text{prob. at least } a+2 \text{ from one parent} \\ &\quad - \dots \} \\ &= \frac{1}{2} \{ q(\alpha) - \sum_{i \in \beta} q(\alpha + i) + \sum_{ij \in \beta} q(\alpha + ij) - \dots \} \\ &= \frac{1}{2} \sum_{\rho \subset \beta} (-1)^{n(\rho)} q(\alpha + \rho). \end{aligned}$$

4.2 In the formal expansion of $l(0, \alpha)$ and $l(i, \alpha)$ terms such as $q(0)$ and $q(i)$ appear. These terms may be given conventional values to preserve the generality of the theorem.

Since

$$\begin{aligned} l(i, 0) &= l(0, i) = \frac{1}{2}, \\ q(i) &= 1, \\ q(0) &= 2. \end{aligned}$$

5. Independent Linkage Parameters

The probabilities $q(\alpha)$ are not independent and this section concerns their interdependence.

5.1 Theorem 4.1 provides an expansion for $q(\alpha)$.

$$\begin{aligned} q(\alpha) &= 2l(\alpha, 0) \\ &= 2l(0, \alpha) && \text{by 3} \\ &= \sum_{\rho \subset \alpha} (-1)^{n(\rho)} q(\rho). \end{aligned}$$

5.2 5.11 may be written in terms of average linkage probabilities. As in 2.3, define

$$5.21 \quad q(a; \kappa, k) = \sum'_{\alpha \subset \kappa} q(\alpha) / \binom{k}{a},$$

where

$$n(\alpha) = a \leq k = n(\kappa).$$

Special cases of this function are

$$\begin{aligned} q(0; \kappa, k) &= q(0) = 2, \\ q(1; \kappa, k) &= 1, \\ q(k; \kappa, k) &= q(\kappa). \end{aligned}$$

Hence

$$q(a; \kappa, k) = \sum'_{a \subset \kappa} \sum_{\rho \subset a} (-1)^{n(\rho)} q(\rho) / \binom{k}{a} \quad \text{by 5.11}$$

$$= \sum_{r=0}^a (-1)^r \binom{a}{r} \sum'_{\rho \subset \kappa} q(\rho) / \binom{k}{r} \quad \text{by symmetry}$$

$$5.22 \quad = \sum_{r=0}^a (-1)^r \binom{a}{r} q(r; \kappa, k) \quad \text{by 5.21}$$

5.3 Theorem

Linkage probabilities of odd order may be expressed in terms of probabilities of lower even order according to

$$q(\alpha) = \sum_{\rho^* \subset \alpha} u[n(\alpha - \rho^*)] q(\rho^*),$$

where the asterisk indicates a set of even size and where $u(r)$ is the coefficient of $x^r/r!$ in the expansion of $\tanh \frac{1}{2}x$ in a power series.

Proof

The coefficients in 5.22 are independent of the set κ of loci so that this relation may be written temporarily

$$5.31 \quad q(a) = \sum_{r=0}^a (-1)^r \binom{a}{r} q(r).$$

Define

$$G_q^*(x) = \sum_{r=0}^{\infty} q(2r) x^{2r} / (2r)!,$$

and

$$G_q(x) = \sum_{r=0}^{\infty} q(2r+1) x^{2r+1} / (2r+1)!$$

Then

$$\begin{aligned} G_q^*(x) + G_q(x) &= \sum_{r=0}^{\infty} q(r) x^r / r! \\ &= \sum_{r=0}^{\infty} \sum_{s=0}^r (-1)^s q(s) x^r / s! (r-s)! \quad \text{by 5.31} \end{aligned}$$

$$= \sum_{s=0}^{\infty} \sum_{r=s}^{\infty} (-1)^s q(s) x^r / s! (r-s)!$$

Put $r = s + t$

$$= \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} (-1)^s q(s) x^{s+t} / s! t!$$

$$= e^x [G_q^*(x) - G_q(x)].$$

Hence

$$5.32 \quad G_q(x) = (\tanh \tfrac{1}{2}x)G_q^*(x).$$

Equate coefficients of odd powers of x

$$q(2a+1) = \sum_{r=0}^a \binom{2a+1}{2r} u(2a-2r+1)q(2r),$$

where

$$\sum_{r=0}^{\infty} u(r)x^r/r! = \tanh \tfrac{1}{2}x.$$

Restore the full notation

$$q(2a+1; \kappa, k) = \sum_{r=0}^a \binom{2a+1}{2r} u(2a-2r+1)q(2r; \kappa, k).$$

In particular, if $n(\kappa) = 2a+1$ and κ is replaced by α ,

$$\begin{aligned} q(\alpha) &= \sum_{r=0}^a \binom{2a+1}{2r} u(2a-2r+1)q(2r; \alpha, 2a+1) \\ &= \sum_{\rho^* \subset \alpha} u[n(\alpha - \rho^*)]q(\rho^*). \end{aligned}$$

5.4 Theorem 5.3 holds for both odd and even α if $u(r)$ is redefined to be the coefficient of $x^r/r!$ in

$$5.41 \quad G_u(x) = 1 + \tanh \tfrac{1}{2}x.$$

Thus

$$u(0) = 1$$

and

$$u(2r) = 0 \quad \text{for } r > 0$$

so that when α is even the formula reduces to

$$q(\alpha^*) = q(\alpha^*).$$

Hence for α of any size we have

$$5.42 \quad q(\alpha) = \sum_{\rho^* \subset \alpha} u[n(\alpha - \rho^*)]q(\rho^*),$$

where $u(r)$ is defined by 5.41. This form of theorem 5.3 is better for reducing functions of general sets to functions of even sets. Table 1 contains some values of $u(r)$.

5.5 Theorem

Probabilities of even order form a complete set of independent linkage parameters.

Proof

By theorem 5.3 odd-ordered linkage probabilities are redundant parameters. The total number of even-ordered parameters for k loci is

$$\binom{k}{2} + \binom{k}{4} + \binom{k}{6} + \dots = 2^{k-1} - 1.$$

There are 2^k possible gametes, homologous gametes have equal probabilities, and the sum of all probabilities is unity so that the number of independent parameters required is $2^k - 1$. Evidently the linkage probabilities of even order provide a complete set of independent parameters.

Jones (1961) obtained the same result by enumeration and without introducing linkage probabilities of odd order. However, they are useful here in providing intermediate steps toward the final expression for $l(\alpha, \beta)$.

TABLE 1
COEFFICIENT $u(r)$ OF $x^r/r!$ IN THE EXPANSION OF $1 + \tanh \frac{1}{2}x$

r	$u(r)$	r	$u(r)$
0	1	13	5461/2
1	1/2	15	-929,569/16
3	-1/4	17	3,202,291/2
5	1/2	19	-221,930,581/4
7	-17/8	21	4,722,116,521/2
9	31/2	23	-968,383,680,827/8
11	-691/4	$2r > 0$	0

6. The Relation between Gametic and Linkage Probabilities

6.1 Define an average of $q(\alpha)$ over two sets of loci

$$6.11 \quad q(a; \kappa, k; b; \lambda, l) = \sum_{a \subset \kappa} \sum_{\beta \subset \lambda} q(\alpha + \beta) / \binom{k}{a} \binom{l}{b},$$

where

$$\begin{aligned} \kappa \lambda &= 0, \\ n(\alpha) &= a \leq k = n(\kappa), \\ n(\beta) &= b \leq l = n(\lambda). \end{aligned}$$

Note that

$$q(a; \kappa, k; 0; 0, 0) = q(a; \kappa, k),$$

and

$$q(k; \kappa, k; l; \lambda, l) = q(\kappa + \lambda).$$

6.2 A further generalization of 5.11, in a similar manner to 5.22, is

$$q(a; \kappa, k; b; \lambda, l) = \sum_{r=0}^a \sum_{s=0}^b (-1)^{r+s} \binom{a}{r} \binom{b}{s} q(r; \kappa, k; s; \lambda, l),$$

and $q(a; \kappa, k; b; \lambda, l)$ may be abbreviated to $q(a, b)$ in this relation.

$$6.3 \quad \text{Define} \quad G_q^*(x, y) = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} q(r, s) x^r y^s / r! s! \quad \text{for even } r+s$$

$$G_q(x, y) = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} q(r, s) x^r y^s / r! s! \quad \text{for odd } r+s$$

Then, as in 5.32,

$$6.31 \quad G_q(x, y) = \{\tanh \tfrac{1}{2}(x+y)\} G_q^*(x, y).$$

Equate coefficients of $x^a y^b$ with odd $a+b$ and restore the full notation

$$q(a; \kappa, k; b; \lambda, l) = \sum_{\substack{r=0 \\ \text{odd } a+b}}^a \sum_{\substack{s=0 \\ \text{even } r+s}}^b \binom{a}{r} \binom{b}{s} u(a+b-r-s) q(r; \kappa, k; s; \lambda, l).$$

6.4 Theorem

The gametic probability, expressed in terms of independent linkage probabilities, is

$$l(\alpha, \beta) = \frac{1}{2} \sum_{\rho^* \subset \alpha + \beta} (-1)^{n(\beta \rho^*)} u[n(\alpha - \rho^*), n(\beta - \rho^*)] q(\rho^*),$$

where $u(r, s)$ is the coefficient of $x^r y^s / r! s!$ in the expansion of $2/(e^{-x} + e^{-y})$ in a power series.

Proof

Define the average gametic probability

$$l(a; \kappa, k; b; \lambda, l) = \sum_{\alpha \subset \kappa} \sum_{\beta \subset \lambda} l(\alpha, \beta) \binom{k}{\alpha} \binom{l}{\beta}$$

with the conventions of 6.11

$$\begin{aligned} &= \frac{1}{2} \sum_{\alpha \subset \kappa} \sum_{\beta \subset \lambda} \sum_{\rho \subset \beta} (-1)^{n(\rho)} q(\alpha + \rho) \binom{k}{\alpha} \binom{l}{\beta} && \text{by theorem 4.1} \\ &= \frac{1}{2} \sum_{r=0}^b (-1)^r \binom{b}{r} \sum_{\alpha \subset \kappa} \sum_{\rho \subset \lambda} q(\alpha + \rho) \binom{k}{\alpha} \binom{l}{\rho} \\ &= \frac{1}{2} \sum_{r=0}^b (-1)^r \binom{b}{r} q(a; \kappa, k; r; \lambda, l). \end{aligned}$$

Once again abbreviated notation may be used

$$l(a, b) = \frac{1}{2} \sum_{r=0}^b (-1)^r \binom{b}{r} q(a, r).$$

Hence

$$\begin{aligned} G_l(x, y) &= \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} l(r, s) x^r y^s / r! s! \\ &= \frac{1}{2} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \sum_{t=0}^s (-1)^t q(r, t) x^r y^s / r! (s-t)! t! \\ &= \frac{1}{2} \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \sum_{s=t}^{\infty} (-1)^t q(r, t) x^r y^s / r! (s-t)! t! \\ &= \frac{1}{2} \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \sum_{u=0}^{\infty} (-1)^t q(r, t) x^r y^{t+u} / r! t! u! \end{aligned}$$

Put $s = t + u$

$$\begin{aligned}
&= \frac{1}{2}e^y \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} q(r, t)x^r(-y)^t/r!t! \\
&= \frac{1}{2}e^y[G_q^*(x, -y) + G_q(x, -y)] && \text{by 6.3} \\
&= \frac{1}{2}e^y[1 + \tanh \frac{1}{2}(x-y)]G_q^*(x, -y) && \text{by 6.31} \\
6.41 \quad &= G_q^*(x, -y)/(e^{-x} + e^{-y}).
\end{aligned}$$

Equate coefficients of $x^a y^b$

$$l(a, b) = \frac{1}{2} \sum_{\substack{r=0 \\ \text{even}}}^a \sum_{s=0}^b (-1)^s \binom{a}{r} \binom{b}{s} u(a-r, b-s) q(r, s),$$

where

$$\begin{aligned}
\sum_{r=0}^{\infty} \sum_{s=0}^{\infty} u(r, s)x^r y^s / r!s! &= 2/(e^{-x} + e^{-y}) \\
&= G_u(x, y).
\end{aligned}$$

Restore the full notation, make $n(\kappa) = a$, $n(\lambda) = b$, and replace κ and λ by α and β :

$$l(\alpha, \beta) = \frac{1}{2} \sum_{\rho^* \subset \alpha + \beta} (-1)^{n(\beta\rho^*)} u[n(\alpha - \rho^*), n(\beta - \rho^*)] q(\rho^*).$$

6.5 Table 2 contains some values of $u(r, s)$. The construction of this table is discussed in Appendix 8.2.

Properties of the $u(r, s)$ are obtained by equating coefficients of $x^r y^s$ in various identities. Since

$$G_u(x, y) = G_u(y, x),$$

$$6.51 \quad u(r, s) = u(s, r).$$

Boundary values of $u(r, s)$ are given by

$$G_u(x, 0) = G_u(x),$$

so that

$$6.52 \quad u(r, 0) = u(r).$$

Similarly,

$$u(0, r) = u(r).$$

6.6 Theorem 5.3 is the special case of theorem 6.4 when $\beta = 0$ and the coefficient is transformed by 6.52.

$$6.7 \quad 2l(\alpha, \beta) = \sum_{\rho^* \subset \alpha + \beta} (-1)^{n(\beta\rho^*)} u[n(\alpha - \rho^*), n(\beta - \rho^*)] q(\rho^*)$$

$$\text{by 6.51} \quad = \sum_{\rho^* \subset \alpha + \beta} (-1)^{n(\rho^* - \beta\rho^*)} u[n(\beta - \rho^*), n(\alpha - \rho^*)] q(\rho^*)$$

$$= \sum_{\rho^* \subset \alpha + \beta} (-1)^{n(\alpha\rho^*)} u[n(\beta - \rho^*), n(\alpha - \rho^*)] q(\rho^*)$$

$$= 2l(\beta, \alpha)$$

as expected.

TABLE 2
COEFFICIENT $u(r, s)$ OF $x^r y^s / r! s!$ IN THE EXPANSION OF $2/(e^{-x} + e^{-y})$

6.8 Particular cases of theorem 6.4 may be written down as follows. For given α and β and given

$$\begin{aligned} r &\leq n(\alpha) = a, \\ s &\leq n(\beta) = b, \end{aligned}$$

define the *total* linkage probability

$$\begin{aligned} q_{r,s} &= \sum'_{\rho \subset \alpha} \sum'_{\sigma \subset \beta} q(\rho + \sigma) && \text{with } n(\rho) = r, n(\sigma) = s \\ &= \binom{a}{r} \binom{b}{s} q(r; \alpha, a; s; \beta, b). \end{aligned}$$

Then

$$l(\alpha, \beta) = \frac{1}{2} \sum_{\substack{r=0 \\ \text{even } r+s}}^a \sum_{s=0}^b (-1)^s u(a-r, b-s) q_{r,s}.$$

In the following examples the sets of genes are specified by the numbers of the genes, $i, j, k, \dots$ and in probabilities of lower order $q_{ij}, q_{ijk}, \dots$ etc. replace the longer $q(ij), q(ijkm), \dots$ etc. To illustrate the expansion take

$$\begin{aligned} l(ijk, mn) &= \frac{1}{2} \sum_{\substack{r=0 \\ \text{even } r+s}}^3 \sum_{s=0}^2 (-1)^s u(3-r, 2-s) q_{r,s} \\ &= \frac{1}{2} [u(3, 2)q_{0,0} + u(3, 0)q_{0,2} - u(2, 1)q_{1,1} \\ &\quad + u(1, 2)q_{2,0} + u(1, 0)q_{2,2} - u(0, 1)q_{3,1}] \\ &= \frac{1}{8}q(0) - \frac{1}{8}q_{mn} - \frac{1}{8}(q_{im} + q_{in} + q_{jm} + q_{jn} + q_{km} + q_{kn}) \\ &\quad + \frac{1}{8}(q_{ij} + q_{ik} + q_{jk}) + \frac{1}{4}(q_{ijmn} + q_{jkmn} + q_{ikmn}) \\ &\quad - \frac{1}{4}(q_{ijkm} + q_{ijkn}) && \text{by Table 2} \\ &= \frac{1}{4} + \frac{1}{8}(q_{ij} + q_{ik} + q_{jk} - q_{im} - q_{in} - q_{jm} - q_{jn} - q_{km} - q_{kn} - q_{mn}) \\ &\quad + \frac{1}{4}(q_{ijmn} + q_{jkmn} + q_{ikmn} - q_{ijkm} - q_{ijkn}). \end{aligned}$$

The expectations of gametic frequency types involving up to seven genes follow:

Two genes:

$$\begin{aligned} l(ij, -) &= \frac{1}{2}q_{2,0} = \frac{1}{2}q_{ij}, \\ l(i, j) &= \frac{1}{2} - \frac{1}{2}q_{1,1} = \frac{1}{2} - \frac{1}{2}q_{ij}. \end{aligned}$$

Three genes:

$$\begin{aligned} l(ijk, -) &= -\frac{1}{4} + \frac{1}{4}q_{2,0} = -\frac{1}{4} + \frac{1}{4}(q_{ij} + q_{ik} + q_{jk}), \\ l(ij, k) &= \frac{1}{4} - \frac{1}{4}q_{1,1} + \frac{1}{4}q_{2,0} = \frac{1}{4} - \frac{1}{4}(q_{ik} + q_{jk} - q_{ij}). \end{aligned}$$

Four genes:

$$\begin{aligned} l(ijkm, -) &= \frac{1}{2}q_{4,0} = \frac{1}{2}q_{ijk}, \\ l(ijk, m) &= -\frac{1}{4} + \frac{1}{4}q_{2,0} - \frac{1}{4}q_{3,1} \\ &= -\frac{1}{4} + \frac{1}{4}(q_{ij} + q_{ik} + q_{jk}) - \frac{1}{2}q_{ijk}, \\ l(ij, km) &= \frac{1}{2} - \frac{1}{4}q_{1,1} + \frac{1}{2}q_{2,2} \\ &= \frac{1}{2} - \frac{1}{4}(q_{ik} + q_{im} + q_{jk} + q_{jm}) + \frac{1}{2}q_{ijk}. \end{aligned}$$

Five genes:

$$\begin{aligned} l(ijkmn, -) &= \frac{1}{2} - \frac{1}{8}q_{2,0} + \frac{1}{4}q_{4,0}, \\ l(ijkm, n) &= -\frac{1}{2} + \frac{1}{8}q_{1,1} + \frac{1}{8}q_{2,0} - \frac{1}{4}q_{3,1} + \frac{1}{4}q_{4,0}, \\ l(ijk, mn) &= \frac{1}{4} - \frac{1}{8}q_{0,2} - \frac{1}{8}q_{1,1} + \frac{1}{8}q_{2,0} + \frac{1}{4}q_{2,2} - \frac{1}{4}q_{3,1}. \end{aligned}$$

Six genes:

$$\begin{aligned} l(ijkmnp, -) &= \frac{1}{2}q_{6,0}, \\ l(ijkmn, p) &= \frac{1}{2} - \frac{1}{8}q_{2,0} + \frac{1}{4}q_{4,0} - \frac{1}{2}q_{5,1}, \\ l(ijkm, np) &= -1 + \frac{1}{8}q_{1,1} + \frac{1}{4}q_{2,0} + \frac{1}{2}q_{4,2}, \\ l(ijk, mnp) &= 5/4 - \frac{1}{8}q_{0,2} - \frac{1}{4}q_{1,1} - \frac{1}{8}q_{2,0} + \frac{1}{4}q_{2,2} - \frac{1}{2}q_{3,3}. \end{aligned}$$

Seven genes:

$$\begin{aligned} l(ijkmnp, -) &= -17/8 + \frac{1}{4}q_{2,0} - \frac{1}{8}q_{4,0} + \frac{1}{4}q_{6,0}, \\ l(ijkmnp, r) &= 17/8 - \frac{1}{4}q_{1,1} - \frac{1}{4}q_{2,0} + \frac{1}{8}q_{3,1} + \frac{1}{8}q_{4,0} - \frac{1}{4}q_{5,1} + \frac{1}{4}q_{6,0}, \\ l(ijkmn, pr) &= -13/8 + \frac{1}{4}q_{0,2} + \frac{1}{4}q_{1,1} + \frac{1}{8}q_{2,0} - \frac{1}{8}q_{2,2} - \frac{1}{8}q_{3,1} + \frac{1}{8}q_{4,0} \\ &\quad + \frac{1}{4}q_{4,2} - \frac{1}{4}q_{5,1}, \\ l(ijkm, npr) &= \frac{5}{8} - \frac{1}{4}q_{0,2} - \frac{1}{8}q_{1,1} + \frac{1}{8}q_{1,3} + \frac{1}{8}q_{2,0} + \frac{1}{8}q_{2,2} - \frac{1}{8}q_{3,1} - \frac{1}{4}q_{3,3} \\ &\quad - \frac{1}{8}q_{4,0} + \frac{1}{4}q_{4,2}. \end{aligned}$$

It should be remembered that in these expectations $q_{r,s}$ is an abbreviation from which two reference sets have been omitted and that these reference sets are different in each expectation. In each expectation $q_{r,s}$ is the sum of $q(\alpha)$ of order $r+s$ over loci taken r at a time from the first set and s at a time from the second set specified on the left of the equation. Jones (1961) has given the expectations involving up to four genes.

7. Schnell's Linkage Parameter

7.1 Schnell (1961) defines a linkage parameter (in our notation)

$$r(\alpha) = 2 \sum_{\rho \subset \kappa} (-1)^{n(\alpha\rho)} l(\rho, \kappa - \rho). \quad \text{for even } \alpha$$

Consider the more general parameter

$$r_c(\alpha) = 2 \sum_{\rho \subset \kappa} c^{n(\alpha\rho)} l(\rho, \kappa - \rho),$$

where c is a numerical constant and α is any subset of κ . Then, by theorem 4.1,

$$r_c(\alpha) = \sum_{\rho \subset \kappa} c^{n(\alpha\rho)} \sum_{\sigma \subset \kappa - \rho} (-1)^{n(\sigma)} q(\rho + \sigma)$$

Put $\rho + \sigma = \beta$

$$= \sum_{\beta \subset \kappa} \sum_{\rho \subset \beta} c^{n(\alpha\rho)} (-1)^{n(\beta) - n(\rho)} q(\beta).$$

Put

$$\rho = \gamma + \delta,$$

where

$$\gamma = \alpha\rho,$$

and

$$\delta = \rho - \alpha.$$

Then

$$\begin{aligned}
 r_c(\alpha) &= \sum_{\beta \subset \kappa} \sum_{\gamma \subset \alpha\beta} \sum_{\delta \subset \beta - \alpha} c^{n(\gamma)} (-1)^{n(\beta) - n(\gamma) - n(\delta)} q(\beta) \\
 &= \sum_{\beta \subset \kappa} \sum_{\gamma \subset \alpha\beta} (-c)^{n(\gamma)} \sum_{\delta \subset \beta - \alpha} (-1)^{n(\delta)} (-1)^{n(\beta)} q(\beta) \\
 &\text{by 2.2 if } c \neq 1, \quad = \sum_{\beta \subset \kappa} (1-c)^{n(\alpha\beta)} d(\beta - \alpha) (-1)^{n(\beta)} q(\beta) \\
 7.11 \quad &= \sum_{\beta \subset \alpha} (c-1)^{n(\beta)} q(\beta).
 \end{aligned}$$

This may be regarded as a generalization of 5.11.

The case of $c = 1$ provides a check on the working:

$$\begin{aligned}
 r_1(\alpha) &= \sum_{\beta \subset \alpha} d(\beta) q(\beta) \\
 &= q(0) \\
 &= 2.
 \end{aligned}$$

by 4.2

Hence

$$\sum_{\rho \subset \kappa} l(\rho, \kappa - \rho) = 1.$$

In other words, the total frequency of gametes arising from a set κ of loci is unity.

7.2 If $r_c(a)$ is the average of $r_c(\alpha)$ over subsets of size a of a set of loci then, from 7.11 and as in 5.31,

$$7.21 \quad r_c(a) = \sum_{r=0}^a (c-1)^r \binom{a}{r} q(r).$$

Hence, as in theorem 5.3,

$$\begin{aligned}
 G_r(x) &= \sum_{s=0}^{\infty} r_c(s) x^s / s! \\
 7.22 \quad &= e^x [G_q^*((c-1)x) + G_q((c-1)x)] \\
 &= e^x [1 + \tanh(\tfrac{1}{2}(c-1)x)] G_q^*((c-1)x) \\
 &= e^{\frac{1}{2}(c+1)x} \{\operatorname{sech}(\tfrac{1}{2}(c-1)x)\} G_q^*((c-1)x).
 \end{aligned}$$

by 5.32

In Schnell's case of $c = -1$, $G_r(x)$ is an even function and it is not even for any other value of c . Take $c = -1$ henceforth. Then

$$7.23 \quad G_r(x) = (\operatorname{sech} x) G_q^*(2x).$$

This means that $r_{-1}(\alpha)$ is identically zero for odd α and Schnell's definition of $r(\alpha)$ may be extended to sets of any size so that

$$\begin{aligned}
 r(\alpha) &= r_{-1}(\alpha) \\
 &= 2 \sum_{\rho \subset \alpha} (-1)^{n(\alpha\rho)} l(\rho, \kappa - \rho) \\
 7.24 \quad &= \sum_{\beta \subset \alpha} (-2)^{n(\beta)} q(\beta).
 \end{aligned}$$

by 7.1

for all α .

Equate even powers of x in 7.23 and restore the full notation

$$r(2a; \kappa, k) = \sum_{s=0}^a 2^{2s} \binom{2a}{2s} v(2a-2s) q(2s; \kappa, k),$$

where

$$\begin{aligned} \sum_{r=0}^{\infty} v(r) x^r / r! &= \operatorname{sech} x \\ &= G_v(x). \end{aligned}$$

In particular, if $n(\kappa) = 2a$ and κ is replaced by α ,

$$7.25 \quad r(\alpha) = \sum_{\rho^* \subset \alpha} 2^{n(\rho^*)} v[n(\alpha - \rho^*)] q(\rho^*),$$

and once again this actually holds for all sizes of α since $v(2r+1)$ is zero. Table 3 contains some values of $v(r)$.

TABLE 3
COEFFICIENT $v(r)$ OF $x^r/r!$ IN THE EXPANSION OF $\operatorname{sech} x$

r	$v(r)$	r	$v(r)$
0	1	12	2,702,765
2	-1	14	-199,360,981
4	5	16	19,391,512,145
6	-61	18	-2,404,879,675,441
8	1385	$2r+1$	0
10	-50,521		

The inverse relation is obtained from 7.22 with $c = -1$.

$$G_q^*(2x) - G_q(2x) = e^{-x} G_r(x),$$

from which it follows that

$$7.26 \quad q(\alpha) = 2^{-n(\alpha)} \sum_{\rho \subset \alpha} r(\rho),$$

The relations, 7.25 and 7.26, between $q(\alpha)$ and $r(\alpha)$ show that these two parameters equally describe the linkage situation. Hence theorem 5.5 applies to $r(\alpha)$ which therefore forms an alternative set of independent linkage parameters.

7.3 As in 6.4, define $r(a, b)$ to be the average of $r(\alpha)$ of order $a+b$ over subsets of size a in one set of loci and over subsets of size b in another set of loci. Then, as in 7.21 with $c = -1$,

$$r(a, b) = \sum_{r=0}^a \sum_{s=0}^b (-2)^{r+s} \binom{a}{r} \binom{b}{s} q(r, s),$$

and

$$\begin{aligned} G_r(x, y) &= \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} r(s, t) x^s y^t / s! t! \\ &= \{\operatorname{sech}(x+y)\} G_q^*(2x, 2y), \end{aligned}$$

as in 7.23

or

$$G_q^*(x, y) = \{\cosh \tfrac{1}{2}(x+y)\} G_r(\tfrac{1}{2}x, \tfrac{1}{2}y).$$

Hence

$$\begin{aligned} G_l(x, y) &= \{\cosh \tfrac{1}{2}(x-y)\} G_r(\tfrac{1}{2}x, -\tfrac{1}{2}y)/(e^{-x} + e^{-y}) && \text{by 6.41} \\ &= \tfrac{1}{2}e^{\frac{1}{2}(x+y)} G_r(\tfrac{1}{2}x, -\tfrac{1}{2}y). \end{aligned}$$

Then

$$l(a, b) = 2^{-1-a-b} \sum_{s=0}^a \sum_{t=0}^b \binom{a}{s} \binom{b}{t} (-1)^t r(s, t),$$

and

$$7.31 \quad l(\alpha, \beta) = 2^{-1-n(\alpha+\beta)} \sum_{\rho \subset \alpha+\beta} (-1)^{n(\beta\rho)} r(\rho).$$

An inverse formula is obtained from the coefficient of x^r in

$$\begin{aligned} G_r(x, y) &= 2e^{-x+y} G_l(2x, -2y), \\ 7.32 \quad r(\alpha) &= 2 \sum_{\rho \subset \alpha} (-2)^{n(\rho)} l(\rho, 0), \end{aligned}$$

which is, of course, just the form 7.24 for $r(\alpha)$.

7.4 Comparison of Jones' and Schnell's Parameters

Since both these parameters may come into general use it is important at the outset to describe their relationship and their relative merits. This may reduce possible confusion over notation such as exists in other parts of biometrical genetics.

Both these parameters provide a complete description of the linkage situation with no redundancy. Both are defined in terms of even sets of loci. The definition of Jones' parameters can be applied also to odd sets of loci (as in this paper) to provide superfluous parameters and Schnell's parameters may be taken to be zero for odd sets of loci. The parameters are linearly related according to 7.24 or 7.25 and 7.26.

The advantage of Jones' parameter is its simple genetical definition— $q(\alpha)$ is the chance of no crossing-over in the set α of loci. Although only parameters of even order are necessary to describe the linkage system and parameters of odd order can be expressed in terms of those of even order, Jones' definition in its wider sense treats all sets of loci equally. On the other hand, Schnell's definition makes a distinction between even and odd sets of loci and is not easy to interpret genetically: it is difficult to imagine a simple linkage property which exists for even sets of loci and vanishes for odd sets.

The advantage of Schnell's parameter can be seen by comparing the coefficients in the two expressions for $l(\alpha, \beta)$:

$$l(\alpha, \beta) = \tfrac{1}{2} \sum_{\rho^* \subset \alpha+\beta} (-1)^{n(\beta\rho^*)} u[n(\alpha-\rho^*), n(\beta-\rho^*)] q(\rho^*), \text{ from theorem 6.4}$$

and

$$l(\alpha, \beta) = 2^{-1-n(\alpha+\beta)} \sum_{\rho \subset \alpha+\beta} (-1)^{n(\beta\rho)} r(\rho). \quad \text{from 7.31}$$

The coefficients in Schnell's expression are only plus and minus unity and the expressions for the various $l(\alpha, \beta)$ with constant $\alpha+\beta$ are orthogonal and would be expected to give simpler expectations of functions of gametic frequencies than Jones' expression. This seems to be borne out by the formulae in their two papers.

To sum up, Jones' parameter seems to be genetically preferable and Schnell's parameter mathematically more convenient. In major gene experiments Jones' parameter is probably the more suitable.

8. Appendix

8.1 This section contains further properties of the coefficient $u(r, s)$ and its relationship to $u(r)$ and $v(r)$. Since

$$\frac{\partial G_u(x, y)}{\partial x} + \frac{\partial G_u(x, y)}{\partial y} = G_u(x, y),$$

$$8.11 \quad u(r, s+1) + u(r+1, s) = u(r, s).$$

In the case of 8.11 with $s = r$,

$$\begin{aligned} u(r, r) &= u(r, r+1) + u(r+1, r) \\ &= 2u(r, r+1). \end{aligned} \quad \text{by 6.51}$$

$$8.12 \quad u(r, r+1) = \frac{1}{2}u(r, r).$$

Since

$$G_u(x, y) = e^y G_u(x-y), \quad \text{by theorem 6.4}$$

$$8.13 \quad u(r, s) = \sum_{t=0}^s (-1)^t \binom{s}{t} u(r+t).$$

The inverse relation is simply

$$u(r) = u(r, 0) = u(0, r). \quad \text{by 6.52}$$

Since

$$G_u(x, y) = e^{\frac{1}{2}(x+y)} G_v\left\{\frac{1}{2}(x-y)\right\},$$

$$8.14 \quad u(r, s) = 2^{-r-s} \sum_{t=0}^r \sum_{u=0}^s (-1)^u \binom{r}{t} \binom{s}{u} v(t+u).$$

Also since

$$G_u(x, -x) = G_v(x),$$

$$8.15 \quad v(r) = \sum_{s=0}^r (-1)^{r-s} \binom{r}{s} u(s, r-s).$$

Note that $(-1)^r v(2r)$ are Euler's numbers ($v(2r+1) = 0$).

8.2 The table of values of $u(r, s)$ may be constructed sequentially by addition, subtraction, and division by two. The following rules of construction arise from the properties of $u(r, s)$ in 6.5 and 8.1.

(a) Put $u(0, 0) = 1$ and follow the arrows in Table 2 to successive entries.

(b) A set of $u(r, s)$ with the same $r+s$ and $r \leq s$ lies perpendicular to the leading diagonal and may be called a *slope* of number $r+s$. The sequence follows increasing slope number and is down the even-numbered slopes using the relation

$$u(r+1, s) = u(r, s) - u(r, s+1), \quad \text{by 8.11}$$

and up the odd-numbered slopes using

$$u(r, s+1) = u(r, s) - u(r+1, s). \quad \text{by 8.11}$$

(c) The even slopes commence at the top with zero (6.52). The odd slopes commence at the bottom with

$$u(r, r+1) = \frac{1}{2}u(r, r), \quad \text{by 8.12}$$

$u(r, r)$ being the diagonal number at the foot of the previous even slope. As a check, the series $u(r)$ is generated by the first row or column (6.52).

(d) The lower part of the table is filled by symmetry using

$$u(s, r) = u(r, s). \quad \text{by 6.51}$$

9. Acknowledgments

I am indebted to two members of this Laboratory, Miss E. Stevenson for computing the tables and noticing the rules of construction of Table 2 and Dr. J. H. Darwin for shortening the proof of theorem 4.1, to Dr. F. W. Schnell for showing me the manuscript of his paper, and also to Prof. W. R. Address of Canterbury University for help with generating functions.

10. References

- ANDERSON, V. L., and KEMPTHORNE, O. (1954).—A model for the study of quantitative inheritance. *Genetics* **39**: 883–98.
- COCKERHAM, C. C. (1954).—An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859–82.
- FISHER, R. A. (1918).—The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edin.* **52**: 399–433.
- GEIRINGER, H. (1944).—On the probability theory on linkage in Mendelian heredity. *Ann. Math. Stat.* **15**: 25–57.
- HAYMAN, B. I. (1954).—A mathematical theory of gene action and interaction. Ph.D. Thesis, University of Birmingham.
- HAYMAN, B. I., and MATHER, K. (1955).—The description of gene interaction in continuous variation. *Biometrics* **11**: 69–82.
- JONES, R. M. (1961).—Linkage distribution and epistacy in quantitative inheritance. *Heredity* **15**: 153–9.
- MATHER, K. (1949).—“Biometrical Genetics.” (Methuen: London.)
- SCHNELL, F. W. (1961).—Some general formulations of linkage effects in inbreeding. *Genetics* **46**: 947–57.
- WALLACE, M. E. (1959).—An experimental test of the hypothesis of affinity. *Genetica* **29**: 243–55.